UNIVERSITÉ DE TOURS

ÉCOLE DOCTORALE : Sciences de la Société : Territoires, Economie et Droit – SSTED

UMR 7324 Cités Territoires Environnement et Sociétés (Citères)

Equipe de recherche Dynamiques et Actions Territoriales et Environnementales

THÈSE Présentée par:

Hedieh SOLTANPOUR

soutenue le : **18 December 2024**

pour obtenir le grade de: **Docteur de l'université de Tours**
Discipline/ Spécialité: Aménagement et environnement

Évaluation de la susceptibilité multi-aléas dans les zones karstiques : étude de cas dans la région d'Orléans, France

THÈSE dirigée par:

M. SERRANO José	Professeur des universités, Université de Tours
M. SERRHINI Kamal	Maître de conférences, Université de Tours
M. C. GILL Joel	Maître de conférences, Cardiff University

RAPPORTEURS:

Mme. BANKS Vanessa	Chercheuse senior, British Geological Survey
M. FUCHS Sven	Chercheur senior, BOKU University

JURY: ***(directeurs, rapporteurs, examinateurs)***

M. SERRANO José	Professeur des universités, Université de Tours
M. SERRHINI Kamal	Maître de conférences, Université de Tours
M. C. GILL Joel	Maître de conférences, Cardiff University
Mme. BANKS Vanessa	Chercheuse senior, British Geological Survey
M. FUCHS Sven	Chercheur senior, BOKU University

Signature du directeur de thése:

Abstract

The core aim of this thesis is to contribute to improving multi-hazard risk management in the context of karst terrains by developing and evaluating an integrated multi-hazard susceptibility map of two interrelated hazards (flood-triggered sinkholes) using a less-explored machine-learning algorithm (Maximum Entropy). Additionally, the effectiveness of the resultant map is assessed through feedback from local experts using an online survey. Therefore, this thesis integrates two research components: (i) multi-hazard susceptibility map modelling, and (ii) stakeholder engagement and evaluation. This research is motivated by the global gap in perceiving karst terrains as multi-hazard areas (evidenced in this thesis) and the need for multi-hazard assessments in French disaster risk reduction planning. The current thesis is divided into six chapters as follows:

Chapter 1, as an introductory chapter, summarizes the purpose and contents of the thesis and discusses briefly the research background, motivation, aims and objectives and research outline.

Chapter 2 presents the introduction to the multi-hazard concept and modelling followed by a critical literature review of the studies on sinkholes constructing a literature database and identifying the current gap in sinkhole hazard assessments. I use my literature database to demonstrate current understanding and perceptions relating to this thesis topic. My critical review identifies **241 papers** published in international scientific journals, following an approach of being representative not exhaustive. These papers were classified into *seven thematic subjects* for the sinkhole hazard studies. It concludes with a review of the current sinkhole studies worldwide and the previous work done in the study area justifying the need for this research.

Chapter 3 presents a detailed research design and methodology focused on applying the Maximum Entropy (MaxEnt) algorithm, alongside introducing Val d'Orléans as the case study area. It consists of four core sections; *Section 3.2* delves into the foundational principles of the MaxEnt algorithm, exploring its historical roots and applications in natural hazard science, emphasizing its effectiveness in modeling natural phenomena using presence-background data. *Section 3.3* describes the unique geographical, geological, and hydrogeological characteristics of Val d'Orléans, highlighting its susceptibility to hazards such as sinkholes and floods and the need for a regional multi-hazard mapping. *Section 3.4* outlines the data collection processes, detailing the implementation of MaxEnt, selecting (geo)-environmental variables, processing data, and configuring the software for optimal performance. *Section 3.5* explains the preparation of data for modeling flood-triggered sinkhole occurrences, including data formatting, parameter tuning, and data partitioning strategies, emphasizing the methodological rigor required for accurate hazard modeling in our research.

Chapter 4 presents the results of Maximum Entropy (MaxEnt) modeling. The sequence of results is structured to build understanding of the model's forecasts progressively. I begin in *Section 4.2.1* by evaluating the model's overall performance through the Receiver Operating Characteristic (ROC) curve, providing a baseline measure of accuracy, and explaining the ROC curve to discuss our findings. *Section 4.2.2* presents the spatial distribution of flood-triggered sinkholes produced by MaxEnt. *Section 4.2.3* delves into the influence of both individual and combined (geo)-environmental variables on the model's forecast, presented through model-specific and marginal response curves. *Section 4.2.4* assesses the robustness and reliability of these predictions with a Jackknife test, highlighting the contribution of each variable. *Section 4.3* spatially compares the official flood hazard map with our multi-hazard susceptibility map, aiming to identify areas of overlap, which is crucial for prioritizing resource allocation, informing policy and

planning, and enhancing emergency preparedness. The comparison details are then presented in *Appendix 4.A* to maintain a streamlined narrative.

Chapter 5 focuses on the stakeholder engagment process aimed at examining expert stakeholder awareness and perceptions of multi-hazard events, evaluating the produced multi-hazard susceptibility map presented in *Chapter 4*, and understanding participants' perspectives on the ownership and pervasive nature of flooding and sinkhole hazards in the area. Additionally, it explores participants' views on establishing a dedicated DICRIM[1] document for multi-hazard interactions. The results of this study are crucial for creating an improvement roadmap, integrating qualitative evidence and expert feedback, with an emphasis on multi-hazard mapping, awareness, and fostering multi-hazard thinking.

Chapter 6 is a concluding chapter. Within this final chapter, we provide a summary of the thesis, encapsulating the main findings and discussions from the earlier chapters. Elements of this summary have been previously mentioned in the concluding sections of *Chapters 2, 3, 4*, and *5*. However, this chapter integrates those summaries to present a unified overview of the research. Additionally, we address the limitations and uncertainties in the modelling process. The chapter concludes with recommendations for future research directions.

[1] A French public awareness document, translated as Municipal Information Document on Major Risks (see: https://www.georisques.gouv.fr/communiquer-aupres-de-mes-administres/dicrim).

French Abstract/ Résumé

L'objectif principal de cette thèse est de contribuer à l'amélioration de la gestion des risques multi-aléas dans le contexte des terrains karstiques, en développant et en évaluant une carte intégrée de susceptibilité multi-aléas pour deux dangers interconnectés (dolines déclenchées par des inondations), en utilisant un algorithme d'apprentissage automatique moins exploré (Maximum Entropy). L'efficacité de la carte produite est également évaluée grâce aux retours d'experts locaux recueillis via un sondage en ligne. Cette thèse intègre donc deux composantes de recherche : la modélisation d'une carte de susceptibilité multi-aléas et l'engagement des parties prenantes pour son évaluation. Cette recherche est motivée par une lacune mondiale dans la perception des terrains karstiques comme zones à risques multiples et par le besoin d'évaluations multi-aléas dans la planification de la réduction des risques de catastrophes en France. La thèse est divisée en six chapitres:

Le **chapitre 1** introduit la thèse en résumant son objectif, son contenu et en présentant brièvement le contexte de la recherche, la motivation, les buts, les objectifs et le plan de la recherche.

Le **chapitre 2** propose une introduction au concept et à la modélisation des aléas multiples, suivie d'une revue critique de la littérature sur les études des dolines, avec la constitution d'une base de données bibliographique et l'identification des lacunes actuelles dans les évaluations des risques de dolines. La revue critique, basée sur 241 articles publiés dans des revues scientifiques internationales (sélection représentative), permet de classer les études sur les dolines en sept thématiques principales. Ce chapitre se conclut par une analyse des recherches actuelles sur les dolines à l'échelle mondiale et des travaux antérieurs réalisés dans la zone d'étude, justifiant la pertinence de cette recherche.

Le **chapitre 3** détaille le design de recherche et la méthodologie appliquée, avec un accent particulier sur l'utilisation de l'algorithme Maximum Entropy (MaxEnt) et la présentation de Val d'Orléans en tant que zone d'étude. Il se compose de quatre sections : la section 3.2 explore les principes fondamentaux de l'algorithme MaxEnt, ses origines historiques et ses applications dans la modélisation des aléas naturels en utilisant des données de présence et de fond. La section 3.3 décrit les caractéristiques géographiques, géologiques et hydrogéologiques du Val d'Orléans, en soulignant sa vulnérabilité aux aléas tels que les dolines et les inondations et le besoin d'une cartographie régionale multi-aléas. La section 3.4 explique les processus de collecte de données, y compris la mise en œuvre de MaxEnt, la sélection des variables (géo)-environnementales, le traitement des données et la configuration du logiciel pour des performances optimales. La section 3.5 décrit la préparation des données pour la modélisation des dolines déclenchées par les inondations, en détaillant le formatage des données, l'ajustement des paramètres et les stratégies de partitionnement, en insistant sur la rigueur méthodologique requise pour une modélisation précise des risques.

Le **chapitre 4** présente les résultats de la modélisation par Maximum Entropy (MaxEnt). La présentation des résultats suit une progression logique pour approfondir la compréhension des prévisions du modèle. La section 4.2.1 commence par évaluer la performance globale du modèle via la courbe ROC (Receiver Operating Characteristic), fournissant une mesure de précision de base. La section 4.2.2 montre la distribution spatiale des dolines déclenchées par les inondations générée par MaxEnt. La section 4.2.3 analyse l'influence des variables (géo)-environnementales individuelles et combinées sur les prévisions du modèle, présentée à travers des courbes de réponse spécifiques et marginales. La section 4.2.4 évalue la robustesse et la fiabilité des prévisions avec un test Jackknife, mettant en lumière la contribution de chaque variable. La section 4.3 compare spatialement la carte officielle des risques d'inondation avec notre carte de susceptibilité multi-aléas, pour identifier les zones de chevauchement, essentielles pour prioriser

l'allocation des ressources, informer les politiques, planifier et renforcer la préparation aux urgences. Les détails de cette comparaison sont présentés en annexe 4.A pour maintenir un fil narratif cohérent.

Le **chapitre 5** se concentre sur le processus d'évaluation, explorant la sensibilisation et les perceptions des experts sur les événements multi-aléas, évaluant la carte de susceptibilité multi-aléas produite au chapitre 4, et recueillant les perspectives des participants sur la gestion des risques d'inondation et de dolines dans la région. Ce chapitre examine également l'intérêt des participants pour l'établissement d'un document DICRIM dédié aux interactions multi-aléas. Les résultats de cette étude sont cruciaux pour élaborer une feuille de route d'amélioration, intégrant des preuves qualitatives et les retours des experts, avec un accent sur la cartographie multi-aléas, la sensibilisation et la promotion d'une approche intégrée des risques.

Le **chapitre 6**, conclusif, offre un résumé de la thèse en regroupant les principales conclusions et discussions des chapitres précédents. Bien que des éléments de ce résumé aient été mentionnés dans les conclusions des chapitres 2, 3, 4 et 5, ce chapitre les intègre pour offrir une vue d'ensemble cohérente de la recherche. En outre, il aborde les limitations et incertitudes du processus de modélisation et se termine par des recommandations pour les futures directions de recherche.

Acknowledgements

I would like to express my deepest gratitude to my internal supervisors, Professor José Serrano and Dr. Kamal Serrhini, for providing me with the opportunity to learn and grow throughout my PhD research. Their unwavering support, guidance, and expertise have been invaluable, and I am deeply appreciative of their mentorship. I am especially grateful to all the insightful meetings I had with Dr. Kamal Serrhini during which I could learn and rethink my approaches.

I am very thankful to Dr. Joel C. Gill (Cardiff University), my external co-supervisor, for the time he invested in me. His passion for science and research has been truly contagious, and I am grateful for the thoughtful comments he patiently made on both my thesis and my paper. The insightful discussions we had were instrumental in shaping my research, and his advice and encouragement supported me immensely along my PhD journey.

I am also deeply grateful to Dr. Solmaz Mohadjer (University of Tübingen), who mentored me through my PhD journey. Her constant support, generosity in sharing her knowledge and experience, and willingness to help at every step were invaluable to me. She was always there, offering guidance and motivation, and I am incredibly fortunate to have had her as a mentor.

During my research, I had the opportunity to spend a few days at BRGM (the French Geological Survey). I would like to thank Mr. Gildas Noury and all the other wonderful people at BRGM who offered warm hospitality and generously helped by providing data, suggesting references, allowing me to view and read unpublished manuscripts. Their assistance greatly enriched my research.

Additionally, I am thankfull and appreciative of all the brilliant experts at BRGM and CEREMA (Centre for Studies and Expertise on Risks, Environment, Mobility, and Urban and Country planning) who willingly participated in my online survey. Moreover, I appreciate all the administrative staff at the University of Tours who kindly and patiently helped me navigate the administrative processes during my PhD years. I also feel compelled to thank my referees for showing their interest and for agreeing to be part of this journey by investing time in examining this thesis.

Last but certainly not least, it is often said that it takes a village to conduct PhD research, and I could not agree more. Many people have walked this journey with me, and I am deeply grateful to them all. I am especially in awe of the unwavering support of my dear partner and my family. Without their love, encouragement, patience, and unyielding support, I would not have been able to complete and submit this research. I would like to sincerely thank them for everything they have done for me; their belief in me has been my greatest strength. I dedicate this work to my beloved partner, my mother, and my father.

Remerciements

Je souhaite exprimer ma plus profonde gratitude à mes directeurs de thèse, le Professeur José Serrano et le Dr Kamal Serrhini, pour m'avoir offert l'opportunité d'apprendre et de grandir tout au long de ma recherche doctorale. Leur soutien indéfectible, leurs conseils et leur expertise ont été inestimables, et je suis profondément reconnaissante pour leur mentorat. Je tiens tout particulièrement à remercier le Dr Kamal Serrhini pour les nombreuses réunions enrichissantes que nous avons eues, au cours desquelles j'ai pu apprendre et repenser mes approches.

Je remercie également chaleureusement le Dr Joel C. Gill (Université de Cardiff), mon co-directeur externe, pour le temps qu'il m'a consacré. Sa passion pour la science et la recherche a été véritablement contagieuse, et je suis reconnaissante pour ses commentaires réfléchis et ses corrections patients sur ma thèse et mes articles. Les discussions perspicaces que nous avons eues ont été déterminantes dans la structuration de ma recherche, et ses conseils et encouragements m'ont immensément soutenue tout au long de mon parcours doctoral.

Je suis aussi profondément reconnaissante envers le Dr Solmaz Mohadjer (Université de Tübingen), qui m'a encadrée tout au long de mon doctorat. Son soutien constant, sa générosité dans le partage de ses connaissances et de son expérience, et sa disponibilité à chaque étape ont été inestimables pour moi. Elle a toujours été présente pour m'offrir ses conseils et sa motivation, et j'ai eu la grande chance de l'avoir comme mentor.

Au cours de ma recherche, j'ai eu l'opportunité de passer quelques jours au BRGM (Bureau de Recherches Géologiques et Minières). Je tiens à remercier M. Gildas Noury ainsi que toutes les autres personnes formidables du BRGM pour leur accueil chaleureux et leur aide précieuse en me fournissant des données, en me suggérant des références, et en me permettant de consulter des manuscrits non publiés. Leur assistance a grandement enrichi ma recherche.

Je souhaite également exprimer ma reconnaissance envers tous les brillants experts du BRGM et du CEREMA (Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement) qui ont participé de manière volontaire à mon enquête en ligne. De plus, je remercie tout le personnel administratif de l'Université de Tours qui m'a aidée avec gentillesse et patience à naviguer dans les démarches administratives pendant mes années de doctorat. Je me sens également obligée de remercier mes rapporteurs pour l'intérêt qu'ils ont manifesté et pour avoir accepté de faire partie de cette aventure en consacrant du temps à l'examen de cette thèse.

Enfin, il est souvent dit qu'il faut un village pour mener à bien une recherche doctorale, et je ne pourrais être plus d'accord. De nombreuses personnes m'ont accompagnée tout au long de ce parcours, et je leur en suis profondément reconnaissante. Je suis particulièrement touchée par le soutien indéfectible de mon cher partenaire et de ma famille. Sans leur amour, leurs encouragements, leur patience et leur soutien inébranlable, je n'aurais pas pu achever et soumettre cette recherche. Je tiens à les remercier sincèrement pour tout ce qu'ils ont fait pour moi ; leur confiance en moi a été ma plus grande force. Je dédie ce travail à mon partenaire bien-aimé, ma mère et mon père.

Table of Contents

Chapter 1. Introduction 1
1.1 The importance of multi-hazard considerations 1
1.2 Initial and ongoing confusion (Multi-layer single hazards Vs Multi-hazard assessment) 3
1.3 Multi-hazard (susceptibility) mapping in Disaster Risk Management Cycle (DRM) 4
1.4 Motivations 6
1.4.1 Multi-hazard forming areas (a globally identified gap) 6
1.4.2 Institutional gaps in natural hazard management in France 8
1.5 Research aims and objectives 9
1.6 Research outline 10

Chapter 2. Literature review (State of the art) 12
2.1 Introduction 12
2.1.1 Exploring the origins and definitions of multi-hazards 12
2.1.2 Multi-hazard interrelationships 16
2.2 Approaches to multi-hazard assessment 18
2.2.1 Summary of qualitative and quantitative approaches for multi-hazard assessment 18
2.2.2 Stochastic approaches 20
2.2.3 Empirical approaches 21
2.2.4 Mechanistic approaches 23
2.3 Karst terrains 29
2.3.1 Karst terrains as a multi-hazard-forming environment 31
2.4 Building a systematic sinkhole-related literature review database 34
2.4.1 Data and Methodology 34
2.4.2 Theme 1. Sinkhole Monitoring and Detection 39
2.4.3 Theme 2: Sinkhole Morphology Studies 41
2.4.4 Theme 3. The interconnectivity of flood hazard and karst sinkholes 43
2.4.5 Theme 4. Mitigation Measures for Sinkholes 48
2.4.6 Theme 5. Sinkhole hazard/susceptibility and risk assessment 49
2.4.6a Geo-technical investigations and geophysical methods 51
2.4.6b Qualitative and semi quantitative methods 52
2.4.6c Quantitative methods 55
2.4.6d Machine Learning methods (ML) 57

2.4.7 Theme 6. Multi-hazard mapping/ modelling in karst focusing on flood and sinkholes 59
2.4.8 Theme. 7 Others 60
2.5 The main previous research in mapping karst hazard done in the Val d'Orléans 61
2.6 Discussions and conclusions 65

Chapter 3. Methodology and Case study 68
3.1 Introduction 68
3.2 Maximum Entropy Modeling (MaxEnt) background 69
3.2.1 The principal of MaxEnt 71
3.2.2 MaxEnt entropy model in natural hazards 73
3.3 Case study: the region of Orléans (Val d'Orléans) 77
3.3.1 General background 77
3.3.2 Geological background 79
3.3.3 Hydrogeological context 81
3.3.4. Val d'Orléans, a multi-hazard forming setting 82
3.3.4a The karst system of the Val d'Orléans 83
3.3.4b Underground cavities and sinkholes in Val d'Orléans 86
3.3.4c Floods of the Loire River 90
3.3.4d Flood risk management and urban planning in the Val d'Orléans 92
3.4 Methodology and data collection 93
3.4.1 MaxEnt implementation (Transition from theory to practical application) 93
3.4.2 Natural sinkhole inventory 95
3.4.3 Geo-environmental conditioning factors for flood-triggered sinkholes 97
3.4.3a Collinearity test 99
3.4.3b Topography data (Elevation and Slope) 104
3.4.3c Surficial geology 107
3.4.3d Alluvium thickness 108
3.4.3e Groundwater spatial data 109
3.4.3f Flood hazard map 114
3.5 Data Processing and Preparation for MaxEnt Modelling 115
3.5.1 Data preparation procedures 115
3.5.1a Creating a Comma-Separated Value (.csv) file 115
3.5.1b Modifying (geo)-environmental layers 116
3.5.1c Conversion of modified geo-environmental layers to ASCII Format 117
3.5.2 Model building process 117

3.5.2a Parameter setting (tuning program settings) 118
3.5.2b Optimal threshold value: 119
3.5.2d Data partitioning for model training and testing (cross validation) 120
3.5.2c MaxEnt model outputs 122
3.6 Discussion and conclusion 124

Chapter 4. The results of multi-hazard susceptibility modelling performed with MaxEnt 125
4.1 Introduction 125
4.2 Modelling Results 127
4.2.1 Model validation 127
4.2.2 Multi-hazard susceptibility mapping (the probability distribution) 131
4.2.3 Analysing the impact of contributing factors on MaxEnt modelling 134
4.2.3a Model-Specific Response Curves (MSRC) 134
4.2.3b Marginal response curves 138
4.2.4 Analysis of variable contributions 140
4.2.4a Jacknife test implication 142
4.3 Comparative analysis of single Vs multi-hazard maps (Overlaps and Divergences) 144
4.4 Discussion and Limitations 148
4.5 Conclusion 152

Chapter 5. Creating a road map to strengthen multi-hazard mapping for decision support 153
5.1 Introduction 153
5.2 Methodology 156
5.2.1) Stakeholder Identification 156
5.2.2) Questionnaire design 159
5.2.2a Ethical considerations (identifying potential ethical dilemmas) 160
5.2.2b Reach out to the selected participants 161
5.3 Data analysis 162
5.3.1 Questionnaire results 162
5.3.1a Awareness of multi-hazard concept/ scenarios 162
5.3.1b Part 2. Evaluating the produced multi-hazard susceptibility map 166
5.3.1c Part 3. Risk acceptability (readiness to embrace the risk) 171
5.3.1d Part 4. The means of awareness and communication on multi-hazards 172
5.4 Discussion 175
5.4.1 Overview and discussion of key findings 175

5.4.2 Limitations 178
5.5 Way forward 179
5.5.1 Recommendations on practices to foster multi-hazard thinking 179
5.5.2 Future improvements in multi-hazard/susceptibility mapping 180
5.5.3 Multi-hazard integration 182
5.6 Conclusion 183

Chapter 6. Discussion and future directions 184
6.1 Introduction 184
6.2 Summary 184
6.3 Research Limitations 188
6.3.1 Limitations on multi-hazard mapping 188
6.4 Future directions for multi-hazard mapping/modelling: 191
6.5 Future directions for multi-hazard integration into the disaster risk reduction plan in the study region 192
6.6 Final thoughts 193

References (APA format) 194

APPENDICES

Appendix 2.A The 241 references used to create the violin plot of sinkhole-related studies 214
Appendix 3.A 3D elevation view and scanned geological map 233
Appendix 3.B Piezometric data for five nearby stations used for interpolation 235
Appendix 3.C Groundwater fluctuation graphs 236
Appendix 3.D R codes (scripts) for data processing to comply with MaxEnt requirements 239
Appendix 4.A The close-up maps' segments in comparing flood vs flood-triggered sinkhole maps 247
Appendix 4.B Insight from applying MaxEnt to a broader geographical area 255
Appendix 4.C The resultant model-specific curves and marginal curves for each (geo)-environmental variable (Side by side comparison) 260
Appendix 5.A Questionnaire 266
Appendix E Additional field visit photos 273

Index of Figures

Figure 1.1 Progression of vulnerability before and after sequential hazard events over time. 2

Figure 1.2 Disaster risk management cycle for consecutive disasters 5

Figure 1.3 Global multi-hazard hotspots. Identification the multi-hazard forming areas 7

Figure 2.1 The database on multi-hazard literature..... 13

Figure 2.2 The diagram shows an overview of the main modelling methods for quantifying different types of interrelations..... 19

Figure 2.3 An overview of the worldwide karst distribution (Chen et al., 2017). 29

Figure 2.4 Genetic classification of sinkholes. 30

Figure 2.5 A simplified sequential diagram of the potential multi-hazard interactions in karst covered terrains.. 33

Figure 2.6 The flowchart of the construction of the systematic literature review database. 35

Figure 2.7 The evolution of sinkhole-related studies. 38

Figure 2.8 An overview of sinkhole hazard/susceptibility modelling classification 50

Figure 3.1 The mind map diagram of main natural hazard modelled by MaxEnt. 75

Figure 3.2 Geographical location of the study area situated within the Loiret department, France. 78

Figure 3.3 Succession of lithological units of the Val d'Orléans. 80

Figure 3.4 The simplified profile of the geological cross-section in Val d'Orléans overlying the Beauce limestone. 80

Figure 3.5 A cross- section diagram of the two superimposed aquifers. 81

Figure 3.6 The flood disturbed the access to the urban areas for several days. 82

Figure 3.7 After the water receded in the 2016 event, numerous sinkhole collapses were observed around roads and buildings. 83

Figure 3.8 The map of karst network in the Val d'Olréans..... 84

Figure 3.9 Map of the Val d'Orléans Aquifer. 85

Figure 3.10 a,b) Close-up view of samples depicting eroded carbonate rocks highlighting the highly karstified bedrock in the area 86

Figure 3.11 Recent sinkhole incident and the evacuated buildings 88

Figure 3.12 Sinkhole karst collapse opened up around mid-March 2021 on the Chevauchée dike..... 89

Figure 3.13 Locations of past sinkhole collapses on the right bank of the Loire River dyke. 89

Figure 3.14 The 19th-century flood picture depicts the extend of the devastating historical flood of the Loire River 92

Figure 3.15 The methodological flowchart. 94

Figure 3.16 Locations of past sinkholes overlaid on the elevation layer. 95
Figure 3.17 Correlation between distance to dyke and sinkholes' frequency. 96
Figure 3.18 The rasterized map of Digital Elevation Model (DEM) of the study area 105
Figure 3.19 The rasterized map of Slope derived from DEM (created by the author). 106
Figure 3.20 Surficial geology consisting of Holocene (new) and Pleistocene (old) alluvium layers 107
Figure 3.21 The rasterized map of alluvium thickness. 108
Figure 3.22 The piezometric map of the Val d 'Orléans. 110
Figure 3.23 The rasterized map of groundwater relative level. 113
Figure 3.24 The official flood hazard map 114
Figure 3.25 A screenshot of CSV format for sinkhole occurrence data in MaxEnt input samples 116
Figure 3.26 A screenshot of R script for unifying the data layers to the MaxEnt format. 117
Figure 4.1 A fundamental ROC graph displaying five distinct classifiers 128
Figure 4.2 The Area Under the Curve (AUC) results. 130
Figure 4.3 Classified multi-hazard susceptibility map generated by MaxEnt model 132
Figure 4.4 Classified multi-hazard susceptibility map generated by MaxEnt model 133
Figure 4.5 MaxEnt model MSRC curves and histogram for flood-triggered sinkhole susceptibility factors 135
Figure 4.6 Marginal response curves. 139
Figure 4.7 The plot of the jackknife test results of the averages over 50 replicate runs contributing factor importance 143
Figure 4.8 Single vs. Multi-Hazard Comparison.. 145
Figure 4.9 Spatially comparing total areas covered by different susceptibility classes for two related yet distinct hazards 146
Figure 5.1 An overview of French regulatory prevention information across the municipality 157
Figure 5.2 Expert responses to the familiarity of the multi-hazard concept. 163
Figure 5.3 Perceived the main multi-hazard scenario in the Val d'Orléans area. 164
Figure 5.4 The likelihood of the identified multi-hazard scenario in the area, both generally and within the context of climate change 165
Figure 5.5 The general perception of the map hazard level based on the colour schemes. 167
Figure 5.6 The hazard level perception of the dark green colour 168
Figure 5.7 Experts' opinions on the colour that best communicates zones with lower susceptibility, not free from susceptibility. 168
Figure 5.8 Participants' perception of probabilistic forecasting nature of the map. 169
Figure 5.9 Expert recommendations for strategic responses to flood-triggered sinkholes. 171
Figure 5.10 Responses to the proposal of providing a separate DICRIM document specifically including multi-hazard information/maps. 173

Figure 5.11 Results of the ranking of main barriers to integrating multi-hazard thinking into local risk and resilience planning strategies. 174

Index of Tables

Table 2.1 Key definitions 15
Table 2.2 Different classification of hazard relationships by various authors. 17
Table 2.3 An updated literature review of the three investigated approaches 24
Table 2.4 Identified causal and triggering factors to sinkhole formations extracted from the literature. 43
Table 2.5 Summary of key publications on the interaction of flooding and sinkholes. 47
Table 2.6 Saaty's scale for pairwise weight assignment (Saaty, 2008) 53
Table 2.6 A summary of the Cerema (2014) research result. 61
Table 2.7 Synthesizing the previous research done in the study area. 64
Table 3.1 Spatial contributing factors included in the MaxEnt model. 98
Table 3.2 Pearson Correlation coefficients among geographical and environmental variables. 101
Table 3.3 Variance Inflation Factor (VIF) matrix. 103
Table 3.4 Five piezometric gauges in the Val d'Orleans 112
Table 4.1 Spatial probability distribution of flood-triggered sinkholes in Val d'Orléans. 132
Table. 4.2 Relative contributions and permutation importance of environmental variables 141
Table 5.1 An overview of the expert definitions of multi-hazard concepts across different fields. 163
Table 6.1 The main sources of model uncertainty 190

Peer review Journal papers

- **Soltanpour, H**., Serrhini, K., Gill, J., Fuchs, S., and Mohadjer, S. (**2024**) Multi-hazard susceptibility mapping in the karst context using a machine-learning method (MaxEnt)., In NHESS., In review.

- Serhinni K, Moy Y, Noury G, Serrano J, **Soltanpour, H** (**2023**). Non-structural vulnerability and territorial resilience in multi-risk situations, Canadian journal of regional science, 71CJRS/RCSR, Volume 46, Number 2. https://doi.org/10.7202/1100216ar (in French).

Conference presentations

- Journées Territoires et Immersions – Lyon – France, **Oral presention** on "Mapping Multi-Hazard Susceptibility Using Maximum Entropy (MaxEnt) Model" 3/4 Octobre **2024**.

- The Needs conference – Enschede - Netherlands, **Poster presentation** on Multi-hazard Susceptibility Mapping in the Karst context Using a Machine-learning Method in Val d'Orléans, October-November **2023**.

- EGU22 – Vienna - Austria, **Oral presentation** on "A Critical Research Gap Study of Sinkhole Hazard Assessments", May **2022**.

PhD schools and Workshops

- PhD school on "Dilemmas and opportunities of fieldwork/in situ research in disaster studies" at the University of Twente, Netherlands - October 30, **2023**.

- Participated in courses related to environmental risks and natural hazards at BOKU University, Vienna, as an exchange student (Feb - June, **2022**). Courses included:

    i) Environmental Risk Analysis and Management
    ii) Simulation Models in Natural Hazards Analysis
    iii) Protection and Mitigation Measures Against Natural Hazards
    iv) Flood Forecasting and Flood Protection

Key machine learning terminology for the thesis:

An algorithm: "A series of steps or set of rules that carries out an action or solves a problem" (Sullivan, 2022)

Bias: "An error and a tendency to consistently learn the same wrong thing". (Domingos, 2012)

Categorical data: "Data that take discrete values and have no ordering" (Nilson, 2005)

Classifier: "A classifier is a system that inputs typically a vector of discrete and/or continuous feature values and outputs a single discrete value, the class." The process of predicting the class or category of a given data point. For example, classifying hazardous areas as "hazardous" or "not hazardous" (Beguería, 2006)

Collinearity: Refers to a situation in which two or more predictor variables in a statistical model are highly correlated with each other. This can lead to problems in the model, such as unstable or biased estimates of the effects of each predictor variable (Dormann et al., 2013)

Continuous data: "Data that take infinite values and are ordered" (Nilson, 2005)

High dimensional data: Refers to data sets that have a large number of features or variables relative to the number of observations (Domingos, 2012)

Machine Learning: Machine Learning involves using mathematical models and algorithms to allow a computer to learn from and make decisions based on data (Nilson, 2005)

Model validation: Refers to "comparing the model predictions with a real-world dataset, for assessing its accuracy or predictive power" (Beguería, 2006)

Overfitting: A modelling error. It occurs when a model is trained on a set of data, and during training, it becomes overly complex and tries to fit the training data very closely to specific data points (Domingos, 2012)

Prediction: Refers to the output of a machine-learning model after it has been trained on a dataset.

Regularization: "Refers to smoothing the model, making it more regular, so as to avoid fitting too complex a model. In MaxEnt the regularization parameters can be changed if required." (Elith et al., 2011)

Supervised learning: "A type of machine learning where the model is trained on labeled data, meaning the input comes with the desired output." (Bi et al., 2019)

Testing set: "A subset of the dataset used to provide an unbiased evaluation of a final model fit on the training dataset."

Threshold: "The level of magnitude of a system process at which sudden or rapid change occurs. A point or level at which new properties emerge in an ecological, economic or other system, invalidating predictions based on mathematical relationships that apply at lower levels" (IPCC, 2007)

Trade-off: Refers to a situation where achieving one desirable quality or attribute in a model requires compromising another (e.g., accuracy vs. Simplicity or Bias vs.Variance).

Training set: “A subset of the dataset used to train a model to understand patterns and features in the data”. (Domingos, 2012).

Variance: is an error, which is a tendency to learn random things regardless to the real signal (Domingos, 2012). (Domingos, 2012).

Chapter 1. Introduction

1.1 The importance of multi-hazard considerations

Natural hazards occur in diverse environments worldwide, encompassing various types of threats. These hazards pose risks, leading to human casualties and substantial economic losses on an annual basis. Based on Munich Re's reports (2023), a leading global re-insurance company, the total economic losses worldwide resulting from disasters caused by natural hazards were approximately USD 247 billion in 2020, USD 320 billion in 2021, and USD 270 billion in 2022. Consequently, the international community has consistently prioritized addressing these concerns (Munich Re, 2023).

The 2023 EM-DAT, the international disaster loss database, report documents a significant year for disasters resulting from natural hazards, recording 399 events that resulted in 86,473 fatalities and affected 93.1 million people, with economic damages of USD 202.7 billion (CRED, 2024). Major disasters included the devastating earthquake in Turkey and Syria, which alone caused 56,683 deaths and USD 42.9 billion in damages, impacting 18 million individuals (CRED, 2024). Other notable events were the Indonesian drought affecting 18.8 million people and Storm Daniel in Libya, which led to 12,352 casualties and USD 6.2 billion in losses (CRED, 2024).

Given our ever-changing and interdependent world, where everything depends on everything else (Saaty, 2008) and also with human interventions being on the rise in the environment, natural hazards can no longer be perceived as independent occurrences but interconnected hazard networks (de Ruiter et al., 2020). According to the World Bank's report by Dilley et al. (2005, p 2), approximately 3.8 million km^2 of land surface and 790 million people worldwide face significant exposure to at least two natural hazards. Additionally, around 0.5 million km^2 and 105 million people are exposed to three or more hazards (Dilley et al., 2005). In 2023, global multi-hazard events affected millions of people worldwide. For instance, in Mozambique, nearly 1.5 million people were impacted by a combination of natural hazards, including cyclones, floods, and disease outbreaks (UNICEF, 2024).

In reality, when faced with the need to address real-life challenges, decision makers usually have to deal with more than one hazard and risk at the same time (Di Mauro, 2005). At this moment, the issues like how to consider the interrelated hazards and their impacts arise (Gill et al., 2022). The traditional approach to addressing natural hazards has treated each hazard individually as single events, which can often create confusion among decision makers, particularly in critical circumstances (e.g., see Kappes et al., 2012). Managing natural hazards in isolation fails to recognize the interconnection and potential cascading effects that can arise from multiple hazards (Gill and Malamud, 2014; 2016). In other words, according to Gill and Malamud (2016), narrowing the focus to a limited portion of the entire Earth system, instead of understanding its dynamic interactions, could lead to decisions that increase vulnerability to overlooked hazards as each hazard event will potentially increase subsequent levels of vulnerability (**Fig. 1.1**). Consequently, the prevalence of single-hazard approaches inevitably leads to a distortion of management priorities (Kappes et al., 2012).

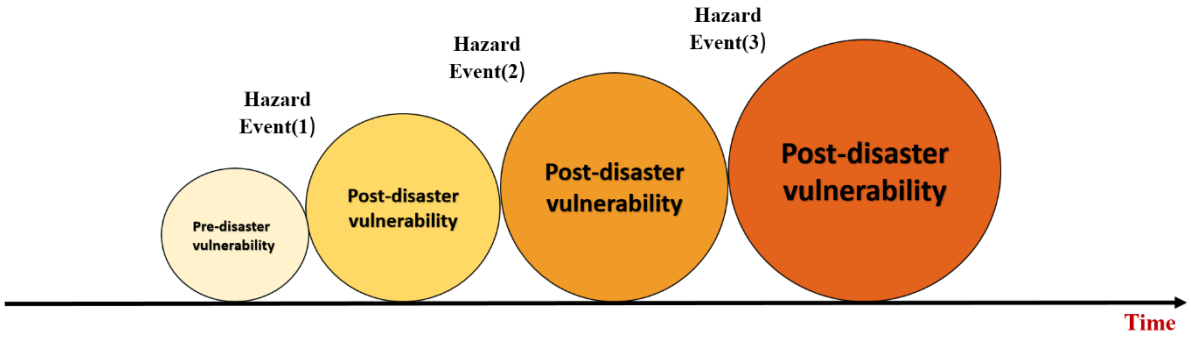


Figure 1.1 Progression of vulnerability before and after sequential hazard events over time. The size and the colour of the circles increases and changes (darker colour bigger in vulnerability) proportionally to represent the growing impact of each hazard event (modified from Gill and Malamud, 2016; Tilloy et al., 2019).

To manage natural hazards sustainably within a community, it is essential to move beyond assessing individual hazards. Equally important is gaining insights into the complex interplay between various hazards and their cascading effects. In this regard, the adoption of a "multi-hazard" methodology serves as a valuable approach to enhance our understanding of the prospective hazards confronting societies (Kappes et al., 2010; Kappes et al., 2012; Scolobig et al., 2017; Nachappa et al., 2020). Today, an increasing number of international and intergovernmental frameworks, as well as regional organizations, have turned to this new approach. This approach posits that if a community is subjected to more than one hazard, resilience and sustainability will be enhanced by considering all hazards and their potential interactions (Padilla et al., 2021; UNEP, 1992; UNISDR, 2005; UNISDR, 2015). The United Nations, through the Sendai Framework for Disaster Risk Reduction 2015-2030, has adopted the concept of multi-hazard approaches within the framework of the Sustainable Development Goals and Agenda 2030 to reduce disaster risk.

In this context, the identification and characterization of multi-hazard regions (i.e., areas affected by multiple, interrelated hazards) and the delineation of potential multi-hazard scenarios can support forward-looking approaches to risk reduction (Kappes et al., 2012; Gill and Malamud, 2014; Liu et al., 2016; Ward et al., 2022). The effectiveness of hazard mitigation measures partially depends on the capacity to collectively consider, visualize, and evaluate the complex nature of hazards (Pourghasemi et al., 2020). As emphasized in the Sendai Framework for Disaster Risk Reduction (UNISDR, 2015) and its predecessor, the Hyogo Framework for Action (UNISDR, 2005), such approaches are critical for a range of stakeholders, including local developers, urban planners, and local and national authorities.

1.2 Initial and ongoing confusion (Multi-layer single hazards Vs Multi-hazard assessment)

Note that this part of the following content has been already posted as a blog post in the EGU Natural Hazard division by the author (Soltanpour, 2024).

Initially, due to the unclear and varied definition of multi-hazards, studies mainly adopted approaches based on their own perceptions of this term, ranging from an "all-hazards-at-a-place approach" (Hewitt and Burton, 1971) to "more-than-one-hazard" approaches (Kappes et al., 2012). Kappes et al. (2012) provided an early definition of "multi-hazard" within the framework of risk reduction as "the totality of relevant hazards in a defined area". This concept aimed to cover all potential environmental or man-made threats that could lead to harm, damage, or degradation. However, the criteria for defining the relevance of a hazard have been subject to debate, ranging from the introduction of a damage-based cut-off point for relevance by Hewitt and Burton (1971), to the European Commission (2011) proposing criteria based on the scale and impact of a hazard.

As mentioned in **Section 1.1**, traditional approaches to hazard management treat each hazard individually as a single event, often confusing decision-makers, particularly in critical circumstances. Looking at hazards as independent phenomena and not considering the interrelationships among them make us vulnerable to other ignored hazards. This approach also can be misleading in setting priorities and underestimate the overall risk. This is because the interaction of different hazards can result in an impact greater than the sum of their individual effects. The occurrence of each hazard in a sequence and the amount of time in between two hazards can substantially change the vulnerability to the next hazard (Gill and Malamud 2016; Marzocchi et al., 2012; Kappes et al., 2012; Terzi et al., 2019).

Subsequently, hazard mapping needs to evolve to address the interrelationships between hazards observed in the real world. Multi-hazard/ or hazard susceptibility maps are used to communicate complex scientific data with stakeholders. These maps serve as a communication tool to support and aid decision-makers by providing valuable information regarding preparedness mitigation measures and informing urban planning and public policy (Komendantova et al., 2014). In addition to single hazard approaches (where hazards are considered in silo), since the definition of multi hazards has been ambiguous, many researchers have produced multi-hazard maps by layering spatially relevant hazards in a given area (Tarvainen et al., 2006; Pia o et al., 2022). Hazards are assessed and mapped separately, then layered, and so can be thought of as '**a multi-layer single-hazard approach**' (Gill and Malamud 2014). Such approaches align with the 'all-hazards-at-a-place' approach set out by Hewitt and Burton (1971), but by not taking into account potential interrelationships could still cause challenges in practice. Considering interrelatinoships between hazards is vital since as quoted by (Gill and Malamud, 2016; Tilloy et al., 2019) "the interaction of different hazards can lead to an impact that is greater than the sum of the single hazard effects". In this regard, a more complete multi-hazard assessment/mapping has been encouraged, where the interrelationships of hazards (triggering, amplifying, compound and consecutive events) (discussed in **Chapter 2**) are also considered. This is now reflected in the definition of 'multi-hazard' adopted by UNISDR, which suggests that *"Multi-hazard means (1) the selection of multiple major hazards that the country faces, and (2) the specific contexts where hazardous events may occur simultaneously, cascadingly or cumulatively over time, and taking into account the potential interrelated effects."* (UNISDR, 2017).

1.3 Multi-hazard (susceptibility) mapping in Disaster Risk Management Cycle (DRM)

According to the UNISDR (2017), Disaster Risk Management (DRM) refers to "*the application of disaster risk reduction policies and strategies to prevent new disaster risk, reduce existing disaster risk and manage residual risk, contributing to the strengthening of resilience and reduction of disaster losses*." The DRM cycle is a continuous process that helps communities prepare for, respond to, and recover from disasters. The DRM cycle typically consists of four key phases (**Fig. 1.2**):

Phase 1: *Mitigation* helps identify risk-prone areas and inform risk reduction strategies.

Phase 2: *Preparedness* guides the development of emergency plans and public awareness.

Phase 3: *During response*, it assists in efficient resource allocation and emergency operations.

Phase 4: In *recovery*, it supports resilient rebuilding and future risk reduction efforts.

(Natural) hazard/susceptibility mapping, as defined by the United Nations Disaster Relief Coordinator (UNDRO), is a process that determines the geographic areas at risk from various natural hazards (e.g., earthquakes, floods and landslides) (UNDRO, 1980). Hazard susceptibility mapping not only identifies these hazards but also visually represents the different degrees of potential threats to people, property, and economic activities. Hazard mapping is a critical tool in disaster management and in fact is the basis for formulating DRM policies. While preventing hazards is often unfeasible, mitigation and risk reduction are achievable and effective hazard mapping can importantly reduce their impact (Rahmati et al., 2016; Pandey et al., 2018; Pourghasemi et al., 2019)

de Ruiter et al. (2020) highlight the critical need to integrate multi-hazard approaches in Disaster Risk Management (DRM). The DRM cycle for multi-hazards and consecutive disasters is shown in **Figure 1.2**. Multi-hazard mapping is positioned as an integral component across all phases of the DRM cycle. For example, in *the planning phase*, when identifying suitable areas for temporary shelters, it is crucial to ensure that these areas are not at increased risk of a cascade of multi-hazards.

Similarly, when developing Disaster Risk Reduction (DRR) policies, it is important to consider not only a single prevailing hazard but also the potential for cascading hazards. Multi-hazard mapping in the context of the DRR cycle is a crucial element for addressing the complexities and interconnectivities of different disaster drivers (Ruiter et al., 2020). Furthermore, van ,mjjjWesten and Greiving. (2017) assert that multi-hazard mapping is crucial for a comprehensive understanding of risk landscape in a region. It helps in better preparing for and managing the risks associated with interconnected and cascading natural hazards, thereby enhancing resilience and safety. This approach is especially important in regions prone to multiple natural hazards, where the impact of one hazard can significantly alter the risk landscape for others.

At the society level, hazard/risk assessments and mapping inform policies across various sectors, helping decision-makers and stakeholders to understand hazards/risks better and plan effective preventative and preparedness strategies (European Comission, 2011). Producing a hazard/risk map fosters a shared understanding among public authorities, businesses, NGOs, and the public about community risks. This shared understanding is vital for prioritizing prevention and mitigation measures and encouraging a broad discussion on these topics (European Comission, 2011). The study done by Komendantova et al. (2014) underscores the significance of multi-hazard maps in spatial planning and disaster prevention, particularly in European contexts. In countries like Norway, these maps are instrumental in restricting developments in areas prone to specific risks by incorporating probabilities of occurrence. The European Flood Directive

further exemplifies this approach by requiring the development of hazard and risk maps for flood-prone areas and formulating Flood Risk Management plans to mitigate flood impacts (Komendantova et al., 2014).

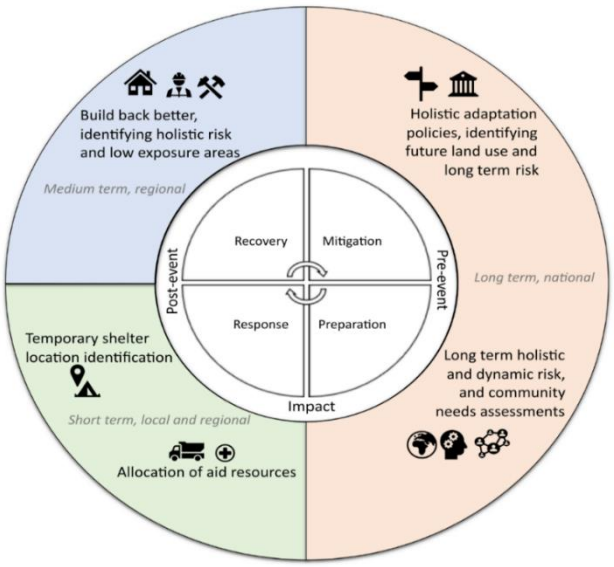


Figure 1.2 Disaster risk management cycle for consecutive disasters. When accounting for the risk of consecutive disasters, the dynamics of the different risk components should become part of all phases of the DRM cycle (de Ruiter et al., 2020).

Multi-hazard maps are dynamic tools that potentially inform and guide various aspects of the DRM cycle, from initial identification to ongoing monitoring and adaptation. Integrating these maps into the cycle enhances the comprehensiveness and effectiveness of risk management efforts.

1.4 Motivations

1.4.1 Multi-hazard forming areas (a globally identified gap)

Understanding and identification of the multi-hazard landscape of a region is crucial as it provides us with a more comprehensive understanding of the overall risks involved for sustainable hazard and risk management (Kappes et al., 2012). This knowledge plays an important role in informing management priorities and guiding decision-making processes. By being aware of the different hazards, we can ensure that the actions taken to reduce vulnerability to one specific hazard do not inadvertently increase vulnerability to others (de Ruiter et al., 2021). Taking multi-hazard interrelationships into account allows for a more effective, sustainable and holistic approach to disaster risk reduction and management (Gill et al., 2021). As emphasized in the Sendai Framework for Disaster Risk Reduction (UNISDR, 2015), multi-hazard approaches are critical for a range of stakeholders including local developers, urban planners, and local and national authorities. The aim of multi-hazard assessments and considerations are to entrench multi-hazard natural risk management in policy and practice, supporting decision makers to consider risk management through the multi-hazard lens (Gill et al., 2020). Planners and experts can use multi-hazard maps to make informed decisions about land use policies.

Across the globe, there are several types of geographical regions affected by multiple natural hazards, which can interrelate and therefore have the intrinsic potential to form a network of hazard interrelationships (**Fig. 1.3)**. Mountainous, coastal and volcanic regions are the commonly recognised multi-hazard-forming zones, where hazards interact with each other in different forms such as simultaneously, cascadingly or cumulatively (Kappes, 2010; Van Westen et al., 2014; Gill and Malamud, 2016; de Ruiter et al., 2020; Gallina et al., 2020).

In mountainous regions, an earthquake or heavy rainfall can trigger flooding, landslides or debris flows, which then by blocking a river or breaching a river's dam can generate further hazards, such as flooding (Kappes et al., 2010; van den Bout et al., 2022). For example, the 2015 Gorkha earthquake event in the Himalayan-nation of Nepal triggered more than 24,000 landslides adding to a death toll of nearly 9000 people (Collins et al., 2015). In the context of mountainous areas, different studies have been carried out to map the typical multi-hazard scenarios (Yousefi et al., 2020; Tsereteli et al., 2018; Pourghasemi et al., 2019; Terzi et al., 2019).

In coastal environments, a hurricane can cause strong wind and storm surge, which triggers flooding, mudslides and landslides (López-Saavedra and Martí, 2023; Appelquist and Balstrøm, 2015). Recently on September 14, 2022, what happened in Puerto Rico demonstrates a perfect example of a chain of multi hazards in coastal environments when the category one hurricane Fiona unleased more than 30 inches of rain resulting in widespread flooding. Consequently, this flooding triggered numerous mudslides and landslides (weather.gov/sju/Fiona, 2022). Throughout the literature, there have been attempts to characterise and map the relevant coastal hazards (see Gallina et al., 2020; Marignani et al., 2017). Similarly, in volcanic regions, during a volcanic unrest, it is expected that a magmatic eruption causes a cascade of hazards including tephra fall, lava flows, lahars, gas emissions, landslides, floods, earthquakes, and in some cases tsunamis like the eruption happened in Eyjafjallajökull eruption (Iceland) (López-Saavedra and Martí, 2023). Several multi-hazard scenario identification and mapping in volcanic contexts can be found in the literature (Neri et al., 2013; López-Saavedra and Martí, 2023).

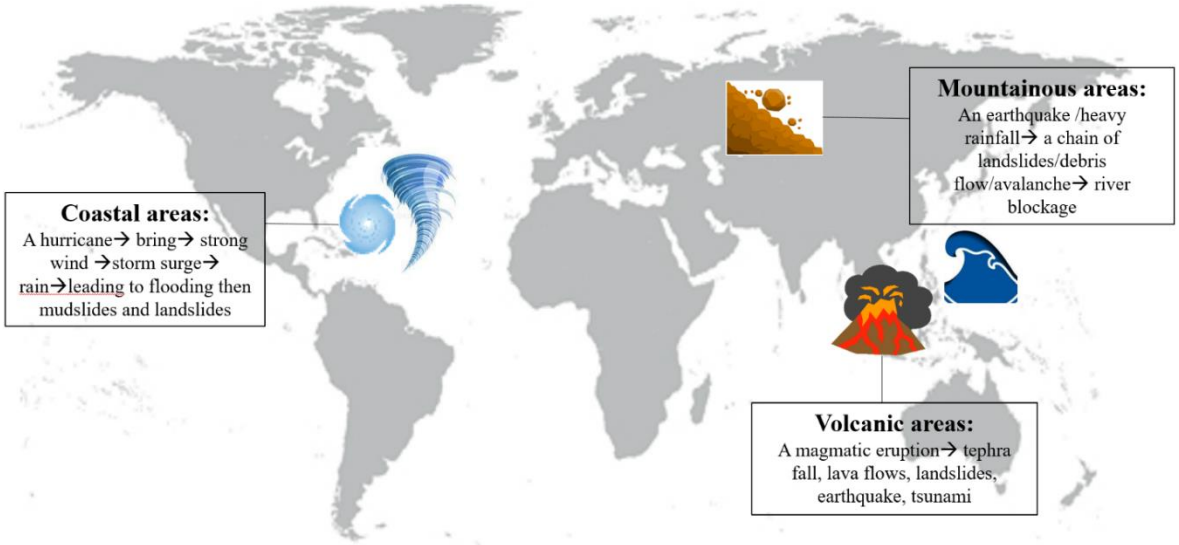


Figure 1.3 Global multi-hazard hotspots. Identification the multi-hazard forming areas, considering hazard interactions are crucial for the disaster risk reduction plans. Note: areas marked are illustrative examples and do not represent the full extent of multi-hazard zones globally. Source: the author

Another multi-hazard forming landscape, often overlooked in the multi-hazard literature, is karst terrain. Karst terrain is created from the dissolution of soluble rocks (carbonates and/or evaporites) and is characterized by distinct landforms such as springs, caves and sinkholes (Waltham et al., 2005). These landscapes, with their distinct geology, geomorphology, hydrogeology, and other environmental characteristics, are prone to multi-hazard events (Gill and Malamud, 2014; Parise et al., 2018). While previous studies recognize the potential for heavy rainfall events and flooding to trigger sinkholes in karst environments and emphasize the need to include karst landscapes in multi-hazard discussions (Lei et al., 2013; Xiao et al., 2016; Xiao et al., 2018; Kwak et al., 2020), sinkhole mapping in karst areas has predominantly focused on single hazards (Subedi et al., 2019; Kim et al., 2022). Therefore, recognition of karst areas as multi-hazard forming environment is a step towards solving the problem. Identifying the potential multi-hazard scenarios and their interconnections can be a wakeup call for local developers, urban planners and local authorities. Additionally, people who live or have properties in karst terrains should be aware of the risks (Geoengineers, 2016).

In this context, the core aim of this research is to contribute to the perception of karst terrains as multi-hazard-forming areas by developing and evaluating an integrated multi-hazard susceptibility map of the two hazards (flood and sinkholes) as a common multi-hazard scenario within the karst context. The aims are outlined further in the following paragraphs.

1.4.2 Institutional gaps in natural hazard management in France

This study aims to develop the above-mentioned multi-hazard susceptibility map for the Val d'Orléans region in France. In France, in particular, the Val d'Orléans region as a case study for this thesis, the Plan de Prévention des Risques d'Inondation (PPRi) is a document that outline flood risk prevention measures and guidelines. Its primary goal is to minimize the impact of flooding on human lives, property, and the environment by regulating urban development in flood-prone areas. The Val d'Orléans has two distinct PPRi due to the specific urban characteristics within its regions. The first PPRi, known as the "Val Amont", which covers the upstream areas of the Val d'Orléans (applies to 16 communes). This document primarily focuses on managing flood risks in more rural or less densely populated areas (DDT du Loiret, 2015a). The second PPRi, titled "Agglomération Orléanaise", addresses the urbanized areas within the heart of urbanization in the Loiret (applies to 13 communes) (DDT du Loiret, 2015b). Given the higher population density and more significant infrastructure in this area, this PPRi includes detailed guidelines for urban planning, construction regulations, and emergency response measures specifically tailored to urban settings.

The PPRi is a regulatory document that municipalities in France must follow. It restricts construction and land use in designated flood hazard zones to reduce vulnerability. However, while the PPRi addresses flood hazards, it does not consider other related natural hazards, particularly in karst areas such as Val d'Orléans. Val d'Orléans is considered as a highly karstified setting, which has been long built on carbonate bedrock (it is discussed further in **Chapter 3**). Flood events in karst areas, which are prone to sinkhole collapses, can result in substantial hazards and consequences. Therefore, this fragmented approach in hazard and risk assessment may cause a limitation on sustainable urban planning in the region.

PPRi and other specific PPRs (Plans de Prévention des Risques) for landslides, earthquakes, wild fire, and coastal erosion have traditionally focused on individual hazards. Each plan targets a specific type of risk and sets regulations accordingly. Subsequently, all the documents relevant to awareness among citizens on the natural hazards they are exposed at and threatened by (e.g., Municipal Information Document on Major Risks (DICRIM) are based on single-hazard approaches (it is discussed more in **Chapter 5**). It is also highlighted in the work of Thiery et al. (2020), which referred to a major gap in the current French regulatory framework. While their work focused on current methods for landslide hazard assessment, it also showed that the advanced approaches are not widely implemented in regulatory documents. It is emphasized that there is a need for integrating new tools and data to improve hazard assessments and make them more comprehensive.

By developing a multi-hazard/ susceptibility map that specifically considers the interrelationship of both flood and sinkhole zones, this research also aims to broaden the scope of the PPRi plan (DDT du Loiret, 2015a; b) and the DICRIM document (DICRIM, 2021) in Val d'Orléans to include a multi-hazard susceptibility assessment that addresses the interconnected nature of multiple hazards in the study area. This integration allows decision-makers, urban planners, and emergency responders to have a more holistic understanding of the potential risks in different areas, thereby enabling them to develop targeted mitigation and response measures.

To address this issue effectively, this research also aims to evaluate the resultant map with expert stakeholders in Val d'Orléans. Prior to that, it will assess the level of awareness among these stakeholders. This part of the research helps to create a roadmap for future improvement of multi-hazard susceptibility maps.

1.5 Research aims and objectives

The core aims of this research are to (i) improve multi-hazard risk mamangement in the context of karst terrains; and (ii) evaluate the potential for machine-learning (the Maximum Entropy, MaxEnt, model) to support this ambition.

To achieve these aims this research has five primary objectives:

1) Critically evaluate the extent to which karst terrains are considered multi-hazard environments, and the opportunities to advance multi-hazard understanding in this context.

2) Review existing approaches used to model both multi-hazard relationships and sinkholes as a single hazard.
3) Examine the effectiveness of Maximum Entropy (MaxEnt) as a machine-learning model in forecasting multi-hazard distributions in karst terrains.

    a) Develop a multi-hazard susceptibility map that identifies the spatial interrelationships of flooding and sinkhole occurrences in the karst environment of the Orléans region.
    b) Analyse the importance of (geo)-environmental factors for flood-triggered sinkholes.
    c) Compare MaxEnt outputs for multi-hazard contexts with current single hazard (i.e., flood) maps for the region.

4) Document local expert stakeholder understanding of the multi-hazard context in the Orléans region and their perspectives on MaxEnt generated multi-hazard maps for flood-triggered sinkholes.
5) Produce a roadmap to guide and strengthen future work on the karst terrain of the Orléans region as multi-hazard environments.

Through these objectives, I hope to improve multi-hazard risk management in the context of karst terrains by **(a)** contributing to the perception of karst terrains as multi-hazard-forming areas, **(b)** advocating for a shift in multi-hazard assessment methodologies, moving from a multi-layer single hazard approach to a multi-hazard approach, **(c)** testing methodologies to characterisise specific multi-hazard interrelationships associated with karst terrains (i.e., evaluate the potential for the MaxEnt model to be used in multi-hazard mapping), and **(d)** supporting decision-makers and authorities to broaden the scope of the flood hazard documents (PPRi) followed by other Departmental Plan for Major Risks (DDRM) and (DICRIM) by incorporating these complex hazard interrelationships.

Collectively, the research and analysis in this thesis is suggested to make the following major contributions to scientific understanding in the context of the multi-hazard risk management:

- Advances understanding of and evidences the importance of multi-hazard analysis in the characterisation of karst terrains.

- Expanded the application of the MaxEnt model, traditionally used in ecological contexts, to a novel multi-hazard framework, specifically modeling hazards, namely flood-triggered sinkholes.

- In the geographic context of the Val d'Orléans, this research contributes to improving local preparedness and decision-making by addressing a significant gap in hazard mapping for the region.

1.6 Research outline

This thesis is divided into **6 Chapters**.

Chapter 1, as an introductory chapter, summarizes the purpose and contents of the thesis and briefly discusses the research background, motivation, aims and objectives and research outline.

Chapter 2 presents the introduction to the multi-hazard concept and modelling followed by a critical literature review of the studies on sinkholes constructing a literature database and identifying the current gap in sinkhole hazard assessments. I use my literature database to demonstrate the current state of this thesis topic. This critical review identifies **241 papers** published in scientific international journals (the attempts were made to be representative not inclusive), which allowed me to classify them into *seven thematic subjects* for the sinkhole hazard studies. It concludes with a review of the current sinkhole studies worldwide and the previous work done in the study area justifying the need for this research.

Chapter 3 presents a detailed research design and methodology focused on applying the Maximum Entropy (MaxEnt) algorithm, alongside introducing Val d'Orléans as the case study area. It consists of four sections; *Section 3.2* delves into the foundational principles of the MaxEnt algorithm, exploring its historical roots and applications in natural hazard science, emphasizing its effectiveness in modeling natural phenomena using presence-background data. *Section 3.3* describes the unique geographical, geological, and hydrogeological characteristics of Val d'Orléans, highlighting its susceptibility to hazards such as sinkholes and floods and the need for a regional multi-hazard mapping. *Section 3.4* outlines the data collection processes, detailing the implementation of MaxEnt, selecting (geo)-environmental variables, processing data, and configuring the software for optimal performance. *Section 3.5* explains the preparation of data for modeling flood-triggered sinkhole occurrences, including data formatting, parameter tuning, and data partitioning strategies, emphasizing the methodological rigor required for accurate hazard modeling in our research.

Chapter 4 presents the results of Maximum Entropy (MaxEnt) modeling. The sequence of results is structured to build understanding of the model's forecasts progressively. We begin in *Section 4.2.1* by evaluating the model's overall performance through the Receiver Operating Characteristic (ROC) curve, providing a baseline measure of accuracy, and explaining the ROC curve to discuss our findings. *Section 4.2.2* presents the spatial distribution of flood-triggered sinkholes produced by MaxEnt. *Section 4.2.3* delves into the influence of both individual and combined (geo)-environmental variables on the model's forecast, presented through model-specific and marginal response curves. *Section 4.2.4* assesses the robustness and reliability of these predictions with a Jackknife test, highlighting the contribution of each variable. *Section 4.3* spatially compares the official flood hazard map with our multi-hazard susceptibility map, aiming to identify areas of overlap, which is crucial for prioritizing resource allocation, informing policy and planning, and enhancing emergency preparedness. The comparison details are then presented in Appendix 4.A to maintain a streamlined narrative.

Chapter 5 focuses on the evaluation process aimed at examining expert stakeholder awareness and perceptions of multi-hazard events, evaluating the produced multi-hazard susceptibility map presented in *Chapter 4*, and understanding participants' perspectives on the ownership and pervasive nature of flooding

and sinkhole hazards in the area. Additionally, it explores participants' views on establishing a dedicated DICRIM document for multi-hazard interactions. The results of this study are crucial for creating an improvement roadmap, integrating qualitative evidence and expert feedback, with an emphasis on multi-hazard mapping, awareness, and fostering multi-hazard thinking.

Chapter 6 is a concluding chapter. Within this last chapter, I provide a summary of the thesis, encapsulating the main findings and discussions from the earlier chapters. Elements of this summary have been previously mentioned in the concluding sections of Chapters 2, 3, 4, and 5. However, this chapter integrates those summaries to present a unified overview of the research. Additionally, we address the limitations and uncertainties in the modelling process. The chapter concludes with recommendations for future research directions.

In this thesis, *Appendices* are organized on a **chapter-based basis**. By linking each appendix with the corresponding chapter content, readers can easily identify and access the materials relevant to their specific interests.

Chapter 2. Literature review (State of the art)

2.1 Introduction

This chapter consists of **two main parts**. The first part provides more general background including **i)** the multi-hazards concept and its historical development (**Section 2.1.1**), **ii)** the importance of understanding the interrelationships between multiple hazards, discussing the existing different interrelationships among natural hazards (**Section 2.1.2**), **iii)** a summary of qualitative and quantitative approaches for multi-hazard assessment (**Section 2.2 and its Sub-sections**).

In the second part of this chapter, I first outline the characteristics of karst terrains as multi-hazard landscapes (**Section 2.3**). In **Section 2.4**, we provide our **critical review of sinkhole-related studies**, presenting relevant data, methodologies, results, and limitations to demonstrate understanding of the thesis topic. **Section 2.5** focuses on key prior research specifically mapping karst hazards in the Val d'Orléans region. The chapter concludes with a discussion and summary of key findings in **Section 2.6**.

2.1.1 Exploring the origins and definitions of multi-hazards

While the importance of the multi-hazard concept has been established in the literature (e.g., Kappes et al., 2010; 2012; Gill and Malamud, 2016; Wang et al., 2020; Dunant et al., 2021; Choi et al., 2021), the discipline is still in an early stage of development. As such terminology and definitions are often used interchangeably (e.g., see Kappes et al., 2010; 2012; Gill and Malamud, 2016; Ciurean et al., 2018, Wang et al., 2020; Dunant et al., 2021; Choi et al., 2021, Gill et al., 2022).

Tilloy et al. (2019) give a helpful perspective on the extent to which the multi-hazard concept is new. In their study, Tilloy et al. (2019) created a multi-hazard literature database consisting of 146 references (mostly from peer-reviewed scientific journals) for a 38-year period (1980-2018). They chose nine journals, which had equal or more than three articles in the multi-hazard topic. The violin plot depicted by Tilloy et al. (2019) (**Fig. 2.1**), demonstrate that attention to the multi-hazard subject began around 1980 and that for most journals, there is an increase in the number of articles published over time. However, when discussing multi-hazards, it is essential to recognize that the encouragement of a multi-hazard approach began before 1980. For example, Hewitt and Burton (1971) highlighted the importance of systematic cross-hazard approaches to comprehensively understand the hazard landscape of a region.

Nonetheless, since the late 1990s, there has been a noticeable increase in the number of published articles on this topic, indicating a growing interest over the past three decades. This surge in publications largely stems from an early policy reference to the multi-hazard concept, in 1992. During the United Nations' Agenda 21 conference on sustainable development in Rio de Janeiro, Brazil, the concept was advocated under the agenda item titled "*complete multi-hazard research*" (UNEP 1992, paragraph 7.61). After that, the Johannesburg Plan proposed a complete multi-hazard approach for disaster management and risk reduction (UN, 2002). Within the Hyogo Framework for Action (2005), the concepts introduced in the

Johannesburg plan were elaborated upon, specifically emphasizing the importance of risk reduction; "*An integrated, multi-hazard approach to disaster risk reduction should be factored into policies, planning and programming related to sustainable development, relief, rehabilitation, and recovery activities in post-disaster and post-conflict situations in disaster-prone countries*" (UNISDR, 2005, p.4).

In a more recent risk management agenda (2015–2030), the Sendai Framework for Disaster Risk Reduction emphasises the need for multi-hazard approaches (UNISDR, 2015). The UNISDR went on to define the term 'multi-hazard' as "*the selection of multiple major hazards that the country faces, and the specific contexts where hazardous events may occur simultaneously, cascadingly, or cumulatively over time, and taking into account the potential interrelated effects*" (UNISDR, 2017). Collectively, these calls and intergovernmental frameworks contributed to a growing interest in multi-hazard studies in academia and practice.

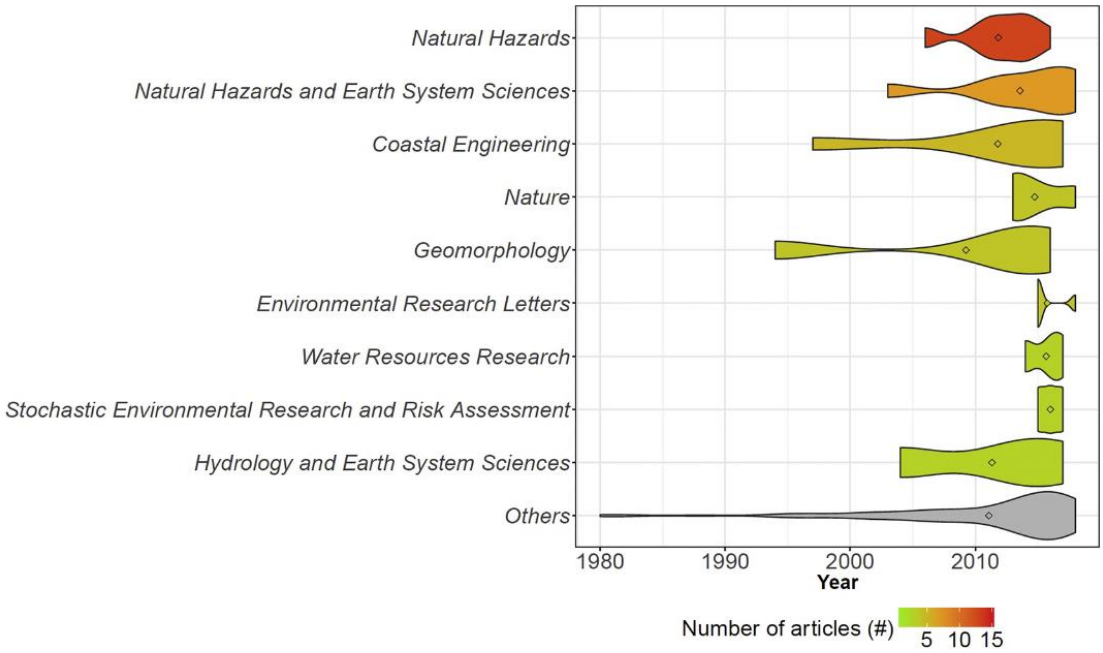


Figure 2.1 The database on multi-hazard literature consisted of 146 references from 83 sources for a 38-year period (1980–2018), (Tilloy et al., 2019).

Given UNISDR (2017) was the first time that the United Nations system had defined the term multi-hazard, it is perhaps understandable that in the academic literature (from both before and after 2017) terms are used interchangeably and there remains a lack of consensus on key definitions. This absence of standardized terminology has been identified as a barrier to effective communication and coordination among various sectors involved in disaster risk management (e.g., see Kappes et al., 2012). Ward et al. (2022) advocates for a strategic research agenda aimed at establishing common definitions to address communication gaps and foster more effective multi-hazard risk management strategies. To support this ambition, Gill et al. (2022) published a handbook of Multi-Hazard, Multi-Risk Definitions and Concepts. In line with the

recommendation of Gill et al. (2022), we adopt the UNISDR (2017) definition of the term ‘multi-hazard’ in this thesis. We summarise the main terminologies in the disaster risk reduction domain in **Table 2.1**. Moreover, the recent literature on multi-hazards has defined and classified distinct types of hazard interrelationship, discussed in **Section 2.1.2**.

Table 2.1 Key definitions

Term	Definition	Source
Natural hazard	"Natural hazards are predominantly associated with natural processes and phenomena."	UNISDR. (2017)
Multi-hazard	"The selection of multiple major hazards that the country faces, and the specific contexts where hazardous events may occur simultaneously, cascadingly, or cumulatively over time, and taking into account the potential interrelated effects."	UNISDR. (2017)
(Cascading) hazard	"Cascading hazard processes refer to an initial hazard followed by a chain of interrelated hazards (e.g., earthquake triggering landslide, landslide triggering flooding, flooding triggering further landslides)".	UNISDR (2019) adopted by Gill et al. (2022)
Hazard interrelationships	"The mode in which one hazard affects another hazard".	Gill et al. (2022)
(Multi)-hazard-forming environment	"Geophysical environmental factors (e.g., climate, topography, geology) are the determining factors for natural hazards, and the geophysical environment which consists of these factors can be defined as the "hazard-forming environment". A geographic area where multiple types of geophysical environmental factors (e.g., climate, topography, geology) have the potential to interact and occur simultaneously, cascadingly, or cumulatively over time.	Adapted from UNISDR (2017)
Multi-layer single hazards	"More than one hazards are considered, but not the interrelationships between these (i.e., they are treated as discrete, independent)".	Gill and Malamud (2014)
Susceptibility assessment	"Susceptibility assessment denotes the relative probability of a hazard event without reference to any specific time interval.	(Gutiérrez et al., 2008).
Multi-hazard susceptibility	Used in this research when referring to the spatial analysis of multi-hazards	The author

2.1.2 Multi-hazard interrelationships

Understanding the interrelations between hazards is essential. It helps us recognise how different hazards can affect each other, giving us a clearer picture of the series of dangers a region might face (Gill and Malamud, 2014). In simpler terms, it helps us understand the overall picture of risk. Neglecting to understand these interrelations, and act on this understanding, can result in an increased overall impact when two or more hazards occur simultaneously (Kappes et al., 2010; Marzocchi et al., 2012; Mignan et al., 2014). This underscores the importance of considering hazard interrelationships to mitigate the potential consequences.

The term 'multi-hazards' includes distinct types of connections between hazards (see **Section 2.1.1**). Following Hochrainer-Stigler et al. (2023), we use 'interrelationships' 'as the umbrella term for these connections. The multi-hazard literature discusses each distinct type of interrelationship (e.g., Gill and Malamud, 2014, 2016; Xua et al., 2014; Mignan et al., 2016; Tilloy et al., 2019; de Ruiter et al., 2020; Wang et al., 2021). Throughout the literature, different terms are used for these distint hazard interrelationships. Several recent classifications of hazard interrelationships are summarized in **Table 2.2**. Typically, different terms refer to similar concepts.

In general, hazard interrelationship types include triggering, amplifying (increased probability), and 'compound' relationships. These interrelationships illustrate the complex interplay between different processes and their effects on various environmental or structural outcomes. Their general definitions are provided below:

a) **Triggering interrelationships:** One form of causal relationship is known as "triggering relationships" where the occurrence of a primary event can lead to the occurrence of secondary events. For example, an earthquake as a primary natural hazard triggers several landslides or tsunami as a second natural hazard (Gill and Malamud, 2014; Tilloy et al., 2019).

b) **Amplifying interrelationships:** The occurrence of one hazard can increase the likelihood and/or magnitude of additional hazards in the future (e.g., forest fires can amplify the triggering of debris flows during heavy rain) Ciurean et al., (2018). Changes in agricultural practices, vegetation removal, surface and subsurface construction, quarrying, and so forth can trigger natural hazards or amplify multi-hazard interrelationships (Gill et al., 2021).

c) **Compound hazard interrelationships:** Compound hazards occur when two or more hazards coincide in space and/or time, leading to impacts greater than the sum of their individual effects. This can happen in two main ways: i) A primary then secondary hazards, and ii) coinciding independent hazards (simultaneously or consecutively). For example, extreme temperatures combined with drought can severely strain water resources (Ciurean et al., 2018).

Table 2.2 Different classification of hazard relationships by various authors.

Published papers	Category of hazard relationships
Kappes et al., 2010	• One process **triggers** the next (producing cascades or domino effects). • **Changes to the disposition** (e.g. frequency, magnitude alterations) of one hazard is altered by another.
Duncan, 2014	• **Causation** (hazard generates a secondary event, which may occur immediately or shortly after the primary hazard, including cascading hazards). • **Association** (hazard increases the probability of secondary event, but it is difficult to quantify this link and therefore confirm causation). • **Amplification or alleviation** (hazard exacerbates or reduce future hazards). • **Coincidence** (hazard occurs in the same place simultaneously, or closely timed, resulting in compounded effects or secondary hazards).
Gill and Malamud, 2014; 2016	• Interactions where a hazard is **triggered**. • Interactions where the probability of a hazard in **increased**. • Interactions where the probability of a hazard is **decreased**. • Events involving the spatial and **temporal coincidence** of natural hazards.
Ciurean et al., 2018	• **Triggering:** One hazard relationship where one hazard causes another to occur. This can result in chains or networks of cascades of hazards. • **Amplification effect:** The occurrence of one hazard can change the likelihood and/or magnitude of additional hazards in the future. • **Compound hazards:** a primary hazard triggering multiple secondary hazards within a given timeframe or the coincidence of two independent hazards.
Tilloy et al.,2019	• Independence • Triggering • Change conditions • Compound hazard • Mutual exclusion (negative dependence)
de Ruiter et al., 2020	• **Consecutive hazards:** two or more disasters that occur in succession, and whose direct *impacts* overlap spatially before the area has recovered from the initial hazard event. This can be split into two types: • **Dependent hazards:** it includes triggering and cascading disasters (e.g. landslides triggered by a flood, or fires caused in the aftermath of an earthquake). • **Independent hazards:** The hazards, which are not triggered by one another and do not influence the probability occurrence of another hazard, but their impacts spatially and/or temporally overlap (e.g. a tropical cyclone followed by an earthquake).
Wang et al., 2020	• Mutually amplified hazards • Mutually exclusive hazards • Non-influential hazards

2.2 Approaches to multi-hazard assessment

2.2.1 Summary of qualitative and quantitative approaches for multi-hazard assessment

Multi-hazard assessment puts different types of hazards into a single system for a joint evaluation. Typically, current approaches are categorized as qualitative, semi-quantitative, or quantitative (Gill and Malamud, 2014; Zschau, 2017; Hochrainer-Stigler et al., 2023). The choice of each approach depends on the purpose of the research, hazard types, the geographical context, and data availability (Wang et al., 2020; Gill et al., 2021). Ciurean et al. (2018) provided an overview of various methodologies in multi-hazard assessments including:

i) **Narrative descriptions** (qualitative): e.g., understanding hazard landscapes and scenarios, reporting hazard events. These are valuable sources of evidence about past events, helping in scenario planning. A prominent example of this category is provided by de Ruiter et al. (2020).

ii) **Hazard wheels** (qualitative): These have been mainly used in coastal hazard management and are more suitable for the Global South where data and institutional capacity may be limited. They classify coastlines, assess their hazard profiles, and determine management options (Appelquist and Halsnæs, 2015).

iii) **Hazard matrices** (qualitative/semi-quantitative): Qualitative or semi-quantitative tools used to examine relationships between various hazards by identifying spatially relevant hazards and determining their interrelations. A prominent example is provided by Gill and Malamud (2014), who produced a 21 × 21 matrix with primary hazards on the vertical axis and secondary hazards on the horizontal axis. As Ciurean et al. (2018) assert, hazard matrices are effective for quickly identifying secondary hazards.

iv) **Network diagrams** (Qualitative/Semi-Quantitative): Network diagrams use nodes (representing hazards or processes) and connectors (depicting triggering or catalyzing relationships) to map out potential multi-hazard interactions. Van Westen et al. (2014) applied this method to illustrate hazards and their relationships in an alpine mountainous setting. These diagrams are visually appealing and effective for documenting potential hazard relationships and identifying scenarios.

v) **Hazard maps:** Cartographic approaches are important tools to depict the spatial overlap of multiple hazards. However, challenges exist in mapping multiple hazards as outlined by Kappes et al. (2012), these include: i) comparability of hazards due to different process characteristics; ii) considering hazard interrelationships; iii) analysing physical vulnerability for multiple hazards. Approaches toward multi-hazard maps will be discussed more in following paragraphs.

vi) **System-based (Physical) modelling** (quantitative): These models represent real-world events based on their underlying physical processes. This is discussed in **Section 2.2.3**.

vii) **Probabilistic and statistical approaches**. They encompasse a wide range of methodologies from stochastic to empirical methods which are discussed in **Sections 2.2.2** and and **2.2.3**.

In recent literature on multi-hazard modelling, Tilloy et al. (2019) classified three approaches applied by different disciplines to quantify hazard interrelations for 14 hazards (**Geophysical**: (*Earthquakes, Landslides, Volcanic eruptions*); **Atmospheric**: (*Lightning, Extreme rainfall, Extreme wind, Extreme temperature, Hail, Tornado*); **Hydrological** (*Sea surge, Extreme waves, River Flood, Tsunami, Drought*). They identified three main quantitative approaches (**Stochastic**, **Empirical** and **Mechanistic**) (**Fig. 2.2**) for modelling the above-mentioned hazards. Note that these models have been used to quantify the interrelationships and not necessarily to spatially mapping interrelationships. These modelling methods are used to understand, predict/forecast, and simulate various types of interrelations, in the context of multi-hazard analysis. In the following sections, each approach will be discussed in turn. We then present a summary of this information in **Table 2.3**.

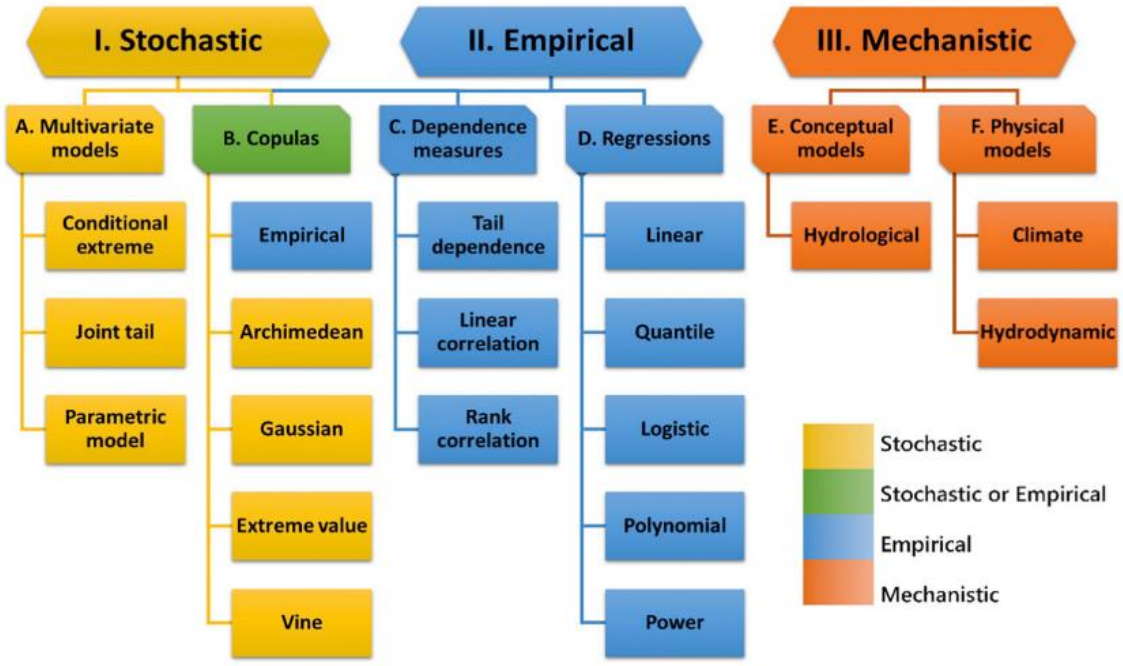


Figure 2.2 The diagram shows an overview of the main modelling methods for quantifying different types of interrelations (Tilloy et al., 2019).

2.2.2 Stochastic approaches

Stochastic refers to random processes or events that are subject to probabilistic outcomes. This approach has two families: *Multivariate* and *Copulas*. Stochastic models include all the methods that generate random data based on statistical patterns and estimate the joint probability of multiple hazards. (Tilloy et al., 2019). They can model the statistical relationship between extreme environmental variables, such as extremely *strong winds* or *heavy rainfall* in the context of coastal flooding. That is why they have been particularly used for climate-related and hydrological hazards and/or compound hazards _various hazards stem from the same primary event which are not necessarily a hazard_ (Gill and Malamud, 2016) in multi-hazard literature. One of the common stochastic models is Monte Carlo simulations, which involve simulating random variations in certain parameters to model and analyse complex systems or processes. Good examples of using this method are firstly to simulate the thickness of ash fall hazards in volcanic regions generating a large number of simulations where parameters such as eruption volume, column height, and wind conditions are varied randomly according to known or assumed probability distributions (Hurst and Smith, 2004).

Moreover, Dunant et al. (2021) present a novel multi-hazard risk assessment framework applied to Franz Josef township in New Zealand. The study stochastically quantifies the interactions and impacts of multiple hazards such as earthquakes, rainfall, landslides, landslide dams, and flooding on the township's infrastructure, including roads, stop banks, and housing. The methodological approach involves two primary phases: 1) network construction, 2) network simulation. In the network construction, relevant elements (e.g., earthquake, house, river, road) are represented as nodes, which are then interconnected based on potential impact footprint, spatial parameters, and expert opinion. This creates a disaster graph. In the network simulation phase, they used Monte Carlo simulation method to run different disaster scenarios (hundreds of thousands) to provide probabilistic outputs and a substantial dataset for analysis. By running a great number of scenarios, a large number of data were obtained on what could happen in different disaster situations. However, the model relies heavily on empirical relationships of individual hazards and the probability/intensity information of the initial hazards that trigger the cascading events (like earthquakes and rainfall). Additionally, Monte Carlo simulations are dependent on the assumptions made in the scenarios and the range of variables included. This randomness inherent in Monte Carlo methods means that while they can provide a broad range of possible outcomes, they may not capture all specific, less probable scenarios.

Stochastic models consider the statistical dependency between the occurrence and intensity of hazards. The purpose is either to model the distribution of each individual hazard or to model the joint probability between hazards. Bayesian networks that have been used in a probabilistic framework are part of the stochastic method. Depending on the quality of the statistical distribution, the result of this model can differ. Therefore, for low frequency events, this method can be challenging (Dunant et al., 2021). The complexity of stochastic models can make them difficult to understand and use correctly without specialized knowledge. This can be a barrier for stakeholders who are not experts in statistical or probabilistic methods (Goda and De Risi, 2023).

2.2.3 Empirical approaches

Empirical models, which are based on observed data and measurements, involve fitting statistical distributions directly to the data. In other words, empirical models are used to identify patterns and correlations based on historical data. Within the realm of empirical models, two main families are identified: *dependence measures* and *regressions* (Tilloy et al., 2019). Dependence measures are used to understand the strength and nature of the link between two or more variables. Common measures include linear correlation (such as Pearson correlation) and rank correlation (such as Spearman or Kendall correlation) (Ward et al., 2018).

Ward et al. (2018) investigated the dependence between high sea levels and high river discharges, particularly in global deltas and estuaries. It aimed to quantify potential causal mechanisms, including atmospheric conditions, and the importance of considering this dependence in flood risk management. It emphasized how this dependence can significantly increase flood hazards. The authors assessed this dependence by analysing sea level data and river discharge data. They use Kendall's rank correlation coefficient and copula models for their analysis, discovering significant dependence in various locations. This study highlights the need for flood risk assessments to consider these joint exceedance probabilities, especially in areas prone to both high sea levels and river discharges.

These measures are straightforward and widely applied to quantify the degree of association between hazards. However, traditional correlation measures may not be adequate when analysing the tail of distributions (the extremes). For extreme value analysis, different dependence measures are more appropriate. These include extremal dependence measures that aim to assess the correlation between variables specifically within the tails of their distributions. Such measures are crucial in understanding the relationship between extreme events, which are often the focus of hazard analysis (Tilloy et al., 2019).

Regressions are another basis of empirical modelling and have been extensively applied to quantify interrelations between natural hazards. As shown in the **Figure 2.2**, various forms of regression exist, including linear, power, logistic, and quantile regressions, which are selected based on the relationship being modelled and the nature of the dependent variable. Linear regressions are the most common form used for estimating relationships between natural hazards. However, in cases where the focus is on the extremes, quantile regression may provide a more accurate modelling approach. For example, quantile regression would be more applicable than linear regression for modelling the relationship between high levels of rainfall and subsequent landslides.

Machine learning, as an advanced empirical method, also falls within this category (it is more discussed in **Section 2.4.6d** with examples of common methods). A growing number of research studies on mapping multi-hazards have applied different machine-learning algorithoms. Recently, Yousefi et al. (2020) focused on using machine learning models for multi-hazard assessments in a mountainous region in Iran. Their aim was to produce risk maps for various mountainous natural hazards, including snow avalanches, landslides, wildfires, land subsidence, and floods. Three machine learning models – Generalized Linear Model (GLM), Support Vector Machine (SVM), and Functional Discriminant Analysis (FDA) – were used to construct risk maps for each hazard. This study combined these individual hazard maps to create an integrated multi-hazard risk map using a Boolean algorithm. These algorithms operate on the principles of Boolean logic, which involves using logical operators such as AND, OR, and NOT. In hazard mapping, for instance, a

Boolean algorithm is used to integrate separate hazard maps for different hazards. It does this by applying logical rules to determine where combined hazards may exist. While the methodology is robust in assessing individual hazards, a potential limitation is the lack of explicit attention to the interrelationships between different hazards. Hazards in a mountainous region are often interrelated; for example, heavy rains can trigger both floods and landslides. The study' approach, focusing on separate models for each hazard, may not fully capture these interdependencies. This could lead to an underestimation of the overall risk in areas where hazards are likely to interact and amplify each other's effects. Moreover, the flood hazard inventory is considered as points, which follows the river line. According to the geometry, a point represents a location in space and has no size. Describing flood hazards only as points could oversimplify the complex spatial nature of flooding, which often extends across areas and may vary in intensity. This assumption (considering floods as points) misleads the system by assuming only the river line is the area where the hazards happened in the past. Flood hazard as a continuous area cannot be treated as points unlike other highly local hazards like sinkholes, landslides, starting point of snow avalanches. That adversely influence the effective land use and disaster management planning.

Kariminejad et al. (2022) spatially mapped three hazards (collapsed pipes, gully heads, and landslides). They explored the use of various data mining techniques and UAV imagery to create and validate multi-hazard susceptibility maps for collapsed pipes, gully heads, and landslides in Golestan Province, NE Iran. It used seven classifiers: Boosted Regression Tree (BRT), Flexible Discriminant Analysis (FDA), Multivariate Adaptive Regression Spline (MARS), Mixture Discriminant Analysis (MDA), Random Forest (RF), Generalized Linear Model (GLM), and Support Vector Machine (SVM). The analysis shows that a major portion of the study area was low susceptible to these hazards, with a slight increase in susceptibility in 2021 compared to 2020. The study found that the Random Forest (RF) model was highly accurate in predicting collapsed pipes, particularly in 2020, with the highest accuracy among the models used. In 2021, the Multivariate Adaptive Regression Spline (MARS) model showed excellent accuracy for mapping gully heads, and the Boosted Regression Tree (BRT) model had the highest accuracy for predicting landslides. The study also noted that the Support Vector Machine (SVM) model had excellent accuracy for gully heads in 2020. Overall, the RF, MARS, and BRT models were highlighted as the most effective for different hazards across the two years, demonstrating their robustness and suitability for multi-hazard susceptibility mapping.

However, this is important to mention that in machine-learning, while ensembles models can reduce bias and improve accuracy, they come with increased complexity, potential for higher variance, and greater computational costs. This is because they try more hypothesis at the same time, which increases the variance, especially if the individual models are overfitting (Domingos, 2012). Therefore, while ensembles may appear more powerful, they are not necessarily better than non-ensemble methods in all cases.

Empirical models are particularly useful for representing cascading hazards as they can include both independent (primary hazard) and dependent (secondary hazard) variables. The main limitation of empirical models is that they cannot extrapolate beyond the range of the observed data, which can be a significant drawback when predicting events outside the historical record. Despite these limitations, empirical models are valuable tools for analyzing the relationships between different hazards and can provide insights that inform risk management and mitigation strategies.

2.2.4 Mechanistic approaches

Devi et al. (2015) describe mechanistic models as idealized mathematical frameworks that represent real-world events based on their underlying physical processes. The ability to develop advanced process-based models that can investigate multiple hazards and their interrelationships is steadily improving (Ciurean et al., 2018). These models, in particular, are used in water bodies, using equations from fluid mechanics (Tilloy et al., 2019). They are broadly classified into two families: conceptual models and physical models, and mainly involve geotechnical and engineering solutions (Lombardo et al., 2020).

Conceptual models, as extensively used in hydrology serve the purpose of describing various components of hydrological processes. These models depict interconnected reservoirs representing different aspects of river flow, such as infiltration, runoff, and snowpack. These models typically require substantial input data, often based on rainfall and temperature records, for parameter assessment through calibration (Tilloy et al., 2019).

Physical models aim to simulate the behaviour of various systems such as the atmosphere, ocean, climate, and hydrological processes. They typically use principles like fluid mechanics, heat transfer equations, or thermodynamic laws. In hydrology, these models represent water movement processes using finite difference equations. A study by Tinti et al. (2003) used a hydrodynamic model to assess the interrelation between a volcanic eruption and tsunamis in the Gulf of Naples. The study simulated a pyroclastic flow's impact from Vesuvius and its potential to generate a tsunami, finding that the maximum amplitude of the tsunami waves would be around 70 cm on the coastline. However, their primary drawback of these models lies in their substantial data requirements for modelling extreme hydrological events (Tinti et al., 2003). Despite their ability to address weaknesses of empirical or stochastic models by using parameters with physical meaning, physical models are often computationally intensive.

2.2.5 Tabulated summary of key literature

Table 2.3 An updated literature review of the three investigated approaches by Tilloy et al. (2019). Note that the empirical section also includes qualitative methods. The objective is to provide a representative overview rather than an exhaustive one.

