People's Democratic Republic of Algeria

Ministry Of Higher Education and Scientific Research

Larbi Tébessi University –Tébessa-
Faculty of Nature, Life and Exact Sciences
Department of Earth Sciences and the Universe
Section of Geology

Ph.D. Dissertation

STATISTICAL AND GIS APPROACHES TO LANDSLIDE SUSCEPTIBILITY ASSESSMENT AND MAPPING IN MILA BASIN (NE ALGERIA)

Submitted In Partial Fulfillment
Of the Requirements for the Degree

Doctor of Philosophy (L.M.D) in Applied Hydrogeology

By

Abdelaziz Merghadi

Defense Committee:

Mohammed Laid Hamila	Professor	University of Tébessa	Chairman
Abderrahmen Boumezbeur	Professor	University of Tébessa	Supervisor
Djamel Athamania	MCA	University of Tébessa	Co- Supervisor
Chamseddine Fehdi	Professor	University of Tébessa	Examiner
Layachi Gouadia	Professor	University of Tébessa	Examiner
Riheb Hadji	Professor	University of Sétif	Examiner
Yacine Achour	MCA	University of Bordj Bou Arreridj	Examiner

Defense Date: *11 February 2020*

Distinction: "*Very Honorable*" (with Committee Praise)

Abstract

The severe landslides affecting Mila Basin (located in the North-East region of Algeria) are serious threats not only to the environment and the local populations but also inflecting economic burdens to local authorities by the non-ending reconditioning and restoration projects. In addition, these landslides affect the current landscape evolution and the geodynamics of the basin. Therefore, predicting and delineating landslides are crucial tasks to reduce their associated damages.

However, landslide risk prevention requires prone areas delineation using an assessment that can integrate into GIS environments and considering the spatial and temporal space component of the basin. This should theoretically provide probabilities for both the spatial and temporal components of this hazard in the form of susceptibility toward landsliding. That being said, no systematic and accountable landslide susceptibility models or even susceptibility maps are available for the basin yet, despite the tremendous losses.

In an attempt to fill this gap, an advanced statistical-based modeling approach (i.e. Machine Learning) was used to provide state-of-the-art models capable of providing the highest landslide prediction capabilities. The main research workflow was rather simplistic as it focuses essentially on elaborating predictive models using some advanced techniques that can be integrated successfully in GIS environments in order to develop customized models for the basin. A partial focus was given to mapping and zoning areas toward landsliding.

The obtained results highlight the overall benefits of using advanced machine learning methods for landslide susceptibility assessment, as the implemented models exhibit reasonably good predictive performance (AUC>0.85, Acc> 78% and kappa>0.56). The generated landslide susceptibility maps were proven to be useful as a technical framework for spatial prediction to develop countermeasures and regulatory policies by decision-makers to minimize the damages introduced by either existing or future landslides.

Keywords: Landslide; Susceptibility mapping; Susceptibility assessment; Machine learning; GIS; Mila basin; Algeria

Résumé

Les glissements de terrain dans le bassin de Mila constituent non seulement une véritable menace aux biens et propriétés de la population de la région mais également au développement socio-économique local. Des sommes importantes d'argent du contribuable sont consommées chaque année par les projets non-finis de réparations, restaurations et remise en service des voix de communications, des canalisations et des habitations. En plus, les glissements de terrain sont directement responsables de l'évolution du paysage et de la géodynamique naturelle du bassin. Par conséquence, prédire et délimiter les zones susceptibles d'être touchées par ce sinistre est une des tâches essentielles pour réduire et limiter l'endommagement résultant suite à la transgression urbaine sur les zones qui étaient, autrefois, jugées marginales.

Toutefois, la prévention et l'allègement du risque au glissement de terrain nécessite, comme première tâche, une délimitation et une bonne reconnaissance des zones exposées à cet aléa. C'est en réalité qu'on appelle l'inventaire, il doit inclure également les caractéristiques géologiques, géomorphologiques, structurales et physico-mécaniques de l'ensemble du bassin. Ces données spatiales et non spatiales sont ensuite statistiquement étudiées pour réduire la redondance puis intégrés dans un environnement SIG. Les séries de calcules dans cet environnement nous donne théoriquement des probabilités d'occurrence de l'aléa à travers tout le bassin d'étude sous forme d'une carte de susceptibilité au glissement. Malgré les pertes et les dommages enregistrées jusqu'à nos jours, aucun model fiable de calcul de la susceptibilité au glissement dans le bassin n'a était établi.

Pour tenter de combler cette lacune et dans un objectif d'estimation de la susceptibilité au glissement dans chaque endroit du bassin, Plusieurs modèles sophistiqués et avancés ont été établis, sur la base de l'application des approches statistiques à savoir LR, GBM, NNET, RF and SVM (i.e. Machine Learning).

Cette recherche a été focalisée essentiellement sur l'élaboration des modèles prédictives dans l'environnement SIG dans le but de choisir un model fiable et efficace pour le calcul de la susceptibilité spécialement pour la région d'étude. Un

intérêt particulier a été donné aux processus de la cartographie de la répartition spatiale des éléments du relief susceptibles aux glissements.

Les résultats obtenus, montrent l'intérêt d'utiliser les méthodes de Machine Learning dans l'évaluation de la susceptibilité aux glissements des terrains, car les modelés montent des performances prédictives intéressantes (AUC>0.85, Acc> 78%, et kappa>0.56). Les cartes de la susceptibilité générées, peuvent être très utiles comme des documents techniques pour la prédiction spatiale de cet aléa. Elles pourront également servir comme document de base dans le but de développer des mesures d'allègement et des politiques réglementaires afin d'orienter les plans d'aménagement avec le minimum de préjudice.

Mots clés: Glissement; Cartographie de la susceptibilité; Evaluation de la susceptibilité; Machine Learning; SIG; Bassin de Mila; Algérie

ملخص

انزلاقات التربة الشديدة التي يتعرض لها حوض ميلة (المتواجد في المنطقة الشمالية الشرقية للجزائر) تشكل خطرا جسيما ليس على المحيط الطبيعي و تجمعات السكانية فقط و لكن ايضا الميزانية العامة للمنطقة تتاثر بالاعباء الاقتصادية عن طريق مشاريع الترميم و الاصلاح. كما ان انزلاقات التربة توثرعلى المحيط العام للمنطقة و جيوديناميكية الحوض. و لهذا السبب فان تنبا وتوقع حدوثها مع تحديد مناطق الانزلاق امر في غاية الأهمية لتقليص الخسائر الناجمة.

عملية الوقاية من اخطار الانزلاقات تُحتم بالضرورة تحديد المناطقة المتضررة عن طريق تقدير الحالة و دمجها في بيئات نظم المعلومات الجغرافية مع اخد الاعتبارخصائص الحوض، فانه يمكن نظريا الحصول على معلومات في شكل احتمالات حول الخصائص الزمانية و المكانية لهذه الظاهرة على شكل احتمالات قابلية الانزلاق. وعلى هذا فان الاخطارالمتوقع حدوثها في المستقبل يمكن الحد منها حسب نتائج احتمالات قابلية الانزلاق. غير انه لا يوجد لحد الساعة اي نموذج اوخريطة لاحتمالات قابلية انزلاق التربة متوفرة للحوض.

لمعالجة هاته المشكلة، هناك مجموعة من التقنيات و نماذج متطورة مبنية على اسس احصائية (التعلم الألي) استخدمت لتوفير نماذج ممتازة تساعد على توفير احسن درجات تنبؤ بانزلاقات التربة. هذا البحث في مجمله بسيط حيث يعتمد بصفة عامة على تحضير نماذج تعتمد بالدرجة الاولى على تقنيات متطورة تسمح بادماجها في نظم المعلومات الجغرافية لتطوير نماذج توقعية خاصة بحوض ميلة مع اهتمام جزي برسم و تقسيم الحوض حسب القابلية لانزلاق التربة.

النتائج المتحصل عليها تسلط الضوء على الفائدة العامة من استخدام تقنيات متطورة من التعلم الألي في مجال تقديرانزلاقات التربة لان نماذج التوقعية المتحصلة عليها ذات جودة عالية و تمتاز بقدرة توقعية ممتازة (المساحة تحت المنحى>0.85، الدقة العامة>78%، معامل كابا>0.56). خرائط احتماليات انزلاق التربة المتحصل عليها انها ذات فاعلية من حيث القدرة التوقعية وذات فائدة كمنصة تقنية للتوقع الجغرافي و تطوير التدابير المضادة و سياسات تنظمية من طرف صناعي القرار لتقليص الخسارر الناجمة عن الانزلاقات الحالية أو المستقبلية.

كلمات مفتاحية: انزلاقات التربة؛ تخطيط قابلية للانزلاق؛ قابلية الانزلاق؛ التعلم الألي؛ نظم المعلومات الجغرافية؛ حوض ميلة؛ الجزائر

TO DEAR GOD

Acknowledgments

First and foremost, I would like to thank Pr. Abderrahmane Boumezbeur, who introduced me to Geomorphology and Geotechnical Engineering. He also suggested and then gave me the opportunity to participate in the Ph.D. program at the Department of Earth Science of Larbi Tébessi University.

I started my research carrier working in Computer science and Software Engineering. Thanks to Boumezbeur, who always sustained the importance of interdisciplinary in the Institute he leads.

This Ph.D. in Computer Science and Engineering Geology represents to me the joining link between my past and my present experience as a researcher.

The unconditioned and careful assistance of my previous professors like Pr. Bali and Pr. Gouadia, whom I ran into regarding logistic and administrative problems and issues.

I cannot forget my previous teachers, a few of them are very important for me to learn how to approach scientific questions such as Dr. Dfaflia who happen to be my Master Dissertation supervisor, Pr. Fahdi, Pr. Bali, Pr. Helima, Pr. Gouadia and Dr. Aoun for their unconditional support.

I would like to express my deepest gratitude to the defense committee for their helpful and constructive comments on the dissertation and everyone provided assistance in the process of making and enhancing this research especially the DTP (Direction des Travaux Publics), both the Mila and Constantine municipalities for providing the necessary data. A special thanks to Dieu Tien Bui and Dou Jie for academic support.

Embarking in such Ph.D. research required my family support. During this period, I received constant encouragement and care from all my family. Among my close relatives, I am particularly indebted to my mom who’s unconditioned and relentless assistance and appreciation were indispensable.

Table of Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1: Introduction

1.1 BACKGROUND

Algeria is the largest country in Africa claiming an area of $2{,}381{,}741\ km^2$ approximately. This vast spatial extent is exposed to a variety of natural hazards such as Earthquakes (e.g. Boumerdes 2003), Floods (spread widely across the country in all four major directions), Landslides (e.g. Mila, Bouira, Medéa), Drought and Desertification. The diversity in natural hazards across the country, mandate a strict and stringent protection strategy against these hazards.

Mila basin is particularly a unique case in Algeria. This basin is considered (and still) the most vulnerable basin in terms of landslides and floods. Landslides are regarded as natural degradation processes produced by natural and human activities [1]. Natural factors such as rainfall, earthquake, and volcanic eruption can trigger landslide occurrences. This hazard might become worse when human activities also contribute to landslide occurrences.

In Mila basin, different locations were surprisingly found in critical states. The economic losses instantiated by the spatial evolution of landslides generate huge burdens on local authorities and thus slowing the local development of the basin. In fact, most local agencies in Mila basin have less experience related to landslides, the preparedness and mitigation activities are not running well in many regions, despite the abundant landslides. The lack of required data, landslide hazard experts, limited budget, lack of reliable susceptibility analyses and the lack of awareness of the local government agencies, are some of the reasons why the mitigation and preparedness activities are far from adequate. On the other hand, these activities are obsolete in order to reduce the effects of the landslides.

In reality, there's exist different nation-wide engineering and hazard mapping projects initiated during the 80's such as ZERMOS (Zones Exposées aux Risques de Mouvements de Sol et de Sous-sol) and PER (Plans d'Exposition aux Risques Naturels) for the sole purpose of implementing hazards susceptibilities (e.g. landslides) by local government agencies in POS (Plans d'Occupation des Sols) and PDAUs (Plan Directeur d'Amenagement et d'Urbanisme).

One of the main purposes of landslide hazard mapping projects (i.e. ZERMOS and PER), is to generate landslide susceptibility maps[1]. These maps, depict the spatial probability of occurrences of landslides based on an empirical assumption of "past and present landslide failures does not occur randomly or by chance, but instead failures follow patterns that share common geotechnical behaviors under similar conditioning factors". This means, correlating all landslides related factors that may influence the landslide occurrences with the past distribution of slope failures. Practically, the only information required is the landslide spatial distribution (i.e. geographical coordinates) and landslide class labels (i.e. from the referent landslide inventory) and the related conditioning factors. However, in order to generate a landslide susceptibility map there exist two possibilities:

- Consider the already published methods and models.
- Build and modify a model from scratch to get better-expected results due to the fact that the model will be tailored for the study area.

In spite of that, the chosen mapping method depends on data availability, financial budget, time available for monitoring and observing landslides, detailed level of the acquired data, scale analysis and the proposed models and methodology for assessing and mapping landslide susceptibility [2]. For decades, models and methodology that rely on deterministic and expert-knowledge heuristic in geology, geomorphology and soil mechanics were the primary focus for understanding landslide phenomenology and its generative process. However, as the hazard get more complex over the years due to human activities influencing landslides in very different and complex ways, large gamut of models that rely computer science and statistic are gaining attention in demystifying and understanding the hazard, especially, when each case study have its unique set of properties that differ from one to another, making the generalization process harder and difficult to achieve. This is practically noticeable, as many local government agencies and researchers tried to implement classic deterministic and expert-knowledge heuristic methods and models for landslide susceptibility assessment in a small and specific area of interest.

[1] Landslide susceptibility maps are used to define the relative degree of instability of the terrain (for more detail see Chapter 2).

However, despite the fact that these models were only implemented for very specific case-studies, these models are known to suffer major drawbacks that can be:

- The relatively high cost.
- Limited scope and spatial extent.
- Relying on subjectivity in defining conditioning factors, weights and scoring values. These given values were based on either user export-opinion or taken values from another location. This kind of subjectivity in any given model is substantial and heavily depends on the expert's familiarity with the area.
- Requiring a significant amount of cooperation and advisement of other experts in the field. Thus, introducing, an infinite amount of subjectivity that can be inconvenient for a preliminary regional assessment and planning[2].

As a result, the generated landslide susceptibility maps by these methods suffer many weaknesses and uncertainties. However, instead of deterministic and expert-knowledge heuristic methods, statistical-based methods and models for mapping and assessing landslide susceptibility, are very useful and convenient for large-medium scale analysis as it rely on fact that "previous, current and future landslide failures do not happen randomly or by chance, but instead, failures follow patterns and share common geotechnical behaviors under similar conditions of the past and the present" [3] and only require, collecting and preparing an accurate database (i.e. a geospatial database of landslide inventory[3] and conditioning factors) with maximum details available. Then, models based on these methods are trained and validated using that database and afterward, the resulting models are used to generate landslide occurrence probabilities [4].

[2] This involve preliminary levels of risk or disaster management, landscape (regional) planning, route selection, insurance management and so forth.

[3] Relying on landslide inventory map to build a landslide susceptibility map is much more convenient to determine the overall landslide pattern that site-specific details and/ or expert opinion.

1.2 PROBLEM STATEMENT

This research is picking interest in landslide susceptibility assessment and mapping at a study area located North East of Algeria, called Mila basin. This basin is well known for the variety of landslides that are (to a certain degree) non-mapped due to the unique heterogeneous properties (i.e. geology, geomorphology and so forth, that vary dramatically). This dramatic variance of landslides in the basin, in terms of spatial repartition and intensity, became a very serious handicap to the urban, local, social and economic development of the basin since 1985. Over the years, the ever-increasing rate of this hazard is, in particular, increasing the number of the element at risk exposed to landslides, especially at urban zones. As a consequence, an increase in the economic burden associated with landslides damages it became a major issue for the local development[4]. Over the year, these burdens trigger a reaction-chain of two separate issues:

- Stable areas are becoming more expensive for landlord and project development.
- Constructions in inadequate terrains and/or soils increase the overall expenses and the project budget by the exposure of such constructions to landslide and land instabilities.

For these reasons, landslide mitigation processes became an absolute necessity that mandate assessing landslides susceptibility in a systematic, fast and evolving ways using models that are capable of anticipating the overall patterns of this phenomena and thus better understanding and evaluating the overall damages and maybe future development projects. Despite the several remedial projects that have been carried out over the years, the effect of landslides in term of damages is still persisting and sometimes even worse due to the fact these remedies actually focus on:

- Treating the symptoms of the landslides instead of the issue itself without considering soil's intrinsic properties.
- Randomly patching site-specific and in-situ related issues.
- Relying on subjective expert-opinion which objectively incomprehensible.

[4] some local experts refer to it as an economic threat.

- Focusing on expensive and classic methods that are limited either in scope or spatial and temporal extent.

Yet, these remedies tend to ignore:

- The overall landslide patterns behaviors.
- Recent advancements in computer science and the new innovative state-of-the-art landslide models.

Overall, this research will try to implement statistical-based modeling for landslide susceptibility assessment and mapping in GIS compatible environment at Mila basin as a case study by relying on Machine learning[5] using state-of-the-art Computer Science models and algorithms instead of well documented and elaborated conventional and traditional approaches for landslide susceptibility. Moreover, it is important to note that this research will focus on the process of modeling and assessing the landslides in the study area, whereas landslide phenomenology, evolution process, triggering mechanism, landslides conditioning and triggering factors are out of the scope of this analysis because of being well presented and discussed in number of research literature and investigation campaigns.

One important research hint is that landslide typologies observed in the study were diverse and entirely different[6], despite landslide occurrences that are mostly linked to Tertiary formations. For this reason alone, it was advised to assess only one landslide type at a time, and eventually combine these separate assessments later on as suggested by Van Westen, Van Asch [5]. Therefore, the focus was substantially giving to "slides" landslides.

1.3 OBJECTIVES

Resting on the above-mentioned motifs, this research was shaped to meet the standardized requirements [6-11] in terms of methodology of data acquisition and manipulation, choices of the advanced modeling approaches for landslide assessment, as well as the model evaluation techniques, and finally, the visualization choices, all via GIS. These objectives could be structured as follows:

[5] This can be considered as one of the most effective methods for solving non-linear geo-spatial problems like landslides susceptibility, using either regression or classification.

[6] Vary from typical deep-seated earth slides, to shallow earth slides and flows, spreads...etc.

1. Address the shortage in literature for Mila basin in term of landslide susceptibility mapping through investigating, implementing, assessing and comparing prediction capability of advanced statistical-based models such as Machine Learning methods and algorithms
2. The production of useful landslide susceptibility mapping and assessment frameworks with a reproducible and unbiased optimization process and exploit the possibility of automating the process of landslide susceptibility mapping or landslide mapping by taking advantage of available resources at the local agencies and open source community.
3. Standardizing the procedure regarding landslide assessment in the study area (i.e. acquisition, scaling, pre-processing, optimization, and evaluation procedures) by preparing custom and reproducible algorithms for specifically the purpose of landslide assessment in the study area using GIS.
4. Implementing a variety of known models and techniques that rely on statistical modeling approaches, but also experimenting with the state-of-the-art techniques, advanced methods and unprecedented solutions for landslide assessment using GIS.
5. Evaluating models performance and the results obtained using the most appropriate procedures and methods, in favor of gaining a qualitative and quantitative descriptors evaluations of the model's performance using GIS in combination with statistical tools.
6. Address the issues of availability, visualization, and publishing of the detailed results in the form of reproducible, reliable, generic landslide susceptibility map per each model using GIS, and web-GIS and estimating their applicability for better environmental management and for reducing the victims and damages caused by future landslide occurrences.

1.4 THESIS OUTLINE

The thesis is structured in seven chapters including this introductory chapter, a Bibliography, and two Appendices:

- **Chapter 1** presents a general introduction, followed by the main problematic which is the subject of the thesis. The outlined Objectives and the initial observations that inspired the research topic were described in the same chapter.
- **Chapter 2** reports theoretical background information on landslides phenomenology and emphasizes the philosophy, techniques, methods, and approaches used for landslide susceptibility assessment and mapping processes. A description of the general knowledge of incorporating spatial information (GIS) in landslide susceptibility is presented. A special section is devoted to the problems and issues encountered in landslide susceptibility processes. These issues will be the main concern of later chapters that deal with the susceptibility analysis.
- **Chapter 3** presents an overview of the used methodologies to assess the landslide susceptibility while concepts and theories used by these methodologies are presented and discussed in details. The discussion is supported by simple mathematical illustration with landslide susceptibility paradigm in mind. These particular algorithms and methodologies were employed in the landslide susceptibility process performed in this research. The presented details will be used later to understand characteristics, behaviors, limitations, advantages, drawbacks, and the predictive capabilities of the landslide susceptibility models. This chapter is also concerned with the most appropriate techniques employed for searching, optimizing, and evaluating hyperparameters and the performance of the susceptibly models. In the last part of the chapter, a detailed research workflow was intended to provide a practical guide to landslide susceptibility assessment and mapping processes.
- **Chapter 4** intended to provide fundamental background information's about the case study (i.e. Mila basin), followed by a description of the main characteristics of landslide typology encountered in the basin. Then, a geospatial database that was used as an input dataset was discussed in terms of reporting and describing the identified landslides hosted in the geodatabase along with the relative conditioning factors used to anticipate the landslide susceptibility process.

- **Chapter 5** reports the results obtained at each step of the implemented research workflow (See Chapter 3). Discussions with critical reviews and comparisons are carried out for the obtained results as recommended by the scientific community. The results are used to set out guidelines for decision-makers and planners to be able to implement landslide susceptibly mapping and assessment effectively.
- **Chapter 6** concentrates on outlining the main achievements the research achieved compared to the primary underlined research objectives.
- **Chapter 7** presents general conclusions drawn from this research. It specifically comprises a number of guidelines for the landslide susceptibility assessment and mapping extracted from the produced results and experience gained during the research. Finally, this chapter gives suggestions for further researchers that might benefit from this specific research topic.
- **The Bibliography** includes an extensive list of references.
- **Appendix A** reports the repositories used to host the source code used to perform the analysis of this research.
- **Appendix B** includes a descriptive statistical table for the landslide conditioning factors used in research.

Chapter 2: Theoretical Background

In order to present the problematic of this thesis systematically, it is first necessary to define and communicate the basic theoretical background behind the landslide phenomenology, the comprehension of qualitative landslide assessment, the impact of available technologies which is in service of landslide assessment and the way in which GIS is enrolled into it.