Approaches (to quantify and/or mapping)	**Assessed multi-hazards**	**Common methods**	**Relevant studies**
Stochastic (considers the statistical dependency between the occurrence and intensity of hazards)	- Statistical relationship of Extreme environmental variables (*winds* or *heavy rainfall*) - Climate-related and hydrological hazards and/or compound hazards. - Extreme fire weather events followed by extreme rainfall (Touma et al., 2022) - Earthquakes and tsunamis (Goda and De Risi, 2023)	-Monte Carlo simulations. -Multivariate statistics	- **Hurst and Smith. (2004)**: Simulated ash fall thickness in volcanic regions generating a large number of simulations where parameters such as *eruption volume*, *column height*, and *wind conditions* are varied randomly according to known or assumed probability distributions. - **Forzieri et al. (2016)** aimed to evaluate the impact of climate change on multiple natural hazards in Europe, specifically examining how the frequency, intensity, and interactions of these hazards might change in the future. - **Dunant et al. (2021)** stochastically quantified the interactions and impacts of multiple hazards (**earthquakes**, **rainfall**, **landslides**, **landslide dams**, and **flooding**) on the township's infrastructure, including roads, stop banks, and housing. - **Touma et al., (2022)** assessed the changes in the frequency and likelihood of temporally compounding **extreme fire weather** and **extreme rainfall events**. It highlights a future with significantly increased post-fire hydrologic risks, such as debris flows, mudslides, and flash floods, due to the increased frequency of extreme rainfall following wildfires. - **Goda and De Risi, (2023)** used **stochastic event sets (**Monte Carlo simulations**)** to model various possible seismic events by characterising earthquake rupture processes by incorporating variability in fault geometry, slip distributions, and other seismic source parameters. They also simulate tsunami wave propagation and run-up, considering heterogeneous slip distributions on the earthquake fault. It involved generating several scenarios of earthquake occurrences and associated tsunamis to understand the range of potential impacts and their

			probabilities. However, the research involves several uncertainties due to lack of historical and empirical data.
Limitations: They are dependent on the assumptions made in the scenarios and the range of variables included. Depending on the quality of the statistical distribution, the result of this model can differ. Therefore, for low frequency events, this method can be challenging (Dunant et al., 2021). One of their key advantages is that they enable predictions beyond the scope of available data (extrapolation). The complexity of stochastic models can make them difficult to understand and use correctly without specialized knowledge. This can be a barrier for stakeholders who are not experts in statistical or probabilistic methods (Goda and De Risi, 2023).			
Empirical (how two or more variables are correlated)	-The extremal dependence between river flow and sea surge (two hazards that combine to form a compound hazard). - Extreme rainfall and high sea level (Lian et al., 2013) - Extreme rainfall and landslides(rainfall-triggered landslides) (Guzzetti et al., (2007) - Landslide, flood, and earthquake (mapped in	-Dependence measures and regressions (linear, power, logistic, and quantile). - Qualitative methods for mapping such as AHP - Machine learning algorithms (Generalized Linear Model, Support Vector Machine, Random Forest (RF))	- **Guzzetti et al. (2007)** expressed the interrelation of extreme rainfall and landslides via a regression between rainfall intensity and duration obtaining a threshold for landslide triggering. (This relationship varies from place to place). - **van Westen et al. (2014)** conducted a multi-hazard risk assessment in the Barcelonnette Basin in the French Alps, focusing on landslides, debris flows, rock falls, snow avalanches, and floods. They applies GIS to compile and analyze various spatial datasets, including topographic data, historical hazard events, and environmental factors. The study involved generating susceptibility maps for different hazards using heuristic and statistical methods, conducting run-out modeling with the Flow-R software, and estimating temporal probabilities based on historical rainfall data and expert opinion. This study primarily focused on individual hazard layers.

	the form of multi single layers) - Snow avalanches, landslides, wildfires, land subsidence, and floods (multi layers of single hazards) - Collapsed pipes, gully heads, and landslides (mapped in the form of multi layers of single hazards) - Flood, landslide, and soil erosion (mapped in the form of multi layers of single) - Landslides, debris flows, rock falls, snow avalanches, and floods Landslides, Floods and wild fires (a multi-layer hazards approach). Droughts and forest fires (a multi-layer hazards approach)		- **Ward et al. (2018)** analysed the dependence between high sea levels and high river discharges to emphasize how this dependence can significantly increase flood hazards. - **Skilodimou et al. (2019)** developed a multi-hazard mapping framework for landslide, flood, and earthquake risks using the analytical hierarchy process (AHP). The study focused on evaluating the relative importance of these hazards independently. They found seismic hazard to be most significant due to frequent activity in the study area and a history of earthquakes triggering secondary events. The study did not account for interactions between these hazards. - **Pourghasemi et al. (2019)** addressed multi-hazard probability assessment and mapping in a mountain area for *landslides*, *flood* and *earthquake* using different machine learning algorithms to evaluate the spatial relationships among these hazards. However, each hazard maps were provided separately. The exact same research has been done by Bordbar et al. (2022) for Kermanshah province in Iran applying ensemble machine learning algorithoms. - **Yousefi et al. (2020)** focused on using machine learning models for multi-hazard assessments in a mountainous region in Iran for hazards including *snow avalanches*, *landslides*, *wildfires*, *land subsidence*, and *floods*. However, there was the lack of explicit attention to the interrelationships between different hazards. - **Kariminejad et al. (2022)** spatially mapped three hazards (collapsed pipes, gully heads, and landslides) using machine learning by sperately mapping each hazards. - **Piao et al. (2022)** developed a multi-layer hazards approach for droughts and forest fires using ensemble machine-learning algorithms Regression Trees Random Forest, and Boosted Regression Trees. The study identified

			high-susceptibility areas by integrating susceptibility maps for both hazards, with the BRT model showing the highest accuracy for predicting forest fire and drought susceptibility. - **Faramarzi et al. (2023)**: Evaluated potential environmental hazards (floods, landslides, and wildfires) in Golestan National Park, Iran, using machine learning techniques. The study applied models like maximum entropy, random forest, boosted regression tree, generalized additive model, and support vector machine to determine the most vulnerable areas. A multi-hazard map was produced by combining the individual hazard maps for wildfires, floods, and landslides. This map identifies areas with the potential for one, two, or all three hazards simultaneously. - **Javidan et al. (2024)** developed a multi risk map for flood, landslide, and soil erosion using different machine learning models in the Gorganrood Watershed, Iran. The final map was the sum up of the individual modelled hazards. Three different hazards (flood, earthquake and landslides) were mapped
Limitations: i) based on observed data and measurements, involve fitting statistical distributions directly to the data. Strongly depends on historical data. Therefore, there is a need for a good amount of data. **ii)** The main drawback compared to other approaches (stochastic and mechanistic) is "the impossibility to extrapolate beyond the range of the data". This can be a significant drawback when predicting events outside the historical record. **iii)** They are site-specific and varies from place to place depending on characteristics of each place like past data, and geography related data). **iv)** Monte Carlo type approaches are challenged by the high dimensionality of uncertainty due to the larger initial condition uncertainty space than non-ensemble models, as well as the computational intensity and structural uncertainty issues (Doyle et al., 2019). Despite these limitations, empirical models are valuable tools for analysing the relationships between different hazards and can provide insights that inform risk management and mitigation strategies.			

Mechanistic (based on the physical processes of a system and mathematical idealizations that represent real-world phenomena)	- Commonly applied on water bodies (hydrology) - a range of meteorological factors (wind, pressure, storm surge, and precipitation resulting in inland flooding) - Rainfall-induced slope failures and debris flows	Conceptual models and physical models (aim to simulate different systems)	- Tinti et al. (2003) used a hydrodynamic model to assess the interrelation between a **volcanic eruption** and **tsunamis** in the Gulf of Naples. They simulated a pyroclastic flow's impact from Vesuvius and its potential to generate a tsunami, finding that the maximum amplitude of the tsunami waves would be around 70 cm on the coastline. Chen et al. (2016) present a physically based model to look at interactions between regional rainfall, slope failures and debris flows, in a seismically active region in China. This intended to use in real time warning systems and large-scale landslide risk management.
Limitations: Substantial data requirements for modelling extreme hydrological events. Despite their ability to address weaknesses of empirical or stochastic models by using parameters with physical meaning, physical models are often computationally intensive.			

Despite advances in moving from single hazard to comprehensive multi-hazard assessment frameworks, natural hazards are still mainly considered independently (mainly flood and sinkholes). Even in the context of multi-hazards, most studies tackle the issue by overlaying individual hazards without considering their interrelationships. This will be discussed in **Section 2.4.9** where we specifically discuss the literature on multi-hazard mapping.

2.3 Karst terrains

Soluble rocks, such as limestone, marble, dolomite, halite, gypsum, and quartzite, form a significant part of Earth's geological composition, covering nearly 12% of the land surface (Lewin and Woodward, 2009). The development of karst is mainly related to the coexistence of soluble rocks, accounting for over 10-15% of Earth's surface (Ford and Williams, 2007). Karst landscapes are shaped by the interaction of soluble rocks with various water forms, including rainwater, rivers, and groundwater.

The distribution of karst landforms varies widely across the globe. Well-known examples of well-developed karst terrains are found in the Mediterranean basin. This region includes substantial areas in Spain, about a third of France (predominantly in the southern regions), and extensive parts of Italy (Lewin and Woodward, 2009). In the United States, Florida is particularly noted for its distinct karst features (Brinkmann, 2013). Apart from these locations, karst terrains are also scattered in various parts of the world, such as the Gauteng Province in South Africa, the Levant region, and the Atlas Mountains in Morocco and Algeria. Other notable karst regions include southern China and northeast England, each offering unique geological insights into karst formation and evolution (Brinkmann et al., 2008; Waltham et al., 2005). **Figure 2.3** shows an overview of the worldwide karst distribution.

The development of karst is different from one geographical setting to another. It is often due to the variety of contributing factors such as geologic (rock type, structure, and texture), topographic, and climatologic factors. Evaporite karst is one of the most sensitive settings with a high propensity to dissolution. Therefore, different interacting processes can lead to different consequences including landforms like caves and rock shelters, karst depressions, disappearing streams, and springs (Brinkmann, 2008; Lewin and Woodward, 2009).

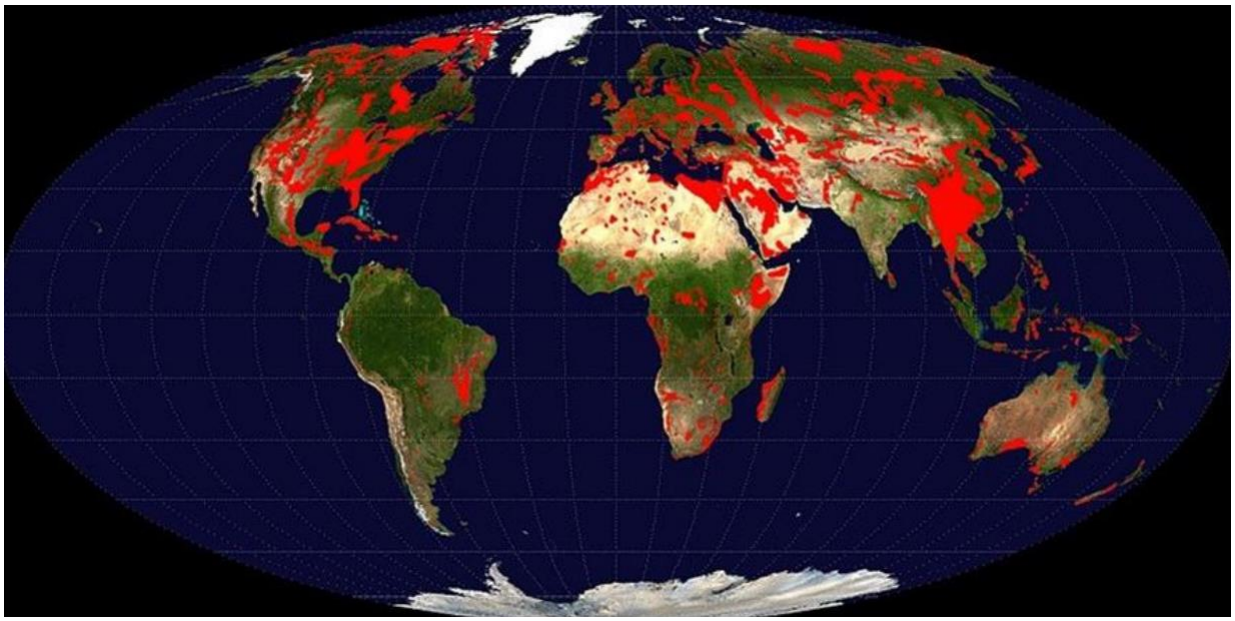


Figure 2.3 An overview of the worldwide karst distribution (Chen et al., 2017).

The chemical reaction of slightly acidic water on soluble rocks such as limestone marble, dolomite, halite, gypsum and quartzite produce a variety of geomorphological karst features namely sinking streams, caves, enclosed depressions, fluted rock outcrops and springs (Ford and Williams, 2007), making them strongly vulnerable to a variety of natural hazards. Flash floods, ground water flooding, back water flooding, slope failures, sinkholes, subsidence are the identified hazards possibly happening in such areas (Gutiérrez et al., 2014; Naughton et al.,2018). The dissolution of rocks in karstic structures poses serious multi-hazards to the communities residing in those areas. These hazards can affect various critical infrastructures such as transportation networks, including highways and high-speed train railways (Lewin and Woodward, 2009). Therefore, worldwide, they are perceived as sensitive environments (De Waele et al., 2011; Parise et al., 2018) requiring more protection and precautionary action. In addition, the damages of karst processes are more likely to increase with 25% of the world's population currently living in these areas (Zhou et al., 2003; Waltham et al., 2005; Gutiérrez et al., 2008).

Numerous classifications for sinkholes have been suggested (e.g., Waltham et al., 2005); however, Gutiérrez et al. (2014) recently introduced a system that identifies six genetic categories of sinkholes, organizing them into two main types. This system includes the processes of ground subsidence across both carbonate and evaporite karst areas. Through this classification framework (**Fig. 2.4**), sinkholes are distinguished into two genetically based categories: solution sinkholes and subsidence sinkholes.

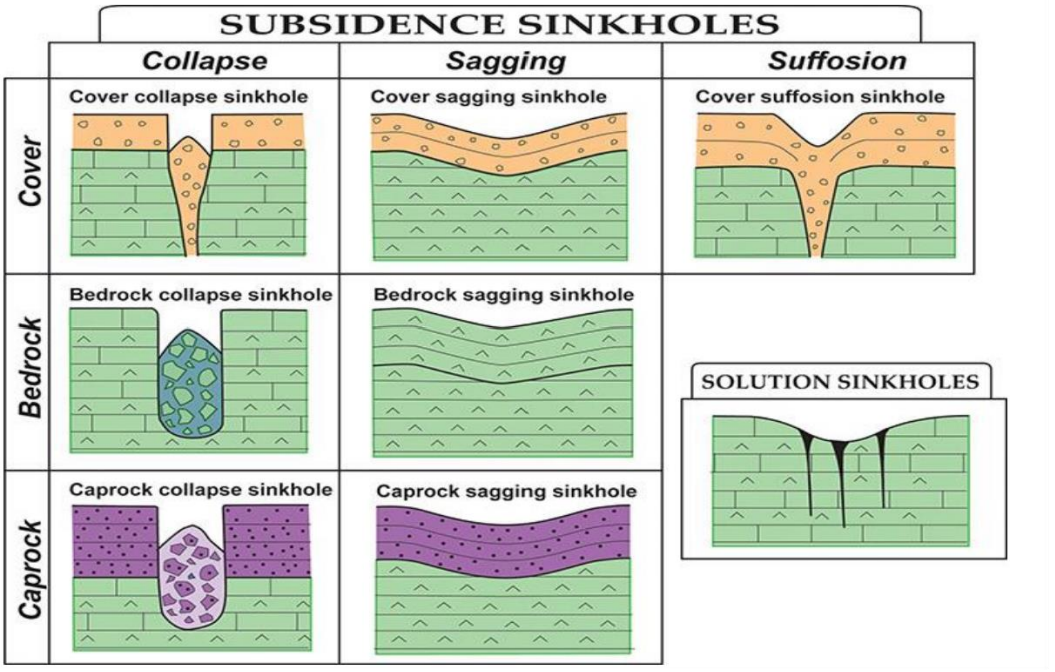


Figure 2.4 Genetic classification of sinkholes: Solution and subsidence sinkholes. The subsidence sinkholes are defined by the material impacted in the downward displacement and subsidence mechanism. The variations in the bedrock's pattern reveal the specific rock types where different subsidence mechanisms predominantly occur (Gutiérrez et al., 2014).

Solution sinkholes develop when the ground surface in exposed karst areas is eroded unevenly due to the dissolving action of water with higher concentrations of carbon dioxide. Water seeping into the epikarst – the uppermost layer of karst regions – moves sideways into areas where the rock is more permeable, which leads to increasingly rapid dissolution. Typically, these sinkholes have floors with a high concentration of clay and may experience flooding from time to time, although they do not often have problems with the ground collapsing or becoming unstable (Gutiérrez, 2016).

Subsidence sinkholes include a diverse range of sinkholes, all formed by the processes of underground dissolution and the deformation and/or internal erosion of the overlying, destabilized material. From a hazard point of view, these types of sinkholes are particularly significant because they involve the ground deforming, which can pose serious hazards. Subsidence sinkholes are classified by the material affected (cover, bedrock, or caprock) and the subsidence mechanism (collapse, sagging, or suffosion) (Gutiérrez, 2014; 2016). The terms "collapse", "sagging", and "suffosion" describe different mechanisms of subsidence sinkholes as outlined by Gutiérrez (2016). "Collapse" occurs when soil or rock deforms due to the creation of failure planes or fragmentation. "Sagging" is the bending or flexing of sediments caused by insufficient support underneath. "Suffosion" involves the downward movement and gradual settling of surface deposits. Collapses are typically sudden and catastrophic events. In contrast, sagging is a more gradual and slow process, whereas suffosion can occur quite swiftly. Identifying the type of sinkhole is crucial for both accurate hazard assessment and the effective design of mitigation strategies (Gutiérrez, 2016).

2.3.1 Karst terrains as a multi-hazard-forming environment

Given the definition of the multi-hazard term as "the specific contexts where hazardous events may occur simultaneously, cascadingly or cumulatively over time" (UNISDR, 2017), (natural) multi-hazard forming environments are where different geophysical processes (the extremes) (e.g., climate, topography, geology) have the potential to overlap in various ways (adopted from UNISDR, 2017). In that sense, it can be argued that karst landscapes, covering about 20 to 25% of the land surface (Waltham et al., 2005), possess such properties.

It is globally known that a sequence of heavy rainfall events can trigger sinkholes in many areas in particular karst environments (Xiao et al., 2016, 2018; Lei et al., 2013; Kwak et al., 2020). Both past and recent examples in karst settings reveal that a great proportion of sinkholes collapsed after a heavy rainfall and/or flooding happened in a given area. In Florida, where almost every citizen has a sinkhole story (Brinkmann, 2013), the tropical storm Debby in June 2012 caused a prolonged rainfall event followed by a period of drought. The flood caused by the storm triggered hundreds of sinkholes collapsing across the state leading to damage to infrastructures, buildings and forcing th e inhabitants to evacuate their homes (Brinkmann, 2013). Similarly, during a 2012 heavy rainfall in Liuzhou city of Guangxi province, south China, a cascade of collapse event occurred in a way that the first collapse resulted in 37 collapse sinkholes within a span of 5 days. (Lei et al., 2016).

The effect of severe rainfall and flood events on sinkhole collapses has been the main research topic of many studies (Hyatt and Jacobs, 1996; Brinkmann et al., 2008; Martinotti et al., 2017; Xiao et al., 2018; Parise et al., 2018b; Noury et al., 2018; Luu et al., 2019; Kwak et al., 2020). Xiao et al (2018), have

quantified the coupled effects of rainfall and groundwater head differences on the timing of sinkhole occurrences using the two proposed sinkhole hazard indexes (precipitation index (PI) and the head difference change index (HDCI) for central Florida. The two PI and HDCI indexes were computed to 10 target sinkholes. The results showed that five out of ten sinkholes happened shortly after heavy rainfalls or storms and the rest occurred due to head differences reaching peak values. This research concluded that both rainfall and groundwater head differences are one of the major factors influencing sinkhole-timing events.

Martinotti et al. (2017) characterised a six-day period of extreme rainfall events in 2014 with the cumulated rainfall exceeding 500 mm causing multiple hazards such as inundations, flash floods, landslides and sinkholes, in the karst terrain of Gargano Promontory, southern Italy. They concluded that there is a little chance that light to heavy rainfall events lead to hazards like landslides, floods and sinkholes, but heavy to torrential rainfall events are highly likely disruptive. To add more, in 1994, flooding of the Flint River triggered at least 312 sinkholes in the karst Dougherty Plain at Albany, Georgia. Flooding caused sinkholes by saturating and liquefying overburden. In a post-disaster research, Hyatt and Jacobs (1996) found that about 88% of the collapse events happened within the flooding limits.

Therefore, in karst terrains, meteorological events could act as a trigger to activate a cascade of hazards. **Figure 2.5** demonstrates a potential multi-hazard chain in an example of a karst terrain. As shown in the conceptual diagram, the slightly acidic rainwater stored in subterranean soluble rocks percolate down through the existing cracks. The exposure of groundwater and easily eroded rocks (carbonate and evaporate rocks) results in the rock dissolution creating small cavities or voids/hollows deep underground (Lewin and Woodward, 2009; Xiao et al., 2016). Over time, those cavities grow larger and the overlying surficial soils move downwards to fill in the cavities, resulting in downward erosion/dissolution of soil particles beginning from top of the overlying surficial soils. This process also can be upward by an upward propagating collapse of rock spanning underlying cavities (to know more about this process please see Noury et al., 2018; Luu et al., 2019). Eventually, when surface soils fall into the subterranean voids due to a loss of compaction, a sinkhole, the most common hazard in karst areas, occurs (Waltham et al., 2005).

Intense rainfall and flooding can apply load to those underground cavities that are close enough to the surface through saturating and liquefying overburden (Hyatt and Jacobs, 1996) and accelerate the internal erosion processes (Noury et al., 2018; Luu et al., 2019). The failure happens when the applied stress by flood load increase and reaches the maximum possible. Additionally, in case of the existence of multiple cavities close to each other, the collapse of one can cause a cascade effect for others creating bigger sinkhole collapses (the case that happened in Liuzhou city of Guangxi province, south China in 2012). Depending on the location of sinkholes, the collapses can increase the probability for another hazard. For example if the collapse of sinkhole break river's dykes another flooding could be added to the hazard chain. In multi-hazard literature, this relationship is called a bi-directional relationship or feedback mechanism where a primary hazard can produce itself (Gill and Malamud, 2016; Ciurean et al., 2018). Furthermore, this chain of hazards can potentially further continue if debris produced by a flood blocks the sinkholes' conduits, resulting in water not being able to infiltrate into the ground at sufficient speed and therefore causing flooding (Zhou, 2007; Xiao et al., 2016).

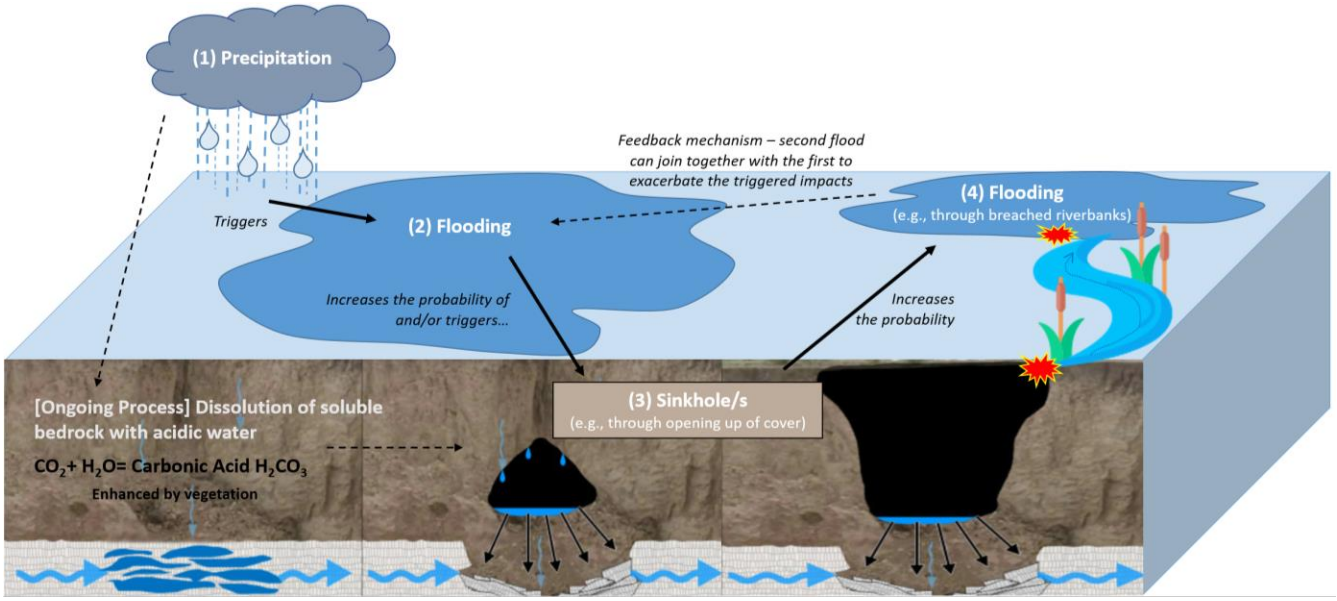


Figure 2.5 A simplified sequential diagram of the potential multi-hazard interactions in karst covered terrains. The diagram illustrates the sequence of events starting with (1) Precipitation, which triggers (2) Flooding (solid arrow). Flooding increases the likelihood of or trigger (3) Sinkhole formation due to the dissolution of soluble bedrock by acidic water (dashed arrow from precipitation to dissolution, an ongoing process over time). Sinkhole collapses can then lead to (4) Secondary flooding (e.g., through breached riverbanks, indicated by the solid arrow from sinkholes to flooding). The dashed bidirectional arrow shows the feedback mechanism where initial flooding can exacerbate subsequent flooding events. The arrows underneath the sinkholes depict the flow of water into underground cavities, contributing to the dissolution process and potential further collapse (Created by the author).

2.4 Building a systematic sinkhole-related literature review database

2.4.1 Data and Methodology

Reviewing 241 peer-review literature, this review outlines the current understanding of sinkhole-related studies and introduces the area of study to give the necessary background, and identifies the research gaps in the sinkhole mapping area. In the process of reviewing the literature, I have classified the studies into seven categories; **Sinkhole Monitoring and Detection (SMD), Sinkhole Morphology Studies (SMS), Interconnectivity of Flood hazard and Karst sinkholes (IFK), Sinkhole Hazard/susceptibility and Risk assessment (SHR), Mitigation Measures for Sinkholes (MMS), Multi-hazard mapping/ modelling,** and **Others.** This classification not only demonstrates the research gap in the field of sinkhole studies but the more favourite research directions. This literature review database begins with a global view in sinkhole-related studies identifying the gap. We visualise this through a violin plot (inspired by Tilloy et al., 2019), a particularly useful for showing the distribution of the data and giving insights into the variability and the density of the data over time.

This database was constructed through a critical review of the existing literature aimed at gaining insights into the current state of sinkhole-related researches and identifying the existing gap in the field of sinkhole hazard mapping. To construct this database, three steps were followed: 1) **Identification**, 2) **Screening**, and 3) **Classification**, illustrated in the flowchart presented in **Figure 2.6**. The following paragraphs elaborate on each of these steps. Before, delving into this systematic literature review, it is worth noting a few points on limitations:

- First of all, I should acknowledge that this database is not exhaustive and definitely, there are other related literature for each subject theme. It is not limited to years since the main aim was to find literature on multi-hazard mapping in karst areas linking with flooding etc. The focus here has been on selecting literature that is most pertinent to the topic. The process of choosing the references has been discussed in the following sections.

- Secondly, it is important to note that this categorization is not just about showing a wide range of research in this area. It is also about carefully choosing themes that are directly relevant and important to the focus of this thesis. Of these seven categories, four themes (SMD, SMS, IFK, SHR) have been identified as particularly relevant to the discussion of this thesis and will be explored in depth with each theme providing insights into the specific objectives and contributions of each research area. While the remaining categories also contribute significantly to the body of knowledge, they extend beyond the specific focus of this research (MMS and Others). Therefore, readers interested in those research themes can find the relevant references provided in the **Appendix 2.A**.

- Thirdly, whiles this review show a much less work done in multi-hazard assessment in karst areas, we also provide a literature review on the main methods applied to sinkhole susceptibility/hazard assessment as a single hazard. The theme "*sinkhole hazard/risk assessment"* delve into more detail categorising existing approaches plus their methods.

Therefore, while this global sinkhole related studies reveal a research gap, it also provide a good foundation to understand what has been done so far in sinkhole mapping.

- Last but not least, it is important to understand that in constructing a classification for sinkhole-related studies, the subjects often cover more than one area. This is because sinkhole research combines different fields like geology, geography, and engineering (being inherently multidisciplinary). Therefore, while our classification serves as a structured framework for organizing the literature, it is usual that some studies do not fit neatly into a single category and may overlaps with other categories as well. We use our categories as a guide to help organize the research literature.

- Once again, we acknowledge that it is a non-exhaustive database targeting representative research work in each theme. In the end of discussing the themes visualised by the violin plot, we also talk about the work has been done in the study area (Val d'Orléans) of this research. I will finish this chapter by introducing the existing gap.

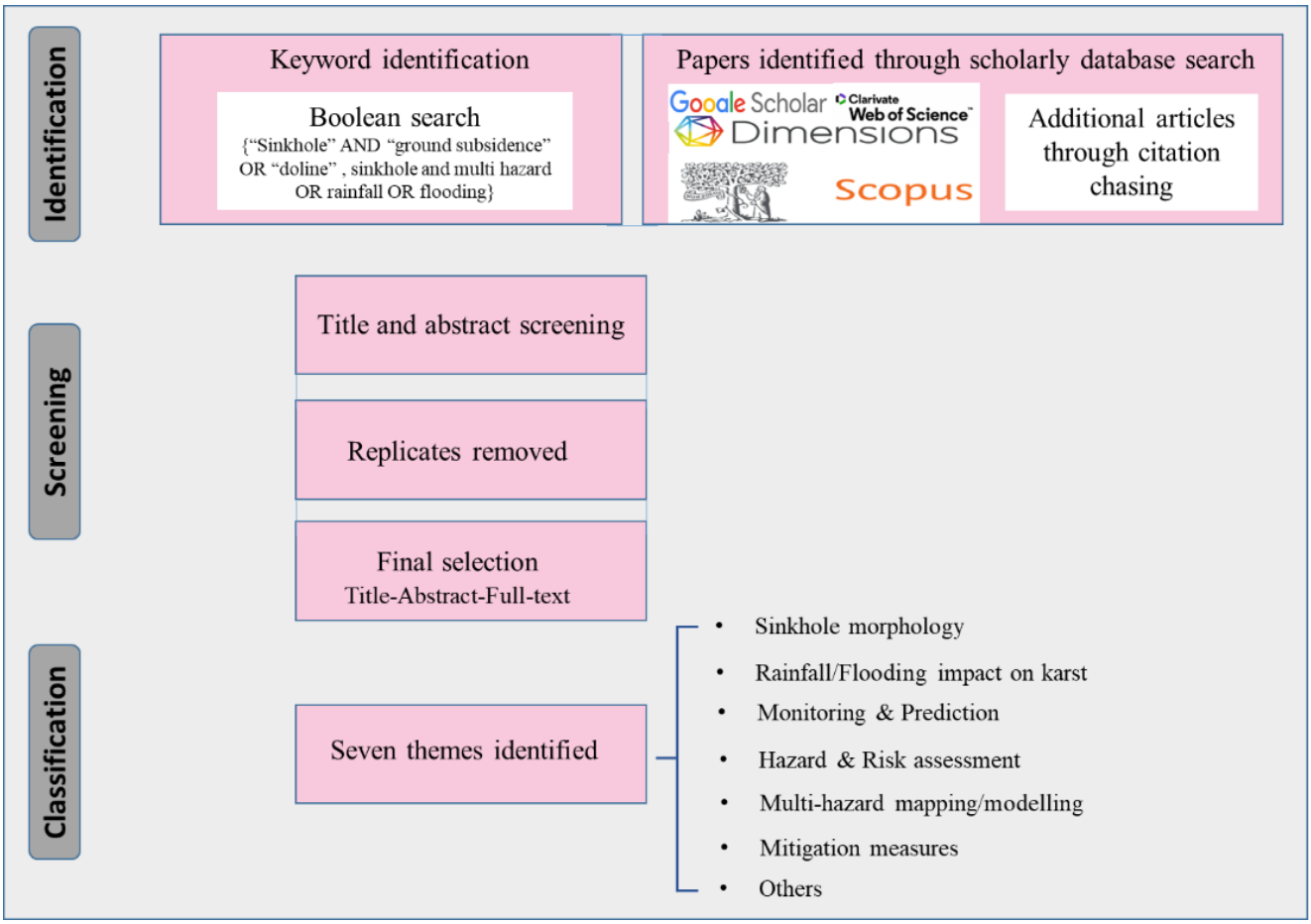


Figure 2.6 The flowchart of the construction of the systematic literature review database (source: the author).

1. Identification: To create our review database, the first step was dedicated to identifying the related international scientific publications through a set of specific keywords. Therefore, this step itself is divided into two steps: **a)** the identification of keywords and **b)** scientific articles.

a) Keyword identification: To efficiently identify potential keywords, I had to consider that not all references have used the same term for the ground collapse phenomenon. In most American publications, it is referred to as a "sinkhole", while European geomorphologists and many other independent authors use the term "doline" (Gutiérrez, 2016). Despite the slight scientific differences between "sinkholes" and "dolines", they are often used interchangeably (Fleury, 2009).

After conducting a text mining process, the resulting synonymous terms were gathered and used for the purpose of this review {'sinkhole', 'doline' 'karst collapse', 'ground subsidence', 'closed depression', and 'karst depression'}. Not to mention that there are also other terms to describe the same phenomenon in different countries (e.g., *cenote* in Mexico, *obruk* in Turkey, *cockpit* in Jamaica, *dahl* in the Arabian Peninsula (Gutiérrez, 2016), and *Foroochaleh* in Iran), but these were not included in this keyword search. Despite the inconsistency in terminology, often based on authors' preferences (Zhou and Beck, 2008), "*sinkhole*" is the most frequently used term in scientific international scientific literature. Thus, in the current research, while we have included the most common terms in our keyword research, we continue to use "s*inkhole"*. Taking the above-mentioned terms into consideration, we started by doing a Boolean search, adding rainfall, flooding and multi-hazard words as key words as follows {'sinkholes' AND 'ground subsidence' AND " karst collapse" OR "dolines", "sinkhole and multi-hazards" OR "rainfall" OR "flooding"} (**Fig. 2.6**)

b) Scientific articles identification: Through the selected keywords, I identified relevant references using various web search engines such as *Google scholar*, *Web of science*, *Scopus*, *Dimensions*, and *Mendeley*. Our primary focus has been on published international literature journals written in English. In any review research, one encounters a vast array of references. This is even more complicated for a phenomenon like a sinkhole being highly multi-disciplinary with different researchers from different majors have been working on. Thus, it is crucial to refine our focus by precisely determining what we aim to understand. For this study, our main emphasis has been on locating relevant research specifically on sinkhole hazard studies, particularly those integrated with flooding. As previously mentioned, we do not claim that this study encompasses all research papers in the areas depicted in the violin plot (**Fig. 2.7**). However, to the best of our knowledge, we have carefully selected and categorized the most pertinent scientific works into themes, representing the most relevant examples in the field.

2. Screening: For the screening part, we followed three concise steps. First, I inspected titles and abstracts of the selected articles. At this stage, the number of papers were gradually narrowed down by eliminating those that fell outside the scope of the current study. In the next step, I excluded articles that were either identical or highly similar, particularly those using the same methodologies or discussing analogous issues but in different geographical contexts. Manual search was further performed to remove unrelated studies or add relevant studies according to the researcher knowledge. That specifically was done for the theme

"Sinkhole hazard/susceptibility" through examining the secondary references (sources that were cited within the primary references, providing an additional layer of information and context to the main topic). In the end, the final literature database was created, including **241 references**.

3. Classification: During the classification phase, I reviewed the titles and abstracts, and in cases of papers' importance, we went through the full texts of the articles in the final database. This allowed us to extract a significant amount of information on the relevant concepts and theories, which then enabled us to identify and define seven thematic subjects within the field of sinkhole studies. This facilitated a more comprehensive and organized review of the literature. After this categorisation, we delved into each theme. Even though, we obtained this categorisation, we are mainly interested in themes closer to the scope of this thesis. Moreover, by delving into the theme "Sinkhole hazard/susceptibility risk assessment", we aimed to provide the literature review for the existing methods used to map sinkholes. Therefore, this theme is also a comprehensive reviewing of existing methods.

In the following paragraphs, I will explore each thematic subject and provide an explanation of the specific topics that each class encompass. Additionally, we will clarify the reasons for categorizing them into their respective categories.

2.4.2 Results

This systematic review was visualised through a violin plot inspired by Tilloy et al (2019) (**Fig. 2.7**). The violin plot is particularly useful for showing the distribution of the data and can give insights into the variability and the density of the data over time. By looking at the plots, we can obtain a quick understanding of how the data is spread out over years. This plot reveals several important keys: i) the overall trend in sinkhole studies; ii) the shift in studies and iii) the gap and future research directions. The most studied topic over a period of 44 years (1977 to 2022) is Hazard/risk evaluation.

The results highlight the importance of the integrated multi-hazard assessment in the areas affected by both flooding and karst related hazards and show that so far sinkhole risk assessment (83 articles – 34.5%) followed by sinkhole morphology (53 articles – 22%) have been the most popular research subject within the discipline. Multi-hazard assessment with only about 1.2 % is the least area studied in karst landscapes.

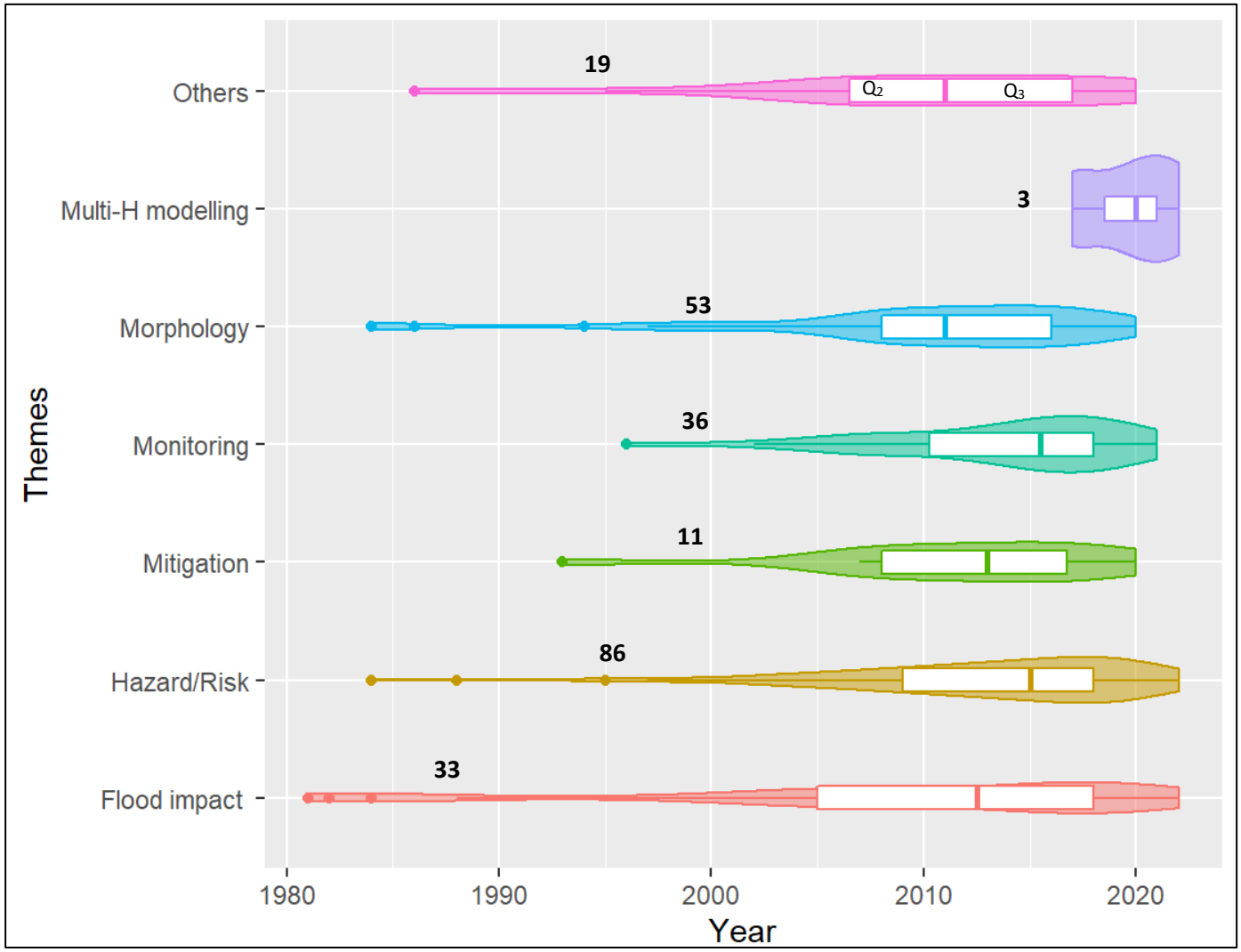


Figure 2.7 The evolution of sinkhole-related studies. The violin plot combined with a box plot overlay of the database of the seven research themes showing the evolution of studies on sinkholes. Q2 and Q3 inside the box plot refer to median (50% of the data falls below this value) and third quartiles (75% of the data falls below this value). This database consists of 241 papers with each plot displaying the number of papers corresponding to a specific theme/plot. (Source: the author plotted using R studio, package "ggplot2").

2.4.2 Theme 1. Sinkhole Monitoring and Detection

The theme of "Sinkhole Monitoring and Detection" represents the studies, which the focus and the scope have been mainly on applying techniques in order to detect and tracking sinkholes mainly in critical infrastructure such as railways, highways etc. (Guerrero et al., 2008; Xua et al., 2017; Pazzi et al., 2018; Williams et al., 2020; Sevil et al., 2021). Of 241 reviews, 36 research papers were identified; all sharing a common goal: to develop and apply new techniques such as *remote sensing, image processing techniques*, to detect sinkhole formation and monitor active sinkholes. Looking at the violin plot, there is an increase in the number of published paper in the category of sinkhole monitoring from 2010 onwards with 75 of the publications done from 2011 until recently. The shape of the violin plot shows that the research interest in monitoring sinkholes increased in the mid-2000s, with a secondary peak a decade later. This suggests the impact of technology advancements on the ability to detect and monitor sinkholes.

Monitoring and detecting sinkholes are vital during the construction and maintenance of critical infrastructure, especially in areas with karst topography (Xu et al., 2017). This advancement of digital and laser photogrammetry has enhanced the ability of understating the evolution and detection of sinhkoles (Al-Halbouni et al., 2017; Zumpano et al., 2019). Moreover, the results of these types of studies enhance the sinkhole early warning systems around the world (Atzori et al., 2015; Intrieri et al., 2015 and 2023).

Traditionally, sinkhole detection relied on visual inspection, a subjective and labour-intensive approach (Kim et al., 2019). However, modern methods now include remote sensing techniques, particularly for conducting morphometric studies of karst landscapes. Sinkhole detection and monitoring is linked to also applying geophysical techniques (discussed in **Section 2.4.3**). Asserting that the detecting and predicting the exact locations of sinkhole is almost impossible, Brinkmann and Parise, (2008) by evaluating climatologic and meteorologic setting under which sinkholes formed have assessed the timing of significant sinkhole occurrences in Tampa and Orlando, Florida. The results of the study show that sinkhole formation in the Tampa and Orlando regions is most common during the summer months, with a peak in August. The authors also found that sinkhole formation is more likely to occur in areas with high levels of rainfall and in areas with a high density of sinkholes.

Several researchers have proposed image-processing techniques to detect and delineate sinkhole boundaries. The main and common techniques extracted from the literature are as follows: Remote Sensing and Aerial Photography, Infrared and Radar Imagery (such as LiDAR - Light Detection and Ranging), Digital Elevation Models (DEMs), and Ground Truthing. Zhu et al. (2014) used LiDAR to map sinkholes in Kentucky as a more reliable alternative for topography maps. The results revealed that Lidar method could detect four times as many sinkholes as the existing database for the same area. Field checking suggested that the success rate of this method was between 80% and 93% for the study area, indicating the method is accurate and reliable. This is while Brinkmann, (2008) has used topography maps to detect the sinkholes in the city of Tampa arguing that air photos or satellite imagery are unable to identify sinkholes correctly and depending on the kind of the study, they cannot differentiate between natural and anthropogenic depressions.

Recent developments in InSAR monitoring systems highlight the evolution of sinkhole early warning systems, utilizing continuous data monitoring and integrating various data sources for enhanced hazard assessment. Intrieri et al. (2015) used GB-InSAR technology in detecting sinkholes and providing early

warnings, which involves radiating an area with microwaves and coherently registering the backscattered signal. From obtaining signals, a complex matrix of values, called a SAR (synthetic aperture radar) image, is obtained. By analysing the displacement and power images, it is possible to detect and track ground deformation, which can be an early warning sign of sinkhole formation. However according to the author, the precision of GB-InSAR technology exhibits limitations as follows; the technology is sensitive to atmospheric conditions, such as changes in temperature and humidity, which can affect the accuracy of the measurements. Moreover, the precision of GB-InSAR technology is limited to a few millimetres within a range of a few tens to hundreds of meters, depending on the instrument, which restricts its applicability.

Nof et al. (2019) describes a semi-automatic InSAR (Interferometric Synthetic Aperture Radar) monitoring system for early warning of sinkhole formation along the Dead Sea shores in Israel. This system utilizes time-series InSAR data for detecting precursory subsidence associated with sinkhole formation. The methodology includes continuous monitoring, data processing, and comparison with previous data and other independent information (like LiDAR-based sinkhole maps). When new subsiding areas are detected, alerts are issued to relevant authorities. The system also involves updating sinkhole susceptibility maps based on InSAR data, enhancing the effectiveness of hazard assessment and mitigation.

In a very recent publication in sinkhole early warning systems, Intrieri et al. (2023) have used a ground-based interferometric radar (GBInSAR) for monitoring ground displacements indicative of potential sinkhole formation. The methodology involves continuous radar surveillance, analysing displacement data to identify sinkhole precursors. By analysing the displacement data, the system could detect subtle ground movements that precede sinkholes, providing an early warning to mitigate risk and inform urban planning in this sinkhole-prone region. The integration of GBInSAR with other data sources enhanced the accuracy and reliability of the early warning system.

Moreover, Zhu and Pierskalla, (2016) used the Random Forests algorithm, applied in a karst region of central Kentucky. The primary objective was to efficiently distinguish sinkholes from other types of depressions in high-resolution LiDAR data. The challenge in processing LiDAR data for sinkhole extraction is the high volume of depressions it generates, necessitating an effective method to separate actual sinkholes from these numerous features. To address this, the researchers developed a Random Forests model, training it on a dataset where LiDAR-derived depressions were manually classified through visual inspection and field verification. Eleven parameters, encompassing both natural and human factors related to sinkholes, were chosen as predictive variables for the model. Due to the imbalanced nature of the training dataset, where the majority were non-sinkholes, a weighted Random Forests method was employed to enhance the accuracy of sinkhole prediction. The study underscores the potential of using automated methods like the Random Forest classifier for large-scale sinkhole mapping using LiDAR data, though it also acknowledges that such methods cannot entirely replace manual inspection and field verification processes.

In conclusion, the field of Sinkhole Monitoring and Detection has seen significant advances, transitioning from traditional visual inspections to sophisticated technologies like GB-InSAR and LiDAR. These developments have not only increased the accuracy of sinkhole detection but also revolutionized early warning systems. By integrating various techniques, from remote sensing to geophysical analysis, researchers have substantially improved the capability to monitor and predict sinkhole formations. This is crucial for ensuring the safety of critical infrastructure and supporting informed urban planning in karst

terrains. The synergy of these technologies represents a pivotal step forward in mitigating the risks associated with sinkholes globally.

2.4.3 Theme 2: Sinkhole Morphology Studies

Reviewing **53** published papers, we have determined to assign a theme named "Sinkhole Morphology". This is the second most focused topic after sinkhole hazard assessment started from 1984 but between 2007 and 2016 studies were more concentrated. The journal papers classified into this theme, are mainly revolved around research on characterising, distribution (spatial analyses), dimensions (morphometric characteristics), and the causal factors of sinkholes for both generation and development phases (Van Den Eeckhaut et al., 2007, Baryakh et al.,2011; Heidari et al., 2011; Lie et al., 2016) in various ways. The results of these types of studies provide a solid basis for further hazard and risk assessments (Parise et al., 2018) as they contribute in building sinkhole inventories. In fact, the quality of predictive models is highly associated with the accuracy and completeness of sinkhole inventory (Galve et al., 2008; Stocks, 2007). Therefore, these kinds of studies are of high importance in mitigation and disaster risk reduction plans. In the following paragraphs, a few studies will be explained. To access the selected papers, see the **Appendix 2.A.**

Many researchers have studied sinkholes analysing their distribution pattern to understand the landscape formation and how humans interact with the landscape (Brinkman, 2008). They have been seeking to answer the question "what factors control sinkhole distributions", (Gao and Alexander, 2008; Van Den Eeckhaut et al., 2007, Brinkmann et al., 2008; Constantinou and Rooy, 2018). This information may be valuable in further analyses of potential sinkhole hazards, decision-making and mitigation planning. In order to answer the question 'what controls sinkhole distribution', Gao et al. (2005) in Minnesota from the statistical results and the sinkhole density distribution concluded that sinkholes tend to form in highly concentrated zones instead of scattered individuals. Brinkmann et al. (2008) have analysed the morphometric characteristics of sinkholes such as size, perimeter, longest axis length, orientation of the longest axis and the circularity index in Tampa Bay area (Florida) in the aim of understanding the pattern of karst evolution in the given area. The examination of three sinkhole clusters namely circular, semi-circular, and not circular revealed that the spatial patterns of sinkholes did not show a regular and meaningful pattern, and that different geological factors influence the pattern.

Galve et al. (2011) assert that the distribution pattern of sinkholes in the Ebro valley is mainly associated with a decrease in the groundwater table, which then intensify the karstification. Nevertheless, it is crucial to differentiate between factors contributing to sinkhole formation from those of triggers sinkhole collapse. In **Table 2.3**, I have classified both causes and triggering factors extracted from the literature. Magdalene and Alexander. (1995) analysed the distribution of sinkholes in Winona County, Minnesota using nearest neighbour analysis and concluded that the sinkhole distribution is clustered. They also noted that new sinkholes do not develop clustered around old, existing sinkholes, but around other newly developed sinkholes. However, on the scale of several kilometres new sinkholes develop around existing old sinkholes.

Among the selected journal papers, there are additional studies that investigate the formation mechanisms of sinkholes, which are also categorized under the subject theme of Sinkhole Morphology (Heidari et al., 2011; Lie et al., 2016; White 1988; Sowers, 1996; Sharp, 2003; Beck, 2004; Gutierrez et al., 2007; Zhou

and Beck., 2011; Parise et al., 2018a). The formation mechanisms of sinkholes vary from site to site depending on site-specific causal factors. Different formation mechanisms lead to sinkhole formation and development and it is not necessarily related to karst phenomena. In fact, there is no single cause leading to sinkhole formations. Sinkholes originate from both natural and human interventions on the environment (Gutierrez et al (2008) (Table 2.3).

Natural sinkholes also called karst sinkholes typically are formed in karst areas where soluble carbonate and evaporate bedrocks are the predominant landscape, or alluvial sediments in special geological situations. In other words, they naturally form due to the chemical and physical erosion processes occurring on the surface and in the subsurface. (White and White, 1984; Waltham and Fookes, 2005, Kovacic and Ravbar, 2010; Liu et al., 2016). Hence, as an initial estimation, the natural sinkhole-prone areas can be zoned based only on geological factors. Natural sinkholes are typically triggered by heavy precipitation events or by changes in the hydrological/hydrogeological regime (Kuniansky et al., 2016; Parise et al., 2018a). These events might have effect on groundwater circulation resulting in an increase in karst sinkhole formation. Direct and indirect human action on the nature accelerate the rate of sinkhole developments (Kuniansky et al., 2016) by ground water over-extraction, pumping or mining, surface loading, urban expansion in sinkhole-prone areas, resulting in increased effective stresses on the supporting rock and drainage of water-filled submerged caverns, all of which can result in increases in sinkhole formation.

Nonetheless, the situation is quite different for anthropogenic sinkholes, as their occurrence is influenced by many other factors, such as historical, cultural, and social issues (Parise, 2015). In this regard, it is important to note that in this thesis, we excluded any events of anthropogenic origin that were attributed to water leakage from pipelines and resulting subsoil erosion. Anthropogenic sinkholes can be more common in some areas. For instance, in Naples-Italy, the main risk to the citizens is generally associated to anthropogenic sinkholes, with the possibility of collapses related to man-made caves (quarries, mines, civil settlements, see Parise and Vennari, 2013; Parise et al., 2015; di Santolo et al., 2016) while natural sinkholes are less reported.

In Nasimshahr (south-west of Tehran, Iran) recently a few ground collapses have been observed which have been attributed to the abandoned Qanats, a series of well-like vertical shafts, connected by hand-dug underground tunnels, invented for sustainable groundwater withdrawal and transfer in ancient Iran. The sinkhole collapses are in close proximity to Qanat wells used for providing drinking water (Beitollahi and Mahdavi., 2021). Those Qanats have buried over time due to changes in surficial soil, but their horizontal underground tunnels, which back in a day acted as underground waterway, remain empty. The underground voids arose from abandoned Qanats can then collapse in the event of stresses due to urbanization etc. (Beitollahi and Mahdavi, 2021). In the far west land located in south Africa, sinkholes and subsidence depressions have formed as a result of the depletion of water table in relation to mining in the gold bearing reefs (De Bruyn and Bell, 2001). Quite the opposite, in some areas, sinkholes develop as a result of the drowning process. It happens when the water table is high up to almost the soil surface (Parise and Vennari, 2013). The acidic gases, which are present in anthropogenic sinkholes. Brezinski (2007) prepared a detailed geologic map and karst features description of Frederick Valley, Maryland, and used this information to point out some factors that contribute to sinkhole formation and to identify the geological units most susceptible to sinkhole development. He noted that human activities such as changing surface stream drainage, storm-water management areas, active quarrying, and soil overburden removal, additionally induce sinkhole development.

Table 2.4 Identified causal and triggering factors to sinkhole formations extracted from the literature.

Triggering Factors			**Sinkhole causes factors**	
Natural	**Anthropogenic**		**Natural**	**Anthropogenic**
Heavy rainfall/ rainstorm/ Flooding	Maintenance works		The interaction of soluble rocks with slightly acidic water	Abandoned underground coal mines
Prolonged drought	Pipeline ruptures		depending on the below-mentioned factors the dissolution rate differs: Soil types/underlying geology/faults and folds/thickness of the overburden	Abandoned Ancient underground or the presence of buried channels (called Qanat in Persian)
Seismic shocks	Groundwater pumping for drinking water and irrigation			Well drilling and mining
Seasonal fluctuation of the water table	Urbanization (land use changes)		The growth and decay of permafrost (cryogenic processes)	leaking water, sewer, or drainage pipes

2.4.4 Theme 3. The interconnectivity of flood hazard and karst sinkholes

In total, we identified about 33 main publications on the interaction of flooding and sinkholes. As shown by the violin plot, the research on this area is quite old staring from early 1980s. It indicates that concerns about the impact of these hazards have been a topic of study since then. The plot shows a progressing trend towards recent years with 75 percent of the studies done between 2005 to 2023. The relationship between rainfall/flooding, and their role as triggers for hazards in karst areas, has been examined from various perspectives across different disciplines. From reviewing the literature, it can be understood that in karst landscapes, the interplay between rainfall, flooding, and sinkhole formation/collapses presents a complex and bidirectional relationship in which both can trigger each other. Previous real-world experiences show that flooding followed by heavy rainfall is the key impact factor, which increases the probability of sinkhole occurrence.

Heavy rainfall/flooding can trigger sinkholes by rapidly infiltrating the subsurface, increasing water pressure and eroding underground cavities (the sediments within). This process can lead to sudden collapses of the ground surface, forming sinkholes. This mechanism has been discussed in geophysical research work (Noury et al., 2018; Luu et al., 2019).

Conversely, existing sinkholes in karst areas can exacerbate flooding so called “sinkhole flooding”. They act as direct conduits for surface water to enter the subsurface, changing the natural drainage patterns and

potentially leading to unexpected flood events (Kovacic and Ravbar., 2010). Understanding this complaex relationship is crucial for assessing the risks and managing the impacts in karst terrains.

In building my database, I initially looked for studies examining how floods lead to the formation/collapse of sinkholes. However, we discovered that a substantial body of research is dedicated to the phenomenon of sinkhole flooding in karst areas in particular in the USA. Therefore, we have expanded the database to include these studies as well. In the following paragraph, two main recent studies on sinkhole flooding in karst regions will be discussed. Subsequently, other studies on the reciprocal relationship of this phenomenon (flooding triggers sinkholes), will be outlined discussing the main five research.

Geologically, sinkholes have served as storm drainage outlets. They are effective collectors of surface runoff, leading to chronic drainage problems. During rainy seasons, some sinkholes may flood, and urbanization generally increases the frequency of such flooding (Kemmerly, 1981). This phenomenon could often be a component of the broader multi-hazard context characteristic of karst regions. Campbell (2005) asserted that most of the flooding problems in Bowling Green, Kentucky are related to sinkholes in a way that most sinkholes experienced spill-over during a 100-year flood event. He discussed that to improve and enhance flood mapping in regions with sinkholes, flood elevations should be accurately determined. By integrating the overflow effects from nearby sinkholes and applyig tools like ArcGIS, 3D Analyse, flood prediction accuracy in karst landscape could be more accurate.

Kovacic and Ravbar (2010) argued that karst areas, particularly karst poljes, are prone to flooding due to their unique geological and hydrological characteristics. Karst landscapes typically feature soluble rock formations like limestone, which over time, develop underground drainage systems with caves and sinkholes. These complex systems result in a rapid and direct transfer of surface water into the subsurface, leading to quick responses to heavy rainfall. Moreover, the topography of karst poljes, being large, flat depressions, makes them natural collecting basins for water, exacerbating their susceptibility to flooding. This inherent vulnerability of karst areas to flooding is a key focus of the study. To prove this, the authors used site-specific data such as daily precipitation, water levels, discharge, and digital elevation models, to analyse the flood dynamics in karst areas of Slovenia. The flood dynamics including water levels, the magnitude of floodwaters, and the flood extend during extreme hydrological conditions in autumn 2008 was measured. This approach demonstrated how karst aquifers quickly react to heavy rainfall, showing behaviour similar to surface water bodies but with limited water retention.

In another study, Fischer et al. (2022) carried out a research to understand how the sinkhole's drainage capacity influences runoff during various rainfall events. It explored, in particular, the threshold rain intensity at which runoff exceeds the sinkhole's capacity, contributing to the risk of flooding in the village of La Vacquerie, France. They applied a 2D hydrological numerical model, Hydrologic Engineering Center's River Analysis System (HEC-RAS). The model simulated responses to four different rainfall events of varying durations and intensities. The research found that sinkhole's drainage varied with rain intensity. It was limited during long and light rains but increased during short and intense rains. The studied sinkhole was submerged when rain intensity exceeds a certain threshold (110 mm/h in this study). By this research, the study highlighted the critical role of the studied sinkhole and other karst features in influencing the severity and occurrence of flash floods. These insights are vital for a deeper understanding of flood mechanisms in karst landscapes and are essential for developing effective strategies to mitigate future flood risks.

More references can be found in the literature selected for this theme (see the **Appendix 2.A**). I continue outlining the five most recent published paper having done research on relationship between flood-triggered sinkhole collapses or sinkhole collapses during flooding.

Following a 6-day period of extreme rainfall events in 2014 in the karst terrain of Gargano Promontory, southern Italy, leading to multiple geohydrological hazards, including landslides, floods, and sinkholes, Martinotti et al. (2017) highlight the lesser-known impact of intense rainfall in causing sinkholes in karst environments, in addition to more commonly recognized hazards like floods and landslides. The study analysed the spatial and temporal distribution of rainfall and the resulting geohydrological hazards. They used hourly rainfall records from 39 rain gauges to study this intense rainfall period, during which the Gargano Promontory experienced cumulative rainfall exceeding 500 mm. The study introduced the Ensemble-Non-Exceedance Probability (E-NEP) algorithm for forecasting the likelihood of rainfall-induced landslides and related geo-hydrological hazards. This algorithm uses rainfall data to calculate the probability of landslide occurrence over time. The study concluded that in karst environments like the Gargano Promontory, geo-hydrological hazards occur in response to extreme meteorological conditions. The region's threshold behaviour to rainfall-driven hazards was highlighted, indicating that hazards are rare for light to heavy rainfall but become significant and disruptive for heavy to torrential rainfall. The study emphasized the unique response of karst landscapes to intense rainfall, including efficient infiltration rates, poor retention capacity in karst aquifers, and the propensity for flash floods and sinkhole formation under extreme conditions.

The study done by Xiao et al. (2018) conducted in central Florida focused on assessing the impact of various hydrologic and hydrogeologic conditions, such as rainfall, groundwater downward leakage, and groundwater head differences, on the development and timing of cover-collapse and cover-suffosion sinkholes. They developed a methodology for predicting sinkhole occurrences, involving the collection of continuous rainfall and groundwater level data. By calculating the Precipitation Index (PI) and Head Difference Change Index (HDCI), researchers could determine the primary factors triggering new sinkholes. The results indicated that heavy rainfall/storms and rapid increases in head differences over short periods are significant factors influencing the timing of sinkhole occurrences. Furthermore, the spatial variation in groundwater downward leakage rate can be used to create sinkhole susceptibility zonation maps, aiding in predicting the likelihood of sinkhole development in specific areas. The study found a direct correlation between sinkhole spatial density and the rate of groundwater downward leakage. Areas with higher leakage rates exhibited greater sinkhole susceptibility, indicating a need for careful monitoring and management in these regions. Analysis of specific sinkholes revealed that both heavy rainfall and rapid head differences played major roles in their formation.

Following a severe drought in 2016–2017, the 2018 long rainy season (MAM) in Kenya was exceptionally wet, leading to extensive flooding with significant impacts on society, environment, and economy. Rateb and Hermas. (2020) incorporated hydrological data (SMAP, VIC, NOAH-3.3 models) and geodetic data (GPS, InSAR) to investigate hydrological changes and subsidence potential. Land deformation related to the 2018 MAM season was analysed using both GPS and InSAR measurements. The study demonstrated varying magnitudes of subsidence, with InSAR measurements showing widespread subsidence ranging from 20.0 mm to 40.0 mm/year across the studied area, particularly in the Kapiti-Athi and Thika plains. The highest subsidence was observed during the peak of the rainy season in April, correlating with the highest hydrologic mass loading. Following the rainy season, there was a notable isostatic rebound, with about 50% recovery of the subsiding land by August 2018. They concluded that there is a potential

relationship between heavy rainfall and ground subsidence in the study area. The occurrence of heavy rainfall events and flooding over volcanoclastic regions, primarily covered by clay soils, likely intensifies sinkhole processes where subterranean channel erosion actively flushes bottom materials downwards, leading to land subsidence. The spatial distribution of land subsidence was closely correlated with the areas that experienced intensive flooding during the 2018 MAM season. The study also highlighted the limitations in validation of InSAR measurements of land deformation due to the absence of dense GPS stations and in situ rainfall and flood gauge stations.

Kwak et al. (2020) investigated the relationship between rainfall intensity and the development of sinkholes, particularly those caused by soil erosion due to groundwater infiltration through damaged sewer pipes. They conducted model experiments simulating urban areas in Seoul, using weathered granite soil commonly used for backfilling sewer pipes. The study focused on three different rainfall intensity conditions: heavy, very heavy, and extremely heavy rainfall. Key methodologies included continuous measurement of ground level changes, ground displacements, and the application of particle image velocimetry (PIV) to analyse displacement at various positions of the model ground. The study found that the height and average width of cavities and sinkholes increased linearly with the applied hydraulic head. Specifically, under extremely heavy rainfall conditions, notable ground subsidence was observed. This implies that the soil's resistance factor becomes smaller than the combined effects of upward force (due to infiltration pressure during water supply) and gravity, particularly when rainfall intensity exceeds a specific threshold. The study also highlighted that after heavy rainfall, leading to a rise in groundwater levels through sewer pipe defects, soil was rapidly discharged from areas above these defects, significantly influencing cavity formation and expansion. These findings underscore the direct impact of rainfall intensity on soil erosion near sewer pipe defects and sinkhole occurrences. They suggest a need for sewage pipe network facility expansion and rehabilitation to mitigate the risks of sinkholes, especially considering the projected increase in torrential rains due to climate change.

Table 2.5 Summary of key publications on the interaction of flooding and sinkholes.

Publication	Key Focus	Key Findings	Location	Methodology
Fischer et al. (2022)	Sinkhole's drainage capacity and its influence on runoff during rainfall events	Sinkhole's drainage varies with rain intensity; limited during long/light rains but increases during short/intense rains. Threshold rain intensity identified as 110 mm/h.	La Vacquerie, France	2D hydrological numerical model (HEC-RAS)
Rateb and Hermas (2020)	Relationship between heavy rainfall and ground subsidence	Heavy rainfall during 2018 MAM season in Kenya led to significant subsidence, highlighting the interaction between flooding and subsidence processes.	Kenya	Hydrological data (SMAP, VIC, NOAH-3.3 models), geodetic data (GPS, InSAR)
Kwak et al. (2020)	Impact of rainfall intensity on sinkholes due to soil erosion and sewer pipe defects	Rainfall intensity directly impacts soil erosion near sewer pipe defects and sinkhole formation, emphasizing the need for infrastructure improvements to mitigate risks.	Seoul, South Korea	Model experiments, continuous ground level measurement, particle image velocimetry (PIV)
Xiao et al. (2018)	Impact of hydrologic conditions on sinkhole development	Heavy rainfall and rapid head differences are significant factors influencing sinkhole timing. Groundwater leakage rates can predict sinkhole susceptibility.	Central Florida	Continuous rainfall and groundwater level data collection, Precipitation Index (PI), Head Difference Change Index (HDCI)
Martinotti et al. (2017)	Impact of intense rainfall on sinkholes and other	Extreme rainfall in 2014 led to multiple hazards including sinkholes; introduced the E-NEP	Gargano Promontory, Italy	Analysis of rainfall records, Ensemble-Non-Exceedance

	geohydrological hazards	algorithm for forecasting rainfall-induced hazards.		Probability (E-NEP) algorithm
Luu et al. (2019); Noury et al. (2018)	Mechanisms of sinkhole formation due to heavy rainfall/flooding	Rapid infiltration of water increases subsurface pressure and erodes underground cavities, leading to ground collapses and sinkhole formation.	Val d'Orléans, France	Geophysical research
Kovacic and Ravbar (2010)	Sinkhole flooding in karst areas	Sinkholes act as conduits for surface water to enter the subsurface, altering drainage patterns and causing unexpected floods.	Slovenia	Analysis of daily precipitation, water levels, discharge, and digital elevation models
Campbell (2005)	Flooding problems in Bowling Green, Kentucky due to sinkholes	Most sinkholes experienced spill-over during a 100-year flood event, suggesting the need for accurate flood mapping and integration of overflow effects from nearby sinkholes.	Bowling Green, Kentucky	Integration of ArcGIS, 3D analysis for flood prediction

2.4.5 Theme 4. Mitigation Measures for Sinkholes

In our database, we have identified 10 papers out of a total of **241** that specifically address the topics of repairing and stabilizing sinkholes, as well as the utilization of lands in the vicinity of karst landforms (land use planning). I have categorised this area of study under the class "Mitigation Planning for Sinkholes". Mitigation can be the next step after sinkhole monitoring and detection or it can be also a post-hazard event step. This section is linked to geophysical approaches and engineering solutions. These publications are devoted to understanding how to repair sinkholes, designing and implementing structural solutions to stabilize the ground, such as underpinning, grouting, and other ground improvement techniques. As explained by Gutierrez et al. (2008), sinkhole mitigation requires a multifaceted approach, which take into consideration geological, geotechnical, hydrological, and risk assessment studies. The methods and measures to be implemented all depend on the local geology, hydrogeology, and the spatial-temporal patterns of sinkhole occurrence. While preventive measures are preferred, in areas where development has already occurred, a combination of monitoring, engineering solutions, and community engagement is necessary to mitigate risk effectively Gutierrez et al. (2008). As mentioned earlier, the nature of this class

is highly multidisciplinary and mainly overlaps with the geotechnical research in the theme No 5. In addition to this, since this topic is not the concern of the current thesis, I do not go into detail. However, the related citations can be found in the appendix for further reference.

2.4.6 Theme 5. Sinkhole hazard/susceptibility and risk assessment

This section provides an overview of previous research on sinkhole hazard and risk assessment, drawing from a total of 83 papers that are classified under this subject. Moreover, an overview of sinkhole hazard/susceptibility modelling classification extracting from existing literature is shown in the **figure 2.8**. Within our database, Hazard/ risk assessment is the most studied topic over a period of 40 years from 1983 to 2023. The plot shows that from 2005 to 2018 the studies were more condensed. Over the recent decades, sinkhole hazard/susceptibility modelling and mapping has been in practical uses for mitigation planning purposes. Such a map is an important tool that informs residents about which locations are at risk and helps in reducing risk posed by sinkholes. Moreover, this map is a beneficial resource for urban planners and technical experts, offering an efficient and economical method for designating zones that are more likely to experience sinkhole formation.

Within this period, researchers from various disciplines have differently approached sinkhole hazard analysing. Geologists, soils engineers, geomorphologists, urban planners, and geophysicists have been working on sinkholes with different perspectives for different spatial scale (local, regional and national). Consequently, depending on a discipline, the scale (spatial extend) and the aim of a given study, methods and approaches may differ. This is also the reason why there are many different modelling classifications in the literature, which often lead to a big confusion for researchers. In this section, I have selected 78 publications on specifically working on sinkhole hazard modelling through different lenses. Within this theme I have categorised the approaches into three broad families. The flow chart depicts this classification

In the following paragraphs, different existing approaches in the literature will be discussed followed by the most common methods in sinkhole susceptibility/hazard mapping.

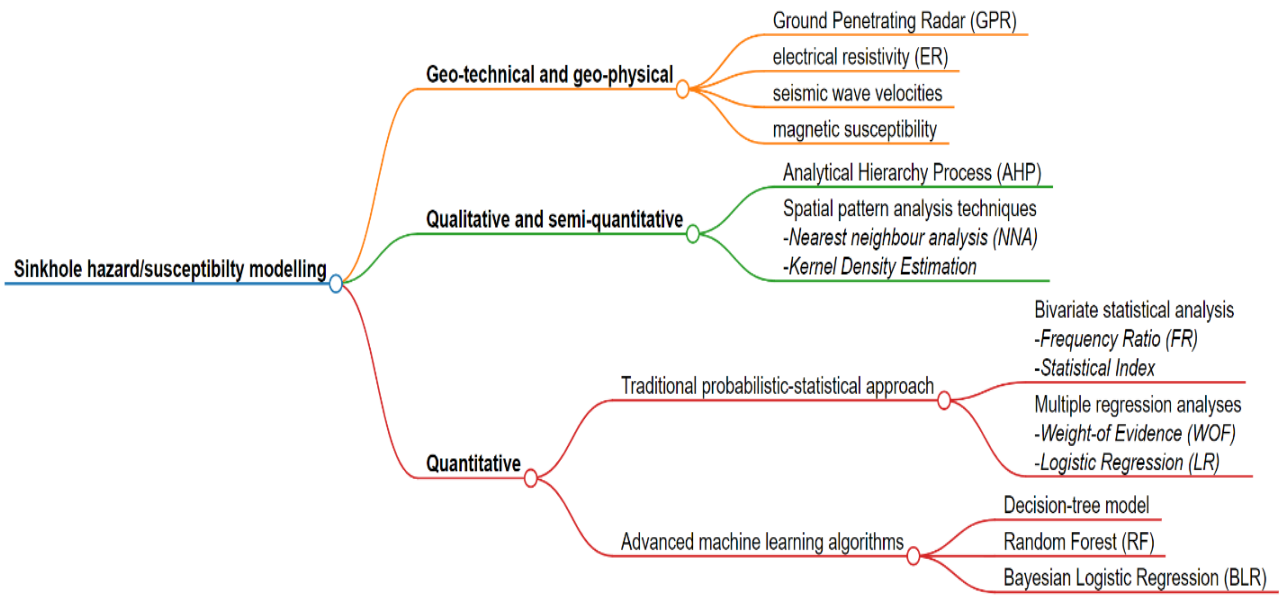


Figure 2.8 An overview of sinkhole hazard/susceptibility modelling classification extracting from existing literature (source: Author, created by markmap.js.org).

2.4.6a Geo-technical investigations and geophysical methods

The first group of approaches for identification of the sinkhole-prone areas belongs to geotechnical investigations and geophysical methods aiming at sinkhole detection and determining sinkhole formation mechanism. The aim of these methods is to understand the subsurface conditions of a location, the cause and the extend of a sinkhole (Ezersky et al.,2013; Kaufmann 2014 ; Alrowaim et al., 2015; Strzałkowski, 2018; Amin et al., 2023). Regarding the importance of a specific location and the degree of stability, these methods are applied to detect subsurface sinkholes or predicting the behaviour of earth materials that may lead to sinkhole formation. They are mainly based changes in physical properties near sinkholes (Kaufmann, 2014). These techniques include gravimetric techniques to identify voids, as well as electrical resistivity (ER), chargeability, seismic wave velocities, magnetic susceptibility, and density which are known for their ability to identify changes in the physical characteristics of the ground (Pazzi et al., 2018). Underground cavities may contain water, air or collapsed/unconsolidated material, leading to detectable changes at the surface (Kaufmann, 2014; Pazzi et al., 2018). The aim of this type of approach also overlaps also with the theme "sinkhole mitigation measures".

One commonly applied geophysical technique is *Ground Penetrating Radar (GPR)*, which uses high-frequency electromagnetic waves to investigate the subsurface. By analysing the reflected signals, GPR can detect changes in subsurface composition, including voids or cavities that may indicate the presence of a sinkhole (Amin et al., 2023). Kaufmann (2014) investigated sinkholes by an integrated geophysical approach to understand the evolution and geophysical signatures of both solution and collapse sinkholes. The methodology combines gravimetry, magnetics, Electrical Resistivity Tomography (ERT), and Self-Potential (SP) measurements to detect and analyse subsurface anomalies associated with sinkholes. The results revealed that these methods can identify potential sinkholes before they fully develop, based on the anomalies or changes in the physical properties of the subsurface. However, once the sinkhole has actually formed and become mature, the significant geophysical signals often disappear. This could be due to the physical properties of the subsurface area becoming more uniform or less distinctive, making it harder for geophysical methods to detect any anomalies.

Even though they usually result in more accurate and detailed investigations, they only zoom in a local scale, which cannot be feasible for larger areas due to being costly requiring high demanding of trained and technical workforce (Taheri et al., 2021), and a huge amount of field work and post-processing corrections required (Pazzi et al., 2018). Furthermore, the effectiveness of identifying cavity location can be limited by the geological context, which may make the technical methods and geological accuracy not always applicable (Tehrani et al., 2022). When choosing a suitable geophysical method, it is important to carefully estimate the benefits they offer, their inherent constraints and uncertainties, and the cost-to-time ratio involved (Pazzi et al., 2018).

In the context of creating sinkhole hazard maps for regional and national scales, approaches can be divided into two broad categories; **1)** qualitative/ hybrid/semi-quantitative, and **2)** quantitative methods (Galve et al., 2009a; Galve et al., 2009b). Selecting each of those is highly dependent on the data availability and the purpose of a given research.

2.4.6b Qualitative and semi quantitative methods:

Many studies addressing sinkhole susceptibility use methodologies largely based on relatively subjective judgements that fall within the area of qualitative and semi quantitative approaches (Galve et al., 2009a; Galve et al., 2009b). Qualitative methods apply heuristic approaches that are descriptive representations built on expert judgments besides field observations (Qiu et al., 2020). It is an expert, who decides whether a given area is sensitive to sinkholes or not, and to what degree it is susceptible. Their judgements come from their long experience in the field, and/or good pattern understanding.

A common and widely used method of this approach is Analytical Hierarchy Process (AHP). The AHP has been frequently applied in the natural hazard science. There has been a great number of studies in the literature, applying the AHP method to assess sinkhole hazard or susceptibility (i.e., Rosdi et al., 2017; Jiang et al. 2005; Taheri et al., 2015; Gao and Alexander, 2008). This approach, which was first introduced by Thomas Saaty in the late 20th century, is one of the most popular and widely used multi-criteria methods addressing complex decisions. Usually, a set of factors contributing to sinkhole formation is chosen by a certain factor depending on the geographical setting. Having selected the criteria and sub-criteria, they are compared to each other by a pairwise comparison matrix according to their relative importance to determine which predisposing factor is more important (Saaty, 2008.) In the following paragraphs, the main methods and their drawbacks will be presented.

Analytical Hierarchy Process (AHP): A high majority of researchers have long focused on direct hazard mapping zonation built on expert criteria (i.e. Rosdi et al., 2017; Jiang et al. 2005; Taheri et al., 2015; Gao and Alexander, 2008). The AHP is one of the most popular and widely used multi-criteria methods addressing complex decisions. The technique is applied for ranking a set of alternatives or for the selection of the best in a set of alternatives. The evaluation process yields opinions based on pair-wise comparisons hence it is an ideal technique for decision making involving multiple choices. In other words, in this approach, you have a set of criteria, which should be priorities by pair-wise comparisons. The importance of the decision is directly linked to the weight of a criterion (**Table 2.6**).

For example, a preference of the weight 1 means that two sub-indicators are equal, whilst 9 implies that the sub-indicator carries nine times more importance with respect to that which is being compared (Saaty, 2008). AHP uses the required weights, and assigns a weight to each criterion against each of the other criteria. The selection of the variables (criteria) considered to be more significant for each study area based on the experts knowledge on a given phenomenon, here sinkholes.

Table 2.6 Saaty's scale for pairwise comparison weight assignment (Saaty, 2008)

Definition	Weight
Equal	1
Equal to moderate	2
Moderate	3
Moderate to strong	4
Strong	5
Strong to very strong	6
Very strong	7
Very strong to extreme	8
Extreme	9

In summary, the following procedure is used for the AHP method:

- Defining scenarios which describe the problem
- Defining effective factors and categorizing them into homogeneous categories
- Structuring the categories into a hierarchical structure
- Creating the matrix for pair-wise comparisons
- Determination of the final way using expert choice software
- Applying consistency and sensitivity analysis over the AHP results
- Hazard evaluation

Research done by Skilodimou et al. (2019), presents a multi-hazard mapping for three hazards (landslide, flood, and earthquake) using the analytical hierarchy process (AHP) approach. The authors used the AHP to evaluate the relative importance between the three hazard maps. They assigned the larger relative importance to the seismic hazard because the study area not only experiences frequent seismic activity but also has a significant history of earthquakes that have triggered secondary events. The landslide hazard ranked second in their relative comparison because the past landslide occurrences within the study area are much more numerous than the flood occurrences. While this research provides a framework for multi-hazard assessment, it primarily maps each hazard landslides, floods, and earthquakes – independently, using distinct contributing factors for each. This approach leaves a gap in understanding the interrelationships among these hazards. Moreover, the reasons of assigning more weights for a hazard for instance earthquake is vague and based on qualitative assessments.

There are several disadvantages to this approach, which include having a high degree of subjectivity requiring knowledgeable experts to be aware of the hazard and the study area (Galve et al., 2009a; Gutiérrez, 2016). The reliance on expert judgement for determining the importance of each variable introduces

variability in the process. As a result, comparing the effectiveness of weightings, which were chosen by one expert against another, becomes problematic, leading to difficulties in achieving consistent and reproducible results (Qiu et al., 2020). Moreover, using AHP for multi-hazard assessments raises questions like why one hazard should receive a higher score than another. Simply referring to a hazard as a primary hazard, which can trigger others, cannot be quite meaningful since their return periods are significantly different. This variation in return periods introduces a complexity in comparing hazards directly. For instance, earthquakes might have longer return periods but potentially catastrophic impacts, while floods could be more frequent yet with variable severity. The other limitations also can be the availability of data on the conditioning factors, some of which are difficult and expensive to obtain due to their subsurface nature. A further constraint should be added to this list is a high degree of simplification in grouping and indexing procedures. Yet, the benefit of the AHP technique is its ability to easily break down a problem into a hierarchy (Taheri et al., 2015).

Spatial pattern analysis techniques: Early sinkhole hazard maps primarily depicted the spatial distribution of sinkholes or transferred information about the density of known sinkholes within specific regions (Parise et al., 2018). Numerous research has been done based on spatial analysis concepts in which the pattern of sinkholes determines the hazard or susceptibility level of a given region (Panno et al., 1994; Dalgleish et al., 1985; Doctor et al, 2008; Galve et al., 2008). The basis for creating such maps is the assumption that in areas where active sinkholes have developed there is a greater possibility that a new sinkhole will form (Brezinski, 2007; Zhou et al., 2003). In this approach, researches are interested in discovering and identifying probable patterns in existing sinkholes, which can help them, examine sinkhole potential behaviours. To provide such maps, sinkhole density, neighbouring effect, different interpolation techniques have been used (Galve et al., 2009a, Galve et al., 2009b).

i. Nearest neighbour analysis (NNA): It is a simple approach that measures the distances between sinkholes and their nearest neighbours. It gives a number called the Nearest Neighbour Index (NNI) Clark and Evans, 1954; Magdalene and Alexander., 1995; Hyatt and Jacobs, 1996; Zhou et al., 2003). This approach hypothesizes that new sinkholes tend to form in the vicinity of existing sinkholes (Galve et al., 2009). Therefore, it seems reasonable to consider that these areas have favourable conditions for the development of sinkholes. Usually, this model is derived from sinkhole inventories. The highest hazard zone corresponds to the minimum distance and the lowest to the maximum distance (Galve et al., 2009). The spatial distribution of many geologic features in a space can be treated as a pattern of points on a map. These patterns can be sorted into three groups: 1) clustered, 2) random, and 3) regular. Magdalene and Alexander. (1995) analyzed the distribution of sinkholes in Winona County, Minnesota using nearest neighbour analysis and concluded that the sinkhole distribution is clustered. They also noted that new sinkholes do not develop clustered around old, existing sinkholes, but around other newly developed sinkholes.

Applying NNA method, Hyatt and Jacobs (1996) observed that the distribution of the sinkholes in Albany, Georgia in the US, tend to be more in clustered which then they used this information for further prediction analyses. NNA analysis provides an objective method for distinguishing among these spatial distributions. This is done by comparing the average observed distance between features and their nearest neighbour to an expected or theoretical average distance between nearest-neighbour points in a distribution created through a random process (Clark and Evans., 1954).

The expected values for the various types of distribution are dependent upon the density of points (ρ) within the study area. If the mean of the measured nearest-neighbour distances (dobs) approaches zero, then the nearest-neighbor index (R) will approach 0.0, indicating tight clustering within the point distribution. If the mean of the nearest-neighbor distances (dobs) approaches the theoretical maximum value (duni), then the nearest-neighbor index (R) will approach 2.15 (1.0745 / 0.5), indicating a uniform distribution of points. For random point distributions the nearest-neighbor index (R) will have a value of 1.0.

The main limitation of this approach is that it is fully dependent on patterns in the data, without any specific "statistical model" that must be evaluated. In fact, the nearest neighbour methods do not rely on any "models" at all but instead simply use the existing data to "predict" new observations (Hachcham, 2023).

ii. Kernel Density Estimation: This method creates a smooth density surface from point data (i.e., sinkhole locations). It helps visualize areas of higher sinkhole density and can assist in identifying potential sinkhole-prone zones. The disadvantages of these maps lie in several factors. The main one is the matter of accuracy and reliability of the produced maps, which not only an exhaustive sinkhole inventory is needed (Perrin et al., 2015) but also the sinkholes need to be visible. This can pose a challenge on also the spatial prediction of other areas, where the sinkholes have not yet reached the surface. Moreover, they may not be feasible in situations where underlying geological feature constraining sinkhole formation is highly directional (e.g., linear) (Galve et al., 2009).

2.4.6c Quantitative methods

'Quantitative methods' is a broad term, which encompass both traditional statistical-probabilistic analyses and advanced Machine Learning algorithms (ML). These methods calculate the statistical relationship between sinkhole occurrences and a set of causal factors (Doctor et al., 2008; Galve et al., 2009). We then classify the traditional statistical-probabilistic analyses into two broader classifications; **Bivariate statistical analysis (**Frequency Ratio (FR), Statistical Index**)** and **Multiple regression analyses (**Weight-of Evidence (WOF) (Ozdemir, 2015; Perrin et al., 2015), Logistic Regression (Ciotoli et al., 2016)). The most commonly applied ML algorithms in this context are Decision Tree (DT), Random Forest (RF). In the following sections, each category and its main methods will be further explained.

Traditional statistical-probabilistic methods are common in sinkhole hazard and susceptibility modelling (Doctor et al., 2008). Galve et al. (2008), have done a susceptibility zoning of sinkholes in the Ebro Valley evaporate karst applying the Favourability Functions approach. The sinkhole inventory and 14 controlling factors were used as input data. In another research, Galve et al. (2009) applied probabilistic approach with the Favourability Functions method, but this time they went a step further by constructing a hazard model in addition to susceptibility models. The Favourability Functions approach is a statistical analysis of the relationships between the dependent variable (in the study is sinkhole occurrence) and independent variables (contributing factors). It involves computing likelihood ratios, which quantify how strongly each factor is associated with the presence of sinkholes. To provide hazard model, the frequency of new sinkholes that appeared in each susceptibility class within a given timeframe (annually) was considered.

There are other studies, which have applied statistical analysis of weighted factors. These studies typically involve integrating multiple datasets, with each representing different hazard factors, and adjusting their

weights based on statistical analyses or expert input to create a hazard assessment model (Todd and Ivey-Burden., 2016; di Santolo et al., 2016). For analysing factors contributing to sinkhole formation in Florida, recently Nam et al. (2020) statistically analysed various factors like hydraulic head difference, soil permeability, and overburden thickness. In their method, they also integrated geo-mechanical factors such as internal soil erosion and overburden soil stability, into a quantitative analysis.

Bivariate statistical analysis: here are three main models in this class. The first one is the frequency ratio (FR) method is a density-based models by which determines the sinkhole hazard index from the sum of the frequency ratio of each contributing factor (Kim et al., 2022; Odzemir, 2015). The value greater than 1 indicates a strong relationship while the value less than 1 represents weak correlation. As a traditional model, FR is popular mainly due to the simple and rapid calculation process (Kim and Nam, 2017).

Multiple regression analyses: one of the most widely used for sinkhole mapping method is *Logistic Regression (LR)* analysis (Zhu and Huang, 2006) which is an extension of the ordinary linear regression model (Kim and Nam, 2017). Simply put, it is a type of regression analysis used for predicting the outcome of a categorical dependent variable based on one or more predictor variables. The result or outcome is usually binary meaning whether a hazard is likely to occur or not in each location (Yes/No, True/False, Presence/Absence). One recent research done by Subedi et al. (2019) compared LR model with AHP model for mapping sinkhole susceptibility. The study results revealed that LR was more accurate than AHP. Particularly, LR showed better performance at predicting areas where sinkholes were less likely to occur and where they were more probable. Unlike AHP, which tended to estimate a larger area as highly susceptible to sinkholes, LR provided a more nuanced view. An important finding was that the proximity to sinkholes significantly influenced sinkhole hazard in the LR model.

In another new study, Kim et al. (2022) produced a sinkhole susceptibility map using two statistical models namely frequency ratio (FR) and logistic regression (LR) for east central Florida (ECF). Six criteria, contributing to sinkhole formation, were selected including annual and monthly rainfall, soil erosion, mechanical stability, distance to karst features, geology, and human activity. The correlation of each factor and the sinkhole density were computed. The result showed no clear relationship between average annual precipitation and sinkhole density in the study area since there is no significant difference between the higher and lower amount of annual precipitation. Nevertheless, the investigation of seasonal rainfall variation demonstrates that May is the peak and June–July exhibits a relatively high frequency of sinkholes. The study output demonstrated that the LR model outperformanced FR in delineating high-susceptibility areas for sinkholes. It was better at reflecting hydrological factors, such as recharge and head difference, in the model. The authors suggested that LR may provide a more nuanced and accurate understanding of sinkhole susceptibility in areas similar to the study region.

The Weights-of-Evidence model is another commonly used methodology in sinkhole hazard mapping. It is a statistical approach rooted in Bayesian probability theory (Oh and Lee, 2010; Perrin et al., 2015). It involves quantifying the spatial correlation between various factors. The main principles of the model are as follows: This model uses dichotomous (binary) maps to represent the presence or absence of certain factors. For each factor, two types of weights are calculated - W+ and W−. W+ is the weight when the factor is present and W− when the factor is absent. The obtained contrast value for each factor comes from the difference between W+ and W−. This value helps in understanding the spatial relationship between the factor and the event occurrence. Additionally, the model tests the conditional independence of each pair of factors to ensure that they contribute uniquely to the predictive model (Oh and Lee, 2010). However, while

these traditional models are effective, they have certain disadvantages compared to advanced machine learning methods: 1) they rely on statistical assumptions and may not capture complex interactions between variables as effectively as machine learning models. 2) they primarily use binary (presence/absence) predictors, limiting its ability to incorporate continuous or multi-class variables directly. More importantly, unlike machine learning models that can learn and adapt from new data, these models lack this dynamic adaptability and requires manual recalibration.

2.4.6d Machine Learning methods (ML)

Spatial data mining is a critical sub-field within the broader domain of machine learning and data science. It focuses on extracting valuable insights/patterns from a large number of spatial data (Gao and Alexander, 2008). This process aims to find and discover complex spatial patterns and relationships that may be hidden within extensive spatial databases. These patterns can encompass a wide range of phenomena, such as the distribution of disease outbreaks, urban planning trends, environmental changes, or natural hazards.

Originally developed in the field of Artificial Intelligence (AI) in 1990s, ML has become popular in knowledge discovery and data mining (Gao and Alexander, 2008). It enables computer systems to automatically learn and improve from a set of datasets without being directly programmed (Taheri et al., 2019; Tehrani et al., 2022). Machine learning algorithms use statistical methods to examine large datasets and discover hidden patterns within them which otherwise would be impossible to detect (Linardos et al., 2022). These patterns are then used to make predictions or help make choices about new data.

The main goal of machine learning is to create algorithms that can progress and improve over time by learning from data. This ability to learn and adapt allows machine learning to improve its performance, making it extremely useful in various areas like recognizing images, understanding language, suggesting things, and finding patterns in spatial data (Pourghasemi et al., 2020; Tehrani et al., 2022). ML is thought to excel in predictive and decision-making tasks compared to humans. It identifies patterns in input data to create rules for generating desired outcomes. Currently, ML might not surpass human abilities in various domains, but its rapid advancements suggest it could revolutionize prediction, decision-making, and impact fields like disaster risk assessment in the future (Linardos et al., 2022; Tehrani et al., 2022). The most common ML algorithms used to model and forecast sinkhole hazards include Decision Trees (DT), Random Forest (RF), Naive Bayes (NB) and Logistic Regression (LR).

a) Decision-tree model

Simply put, a decision tree is a machine learning method that helps a computer to make decisions or predictions. The decision tree model is applied in spatial data mining for discovering spatial patterns that exist within extensive spatial databases. In natural hazard mapping, this method is used to quantify the map-making process and reduce the potential for subjective biases (Gao and Alexander, 2008). Gao and Alexander (2008), produced a sinkhole probability map assessing four elements namely sinkhole density, neighboring effects, bedrock geology and depth to bedrock in south-eastern Minnesota and north western Iowa. The decision tree model was developed using the collected data. This involved defining criteria and thresholds for different variables that influence sinkhole formation. Each variable formed a “node” in the

tree, where a decision (or classification) was made based on the value of that variable. For instance, a node might classify an area as high risk if the depth to bedrock is shallow. The decision tree classified different regions based on the set criteria. Each classification path in the tree led to a conclusion about the sinkhole probability in that area (e.g., high, moderate, low probability).

However, this method has a few limitations (Ben-Gal and Trister, 2015):

Overfitting: One of the most critical drawbacks of decision trees is their tendency to overfit the data. This means they can create a model that fits the training data very closely, including the noise and errors, leading to poor performance on unseen or test data.

Limited extrapolation ability: They cannot extrapolate trends which are not in the training data. Decision trees are inherently limited by the range of the input variables they were trained on, therefore making predictions/forecasts for values outside this range cannot be reliable.

Handling of continuous variables: Decision trees may struggle with continuous variables. They typically handle these by discretizing them into categories, which can lead to loss of information and model accuracy.

Complexity with large datasets: For large datasets, especially those with many features (high dimensionality), the decision tree can become overly complex. This complexity can make the model difficult to interpret and increase the risk of overfitting.

b) Random Forest (RF)

An extension to decision tree is the Random Forest algorithm, originally proposed by Breiman in the late 1990s and early 2000s (Breiman, 2001). It operates by constructing a multitude of decision trees (ensemble) at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees (Piao et al., 2022). Each decision tree is built using a random subset of the training data and a random subset of features. This randomness helps to reduce overfitting and increases the diversity of the individual trees (Soroush et al. 2018). FR has recently gained widespread popularity for multi-hazard analysis due to its rapid analysis speed and excellent performance (Pourghasemi et al., 2020; Piao et al., 2022).