2.1 LANDSLIDE PHENOMENOLOGY

2.1.1 Definitions and Scope

The term landslide had evolved over the years gaining various interpretations and conceptions depending on the scientific discipline or the research-schools. As a result, "landslide" as term expresses more or less either specific or general phenomena. From the first-person perspective, a general definition (term, conception or description) that encompasses the phenomena would be highly feasible and recommended as it offers a simple solution, but overall lacking details. Thus, it makes it even more difficult to really understand the landslide phenomena. However, the abundant information's on the aforementioned issue is widely available on literature reviews and will not serve any particular interest in this research.

The following paragraphs provide a brief summary explaining, discussing and illustrating the essential landslide terminologies used in the multilingual landslide glossary, which should (for uniformity of practice) be adopted when classifying and describing a landslide to fully comprehend and understand landslide phenomena based on Varnes [10], Cruden and Varnes [12], and Cruden and Couture [13] definitions and interpretations, which are one of the most broaden and highly endorsed landslide interpretations and definitions worldwide by lead scientific communities and consortium such as WP/WLI[7].

Landslides are denoted as a gravitational (i.e. under the influential of gravity) downward and outward mass-movements of different slope-forming materials (i.e.

[7] The International Geotechnical Societies' UNESCO Working Party on World Landslide Inventory (WP/WLI) was formed for the International decade for Natural Disaster Reduction (1990 to 2000).

rocks, debris, earth masses (soils), artificial fills or mix of everything), developed along predefined surfaces (i.e. stratification joints, schistosity plan, discontinuities…etc.), that are widely known as slip-surfaces[8] (can vary from simple geometry such as planar slip-surface to a higher-order geometry resulting in very complex slip-surfaces), on which propagate throughout the mass and clearly separate intact bedrock material from the moved material above.

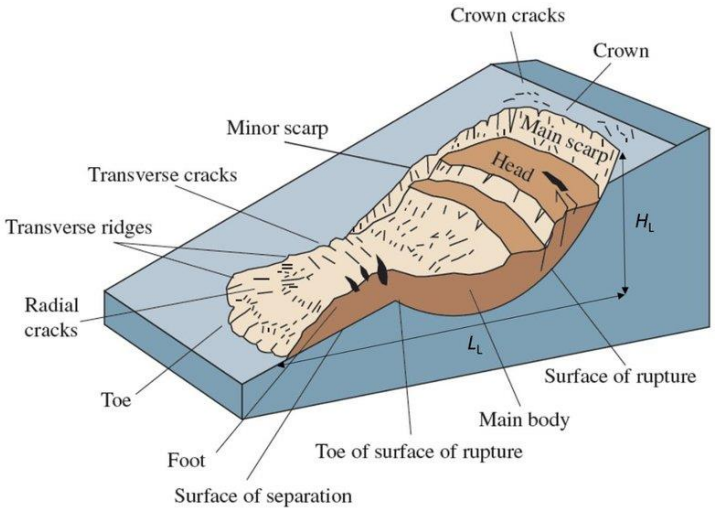


Figure 2.1 Nomenclature for labeling the parts of a landslide adopted in Cruden and Couture [13] and Varnes [10].

Morphologically an ideal landslide consists of (from top to bottom) crown and head separated by scrap; the main body, channeled by flanks; foot terminated by a toe; depletion zone and accumulation zone capturing upper-lower portions of the landslide (Figure 2.1). Landslides develop mostly in slopes either natural or engineered and can vary in terms of size and the affected area, which make landslides a hot-topic for various fields such as Geology, Geomorphology, engineering fields (e.g. Geological, Geotechnical, Civil Engineering), and Computer Science [10, 14-18]. Therefore, detailed descriptions and definitions of landslides related terminologies, geometries, and dimensions are shown in Figure 2.2 and Figure 2.3 and explained in Table 2.1 and Table 2.2, respectively.

[8] Also known as the surface of rupture.

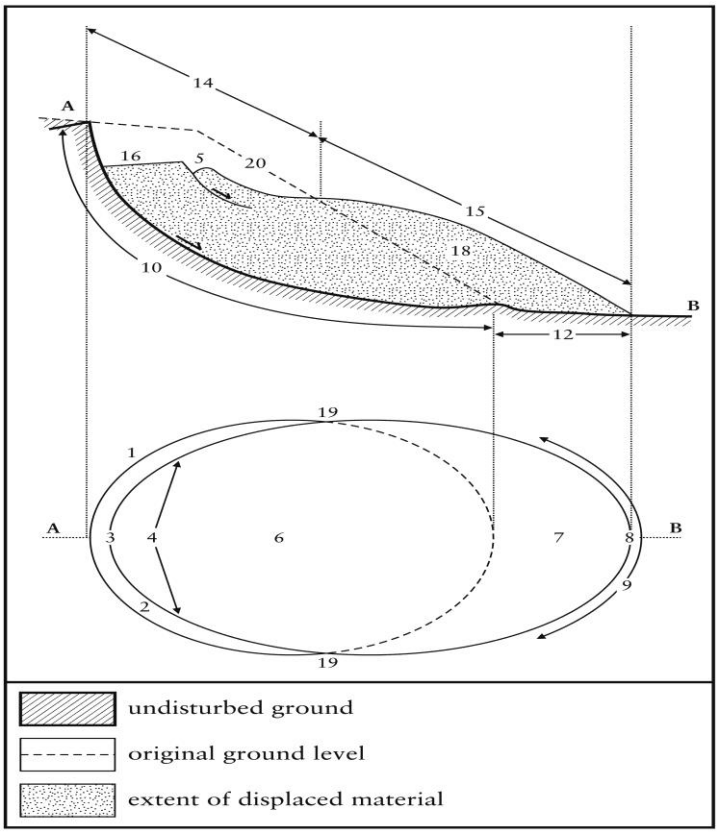


Figure 2.2 Description of landslide parts in profile and plan views.

(See Table 2.1 for an explanation of the numbers, Source: After WP/WLI [19])

Table 2.1 Description of landslide parts in profile and plan views.

No.	***Material type***	***Description***
1	Crown	The practically non-displaced material still in place and adjacent to the highest parts of the main scarp (2).
2	Main Scarp	A steep surface on the undisturbed ground at the upper edge of the landslide, caused by the movement of the displaced material (13) away from the undisturbed ground. It is the visible part of the surface of rupture (10).
3	Top	The highest point of contact between the displaced material (13) and the main scarp (2).
4	Head	The upper parts of the landslide along the contact between the displaced material and the main scarp (2).
5	Minor Scarp	A steep surface on the displaced material of the landslide produced by differential movements within the displaced material.
6	Main Body	The part of the displaced material of the landslide that overlies the surface of rupture (10) between the main scarp (2) and the toe of the surface of rupture (11).

7	Foot	The portion of the landslide that has moved beyond the toe of the surface of rupture (11) and overlies the original ground surface (20).
8	Tip	The point of the toe (9) farthest from the top (3) of the landslide.
9	Toe	The lower, usually curved margin of the displaced material of a landslide, it is the most distant from the main scarp (2).
10	Surface of Rupture	The surface which forms (or which has formed) the lower boundary of the displaced material (13) below the original ground surface (20).
11	Toe of the Surface of Rupture	The intersection (usually buried) between the lower part of the surface of rupture (10) of a landslide and the original ground surface (20).
12	Surface of Separation	The part of the original surface (20) overlain by the foot (7) of the landslide.
13	Displaced Material	Material displaced from its original position on the slope by movement in the landslide. It forms the depleted mass (17) and the accumulation (18).
14	Zone of Depletion	The area of the landslide within which the displaced material lies below the original ground surface (20).
15	Zone of Accumulation	The area of the landslide within which the displaced material lies above the original ground surface (20).
16	Depletion	The volume bounded by the main scrap (2), the depleted mass (17) and the original ground surface (20).
17	Depleted Mass	The volume of the displaced material which overlies the rupture surface (10), but underlies the original ground surface (20).
18	Accumulation	The volume of the displaced material (13) which lies above the original ground surface (20).
19	Flank	The non-displaced material is adjacent to the side of the rupture surface. Compass directions are preferable in describing the flanks but if left and right are used, they refer to the flanks as viewed from the crown (1).
20	Original Ground Surface	The surface of the slope that existed before the landslide took place.

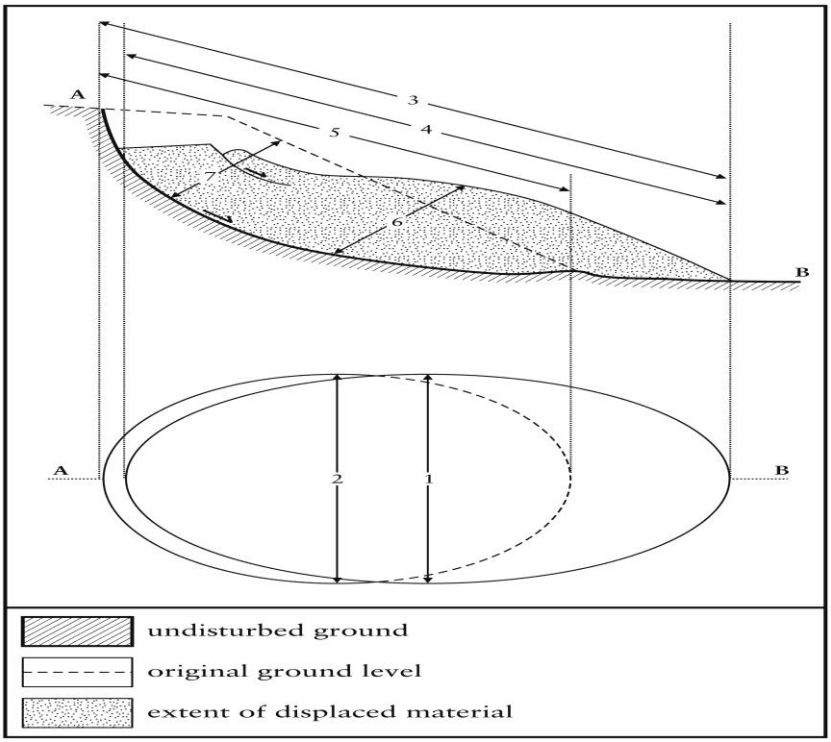


Figure 2.3 Description of landslide dimensions in profile and plan views.

(See Table 2.2 for an explanation of the numbers, Source: Based on Cruden and Varnes [12] and WP/WLI [11])

Table 2.2 Description of landslide dimensions in profile and plan views.

No.	***Material type***	***Description***
1	The width of the Displaced Mass	The width of the displaced mass, $\boldsymbol{W_d}$ is the maximum breadth of the displaced mass perpendicular to the length of the displaced mass, $\boldsymbol{L_d}$.
2	The width of the Rupture Surface	The width of the rupture surface, $\boldsymbol{W_r}$, is the maximum width between the flanks of the landslide, perpendicular to the length of the rupture surface, $\boldsymbol{L_r}$.
3	Total length	The total length, $\boldsymbol{L}$, is the minimum from the tip of the landslide to the crown.
4	Length of the Displaced Mass	The length of the displaced mass, $\boldsymbol{L_d}$, is the minimum distance from the tip to the top.
5	Length of the Rupture Surface	The length of the rupture surface, Lr, is the minimum distance from the toe of the surface of rupture to the crown.
6	The depth of the Displaced Mass	The depth of the displaced mass, $\boldsymbol{D_d}$, is the maximum depth of the displaced mass, measured perpendicular to the plane containing $\boldsymbol{W_d}$ and $\boldsymbol{L_d}$.
7	The depth of the Rupture Surface	The depth of the rupture surface, $\boldsymbol{D_r}$, is the maximum depth of the rupture surface below the original ground surface measured perpendicular to the plane containing $\boldsymbol{W_r}$ and $\boldsymbol{L_r}$.

Slopes stability is maintained and determined by the equilibrium of two driving forces that act upon the slope. Displacement is irreversible deformation that will take place if resisting forces are succumbed by driving forces. The driving forces involve: increasing slope weight or shear stress (i.e. via water saturation, adding load or rearranging of the slope geometry), loss of support (i.e. via erosion and rearranging of the slope geometry) or dynamic influences. On the opposite, resisting forces are represented by shear strength and cohesion of slope materials, as well as friction along a slip-surfaces, which all further depend on the nature and condition of the slope materials[9], but as well on the slope morphology and geometry (i.e. steepness, elevation, curvature etc.).

Conditioning factors are the feature sets, variables or parameters that may influence slopes stability by influencing either directly or indirectly the driving and/or resisting forces. These factors are able to provide technical background on the landslides occurrences. On the other hand, triggering factors are the feature sets (variables or parameters) that once the terms of slopes failure are reached and satisfied, the landslide process unfolds under the influence of one of these different factors or by their combination. It's important to note that according to Cruden and Couture [13], conditioning and triggering factors are regrouped into four categories[10], i.e. geological, morphological, physical and human-induced factors, as given in Table 2.3. However, it may not be mandatory to include all factors in each landslide assessment it solely depends on a variety of parameters (e.g. the case study where few or more factors can be used), as explained by Soeters and Van Westen [2].

[9] Refer to: (a) freshness such as weathering degree; (b) structural elements such as the presence of joints and fissures; (c) heterogeneity such as contrasts of water permeability or deformability; (d) presence and/or absence of vegetation.

[10] Only a brief outline is given since description of landslide causal factors is not the focus of this study.

Table 2.3 A brief list of landslide conditioning and triggering factors.

Geological Factors	***Morphological factors***
Plastic weak material	Tectonic uplift
Sensitive material	Volcanic uplift
Collapsible material	Glacial rebound
Weathered material	Fluvial erosion of the slope toe
Sheared material	Wave erosion of the slope toe
Jointed or fissured material	Glacial erosion of the slope toe
Subterranean erosion (solution, piping)	Erosion of the lateral margins
	Deposition loading of the slope or its crest
Adversely oriented mass discontinuities (including bedding, schistosity, cleavage)	Vegetation removal (by erosion, forest fire, drought)
The contrast in permeability and its effects on groundwater contrast in stiffness (stiff, dense material over plastic material)	Adversely oriented structural discontinuities (including faults, unconformities, flexural, shears, sedimentary contacts)
Physical factors	***Human-Induced Factors***
Intense, short-period rainfall	Excavation of the slope or its toe
Rapid melt of deep snow	Loading of the slope or its crest
Prolonged high precipitation	Drawdown (or reservoirs)
Rapid drawdown following floods, high tides or breaching of natural dams	Mining and quarrying (open pits or underground galleries)
Earthquake	Defective maintenance of drainage systems
Volcanic eruption	Water leakage from services
Breaching of crater lakes	Vegetation removal (deforestation)
Thawing of permafrost	Irrigation
Freeze and thaw weathering	Creation of dumps of very loose wastes
Shrink and swell weathering of expansive soils	Artificial vibration (including traffic, pile driving, heavy machinery)

Landslide activity is a high valuable parameter that aids in estimate the current state (Figure 2.4 and Table 2.4), style (Figure 2.6 and Table 2.6) and distribution (Figure 2.5 and Table 2.5) of the activity of the landslide development process. Essential, slopes develop progressively and become cyclical as soon as they enter the landslide process, which implicates that the relative displacements are at the highest peak during the initial activation and tend to decrease per each reactivation cycle, but on opposite side, the frequency of the events increases as a landslide progresses

toward an active stage and therefore it is highly important to estimate the state, style and the distribution of their current activity in order to scale the future displacement rates. Slopes usually have a life cycle that loop (repeatedly cycle) through the first failure stage followed by a reactivation stage, which is separated by suspended and dormant stages and this repeats per every cycle until the active stage is reached.

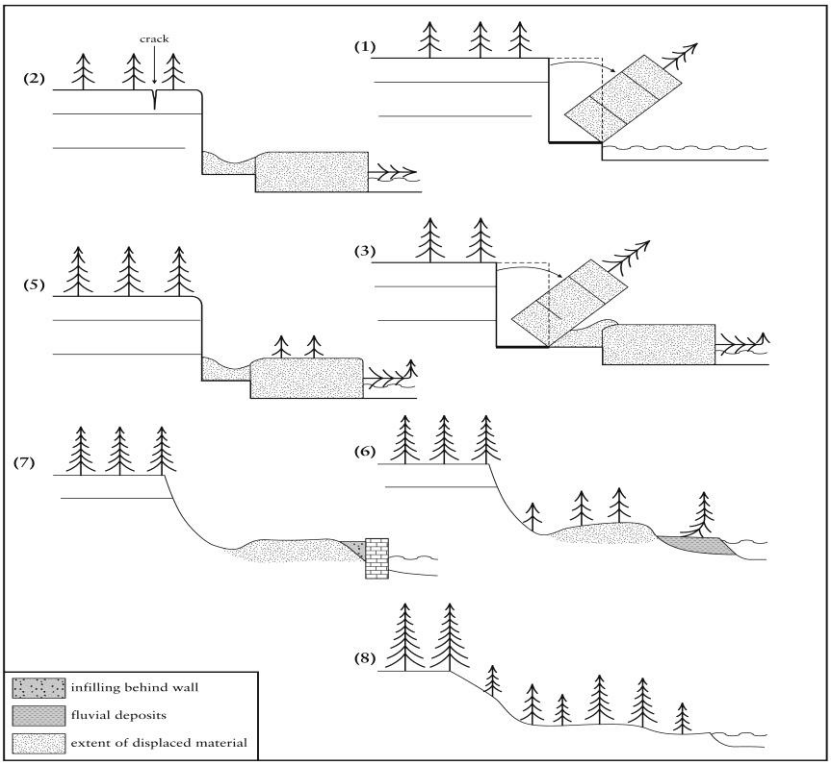


Figure 2.4 Classification of the states of activity of landslides.

(See Table 2.4 for an explanation of the numbers and states, Source: after WP/WLI [19])

Table 2.4 Classification of the states of activity of landslides.

No.	Activity state	Description
1	Active	An active landslide is currently moving. In the example shown erosion at the toe causes a block to topple.
2	Suspended	A suspended landslide has moved within the last 12 months but is not active at present. In the example shown local cracking can be seen in the crown of the topples.
3	Reactivated	A reactivated landslide is an active landslide that has been inactive. In the example shown another block topples and disturbs the previously displaced material.

4	Inactive[11]	An inactive landslide has not moved within the last 12 months and can be divided into 4 states: Dormant, Abandoned, Stabilized and Relict.
5	Dormant	A dormant landslide is an inactive landslide which can be reactivated by its original causes or other causes. In the example shown the displaced mass begins to regain its tree cover and scarps are modified by weathering.
6	Abandoned	An abandoned landslide is an inactive landslide which is no longer affected by its original causes. In the example shown the fluvial deposition has protected the toe of the slope, the scarp begins to regain its tree cover.
7	Stabilized	A stabilized landslide is an inactive landslide which has been protected from its original causes by remedial measures. In the example shown a retaining wall protects the toe of the slope.
8	Relict	A relict landslide is an inactive landslide which developed under climatic or geomorphological conditions considerably different from those at present. In the example shown uniform tree cover has been established.

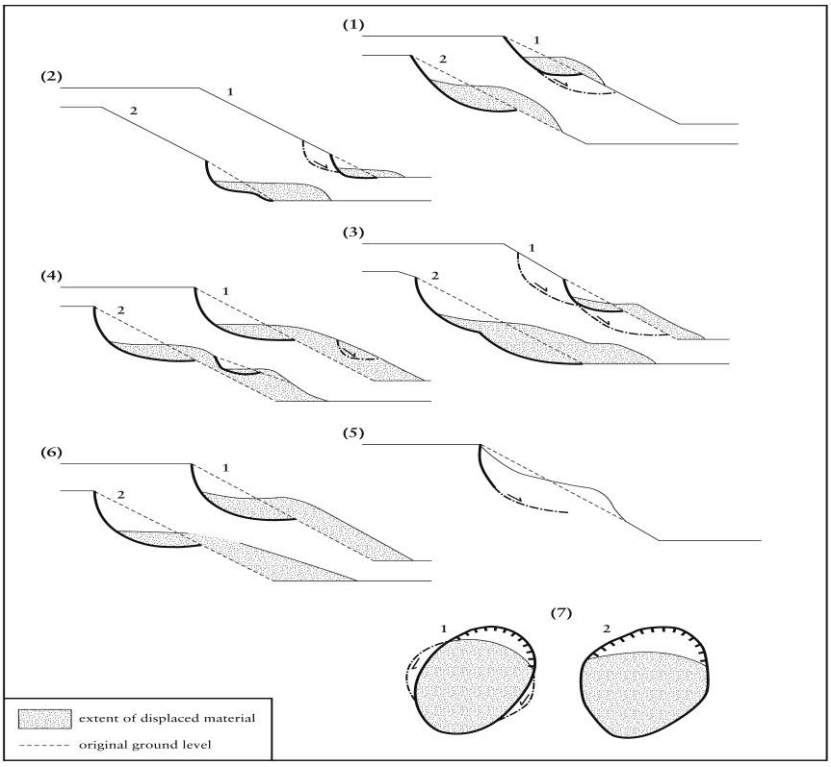


Figure 2.5 Distribution of the activity of landslides.

(See Table 2.5 for an explanation of the numbers and distribution terms, Source: after WP/WLI [19])

[11] State (4) inactive is divided into states (5)-(8).

Table 2.5 Distribution of the activity of landslides (Source: after WP/WLI [19]).

No.	Activity style	Description
1	Complex	A complex landslide exhibits at least two types of movement (falling, toppling, sliding, spreading and flowing) in sequence. In the example shown gneiss and a pegmatite vein toppled with valley incision. Alluvial deposits fill the valley bottom. After weathering had weakened the toppled material some of the displaced mass slid further downslope.
2	Composite	A composite landslide exhibits at least two types of movement simultaneously in different parts of the displacing mass. In the example shown the limestones have slid on the underlying shales causing toppling below the toe of the slide rupture surface.
3	Successive	A successive landslide is the same type as a nearby, earlier landslide, but does not share displaced material or rupture surface with it. In the example shown the later slide $\boldsymbol{AB}$ is the same type as $\boldsymbol{CD}$ but does not share displaced material or a rupture surface with it.
4	Single	A single landslide is a single movement of displaced material.
5	Multiple	A multiple landslide shows repeated development of the same type of movement.

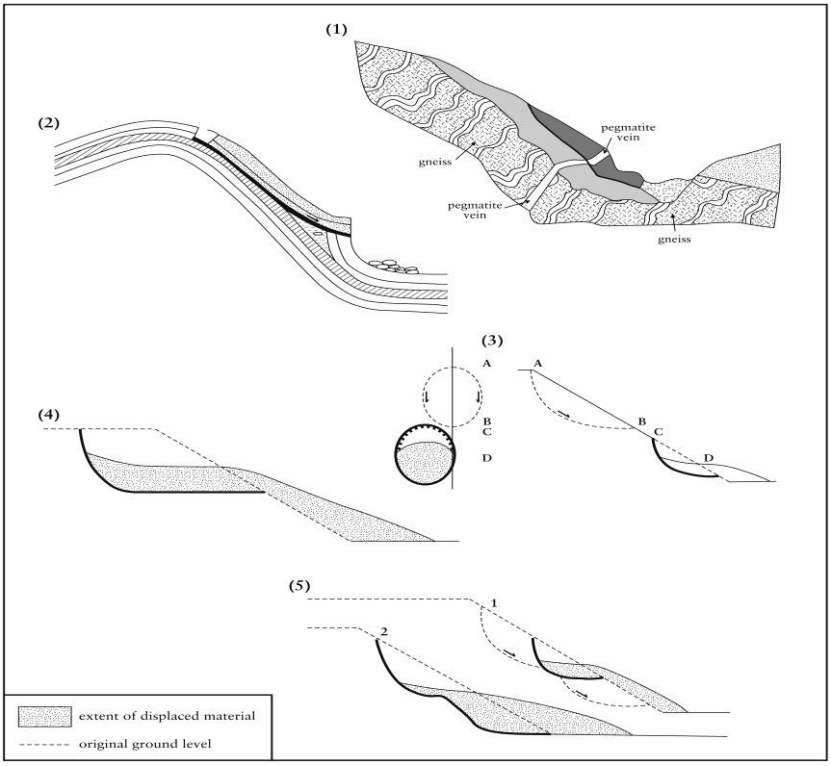


Figure 2.6 Styles of landslide activity.

(See Table 2.6 for an explanation of the numbers and styles terms, Source: after WP/WLI [19]).

Table 2.6 Styles of landslide activity.

No.	Activity style	Description
1	Complex	A complex landslide exhibits at least two types of movement (falling, toppling, sliding, spreading and flowing) in sequence. In the example shown gneiss and a pegmatite vein toppled with valley incision. Alluvial deposits fill the valley bottom. After weathering had weakened the toppled material some of the displaced mass slid further downslope.
2	Composite	A composite landslide exhibits at least two types of movement simultaneously in different parts of the displacing mass. In the example shown the limestones have slid on the underlying shales causing toppling below the toe of the slide rupture surface.
3	Successive	A successive landslide is the same type as a nearby, earlier landslide, but does not share displaced material or rupture surface with it. In the example shown the later slide $\boldsymbol{AB}$ is the same type as $\boldsymbol{CD}$ but does not share displaced material or a rupture surface with it.
4	Single	A single landslide is a single movement of displaced material.
5	Multiple	A multiple landslide shows repeated development of the same type of movement.

Landslide classification in accordance with a system (Figure 2.7 and Table 2.7 - Table 2.8) by combining principally material and movement type, complemented with the estimation of the activity state and velocity.

The existing landslide classification system (is convention based on the accordance that every landslide could be classified) rely on the process, morphology, geometry, movement type and rate, type of material and activity [12, 13, 20]. The most widely used classification scheme was formulated by Varnes [20] and is based on combining material and displacement mechanism (movement type), complemented with the estimation of the activity state and velocity. The scheme was set up according to features that may be observed at once or with the minimum investigation, and without any reference to the causes of the landslide. However, there are exceptions, which make this system more complex (this is the reason behind which the local classifications are occasionally preferred), and encourage its further refinement, since it suffers from a certain simplification and subjectivity, just as any other classification system [21].

Landslide materials define and describe the type of the displaced material in the landslide before it was displaced and they are being classified as follows:

- Rock is a hard or firm mass that was intact and in its natural place before the initiation of movement.
- Soil is an aggregate of solid particles, generally of minerals and rocks that either were transported or were formed by the weathering of rock in place. Gases or liquids filling the pores of the soil form part of the soil.
- Earth describes the material in which 80% or more of the particles are smaller than 2 mm, the upper limit of sand-sized particles.
- Mud describes the material in which 80% or more of the particles are smaller than 0.06 mm, the upper limit of silt-sized particles.
- Debris contains a significant proportion of coarse material; 20% to 80% of the particles are larger than 2 mm, and the remainder is less than 2 mm.

Landslide Velocity is a landslide descriptor that is also governed by material type and movement mechanism and can vary from extremely slow (mm per year in creep) to extremely rapid (m/s in debris flows) (See Table 2.7).

Table 2.7 Landslide velocity scale according to WP/WLI [11] and Cruden and Couture [13].

class	Description	Velocity ($\boldsymbol{mm/s}$)	Typical velocity
7	Extremely Rapid	$5 * 10^{3}$	5 m/s
6	Very Rapid	$5 * 10^{1}$	3 m/min
5	Rapid	$5 * 10^{-1}$	1.8 m/h
4	Moderate	$5 * 10^{-3}$	13 $m/month$
3	Slow	$5 * 10^{-5}$	1.6 $m/year$
2	Very Slow	$5 * 10^{-7}$	16 $mm/year$
1	Extremely Slow		

Displacement mechanism defines the landslide movement typology, which identified according to Varnes [20] into six kinematical distinct types:

- Falls starts with the detachment of soil or rock from a steep slope along a surface on which little or no shear displacement takes place. The material then descends largely by falling, bouncing or rolling.
- Topples are a forward rotation, out of the slope, of a mass of soil and rock about a point or axis below the center of gravity of the displaced mass
- Slides are denoted by slopes failure along one or more, continuous or discrete slip surfaces (i.e. the surface of rupture) where all slope motions are parallel to these surfaces. Depending on slip surface geometry, there exist two types of slides, transitional and rotational:
    - Transitional slides are failures denoted by one or more planar slipping surfaces, along which the slope mass is deformed and usually cease into separated units while moving downward (i.e. downslope). The transitional slide movement is mostly influenced by weak surfaces such as bedding, joints, foliations, faults, shear zones and so forth
    - Rotational slides are unlike transitional slides; develop mostly incoherent, fine and/or homogeneous soil formations (i.e. clayey, shale, marly soil formations, loamy and sandy formations, weathered rocks and soils) along concavely upward curved slip surface in more or less rotational movements on an axis parallel to the slope geometric contour. This usually results in single, multi or even successive rotational slides.
- Flows can be separated into two categories depending on the displacement velocity:
    - Rapid movements of slopes material as a viscous mass on which inter-granular movements predominately overcome shear surface movements.
    - Slow, persisting, and spatially continuous deformation of rocks and soils of the slope.

- There exists a gradual transition from slow to rapid movements depends on various conditions such as water content, materials mobility, and the characteristic of the initial movement[12] .
- Spreads involve liquefaction of slope materials and overall mass by saturating less-coherent sediments to liquefaction (i.e. Liquid state). Spreads are mostly an extension of cohesive soil or rock mass combined with general subsidence of the fractured mass of cohesive material into the softer underlying material. The rupture surface is not a surface of intense shear. Spreads may result from liquefaction or flow (and extrusion) of the softer material.
- Composite or Complex involves a combination of one more displacement mechanism or failure movement types developed within various parts of the slope or at different times.

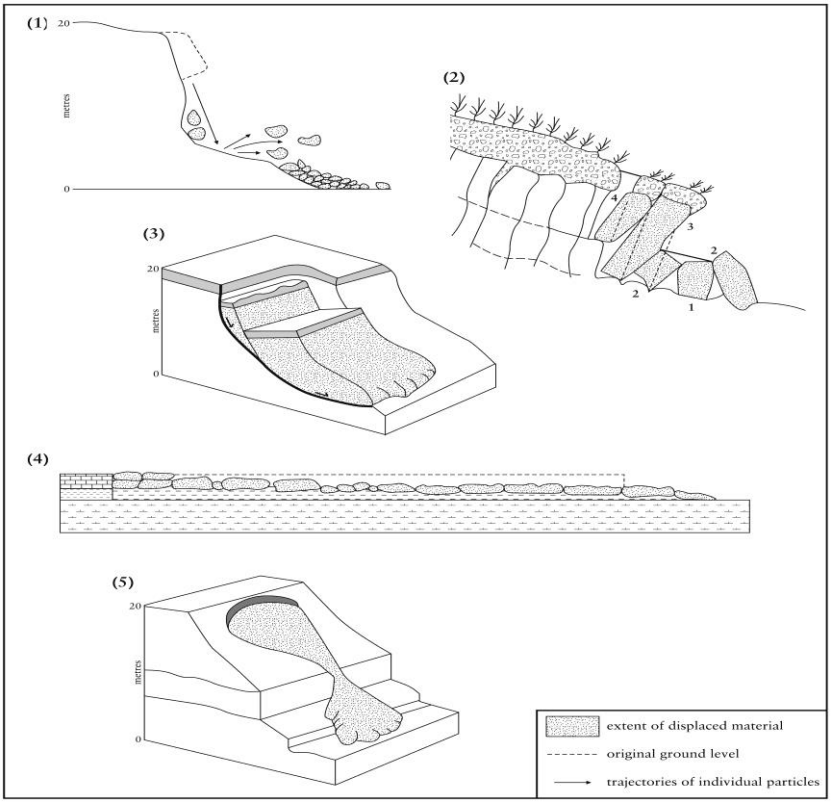


Figure 2.7 Types of landslides.

(Source: after WP/WLI [19]).

[12] For example, debris flow develops from existing slumps that usually generated during slopes mass failure, while advancing downslope.

In the end, by combining the two or more terms the overall classifications (Table 2.8) would be, for example, Rock-fall, Rock topples, Debris-slide, Debris-flow, Earth-slide, Earth-spread…etc.

Table 2.8. Abbreviated types of landslides according to Varnes classification of slope movements [10].

Type of movement		***TYPE OF MATERIAL***		
		Bedrock	***Engineering Soils***	
			Predominantly coarse	*Predominantly fine*
Fall		Rock fall	Debris fall	Earth fall
Topples		Rock topple	Debris topple	Earth topple
Slides	*Rotational*	Rock slide	Debris slide	Earth slide
	Translational			
Lateral spreads		Rock spread	Debris spread	Earth spread
Flows		Rock flow	Debris flow	Earth flow
		(deep creep)	(soil creep)	
Complex		Combination of two or more principal types of movement		

2.2 SUSCEPTIBILITY, HAZARD, VULNERABILITY, AND RISK

2.2.1 Definitions and Scope

Similar to the term landslide the terms hazard, susceptibility, susceptibility map, and risk are intuitively similar and interchangeable with linguistically flexibilities. Therefore, it is pertinent to elaborate and articulate their meaning in the analytical, quantitative framework, which is consistent with the internationally approved terminology of Geo-Engineering communities. From this section and forward, their international approved terminology and articulation will be used only as such in order to avoid any misunderstandings.

Landslide susceptibility (M) is depicted as “a spatial probability of landslide occurrence which analyses and handle the spatial distribution and the magnitude estimation of a landslide which may or may not potentially occurs in a given area” [7]. The landslide magnitude can be expressed by variant landslide descriptors such as total area, volume, relative displacement or velocity...etc. However, the susceptibility is explicitly one single component that handles the spatial probability

in a purely spatial frame, with no temporal component[13]. This introduces other terms as potential complements (sometimes partly match) to the term "landslide susceptibility", such as landslide potential, sensitivity, relative hazard, total landslide density and so forth [18, 22].

Landslide hazard (H) can be explained as "the probability of damaging and/or landslide occurrence in given an area within a given period of time resulting in temporal-spatial probability which also known as the probability of recurrence (P_t)" [10]. Landslide hazard usually considered as the temporal extension to the space component of the susceptibility. This explains the confusion between hazard and susceptibility. However, as long as the temporal component is noticeable a distinction can be made.

Using both landslide magnitudes (i.e. area, volume...etc.) and the probability of its recurrence (P_t), it's possible to estimate using Equation (2.1) [7]:

$$H = M * P_t \tag{2.1}$$

Where: H is Landslide hazard; M is the landslide susceptibility (i.e. landslide magnitude); and P_t is the probability of landslide recurrence.

It should be noted that according to Lee and Jones [18], susceptibility can be depicted as a special case of the hazard in form of a single stack of single dimensions instead of a stack of dimensions (time-series).

The Element at Risk (ER) is the ensemble of any entity or component of the terrain such as human personal, settlements, goods, equipment, infrastructure or even the environment that are potentially affected or endangered by a susceptible damaging landslide hazard. It may involve the following (but not everything):

- Human personal, population, settlements.
- Goods, equipment and objects of personal property.
- Infrastructure and engineering artwork-crafts
- Economic activities,

[13] For example, susceptible slopes will be affected more frequently than less susceptible ones and this may change in future. The less susceptible may became more highly susceptible

- Public services
- The surrounding environment and environmental valuables.

The natural hazard exists only is there exist an element at risk pre-subjugated, or exposed to a potentially by damaging natural event or phenomena. However, natural phenomena remain natural events and/or phenomena unless there's an element at risk present endangered.

Vulnerability (V) is the measure of exposure toward the hazard by expressing the potential exposure damages, which gives the possibility to quantify the degree of loss of an element at risk within the affected or endangered area. Depending on the element at risk, the vulnerability can vary spatially, temporally and individually and even subtypes could be derived for vulnerability at hand [7, 18]. The vulnerability is important parameters to estimate the risk using Equation (2.2) [7]:

$$R = H * V \tag{2.2}$$

Where: R is the risk; H is landslide hazard; V is landslide vulnerability.

Risk (R) can be formulated using Equation (2.2) and it's denoted as a measure of landslide occurrence probability by taking into account the severity of its effects. In reality, risk comprehension is difficult to conceptualize considering the fact that it resides in the future, especially for planning and decision making processes as those require a pre-planning estimation to be able to act upon the risk in advance before disasters strike. Additionally, while susceptibility and hazard analysis are not influenced by the choice of the element at risk, Risk itself is and could be categorized and further segregated into different categories according to the element at risk or even the process at hand[14] [7, 18].

Landslide susceptibility map (LSM) contains a subdivision of the terrain in zones that have a different spatial likelihood that landslides may occur. The likelihood may be indicated either qualitatively (as high, moderate-low, and not susceptible) or quantitatively (e.g. as the density in number per square kilometers, or area affected per square kilometer). Landslide susceptibility maps should indicate the

[14] For example, decision-making require additional subgrouping such as acceptable, tolerable...etc.

zones where landslides have occurred in the past and where they may occur in the future and possibly also the run-out zones.

Landslide inventory is a collection of landslide features in a certain area for a certain period, preferably in digital form with spatial information related to the location (as points or polygons) combined with attribute information. These attributes should ideally contain information on the type of landslide, date of occurrence or relative age, size and/or volume, current activity, and causes. Landslide inventories are either continuous in time or provide so-called event-based landslide inventories, which are inventories of landslides that happened as a result of a particular triggering event (rainfall, earthquake).

All terminology is in accordance with the appropriate conventions[15] [6-11, 16], and due to the nature of the subsumed research work, this thesis will mostly concentrate on susceptibility assessment, while the hazard and risk will be only speculated by their feasibility for further extensions of the research.

2.3 LANDSLIDE ASSESSMENT

2.3.1 Concepts, Principles, and Issues

Landslide assessment can be denoted as "a systematic process of gathering of the available or potential information's, processing and/or modeling using those information's and formulate (forming) a judgment about landslides in a transient work-flow" [18] (Lee & Jones 2004). According to Gerath, Jakob [8] that work-flow usually consists of:

- Initiation (i.e. the definition of the objectives, fundamental details, the scale of the analysis, assessment type and study area).
- Acquisition (i.e. gathering of the required information, data and background information).
- Analysis and modeling (of landslide susceptibility and hazard),
- Evaluation.
- Recommending and advising (i.e. usually optional).

[15] WP/WLI

- Reporting, publishing, and visualizing.

Despite the fact the all Landslide susceptibility studies share this common work-flow template, the choice of assessment approach remain different due to the fact that each case study have its unique set of properties that are different from the others such as the aspect of the problematic to be solved that highly influence choices like modeling approach, micro-processes and sub-stages of the data acquisition and analysis.

Moreover, landslide investigations and landslide assessment revolve around several empirical principles and assumptions articulated in the works of (e.g. Chacón, Irigaray [22] and Guzzetti, Mondini [21]) such as:

- Assumptions:
    - Slope failures do not occur randomly or by chance, but as a result of the conjunction of different conditions, governed by different physical and Geotechnical processes and behaviors.
    - Landslides leave more-or-less distinct footprints (upon activation or after a reasonable period of inactivity) that could be mapped in the field or remotely.
    - Similar landslides may similarly and share common results (i.e. footprints).
    - Implicitly, conditions that are not taken into account in the model do not change systematically in time or space (time-space invariant).
- Principles:
    - The principle of historical recurrence of landslides implies that the landslides share common reoccurring behaviors, patterns, and locations of the one that got activated in the past.
    - The principle of uniformitarianism (past and present are keys for the future) implies that the slope is more likely to fail under the same conditions (that resulted in instability) that led to slopes failure in the past, the present or even the future, at other, environmentally similar locations.
    - Knowledge of landslides of some areas can be generalized and expanded to other areas where similar conditions apply.

It is crucial to understand the limitations and conditions under which all these assumptions apply, and to single-out special cases and exceptions, to reach a common and standardized level for the resulting products: landslide inventory maps, landslide susceptibility maps, landslide hazard maps and eventually, landslide risk maps. These postulates are approved by conventions [6-11, 16], as well as the concepts and methodology which are further to be described.

2.3.2 Issues in Landslide Assessment

Data Acquisition Issues

Hitherto, acquiring data for landslide assessment framework is challenging task highly affected by the initial case study definition, i.e. required scale, level of detail, landslide size, mechanism type, configuration of the terrain, availability of the repositories, and they all bring about specific problematic, precision, accuracy and certainty issues and so forth. Data acquisition sources usually classified in accordance with the type of investigation, methodology, and technology. However, experts, e.g. Guzzetti, Mondini [21], speculate and argue between an old, conventional and new methods for data acquisition classifications.

7. conventional methods:

The methods under this category have been established for a long time and have been proven in practice, but yet suffer from specific limitations. We have:

- Mining and/or investigating of the historical records, this one of the early and necessary steps of any landslide-related endeavor and usually used to get familiar with the facts on the landslides in specific study area and its soundings (i.e. geology, geomorphology, climatology, seismicity, land Use, history of disasters and so forth.), in order to, encompass a wider and better perspective on regional and local conditions in action. Principal investigation of historical records includes analyses of historical topographic and geological repositories, where applicable. Surprisingly, newspaper and diary reports[16] on disastrous events can also be very resourceful, especially for hazard analysis. They can contribute to the existing databases but must be treated with caution and criticism in order

[16] Social media in particular like Facebook groups, Twitter...etc.

to avoid misconceptions, and where applicable, to be confirmed by other plausible resources.

- Field mapping techniques, albeit this being a viable source of information by experts and specialist on slope processes (i.e. geologist, geomorphologist or engineering-geologist), it remains confined by the practitioner's observational field of view such as perspective and point of view[17] which make it rather difficult for a reliable interpretation of larger landslide sites. There exist a high level of uncertainties behind this process as the interpretational subjectivity during map design (estimation of the landslide shapes and spread and their compilation at different scales), which leaves the final result somewhat uncertain. Despite, this issue can be tackled by surveying (i.e. boreholes, laboratory testing, in-situ testing, geophysical surveys and probing), but this additional counter-measure usually introduces additional expenses and constraints to the research budget.
- Visual interpretation of aerial photographs this method relies on using stereoscopic-techniques and equipment which will allow for wider area coverage with a better perspective with the possibility of analyzing different time series and scales. However, it obvious that it will fail under the field of vision obstruction such as vegetation, infrastructural and urban objects (especially for shallow landslides and debris flows). Similar to field mapping, the subjectivity is highly pronounced and noticeable due to the individual visual perception capabilities of the practitioner. Luckily, standard guidelines (i.e. criteria for landslide recognition) do exist[18] , and it's able to limit and reduce these uncertainties to a certain extent.

8. New methods:

[17] In the case of larger or older landslides, might be obscured by the urban or vegetation cover or by more recent geomorphological entities.

[18] Criteria for geomorphological landslide signature exist and usually include: shape, size, tone, color, texture, and pattern of shadows, pattern of objects, overall topography and setting. It is assumed that occurrence of landslides cause's characteristic optical properties of mentioned elements.

9. The methods under this category involve novel complicated approaches complemented by software's and/or hardware in order to exploit the maximum potential, we have:

    - Instrumental monitoring techniques or Field instrumentation (in-situ), often known as geotechnical instrumentation, which undergone a huge leap of enhancements, benefiting from the technological advancements achieved in recent years to achieve near-real-time to real-time data acquisition and distribution of In-situ measurements of displacements with the highest precision (reaching Millimeter precision) providing valuable information's for the assessment of landslide activity such as monitoring the physical parameters of the triggering event of landslides. This highly valuable information's if combined with real-time data distribution, it possible to develop an Early-Warning Systems, crucial for the suppression and mitigation of the landslide risk. The major drawback is the equipment cost, together with the installation and maintenance requirements, and localized information, rarely transferable from one study area to another.
    - Contemporary Remote Sensing techniques (RS), from the most recent perspective, RS was the result of the high technological advancements achieved in satellite technologies thought several Earth Observation programs missions. These missions involve global coverage by multi-channeled sensors, (i.e. multispectral and hyperspectral sensors for visible), infra-red, thermal spectral, as well as microwave sensors with unprecedented precision of spectral, temporal and spatial resolution reaching sub-metric resolution allowing near-real-time tracking of surface deformation and promoting surface-based to sub-surface-based monitoring at desired temporal frequency (i.e. temporal resolution). Thus, allowing gaining more diverse details on the geological and physical conditions of the terrain at hand.
    - Benefits of using RS techniques in landslide assessment are multiple, including, but not limited to:
    - Synoptic view,
    - Geo-referenced data,

- Lower expense of research,
- Encouraged raster modeling approach,
- The possibility of quantitative modeling method implementation (pixel and object-based classifications implementation,
- Pixel and object-based classifications through a combination of advanced statistics and Machine Learning with GIS) and therefore reduced subjectivity in design, the possibility for urgent response and Early-Warning Systems for disastrous landslide events, even enabling on-screen visual 2-3D analysis, via special hardware and/or software configurations [21].
- Special attention in the most recent technology is drawn by the unmanned vehicles and micro- vehicles, which are capable of producing high-resolution imagery at extremely low cost.
- Limitations, on the other hand, are mostly technical:
- Unavailability of the specific sensor at the site (particularly, pricey and rare airborne and/or terrestrial LiDAR and SAR data),
- The relatively short operational history of RS programs (only several decades, through which the data are not entirely consistent in terms of resolution and other technical features), and therefore limited applicability for temporal (hazard) framework.