Elmahdy et al. (2022) aimed to develop a susceptibility map for land subsidence and sinkholes in Al Ain, UAE, and analyse the spatial relationship between these phenomena and fault zone trends in areas of intensive human activity. The study employed an integrated approach using **Frequency Ratio (FR)** and **Random Forest (RF)** models. The FR model was applied to analyse the relationship between the locations of land subsidence and sinkholes and eight essential conditioning factors (CFs) including topography, lithology, distance to faults and land use. Then, the values of the CFs were employed as inputs to the RF model used to generate a land subsidence and sinkhole susceptibility map. The map showed varying susceptibility levels across the study area.

c) Bayesian Logistic Regression (BLR)

Bayesian Logistic Regression (BLR) is a form of logistic regression that includes Bayesian methods. BLR is used to model binary outcomes (two categoried). In case of natural hazards, it is used to predict the occurrence of a given hazard based on various factors (Nhu et al., 2020). In traditional logistic regression, the model estimates the probability of a binary response based on one or more predictor (or independent) variables. It is typically used for classification tasks (Nhu et al., 2020). In logistic regression, the odds of the target variable are modelled as a linear combination of the predictor variables. On the other hand, the Bayesian approach to logistic regression involves updating the probability estimates as more data becomes available.

In a recent attempt to sinkhole susceptibility mapping, Taheri et al. (2019) did a comparison between Bayes-based machine learning algorithms (Bayes Net (BN), Naïve Bayes (NB), Bayesian Logistic Regression (BLR) and Logistic Regression (LR). This study aimed to evaluate the performance of these models using both training and validation datasets. The research revealed that the Bayes Net (BN) model outperformed the others having a strong fit to the training data. The NB model was superior in correctly identifying sinkhole and no-sinkhole areas, as shown by its high sensitivity, specificity, accuracy. In the validation phase, all models demonstrated strong predictive capabilities. However, the Bayesian models (BN and NB) generally outperformed the BLR and traditional LR models in predicting sinkhole susceptibility.

2.4.7 Theme 6. Multi-hazard mapping/ modelling in karst focusing on flood and sinkholes

Research on multi-hazard assessments has primarily focused on the hazards that have the potential to turn into a disaster such as earthquakes, floods, landslides, volcanic eruptions. According to UNISDR, (2017) the disaster is a "*serious disruption of the functioning of a community or a society involving widespread human, material, economic or environmental losses and impacts, which exceeds the ability of the affected community or society to cope using its own resources*".

Over the past a few years, researchers chiefly focused on multi-hazard-forming areas such as mountainous, coastal, and volcanic areas which potentially could lead to a complex interrelationships of hazards (this was discussed more in **Chapter 1**). These events have a high profile due to their immediate and often dramatic impact on human settlements and the environment. Consequently, it comes as no surprise that karst regions have been almost overshadowed by the visible importance of the other areas. As it already discussed in the **Section 2.3**, karst areas are where the substrates are subject to solution and erosion of soluable bedrock. It takes geological time period for the soluble rocks to dissolve with the underground water finally leading to underground voids. This is a hidden hazard and no one can see the progressing grow of the underground voids until they collapse leaving economic damages and human life fatalities. So far, in literature, studies on the multi-hazard mapping specifically targeting the interrelationships between karst hazards have been overlooked. There are attempts to understand the relationships between specific environmental hazards such as drought and sinkholes. Linares et al. (2017) assessed the potential relationships of droughts and the associated declines in the piezometric level of karst aquifers, and the development of collapse sinkholes.

This study established a significant correlation between drought periods and increased sinkhole activity. A detailed sinkhole inventory, constructed using historical aerial photographs, ortho-images, and field surveys, reveals temporal clusters of sinkhole formation that correspond with drought periods. These types of studies are useful to understand the different triggers for sinkholes. However, multi-hazard mapping in karst setting is still limited.

2.4.8 Theme. 7 Others

Not all the aims of sinkhole studies are along with hazard/risk assessment in which sinkholes are viewed as a collapse that destroys buildings and roadways. There have been many other studies in the literature, which have looked at karst sinkholes differently (Liu et al., 2019). In my database all the research done on sinkholes and their environmental, economic, wildlife impacts were categorised as theme "Others". Since this theme is not the focus of this research, I will briefly give an overview of this theme which was done by reviewing the main research topic on this theme.

A sinkhole can be investigated as a natural drainage point into a subsurface conduit system, which then raise concerns on groundwater pollution in way that they are able to transmit contaminants into aquafers without adequate filtration (Chen, 1993; Lindsey et al., 2010). Water from the surface that flows into sinkholes can quickly reach the groundwater, often within minutes. As it travels, it can carry a range of toxins and contaminants picked up from urban areas like streets and parking lots, as well as from industrial sites and landfills (White 1988; Ford and Williams 2007).

Karst aquifers are valuable resources in terms of providing drinking water. Almost half of the drinking water in countries such as Slovenia, Croatia and Austria and over 10 million people in Florida are dependent on karst water sources (Parise et al., 2018). Therefore, due to the increasing concerns of groundwater pollution, to understand and protect the underground water in karst regions, scientists often include a detailed mapping of sinkholes in their studies. Therefore, numerous studies have focused on the crucial role of sinkholes in the hydrologic regime and their environmental significance in water resource management (Petersen and Vondracek, 2006; Hackley, 2007; Lindseyet al.2010)

In some settings, sinkholes also can be viewed as an entrance to a cave being home to wildlife species. In ecological studies, sinkholes are crucial in maintaining biodiversity known as key habitats for scientifically important plant species (Bátori et al., 2014; Liu et al., 2019). Moreover, karst features like caves served as shelters as well as burial and sacred places for our ancestors, which they can called 'supernatural heritage' (De Waele et al., 2011). In another study done by Klopmeier et al (2018), impacts of sinkholes on wildlife mortalities was discussed. The findings showed that beyond the fact that sinkholes cause damages to human lives and properties, they also cause wildlife mortalities. Animals can be trapped into the opened sinkholes or fall into them. Additionally, it is interesting to know that some well-developed sinkholes are being used as touristic spots

2.5 The main previous research in mapping karst hazard done in the Val d'Orléans

In the Val d'Orléans region, numerous studies and research projects have been conducted over the past decade. In the following paragraphs, I have provided a chronological summary of the key studies in the area. I have divided the literature into two parts: 1) studies done before the 2016 multi-hazard event and 2) studies done after the occurrence of the 2016 multi-hazard event. A brief summary of these research is provided in the following paragraphs. In the following, I delve into detail of each study discussing the scope, the method and the criteria used in those.

Cerema (2014) conducted a study on the distribution of karstic events to identify factors influencing their occurrence, thereby enhancing the understanding of areas susceptible to karstic hazards. The methodology for developing the karst hazard susceptibility map centred on five key criteria regarding to the karst hazard locations: proximity to the Loire River, density of sinkholes, thickness of alluvium, presence of paleochannels, and the presence of drainage axes in the Beauce Aquifer. These criteria were selected based on specific hypotheses. The chosen criteria, hypothesis, and final results are summarised in **Table 2.6**. This work has encouraged further research to collect not only a more exhaustive list of sinkhole locations but also to document their other characteristics (such as size, depth, and date of appearance), which are crucial for hazard evaluation.

Table 2.6 A summary of the Cerema (2014) research result.

Criteria	Hypothesis	Results
Proximity to the Loire River	Closer proximity to the Loire River correlates with a higher frequency of karstic events.	Half of the events occurred within a 1 km band from the river. There is a higher event frequency up to 1500 meters from the Loire. This indicates a significantly higher karstic hazard near the Loire.
Density of Sinkholes	Areas with a higher density of sinkholes are more prone to future karstic events.	Three-quarters of the events are located within 300 meters of another event, and less than 10% of the events are more than 800 meters apart. This clearly indicates a concentration of karstic hazard in certain areas.
Thickness of Alluvium	Thinner alluvial cover leads to a greater karstic hazard.	The occurrence of karstic events is proportionate to the extent of the valley covered by specific alluvium thicknesses. For instance, about 70% of events occur in areas with 5-10 meters of alluvial cover, precisely because this thickness range covers about 70% of the valley's surface. Therefore, it is not a significant factor

Presence of Paleochannels	Paleochannels influence the occurrence of karstic events.	Only 13.6% of events are on these paleochannels, suggesting limited relevance of this criterion for overall karstic hazard evaluation in the Val d'Orléans. Although some paleochannels showed higher events, their proximity to the Loire suggests an overlap with the criterion related to distance from the Loire.
Presence of Drainage Axes of the Beauce Aquifer	Drainage axes from the Beauce limestone aquifer correlate with the occurrence of karstic events.	Karstic events are significant even up to a kilometer from drainage axes, but their frequency decreases with distance. The widespread coverage of these axes across the valley contributes to this pattern. Due to the incomplete and imprecise mapping of these axes, they provide limited insight into the configuration of karstic voids, making them a less reliable factor for assessing karstic hazards in the Val d'Orléans.

Perrin et al. (2014) conducted a study commissioned by BRGM to map karst hazards in the Val d'Orléans region. Using an inventory of all known karstic forms, this study first delimitated the karstified areas. Utilizing the area's geological map, they developed a detailed map of karst typologies, which differentiated between two types of karsts in the region: karsts under a permeable or slightly permeable cover and bare karst areas. The susceptibility map of the area was generated by analysing three explanatory criteria in the context of karst under cover (Distance to Watercourses, thickness of low permeable and thickness of saturated layers), using the Weight of Evidence theory. The results shows a higher susceptibility for distances < 200 meters from the watercourse, the thickness of the less permeable formation plays a very significant protective role > 8 meters thickness, and areas where the limestone is not completely saturated (saturated thickness of the cover less than 0) are also less susceptible. In addition to susceptibility map, the sinkhole intensity map was also created using collapse diameter considering three intensity classes: low (diameter < 3m), moderate (between 3 to 10 m), and high (greater than 10m). This intensity map provides a visual representation of collapse intensities, derived using Sowers' empirical formula. The formula establishes a relationship between the collapse diameter and the thicknesses of different geological formations – cohesive (less permeable) and non-cohesive (permeable). The map also incorporates real-world data, comparing the predicted collapse diameters with those actually observed in documented sinkholes.

In a similar study which seems to be a continuation of the previous work, Perrin et al. (2015) produced a karst susceptibility map applying the same statistical method "weight-of-evidence theory" to select the most relevant and independent contributing factors. Weights-of-Evidence (WofE) is a statistical method used in geoscience for spatial data analysis. It works by quantifying the correlation between the presence of a specific event (like karst subsidence) and various geospatial data layers or factors. Essentially, each factor is given a weight based on how strongly it predicts the event. These weights are then combined to create a predictive map showing the likelihood of the event occurring in different areas. Despite of the previous research, here four potential factors were evaluated (thickness of low permeability and saturated layers, distance to streams, and alluvium thickness), among which the last three factors were assessed for the final integration. Binary patterns (presence or absence) were extracted for each factor and compared with sinkhole occurrences. Positive (W+) and negative (W-) weights were computed for each pattern, using

a formula involving the probability of event occurrence with and without the factor. The contrast (difference between W+ and W-) for each factor was calculated and its statistical significance assessed. The results demonstrated higher susceptibility to karst subsidence in mantled karst settings. It was found that factors like the presence of low-permeability beds in the overburden and the position of the water table significantly influenced the subsidence hazard. The study also concluded that total overburden thickness was not a major conditioning factor in the area studied.

However, while geology and hydrogeology are fundamental to sinkhole development, additional contributing factors are necessary to pinpoint the areas most susceptible to sinkholes in a region characterized by mantled karst. From a practical standpoint, the resulting map indicates that nearly the entire region falls into high hazard zones. This poses challenges for decision-makers in terms of prioritization and practical implementation. The difficulties may stem from both the selection of contributing factors giving more weight to the mantled karst while the whole area is overlaid by that and the model's potential failure to capture non-linear interrelationships.

In a much smaller area of the Loire valley (Val de Bou) covering 312.5 km², Cerema. (2017) has adopted an approach for mapping based on the density of historical karst events in relation to three criteria (proximity to the Loire River, presence of paleo channels, and water flow on the hillside). Each of these criteria was used to categorize different zones in the study area based on their risk levels. This included areas near the Loire River, within paleochannels, and in topographic depressions on the hillside. The study calculated the density of karst events in each of these zones. This involved analysing the number of events per square kilometre in different areas and comparing them to determine relative hazard levels. The results revealed that a significant majority of sinkhole events (over 80 %) occurred within a 600-meter band of the river. However, the study suggested that since paleochannels, which are old watercourses fossilized by sedimentary and are linked to the past flow of the Loire River, are mainly located in less than 600 meters from the Loire. Therefore, it cannot be a significant criterion in mapping since it overlaps with the distance to the Loire River criteria. The sinkhole density in the water flow on the hillside shows that due to the topographic configuration and the presence of an impermeable clay layer from the Burdigalian age, the coteau/hillside area is relatively protected from the influence of the Loire River. Karst events are significantly less numerous there. Therefore, the density of karst events per km^2 determine the hazard level. In summary, the areas situated in less than 600 meters from the Loire with 58.8 events per km^2 considered as very high hazard zones and the part of the hillside without topographic depressions, with a density of only 1.75 events per km² was considered low hazard.

A notable limitation of this study on karst hazards in the Val de Bou lies in the fragmentary nature of the data concerning the karst events themselves. In other words, most data collected concerns only the location of the events. This traditional method relies on heavily available historical data. Therefore, primarily descriptive, focusing on identifying patterns based on historical occurrences. It may not effectively predict future events, especially in areas without historical data.

Studies after the 2016 flooding and sinkhole event: Following the 2016 flood event in the Val d'Orléans, Noury et al. (2018) emphasizing the role of flooding in sinkhole collapses due to accelerated internal erosion during the 2016 flood event in the region of Orléans, aimed to understand the triggers of the erosion process. The study simulated the upward propagation of cavities in covered/mantled karst terrains using numerical models combining Lattice Boltzmann and Discrete Element methods. Two scenarios were simulated, considering three parameters: conduit size, soil cohesion, and pressure gradient, which influence the erosion

threshold. They concluded that intense rainfall and floods significantly increased sinkhole formation in this region with rates up to 24,000 times higher than normal. This work is considered a physical modelling research where numerical modelling and laboratory experimentation are used to study the internal erosion process in covered karst terrains, with the aim of improving the prediction of sinkhole hazards and risks.

Building on this work, Luu et al. (2019) focused on simulating the internal soil erosion during floods in particular during the 2016 floods in Val d'Orléan to underestand different erosion regimes which lead to different sinkhole shapes (hourglass and inverted bowl sinkholes). Applying hydromechanical models, three phases were simulated as follows: 1) the formation of a stable cavity within the cover material; 2) the upward propagation of a cavity leading to a dropout sinkhole; 3) the downward discharge of the granular media, called the subsidence sinkhole. They also conducted a parametric analysis of the dropout sinkhole, revealing that the collapse size increases linearly with cover thickness. An experimental model was created using sand with artificial cohesion to replicate different erosion regimes predicted by the numerical simulations. The experimental results qualitatively reproduced the processes observed in simulations, supporting the numerical findings. The results provide insights into the micro-mechanical features of internal soil erosion in a flood situation and contribute to a better understanding of sinkhole occurrences in covered karst terrains during a flood.

A Synthesizing the previous research done in the study area is provided in **Table 2.7.**

Table 2.7 Synthesizing the previous research done in the study area.

Author(s)	Year	Focus Area	Methodology	Criteria	Key Findings
Cerema	2014	Karst hazard mapping	Based on distance of the selected criteria to karst events	Proximity to Loire River, sinkhole density, alluvium thickness, paleochannels, drainage axes in Beauce Aquifer	High karstic hazard near Loire, sinkhole concentration in certain areas, limited relevance of paleochannels and drainage axes
Perrin et al.	2014	Karst hazard mapping	Inventory of karstic forms, geological mapping	Distance to watercourses, permeable cover thickness, saturated layer thickness	High susceptibility near watercourses, protective role of permeable formation beyond 8 meters, less susceptibility in non-saturated areas.
Perrin et al.	2015	Karst susceptibility mapping	Weight-of-Evidence theory	Thickness of low permeability and saturated layers, distance to streams, alluvium thickness	Higher subsidence hazard in mantled karst settings, influence of low-permeability beds and water table position

Cerema	2017	Karst hazard mapping in Val de Bou (An area smaller the Val d'Orléans)	Historical event density analysis	Proximity to Loire River, paleo channels, hillside water flow	High hazard within 600 meters of Loire, limited influence of paleochannels, hillside areas less susceptible
Studies done after the 2016 multi-hazard event					
Noury et al.	2018	Simulation of sinkhole formation during floods	Numerical modeling (Lattice Boltzmann and Discrete Element methods)	Conduit size, soil cohesion, pressure gradient	Intense rainfall and floods increase sinkhole formation rate
Luu et al.	2019	Soil erosion simulation during floods	Hydromechanical models, experimental modelling	Cover thickness, cohesion, erosion regimes	Collapse width correlation with cover thickness, insights into soil erosion in flood situations.

2.6 Discussions and conclusions

Within the current chapter, I have first provided an overview of the multi-hazards concept and its historical development. I also outlined the importance of understanding the interrelationships between multiple hazards as discussed in various studies, highlighting that recognizing these interactions is crucial for assessing the hazard landscapes. I noted that failing to understand these relationships can amplify the impact when multiple hazards occur simultaneously. The literature defines several types of hazard interrelations, which are commonly understood but sometimes referred to with different terminologies. In that regard, I summarised the most recent classifications describing the interactions as triggering, amplifying, condition changing, mutually exclusive. Further, the main existing approaches to quantify and mapping multi-hazards were discussed within a wide range of hazards in a global context. This part of the literature review was done to only discuss the multi-hazard literature. However, since the focus of the study has been on flood and sinkhole, I have done a systematic literature review on this subject.

In **Section 2.3**, I not only provided an explanation of the karst areas, but also I argued that karst hazards due their special characterizations in geology and hydrogeology should be viewed within the broader category of multi-hazard zones. I then, based on historical real world examples, elaborated the main potential multi-hazard chain that could happen in those areas trying to introduce and highlight a new concept in understanding and managing hazards in karst environments, taking into account their distinct properties and the complex interactions of multiple hazard factors.

In **Section 2.4**, in an attempt to highlight the existing gap related to multi-hazard mapping in karst areas in particular flood and sinkholes, I have done a systematic literature review in which I reviewed 241 peer reviewed publications. The construction of our database involved a process with a great attention to selecting the representative peer review papers. It started with the identification of scientific publications

through specific keywords. This was a two-step process: first, identifying relevant keywords, in which I observed variations in the terms such as “sinkhole” and “doline” applied in various regions, and secondly, scientific articles selection. Despite the terminological variances, “sinkhole” was the most frequent terminology found in the international literature to guide our keyword search.

The application of these key words was essential in sourcing our relevant articles from different web search engines focusing on English-written international literature journals with a specific emphasis on sinkhole hazard studies especially on flooding. The screening process involved a careful examination of titles and abstracts to narrow down relevant papers. During this process, I excluded duplicates or similar studies. Finally, in the classification phase, I selected 241 references, extracting key information to identify and define seven thematic subjects within the field of sinkhole studies. This structured approach enabled a comprehensive and organized literature review.

However, there are some inherent limitations in classifying the literature. They are important considerations, as it reflects on the inherent complexities of conducting a literature review. The challenges are listed in the following:

i) **Overlapping references**: This challenge arises from the multi-disciplinary nature of sinkhole studies involving a vast range of domains. Therefore, the publications might overlap and fit into several categories (themes) at the same time with the same study supporting different themes. Even though efforts have been made to place them in the closest relevant class, this limitation remains.

ii) **Acknowledging non-exhaustive literature:** In our approach to exploring sinkhole studies, particularly those integrated with flooding, I have carefully selected and categorized related scientific research from published international literature journals written in English. While this process has enabled us to construct a focused and detailed review, it is important to acknowledge that our study does not claim to be exhaustive but representative. The field of sinkhole research is vast and continuously evolving, being highly multi-disciplinary. Despite our search and categorization efforts, there might be relevant studies, particularly those published in non-English languages or in less prominent journals that have not been included. Additionally, given the rapid advancement of research in this area, recent studies postdating our review might provide new insights that are not captured in our current analysis. Therefore, our work is a significant effort to gather and synthesize existing knowledge in the field. However, it should be viewed within the context of these constraints, recognizing that it is a small part of a dynamic and expansive body of research.

While numerous research has made attempts to study sinkholes and their contributing factors, to date, few studies have perceived and assessed flooding and sinkhole as a multi-hazard event. Since globally, a shift from single to multi-hazard assessment is being encouraged by international risk communities, as already mentioned in **Chapter 1,** the present research is to provide a multi-hazard susceptibility map of the flood-triggered sinkholes emphasising multi-hazard approaches.

Two related hazards: flood and sinkholes, despite their frequent occurrence and potential for significant impact, often receive less scientific attention compared to other natural hazards and multi-hazard forming areas such as mountainous, volcanic and coastal areas. This lack is partly due to the typically lower death

tolls associated with sinkholes and their often localized, remote nature (Intrieri et al., 2018; Fleury, 2009). Additionally, the slow, subterranean formation process of sinkholes, which occurs over thousands of years and is invisible to the naked eye, contributes to their neglect in both public consciousness and scientific study. It is only when a collapse occurs, leading to dramatic, sudden openings, that sinkholes gain momentary focus. This situation underscores the need for increased attention and research into sinkhole formation and impacts, particularly in areas with karst bedrock.

Recent events underscore the need to shift from single-hazard assessments in karst areas to a multi-hazard approach that recognizes the simultaneous potential for both flooding and sinkholes. This change is crucial because traditional, isolated strategies are inadequate for these complex, interconnected natural phenomena. An integrated perspective is essential for effective hazard management and mitigation in these geologically diverse regions.

As the literature discuss, ML algorithms all excel at forecasting and dealing with big and dimensional data. They are powerful in finding subtle patterns in data, which are not linear but non-linear relationships being more advanced at handling and interpreting complex, multi-dimensional data. These patterns and relationships among data can be overlooked in traditional methods. Throughout the literature the application of these new ML predictive models has been explored in different geoscience fields including sinkholes. They yield hazard susceptibility evaluations that are both less biased and consistent, with the potential for reaching high spatial resolution. However, more algorithoms need to be explored and examined.

In advancing my research, I am now turning towards the application of the Maximum Entropy modelling approach, commonly referred to as MaxEnt, to model flood-triggered sinkholes, an area that despite extensive studies globally and regionally in both multi-hazards and sinkhole hazards, remains largely unexplored in terms of integrated assessment. This integration is pivotal in karst terrains, where the unique geological features considerably influence both flood dynamics and sinkhole formation. Karst regions are particularly susceptible to these hydro-geo hazards, yet the interplay between flooding and sinkhole occurrences has not been sufficiently investigated and mapped. Moreover, despite of the past occurrence of this multi-hazard event in Val d'Orléans, the region lacks this integration in a form of multi-hazard susceptibility map, which demonstrate the area susceptible to both hazards. This adds geographical specificity to the research gap.

Chapter 3. Methodology and Case study

3.1 Introduction

This chapter presents a detailed research design and methodology focused on applying the Maximum Entropy (MaxEnt) algorithm to develop a multi-hazard susceptibility map. It also introduces the case study. The organization of this chapter is as follows:

Section 3.2 delves into the foundational principles of the MaxEnt algorithm, tracing its historical roots and exploring its wide-ranging applications across various disciplines, notably in natural hazard science. This section outlines the theoretical foundations of MaxEnt, its advantages in handling high-dimensional data, and its effectiveness in modelling the distribution of natural phenomena, such as species, stars, and hazards, using presence-background data.

Section 3.3 introduces Val d'Orléans as the case study area. The section describes the unique geographical, geological, and hydrogeological characteristics that make this region particularly susceptible to natural hazards like sinkholes and floods. It also discusses past occurrences and the interrelationship between different hazard types, highlighting the need for advanced predictive modeling to better understand and mitigate these risks.

Section 3.4 outlines the data collection processes used to apply the MaxEnt for mapping the susceptibility of flood-triggered sinkholes in Val d'Orléans. The section begins by describing the implementation of MaxEnt to translate theoretical models into actionable hazard maps. It includes a step-by-step account of selecting appropriate environmental variables, preparing and processing the data, and configuring the MaxEnt software for optimal model performance.

Section 3.5 provides detailed explanation of preparing data for the MaxEnt application to model flood-triggered sinkhole occurrences. This includes formatting data to meet MaxEnt's specific requirements. The model building phase involves tuning of parameters to balance model complexity and accuracy, including setting regularization parameters and selecting optimal threshold values. Lastly, I discuss data partitioning strategies, emphasizing the unique considerations for modelling natural hazards like sinkholes compared to biological species. This section outlines the essential steps and methodological rigor involved in preparing and modeling data with MaxEnt in our research.

3.2 Maximum Entropy Modeling (MaxEnt) background

Initially delineated by Jaynes in 1957, MaxEnt principle has found a huge amount of use in machine learning and in the modelling of real-world processes. Edwin Jaynes (1982) in a very brief statement defined MaxEnt as making statistical inferences based on incomplete information through a set of known information. In fact, Jaynes formally introduced the Maximum Entropy principle in the context of statistical mechanics and information theory in the 1950s (Jaynes, 1957a; b). His work was foundational which later on, he elaborated on the principle further in the 1980s (Jaynes, 1982).

However, it is important to acknowledge that ecologists started using this principle independently, particularly for predicting species distribution and modelling ecological patterns. This application became especially popular in the early 21st century by the work of Phillips et al, (2006). Ecologists adopted MaxEnt to estimate probability distributions of species occurrences based on limited or incomplete data, making it an important tool in conservation biology and biodiversity research. Therefore, it has been known as an *ecological* model. The arrival of MaxEnt model revolutionised the field of ecological niche modelling (Sillero et al., 2021). The potential distributions are forecasted by presence-only species within the region where conditions fit for the species presence.

Before the introduction of the Maximum Entropy approach in ecology, particularly in species distribution modeling, researchers faced challenges in modelling the probability distributions of species. One of the main issues was the need for (geo-referenced) data on both species presence and absence to accurately model distributions (Sillero et al., 2021). Absence locations are geographic locations where a species is not observed. Nevertheless, not observing a species in a location does not necessarily mean the species cannot exist there; it might simply mean the species was not detected during the time of observation. Collecting reliable absence data is challenging. It requires extensive fieldwork and often leads to uncertainties because a species' absence could be due to various factors, including the real absence of the species or the failure to detect it during sampling. Since absence data could be unreliable or hard to obtain, models based on this approach could be biased or less accurate. Here is where MaxEnt comes practical allowing modelling distributions using only presence data. The growing access to substantial amounts of occurrence data has made this model a desirable method for more exploration into not only ecological studies, but also other fields. Researchers from different disciplines, who are interested in forecasting a particular system (social, economic, natural environment, climate change, and natural language processing), have widely used MaxEnt for various applications.

However, I would like to mention that in the context of natural hazards this "absence data" functions differently in comparison with ecological applications. One of the main reasons behind this is that for natural hazards data typically consists of historical records of where and when hazards occurred and their severity. The absence of a hazard in a location generally means it has not been recorded there yet, but it does not eliminate the potential for future occurrences. In fact, for natural hazards the aim is slightly different because by applying the method, modellers aim to have a forecast the areas where hazards might occur in the future, rather than focusing on areas of absence.

The Maximum Entropy model is a statistical–probabilistic machine learning method, which excels at prediction of high dimensional data (Phillips et al., 2006). The goal of MaxEnt is to find the probability distribution (p) of target occurrences over the set locations (X) within the study area. MaxEnt does it by finding patterns in existing data (so called presence-data) and applies it to new data. This is discussed more in the next **Section 3.2.1**. Predicting where target occurrences (species, stars, hazards., etc.) are found often uses both presence and absence data, with general methods available for both. While presence data is common and easily accessible, absence data is scarce, making methods that rely only on presence data valuable for species distribution modelling (Phillips et al., 2006).

It is important to clarify a common misconception about MaxEnt. It has often been incorrectly categorized as a presence-only method, while recent literature has explained that this is a presence-background method in which the available environmental conditions are compared in the study area namely background with the conditions used by a given phenomenon (plants, hazards., etc.) as represented by its occurrence (Sillero et al., 2021)

MaxEnt has proved a powerful framework for probabilistic modelling leading to highly accurate predictions in many cases. It is a model that offers many advantages and a few drawbacks. The popularity of MaxEnt model lies in several main reasons as follows:

Advantages from machine learning perspectives (mainly exctraceted from Phillips et al., 2006):

1) Unlike other machine learning algorithms, which usually happen in a black box, the model has a clear mathematical definition and is therefore open to analysis (Phillips et al., 2006).
2) It is a powerful algorithm due to its regularization techniques and the use of cross-validation that makes it powerful in balancing overfitting (it is discussed in **Chapter 4**).
3) Availability: It is an open source software and is easily available to public users.
4) Cost effective: Nowadays, with the technological advancements, the environmental data are being more and more accessible to everyone.
5) Suitable for a general prediction for the applications involving presence-only datasets (Phillips et al., 2006)
6) It only needs a set of known data coupled with some environmental variables for the entire area of interest.
7) It allows the use of diverse features selected by the user (Mohri et al., 2018). Both type of environmental data can be used, *continuous* and *categorical* data.
8) Powerful for multivariate analyses: As a machine learning method, MaxEnt ables to iterate the data thousands times which it enhances the model's predictive accuracy.
9) Efficiency: MaxEnt enables simultaneous analysis of multiple factors.
10) Output is continuous, allowing fine distinctions between suitability of different areas.

However, it is important to be aware of the model limitations that may affect the model accuracy. One of the main limitations is that a complete sinkhole inventory is required to increase the accuracy of the models, which is challenging due to the difficulty of obtaining a comprehensive inventory.

3.2.1 The principal of MaxEnt

When there is incomplete and unknown information about a probability distribution of a system, according to Jaynes (1957) the best estimation is that we choose the distribution that maximizes entropy subject to the constraints (environmental variables) that we know (Jaynes, 1957; Phillips et al., 2006). This concept is known as the maximum entropy principle (Phillips et al., 2006). It is a totally natural procedure for determining distributions which match the information one brings to the table without introducing any extra assumptions. In other words, the idea is to estimate a target probability distribution by finding the distribution with maximum entropy (most spread out or uniform) subject to a set of constraints that represent our incomplete information about the target. This method is particularly effective for presence-only species distribution modelling (Phillips et al., 2006).

To apply maximum entropy modelling, understanding the concept of Shannon entropy is a prerequisite as it is an important concept used in the MaxEnt algorithm. Entropy as a key component in Maximum Entropy is a measure of the randomness or uncertainty of a distribution, or as Shannon (1948) describes entropy as "a measure of how much 'choice' is involved in the selection of an event". Therefore, maximizing entropy means that I are choosing the most unbiased or least restrictive distribution possible. A distribution with higher entropy involves more choices (i.e., it is less constrained) (Philippe et al., 2006). Entropy or information entropy is calculated as follows:

$$H(\hat{\pi}) = -\sum x \in X \hat{\pi}(x) \ln \hat{\pi}(x) \qquad \text{(q 3.1)}$$

Where **H (ˆπ)** is the entropy of the probability distribution **ˆπ**. **X** is the set of all possible outcomes or events. **ˆπ(x)** is the probability of event **x** occurring, according to the distribution **ˆπ**. log2ˆπ(x) is the logarithm base 2 of ˆπ(x). In particular, the entropy is highest when all outcomes are equally likely (i.e., the distribution is "uniform"), and lowest when one outcome is certain to occur (i.e., the distribution is "deterministic").

Therefore, the result will be a distribution with higher entropy involves more randomness (i.e., it is less constrained). The maximum entropy distribution that I obtain represents the most unbiased distribution that is consistent with our available information. It neither assumes anything that is not known nor ignores any available information (Jaynes, 1990; Philipe, 2006). In other words, this concept revolves around the idea that a particular probability distribution, while maximizing entropy under specific limitations that reflect our limited knowledge, serves as the basis for drawing informed conclusions. This approach aligns with all available information that is known but avoids making unjustified assumptions.

The maximum entropy principle has gained attention in machine learning, particularly with the development of efficient algorithms for finding the MaxEnt distribution. In other words, MaxEnt is a statistical–probabilistic machine learning algorithm which quantitatively estimates the probability distribution (**P**) of target occurrences over the set locations (pixels) **X** within the study area based on known contributing factors (also referred to as features in maxEnt technical language) (Phillips et al., 2006; Elith et al., 2011). Although MaxEnt is often associated with the principle of maximum entropy, it does not

directly use the Bayesian rule in its standard form. However, MaxEnt and Bayesian approaches share conceptual similarities, as both aim to infer probability distributions based on constraints or prior knowledge (Rahmati et al., 2016; Shi, 2022). MaxEnt does it by dividing the study area into locations or pixels **X**, each representing a computational unit with a probability value (**π(x)**) indicating the likelihood of a target event occurrence. The distribution **π** assigns a positive probability **π(x)** to every point **x**, and the sum of these probabilities equals 1.

The response variable (**y**) represents whether a computing unit experiences a target event (here, flood-triggered sinkholes) (**y =1**) or not (**y = 0**). To calculate the probability of a given event occurring at a particular point (**P (y=1|x)**), Rahmati et al (2016) summarises the main essential aspect of the model extracted from (Phillips et al., 2006; Elith et al., 2011) as follows:

$$P(y=1|x) = \frac{\mathrm{P(y=1)P(x|y=1)}}{\mathrm{P(x)}} = \frac{\mathrm{P(y=1)\Phi(x)}}{\frac{1}{|x|}} \qquad \text{(q 3.2)}$$

Where:

- P (y = 1|x) is the probability of a hazard at a specific point.
- P (y = 1) is the probability of event y being equal to 1, without considering any other variables. It is the marginal probability of variable y taking the specific value 1.
- P (x | y = 1) is the conditional probability of event x occurring, given that event y is equal to 1.
- P(x) is the probability of event x occurring without considering any other variables. It is the marginal probability of variable x.

The MaxEnt algorithm estimates Φ(x) as equivalent to a Gibbs probability distribution. The Gibbs probability distribution is denoted as Eq (2):

$$q\lambda\,(\mathbf{x}) = \frac{1}{Z\lambda\,(x)} \exp\left(\sum_{i=1}^{n} \lambda_i\, f_i(\boldsymbol{x})\right) \qquad \text{(q 3.3)}$$

Where:

- $Z\lambda$ (x) is a normalization constant (ensuring q_λ (x) sums to one across the study area).
- λ_i represents the vector of weights assigned to the features.

During the estimation phase of qλ (x), MaxEnt modeling seeks to find the distribution that closely follows the constraints. To prevent overfitting, it utilizes regularization, which was discussed in the previous section. Therefore, the MaxEnt model aims to discover the Gibbs probability distribution that maximizes penalized log-likelihood values.

Additionally, if there are *m* occurrences in the study area, the objective function to be maximized, which balances regularization and log-likelihood, is:

$$\Psi(\lambda) = 1/\mathrm{m} \sum_{i=1}^{m} In\ (\ \mathrm{q}\lambda\ (xi)) - \sum_{j=1}^{n} \beta j|\lambda j \qquad \text{(q 3.4)}$$

Where:

- βj represents the regularization parameter for the j-th feature (fj).

All input conditioning factors are treated as random variables in the model, following the MaxEnt algorithm.

3.2.2 MaxEnt entropy model in natural hazards

Over the past decade, the maximum entropy model has received significant attention among experts who are working on different aspect of natural hazards. This growing interest has led many scholars across the globe to apply and evaluate the performance of the MaxEnt specifically in relation to their target hazards (Azareh et al., 2019; Mokhtari and Abedian 2019; Park 2015). Among natural hazards, landslides have been the more favourite research area to apply the method on (Yang et al., 2010; Davis and Blesius.. 2015; Park, 2015; Kornejady et al., 2017; Kerekes et al., 2018; Mokhrari and Abedian., 2019; Sardooi et al., 2021; Liu et al., 2022; Boussouf et al., 2023) followed by floods (Siahkamari et al., 2018; Mobley et al., 2019), erosion (Pournader et at., 2018; Azareh et al., 2019), debris flow (Lombardo et al., 2016), forest fire (Banerjee, 2021), multi-hazards (Javidan et al., 2021; Rusk et al., 2022) and more recently, sinkholes (Bianchini et al., 2022) (**Fig 3.1**).

In recent research done in Taleghan basin in Iran, to examine the efficiency of the existing landslide models, Mokhtari and Abedian (2019) have compared the performance of seven probabilistic prediction models: frequency ratio, logistic regression, maximum entropy, and support vector machine with radial base, sigmoid, linear, and polynomial kernel functions. The validation results of the landslide susceptibility maps proved the most successful performance from both maximum entropy and logistic regression models. However, the result of another research done by Park (2015), reveals that in a comparison just between the two well-performed models (MaxEnt and logistic regression), the performance of MaxEnt modeling was rather more effective for landslide susceptibility mapping. In the mentioned study, Park (2015) tested the applicability of MaxEnt modeling producing a landslide susceptibility map in Korea. Six causal factors were selected as inputs for the model such as elevation, slope, distance from lineaments, forest type, soil drainage and lithology. The result showed that the areas with high degree of slope, weathered with excessively drained granite residuum soils were very susceptible to landslides.

Similarly, other authors who have used the MaxEnt method in different regions and hazards have yielded the successful results. Pournader et al. (2018) have tested the performance of MaxEnt model aiming at producing the susceptibility map of soil erosion in the Golgol watershed, Ilam province-Iran. With the use of the receiver operating characteristic (ROC) curve (AUC), they evaluated the model performance. The

result for the training and validation datasets (0.867, and 0.794) placed the model in the category ‘very good’. Therefore, they have successfully proved that the MaxEnt model can be used for mapping erosion susceptibilities.

In addition, Abdollahi et al. (2019) implemented MaxEnt model to map the land subsidence susceptibility of the two counties in Kerman Province in Iran. The researchers have identified several environmental factors locally contributing to the subsidence events in the region, which are as follows: slope degree, aspect, elevation, profile curvature, plan curvature, topographic wetness index (TWI), distance of drainage, lithology units, piezometric data, land use, and normalized difference vegetation index (NDVI). The final map included four susceptibility classes namely, low, moderate, high, severe and showing the ROC about 85.9%, the predictive models proved a high accuracy, which makes MaxEnt an effective prediction model.

Another example of the MaxEnt application can be found in mapping the distribution of flood hazard. Throughout the literature, there has been a number of research carried out on prediction the spatial distribution of flooding (see Eini et al., 2020; Mobley et al., 2019; Norallahi and Kaboli, 2021). Eini et al. (2020) compared the two methods: MaxEnt and Genetic Algorithm Rule-Set Production (GARP) to provide a flood risk map for the city of Kermanshah in Iran. The results demonstrated that the MaxEnt model had a better performance than the GARP model based on two common indices which are the area under the receiver-operator characteristic curve (AUC-ROC ¼ 96.76–98.32%) and kappa statistic (Kappa ¼ 0.82-0.86). These findings showed that machine-learning models provide reliable results especially for areas where data availability is a major issue.

Recently, the ability of MaxEnt model has been also explored in multi-hazard assessments (Javidan et al., 2021; Rusk et al., 2022). Javidan et al. (2021) in an attempt to provide a multi-hazard map, have applied MaxEnt assessing three natural hazards namely, flood, landslide, gully erosion in the Golestan Province-Iran. To meet this aim, they have selected 17 geo-environmental factors contributing to the aforementioned hazards plus the presence-data of the spatially settings of the hazards. The three hazards were produced individually then from their combination, a multi-hazard probability map was generated. Moreover, in a large-scale study covering the Hindu Kush Himalaya regions (608,100 km^2), Rusk et al. (2022) applied the MaxEnt model to map the pattern of the multi-hazard distribution susceptibility and the population exposure to three natural hazards (c). However, all the selected hazards were generated separately and from a sum of those the multi hazard hazard susceptibility

A new study by Bianchin et al (2022) recently explored the MaxEnt method on mapping sinkhole susceptibility in Guidonia-Bagni di Tivoli plain (Rome), Italy. According to the authors, this work is the first attempt in using MaxEnt modelling in sinkhole hazard. Eight factors were considered for sinkhole probability distribution including lithology, land use, soft soil thickness, travertine thickness, fault setting, hydrography, groundwater, and satellite InSAR data. Obtaining the AUC value (0.857), this study proved the maxent model as a good prediction method.

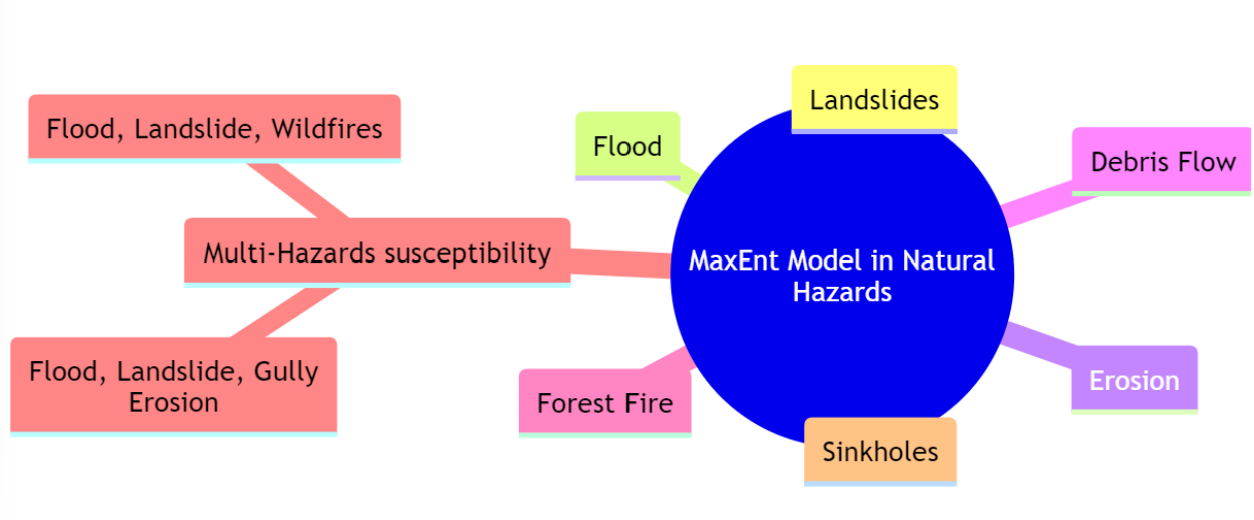


Figure 3.1 The mind map diagram of main natural hazard modelled by MaxEnt. (Source: the author; generated by Mermaid chart tool).

To conclude, while there are efforts to apply MaxEnt on various natural hazards, MaxEnt application on phenomena like flood- triggered sinkholes and multi-hazards are still very limited. The current research is an attempt to explore MaxEnt for spatially mapping flood-triggered sinkholes as a multi-hazard modelling acknowledging the interrelationship between flood and sinkholes. I use this advantage of MaxEnt in handling both categorical and continuous by introducing the flood hazard map as a contributing factor to the sinkhole spatial map. To apply the method, Val d'Orléans in France was chosen as a case study. The **Section 3.3** is dedicated to introduce the area.

However, in applying MaxEnt and other machine learning methods, a notable limitation arises from the inherent variability of model outputs when the same dataset is applied repeatedly (Phillips et al., 2006; 2017; Sillero and Barbosa, 2021). This issue can be mitigated by running the model multiple times on the same dataset and calculating the average and standard deviation of the results to evaluate if outcomes are consistent across datasets. Ideally, models should be replicated a minimum of 10 to higher numbers like 50 or 100 times, depending on the available computer time and storage (Phillips et al., 2006; Sillero and Barbosa, 2021). Yet, this crucial step is frequently overlooked in the literature.

Another critical aspect that requires attention is the issue of collinearity among variables used in the model (Sillero and Barbosa, 2021; Sillero et al., 2021). Often, modellers tend to overload the model with a wide range of variables without assessing the potential collinearity among them, resulting in outputs that may not accurately reflect the underlying contributing factors. This step is critical because high collinearity among variables could lead to unstable estimates of the model parameters, making it difficult to distinguish the individual effect of each predictor on the response variable. In addition, this overloading can result in models that are less generalizable to new data, as they may capture noise rather than meaningful patterns.

MaxEnt modelling, as a machine learning method, is relatively new and even more so when applied to natural hazards. Consequently, many studies have applied the model without proper adjustments in the

modelling process, which can result in misleading outcomes and not accurate interpretations. The novelty of this approach in the context of natural hazards means that best practices are still being developed, and there is a critical need for researchers to carefully tailor their models to the specific requirements of their study area and data. It is mainly due to the inherent differences in natural hazard nature and ecology.

3.3 Case study: the region of Orléans (Val d'Orléans)

In this section, I begin by introducing the study area with a broad overview and general background of its geography (**Section 3.3.1**). Following this, I provide information on the area's geology (**Section 3.3.2**) and hydrogeology **(Section 3.3.3**). This leads us to **Section 3.3.4**, where I explain why I consider the area as a multi-hazard zone, supported by real examples of incidents in the region. I then discuss the karst system (**3.3.4a**), including sinkhole events (**3.3.4b**), and the history and causes of flooding in the area (**3.3.4c**). Finally, I conclude by exploring the evolution of flood risk management and urban planning in the Val d'Orléans (**Section 3.3.4d**).

3.3.1 General background

The Val d'Orléans, located in the Loiret department within the Centre-Val de Loire region of France, encompasses an area of 260 km² and includes 31 municipalities. This region is represented by the floodplain of the Loire River and has approximately 70,000 inhabitants. The geography of Val d'Orléans is characterized by an extended basin (a 16,917-hectare floodplain on the left bank of the river) that stretches for approximately 35-36 km from east to west and spans 7 km from north to south (the width) (Blanchard, 1903; Auterives et al., 2014; Dournel et al., 2019) (**Fig 3.2**). This basin is bordered by the Forest of Orléans to the north and the plateau of Sologne to the south. Flowing through the entire length of the basin is the Loire River (the longest river in France with 1012 km length, which meanders with sinuous curves. It extends over a length of 33 km, from Sigloy, in the east, to the confluence of the Loire-Loiret in the west. Its average width of 6 km can reach 8 km in Sandillon. The river has eroded the surrounding hills and deposited sediment in the basin, creating a flat and fertile plain.

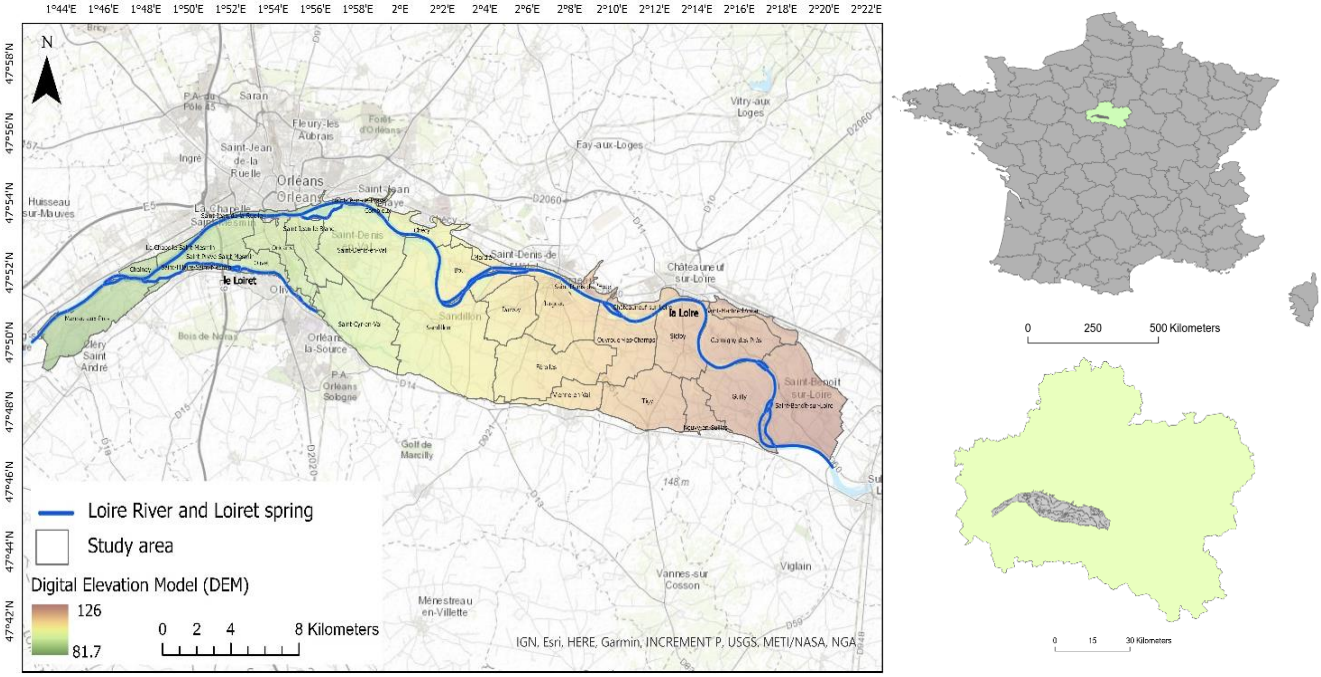


Figure 3.2 Geographical location of the study area situated within the Loiret department, France.

The alluvial deposits in the basin are composed of a variety of materials, including gravel, sand, silt, and clay, which were transported by the river from upstream areas (Blanchard, 1903). In addition to the Loire, which delimits the northern part of the valley, the surface waters of the Val d'Orléans are drained by two main rivers, the Dhuy, a tributary of the Loiret, and the latter that flows into the Loire at the tip of Courpain. The area is known for its karstified rocks covered by the alluvial deposits (in the Val and on the Plateau Sologne). Built on a highly karstified marly-limestone substratum, the two banks of the Loire Valley are covered by the Beauce limestone, which is overlain by other Tertiary sediments at shallow depth. The limestone is mantled by Quaternary alluvium (5 to 15 m thick). The pH level of the Loire River is mildly acidic thus leading to a chemical reaction with the ground of its surroundings resulting in a karst network (Gutierrez and Binet., 2010). Ground collapse and sinkholes regularly occur in this vicinity. These topographic depressions were also the cause of unforeseen expenses during the construction of a tramway line and a bridge over the Loire (Samyn et al., 2014).

The topography of the region maintains a generally flat profile, with a noticeable elevation drop at the basin's edges. In fact, it is considered as a large depression within the major bed of the Loire River (Auterives et al., 2014). On the northern side, altitudes range between 125 m and 109 m, while on the southern side, they vary from 130 m to 105 m. Within the basin itself, altitudes fluctuate between 110 m and 90 m, resulting in an average elevation difference of 15 m (Blanchard, 1903).

3.3.2 Geological background

Located south of the Paris Basin, the geology of the Val d'Orléans presents marine and then lake formations deposited during the successive Cenozoic regressions (Blanchard, 1903). The geological bedrock in the region consists of the Beauce limestone. This particular limestone, classified as a Tertiary lacustrine (lake) limestone. It has a considerable thickness that ranges from 50 to 90 meters. The process of karstification in the Orléans region is believed to have begun primarily during the Würm glacial period (Perrin et al., 2015) (the last glacial period of the Pleistocene).

This period was characterized by lower sea levels, which led to increased erosion potential for the Loire River. As a result, the river began to erode the low-permeability sediments that covered the limestone in Val d'Orléans. This erosion event caused a significant influx of aggressive water from the Loire River and its tributaries into the limestone formations (Perrin et al., 2015). The **Beauce limestones** constitute one of the most extensive aquifers in France. It supplies a large part of the Paris basin. These aquifer formations extended over nearly 9000 km², contain an unconfined aquifer to the north of the Loire and captive under the clay-sandy covering of the *Burdigalien* de la Sologne, to the south of the department. The Beauce limestone aquifer under the Val d'Orléans is more extensive than the alluvial aquifer. It extends on both sides under the forest of Orléans, under the Sologne and under the Val d'Orléans. All of the Beauce limestone consists mainly of permeable levels of *Pithiviers* limestone on its upper part, and *Etampes* limestone on its lower part. These two levels are separated by the less permeable formations of the *Gâtinais molasses*. The Loire floodplain's limestone is covered by a layer of Quaternary fluvial alluvium consisting, which is up to 20 meters thick. The Loire River's smaller bed contains recent alluvium consisting of sands and gravels, while the larger bed has flood silts, sandy and micaceous materials (Desprez, 1976; Martin et al., 2003). Therefore, in general, the geology of Val d'Orléans is characterized by a variety of sedimentary rocks, including limestone, marl, and sandstone (**Fig 3.3**). The theoretical cross-section of the alluvial deposits is show in **Figure 3.4**, which are as follows (from the surface to depth).

-Upper layer: Silty sands, sandy clay, and fine sandy silt.

-Middle layer: Sandy gravels.

-Lower layer: Gravelly sands.

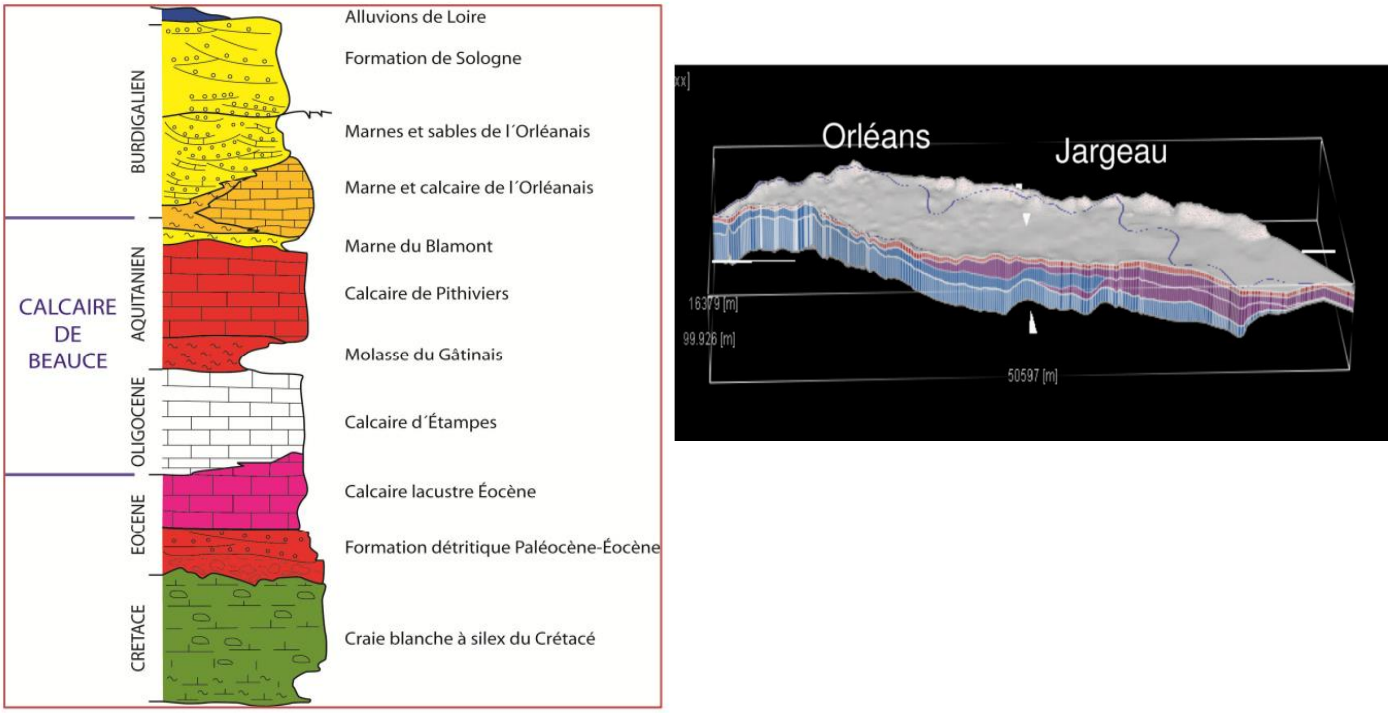


Figure 3.3 Left. Succession of lithological units of the Val d'Orléans showing the Burdigalien formation on top of the Beauce limestone, (Martin et al., 2003). **Right.** 3D representation of the geology of the Val d'Orleans; Red= Alluvium, Blue= Limestone, Purple= Marls (Auterives et al., 2011).

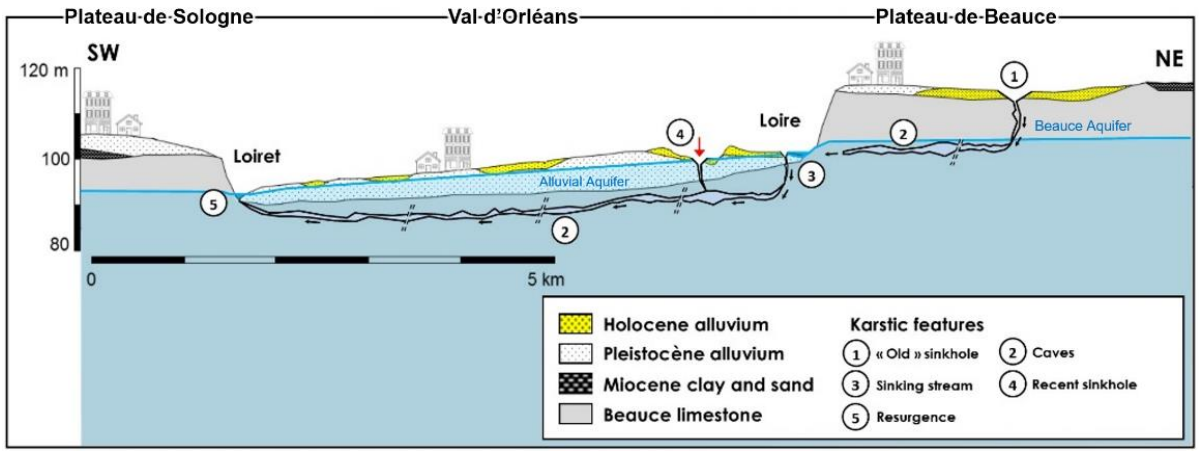


Figure 3.4 The simplified profile of the geological cross-section in Val d'Orléans overlying the Beauce limestone (Perrin et al., 2015 modified from Caudron, 1964).

3.3.3 Hydrogeological context

According to Cerema (2014), Desprez conducted a comprehensive hydrogeological reference study on the Val d'Orléans in 1967. The study delineates the presence of two superimposed aquifers: one formed by the layer of alluvial deposits and the other by the karstified Beauce Limestone (**Fig 3.5**). As previously mentioned, the aquifer in the region is composed of highly-fissured karstic limestone, known as Beauce limestone, and a carbonate lacustrine deposit, impacting the area through a network of underground caves (animated by BRGM, 2014, available at sigescen.brgm.fr/-Hydrogeologie-.html).

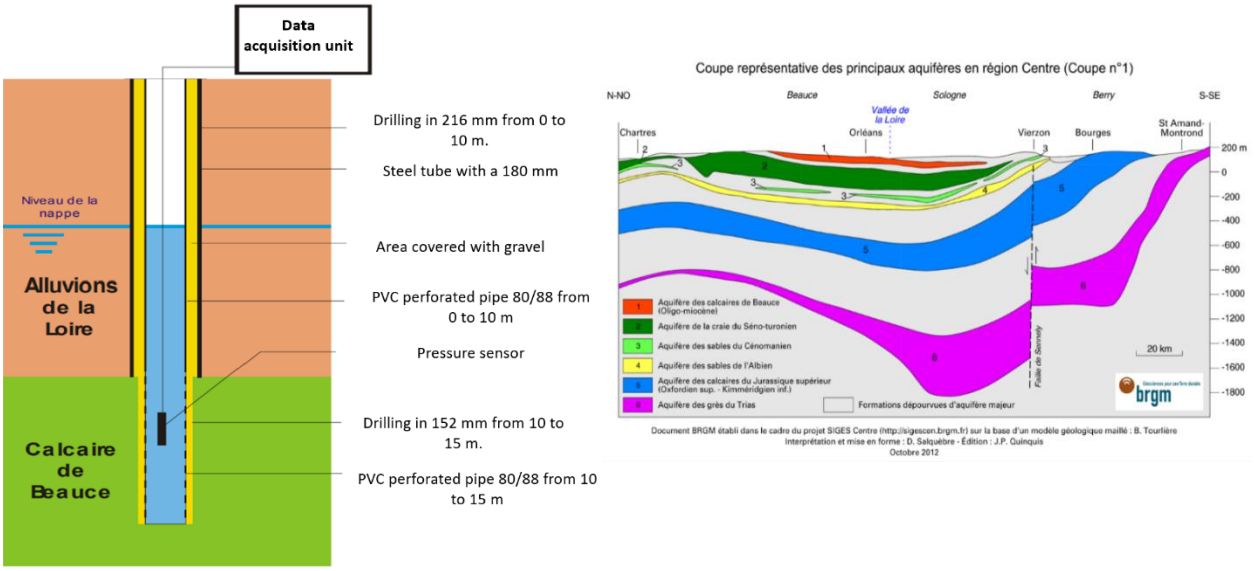


Figure 3.5 Left. A cross- section diagram of the two superimposed aquifers: i) formed by the layer of alluvial deposits (in orange) and ii) by the karstified Beauce Limestone (in green); **Right.** A representative profile of the main aquifers in the central area. The red layer shows the Beauce limestone quifier (BRGM, sigescen.brgm.fr, n.d.).

The karst aquifer, overlain by the Quaternary alluvial deposits of the Loire River (Auterives et al., 2014), is characterized by a complex geological setting. This setting consists of a multi-layered system of clay interbedded with limestone, creating a confined aquifer system with no flow between alluvia and limestone (Lepiller, 2006). However, communication between the two aquifers exists across the entire Val d'Orléans, facilitated through sinkholes/ swallow holes. The relationships between the two groundwater systems are thus intricate and play a significant role in the occurrence of collapses (Cerema, 2014).

This karst aquifer features numerous swallow holes along the Loire River and its tributaries upstream of the city of Orléans. These swallow holes cause significant water losses from the river to the aquifer, historically estimated at 15 to 20 cubic meters per second (m^3/s). Recent research, based on 46 years of data, suggests a lower average water loss of 11.5 m^3/s (Perrin et al., 2016). The lost water is brought back to the Loire River by a few springs among which the main and most important one is the Bouillon spring, which is the origin of the Loiret River (Perrin et al., 2016)_ this will be discussed more in **Section 3.3.4.a.**

3.3.4. Val d'Orléans, a multi-hazard forming setting

In June 2016, the special characteristics in terms of geology and hydrogeology of Val d'Orléans manifested in the form of a multi-hazard event. Five consecutive days of heavy rainfall led to flooding across a 30-km^2 area and 1 km nearby the Loire River due to the overflow of an artificial canal (Noury et al., 2018). The flood disrupted access to urban areas for several days, damaging the A10 and multiple departmental roads (**Fig 3.6**). Once the floodwater had receded, a significant number of sinkholes were noted near roads and buildings (**Fig 3.7**).

The floodwater as a primary hazard triggered more than 100 sinkholes among which 12 collapsed either immediately or shortly after the meteorological event (Noury et al., 2018; Luu et al., 2019). The flood-induced sinkholes happened nearby the Loir River breaching the river dyke, which caused another flooding. In multi-hazard literature, this relationship is called a bi-directional relationship or feedback mechanism where a primary hazard can produce itself (Gill and Malamud, 2016). Furthermore, this chain of hazards can potentially further continue if debris produced by a flood blocks the sinkholes' conduits, resulting in water not being able to infiltrate into the ground at sufficient speed and therefore causing flooding (Zhou, 2007; Xiao et al., 2016). In the following, I provide explanation of the origin of sinkholes and floods in the region.



Figure 3.6 The flood disturbed the access to the urban areas for several days by cutting off the A10 and several departmental roads (BRGM.fr, 2024).



Figure 3.7 After the water receded in the 2016 event, numerous sinkhole collapses were observed around roads and buildings (BRGM, 2016).

3.3.4a The karst system of the Val d'Orléans

When discussing the karst system in the Val d'Orléans, it is important to note that it is equivalent to the hydrogeological situation described in the previous section. However, in the following, I will emphasize this situation through the lens of the karst system to provide a better understanding for readers.

The Val d'Orléans is home to a distinct karst system, influenced by the contributions of the Loire River's losses (pertes in French). The development of this karst landscape began when the main channel of the Loire River gradually cut deeper into the Beauce limestone. This process started during the Wurm glaciation (the last glacial period of the Pleistocene, approximately 115,000 to 11,700 years ago), a time when the sea level dropped significantly, around 100 m lower than today. This change led to a more forceful flow of the river, causing substantial erosion and creating favourable conditions for karst formation. This process was particularly prominent in the interaction zone between the river and the upper Beauce limestone, referred to as the Pithiviers limestone (Lepiller, 2006; Jozja et al., 2010).

The overall direction of the karst is southwestward (**Fig 3.8**). Water losses from the Loire feed the karst from Jargeau in the north-east of Val d'Orléans, and generally between Châteauneuf-sur-Loire and Sandillon (**Fig 3.8**). Large quantities of the Loire's water (more than 80%) seep into the aquifer notably in the Jargeau area (Gutierrez and Binet 2010).

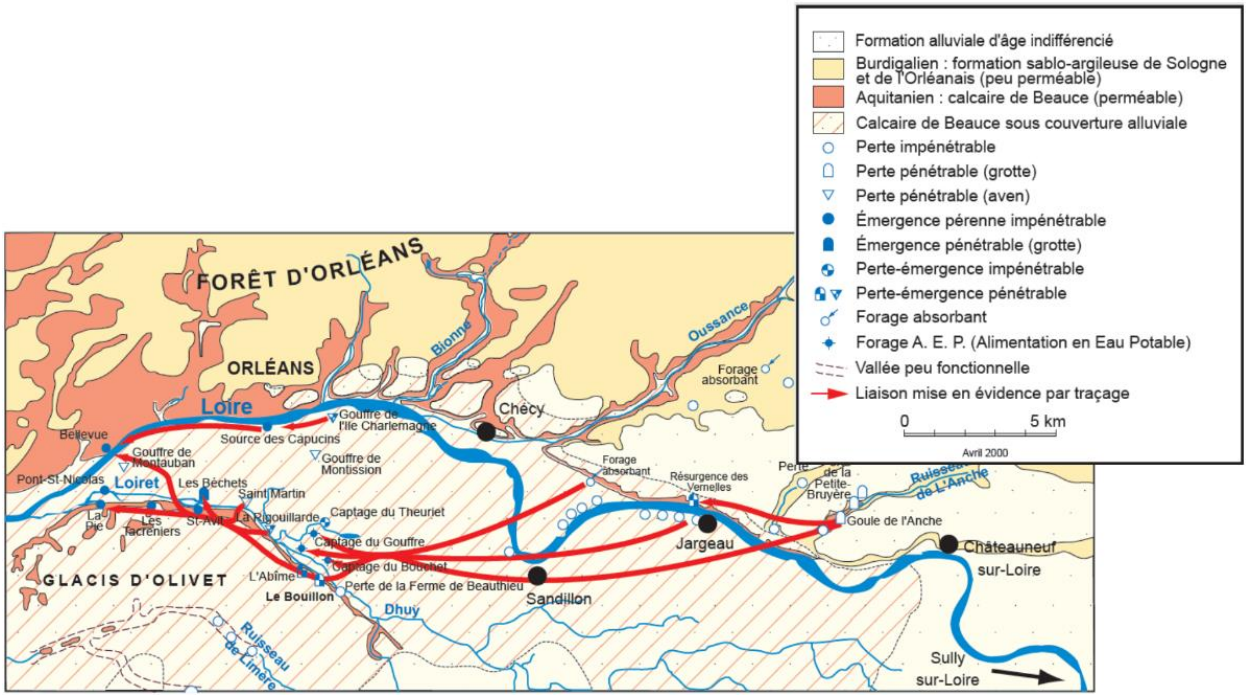


Figure 3.8 The map of karst network in the Val d'Olréans (Lepiller et al., 2000). The red arrows shows the overall direction of the karst networks which is south-westward.

The waters that have disappeared further upstream subsequently reappear at various sources of the Loiret River further south-west downstream, the first being that of Bouillon (Cerema, 2014). The Loiret River, from which the department takes its name, is a 12-kilometer watercourse that is part of the Loire hydrosystem. Its origins are the re-emergences (resurgences) of waters from the Loire that have disappeared into the karst system of the Val d'Orléans. To explain more, the Loiret River, forming the southern boundary of the Val, receives some of its water from the Loire River through an underground karst network. This includes contributions from tributaries like the Dhuy. The source of the Loiret, located in Orléans-la-Source (floral park), acts as a resurgence of the Loire. The connection between the upstream losses of the Loire, the water table within the Val, and the downstream resurgences demonstrates the existence of an underground karstic network across the entire area. During the summer, with minimal local precipitation, the waters of the Loire directly supply the Loiret through this underground network (**Fig 3.8 and 3.9).** The losses in the Loire River and the main resurgence of the Loiret, called Bouillon, situated in Orléans-la-Source (Parc Floral) are shown in **Figures 3.10b** and **3.10c**.

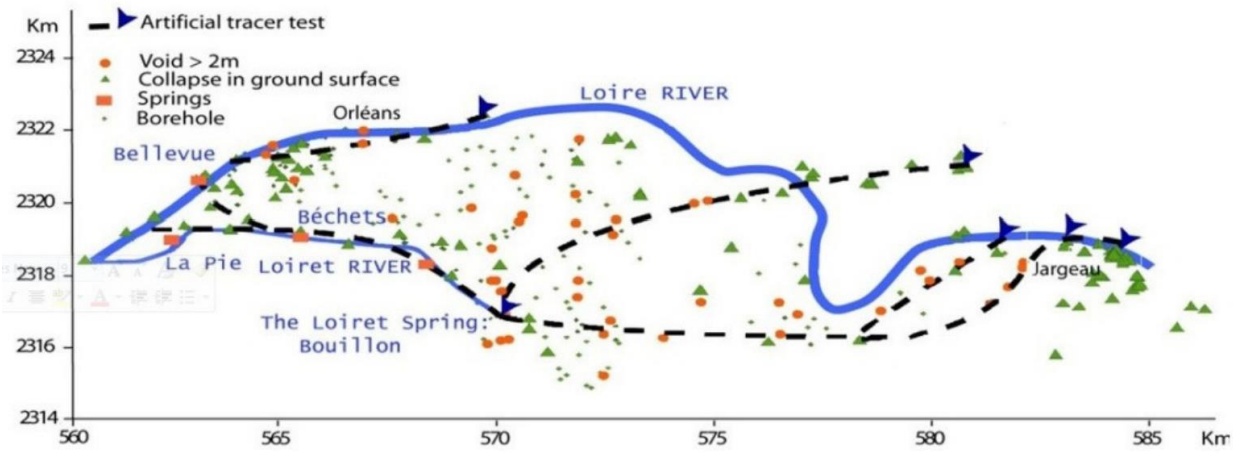


Figure 3.9 Map of the Val d'Orléans Aquifer. Observed voids in the boreholes (circles) and springs (squares), land-surface collapses (triangles) and the tracer tests (dashed lines) are indicated (Auterives et al., 2011).



Figure 3.10 a,b) Close-up view of samples depicting eroded carbonate rocks highlighting the highly karstified bedrock in the area (Credits: the author, August 2022). **c,d)** The water loss of Moulins in Jargeau; from this water loss the Loire water goes into the ground and comes up from the main resurgence of the Loiret called Bouillon situated in Orléans-la-Source (Parc Floral) (Jozja et al., 2010).

This underground network (exchanging water between Loire and Loiret) is linked to surface karst formations, some of which cause deformation of the ground typically in the forms of collapses, subsidence, or decompressed areas. The flow of ground water and the infiltration of surface water (river, precipitation) act as exacerbating factors in the emergence of surface karstic phenomena (Cerema, 2014).

As discussed earlier in this thesis (**Chapter 2**), karst is a type of landscape that can form in soluble formations such as carbonate, evaporitic rocks. In the study area, the karst formation is carbonate-based. The solubility of calcium carbonate in water is relatively low, but it becomes significantly higher in the presence of carbon dioxide. Biomass and micro-organisms release carbon dioxide through different natural processes. In the specific context of the Val d'Orléans region, the presence of micro-plankton in the alluvial aquifer contributes to water acidification due to the release of carbon dioxide and the subsequent formation of carbonic acid. The various chemical reactions involved can be summarized as follows:

$$CaCO3 + CO2 + H2O = Ca (HCO3)2 \text{ (soluble)}. \quad (q\ 3.5)$$

Understanding the rate at which this process occurs is complex and influenced by a range of factors. These factors include the water's aggressiveness, which refers to its capacity to dissolve limestone, as well as the frequency and speed of water flow replenishment (Cerema, 2014). According to Cerema, (2014), in humid-temperate regions, specific dissolution rates (which represent the proportion of dissolved rock volume to exposed surface area) have been documented in the range of 25 to 81 mm/1000 years. However, these rates may need to be adjusted higher in the case of the Val d'Orléans, where the water flow replenishment is very significant.

3.3.4b Underground cavities and sinkholes in Val d'Orléans

Generally, there are three main types of cavities in the Val d'Orléans (Blanchard, 1903):

1. *Natural karst cavities.* These are formed due to the dissolution of the limestone substratum, typical of karst landscapes. These cavities, often unseen by the public except for naturalists, contribute to the area's unique geological features and pose risks for collapses and subsidence.

2. *Quarries.* These are man-made (anthropogenic) excavations created for extracting limestone used in construction, roadworks, and agricultural modifications. Over time, as these quarries became obsolete, their entrances were filled in, but they still present potential hazards due to the possibility of collapses.

3. *Cellars.* Historically linked with dwellings for storing food and drinks, these underground spaces were common in the region's farms and homes. However, with changes in contemporary usage and needs, these cellars have been increasingly sealed off and forgotten, leading to concerns about the integrity of their masonry and potential safety risks.

The focus of the present research is on the natural karst cavities. In the study area, sinkholes refer to circular depressions in the ground, reaching sizes of up to 35 m in diameter and 15 m in depth. These depressions are formed when the water that fills the caverns in the Beauce limestone of the region becomes overloaded due to a significant increase in water from the Loire River during a flood (Blanchard, 1903). The water then seeps into the marl that forms the roof of the caverns, gradually eroding it until it collapses under the weight of the soil above. Once formed, these sinkholes can transform into either sources or absorption sinkholes, depending on the water level in the Loire River. There are many sinkholes in the Val d'Orléans region, and that they can have a significant impact on the local landscape and hydrology (Blanchard, 1903). In the study area, the karst is covered by the Loire's alluvium. A complex phenomenon occurs here: the karstic network can draw away the covering materials, creating a new cavity at the alluvium-limestone interface.

In general, geophysics and geologists have identified two mechanisms leading to sinkholes' formation in the study area (Perrin et al., 2015):

1) The process of **suffosion** involves the erosion and movement of alluvial deposits towards karstic conduits, creating a void at the base of the alluvium. This void gradually enlarges until it eventually reaches the surface.

2) The ceiling of a karst conduit can collapse due to changes in the conduit size or variations in hydrostatic pressure within the karst aquifer.

The water levels in alluvial and karstic aquifers near the Loire River fluctuate with the river levels, particularly during floods. This can cause slight differences in water levels between the aquifers, which may lead to processes like suffosion, contributing to sinkhole formation. Additionally, hydraulic fracturing in the soil at the contact with the rock head can cause sinkholes when water pressure gradients exceed normal stress gradients. These conditions are especially problematic when the water table drops, reducing support and increasing collapse risk (Perrin et al., 2016).

The study done by Cerema (2014) explains that the primary source of data regarding the location of karstic events is a study of the Val d'Orléans conducted in 1967 by N. DESPREZ, in which 170 incidents were documented. Studies by Blanchard (1963) and the BRGM (*Détermination de la présence de karsts sous les levées domaniales du bassin de la Loire, 2009*), the "BD cavités" database, and the archives of LR Blois have also contributed to locating numerous karstic events. However, in most cases, only the locations of these events were documented. In total, 436 events have been documented throughout the valley, between 1856 and 2013. Therefore, the theoretical frequency is estimated at 2.8 events per year (this is a minimum estimate, as many minor events go unreported) (Cerema, 2014).

Despite the region's apparent stability, the complex underground network of natural karst cavities poses safety risks to property and individuals. The formation of sinkholes beneath buildings and the Loire River's dyke pose a significant risk to both human lives and properties (see examples **in Figures 3.11** and **3.12**).

Figure 3.13 shows a sagging ground one the right side of the Loire River showing the past sinkhole activation in this location. These sinkholes had been filled and repaired before. However, the sagging ground is a strong warning for reactivation. The local authors have warned that it is highly probable that this location collapses again during an intensive rainfall. Another recent example in the area is the activation of sinkhole beneath buildings. An alarming event on 12th January 2022, in the centre of Orléans reminds the authorities how dangerous sinkholes can be. Residents of the two buildings were forced to evacuate due to the growing sinkhole formation underneath the buildings (**Fig 3.11**). At the time of writing this thesis, the collapse is predicted to be imminent.



Figure 3.11 Recent sinkhole incident and the evacuated buildings in the Bourgogne street, the city of Orléans (Credit: France 3, 2022)



Figure 3.12 Sinkhole karst collapse opened up around mid-March 2021 on the Chevauchée dike in Saint-Jean-le-Blanc (45). Drone shot by B. Chevrier (BRGM).



Figure 3.13 Locations of past sinkhole collapses on the right bank of the Loire River dyke in the commune of Bou, Loiret. These sinkholes have been filled and repaired twice. Due to sagging ground, they have been identified as re-activated sinkholes with the potential to collapse again, especially during intense heavy rainfall. (Source: Personal communications with field geologists; photos by the author, August 2022).

3.3.4c Floods of the Loire River

The following information about the Loire River and flooding is a mixed information extracted and translated from the PPRi document (2015) and a historical document on Loire flood references (Le Plan Loire Grandeur Nature, 1996).

The Loire River is the longest river in France, with a length of 1,020 km. The river crosses the department of Loiret on a length of about 132 km from east to west. The size of its basin and its alignment with the prevailing winds results in significant irregularity. The Loire River can be volatile during flood periods and may transform into a shallow body of water during dry spells (PPRi, 2015). To protect the lands and residents from floods, many structures have been built along the river, such as dikes, dams, and locks. However, despite these protection measures, catastrophic floods have always occurred.

The floods of the Loire River are a special case due to i) the considerable volume of water they can bring, ii) the surface area of the zones they can submerge, and iii) the extent of the damage they can cause. A major flood of the Loire River that would occur today would represent, in terms of damage, the most significant national disaster after the Paris flood and a significant earthquake in Nice.

Floods of the Loire River are caused by two different meteorological phenomena:

1. The flood of oceanic origin: It is caused by west winds pouring abundant rain on the Paris Basin, Limousin, and Auvergne. It is a slow flood that represents significant volumes of water and affects the tributaries. It happens mostly in winter and spring.

2. The Cévennes flood: It is caused by intense rain and storms in the eastern region of the Massif Central, mainly affecting the basins of Allier and Haute-Loire. It is a locally faster and more abrupt flood that spreads and reaches the Loiret department in a few days. It happens mostly in autumn.

When both the oceanic and Cévennes phenomena occur simultaneously, they can lead to catastrophic floods, as witnessed in the three major floods that occurred in the 19th century (in 1846, 1856, and 1866 , causing more than 1,000 deaths and considerable damage), each more than a century in magnitude, occurring in less than thirty years. These mixed floods are characterized by a significant rise in water levels throughout the basin and are the most dangerous for the Loiret department.

The 1846 Loire River flood was a catastrophic event that significantly impacted the Loire Valley. On October 22, the river's waters rose sharply, leading to 100 dike breaches between *Briare* (a town located approximately 150 km south of Paris), and *Langeais*, (situated west of Tours in central France) and the destruction of numerous infrastructure elements, including parts of the newly built Orléans-Tours railway line. This flood inundated cities like Nevers (located in the Bourgogne-Franche-Comté region, about 260 km south of Paris, along the Loire River), and damaged towns and villages along the river, causing substantial financial losses and highlighting the inadequacies of existing flood defense systems. It marked a critical moment in the history of flood management in the region, underscoring the need for more effective and resilient flood control measures.

The 1856 Loire River flood, known as "the reference flood," was a devastating natural disaster impacting the entire Loire Valley. This event was unique due to the simultaneous flooding of all Loire tributaries, from Velay to Nantes, highlighting the effects of embankment on flood regimes. The flood caused substantial destruction, with 160 levee breaches over a 23 km stretch, inundating 100,000 hectares, and

destroying thousands of hectares of farmland. In Nevers, the water flow reached about 4,000 cubic meters per second (m^3/ s), and at Tours, 6,000 m^3/ s. The flood severely damaged infrastructure, including bridges and 98 km of railway tracks, and significantly impacted the regions of Indre-et-Loire and Maine-et-Loire. This event underscored the limitations of existing flood defences and the need for improved flood management strategies in the Loire Valley.

The 1866 flood of the Loire River, striking just ten years after the disastrous 1856 flood, marked another significant natural disaster in the region's history. Occurring in September 1866, this flood was predominantly caused by the Loire itself, with minimal contribution from its tributaries, and was particularly severe on the Allier River. Notably, it was the worst flood ever recorded in Nevers, a consequence of a very humid summer that led to rapidly rising water levels starting September 23. This event further exposed the challenges and inadequacies of the existing flood management systems in the Loire Valley, underscoring the need for more effective flood defense strategies.

Flooding of the Loire River occurs due to a combination of high water levels (resulting from meteorological events) and the constraints of the local flood defense systems. The dikes along the river have been constructed over several centuries with successive layers, which increases the risk of failure due to:

i) Vertical heterogeneity in the structure of the embankment, which can cause internal erosion and collapses due to water infiltration.

ii) The presence of karst in the underlying limestone beneath the dikes. In the 19th century, several breaches were attributed to this phenomenon (**Fig. 3.14**). In January 2009, this occurred on the top of the dike itself in Jargeau, and in 2010, a multi-story house collapsed in a karst near the levees in Saint-Pryvé-Saint-Mesmin. It is also believed that the breach in the Orleans channel during the 2016 flood event may be due to the collapsing of the underground cavities (Noury et al., 2018).

iii) Pipelines passing through the dikes, weakening the structure of the dike and creating preferential infiltration paths for water.

Moreover, water levels can exceed the height of the dike and overflow it. Such floods are particularly hazardous because when a dike breaks, water rapidly pours through the breach, attaining high speeds and flowing into the valley with significant destructive force (with flow rates around 2,000 m^3/s).

During the catastrophic floods of the 19th century, spillways were constructed where the levees had failed. Beyond a certain flow rate, these spillways allow a limited volume of water to flow into the valley to prevent a widespread rupture of the levee system (flow rate of about 800 to 900 m3/s). Today, there are five spillways in the Loiret department, located in the municipalities of *Saint-Martin-sur-Ocre, Pierrelaye* in *Dampierre-en-Burly, Ouzouer-sur-Loire, Jargeau,* and *Marzan* in Dry.



Figure 3.14 The 19th-century flood picture depicts the extend of the devastating historical flood of the Loire River which resulted in a breach of the dyke. The breaches of the dyke is attributed to the presence of karst in the underlying limestone of the dikes. (DREAL, obtained from the DICRIM Document).

3.3.4d Flood risk management and urban planning in the Val d'Orléans

The flood risk management in Val d'Orléans involves a collaborative approach primarily led by the state, with active support from various intercommunal entities. From the **1960s** to **1990s**, urbanization in the Val d'Orléans expanded, driven by state confidence in flood control through high dikes/levees and planned flood control dams since initial flood protection relied on high dikes and planned flood control reservoirs (Dournel, 2019). However, changing national policies and opposition from environmental groups led to a shift towards stronger flood prevention measures, including the **1995 law** promoting flood risk prevention plans (PPRi).

From **2003** to **2009**, the local government and its partners implemented awareness campaigns, housing adaptation programs, and distributed informational materials to reduce flood vulnerability. These efforts were coordinated with regional, national, and European projects, leading to the establishment of the European Center for Flood Risk Prevention and Management in Orléans in 2006.

By 2015, revised PPRi plans distinguished between urban and agricultural/natural areas to manage floodwaters better, sparking local debate. These integrated actions across various levels of governance demonstrate a coordinated approach to flood risk management, environmental protection, and community resilience (Dournel, 2019).

3.4 Methodology and data collection

3.4.1 MaxEnt implementation (Transition from theory to practical application)

After discussing the theoretical framework of Maximum Entropy (MaxEnt) and its foundational principles in the beginning of this chapter (**Section 3.2**), the next phase involves the application of this theory to model the probability of hazard distribution. This section will delve into the practical implementation of MaxEnt, using specialized MaxEnt software to translate theoretical models into hazard map. I first will outline the process of selecting appropriate environmental variables, which serve as input data for the MaxEnt model. Following this, the method of data preparation and processing will be detailed, to make sure that the inputs align with the requirements of the MaxEnt algorithm. After that, our focus will shift to setting up the MaxEnt software, including tuning the model by choosing the optimal parameters for this analysis. **Figure 3.15** shows a methodological flowchart to give a whole overview of the process used to generate and explore outputs from MaxEnt.

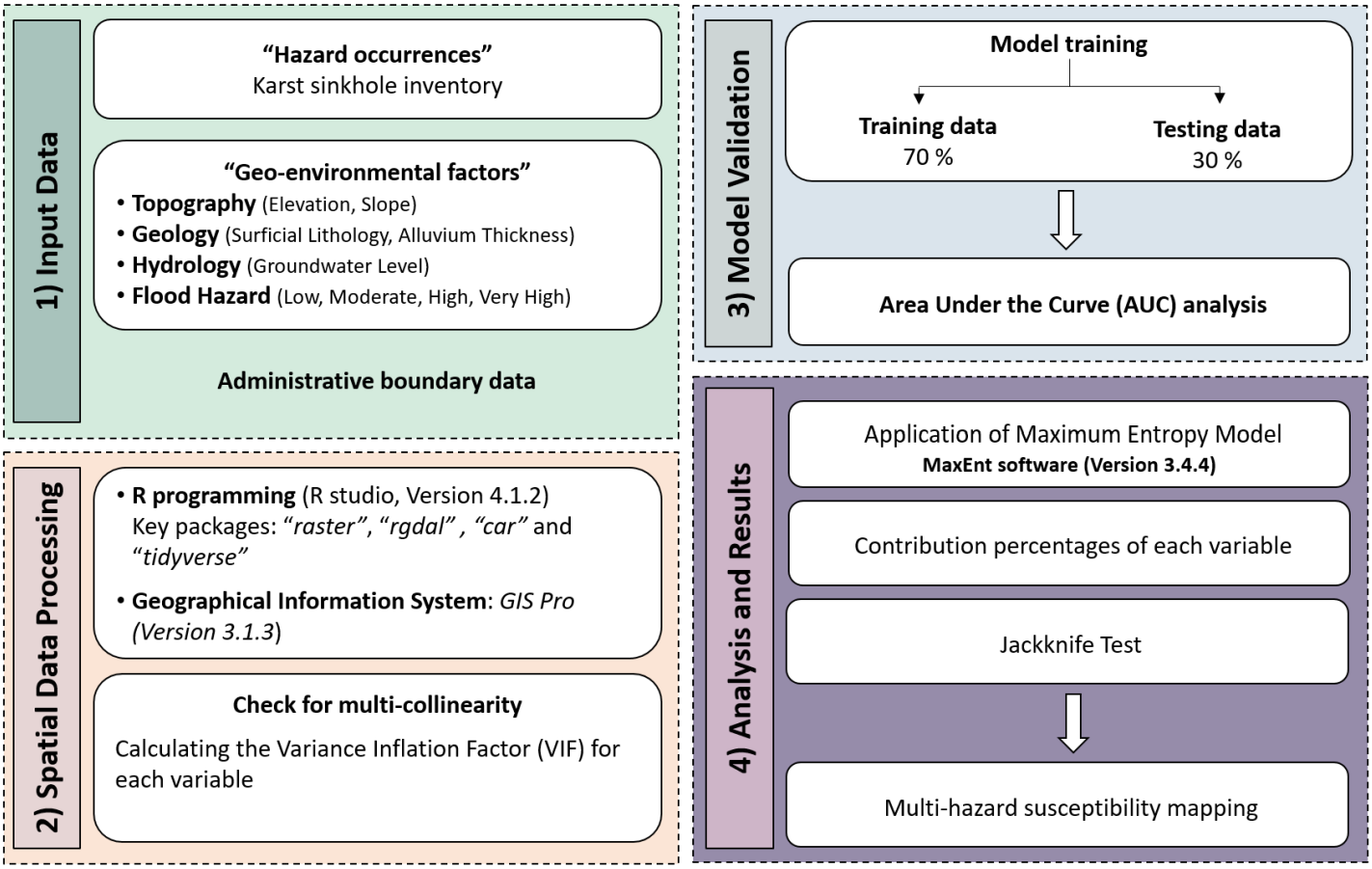


Figure 3.15 The methodological flowchart. The modelling process is divided into four stages: **1)** input data (data collection); **2)** spatial data processing, **3)** model validation; and **4)** model application (analysis and results) which is discussed in Chapter 4.

3.4.2 Natural sinkhole inventory

The first step in every sinkhole-related hazard mapping is to create a sinkhole inventory map showing the locations of the previously identified sinkholes (Kim and Nam, 2018). In France, sinkhole events under the title of ground movements (the file's name: BD Cavité, BDMVT, BSS) are recorded by the French Geological Survey (BRGM) and are publicly available on the national cavity database website (www.georisques.gouv.fr/). In this study, 915 sinkhole locations with natural origins are considered (**Fig. 3.16**). It is important to note that it is unclear how many are triggered by floods, but relying solely on that data would likely be insufficient for modelling.

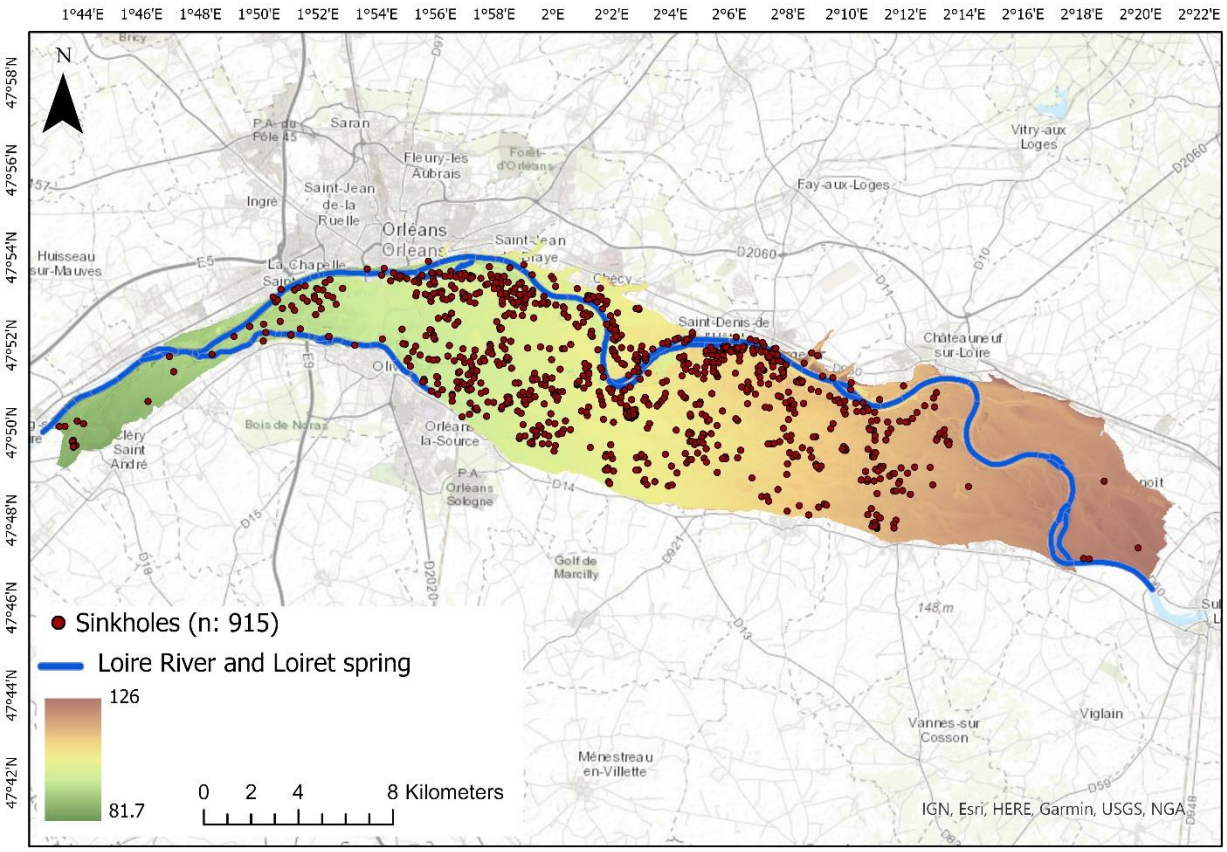


Figure 3.16 Locations of past sinkholes overlaid on the elevation layer; the red dots present sinkholes' locations; the blue lines show the watercourse of the Loire River and Loiret spring (created by the author).

The location of sinkholes suggests that more than 50% of sinkholes have occurred within approximately one kilometre of the Loire River (**Fig 3.17**). This observation gives a rise to the concern of more collapses in case of a flood event.

The associated process is still poorly understood, particularly as to whether flood directly triggers sinkhole formations or triggers the pre-existing ones to collapse. Therefore, here I use flood-triggered sinkholes term for both scenarios to hypothesize a potential trigger based on observed correlations, rather than definitively stating causation.

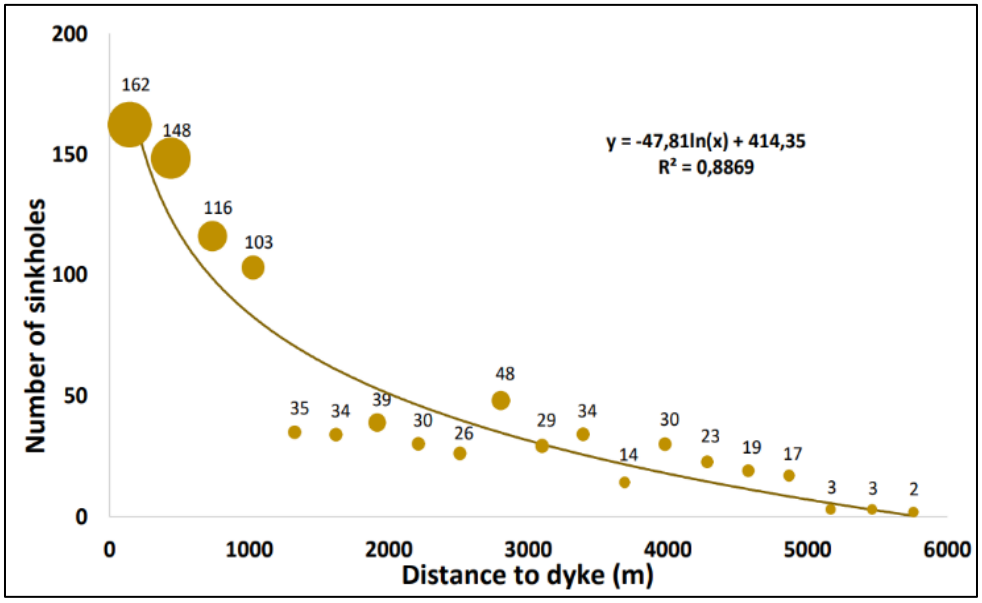


Figure 3.17 Correlation between distance to dyke and sinkholes' frequency. The plot reveals the closer to the river's dyke, the more frequent sinkholes are. R^2 value (coefficient of determination) equals ~ 0.88 suggesting a high correlation between distance to the river and the sinkhole frequency.

3.4.3 Geo-environmental conditioning factors for flood-triggered sinkholes

In this study, I selected contributing factors (also referred to (geo)- environmental contributing factors) based on the principle of parsimony, which seeks to balance model simplicity with explanatory power (Shatz, 2019). This approach, rooted in Occam's razor, prioritizes using the minimum necessary elements for effective explanation (Baker, 2003; Shatz, 2019). More importantly, modellers need to know that it is of high importance of thinking of two main issues in machine learning when selecting features (contributing factors):

1. One main issue that is not only limited to MaxEnt but also prevalent in many machine learning models is the issue of "Spatial Autocorrelation". In the MaxEnt literature, the issue of *Spatial Autocorrelation* has been highlighted by Sillero and Barbosa (2021) and Sillero et al. (2021). Spatial autocorrelation is a property of the first law of the Geography known as Tobler's Law "everything is related to everything else, but near things are more related than distant things" (Tobler 1970). This means that closer locations tend to have more similar values. As for selected contributing variables, when they are spatially auto-correlated, their values are not independent across space, which can result in misleading variable importance. For example, precipitation and temperature levels tend to be more similar between nearby locations than those further apart. This phenomenon, where the values of a variable are correlated based on spatial proximity, is known as spatial autocorrelation, a spacial form of correlation occurring in space (Sillero and Barbosa, 2021).

2. The second issue which is often overlooked by modellers is the variability which refers to the range and diversity of data within each feature. In a simple term, I need to have a wide range and sufficient variability in our features ensures that the model can learn meaningful patterns and relationships. Features with little variation provide limited information and can weaken the model's ability to generalize. For example, if the rainfall value in an area does not show any significant gradient, it means that rainfall is relatively uniform across the region. This lack of variability means that the model cannot effectively use rainfall data to distinguish between different locations or predict outcomes based on this feature. Consequently, incorporating such a feature would not enhance the model's predictive power and might even introduce noise, leading to less accurate predictions. Therefore, ensuring that selected features have sufficient variability is crucial for building robust and generalizable models.

Accordingly, we carefully selected six key factors that offer significant explanatory and predictive capabilities while avoiding excessive complexity. The six factors chosen are as follows:

- *topography* (elevation and slope),
- *geology* (surficial geology and alluvium thickness),
- *hydrology* (groundwater level and flood hazard zones).

The selection of factors for our study was based on literature review, which included results from physical modelling and geo-technical research conducted in the Val d'Orléans area (Cerema, 2014; Noury et al., 2018; Luu et al., 2019), as well as universally accepted factors identified in the scientific literature, such as topography. I introduced flood hazard zones as a key factor, acknowledging their dynamic impact on

sinkhole collapses. I also applied collinearity test to ensure there is not any significant collinearity among selected contributing factors (discussed in the next Section **3.4.3a**).

Each thematic map is presented and analyzed in the subsequent Sections (from **3.4.b to 3.4f**). It is important to note that all thematic layers were created using the most recent version of ArcGIS Pro software available at the time of writing, version 3.1.3. The data type and their respective sources are demonstrated in **Table 3.1.** All the data were extracted from their sources mentioned in Table. 3.1 and after conducting data manipulation, their respective GIS maps were produced.

Table 3.1 Spatial contributing factors included in the MaxEnt model.

Variables	Data type and units	Data Source
Digital Elevation Model (DEM)	Continuous (Meter)	BD ALTI® (géoservices.ign.fr)
Slope	Continuous (Degree)	Extracted from DEM layer
Surficial geology	Categorical (geological period)	BD – Charm-50, the geological map (1/50,000), (BRGM)
Groundwater	Continuous (Meter)	Groundwater national portal (ades.eaufrance.fr)
Flood hazard zones	Continuous (Meter)	Plan de Prévention des Risques d'Inondation (PPRi), 2015
Alluvium thickness	Continuous (Meter)	Extracted from Perrin et al. (2015)

3.4.3a Collinearity test

Discussing the concept of collinearity in statistical modelling, it is important to first provide a clear understanding of its meaning and application. Collinearity, in this context, refers to the situation where two or more predictor variables (also known as independent or explanatory variables) in a statistical model are highly correlated (Dormann et al., 2013). This condition can affect the model's ability to accurately estimate the relationship between the response variables and the set of predictor variables. This is a main concept across various types of statistical models, including regression models, classification and regression trees, and neural networks (Dormann et al., 2013). However, as discussed earlier in this chapter, this has often been ignored in natural hazard modeling, which has frequently led to distorted results.

High collinearity among predictors can lead to unstable estimates of the model parameters, making it difficult to distinguish the individual effect of each predictor on the response variable. This can lead to inflated standard errors, reduced statistical power, and potentially misleading conclusions about the importance or significance of predictors (see Dormann et al., 2013). Therefore, by examining collinearity within the framework of MaxEnt model, I aim to identify and address these issues prior to running the model, thereby enhancing the model's predictive accuracy and interpretability.

I did a comprehensive analysis to identify potential collinearity among the key geo-environmental variables: Elevation, Slope, Geology, Thickness, Groundwater and Flood hazard layers. The analysis involved two main statistical methods: 1) Pearson correlation coefficients and 2) Variance Inflation Factor (VIF) calculations. Both calculations were done using R studio with relevant packages namely "car" package (the screenshot of the process and the R codes have been placed in the **Appendix 3.D**). The methodology involved converting raster data into a vector form, enabling the calculation of Pearson correlation and VIF between each pair of variables.

i) Pearson correlation coefficients: This statistical approach, often denoted as "r" identifies the degree of linear association between variables, with coefficients ranging from -1 (perfect negative correlation) to 1 (perfect positive correlation) (Mukaka, 2012), where:

r = 1 indicates a perfect positive linear relationship,

r = −1 indicates a perfect negative linear relationship,

r = 0 indicates no linear relationship.

It is calculated using the following formula:

$$r = \frac{\sum_i (xi - \bar{x})(yi - \bar{y})}{\sqrt{\sum_1 (xi - \bar{x})2 \sum_1 (yi - \bar{y})2}} \qquad \text{(q 3.6)}$$

where x_i and y_i are the individual sample points, and $\bar{x}$ and $\bar{y}$ are the means of the x and y samples, respectively (Mukaka, 2012).

The data, derived from high-resolution spatial raster files, provides full coverage of the study area, which ensures a robust basis for the subsequent correlation analysis. **Table 3.2** demonstrates the results of the Pearson correlation matrix, which shows the pairwise correlation coefficients between variables.

However, collinearity or multi-collinearity assessment often involves looking beyond just the correlation between independent variables; it typically involves calculating the Variance Inflation Factor (VIF) to quantify how much the variance of an estimated regression coefficient increases if our predictors are correlated. Having calculated the Pearson correlation, I further calculated the VIF of each variable shown in the **Table 3.3** (Chatterjee and Hadi, 2013). The screenshots of the calculations done in R programming can be seen in **Appendix 3.D.**

The following explores the linear relationships between various geographical and environmental variables, including topography (elevation, slope), geology, groundwater, alluvium thickness and flood hazard through Pearson correlation analysis. Understanding these interrelationships is crucial for interpreting environmental patterns and for the predictive modelling of geographical phenomena. Pearson correlation coefficients offer insights into the strength and direction of linear associations between pairs of variables, facilitating a deeper understanding of their interplay within the studied region.

Table 3.2 Pearson Correlation coefficients among geographical and environmental variables.

Variable	Elevation	Slope	Geology	Thickness	Groundwater	Flood
Elevation	1	-0.053	-0.147	0.186	-0.467	-0.300
Slope	-0.053	1	0.0262	-0.023	0.120	0.211
Geology	-0.147	0.0262	1	-0.033	0.109	0.350
Thickness	0.186	-0.023	-0.033	1	0.074	-0.037
Groundwater	-0.467	0.120	0.109	0.074	1	0.237
Flood	-0.300	0.211	0.350	-0.037	0.237	1

The result of the Pearson correlation assessment (**Table 3.2**) suggest that:

i) *Collinearity of Elevation with other variables*: **Elevation and slope**: The correlation coefficient is -0.0535, indicating a very weak negative linear relationship between elevation and slope in the area. This suggests that changes in elevation have little to no linear association with changes in slope. **Elevation and flood**: The coefficient is -0.3008, suggesting a weak to moderate negative linear relationship. As elevation increases, Flood layer tends to decrease, but the relationship is not very strong. **Elevation and Geology:** The coefficient is -0.1471, indicating a weak negative linear relationship between elevation and geology. **Elevation and Groundwater**: The coefficient is -0.4670, suggesting a moderate negative linear relationship. This indicates that as elevation increases, Groundwater tends to decrease. **Elevation and thickness**: The coefficient is 0.1864, indicating a weak positive linear relationship. This suggests that as elevation increases, thickness also tends to increase slightly.

ii) *Collinearity of Slope with the other variables:* The correlation coefficients between slope and the other variables are all relatively low (absolute values less than 0.25), indicating weak linear relationships. **Slope and geology**: Very weak positive correlation (0.0262). Slope and geology are nearly independent. **Slope and thickness**: Very weak negative correlation (-0.023). **Slope and groundwater**: Weak positive correlation (0.120). **Slope and Flood**: Weak positive correlation (0.211).

iii) *Collinearity of Geology with other variables:* Geology shows a moderate positive correlation of 0.3503 with flood layer, indicating that areas prone to flooding may have certain geological characteristics. **Geology and thickness**: Very weak negative correlation (-0.033). Geology and thickness are nearly independent. **Geology and Groundwater:** Weak positive correlation

(0.109). Geology has a slight positive relationship with groundwater presence. **Geology and Flood:** Moderate positive correlation (0.350).

iv) *Collinearity of Thickness with other variables:* **Thickness and groundwater**: Very weak positive correlation (0.074). **Thickness and Flood**: Very weak negative correlation (-0.037). Thickness and flooding are almost independent.

v) *Collinearity of Groundwater with other variables:* **Groundwater and Flood**: Weak positive correlation (0.237). Areas with more groundwater tend to have a slightly higher likelihood of flooding.

ii) Variance Inflation Factor analysis (VIF): To further assess collinearity, the VIF analysis provided a more nuanced understanding of collinearity among the variables within the context of a multiple regression model. The following table summarizes the VIF results for elevation vs other variables:

$$\text{VIF}_i = \frac{1}{1-R_i^2} \qquad \text{(q 3.7)}$$

Where:

VIF*i* is the variance inflation Factor for the predictor variable and R_i^2 is the coefficient of determination (R-squared) of a regression model where the predictor is regressed on all the other predictors. According to Chatterjee and Hadi (2013), the authors of the book *Regression analysis by example*, if there is no linear relationship between predictor variables_meaning they are orthogonal (independent of each other) VIF*i* would be 1. A VIF*i* value that deviates from 1 suggests a movement away from orthogonality and an inclination towards collinearity. A VIF above 10 indicates that collinearity could be problematic for the estimation accuracy.

In summary:

VIF = 1: No correlation among the predictor and the other predictors.

1 < VIF < 5: Moderate correlation, but often not a concern.

VIF >= 5 (or 10, by more conservative standards): Indicates a high level of multi-collinearity that may warrant further investigation or action.

A clear overview of the VIF values for each variable in relation to the other factors are shown in **Table 3.3 (**The screenshots of the calculations done in R programming can be seen in **Appendix 3.D).** It indicates that all variables have VIF values well below the commonly used threshold of 5, suggesting that there is no severe multi-collinearity issue among these variables. Thus, they can be considered for inclusion in our model without the need for concern over collinearity adversely affecting the regression analysis.

Table 3.3 Variance Inflation Factor (VIF) matrix for the six variables obtained by (R studio, package "car").

Variable	Elevation	Slope	Geology	Thickness	Groundwater	Flood
Elevation	1.00	1.056	1.145	1.010	1.075	1.245
Slope	1.432	1.00	1.143	1.073	1.342	1.243
Surficial Geology	1.435	1.056	1.00	1.074	1.353	1.158
Thickness	1.349	1.058	1.146	1.00	1.306	1.295
Groundwater	1.140	1.050	1.146	1.037	1.00	1.287
Flood	1.379	1.016	1.025	1.075	1.345	1.00

3.4.3b Topography data (Elevation and Slope)

Topography plays an important role in flood-triggered sinkholes in several aspects and have been long considered as one of the main important contributing factors for both flood and sinkholes studies (Ozdemir, 2015). Both flood and sinkholes normally occur in low-lying areas. Low elevation areas can contribute to water accumulation and facilitate seepage into underground cavities, thereby weakening their stability. Furthermore, in flat terrains, the slower infiltration of water can result in elevated groundwater levels, exerting increased pressure on subsurface voids. These combined effects significantly increase the risk of sinkhole formation during flooding events (Brinkmann et al., 2008).

The Elevation data was obtained from the Digital Elevation Model (DEM), which is available on the national website BD ALTI® (géoservices.ign.fr). The region of Orléans is a relatively low-lying region with the elevation ranging from nearly 82 m to 126 m (**Fig 3.18**). Notably, the elevation gradually decreases from east to west in the area where more sinkholes concentrated. Another important topographic factor is slope. The slope gradient is the measure of the ground surface's steepness, expressed in degrees or percent (Galve et al., 2009a). It has been derived from the digital elevation layer (DEM) using GIS pro software. The slope map of the area reveals a predominantly gentle topography, with slope degrees ranging from 0 to 71 degrees (**Fig 3.19**). The terrain is characterised by a generally moderate to low slope gradient, suggesting a relatively flat landscape. However, a noticeable variation can be observed alongside the river, where the slope becomes relatively steeper.

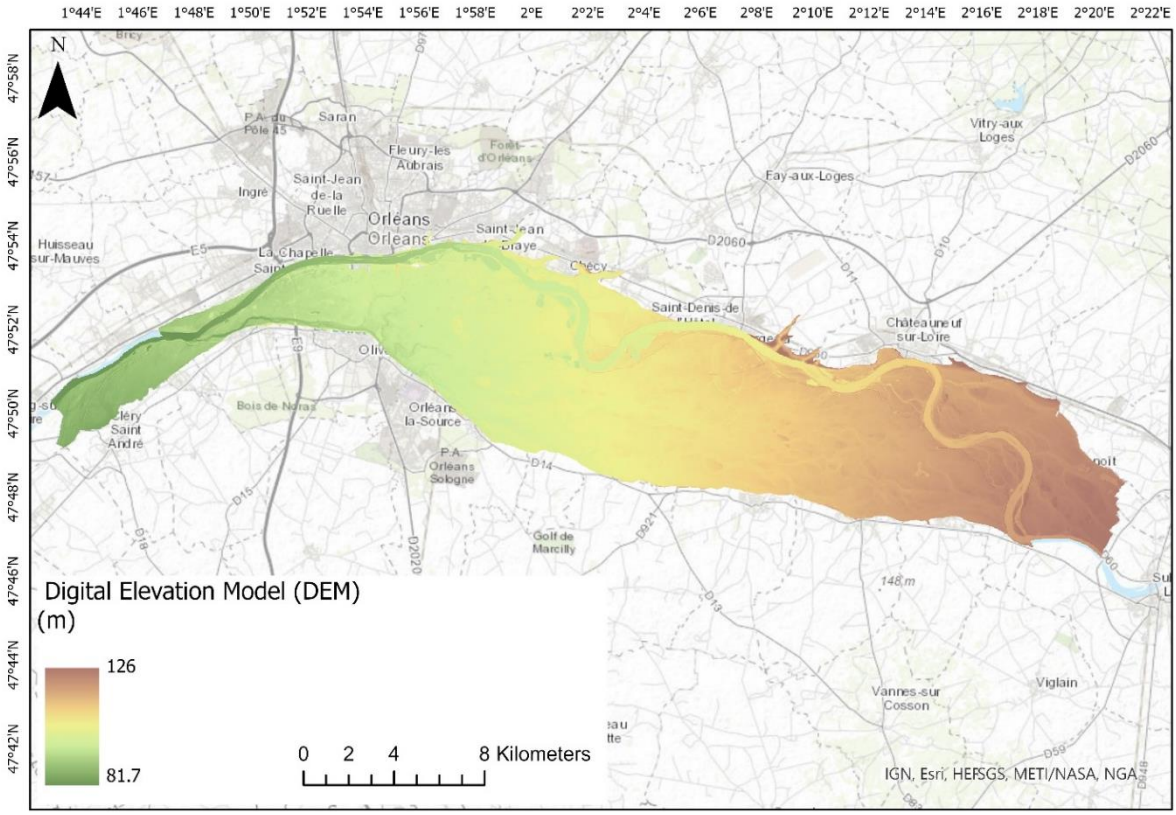


Figure 3.18 The rasterized map of Digital Elevation Model (DEM) of the study area (created by the author)

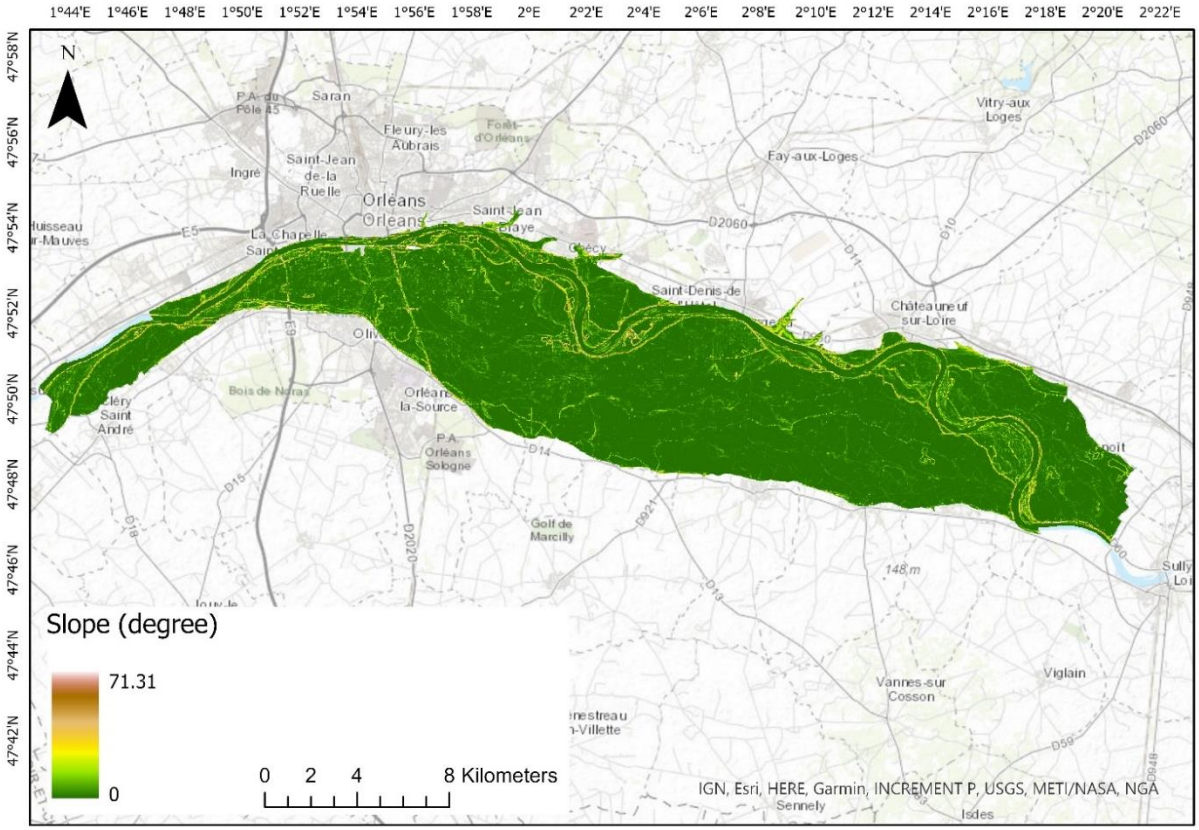


Figure 3.19 The rasterized map of Slope derived from DEM (created by the author).

3.4.3c Surficial geology

In karst areas, in most cases, sinkholes do not usually result from the collapse of the soluble rock. Instead, they predominantly occur due to the ground's failure where unconsolidated sediments/soil rapidly washes into bedrock cavities that have developed over geological timescales (Waltham et al., 2005). Therefore, in our analysis, I focused on the surficial geology of the area, recognizing that it is situated above Beauce limestone. This underscores the importance of the surficial layer in our analyses, rather than the bedrock.

The geological map of the area has been published by BRGEM at the scale of 1/50,000. The spatial data was extracted and mapped using the database BD-charme-50 being available in the BRGEM official website (**Fig 3.20**). The area is mainly covered by recent and older alluvium (Holocene and Pleistocene respectively) which are mostly permeable.

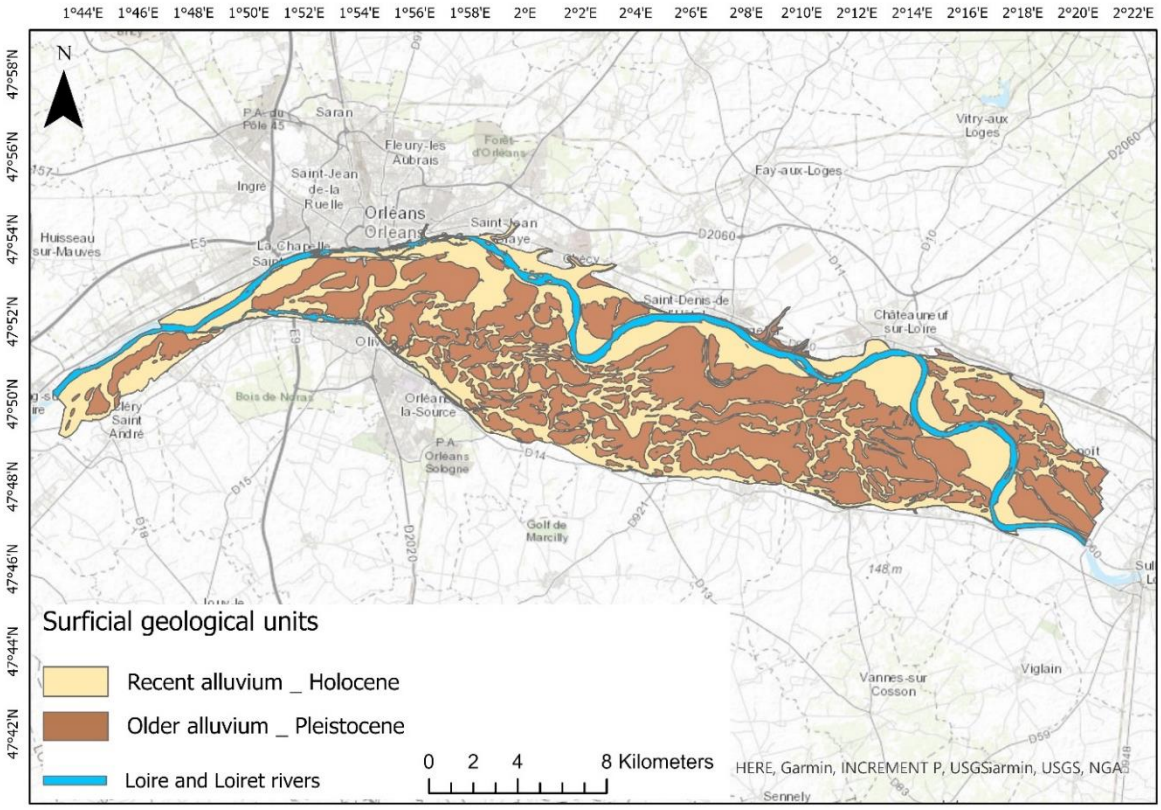


Figure 3.20 Surficial geology consisting of Holocene (new) and Pleistocene (old) alluvium layers (Created by the author).

3.4.3d Alluvium thickness

The accumulation of river sediments over time contributes to the increased thickness of alluvium in the central part of the region (2°2'-2°10'E), where also has a low elevation (81 m to 95 m). In these areas, the Beauce limestone is covered by Quaternary alluvial deposits with a thickness up to 20 meters. According to Noury et al. (2018), the sinkholes mainly occur due to suffosion or down washing of the cover material. Therefore, the hypothesis is that the thinner the alluvium is, the more likely cavities reach the surface (Perrin et al., 2015). Originally generated by BRGM, the data was extracted from Perrin et al. (2015) applying the IDW method to create a raster map (**Fig 3.21**). Sinkholes in the area mainly occur in the alluvial deposits with their sizes ranging from 1m to 5 m (median 2.8m) depths between 1 m and 4 m(median 2.0) (Perrin et al., 2015).

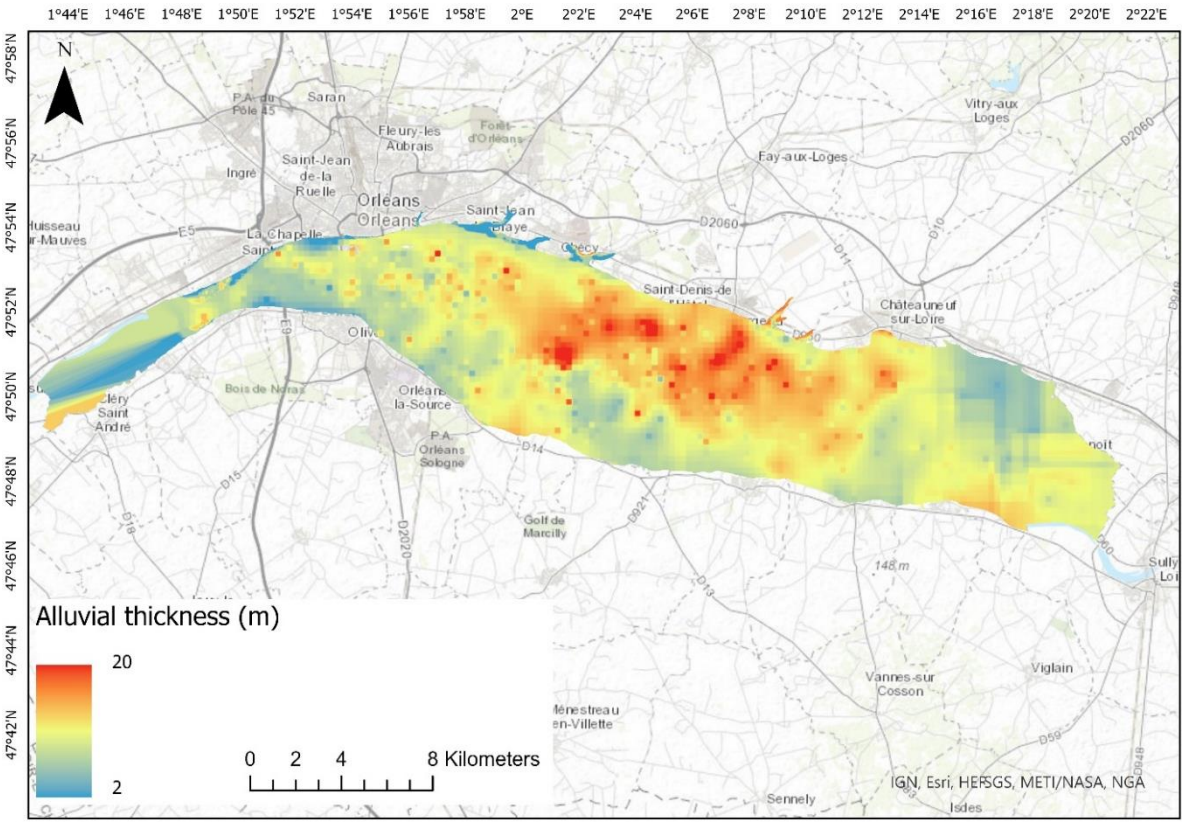


Figure 3.21 The rasterized map of alluvium thickness; Quaternary alluvial deposits with a thickness up to 20 meters (created by the author).

3.4.3e Groundwater spatial data

Subsurface cavities, when exposed to groundwater flow, can enlarge and migrate towards the surface, driven by groundwater circulation, which facilitates the dissolution of specific rock formations (Linares et al., 2017; Pazzi et al., 2018). In karst regions underlain by soluble bedrock, continuous exposure to both subterranean and surface water can increase the process of dissolution (Linares et al., 2017; Intrieri et al., 2018; Xiao et al., 2018). When groundwater increases, it can lead to soil erosion by weakening the soil structure and facilitating the process of soil structure ravelling (Kim and Nam, 2018). Recent studies on Val d'Orléans, simulating the internal erosion during flood revealed the role of groundwater leading to different erosion regimes in the area (discussed in **Chapter 2**). Therefore, the level of and the variation in piezometric level is one of the main factors to assess when evaluating the flood triggered sinkhole susceptibility.

The groundwater in the region is closely connected to the Loire River. Changes in the rivers water level can affect the water table in the aquifers and therefore the level of the Loire is an essential reference data for understanding the level of the groundwater at the piezometers, mainly fed by the water losses from the Loire (BRGM RP/63864-FR; Blanchard, 1903). The large quantities of the Loire's water (more than 80%) seep into the aquifer notably in the Jargeau area (Gutierrez and Binet, 2010). An earlier study conducted by Desprez (1967) generated piezometric maps of the study area, using data from over 700 boreholes. These maps represented periods of low and high water levels (**Fig 3.22**). The results indicated that the annual difference between these two levels is approximately one meter. More recently, to assess water heights, a new map was created at 100 points in 2008. It does not show significant differences (Joodi, 2009). As such, for this study, I have considered the annual average. Moreover, the piezometric monitoring set up for the 2012-2013 period provided precise data on the groundwater level revealing that the fluctuations in the groundwater level are directly linked to those of the Loire.

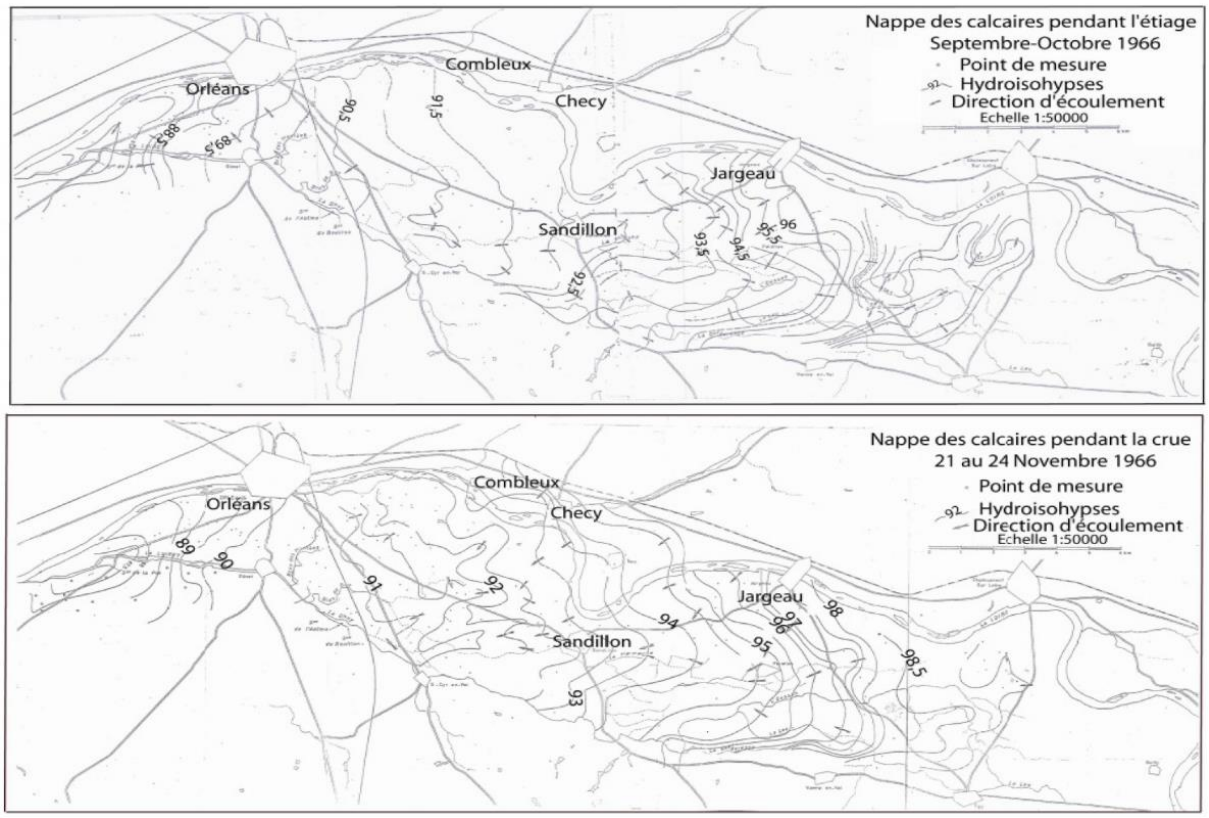


Figure 3.22 The piezometric map of the Val d 'Orléans (Desprez, 1967). **a)** September- October 1966 (on 712 wells) during the period of low flow. **b)** November 1966 during high flow. The annual difference between these two levels is 1 m (Auterives et al., 2014).

The data related to the groundwater was obtained from the groundwater national portal (*ades.eaufrance.fr*). According to the ades website, there are five active piezometer water wells in the scale of Val d'Orléans. Each station/gauge has start and end date with the total number of measurements (**Table 3.4**). Moreover, to improve the accuracy of the interpolation, five nearest stations close to the border from north, west and northwest were selected. In **Table 3.4** the start and end date of data collection is provided with the number of measurements. The latest start date among the water gauges is 24/11/2011 (for BSS001ACSC) and the earliest end date is 31/05/2019 (for BSS001AEQB). Therefore, the common operational period for all the gauges is from 24/11/2011 to 31/05/2019. The graph of each water gauge is provided in **Appendix 3.C**.

To interpolate and rasterize the data, I used the Inverse Distance Weighting (IDW) method integrated in Arc GIS Pro (**Fig 3.23**). (Spatial) Interpolation is a method used in mathematics, (geo) statistics, and various scientific fields to estimate or predict a value within a range (domain) that is bounded by known observations/sampled locations (Li and Heap, 2008; 2014). It is based on the assumption that there is a smooth transition or relationship between the known data points often nearby (Li and Heap, 2008). In geo-statistics, two common methods are used for non-linear interpolations: i) Inverse Distance Weighting

(IDW) and ii) Kriging. They are both geo-statistical methods used for spatial interpolation, but they differ in their approach and underlying assumptions:

i) **IDW (Inverse Distance Weighting)**: This method assumes that the value of a sampled point is a weighted average of the values of nearby sampled points. The weights decrease as the distance increases, based on the idea that points closer to the unknown point are more influential in predicting its value. The user defines the power of the distance in the weighting formula.

The general form of the IDW method is given by:

$$f(x,y) = \frac{\sum_{i=1}^{N} w\,(di)zi}{\sum_{i=1}^{N} w(di)} \qquad \text{(q 3.8)}$$

where $w(d)$ is the weighting function, zi is the value at known point i, and di is the distance from point i to the point (x,y) ((Isaaks and Srivastava, 1989 as cited in Li and Heap, 2008).

ii) **Kriging**: this method also known as "Gaussian process regression" was developed by D. Krige in mining application in South Africa. Kriging is not only considers the distance between points but also the spatial correlation between observed data points. Kriging uses variograms to model this spatial correlation and provides estimates along with a measure of uncertainty or variance (Li and Heap, 2008).

The two methods have been integrated into all forms of ArcGIS software and selecting one depends on the data type and purpose. I have used IDW method due to the even distribution of piezometer wells across the study area, which makes IDW a simple and straightforward method for our dataset, providing accurate local interpolations without requiring complex spatial modelling.

Table 3.4 Five piezometric gauges in the Val d'Orleans, Loiret department (extracted from ades website; ades.eaufrance.fr).

Water gauge	Start date	End date	Number of measurements
BSS001ACSC (03635X0545/PZ1)	24/11/2011	18/10/2021	3609
BSS001AEQB (03636X1060/PZ2)	24/05/2010	31/05/2019	3135
BSS001AEQC (03636X1061/PZ3)	24/05/2010	18/10/2021	3679
BSS001AEQD (03636X1062/PZ4)	23/05/2010	01/10/2022	4029
BSS001CBUA ((03982X1045/F))	31/03/2010	30/09/2022	4127

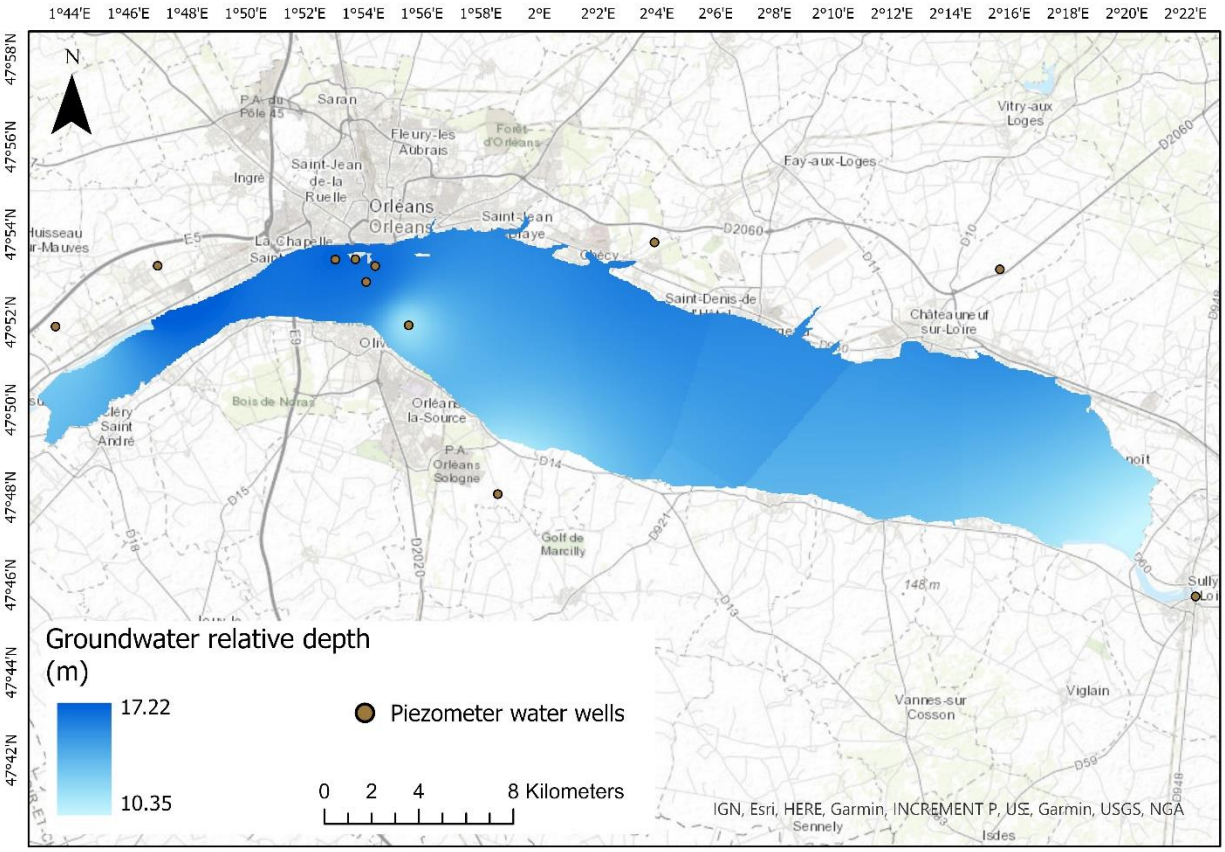


Figure 3.23 The rasterized map of groundwater relative level. The annual average level of groundwater based on 11 years data; ranging from about 10 m to 17 m. Five water points in the study area and six out of the area were considered for Inverse Distance Weighting (IDW) Interpolation. (Created by the author).

3.4.3f Flood hazard map

The effect of flooding in sinkhole occurrences was previously discussed. The official flood hazard map for a 100-return period produced by PPRi (2015) (*Plan de Prévention des Risques d'Inondation*) was used. The map has been built based on topographic data from the airborne laser digital terrain model (DTM, represents only the bare ground surface) and the analysis of the three past floods (for frequency and intensity calculations). Three major floods from the 19th century (**1846**, **1856**, and **1866**) were chosen as the **reference floods**. Historical markers of flood evidence enabled the calculation of peak flood levels, called the highest known waters (Document DICRIM; PPRi, 2015). The spatial modelling then analysed these floods' hydraulic behaviour to determine water flow rates in the valley.

The hazard zones represents the water height (**Fig 3.24**). There are four hazard zones which each were incorporated into the model (**Zone 1**: Low (0 to 0.5 m), 2: **Moderate** (0.5 to 1 m), 3: **High** (1 to 2.5m), 4: **Very high** (greater than 2.5)).

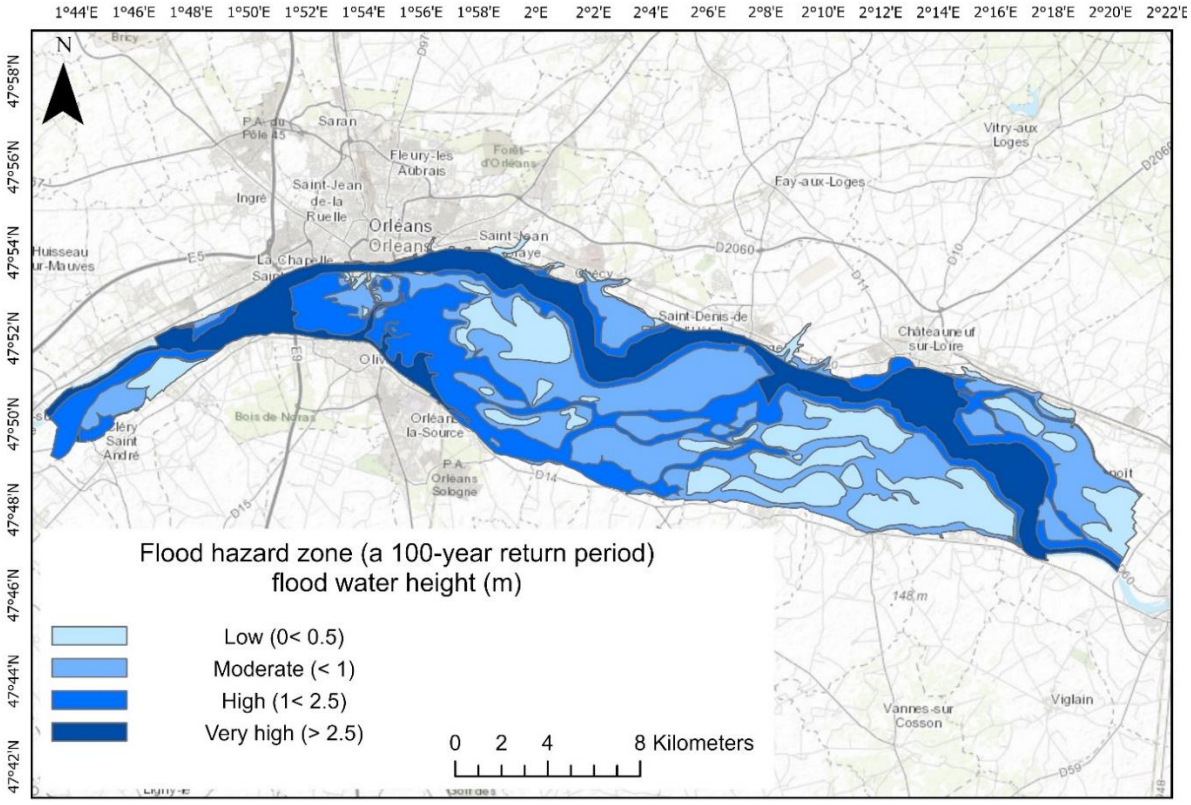


Figure 3.24 The official flood hazard map: classified into four level of hazard (low, moderate, high and very high). (PPRi, 2015).

3.5 Data Processing and Preparation for MaxEnt Modelling

3.5.1 Data preparation procedures

Data preparation is the process of changing, reorganizing, or transforming data to make it suitable for analysis or for a specific purpose. This can involve tasks such as filtering, sorting, merging, aggregating, or transforming data in some way. One of the most important and fundamental components in the MaxEnt model application is the data preparation.

To produce the forecasting model, MaxEnt only accepts specific data formats and structures. To prepare the data for analysis, I used a combination of software tools, namely **Microsoft Excel**, **ArcGIS Pro**, and **R Studio** (See the flowchart **Fig 3.15**). I will detail the process step by step, illustrating how the data was transformed into the appropriate format for use with the MaxEnt application. This will provide a clear understanding of the processes involved in making the data ready for MaxEnt analysis.

3.5.1a Creating a Comma-Separated Value (.csv) file

To comply with the required format in MaxEnt for the input samples (presence data or sinkhole location data), a CSV file was generated using the coordinates of sinkhole occurrences (georeferenced data). This file is structured into three distinct columns, which include the name of each occurrence, its latitude, and longitude. **Figure 3.25** provides a glimpse of this CSV file, illustrating a representative segment of the sinkhole occurrence data as one of the main inputs for MaxEnt. It is importat to note that MaxEnt is sensitive to the order of the columns, therefore, I made sure to respect the order, from left to right: occurrences, latitude and longitude respectively.

	A	B	C
1	Occurences	X	Y
2	Sinks	634565.01	6749953
3	Sinks	634398	6749970
4	Sinks	630945	6751740
5	Sinks	631901.01	6752248
6	Sinks	633225.99	6749407
7	Sinks	633450	6752406
8	Sinks	633774.99	6750749
9	Sinks	634076.01	6751410
10	Sinks	634341	6751540
11	Sinks	634535.01	6750956
12	Sinks	634632	6748417
13	Sinks	635003.01	6748519
14	Sinks	635414.01	6750208

Figure 3.25 A screenshot of CSV format for sinkhole occurrence data in MaxEnt input samples.

3.5.1b Modifying (geo)-environmental layers

Following the initial preparation of the CSV file containing sinkhole occurrences, the next crucial step in our analysis for MaxEnt involved modifying the geo-environmental layers. This step is integral in ensuring that all predictors/ (geo)-environmental layers are compatible with the MaxEnt model. To achieve this, each (geo)-environmental layer was converted into a rasterized format, as it was done and show in **Section 3.4.3.** It was essential to ensure that all layers shared the same *raster format, geographic boundaries, cell size,* and *map projections* for consistency and accuracy in the MaxEnt analysis.

This process of standardization and alignment of the layers was conducted using *R Studio version 4.1.2*. Several packages (libraries) were used such as *Raster, rgdal and Tydyvers* packages. All the rasters were projected to the French coordinate system "*RGF 1993 Lambert-93*". Furthermore, a 10-meter pixel size was used for the entire raster layers. This small pixel size, indicative of a higher resolution, is instrumental in capturing vital spatial details that are crucial for accurate environmental modelling. Such a resolution allows for a more nuanced and precise representation of the environmental factors in our analysis. **Figure 3.26** shows a screenshot of one of the R scripts when converting each layer to specified formats. Each predictor layer separately modified to be the same extent, cell size and coordinate system.

```
raster_refCval ← raster(xmn = xmin(EXT_CVal), xmx = xmax(EXT_CVal),
                        ymn = ymin(EXT_CVal), ymx = ymax(EXT_CVal),
                        res = 10)
```

Figure 3.26 A screenshot of R script for unifying the data layers to the MaxEnt format.

3.5.1c Conversion of modified geo-environmental layers to ASCII Format

After successfully modifying all the environmental layers to ensure uniformity in their format, extent, and resolution, the subsequent phase involved converting these layers into ASCII format. ASCII files which stand for "American Standard Code for Information Interchange" are simple text file formats representing grid-based raster geographic data (Phillips et al., 2006). They are the designated raster format compatible with MaxEnt, a critical requirement for running the MaxEnt model effectively. This conversion to ASCII format was done using the R programming environment.

3.5.2 Model building process

A critical step in the context of machine learning is building an effective model which requires careful attentions to both preparation data and the configuration of model parameters. In the previous section, I already talked about the data preparation and now in this section I aim to delineate the tuning steps I followed to build our model. Moreover, tuning the parameters for best fit to the data is important but it will not be all. After running the model the main question that arise is "How do we estimate a performance of a machine learning model?" (Raschka, 2018). This will be discussed after running the model and obtaining the results in **Chapter 5.**

3.5.2a Parameter setting (tuning program settings)

In MaxEnt, there are several settings that one can adjust. These settings directly influence how accurately the model (also referred to as a classifier in machine-learning literature (Domingos, 2012)) can forecast distributions. The settings determine how the model perceives and integrates (geo)-environmental factors, or dependencies, into its calculations (Phillips and Dudik, 2008). MaxEnt models these dependencies using "features". Features are mathematical functions derived from (geo)-environmental variables. For example, a feature might describe how temperature affects the likelihood of a species being found in a particular area. More complex features could allow the model to fit more nuanced environmental dependencies, potentially improving model accuracy if the real-world relationship is indeed complex.

Nevertheless, fitting more complex features usually requires more data. Without sufficient data, a model using complex features might overfit, meaning it becomes too tailored to the specific dataset and performs poorly on new, unseen data. From a machine learning perspective, to understand overfitting we need to know about two main errors: bias and variance. According to Domingos (2012), bais is a tendency to always learn a wrong thing/pattern and variance is a tendency to learn random patterns. High bias can make a model too simple, missing important patterns in the data (under-fitting), while high variance can cause a model too sensitive to the training data, capturing random noise instead of the actual patterns (overfitting). It is easy to avoid overfitting (variance) by falling into the opposite error of under fitting (bias). Simultaneously avoiding both requires a better model that can trade off better, admitting the fact that there is no single perfect model that will always do the best (Domingos, 2012).

These collective effect can result in poor generalization to new unseen data. According to Radosavljevic and Anderson (2014), overfit models tend to perform well on the training data but poorly on testing data because they have essentially memorized the training data instead of learning the underlying patterns. Therefore, what can be done to reduce overfitting?. In this situation, the complexity is controlled by one main mechanism, regularization parameters (Phillips and Dudik, 2008). The ultimate aim of "Regularization" in machine learning, including in MaxEnt, is to prevent overfitting.

Regularization in MaxEnt typically involves adding a penalty to the complexity of model. This penalty is in the form of a **regularization multiplier,** which is a coefficient applied to the value of the respective beta parameter of each feature class (Radosavljevic and Anderson, 2014). A higher value of the regularization multiplier increases the penalty, leading to smaller coefficients, while a lower value decreases the penalty, allowing larger coefficients (Radosavljevic and Anderson, 2014). To explain more, I penalize the model with more structures, thereby favouring smaller ones less room to overfit the training data (Domingos, 2012). When I talk about **penalizing a model with more structures**, I are in fact referring to the practice of **regularization**.

The goal is to balance the fit of the model to the data with its complexity. In MaxEnt, complexity often relates to the weights assigned to the different (geo)-environmental features used in the model. The regularization parameter controls the extent of regularization. A higher value places more emphasis on reducing model complexity (leading to a simpler model), while a lower value allows the model to fit more closely to the data (potentially making it more complex) (Phillipe et al., 2006; Radosavljevic and Anderson, 2014). Therefore, selecting the right regularization parameter is crucial. If the parameter is set too high, the model may become too simple and fail to capture important patterns (under fitting). If it is set too low, the

model may become too complex and fit to noise in the data (overfitting). MaxEnt software incorporates features to address the issue of regularisation through built-in regularization, adjustable parameters, and evaluation tools.

As MaxEnt is used for tackling more complex issues rather than simple exploratory analyses, it becomes important to ensure that modelling decisions are grounded in study objectives, and hazard-specific considerations. It is essential that these decisions align with the intended a priori assumptions and reflect the complexity of the problem at hand (Merow et al., 2013). The MaxEnt model has several parameters that need to be tuned, such as the *regularization parameter* and *the convergence threshold*. As discussed, the choice of these parameters can affect the performance and robustness of the model, and may require careful calibration or cross-validation.

Therefore, in this thesis, I have made informed decisions of choosing the optimal settings rather than just following what is in default. In MaxEnt, the regularization strength is controlled by a parameter called the **regularization multiplier**. To prevent overfitting in my MaxEnt model, I used regularization, which is implemented in the form of a regularization multiplier. I applied a value of 0.5 to apply a moderate level of regularization which helped in preventing overfitting. This number is also a trade-off value suggested by Sillero and Barbosa (2020).

3.5.2b Optimal threshold value:

A notable limitation in using MaxEnt and other ML learning methods arises from the inherent variability of model outputs when the same dataset is applied repeatedly (Phillips et al., 2006; 2017; Sillero and Barbosa, 2021). This issue can be mitigated by running the model multiple times on the same dataset and calculating the average and standard deviation of the results to evaluate if outcomes are consistent across datasets. Ideally, models should be replicated a minimum of 10 to higher numbers like 50 or 100 times, depending on the available computer time and storage (Phillips et al., 2006; Sillero and Barbosa, 2021). Yet, this crucial step is frequently overlooked in the literature (Kornejady et al., 2017; Norallahi et al., 2021; Boussouf et al., 2023 among many). Moreover, within each replicate there is a need to iterate the process. In the current thesis, I have replicated 50 times the entire algorithm with 1000 iterations for each replicate (it is more discussed in **Chapter 4**).

To ensure the convergence of our model towards an optimal solution and a good model fit, I set the maximum iterations parameter to 1,000. This allowed the MaxEnt algorithm sufficient opportunity to adjust its parameters iteratively, seeking to minimize prediction error and accurately reflect the observed distribution patterns of the hazard. This relatively high number of iterations was chosen to make sure that the model had ample scope to reach a stable state, where further iterations would not significantly alter the outcome, thereby indicating a robust fitting of the model to our data.

The combination of 1000 maximum iterations and 50 replicates forms the foundation of our modelling strategy, ensuring that each model run is both thoroughly optimised and that the ensemble of models reflects the inherent variability and uncertainty in hazard distribution predictions. This methodology provides a robust framework for understanding the potential distribution of the hazard under study, contributing valuable insights into its ecological preferences and resilience to environmental changes.

3.5.2d Data partitioning for model training and testing (cross validation)

Forecasting outcomes based on known data is not particularly valuable. Thus, it is essential to test how well models can forecast new, unseen data. A common challenge with data in machine learning is the limited amount of data available for model training. To address this, it is beneficial to allocate as much data as possible for training while minimizing the data reserved for validation (Nilsen, 2005). I used a data partitioning strategy. Specifically, I partitioned the dataset into two subsets:

1. **Training set (70%)**: This subset, including 70% of the data, was used to train the model. During training, the model learns the underlying patterns and relationships within the data.
2. **Testing set (30%)**: The remaining 30% of the data was saved for testing the model. After the training phase, the model's performance was evaluated on this testing set to assess how well it generalizes to unseen data.

By using a 70/30 split, I ensured that a substantial portion of the data was available for model training, while still keeping a good portion for unbiased evaluation. This is widely suggested share that keeps the balance and avoid biases (Rahmati et al., 2016; Yousefi et al., 2020; Bordbar et al., 2022 Rusk et al., 2022).

On the other hand, how I partition the training and testing data is important as different data partitioning methods can affect model overfitting and performance. In general, there are two partitioning methods: 1) randomly partitioning, and 2) masked geographically structured, which restricts training data to the regions where the model is being calibrated/trained (Radosavljevic and Anderson, 2014).

In modelling, it is important to pay attention to this fact that random partitioning, although it sometimes inflates performance estimates (Radosavljevic and Anderson, 2014), provides a straightforward and unbiased method to split data for model validation and can be beneficial when combined with proper tuning of regularization settings to manage model complexity.

Moreover, random partitioning is simple to implement and often used as a baseline to compare more complex partitioning strategies. This approach allows for the assessment of the basic model's discriminatory ability before applying more structured approaches, which may introduce their own biases or complexity (Radosavljevic and Anderson, 2014). Random partitioning offers a more general approach that can be applicable across various types of data and study designs.

Although the research done by Radosavljevic and Anderson (2014) encouraged more complex methods such as geographically structured partitioning, these require detailed understanding of the spatial distribution of data and may not be necessary or practical for all types target occurrences, in our case, hazards distribution models. This study discusses the existing **auto-correction issue** when I randomly divide the data. However, in the current study, I should consider the difference between natural hazards like sinkholes with biological units like species (As discussed in **Section 3.2**).

This is important because MaxEnt model is initially an ecological model dealing with living things. The context and the nature of the data affect how we should approach MaxEnt modelling and the significance of spatial autocorrelation. The flood-triggered sinkhole occurrences are linked to specific geological and hydrological conditions, and their occurrences are not influenced by biological interactions but by physical

processes. Therefore, each observation of a sinkhole provides direct information about the conditions conducive to sinkhole formation in that area.

On the contrary, in species distributions, observations can be influenced by a multitude of ecological interactions, mobility, and adaptive behaviours which adds complexities in data inventories. Generally, species like birds (as only one example among many) show complex ecological interactions. Their locations (presence) can be influenced by factors such as competition for resources with other species, predation pressure, availability of suitable nesting sites, and the abundance of food (Elith and Leathwick, 2009) which complicate their distribution patterns.

In species distribution modelling, spatial autocorrelation can often mask the underlying ecological processes if not addressed properly. However, for sinkholes, proximity and clustering of occurrences are highly informative because they can indicate an underlying susceptibility in the area's geology. This contrasts with species data, where proximity might just reflect habitat preference or migration patterns rather than direct ecological causality.

In sinkhole modelling, treating spatial autocorrelation as a factor to be minimized might overlook the importance of localized physical conditions that are prone to causing sinkholes.

Given these differences, the modelling approach for sinkholes should indeed embrace spatial autocorrelation rather than attempt to mitigate it. To summarize:

- **Cluster analysis**: Identifying clusters of sinkholes can help forecast where similar conditions might lead to future occurrences.

- **Hazard susceptibility zoning:** By acknowledging the spatial clustering of sinkholes, models can be used to define susceptibility zones more accurately.

- **Geotechnical investigation**: Proximity in sinkhole data could guide detailed geotechnical investigations to assess the stability and risk factors of specific areas.

This observation highlights a crucial distinction in the application of geospatial data analysis between biological and geological phenomena. For sinkholes, each observation point is not just a record but an important indicator of potential risk, making the spatial characteristics of the data inherently valuable for modeling. This approach contrasts with biological models where the goal is often to separate broader ecological trends and generalizations across a species' range. Embracing the spatial characteristics inherent to sinkholes can lead to more accurate and useful models for hazard assessment and prevention.

Therefore, to avoid sample selection bias and achieve a representative overview of the landscape, I applied *a random sampling method*. This approach involved randomly partitioning the landscape to capture diverse, unbiased data points. This made our findings more reliable and accurate.

3.5.2c MaxEnt model outputs

MaxEnt has the capability to produce model values in four different output formats: raw, cumulative, logistic and cloglog with each having different theoretical justification. In the following, I will briefly explain the reason of choosing the **logistic output format** as the recent extention to MaxEnt according to the paper Phillips and Dudı´k (2008).

MaxEnt generates a function, referred to as **qλ (x)**, that assigns a probability value to each location (site) based on the model's training data. These values are referred to as "**raw**" values. Raw values are challenging to interpret, especially when they are applied to (geo)-environmental conditions at new sites that were not included in the training data. Moreover, raw outputs are scale-dependent meaning they depend on the amount of background data used; using more background data results in smaller raw values because the total sum of probabilities must still equal one, but across more points. Consequently, to make the raw values more understandable and practical, they are often converted into what is called a "**cumulative**" format. This format involves calculating omission rates based on thresholds applied to the raw values. An omission rate, in this context, is the percentage of actual presence points that are predicted to be absent when a certain threshold (p) is applied to the raw output. The formula transforms each raw threshold into a corresponding omission percentage (**c(p)**).

$$c(p) = \frac{\sum \text{of raw values less than } p}{Total\ sum\ of\ all\ raw\ values} \qquad \text{(q 3.9)}$$

The cumulative format is beneficial because it is scale-independent (not affected by the amount of background data used) and easier to interpret, particulary when applied to new (geo)-environmental conditions or sites. However, it is important to note that this format does not directly correspond to the probability of an occurrence being present; rather, it reflects omission rates at different thresholds. It means when using a specific cumulative threshold **c**, the omission rate for test sites is exactly **c%**, meaning that **c%** of presence points are predicted to be absent. The main limitation of cumulative outputs is that for an occurrence that is generally found everywhere in the study area (high presence probability everywhere), the cumulative values will only reflect the ranking of the sites rather than significant differences in (geo) environmental suitability. This can lead to misleading interpretations where large differences in cumulative values do not necessarily indicate large differences in the actual probability of presence or habitat suitability.

To address the limitations of the cumulative format in reflecting true probabilities of presence, Phillips and Dudı´k (2008) introduced a logistic output format. This format is comparable to logistic regression in traditional statistics, which estimates the probability of an outcome (in our case, presence of a hazard) given a set of predictors (geo-environmental variables). Traditional logistic regression estimates the conditional probability of presence $P\ (y = 1 \mid z)$ given the (geo) environmental conditions z. MaxEnt uses (geo) environmental variables to create a similar logistic output that estimates $P\ (y=1|x)$, where x refers to specific

site conditions. The formulation uses a Bayesian approach to relate the conditional probability of presence to the raw Maxent output: (for more information, see Phillips and Dudı´k (2008)).

$$P(y=1|z) = \frac{P(z|y=1)\,P(y=1)}{P(z)} = \frac{\sum x \in X(z) P(x|y=1)\,P(y=1)}{\sum x \in X(z) P(x)} \quad \text{(q 3.10)}$$

The logistic regression model implemented by MaxEnt uses the maximum entropy principle to derive the logistic output format:

$$P(y=1|z) = \frac{e^{H.q\boldsymbol{\lambda}(\mathbf{x}(\mathbf{z}))}}{1+ e^{H.q\boldsymbol{\lambda}(\mathbf{x}(\mathbf{z}))}} \quad \text{(q 3.11)}$$

Here, *H* represents the entropy term, which is derived from the MaxEnt model, and qλ is the raw output from the MaxEnt model at site *x* corresponding to (geo) environmental condition *z*.

Here are the key points of the advantages of logistic format (Phillips and Dudı´k, 2008):

i) The logistic format enhances the MaxEnt model by providing a more intuitive and statistically robust method for forecasting the presence of hazards. It adapts well to varying data availability and sampling efforts, making it versatile across different scenarios.

ii) The logistic format offers a probabilistic interpretation of presence data based on (geo) environmental variables. This approach delivers a more meaningful and precise measure for modeling and planning, particularly in contexts like hazard modeling and mitigation planning.

iii) By adjusting for the entropy or uncertainty inherent in model predictions, the logistic output refines the probability estimates of the model. This adjustment ensures that the predictions are statistically sound and reflect realistic uncertainty levels in the model's output.

3.6 Discussion and conclusion

In the current chapter, I started by introducing the Maximum Entropy (MaxEnt) model for developing a multi-hazard susceptibility map, focusing on the Val d'Orléans region. In addition to this, the methodology, data collection, data processing tailored for MaxEnt application were shown and discussed. This sets the stage for **Chapter 4** where I will present the results of the MaxEnt modeling. Our aim in the current chapter was to cover all the necessary data processing steps before applying the MaxEnt model.

Here, I briefly summarize the key steps and findings from the data collection and processing research.

- The MaxEnt model, when properly tuned and optimized, provides reliable predictions of (hazard) distributions. The VIF values obtained are all significantly below the threshold of concern (VIF > 5), indicating minimal collinearity among the predictors when modelling. These results suggest that each variable can independently contribute to our model without causing unnecessary inflation in the variance of the estimated coefficients due to collinearity.
- The model's performance improved by setting the maximum iterations to 1000 and running 50 replicates. This thorough optimization allows the MaxEnt algorithm to accurately reflect the observed distribution patterns of the hazards.
- Random partitioning was chosen for model training and validation which as argued in **Section 3.4.2d**, it facilitates a balanced assessment of the model's discriminatory ability.
- I also argued unlike species distribution models where spatial autocorrelation is often minimized to avoid masking ecological processes, the proximity and clustering of sinkhole occurrences were highly informative. These spatial characteristics were crucial for understanding the conditions conducive to sinkhole formation and for defining accurate risk zones.

Chapter 4. The results of multi-hazard susceptibility modelling performed with MaxEnt

4.1 Introduction

This chapter explores the results of Maximum Entropy (MaxEnt) modelling, structured to convey the findings clearly and comprehensively. The sequence in which these results are presented is designed to gradually build understanding in the model's forecasts. In **Section 4.2.1,** I begin by evaluating the model's overall performance through the Receiver Operating Characteristic (ROC) curve, providing a baseline measure of accuracy. To be able to make the result understandable, I first start by explaining the ROC curve discussing our findings simultaneously.

Following this in **Section 4.2.2**, I present the spatial distribution visualisation of forecasted outcomes using the MaxEnt produced the susceptibility map. Later, in **Section. 4.2.3,** I delve deeper into the influence of both individual (geo)-environmental variables on the models predictions and their influence when all the variables are considered together, presented through (model-specific, marginal) response curves. This is essential to

i) Understand how each environmental factor influences multi-hazard distribution predictions once when assessed considering all the others variable (model-specific curves) and once when assessed alone (marginal curves);
ii) Identify potential optimal ranges of each variables for hazard presence.

This understanding helps make informed decisions based on how changes in the (geo)-environmental variables might affect flood-triggered sinkholes.

Section 4.2.4 assesses the robustness and reliability of these predictions with a Jackknife test (a statistical method that will be explained in detail in that section), highlighting the contribution of each variable. The last but not least, **Section 4.3** spatially compares the official flood hazard map as a single hazard (produced by PPRi, 2015) with our produced multi-hazard susceptibly map. Acknowledging the differences between hazard and susceptibility maps, the rationale behind comparing these two types of maps lies in the interconnected nature of the hazards they represent. Floods do not occur in isolation spatially in karst terrains; they interact with the geology and hydrology of the area, potentially leading to other hazards such as sinkholes. By examining the flood hazard map alongside the multi-hazard susceptibility map, we aim identify areas of overlaps, which is of high importance for:

i) Prioritize resource allocation;
ii) Inform policy and planning;
iii) Enhance Emergency Preparedness.

It should be noted that, to clarify the comparison of hazard maps in this research, areas of overlap and convergence have been extracted from each map (both single and multi-hazard) and are presented in the **Appendix 4.A**. The interpretation of these zones is discussed in **Section 4.3**, but to view the map segments

side by side, please refer to the **Appendix 4.A**. This approach maintains a streamlined narrative in the main text and avoids overcomplicating the flow of result presentation.

Furthermore, the existing body of literature on MaxEnt modelling discuss that that the size of the study area applies a significant influence on the metrics of model performance, notably the AUC score (Sillero et al., 2021). This sensitivity to geographic extent is particularly pivotal as MaxEnt continues to emerge as a critical tool in the modelling domain, especially in applications related to natural hazard forecasts. Given the model's relative novelty and the critical role it plays, it is essential to provide more empirical examples that test its robustness and reliability across different spatial scales.

By doing so, we contribute to a deeper, shared understanding of MaxEnt modeling, which is important not only for the scientific community but also for decision-makers who rely on these forecasts to formulate risk management and mitigation strategies. This comprehensive approach ensures that stakeholders at all levels are equipped with reliable, actionable insights derived from robust scientific evidence. We expanded our analysis to a larger area encompassing the Val d'Orléans and the surrounding 31 communities. For the same above-mentioned reason, to see the result of this refer to **Appendix 4.B**.

4.2 Modelling Results

4.2.1 Model validation

As mentioned in the previous chapter, model validation is essential for determination of the accuracy and predictive power of the produced model in any natural hazard studies (Beguería, 2006). In other words, model validation refers to comparison of the model predictions (forecasts) with the real world data to see how well the model could identify the areas with actual events, which is very important to practitioners (Beguería, 2006).

This process involves several statistical methods (e.g. confusion matrices described by Beguería, 2006) among which the Receiver Operating Characteristic (ROC) graph being one of the most widely used and recommended. ROC analysis was originally developed in signal processing and has become a standard method for evaluation and benchmarking, mainly in medical sciences (Powers, 2011) and recently in natural hazards. Unlike traditional error analysis such as confusion matrices, which provide an exact breakdown of correct and incorrect classifications, the ROC curve offers a broader view of the model's performance across all thresholds_ the threshold (cut-off value) is a value that determines the minimum probability required for a data point to be classified as positive (Beguería, 2006).

ROC summarizes the trade-offs between the **true positive rate** (benefits) and **false positive rate** (costs) (Beguería, 2006; Fawcett, 2006). In other words, the ROC curve does not rely on a single decision threshold to evaluate model performance. Instead, it assesses model performance across all possible classification thresholds. This is particularly useful because it provides a comprehensive view of a model's effectiveness, independent of any specific cut-off (threshold) point that divides predictions into different classes, such as "safe" vs. "unsafe." Therefore, ROC is highly recommended as a suitable tool for decision-making in risk management context (Beguería, 2006).

To add more, the ROC graph is a two-dimensional graph, commonly used method for quantitatively assessing the performance of a diagnostic test (Beguería, 2006; Phillips et al., 2006; Sillero et al., 2021). Therefore, in a simpler term, ROC measures the performance of a model (how well a predictive model distinguishes between the occurrence of a hazardous event and its non-occurrence). In the context of MaxEnt, it represents relative trade-offs between the true positive rate (also called *sensitivity*) and the false positive rate (also called *1 - specificity*) across threshold values ranging from 0.5 to 1.0. In ROC graph, **the vertical axis (y)** illustrates the model's **the true positive** rate which refers to the ability of the model to correctly identify actual positives to the overall actual positives in the data.

It is calculated as:

$$\textbf{True positive rate} \approx \frac{Positives\ correctly\ classified}{Total\ positives} \qquad \text{(q 4.1)}$$

The **horizontal axis (x)** represents **the false positive** rate representing a ratio of incorrectly identified negatives in the data (**Fig 4.1**).

False positive rate $\approx \frac{Negatives\ incorrectly\ classified}{Total\ Negatives}$ (q 4.2)

A simple ROC graph on which each point (A, B, C, D, E) is show in **Figure 4.1** which represents the performance of a classifier at a specific threshold. It can be summarised in the following (Fawcett, 2006):

- **Points A:** shows a performance of a model with *low false positive rate* and *true positive rate*, indicating it is conservative in predicting positive classes.
- **Point B:** a performance *with a higher true positive*. It is thought as satisfactory performance.
- **Point C:** there is a balance between true and false positive rates, representing a more balanced approach between sensitivity and specificity.
- **Point D:** An excellent model performance with a high true positive rate and a low false positive rate
- **Point E:** Both low true positive and high false positive rates, indicating poor performance.

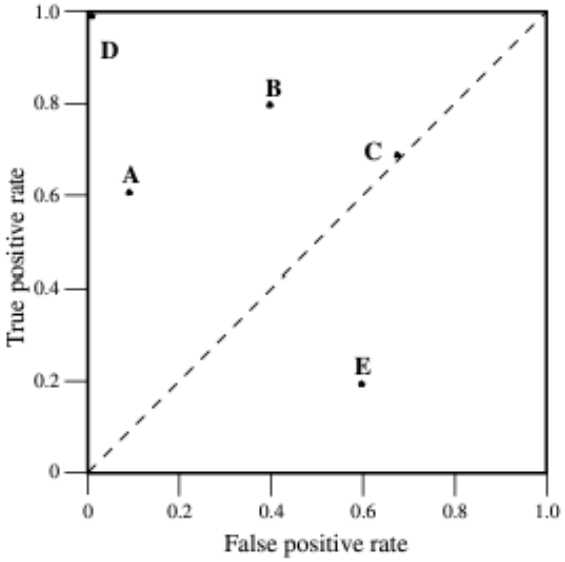


Figure 4.1 A fundamental ROC graph displaying five distinct classifiers (Fawcett, 2006). Each point represents the performance of a classifier with the dashed line representing a random chance. **Point A** shows a performance of the model with low false positive rate and true positive rate; **Point B** represents a balance performance between true and false positive rates; **Point C** shows an excellent performance with a higher true positive rate and a low false positive rate; and Point E represents a poor performance with low true positive and high false positive rates.

The calculation of the AUC for the ROC curve goes as follows (Powers, 2011 take from Schneider and Pontius Jr, 2001):

AUC = $\sum_{i=1}^{n}([x_{i+1} + 1 - x_i] * [y_i + (y_{i+1} - y_i)/2])$ (q 4.3)

Where:

$\mathbf{x_i}$ is the false positive rate for i threshold; $\mathbf{y_i}$ is the true positive rate for i threshold; and n the number of thresholds.

The number of replications is highly depends on the computational resources (the processing power available to execute tasks and the storage space for data) varying from 10 to 50 even 100 replicates (Phillips et al., 2006; Sillero et al., 2021). Here, we executed 50 replications of our model, which, to be more precise, required nearly *three hours* of computing time to complete. In addition to these 50 replications, we also run each replication with 1000 iterations. Iterations refer to the number of times the algorithm processes data within a single run (or model building process). By doing this number of iterations as a modeller, we allow the algorithm takes to converge on a solution, refine its parameters minimising its error. In summary, the entire model was executed 50 times (from the beginning to end with potentially different random seed) and each time was iterated 1000 times. This positively affects the model's stability and the final interpretation of the model output.

The ROC curve for flood-triggered sinkholes is shown in **Figure 4.2**. This curve was plotted using the validation data including training (70%) and testing (30%) datasets. As mentioned in previous chapters, this is important to notice that machine-learning models produce slightly different outcomes each time a model is run with the same data (Sillero et al., 2021; Sillero and Barbosa, 2021). This has been largely overlooked in previous MaxEnt modelling attempts (Mobley et al., 2019; Javidan et al., 2021; Rusk et al., 2022, among many). Therefore, to make more reliable modelling outcomes, it is suggested to replicate the model. The average test AUC over 50 MaxEnt replicate runs is **0.702**, with a standard deviation of **0.055.** This value indicates the ability of the model to accurately distinguish between positive and negative observations within the validation sample. This values means that the model is reasonably good at ranking or distinguishing between places that are in hazard zones and those that are not according to the threshold assessment explained earlier. In **Appendix 4.B** we discuss the reliability of AUC by applying this to a bigger area.

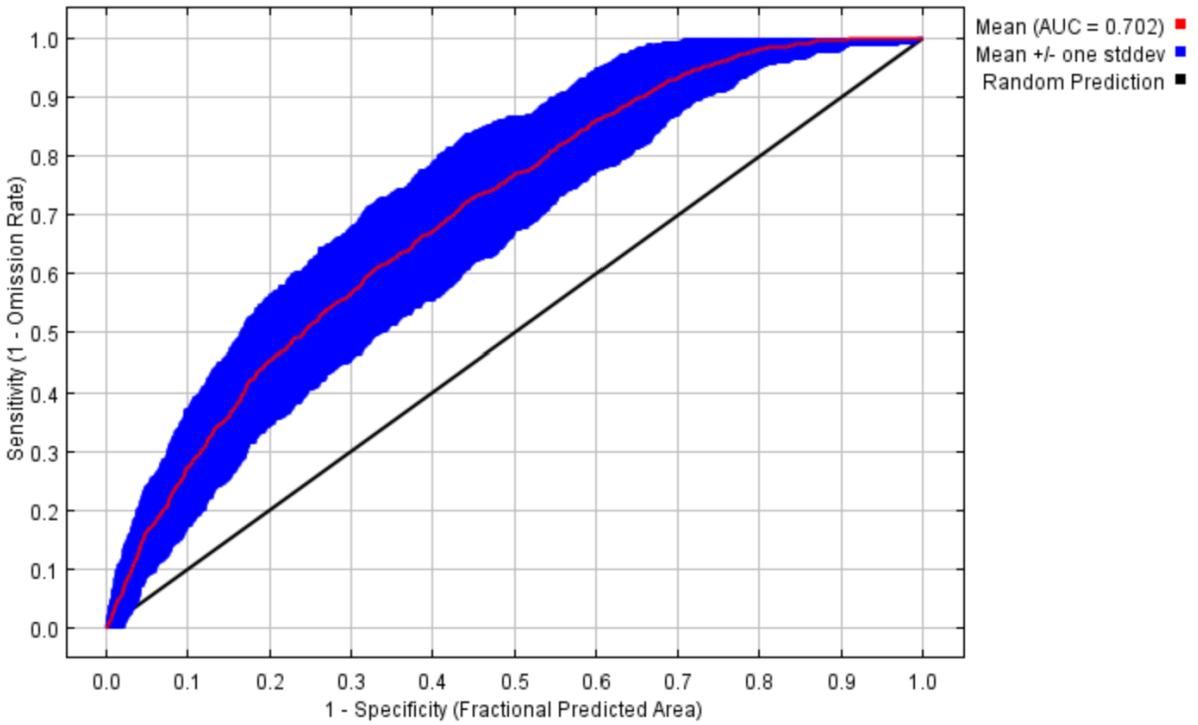


Figure 4.2 The Area Under the Curve (AUC) results of the Receiver Operating Characteristic (ROC) analysis applied to both training and testing data. The black diagonal line depicts all potential cut-off thresholds, which correspond to the cases within the sample. The red curve represents the mean response of the 50 replicate MaxEnt runs, which is **0.702**. The blue shaded area around the mean ROC curve represents the variability of the AUC (+/- one standard deviation).

4.2.2 Multi-hazard susceptibility mapping (the probability distribution)

Based on MaxEnt logistic output, the calculated value of the model is in a range from 0 to 1, with higher values indicating a higher probability of multi-hazard susceptibility. Before discussing the resultant map, it is important to clarify on the hazard level classification and the threshold classifying method.

Once the model is built, a decision threshold or a cut-off value is usually set to divide the continuous response variable into hazard/susceptibility classes. According to Beguería (2006) a continuous map is more informative than a classified categorical scale since final users (decision- makers) could potentially handle better these types of map labelled the most and the least relatively susceptible or safe vs unsafe areas. Having said that, different regulations require different categories. In this study, the map generated by MaxEnt model was reclassified into four susceptible zones: "Low", "Moderate", "High", and "Very high". This classification into four distinct classes was done in accordance with the map classification guidelines outlined in the Plan de Prévention des Risques d'Inondation (PPRi, 2015), the official French document for flood risk prevention planning. Moreover, for the purpose of natural hazards risk management and policy changing encouraging a transition from single hazard to multi-hazard approaches, this classification helps us to have a meaningful comparison between this multi-hazard map with the official flood hazard map identifying the area of overlaps (this will be discussed further in **Section 4.3**).

In terms of identifying thresholds for hazards, generally, two approaches are employed: **i)** an **impact-based criterion**, which include frameworks proposed by entities such as the European Commission in 2011. This method assesses thresholds based on the scale and impacts of a given hazard on various aspects, including societal impacts; and **ii)** data **distribution methods**, such as percentile-based approaches or **clustering methods**. These are based on empirical observations of the data, where thresholds are determined by analyzing the distribution of extreme values, often focusing on the upper percentiles or distinct groupings within the data (Kappes et al., 2012).

In classification of the susceptibility threshold in our study, the second approach was chosen. This is simply due to the nature of this study, which aims for spatial likelihood of the distribution of the hazard (susceptibility maps). Therefore, the Natural Breaks classification was applied. Natural Breaks classification, also known as *Jenks optimization* method, is a data clustering/distribution method that identifies natural groupings inherent in the data for setting thresholds. This method seeks to reduce the variance within classes and maximise the variance between classes (Herries, 2018). It is important to differ this method from impact-based which often rely on the potential consequences of events.

In our classification, the thresholds at 0.19, 0.37, 0.52, and 0.89 ~ 9 (shown in **Figure 4.3 and 4.4**) are points where the algorithm has found a significant increase in variance, indicating a natural break or grouping in our data. In other words, these thresholds are where the data shifts from one cluster to another, creating the most statistically significant classes. The reason of choosing this method is that we are interested in the existing natural patterns of the susceptibility data while also highlighting outliers (Herries, 2018).

The potential (relative) spatial probability distribution of the flood-triggered sinkholes across Val d'Orléans, illustrated in **Figure 4.3 and 4.4,** presents a raster surface with a resolution of 10 m per cell.

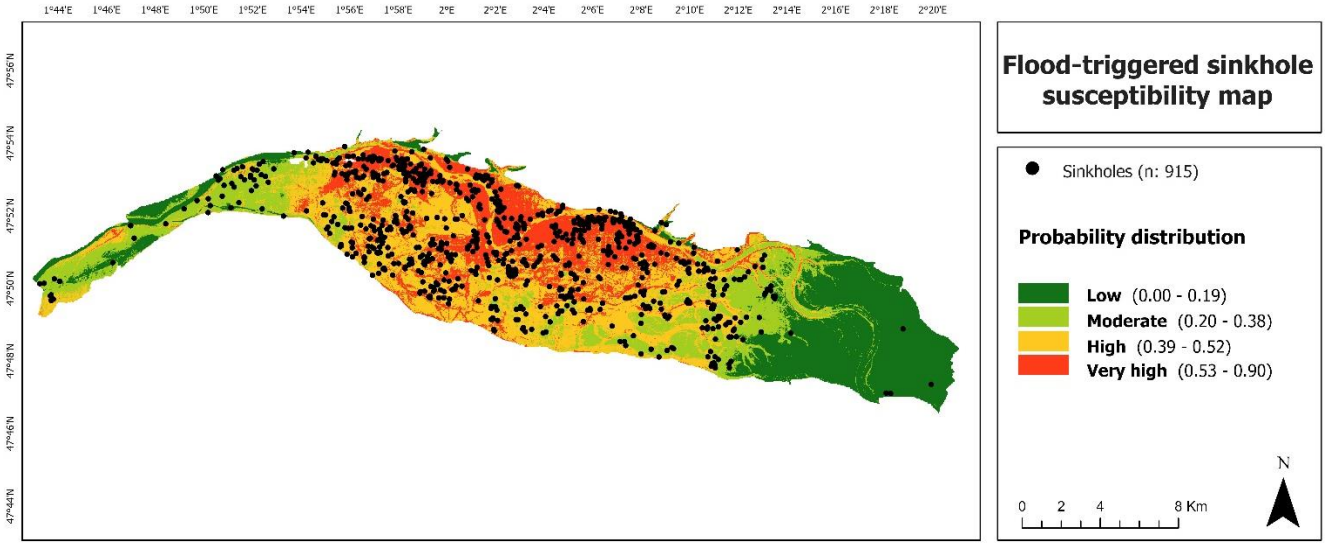


Figure 4.3 Classified multi-hazard susceptibility map generated by MaxEnt model using the Natural Breaks method overlaid with the sinkhole locations. Dark green indicates comparatively low hazard susceptibility, with a relative index of probability distribution ranging from 0.00 to 0.19; Light green corresponds to a moderate hazard level ranging from 0.29 to 0.38. The orange and red colors represent the high and very high susceptible areas with probability distribution of 0.39 to 0.52 and 0.53 to 0.90, respectively.

Table 4.1 Spatial probability distribution of flood-triggered sinkholes in Val d'Orléans.

Susceptibly level and colour	**Proportion of the total area (%)**	**Relative probability distribution**	**Covered areas in Km^2**
Low (Dark green)	23.2	0.00- 0.19	60.4
Moderate (Light green)	21.6	0.20- 0.38	53.6
High (Orange)	33.5	0.39- 0.52	89.6
Very High (red)	21.7	0.53- 0.90	56.5

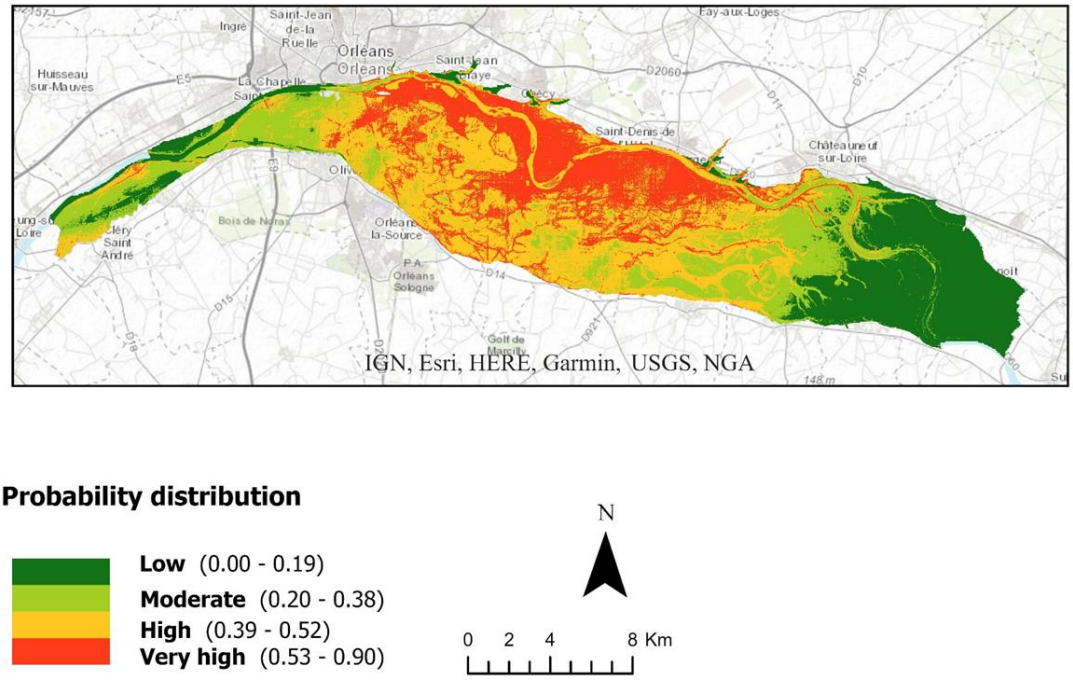


Figure 4.4 Classified multi-hazard susceptibility map generated by MaxEnt model using the Natural Breaks method without the sinkhole locations.

4.2.3 Analysing the impact of contributing factors on MaxEnt modelling

This section provides two sets of response curves: **1)** model-specific response and **2)** marginal response curves. Model-specific response curves and marginal curves are tools used in machine learning to interpret the influence of individual variables on the mode's output (Elith and Leathwick, 2009). While they are not specific to MaxEnt, they are widely used in various machine learning models (linear regression, decision trees, random forests, and neural networks), including MaxEnt.

These response curves are to understand what and how variables contributed to the overall results. The **model-specific response** curves include curves derived from MaxEnt models created using only one variable at a time. These reflect not just the isolated effect of each variable but also capture any dependencies due to correlations with other variables.

On the other hand, **marginal response curves** assess the impact of each variable in isolation, keeping all other variables at their average sample value. They show how the predicted probability of presence changes as we vary one specific contributing factor while holding the rest constant. These curves can help us understand the direct effect of each variable on the model's output. In interpreting the modelling results, the two sets of curves are important to gain a full understanding of the factors influencing the model's forecasting. However, the model-specific responses provide a more specific and realistic results as they reflect the combined effects of multiple variables as they occur in the model. In the following paragraphs, each set of curves will be discussed presenting the resultant curves.

4.2.3a Model-Specific Response Curves (MSRC)

The result of the response curves and a histogram for six conditioning factors used for flood-triggered sinkhole susceptibility assessment is presented in **Figure 4.5**. The curves are derived from MaxEnt models, each built considering only one factor at a time. This approach allows for understanding how a single variable impacts the model, taking into account its correlations with other variables. Please note that to enhance the clarity and readability, the detailed graphical representations of the (geo)-environmental factors have been placed in the **Appendix 4.C**. Each factor is represented by the two curves (MSRC and marginal curves) side by side.

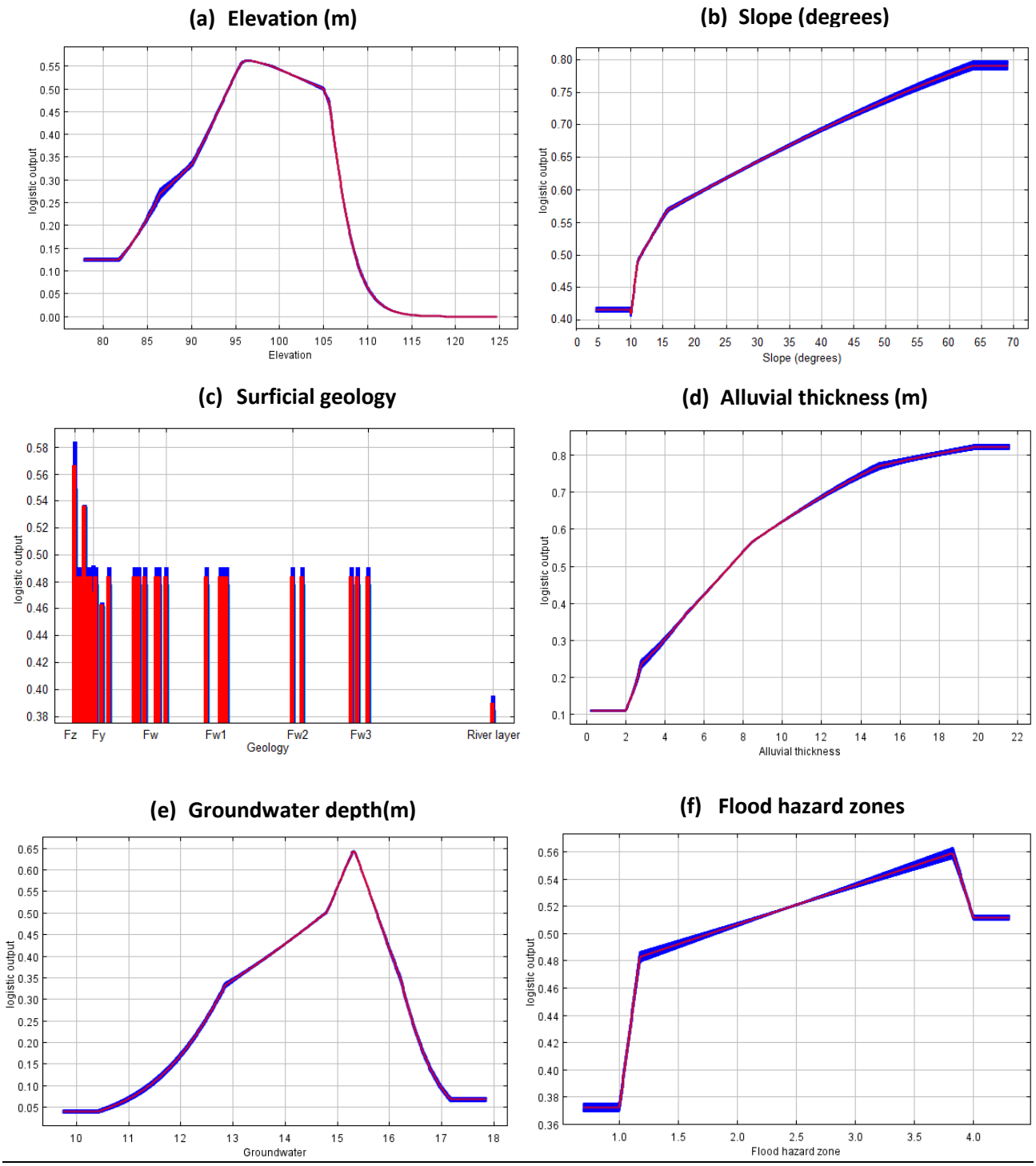


Figure 4.5 MaxEnt model MSRC curves and histogram for flood-triggered sinkhole susceptibility factors (the full figure caption is on the following page)

Figure 4.5 (**expanded version**) MaxEnt model MSRC curves and histogram for flood-triggered sinkhole susceptibility factors. The curves and the histogram show the mean response of the 50 replicate MaxEnt runs (red) and the mean +/- one standard deviation (blue); each graph provides information of the impact of each factor and their correlations with other variables. **a)** Elevation Response Curve: Indicates a peak sinkhole occurrence probability between elevations of 90 m to 105 m. **b)** Slope Response Curve: Demonstrates an increasing probability of sinkholes with increasing slope values, suggesting higher susceptibility in relatively sloped areas. **c)** Geology Histogram: Displays the distribution of sinkhole probability across various geological units, highlighting that the highest value corresponds to the recent alluvium (Holocene) signified the letter "fz"; the other letters are associated with different GIS layers of older alluvium (Pleistocene). **d)** Alluvial Thickness Response Curve: Shows a sharp and steady increase in the suitability of the occurrences as the thickness increases. **e)** Groundwater Response Curve: Reveals a distinct peak indicating the optimal groundwater level range associated with higher sinkhole probability. **f)** Flood Hazard Response Curve: Illustrates the relationship between increasing flood hazard and the probability of sinkhole occurrence, suggesting a positive correlation between flood hazard and sinkhole formation potential.