- Surveying is rather experimental for now despite equipment's has experienced faster acquisition time with sufficient precision and tends to provide improvements from several aspects. Mostly via Global Navigation Satellite System (GNSS) and synergy of Photogrammetric and high-resolution optical imaging (terrestrial, airborne and satellite). A huge benefit for such a method is allowing very precise systematic surveys by easily assembled alongside the standard in-situ instrumentations. Despite being experimental, the main drawback or limitation of this method[19] is being highly dependent on the terrain physiographic condition (i.e. configuration and setting, screening by vegetation cover and urban

[19] In the context of surveying framework.

objects), and the engagement of the practitioner, making it time-consuming, resource-intensive and terrain specific (change this).

It is probably the most advisable to combine as many of the acquisition techniques as possible and never to rely entirely on a single one. Those older, conventional methods, especially aerial photography interpretation, are not to be neglected among acquisition techniques and should be cherished in the landslide assessment practice [21]. Novel techniques, which are developing toward automatic (semi-supervised) landslide mapping, will hardly reach sufficient levels of certainty since they face different, non-compensable limitations, unlike visual, expert-driven interpretation.

Modeling Approach Issues

Once the data are fully prepared and preprocessed (i.e. selection and structure), they fed directly the proper modeling method. However, the choice of the method strongly influences the quality and type of outcomes. Based on the model's predictability we can separate models into two distinct and different cases:

- Predictive models are the type of models that are generally based on non-linear supervised classification problems upon spatial and/or temporal references (i.e. spatial and/or temporal conditions) that can be related to past landslide occurrence (and even several generations of past occurrences) within given area to predict future events occurrences and localize endangered susceptible zones. These models usually require fulfilling general assumptions principles and to apply noted in (see the postulates in Chapter 2.3) and specific structure and type of the input data such as the availability of thematic preferences called variables or conditioning factors and reliable landslide inventory or multi-temporal inventory. Even though the resulting model provides numerical, i.e. quantitative measurements (usually in form of spatial and/or temporal probability of occurrence), relative scoring is yet preferred due to many

assumptions that trouble the quantitative way of expressing the landslide susceptibility, risk or hazard[20].

- Non-Predictive models, this kind of models are very different than predictive models as they tend to spatially analyze the relationships among the different thematic variables and analyze their overall influence and contribution to the landslide susceptibility, hazard or risk by figuring out the relation between the condition factors and the landslide occurrence in a statistical manner using various statistical relations and methods. These models are very simpler, i.e. less computation cost and time compared to predictive techniques, but comes up with quantified values of the individual impact of each factor. However, these methods tend to decrease the certainty by surpassing some of the assumptions that are commonly made in the predictive modeling, as they tend to introduce additional uncertainties through the data preparation, due to empirical and/or arbitrary rearrangement of the raw data (slicing, ranging of continuous data into intervals, transforming the data, quantifying non-numerical data and so forth). The most important advantage of this approach is the quantitative nature, which is relatively easy for comprehension to non-landslide experts, planners and decision-makers [6].

One can alternatively discuss the modeling choice and briefly the problems that come with it by accommodating a more conventional perspective. For example, based on the method of treating and handling the landslide assessment, we can denote:

- Direct methods are the expert-opinion-driven approach that relies on expert evaluation of the direct relationship between conditioning factors and landslides occurrences during a survey campaign on the site of failure.
- Indirect methods rely on mapping and analyzing a different set of conditioning factors and their relative contribution to the occurrence of

[20] Some assumptions are taken into account but some of the uncertainties usually remain unconsidered, and it is therefore disputable to measure susceptibility and/or hazard in absolute quantitative scale.

slopes failure resulting in a relationship between the landslide condition factors and landslides occurrences.

In reality, is a combination of both methods is made in order to determine the conditions under slopes failures occur. However, the most usual classification of the methodological model approaches (Table 2.9) is the following:

- Heuristic or Expert-driven approach is an expert's opinion-driven approach of weighting conditioning factors that relate to a landslide inventory in order to determine landslide zonation. The weighting process is achieved through the hierarchical leveling process of the landslide conditioning factors [23]. Usually, this process is a combination of direct mapping analysis and qualitative map combination. The former, determine the susceptibility straightforward in the field which is based on individual experience. The Latter, experts use their knowledge to determine the weighting value for each class parameter in each conditioning factor.
- Physically-based or Deterministic approach, highly focus on quantitatively generating an index called "stability index" by calculating the "safety factors". This involves some complicated evaluation of safety factors that required detailed measurement of a handful of parameters that influence slopes. On top of the measurement being in-situ specific, the calculation is made on the assumption that whole research area is moderately homogeneous and the existing landslide types are simples making this approach relatively in-situ specific and only validate over small areas [24].
- Overall, the deterministic approach in landslide assessment has been pioneered relatively early by Montgomery and Dietrich [25] and there have been several succeeding developments involved. They all gradually perplexed the model and introduced more variables, by decreasing the number of approximations, but their reach inapplicability has been disputed, i.e. limited to a very specific, homogeneous ambient and conditions, very rarely present in actual terrains.
- Statistical approach is rapidly evolving and expanding in terms of producing an objective landslide hazard assessment [26]. Methods and models based on this approach are based on the assumption that "previous,

current and future landslide failures do not happen randomly or by chance, but instead, failures follow patterns and share common geotechnical behaviors under similar conditions of the past and the present" [3], which require, collecting and preparing an accurate database, i.e. a geospatial database consist of an inventory map of past and present landslide occurrences and set and/or combination of conditioning factors, with maximum details available. Then, models based on these methods are trained and validated using that database and the resulting models are used to generate landslide occurrence probabilities in order to forecast the future landslide's areas using past and present landslide occurrences [2, 4]. Unlike other landslide assessment approaches[21], statistical approach, in particular, is able to extracts and obtain a relationship that relates landslide occurrence to the conditioning factors very efficiently for large scale analysis (depending on data availability it may relate the values, distributions, aggregations and other data features), which introduce an objective prognostic dimension to the implemented model[22], especially if an advanced methods such as Machine Learning (ML) is implemented which, can introduce more depth to statistical approach by incorporates a broad range of complex learning procedures that are effective in solving problems of landslides such as susceptibility assessment. This modeling capability can be highlighted in three main areas [27]:

  - The system's deterministic model is computationally expensive and ML can be used as a code accelerator tool.
  - There is no deterministic model but an empirical ML-based model can be derived using the existing data.
  - Classification problems.

- Despite that, it is critical to understand that with recent advancements in computer science in the last decade or so, it became much difficult to subcategorizing ML under a more general approach such statistical approach as the boundaries became more blurry due the fact the ML model

[21] As explained in Chapter 2.3.2.

[22] Although the prognosis only spatial and not temporal.

became interdisciplinary and built on many different concepts, such as probability and statistics, artificial intelligence, information theory, as well as philosophy, psychology, neurobiology and so forth [28-30]. As this debate is totally out of the scope of this thesis research as it depends on the field and the problem to solve. However, for Geoscience problems in general and landslide susceptibility assessment problems in particular, ML are still considered under the umbrella of statistical approach, due to the common requirements, limitations and more importantly the lack of clear evidence in literature about ML singled-out as a separate independent approach by itself or even deserves a slightly higher hierarchical position among the other approaches.

Table 2.9 A brief summary of the available landslide assessment modeling approaches.

Modeling approach	Description Summary
Heuristic or expert-driven	Use thematic data (variables such as geological, geomorphological, Land use, infrastructure and so forth) and suffer from uncertainty related to the subjectivity of the practitioner in both, data preparation and modeling itself rending the approach more of "opinion-oriented"23method.
Statistical	They can suffer from uncertainty due to the data preparation, but the tendency of using advanced techniques, such as Machine Learning algorithms, might be helpful due to their capability of canceling-out these sources of uncertainty.
Physically-based or Deterministic	Regard only the simplest mechanisms and introduce numerous assumptions into the modeling [25], thus their uncertainty is relatively high. In regional scales implement such an approach is not feasible.

In respect to the preceding passages, only heuristic approach methods can be qualified as a direct method (it can be an indirect method but to a limited extent). On the other hand, only statistical approach methods can be qualified as a predictive approach, but could also be enlisted among non-predictive, while the remaining three only qualify as non-predictive approaches. Only statistical approach based methods seems to be a viable option to use, especially the more advanced (predictive) models such as ML turn out to be the most promising and least limited for the exploration because physical-based models are capable of delivering the highest prediction

[23] The main problem of this approach is in determining the exactly weighting value.

accuracy require a fair amount of detailed data information's to provide reliable results, which is unbelievably expensive and heuristic or expert-driven approach is limited and very controversial as it requires expert-opinion, which make the model unreliable due to uncertainties and subjectiveness.

GIS Issues

During the past few decades, a huge advancements and improvements were achieved in computer science, specifically, the computational capabilities making GIS more affordable and widely available, especially, when GIS offers more to the plate by introducing new unparalleled tools and possibilities for better data manipulation and advanced modeling opportunities such as surface features extraction (e.g. morphometric evaluation) and creation of novel thematic spatial layers, unachievable through conventional and analog practices [31]. This can be particularly handful, in large scale landslide analysis[24] as opposed to site-specific analysis[25].

This breakthrough in GIS, in particular with the continuous evolution of raster formats[26] led to numerous advanced GIS platforms and modules that speed up the process of implementing and introducing newer novel advanced and hardware-demanding algorithms and techniques such as ML algorithms. These frameworks and/or modules are so beneficial as they guarantee the option of working under the familiar environment for the single practitioner or assembling a cross-disciplinary team of practitioners which result in ensuring better communication, better interoperability, better results and eventually extending the overall analytical power (of the team). One of the main benefits of GIS is the ability of visualization of both, the input data and the resulting models making landslide related information's became much better disseminated and visualized by offering the possibility of not only visualizing locally dense spatial information's in multidimensional 2/3/4-D

[24] Where some conditioning factors are preferred than others as they fully benefit from GIS capabilities, e.g. geology, altitude derivatives, landuse...etc.), as opposed to site-specific analyses.

[25] Because geotechnical parameters are required and sampled through a series of instrumental measurements and laboratory tests, it doesn't not fully benefit from GIS capabilities,

[26] Made a major breakthrough for implementation of these advanced methods due the fact that raster formats are analyzable, synthesizable, decomposable, combinable, scalable, in other words, fully spatially operable 32. Bonham-Carter, G.F., *Geographic Information Systems for geoscientists-modeling with GIS*. Vol. 13. 1994: Elsevier BV. 398-398..

displays but also the possibility of implementing and using global web-GIS systems (i.e. Google Earth, Bing Maps,...etc.) or local GIS portals[27] [21]. Moreover, the possibility of performing geostatistical analysis and database pre-processing in a GIS-ready environment helps in determining the relationship between slope failures and conditioning factors that used to generate landslide susceptibility maps.

Furthermore, the usage of different data resources, types and scales introduces a lot of issues concerning data quality and compatibility. Thus, makes the process of fitting the data for research with specific interest difficult to achieve. Combined with the increasing "user-friendliness" of GIS platforms and modules that tend to neglect the input data quality issues and rather focus on introducing complex and sophisticated data manipulation and model implementation [31]. In reality, no matter how the data manipulation or modeling technique became sophisticated, they can never truly compensate for the inadequate scale and quality of the input data, due to an intrinsic error that is continually replicated within the model. On the other hand, data availability and open-source policies are one of the most significant issues in research budget design, and lack of affordable data could lead to decreasing in the assessment quality, but this is rather financial than a scientific issue to discuss.

Other Issues

Despite, the most critical concerns in the landslide assessment framework are discussed in the past couple of preceding passages, there is still a suffice of other issues in landslide susceptibility assessment problematic, ranging from scientific, practical, technical, to social speculations. For example, uncertainty is definitely one of the critical issues for reliable landslide models as it can be related to the data, the modeling procedure and approach choices event from the case study local environmental conditions (i.e. real-world conditions). The former two were partly discussed before (see Chapter 2.3.1 and 2.3.2), but some specific details are to be emphasized:

- Fuzziness and randomness are technically omnipresent in the input dataset and can be highly pronounced in noised, biased or skewed dataset resulting in introducing more uncertainties to the overall uncertainty of the landslide susceptibility model. Fuzziness is known for adding local imprecision,

[27] Usually portals of governmental agencies and administrations.

while randomness prevents generalizing the regular landslide patterns available in the input data distribution.

- Incompleteness depends on the level of oversimplification in the modeling stage. As an example, landslide assessment relies on many fundamental assumptions (see the postulates in Chapter 2.3) that highly subjugated to different degrees of simplification[28]. Additionally, the inclusion and/or exclusion of unimportant conditioning factors in the input dataset highly affect the incompleteness. Yet this issue is inevitable from a technical standpoint, because till during date, no agreement on whether any data shall be excluded, even if biased. On the other hand, some data are not excluded on purpose, but due to the lack of resources for the corresponding phenomenon or it is simply unforeseen as a pertinent factor by mistake or insufficient knowledge [5, 31, 33].
- Environmental or real-world uncertainty, are theoretically unpredictable as it involves various consequential actions in the past, present or future either consciously, subconsciously or unconsciously of different entities such as agencies, administrations, public, private, collectives, or individuals that may directly or indirectly drive even high-quality predictions off the course [18].

Another peculiarity due to the resourcefulness of the data comes along with the rising popularity of RS products in landslide assessment. Usage of raw products is the easiest but irresponsible solution since each one of them contains intrinsic noise, which foremost requires determination of its type, quantity, and propagation. Subsequently, noise filtering is managed through image preprocessing, i.e. pan-sharpening, ortho-rectification, co-registration, and radiometric correction, in this respective order [34]. Working with initial noise is qualified as a systematic error and will affect, perhaps even sabotage the model. However, that leads to a very important issue that is concerning obtaining reliable landslide susceptibility maps is the quality check of input dataset and output results. Data quality check is obviously the first step, particularly due to the plenitude and abundance of resources that limit the possibility of standardizing and objectifying the quality check. Carrara and Pike [31]

[28] Highly noticeable in the deterministic approach models.

suggest at least two basic quality check requirements should be met for successful landslide susceptibility modeling:

- The appropriate strategy for model performance evaluation.
- The actual valorization of the model.

In the end, most of the times practitioners are more involved in their modeling choices in order to suit the universal circumstances by assuming that the best model is the most complex and robust one [5, 31, 33]. Instead, the above-mentioned result and data quality check might be an apt response to the particular problem posed before them.

Chapter 3: Methods & Procedures

This chapter is structured into several sections that will focus and depict different methodologies and their implementation at different stages of the research, i.e. conditioning factors analysis methods, landslide assessment methods, resampling strategy methods, model optimization and tuning methods and model performance evaluation methods.

Instead of a general style that could be found elsewhere, in various textbooks and articles, this chapter is explaining these different methods in the light of the GIS landslide assessment, using respective examples and descriptions, which brings the topic closer and more comprehensible, especially outlining the unique features of each model and technique used in this research with a specific attention to landslide susceptibility paradigm. Furthermore, a special section is devoted to the research workflow at the end of this chapter, in order to present and describe additional details of the used methodology.

3.1 CONDITIONING FACTORS ANALYSIS METHODS

This section of methods is featuring Objective 3 (see Chapter 1.3).

It's widely disputable whether a conditioning factor is actually contributing to the landslide susceptibility model by enhancing it or biasing it. To solve such an issue usually attribute screening or optimization is the perfect candidate for such a problem, especially is it's done correctly[29]. For that reason, Pearson Correlation Coefficient analysis and Variable Inflation Factor analysis was opted for this case study, for the purpose of evaluating, demonstrating and highlighting the suitability of the underlying assumption used to select the conditioning factors based on the non-independence among factors[30] and exclude, if exist, the problematic conditioning factors.

[29] It is arguable whether it should be used, especially in multivariate framework.

[30] Mostly to prove their statistical independency to the landslide inventory.

3.1.1 Pearson's Correlation Coefficient

Pearson's correlation coefficient (PCC), also referred to as the Pearson Product-Moment Correlation Coefficient (PPMCC) or the bivariate correlation, can be denoted as the covariance of each pair of conditioning factors divided by the product of their standard deviations. For each given paired data $\langle (x_1, y_1), \dots, (x_n, y_n) \rangle$ consisting of n pairs, the PCC is defined according to Equation (3.1) as:

$$r_{x.y} = \sum_{i=1}^{n} \frac{x - \overline{x}}{\sqrt{\sum_{k=1}^{n} (x - \overline{x})^2}} * \frac{y - \overline{y}}{\sqrt{\sum_{k=1}^{n} (y - \overline{y})^2}} \tag{3.1}$$

Where: r_{xy} is the sample correlation coefficient (also known as the sample Pearson Correlation Coefficient); n is the sample size; x_i, y_i are the individual sample points indexed with i; $\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ (the sample mean); and analogously for $\overline{y}$.

The obtained values of PCC indicate the extent to which two conditioning factors x and y are linearly related. This value varies between -1 and 1. A value of $PCC = 1$ indicates there's is a total positive linear correlation that implies a linear equation that describes the relationship between x and y perfectly, with all data points laying on a line for which y increases as x increases. On the other hand, a value of $PCC = -1$ indicates there's a total negative linear correlation, which implies that all data points lay on a line of which y decreases as x increases and value of $PCC = 0$ implies that there is no linear correlation between the two conditioning factors. Generally, note that $PCC > 0$ if and only if x and y lay on the same side of their respective means. Thus, the correlation coefficient is positive if x and y tend to be simultaneously greater than, or simultaneously less than, their respective means. The correlation coefficient is negative (anti-correlation) if x and y tend to lie on opposite sides of their respective means. Researches widely agree that $PCC \geq 0.7$ which indicates a high correlation between each pair of data [35].

3.1.2 Variance Inflation Factor

The Variance Inflation Factor (VIF) is the ratio of variance in a model with multiple conditioning factors, divided by the variance of a model with one

conditioning factor alone. It quantifies and then detects the severity of multicollinearity in regression analysis. The VIF estimates how much the variance of a regression coefficient is inflated due to multicollinearity in the model according to Equation (3.2):

$$VIF_i = \frac{1}{1 - R_i^2} \tag{3.2}$$

Where: R^2 is the R-squared value and i is the predictor of interest (i.e. conditioning factor). Some statisticians suggest using the tolerance (TOL) instead of VIF, where TOL is:

$$TOL_i = \frac{1}{VIF_i} = 1 - R_i^2 \tag{3.3}$$

VIF values range from 1 to $+\infty$, and according to Marquaridt [36] VIF values can be interpreted as not correlated variables if $VIF = 1$; moderately correlated variable if $1 \geq VIF < 5$ and highly correlated variable if $VIF \geq 5$. Predictors or conditioning factors with $VIF \geq 5$ are not safe to use and highly indicate a severe multicollinearity [35-38].

Sometimes, there's no reason for concern at all if VIF is too high. For example, you can get a high VIF by including products or powers of other variables as conditioning factor, say x and x^2. Usually, dummy variables representing categorical variables with three or more categories show high VIFs, but those are usually not a problem. However, if VIF is regarded as being too high for variables, the solutions are to:

- Obtain more data, so as to reduce the standard errors.
- Use techniques designed to work better with high VIFs, such as Shapley regression (note that such techniques do not actually solve the VIF problem but instead ensure that the estimates are more reliable, i.e., consistent).
- Obtain better data, where the predictors are less correlated (e.g., by conducting an experiment)

- Recode the predictors in a way that reduces correlations (e.g., using orthogonal polynomials instead of polynomials).

3.2 LANDSLIDE ASSESSMENT METHODS

This section of methods is featuring Objective 4 (see Chapter 1.3).

Landslide assessment methods used in this thesis and to be presented in detail involve numerous statistical techniques with a particular focus on the ML techniques.

First of all Machine Learning (ML), represents an emerging field of computer science which studies computer algorithms that improve automatically through experience [28-30]. Technically, ML models and algorithms[31] are considered "universal approximators" that learn from machine-readable data in order to provide multivariate, nonlinear, nonparametric regression or classification analysis. This imply that ML-based models are capable to learn the underlying patterns and behaviors of a system, i.e. landslides susceptibility, from a one or more sets of constructive comprehensive examples that cover the input space (i.e. input dataset) called "training dataset" and validate these models against a completely independent random subset of the input dataset. Thus, it makes ML as one of the most effective methods for solving non-linear Geo-spatial problems like landslides susceptibility using either regression or classification. This learning concept is different than any of the mentioned modeling approaches (as explained in Chapter 2.3.2),

Since ML-based models thrive on the benefits of statistical elements and depend on the statistical approach foundations. Therefore, it is necessary to acquire a significant number of conditioning factors to obtain reliable results. So this makes ML models capable of introducing an objective prognostic dimension[32] to the implemented landslide susceptibility model, and depending on the model, it may empower some additional predictability to the spatial domain. After all, ML has proven to be ideal for addressing large-scale analysis problems where theoretical knowledge about the problem is still incomplete [27], especially when prior knowledge about the nature of the relationships between the input data and

[31] For example, neural networks, support vector machines, self-organizing map, decision trees, random forests, case-based reasoning, genetic programming…etc.

[32] Although the prognosis only spatial and not temporal.

conditioning factors are not required by the ML-based techniques. That being said, in a perfect situation where we had a complete theoretical understanding of landslides, ML would be superfluous.

In literature, several studies have been able to implement and compare ML models in landslide susceptibility modeling. Nevertheless, no solid agreement about which method or technique is the best for landslide-prone areas prediction has been identified [31]. Thus, the prediction accuracy of landslide modeling is influenced by not only quality of landslide inventories and influencing factors, but also the fundamental quality of the ML algorithm used [4, 39-41]. Therefore, assessing and comparing the prediction capabilities of advanced ML methods for landslide susceptibility should be carried out.

3.2.1 Learning Problem

There's exist a plethora of algorithms, models and hybrid combinations of these techniques, that can be successfully implemented in ML to predict landslides and even understand the triggering mechanism behind it due to the modeling capabilities offered by ML. Yet, that depends on the learning problem (or task) at hand. This can be clustering, classification or regression. Herein, for landslide susceptibility analysis the learning problem is strictly classification issue which limits the possibilities only to classification-related algorithms that are capable to produce classification models.

First and foremost, explaining the learning problem in more comprehensive mathematical details helps in illustrating the landslide assessment framework. This will be crucial as it will be helpful in taking full grasp of the upcoming descriptions about the different ML algorithms. Essentially, the learning problem in the landslide susceptibility analysis framework is an automated procedure that assumes that after the initial acquisition of the necessary spatial data, i.e. input dataset, a set of data (from the input dataset) that represent the study area called "training dataset" is fed to a model that rely on supervised learning approach in order to learn landslide patterns in the representative training dataset of study area by relating the learning instances found in the training dataset, i.e. landslide presence, and the set of conditioning factors prepared for the case study. Afterward, the model generates a learning rule that can be extrapolated to the rest of the study area and thus resulting in an automated prognosis of the spatial distribution of landslides.