Looking at the resultant **MSRC graphs** (**Fig 4.5**) it is observed that:

i) Elevation (Fig 4.5a):

- Starting at the lower elevation range on the graph, there is a relatively low probability of susceptiblity, which indicates that flood-triggered sinkholes are less likely to occur at lower elevations.
- There is a distinct peak in the curve at around **95 m** to **105 m**. This indicates the relative optimal elevation range for the hazard presence, where the model forecasts the highest relative probability of finding flood-triggered sinkholes.
- After this peak, the curve drops sharply (from **105 m** to **126 m**), showing that the probability of the hazard presence decreases significantly as the elevation continues to increase beyond the optimal range.
- As the elevation reaches the higher end of **110 m** and above, the probability of presence flattens out to a very low value, suggesting that hazard are very unlikely to be present at these elevations (which are the relatively highest elevation in the area).

ii) Slope (Fig 4.5b):

- The slope curve exhibits a typical logistic (S-shaped) pattern, indicating an increasing probability of hazard occurrences with rising slope values (measured in degrees).
- Slopes up to 10 degrees show a very low probability of flood-triggered sinkhole occurrences.
- Between approximately 10 to 50 degrees, there is a sharp, consistent increase in the probability of hazard occurrence.
- Beyond 50 degrees up to 72 degrees, the probability remains high but stabilized.
- This suggests higher susceptibility in areas with higher slopes, which are primarily located near the river (as observed in **Figure 3.17**).

iii) Geology (**Fig 4.5c**):

- The geological response histogram, presented as a categorical layer, indicates that the highest value corresponds to the recent alluvium (Holocene). The predominant surficial geological layers in the area consist of both recent and older alluvium (Holocene and Pleistocene, respectively).

iv) Alluvial Thickness (**Fig 4.5 d**):

- The curve starts with a very low probability of flood-triggered sinkhole occurrences when the alluvial thickness is minimal (**close to 0**).
- As the alluvial thickness increases, there is a sharp and steady increase in the logistic output up to about **10 m**, indicating a rising probability of hazard occurrence with increased alluvial deposits.
- Beyond this point, the curve shows a more gradual yet consistent increase in the probability of sinkhole occurrence, which continues as the alluvial thickness extends towards **20 m**.

v) Groundwater (Fig 4.5 e):

- There is a low predicted susceptibility for flood-triggered sinkholes when groundwater levels are low (around **10-12 m**), indicated by the logistic output. This result might reflect that insufficient groundwater in the area does not favour the formation or presence of sinkholes. One possible explanation for this threshold could potentially be explained by buoyant effect in geology (relates to the principle of isostasy (**see** Beuthe, 2021); where the presence of water can exert a supportive force, counteracting gravitational pull that might otherwise lead to the formation of sinkholes.
- As groundwater levels increase from **12 m** to around **15 m**, the probability increases sharply, indicating that the probability of the hazard presence increases with higher groundwater levels. This could be due to higher groundwater levels eroding subterranean materials, which may lead to the collapse of the surface and the formation of sinkholes.
- There is a peak at a specific groundwater level (around **15.5 m**), which represents the optimal condition for flood-triggered sinkhole formation in the study area according to the model. It may correspond to the point where the subsurface conditions are most vulnerable to erosion or collapse due to the level of groundwater.
- After this peak (roughly **15.5 m**), as groundwater levels increase further (from **15.5 m to about 17 m**), there is a sharp decline in the logistic output, indicating a rapid decrease in the probability of flood-triggered sinkhole presence at higher groundwater levels.
- At the deepest groundwater levels shown (around **17-18 m**), the probability of presence drops to a level similar to that of the lowest groundwater levels, suggesting that sinkholes are relatively unlikely to be present at these levels.

vi) Flood hazard zones (Fig 4.5 f):

- Starting at **flood hazard zone 1** (Low), there is a sharp increase in the forecasted probability of sinkholes occurring as the flood hazard zone level increases. This suggests that sinkholes are more likely to occur as you move from the lowest flood hazard zone to higher zones and by that it means flood water height over a 100-return period.
- This increase is particularly notable as the flood hazard shifts from **zone 2** (Moderate) up to just below **zone 3** (High).
- After that a continued increase in the probability of sinkhole occurrence as the flood hazard zone moves from **zone 3** (High) to **zone 4** (Very High) is observed.
- There is a slight decline in the sinkhole occurrence probability as the flood hazard zone level increases to **zones between 3.5 to 4** (High to Very high).

4.2.3b Marginal response curves

One key output from MaxEnt is the set of marginal response curves. These curves demonstrate how variations in each geo-environmental variable independently affect the model's predictions for hazard presence, by adjusting one variable while keeping others at their mean values. This provides clear insight into the independent effects of each contributing factor.

Consequently, these curves are essential for understanding how each variable individually contributes to the overall model. We present the marginal response curve for each variable in **Figure 4.6**. Instead of detailing each curve individually, which follow the more or less the patterns observed in the model-specific graphs, we focus on highlighting the distinct differences and unique characteristics between the model-specific and the marginal response curves.

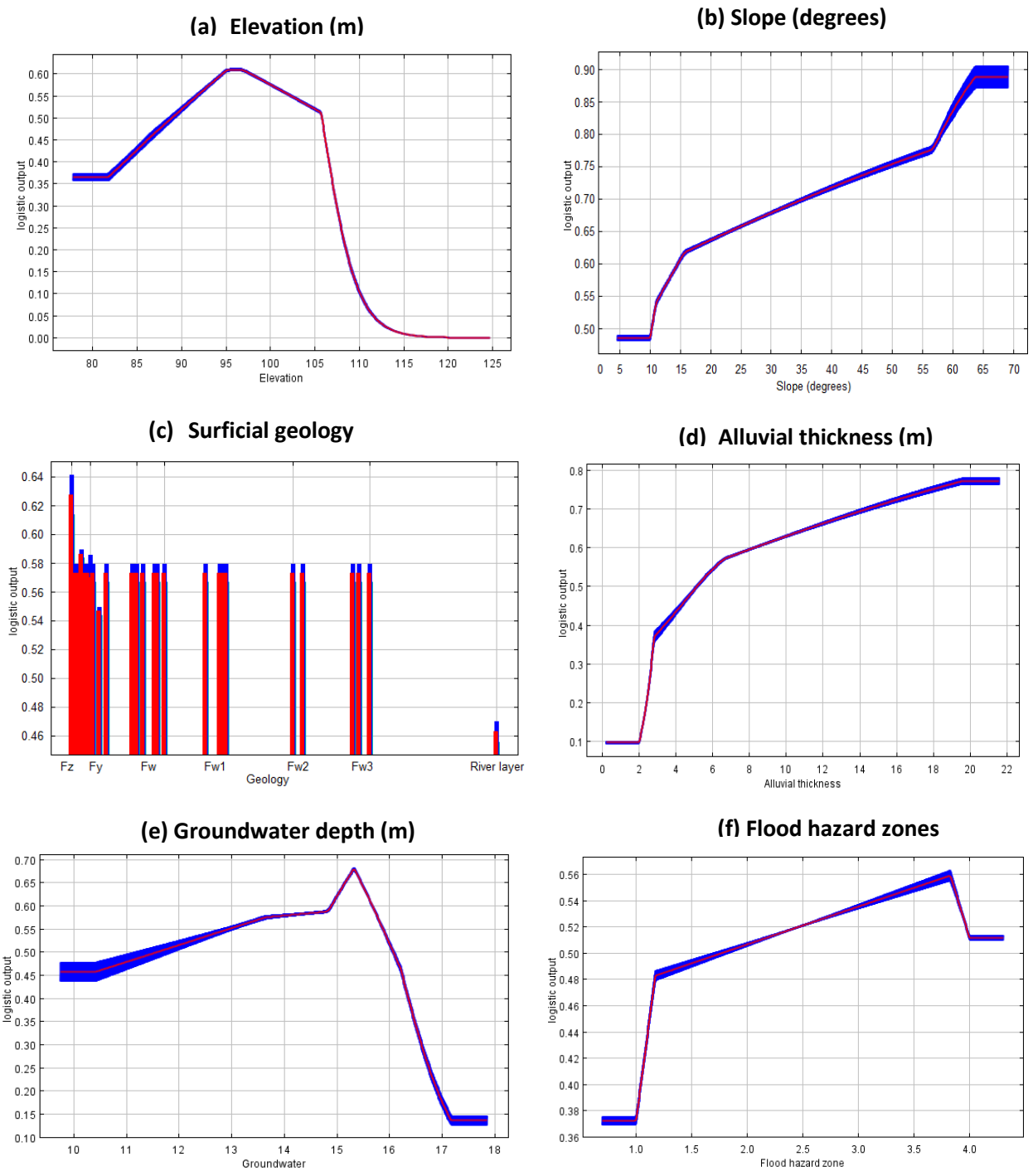


Figure 4.6 Marginal response curves illustrate the variation in the predicted probability of presence when each environmental variable changes, while keeping all other environmental variables fixed at their average values within the sample (with prevalence established at 0.5). a) elevation: showing a low-lying area with elevations ranging from 81.7 m to 126 m; b) slope: depicting a degree of slopes from 0 to 71; c) alluvium

thickness: Quaternary alluvial deposits with a thickness up to 20 m; d) groundwater level: the annual average level of groundwater based on 11 years data; ranging from about 10 m to 17 m; e) surficial geology: consisting of new and old alluvium layers; f) official flood hazard map: classified into four levels of hazard (low, moderate, high and very high) (PPRi, 2015).

In examining this set of graphs (**Fig 4.2**), we do not explain each one by one an as the results look very similar to the model-specific graphs. Instead, we detected a few key differences and characteristics between the model-specific curves and the marginal curves:

In comparison of both sets of curves, a consistency in trends is observed which suggest that their interactions with other variables do not drastically change the result

- Both show a similar trend of increasing thickness, but the marginal curve is more smoothed out, with less variability at smaller thickness values.

The model's peak logistic output, which occurs within the **90-105** m elevation range, suggests that this is the most probable elevation at which sinkholes might potentially occur in the area. These areas fall into the centre of the region.

4.2.4 Analysis of variable contributions

The following table gives estimates of relative contributions of the (geo)-environmental variables to the MaxEnt model. MaxEnt calculates the relative contribution of each (geo)-environmental variable by measuring how changes in each variable affect the model's performance during the iterative training process. Accordingly, MaxEnt provides this information based on two main process: **1)** Regularized gain, and **2)** Lambada Adjustment. In the first process, during each round of iteration, MaxEnt changes the weights of each variable based on whether they increase or decrease the regularized gain. Regularized gain measures how well the model fits the data while penalizing overly complex models. In the second process, if the changes increase the regularized gain, they are added to the variable's contribution score. If the changes decrease the regularized gain, they are subtracted. If Lambda (the regularization parameter) decreases, it means the variable might be adding noise or overfitting, so its contribution is adjusted.

MaxEnt also provides another method to evaluate the importance of each variable called *Permutation importance*. To simply explain this, it is a method used to understand the contribution of each variable to the accuracy of the model by assessing the impact of shuffling or randomly permuting the values of each feature (Phillips et al., 2006).

Permutation Importance works by first training a model on the original dataset and evaluating its performance using a specific metric, such as accuracy, AUC score. Next, the values of one feature are randomly shuffled while keeping all other features unchanged, disrupting the relationship between that feature and the target. The model's performance is then re-evaluated on this modified dataset normalized to percentages. By comparing the performance metrics of the original and perturbed datasets, the importance of the feature is determined (Phillips et al., 2006).

As with the variable jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated.

In **Table 4.2** the percent, contribution and permutation importance is presented. To clarify more, the percent contribution indicates how much each variable contributes to the model's predictive power, while the permutation importance measures the impact on model accuracy (AUC) when the values of that variable are randomly shuffled. High permutation importance means that the model's prediction accuracy is highly dependent on that variable, and thus it is an important predictor within the model.

Table. 4.2 Relative contributions and permutation importance of environmental variables in the MaxEnt Model

Variable	Percent contribution (%)	Permutation importance
Elevation	58.3	63.9
Alluvial thickness	11.7	8.3
Slope	8.4	4.5
Groundwater level	8.3	14.9
Flood hazard zones	7.3	3.1
Geology	6	5.3

- **Elevation**: With the highest percent contribution (58.3%) and permutation importance (63.9%), Elevation is the most influential factor in predicting sinkhole occurrences. This means that changes in elevation considerably impact the likelihood of flood-triggered sinkholes.
- **Alluvial Thickness**: The second most important variable, with an 11.7% contribution and 8.3 permutation importance. Alluvial thickness is a notable predictor, but less critical than elevation.
- **Slope**: Contributes 8.4% to the model with a 4.5 permutation importance. Slope has a moderate impact on the predictions.
- **Groundwater**: Contributes 8.3% with a higher permutation importance of 14.9. Interestingly, its permutation importance is relatively high (higher than Alluvial thickness). This indicates that groundwater levels are relatively important in predicting sinkholes, as changing these values notably impacts the model's performance.
- **Flood Hazard Zone**: Contributes 7.3% with a permutation importance of 3.1. It has a lower impact compared to the above variables but still plays a role in the model.
- **Geology**: With the lowest percent contribution (6%) and permutation importance (5.3), Geology is the least influential predictor among those listed, but it still contributes to the model.

4.2.4a Jacknife test implication

In **Figure 4.7,** the results of the Jacknife test is presented. Jacknife test is a statistical method also known as the "jackknife resampling" or "leave one out method". This test is primary used for estimating the bias and variance of a statistical estimate when the sample size is small. It works by systematically recalculating the estimate leaving out one or more observations at a time from the sample set (Abdi and Williams, 2010). In the context of machine learning and specifically in MaxEnt modelling, it is applied to feature importance estimation, which is done by evaluating how sensitive the model is to the sequential removal of each feature (Philips et al., 2006). It is also similar to permutation importance. MaxEnt provides both *permutation importance* and *jackknife test* results because these metrics offer complementary insights into the importance and stability of features within the model. Each method evaluates feature importance in a different way, and using both can provide a more broad understanding of the model's behaviour and the role of each variable.

It involves: **i)** running the model multiple times, each time a variable is left out to see how the model performance metrics like AUC change (that variable will be used to run a single model); **ii)** comparing the performance of these models to the one using all variables. This test helps in understanding which variables have the most influence on the predictive ability of the model which is particularly important for disaster risk management and urban planning. Abdi and Williams (2010) explain the basic definition of jacknife application through several formulas. Here, to understand general formulas, they are briefly explained.

1) **Full sample estimate (T)**:

$T = f\ (X_1, X_2, ..., X_N)$

This represents the parameter estimate using all observations in the sample.

2) **Leave-one-out partial estimate (T_n)**:

$T_{-n} = (X_1,..., X_{n-1}, X_{n+1},...,X_N)$

Here, each T_{-n} is computed by excluding the *nn*-th observation, providing an estimate based on all data except one point.

3) **Pseudo-values (T^*_n):**

$Tn* = N \times T - (N-1) \times T-n$

Pseudo-values are calculated for each observation as the difference between the scaled full sample estimate and the scaled leave-one-out estimate. These values adjust for bias introduced by omitting each data point.

4) **Jackknife estimate (T*) :**

$T* = \frac{1}{N} \sum_{n=1}^{N} T * n$

The overall Jackknife estimate is the average of the pseudo-values, providing a more robust estimate of the parameter.

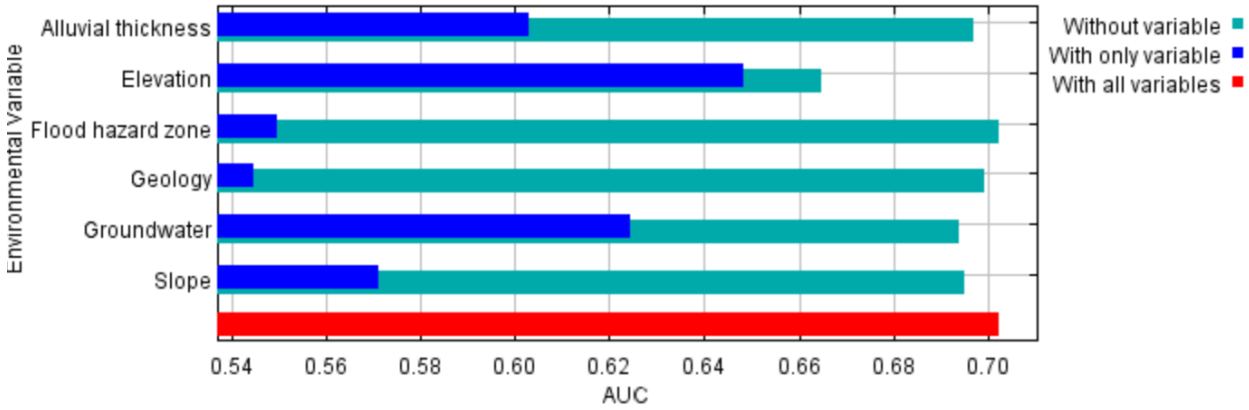


Figure 4.7 The plot of the jackknife test results of the averages over 50 replicate runs contributing factor importance. The green bar represents the AUC of the model without the corresponding variable, while the blue bar shows the AUC processed with only the corresponding variable. The longer the blue bar, the more important is the variable for model probability distribution. The red bar represents the AUC of the final model considering all the variables.

4.3 Comparative analysis of single Vs multi-hazard maps (Overlaps and Divergences)

For the purpose of sustainable land use and natural hazard mitigation planning in Val d'Orléans, we compared and visualized the spatial extents affected by flood as a single hazard and those susceptible to flood-induced sinkholes, representing a complex multi-hazard scenario. It is important to acknowledge the inherent differences in the nature and methodologies employed in hazard mapping (considering magnitude and frequency) and susceptibility assessments (spatial probability distribution). However, we have focused on **comparing the spatial overlaps** by the means of classification approach representing an effort towards standardization, which has been introduced as a common approach to partially overcome the existing challenges in comparison hazards within multi-hazard studies (Kappes et al., 2012).

The extent to which different areas are affected by flood as a single hazard (the official map produced by PPRi, (2015) and susceptible to sinkholes triggered by flooding as a multi-hazard is visualised in **Figure 4.8**. This side-by-side presentation of maps is crucial for a direct visual comparison, allowing stakeholders to observe not just the flood hazard map but also the **potential overlaps and divergences** between the two (flood-triggered sinkholes).

In **Figure 4.9** I further extend this analysis by providing a bar chart that displays the percentage of the total area each hazard class occupies, alongside the corresponding area in square kilometres (km^2). This visual representation is designed to make the data more accessible and interpretable. For each hazard classification (Low, Moderate, High, and Very high) the bar chart displays the proportion of the area affected by floods and the area susceptible to flood-triggered sinkholes. By examining the bar chart, stakeholders can distinguish at a glance where the important discrepancies lie, such as the marked difference in the Moderate category, where flood hazards cover a considerably larger proportion of the region compared to areas susceptible to sinkholes. Conversely, the bar chart also reveals where there is a close alignment in percentages, particularly in the Very High category, suggesting a potential overlap in the spatial distribution of the two hazards.

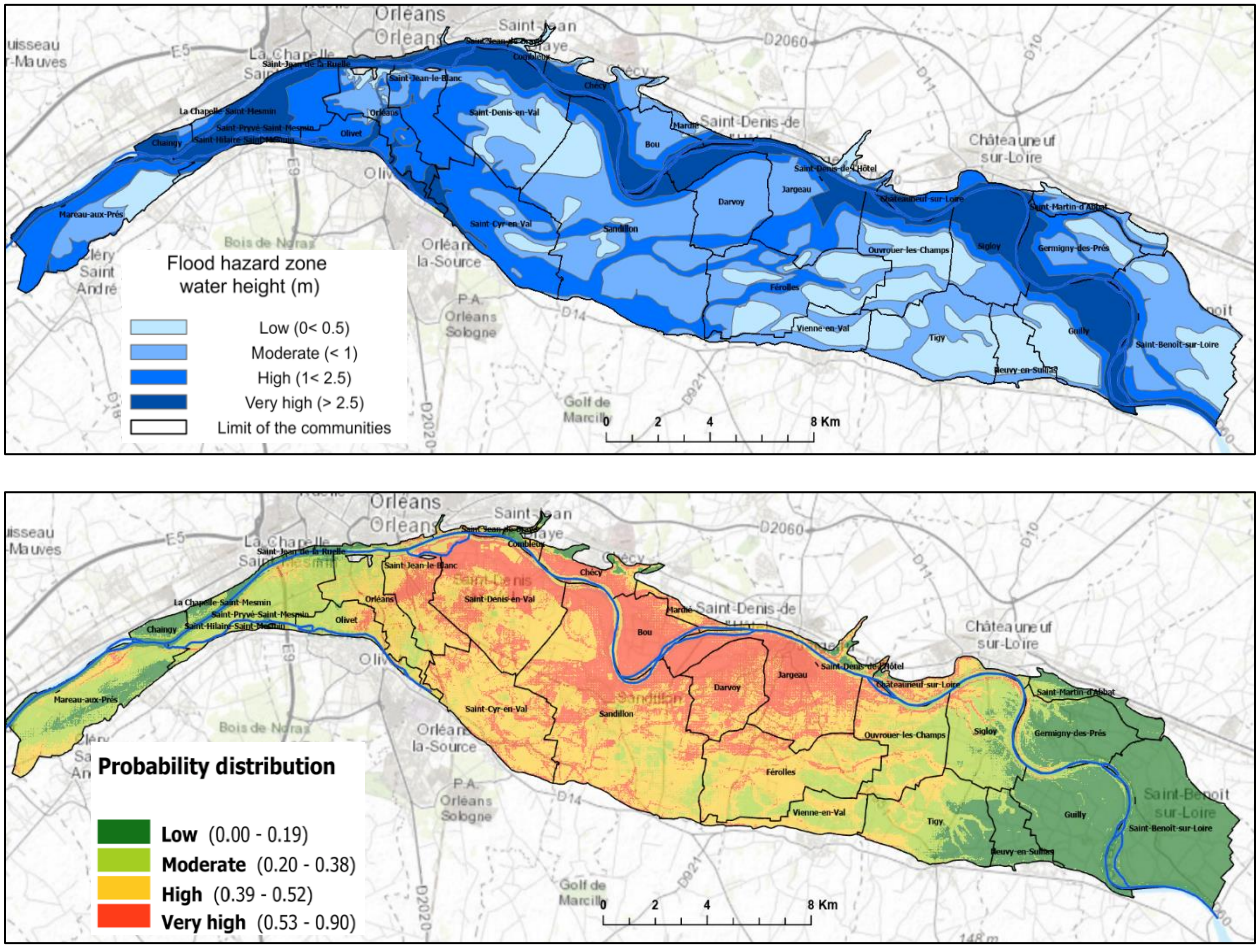


Figure 4.8 Single vs. Multi-Hazard Comparison. The top map shows the extent to which different areas are affected by flooding as a single hazard (the official map produced by PPRI, 2015), while the bottom map indicates areas susceptible to sinkholes triggered by flooding, illustrating a multi-hazard scenario.

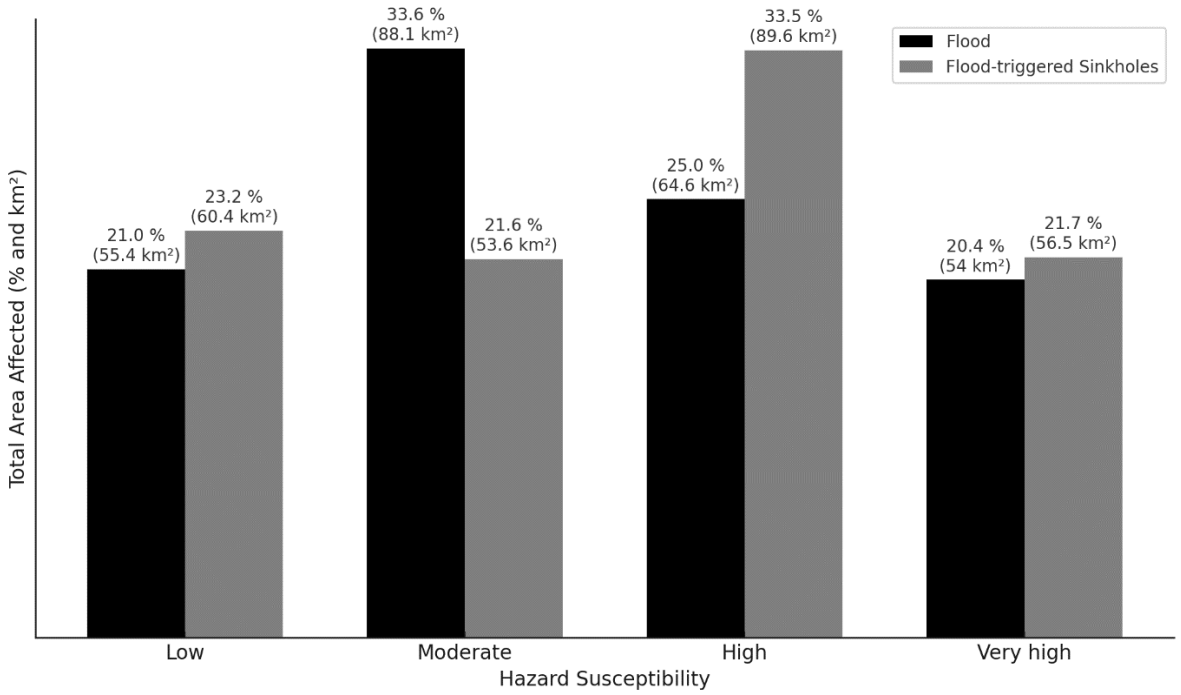


Figure 4.9 Spatially comparing total areas covered by different susceptibility classes for two related yet distinct hazards in the study area: flood and flood-triggered sinkholes.

Analyzing the maps (**Fig 4.8**) and the bar chart (**Fig 4.9**), we observe:

i) **Areas of High and Very high overlaps.** Areas exhibiting both High and Very high overlaps in flood hazard and flood-triggered sinkhole susceptibilities, such as those observed along the Loire River, require heightened attention. These zones are not only threatened by severe flooding but are equally prone to sinkhole occurrences. Such dual-risk areas, exemplified by the commune of *Chécy*, which was severely impacted by both floods and sinkholes in 2016, necessitate prioritized interventions. In *Chécy*, it is evident that the risks extend beyond high flood levels to include high susceptibility to flood-triggered sinkholes. Authorities should use these insights to prioritize and tailor mitigation strategies, particularly in areas along the Loire River where such overlaps are pronounced. Details of these observations and specific comparisons for Chécy are further elaborated in the appendix with a distinct zoomed comparison (See **Figure 4.A.1**). Moreover, in the municipality of *Saint-Jean-le-Blanc* a High to Very high level of multi hazard is observed. This area is particularly highlighted by the Montission sinkhole (**Appendix E.3**), which has been known for many years but was the subject of a detailed investigation in 2019. The sinkhole poses a significant sinkhole hazard, as detailed in the BRGM expert report (Noury, 2020).The proximity of the sinkhole to urban structures and critical infrastructure adds to the potential severity of its impact in the event of extreme events such as heavy rainfall and flooding.

ii) **Discrepancies in Susceptibility/hazard Classes.** For areas of "**Very High**" susceptibility, the proportion of total areas covered by floods and sinkholes converges (**20.4% for floods** and **21.7% for flood-triggered sinkholes**). The "**High**" susceptibility class shows a considerable

difference between covered areas by the two hazards, with a higher percentage of areas susceptible to **flood-triggered sinkholes (33.5%)** than to **flooding (25%)**. This divergence helps to clarify that not all flood-prone areas with High and Very high hazard levels face equal susceptibility to flood-triggered sinkholes, emphasizing the need for localized assessments. For example while the majority of Saint-Pryvé-saint-Mesmin falls into High and Very high flood hazard, it is mainly fall into Moderate level of multi-hazard (See **Figure 4.A.2**). This also takes us to the following point:

iii) Despite the interrelationship between floods and sinkholes, **they do not always coincide spatially**. Both maps and the bar chart reveals that these hazards do not always follow **a spatially linear trend but it is also clear that there are parts where the non-linear relationship exist**. The map indicates that Low flood hazard areas may still be Highly susceptible to sinkholes. For instance, while *Saint-Denis-en-Val* lies in a mainly Low and Moderate flood hazard zone, it has a High to Very high susceptible level for sinkholes, particularly in areas not necessarily consistent with the highest flood risk zones (See **Figure 4.A.3**). This suggests that while some regions may appear less prone to flooding, they still possess a significant risk of ground instability due to other geological factors. Similarly, *Ouvrouer-les-Champs*, though in a low flood hazard area, is Very highly to Highly prone to sinkholes (See **Figure 4.A4**). This underscores the fact that even areas with a lower perceived flood hazard can be susceptible to other hazards, in this case, sinkholes. The bar chart reveals that in total **21%** of the region has a **Low flood** hazard, with **23.2% Low susceptible to multi-hazard**.

iv) **Moderate flood hazard** zones encompass **33.6%** of the region, while **21.6%** is **moderately susceptible to flood-induced sinkholes.** Notable examples include *Darvory* and the town of *Bou*. The former is classified under Moderate flood hazard yet exhibits a Very high susceptibility to flood-triggered sinkholes (See **Figure 4.A.5 (top panels)**). The latter is located in the east of the *Orléans Metropolis* within the Loire's last main meander, where it is entirely subject to a Very high multi-hazard susceptibility (See **Figure 4.A.5 (bottom panels)**). In contrast, most parts of *Jargeau* to the north and west are Moderately at risk of flooding, but the whole area ranks High to Very high for combined flood and sinkhole susceptibility (See **Figure 4.A.6**).

v) The **spatial overlap** between the flood hazard zones and areas with higher susceptibility to multi-hazard/susceptibility (flood-triggered sinkholes). A large parts of the area, which are mainly along the Loire River and the central part of the region are relatively spatially overlapped with flood hazard zones in particular the high and very high hazard level with multi-hazard/susceptibly map covering more areas with high and very high susceptible level.

vi) Even though across all the study area, flood hazard is High to Very high alongside the Loire River, the results shows that not all parts all along the river are necessarily Highly susceptible to flood-triggered sinkholes. This highlight the role of other factors such as elevation. In higher elevations as seen in the eastern part, the flood-triggered sinkholes are potentially less likely to happen. Towns/communes including *Saint-Martin-d'Abbat*, *Germingy-des-Prés*, *Guilly*, *Neuvy-en-Sullias* and *Saint-Benoit-sur-Loire* are situated in the east part of the region with a mixture flood hazard level from Low to Very high, while the model result shows they are

situated mainly in Low susceptibility level of flood-triggered sinkholes (See **Figure 4.A.7).** Similarly the exact difference can be found along the river in the far west of the area (See **Figure 4.A.8**). Having said that, the data indicates that areas with a High level of flood hazard should be closely evaluated for their sinkhole potential, emphasizing the need for integrated hazard assessments and mitigation strategies that consider both surface and subsurface conditions. It is an indicator that susceptibility to one type of hazard does not automatically imply a similar level of susceptibility to another, even if they are related. This nuanced understanding can be used to refine emergency planning and public awareness campaigns, ensuring they are tailored to address the specific risks present in these moderate-risk areas.

4.4 Discussion and Limitations

The MaxEnt method was applied to model susceptibility to flood-triggered sinkholes, achieving a satisfactory result with a predictive accuracy of 0.702 for AUC. The model's spatial distribution of relative probabilities indicates that approximately 55% of the study area (about 146 km^2) falls within the High and Very high susceptibility categories.

Elevation appears to be the main variable in detecting areas of higher susceptibility, particularly at elevations ranging from 90 to 105 m. Additionally, areas with groundwater levels of about 15 m appear to be relatively more prone to sinkholes. This underlines that in the Val d'Orléans, characterized by its flood plains overlying karstic bedrock, factors such as elevation and groundwater levels significantly contribute to the likelihood of sinkholes occurring during flood events. Furthermore, areas within the region that have relatively steeper slopes (above 20 degrees), greater alluvial thickness (about 14 m to 16 m), covered by recent alluvial (Holocene) units, and higher flood hazard scores, which correspond to higher water levels are more susceptible to flood-triggered sinkholes.

The resultant jackknife test in MaxEnt outputs should be interpreted with caution as this test provides relative importance values that may not fully capture the multi-dimensional interactions among variables or the model's transferability to different areas and conditions. Therefore, in our modelling, flood hazard zones and surficial geology datasets emerged as the least impactful factors compared to others. This cannot be directly interpreted as less important for flood-triggered sinkhole susceptibility. Instead, it may indicate that within the context of our dataset and model structure, these variables are overshadowed by more dominant variables. Moreover, this observation could be partly attributed to the spatial overlap of flood hazard zones and surficial geology factors as indicated by their slightly moderate correlation in the Variance Inflation Factor (VIF) test (discussed in **Chapter 3**). However, as the model response curves (both model specific (**Fig 4.5**) and marginal curves (**Fig 4.6**)) show, these two factors contribute the entire model performance.

Our findings of the spatial distribution of flood-triggered sinkholes reveal the largest part of the area falls into High and Very high multi-hazard susceptibility (55.2 %). Areas with the highest susceptibility are primarily located along the Loire River and close to the Loiret spring, pinpointing these zones as critical regions of concern for potential multi-hazard occurrences. **This pattern of susceptibility is underpinned by historical hydrological events; notably, Martin et al. (2003) documented a significant event in 1907 where flooding led to the Loire River's backflow into the Loiret.** Consequently, the spatial analysis delineates the historically affected areas by this backflow between the Loire and Loiret as areas with Very

high to High susceptibility, which also shows the traces an extensive karstic cave network. Meanwhile, the western sections of the study area (approximately ranging from longitude 1°43'E to 1°54'E between latitudes 47°50'N to 47°54'N) are classified as having moderate to Low susceptibility, although sporadic areas of Very high and High susceptibility exist near the Loire River, suggesting localized areas of susceptibility. In contrast, the eastern parts (approximately ranging from longitude 2°14'E to 2°20'E between latitudes 47°53'N and 47°48'N) are characterized by relatively Low susceptibility, largely attributed to higher elevations (about 110 to 126 m) which likely mitigate the impact of flood and sinkhole formations. The highest susceptibility probability in the most susceptible areas is a relative probability value of 9 and the lowest susceptibility probability is a relative value of 0.0001 (~ 0.00) suggesting that the model forecast a very low, but not necessarily zero of susceptibility. This information has relevance for policymakers and disaster risk management officials when prioritizing resources and efforts for areas with the highest risk, while not neglecting the lower risk areas, which still face a non-negligible probability of hazard susceptibility occurrence.

Upon analysing the comparative chart of the flood hazard and multi-hazard susceptibility maps, it becomes evident that the areas susceptible to flooding do not always correlate directly with those susceptible to flood-triggered sinkholes, which suggests a spatially non-linear relationship. It is an indicator that susceptibility to flood hazard does not automatically imply a similar level of susceptibility to flood-triggered sinkholes, even when they are related. A key observation that the data indicates is that areas with a high and very high level of flood hazard should be closely evaluated for their sinkhole potential, emphasizing the need for integrated hazard assessments and mitigation strategies that consider both surface and subsurface conditions. This nuanced understanding can be used to refine emergency planning and public awareness campaigns. This work strongly advocate for a multi-hazard framework in risk assessment and disaster preparedness within the karst context. The data indicates that while flood management is essential, it must be coupled with an understanding and anticipation of secondary hazards, such as sinkholes in kart terrains that may be triggered or exacerbated by flooding. This is particularly crucial in areas of “High” and “Very high” flood hazard zone, where the hazard of sinkholes is evidently increased.

The data indicates that areas like Chécy, with a history of severe multi-hazards, should be the focus of integrated hazard and risk assessment. This approach combines flood mitigation with geo-technical assessments to reduce the overall risk profile. With limited resources, authorities need to prioritize intervention where the overlap of hazards is greatest. Historical data, such as the 2016 events in Chécy, should guide future preventative measures.

In summary, areas along the Loire River showing High and Very high overlaps of flood hazard and sinkhole susceptibility require critical attention for integrated risk management. However, not all flood-prone areas exhibit a similar susceptibility to sinkholes, emphasizing the importance of localized ground investigations/ assessments. Discrepancies between the flood hazard zones and sinkhole susceptibility highlight the need for nuanced strategies that consider both surface and subsurface conditions. In communes where these hazards do not align spatially, such as areas in the far East and West, policies should address the unique geological risks of each area before any further development. Overall, this study underscores the importance of integrated hazard assessments to effectively manage and mitigate the diverse risks present across the region.

However, it is important to reflect on **a few limitations** that introduce some degree of uncertainties into our modelling of susceptibility.

i) The quality of predictive models depends heavily on the completeness and representativeness of the sinkhole inventory (epistemic uncertainty) (Galve et al., 2009; Taheri et al., 2019). Specifically, for mapping the susceptibility of flood-triggered sinkholes, having an inventory that documents only sinkholes opened up by flood hazards might yield more accurate results. Nevertheless, obtaining this information typically presents challenges, such as data scarcity and potential limitations in coverage. This also underscores the importance of recording the triggers along with their respective locations.

ii) Another limitation worth mentioning is that in an area characterised by a high flood hazard, particularly in the context of flood-triggered sinkholes, it becomes evident that there exist localized geological features that pose significant challenges for inclusion in modelling efforts. These localized factors are primarily related to the complex geological characteristics of the region. For instance, certain areas may exhibit distinct lines of geological weakness and fractures that facilitate the movement of water. Consequently, water is more likely to erode materials in these areas, increasing the risk of sinkhole formation. Integrating such fine-scale geological information into a model can be challenging, as it requires highly detailed geological maps with fine resolution. These features represent specific susceptibilities that may not be adequately accounted for in standard modelling approaches like MaxEnt. As a result, a more in-depth investigation, such as detailed mapping and site-specific assessments, becomes essential for gaining a comprehensive understanding of the likelihood and precise locations of flood-triggered sinkhole occurrences, especially in areas with complex geological characteristics.

iii) As the spatial overlaps do not guarantee that the secondary hazard will be triggered, the susceptibility maps might be seen as a limitation. However, paying attention to a couple of points is important:

1) Sinkholes are hidden hazards and detecting them in a regional scale, if not possible, is a very hard work, therefore, in the literature, sinkholes are commonly assessed in terms of their probability distributions for regional mapping purposes.

2) In the case of flood and sinkholes, (geo)-environmental conditions such as elevation, slope and geology coincide. These conditions should be met so that a secondary hazard happen. In case of karst areas in particular Val d'Orleans where is situated in a low-lying and flood plain areas, the fact of being karstistic can more or less increase the probability that in case of flooding we should expect sinkhole collapses because of the (geo)-environmental conditions overlap.

iv) MaxEnt Limitations:

- **Influence of study area extent on AUC values:** As observed by comparing two different geographical scales (**Appendix 4.B**), the AUC value can be significantly influenced by the extent of the study area. This can present a notable challenge, as larger areas tend to inflate AUC values, potentially leading to misleading interpretations of model accuracy. For instance, a high AUC value might be perceived as indicative of excellent model performance, while this may not accurately reflect the predictive capability of the model. The issue arises particularly in vast study areas where those hazards of interest are known to be present only in a specific region. As seen in evaluating the model across the entire Val d'Orléans plus 31 communites (cities) around obtained an inflated AUC value compared to a more focused assessment within the Val d'Orléans. The higher AUC value likely results from having more data points, greater environmental diversity, and better input representation, all of which contribute to the model's enhanced ability to generalize and make accurate predictions.

- To mitigate this, it is advisable to tailor the study area closely to the known distribution of given occurrences. This can provide a more accurate evaluation of the model's performance, while ensuring the AUC value reflects a realistic measure of predictive accuracy.

- We also acknowledge that despite our careful and informed approach in tuning parameters for the MaxEnt model, there remains an inherent uncertainty associated with tuning the model parameter settings.

- One main issue that is not only limited to MaxEnt but also prevalent in many machine learning models is the issue of "Spatial Autocorrelation". In the MaxEnt literature the issue of *Spatial Autocorrelation* has been highlighted by Sillero and Barbosa (2021) and Sillero et al. (2021). Spatial autocorrelation is a property of the first law of the Geography known as Tobler's Law "everything is related to everything else, but near things are more related than distant things" (Tobler 1970). This means that closer locations tend to have more similar values. However, for natural hazards such as sinkholes the closer locations give us better information. As for selected contributing variables, when they are spatially auto-correlated, their values are not independent across space which can result in misleading variable importance.

- Finally, it is important to mention that in machine learning, well-constructed data is crucial for calibrating (training) and validating (testing) models. Previous research suggests that no single model or algorithm consistently outperforms others across all scenarios (Tilloy et al., 2019; Piao et al., 2022), due to the inherent assumptions and simplifications each model makes. However, it is still beneficial to apply multiple machine learning models for the specific area to gain a better understanding of their performance and help identify the best-performing ones for specific applications.

4.5 Conclusion

In this study, we produced a forecasting map applying a less explored machine-learning model called Maximum Entropy (MaxEnt) in natural hazards specifically in multi-hazard application to identify the area susceptible to flood-triggered sinkholes. To build our model, we have chosen six main contributing factors/variables explaining both flood and sinkhole contributing factors. To advocate an approach in which we do not simply overlay different hazard on top of each other, we integrate the official flood hazard zones into our modelling identifying recognising the flood hazard as a triggering contributing factor to sinkhole collapses.

The multi-hazard map generated through this study serves as a communication tool to support decision makers involved in resource allocation and risk mitigation efforts. Additionally, it can be used for raising awareness among the public about the diverse risks they may encounter, fostering a culture of preparedness and resilience within community.

It should be mentioned that while the probability of simultaneous occurrences may be low, the consequences could be disproportionately high, requiring a proactive and preventive approach to community safety and disaster risk reduction. However, since multi hazards are rather recent practices, along with more in depth investigations, further research is essential to seek stakeholders' perspectives on the utility of multi hazard maps. This is also necessary to ensure that the generated maps can effectively address challenges in decision-making processes.

The analysis of the presented data underscores the importance of a multi-hazard approach to risk assessment in the studied region. While traditional flood maps provide critical information on water-related risk, the inclusion of sinkhole susceptibility reveals a more complex risk landscape that could have been overlooked otherwise.

We also applied MaxEnt model for a broader area to test the mentioned by hypothesis by experts which says AUC value will be inflated in bigger areas. We ran the model for the Val d'Orléans and the surrounding areas (613.53km^2) using the same spatial data however by extending the area the geographical conditions will become different meaning that (geo)-environmental spatial data change which resulted in a different evaluation of the factor importance. Therefore, from this exploration it became evident that the more you zoom out, the more variety in (geo)-environemtal conditions appears. That can lead to the inflation in the AUC result and this could be misleading as the modeller gets the high score result but in fact it is inflated by the extend of the area. The higher AUC value likely results from having more data points, greater environmental diversity, and better input representation, all of which contribute to the model's enhanced ability to generalize and make accurate predictions.

The solution came up by discussing with experts is that if you know a given hazard point does not exist for the entire area or they are restricted in particular areas it is better to model the focused area. For example the whole part of the department vs only Val d'Orléans.

Finally, in terms of mitigation planning purposes, our findings underscore the critical importance of considering both flood and sinkhole hazards in the Val d'Orléans. The multi-hazard susceptibility map provides essential insights that can transform how we approach risk reduction and resilience building in this region.

Chapter 5. Creating a road map to strengthen multi-hazard mapping for decision support

5.1 Introduction

(Multi-) hazard/susceptibility maps are used to communicate complex scientific data with stakeholders. These maps serve as a communication tool to support and aid decision makers by providing valuable information regarding mitigation measures (Komendantova et al., 2014). They pinpoint areas where events of a particular probability are anticipated, enabling authorities to proactively direct community hazard mitigation and emergency preparedness measures including evacuation plans well in advance of any imminent threat (Lindell, 2020; Clive et al., 2021).

Synthesizing complex scientific data, including interconnected hazards, and visually representing these through zones of low to high hazard levels on a map requires considerable expertise and sophisticated analytical tools. Achieving a comprehensive understanding of hazard chain assessments, such as multi-hazard maps, is crucial. Equally important is the ability to effectively convey and transfer the information embedded in these maps to stakeholders. In essence, it is not only about the creation of complex multi-hazard maps but also about ensuring that the insights derived from these assessments meet the needs of decision makers and are successfully communicated to those who need them (Gill et al., 2021). Therefore, at the heart of this process lie the stakeholders, who vary from organizations primarily engaged in spatial planning, risk reduction strategies, and emergency preparedness and response, to different level of decision-makers and citizens. These entities operate within a country governed by distinct legislative and planning procedures, each with its own unique mandate (van Westen and Greiving, 2017).

However, the impact of multi-hazard visualizations on decision-making contexts remains unclear. Moreover, the perceptions of various stakeholders (including decision-makers, experts, and societal stakeholders) towards these maps are crucial. These perceptions affect their understanding of spatial hazard distribution and, consequently, their decisions regarding resource allocation and risk mitigation. Engaging with stakeholders entails a series of interactions that enable us to gain a deeper insight into various perspectives, criteria, preferences, and the choices we must make when determining a course of action (Komendantova et al., 2014). Translating hazard science to policy and practice is a recognised challenge for the hazard community (Šakic Trogrlic et al., 2022). Participatory approaches, such as collecting feedback from stakeholders and including them in the evaluation process of an output such as hazard map, facilitates the translation of theory into practice, bridging the gap between scientific research and practical application (Komendantova et al., 2014).

In the region of Orléans (Val d'Orléans), research by Serrhini et al. (2023) examined the main stakeholders' perception on flood and sinkholes related to the Val de Loire flooding. The aim was to improve institutional communication document on major risks (DICRIM) and highlight karst risks in connection with flooding.

A survey was conducted, seeking input from key stakeholders, including local authorities and governmental bodies, in three main sections: (1) the perceptions and considerations regarding the karst collapse hazard, (2) obtaining an overall assessment of the DICRIM as a communication tool for the population, and (3) evaluating municipal maps involved overlaying the locations of past sinkhole events on different backgrounds. The results from 26 respondents identified a need for improved institutional communication and non-structural preventive measures to enhance territorial resilience against flood and sinkholes. The results also highlighted a general awareness of flood risks but a lack of specific focus on karst collapse risks among key actors. The research underlined the importance of integrating multi-hazard risks into local planning and communication strategies to better inform and protect the community.

Building upon these findings, this thesis seeks to delve deeper into the realm of multi-hazard risk management by not only assessing the effectiveness of multi-hazard susceptibility maps but also exploring the awareness levels of local expert stakeholders regarding multi-hazard scenarios. The first part of the current thesis contributes to our understanding of multi-hazard scenarios (mainly flood-triggered sinkholes) in karst environments, particularly in the Val d'Orléans. It represents an effort to position karst landscapes among other multi-hazard-forming areas where complex interrelationships among hazards are expected. A statistical-probabilistic machine-learning algorithm (MaxEnt) was used to map the susceptibility of interconnected hazards: flood-triggered sinkholes. The probabilistic spatial distribution map identified the areas with the low to high level of susceptibility to flood-triggered events. The multi-hazard map generated through this study potentially serves as a communication tool to support and aid decision makers involved in resource allocation and risk mitigation efforts. It can be placed among the first attempts to model the two interlinked hazards (flood and sinkholes) affecting the same area in the Val d'Orléans. Nevertheless, further research is essential to seek expert stakeholders' perspectives on the utility of multi-hazard maps. This is necessary to ensure that the generated map can effectively address challenges in decision-making processes.

The next step of research in this thesis focuses on local experts/scientists as stakeholders, since in the first place producing such maps is primarily a scientific task (Dallo et al., 2024). We seek the feedback and opinions of experts who are: i) science providers and are actively involved in the field of natural risks and urban planning, and who advise decision makers and report specifically on the Val d'Orléans. Since their influence impacts directly and indirectly on the decision outcome, their opinions matter. Therefore, their views on the map and perceptions of multi-hazard concept mirror the policy making processes; ii) experts (in natural hazard and urban planning domains) who can provide valuable insight on multi-hazard perceptions and attempts to map them since this is a relatively new research area.; iii) scientists and researchers who can provide unbiased opinions purely on the scientific aspects of the subject. This will be discussed more in **Section 5.2.1**.

The first step in the development and adoption of multi-hazard approaches in a community is to be aware of multi hazards and their consequences. In fact, an understanding of the level of preparedness of communities requires an analysis of stakeholders' awareness of the potential risks. Therefore, in this study, we not only assess the multi-hazard susceptibility map through the involvement of local expert stakeholders but also explore their awareness of multi-hazard concepts more generally. This shall lead to valuable insights into their current understanding of multi-hazards, which shape their perception and preparedness towards enhancing the resilience of their community (Mărgărint et al., 2021). Moreover, this part of the research aims at providing a roadmap for future improvements in terms of both mapping and a strategy for integrating multi-hazard thinking into local strategy.

To achieve this, it is imperative to first assess the stakeholders' knowledge of multi-hazard concept. Currently, multi-hazard thinking is poorly reflected, if at all, in policy and practice in the Val d'Orléans as outlined by Serrhini et al., (2023) since the only tool available to the Val d'Orléans is the Flood Risk Prevention Plan (PPRi, 2015). PPRi(s) introduced by law in 1995 as a key measure for preventing natural risks, the PPRi's goal is to minimise flood damage by regulating development and land use in vulnerable areas. Despite the area situated in a multi-hazard forming setting (karst landscape), the PPRi together with DICRIM documents are based on single hazard approaches.

Therefore, this evaluation process aims to:

- Examine stakeholder's awareness and perceptions of multi-hazard events.
- Evaluate the produced multi-hazard susceptibility map.
- Understand participants' perspectives on ownership and pervasive nature of flooding and sinkhole hazards in the area.
- Understand participants' perspectives on establishing a dedicated DICRIM document for multi-hazard interactions.

The results of this study are key for creating an improvement roadmap, integrating qualitative evidence and expert feedback, focusing on multi-hazard mapping, awareness, and fostering multi-hazard thinking.

This chapter is organised as follows: **Section 5.2** outlines the research methodology including the research approach: stakeholders' identification, the questionnaire design, highlighting the potential ethical considerations, and the means to reach out the selected participants. **Section 5.3** presents the results of the data analyses and the visualisation of the resulting data. **Section 5.4** presents the discussion section, which provides the interpretation of the key findings and the research limitations. **Section 5.5** presents a way forward as a roadmap for future research. **Section 5.6** is the final concluding chapter.

5.2 Methodology

This research employs a mixed-method approach, integrating both quantitative and qualitative methods, to collect and analyse feedback from local expert stakeholders (e.g., Smith and Shashkina, 2023). Our approach unfolds in two critical stages: 1) the identification of local expert stakeholders pivotal to our investigation, and 2), the design of a questionnaire. This questionnaire encompasses both closed and open-ended questions, enabling a comprehensive exploration of stakeholders' insights. The following sub-sections detail the stakeholder identification process and design of the questionnaire.

5.2.1) Stakeholder Identification

To accurately identify stakeholders, a thorough understanding of the French disaster risk management organizational structure is crucial. We will start by detailing the hierarchy within this system. As outlined in **Section 5.1**, our survey specifically targets scientists and experts in the fields of natural risk or urban planning, with a focus on the Val d'Orléans area. By explaining the system's hierarchy, we aim to pinpoint the key actor organisations that will aid how to approach potential experts.

In France, various organisations manage natural and technological hazards, involving multiple stakeholders, including the state, local authorities, experts, infrastructure managers, and institutional actors. Policies and preventive measures are established and enforced through a range of different legal documents. (Ferrer et al, 2018). The flowchart shown in **Figure 5.1** created by Ferrer et al. (2018) depicts the regulatory tools associated with urban planning which are used to disseminate and facilitate the exchange of preventative information among different stakeholders. The following summary is based on both the research done by Ferrer et al. (2018) and the information documented in the prefecture website (https://www.loiret.gouv.fr/).

According to the flowchart (**Fig 5.1**), there are two key governmental stakeholders namely the prefecture representing the government at a departmental (country) level and the Mayor who heads the processes.

In the context of a flood-prone town, the Mayor is responsible for developing a Flood Risk Prevention Plan, known as PPRi, which must receive approval from the prefect. Subsequently, the public is invited to express their views on this plan. A variety of communication methods, including information packages, such as DICRIM, is used not only to disseminate safety guidelines but also to enhance the public's understanding of the risks involved. DICRIM is a communal or/and departmental document on major risks in France aiming to raise awareness among residents and serving as a primary reference to inform the public on their hazard exposure (this will be more discussed further in this section). To establish DICRIM, first, the prefect compiles essential information on major risks within their department and documents them in the DDRM (Departmental Document on Major Risks) (more details provided in subsequent paragraphs). The Mayor then develops the DICRIM based on the information documented in DDRM. Moreover, the Mayor is responsible for keeping citizens informed about the major hazards at least every two years.

Therefore, here we identified two key actors namely the **national/state** and **local**. The **authority of the state** (Les services de l'État) is distributed among its various constituent services, such as the prefecture, the departmental directorate of territories (DDT) and the Regional Directorate of the Environment, Planning, and Housing (DREAL). These departments have the task of implementing and overseeing civil defence and territorial management policies. They play a crucial role in identifying towns and cities that are at risk of various hazards. Once identified, these municipalities are listed in the document called the DDRM or the Departmental document of Major Risks. This document details the risks each area faces, and is a key tool for planning and managing risk. Essentially, these state services are dedicated to pinpointing potential hazard areas and documenting this information for effective risk management and precautionary measures (Douvinet et al., 2013).

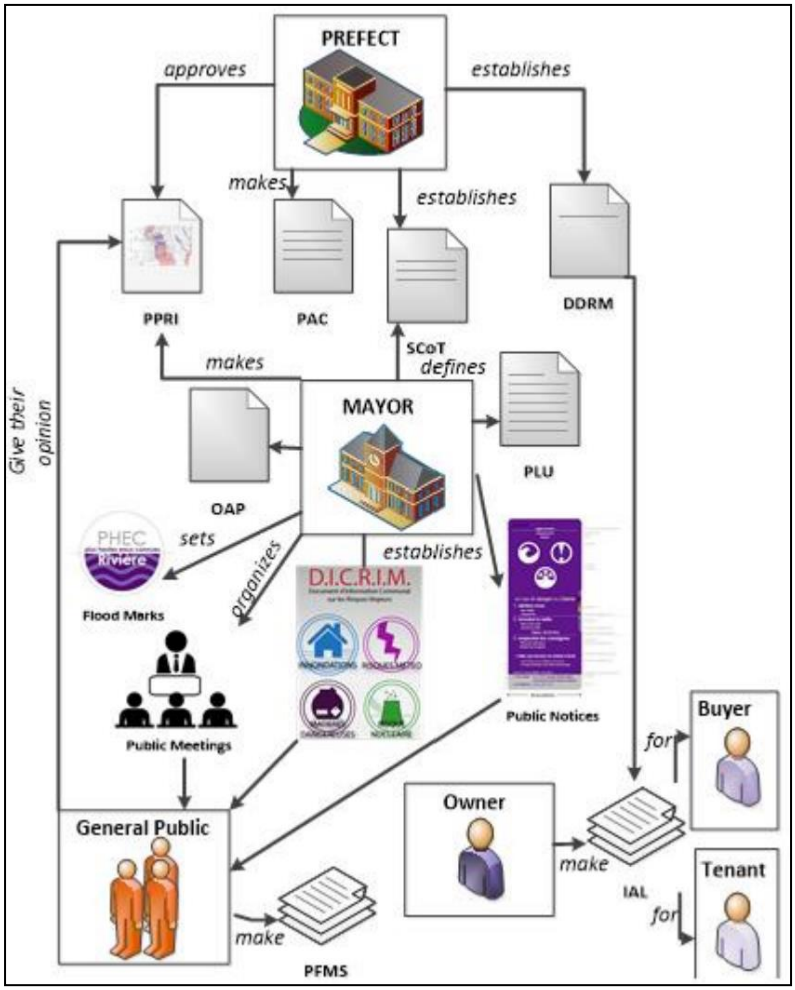


Figure 5.1 An overview of French regulatory prevention information across the municipality (Ferrer et al., 2018).

The **local actors** include Mayors, inter-municipalities and municipalities and the regional council. As mentioned earlier, the Mayor of the municipality is responsible to create the DICRIM, see below. In the current thesis, we are interested to identify the scientific team behind this document. In the following paragraphs we briefly discuss this document and we will give an overview of the scientific team behind this document.

i) What is a DICRIM. When a town or city faces even a single significant hazard, it is the city responsibility to alert the residents. Since 1990, any French municipality facing one or more risks is required to develop and distribute a document known as the Municipal Information Document on Major Risks (DICRIM) (The decree number 90-918, issued on October 11, 1990[2]). This document serves as a crucial means of preventive information, aimed at equipping the community to respond effectively during emergencies. The DICRIM includes the information transmitted by the prefect as part of the departmental major risks file (DDRM), in particular:

- the list of major risks to which the municipality is exposed;
- the description of each of these risks and their foreseeable consequences for people, property and the environment;
- prevention, protection and safeguarding measures for each of these risks;
- the individual safety instructions to be implemented (https://www.georisques.gouv.fr/).

The DICRIM is customised for the specific area it serves. It is suggested to make it more user-friendly by adding illustrations, brief lists of important actions, and simplified maps as guides. A brief introduction from the Mayor can also be included. This introduction can highlight the local risks and stress the importance of everyone being proactive in ensuring their own safety. Every risk should be described in a concise and specific way. For instance, in case of a flood risk, it is important to name the exact water bodies involved to help people easily identify them. Including archival materials like photographs and newspaper articles helps underscore the reality of these risks and aids in preserving a shared historical awareness. It is also highly advised to use maps instead of just lists when showing the level of hazards, evacuation routes and shelter locations etc. This document is in the form of an A5-sized booklet, foldable and about fifteen pages long, working well for distributing to residents in print form. This municipal information document is implemented under the political responsibility of the Mayor, yet, like other risk prevention tools, it is framed in advance by the State.

ii) Members of the scientific team behind the creation of DICRM. In this research, we are focusing on the technical science advisors (e.g. geologists, geophysicists, engineers, and social scientists) that provides support and scientific information to state and local authorities in the development, deployment, and evaluation of public planning and transport policies. Typically, the officials such as Prefects and Mayors, who issue documents related to natural hazards, heavily depend on the expert teams for generating comprehensive information. This interdisciplinary group includes geologists, meteorologists, hydrogeologists, civil engineers, and other specialists in risk assessment and management. Their expertise is pivotal in shaping decisions as they are directly or indirectly connected to the final decision makers and

[2] https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000535490

impact decision making processes in the Centre-Val de Loire regions[3] and more specifically in the Val d'Orléans. Two key institutions in this context are **BRGM** and **CEREMA.** In the following paragraphs, the expertise and services and contribution to public policy and planning of each will be briefly discussed (more information can be find in their websites: cerema.fr; brgm.fr).

a) ***BRGM (****the French Geological Survey***)**. BRGM is France's reference public institution for Earth Science applications in the management of surface and subsurface resources and risks. BRGM provides expertise in geology, hydrogeology, and risk assessment. It conducts research, provides data and assessments, and offers consultancy services in fields like mineral resources, environmental risks (earthquakes, landslides, flooding), and groundwater management. BRGM supports public authorities with scientific data and analyses, informing decision-making in urban planning, natural hazard prevention, and environmental protection. The institution is in partnership with CEREME (outlined below).

b) ***CEREMA (****Centre for Studies and Expertise on Risks, Environment, Mobility, and Urban and Country planning****)***. CEREMA is a French public agency that operates under the supervision of the Ministry in charge of sustainable development, urban planning, transports, and the sea. It focuses on providing expertise and studies in the fields of risk management, environmental planning, urban development, and transportation and sustainable mobility solutions. As it has mentioned in its website, one of the main missions of CEREMA is advising and supporting public authorities at both local and national levels. It assists in the development and implementation of public policies in its areas of expertise. This includes providing technical support, conducting studies and research, and developing methodologies and tools.

The contributions of both BRGM and CEREMA are vital in enhancing the resilience of Centre-Val-de-Loire regions against potential natural hazards. Their collaboration with local authorities results in DICRIM documents that are scientifically proven providing a solid foundation for informed decision-making and effective risk communication to the public.

5.2.2) Questionnaire design

To collect feedback from selected participants, a structured questionnaire was designed and influenced by the literature (Pescaroli, 2018; Margarint et al., 2021; Clive et al., 2021; Smith and Shashkina, 2023). The questionnaire was formulated into four thematic sections spanning six pages including the introductory section (See **Appendix 5.A**). In the introduction section, the goal of the survey and the structure of the questions were clearly outlined, and anonymity and privacy of the respondents were assured. The questions designed to have all forms of questions such as multiple choices, rating scales and ranking answers alongside open questions.

[3] Val d'Orléans is part of the larger Centre-Val de Loire region, an administrative division that encompasses six departments, including Loiret, where Orléans serves as the prefecture (centre-Val de loire.fr).

The four parts of the questionnaire are as follows:

- Part 1. *Awareness of the multi-hazard concept and the potential multi-hazard scenario in the area.* This section examines the participants' awareness and understanding of multi-hazards and the multi-hazard scenario and their relevance to the locality, resilience, and climate change, which as mentioned in Section 5.1 could roughly reflect their preparedness and resilience level.

- Part 2. *Evaluating the produced multi- hazard map (interpretation and understanding)*. This section assessed the perceived usefulness of the map, understanding the strengths and weaknesses of the generated map. To protect sensitive information, only a representative portion of the map was displayed as a draft map on the online platform. This strategy ensured data confidentiality while allowing participants to interact effectively with the relevant aspects of the research.

- Part 3. *Risk acceptability (readiness to embrace the risk)*. This section explored the perspectives on ownership and tolerance for risk (ownership and risk acceptability).

- Part 4. *The means of awareness and communication on multi-hazards.* This section gathered insights on the creation of a dedicated DICRIM document focused specifically on the interaction of multiple hazards. At the end of the questionnaire, the participants were asked to recommend any possible practices to foster the multi-hazard thinking among decision-makers and citizens, and rank main barriers to integrate multi hazard practices in local resilience strategies.

5.2.2a Ethical considerations (identifying potential ethical dilemmas)

This section was discussed with peers while attending to a PhD school on "Dilemmas and opportunities of fieldwork/in situ research in disaster studies" at the University of Twente, Netherlands - October 30, 2023. This discussion took place two months before conducting the survey. Therefore, it provided us with enough knowledge before approaching the potential participants.

In the realm of ethical considerations within the current research, we identified the potential ethical dilemmas, particularly when engaging expert stakeholders in disaster management. Given that the focus of this thesis does not involve direct interaction with affected and vulnerable populations, this work was generally categorised as a *low-risk* survey. Nevertheless, it still presents a range of ethical dilemmas and risks specific to our context. Two of the key aspects of ethical standards concerns the principle of "informed consent" from all research participants and taking into consideration power dynamics. In the following, the identified potential ethical dilemmas are briefly discussed.

- *Privacy and consent.* Ethical concerns arise regarding privacy and informed consent namely obtaining permission from the participants to use their views and responses. To address this, at the end of the questionnaire, we asked the participants for their consent to use and analyse the information they provided.

- *Power dynamics.* Approaching experts in the field seeking their perspectives always should come with proper respect. As such, adhering these points can help build trust among interviewees and the interviewers; clearly explaining the objectives of the study, how data will be treated, analysed and used, and stressing the right not to participate and the right to withdraw at any moment, and that those decisions will also be confidential.

Moreover, in designing questions to evaluate experts' insights on the multi-hazard concept, it was vital to design them in a way that avoided triggering their defensiveness or bias, ensuring that the questions were neutral and conducive to an accurate assessment of their knowledge. In the course of this research, ethical considerations were followed according to the guidance provided by my internal supervisors, and and all necessary steps were taken to ensure the ethical integrity of the research. I approached each participant with the approval letter provided by my internal supervisor.

5.2.2b Reach out to the selected participants

This section outlines the methodology employed to engage participants. The research survey was designed in an online **Google Form** in both English and French. This online approach offered a convenient interaction, ensuring wider accessibility and ease of participation mainly due to the language barriers. However, while the questionnaire was a sufficient way to collect participants' responses, we also provided them with a choice of being interviewed in person or through platforms such as Microsoft Teams.

We first provided a primary list of potential participants. To identify these experts, we used papers, national reports and experts' LinkedIn profiles. We reviewed their LinkedIn profiles to make sure that they are currently working in the region of Orléans or in the Loiret department. That was specifically important for questions regarding Part 1 in which we aimed to understand the experts' awareness on multi-hazard perception and the relevant multi-hazard scenarios threatening their areas. In total, we contacted 16 main experts via individual emails, offering them two options: i) participation through an online survey, or ii) a semi-structured interview conducted over online platforms. This approach was done to respect the varied backgrounds of the experts and to clearly communicate why their specific insights were particularly valuable. Of the 16 contacted participants, 7 experts responded (6 online-form responses and 1 interviewed via Microsoft Teams). The experts' profiles are varied namely geologists, hydrologist, hydrogeologists, geomorphologists, urban planners, and geophysicists.

5.3 Data analysis

To analyse the data, we began by collecting all responses received via Google Forms. In addition to these, we conducted an interview with one expert through Microsoft Teams, and the insights gathered were subsequently integrated into the form data. This dataset was then exported to Microsoft Excel for basic data cleaning, ensuring both consistency and accuracy across the responses. Given the bilingual nature of our data collection, the four responses that were originally in French were carefully translated into English to maintain uniformity across the dataset.

Given the manageable number of responses, we were able to undertake a detailed examination of the qualitative data derived from both the open-ended questions and the expert interview. This examination involved direct quotations where appropriate, as well as a separate analysis of each response. For certain questions, we identified recurring themes and patterns, which allowed us to conduct a thematic analysis. This involved a detailed exploration of the data through critical reading; specifically focusing on identifying and understanding the most frequently used phrases. In the following sections, the results of each part of the survey will be shown by relevant tables and graphs.

5.3.1 Questionnaire results

In this section, we present the results of the data collected from our **seven** participants. We maintain the same structure for presenting the results as was used in the questionnaire, organizing them by each part.

5.3.1a Awareness of multi-hazard concept/ scenarios

This section consisted of **four** questions. Participants were asked to give their understanding and definitions of the multi-hazard concept (**Q1b**). Prior to this question, we first asked them to specify whether they have come across the multi-hazard term or not (**Q1a**). To be able to gain more insight, they were further asked to note the main multi-hazard scenario that could be a threat in Val d'Orléans (**Q2**). In **Q3** and **Q4**, we assessed whether they are aware of these hazards and to what extend they are taking it seriously.

The results show that five out seven experts were unsure whether they are familiar with the term of multi-hazard (**Fig 5.2**). **Table 5.1** presents an overview of the expert opinions on multi-hazard concepts, delineating the different perspectives across various fields of specialisation. We have captured their actual insights and understanding of the multi-hazard perceptions. Of seven respondents, four answered in French which then were translated accurately into English to keep the format consistent.

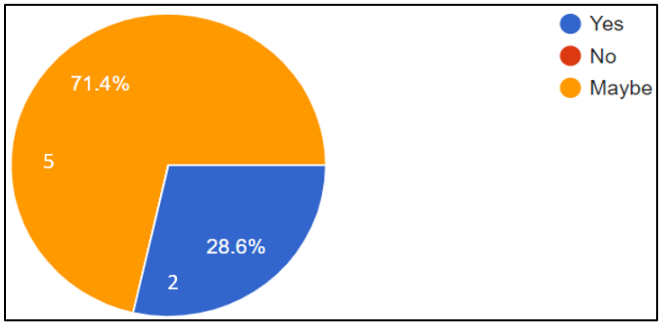


Figure 5.2 Expert responses to the familiarity of the multi-hazard concept.

Table 5.1 An overview of the expert definitions of multi-hazard concepts across different fields.

Expert profile	**Definitions by Experts**	**Working Institutions**
Geologist	**"**More than one hazard occurring in the same area and around the same time**"**	BRGM
Geologist	**"**This term refers to sectors where several hazards are encountered, one of which can influence the other (the reverse is not always systematic). At CEREMA, we mainly talk about these subjects for natural hazards. Working on the issues of ground movement, this subject often relates the flood hazard or the runoff hazard with the ground movement hazard. **"**	CEREMA
Hydrologist	**"**Several risks in the same areas; with more or less the same origin; the same driver**"**	BRGM
Hydrogeologist	**"**Several types of unpredictable events that can interact with each other**"**	BRGM
Geophysicist	**"**Combination of several hazards**"**	BRGM
Geomorphologist	**"**Several hazards in the same area**"**	BRGM
Urban planner/ science communicator	**"**A territory subject to multiple risks**"**	CEREMA

Q2 targeted the experts' awareness of the existing multi-hazard scenario threatening the area. It should be noted that due to the nature of the semi-structured interview, which provided a more fluid and two-way conversations, we could not capture the response for this question since before the interview, we have already discussed the multi-hazard scenario. Therefore, we have only obtained six responses specifically for **Q2**. The responses were asked to define a main multi-hazard scenario threatening the area.

As demonstrated in **Figure 5.3**, responses based on the disciplines are varied namely:

- The **two geologists** outlined flood and subsidence/sinkholes with one of whom describing a multi-hazard scenario explaining the interrelationships between them.
- The **hydrologist** mentioned flood risk as a multi-hazard scenario.
- The **geophysicist** described a scenario starting from flooding associated with periods of intense precipitation and periods of drought, then linking these hazards to swelling of clays leading to gravitational collapses (landslides) with the natural and human-made origins.
- The **hydrogeologist** outlined a set of single hazards from drought, ground movement to flooding.
- Finally, the **urban planner** outlined flood and storm.

Only two experts presented a multi-hazard scenario with the integration (the geologist and the geophysicist).

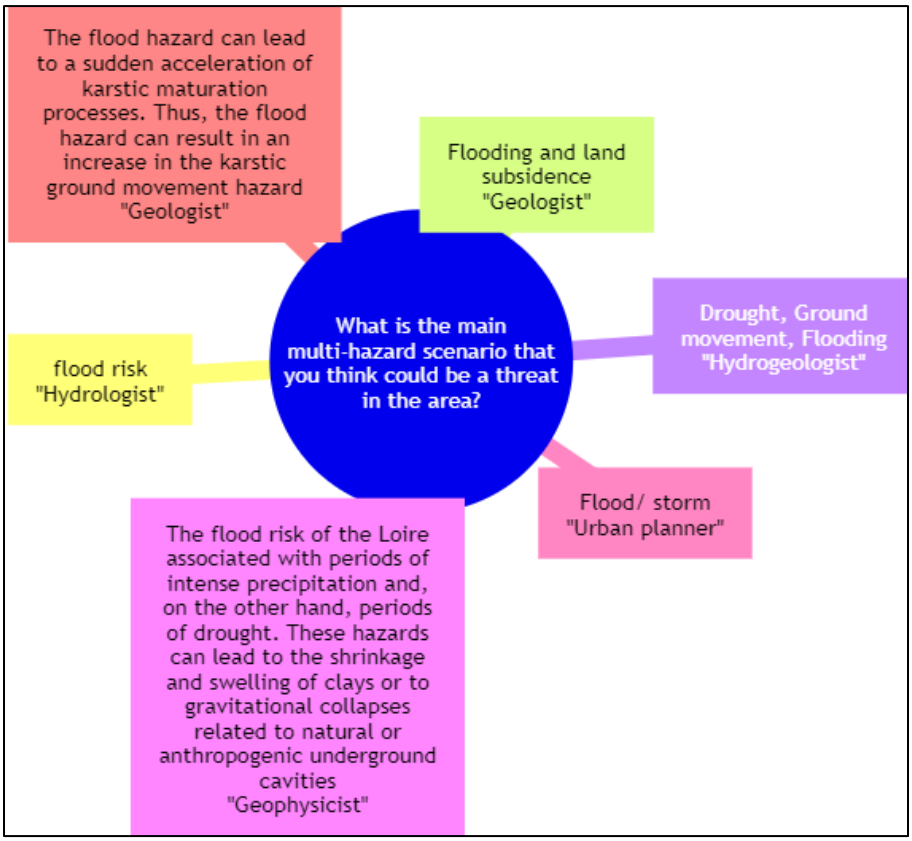


Figure 5.3 Perceived the main multi-hazard scenario in the Val d'Orléans area.

In response to **Q3**, all respondents rated the probability/ likelihood of the potential multi-hazard scenarios they identified as ranging from likely to very likely, with five rating them as very likely and two as likely. Specifically, in response to **Q4**, five experts believed that, due to climate change, the occurrence of the multi-hazard scenarios outlined in **Q2** will become increasingly frequent and very likely, with two experts rating them as likely. The visualizations for both **Q3** and **Q4** are combined in a single figure (Fig 5.4), as they share the same format and similar responses. This indicates that all the scenarios mentioned by the experts in **Q2** are valid and are likely to very likely to occur in the area.

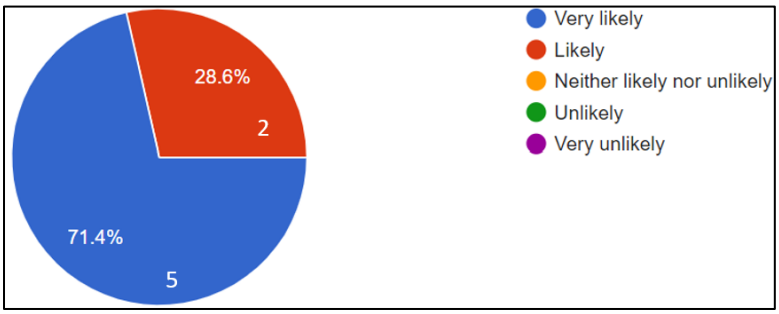


Figure 5.4 The likelihood of the identified multi-hazard scenario in the area, both generally and within the context of climate change (Q3 and Q4).

5.3.1b Part 2. Evaluating the produced multi-hazard susceptibility map

Prior to showing the results of this part, we briefly explain the importance of colour in mapping and the reasons behind the question related to map colour scheme. One of the key elements in natural hazard maps that facilitate communication is the use of colours. In fact, colour-coding scheme impact how users interpret maps (Lindell, 2020). Humans rely on their visual senses to quickly evaluate their environment, including danger identification. Dallo et al. (2024) in their new research assert that inappropriate colour scales on maps can lead to a variety of problems. Firstly, they can cause data distortion making it difficult to interpret accurately. Moreover, if colours are not well chosen, maps could be inaccessible to those with colour vision deficiencies. By that, it means that maps would not be inclusive anymore whilst it is important to not exclude anyone. The choice of colour scheme is vital, not only for ensuring the map's accuracy and accessibility but also because it facilitates rapid and intuitive data interpretation by leveraging the pre-attentive processing capabilities of the human visual system (more references will be found in the paper Dallo et al., 2024).

When it comes to conveying danger, warm colours like red, orange, and yellow might be more effective at communicating danger (Engeset et al., 2022). Our attention is immediately drawn to them thus sending a clear warning signal. On the other hand, when it comes to communicating safety, cool colours like green, have a soothing effect and can convey a sense of calm and security more efficiently (Engeset et al., 2022). However, in the current map, the use of green does not imply that the area is entirely free of hazards (so that it prevents people from taking actions); rather, it indicates a proportionately low hazard existing, ensuring that individuals remain aware and take appropriate actions (Dallo et al., 2020). Given this fact, we are interested in understanding what the colour green immediately communicates to viewers. To achieve this, we will show them two maps with different colour schemes: one using red, orange, light, and dark green, while the other transitions from dark purple to yellow as the hazard level decreases (see **the appendix C).** We aim to determine which map effectively conveys the idea that a low hazard level does not equate to zero hazard (**Q8**). It is important to remember that this survey targeted the perspectives of various experts; therefore, the results are not immediately transferable to a wider public audience.

This section focuses on evaluating the effectiveness of the produced map, specifically regarding its interpretability and utility. Our objectives are to:

i) Determine if the map's colour scheme effectively communicates its intended message, assessing if the visual design aligns with the communication goals (**Q5**, **Q6**, **Q7**, and **Q8**).
ii) Evaluate the understanding of the map's purpose as a tool for probabilistic forecasting of hazards, rather than precise predictions, and evaluate how well it conveys the inherent uncertainties associated with these forecasts (**Q6**).
iii) Explore how the map's colour choices influence readers' perceptions of risk prevalence across different areas, examining the impact of colour on the interpretation of risk distribution (**Q8**).
iv) Assess the map's practical utility for experts, including whether users find it informative and helpful for daily decision-making, and gather any suggestions for improvement they might have (**Q9a, Q9b, and Q10**).

These data were gathered through asking five targeted questions. The first question (**Q5**) posed to the participants aimed to explore whether by looking at the map one can accurately identify areas from low to high considering the colour scheme. The result presented in **Figure 5.5** indicate that the majority felt confident to identify the susceptible areas from low to high. This confidence was further supported by the responses to the following questions. As previously mentioned, in the current map, dark green does not denote an absence of hazard. Participants' perceptions of the dark green colour were examined, with results indicating a majority associated it with areas of lower risk, with none interpreting it as indicating no risk (**Fig 5.6**).

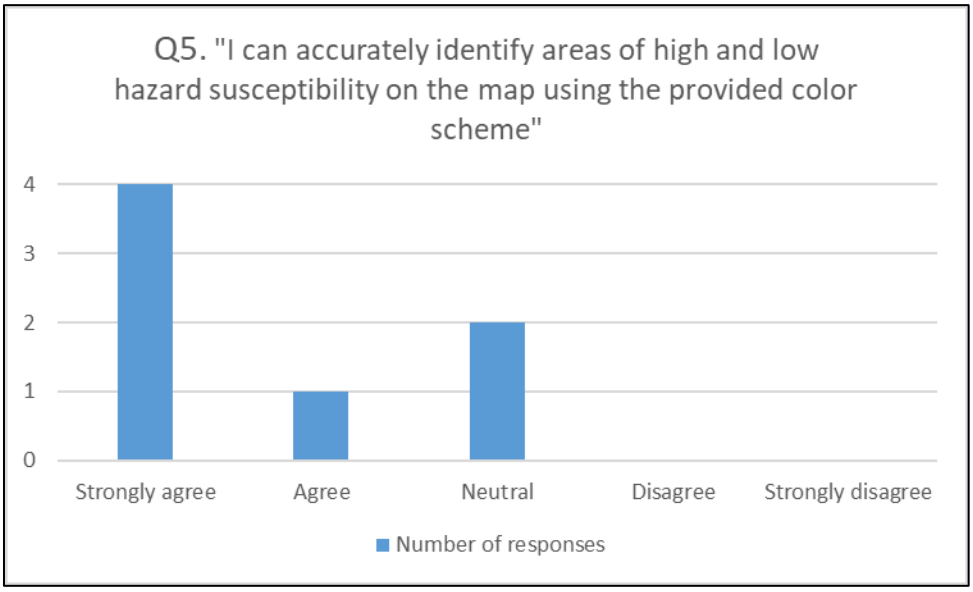


Figure 5.5 The general perception of the map hazard level based on the colour schemes.

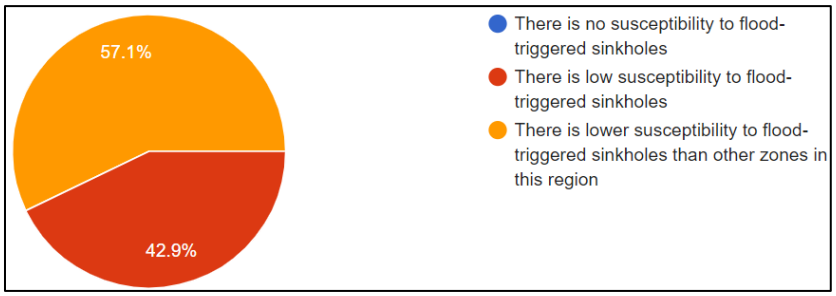


Figure 5.6 The hazard level perception of the dark green colour.

Participants were then shown two maps differing only in the colours representing lower-hazard areas, and were asked which colour (dark green or yellow) better communicated that these areas, while lower in hazard level, were not hazard-free. They were also asked on alternative colour choices for this purpose. One participant highlighted that neither yellow nor dark green is optimal for conveying lower hazard, citing yellow's lack of clarity and dark green's association with safety, suggesting blue or purple as preferable alternatives. During a semi-structured interview, another expert rationalized dark green as the superior choice, explaining that unassessed areas in the study zone would be rendered in grey, white, or transparent to indicate the absence of hazard data. They further referred to the standard colour spectrum progression from violet to red, commonly used in risk mapping.

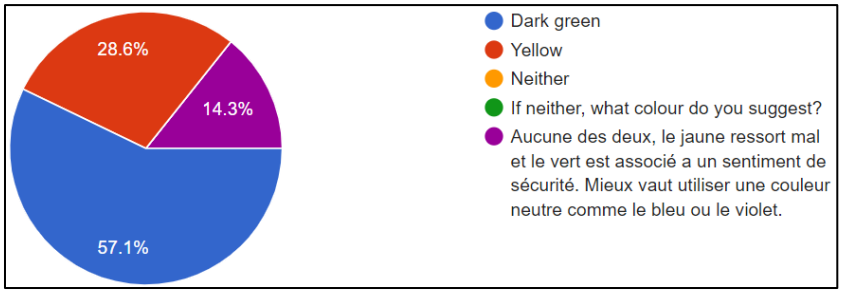


Figure 5.7 Experts' opinions on the colour that best communicates zones with lower susceptibility, not free from susceptibility.

To evaluate the understanding of the map's function as a probabilistic hazard-forecasting tool and its ability to express the associated uncertainties, participants were questioned on the implication of red areas on the map. As depicted in **Figure 5.8**, nearly all participants (six out of seven) showed a solid grasp of the map's intent. The majority accurately interpreted the red zones as regions where flood-triggered sinkholes potentially might occur in future events. These responses imply that participants were adept at not only distinguishing areas by colour scales but also at understanding the probabilistic nature of the forecast provided by the map.

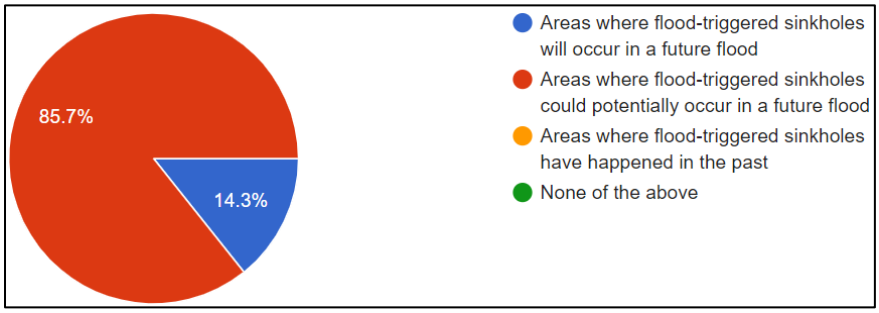


Figure 5.8 Participants' perception of probabilistic forecasting nature of the map.

Moving from colour related questions, in **Q9a, Q9b** and **Q10**, the experts were asked to explain the utility of this multi-hazard susceptibility map into their own work. Four responses were collected. The first respondent who is a geophysicist working on groundwater sections affirms the utility of the multi-hazard map, particularly in assessing the vulnerability of sectors related to water quality. This revealed the importance of water contamination associated with sinkholes, which could cause concerns in relation to the environmental hazard impact caused by sinkholes in karst areas.

The second response, which is from the urban planner who is also a science communicator, introduces a governance aspect, revealing that these maps are not just technical tools but also could play a part in informing elected officials. This suggests their utility extends into the realm of policy-making, where they may be used to guide decisions and allocate resources for environmental protection.

The other response reinforces the maps' practical use in identifying areas where protective measures are required. This respondent's view, although not from a decision-maker's perspective, underscores the maps' role in operational planning and risk management. It implies that these tools are instrumental in visualizing and prioritizing interventions in the environmental sector. However, the final response highlights the limitations of the current state of these maps, pointing out that the lack of a detailed methodology for creating such maps as multi-hazards make it difficult for decision makers to determine their usability.

At the end of the Part 2, participants were asked for suggestions to enhance the utility of the map or any other improvements (**Q10**). Seven responses were received, which provided a variety of perspectives on potential improvements. Through textual analyses, themes were extracted as follows:

- **resolution and detail**,
- **accessibility and inclusivity**,
- **functionality improvements**,
- **uncertainty communication**,
- **clarity of information,**
- **methodological transparency**.

Moreover, respondents indicated that conveying the uncertainties associated with the map's data could cultivate trust and clarity among users. The legend's clarity was another point of discussion, with emphasis to differentiate more clearly between "Lower" and "Low" susceptible areas, suggesting that apart from colour communication considerations, users also need to read a clear and informative guidance to navigate the map effectively.

A participant pointed the need for methodological transparency, particularly regarding the scenarios and parameters underlying the map. This aspect of feedback highlights the importance of not only how the map is presented but also how it is understood in terms of the data that informs it. Such transparency is key to fostering an informed user base capable of leveraging the map to its fullest potential.

Lastly, one respondent recommended the inclusion of a warning or advisory text beneath the map. This would be a good practice to clearly communicate any uncertainties or critical considerations associated with the map's data and interpretation. Such a cautionary note is an effective descriptive disclaimer, as it guides the user in understanding the map's limitations and the contexts in which it should be used, ensuring a more informed and cautious application of the information provided. The feedback from the respondents provides a roadmap for improving the map's design and functionality. These enhancements have the potential to increase the map's effectiveness as a tool for local risk and resilience planning strategies.

5.3.1c Part 3. Risk acceptability (readiness to embrace the risk)

Building upon the previous questions related to Part 2, which highlighted the widespread presence of hazard, this part is particularly interested to explore the perceived responsibility regarding this pervasive (and hidden) nature of hazard and what may be the right response. We achieved this through two main questions with one being a multiple choice and the other one being an open ended. **Q11** directly asked the participant that in the context of sinkholes triggered by flood, how they think one should respond. They were given three options with participants being able to choose more than one response, since the aim was to obtain collective answers rather than individual.

As presented in **Figure 5.9**, the majority of the respondents (6 out of 7) support the promotion of multi-hazard thinking among decision-makers and citizens. Investing money and time in completing subsurface investigations also came out as an important and preferred action. Two experts also indicated that people also need to accept the risk. One additional comment related to acceptance was "we also need to accept the fact that there is risk, but in France the law makes Mayors responsible. Therefore, accepting the existing risk can also open the door to build everywhere, which is not possible since certain people are still accountable. Thus, the acceptance needs to be intertwined with education and teaching people about the potential risks". In general, it shows that the majority are in favour of forward-looking and proactive actions in reducing the risk rather than the passive acceptance.

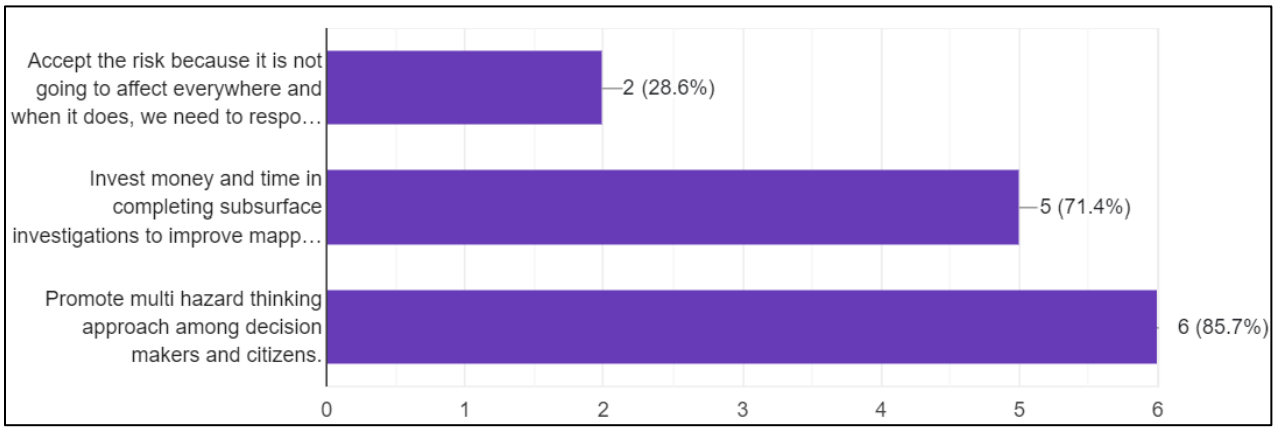


Figure 5.9 Expert recommendations for strategic responses to flood-triggered sinkholes.

The responses to question **Q12** about who should take the responsibility of flood triggering sinkholes in the area showed a common theme, which is a shared responsibility among different levels of government and experts, with responses indicating:

- The **local authorities** (cities, local governments) are mentioned, suggesting that respondents believe that the responsibility should be a more localized approach to hazard/risk management.

- **Regional and departmental authorities** in collaboration with the state (national government) are also seen as responsible parties. This shows that the respondents expect coordination between different levels of government.

- Some responses indicate that **organizations like BRGM, CEREMA and DREAL** in France can provide support to understand and reduce the multi-hazard risk. However, they cannot be held "responsible" in the strict sense. They only can assist and give advice but not take full responsibility for risk management. Ultimately, the overarching responsibility for managing and addressing these hazards/risks falls on the state.

In general, the responses suggest a preference for a multi-layered approach, where local governance structures such as regions, departments, and communes work in tandem, supported by the national government. While expert input is valuable, the final responsibility might be held by governmental structures capable of policy-making and enforcement.

5.3.1d Part 4. The means of awareness and communication on multi-hazards

This final theme is dedicated to exploring effective strategies for communicating the multi-hazard map to the public. As explained in **Section 5.2**, France employs an awareness tool known as DICRIM to inform its citizens. We were in particular interested to understand expert opinions on the best methods for incorporating multi-hazard scenarios into these single-hazard focused documents. Accordingly, in the survey, respondents were asked about the advisability of creating a separate document for multi-hazard scenarios in addition to the current single-hazard documents, or whether it would be more effective to integrate them into the existing publications.

The responses (**Fig 5.10**) demonstrate a strong consensus for the integration of multi-hazard scenario information into current publications. Almost all the experts (6 out of 7) agree that it would be beneficial to include multi-hazard information in existing publications for easy access to all information.

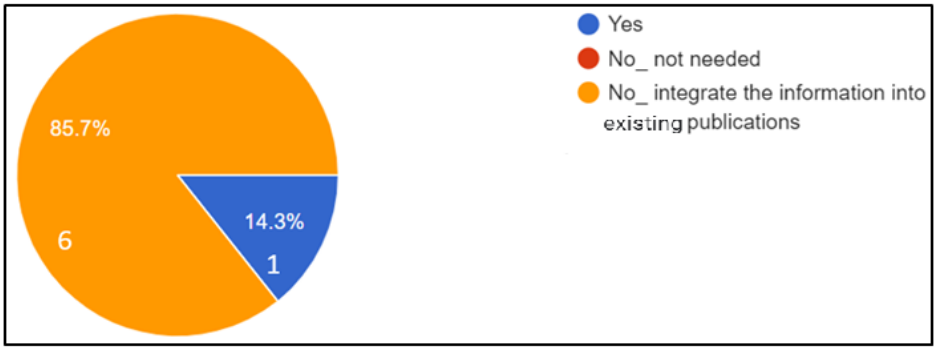


Figure 5.10 Responses to the proposal of providing a separate DICRIM document specifically including multi-hazard information/maps.

In **Q14**, the experts provided several recommendations for potential practices to encourage multi-hazard thinking among decision makers and citizens:

- **Le géoportail**: which is a French government initiative providing access to geographic and environmental data. One suggested the use of this platform to disseminate information to both decision makers and citizens.

- **Information sessions**: This traditional yet effective approach involves organizing informational meetings or workshops. Such sessions can be tailored to diverse audiences, providing them with the relevant knowledge about hazards and risk mitigation strategies. It is a direct way to raise the awareness and engage communities, ensuring that they are informed about potential risks and the importance of multi-hazard preparedness.

- **Organizing events involving key stakeholders**: This suggestion emphasizes the importance of bringing together various stakeholders, including decision-makers, experts, community leaders, and citizens, to communicate and collaborate on multi-hazard issues. Such events can foster a sense of community, enhance mutual understanding, and encourage the exchange of ideas and strategies for effective hazard management.

- **Creating models explaining phenomena:** The creation of physical or digital models can help visualize and understand complex phenomena, making the information more accessible to non-experts.

- **Digital exhibitions:** This suggests using digital platforms to create exhibitions that can reach a wider audience, providing an engaging way to raise the awareness among the public about multi-hazard risks.

Q15 asked participants to rank the barriers to integrating multi-hazard strategies into local resilience strategies ranking them on a scale from 1 for the most critical barrier to 5 for the least critical. The results depicted in the **Figure 5.11** show a strong consensus among participants, identifying **financial constraints** as the most critical barrier.

However, there was some variation in opinions regarding the importance of other barriers. The second most important barrier, according to the data, was the **lack of clear mandates for resilience**, which four participants rated as such. Finally, **lack of political will** and **limited decision-making power at local level** were considered as the least constraints.

In addition to this, one additional barrier was commented by one expert mentioning "*the main obstacles also include understanding the concept of multi-hazard and the existence of a proven method to characterise it. Furthermore, characterising the ground movement hazard at the scale of the Val d'Orléans does not have consensus, and discussing multi-hazard concepts with elected officials seems premature to me*".

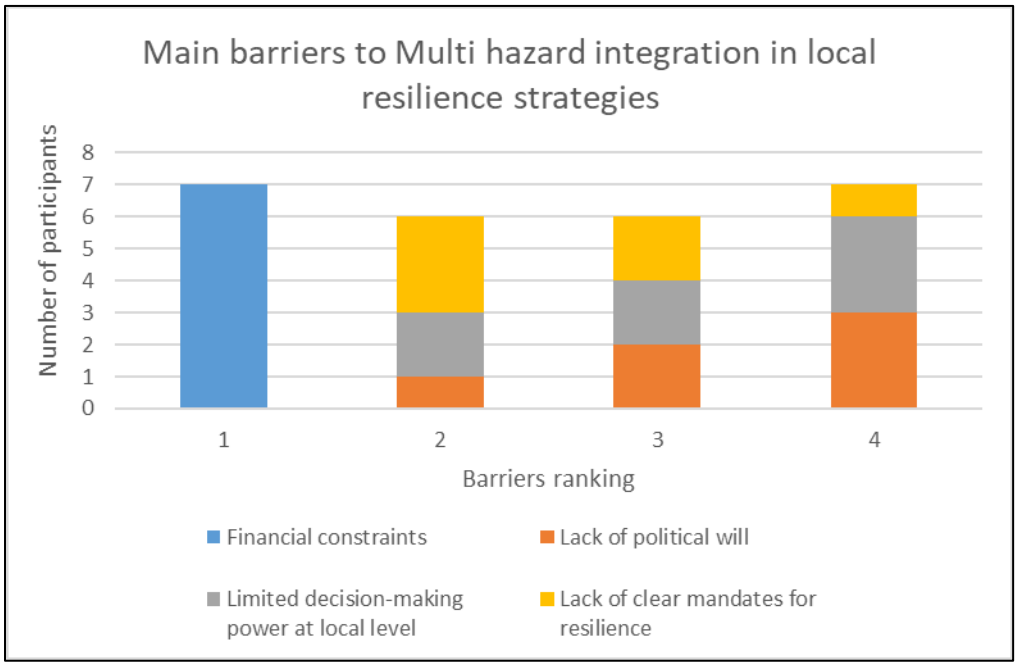


Figure 5.11 Results of the ranking of main barriers to integrating multi-hazard thinking into local risk and resilience planning strategies.

5.4 Discussion

In this discussion section, we first give our interpretation of the key findings for each part in **Section 5.4.1.** Then, in **Section 5.4.2**, the limitations will be discussed.

5.4.1 Overview and discussion of key findings

i. ***Awareness of the multi-hazard concept.*** Analysis of the responses reveals a prevailing uncertainty about the experts' knowledge of the multi-hazard concept, as evidenced by the frequent "maybe" answers. This suggests that the experts might be familiar with the concept to some extent but do not feel confident enough in their knowledge to give a definitive "yes". Such uncertainty likely stems from several factors: **i)** the concept of multi-hazards being relatively recent; **ii)** its perception as specialised terminology (being as jargon); and **iii)** a lack of direct, hands-on experience.

Despite this, experts demonstrate a basic level of awareness regarding multi-hazards, particularly acknowledging that multiple hazards can co-occur in the same geographical region or timeframe. Yet, recognising the presence of multiple hazards does not equate to a deep understanding of their complex interrelations and combined impact, which is essential for thorough multi-hazard analysis and effective strategy development. The responses do not typically highlight the systemic interactions among these hazards or how one hazard may exacerbate another, hinting at a superficial grasp of the concept's broader implications.

Furthermore, when asked to identify the primary multi-hazard scenarios threatening the area of Val d'Orléans, most experts pinpointed single/individual hazards rather than a scenario where interrelationships their synergistic/amplification effects are implied. This oversight suggests a limited consideration of hazard interplay. However, the insights from a hydrogeologist and two geologists provide a clearer understanding of potential multi-hazard scenarios in the area.

For instance, the hydrogeologist connected flooding from the Loire River with cyclic patterns of intense precipitation and drought, which could subsequently lead to soil instability or the development of natural and man-made underground cavities. Similarly, the geologists identified floods and sinkholes as key multi-hazard scenario, with one elaborating on the relationship between flooding and accelerated karst processes, which can further intensify karst ground movement hazards.

Overall, the varied responses indicate several key points:

1. The majority responses **identified individual hazards without addressing how these hazards might interact with one another**. This suggests a lack of emphasis on the interconnected nature of multi-hazard scenarios.

2. There appears to be a **distinct difference in the understanding of multi-hazards across different disciplines,** with professionals tending to concentrate on hazards within their areas of expertise. For instance, a hydrologist primarily identified flooding as a main multi-hazard, whereas geologists and a hydrogeologist pointed out both floods and sinkholes. Another hydrologist considered drought, landslides, and floods, and an urban planner highlighted floods and storms. This diversity in focus

underscores the disciplinary silos that can impact the comprehensive assessment of multi-hazard scenarios.

3. The different range of hazards identified by the experts (with different background) collectively provides a more holistic view on potential multi-hazard scenarios, illustrating the **complex dynamics at play in the area.** Drought periods followed by heavy rainfall could lead to flooding of the Loire River. Additionally, storms can also trigger flooding. Flooding can activate pre-existing underground cavities in the karst landscape or trigger new sinkholes, illustrating the layered and complex nature of multi-hazard risks. An overlooked aspect in the responses was the potential bidirectional relationship involving river dike breaches -whether from flooding or the collapse of sinkholes beneath the dike- possibly leading to further flooding. This gap **highlights the need for a more integrated approach to understanding and managing multi-hazard scenarios, emphasising the interplay between various hazard processes.**

4. The fact that **all the respondents agreed that these hazards are imminent in the area** shows that not only the flood triggering sinkholes is a quite relevant scenario in the area but also there is a wider picture of multi-hazard scenario in the area which are also relevant and are in need of collective investigation.

ii. ***The multi-hazard/susceptibility map evaluation.*** Overall, the map colour assessment showed a high understating of the map purpose. There is an understanding that the map shows probabilistic forecasting of the distribution of these hazards rather than historical data or definitive future events. This insight indicates a risk-aware interpretation of the map, where red zones are seen as areas of potential hazard, not certainties or records of past events. As for a colour to depict a lower hazard level but not free of hazard, participants mainly were in favour of the dark green colour, as the majority perceived it as low to lower hazard zones. It was mentioned that in case of a free-hazard zone, risk mappers usually allocate white or grey colour or no data text.

Furthermore, in the evaluation of the multi-hazard susceptibility map, several aspects of feedback were collected as detailed in **Section 5.3.1b**. Each piece of feedback and suggestions provided us with valuable insights for future considerations and refinement of the map. Here, we will briefly discuss each, focusing on how they can help and enhance the utility of the map:

- **Accessibility and inclusivity.** The responses emphasised on considering the colour blindness issues. The fact that individuals with the condition of colour vision deficiency (CVD) see and subsequently perceive colours differently.

- **Uncertainty communication and methodological transparency.** The experts highlighted that it would always be beneficial to build trust by admitting the potential uncertainties in creating the map and the way how to best do it is to add this in the map frame so that it will be an integral part of the map. Individuals observe the map while also being aware of the uncertainties. The transparency in terms of the used method was also another point coming out from the survey. This combined focus on both acknowledging uncertainties and being transparent about the methods used

enhances the map's credibility and utility, fostering a well-informed and trustful engagement with the map by its users.

- **Clarity of information.** This pertains to the legend's clarity, which is crucial for outlining the modellers' intentions. For instance, it is important to differentiate more clearly between "Lower" and "Low" susceptibility areas. This suggests that, beyond colour communication, users also require clear and informative guidance to navigate the map effectively. This finding also aligns with research by Engeset et al. (2022) which found that simplifying the legend and adding annotated numbers or proportional sizes helped both color-vision-deficient (CVD) and full-color-vision (FCV) users better interpret danger levels.

- **Resolution and detail.** The ability to zoom in the map came from the science communicator who is also an urban planner. This comment reminds us to highlight the importance of map scales (local, regional, national). The produced map for the current thesis has been provided in a regional scale, which does not allow us to either zoom in or zoom out. That means either zooming in for a more detailed view or zooming out for a broader perspective; the perceived importance of the factors used to create the map may shift accordingly. Therefore, this dynamic relationship underscores the need to consider scale as a critical variable in our analysis, influencing both the interpretation and applicability of our results. To clearly communicate this purpose of the map, model developers/ map providers need to emphasis on this somewhere in the map. A regional-scale map generally provides a balance between overview and detail. It will not have the fine detail of street-level maps, but it offers more detail than a national or global scale map. This means it might show major rivers, cities, and natural landmarks.

From the results, we also learnt that multi-hazard maps could be a critical input for other professional users (scientists/experts) involved in different part of natural hazard studies. As shown by evidence of the current research, often other disciplines may use the outcome of others. This key finding highlights the importance of participatory multi-hazard mapping.

The initial multi-hazard map as outcome of the current thesis has the potential to bring other expertise around the table and start conversations on different scenarios and how those could be relevant to other phenomena, which could be overlooked by experts working separately. It is also highly important to launch these meetings/workshops in the presence of science communicators on disaster risk strategies.

The feedback provided here is invaluable for the iterative development of such tools, emphasizing the need for continuous improvement to meet the diverse needs of stakeholders. These results highlight the need for clearer communication, standardized definitions, and more demonstrable evidence of the effectiveness of multi-hazard and multi-risk decision-support tools in improving disaster risk management outcomes.

iii. Risk acceptability and responsibility. In managing risk, understanding and accepting the pervasive natural risks help the community to prepare more effectively. In response to these pervasive flood-triggering sinkholes, the experts agreed more on promoting multi-hazard thinking approach among decision makers and citizens. This approach is complemented by investments in money and effort towards completing subsurface investigations to enhance our understanding and mapping of these risks. However, it has also been emphasised that accepting these risks should go hand-in-hand with education and awareness-raising efforts.

iv. Integration into current DICRIM. Respondents emphasised the importance of including this information, particularly in the final pages of the documents. This recommendation stems from the understanding that people often exercise choice in what they read, and therefore, making the information readily accessible is crucial. The consensus underscores a strategic approach to foster public awareness, one must provide citizens with easy access to information. By embedding multi-hazard scenarios within the familiar format of existing documents, we can ensure that vital information reaches a wider audience, equipping them with the knowledge to make informed decisions in the face of multiple interrelated risks.

v. Barriers to integration. While understanding the need for the multi-hazard integration is one thing, the actual process of incorporating this complex concept into local strategy comes with unique challenges. Identifying these obstacles is crucial to prioritise and overcome them. The primary obstacles were identified as financial constraints and lack of clear mandates. Not having a clear understanding of objectives and responsibilities can hinder the progress. As respondents also mentioned, there is a lack of understanding of the multi-hazard terms and detailed methodologies, which can lead to confusion. Limited decision-making power at local level and lack of political will came out as the least important barriers. This does not imply that they are not barriers at all; rather, compared to other factors, they are considered lower priorities.

5.4.2 Limitations

Accurately identifying current experts in the area proved to be particularly challenging and time consuming. Several individuals listed as active professionals in the area were, upon contact, no longer working in the region. Especially in the dynamic fields where professionals frequently shift focus or change projects. Therefore, due to this discrepancy, some potential experts refused to participate in the survey, as they were no longer working in the study area.

The relatively small number of interviews might be also perceived as a limitation. Nevertheless, it is important to consider the significant challenge in identifying experts actively engaged in the specific area of interest. Despite this constraint, the participants selected for the study are authoritative figures within their respective domains. This is evidenced by the inclusion of two experts who hold leadership positions in urban planning and natural hazard studies, underscoring the credibility and relevance of their insights. Besides, we believe that the diversity within the experts' professional backgrounds has enriched the study introducing a broader spectrum of perspectives. Thus, although the number of participants was relatively small, the different level and background of expertise of the participants still gives great value and more importantly credibility to the results.

5.5 Way forward

The importance of the current research lies in the intrinsic characteristics of Val d'Orléans being exposed to multiple hazards, particularly flood-triggered sinkholes. Moving forward towards promoting and integrating multi-hazard thinking into local resilience management and risk reduction, first a foundational understanding of its importance is required. Our observations of the results from this survey has enabled us to propose practical suggestions for future improvements. We have defined a list of actions as strategic priorities for enhancing both multi-hazard mapping and the integration of multi-hazard perspectives into local resilience and risk management programs in the Val d'Orléans, which could also benefit other departments.

This roadmap was shaped by a combination of local experts' opinions on what they consider important, as gathered from their responses in the survey, as well as the overall survey results (**Section 5.3**). The deliberate sequence of the outlined actions aims to establish a logical progression. This approach underscores the importance of raising awareness and enhancing understanding as fundamental initial steps before any substantial action can be effectively implemented. The list of actions provided is not exhaustive; Rather, it represents a set of priorities identified by local experts involved in urban planning and natural risk research. These experts have a wide range of experience and expertise in their area, with two of them leading the urban planning and risk studies at CEREMA. The ultimate goal is to encourage future research and action that will raise awareness of and resilience to multi-hazard risks, thereby contributing to more effective risk management and local resilience strategies.

5.5.1 Recommendations on practices to foster multi-hazard thinking

- *Establishing experts'/scientists' consensus.* The level of multi-hazard perception among local experts can highlight the fact that in order to have meaningful transformation and transformative actions for fostering multi-hazard thinking among decision makers and citizens, the focus should most likely first need to happen within the expert community and then pass into other community levels. Experts in their daily practices/work learn and exchange information and can act upon them. It would not be enough if only geologists or hydrogeologist know about the existing multi-hazard scenarios, but there is a need for a so-called collective brain among different experts to discuss the multi-hazard concept.

- Regarding the above-mentioned point, there is a need for experts to understand the multi-hazard concept and their amplification effects. As supported by local experts' feedback and suggestions in the recent literature (Gill et al., 2021), in order to foster the multi-hazard thinking and training among experts, more space such as physical office space and space on a curriculum should be dedicated to working across disciplines to discuss and identify potential link between single hazards. This also reinforce cross-sectoral coordination, encouraging knowledge and data sharing.

- *Policy advocacy based on expert consensus.* High-level advocacy will be necessary. Once there is a stronger consensus among local experts on the need for a multi-hazard approach, this unified stance can be leveraged to advocate for policy changes (e.g., a mandated multi-hazard focus to DICRIM outputs)

5.5.2 Future improvements in multi-hazard/susceptibility mapping

- *Map contents and structure.* To clearly communicate the purpose of the map, the following is required to be included in the map: i) the methodology applied to model the map; ii) any uncertainties relevant to the map.
- To make the map more inclusive, consideration should be given to individuals with colour vision deficiency (CVD).
- To clearly communicate the extent of 'hazard', the legend should be clear that lower hazard areas mean that there is still hazard data in the area not that it is free of hazard.
- More detailed investigations into surface and subsurface processes are needed to generate more local digital data to improve local multi-hazard mapping.
- *Enhance interdisciplinary multi-hazard mapping.* Multi-hazard approaches are typically interdisciplinary, beyond the scope of a single discipline. Looking at the results, it was revealed that the map could be used by other specialists working in groundwater and water quality sectors to identify vulnerable areas where they expect water contamination due to sinkholes. Multi-hazard maps have the potential to encourage collaboration between geologists, hydrologists, environmental scientists, and climate change researchers. By sharing insights and data, these professionals can deepen their understanding of how floods and sinkholes interact and potentially trigger other hazards. This collective effort can lead to innovative approaches to hazard forecasting and risk assessment.

- *Strengthen co-design mapping approach among experts.* Multi-hazard maps should not be produced in isolation dominated by single disciplines, but with a group of experts from different backgrounds. The interrelationships and impacts should be communicated by scientists in different domains from physical geography, natural hazards to urban planners. This takes us to this point that prior to policy makers (politics) and citizens (society), integrating multi-hazard thinking into local resilience and risk strategies should happen at the expert level (scientific community). They need to collaborate and be convinced about the amplification of the potential multi-hazard scenarios. Moreover, the detailed methodology for creating such maps should be discussed by different experts. The collaborative multi-hazard map can then be used as a communication tool to promote multi-hazard household preparedness and emergency management measures.

- Some of these recommendations could be delivered through (for example) *interdisciplinary workshops and seminars*. Organising workshops and seminars that bring experts together from diverse domains such as geology, meteorology, hydrology, urban planning, and emergency management, could be useful as platforms for sharing insights and feedback derived from initial multi-hazard maps and discussing their implications across different domains. As also supported by the experts interviewed in the current research, models and maps should be discussed during public meetings and digital exhibitions. In these events, experts can comment on the initial maps and provide practical insights, adding to the map's effectiveness. Moreover, participation of decision makers in experts' meeting can provide a ground where also experts can realistically discuss all limitations and uncertainties they face when modelling/mapping. Organising events bringing together actors communication on the topic is a highlight of the local experts' view. The learning from these events can open further discussion on the subject. Public release of the multi-hazard models should then happen by a consensus of a group of interdisciplinary experts.

- Building on the recent innovative research done by Dallo et al. (2024), we also support their recommendations for adopting similar approaches in multi-hazard risk modeling. Their research provides a comprehensive communication strategy that emphasises an iterative design approach, ensuring that the end products are user-centered and of high quality, catering to the needs of both the scientific community and broader societal stakeholders. To effectively disseminate official multi-hazard maps to the public, it is beneficial to adhere to the strategy proposed in their study, which involves:

i. A preparation phase that includes developing a communication concept, conducting end-user testing, seeking expert feedback, and establishing a network of outreach specialists;

ii. The dissemination phase, where information materials and model data are made available at events and through various distribution channels;

iii. A follow-up phase focused on responding to inquiries, providing training and webinars, and assessing the impact of the outreach activities.

This structured approach ensures that critical information reaches the intended audiences efficiently and effectively, fostering greater understanding and preparedness for multi-hazard risks.

5.5.3 Multi-hazard integration

- *Climate change*. Considering long-term climate change considerations into all aspects of planning and risk reduction measures.

- *Coordinated, multi-level risk governance.* There is a clear expectation for a coordinated, multi-level governmental approach, supported by expert agencies, in managing the pervasive risk of flood-triggered sinkholes. Therefore, to effectively coordinate and manage the risk, it is beneficial to clearly define the roles and responsibilities of each entity. This analysis should inform policymakers and community planners in designing more integrated and effective risk management frameworks that clarify roles, leverage expertise, and actively involve communities in the process.

- *Enhancement of the DICRIM information documents.* Incorporate multi-hazard scenario information prominently in the final sections of the current single-hazard information documents (DICRIM) to enhance accessibility. This strategy is critical for raising public awareness and supporting informed decision-making by ensuring that essential information is readily available to everyone. Integrating this content into familiar document formats will broaden its reach and impact. A similar activity is currently underway for the UNISDR / International Science Councile Hazard Information Profiles (UNISDR, 2024).

- *Public engagement.* Other forms of risk communication need to go alongside wider efforts to strengthen hazards education and teach the public about potential multi-hazard scenarios.

- *Iterative approaches.* The multi-hazard susceptibility map needs to be continually improved and updated based on models and experts' feedback.

5.6 Conclusion

This study contributes to the debate around the shift from a single-hazard basis approach to multiple hazards' approach and catalyse multi-hazard thinking in decision-making context. To be able to make a difference in the relatively new field of multi-hazard mapping, collaboration between disciplines is required. Exactly like natural hazards that do not happen in silo, multi-hazard mapping also should not be done in isolation by single disciplines. It is beneficial that a large group of experts in natural hazards and urban planning co-design the product for the public release. We believe that our produced multi-hazard susceptibility map in the current thesis as an initial output has the potential to bring the scientists/experts with different backgrounds around the table to foster multi-hazard thinking through science experts' dialogue. We conclude that local scientists first need to reach an understanding level of multi-hazards and then concentrate on providing a participatory mapping and the communicate these with decision makers and citizens.

Chapter 6. Discussion and Future directions

6.1 Introduction

In this final chapter, I aim to provide a comprehensive recap of the entire thesis, summarizing the key findings and discussions presented in the previous chapters. Notably, parts of this recap have already been included in the concluding Sections of **Chapters 2, 3, 4**, and **5**. However, this Chapter combines those summaries to present a cohesive overview of the research. We also discuss the general limitations, uncertainties in modelling. This chapter is concluded by providing suggestions for future research directions.

6.2 Summary

This thesis began with the presentation of the research rationale, along with its motivation, aims and objectives in **Chapter 1.** We highlighted the importance of understanding the interrelationships between multiple hazards, as discussed in various studies, emphasizing that recognizing these interactions is crucial for assessing multi-hazard landscapes. We noted that failing to understand these relationships can amplify the impact when multiple hazards occur simultaneously. Emphasizing on the importance of multi-hazard mapping along with characterisations of multi-hazard-forming areas, we clearly identified that there is a lack of multi-hazard characterisation in karst environments (**Chapter 2** visualised this lack by providing a systematic literature review). This research was driven by two motivations:

- First, there is a global lack of assessment of multi-hazards in karst areas, particularly floods and sinkholes (flood-triggered sinkholes).
- Second, there is an institutional gap in natural hazard management practices in France, where planning for mitigating natural hazards still happens in silos (single hazard assessment form). This fragmented approach is unsustainable.

Therefore, this research work can be viewed as encouraging future planning and mitigating hazards in a more inclusive/holistic approach. In summary, if we are to shorten the main core of the current research it will be as follows:

The core aims of this research was to (i) improve multi-hazard risk mamangement in the context of karst terrains; and (ii) evaluate the potential for machine-learning (the Maximum Entropy, MaxEnt, model) to support this ambition.

In **Chapter 2**, we first provided an overview of the multi-hazards concept and its historical development. The literature defines several types of hazard interrelations, which are commonly understood but sometimes referred to with different terminologies. In that regard, at the time of writing, we summarised the most recent classifications describing the interactions as triggering, amplifying, condition changing, mutually exclusive. Further, the main existing approaches to quantify (measuring the interrelationships, joint frequency and intensity of hazards) and map (in the form of multi-layers of individual hazards to show spatial multi-hazards) were discussed within a wide range of hazards in a global context.

In **Section 2.3**, we provided an explanation of hazardous processes associated with karst areas, and argued that karst terrains should be viewed within the broader category of multi-hazard zones. We then used historical real-world examples to elaborate the main potential multi-hazard chain that could happen in those regions, introducing and highlighting a new concept in understanding and managing hazards in karst environments, taking into account their distinct properties and the complex interactions of multiple hazard factors.

In **Section 2.4**, to highlight the existing gap related to multi-hazard mapping in karst areas in particular flood and sinkholes, we did a systematic literature review in which we reviewed 241 reviewed publications. The construction of our database involved a process with a great attention to selecting the representative peer review papers. It started with the identification of scientific publications through specific keywords. This was a two-step process: first, identifying relevant keywords, in which we observed variations in the terms such as “sinkhole” and “doline” applied in various regions, and secondly, scientific articles selection. Despite the terminological variances, “sinkhole” was the most frequent terminology found in the international literature to guide our keyword search.

The application of these key words was essential in sourcing our relevant articles from different web search engines focusing on English-written international literature journals with a specific emphasis on sinkhole hazard studies especially on flooding. The screening process involved a careful examination of titles and abstracts to narrow down relevant papers. During this process, we excluded duplicates or similar studies. Finally, in the classification phase, we selected 241 references, extracting key information to identify and define seven thematic subjects within the field of sinkhole studies. This structured approach enabled a comprehensive and organized literature review (with key limitations discussed in **Section 2.6**).

The literature review revealed that while exsiting research has studied sinkholes and their contributing factors, few studies (if any) have perceived and mapped multi-hazard scenarios relating to flooding and sinkholes in karst areas. Since globally, a shift from single to multi-hazard assessment is being encouraged by international risk communities, as already mentioned in **Chapter 1,** the present research was to provide a multi-hazard susceptibility map of the flood-triggered sinkholes emphasising multi-hazard approaches.

As the literature discussed, ML algorithms generally excel at forecasting and dealing with big and dimensional data, compared to traditional methods. They are powerful in finding subtle and non-linear patterns in data, and more advanced at handling and interpreting complex, multi-dimensional data. Such patterns and relationships among data can be overlooked in traditional methods. Throughout the literature, the application of these new ML predictive models has been explored in different geoscience fields including sinkholes. They yield hazard susceptibility evaluations that are both less biased and consistent, with the potential for reaching high spatial resolution. However, more algorithoms need to be explored and examined.

In advancing my research, I applied the Maximum Entropy modelling approach, commonly abbreviated to MaxEnt, to model flood-triggered sinkholes, an area that despite many studies globally and regionally in ecological studies, remains largely unexplored in terms of integrated natural hazard assessment. This

integration is essential in karst terrains, where the unique geological features considerably influence both flood dynamics and sinkhole formation. Karst regions are particularly susceptible to these geo-hazards, yet the interplay between flooding and sinkhole occurrences has not been sufficiently investigated and mapped. Moreover, despite of the past occurrence of this multi-hazard event in Val d'Orléans, the region lacks this integration in a form of multi-hazard susceptibility map, which demonstrate the area susceptible to both hazards.

In **Chapter 3**, we began by outlining the research design and methodology, emphasizing the MaxEnt model's application to hazard mapping. It introduced the Val d'Orléans region as a case study area, highlighting its unique geological and hydrogeological characteristics that make it susceptible to natural hazards such as sinkholes and floods. Moreover, detailed explanations of the MaxEnt model were provided, tracing its origins from statistical mechanics and its adoption in ecological modeling. The principles of MaxEnt, including Shannon entropy and the use of presence-background data, were thoroughly discussed to establish a theoretical foundation.

The geographical, geological, and hydrogeological contexts of Val d'Orléans were described, setting the stage for understanding the region's susceptibility to sinkholes and floods. This section included the historical context and the interaction between different hazard types. Further, in the same chapter, the data collection process was outlined, including the preparation of (geo) environmental variables and the sinkhole inventory. This section detailed the selection of geo-environmental factors such as topography, surficial geology, alluvium thickness, groundwater levels, and flood hazard zones. The practical application of the MaxEnt model was described, including data preparation (cleaning, manipulation), parameter tuning, and model building. We also provided a methodological flowchart to illustrated the step from data collection to model application and validation.

In **Chapter 4**, the results of the application of Maximum Entropy (MaxEnt) model were presented and discussed. By incorporating six key contributing factors, we aimed to provide a spatial forecasting map of multi-hazard areas targeting flood-triggered sinkholes.

Unlike traditional approaches that simply overlay different hazards, our model integrated official flood hazard zones, recognizing flooding as a critical trigger for sinkhole formation in our case study. To support land use and natural hazard mitigation planning in Val d'Orléans, we also compared and visualized the spatial extents affected by flood as a single hazard and those susceptible to flood-induced sinkholes, representing a complex multi-hazard scenario. Additionally, we tested the MaxEnt model over a broader area, confirming that larger study areas can inflate AUC values, potentially leading to misleading interpretations of model performance (discussed in **Appendix 3**).

Chapter 5 transitions from generating scientific insights to evaluating our product (the resultant multi-hazard susceptibility map) by engaging local experts/scientists as stakeholders. Their feedback and opinions were crucial in shaping the future direction of this thesis topic. The methodology involves engaging local experts through an online survey and one interview via Teams application. To gather their input, we provided a questionnaire accessible online to the selected participants. While the primary aim was to assess our map, we also pursued additional objectives:

- To examine stakeholder's awareness and perceptions of multi-hazard events
- To evaluate the produced multi-hazard susceptibility map.
- To understand participants' perspectives on ownership and pervasive nature of flooding and sinkhole hazards in the area.

- To understand participants' perspectives on establishing a dedicated DICRIM document for multi-hazard interactions.

The findings reveal the need for improved institutional communication and participatory approaches in multi-hazard mapping. The study concludes with recommendations for fostering multi-hazard thinking among experts and policymakers, advocating for collaborative and transparent mapping processes to enhance community resilience. In the following sections, we will discuss the limitations of this research, and then summarize the findings and recommendations previously mentioned in **Chapters 4** and **5**.

6.3 Research Limitations

6.3.1 Limitations on multi-hazard mapping

Uncertainty is a global challenge in the natural hazard context and identifying them could set a realistic expectation to see whether they can be reduced in future (Doyle et al., 2019). According to the research findings done by Doyle et al. (2019), communicating uncertainties is an ongoing research and so far, what is certain is that we cannot overwhelm decision makers by transferring all uncertainties related to modelling but rather it is important to know the decision maker needs and timeframe then based on that we communicate the relevant uncertainties.

There are numerous levels of uncertainty in science advice in natural hazards. In general, there are two general sources of uncertainty: *aleatory* and *epistemic* uncertainties. Aleatory uncertainty (natural stochastic uncertainty) stems from the randomness of natural events in their time, location and magnitude of occurrences. This type of uncertainty cannot be eliminated since it is a fundamental aspect of natural variability (Rougier and Beven, 2013).

In the current research, discussions about uncertainty mainly focus on epistemic uncertainty. Epistemic uncertainty arises from limited current knowledge, and is characterized by their uncertainty regarding inadequate understanding, incomplete information, and undifferentiated alternatives. It is important to reflect on a few limitations that introduce some degree of uncertainties into the modelling of multi-hazard susceptibility in this study.

i) As quoted by Galve et al. (2009) and Taheri et al. (2019), the quality of predictive/forecasting models depends heavily on the **completeness and representativeness of the sinkhole inventory**. Specifically, for mapping the susceptibility of flood-triggered sinkholes, having an inventory that documents only sinkhole collapses triggered by flood events might yield more accurate results. Nevertheless, obtaining this information typically presents challenges, such as data scarcity and potential limitations in coverage. This also underscores the importance of recording sinkhole triggers along with their respective locations, itself requiring sufficient monitoring to support identification of newly triggered sinkholes.

ii) **The challenge of integrating fine-scale geological features in modelling for regional mapping.** Another limitation is that in an area characterised by a high flood hazard, particularly in the context of flood-triggered sinkholes, it becomes evident that there exist localized geological features and underground karst conduits that pose major challenges for inclusion in modelling efforts. These localized factors are primarily related to the complex geological characteristics of the region. For instance, certain areas may exhibit distinct lines of geological weakness and fractures that facilitate the movement of water. Consequently, water is more likely to erode materials in these areas, increasing the risk of sinkhole formation. Integrating such fine-scale geological information into a model (more importantly regional mapping) can be challenging, as it requires highly detailed geological maps with fine resolution. These features represent specific susceptibilities that may not be adequately accounted for in standard modelling approaches like MaxEnt. As a result, a more in-depth investigation, such as detailed mapping and site-specific assessments, becomes essential for gaining a comprehensive

understanding of the likelihood and precise locations of flood-triggered sinkhole occurrences, especially in areas with complex geological characteristics. We also acknowledge that despite our careful and informed approach in choosing parameters for the MaxEnt model, there remains an inherent uncertainty associated with tuning the model parameter settings.

iii) **MaxEnt modelling for Natural hazards.** MaxEnt is a recent modelling approach, integrated into machine-learning by ecologists since 2006. It has been typically used in ecological contexts. Exploring its application for natural hazards is still an ongoing research, and in some areas, it requires a different approach when applied to natural phenomena such as flood-triggered sinkholes. Unlike species distribution, where spatial autocorrelation can complicate the understanding of ecological processes and in some cases where techniques like filtering are required to address spatial bias (Paradinas et al., 2023), in sinkhole modelling, spatial clustering is informative and should be embraced. This is because each sinkhole observation provides direct information about the local geological and hydrological conditions conducive to sinkhole formation. Recognizing spatial autocorrelation in sinkhole data can improve hazard susceptibility zoning and guide detailed geotechnical investigations, thus enhancing the accuracy of hazard assessment and prevention. Therefore, the spatial effect which can be misleading for living species is in fact informative for natural hazards. This nuance understanding will save time and provide more accuracy for future modellers when it comes to data portioning or field data collection.

Moreover, despite its many advantages as a machine learning tool, MaxEnt's performance is strongly affected by how complex the model is and how well the data fits its assumptions. The size of the study area also highly matters. If the area is too large it can make the AUC value seem better than it really is, which can be misleading. This is a big issue for ensuring the accuracy and reliability of MaxEnt models in predicting natural hazards.

Table 6.1 The main sources of model uncertainty (adapted from Kim, 2019)

Source of uncertainty	Type	Reason
Hazard identification	Aleatory	The inherently unpredictable variation of certain environmental elements
	Epistemic: data	i) Inaccurate measurement techniques of some environmental factors. ii) Lack of data on sinkholes specifically triggered by flooding.
	Epistemic: model parameters	The choice of classification methods standard for susceptibility to each hazard.
	Epistemic: science	Certain unidentified environmental elements that may affect the spread of natural hazards (contributing factors in fine-scales)
Hazard/ Susceptibility analysis	Aleatory	The intrinsically random variability of hazard occurrence.
	Epistemic: data	The precision of data collection methods and the availability of comprehensive historical data. Limited historical data can lead to less accurate hazard analysis.
	Epistemic: model parameters	When using machine learning models to analyze hazards, the choices made during parameter tuning (like learning rates, regularization factors, etc.) can introduce uncertainty into the model's predictions.
	Epistemic: science	Other known and unknown trigger factors, which could induce other hazards.
Hazard interaction analysis	Aleatory	Similar to hazard identification, this reflects the random nature of when and where hazards, such as floods and sinkholes, might occur.
	Epistemic: science	Other unknown relationships between different natural hazards. For instance, how a flood might trigger a sinkhole or vice versa in reality, and the underlying mechanisms of these interactions might not be fully understood.

6.4 Future directions for multi-hazard mapping/modelling:

- *Detailed hazard analysis:* The produced map is useful for identifying regions that require further detailed hazard analysis. By using the resultant map from the current research, local zones can be pinpointed for closer examination of subsurface conditions. This approach considers the potential hazard interactions when planning urban environments and evacuation programs, helping to avoid mistakes made in the past.

- *Temporal likelihood and early warning systems:* Further research can explore the topic of temporal likelihood and if any relevant thresholds or tipping points can be identified. Given the susceptibly of the area to flooding and sinkholes, doing research on establishing sinkhole warning systems based on the rainfall amount could be an effective measure towards mitigating multi-hazards.

- *Multi-hazard scenario expansion:* Regarding the area being karstified and its susceptibility to flood triggered sinkholes, the multi-hazard scenario is needed to go further to include other different potential hazard chain scenarios might be happening in the karst areas. For instance, the reverse multi-hazard chain where the old sinkholes (already collapsed) cause flooding by not being able to let the runoff water from heavy rainfall go into the ground.

- *Risk assessment (vulnerability * exposure)*: To have a more comprehensive view of risk map, vulnerability and exposure assessment and mapping is required.

- *Regular updates of multi-hazard maps*: Multi-hazard/susceptibility and future risk maps need to be kept up-to date. It is important to regularly reassess methods as inventory data being completed.

- The use of MaxEnt modelling in natural hazards as a rather new modelling will continue in future. Therefore, parameters tuning, dealing with biases and evaluating the outcomes need to be developed more by natural hazard modellers and developers.

- Reducing, characterizing, and exploring the impacts of the map uncertainties on decision-making are critical steps forward.

6.5 Future directions for multi-hazard integration into the disaster risk reduction plan in the study region

The current research highlights the susceptibility of Val d'Orléans to multiple hazards, especially flood-triggered sinkholes. To enhance local resilience and risk management, it is essential to understand the importance of integrating multi-hazard thinking. The study proposes practical suggestions for improving multi-hazard mapping and integration into local strategies, based on expert opinions and survey results.

The future directions were thoroughly discussed in **Chapter 5** (***Section 5.5***) under the title of "A way forward". Here is a general recap of key points:

i) Recommendations on practices to foster multi-hazard thinking

- Expert consensus: Foster consensus among experts to transform multi-hazard thinking, emphasizing collaboration among diverse disciplines.
- Amplification effects: Dedicate resources to cross-disciplinary discussions to understand and address the interconnectedness of hazards.
- Policy advocacy: Leverage expert consensus to advocate for policy changes supporting a multi-hazard approach (e.g., a mandated multi-hazard focus to DICRIM outputs).

ii) Future Improvements in Multi-Hazard/Susceptibility Mapping

- *Map Contents and Structure*. To clearly communicate the purpose of the map, highlight uncertainties, and make maps inclusive for individuals with colour vision deficiency (CVD).
- Broad investigations: Invest in comprehensive studies of surface and subsurface processes to improve local multi-hazard mapping.
- Interdisciplinary collaboration: Encourage collaboration among geologists, hydrologists, environmental scientists, and climate researchers to enhance hazard forecasting and risk assessment.
- Strengthen co-design mapping: Promote co-design mapping approaches among experts from various fields to create comprehensive and effective multi-hazard maps.

iii) Multi-Hazard Integration in Disaster risk reduction documents (DICRIM)

- Incorporate multi-hazard information into existing single-hazard documents to raise public awareness and support informed decision-making.
- Define clear roles and responsibilities for entities involved in managing flood-triggered sinkholes, supported by expert agencies.
- Climate change considerations: Integrate long-term climate change considerations into all aspect of planning and risk reduction measures.

- Regularly update the multi-hazard susceptibility map based on models and expert feedback (Iterative approach).

6.6 Final thoughts

This thesis has built on existing sinkhole studies from a multi-hazard perspective. The results confirm that future risk reduction management in the case study, and more generally in other karst settings should be addressed from a multi-hazard approach lens. It advocates for multi-hazard assessment in karst areas rather than viewing relevant hazards separately. This approach allows for more efficient use of resources by identifying and mitigating overlapping areas where interlinked hazards overlap, such as in the communes of *Bou*, *Darvoy*, *Saint-Denis-en-Val*, and *Chécy*. Our multi-hazard susceptibility map focused on a two-hazard scenario where floods could potentially trigger sinkholes. This was based on integrating historical occurrences of sinkholes, hydrological, geological, and topographical factors, in addition to flood hazard zones. The results show that elevation is the main contributing factor, with low-elevated zones being more exposed to flooding and sinkholes in an area characterized by karst and floodplain. Therefore, this research provided empirical evidence for the institutional rules and administrative structures of response organisations to be redesigned and moving towards multi-hazard approach particular in karst environments.

This exploratory research applied the Maximum Entropy (MaxEnt) model due to its less explored attributes in the realm of natural hazards, particularly multi-hazards such as flood-triggered sinkholes. Focusing on the MaxEnt model, the thesis aimed to demonstrate its capabilities and contribute new insights to the field. The use of MaxEnt modelling in natural hazards has shown potential for identifying susceptible areas and improving predictive accuracy, thereby enhancing the overall hazard assessment process.

In a broad implication, the produced multi-hazard map can inform residents, city planners, and emergency services about potential risks and help in planning mitigation strategies. This practical application underscores the importance of integrating multi-hazard considerations into urban planning and emergency preparedness. Moreover, the map can be used by researchers working in water quality management and other potential relavent field to understand the areas more vulnerable to water contaminations. Finally, I would like to emphesise that the learning from this thesis (both modelling and stakeholder engagement parts) can be beneficial to a wider set of locations and even other multi-hazard scenrarios.

References (APA format)

Abdi, H., Williams, L.J. (2010). Jacknife. In Neil Salkind (Ed.), Encyclopedia of Research Design. Thousand Oaks, CA: Sage, Retrieved from: efaidnbmnnnibpcajpcglclefindmkaj/https://web.utk.edu/~ewbrewer/pdf/encylopedia/Encyclopedia%20of%20Research%20Design_Volume%202.pdf

Abdollahi, S., Pourghasemi, H. R., Ghanbarian, G. A., and Safaeian, R. (2019). Spatial simulation and land-subsidence susceptibility mapping using Maximum Entropy model. Journal of Watershed Management Research, 10(20), 133-144, https://jwmr.sanru.ac.ir/article-1-924-en.pdf.

Alrowaim, Galve, J.P., Bonachea, J., Remondo, J., Gutiérrez, F., Guerrero, J., Lucha, P., Cendrero, A., Gutiérrez, M., and Sánchez, J.A. (2008). Development and validation of sinkhole susceptibility models in mantled karst settings: A case study from the Ebro valley evaporite karst (NE Spain): Engineering Geology, v. 99, no. 3−4, p. 185−197, doi.org/10.1016/j.enggeo.2007.11.011.

Amin, P., Ghalibaf, M. A., Mermut, A. R., and Delavarkhalafi, A. (2023). Assessment of zones prone to sinkhole using ground penetrating radar and soil properties in Central Iran. *Geoderma Regional*, e00630, doi.org/10.1016/j.geodrs.2023.e00630.

Appelquist, L. R., and Balstrøm, T. (2015). Application of a new methodology for coastal multi-hazard- assessment and management on the state of Karnataka, India. Journal of Environmental Management, 152, 1-10, doi.org/10.1016/j.jenvman.2014.12.017.

Atzori, S., Baer, G., Antonioli, A., and Salvi, S. (2015). InSAR-based modeling and analysis of sinkholes along the Dead Sea coastline. Geophysical Research Letters, 42(20), 8383-8390, doi.org/10.1002/2015GL066053.

Auterives, C., Binet, S., and Albéric, A. (2011). Inferred conduit network geometry from geological evidence and water head measurements in a fluvio-karstic system (Val d'Orleans, France). In 9th Conference on Limestone Hydrogeology H2KARST – Besançon – France (communication orale et actes).

Azareh, A., Rahmati, O., Rafiei-Sardooi, E., Sankey, J. B., Lee, S., Shahabi, H., and Ahmad, B. B. (2019). Modelling gully-erosion susceptibility in a semi-arid region, Iran: Investigation of applicability of certainty factor and maximum entropy models. Science of the Total Environment, 655, 684-696, doi.org/10.1016/j.scitotenv.2018.11.235.

Banerjee, P. (2021). Maximum entropy-based forest fire likelihood mapping: Analysing the trends, distribution, and drivers of forest fires in Sikkim Himalaya. Research Square. doi.org/10.21203/rs.3.rs-22646/v3.

Banks, V J, Bricker, S H, Royse, K R, and Collins, P E F. (2015). Anomalous buried hollows in London: development of a hazard susceptibility map. Quarterly Journal of Engineering Geology and Hydrogeology 48, 1, 55-70.

Baryakh, A. A., and Fedoseev, A. K. (2011). Sinkhole formation mechanism. *Journal of Mining Science*, *47*(4), 404-412, doi.org/10.1134/S1062739147040022.

Bátori Z., Csiky J., Farkas T., Vojtkó A.E., Erdős L., Kovács D., Wirth T., Körmöczi L. and Vojtkó A. (2014). The conservation value of karst dolines for vascular plants in woodland habitats of Hungary: refugia and climate

change. International Journal of Speleology, 43 (1), 15-26. Tampa, FL (USA) ISSN 0392-6672 doi.org/10.5038/1827-806X.43.1.2.

Beguería, S. (2006). Validation and evaluation of predictive models in hazard assessment and risk management. Nat. Hazards 37, 315–329, doi.org/10.1007/s11069-005-5182-6.

Beitollahi, A, Mahdavi, M. (2021). A report of Nasimshahr sinkholes in sout-west of Tehran,Iran, Road, Housing and Urban Development Research Center. In Persian.

Ben-Gal I. and Trister C. (2015). Parallel Construction of Decision Trees with Consistently Non Increasing Expected Number of Tests., Applied Stochastic Models in Business and Industry, Vol. 31(1) 64-78. 10.1002/asmb.2086, doi.org/10.1002/asmb.2086

Beuthe, M. (2021). Isostasy with Love–I: elastic equilibrium. Geophysical Journal International, 225(3), 2157-2193, doi: 10.1093/gji/ggab073.

Bianchini, S., Confuorto, P., Intrieri, E., Sbarra, P., Di Martire, D., Calcaterra, D., and Fanti, R. (2022). Machine learning for sinkhole risk mapping in Guidonia-Bagni di Tivoli plain (Rome), Italy. Geocarto International, doi.org/10.1080/10106049.2022.2113455.

Blanchard, R. (1903). Le Val d'Orléans. Annales de Géographie, 12(64), 307-323. doi.org/10.3406/geo.1903.6348.

Boussouf, S., Fernández, T. and Hart, A.B.(2023) Landslide susceptibility mapping using maximum entropy (MaxEnt) and geographically weighted logistic regression (GWLR) models in the Río Aguas catchment (Almería, SE Spain). *Nat Hazards* **117**, 207–235, doi.org/10.1007/s11069-023-05857-7.

Brezinski, D. K. (2007). Geologic and anthropogenic factors influencing karst development in the Frederick region of Maryland. *Environmental Geosciences*, 14(1), 31-48, doi.org/10.1306/eg.01050605014.

BRGM. (n.d.). Hydrogeology. Retrieved May 16, 2024, from https://sigescen.brgm.fr/-Hydrogeologie-.html.

BRGM. (7 February 2024). Vous vous souvenez de l'A10, sous l'eau, en 2016?. X (former Tweeter), https://x.com/BRGM_fr/status/1755238103885742434.

Brinkmann R, Parise M, Dye MD. (2008). Sinkhole distribution in a rapidly developing urban environment: Hillsborough County, Tampa Bay, Florida. Eng Geol 99(3–4):169–184, doi.org/10.1016/j.enggeo.2007.11.020.

Brinkmann, R. (2013). Florida Sinkholes. University Press of Florida, ISBN 978-0-8130-4495-8.

Breiman, L. (2001). Random forests,"Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.

Bordbar, M., Aghamohammadi, H., Pourghasemi, H. R., and Azizi, Z. (2022). Multi-hazard spatial modeling via ensembles of machine learning and meta-heuristic techniques. Scientific Reports, 12(1), 1451, doi.org/10.1038/s41598-022-05364-y.

Campbell, C. W. (2005). Complexities of flood mapping in a sinkhole area. In Sinkholes and the Engineering and Environmental Impacts of Karst (pp. 470-478), doi.org/10.1061/40796(177)50.

Centre for Research on the Epidemiology of Disasters (CRED). (2024). 2023: Disasters in Numbers. Brussels: CRED. Retrieved from https://files.emdat.be/reports/2023_EMDAT_report.pdf.

Chatterjee, S., and Hadi, A. S. (2013). Regression Analysis by Example, John Wiley and Sons, Hoboken, New Jersey, Fifth edition, doi.org/10.1002/9781118532843.

Chen Z, Auler AS, Bakalowicz M, Drew D, Griger F, Hartmann J, Jiang G, Moosdorf N, Richts A, Stevanovic Z, Veni G, Goldscheider N. (2017). The world karst aquifer mapping project: concept, mapping procedure and map of Europe. Hydrogeol J 25(3):771–785, doi.org/ 10.1007/s10040-016-1519-3.

Chen, H. X., Zhang, S., Peng, M., and Zhang, l. M. (2016). A physically-based multi-hazard risk assessment platform for regional rainfall-induced slope failures and debris flows. engineering geology, 203, 15-29, doi.org/10.1016/j.enggeo.2015.12.009.

Choi, E., Ha, J. G., Hahm, D., and Kim, M. K. (2021). A review of multihazard risk assessment: Progress, potential, and challenges in the application to nuclear power plants. International Journal of Disaster Risk Reduction, 53, 101933, doi.org/10.1016/j.ijdrr.2020.101933.

Ciotoli, G., Di Loreto, E., Finoia, M. G., Liperi, L., Meloni, F., Nisio, S., and Sericola, A. (2016). Sinkhole susceptibility, Lazio region, central Italy. Journal of Maps, 12(2), 287-294, doi.org/10.1080/17445647.2015.1014939.

Ciurean, R., Gill, J., Reeves, H. J., O'Grady, S., and Aldridge, T. (2018). Review of environmental multi-hazards research and risk assessments (OR/18/057). British Geological Survey, Nottingham, UK. http://nora.nerc.ac.uk/id/eprint/524399/1/OR18057.pdf (last accessed April 13, 2023).

Clark, P.J., and Evans, F.C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations: Ecology, v. 35, p. 445-453, doi.org/10.2307/1931034.

Clive, M. A. T., Lindsay, J. M., Leonard, G. S., Lutteroth, C., Bostrom, A., and Corballis, P. (2021). Volcanic hazard map visualisation affects cognition and crisis decision-making. International journal of disaster risk reduction, 55, 102102, doi.org/10.1016/j.ijdrr.2021.102102.

Constantinou, S., and Van Rooy, J. L. (2018). Sinkhole and subsidence size distribution across dolomitic land in Gauteng. Journal of the South african institution of civil engineering, 60(2), 2-8, doi.org/10.17159/2309-8775/2018/v60n2a1.

Dalgleish, Janet B; Alexander Jr., E Calvin. (1984). Sinkhole Distribution in Winona County, Minnesota. Proceedings of the First Multidisciplinary Conference on Sinkholes. Orlando, Florida 15-17 October 1984.

Dallo, I., Marti, M., Valenzuela, N., Crowley, H., Dabbeek, J., Danciu, L., Zaugg, S., Cotton, F., Giardini, D., Pinho, R., Schneider, J. F., Beauval, C., Correia, A. A., Ktenidou, O.-J., Mäntyniemi, P., Pagani, M., Silva, V., Weatherill, G., and Wiemer, S.(2024). The communication strategy for the release of the first European Seismic Risk Model and the updated European Seismic Hazard Model, Nat. Hazards Earth Syst. Sci., 24, 291–307, doi.org/10.5194/nhess-24-291-2024.

Dallo, I., Stauffacher, M., and Marti, M. (2020). What defines the success of maps and additional information on a multi-hazard platform?. International Journal of Disaster Risk Reduction, 49, 101761. doi.org/10.1016/j.ijdrr.2020.101761

Davis, J., and Blesius, L. (2015). A hybrid physical and maximum-entropy landslide susceptibility model. Entropy, 17(6), 4271-4292.

De Bruyn, I. A., and Bell, F. G. (2001). The occurrence of sinkholes and subsidence depressions in the Far West Rand and Gauteng Province, South Africa, and their engineering implications. Environmental and Engineering Geoscience, 7(3), 281-295, doi.org/10.2113/gseegeosci.7.3.281.

de Ruiter, M.C., Couasnon, A., van den Homberg, M.J.C., Daniell, J.E., Gill, J.C., and Ward, P.J. (2020). Why we can no longer ignore consecutive disasters, Earth's Future, 8, e2019EF001425, doi.org/10.1029/2019EF001425.

de Ruiter, M. C., de Bruijn, J. A., Englhardt, J., Daniell, J. E., de Moel, H., & Ward, P. J. (2021). The asynergies of structural disaster risk reduction measures: Comparing floods and earthquakes. *Earth's Future*, *9*(1), e2020EF001531.

De Waele, J., Gutiérrez, F., Parise, M., and Plan, L. (2011). Geomorphology and natural hazards in karst areas: a review. Geomorphology, 134(1-2), 1-8, doi:10.1016/j.geomorph.2011.08.001.

Desprez, N. (1967). Etude hydrogéologique du calcaire de Beauce – Inventaire du bassin versant de la Loire en amont d'Orléans, Retreived from: https://france3-regions.francetvinfo.fr/centre-val-de-loire/loiret/orleans/orleans-effondrement-imminent-de-deux-immeubles-rue-de-bourgogne-2684636.html.

Devi, G. K., Ganasri, B. P., and Dwarakish, G. S. (2015). A Review on Hydrological Models. *Aquatic Procedia*, 4, 1001-1007. doi:10.1016/j.aqpro.2015.02.126.

di Santolo, A. S., Forte, G., De Falco, M., and Santo, A. (2016). Sinkhole Risk Assessment in the Metropolitan Area of Napoli, Italy. Procedia Engineering, 158, 458-463, doi:10.1016/j.proeng.2016.08.472.

Di Mauro, C., Bouchon, S., Carpignano, A., Golia, E., and Peressin, S. (2006). Definition of multi-risk maps at regional level as management tool: Experience gained by civil protection authorities of Piemonte region. In *Proceedings of the 3rd International Conference on Early Warning* (EWC III), Bonn, Germany., Retrieved from conference.ing.unipi.it/vgr2006/archivio/Archivio/2008/Archivio/2006/Articoli/700196.pdf.

DICRIM. (2021). Document d'Information Communal sur les Risques Majeurs. Document conçu par la Mairie d'Orléans - Création graphique : Laure Scipion - Septembre 2021. Orléans Métropole. Retrieved from https://www.orleans metropole.fr/fileadmin/orleans/MEDIA/kiosque/environnement/risques_majeurs/dicrim.pdf.

Dilley, M., Chen, U., Deichmann, R.S., Lerner-Lam, A., Arnold, M. (2005). Natural disaster hotspots: a global risk analysis. In: Disaster Risk Management Series, 5.The World Bank. Retrieved from efaidnbmnnnibpcajpcglclefindmkaj/https://documents1.worldbank.org/curated/en/621711468175150317/pdf/344230PAPER0Na101official0use0only1.pdf.

Cerema Normandie-Centre , DDT45. (2014). L'aléa effondrement lié à des phénomènes karstiques dans le Val d'Orléans _ Approche bibliographique. Rapport n°131320/A.

Doctor, K. Z., Doctor, D. H., Kronenfeld, B., Wong, D. W., and Brezinski, D. K. (2008a). Predicting sinkhole susceptibility in Frederick Valley, Maryland, using geographically weighted regression. In 11th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst. Tallahassee, Florida, United States: American Society of Civil Engineers (pp. 100-102).

Doctor, K. Z. (2008b). GIS and Spatial Statistical Methods for Determining Sinkhole Potential in Frederick Valley, Maryland. Conference: U.S. Geological Survey Karst Interest Group Proceedings at: Bowling Green, KentuckyVolume: U.S. Geological Survey Scientific Investigations Report 2008-5023.

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78-87. doi:10.1145/2347736.2347755.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., and Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography, 36(1), 27-46, doi.org/10.1111/j.1600-0587.2012.07348.x.

Dournel, P. (2019). L'eau, vecteur de projets communs, à contrecourant de politiques sectorielles, concurrentielles et autocentrées. Gestion et requalification des milieux fluviaux dans le département du Loiret. Sciences Eaux and Territoires, 2019(27). Retrieved from https://journals.openedition.org/soe/5308#tocto1n2.

Douvinet, J., Pallares, R., Genre-Grandpierre, C., Gralepois, M., Rode, S., and Servain-Courant, S. (2013). L'information sur les risques majeurs à l'échelle communale. Occurrence et facteurs explicatifs du DICRIM, un outil préventif sous-utilisé. Cybergeo: European Journal of Geography, doi.org/10.4000/cybergeo.2611.

Doyle, E. E., Johnston, D. M., Smith, R., and Paton, D. (2019). Communicating model uncertainty for natural hazards: A qualitative systematic thematic review. *International journal of disaster risk reduction*, *33*, 449-476, doi.org/10.1016/j.ijdrr.2018.10.023.

Dunant, A., Bebbington, M., and Davies, T. (2021). Probabilistic cascading multi-hazard risk assessment methodology using graph theory, a New Zealand trial. *International Journal of Disaster Risk Reduction*, *54*, 102018, doi.org/10.1016/j.ijdrr.2020.102018.

Durlević, U., Novković, I., Lukić, T., Valjarević, A., Samardžić, I., Krstić, F., and Ćurić, V. (2021). Multihazard susceptibility assessment: A case study–Municipality of Štrpce (Southern Serbia). Open Geosciences, 13(1), 1414-1431.

Eini, M., Kaboli, H. S., Rashidian, M., and Hedayat, H. (2020). Hazard and vulnerability in urban flood risk mapping: Machine learning techniques and considering the role of urban districts. International Journal of Disaster Risk Reduction, 50, 101687. https://www.sciencedirect.com/science/article/pii/S2212420920303484.

Elith, J., and Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual review of ecology, evolution, and systematics*, *40*, 677-697. 10.1146/annurev.ecolsys.110308.120159.

Elith, J., Phillips, S.J., Hastie, T., Dudık, M., Chee, Y.E. and Yates, C.J.(2011). A statistical explanation of MaxEnt for ecologists. Diversity and Distributions, 17, 43–57, https://doi.org/10.1111/j.14724642.2010.00725.x.

Elmahdy, S. I., Mohamed, M. M., Ali, T. A., Abdalla, J. E. D., and Abouleish, M. (2022). Land subsidence and sinkholes susceptibility mapping and analysis using random forest and frequency ratio models in Al Ain, UAE. *Geocarto International*, *37*(1), 315-331.

Engeset, R. V., Pfuhl, G., Orten, C., Hendrikx, J., and Hetland, A. (2022). Colours and maps for communicating natural hazards to users with and without colour vision deficiency. International Journal of Disaster Risk Reduction, 76, 103034. doi.org/10.1016/j.ijdrr.2022.103034.

European Commission. (2011). Risk assessment and mapping guidelines for disaster management. Commission staff working paper, European Union.

Ezersky, M. G., Bodet, L., Akawwi, E., Al-Zoubi, A. S., Camerlynck, C., Dhemaied, A., and Galibert, P. Y. (2013). Seismic surface-wave prospecting methods for sinkhole hazard assessment along the Dead Sea

shoreline. Journal of Environmental and Engineering Geophysics, 18(4), 233–253, doi.org/10.2113/JEEG18.4.233.

Faramarzi, H., Hosseini, S. M., Pourghasemi, H. R., and Farnaghi, M. (2023). Using machine learning techniques in multi-hazards assessment of Golestan National Park, Iran. Natural Hazards, 117(3), 3231-3255, doi.org/10.1007/s11069-023-05984-1.

Ferrer, L., Curt, C., and Tacnet, J. M. (2018). Analysis of a risk prevention document using dependability techniques: a first step towards an effectiveness model. *Natural Hazards and Earth System Sciences*, *18*(4), 1201-1221, doi.org/10.5194/nhess-18-1201-2018.

Fischer, P., Pistre, S., and Marchand, P. (2022). Effect of fast drainage in karst sinkholes on surface runoff in Larzac Plateau, France. Journal of Hydrology: Regional Studies, 43, 101206, doi.org/10.1016/j.ejrh.2022.101206.

Fleury, S. (2009). Land use policy and practice on karst terrains: Living on limestone. Springer Science & Business Media. doi.org/10.1007/978-1-4020-9670-9.

Ford, D.C. and Williams P.W. (2007). Karst geomorphology and hydrology. 2nd ed. John Wiley and Sons, Chichester,U.K, available at: https://link.springer.com/book/9789401177801.

Forzieri, G., Feyen, L., Russo, S., Vousdoukas, M., Alfieri, L., Outten, S., and Cid, A. (2016). Multi-hazard assessment in Europe under climate change. *Climatic Change*, *137*, 105-119, doi.org/10.1007/s10584-016-1661-x.

France 3. (2022). Orléans: Effondrement imminent de deux immeubles rue de Bourgogne. France 3 Régions. Retrieved from https://france3-regions.francetvinfo.fr/centre-val-de-loire/loiret/orleans/orleans-effondrement-imminent-de-deux-immeubles-rue-de-bourgogne-2684636.html.

Fawcett, T. (2006). An introduction to ROC analysis. Pattern recognition letters, 27(8), 861-874. doi:10.1016/j.patrec.2005.10.010.

Gallina, V., Torresan, S., Critto, A., Sperotto, A., Glade, T., and Marcomini, A. (2016). A review of multi-risk 945 methodologies for natural hazards: Consequences and challenges for a climate change impact assessment, J. Environ. Manag., 168, 123–132, doi.org/10.1016/j.jenvman.2015.11.011.

Galve JP, Remondo J, Gutiérrez F. (2011). Improving sinkhole hazard models incorporating magnitude-frequency relationships and nearest neighbor analysis. Geomorphology. 2011; 134: 157–170. doi.org/10.1016/j.geomorph.2011.05.020.

Galve, J.P., Gutierrez, F., Lucha, P., Guerrero, J., Bonachea, J., Remondo, J., Cendrero, A., (2009a). Probabilistic sinkhole modelling for hazard assessment. Earth Surface Processes and Landforms 34, 437–452. doi.org/10.1002/esp.1753.

Galve, J.P., Gutiérrez, F., Remondo, J., Bonachea, J., Lucha, P., Cendrero, A. (2009b). Evaluating and comparing methods of sinkhole susceptibility mapping in the Ebro Valley evaporite karst (NE Spain). Geomorphology 111, 160–172

Gao Yongli, Alexander, E.C., and Barnes, R. (2005). Karst database implementation in Minnesota: analysis of sinkhole distribution: Environmental Geology, v. 47, no. 8, p. 1083–1098. doi.org/10.1007/s00254-005-1241-2.

Gao, Y., and Alexander, E. C. (2008). Sinkhole hazard assessment in Minnesota using a decision tree model. *Environmental Geology*, *54*(5), 945-956. https://link.springer.com/article/10.1007/s00254-007-0897-1.

GeoEngineers. (2016, January 13). Sinkholes: A threat after heavy rains. Retrieved from https://www.geoengineers.com/news/sinkholes-a-threat-after-heavy-rains/

Gill, J. C., and Malamud, B. D.(2016). Hazard interactions and interaction networks (cascades) within multi- hazard methodologies, Earth Syst. Dynam., 7, 659–679, doi.org/10.5194/esd-7-659-2016.

Gill, J. C., and Malamud, B. D. (2014). Reviewing and visualizing the interactions of natural hazards, Rev. Geophys., 52, 680–722, doi.org/10.1002/2013RG000445.

Gill, J. C., Malamud, B. D., Barillas, E. M., and Guerra Noriega, A. (2020). Construction of regional multi- hazard interaction frameworks, with an application to Guatemala, Nat. Hazards Earth Syst. Sci., 20, 149– 180, https://doi.org/10.5194/nhess-20-149-2020.

Gill, J. C., Taylor, F. E., Duncan, M. J., Mohadjer, S., Budimir, M., Mdala, H., and Bukachi, V. (2021). Invited perspectives: Building sustainable and resilient communities – recommended actions for natural hazard scientists, Nat. Hazards Earth Syst. Sci., 21, 187–202, https://doi.org/10.5194/nhess-21-187-2021, 2021.

Gill, J., Duncan, M., Ciurean, R., Smale, L., Stuparu, D., Schlumberger, J. (2022). MYRIAD-EU D1.2 Handbook of Multi-hazard, Multi-Risk Definitions and Concepts, H2020 MYRIAD-EU Project, grant agreement No. 101003276, pp75, Retrived from: https://nora.nerc.ac.uk/id/eprint/533237/ (last access: 24 April, 2024).

Goda, K., and De Risi, R. (2023). Future perspectives of earthquake-tsunami catastrophe modelling: From single-hazards to cascading and compounding multi-hazards. *Frontiers in Built Environment*, *8*, 1022736, doi.org/10.3389/fbuil.2022.1022736.

Greiving, Y. J., and Nam, B. H. (2017). Sinkhole Hazard Mapping Using Frequency Ratio and Logistic Regression Models for Central Florida. Proc., Geo-Risk 2017, American Society of Civil Engineers, Reston, VA, 246-256.

Gutiérrez, F. (2016). Sinkhole hazards. In Oxford Research Encyclopedia of Natural Hazard Science. doi.org/10.1093/acrefore/9780199389407.013.40.

Gutiérrez, F., Guerrero, J., and Lucha, P. (2008). A genetic classification of sinkholes illustrated from evaporite paleokarst exposures in Spain. *Environmental Geology*, *53*(5), 993–1006. doi.org/10.1007/s00254-007-0727-5.

Gutiérrez, F., Parise, M., De Waele, J., and Jourde, H. (2014). A review on natural and human-induced geohazards and impacts in karst. *Earth-Science Reviews*, *138*, 61-88, doi.org/10.1016/j.earscirev.2014.08.002.

Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P. (2007). Rainfall thresholds for the initiation of landslides in central and southern Europe. Meteorog. Atmos. Phys. 98, –267, doi.org/10.1007/s00703-007-0262-7.

Gutierrez A, Binet S (2010) La Loire souterraine, circulations karstiques dans le Val d'Orléans. Géosciences. Ed. BRGM.

Hachcham., A (2023, August 11). *The KNN Algorithm - Explanation, Opportunities, Limitations*. Retrieved from https://neptune.ai/blog/knn-algorithm-explanation-opportunities-limitations.

Heidari, M., Khanlari, G. R., Beydokhti, A. T., and Momeni, A. A. (2011). The formation of cover collapse sinkholes in North of Hamedan, Iran. *Geomorphology*, *132*(3-4), 76-86, doi.org/10.1016/j.geomorph.2011.04.025.

Herries, J. (June 26, 2018). Better Breaks Define Your Thematic Map's Purpose. Retrieved form: https://www.esri.com/arcgis-blog/products/arcgis-online/mapping/better-breaks-define-your-thematic-maps-purpose/?rsource=https%3A%2F%2Flinks.esri.com%2Fbetter-breaks.

Hewitt K, Burton I. (1971). Hazardousness of a place: a regional ecology of damaging events. Toronto Press, Toronto, Retrieved from: www.ilankelman.org/miscellany/hewittetal.1971.pdf.

Hochrainer-Stigler, S., Trogrlić, R. Š., Reiter, K., Ward, P. J., de Ruiter, M. C., Duncan, M. J., and Gottardo, S. (2023). Toward a framework for systemic multi-hazard and multi-risk assessment and management. *IScience*, *26*(5). http://creativecommons.org/licenses/by/4.0/.

Hurst, T., and Smith, W. (2004). A Monte Carlo methodology for modelling ashfall hazards. Journal of volcanology and geothermal research, 138(3-4), 393-403, doi:10.1016/j.jvolgeores.2004.08.001.

Hyatt, J. A., and Jacobs, P. M. (1996). Distribution and morphology of sinkholes triggered by flooding following Tropical Storm Alberto at Albany, Georgia, USA, Geomorphology, 17, 305–316, doi.org/10.1016/0169-555X(96)00014-1.

Intrieri, E., Confuorto, P., Bianchini, S., Rivolta, C., Leva, D., Gregolon, S., and Fanti, R. (2023). Sinkhole risk mapping and early warning: the case of Camaiore (Italy). Frontiers in Earth Science, 11, 1172727, doi.org/10.3389/feart.2023.1172727.

Intrieri, E., Fontanelli, K., Bardi, F., Marini, F., Carlà, T., Pazzi, V., and Fanti, R. (2018). Definition of sinkhole triggers and susceptibility based on hydrogeomorphological analyses. Environmental earth sciences, 77(1), 4. doi.org/10.1007/s12665-017-7179-3.

Intrieri, E., Gigli, G., Nocentini, M., Lombardi, L., Mugnai, F., Fidolini, F., and Casagli, N. (2015). Sinkhole monitoring and early warning: An experimental and successful GB-InSAR application. Geomorphology, 241, 304-314.

IPCC. (2007). Appendix I - Glossary. In IPCC (2007) AR4 Climate Change 2007: Impacts, Adaptation, and Vulnerability. Available at: https://www.ipcc.ch/site/assets/uploads/2018/03/ar4_wg2_full_report.pdf (last accessed 16 March 2022).

Javidan, N., Kavian, A., Conoscenti, C., Jafarian, Z., Kalehhouei, M., and Javidan, R. (2024). Development of risk maps for flood, landslide, and soil erosion using machine learning model. Natural Hazards, 1-24. doi.org/10.1007/s11069-024-06670-6.

Javidan, N., Kavian, A., Pourghasemi, H. R., Conoscenti, C., Jafarian, Z., and Rodrigo-Comino, J. (2021). Evaluation of multi-hazard map produced using MaxEnt machine learning technique. Scientific Reports, 11(1), 1-20. https://www.nature.com/articles/s41598-021-85862-7.

Jayasekara, R., Siriwardana, C., Amaratunga, D., and Haigh, R. (2022). Evaluating the network of stakeholders in Multi-Hazard Early Warning Systems for multiple hazards amidst biological outbreaks: Sri Lanka as a case in point. Progress in disaster science, 14, 100228, doi.org/10.1016/j.pdisas.2022.100228.

Jaynes, E. T. (1982). On the rationale of maximum-entropy methods. Proceedings of the IEEE, 70(9), 939-952.

Jaynes, E.T. (1957). Information theory and statistical mechanics, Phys. Rev., 106, 620–630, https://doi.org/10.1103/PhysRev.106.620.

Jiang, X., Lei, M., Li, Y., and Dai, J. (2005). National-Scale Risk Assessment of sinkhole Hazard in China (pp. 649–658). American Society of Civil Engineers (ASCE), doi.org/10.1061/40796(177)69.

Joodi, A. (2009). Apport de l'équation de Brinkman à la modélisation de l'écoulement d'eau et du transport de soluté dans l'aquifère karstique: application au système karstique du Val d'Orléans. Hydrologie. Université d'Orléans, Français. tel-00417805.

Jozja, N., Maget, P., Défarge, C., Mouget, C., Pidon, A., et al. (2010). Apport des traçages à la connaissance du système karstique du Val d'Orléans. Historique et nouvelles avancées. Géologues, pp.1-5.

Kappes, M. S., Keiler, M and. Glade, T. (2010). From single- to multi-hazard risk analyses: A concept addressing emerging challenges, in Mountain Risks: Bringing Science to Society, edited by J. P. Malet, T. Glade, and N. Casagli, pp. 351–356, CERG Editions, Strasbourg, France.

Kappes, M.S., Keiler, M., von Elverfeldt, K. and Glad, T. (2012). Challenges of analyzing multi-hazard risk: a review. Nat Hazards 64, 1925–1958, doi.org/10.1007/s11069-012-0294-2.

Kariminejad, N., Pourghasemi, H. R., Maleki, S., and Hosseinalizadeh, M. (2022). Digital soil mapping and modeling in Loess-derived soils of Iranian Loess Plateau. *Geocarto International*, *37*(26), 11633-11651, doi.org/10.1080/10106049.2022.2060317.

Kaufmann, G. (2014). Geophysical mapping of solution and collapse sinkholes: Journal of Applied Geophysics, v. 111, p. 271−288, doi.org/10.1016/j.jappgeo.2014.10.011.

Kemmerly, P. (1981). The need for recognition and implementation of a sinkhole-floodplain hazard designation in urban karst terrains. Environmental Geology, 3(5), 281-292, doi.org/10.1007/BF02473519.

Kerekes, A. H., Poszet, S. L., and Andrea, G. Á. L. (2018). Landslide susceptibility assessment using the maximum entropy model in a sector of the Cluj–Napoca Municipality, Romania. Revista de Geomorfologie, 20(1), 130-146.

Khanlari, G., Heidari, M., Momeni, A. A., Ahmadi, M., and Beydokhti, A. T. (2012). The effect of groundwater overexploitation on land subsidence and sinkhole occurrences, western Iran. Quarterly Journal of Engineering Geology and Hydrogeology, 45(4), 447-456, doi.org/10.1144/qjegh2010-069.

Kim, K., Kim, J., Kwak, T.-Y., and Chung, C.-K. (2018). Logistic regression model for sinkhole susceptibility due to damaged sewer pipes, Nat. Hazards, 93, 765–785, doi.org/10.1007/s11069-018-3323-y.

Kim, Y. J., and Nam, B. H. (2017). Sinkhole Hazard Mapping Using Frequency Ratio and Logistic Regression Models for Central Florida (pp. 246–256). American Society of Civil Engineers (ASCE). https://doi.org/10.1061/9780784480717.023.

Kim, Y. (2019). Probabilistic Spatio-Magnitude Sinkhole Hazard Analysis for East Central Flrodia (Doctoral disseration), University of Central Florida, available at: https://stars.library.ucf.edu/etd/6878/.

Kim, Y. J., Nam, B. H., Jung, Y. H., Liu, X., Choi, S., Kim, D., and Kim, S. (2022). Probabilistic spatial susceptibility modeling of carbonate karst sinkhole. Engineering Geology, 106728, doi.org/10.1016/j.enggeo.2022.106728.

Klopmeier, N. W., Pesi, S. M., Morris, G., and Conner, M. (2018). Sinkholes as a Source of Wildlife Mortality. Southeastern Naturalist, 17(4).

Komendantova, N., Mrzyglocki, R., Mignan, A., Khazai, B., Wenzel, F., Patt, A., and Fleming, K. (2014). Multi-hazard and multi-risk decision-support tools as a part of participatory risk governance: Feedback from civil protection stakeholders. International Journal of disaster risk reduction, 8, 50-67, doi.org/10.1016/j.ijdrr.2013.12.006.

Kornejady, A., Ownegh, M., and Bahremand, A. (2017). Landslide susceptibility assessment using maximum entropy model with two different data sampling methods. Catena, 152, 144-162, doi.org/10.1016/j.catena.2017.01.010.

Kovacic, G., and Ravbar, N. (2010). Extreme hydrological events in karst areas of Slovenia: The case of the Unica River basin. *Geodinamica Acta, 23*(1-3), 89-100. doi.org/10.3166/ga.23.89-100.

Krige, D. G. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. Journal of the Chemical, Metallurgical and Mining Society of South Africa, 52(6), 119-139. https://hdl.handle.net/10520/AJA0038223X_4792.

Kuniansky, E. L., Weary, D. J., and Kaufmann, J. E. (2016). The current status of mapping karst areas and availability of public sinkhole-risk resources in karst terrains of the United States. *Hydrogeology Journal*, *24*(3), 613, doi.org/10.1007/s10040-015-1333-3.

Kwak, T. Y., Woo, S. I., Chung, C. K., and Kim, J. (2020). Experimental assessment of the relationship between rainfall intensity and sinkholes caused by damaged sewer pipes. *Natural Hazards and Earth System Sciences*, *20*(12), 3343-3359, doi.org/10.5194/nhess-20-3343-2020.

Le Préfet, Coordonnateur du Plan Loire Grandeur Nature. (1996). Histoire des protections contre les crues. DREAL Centre-Val de Loire, Cartophiles du Loiret, Musée de la Marine de Châteauneuf, AFP, Météo France Conception graphique. Crues de Loire. Imprimé sur du papier recyclé. Troisième édition.

Lei, M., Gao, Y., Jiang, X., and Guan, Z. (2013). Emergency investigation of extremely large sinkholes, Maohe, Guangxi, China. In *Proceedings of the Thirteenth Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst* (pp. 293–297), doi.org/10.5038/9780979542275.1139.

Lei, M., Gao, Y., Jiang, X., and Guan, Z. (2016). Mechanism analysis of sinkhole formation at Maohe village, Liuzhou city, Guangxi province, China. *Environmental Earth Sciences*, *75*(1), 1–11, doi.org/10.1007/s12665-015-5100-5.

Lepiller, M.: Hydrogéologie du Val d'Orléans. In les aquifères de France. ed. BRGM— AIH 2006, 956 p, 2006.

LepilerM., LeclercB, AlbéricP., Boismoreau P.(2000).Cartedescirculations karstiques du val d'Orléans.

Lewin, J., and Woodward, J. C. (2009). Karst geomorphology and environmental change. The physical geography of the mediterranean, 287-317.

Li, J. and Heap, A.D. (2008). A Review of Spatial Interpolation Methods for Environmental Scientists. Geoscience Australia, Record 2008/23, 137 pp.

Li, J., and Heap, A. D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. Environmental Modelling & Software, 53, 173-189.

Lian, J. J., Xu, K., and Ma, C. (2013). Joint impact of rainfall and tidal level on flood risk in a coastal city with a complex river network: a case study of Fuzhou City, China. Hydrology and Earth System Sciences, 17(2), 679-689. doi:10.5194/hess-17-679-2013.

Linardos, V., Drakaki, M., Tzionas, P., and Karnavas, Y. L. (2022). Machine learning in disaster management: recent developments in methods and applications. *Machine Learning and Knowledge Extraction*, *4*(2), doi.org/10.3390/ make4020020. doi.org/10.1016/j.envsoft.2013.12.008.

Linares, R., Roqué, C., Gutiérrez, F., Zarroca, M., Carbonel, D., Bach, J., and Fabregat, I. (2017). The impact of droughts and climate change on sinkhole occurrence. A case study from the evaporite karst of the Fluvia Valley, NE Spain. Science of the total environment, 579, 345-358, doi.org/10.1016/j.scitotenv.2016.11.091.

Lindell, M. K. (2020). Improving hazard map comprehension for protective action decision making. Frontiers in Computer Science, 2, 27, doi.org/10.3389/fcomp.2020.0002.

Lindsey, B. D., Katz, B. G., Berndt, M. P., Ardis, A. F., and Skach, K. A. (2010). Relations between sinkhole density and anthropogenic contaminants in selected carbonate aquifers in the eastern United States. Environmental Earth Sciences, 60, 1073-1090.

Liu B, Siu Y. L, and Mitchel G. (2016). Hazard interaction analysis for multi-hazard risk assessment: a systematic classification based on hazard-forming environment. *Natural Hazards and Earth System Sciences*, *16*(2), 629-642, doi:10.5194/nhess-16-629-2016.

Liu, R., Zhang, Z., Shen, J., and Wang, Z. (2019). Bryophyte diversity in karst sinkholes affected by different degrees of human disturbance. Acta Soc. Bot. Pol. 88 (2), 3620, doi:10.5586/asbp.3620.

Liu, Y., Zhao, L., Bao, A., Li, J., and Yan, X. (2022). Chinese high resolution satellite data and GIS-based assessment of landslide susceptibility along Highway G30 in Guozigou Valley using logistic regression and MaxEnt model. Remote Sensing, 14(15), 3620, doi.org/10.3390/rs14153620.

Lombardo, L., Bachofer, F., Cama, M., Märker, M., and Rotigliano, E. (2016). Exploiting Maximum Entropy method and ASTER data for assessing debris flow and debris slide susceptibility for the Giampilieri catchment (north-eastern Sicily, Italy). Earth Surface Processes and Landforms, 41(12), 1776-1789.

Lombardo, L., Tanyas, H., and Nicu, I. C. (2020). Spatial modeling of multi-hazard threat to cultural heritage sites. *Engineering Geol*ogy, 277, 105776, doi.org/10.1016/j.enggeo.2020.105776.

López-Saavedra, M., and Martí, J. (2023). Reviewing the multi-hazard concept. Application to volcanic islands. Earth-Science Reviews, 104286, doi.org/10.1016/j.earscirev.2022.104286.

Luu, L. H., Noury, G., Benseghier, Z., and Philippe, P. (2019). Hydro-mechanical modeling of sinkhole occurrence processes in covered karst terrains during a flood. Engineering Geology, 260, 105249, doi.org/10.1016/j.enggeo.2019.105249.

Magdalene, S., and Alexander, E. C. (2018). Sinkhole distribution in Winona County, Minnesota revisited. In *Karst geohazards* (pp. 43-51). Routledge.

Martin, J.C., Amraoui, N. and Pathirana, N. (Janvier 2003). Caractérisation du risque d'inondation par remontée de nappe sur le val d'Orléans (No. BRGM/RP-52121-FR).

Marignani, M., Bruschi, D., Garcia, D.A., Frondoni, R., Carli, E., Pinna, M.S., Cumo, F., Gugliermetti, F., Saatkamp, A., and Doxa, A. (2017). Identification and prioritization of areas with high environmental risk in

Mediterranean coastal areas: A flexible approach, Sci. Total Environ., 590, 566–578, https://doi.org/10.1016/j.scitotenv.2017.02.221.

Martinotti, M. E., Pisano, L., Marchesini, I., Rossi, M., Peruccacci, S., Brunetti, M. T., Melillo, M., Amoruso, G., Loiacono, P., Vennari, C., Vessia, G., Trabace, M., Parise, M., and Guzzetti, F. (2017). Landslides, floods and sinkholes in a karst environment: The 1–6 September 2014 Gargano event, southern Italy. *Natural Hazards and Earth System Sciences, 17*(3), 467-480, doi.org/10.5194/nhess-17-467-2017.

Mărgărint, M. C., Niculiţă, M., Roder, G., & Tarolli, P. (2021). Risk perception of local stakeholders on natural hazards: implications for theory and practice. Natural Hazards and Earth System Sciences Discussions, 2021, 1-45.

Marzocchi, W., Garcia-Aristizabal, A., Gasparini, P., Mastellone, M. L., and Di Ruocco, A. (2012). Basic principles of multi-risk assessment: a case study in Italy. Natural hazards, 62, 551-573, doi.org/10.1007/s11069-012-0092-x.

Merow, C., Smith, M. J., and Silander Jr, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography, 36(10), 1058-1069. doi: 10.1111/j.1600-0587.2013.07872.x.

Mobley, W., Sebastian, A., Highfield, W., and Brody, S. D. (2019). Estimating flood extent during Hurricane Harvey using maximum entropy to build a hazard distribution model. Journal of Flood Risk Management, 12, e12549.

Mokhtari, M., and Abedian, S. (2019). Spatial prediction of landslide susceptibility in Taleghan basin, Iran. Stochastic Environmental Research and Risk Assessment, 33(7), 1297-1325.

Mignan, A., S. Wiemer, and D. Giardini (2014), The quantification of low-probability–high-consequences events: Part I. A generic multi-risk approach, Nat. Hazards, 73, 1999–2022.

Munich Re. (2023). NatCatService. Retrieved from https://www.munichre.com/en/risks/natural-disasters.html

Nachappa, T. G., Ghorbanzadeh, O., Gholamnia, K., and Blaschke, T. (2020). Multi-hazard exposure mapping using machine learning for the State of Salzburg, Austria. Remote Sensing, 12(17), 2757.

Naughton, O., McCormack, T., Gill, L., and Johnston, P. (2018). Groundwater flood hazards and mechanisms in lowland karst terrains. *Geological Society, London, Special Publications*, *466*(1), 397-410.

Neri, M., et al. (2012). A semi-quantitative method based on event tree: Modeling multi-hazard in areas with data scarcity for Kan

Neri, M., Le Cozannet, G., Thierry, P., Bignami, C., and Ruch, J. (2013). A method for multi-hazard mapping in poorly known volcanic areas: an example from Kanlaon (Philippines), Nat. Hazards Earth Syst. Sci., 13, 1929–1943, doi.org/10.5194/nhess-13-1929-2013.

Nhu, V. H., Zandi, D., Shahabi, H., Chapi, K., Shirzadi, A., Al-Ansari, N and Nguyen, H. (2020). Comparison of support vector machine, Bayesian logistic regression, and alternating decision tree algorithms for shallow landslide susceptibility mapping along a mountainous road in the west of Iran. *Applied Sciences*, *10*(15), 5047.

Nilson, N. J. (2005). Introduction to machine learning. Stanford University, USA, 179. Retrieved from: https://bafflerbach.github.io/test_software_carpentry/files/Module%201%20Lecture%20Notes.pdf

Nof, R. N., Abelson, M., Raz, E., Magen, Y., Atzori, S., Salvi, S., and Baer, G. (2019). SAR interferometry for sinkhole early warning and susceptibility assessment along the Dead Sea, Israel. *Remote Sensing*, *11*(1), 89.

Norallahi, M., Seyed Kaboli, H. (2021). Urban flood hazard mapping using machine learning models: GARP, RF, MaxEnt and NB. Nat Hazards 106, 119–137, doi.org/10.1007/s11069-020-04453-3.

Noury, G. (2020). *Rapport d'expertise: Saint-Jean-le-Blanc (45) – Risque mouvements de terrain au niveau du « gouffre de Montission »*. BRGM/RP-70260-FR. Orléans Métropole, Direction de l'Environnement et de la Prévention des Risques, Service Prévention des Risques Majeurs.

Noury, G., Perrin, J., Luu, L. H., Philippe, P., and Gourdier, S. (2018). Role of floods on sinkholes occurrence in covered karst terrains: Case study of Orléans area (France) during the 2016 meteorological event and perspectives for other karst environments. In The 15th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst, Shepherdstown, West Virginia, United States.

Office of the United Nations Disaster Relief Coordinator (UNDRO) (1980). Natural disasters and vulnerability analysis: report of Expert Group Meeting, 9-12 July 1979. Available at: https://digitallibrary.un.org/record/95986?ln=en&v=pdf.

Oh, H. J., and Lee, S. (2010). Assessment of ground subsidence using GIS and the weights-of-evidence model. *Engineering Geology*, *115*(1-2), 36-48, doi.org/10.1016/j.enggeo.2010.06.015.

Ozdemir, A. (2015). Investigation of sinkholes spatial distribution using the weights of evidence method and GIS in the vicinity of Karapinar (Konya, Turkey). *Geomorphology*, *245*, 40-50.

Padilla, L., Dryhurst, S., Hosseinpour, H., and Kruczkiewicz, A. (2021). Multiple Hazard Uncertainty Visualization Challenges and Paths Forward. *Frontiers in Psychology*, 1993, doi.org/10.3389/fpsyg.2021.579207.

Pandey, V. K., Pourghasemi, H. R., and Sharma, M. C. (2018). Landslide susceptibility mapping using maximum entropy and support vector machine models along the Highway Corridor, Garhwal Himalaya. Geocarto International, doi.org/10.1080/10106049.2018.1510038.

Panno, S. V., Wiebel, C. P., Heigold, P. C., and Reed, P. C. (1994). Formation of regolith-collapse sinkholes in southern Illinois: Interpretation and identification of associated buried cavities. Environmental Geology, 23(3), 214-220. doi.org/10.1007/BF00771791.

Paradinas I, Illian J, Smout S. (2023). Understanding spatial effects in species distribution models. PLoS ONE 18(5): e0285463, doi.org/10.1371/journal.pone.0285463.

Parise, M., and Vennari, C. (2013). A chronological catalogue of sinkholes in Italy: the first step toward a real evaluation of the sinkhole hazard.

Parise, M. (2015a). A procedure for evaluating the susceptibility to natural and anthropogenic sinkholes: Georisk, v. 9, no. 4, p. 272−285, doi.org/10.1080/17499518.2015.1045002.

Parise, M., Gabrovsek, F., Kaufmann, G., and Ravbar, N. (2018a). Recent advances in karst research: from theory to fieldwork and applications. *Geological Society, London, Special Publications*, *466*(1), 1-24.

Parise, M., Pisano, L., and Vennari, C. (2018b). Sinkhole clusters after heavy rainstorms, J. Cave Karst Stud., 80, doi.org/10.4311/2017ES0105.

Park, N. W. (2015). Using maximum entropy modeling for landslide susceptibility mapping with multiple geoenvironmental data sets. Environmental Earth Sciences, 73(3), 937-949. https://link.springer.com/article/10.1007/s12665-014-3442-z.

Pazzi, V., Ceccatelli, M., Gracchi, T., Masi, E. B., and Fanti, R. (2018). Assessing subsoil void hazards along a road system using H/V measurements, ERTs and IPTs to support local decision makers. *Near Surface Geophysics*, *16*(3), 282-297.

Perrin J., Noury G., Cartannaz C., Gutierrez T., Van Laere A. (2014). ALEA-Karst : vers une approche multicritères de l'aléa lié aux cavités karstiques en contexte calcaire. Volume 2: étude de sites - Rapport BRGM/RP-63771-FR, 119 pages.

Perrin, J., Cartannaz, C., Noury, G., and Vanoudheusden, E. (2015). A multicriteria approach to karst subsidence hazard mapping supported by weights-of-evidence analysis. Engineering Geology, 197, 296-305.

Perrin, J., Pasquier, S., Gutierrez, A., Salquebre, D., and Vandoudheusden, E. (2016). Investigating physical processes leading to sinkhole occurrence in Val d'Orléans (France). EuroKarst 2016, Neuchatel, Switzerland. ffhal-01355612f.

Petersen, A., and Vondracek, B. (2006). Water quality in relation to vegetative buffers around sinkholes in karst terrain. Journal of soil and water conservation, 61(6), 380-390, Available at: https://www.jswconline.org/content/61/6/380.

Pescaroli G, Alexander D. Understanding compound, interconnected, interacting, and cascading risks: a holistic framework. Risk Anal 2018;38(11):2245–57.

Phillips, S. J. (2005). A brief tutorial on MaxEnt. ATandT Research, 190(4), 231-259.

Phillips, S.J., Anderson, R.P. and Schapire, R.E. (2006). Maximum entropy modeling of species geographic distributions. Ecological Modelling, 190, 231–259, https://doi.org/10.1016/j.ecolmodel.2005.03.026.

Phillips, S. J., and Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography, 31(2), 161-175.

Phillips, S. J., Dudík, M., and Schapire, R. E. (2004, July). A maximum entropy approach to species distribution modeling. In Proceedings of the Twenty-First International Conference on Machine Learning (p. 83).

Piao, Y., Lee, D., Park, S., Kim, H. G., and Jin, Y. (2022). Multi-hazard mapping of droughts and forest fires using a multi-layer hazards approach with machine learning algorithms. *Geomatics, Natural Hazards and Risk*, *13*(1), 2649-2673.

Plans de prévention des risques d'inondation du Val d'Orléans. (2015). Approuvé par arrêté préfectoral du 20 janvier. Available at: https://www.loiret.gouv.fr/contenu/telechargement/39868/285153/file/1-%20Note-pr%C3%A9sentation.pdf.

Pourghasemi, H. R., Gayen, A., Panahi, M., Rezaie, F., and Blaschke, T. (2019). Multi-hazard probability assessment and mapping in Iran. *Science of the total environment*, *692*, 556-571, doi.org/10.1016/j.scitotenv.2019.07.203.

Pournader, M., Ahmadi, H., Feiznia, S., Karimi, H., and Peirovan, H. R. (2018). Spatial prediction of soil erosion susceptibility: An evaluation of the maximum entropy model. Earth Science Informatics, 11(3), 389-401. https://link.springer.com/article/10.1007%2Fs12145-018-0338-6.

Powers, D.M.W. (2011). Evaluation: from Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation, Journal of Machine Learning Technologies, 2(1), 37–63.

Qiu, X., Wu, S. S., and Chen, Y. (2020). Sinkhole susceptibility assessment based on morphological, imagery, and contextual attributes derived from gis and imagery data. *Journal of Cave and Karst Studies*, *82*(1), 1-17.

Radosavljevic, A., and Anderson, R. P. (2014). Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography*, 41(4), 629-643. http://wileyonlinelibrary.com/journal/jbi doi:10.1111/jbi.12227.

Rahmati, O., Pourghasemi, H. R., and Melesse, A. M. (2016). Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: A case study at Mehran Region, Iran. Catena, 137, 360-372.

Raschka, S. (2018). Model evaluation, model selection, and algorithm selection in machine learning. arXiv preprint arXiv:1811.12808.

Rateb, A., and Hermas, E. (2020). The 2018 long rainy season in Kenya: Hydrological changes and correlated land subsidence. Remote Sensing, 12(9), 1390, doi.org/10.3390/rs12091390.

Rehman, A., Song, J., Haq, F., Mahmood, S., Ahamad, M. I., Basharat, M., and Mehmood, M. S. (2022). Multi-Hazard Susceptibility Assessment Using the Analytical Hierarchy Process and Frequency Ratio Techniques in the Northwest Himalayas, Pakistan. Remote Sensing, 14(3), 554.

Rosdi, M. A. H. M., Othman, A. N., Zubir, M. A. M., Latif, Z. A., and Yusoff, Z. M. (2017). Sinkhole susceptibility hazard zones using GIS and analytical hierarchical process (AHP): A case study of Kuala Lumpur and Ampang Jaya. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 145-151.

Rougier, J., and Beven, K. J. (2013). Model and data limitations: the sources and implications of epistemic uncertainty. Risk and uncertainty assessment for natural hazards, 40.

Rusk, J., Maharjan, A., Tiwari, P., Chen, T. H. K., Shneiderman, S., Turin, M., and Seto, K. C. (2022). Multi-hazard susceptibility and exposure assessment of the Hindu Kush Himalaya. Science of the Total Environment, 804, 150039, doi.org/10.1016/j.scitotenv.2021.150039.

Saaty, T. L. (2004). Decision making—the analytic hierarchy and network processes (AHP/ANP). *Journal of systems science and systems engineering*, *13*, 1-35.

Saaty, T.L. (1988). Multicriteria Decision Making: The Analytic Hierarchy Process. InPlanning, Priority Setting, Resource Allocation;McGraw-Hill: New York, NY, USA,.

Saaty, TL. (2008). Relative Measurement and Its Generalization in Decision Making Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors The Analytic Hierarchy/Network Process, RACSAM, VOL. 102 (2), pp. 251–318. https://link.springer.com/article/10.1007/BF03191825.