That being said, it is necessary to assume that the input data, i.e. appropriate conditioning factors (geological, morphometric, environmental) and the referent landslide inventory map, are presented in a 2-D raster format. The inventory is used as a reference in the evaluation process of the models. The input rasters are spatially overlapped in the way that each grid element (i.e. pixel), represents a data instance at a given point location in the study area. This obviously initiates a classification task classifies that classifies and assigns each pixel into an appropriate landslide category using the conditioning factors values associated with that pixel. The task applicable only to the remaining area usually called the testing area (the area that has not been assessed by an expert).

To simply things up, if ML models rely on an input of 2D rasters that represent the conditioning factors and the referent landslide inventory map[33] (as required by all statistical-based approach methods), so by overlapping the set of conditioning factors on top of each other along with the inventory map, it is possible to obtain grid of pixels, that for each pixel have set of data instances at given point in the study area. This will be the foundation for the supervised classification task that will classify each pixel to the appropriate landslide category class[34] based on the learned landslide patterns[35] from the training dataset. Then, the generated model is used to predict the remainder of the study area[36]. Therefore, the corresponding learning problem could be formulated mathematically as follows:

Suppose $x \Leftrightarrow x \in \mathbb{R}^n$, where $\mathcal{P}$ are all pixels instances extracted from all conditioning factor rasters for the study area and x an n-dimensional real vector of $x = \{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n\}$ and $\mathcal{X}_i$ represents the value of the i_{th} conditioning factor associated with the pixel x_i. Further, let $C = c_1, c_2, \dots c_j$ be the set of j disjunctive, predefined landslide classes[37]. A function $f_c(\mathcal{P}) \rightarrow C$ is called a classification if for

[33] Hypothetically required, only during the training stage, but also required as a reference during the validation stage (i.e. model performance evaluation).

[34] Depending on the underlined objectives of the landslide susceptibility analysis, landslide classes and groups can vary according measurement, activity, state etc....In this case study its binary of "Yes" or "No" for landslide presence and absence, respectively.

[35] Patterns learned from the data instances sets of each landslide pixel.

[36] Unseen data usually called testing data or area.

[37] In binary classification (similar to this case), $j = 2$.

each $x \in \mathcal{P}$ it holds that $f_c(x) = c_j$ whenever a pixel x belongs to the landslide susceptibility class c_j. If, $\mathcal{P}_{train}$ is the training set and $\mathcal{P}_{test}$ testing set, then the learning problem (ML model) has the objective to approximate a function $\hat{f}_c$ using only the samples instances available in the training set $\mathcal{P}_{train}$ and a specific learning method to approximate as closely as possible to a real, unknown function f_c.

3.2.2 Learning Methods

As mentioned before, ML includes a variety of algorithms and over the last decade or so, there has been considerable progress in developing ML-based methodologies for many of landslide susceptibility modeling. In fact, in this thesis, some advanced state-of-the-art methods and techniques and models regarding landslide susceptibility will be presented hereinafter. Furthermore, models and methods are presented to demonstrate the efficiency of ML for tackling the landslide susceptibility assessment problems.

Random Forest

Random forest (RF) is an ensemble approach of decision trees such that each tree fits a data subset sampled independently using bootstrapping [42, 43]. In fact, RF is able to perform binary classification tasks by growing trees using an input vector or dataset. Each tree provides a "vote" for either as "Yes" or "No" class. Then, the final classification decision is voted based on overall forest trees votes. Yet, the rationale behind trees growing process is rather simplistic. If by assuming that N is the number of cases in the training set; M is the number of variables or conditioning factors available in the input dataset and m is the number of variables[38] drawn randomly out of M, only if "$m \ll M$" is satisfied. Then, the training sets for growing the trees are generated by bootstrapping the original input dataset (i.e. random sampling with replacement) to obtain N sample cases, which will ensure growing each tree to the fullest and largest extent. Therefore, "no pruning" will be available during the trees growing process [44].

As mentioned above, Bootstrapping [45, 46] formulate a fundamental building block for RF, that can be defined as: "a general-purpose sample-based statistical method in which several (non-disjoint) training sets are obtained by drawing

[38] Held constant during the forest growing.

randomly, with replacement, from a single base dataset" [47], where the general procedure is described in Algorithm 3.1.

Algorithm 3.1 Bootstrap procedure for classification

input : Size-N sample $Z = z_1, z_2, ..., z_N$ of a (potentially infinite) population P.
B, number of bootstrap samples.
output: Estimate $\hat{T}(P)$ of the population statistic.

```
for b = 1 to B do
    Draw, with replacement, N samples from Z, obtaining the b-th
     bootstrap sample Z*_b;
    foreach sample Z*_b do
        Compute, the estimate of the statistic T̂(Z*_b)
    end
end
Compute the bootstrap estimate T̂(P) as the average of
  T̂(Z*_1), ..., T̂(Z*_B);
Compute the accuracy of the estimate, using e.g., the variance of
  T̂(Z*_1), ..., T̂(Z*_B);
```

Using Bootstrap on input dataset of N samples, the probability of each instance being selected is $1/N$, this implies that after N draws a given instance have the probability 0.368% not being selected (following Equation (3.4)), which mean that each sample contains roughly 63.2% of the instances [47].

$$(1-\frac{1}{N})^N \approx exp(-1) \approx 0.368 \tag{3.4}$$

In literature, RF is able to provide a robust error rate with respect to outliers in predictors due to features random selection at each split node by depending on two data objects namely, Out-Of-Bag (OOB) and proximities [44]:

- OOB data is used to get both variable importance estimations and an internal unbiased OOB error (the classification error) as trees are added to the forest, while bagging is used to randomly select samples of variables as the training dataset for model calibration. For each variable, the function determines model prediction error if the values of that variable are permuted across the OOB observations [48].
- Proximities, on the other hand, are used to replace missing data, locating outliers, and producing illuminating low-dimensional views of the data and

can be only calculated only after each tree is fitted on for each pair of cases then normalized by dividing over the total number of fitted trees [42, 44, 49].

In the original paper on RF, it was shown that the forest error rate depends on the correlation between any two trees in the forest and the strength of each individual tree in the forest[39]. As a result, reducing m reduces both the correlation and the strength. Increasing it increases both. Somewhere in between is an "optimal" range of m. Using the OOB error rate, a value of m in the optimal range can quickly be found. This is the only adjustable parameter to which random forests are somewhat sensitive.

Gradient Boosting Machine

Gradient Boosting Machine (GBM) or simply Gradient boosting is an ensemble of weak learners (WL)[40] typically regression trees or decision trees, that cast boosting as a numerical optimization problem by adding weak learners using a functional gradient descent associated with the whole ensemble to minimize the loss function [51-55]. The rationale behind GBM is that the learning process consecutively introduces weak learners using a functional gradient descent in a stage-wise additive approach sequentially allowing the algorithm to enhance the overall accuracy simply by readjusting previous error terms when new weak learners are added [54]. Thus, it makes GBM particularly attractive and compelling not only because of the practical performance, but also the several theoretical and algorithmic features introduced such as the freedom of choice of base learners and criterion for updating the weights of the training samples. Thus, introducing different boosting algorithms models platforms [56-58].

Boosting which is the essence of GBM is in fact repeatedly using WL algorithms and models as a base on differently weighted versions of the training data

[39] Increasing the correlation increases the forest error rate and a tree with a low error rate is a strong classifier. Therefore, increasing the strength of the individual trees decreases the forest error rate.

[40] A weak learner (WL) is a learning algorithm capable of producing classifiers with probability of error strictly (but only slightly) less than that of random guessing (0.5, in the binary case). These concepts are rooted in the theory of PAC (probably approximately correct) learning 50. Valiant, L.G., *A Theory of the Learnable.* Commun. ACM, 1984. **27**(11): p. 1134-1142.. On the other hand, a strong learner (SL) is formally defined in a similar way as weak learner, is able (given enough training data) to yield classifiers with arbitrarily small error probability.

and yielding a sequence of WL that are combined into an ensemble. The weighting of each instance in the training data at each round of the algorithm depends on the accuracy of the previous classifiers. Thus, allowing the algorithm to focus its attention on those samples that are still incorrectly classified. This was proved by Schapire [58] in Algorithm 3.2.

Algorithm 3.2 Boosting procedure for classification.

---

input : Dataset $Z = z_1, z_2, ..., z_N$, with $z_i = (\boldsymbol{x_i}, y_i)$, where $\boldsymbol{x_i} \in \mathcal{X}$ and $y_i \in \{1, +1\}$.

output: A classifier $H : \mathcal{X} \to \{1, +1\}$.

1 Randomly select, without replacement, $L_1 < N$ samples from Z to obtain Z_1^*;

2 Run the weak learner on Z_1^*, yielding classifier H_1;

3 Select $L_2 < N$ samples from Z, with half of the samples misclassified by H_1, to obtain Z_2^*;

4 Run the weak learner on Z_2^*, yielding classifier H_2;

5 Select all samples from Z on which H_1 and H_2 disagree, producing Z_3^*;

6 Run the weak learner on Z_3^*, yielding classifier H_3;

7 Produce the final classifier as a majority vote:

$$\boldsymbol{H(x)} = \text{sign}(\sum_{\boldsymbol{b=1}}^{\boldsymbol{3}} \boldsymbol{H_b(x)});$$

Historically, GBM was recast in a statistical framework first by under the name of ARCing algorithms [59] involving three elements:

- (1) Loss function to be optimized based on the objective function to be solved;
- (2) Weak learner to make predictions specifically a decision trees that are constructed in a greedy manner by choosing the best split points based on specific scores; and
- (3) an additive model to add weak learners to minimize the loss function, therefore a weighted combination of classifiers that optimizes the cost using gradient descent in function space [60, 61].

Logistic Regression

Logistic regression (LR) is a particular case of the generalized linear model [62] configured to provide a binary form of result. The ability to find the best fitting

function to describe the nonlinear relationship between the presence or absence of landslides and a set of conditioning factors combined with practically zero hyperparameters to tune in makes LR so compelling to be a baseline model in susceptibility analysis mapping. Basically, logistic regression relates the probability of landslide occurrence to a link function (in this case "logit") assumed to contain the conditioning factors on which landslide occurrence may depend, where the relationship between the occurrence and its dependency on conditioning factors can be expressed by Equation (3.7):

$$\hat{P} = \frac{1}{1 + e^{-z}} = \frac{e^{z}}{1 + e^{z}} \tag{3.5}$$

where $\hat{P}$ is the probability of a landslide occurrence and has a range of [0, 1] on an S-shaped curve; z is a linear fitting equation that involves the supplied set of landslide-related variables in the form of the following Equation (3.8):

$$Z = b_0 + b_1 X_1 + b_2 X_2 + \cdots + b_n X_n \tag{3.6}$$

where b_0 is the intercept of the model; b_n is the partial regression coefficients; and X_n is the conditioning variable.

Artificial Neural Network

An artificial neural network or shortly neural network (NNET) is a black-box model defined as a "computational mechanism able to acquire, represent, and compute a mapping from one multivariate space of information to another, given a set of data representing that mapping" [63].

Most NNET models are composed of simple and highly interrelated processing units (neurons) that are in permanent connection with each other. Generally, neurons are located in different layers, and NNET are characterized on the basis of the number of layers and the training procedures (depending on the characteristics and performance the training procedure used to carry out the learning process in a neural network, can vary widely). Connections between processing units are physically represented by weights, and each neuron has a rule for summing the input weights and a rule for calculating an output value. More than one layer of neurons can be

included in the perceptron in order to cope with non-linearly separable problems, and a multilayer perceptron (MLP) can be obtained (Figure 3.1).

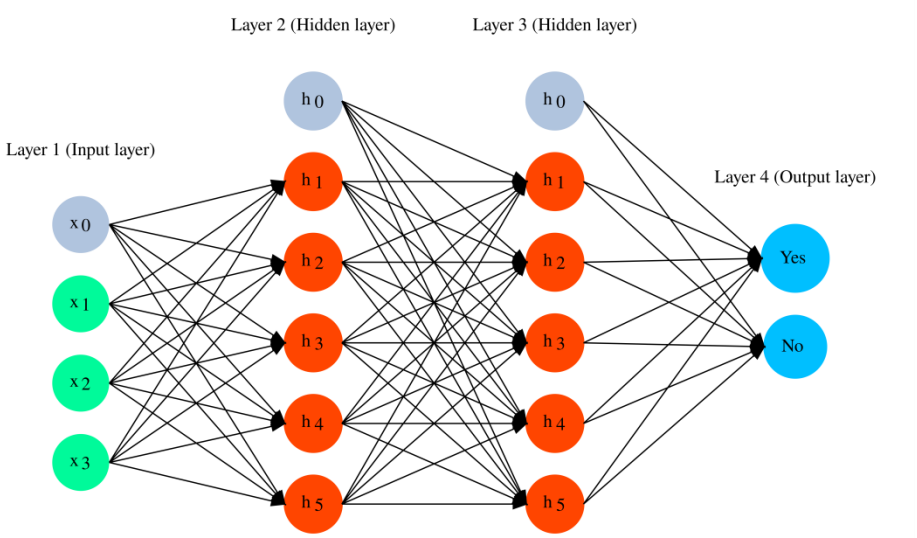


Figure 3.1 General architecture of NNET.

The learning problem in neural networks is formulated in terms of the minimization of a loss function f. This function is in general, composed of an error and regularization terms. The error term evaluates how a neural network fits the data set. On the other hand, the regularization term is used to prevent overfitting, by controlling the effective complexity of the neural network. The loss function f is, in general, a non-linear function of adaptive parameters such as biases and synaptic weights, (which can be conveniently grouped together into a single n-dimensional weight vector w).

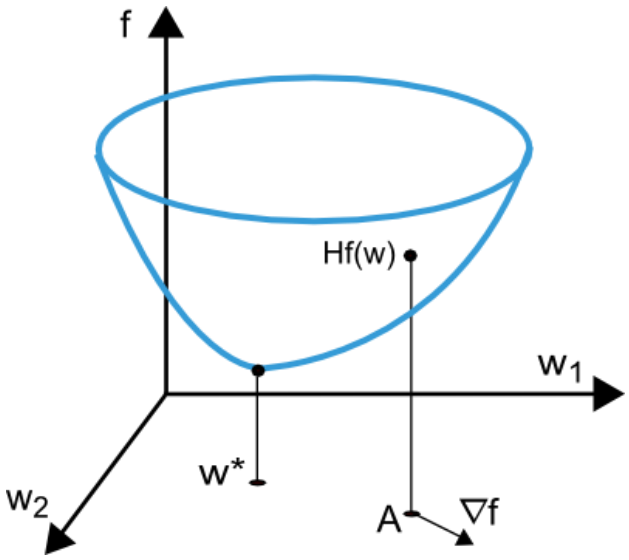


Figure 3.2 Representation of the loss function $f(w)$.

As we can see in Figure 3.2, the point w^* is minima of the loss function. At any point A, we can calculate the first and second derivatives of the loss function. The first derivatives are grouped in the gradient vector, whose elements can be written as:

$$\nabla_i f(w) = \frac{\partial f}{\partial w_i} \quad (i = 1, \dots, n) \tag{3.7}$$

Similarly, the second derivatives of the loss function can be grouped in the Hessian matrix:

$$H_i f(w) = \frac{\partial^2 f}{\partial w_i \cdot \partial w_j} \quad (i, j = 1, \dots, n) \tag{3.8}$$

The problem of minimizing the continuous and differentiable functions of many variables has been widely studied. Many of the conventional approaches to this problem are directly applicable to that of training neural networks, which means, the learning problem for neural networks is simply searching for the parameter vector w^* at which the loss function f takes a minimum value. The necessary condition mandate, that if the loss function of the neural network is at a minimum, then the gradient is the "zero vector". As a consequence, it is not possible to find closed training algorithms for the minima. Instead, we consider a search through the parameter space consisting of a succession of steps. At each step, the loss will decrease by adjusting the neural network parameters. This way of training NNET we start with some parameter vector (often chosen at random). Then, we generate a sequence of parameters, so that the loss function is reduced at each iteration of the algorithm. The change of loss between two steps is called the loss decrement. The training algorithm stops when a specified condition, or stopping criterion, is satisfied.

In this study, we are considering the "hill-climbing" procedure that belongs to a class of algorithms that are based on Newton's method but does not require the Hessian matrix of second derivatives of the objective function to be computed. Instead, it is updated by using gradient vectors; these are called "quasi-Newton" (or secant) methods. Newton's method is a second-order algorithm because it makes use of the Hessian matrix. The objective of this method is to find better training directions by using the second derivatives of the loss function. However, the Application of Newton's method is computationally expensive, since it requires many operations to evaluate the Hessian matrix and compute its inverse. Alternative

approaches, known as “quasi-Newton” or variable matrix methods are developed to solve that drawback.

These methods, instead of calculating the Hessian directly and then evaluating its inverse, build up an approximation to the inverse Hessian at each iteration of the algorithm. This approximation is computed using only information on the first derivatives of the loss function. The Hessian matrix is composed of the second partial derivatives of the loss function. The main idea behind the quasi-Newton method is to approximate the inverse Hessian by another matrix G, using only the first partial derivatives of the loss function. Then, the quasi-Newton formula can be expressed as:

$$w_{i+1} = w_i - (G_i - g_i) \cdot \eta_i \quad (i = 0,1 \dots, n) \tag{3.9}$$

The training rate η can either be set to a fixed value or found by line minimization. The inverse Hessian approximation G has different flavors. Two of the most used are the Davidon–Fletcher–Powell formula (DFP) and the Broyden–Fletcher–Goldfarb–Shanno formula (BFGS). Yet, BFGS is regarded as one of the best procedures for solving nonlinear optimization problems (in the absence of constraints) and weight adjustment [64], because using a general algorithms from unconstrained optimization seems the most fruitful approach [65], which lead to a faster convergence and provide a better results with less complication and parameters to tune in.

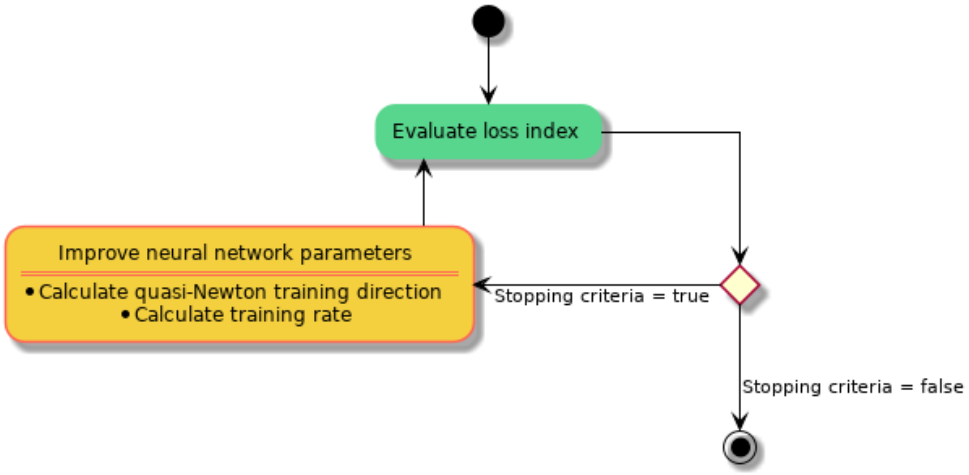


Figure 3.3 The activity diagram of the quasi-Newton BFGS NNET training process.

Improvement of the parameters is performed by first obtaining the quasi-Newton training direction and then finding a satisfactory training rate.

Compared to conjugate gradient and the popular gradient descent coupled with vanilla backpropagation or one of its variants used in most landslide susceptibility studies, the quasi-newton BFGS NNET had proven to be way faster to converge and provide a better results with fewer complications and parameters to tune in and the exact Hessian does not need to be computed and inverted like Newton methods.

Support Vector Machine

Support vector machine (SVM) is one of the new mathematical tools, which is used as a universal constructive learning procedure based on the statistical learning theory rather than loose analogies with natural learning systems developed by Corinna Cortes and V.Vapnik [66]. SVM provides non-linear solutions to regression and classification problems by transforming the input variables in a large-dimension space, whose inner product is given by positive definite kernel functions, then trained using dual optimization techniques with constraints [67]. Recently several researches have shown very competitive results if not excellent performance of SVM on different problems of regression and classification; with just a minimum amount tuning required [e.g. 40, 66, 67, 68, 69-76]

More details of SVM mathematical classification description can be found in [77] [78-83]. However, we are outlining the basics following Scholkopf, Mika [81] notations

First, Let's consider a set of training points $x_i(i = 1,2, \dots, l)$ where $x_i \in \mathbb{R}^n$ be the input vectors in input space, with corresponding binary labels which $y_i \in \{-1,1\}^l$ (i.e. y_i takes 1 if x_i is in class 1 and takes -1 if x_i is in class 2). Typically, SVM is designed for two-class problems where both positive and negative objects exist. For two-classes classification problems SVM seek to find a hyperplane in the feature space that maximally separates the two target classes [73]. The goal of the two-class SVMs is to find an optimal separating hyperplane with the maximal margin between the training points for class -1 and class $+1$. Which means, a discriminant function can be easily defined as :

$$g(x) = (w \cdot x) + b \tag{3.10}$$

Where: $w = (w1, \dots, w_n)$ is a vector of n elements, n is the feature space dimension; b is a scalar; $(w \cdot x)$ is the inner dot $(\cdot)$ product of w and x. Therefore, the classification rule is:

$$f(x) = sign\big((w \cdot x) + b\big) \Leftrightarrow \begin{cases} f(x) > 0 & \Rightarrow x \in class\ y_i = +1 \\ f(x) < 0 & \Rightarrow x \in class\ y_i = -1 \end{cases} \tag{3.11}$$

Second, let $\phi(x_i)$ be the corresponding vectors in feature space, where $\phi(x_i)$ is the implicit kernel mapping or precisely the feature function that map training vectors x_i into a higher (maybe infinite) dimensional space, and let $K(x_i x_j) = \phi(x_i)^T \cdot \phi(x_j)$ be the kernel function, implying a dot $(\cdot)$ product in the feature space.

The optimization problem for an SVM is:

$$\min_{w,b}\{\frac{1}{2} \| w \|^2 + C \sum_{i=1}^{l} \xi_i\} \tag{3.12}$$

Subject to constraints:

$$y_i(w \cdot x) + b \geq 1 - \xi_i \quad and \quad \xi_i \geq 0 \tag{3.13}$$

Where: w is the separating hyperplane normal vector in feature space and $C > 0$ is a regularization parameter (penalty parameter) controlling the penalty for misclassification. Formally, the Equation (3.14) is referred to as "the primal equation" that can be solved by the Lagrangian, so a dual problem can be derived into Equation (3.16), which is a quadratic optimization problem that can be efficiently solved using algorithms such as "Sequential Minimal Optimization" (SMO) [80].

$$w(\alpha) = \max_{\alpha}\{\sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i\, \alpha_j y_i y_j K(x_i x_j)\} \tag{3.14}$$

Subject to constraints:

$$0 \leq \alpha_i \leq C \tag{3.15}$$

During the SVM optimization process, many α_i converge to zero and the remaining x_i instance whose α_i satisfying $\alpha_i > 0$ are called support vectors N_s . However, to simplify SVM general procedure, let's assume that all non-support

vectors have been eliminated, so that $N_s = l$ is now the number of support vectors and $alpha_i > 0$ is for all i. With this formulation, the normal vector of the separating plane w can be expressed as:

$$w = \sum_{i=1}^{l} \alpha_i \, y_i x_i \tag{3.16}$$

Since $\phi(x_i)$ is defined implicitly, w can only exist in feature space and cannot be computed directly. As a result, the classification rule $f(x)$ of a new query vector x can only be determined by computing the kernel function of x with every support vector following a decision rule expressed as:

$$f(x) = sign(\sum_{i=1}^{N_s} \alpha_i \, y_i K(x_i, x_j) + b) \tag{3.17}$$

Where: N_s is the number of support vectors, b is the bias term representing the offset of the hyperplane along its normal vector, determined during SVM training [84], and $K(x_i x_j)$ is the kernel function being one of the following four basic kernels:

$$\begin{array}{ll} Linear & K(x_i x_j) = x_i^T x_j \\ Polynomial & K(x_i x_j) = (\gamma x_i^T x_j + r)^d \\ Radial\ Basis\ Function\ (RBF) & K(x_i x_j) = exp(-\gamma \parallel x_i - x_j \parallel^2) \\ Sigmoid & K(x_i x_j) = tanh(\gamma x_i^T x_j + r) \end{array} \tag{3.18}$$

Where: γ is an appositive parameter controlling the radius r (i.e. kernel width), and d is the polynomial degree.

3.3 RESAMPLING STRATEGY METHODS

This section of methods is featuring Objective 3 & 4 (see Chapter 1.3).

A modeling approach of certain problem is an absolute approximation of the problem that may not capture the true underlying model behind the data, which means that models are subjected to bias and errors during different stages of the modeling process and since modeling is a data-driven approach, so a proper assessment a model predictive capabilities, obligate supplying an independent testing

dataset to ensure model correctness, but as in most modeling cases (especially in landslide studies as observations are scares and hard to obtain), a common approach is performing a resampling on the input dataset, which divide the input data into a training dataset for fitting models and testing dataset to validate the models depending on the used strategy [85].

Resampling methods can be efficiently used for:

- Tuning of Hyperparameters, as most models (modeling techniques) require some sort of fine-tuning to certain parameters called hyperparameters. Thus, optimizing those hyperparameters to an optimum will highly lead to achieving a model with the highest quality possible.
- Accuracy Assessment, it's commonly known that accuracy plays a detrimental effect on the modeling process and its optimization so the introduction of error and bias during the modeling process will render the obtained model (or results) irrelevant (or non-reliable to a certain degree).
- Model Selection, selecting the most appropriate model is highly related to how the assessment was performed, in which the most reliable is screened based on assessment results during the modeling process (by favoring certain candidate model over another).

3.3.1 Basics and Statistical Properties

Assuming, we have a fitness function f, a set of input points $D = \{d_1, \dots, d_n\}$ with a decision variable $d_i = (x_i, y_i)'$, where parameters $x_i \in \mathbb{R}^m$ and the associated function values $y_i = f(x_i) \in \mathbb{R}$. The goal is the approximation of f by finding a meta-model using the information either contained in or extracted from D. Therefore, fitting function rules $\hat{f}_D \ D$ can be achieved by Equation (3.19). However, the approximation degree of $\hat{f}_D$ to the real fitness function f is questionable. In fact, a proper definition of loss function $L(y, \hat{f}(x))$ must be introduced, but in general, aggregation of these loss values is reported using functions like “mean” or “median”. If we assume that the loss function is defined (which depend on the problem to be solved), the risk associated with fitness model function is expressed as in Equation (3.20):

$$\hat{y} = \hat{f}_D(x, \varepsilon) \tag{3.19}$$

$$R(f,p) = \int_{\mathbb{R}} \int_{\mathbb{R}^m} L\,(y, f(x))\, p(x,y)\, dx\, dy \tag{3.20}$$

$$GE(\hat{f}_D, p) = \int_{\mathbb{R}} \int_{\mathbb{R}^m} L\,(y, \hat{f}_D(x))\, p(x,y)\, dx\, dy \tag{3.21}$$

$$\widehat{GE}(\hat{f}_D, D^*) = \sum_{(x,y)' \in D^*} \frac{L(y, \hat{f}_D(x))}{\mid D^* \mid} \tag{3.22}$$

$$\widehat{GE}(\hat{f}_D, D^*) = \sum_{(x,y)' \in D^*} \frac{L(y, \hat{f}_D(x))}{\mid D^* \mid} \tag{3.23}$$

In Equation (3.20), $p(x,y)$ is the joint Probability Density Function (PDF) of decision space and function values, and by incorporating the estimator $\hat{f}_D$ of f into Equation (3.20) (because models are often based on data) to get what's called "generalization error" (GE) or "conditional risk associated with the predictor" (See Equation (3.21)), and since the GE directly depends on the data used to fit the models then the underlying distribution p of the input dataset would be difficult to know or even efficiently estimated so we can replace it by either a test subset D^* to get Equation (3.22), but if we incorporate the input dataset D itself we can get resubstitution error (Equation (3.23)).

It should be noted that using the input dataset to train the model and estimate the GE (as detailed in Equations (3.19) ~ (3.23)) is inconvenient, because of the biased estimation of the generalization error. In that case, a model selection will be unintentionally biased toward adaptive and complex models. To solve this issue, splitting the input dataset into training set D_{train} and of course a testing set D_{test} so that $D_{train} \cup D_{test} = D \quad D_{train} \cap D_{test} = \emptyset,$ will ensure fitting the model using D_{train} to obtain $\hat{f} D_{train}$ and at same time calculate the GE using the test subset D_{test}. This approach of training-testing subset is well known as "Hold-out" or "Train-Test Split".

$$\widehat{GE}_{Hold-out} = \widehat{GE}(\hat{f}_{D_{train}}, D_{test}) \tag{3.24}$$

$$\widehat{GE}_{resamp} = \frac{1}{k} \sum_{i=1}^{k} \widehat{GE}\,(\hat{f}_{D^i}, D/D^i) \tag{3.25}$$

From a statistical standpoint, this approach is trivial, because the test subset observations are independent of those in the training subset, so in this case, the GE can be estimated by Equation (3.24). However, splitting the input dataset comes with two important shortcomings:

- First, a large input dataset D must be supplied so we can have enough observations in both the training subset to build an adequate model and the test subset to fully obtain statistical valid performance results.
- Second, a large number of sample observations are difficult to obtain in most cases let alone landslide inventory samples.

Solving those issues can be possible by resampling the input of the dataset using resampling techniques (i.e. cross-validation, bootstrapping, and subsampling). These resampling techniques and strategies partition[41] the input dataset (D) to generate training sets D^i and testing sets (D/D^i, $i = 1, \dots, k$), in such way that a single model is trained for each training set, then predictions are made based on the corresponding testing set resmaple2 and the loss function value $s^i = s(D^i, D/D^i)$ of each model is calculated. Later, the k individual loss function values are aggregated into a performance indicator S by performance measures p^i (i.e. mean, median…etc.). As a result, the GE quality will mostly depend on:

- The training-testing sets size compared to the original input dataset D.
- The number of subsamples k drawn by the resample strategy.
- The dependency structure between the subsamples D^i.

All resampling strategies share the common general framework procedure detailed above and depicted in Figure xx, where the GE estimation for Hold-Out method in Equation (3.24) can be generalized to Equation (3.25). As explained before, the GE estimation is data-driven based on the input dataset, which implies that both training and testing sets sample size must reasonable, considering the totality of samples available to ensure the GE estimation cant neither be pessimistic nor optimistic.

[41] Maybe repeatedly depend on the method.

3.3.2 Cross-Validation

Cross-validation (CV) [86], is one of the well-established and commonly used resampling strategies. By translating Equations (3.19) to (3.23) into Algorithm 3.3, we can obtain an implementable common generic procedure layout shared between most resampling strategies, but what differs in most of the cases is the training-testing subsets (k subsets) generation process (line 1 of Algorithm 3.3). In fact, CV uses a simple procedure to generate the k subsets (Algorithm 3.4). Essentially, the input dataset is partitioned into k equal (or nearly equally) sized subsets called "folds" and then a $k-1$ folds are used to fit the model and the remaining fold are used to validate the model. This process is repeated k times for all possible $k-1$ combinations and this ensures that each $k-fold$ is used precisely once as validation subset.

Algorithm 3.3 Generic resampling procedure.

input : A dataset D of n observations d_1 to d_n, the number of subsets k to generate and a loss function L.
output: Summary of the validation statistics.

1 Generate k subsets of D named D^1 to D^k ;
2 $S \leftarrow \emptyset$;
3 **for** $i \leftarrow 1$ **to** k **do**
4 $\quad D^i \leftarrow D/D^i$;
5 $\quad \hat{f} \leftarrow \texttt{FitModel}(D^i)$;
6 $\quad s_i \leftarrow \sum_{(x,y)' \in D^i} L(y, \hat{f}(x))$;
7 $\quad S \leftarrow S \cup \{s_i\}$
8 **end**
9 Summarize S, for example: $mean(S)$, $median(S)$

Algorithm 3.4 Subsets procedure for k-fold CV.

input : A dataset D of n observations d_1 to d_n and the number of subsets k to generate.
output: k subsets of D named D^1 to D^k.

1 $D \leftarrow \texttt{Shuffle}(D)$;
2 **for** $i \leftarrow 1$ **to** k **do**
3 $\quad D^i \leftarrow D$
4 **end**
5 **for** $j \leftarrow 1$ **to** n **do**
6 $\quad i \leftarrow (j \bmod k) + 1$;
7 $\quad D^i \leftarrow D^i/\{D_j\}$
8 **end**

Generally, for problems where the input dataset may contain categorical variables or the target variable is also categorical (classification problems), Stratified

Cross-Validation (SCV) is used. The rationale behind SCV is the insurance a good data-representation during resampling by rearranging the input dataset data in a way that the sample's distribution of each fold matches (or as close as possible to) the distribution of the input dataset using a process called “Stratification”.

3.3.3 Overfitting

One of the most common and troublesome issues in all ML algorithms and models is the random error or noise that can be also known as “Overfitting”, which is directly related to overall GE of the classier model and the training dataset produced by resampling strategy [29]. This noise can be explained as, the problem of underperforming (decrease in overall performance) during the validation-testing stage while gaining high performance during the training stage (Figure 3.4).

Overfitting is highly pronounced when:

- The overall performance is decreasing while the model complexity is increased (Figure 3.4b).
- A large amount of data is fed and used to build the model.

This means the model is rather learning the noise than generalizing the problem at hand (i.e. the learning becomes too specialized and the algorithm does not generalize well enough). Two possible ways of dealing with overfitting are:

10. First, optimizing the generalization power of the algorithm.
11. Second, generating training and testing splits which will have balanced class distributions, i.e. the sizes of all classes will remain proportional in both splits (i.e. Stratification).

The former (case 1), is hard to achieve, due to various reasons that can vary depending on the implemented algorithms (i.e. limitation and drawbacks). The latter (case 2), is not always feasible in spatial modeling, due to the abundance of one class and scarcity of another or several other classes[42]. This is especially pronounced if the Test-Train splitting sampling strategy (i.e. Hold-Out split) is adopted. Therefore, selecting a resampling strategy requires a technique that takes overfitting into consideration and minimizes (even partially) the overfitting effect by involving a

[42] That’s usually the case for landslide susceptibility studies.

specific resampling strategy that balances the trade-off model complexity for its fitness, i.e. the model's variance against its bias (Figure 3.4b).

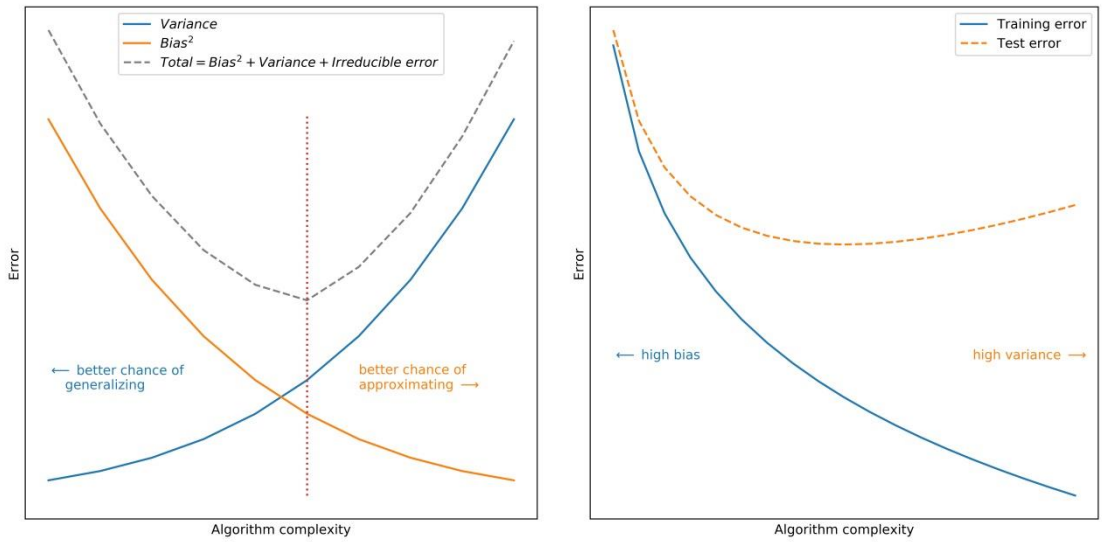


Figure 3.4 The overfitting problem of ML models.

The functions are showing an evident rise of the erroneous returns in testing mode despite the rise in data feed (amount of training and testing data or complexity of the model).

3.4 MODEL OPTIMIZATION AND TUNING METHODS

This section of methods is featuring Objective 2 (see Chapter 1.3).

Models and algorithms tend to vary dramatically depending on the approach. Yet, exploring model full potential requires correctly tuning a variety of incidental parameter choices and settings [87]. In rare cases, the hand-tuning of the optimal hyperparameters is enough to rely on trial-and-error methods such (i.e. Grid search, Random search, Gradient-Based Optimization). However these methods and techniques were considerably simplistic and easy to implement, and yet they produce very poor results that lead to:

- Costly evaluations. Especially, if the computational budget is limited.
- Wrong assessments about the implemented models whether they are genuinely bad or simply badly tuned [88].

To avoid such common problems, a state of art technique called Sequential Model-Based Optimization (SMBO), also known as Bayesian optimization that can efficiently optimize and work on a strictly reduced budget for function evaluations

and hyperparameter optimization of expensive black-box models and obtain better results in fewer experiments than traditional techniques, was considered. The excellence in performance in SMBO is related to:

- The ability to reason about the quality of experiments before they are run [89-92].
- Benefiting from the "adaptive capping" to avoid long runs [93].

First to explain how SMBO work, let's assume that an expensive black-box function $f: x \in \mathbb{R}$ with d-dimensional input space $\mathcal{X} = \mathcal{X}_i; \dots; \mathcal{X}_d$ and a deterministic output $y = f(x)$. Each $\mathcal{X}_i$ is a parameter with constraints box that can be bounded for numeric values (i.e. $\mathcal{X}_i = [l_i;\ u_i] \in \mathbb{R}$ where: l_i and u_i are the lower and upper bounds), or finite set s of categorical values (i.e. $\mathcal{X}_i = \{v_i, \dots, v_s\}$). These constraint boxes, formulate the $\mathcal{X}_i$ parameter space. The aim is to minimize the target value y (i.e. $f(x)$) by fulfilling Equation (3.26):

$$x^* = \underset{x \subset \mathcal{X}}{argmin} f(x) \tag{3.26}$$

SMBO optimization, basically approximate the expensive black-box function $f(x)$ and iteratively update and refine using meta-models called "surrogate models" in-lieu-of expensive optimizing stimulator of the expensive function f by regression models, which are computationally cheap to evaluate. Since the surrogate models are regression models, they are capable of direct estimation $\hat{f}(x)$ of the true value $f(x)$ and an estimation of the prediction standard error (i.e. $\hat{s}(x)$ or $\sigma(x)$ or $\sigma^2(x)$, which called spread posterior distribution[43] of $\hat{f}(x)$. Obviously, since $f(x)$ is expensive to evaluation (i.e. according to the assumption above), which means the evaluations of $f(x)$, if (for sure) have a budget constraint that can be:

- Termination criteria (i.e. time elapsed, an optimization threshold was reached).
- The total number of evaluations is exhausted.

[43] In statistic, $\sigma(x)$ and $\sigma^2(x)$ also called local uncertainty estimators and used to measure "The trustworthiness" of the prediction.

The general approach of SMBO illustrated in details in Algorithm 3.5, and can be summarized into six main steps procedures:

1. Vector of an indexed set (i.e. called initial design) $\mathcal{D} = x_1, x_2, \dots, x_n$ of n_{initial} points x_i sampled from $\mathcal{X}$ (i.e. $\mathcal{X} \in \mathbb{R}$) and $y_i = f(x_i)$ is the associated target value of f. At each point of the indexed set, f is evaluated and yield the outcome of the target value y in a vector of (y_i, x_i). This vector will be the input for the initial surrogate model $\hat{f}$ (i.e. in the next step (2)).
2. Fit the surrogate model at each evaluation point x_i and its target value y_i using the previously indexed vector $\mathcal{D}$.
3. An acquisition function α (i.e. called infill criterion) suggest a set of point $x_{\{i+j\}} \in \mathcal{X}$ $(j = 1, \dots, m)$ defined on $\mathcal{X}$. The acquisition function α, operate on the surrogate model $\hat{f}$ to propose and determine points that are "promising" for optimization based on $\hat{f}$. The proposed points can have:
    a) Good "exploitation" values (i.e. improving the expected objective function f value);
    b) Good "exploration" values (i.e. high potential to improve the quality of the surrogate model);
    c) A balanced combination of the above (i.e. (a) and (b)).
4. The proposed points by α are evaluated using f and the new vector $(y_{\{i+j\}}, x_{\{i+j\}})$ added to the design $\mathcal{D}$.
5. Check the budget whether if the budget was exhausted. If "yes", move to step 6, otherwise return to step 2.
6. Return the proposed solution of the optimization problem in form of vector contains the best points.

Algorithm 3.5 General procedure of SMBO optimization approach.

1 Generate an initial design $\mathcal{D} \subset \mathcal{X}$;
2 Compute $\boldsymbol{y} = f(\mathcal{D})$;
3 **while** *total evaluation budget is not exceeded* **do**
4 | Fit surrogate on $\mathcal{D}$ and obtain $\hat{f}, \hat{s}$;
5 | Get new design point $\boldsymbol{x}^*$ by optimizing the infill criterion based on $\hat{f}, \hat{s}$;
6 | Evaluate new point $y^* = f(\boldsymbol{x}^*)$;
7 | Update: $\mathcal{D} \leftarrow (\mathcal{D}, \boldsymbol{x}^*)$ and $\boldsymbol{y} \leftarrow (\boldsymbol{y}, \boldsymbol{y}^*)$;
8 **end**
9 **return** $y_{min} = \min(\boldsymbol{y})$ and the associated $\boldsymbol{x}_{min}$

3.4.1 Initial Design

The initial design is an indexed set of points carefully sampled from the input search space $\mathcal{X}$ for the purpose to evaluate against the expensive function f and generate the initial surrogate model $\hat{f}$. The total number of points in the initial design $\mathcal{D}$ needs to be optimum. In other words, if the number of samples in $\mathcal{D}$ doesn't cover $\mathcal{X}$, the fitting the surrogate model would be poor in best-case scenarios which subsequently lead to suboptimal point proposition by $\hat{f}$that would negatively influence the progress of the optimization. In some cases, building the surrogate model $\hat{f}$ is impossible, if $\hat{f}$ is too low. On the other hand, a large initial design helpful to get insight into the search space landscape, but the overall budget must be reduced to cope with the large design $\mathcal{D}$, which is in some cases a bad practice. There exist multiple methods to generate an initial design sample from the search space like manual sampling, random sampling, coarse grid designs, or by the method used in this case study, the space-filling fashion of Latin Hypercube designs (lhs).

3.4.2 Surrogate Model

Surrogate models are meta-models are actually the sampling algorithm to propose new points x_i that lead to "optima". Technically, surrogate models are replacement for more expensive stimulators of the expensive black-box function f. In fact, optimizing the surrogate model is in lieu of f simulator much cheaper (i.e. computational and time budgets). Selecting the appropriate surrogate model $\hat{f}$ depend not only on the computational budget available but also the current structure of input space $\mathcal{X}$. There exist three cases for $\mathcal{X}$:

1. $\mathcal{X}$ is purely numeric ($\mathcal{X} \in \mathbb{R}$);

2. $\mathcal{X}$ is purely categorical, hierarchical or Boolean space;

3. $\mathcal{X}$ is a mixed search space (a combination of cases 1 and 2).

For case (1), kriging (i.e. Gaussian process) is recommended and should be used because it provides state-of-the-art performance; for case (2) and (3) dummying the input space $\mathcal{X}$ would be great solution and therefore use "Kriging", but with recent years, Random forest (RF) prove to be a viable alternative (option) due to its ability to digest most types of the input spaces directly without any pre-processing.

3.4.3 Infill Criteria

The infill criteria α is actually well known as the acquisition function, which used to guide the SMBO optimization process by evaluating the goodness of fit of each "candidate" point, then if the "candidate" point if good enough, this point will be evaluated against the objective function $\hat{f}$. Generally, the acquisition function α is contrasted in pair-wise fashion by combining the posterior mean ($\hat{\mu}$) and the posterior spread ($\hat{\sigma}$) [44] in single "well balanced" numeric formula. Both $\hat{\sigma}$ and $\hat{\mu}$ are directly estimated by the surrogate model $\hat{f}$.

To better understand the purpose behind combining $\hat{\sigma}$ and $\hat{\mu}$, we need to understand the functionality of each term. "Local uncertainty" estimators (i.e. $\hat{\sigma}$ and $\widehat{\sigma^2}$) are used as an "exploration" indicator. In fact, higher values for (i.e. $\hat{\sigma}$ or $\widehat{\sigma^2}$) highly indicate less explored regions in $\mathcal{X}$ the search space landscape which means less certainty about the true landscape of the search space due to either lack or presence of few points nearby or close to those regions. On the other hand, $\hat{u}$ is used as an "exploitation" indicator, lower values are more promising as they indicate a low true function value of f. So α is a balanced trade-off between two conflicting criterion's (i.e. "exploitation" and "exploration") by considering promising regions (i.e. points) with a low posterior mean (i.e. $\underset{x \subset \mathcal{X}}{\operatorname{argmin}} \hat{\mu}(x)$) and high posterior spread (i.e. $\underset{x \subset \mathcal{X}}{\operatorname{argmax}} \hat{\sigma}(x)$).