Samyn, K., Mathieu, F., Bitri, A., Nachbaur, A., and Closset, L. (2014). Integrated geophysical approach in assessing karst presence and sinkhole susceptibility along flood-protection dykes of the Loire River, Orléans, France. Engineering geology, 183, 170-184, doi.org/10.1016/j.enggeo.2014.10.013.

Santo, A., Ascione, A., Del Prete, S., Di Crescenzo, G., and Santangelo, N. (2011). Collapse sinkholes in the carbonate massifs of Central and Southern Apennines. Acta Carsologica, 40(1), doi.org/10.3986/ac.v40i1.31.

Sardooi, E. R., Azareh, A., Mesbahzadeh, T., Sardoo, F. S., Parteli, E. J., and Pradhan, B. (2021). A hybrid model using data mining and multi-criteria decision-making methods for landslide risk mapping at Golestan Province, Iran. Environmental Earth Sciences. https://link.springer.com/article/10.1007/s12665-021-09788-z.

Scolobig, A., Komendantova, N., and Mignan, A. (2017). Mainstreaming multi-risk approaches into policy. *Geosciences*, *7*(4), 129, doi.org/10.3390/geosciences7040129.

Serrhini, K., Moy, Y., Noury, G., Serrano, J. and Soltanpour, H. (2023). Vulnérabilité non-structurelle et résilience territoriale en situation de multi-risque. Perception et préférences cartographiques des aléas effondrement et inondation (val d'Orléans, Loiret, France). Canadian Journal of Regional Science / Revue canadienne des sciences régionales, 46(2), 73–84, doi.org/10.7202/1100216ar.

Siahkamari, S., Haghizadeh, A., Zeinivand, H., Tahmasebipour, N., and Rahmati, O. (2018). Spatial prediction of flood-susceptible areas using frequency ratio and maximum entropy models. Geocarto International, 33(9), 927-94.

Sillero, N. and Barbosa, A. M.(2021). Common mistakes in ecological niche models, Int. J. Geogr. Inf. Sci., 35, 213-226, doi.org/10.1080/13658816.2020.1798968.

Sillero, N., Arenas-Castro, S., Enriquez-Urzelai, U., Gomes Vale, C., Sousa-Guedes, D., Martínez-Freiría, F., Real, R., and Barbosa, A. M. (2021). Want to model a species niche? A step-by-step guideline on correlative ecological niche modelling, Ecol. Model., 456, 109671, https://doi.org/10.1016/j.ecolmodel.2021.109671.

Shi, J. (2022). Engineering Application of Landslide Susceptibility Mapping in Linear Engineering Optimization. Front. Earth Sci. 10:861057. https://doi.org/ 10.3389/feart.2022.861057.

Skilodimou, H. D., Bathrellos, G. D., Chousianitis, K., Youssef, A. M., and Pradhan, B. (2019). Multi-hazard assessment modeling via multi-criteria analysis and GIS: a case study. *Environmental Earth Sciences*, *78*(2), 1-21.

Soltanpour, H. (2024, May 20). Evolving multi-hazard paradigms in a nutshell. Natural Hazards Blog. European Geosciences Union. Available at: https://blogs.egu.eu/divisions/nh/2024/05/20/evolving-multi-hazard-paradigms-in-a-nutshell/.

Smith, R. L., and Shashkina, O. (2023). A way forward to integrating cascading risk into local resilience strategies in Central Asia. Progress in disaster science, 17, 100273. doi.org/10.1016/j.pdisas.2022.100273.

Steven, J. P. (2017). A brief tutorial on MaxEnt. Available from: http://biodiversityinformatics.amnh.org/open_source/MaxEnt/. Accessed on 2022-7-21.

Stocks, L. (2007). Image-object analysis for sinkhole inventory, in Papers and Proceedings of Applied Geography Conferences, v. 30, p. 447.

Strzałkowski P, Tomiczek K. (2015). Proposal of a methodology assessing the risk of sink holes formation in mining areas. Int J Min Sci Technol 25(1):85–89

Subedi, P., Subedi, K., Thapa, B. et al (2019). Sinkhole susceptibility mapping in Marion County, Florida: Evaluation and comparison between analytical hierarchy process and logistic regression-based approaches. Sci Rep 9, 7140, doi.org/10.1038/s41598-019-437056.

Sullivan, E. (2022). Understanding from machine learning models. The British Journal for the Philosophy of Science. 73 (1):109-133. Retrieved from:

https://philarchive.org/go.pl?id=SULUFMandproxyId=andu=https%3A%2F%2Fphilpapers.org%2Farchive%2FSULUFM.pdf.

Taheri K, Gutiérrez F, Mohseni H, Raeisi E, Taheri M. (2015). Sinkhole susceptibility mapping using the analytical hierarchy process (AHP) and magnitude–frequency relationships: A case study in Hamadan province, Iran. Geomorphology. 2015; 234: 64–79, doi: 10.1016/j.geomorph.2015.01.005.

Taheri, K., Shahabi, H., Chapi, K., Shirzadi, A., Gutiérrez, F., and Khosravi, K. (2019). Sinkhole susceptibility mapping: A comparison between Bayes-based machine learning algorithms. *Land Degradation and Development*, *30*(7), 730-745.

Tarvainen, T., Jarva, J., Greiving, S. (2006). Spatial pattern of hazards and hazard interactions in Europe. In: Schmidt-Thomé, P. (Ed.), Natural and Technological Hazards and Risks Affecting the Spatial Development of European Regions. Geological Survey of Finland Special Paper 6, vol. 42. pp. 83–91.

Tehrani, F. S., Calvello, M., Liu, Z., et al. (2022). Machine learning and landslide studies: Recent advances and applications. *Natural Hazards*, *114*(3), 1197–1245, doi.org/10.1007/s11069-022-05423-7.

Terzi, S., Torresan, S., Schneiderbauer, S., Critto, A., Zebisch, M., and Marcomini, A. (2019). Multi-risk assessment in mountain regions: A review of modelling approaches for climate change adaptation. Journal of Environmental Management, 232, 759-771, doi:10.1016/j.jenvman.2018.11.100.

Thiery, Y., Terrier, M., Colas, B., Fressard, M., Maquaire, O., Grandjean, G., & Gourdier, S. (2020). Improvement of landslide hazard assessments for regulatory zoning in France: state–of–the-art perspectives and considerations. International journal of disaster risk reduction, 47, 101562. doi.org/10.1016/j.ijdrr.2020.101562.

Tilloy, A., Malamud, B. D., Winter, H., and Joly-Laugel, A. (2019). A review of quantification methodologies for multi-hazard interrelationships. *Earth-Science Reviews*, *196*, 102881. doi.org/10.1016/j.earscirev.2019.102881.

Tinti, S., Pagnoni, G., Piatanesi, A. (2003). Simulation of tsunamis induced by volcanic activity in the Gulf of Naples (Italy). Nat. Hazards Earth Syst. Sci. 3, 311–320, doi.org/10.5194/nhess-3-311-2003.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. Economic Geography 46: 234-40.

Todd, A., and Ivey-Burden, L. (2016). A Method of Mapping Sinkhole Susceptibility Using a Geographic Information System: A Case Study for Interstates in the Karst Counties of Virginia (pp. 299–306). University of South Florida Libraries. https://doi.org/10.5038/9780991000951.1048.

Touma, D., Stevenson, S., Swain, D. L., Singh, D., Kalashnikov, D. A., and Huang, X. (2022). Climate change increases risk of extreme rainfall following wildfire in the western United States. *Science advances*, *8*(13), eabm0320, doi.org/10.1126/sciadv.abm0320.

Trogrlić, R. Š., Reiter, K., Ciurean, R. L., Gottardo, S., Torresan, S., Daloz, A. S., and Ward, P. J. (2024). Challenges in assessing and managing multi-hazard risks: A European stakeholders perspective. Environmental Science and Policy, 157, 103774.

Trzałkowski P. (2017). The proposal of predicting formation of sinkholes with an exemplary application. J Min Sci 1:53–58.

Tsereteli, E., Gaprindashvili, G., Gaprindashvili, M., Bolashvili, N., and Gongadze, M.(2019). Hazard risk of 1185 debris/mud flow events in Georgia and methodological approaches for management, In IAEG/AEG Annual Meeting Proceedings, San Francisco, California, 2018-Volume 5 (pp. 153-160), doi.org/10.1007/978-3-319-93136-4_19.

UNEP. (1992). Agenda 21. Tech. rep., United Nations Environment Programme., https://sustainabledevelopment.un.org/outcomedocuments/agenda21 (last access: 12 April 2024).

UNICEF. (2024). Mozambique Humanitarian Situation Report 2023. Retrieved from: https://www.unicef.org/mozambique/en/reports/mozambique-humanitarian-situation-report-2023.

UNISDR. (2005). Hyogo Framework for Action 2005-2015: Building the Resilience of Nations and Communities to Disasters. United Nations, Hyogo, Japan.

UNISDR (2015). Sendai Framework for Disaster Risk Reduction 2015–2030. In UN Report UNISDR/GE/015; United Nations Office for Disaster Risk Reduction: Geneva, Switzerland.

United Nations (2002). Report of the World Summit on Sustainable Development. Johannesburg, South Africa, 26 August-4 September 2002 Available at: https://documents.un.org/doc/undoc/gen/n02/636/93/pdf/n0263693.pdf.

United Nations Office for Disaster Risk Reduction (UNISDR). UNISDR Terminology on Disaster Risk Reduction; UNISDR: Geneva, Switzerland, 2009. Available online: https://www.unisdr.org/files/7817_UNISDRTerminologyEnglish.pdf (accessed on 4 October 2021).

United Nations Office for Disaster Risk Reduction (UNISDR). (2017) Report of the open-ended intergovernmental expert working group on indicators and terminology relating to disaster risk reduction. Available at: https://www.preventionweb.net/files/50683_oiewgreportenglish.pdf. (last access: 14 Septemebr, 2024).

United Nations Office for Disaster Risk Reduction (UNISDR). (2024) Disaster risk reduction: UNISDR and ISC to review Hazard Information Profiles ahead of 2025 Global Platform. Available at: https://www.UNISDR.org/news/disaster-risk-reduction-UNISDR-and-isc-review-hazard-information-profiles-ahead-2025-global (las access: 14 Septemeber 2024).

van den Bout, B., Tang, C., van Westen, C., and Jetten, V. (2022). Physically based modeling of co-seismic landslide, debris flow, and flood cascade, Nat. Hazards Earth Syst. Sci., 22, 3183–3209, doi.org/10.5194/nhess-22-3183-2022.

van Den Eeckhaut, M., Poesen, J., Dusar, M., Martens, V., and Duchateau, P. (2007). Sinkhole formation above underground limestone quarries: A case study in South Limburg (Belgium). *Geomorphology*, *91*(1-2), 19-37.

van Westen, C. J., and Greiving, S. (2017). Multi-hazard risk assessment and decision making. Environmental hazards methodologies for risk assessment and management, 31.

van Westen, C.J., Kappes, M.S., Luna, B.Q., Frigerio, S., Glade, T., and Malet, J-P. (2014). Medium-scale multi-hazard risk assessment of gravitational processes, in: T. van Asch, J. Corominas, S. Greiving, J-P. Malet and S. Sterlacchini (eds.) Mountain risks: from prediction to management and governance, Springer, Netherlands, 201–231, doi:10.1007/978-94-007-6769-0_7.

Waltham, A.C. and Fookes, P.G. (2005). Engineering classification of karst ground conditions, Q. J. Eng. Geol. Hydrogeol., 3(1), 1–20, doi:10.1144/1470-9236/2002-33.

Waltham, T. Bell, F. G., Culshaw, M. G., Knez, M., and Slabe, T.((2005). Sinkholes and subsidence: karst and cavernous rocks in engineering and construction, (Vol. 382), Berlin: Springer,ISBN : 978-3-540-20725-2.

Wang, J., He, Z., and Weng, W. (2020). A review of the research into the relations between hazards in multi-hazard risk analysis. *Natural Hazards*, *104*(3), 2003-2026, doi.org/10.1007/s11069-020-04259-3.

Ward, P. J., Daniell, J., Duncan, M., Dunne, A., Hananel, C., Hochrainer-Stigler, S., and de Ruiter, M. (2022). Invited perspectives: A research agenda towards disaster risk management pathways in multi-(hazard-) risk assessment. *Natural Hazards and Earth System Sciences*, *22*(5), 1487–1497, doi.org/10.5194/nhess-22-1487-2022.

Ward, P.J., Couasnon, A., Eilander, D., Haigh, I.D., Hendry, A., Muis, S., Veldkamp, T.I.E., Winsemius, H.C., Wahl, T. (2018). Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries. Environ. Res. Lett. 13,084012, doi.org/10.1088/1748-9326/aad400.

White, E.L., White, W.B., (1979). Quantitative morphology of landforms in carbonate rock basins in the Appalachian Highlands. Bull. Geol. Soc. Am. 90, 385–396.

Williams, G. B., Banks, V. J., Bowman, E. T., Cooper, A. H., Jones, L. D., Kirkham, M. P., and Shepley, P. (2020). Integrating monitoring techniques for a buried sinkhole in an urban environment. doi.org/10.5038/9781733375313.1020.

Xiao, H., Kim, Y. J., Nam, B. H., and Wang, D. (2016). Investigation of the impacts of local-scale hydrogeologic conditions on sinkhole occurrence in East-Central Florida, USA. *Environmental Earth Sciences*, *75*, 1-16.

Xiao, H., Li, H., and Tang, Y. (2018). Assessing the effects of rainfall, groundwater downward leakage, and groundwater head differences on the development of cover-collapse and cover-suffosion sinkholes *in central Florida (USA). Science of the total environment, 644, 274-286.*

Xua J, Heb J, Zhang L. (2017). Collapse prediction of karst sinkhole via distributed Brillouin optical fiber sensor. Measurement 100:68–71.

Yang, Z., and Qiao, J. (2010). Regional landslide zonation based on entropy method in Three Gorges area, China. 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery (Vol. 3). IEEE.

Yousefi S, Pourghasemi HR, Emami SN, Pouyan S, Eskandari S, Tiefenbacher JP. (2020). A machine learning framework for multi-hazards modeling and mapping in a mountainous area. Sci Rep. 2020 Jul 22;10(1):12144, doi: 10.1038/s41598-020-69233-2. PMID: 32699313; PMCID: PMC7376103.

Zhang, T., Quevedo, R. P., Wang, H., Fu, Q., Luo, D., Wang, T,and Renno, C. D. (2022). Improved tree-based machine learning algorithms combining with bagging strategy for landslide susceptibility modeling. Arabian Journal of Geosciences, 15(2), 1-19, doi.org/10.1007/s12517-022-09488-3.

Zhou W, Beck BF, Adams AL .(2003). Sinkhole risk assessment along highway I-70 near Frederick, Maryland. In: Beck BF (ed) Sinkholes and the engineering and environmental impacts of karst. ASCE Geotechnical Special Publication No. 122:591–601.

Zhou, W. (2007). Drainage and flooding in karst terranes. *Environmental Geology*, *51*(6), 963-973, doi.org/10.1007/s00254-006-0365-3.

Zhou, W., and Beck, B. F. (2005). Roadway construction in karst areas: management of stormwater runoff and sinkhole risk assessment. *Environmental Geology*, *47*(8), 1138-1149, doi.org/10.1007/s00254-005-1252-z.

Zhou, W., Beck, B. F., and Adams, A. L. (2003). Sinkhole Risk Assessment along Highway I-70 near Frederick, Maryland (pp. 591–604). American Society of Civil Engineers (ASCE), doi.org/10.1061/40698(2003)54.

Zhu JunFeng, Z. J., and Pierskalla, W. J. (2016). Applying a weighted random forests method to extract karst sinkholes from LiDAR data, doi.org/10.1016/j.jhydrol.2015.12.012.

Zhu, J., Taylor, T. P., Currens, J. C., and Crawford, M. M. (2014). improved karst sinkhole mapping in kentucky using lidar techniques: a pilot study in floyds fork watershed. *Journal of Cave and Karst Studies*, *76*(3), doi.org/10.4311/2013ES0135.

Zhu, L., and Huang, J. F. (2006). GIS-based logistic regression method for landslide susceptibility mapping in regional scale. *Journal of Zhejiang University-Science A*, *7*(12), 2007-2017. ISSN 1009-3095 (Print); ISSN 1862-1775 (Online).

Zschau, J. (2017). Where are we with multihazards, multirisks assessment capacities?. in: Science for disaster risk management 2017: knowing better and losing less, edited by: Poljansek, K., Marin Ferrer, M., De Groeve, T., and Clark, I., European Union, Brussels, Belgium, https://drmkc.jrc.ec.europa.eu/knowledge/science-for-drm/science-for-disaster-risk-management-2017 (last access: 13 April 2024).

Zumpano, V., Pisano, L., and Parise, M. (2019). An integrated framework to identify and analyze karst sinkholes. Geomorphology, 332, 213-225, doi.org/10.1016/j.geomorph.2019.02.013.

Appendix 2.A

This appendix is dedicated to the 241 citations used in the systematic literature review for creating the violin plot of sinkhole-related studies.

References theme "Hazard/risk sinkhole assessments"

1. Ahmed M. (2013). Experimental and Numerical Modelling of Sinkhole Collapse. In Proceedings of the Transportation Research Board 92nd Annual Meeting (pp. 13–17). Washington, D.C.
2. Ali, H., and Choi, J. H. (2019). Risk prediction of sinkhole occurrence for different subsurface soil profiles due to leakage from underground sewer and water pipelines. Sustainability, 12(1), 310.
3. Al-Kouri, O., Al-Rawashdeh, S., Sadoun, B., Sadoun, B., and Pradhan, B. (2013). Geospatial modeling for sinkholes hazard map based on GIS and RS data. Journal of geographic information system, 2013.
4. Alrowaim Galve, J.P., Bonachea, J., Remondo, J., Gutiérrez, F., Guerrero, J., Lucha, P., Cendrero, A., Gutiérrez, M., and Sánchez, J.A., (2008), Development and validation of sinkhole susceptibility models in mantled karst settings: A case study from the Ebro valley evaporite karst (NE Spain): Engineering Geology, v. 99, no. 3−4, p. 185−197. https://doi.org/10.1016/j.enggeo.2007.11.011.
5. Angel, J.C., Nelson, D.O., and Panno, S.V., (2004), Comparison of a new GIS-based technique and a manual method for determining sinkhole density: An example from Illinois' sinkhole plain: Journal of Cave and Karst Studies, v. 66, no. 1, p. 9−17.
6. Atzori, S., Baer, G., Antonioli, A., and Salvi, S. (2015). InSAR-based modeling and analysis of sinkholes along the Dead Sea coastline. Geophysical Research Letters, 42(20), 8383-8390.
7. Buttrick, D., and Van Schalkwyk, A. (1998). Hazard and risk assessment for sinkhole formation on dolomite land in South Africa. Environmental Geology, 36(1), 170-178.
8. Cahalan, M. D., and Milewski, A. M. (2018). Sinkhole formation mechanisms and geostatistical-based prediction analysis in a mantled karst terrain. Catena, 165, 333-344.
9. Capp, J. A., and Ebersole, S. (2020). Spatial and stratigraphic distribution of sinkholes in Blount County, Alabama: a study using lidar for mapping sinkhole hazards and risk zones. Geological Society of America. https://doi.org/10.1130/abs/2020se-344021.
10. Castellanza, R., Lollino, P., and Ciantia, M. (2018). A methodological approach to assess the hazard of underground cavities subjected to environmental weathering. Tunnelling and Underground Space Technology, 82, 278–292. https://doi.org/10.1016/j.tust.2018.08.041

11. Ciotoli, G., Di Loreto, E., Finoia, M.G., Liperi, L., Meloni, F., Nisio, S., and Sericola, A., (2016), Sinkhole susceptibility, Lazio Region, central Italy: Journal of Maps, v. 12, no. 2, p. 287−294. https://doi.org/10.1080/17445647.2015.1014939

12. Ciotoli, G., Nisio, S., Finoia, M. G., and Corazza, A. (2012). Susceptibility maps of collapse sinkholes in urban areas by using geospatial analysis. In Proceeding of 7thEUREGEO, European Congress on Regional Geoscientific Cartography and Information Systems (Vol. 2, pp. 480-481).

13. Closson, D., Karaki, N. A., Klinger, Y., and Hussein, M. J. (2005). Subsidence and sinkhole hazard assessment in the southern Dead Sea area, Jordan. Pure and Applied geophysics, 162(2), 221-248.

14. D'Angella, A., Canora, F., and Spilotro, G. (2015). Sinkholes susceptibility assessment in urban environment using heuristic, statistical and Artificial Neural Network (ANN) models in evaporite karst system: A case study from Lesina Marina (southern Italy). In Engineering Geology for Society and Territory-Volume 5 (pp. 411-414). Springer, Cham.

15. de Castro Tayer, T., and Rodrigues, P. C. H. (2021). Assessment of a semi-automatic spatial analysis method to identify and map sinkholes in the Carste Lagoa Santa environmental protection unit, Brazil. Environmental Earth Sciences, 80(3). https://doi.org/10.1007/s12665-020-09354-z

16. di Santolo, A. S., Forte, G., De Falco, M., and Santo, A. (2016). Sinkhole Risk Assessment in the Metropolitan Area of Napoli, Italy. Procedia Engineering, 158, 458-463.

17. Doctor, D.H., and Young, J.A., (2013), An evaluation of automated GIS tools for delineating karst sinkholes and closed depressions from 1-meter LIDAR-derived digital elevation data, https://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=1156andcontext=sinkhole_2013 [accessed August 18, 2019]. https://doi.org/10.5038/9780979542275.1156.

18. Doctor, K. Z., Doctor, D. H., Kronenfeld, B., Wong, D. W., and Brezinski, D. K. (2008a). Predicting sinkhole susceptibility in Frederick Valley, Maryland, using geographically weighted regression. In 11th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst. Tallahassee, Florida, United States: American Society of Civil Engineers (pp. 100-102).

19. Doctor, K. Z.,(2008b), GIS and Spatial Statistical Methods for Determining Sinkhole Potential in Frederick Valley, Maryland.

20. Dursun, A. E. (2022). Risk analysis of natural sinkholes hazards in Karapınar basin (Konya, Turkey). Arabian Journal of Geosciences, 15(3), 1-20.

21. Elmahdy, S. I., Mohamed, M. M., Ali, T. A., Abdalla, J. E. D., and Abouleish, M. (2020). Land subsidence and sinkholes susceptibility mapping and analysis using random forest and frequency ratio models in Al Ain, UAE. Geocarto International, 37(1), 315-331.

22. Evaluation of susceptibility for terrain collapse and subsidence in karst areas, municipality of Iraquara, Chapada Diamantina (BA), Brazil, 2018. de Queiroz Salles, L., Galvão, P., Leal, L.R.B. et al. Evaluation of susceptibility for terrain collapse and subsidence in karst areas, municipality of Iraquara, Chapada Diamantina (BA), Brazil. Environ Earth Sci **77**, 593 (2018). https://doi.org/10.1007/s12665-018-7769-8

23. Galve, J.P., Gutiérrez, F., Cendrero, A., Remondo, J., Bonachea, J., Guerrero, J., and Lucha, P., (2009b), Predicting sinkholes by means of probabilistic models: Quarterly Journal of Engineering Geology and Hydrogeology, v. 42, no. 2, p. 139−144. https://doi.org/10.1144/1470-9236/08- 039.
24. Galve, J.P., Gutiérrez, F., Lucha, P., Guerrero, J., Bonachea, J., Remondo, J., and Cendrero, A., (2009a), Probabilistic sinkhole modelling for hazard assessment: Earth Surface Processes and Landforms, v. 34, no. 3, p. 437−452. https://doi.org/10.1002/esp.1753.
25. Galve, J.P., Gutiérrez, F., Remondo, J., Bonachea, J., Lucha, P., and Cendrero, A., (2009c), Evaluating and comparing methods of sinkhole susceptibility mapping in the Ebro Valley evaporite karst (NE Spain): Geomorphology, v. 111, no. 3−4, p. 160−172. https://doi.org/10.1016/j. geomorph.2009.04.017
26. Galve, J.P., Remondo, J., and Gutiérrez, F., 2011, improving sinkhole hazard models incorporating magnitude-frequency relationships and nearest neighbour analysis: Geomorphology, v. 134, p. 157−170. https://doi.org/10.1016/j.geomorph.2011.05.020
27. Gao, Y, Alexander EC Jr, (2003) A mathematical model for a sinkhole probability map in fillmore county, minnesota. In: Beck BF (ed) Sinkholes and the engineering and environmental impacts of karsts. Proceedings of the 9th multidisciplinary conference. Huntsville, Alabama, September 6–10, ASCE Geotechnical Special Publication, 122: pp 439–449
28. Gao, Y., and Alexander, E. C. (2008). Sinkhole hazard assessment in Minnesota using a decision tree model. Environmental Geology, 54(5), 945-956.
29. Guerrero, J., Gutiérrez, F., Bonachea, J., and Lucha, P. (2008). A sinkhole susceptibility zonation based on paleokarst analysis along a stretch of the Madrid-Barcelona high-speed railway built over gypsum- and salt-bearing evaporites (NE Spain). Engineering Geology, 102(1–2), 62–73. https://doi.org/10.1016/j.enggeo.2008.07.010
30. Gutiérrez, F., Galve, J. P., Lucha, P., Castañeda, C., Bonachea, J., and Guerrero, J. (2011). Integrating geomorphological mapping, trenching, InSAR and GPR for the identification and characterization of sinkholes: A review and application in the mantled evaporite karst of the Ebro Valley (NE Spain). Geomorphology, 134(1-2), 144-156.
31. Gutiérrez, F., Guerrero, J., and Lucha, P. (2008). Quantitative sinkhole hazard assessment. A case study from the Ebro Valley evaporite alluvial karst (NE Spain). Natural Hazards, 45(2), 211-233.
32. Hubbard, Jr., D. A. (2003). Use of Regional Sinkhole Mapping for Sinkhole Susceptibility Maps (pp. 61–71). American Society of Civil Engineers (ASCE). https://doi.org/10.1061/40698(2003)5
33. Jeannin, P.-Y. L., and Kopp. (2012). Inception horizon concept as a basis for sinkhole hazard mapping. Geophysical Research Abstracts EGU General Assembly, 14(1122), 2012–5226.
34. Jennings, D. N., Woodmansey, P. R., and Cavers, L. D. (2018). Risk assessment of sinkhole collapse associated with mining, Waihi, New Zealand. In ISRM International Symposium 2000, IS 2000. International Society for Rock Mechanics.
35. Jiang, X., Lei, M., Li, Y., and Dai, J. (2005). National-Scale Risk Assessment of sinkhole Hazard in China (pp. 649–658). American Society of Civil Engineers (ASCE). https://doi.org/10.1061/40796(177)69

36. Kaufmann, G., 2014, Geophysical mapping of solution and collapse sinkholes: Journal of Applied Geophysics, v. 111, p. 271−288. https://doi. org/10.1016/j.jappgeo.2014.10.011.
37. Keaton, J. R., and Boudra, L. H. (2014). Development of sinkhole hazard for pipeline risk assessment in Northern Florida. In Proceedings of the Biennial International Pipeline Conference, IPC (Vol. 3). American Society of Mechanical Engineers (ASME). https://doi.org/10.1115/IPC2014-33117
38. Kemmerly, P.R., 2006, Modelling doline populations with logistic growth functions: Earth Surface Processes and Landforms, v. 32, p. 587–601.
39. Kesseru, Z. (1997). Assessing the risk of cave-collapse sinkholes using analogous information from mining. The Engineering Geology and Hydrogeology of Karst Terrances. Proc. 6th
40. Kim, K., Kim, J., Kwak, T. Y., and Chung, C. K.: Logistic Regression Model for Sinkhole Susceptibility due to Damaged Sewer Pipes, Nat. Hazards, 93, 765–785, https://doi.org/10.1007/s11069-018-3323-y, 2018.
41. Kim, Y. J., and Nam, B. H. (2017). Sinkhole Hazard Mapping Using Frequency Ratio and Logistic Regression Models for Central Florida (pp. 246–256). American Society of Civil Engineers (ASCE). https://doi.org/10.1061/9780784480717.023.
42. Kim, Y. J., Nam, B. H., and Zheng, Q. P. (2020). An artificial neural network approach to sinkhole hazard assessment for East Central Florida. doi: https://doi.org/10.5038/9781733375313.1031
43. Kim, Y. J., Nam, B. H., Jung, Y. H., Liu, X., Choi, S., Kim, D., and Kim, S. (2022). Probabilistic spatial susceptibility modeling of carbonate karst sinkhole. Engineering Geology, 106728.
44. Kim, Y. J., Nam, B. H., Shamet, R., Soliman, M., and Youn, H. (2020). Development of sinkhole susceptibility map of east central Florida. Natural Hazards Review, 21(4), 04020035.
45. Kim, Y. J., Xiao, H., Wang, D., Choi, Y. W., and Nam, B. H. (2017). Development of Sinkhole Hazard Mapping for Central Florida. In Geotechnical Frontiers 2017 (pp. 459-468).
46. Kim, Y., and Nam, B. H. (2018). A Comparative Study of Karst Sinkhole Hazard Mapping Using Frequency Ratio and Artificial Neural Network for East Central Florida. University of South Florida Libraries. https://doi.org/10.5038/9780991000982.1035
47. Kim, Y., Nam, B. H., Lim, C., Jung, H. S., and Moon, J. S. (2018). A Decision Tree Based Hazard Assessment of Karst Sinkholes. In IFCEE 2018 (pp. 43-53).
48. Kotyrba, A., and Kortas, Ł. (2016). Sinkhole hazard assessment in the area of abandoned mining shaft basing on microgravity survey and modelling—Case study from the Upper Silesia Coal Basin in Poland. Journal of applied geophysics, 130, 62-70.
49. Kropotkin, M. P. (2019). ESTIMATION OF KARST SINKHOLE HAZARD — DETERMINISTIC OR STOCHASTIC? Engineering Survey, 12(9–10), 12–26. https://doi.org/10.25296/1997-8650-2018-12-9-10-12-26

50. Lamelas, M.T., Marinoni, O., Hoppe, A., and De La Riva, J., 2008, Doline probability map using logistic regression and GIS technology in the central Ebro Basin (Spain): Environmental Geology, v. 54, no. 5, p. 963−977. https://doi.org/10.1007/s00254-007-0895-3.

51. Lollino, P., Perrotti, M., Fazio, N. L., and Parise, M. (2019, June). Sinkhole susceptibility assessment of underground caves in soft rocks by means of FEM-based charts. In 53rd US Rock Mechanics/Geomechanics Symposium. OnePetro.

52. Malinowska, A. A., and Dziarek, K. (2014). Modelling of cave-in occurrence using AHP and GIS. Natural Hazards and Earth System Sciences, 14(8), 1945-1951.

53. Malinowska, A. A., and Matonóg, A. (2017). Sinkhole hazard maping with the use of spatial analysis and analytical hierarchy process in the light of mining-geological factors. Acta Geodynamica et Geomaterialia, 14(2), 159–172. https://doi.org/10.13168/AGG.2016.0037

54. Margiotta, S., Negri, S., Parise, M., and Valloni, R.,(2012), Mapping the susceptibility to sinkholes in coastal areas, based on stratigraphy, geomorphology and geophysics: Natural Hazards, v. 62, no. 2, p. 657−676. https://doi.org/10.1007/s11069-012-0100-1

55. Motlaq, K. S., Pirasteh, S., Norozi, O. A., Pradhan, B., and Mahmoodzadeh, A. (2011). Sinkholes mapping using gis: vulnarebility and risk assessment emphasis on development areas. International Geoinformatics Research and Development Journal, 2(2), 49-61.

56. Multidisciplinary Conference, Springfield, 1997, 55–60. https://doi.org/10.1201/9781003078128-8

57. Nam, B. H., Kim, Y. J., and Youn, H. (2019). Probabilistic spatiotemporal-magnitude sinkhole hazard map for east central florida. In Proceedings of the Symposium on the Application of Geophyics to Engineering and Environmental Problems, SAGEEP (Vol. 2019-March, p. 156). J and N Group, Ltd.

58. Nefeslioglu, H. A., Tavus, B., Er, M., Ertugrul, G., Ozdemir, A., Kaya, A., and Kocaman, S. (2021). Integration of an InSAR and ANN for sinkhole susceptibility mapping: A case study from Kirikkale-Delice (Turkey). ISPRS International Journal of Geo-Information, 10(3), 119.

59. Ogden, A.E., 1984. Methods for describing and predicting the mechanical strength of the rock and vegetation occurrence of sinkholes. In: Beck, B.F., (Ed.), Sinkholes: Their Geology Engineering and Environmental Impact. Proceedings of the 1st Multidisciplinary Conference on Sinkholes holes, Orlando, FL, pp. 177–182.

60. Oh, H. J., and Lee, S. (2010). Assessment of ground subsidence using GIS and the weights-of-evidence model. Engineering Geology, 115(1-2), 36-48.

61. Orhan, O., Yakar, M., and Ekercin, S. (2020). An application on sinkhole susceptibility mapping by integrating remote sensing and geographic information systems. Arabian J. Geosciences 13 (17), 1–17. doi:10.1007/ s12517-020-05841-6

62. Ozdemir, A. (2015). Sinkhole susceptibility mapping using a frequency ratio method and GIS technology near Karapınar, Konya-Turkey. Procedia Earth and Planetary Science, 15, 502-506.

63. Papadopoulou-Vrynioti, K., Bathrellos, G. D., Skilodimou, H. D., Kaviris, G., and Makropoulos, K. (2013). Karst collapse susceptibility mapping considering peak ground acceleration in a rapidly growing urban area. Engineering Geology, 158, 77-88.

64. Perlow Jr, M., Perlow, M. W., and Wagner, J. (2017). Limit State Sinkhole Risk Assessment. In Geotechnical Frontiers 2017 (pp. 495-504). DOI: 10.1061/9780784480441.052.

65. Perrin, J., Cartannaz, C., Noury, G., and Vanoudheusden, E. (2015). A multicriteria approach to karst subsidence hazard mapping supported by weights-of-evidence analysis. Engineering Geology, 197, 296-305.

66. Qiu, X., Wu, S. S., and Chen, Y. (2020). Sinkhole susceptibility assessment based on morphological, imagery, and contextual attributes derived from gis and imagery data. Journal of Cave and Karst Studies, 82(1), 1-17.

67. Sartain, N., Mian, J., O'Riordan, N., and Storry, R. (2011). Case study on the assessment of sinkhole risk for the development of infrastructure over karstic ground. Geotechnical Safety and Risk. ISGSR 2011, 635-642.

68. Ścigała, R., Duży, S., Szafulera, K., Kruczkowski, M., Dyduch, G., and Jendryś, M. (2021). Assessment of Sinkhole Hazard in the Area of Shallow Mining Workings Using Electrical Resistivity Tomography. Energies, 14(24), 8197.

69. Shamet, R., and Nam, B. H. (2020). Probabilistic Field Assessment of Sinkhole Occurrence Using the Raveling Index. In Geo-Congress 2020: Modeling, Geomaterials, and Site Characterization (pp. 602-611). Reston, VA: American Society of Civil Engineers.

70. Shamet, R., Nam, B. H., Tu, T., and Horhota, D. (2019, March). Sinkhole Vulnerability Assessment Using Groundwater Monitoring and Internal Soil Raveling Analysis—A Central Florida Case Study. In Geo-Congress 2019: Engineering Geology, Site Characterization, and Geophysics (pp. 289-299). Reston, VA: American Society of Civil Engineers.

71. Strzałkowski, P. (2015). Mathematical model of forecasting the formation of sinkhole using Salustowicz's theory. Archives of Mining Sciences, 60(1), 63-71.

72. Subedi, P., Subedi, K., Thapa, B., and Subedi, P. (2019). Sinkhole susceptibility mapping in Marion County, Florida: Evaluation and comparison between analytical hierarchy process and logistic regression based approaches. Scientific Reports, 9(1). https://doi.org/10.1038/s41598-019-43705-6

73. Taheri, K., Gutiérrez, F., Mohseni, H., Raeisi, E., and Taheri, M. (2015). Sinkhole susceptibility mapping using the analytical hierarchy process (AHP) and magnitude–frequency relationships: A case study in Hamadan province, Iran. Geomorphology, 234, 64-79.

74. Taheri, K., Shahabi, H., Chapi, K., Shirzadi, A., Gutiérrez, F., and Khosravi, K. (2019). Sinkhole susceptibility mapping: A comparison between Bayes-based machine learning algorithms. Land Degradation and Development, 30(7), 730–745. https://doi.org/10.1002/ldr.3255

75. Theron, A., and Engelbrecht, J. (2018). The role of earth observation, with a focus on SAR Interferometry, for sinkhole hazard assessment. Remote Sensing, 10(10), 1506.

76. Todd, A., and Ivey-Burden, L. (2016). A Method of Mapping Sinkhole Susceptibility Using a Geographic Information System: A Case Study for Interstates in the Karst Counties of Virginia (pp. 299–306). University of South Florida Libraries. https://doi.org/10.5038/9780991000951.1048

77. Upchurch, S. B., and Littlefield, J. R. (1988). Evaluation of data for Sinkhole-development risk models. Environmental Geology and Water Sciences, 12(2), 135–140. https://doi.org/10.1007/BF02574800

78. Witthuhn MK, Alexander EC, Jr. (1995) Sinkholes and sinkhole probability. Geologic Atlas Fillmore County, Minnesota, County Atlas Series C-8, Part B, Plate 8 (1:100,000). Minnesota Department of Natural Resources, Division of Waters

79. Yilmaz, I., Marschalko, M., and Bednarik, M. (2013). An assessment on the use of bivariate, multivariate and soft computing techniques for collapse susceptibility in GIS environ. Journal of earth system science, 122(2), 371-388.

80. Zhou, W., Beck, B. F., and Adams, A. L. (2003). Application of matrix analysis in delineating sinkhole risk areas along highway (I-70 near Frederick, Maryland). Environmental Geology, 44(7), 834–842. https://doi.org/10.1007/s00254-003-0828-8.

81. Zhou, W., Beck, B. F., and Adams, A. L. (2003). Sinkhole Risk Assessment along Highway I-70 near Frederick, Maryland (pp. 591–604). American Society of Civil Engineers (ASCE). https://doi.org/10.1061/40698(2003)54.

82. Zini, L., Calligaris, C., Forte, E., Petronio, L., Zavagno, E., Boccali, C., and Cucchi, F., 2015, A multidisciplinary approach in sinkhole analysis: The Quinis village case study (NE-Italy) Engineering Geology, v. 197, p. 132−144. https://doi.org/10.1016/j.enggeo.2015.07.004.

References theme "Sinkhole Morphology Studies"

1. Al-Halbouni, D., Emam, S., Holohan, E. P., Taheri, A., Schöpfer, M. P., and Dahm, T. (2020). PFC2D modelling of sinkhole cluster in karstic depressions. In *Itasca Symposium 2020: Extended Abstracts* (pp. 1-7). Itasca Consulting Group, Inc. Retrieved from https://itasca-downloads.s3.amazonaws.com/documents/Itasca%20symposia/2020/ExtendedAbstracts/01-07.pdf.
2. Avutia, D., and Kalumba, D. (2014). Analytical study of dolomite sinkholes in Centurion, South Africa. In *Soil Behavior and Geomechanics* (pp. 633-641), doi: 10.1061/9780784413388.066.

3. Angel, J.C., Nelson, D.O., and Panno, S.V., 2004, Comparison of a new GIS-based technique and a manual method for determining sinkhole density: An example from Illinois' sinkhole plain: Journal of Cave and Karst Studies, v. 66, p. 9−17.

4. Ardau, F., Balia, R., Bianco, M., and De Waele, J. (2007). Assessment of cover-collapse sinkholes in SW Sardinia (Italy). Geological Society, London, Special Publications, 279(1), 47-57.

5. Argentieri, A., Carluccio, R., Cecchini, F., Chiappini, M., Ciotoli, G., De Ritis, R., ... and Venuti, A. (2015). Early stage sinkhole formation in the Acque Albule basin of central Italy from geophysical and geochemical observations. Engineering geology, 191, 36-47.

6. Baryakh, A. A., and Fedoseev, A. K. (2011). Sinkhole formation mechanism. Journal of Mining Science, 47, 404-412.

7. Basso, A., Bruno, E., Parise, M., and Pepe, M. (2013). Morphometric analysis of sinkholes in a karst coastal area of southern Apulia (Italy). Environmental earth sciences, 70, 2545-2559.

8. Bastos, A. C., Moura, R. L., Amado-Filho, G. M., D'Agostini, D. P., Secchin, N. A., Francini-Filho, R. B., ... and Thompson, F. L. (2013). Buracas: novel and unusual sinkhole-like features in the Abrolhos Bank. Continental Shelf Research, 70, 118-125.

9. Brinkmann, R., Parise, M., and Dye, D. (2008). Sinkhole distribution in a rapidly developing urban environment: Hillsborough County, Tampa Bay area, Florida. Engineering Geology, 99(3-4), 169-184.
10. Bruno, E., Calcaterra, D., and Parise, M. (2008). Development and morphometry of sinkholes in coastal plains of Apulia, southern Italy. Preliminary sinkhole susceptibility assessment. Engineering Geology, 99(3-4), 198-209.

11. Cahalan, M. D., and Milewski, A. M. (2018). Sinkhole formation mechanisms and geostatistical-based prediction analysis in a mantled karst terrain. Catena, 165, 333-344.

12. Caramanna, G., Ciotoli, G., and Nisio, S. (2008). A review of natural sinkhole phenomena in Italian plain areas. Natural hazards, 45(2), 145-172.

13. Chenoweth, M. S. (1997). The spatial distribution and morphometric analysis of dolines, Buffalo National River, Newton County, Arkansas (Doctoral dissertation, University of Arkansas).

14. Constantinou, S., and Van Rooy, J. L. (2018). Sinkhole and subsidence size distribution across dolomitic land in Gauteng. Journal of the South african institution of civil engineering, 60(2), 2-8.

15. Del Prete, S., Iovine, G., Parise, M., and Santo, A. (2010). Origin and distribution of different types of sinkholes in the plain areas of Southern Italy. Geodinamica Acta, 23(1-3), 113-127.

16. Eppelbaum, L. V., Ezersky, M., Al-Zoubi, A., Goldshmidt, V., and Legchenko, A. (2008). Study of the factors affecting the karst volume assessment in the Dead Sea sinkhole problem using microgravity field analysis and 3-D modeling. Advances in Geosciences, 19, 97-115.

17. Ezersky, M., Legchenko, A., Camerlynck, C., and Al-Zoubi, A. (2009). Identification of sinkhole development mechanism based on a combined geophysical study in Nahal Hever South area (Dead Sea coast of Israel). Environmental Geology, 58, 1123-1141.

18. Ezersky, M., Legchenko, A., Camerlynck, C., and Al-Zoubi, A. (2009). Identification of sinkhole development mechanism based on a combined geophysical study in Nahal Hever South area (Dead Sea coast of Israel). Environmental Geology, 58, 1123-1141.

19. Festa, V., Fiore, A., Parise, M., and Siniscalchi, A. (2012). Sinkhole evolution in the Apulian karst of southern Italy: a case study, with some considerations on sinkhole hazards. Journal of Cave and Karst Studies, 74(2), 137-147.

20. Florea, L. J. (2005). Using State-wide GIS data to identify the coincidence betwen sinkholes and geologic structure. Geography/Geology Faculty Publications, 14.
21. Gao Yongli, Alexander, E.C., and Barnes, R., 2005, Karst database implementation in Minnesota: analysis of sinkhole distribution: Environmental Geology, v. 47, no. 8, p. 1083–1098. https://doi.org/10.1007/s00254-005-1241-2.
22. Ezersky, M. G., Frumkin, A., Al-Zoubi, A., Abueladas, A., Keydar, S., Eppelbaum, L., and Levi, E. (2011, September). Geophysical assessment of salt sinkholes hazard along the Dead Sea shore in Israel and Jordan. In Near Surface 2011-17th EAGE European Meeting of Environmental and Engineering Geophysics (pp. cp-253). European Association of Geoscientists and Engineers. https://doi.org/10.3997/2214-4609.20144466.
23. Gutiérrez, F. (2016). Sinkhole hazards. In Oxford research encyclopedia of natural hazard science. DOI: doi.org/10.1093/acrefore/9780199389407.013.40

24. Gutiérrez, F., Guerrero, J., and Lucha, P.: A genetic classification of sinkholes illustrated from evaporite paleokarst exposures in Spain, Environ. Geol., 53, 993–1006, https://doi.org/10.1007/s00254-007-0727-5, 2008.

25. Heidari, M., Khanlari, G. R., Beydokhti, A. T., and Momeni, A. A. (2011). The formation of cover collapse sinkholes in North of Hamedan, Iran. Geomorphology, 132(3-4), 76-86.

26. Hubbard DA Jr (1984) Sinkhole distribution in the central and northern Valley and Ridge province, Virginia. In: Beck BF (eds) sinkholes: their geology,engineering and environmental impact. Proceedings of the 1st multidisciplinary conference on sinkholes. Orlando, Florida, 15–17 October, A.A.Balkema, Rotterdam, pp 75–78.
27. Hubbard, D. A. (2001). Sinkhole distribution of the Valley and Ridge province, Virginia. In Geotechnical and environmental applications of karst geology and hydrology (pp. 33-36).

28. Hyatt, J. A., Wilkes, H. P., and Jacobs, P. M. (1999). Spatial relationships between new and old sinkholes in covered karst, Albany, Georgia, USA. In Hydrogeology and engineering geology of sinkholes and karst: 1999 (pp. 37-44).

29. Jiang, F. (2020). Criteria of sinkhole formation in soils from physical models. Bulletin of Engineering Geology and the Environment, 79(7), 3833-3841.
30. Kidanu, S. T., Anderson, N. L., and Rogers, J. D. (2018). Using GIS-based spatial analysis to determine factors influencing the formation of sinkholes in Greene County, Missouri. Environmental and Engineering Geoscience, 24(3), 251-261.
31. Kim, J.-W., Lu, Z., and Degrandpre, K. (2016). Ongoing deformation of sinkholes in Wink, Texas, observed by time-series Sentinel-1A SAR interferometry (preliminary results). Remote Sensing, 8(4), 313.

32. Kobal, M., Bertoncelj, I., Pirotti, F., and Kutnar, L. (2014). Lidar processing for defining sinkhole characteristics under dense forest cover: A case study in the Dinaric mountains. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 40, 113-118.
33. La Rosa, A., Pagli, C., Molli, G., Casu, F., De Luca, C., Pieroni, A., and D'Amato Avanzi, G. (2018). Growth of a sinkhole in a seismic zone of the northern Apennines (Italy). Natural Hazards and Earth System Sciences, 18(9), 2355-2366.

34. Lei, M., Gao, Y., and Jiang, X. (2015). Current status and strategic planning of sinkhole collapses in China. In Engineering Geology for Society and Territory-Volume 5: Urban Geology, Sustainable Planning and Landscape Exploitation (pp. 529-533). Springer International Publishing.
35. Lei, M., Gao, Y., Jiang, X., and Guan, Z., 2016, Mechanism analysis of sinkhole formation at Maohe village, Liuzhou city, Guangxi province, China: Environmental Earth Sciences, v. 75, 542. https://doi.org/10.1007/s12665-015-5100-5.

36. Lindsey, B. D., Katz, B. G., Berndt, M. P., Ardis, A. F., and Skach, K. A. (2010). Relations between sinkhole density and anthropogenic contaminants in selected carbonate aquifers in the eastern United States. Environmental Earth Sciences, 60, 1073-1090.

37. Margiotta, S., Negri, S., Parise, M., and Quarta, T. A. M. (2016). Karst geosites at risk of collapse: the sinkholes at Nociglia (Apulia, SE Italy). Environmental Earth Sciences, 75, 1-10.

38. McDowell, P. W., Coulton, J., Edmonds, C. N., and Poulsom, A. J. (2008). The nature, formation and engineering significance of sinkholes related to dissolution of chalk in SE Hampshire, England. Quarterly Journal of Engineering Geology and Hydrogeology, 41(3), 279-290.

39. McDowell, P. W., Coulton, J., Edmonds, C. N., and Poulsom, A. J. (2008). The nature, formation and engineering significance of sinkholes related to dissolution of chalk in SE Hampshire, England. Quarterly Journal of Engineering Geology and Hydrogeology, 41(3), 279-290.
40. Mouici, R., Baali, F., Hadji, R., Boubaya, D., Audra, P., Fehdi, C. É., ... and Arfib, B. (2017). Geophysical, geotechnical, and speleologic assessment for karst-sinkhole collapse genesis in Cheria plateau (NE Algeria). Mining Science, 24, 59-71.
41. Panno, S. V., and Luman, D. E. (2018). Characterization of cover-collapse sinkhole morphology on a groundwater basin-wide scale using lidar elevation data: A new conceptual model for sinkhole evolution. Geomorphology, 318, 1-17.

42. Panno, S. V., Wiebel, C. P., Heigold, P. C., and Reed, P. C. (1994). Formation of regolith-collapse sinkholes in southern Illinois: Interpretation and identification of associated buried cavities. Environmental Geology, 23(3), 214-220.
43. Parise, M., 2012, A present risk from past activities: sinkhole occurrence above underground quarries: Carbonates and Evaporites, v. 27, no. 2, p. 109−118. https://doi.org/10.1007/s13146-012-0088-3.

44. Parise, M., and Vennari, C., 2013, A chronological catalogue of sinkholes in Italy: The first step toward a real evaluation of the sinkhole hazard, in Land, L., Doctor, D.H., and Stephenson, B., eds., Proceedings of the 13th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst, Carlsbad (New Mexico, USA), May 6−10, 2013, National Cave and Karst Research Institute, p. 383−392.
45. Qiu, X. and Wu, S., 2017, A Knowledge-based computerized approach to the development of a sinkhole database: The Professional Geographer, v. 69, no. 2, p. 239−250. https://doi.org/10.1080/00330124.2016.1229622.

46. Rodriguez, V., Gutiérrez, F., Green, A. G., Carbonel, D., Horstmeyer, H., and Schmelzbach, C. (2014). Characterizing sagging and collapse sinkholes in a mantled karst by means of ground penetrating radar (GPR). Environmental and Engineering Geoscience, 20(2), 109-132.
47. Southworth, C. S. (1984). STRUCTURAL AND HYDROGEOLOGIC APPLICATIONS OF REMOTE SENSING DATA, EASTERN YUCATAN PENINSULA, MEXICO. In Sinkholes: Their Geology, Engineering and Environmental Impact, Proceedings of the First Multidisciplinary Conference on Sinkholes. (pp. 59-64).

48. Tihansky, A. B. (1999). Sinkholes, west-central Florida. Land subsidence in the United States: US geological survey circular, 1182, 121-140.
49. Valois, R., Camerlynck, C., Dhemaied, A., Guerin, R., Hovhannissian, G., Plagnes, V., ... and Robain, H. (2011). Assessment of doline geometry using geophysics on the Quercy plateau karst (South France). Earth surface processes and landforms, 36(9), 1183-1192. doi: 10.1002/esp.2144.

50. Van Den Eeckhaut, M., Poesen, J., Dusar, M., Martens, V., and Duchateau, P. (2007). Sinkhole formation above underground limestone quarries: A case study in South Limburg (Belgium). Geomorphology, 91(1-2), 19-37.
51. White, E.L., Aron, G., and White, W.B., 1986, The influence of urbanization on sinkholes development in Central Pennsylvania: Environmental Geology and Water Sciences., v. 8, p. 91–97. https://doi.org/10.1007/BF02525562.
52. Zumpano, V., Pisano, L., and Parise, M. (2019). An integrated framework to identify and analyze karst sinkholes. Geomorphology, 332, 213-225.

References of the theme "Sinkhole Monitoring and Detection"

1. Ahmed, M. (2013). Application of digital image cross correlation to study sinkhole collapse. International Scholarly Research Notices, 2013.
2. Alfuqara, D. A. A. (2017). 2-D and 3-D mapping of soil and rock structure in karst terrain of southwestern Missouri using the non-invasive electrical resistivity tomography method (ERT). Missouri University of Science and Technology.
3. Al-Halbouni, D., Holohan, E. P., Saberi, L., Alrshdan, H., Sawarieh, A., Closson, D., ... and Dahm, T. (2017). Sinkholes, subsidence and subrosion on the eastern shore of the Dead Sea as revealed by a close-range photogrammetric survey. Geomorphology, 285, 305-324.
4. Al-Zoubi, A., Abueladas, A. R., Camerlynck, C., Al-Ruzouq, R., Al-Rawashdeh, S., Ezersky, M., and Ali, W. (2007). Use of 2D multi electrodes resistivity imagining for sinkholes hazard assessment along the eastern part of the Dead Sea, Jordan. American Journal of Environmental Sciences, 3(4), 229-233.
5. Batayneh, A. T., Abueladas, A. A., and Moumani, K. A. (2002). Use of ground-penetrating radar for assessment of potential sinkhole conditions: an example from Ghor al Haditha area, Jordan. Environmental Geology, 41(8), 977-983.
6. Billi, A., De Filippis, L., Poncia, P. P., Sella, P., and Faccenna, C. (2016). Hidden sinkholes and karst cavities in the travertine plateau of a highly-populated geothermal seismic territory (Tivoli, central Italy). Geomorphology, 255, 63-80.
7. Bodet, L., Galibert, P. Y., Dhemaied, A., Camerlynck, C., and Al-Zoubi, A. (2010, June). Surface-wave profiling for sinkhole hazard assessment along the eastern Dead Sea shoreline, Ghor Al–Haditha, Jordan. In 72nd EAGE Conference and Exhibition incorporating SPE EUROPEC 2010 (pp. cp-161). European Association of Geoscientists and Engineers.
8. Brinkmann, R., and Parise, M., 2010, The timing of sinkhole formation in Tampa and Orlando, Florida: The Florida Geographer, v. 41, p. 22−38.
9. Brinkmann, R., Wilson, K., Elko, N., Seale, L.D., Florea, L., and Vacher, H.L., 2007, Sinkhole distribution based on pre-development mapping in urbanized Pinellas County, Florida, USA, in Parise, M., and Gunn, J., eds., Natural and Anthropogenic Hazards in Karst Areas: Recognition, Analysis and Mitigation. Geological Society, London, Special Publication no. 279, p. 5–11. doi.org/10.1144/SP279.2
10. Chang, L., and Hanssen, R. F. (2014). Detection of cavity migration and sinkhole risk using radar interferometric time series. Remote sensing of environment, 147, 56-64.
11. Ékes, C. (2017). Sinkhole locating and corrosion quantification with pipe penetrating radar. In Pipelines 2017 (pp. 1-11).

12. Gutiérrez, F., Benito-Calvo, A., Carbonel, D., Desir, G., Sevil, J., Guerrero, J., and Fabregat, I. (2019). Review on sinkhole monitoring and performance of remediation measures by high-precision leveling and terrestrial laser scanner in the salt karst of the Ebro Valley, Spain. Engineering Geology, 248, 283-308.
13. Higuera-Díaz, I. C., Carpenter, P. J., and Thompson, M. D. (2007). Identification of buried sinkholes using refraction tomography at Ft. Campbell Army Airfield, Kentucky. Environmental Geology, 53, 805-812.
14. Intrieri, E., Gigli, G., Nocentini, M., Lombardi, L., Mugnai, F., Fidolini, F., and Casagli, N. (2015). Sinkhole monitoring and early warning: An experimental and successful GB-InSAR application. Geomorphology, 241, 304-314.
15. Khoshlahjeh Azar, M.; Hamedpour, A.; Maghsoudi, Y.; Perissin, D. Analysis of the Deformation Behavior and Sinkhole Risk in Kerdabad, Iran Using the PS-InSAR Method. Remote Sens. 2021, 13, 2696. https://doi.org/10.3390/rs13142696.
16. Kim, Y. J., Nam, B. H., and Youn, H. (2019). Sinkhole detection and characterization using LiDAR-derived DEM with logistic regression. Remote Sensing, 11(13), 1592.
17. Kromhout, C., Alfieri, M. C. (2018). Assessment of Historical Aerial Photography as Initial Screening Tool to Identify Areas at Possible Risk to Sinkhole Development.
18. Kromhout, C., Alfieri, M. C. (2018). Assessment of Historical Aerial Photography as Initial Screening Tool to Identify Areas at Possible Risk to Sinkhole Development.
19. Lei, M., Gao, Y., Li, Y., Meng, Y., Yu, L., Gan, F. (2008). Detection and treatment of sinkholes and subsurface voids along Guilin-Yangshuo highway, Guangxi, China. In Sinkholes and the engineering and environmental impacts of Karst (pp. 632-639).
20. Margiotta, S., Negri, S., Parise, M., and Quarta, T. A. (2015). Sinkhole field above karst caves: detection and analysis through integrated techniques. In Engineering Geology for Society and Territory-Volume 5: Urban Geology, Sustainable Planning and Landscape Exploitation (pp. 553-557). Springer International Publishing.
21. Mellett JS, MacCarillo BJ (1996) A model for sinkhole formation on interstate and limited access highways. with suggestions on reme-diation. Int J Rock Mech Min Sci Geomech Abst 33(2):A85
22. Miao Xin, Qiu Xiaomin, Wu Shuo-Sheng, Luo Jun, Gouzie, D.R., and Xie Hongjie, 2013, Developing efficient procedures for automated sinkhole extraction from Lidar DEMs: Photogrammetric Engineering Remote Sensing, v. 79, no. 6, p. 545–554.
23. Miao, X., Qiu, X., Wu, S. S., Luo, J., Gouzie, D. R., and Xie, H. (2013). Developing efficient procedures for automated sinkhole extraction from Lidar DEMs. Photogrammetric Engineering and Remote Sensing, 79(6), 545-554.

24. Nof, R. N., Abelson, M., Raz, E., Magen, Y., Atzori, S., Salvi, S., and Baer, G. (2019). SAR interferometry for sinkhole early warning and susceptibility assessment along the Dead Sea, Israel. Remote Sensing, 11(1), 89.
25. Nwokebuihe, S., Torgashov, E., and Anderson, N. (2014, March). A description of an effective sinkhole investigation approach: a case study of a site in Greene County, Missouri. In Symposium on the Application of Geophysics to Engineering and Environmental Problems 2014 (pp. 304-318). Society of Exploration Geophysicists and Environment and Engineering Geophysical Society.
26. Pazzi, V., Ceccatelli, M., Gracchi, T., Masi, E. B., and Fanti, R. (2018). Assessing subsoil void hazards along a road system using H/V measurements, ERTs and IPTs to support local decision makers. Near Surface Geophysics, 16(3), 282-297.
27. Pazzi, Veronica, Michele Di Filippo, Maria Di Nezza, Tommaso Carlà, Federica Bardi, Federico Marini, Katia Fontanelli, Emanuele Intrieri, and Riccardo Fanti. "Integrated geophysical survey in a sinkhole-prone area: Microgravity, electrical resistivity tomographies, and seismic noise measurements to delimit its extension." Engineering Geology 243 (2018): 282-293.
28. Rigbey, S. (2004). Monitoring upstream sinkhole development by detailed sonar profiling. Available from Golder Associates, 1000, 940 - 6th Avenue SW, Calgary, Alberta, T2P 3T1
29. Sevil, J., Benito-Calvo, A., and Gutiérrez, F. (2021). Sinkhole subsidence monitoring combining terrestrial laser scanner and high-precision levelling. Earth Surface Processes and Landforms, 46(8), 1431-1444.
30. Sevil, J., Gutiérrez, F., Zarroca, M., Desir, G., Carbonel, D., Guerrero, J., ... and Fabregat, I. (2017). Sinkhole investigation in an urban area by trenching in combination with GPR, ERT and high-precision leveling. Mantled evaporite karst of Zaragoza city, NE Spain. Engineering Geology, 231, 9-20.
31. Soliman, M. (2020). Numerical Analysis on the Geomechanical Mechanisms and Stability of Sinkholes in Central Florida.
32. Williams, G. B., Banks, V. J., Bowman, E. T., Cooper, A. H., Jones, L. D., Kirkham, M. P., ... and Shepley, P. (2020). Integrating monitoring techniques for a buried sinkhole in an urban environment.
33. Wu, Q., Deng, C., and Chen, Z. (2016). Automated delineation of karst sinkholes from LiDAR-derived digital elevation models. Geomorphology, 266, 1-10.
34. Xua J, Heb J, Zhang L (2017) Collapse prediction of karst sinkhole via distributed Brillouin optical fiber sensor. Measurement 100:68–71

35. Zhu, J., and Pierskalla Jr, W. P. (2016). Applying a weighted random forests method to extract karst sinkholes from LiDAR data. Journal of Hydrology, 533, 343-352.
36. Zhu, J., Taylor, T. P., Currens, J. C., and Crawford, M. M. (2014). Improved karst sinkhole mapping in kentucky using lidar techniques: a pilot study in floyds fork watershed. Journal of Cave and Karst Studies, 76(3).

References on Mitigation measures for Sinkholes

1. Carbonel, D., Rodríguez, V., Gutiérrez, F., McCalpin, J. P., Linares, R., Roqué, C., and Sasowsky, I. (2014). Evaluation of trenching, ground penetrating radar (GPR) and electrical resistivity tomography (ERT) for sinkhole characterization. Earth Surface Processes and Landforms, 39(2), 214-227.
2. Fadaei, H., Saeideh Sahebi, V., Esmaili Sarteshnizi, R., and Ghasemkhani, N. (2020). Using GIS and DANP in detecting potential areas for Sinkholes. Earth Observation and Geomatics Engineering, 4(2), 132-147.
3. Galve, J. P., Gutiérrez, F., Guerrero, J., Alonso, J., and Diego, I. (2012). Application of risk, cost–benefit and acceptability analyses to identify the most appropriate geosynthetic solution to mitigate sinkhole damage on roads. Engineering geology, 145, 65-77.
4. Gunn, J., 2007, Contributory area definition for groundwater source protection and hazard mitigation in carbonate aquifers, in Parise, M., and Gunn, J., eds., Natural and Anthropogenic Hazards in Karst Areas: Recognition, Analysis, and Mitigation. Geological Society, London. 279, p. 97–109. https://doi.org/10.1144/SP279.9
5. Gutierrez, F., Cooper, A. H., and Johnson, K. S. (2008). Identification, prediction, and mitigation of sinkhole hazards in evaporite karst areas. Environmental Geology, 53, 1007-1022. DOI10.1007/s00254-007-0728-4
6. Kemmerly, P. R. (1993). Sinkhole hazards and risk assessment in a planning context. Journal of the American Planning Association, 59(2), 221-229.
7. Moshodi, T., Coetzee, C., and Fourie, K. (2016). Inadequate stakeholder management and its effect on a coherent sinkhole risk management strategy: The case of the Merafong Local Municipality, South Africa. Jàmbá: Journal of Disaster Risk Studies, 8(1), 1-8.
8. Sahu, P., and Lokhande, R. D. (2015). An investigation of sinkhole subsidence and its preventive measures in underground coal mining. Procedia Earth and Planetary Science, 11, 63-75.
9. Sevil, J., Gutiérrez, F., Zarroca, M., Desir, G., Carbonel, D., Guerrero, J., ... and Fabregat, I. (2017). Sinkhole investigation in an urban area by trenching in combination with GPR, ERT and high-precision leveling. Mantled evaporite karst of Zaragoza city, NE Spain. Engineering Geology, 231, 9-20.
10. Zhou, W. F., and Lei, M. T. (2017). Conceptual site models for sinkhole formation and remediation. Environ Earth Sci 76: 818.
11. Zhou, W., and Beck, B. F. (2008). Management and mitigation of sinkholes on karst lands: an overview of practical applications. Environmental Geology, 55(4), 837-851.

References on the interconnectivity of flood hazard and karst sinkholes

1. Baryakh, A. A., Devyatkov, S. Y., and Samodelkina, N. A. (2016). Theoretical explanation of conditions for sinkholes after emergency flooding of potash mines. Journal of Mining Science, 52, 36-45.
2. Bonacci, O., Ljubenkov, I., and Roje-Bonacci, T., 2006, Karst flash floods: an example from the Dinaric karst, Croatia. Natural Hazards and Earth System Sciences, v. 6, p. 195–203. https://doi.org/10.5194/nhess-6-195-2006
3. Campbell, C. W. (2005). Complexities of flood mapping in a sinkhole area. In Sinkholes and the Engineering and Environmental Impacts of Karst (pp. 470-478).
4. Cardoso de Salis, H. H., Monteiro da Costa, A., Moreira Vianna, J. H., Azeneth Schuler, M., Künne, A., Sanches Fernandes, L. F., and Leal Pacheco, F. A. (2019). Hydrologic modelling for sustainable water resources management in urbanized karst areas. International journal of environmental research and public health, 16(14), 2542.
5. Crawford, N. C. (1984). Sinkhole flooding associated with urban development upon karst terrain: Bowling Green, Kentucky. In Multidisciplinary conference on sinkholes. 1 (pp. 283-292).
6. Fabregat, I., Gutiérrez, F., Roqué, C., Zarroca, M., Linares, R., Comas, X., ... and Carbonel, D. (2019). Subsidence mechanisms and sedimentation in alluvial sinkholes inferred from trenching and ground penetrating radar (GPR). Implications for subsidence and flooding hazard assessment. Quaternary International, 525, 1-15.
7. Field, M. S. (2010). Simulating drainage from a flooded sinkhole. Acta Carsologica, 39(2).
8. Fischer, P., Pistre, S., and Marchand, P. (2022). Effect of fast drainage in karst sinkholes on surface runoff in Larzac Plateau, France. Journal of Hydrology: Regional Studies, 43, 101206.
9. Gregor Kovacic and Nataša x Nataša Ravbar (2010) Extreme hydrological events in karst areas of Slovenia, the case of the Unica River basin, Geodinamica Acta, 23:1-3, 89-100, DOI: 10.3166/ga.23.89-100
10. Hallberg GR, Hoyer BE (1982) Sinkholes, hydrogeology, and groundwater quality in northeast IOWA, Iowa Department of Natural Resources, Geological Survey Bureau, Open File Report 82–3, 120 pp
11. Hart, E. A. (2006). Land use change and sinkhole flooding in Cookeville, Tennessee. southeastern geographer, 46(1), 35-50. https://www.jstor.org/stable/10.2307/26222220.
12. Hart, E. A., Mills, H. H., and Peter, L. (2009). Report on Sinkhole Flooding and Determination of 100-year Sinkhole Floodplains for the City of Cookeville. TN, Department of Earth Sciences, Tennessee Tech University.
13. Hyatt, J. A., and Jacobs, P. M. (1996). Distribution and morphology of sinkholes triggered by flooding following Tropical Storm Alberto at Albany, Georgia, USA. Geomorphology, 17(4), 305-316.
14. Intrieri, E., Fontanelli, K., Bardi, F., Marini, F., Carlà, T., Pazzi, V., ... and Fanti, R. (2018). Definition of sinkhole triggers and susceptibility based on hydrogeomorphological analyses. Environmental earth sciences, 77(1), 4. DOI: 10.1007/s12665-017-7179-3

15. Kemmerly, P. (1981). The need for recognition and implementation of a sinkhole-floodplain hazard designation in urban karst terrains. Environmental Geology, 3(5), 281-292.
16. Kwak, T. Y., Woo, S. I., Chung, C. K., and Kim, J. (2020). Experimental assessment of the relationship between rainfall intensity and sinkholes caused by damaged sewer pipes. Natural Hazards and Earth System Sciences, 20(12), 3343-3359.
17. Lolcama, J. L., Cohen, H. A., and Tonkin, M. J. (2002). Deep karst conduits, flooding, and sinkholes: lessons for the aggregates industry. Engineering Geology, 65(2-3), 151-157.
18. Luu, L. H., Noury, G., Benseghier, Z., and Philippe, P. (2019). Hydro-mechanical modeling of sinkhole occurrence processes in covered karst terrains during a flood. Engineering geology, 260, 105249.
19. Martinotti, M. E., Pisano, L., Marchesini, I., Rossi, M., Peruccacci, S., Brunetti, M. T., Melillo, M., Amoruso, G., Loiacono, P., Vennari, C., Vessia, G., Trabace, M., Parise, M., and Guzzetti, F. (2017): Landslides, floods and sinkholes in a karst environment: the 1–6 September 2014 Gargano event, southern Italy, Nat. Hazards Earth Syst. Sci., 17, 467–480, https://doi.org/10.5194/nhess-17-467-2017.
20. Molina, M., and McDonald, F. (1988). Sinkhole management and flooding in Jamaica. Environmental Geology and Water Sciences, 12(2), 113-116.
21. Morton, L. W., and Olson, K. R. (2015). Sinkholes and sand boils during 2011 record flooding in Cairo, Illinois. Journal of Soil and Water Conservation, 70(3), 49A-54A.
22. Noury, G., Perrin, J., Luu, L. H., Philippe, P., and Gourdier, S. (2018, April). role of floods on sinkholes occurrence in covered karst terrains: case study of orléans area (france) during the 2016 meteorological event and perspectives for other karst environments. In 15th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst.
23. Ogden, A. E. (2018). Investigation of sinkhole flooding problems in Knoxville, Tennessee. In Karst Geohazards (pp. 291-296). Routledge.
24. Palanisamy, B., and Workman, S. R. (2015). Hydrologic modeling of flow through sinkholes located in streambeds of Cane Run Stream, Kentucky. Journal of Hydrologic Engineering, 20(5), 04014066.
25. Parise, M. (2003). Flood history in the karst environment of Castellana-Grotte (Apulia, southern Italy). Natural hazards and Earth system sciences, 3(6), 593-604.
26. Parise, M., Pisano, L., and Vennari, C.: Sinkhole clusters after heavy rainstorms, Journal of Cave and Karst Studies, 80, 28–38, https://doi.org/10.4311/2017ES0105, 2018. Influence (morpho mitune bacshe) khundam tick
27. Rateb, A., and Hermas, E. (2020). The 2018 long rainy season in Kenya: Hydrological changes and correlated land subsidence. Remote Sensing, 12(9), 1390.
28. White, E.L., and White, W.B., 1984, Flood hazards in karst terrains: lessons from the Hurricane Agnes storm, in Burger, A., and Dubertret, L., eds., Hydrogeology of Karst Terrains, v. 1, p. 261−264.
29. Xiao, H., Kim, Y. J., Nam, B. H., and Wang, D. (2016). Investigation of the impacts of local-scale hydrogeologic conditions on sinkhole occurrence in East-Central Florida, USA. Environmental Earth Sciences, 75, 1-16.
30. Xiao, H., Li, H., and Tang, Y. (2018). Assessing the effects of rainfall, groundwater downward leakage, and groundwater head differences on the development of cover-collapse and cover-suffosion sinkholes in central Florida (USA). Science of the total environment, 644, 274-286.

31. Zhou, W. (2007). Drainage and flooding in karst terranes. Environmental Geology, 51(6), 963-973.
32. Zhou, W., and Beck, B. F. (2005). Roadway construction in karst areas: management of stormwater runoff and sinkhole risk assessment. Environmental Geology, 47(8), 1138-1149.

References for the theme "Others"

1. Aurit, M. D., Peterson, R. O., and Blanford, J. I. (2013). A GIS analysis of the relationship between sinkholes, dry-well complaints and groundwater pumping for frost-freeze protection of winter strawberry production in Florida. PLoS One, 8(1), e53832.
2. Baffaut, C., and Benson, V. W. (2008). Possibilities and challenges for modeling flow and pollutant transport in a karst watershed with SWAT. In 2008 Providence, Rhode Island, June 29–July 2, 2008 (p. 1). American Society of Agricultural and Biological Engineers.
3. Bátori, Z., Csiky, J., Farkas, T., Vojtkó, A., Erdős, L., Kovács, D., et al. (2014). The conservation value of karst dolinesfor vascular plants in woodland habitats of Hungary: refugia and climate change. Ijs 43 (1), 15–26. doi:10.5038/1827-806X.43.1.2
4. Boyer, D. G. (2008). Assessment of a sinkhole filter for removing agricultural contaminants. Journal of soil and water conservation, 63(1), 47-52.
5. Davraz, A., Karaguzel, R., Soyaslan, I., Sener, E., Seyman, F., and Sener, S. (2009). Hydrogeology of karst aquifer systems in SW Turkey and an assessment of water quality and contamination problems. Environmental geology, 58, 973-988.
6. Gutiérrez, F., Parise, M., De Waele, J., and Jourde, H. (2014). A review on natural and human-induced geohazards and impacts in karst. Earth-Science Reviews, 138, 61-88.
7. Hackley, K. C. (2007). Groundwater quality of springs and wells of the sinkhole plain in southwestern Illinois: Determination of the dominant sources of nitrate. Circular no. 570.
8. Kalmes, A., and Mohring, E. (1995). Sinkhole treatment to improve water quality and control erosion in southeastern Minnesota. In Karst Geohazards (pp. 265-272). Routledge.
9. Klopmeier, N. W., Pesi, S. M., Morris, G., and Conner, M. (2018). Sinkholes as a Source of Wildlife Mortality. Southeastern Naturalist, 17(4).
10. Lewis, J. J., Milne, M. A., Stephen, C. D., and Dourson, D. C. (2020). Karst sinkholes as foci of biodiversity in the hoosier national forest. Journal of Cave and Karst Studies, 82(4).
11. Liu, R., Zhang, Z., Shen, J., and Wang, Z. (2019). Bryophyte diversity in karst sinkholes afected by diferent degrees of human disturbance. Acta Societatis Botanicorum Poloniae, 88(2).
12. Lolcama, J. L., Cohen, H. A., and Tonkin, M. J. (2002). Deep karst conduits, flooding, and sinkholes: lessons for the aggregates industry. Engineering Geology, 65(2-3), 151-157. https://doi.org/10.1016/S0013-7952(01)00122-3.
13. Petersen, A., and Vondracek, B. (2006). Water quality in relation to vegetative buffers around sinkholes in karst terrain. Journal of soil and water conservation, 61(6), 380-390.
14. Quinlan, J. F. (1986). Legal aspects of sinkhole development and flooding in karst terranes: 1. Review and synthesis. Environmental Geology and Water Sciences, 8(1-2), 41-61.
15. Šaulienė, A., and Šukys, P. (2007). Ground water regime and pollution in the vicinity of Karst sinkholes. In Proceedings of the 13th International Scientific Conference" Research for Rural

Development 2007", Jelgava, Latvia, 16-18 May 2007 (pp. 258-263). Latvia University of Agriculture.
16. Scheidt, J., Lerche, I., and Paleologos, E. (2005). Environmental and economic risks from sinkholes in west-central Florida. Environmental Geosciences, 12(3), 207-217.
17. Shaban, A., and Darwich, T. (2011). The role of sinkholes in groundwater recharge in the high mountains of Lebanon. Journal of Environmental Hydrology, 19.
18. Smithbci, T. J. (2020). Sinkhole damage investigations for the insurance industry in west-central Florida. In The Engineering Geology and Hydrology of Karst Terrains (pp. 299-304). CRC Press.
19. Toulkeridis, T., Rodríguez, F., Arias Jiménez, N., Baile, D. S., Martínez, R. S., Addison, A., ... and Díaz Perez, C. (2016). Causes and consequences of the sinkhole at El Trébol of Quito, Ecuador–implications for economic damage and risk assessment. Natural Hazards and Earth System Sciences, 16(9), 2031-2041.

Multi hazard integrating sinkholes

1. Linares, R., Roqué, C., Gutiérrez, F., Zarroca, M., Carbonel, D., Bach, J., and Fabregat, I. (2017). The impact of droughts and climate change on sinkhole occurrence. A case study from the evaporite karst of the Fluvia Valley, NE Spain. *Science of the total environment*, *579*, 345-358.
2. Yousefi, S., Pourghasemi, H. R., Emami, S. N., Pouyan, S., Eskandari, S., and Tiefenbacher, J. P. (2020). A machine learning framework for multi-hazards modeling and mapping in a mountainous area. *Scientific Reports*, *10*(1), 12144.
3. Kariminejad, N., Pourghasemi, H. R., and Hosseinalizadeh, M. (2022). Analytical techniques for mapping multi-hazard with geo-environmental modeling approaches and UAV images. *Scientific Reports*, *12*(1), 14946.

Appendix 3.A

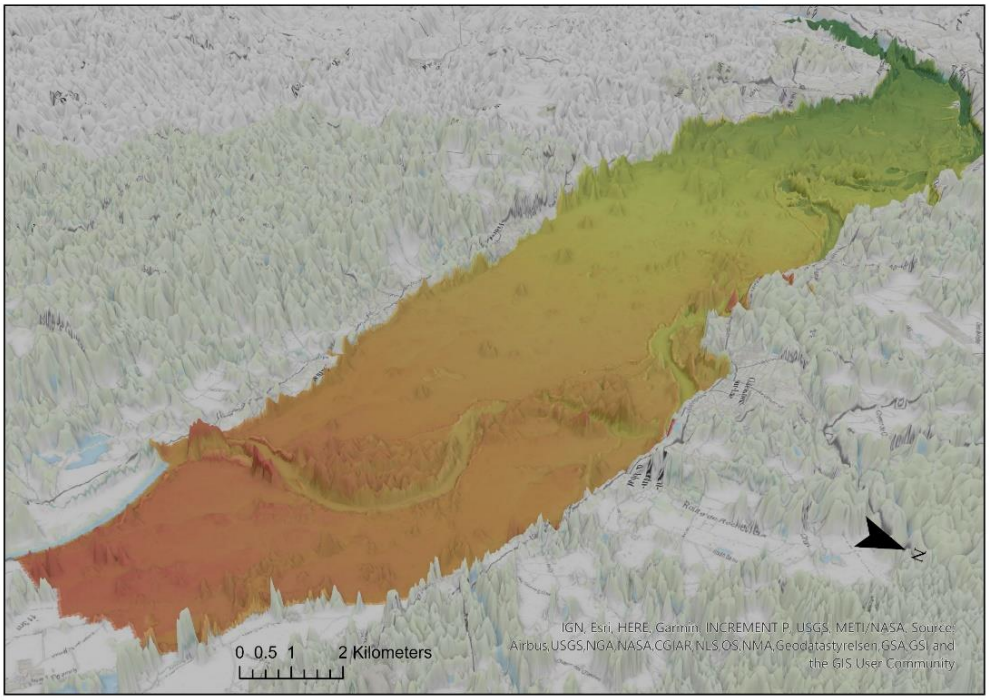


Figure 3.A.1 3D view of the Elevation (Vertical exaggeration of 50%). It shows the low situation of the area compared to the surrounding.

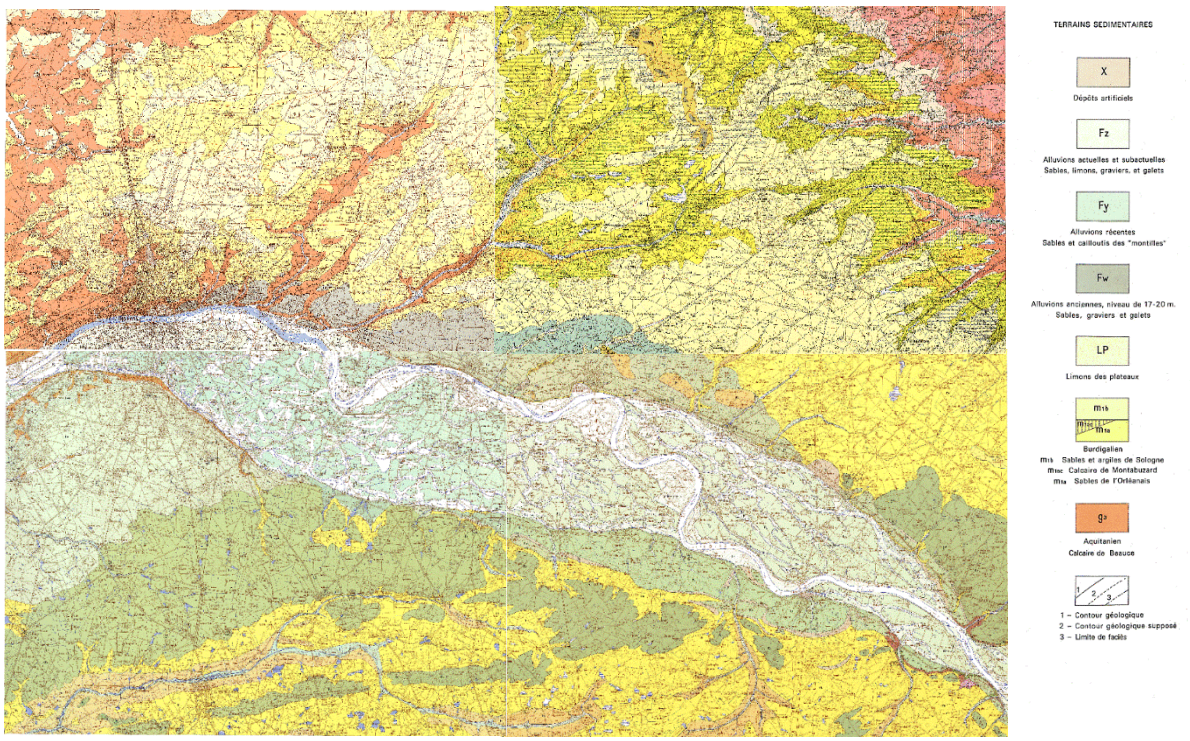


Figure 3.A.2 A scanned (digitized) geological map of the Val d'Orléans (Scale 1:50,000) produced by BRGM. The orange colour (g3) is the Beauce Limestone under which the area is lying. In the map the outcrops of this layer is shown.

Appendix 3.B

Table 3.B.1 Five piezometric gauges in the Val d'Orléans and Five gauges used as interpolations, Loiret department (extracted from ades website access data: 29 December 2022).

Water gauge	Start date	End date	Number of measurements
BSS001ACSC (03635X0545/PZ1)	24/11/2011	18/10/2021	3609
BSS001AEQB (03636X1060/PZ2)	24/05/2010	31/05/2019	3135
BSS001AEQC (03636X1061/PZ3)	24/05/2010	18/10/2021	3679
BSS001AEQD (03636X1062/PZ4)	23/05/2010	01/10/2022	4029
BSS001CBUA (03982X1045/F)	31/03/2010	30/09/2022	4127
5 nearest stations outside the area selected for interpolation			
Water gauge	**Start date**	**End date**	**Number of measurements**
BSS001CHNQ (03997X0004/F)	03/03/2005	23/04/2022	6974
BSS001CCTX (03983X0267/PZ3)	03/08/1994	01/02/2022	18279
BSS001BWKE (03974X0007/F)	04/11/1974	07/02/2004	1566
BSS001AFUZ (03638X0018/P)	05/11/1974	07/02/2004	1379
BSS001AAMB (03628X0038/P)	28/10/1974	04/01/2007	1709
BSS001AHRM (03646X0083/F)	17/01/1997	31/08/2006	534

Appendix 3.C

The graphs depict the fluctuations in groundwater depth over time for the five main stations in the case study area. (Retrieved data from ades.fr). Their annual average of spatially mapped in GISpro for the groundwater level.

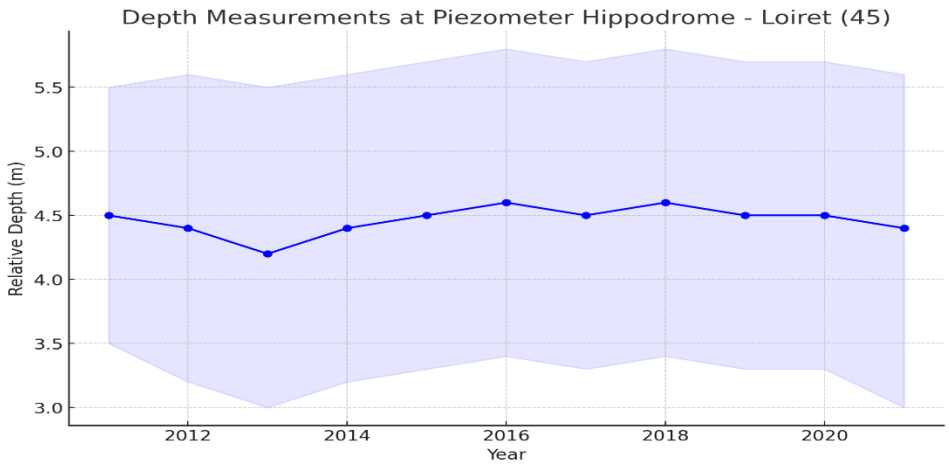


Figure 3.C.1 The relative depth measurements at the Piezometer Hippodrome (03635X0545/PZ1) - Loiret (45) from 2011 to 2021.

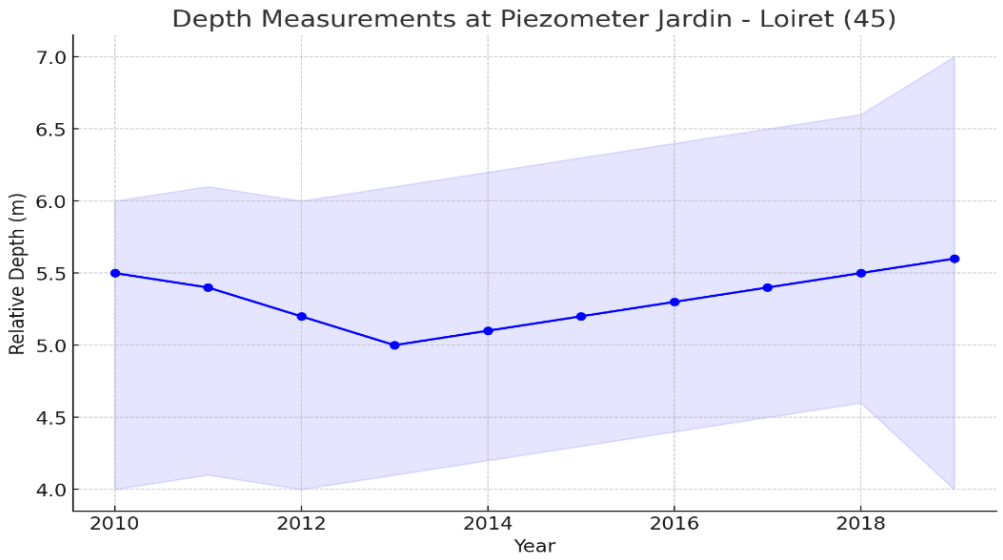


Figure 3.B.2 The relative depth measurements at the Piezometer Jardin (03636X1060/PZ2) - Loiret (45) from 2010 to 2019.

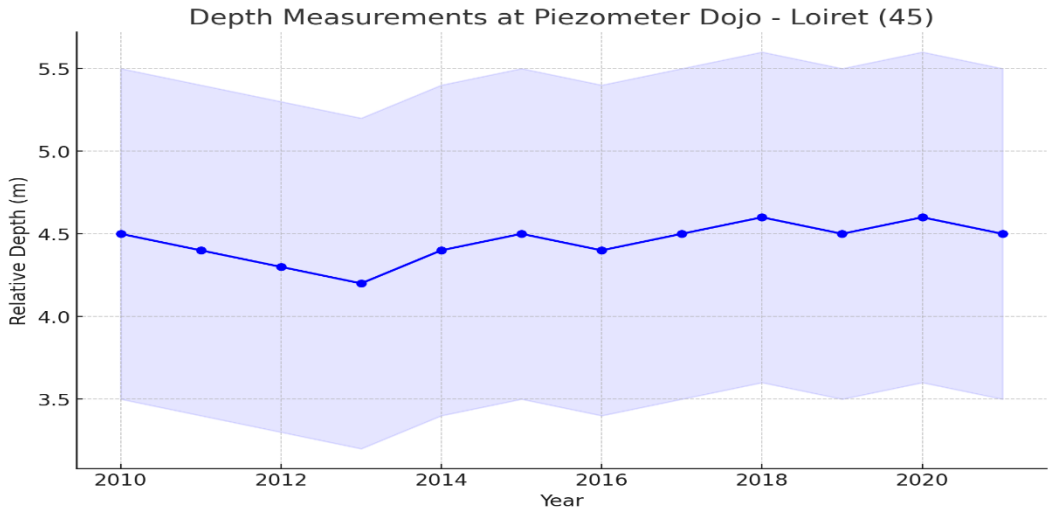


Figure **3.C.3** The depth measurements at the Piezometer Dojo - Loiret (45) from 2010 to 2021.

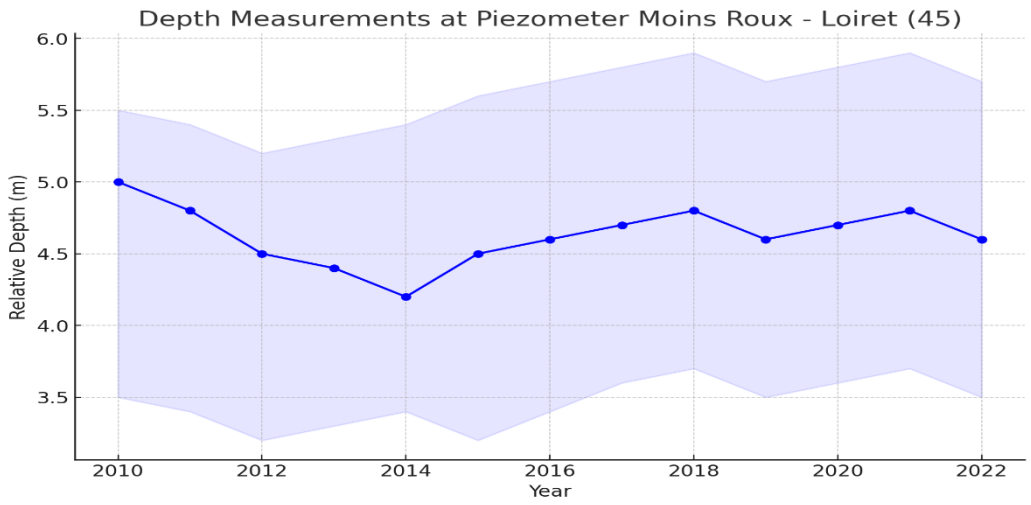


Figure 3.C.4 The depth measurements at the Piezometer Moins Roux - Loiret (45) from 2010 to 2022.

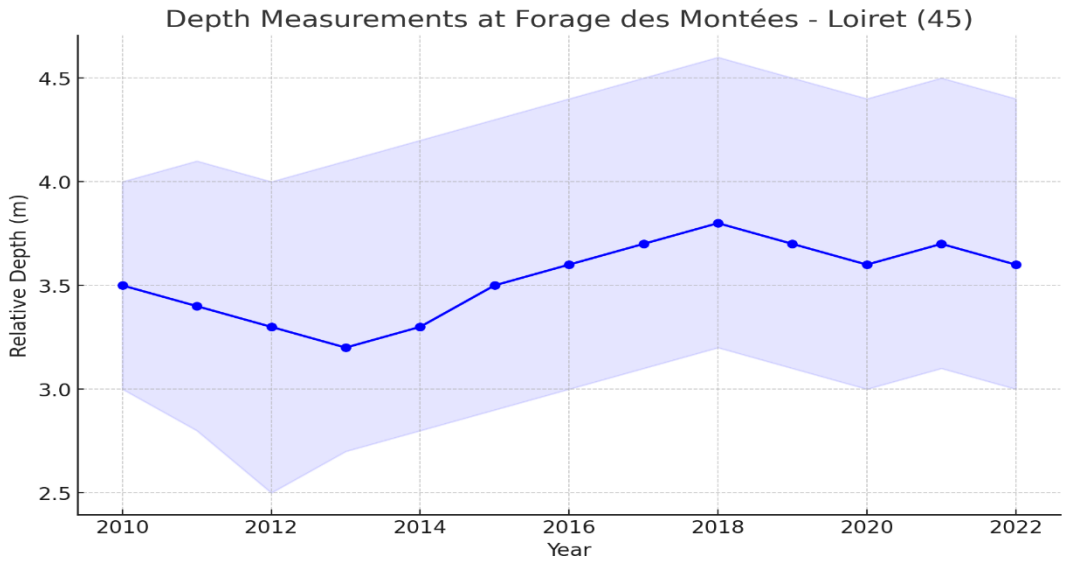


Figure 3.C.5 The depth measurements at the Piezometer Forage des Montées - Loiret (45) from 2010 to 2022.