Acquisition functions α are called "utility" function and can be interpreted in Bayesian Decision Theory as "evaluating an expected loss associated with f at

[44] Also known as the posterior standard deviation and sometimes, posterior variance $\widehat{\sigma^2}$ is used instead, but both are usually considered as "local uncertainty estimator".

point$x$. The best point with the lowest expected loss is returned and selected as an optimization solution for f". Essentially the role of the acquisition function α is to guide the optimization process for global convergence (i.e. finding the optimum). Typically, α are defined with maximization perspective (i.e. positive pure maximization (argmax); or negative pure maximization (argmin) where these promising regions, correspond to "potentially" better objective function values. The maximized α is used to propose (select) the next point at which to evaluate against f.

Probability of Improvement (PI):

Probability of improvement (PI) is considered as the first utility function α designed for the Bayesian Optimization Framework Thanks to the early work of (Kushner). Assuming that $y_{\min} = \text{argmin} f(x)$ is the lowest (minimal) value of f recorded, so far, which means α_{PI} (PI Acquisition function) will likely evaluate points x_i that most likely improve upon this value (i.e. $y_{\min}$). This corresponds, to a utility function $\vartheta(x)$ (i.e. utility function is simply negative loss function) associated with evaluating f at given point x and x_i:

$$\vartheta(x) = \begin{cases} 0, & f(x) > y_{min} \\ 1, & f(x) \leq y_{min} \end{cases} \tag{3.27}$$

By Equation (3.27), it's clear that the reward unit reception in case of $f(x)$ turn out to be for $\leq y_{\min}$and no reward otherwise. Therefore, the probability of improvement acquisition functions the expected utility function of x.

$$\alpha_{PI} = \begin{cases} E[\, I(x) \mid x, \mathcal{D}\,] \\ \Phi(\gamma(x)) \end{cases} \tag{3.28}$$

Where $\gamma(x) = \frac{y_{\min} - \hat{\mu}(x)}{\sigma(x)}$ and $\Phi(\cdot)$ is denoted as the Cumulative Distribution Function of the Standard Normal Distribution (CDF).

The point x^* with the best-expected utility function (i.e. highest probability of improvement) is chosen as "global optima". However, this can lead to bad results as α_{PI} is highly biased toward exploitation (pure minimizations of the posterior mean $\hat{\mu}(x)$). On top of that, the odd system behavior of the utility function is indicative odd behaviors and limitation like stuck in "local optima" and under-explored landscape region globally due to the odd of the reward system of

improvement upon the current minimum (i.e. $y_{\min}$) independent of the size of improvement.

Expected Improvement (EI):

An alternative solution to α_{PI} would be an acquisition function that takes into account not only the probability of improvements but also the overall magnitude (i.e. size) of the improvement that a point can probably yield. B. Mockus and Mockus [94], Frean and Boyle [95] proposed a maximization by taking into account the maximal value of y observed so far (i.e. $y_{\min}$) where α_{EI} (i.e. Expected Improvement acquisition function) evaluate f at each point x_i in expectation to improve upon $y_{\min}$ the most. This corresponds to the following utility function:

$$\vartheta(x) = max(y_{min} - f(x), \qquad 0) \tag{3.29}$$

Where: the reward reception is strictly equal to the "improvement" $y_{\min} - f(x)$ only if $f(x) < y_{\min}$, otherwise no reward reception. Therefore, the expected improvement acquisition function is then the expected utility as a function of x:

$$\alpha_{EI} = E\big(I(x)\big)\begin{cases} f(x) > 0, & \sigma(x)(\gamma(x)\Phi(\gamma(x))) + \sigma(x)\phi(\gamma(x)) \\ f(x) = 0, & 0 \end{cases} \tag{3.30}$$

Where: $\gamma(x) = \frac{y_{\min} - \hat{\mu}(x)}{\sigma(x)}$ and $\Phi(\cdot)$ and $\phi(\cdot)$ are denoted by B. Mockus and Mockus [94] as the Cumulative Distribution Function (CDF) and Probability Density Function (PDF) of the Standard Normal Distribution, respectively. The point x_i with maximal α_{EI} will be selected.

The two conflicting terms of CDF and PDF are balanced and can be interpreted explicitly encoding trade-off between exploration and exploration. The former can be maximized by minimizing the posterior mean $\hat{\mu}(x)$. The second latter can be maximized by maximizing the posterior spread $\sigma(x)$.

Upper & Lower Confidence Bounds (UCB/LCB):

The early work of Cox and John [96] introduced the Sequential Design Optimization (SDO) algorithm that proposes and selected points based on the confidence bounds of the posterior mean $\hat{\mu}(x)$ and weighted posterior spread $\sigma(x)$.

The upper confidence bound (UCB) is described in terms of maximization of f rather than minimizing like the lower confidence bound (LCB) (Equation (3.31)). However, in most literature, LCB is used instead of UCB as the term is ingrained in literature as a standard term.

$$\lambda > 0 \begin{cases} \alpha_{LCB} = \hat{\mu}(x) + \lambda\sigma(x) \\ \alpha_{UCB} = \hat{u}(x) - \lambda\sigma(x) \end{cases} \tag{3.31}$$

Where: λ is constant the denoted as parameter that control "Exploration vs. Exploration" trade-off, and should be $\lambda > 0$. If $\lambda = 0$, α_{LCB} and α_{UCB} coincides with the predicted mean value. The larger λ is chosen, the more attractive unexplored regions of the search space become.

Surprisingly, the confidence-bound acquisition functions cannot be interpreted similarly to computing an actual expected utility function like α_{PI} and α_{EI}. Nonetheless, the confidence bound acquisition functions are known for the strong theoretical results that under certain conditions the iterative application of this acquisition function will converge to true global optima of f.

3.4.4 Infill Optimization

Optimizing the acquisition function infill criterion α is technically the process of points x_i that yield the best value according to α. In reality, optimizing α is inexpensive especially with cheap surrogate models $\hat{f}$ compared to optimizing the original function f directly so the evaluations can be spent lavishly. Some branches and bounds are proposed for this task by various researches e.g. jones. However, a very generic approach called "Focus Search" proposed by Bischl, Richter [88] outlined in Algorithm 3.6, has proven to be efficient and effective since it is able to handle and digest all the available sorts of the search space $\mathcal{X}$ (i.e. numeric, categorical or mix search space).

Algorithm 3.6 Focus Search infill optimization procedure.

Require: Infill criterion $c : \mathcal{X} \rightarrow \mathbb{R}$, control parameters $n_{restart},\ n_{iters},\ n_{points}$

for $u \in \{1, \ldots, n_{restart}\}$ **do**
  Set $\tilde{\mathcal{X}} = \mathcal{X}$*;*
  for $v \in \{1, \ldots, n_{iters}\}$ **do**
    Generate random design $\mathcal{D} \subset \tilde{\mathcal{X}}$ *of size* n_{points}*;*
    Compute $\boldsymbol{x}^*_{u,v} = (x^*_1, \ldots, x^*_d) = \arg\min_{\mathbf{x} \in \mathcal{D}} c(\boldsymbol{x})$*;*
    Shrink $\tilde{\mathcal{X}}$by focusing on $\boldsymbol{x}^*$*: ;*
    foreach search space dimension $\tilde{X}_i \in \tilde{X}$ **do**
      if $\tilde{\mathcal{X}}_i$numeric: $\tilde{\mathcal{X}}_i = [l_i, u_i]$ **then**
        $l_i = \max\{l_i, x^*_i - \frac{1}{4}(u_i - l_i)\}$*;*
        $u_i = \min\{u_i, x^*_i + \frac{1}{4}(u_i - l_i)\}$*;*
      end
      if $\tilde{\mathcal{X}}_i$categorical: $\tilde{\mathcal{X}}_i = \{v_{il}, \ldots, v_{is}\},\ s > 2$ **then**
        $\bar{x}_i$ = sample one category uniformly from $\tilde{\mathcal{X}}_i / x^*_i$*;*
        $\tilde{\mathcal{X}}_i = \tilde{\mathcal{X}}_i / \bar{x}_i$*;*
      end
    end
  end
end
return $\mathbf{x}^* = \arg\min_{\{u \in 1, \ldots, n_{restart}, v \in 1, \ldots, n_{iters}\}} c(\boldsymbol{x}^*_{u,v})$

The algorithm first start by generating moderate to large random design than the surrogate model $\hat{f}$ is used to evaluate all design points to determine the most "promising" points. Next, with help of these "promising" points, the focus search is able to "focus" and "shrink" the search space $\mathcal{X}$ around these points to randomly sample new points out of the focused search space. The shrinkage procedure (i.e. focus procedure) is iterated n_{iters} times and the whole procedure is restarted $n_{restart}$ to avoid "local optima". Finally, the best point x^* from all iterations and restarts is returned.

3.4.5 Termination

The termination criterion for SMBO can vary dramatically depending on the user, used approach and the computational budget available. Usually, the total number of evaluations of the objective function f or the total number of the iteration is used. However, other criterions that depend on the available time budget can be used to point where the termination of the optimization is ended only when the available time budget is exhausted. Depending on the case, a predefined objective value can be set as a termination criterion and even a combination of the criterions can be implemented at once.

3.4.6 Final Best Points Returning

At the end of the optimization where the budget is exhausted and/or met the final solution x^* is returned either the best-observed point during optimization, or a final surrogate model $\hat{f}$ is performed on all the evaluated points (i.e. points in $\mathcal{D}$ and the proposed points by α) is $\hat{f}$ is known to be noisy.

3.5 MODEL PERFORMANCE EVALUATION METHODS

This section of methods is featuring Objective 5 (see Chapter 1.3).

Evaluating model's performance is a delicate subject, as it depends on the learning problem, task at hand and most importantly the underlined goals of the analysis. Luckily, most performance metrics (at least the one implemented in classification problems) rely on correctly and incorrectly classified landslide instances and the relation between them. Correctly and incorrectly classified landslide instances, are usually depicted using what we call "confusion matrix"[45] or "contingency table" (Table 3.1).

Confusion matrix is a specific table layout[46] that allows visualization of the performance of an algorithm by introducing different classification hits and errors metrics or measures such as True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN) that are based the positive class (the landslide presence i.e. landslide class of "Yes" or 1) and the negative class (landslide absence i.e. Non-landslide class of "No" or 0). The sum of pixels instances that have been "correctly" classified in the positive class are known as the True Positives (TP) and the sum of pixels instances that have been "falsely" classified in the positive class is known as False Positives (FP). On the other hand, the sum of pixels instances that have been "correctly" classified in the negative class and the sum of pixels instances that have been "falsely" classified in the positive class are known as True Negatives (TN) and False Negatives (FN), respectively.

[45] The name convention, stems from the fact that it makes it easy to see if the system is confusing two classes (i.e. commonly mislabeling one as another).

[46] The confusion matrix layout consist of two dimensions ("actual" and "predicted"), and identical sets of "classes" in both dimensions (each combination of dimension and class is a variable in the contingency table). Each row of the matrix represents the instances in a predicted class while each column represents the instances in an actual class (or vice versa).

Table 3.1 Confusion matrix and appropriate error measures.

		Landslide Inventory	
		True	False
model	Positive	TP (True Positive)	FP (False Positive)
	Negative	FN (False Negative)	TN (True Negative)

Herein, the following subsections will describe the performance metrics that will help to highlight the spatial predictive capabilities and express models full potential in landslide susceptibility assessment:

3.5.1 Accuracy

The Overall Accuracy (ACC) is used to assess models accuracy and can be denoted as the fraction (ratio) or the counts of correctly classified events of both landslide and non-landslide instances[47], following Equation (3.32):

$$Accuracy(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} 1\,(\hat{y}_i = y_i) \tag{3.32}$$

Where: $\hat{y}_i$ is the predicted value of the i_{th} sample, y_i is the corresponding true value, and n_{samples} is the total number of samples. Because ACC, focus on True values reported in the confusion matrix (i.e. TP and TN), Equation (3.32) can be simplified into Equation (3.33):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{TP + TN}{n_{samples}} \tag{3.33}$$

That being said, ACC usually expressed in float (decimal) format ranging from 1 (correctly classify all events) to 0 (fail to classify any events).

3.5.2 The area under the ROC Curves

The Area under the ROC Curves (AUC) is the probability of a classifier to correctly anticipate the occurrence or non-occurrence of predefined events [97-100].

[47] Model predictions instances of both landslide and non-landslide classes claims to be different than the existing in the reference (i.e. the inventory).

This is proven to be convenient because maximizing AUC is basically equivalent to maximizing the ACC of the classifier. AUC can be computed mathematically by the trapezoidal rule of integral calculus as shown in Equation (3.34):

$$AUC = \sum_{k=1}^{n} (X_{k+1} - X_k)(S_{k+1} - S_k - S_k/2) \tag{3.34}$$

Where, X_k indicates 1-specificity and S_k is the sensitivity.

This value of AUC varies from 0.5 (very poor performance) and 1.0 (perfect performance). Kantardzic [101] and Tien Bui, Tuan [102] a value of $0.8 \leq AUC \leq 0.9$, indicate a very good discrimination ability of the model, and values of $AUC > 0.9$ suggest an excellent classification models. On the other hand, $0.7 \leq AUC \leq 0.8$ indicates a good predictive model, and $0.6 \leq AUC \leq 0.7$ suggest an average classification model. A value of $AUC \leq 0.5$ signifying that the ability of the performance used the models has no power to distinguish.

It is important to highlight that the ROC curves are used explicitly to calculate the area under the ROC curve (AUC), but can be used to diagnose both the sensitivity-specificity trade-off and the classifier ability when cut-off threshold varies.

3.5.3 Cohen Kappa Index

Cohen Kappa Index (κ-index), represents the measure of agreement between compared entities, rather than the measure of classification performance [103]. κ-index is very convenient for comparison of maps with the same classes[48] [32], as it is able to measure landslide models reliability by calculating the proportion of observed agreement beyond that expected by chance, using Equation (3.35):

$$\kappa = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e} \tag{3.35}$$

Where: $p_o$is the empirical probability of agreement on the label assigned to any sample (the observed agreement ratio), and p_e is the expected agreement when both

[48] As it commonly the case in ML-based classification experiments.

annotators assign labels randomly. p_e is estimated using a per-annotator empirical prior to the class labels. p_o and $p_e$the are easily calculated from confusion matrix according to Equations (3.36) and (3.37):

$$\begin{aligned} p_o &= \frac{TP + TN}{n_{samples}} \\ p_e &= p_{yes} + p_{no} \end{aligned} \tag{3.36}$$

$$\begin{aligned} p_{yes} &= \frac{TP + FP}{n_{samples}} + \frac{TP + FN}{n_{samples}} \\ p_{no} &= \frac{TN + FP}{n_{samples}} + \frac{TN + FN}{n_{samples}} \end{aligned} \tag{3.37}$$

The possible values of κ-index are ranging between -1 and 1. A value of κ-index $= 1$ indicates a complete agreement then, If there is no agreement among the raters other than what would be expected by chance (as given by p_e), κ-index $= 0$. It is possible for the κ-index to be negative, which implies that there is no effective agreement between the two raters or the agreement is worse than random. However, according to Landis and Koch [103], the strength of agreement given the κ-index magnitude is for 0.8–1.0 almost perfect, 0.6–0.8 substantial, 0.4–0.6 moderate, 0.2–0.4 fair, 0–0.2 slight, and ≤ 0 poor.

3.6 RESEARCH WORKFLOW

This section is featuring Objective 3 (see Chapter 1.3).

This section, focus on presenting the proposed methodology used to conduct this research. The research was performed using five ML models (i.e. Gradient Boosting Machine (GBM), Logistic Regression (LR), Artificial Neural Network (NNET), Random Forest (RF), and Support Vector Machine (SVM)). Model's hyperparameters were tuned and configured using Sequential Model-Based Optimization (SMBO). The analysis was programmed from scratch in an R environment[49] because:

- The high flexibility R offer.
- Reduction of the error and bias that can be introduced either by evaluating models in different software or platforms that may respond differently.

[49] The source code is available at Github (See Appendix B).

The upcoming subsections are dedicated to explain and describe the details and procedures used in this each step of the research rather informally, according to the overall concept of the proposed methodology of this research shown in Figure 3.5.

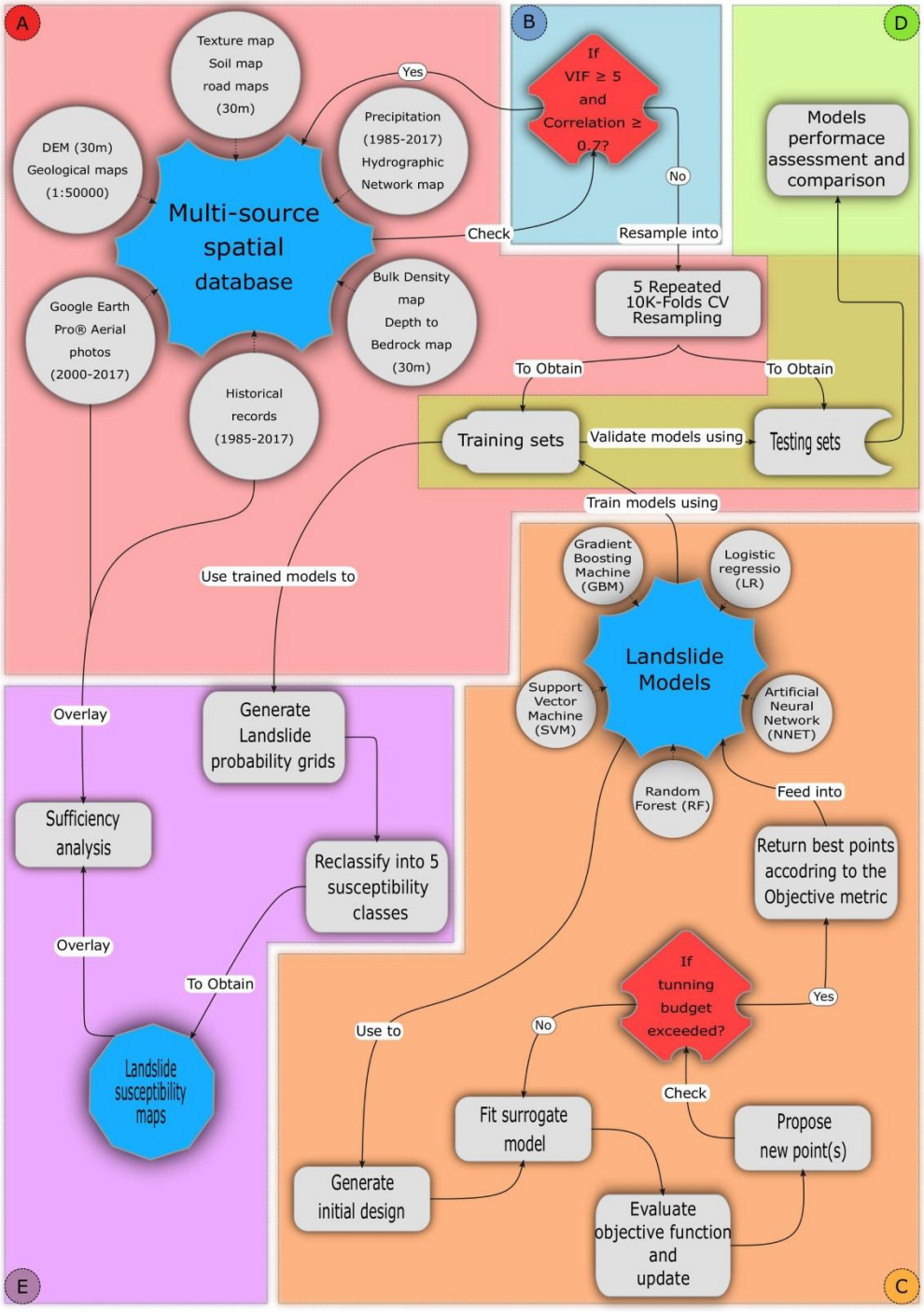


Figure 3.5 The overall concept of the proposed methodology for this research.

(A) Construct a spatial database that will serve as an input dataset for the study from the landslide inventory map and the landslide conditioning factors; (B) Analyzing and optimizing landslide conditioning factor based on Pearson Correlation Coefficients and Variance Inflation Factors analyses results; (C) Model configuration and implementation using the appropriate model hyperparameters optimization strategy;(D) Model training, validation, and comparison using the appropriate performance indicator metrics; (E) Landslide susceptibility maps generation, assessment, and evaluation based on the appropriate assessment-evaluation strategy.

3.7 TRAINING AND TESTING DATASETS PARTITIONING

First and foremost, a geospatial database (see Chapter 4.3) was constructed from 16 factors and a landslide inventory map using various sources in QGIS, Saga. Since the implemented models can handle mixed space variables (i.e. numeric and categorical) efficiently, there was no need to dummy the data (i.e. numeric decoding of categorical variables), only the target class (i.e. landslides) was set to "Yes" label if samples are landslide positive, otherwise, it's set to "No". While this database is mainly used to as input dataset to train landslide susceptibility models, an independent testing dataset must be used to properly assess and validate the trained models, otherwise, the trained models will have no scientific meaning (Tien Bui et al. 2016a). However, selecting the size of the training area is very delicate, and requires particular strategies. An optimal approach is to build a sufficiently accurate model with a smaller number of training examples. Thus, lead to reducing expert engagement. On the other hand, the practical value of a model in the landslide assessment framework lies in the model's prediction power, which implies a more meaningful training sampling strategy. On top of that, landslide samples are scares and hard to obtain. Therefore, resampling the input dataset into training and testing sets would be mandatory to obtain reliable results. Additionally, the implemented models and algorithms require fine-tuning some of its hyperparameters[50]. For that purpose, repeatedly cross-validating the input dataset would be effective since the instantiation is done once, so the same training and testing sets are used for (1) hyperparameters tuning, (2) model selection, and (3) performance assessment.

In this case study, the input dataset was randomly resampled into 5 times repeated 10 k-folds cross-validation approach (Figure 3.5A), aimed at optimizing models hyperparameters and optimizing the final models. It's important to understand that the implemented resampling approach is a trade-off in term of speed, accuracy, computational cost and complexity, but also effective it reduces:

- The variance introduces by simple k fold cross-validation.
- The split randomness that comes with holdout-split resampling (test-train split).

[50] Expect for LR.

This would allow the input dataset to be used for three different purposes:

- Tuning models hyperparameters.
- Train models with this subset using after optimal parameters are found.
- Models validation, assessment, and comparison.

It is important to mention that the training area has been selected by sampling instances randomly and uniformly throughout the area.

3.8 ANALYZING AND OPTIMIZING LANDSLIDE CONDITIONING FACTORS

It's common for input datasets used in landslide susceptibility analysis to have high correlation among certain conditioning factors that lead to a faulty modeling with erroneous system analysis [104], so a possible solution can be performing a multicollinearity analysis to evaluate the suitability of the underlying assumption used to select the conditioning factors based on the non-independence among them. To detect and quantify multicollinearity among the chosen variables, PCC [105] can be performed, but in most cases, PCC is not usually sufficient, then VIF is implemented.

3.9 MODELS CONFIGURATION AND IMPLEMENTATION

As mentioned before, the experiment has not been too detailed, which has also reflected the optimization of the modeling parameters. Practically, A simple procedure was conducted for estimating "mtry", "interaction.depth", "n.trees", "num.trees" and "size" hyperparameters, as these hyperparameters are the only with the option of user-parameter estimation according to specific instructions and guidelines. Otherwise, the remaining hyperparameters are exactly bounded to the allowed (or default) values (or range of values) by each package used to implement each model. Overall, these hyperparameters were summarized along with values, short descriptions, and the package used to implement each model in Table 3.2.

Table 3.2 The overall hyperparameters set used by each model along with its respective values.

Model	*Package*	*Parameter*	*Definition*	*Value*
GBM	"Generalized Boosted Regression Models" Formerly: "gbm" package, [106]	distribution	The loss function	Bernoulli
		Shrinkage	Learning rate	From 0 to 1
		bag.fraction	The fraction of the training set observations randomly selected to propose the next tree	0.5 (default)
		train.fraction	Observations fraction that is used to fit the GBM	1 (default)
		n.trees	Total number of trees	From 25 to 210
		interaction.depth	Maximum depth of variable interactions	From 1 to 8
		n.minobsinnode	Minimum number of observations in the trees terminal nodes	20 (default)
LR	"stats" package, [107]	link	Model link function	logit
NNET	"Feed-Forward Neural Networks and Multinomial Log-Linear" Formerly: "nnet" package, [108]	Maxit	Maximum number of iterations	150 (default)
		MaxNWts	The maximum allowable number of weights	10000 (default)
		Rang	Initial random weights on [-rang, rang]	0.5 (default)
		Hess	Find the Hessian of the measure of fit at the best set of weights	TRUE (default)
		Size	Number of units in the hidden layer	From 4 to 33
		Decay	Penalty term or weight decay	From 0 to 1
RF	"A Fast Implementation of Random Forests ranger" Formerly: "ranger" package, [109]	Replace	Sample with replacement	FALSE or TRUE
		respect.unordered. factors	Handling of unordered factor covariates	TRUE (default)
		sample.fraction	The fraction of observations to sample	From 0.632 to 1
		num.trees	Number of trees	From 25 to 210
		mtry	Number of variables	From 2

				to 8
SVM	"Misc Functions of the Department of Statistics, Probability Theory Group, TU Wien" Formerly: "E1071" package, [110]	kernel	kernel function	radial or polynomial
		Cost	regularization cost	From 2-15 to 215 (default)
		gamma (if kernel =: "radial")	kernel width	From 2-15 to 215 (default)
		degree (if kernel =: "polynomial")	Polynomial degree	From 1 to 16 (default)

For the number of variables in each tree (interaction.depth and mtry), various heuristics suggested by packages that provide GBM and RF were used to set the optimum value (Table 3.3). These heuristics suggest that ranges of 1 to 8 and 2 to 8 would be accurate for "interaction.depth" and mtry. The additive nature of GBM, allows for the one-way interaction variable in each tree ($interaction.depth = 1$), on the contrary, RF does not allow one-way interactions, only two-way interactions or more ($mtry \geq 2$). On the other hand, instructions of the used packages and some experimental researches, e.g. [e.g. 111, 112], the total number of trees to fit (n.trees for GBM and num.trees for RF), was set to an exponential value using a base of 2 ($2^i,\ i = 5, \dots, 11$) on which the optimal value is between 2^5 and 2^{10}.

Table 3.3 The heuristics proposed by the package instructions to set the optimum number of variables for GBM and RF.

Package	***Suggested Value***	
	mtry	*interaction.depth*
gbm	N.A	$\sqrt{\boldsymbol{N_i}}$, but often the search space is set between 1 and $\sqrt{\boldsymbol{N_i}}$
ranger	$\sqrt{\boldsymbol{N_i}}$ = 4	N.A
xgboost	6	6
h2o	2 to 8	2 to 8
randomForest	$\sqrt{\boldsymbol{N_i}}$= 4	N.A

$\boldsymbol{N_i}$: the total number of variables (i.e., 16 in this research)

Furthermore, the number of nodes in the hidden layer ("size") for NNET was set in a range of 4 to 33 according to empirical suggestions proposed by different authors summarized in Table 3.4.

Table 3.4 The heuristics proposed to compute the optimum number of hidden layer nodes for NNET (modified from and Kavzoĝlu [113]).

Proposed by	***Heuristic***	***hidden nodes***
Hecht [114]	$2N_i + 1$	33
Ripley [115]	$(N_i + N_o)/2$	8 or 9
Paola and Schowengerdt [63]	$\frac{2 + (N_i * N_o) + \frac{1}{2}N_o(N_i{}^2 + N_i) - 3}{N_i + N_o}$	9
Wang [116]	$2 * N_i /3$	11
Aldrich, Van Deventer [117]	$\frac{N_p}{k(N_i + N_o)}$ $(k = 10)$	7
Aldrich, Van Deventer [117]	$\frac{N_p}{k(N_i + N_o)}$ $(k = 7)$	10
Kaastra and Boyd [118]	$\sqrt{N_i * N_o}$	4
Kanellopoulos and Wilkinson [119]	$2N_i$	32

$\mathbf{N_i}$: number of input nodes (i.e., the total number of variables of 16 in this study); $\boldsymbol{N_o}$: number of output nodes ; $\boldsymbol{N_p}$: Number of training samples; $\boldsymbol{k}$: the noise factor (varies between 4 and 10) is an index number representing the percentage of false measurements in the data or degree of error

In the end, SMBO was implemented using an initial design grid of size 40 with 30 iterations budget, and the lower confidence bound (α_{LCB}) as infill criterion for optimizing the implemented models and they respective hyperparameters (Table 3.2).

3.10 MODELS VALIDATION AND EVALUATION

Different performance metrics can be implemented for quantitative comparison; however, landslide susceptibility problems are strictly classification problems where quality and confidence in probabilities toward land sliding is critical. Therefore using a performance metric to assess prediction robustness is necessary and for this reason, the area under the ROC curves (AUC) will be implemented as the only metric for the objective functions in hyperparameters tuning and one of

three overall performance indicators of the landslides predictive models. Additionally, the overall performance and the predictive capabilities of the tuned models was assessed using not only the robustness of the prediction using AUC as it is not enough, the accuracy and reliability of trained models also assessed using the overall Accuracy (ACC) and Cohen kappa index (κ-index), respectively.

Moreover, model performance results were tested using nonparametric statistical procedures for statistical significance to evaluate and compare landslide susceptibility models against each other using the Wilcoxon signed-rank test at the 5 % significance level for each pair of models to individually detect differences in model performances. Basically, the Wilcoxon signed-rank test relies on a null hypothesis (i.e., there are no differences between the performances of the landslide models), on which values called $p.value$ and $z.value$ are used to determine the probability of rejecting or accepting the null hypothesis [102]. If $p.value$ is lower than the significance threshold (i.e. $p.value < 0.05$) and z.value exceed its critical values (i.e. $z.value < -1.96$ or $z.value > +1.96$), it's safe to assume that the null hypothesis is not valid and can be rejected and therefore a significant difference between the two compared models exist, otherwise (i.e. $p.value \geq 0.05$ and $-1.96 \geq z.value \geq +1.96$) it's safe to assume the opposite.

3.11 LANDSLIDE SUSCEPTIBILITY MAP GENERATION AND ASSESSMENT

Apart from the performance metrics, sufficiency analysis must perform to assess the sufficiency and accuracy of predictive models that produce landslide susceptibility maps. This analysis is based on the assumption that: "A model is sufficient and accurate when there is an increase in the landslide density ratio when moving from low to high susceptible classes and high susceptibility classes covers small areas extent" [4, 40, 120-122].

The sufficiency analysis can be performed based on reclassifying the continuous probability grids (ranging from 0 to 1) generated for the study area generated by each susceptibility model into five standard categories of relative susceptibility as described in Table 3.5. The Standard deviation method was chosen to classify the susceptibility. The class breaks were determined by the supported mean value. Subsequently, by overlying the landslide inventory; it's possible to generate a summary statistic for each class (landslide density and area extent).

Table 3.5 Probability intervals for landslide susceptibility classes.

Susceptibility Class	*Very Low*	*Low*	*Moderate*	*High*	*Very High*
Probability Range	From 0 to 0.05	From 0.05 to 0.30	From 0.30 to 0.60	From 0.60 to 0.75	From 0.75 to 1

It's important to understand that only the highest and lowest susceptibility classes, i.e. Very High and Very Low, were regarded for sufficiency evaluation against the referent Landslide Inventory. This was inspired by the fact that existing landslides and Non-landslides should be marked as a priority zone (preferably as Very High and Very Low susceptibility class).

Chapter 4: Case Study

This chapter is featuring the Objectives 2 & 3, and indirectly all the others (see Chapter 1.3).

The problematic of landslide susceptibility, in the context established throughout this thesis, was practically unattended in this study area in the past. There has been a host of practical considerations, mainly small geotechnical projects and reports, tightly related to the landslide problematic for various purposes, mainly site-specific ones, for construction design, few studies at regional scales for urban and regional planning, or just some plenary researches targeted at different geological aspects were carried. Thus indicate, that this research is just barely scratching the surface in term of the landslide hazard and susceptibility problematic.

Nevertheless, there was a national plan for nation-wide engineering and hazard mapping in a scale[51] of 1:200000 by the end of the 20th century. These maps should have matched the existing geological map on the same scale, but the idea was not realized to date. Such a situation with data availability affected the initial stage of this research, but the case study area turned resourceful after recompiling and merging separate patches and sheets of the data from CAD files to GIS Layers.

It is important to outline, that this study area has been researched for three years and there have been different aspects where the elaborated work was published.

4.1 BACKGROUNDS

The following paragraphs will focus on presenting the essential backgrounds about Mila basin rather informally; in order to stay committed to the main problematic of this research. Details such as geology, hydrogeology and so forth, about the study area are well documented and presented in comprehensive details in different literature (e.g. [123-137]).

[51] PER and ZERMOS.

Geography

Mila Basin is situated in the northeastern part of Algeria between longitudes of $5.921°E$ and $6.828°E$, and latitudes of $36.185°$ N and $36.611°$ N and covering an area of approximately $\approx 2760\ km^2$ distributed over 42 municipalities (mostly over the central parts of the Mila and Constantine provinces). Geographically, the study area is fully surrounded by mountain ranges such as M'Cid Aicha and Sidi Driss from the North; Djebel Ossmane and Grouz by the South; Djebel Akhal, Chettaba and Kheneg from the East; and Djebel Boucherf and Oukissene by the West (Figure 4.1).

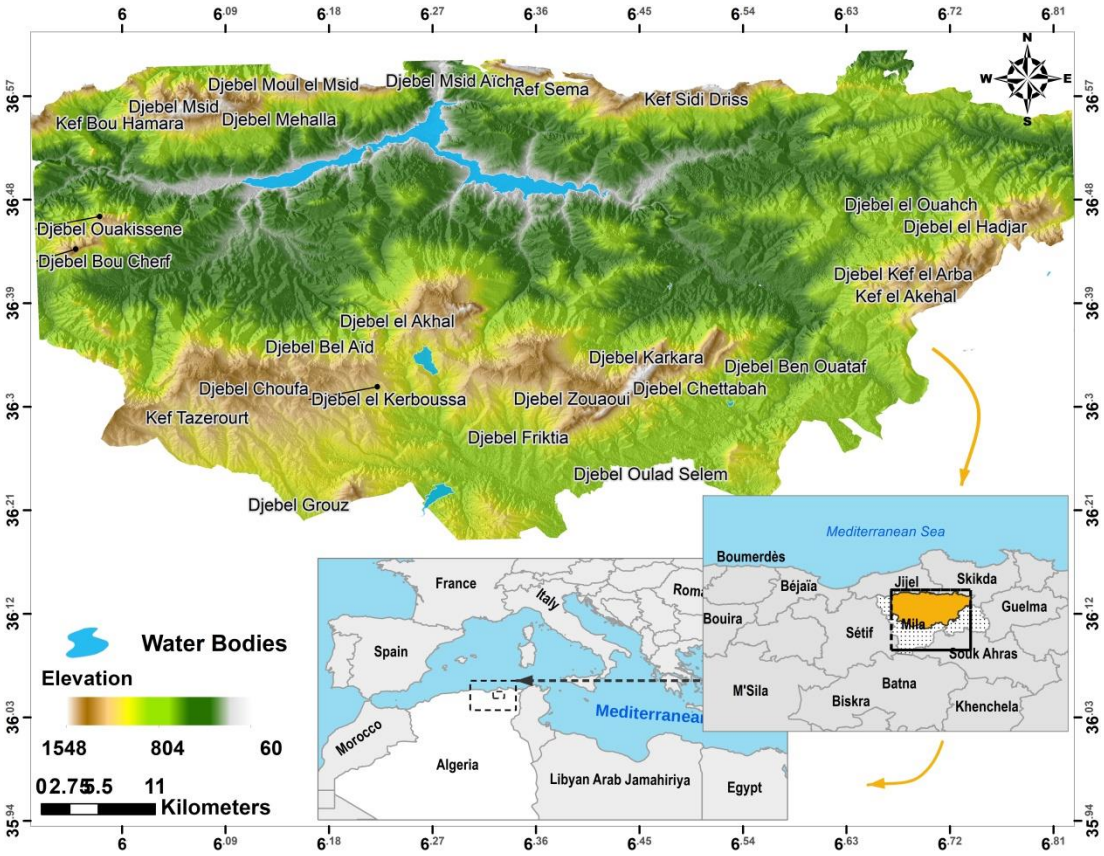


Figure 4.1 The geographical location of the study area.

Landuse and Vegetation

Landuse is mostly for bare lands, cereals crops or wild herbs. This low-density vegetation is making the basin a hotspot for agriculture investments and farming industry (i.e. cattle breeding, poultry, stock farming...etc.). However, such vegetation accelerates land degradation and instabilities by soil erosions.

Climate

In Mila basin, usually, the wet season is relatively short compared to a long dry season. The local climate can be divided into two separate entities:

- Semi-arid with a mild winter denoted by significance difference in temperature (reaching $40°C$ and below $0°C$ during summer and winter, respectively) and reaching an average of 500 mm/year (Figure 4.2).
- Sub-humid fresh climate (typical for a mountainous landscape) surrounding the first entity and denoted by relatively dry and hot dry season, fresh and humid wet season. The precipitation mean is fluctuating between 900 and 1200 mm/year [127, 135].

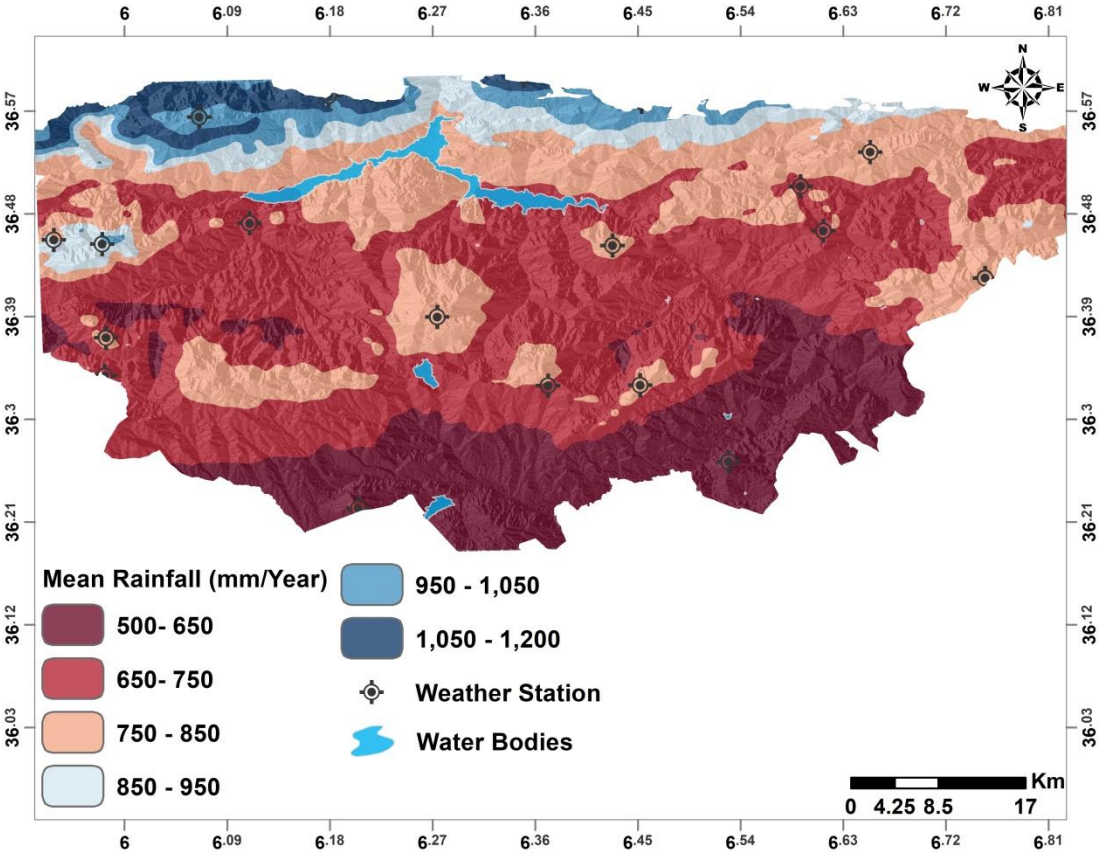


Figure 4.2 The mean rainfall map of the study area.

Hydrology and Hydrogeology

The study area is technically a high elevated basin (mean elevation surpasses $500\ m$), which is part of a much larger watershed called "Kébir Rhumel". This basin, i.e. Mila basin, is characterized by asymmetrical elongated geometrical

form (along the East-West direction) drained by a dense and hierarchical hydrographic network in N-S direction depending on the stream (Figure 4.3) [131].

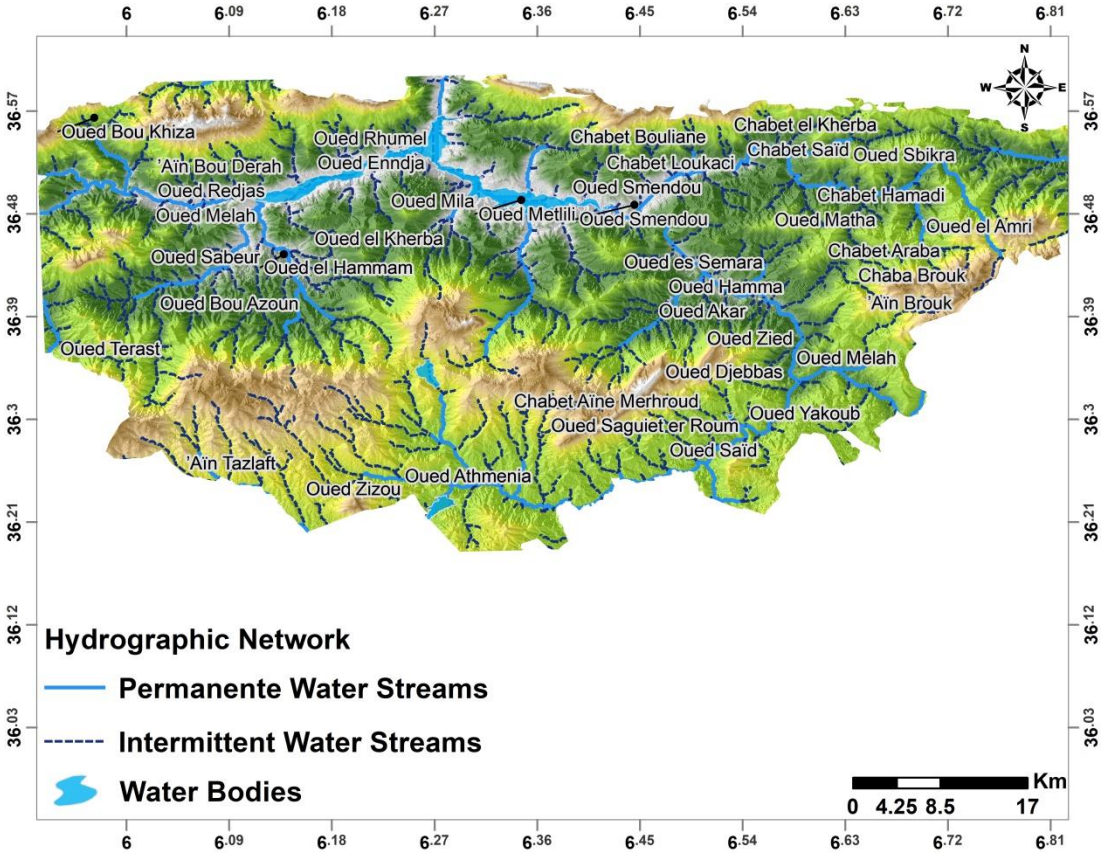


Figure 4.3 The hydrographic network map of Mila basin.

In terms of hydrology, the hydrographic network streams are mostly depleted of water during the dry season. This renders the streams with less useful to the overall local economy. However, during wet season everything change and the overwhelming flow[52] highly contribute to soils erosion. This case, in particular, exposes human settlements and constructions near the hydrographic network streams to the risk of land instabilities.

On the other hand, from a hydrogeological perspective, the study area does not possess any important aquifer. However, theoretically,[53] there exist formations with to formulate and (or be) aquifers for groundwater bodies such as:

- Quaternary formations with mainly Alluvium deposits.

[52] Especially during intense rainfall.

[53] Mila basin highly suffers from lack of underground data such as hydrogeology due to the nature of geology.

- Sand, sandy and/or sandstone deposits available in the Mio-Pliocene formations, especially if it is deposited in lenses.
- Lacustrine limestones have a high potential of retaining high capacities of water[54].
- High infiltration zones such as shear zones (Major tectonic accident like