Appendix 3.D R codes (scripts) for data processing to comply with MaxEnt requirements

This appendix includes critical and representative examples of each step in the data processing workflow. The provided R codes and scripts demonstrate key operations, with screenshots of the code and outputs to illustrate the processes. Below are the details for each step:

(1) library_loading_setup.R:

 # Load necessary libraries for spatial data processing and analysis

```{r}

library (**raster**)

library (**rgdal**)

library (**tidyverse**)

library (**sp**)

library (**sf**)

"**raster**" is used for reading, writing, manipulating, analyzing, and modeling of gridded spatial data.

library (raster)

"**rgdal**" provides bindings to the Geospatial Data Abstraction Library (GDAL) for reading and writing vector data.

Library (rgdal)

"tidyverse" is a collection of R packages designed for data science, making it easy to work with data.

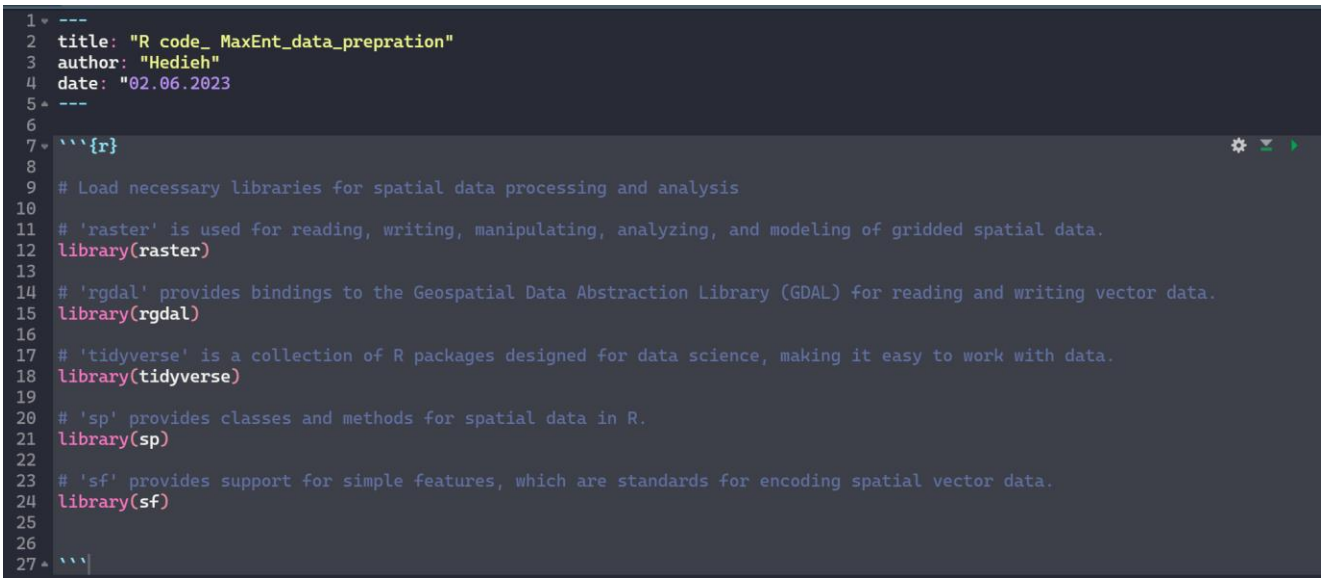
library(tidyverse)

“**sp**” provides classes and methods for spatial data in R.

library(sp)

“**sf** ” provides support for simple features, which are standards for encoding spatial vector data.

library(sf)



```
---
title: "R code_ MaxEnt_data_prepration"
author: "Hedieh"
date: "02.06.2023
---

```{r}

# Load necessary libraries for spatial data processing and analysis

# 'raster' is used for reading, writing, manipulating, analyzing, and modeling of gridded spatial data.
library(raster)

# 'rgdal' provides bindings to the Geospatial Data Abstraction Library (GDAL) for reading and writing vector data.
library(rgdal)

# 'tidyverse' is a collection of R packages designed for data science, making it easy to work with data.
library(tidyverse)

# 'sp' provides classes and methods for spatial data in R.
library(sp)

# 'sf' provides support for simple features, which are standards for encoding spatial vector data.
library(sf)


```
```

Figure 3.D.1 A screenshot of the script includes the loading of necessary libraries for spatial data processing and analysis.

(2) data_import_preparation.R:

 # Load raster data

EXT_DEM_31 ← raster ("C:/Users/hhsol/Desktop/MH-Val/asciiFiles/Elevation.asc")

 # Load Vector data

EXT_Study_boundary ←

readOGR("C:/Users/hhsol/Desktop/thesis/GISthesis/TD3/Com/Com31VdO/Com31_VdO.shp")

 # Set coordinate reference systems

crs(raster_ Study_boundary) <- CRS('+init=EPSG:2154')

crs(EXT_DEM_31) <- CRS('+init=EPSG:2154')

 # Set the extent and resolution of all rasters to be the same as the first raster

raster_reff <- raster(xmn = xmin(EXT_CVal), xmx = xmax(EXT_CVal),

ymn = ymin(EXT_CVal), ymx = ymax(EXT_CVal),

res = 10)

```
13
14  #New try with different extend
15
16  EXT_DEM_31 ← raster ("C:/Users/hhsol/Desktop/MH-Val/asciiFiles/Elevation.asc")
17
18  EXT_CVal← readOGR("C:/Users/hhsol/Desktop/thesis/GIS thesis/TD3/Com/Com31VdO/Com31_VdO.shp")
19
20  plot(EXT_DEM_31)
21
22  Flood← readOGR ("C:/Users/hhsol/Desktop/thesis/GIS thesis/TD3/Aléa_I/A_I_Simpl.shp")
23  GeoVal← readOGR ("C:/Users/hhsol/Desktop/MH-Val/Clip_Geo_Val/GeoVal.shp")
24
25  Thickness ← readOGR ("C:/Users/hhsol/Desktop/Data_collection_H/onlyVldO/thickVal.shp")
26  plot(Thickness)
27
28  GW ← raster("C:/Users/hhsol/Desktop/MH-Val/asciiS/Groundwater.asc")
29  Elevation ← raster("C:/Users/hhsol/Desktop/Data_collection_H/Rainfall/Topo/cilpped_DEM_31.tif")
30  Slope ← raster("C:/Users/hhsol/OneDrive/Documents/ArcGIS/Projects/ARcSCene/Slope_cilppe1.tif")
31
32  floodRaster ← raster("C:/Users/hhsol/Desktop/Data_collection_H/Rainfall/Topo/flood_raster.tif")
33  plot(floodRaster)
34
35  floodRaster1 ← projectRaster(floodRaster, raster_refCval)
36
37
38  crs(raster_refCval) ← CRS('+init=EPSG:2154')
39  crs(floodRaster) ← CRS('+init=EPSG:2154')
40
41  raster_reff ← raster(xmn = xmin(EXT_CVal), xmx = xmax(EXT_CVal),
42                       ymn = ymin(EXT_CVal), ymx = ymax(EXT_CVal),
43                       res = 10)
```

Figure 3.D.2 A screenshot of the R script showing the preparation of spatial data for MaxEnt analysis.

 # Exporting_rasters_in_ASCII_files

writeRaster(EXT_DEM_31,"C:/Users/hhsol/Desktop/MHVal/asciiFiles/writeRasterF1.asc", overwrite = T)

```
```{r}
writeRaster(GeoVal_1_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval1.asc", overwrite = T)
writeRaster(GeoVal_2_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval2.asc", overwrite = T)
writeRaster(GeoVal_3_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval3.asc", overwrite = T)
writeRaster(GeoVal_4_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval4.asc", overwrite = T)
writeRaster(GeoVal_8_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval8.asc", overwrite = T)
writeRaster(GeoVal_10_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval10.asc", overwrite = T)
writeRaster(GeoVal_11_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval11.asc", overwrite = T)
writeRaster(GeoVal_21_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval21.asc", overwrite = T)
writeRaster(GeoVal_24_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval24.asc", overwrite = T)
writeRaster(GeoVal_25_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval25.asc", overwrite = T)
writeRaster(GeoVal_26_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval26.asc", overwrite = T)
writeRaster(GeoVal_27_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval27.asc", overwrite = T)
writeRaster(GeoVal_28_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval28.asc", overwrite = T)
writeRaster(GeoVal_47_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterGeoval47.asc", overwrite = T)

writeRaster(thick1_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterthickness_5-10.asc", overwrite = T)
writeRaster(thick2_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterthickness_10-13.asc", overwrite = T)
writeRaster(thick3_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterthickness_5.asc", overwrite = T)

writeRaster(g3diss_ras,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/writeRasterg3.asc", overwrite = T)

writeRaster(Elevation,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/Elevation.asc", overwrite = T)

writeRaster(Slope,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/Slope.asc", overwrite = T)


writeRaster(floodRaster1,"C:/Users/hhsol/Desktop/MH-Val/asciiFiles/Flood_hazard.asc", overwrite = T)
```
```

Figure 3.D.3 A screenshot of the R script showing exporting prepared raster data in ASCII format.

(3) Correlation_analysis.R

 # Load necessary libraries for correlation analysis

library(carData) # Provides datasets for regression analysis

library(car) # Companion to Applied Regression, provides functions for regression analysis

 # Load the raster data

Elevation ←raster("C:/Users/hhsol/Desktop/NewExperimentswith Maxent/asciis_val_corrected/Elevation.asc")

 # Convert the raster data to vectors

Elevation_vec ← getValues(elevation)

 # Calculate Pearson correlation coefficient

correlation_result <- cor(elevation_vec, slope_vec, method = "pearson")

 # Print the result

print(correlation_result)

```{r}

```r
```{r}
# Load necessary library
library(raster)
library(carData)
library(car)
library(tidyverse)

# Load the data
Elevation ← raster("C:/Users/hhsol/Desktop/New Experiments with Maxent/asciis_val_corrected/Elevation.asc")
Slope ← raster("C:/Users/hhsol/Desktop/New Experiments with Maxent/asciis_val_corrected/Slope.asc")
Flood ← raster("C:/Users/hhsol/Desktop/New Experiments with Maxent/asciis_val_corrected/Flood hazard zone.asc")
Geology ← raster("C:/Users/hhsol/Desktop/New Experiments with Maxent/asciis_val_corrected/Geology.asc")
Groundwater ← raster("C:/Users/hhsol/Desktop/New Experiments with Maxent/asciis_val_corrected/Groundwater.asc")
Thickness ← raster("C:/Users/hhsol/Desktop/New Experiments with Maxent/asciis_val_corrected/Alluvial thickness.asc")

# Convert the raster data to vectors
Elevation_vec ← getValues(elevation)
Slope_vec ← getValues(slope)
Flood_vec ← getValues(f)
Geology_vec ← getValues(g)
Groundwater_vec ← getValues(gw)
Thickness_vec ← getValues(th)


data_df ← data.frame(Elevation = Elevation_vec, Slope = Slope_vec, Flood = Flood_vec , Groundwater = Groundwater_vec, Thickness =
Thickness_vec)

correlation_matrix ← cor(data_df, method = "pearson", use = "complete.obs")


# Calculate Pearson correlation coefficient
correlation_result ← cor(elevation_vec, slope_vec, method = "pearson")

# Print the result
print(correlation_result)

print(correlation_matrix)

```
```

```
            elevation       slope           f           g          gw          th
elevation  1.00000000 -0.05350155 -0.30080644 -0.14706828 -0.46701300  0.18641947
slope     -0.05350155  1.00000000  0.21163699  0.02625806  0.12051301 -0.02330127
f         -0.30080644  0.21163699  1.00000000  0.35027043  0.23749969 -0.03742922
g         -0.14706828  0.02625806  0.35027043  1.00000000  0.10952166 -0.03380872
gw        -0.46701300  0.12051301  0.23749969  0.10952166  1.00000000  0.07447209
th         0.18641947 -0.02330127 -0.03742922 -0.03380872  0.07447209  1.00000000
```

Figure 3.D.3 A screenshot of the R script of the correlation analysis and the results.

model_fitting_and_VIF.R

 # Correct the vector conversion lines

Elevation_vec <- getValues(Elevation)

Slope_vec <- getValues(Slope)

Flood_vec <- getValues(Flood)

Geology_vec <- getValues(Geology)

Groundwater_vec <- getValues(Groundwater)

Thickness_vec <- getValues(Thickness)

 # Correct the data frame creation line

data_df <- data.frame(Elevation = Elevation_vec, Slope = Slope_vec, Flood = Flood_vec, Geology = Geology_vec, Groundwater = Groundwater_vec, Thickness = Thickness_vec)

 # Now, retry fitting the model

model <- lm(Thickness ~ Elevation + Slope + Flood + Geology + Groundwater, data = data_df)

 # Calculate Variance Inflation Factor (VIF) to check for multi-collinearity

vif_values <- vif(model)

 # Print the VIF values

print(vif_values)

```

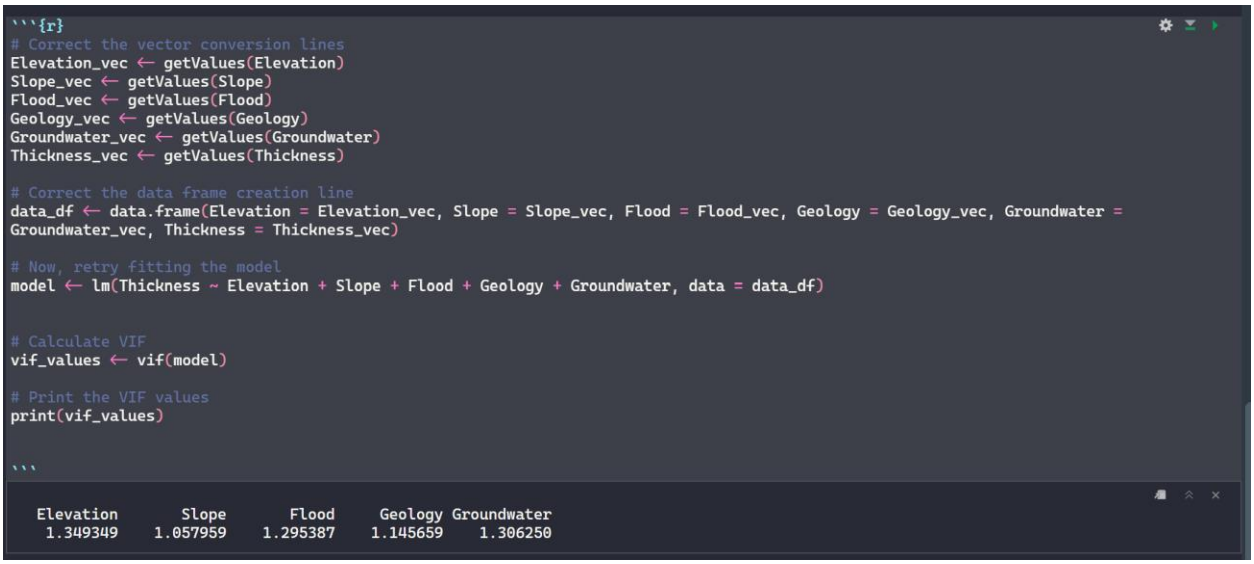



**Figure 3.D.4** A screenshot of the R script of the correlation analysis and the results.
```

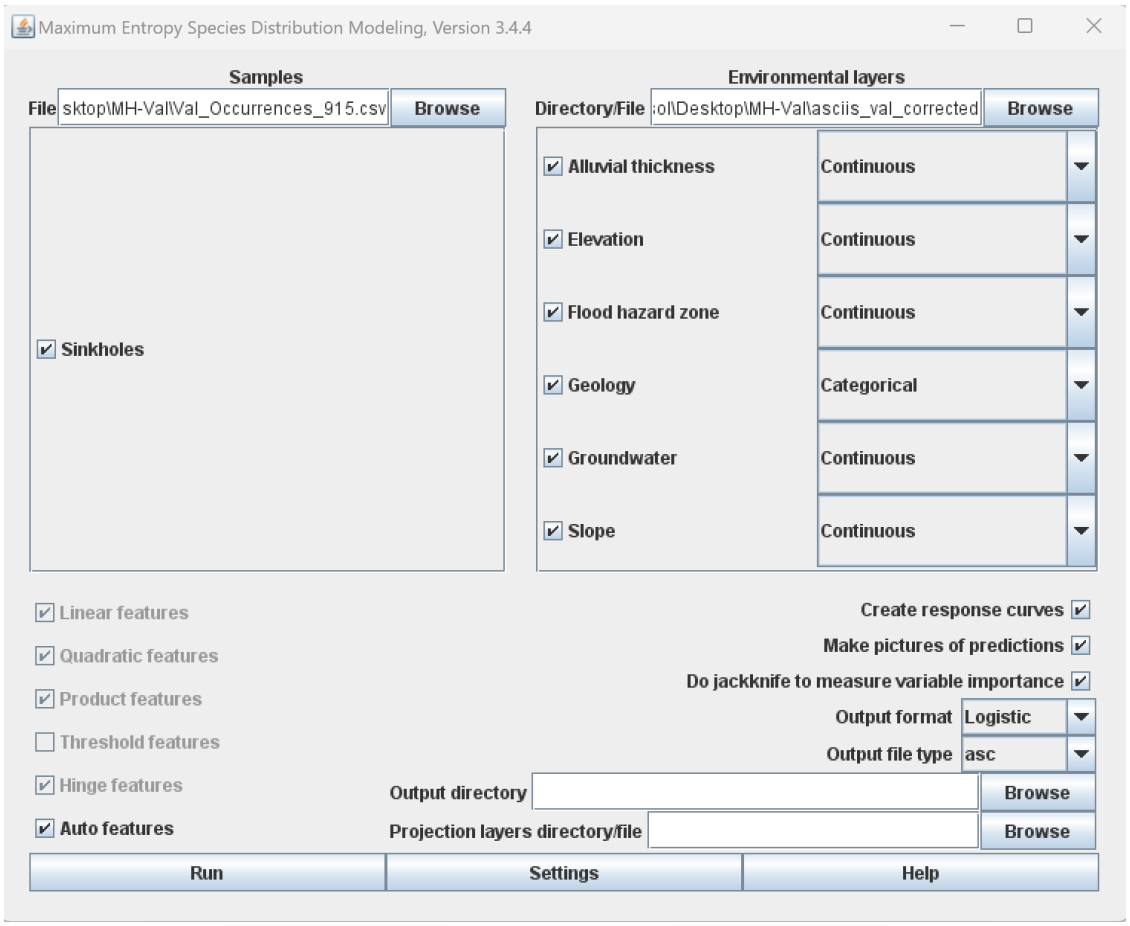


Figure 3.D.5 A screenshot of MaxEnt modelling software.

Appendix 4.A

The close-up maps' segments in comparing flood vs flood-triggered sinkhole maps

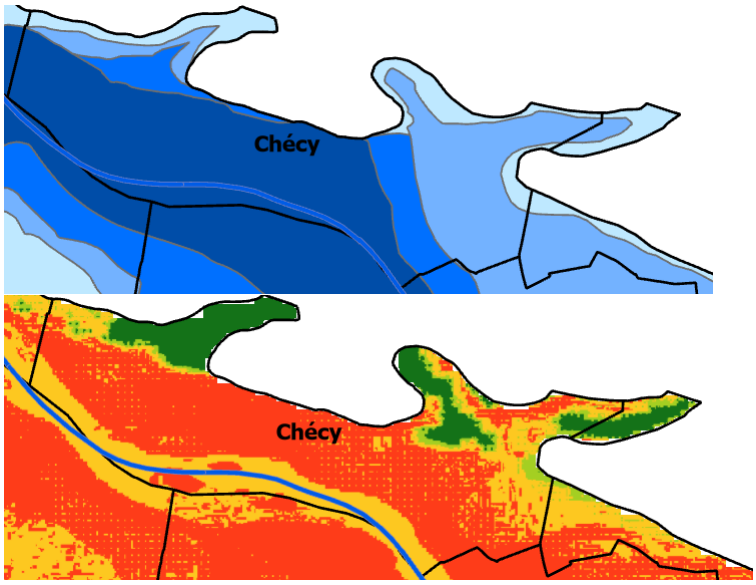


Figure 4.A.1. This figure presents a close-up view of *Chécy*, extracted from the regional single and multi-hazard and maps, to illustrate the spatial distribution and severity of flood hazards (top panel) and flood-triggered sinkhole susceptibility (bottom panel). **Top panel:** The flood hazard map categorizes the area into zones of varying risk levels, with darker shades of blue indicating higher water heights and greater flood hazard. Chécy is predominantly in a high to very high flood hazard zone. **Bottom panel:** It shows that Chécy, while predominantly under high risk from floods, also faces a considerable risk from sinkholes, especially in areas highlighted by darker red shades.

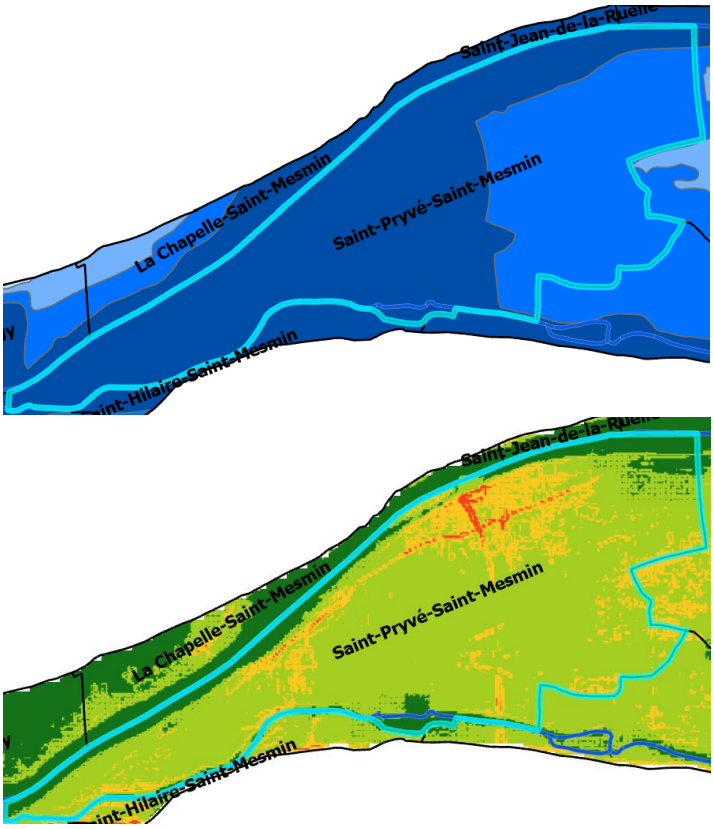


Figure 4.A.2 A close-up view of the Saint-Pryvé-Saint-Mesmin area, exctracting from the broader regional maps to focus on disparency of single vs multi-hazard levels. **Top panel:** Displays the flood hazard with High and Very high level (two shades of dark blue). **Botton panel:** Shows the flood-triggered sinkholes with a mixed level of susceptibility, predominantly at a moderate level (light green).

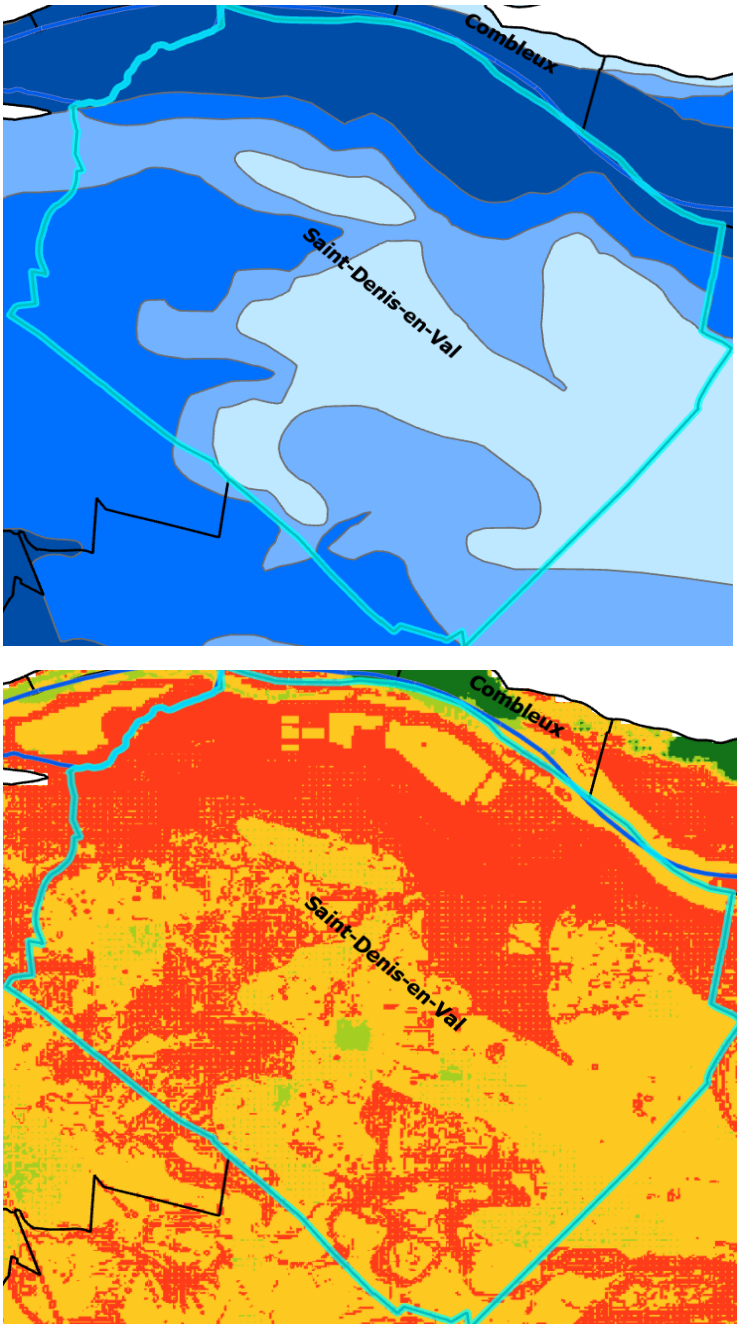


Figure 4.A.3 A close-up view of the Saint-Denis-en-Val area, exctracting from the broader regional maps to focus on disparency of single vs multi-hazard levels. **Top panel**: Illustrates the flood hazard levels marked by varing flood hazard level dominately at low and moderate levels. **Bottom panel:** Shows that there is a notable presence of High to Very high Flood-triggered sinkhole susceptibility, particularly in areas not necessarily consistent with the highest flood risk zones.

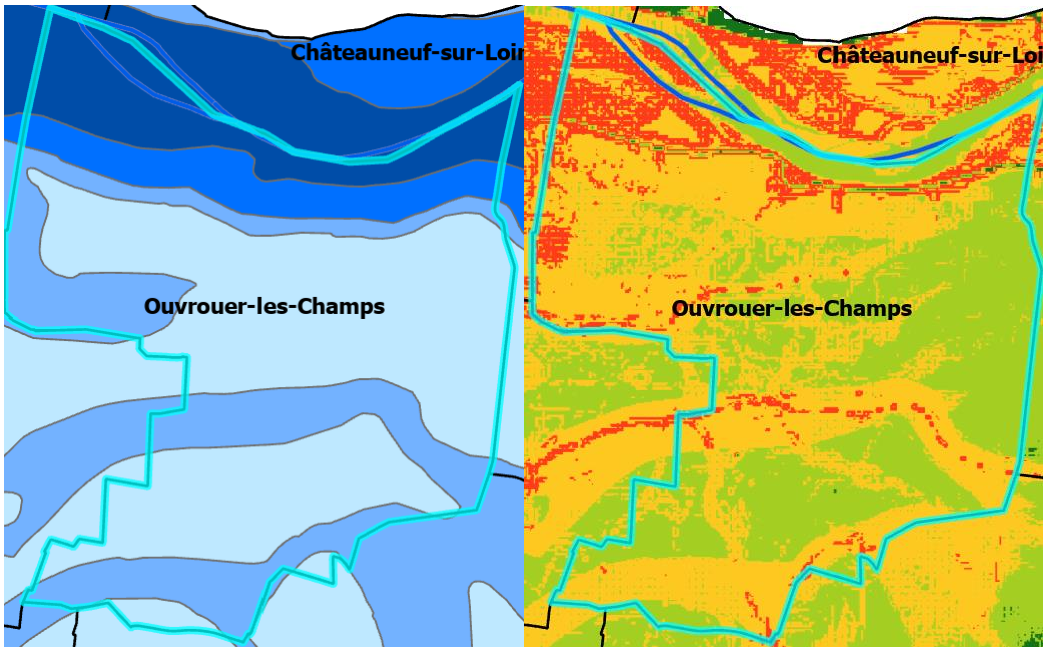


Figure 4.A.4 A close-up view of the Ouvrouer-les-Champs area, exctracting from the broader regional maps to focus on disparency of single vs multi-hazard levels. **The left figure:** Displays the flood hazard levels, indicating that the area is primarily categorized under low to moderate flood hazard. **The right figure:** Shows a mixture of moderate to very high susceptibility, indicating that while flood risks might be lower, the potential for other hazards, such as sinkholes, is significantly varied and can be severe in certain zones.

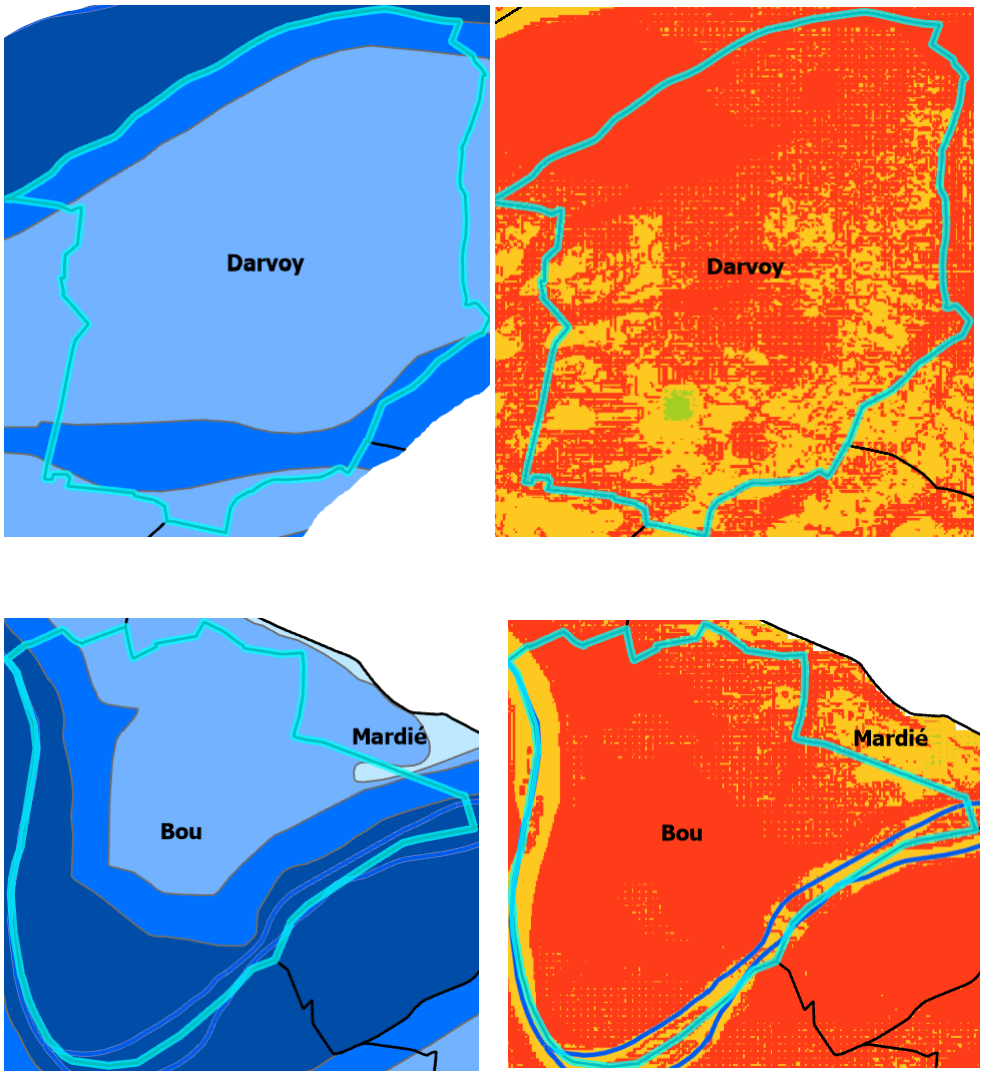


Figure 4.A.5 A close-up view of the Darvoy area (top panels) and the Bou area (bottom panels), extracted from the broader regional maps to focus on the disparity between single vs multi-hazard levels. **The top left panel** displays the Darvoy area predominantly exhibiting moderate flood hazard levels, while **the top right panel** demonstrates that almost the entire area falls into High and Very high multi-hazard level. **The bottom left panel** shows the Bou area with varying in flood hazard levels, with a predominance of moderate hazard while **the bottom right panel** reveals that almost the entire Bou area is under Very high level of multi-hazard.

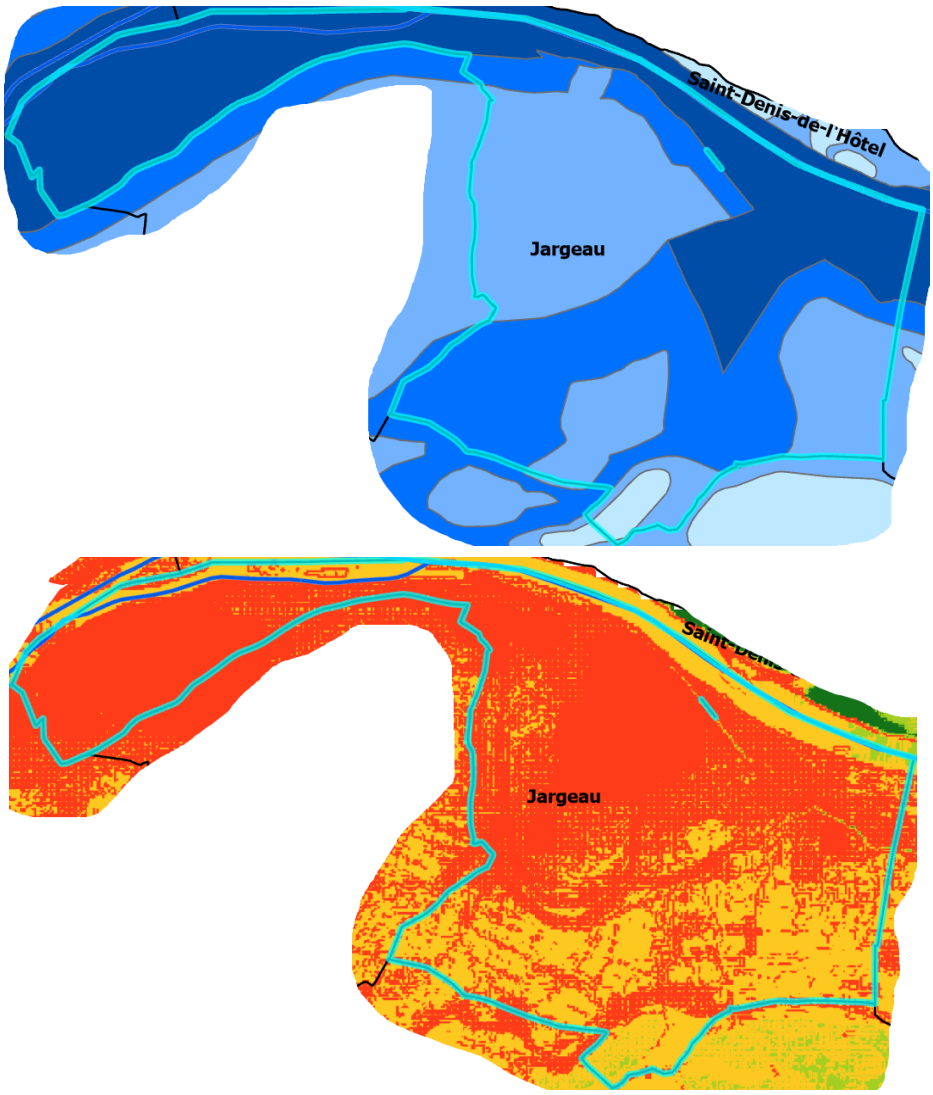


Figure 4.A.6 A close-up view of the Jargeau area, extracted from the broader regional maps to focus on the disparity between single vs multi-hazard levels. The top panel depicts that most parts of *Jargeau* to the north and west are moderately at risk of flooding while **the bottom panel** reveals that the whole area ranks High to Very high for combined flood and sinkhole susceptibility.

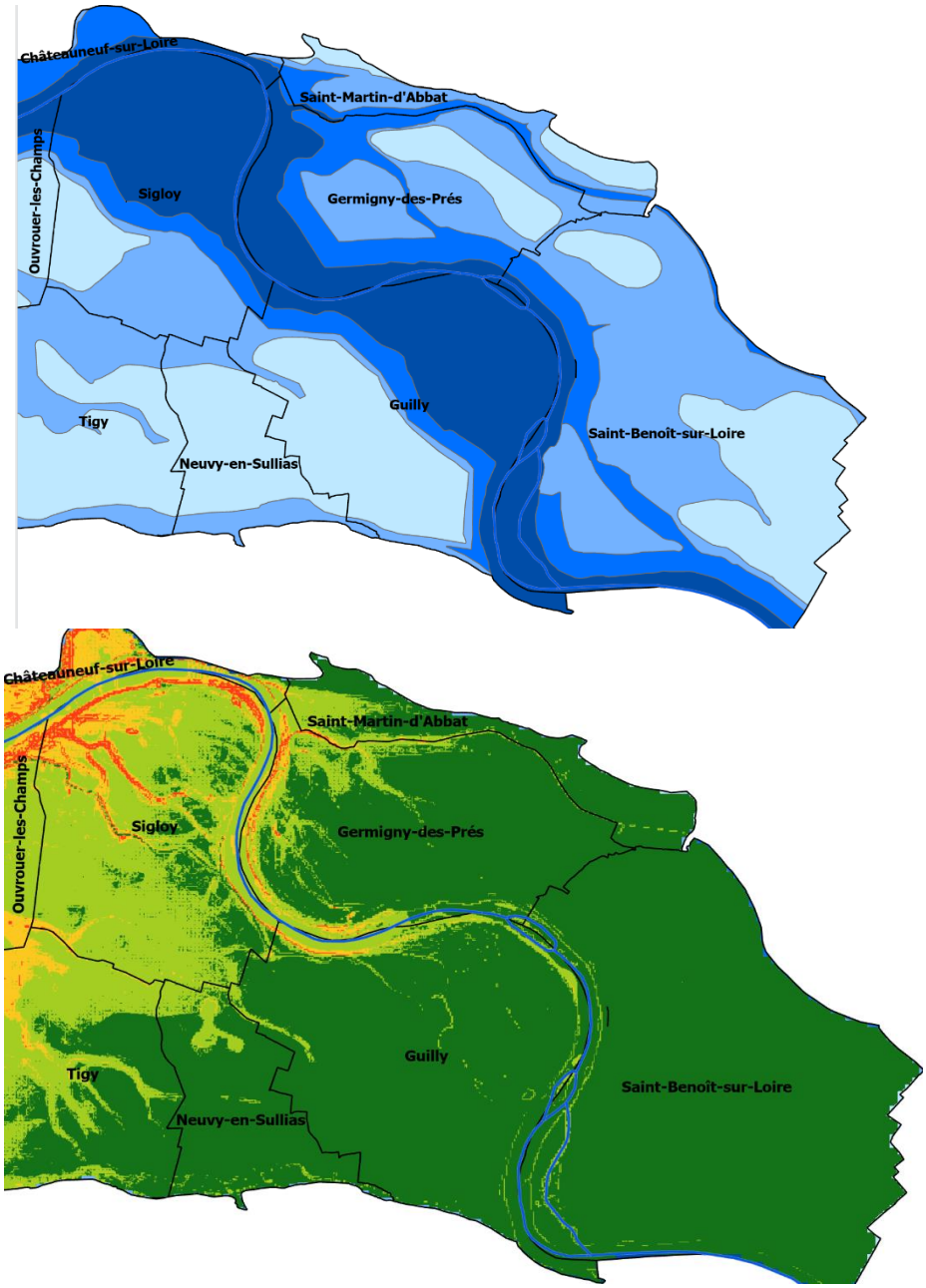


Figure 4.A.7 A close-up view of the Sigloy, *Saint-Martin-d'Abbat*, *Germingy-des-Prés*, *Guilly*, *Neuvy-en-Sullias* and *Saint-Benoit-sur-Loire*, extracted from the broader regional maps to focus on the disparity between single vs multi-hazard levels. **The top panel** shows that these areas are situated in the east part of the region with a mixture flood hazard level from Low to Very high (a more attention to the meanders),

while **the bottom panel** shows that they are situated mainly in Low susceptibility level of flood-triggered sinkholes.

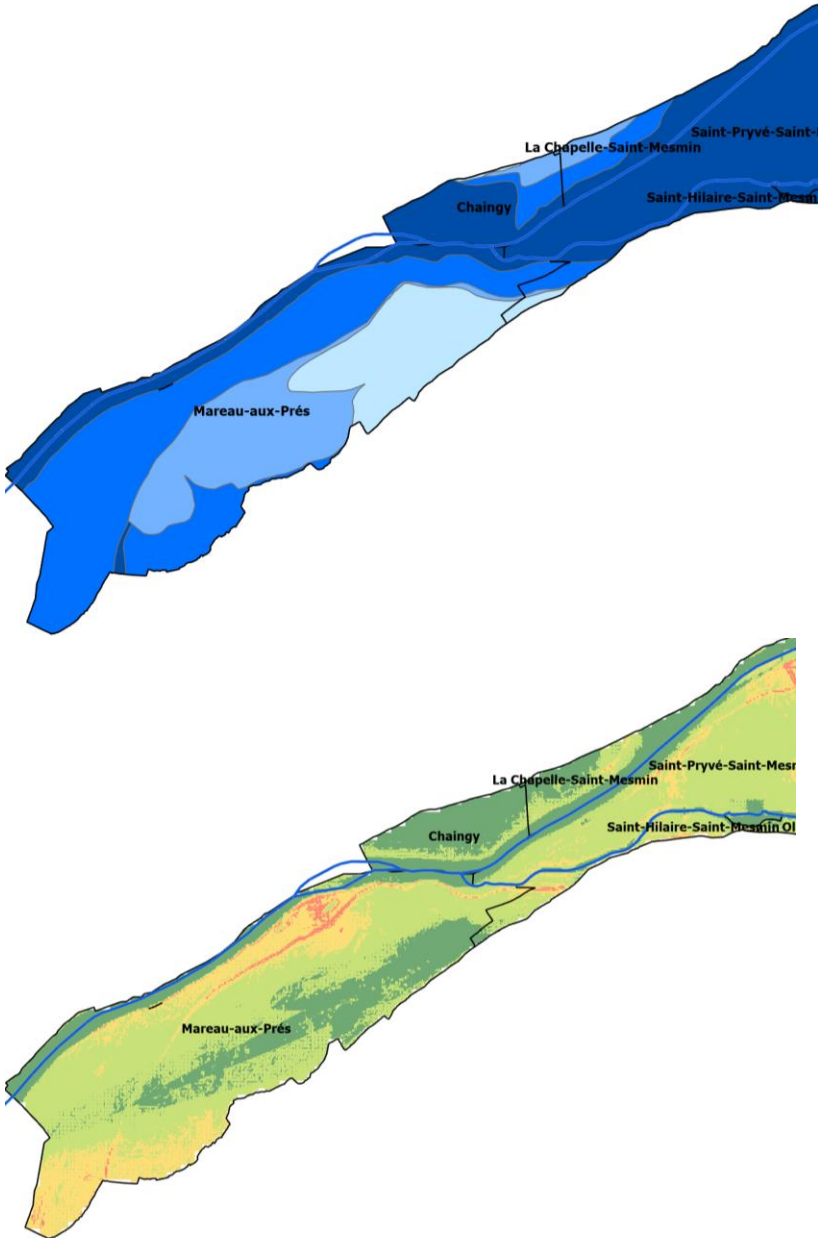


Figure 4.A.8 A close-up view of the Chaingy, *La Chapelle-Saint-Mesim, Mreau-aux-Prés* (in the far west of the area) extracted from the broader regional maps to focus on the disparity between single vs multi-hazard levels. The top panel shows that these areas are situated in the east part of the region with a mixture flood hazard level from Low to Very high (a more attention to the meanders), while the bottom panel shows that they are situated mainly in Low susceptibility level of flood-triggered sinkholes.

Appendix 4.B Insight from applying MaxEnt to a broader geographical Area (The region of Orléans with 31 communities around)

In extending the application of MaxEnt to the larger region of Orléans, including 31 communities, several critical insights were derived that have implications for the model's performance and applicability in spatial data mining, particularly in hazard assessment.

We accurately replicated all initial processes for this broader area to ensure a comprehensive evaluation of how study area dimensions influences MaxEnt AUC scores. In the following, the spatial rasterized maps of each contributing factors were produced. Then, all the same processes were done to obtain the model. The same contributing factors were used. However, broader area means that the information on each contributing factors (i.e., elevation, slope, etc.) will become accordingly different. This relationship is important, as the spatial extent of the study area directly impacts the environmental gradient and the diversity of habitats included in the model, which in turn affects the model's ability to accurately predict species distributions.

Impact of study area size on MaxEnt Performance:

AUC sensitivity to scale: The analysis revealed that MaxEnt's Area Under the Curve (AUC) scores are sensitive to the scale of the study area. Larger areas tend to produce higher AUC values, which might misleadingly suggest better model performance. This phenomenon occurs because larger areas introduce greater (geo) environmental gradients and diversity, which the model captures, thereby enhancing its predictive capabilities.

Homogeneity vs. variability: In smaller, more homogeneous areas like Val d'Orléans, the data tends to be less variable, making it more challenging to distinguish between influential factors. This can result in under-fitted models that do not generalize well beyond the training data. Conversely, larger areas with more variability allow MaxEnt to identify more distinct patterns and relationships, improving its performance.

Guidelines for study area definition:

Optimal spatial extent: It is crucial to define the study area based on scientific criteria, ensuring it encompasses the relevant environmental gradients without becoming excessively large. An unnecessarily large study area can inflate the AUC values and lead to overestimation of model accuracy. Hence, the study area should ideally be limited to regions where the phenomenon of interest is known to occur, such as focusing on the Amazon basin rather than the entire South America for certain species distributions.

Practical implications:

- **Urban Planning and zoning:** The findings from this broader application provide empirical evidence to inform urban planning, zoning laws, and building codes. Understanding how and why scale affects model performance can help in designing more accurate and reliable hazard maps is important for mitigating risks in urban development. and ease the communication of model results with stakeholders

- **Model complexity and assumptions:** The performance of MaxEnt is heavily influenced by the complexity of the model and the degree to which the data conforms to its assumptions. It is important to carefully select contributing factors and ensure that the study region is appropriately defined to avoid misleading results.

Limitations and Future Directions:

Scale-Dependent Performance: The study highlighted a limitation of MaxEnt in that its performance metrics, such as AUC, can be misleading if the scale is not appropriately managed. Future research should focus on establishing standardized criteria for defining study areas to ensure the robustness and accuracy of MaxEnt models.

Refining Contributing Factors: Further exploration is needed to refine the selection and weighting of contributing factors, particularly in more homogeneous areas, to improve the model's predictive power and reliability.

By applying MaxEnt to a broader geographical area, this study underscores the importance of carefully considering the scale of the study area in spatial data mining and hazard prediction. These insights are crucial for enhancing the practical application of MaxEnt and ensuring that its predictions are both accurate and actionable in real-world scenarios.

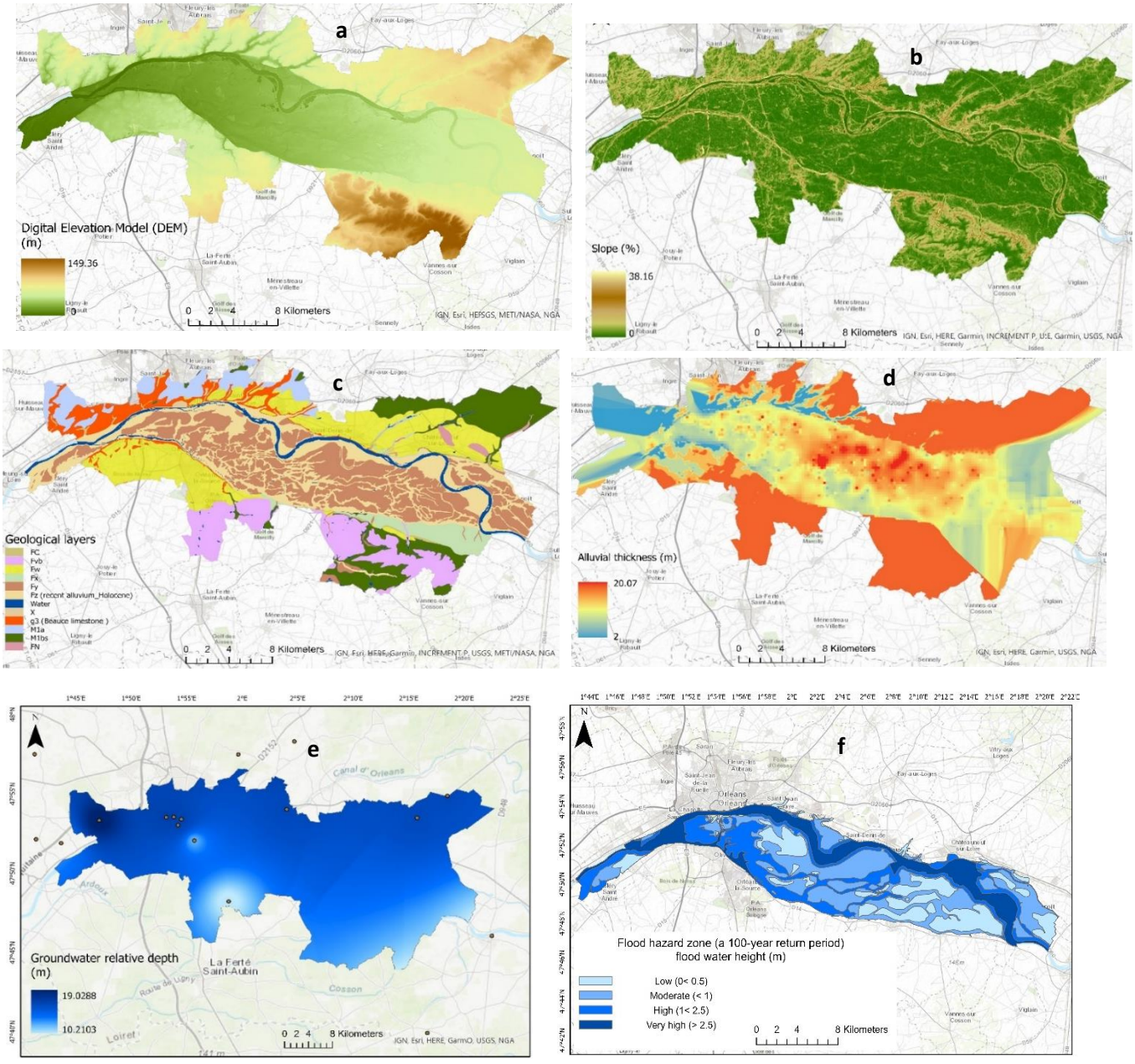


Figure 4.B.1 The six thematic maps of geo-environmental factors for assessing the spatial distribution of flood-triggered sinkholes over the extended area of Val d'Orléans. a) elevation, b) slope, c) surficial geology, d) Alluvial thickness, e) groundwater level and f) flood zones.

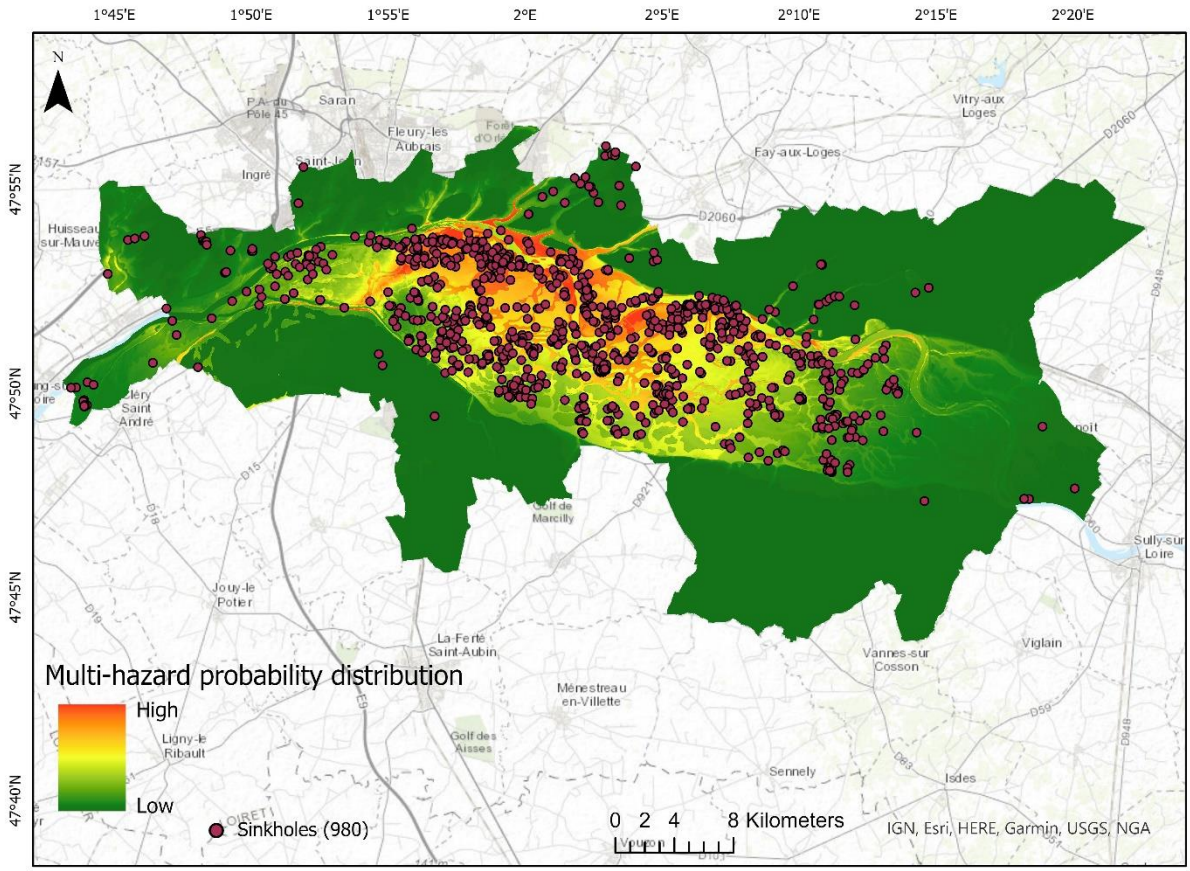


Figure 4.B.2 The multi-hazard susceptibility map generated by MaxEnt model using the natural breaks method for an extended area.

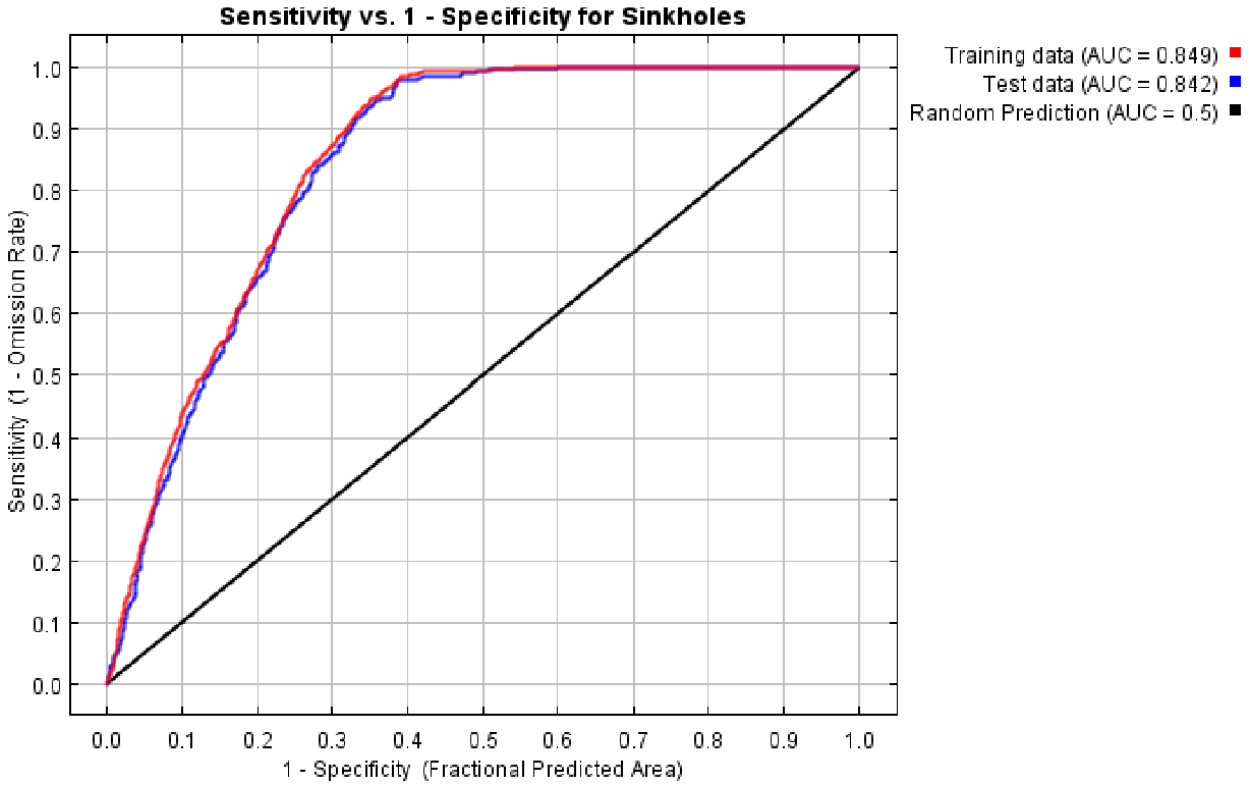


Figure 4.B.3 The Area Under the Curve (AUC) results of the Receiver Operating Characteristic (ROC) analysis applied to both training and testing data.

Appendix 4.C The resultant Model-Specific curves and Marginal curves for each (geo)-environmental variable (Side by side comparison)

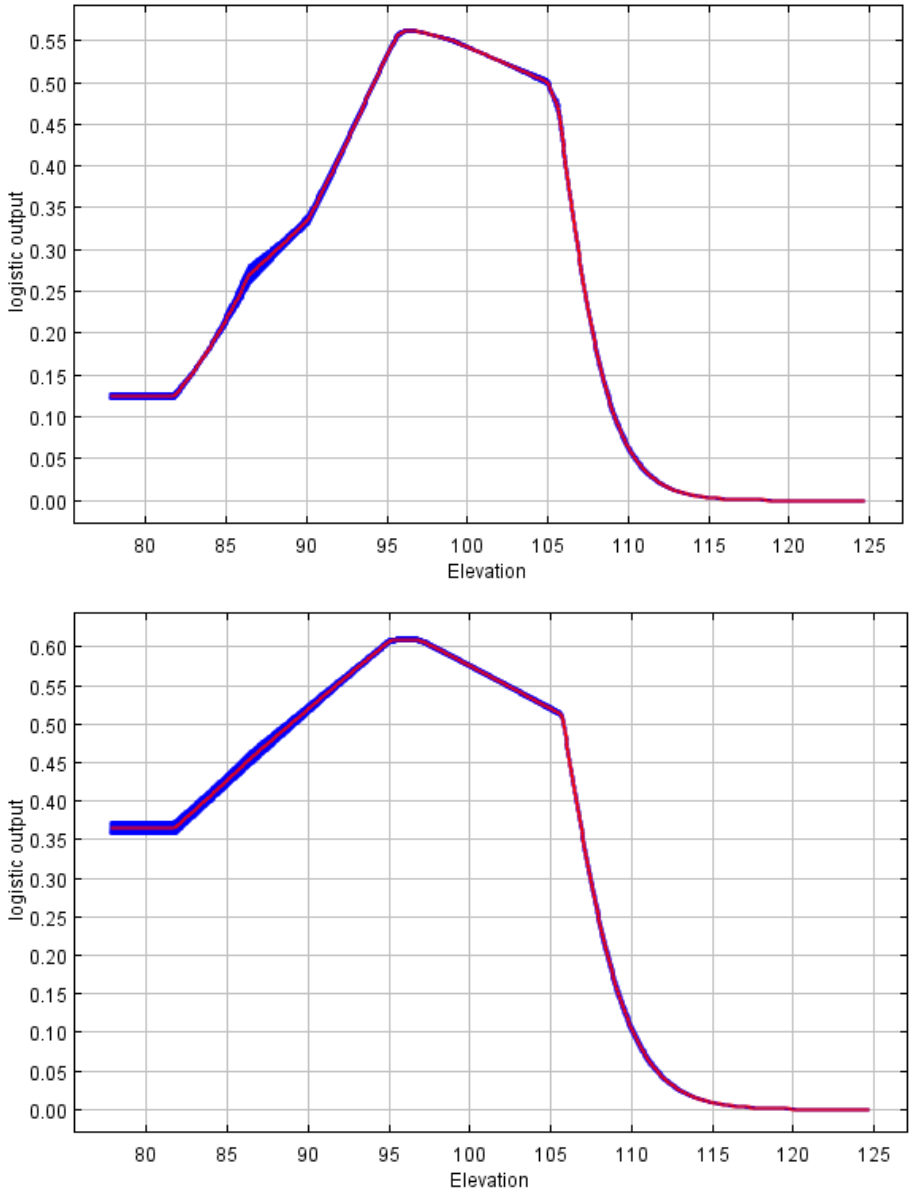


Figure 4.C.1 A side-by-side comparison of model specific (top) and marginal (bottom) curves for elevation (m).

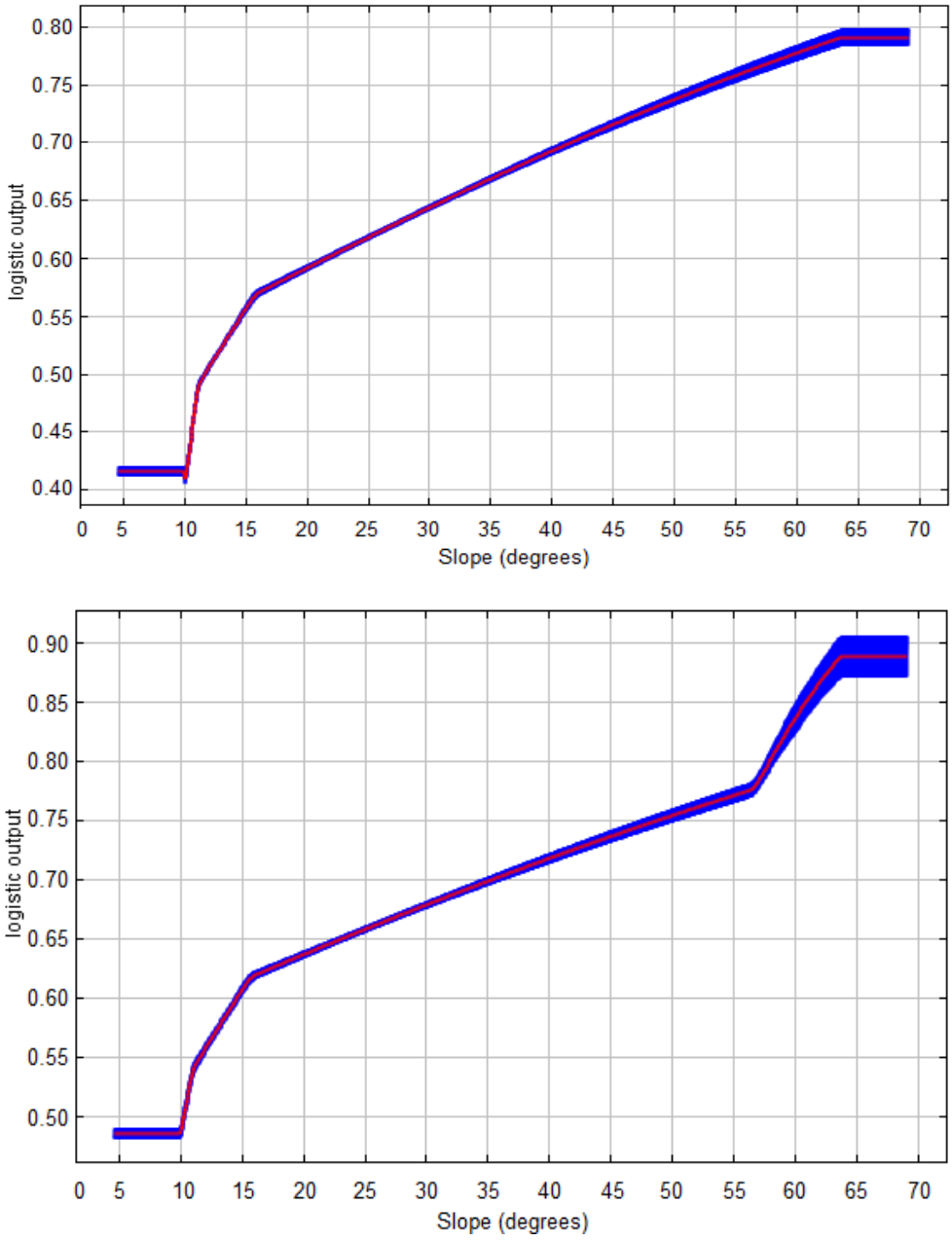


Figure 4.C.2 A side-by-side comparison of model specific (top) and marginal (bottom) curves for Slope (degrees).

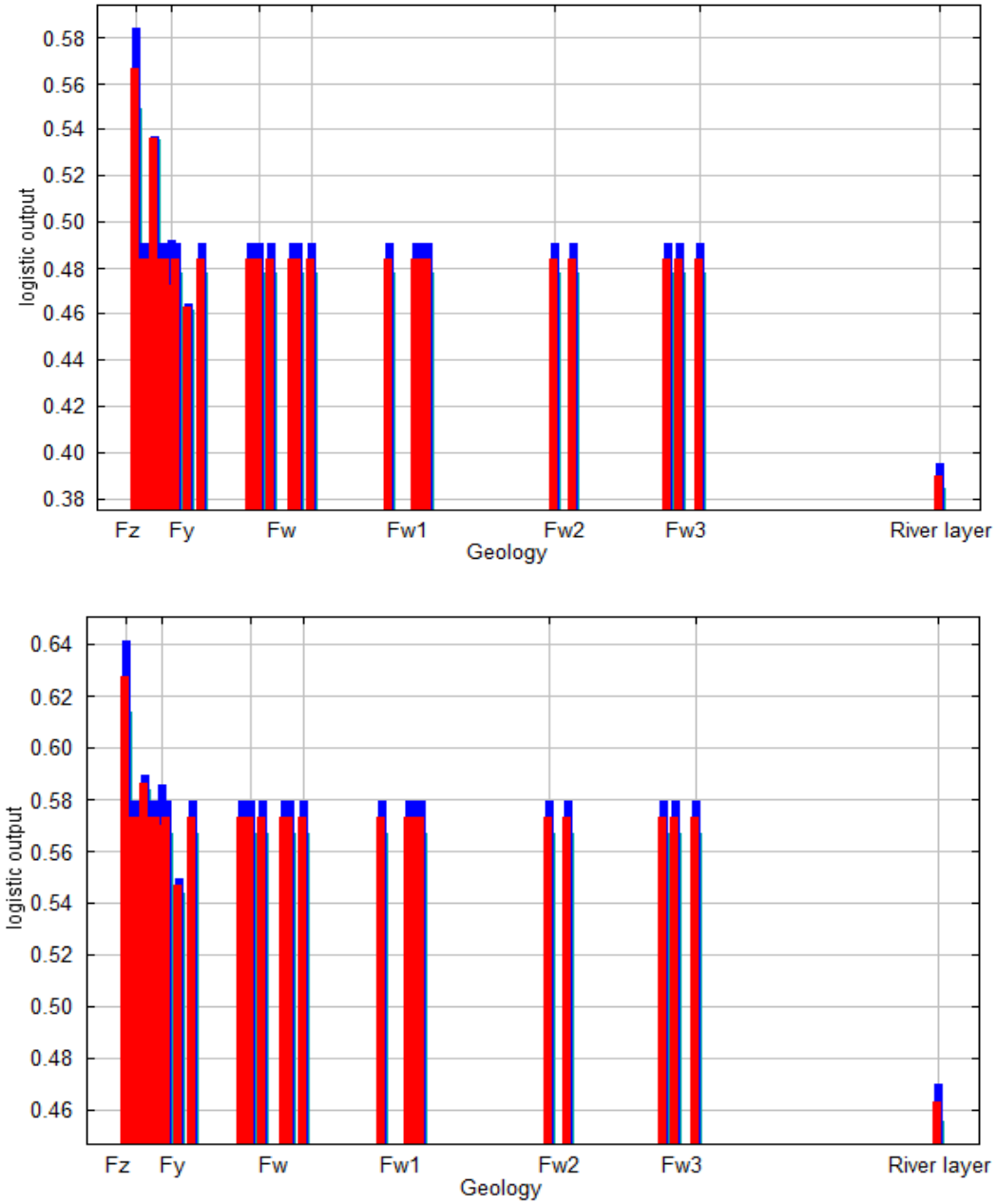


Figure 4.C.3 A side-by-side comparison of model specific (top) and marginal (bottom) curves for Surficial geology.

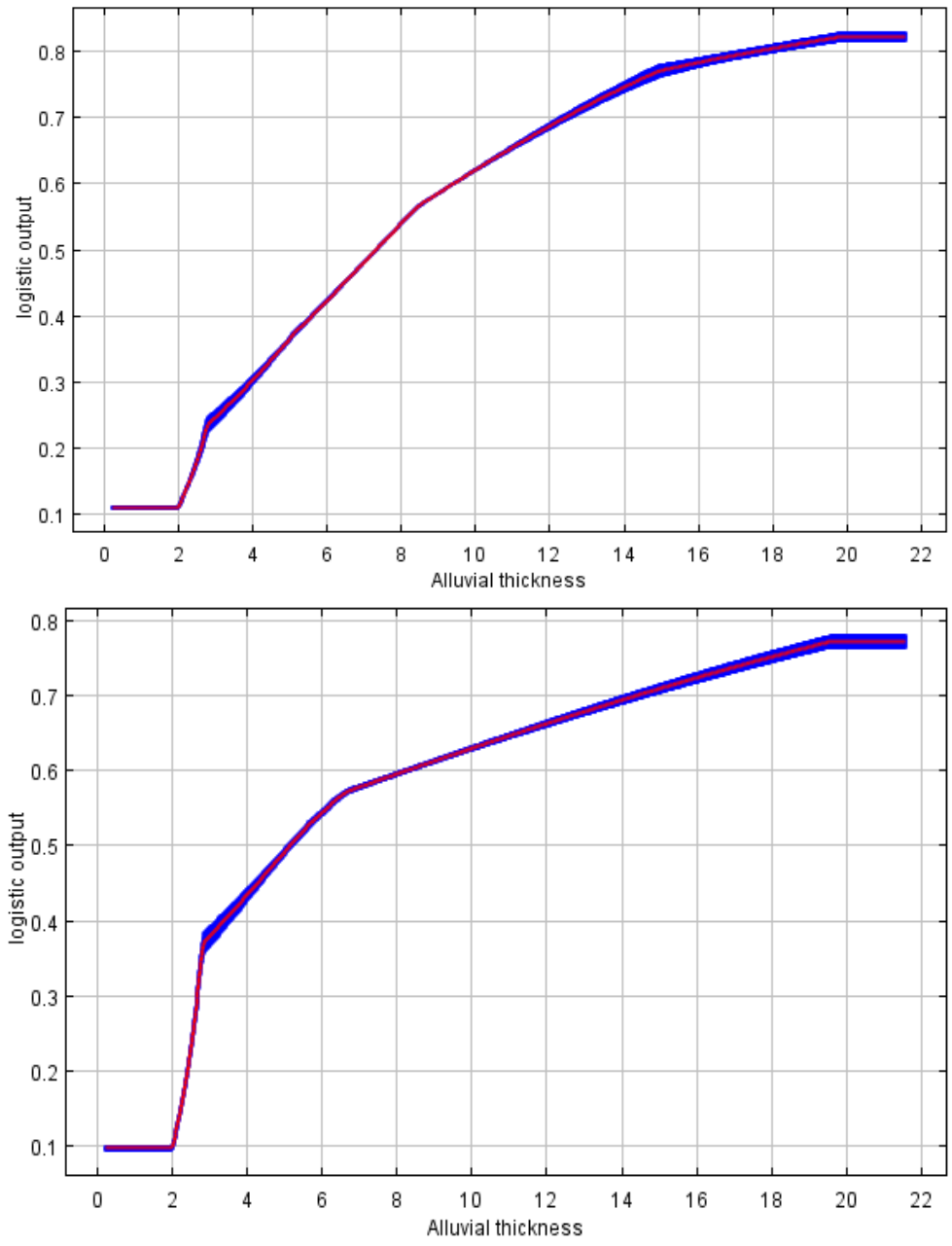


Figure 4.C.4 A side-by-side comparison of model specific (top) and marginal (bottom) curves for Alluvial thickness (m).

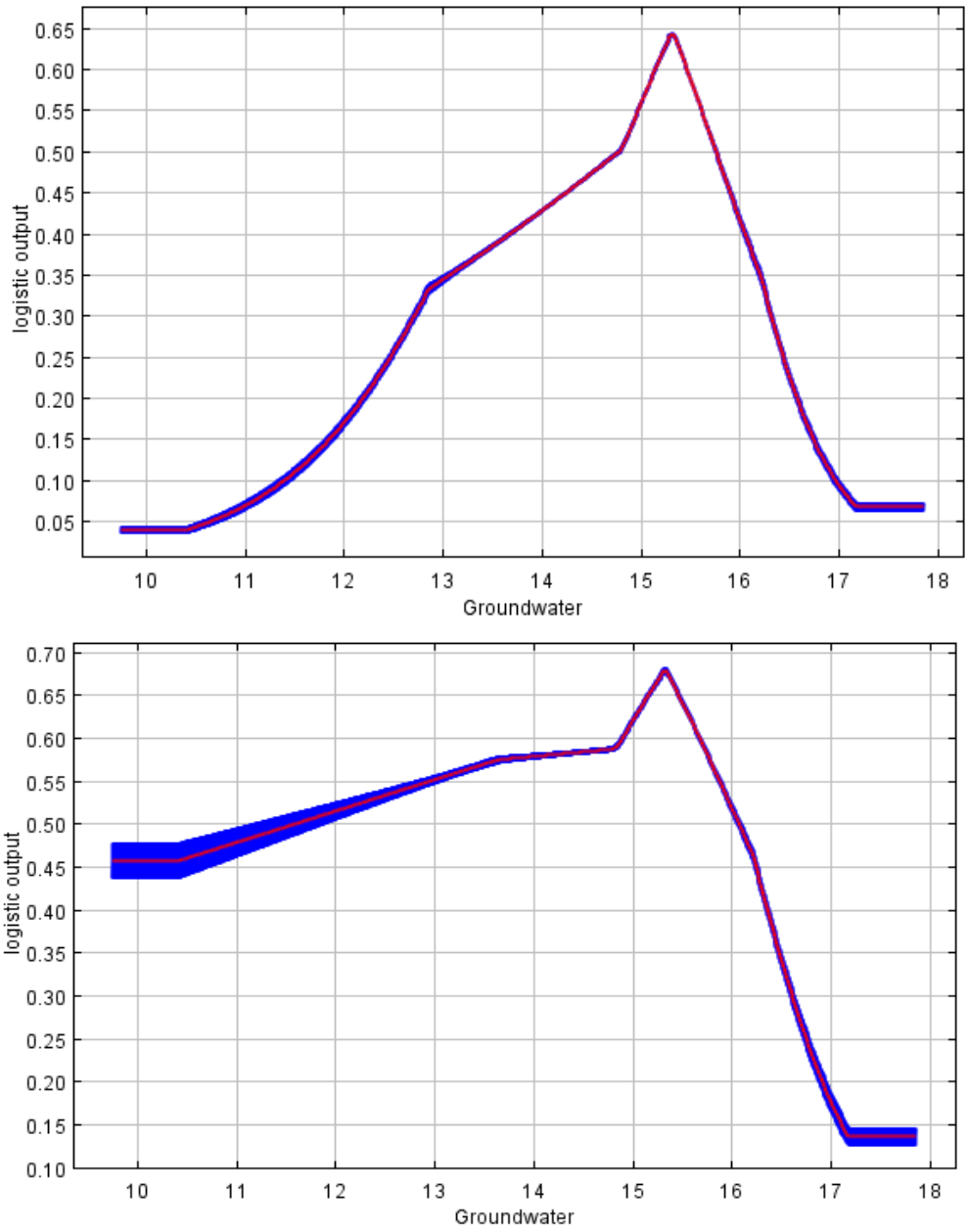


Figure 4.C.5 A side-by-side comparison of model specific (top) and marginal (bottom) curves for groundwater level (m).

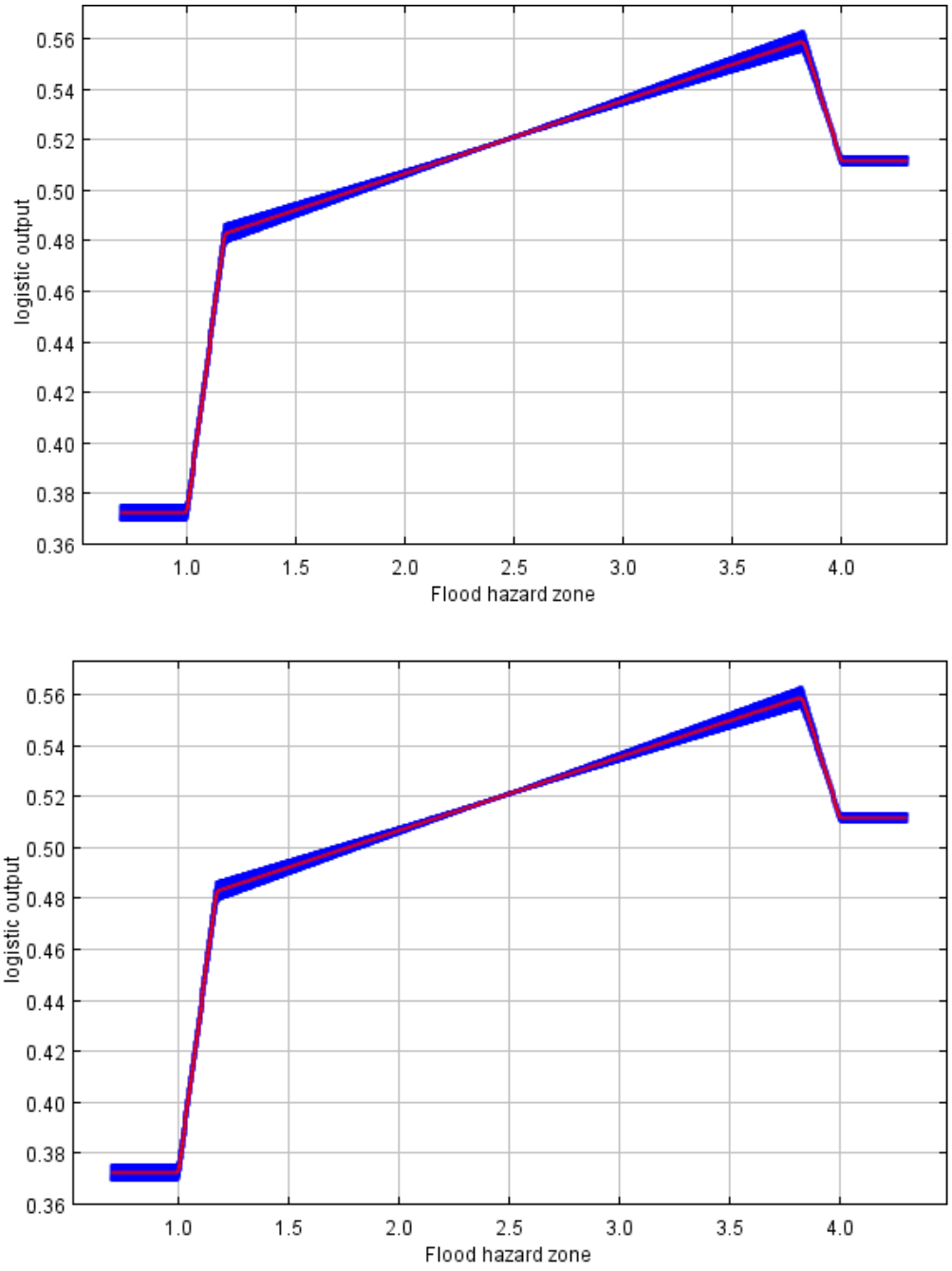


Figure 4.C.6 A side-by-side comparison of model specific (top) and marginal (bottom) curves flood hazard zones.

Appendix 5.A

Questionnaire

The Perception of Multi-hazard Events in the Val d'Orléans

In the context of this PhD research, we consider a ‘multi-hazard scenario’ to be a potential event involving two or more natural hazards, that may or may not be connected.

We have generated a draft multi hazad map as a tool to support decision-making process. Thus, within the framework of the doctoral thesis project, we aim to explore four key areas that are integral to our research on multi-hazard scenarios. Your responses will provide essential insights into each of these areas, helping us to better understand and address the complexities of multi-hazard risks.

The areas are as follows:

- The perception of the multi-hazard concept
- Understanding the strengths and weaknesses of the generated map
- Risk acceptability
- The means of awareness and communication on multi-hazards

This questionnaire is structured into four main parts:

Part 1. Perception and consideration of multi hazards,

Part 2. Utility of the produced map (Interpretation and understanding),

Part 3. Acceptability of risk,

Part 4. Recommendations for an exclusive DIRIM document.

Your participation will significantly aid in improving local risk management strategies. Thank you for contributing to this important effort.

Part 1. The awareness of the multi-hazard concept and the potential multi-hazard scenario in the area

This section consists of **five** questions.

Q1a. Have you come across the term ‘multi-hazard’ ? [Yes/Maybe/ No]

Q1b. If Yes/ Maybe, what is your underestanding (or prefered definition) of the term ‘multi-hazard’?

* please note that there is no fixed definition.

Q2. What is the main multi-hazard scenario that you think could be a threat in Val d’Orléans?

Q3. What is the likelihood of this scenario occurring in your local area or nearby?

1. Very likely
2. Likely
3. Neither likely nor unlikely
4. Unlikely
5. Very unlikely

Q4. In the context of a changing climate, do you believe this scenario will become more or less likely to occur?

1. Very likely
2. Likely
3. Neither likely nor unlikely
4. Unlikely
5. Very unlikely

Part 2. Evaluating the produced multi-hazard map (Interpretation and Understanding)

We are now going to ask you some questions on a multi-hazard map, for one multi-hazard scenario: sinkholes triggered by floods. This map was generated using **geological**, **hydrogeological**, **topography**, and **historical data** applying standard statistical modelling called the **Maximum Entropy Model**.

This map visualises multi-hazard *susceptibility (the relative probability of a hazard event without reference to any specific time interval).* Hazard zoning represents the regions susceptible to sinkhole collapses triggered by flooding and its purpose is to support decision-makers.

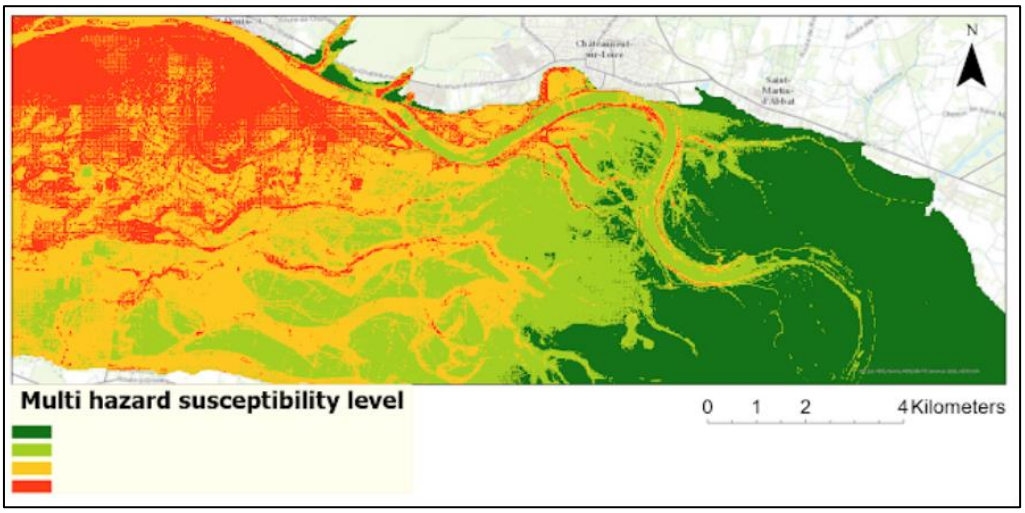
Please note that this is a draft version. It is a working document intended for review and feedback, not a final or definitive representation of risk. For this survey purposes, we only present a selected segment of the map.

Our goal is to collaborate with you to refine this map. We seek your insights to ensure the accuracy, relevance, and clarity of the information aiming for continuous improvements.

This part consists of **five** questions.

Q5. Please indicate whether you agree or disagree with the following statement:

'I can accurately identify areas of high and low hazard susceptibility on the map using the provided color scheme.'



1. Strongly agree
2. Agree
3. Neither agree nor disagree
4. Disagree
5. Strongly disagree

Q6. The map includes some zones coloured in red, what do these zones represent?

1. Areas where flood-triggered sinkholes will occur in a future flood
2. Areas where flood-triggered sinkholes could POTENTIALLY occur in a future flood
3. Areas where flood-triggered sinkholes have happened in the past
4. None of the above

Q7. What do you believe the colour dark green represents on the map?

1. There is **no** susceptibility to flood-triggered sinkholes
2. There is **low** susceptibility to flood-triggered sinkholes
3. There is **lower** susceptibility to flood-triggered sinkholes than other zones in this region

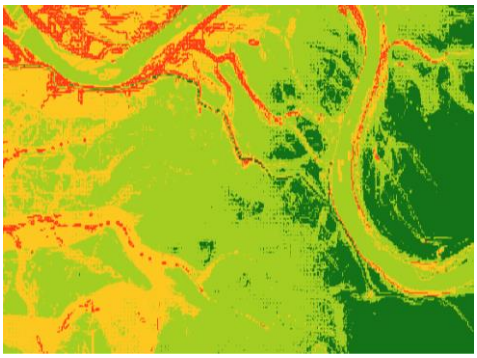
Q8. Regarding the two maps below, which colour do you think best communicates that there are regions with lower susceptibility to flood-triggered sinkholes (but that susceptibility still exists)

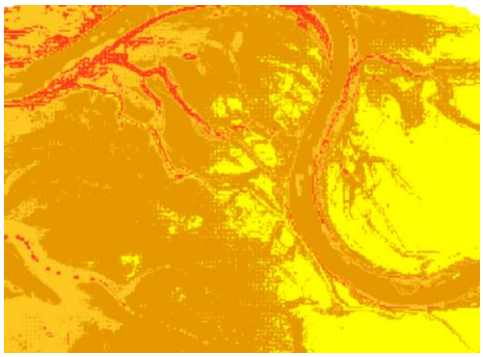
1. Dark Green
2. Yellow

3. Neither

If neither, what colour do you suggest?

……





Q9a. Do you think the type of maps shown above could enhance your professional work and/or decision-making? [Yes/ Mybe/No]

Q9b. If yes/ maybe, please could you provide more detail (e.g., specific examples of its practical utility)

Q10. Do you have any suggestions for improvements to the map to increase its usefulness?

Part 3. Risk acceptability (readiness to embrace the risk)

This part consists of **two** questions.

Q11. In the context of sinkholes triggered by flood, how do you think we should respond? (multiple options possible)

1. Accept the risk because it is not going to affect everywhere and when it does we need to response.
2. Invest money and time in completing subsurface investigations to improve mapping of this hazard.
3. Promote multi hazard thinking approach among decision makers and citizens.
4. Other responses

Other: ……..

Q12. Who do you think should take responsibility for understanding this risk or reducing this risk?

Part 4. Publishing a document (DICRIM/DDRM) on specifically multi hazards

This final theme is dedicated to exploring effective strategies for communicating the multi-hazard map to the public.

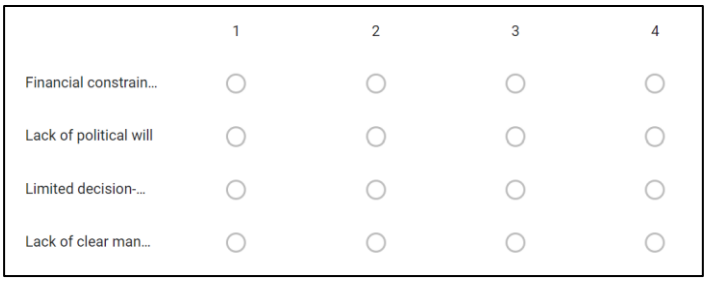
It consists of **five** questions.

Q13. Do you think it would be beneficial to establish a single Disaster Risk Communication (DICRIM/DDRM) document covering the multi-hazard concept (i.e., events involving two or more natural hazards, that may or may not be connected) and their relevance in the area? (Yes/ No_ not needed/ No_ integrate the information into existing publications)

Q14. What other practices do you know/recommend to foster multi-hazard thinking among decision-makers and citizens?

Q15. What do you think are the main barriers to integrating multi hazard thinking/ strategies in local resilience strategies? (Please rank)

1. Financial constraints
2. Lack of skills and expertise
3. Lack of political will
4. Limited decision-making power at local level
5. Lack of clear mandates for resilience



Others:…….

Do you have any additional comments or suggestions?

Consent form

I consent to the data being used to this research project. (Yes/ No)

Thank you for your collaboration!

Appendix E

Additional Field Visit Photos



Figure E.1 A visit from Val d'Orléans observing the sensitivity of the bedrock. (Credit: taken by the author, 2022).



Figure E.2 One of the many sinkholes that formed near the Loire River, hypothesized to have resulted from water receding after the 2016 flooding event. (Credit: taken by the author, 2022).



Figure E.3 A protected sinkhole called “Montission sinkhole” located in the *Saint-Jean-le-Blanc* municipality (Credit: taken by the author, 2022).

Évaluation de la susceptibilité multi-aléas dans les zones karstiques : étude de cas dans la région d'Orléans, France

L'objectif principal de cette thèse est de contribuer à l'amélioration de la gestion des risques multi-aléas dans le contexte des terrains karstiques, en développant et en évaluant une carte intégrée de susceptibilité multi-aléas pour deux dangers interconnectés (dolines déclenchées par des inondations), en utilisant un algorithme d'apprentissage automatique moins exploré (Maximum Entropy). L'efficacité de la carte produite est également évaluée grâce aux retours d'experts locaux recueillis via un sondage en ligne. Cette thèse intègre donc deux composantes de recherche : la modélisation d'une carte de susceptibilité multi-aléas et l'engagement des parties prenantes pour son évaluation. Cette recherche est motivée par une lacune mondiale dans la perception des terrains karstiques comme zones à risques multiples et par le besoin d'évaluations multi-aléas dans la planification de la réduction des risques de catastrophes en France. Cette thèse s'appuie sur des études existantes sur les dolines, adoptant une perspective multi-aléas. Les résultats confirment que la gestion future de la réduction des risques dans l'étude de cas doit être abordée sous l'angle d'une approche multi-aléas. Elle plaide en faveur d'une évaluation multi-aléas dans les zones karstiques, plutôt que de considérer les aléas pertinents de manière isolée. Cette approche permet une utilisation plus efficace des ressources en identifiant et en atténuant les zones de chevauchement où les aléas interconnectés se superposent, comme dans les communes de Bou, Darvoy, Saint-Denis-en-Val et Chécy. Notre carte de susceptibilité multi-aléas intègre les occurrences historiques d'effondrements, ainsi que des facteurs hydrologiques, géologiques, topographiques et les zones à risque d'inondation. Les résultats révèlent que les zones de basse altitude, dans un contexte karstique et de plaine inondable, sont les plus exposées. Cette recherche soutient la révision des règles institutionnelles et administratives pour adopter une approche multi-aléas, particulièrement dans les environnements karstiques. Mots clés: Multi-aléas, Interrelations entre les aléas, Zones karstiques, Maximum Entropy model

The core aim of this thesis is to contribute to improving multi-hazard risk management in the context of karst terrains by developing and evaluating an integrated multi-hazard susceptibility map of two interrelated hazards (flood-triggered sinkholes) using a less-explored machine-learning algorithm (Maximum Entropy). Additionally, the effectiveness of the resultant map is assessed through feedback from local experts using an online survey. Therefore, this thesis integrates two research components: (i) multi-hazard susceptibility map modelling, and (ii) stakeholder engagement and evaluation. This research is motivated by the global gap in perceiving karst terrains as multi-hazard areas (evidenced in this thesis) and the need for multi-hazard assessments in French disaster risk reduction planning.This thesis has built on existing sinkhole studies from a multi-hazard perspective. The results confirm that future risk reduction management in the case study should be addressed from a multi-hazard approach lens. It advocates for multi-hazard assessment in karst areas rather than viewing relevant hazards separately. This approach allows for more efficient use of resources by identifying and mitigating overlapping areas where interlinked hazards overlap, such as in the communes of Bou, Darvoy, Saint-Denis-en-Val, and Chécy. Our multi-hazard susceptibility map was based on integrating historical occurrences of sinkholes, hydrological, geological, and topographical factors, in addition to flood hazard zones. The results show that elevation is the main contributing factor, with low-elevated zones being more exposed to flooding and sinkholes in an area characterized by karst and floodplain. Therefore, this research provided empirical evidence for the institutional rules and administrative structures of response organisations to be redesigned and moving towards multi-hazard approach particular in karst environments. Keywords: Natural multi-hazards, Multi-hazard interrelationships, Karst areas, Maximum Entropy model

Signature du directeur de these: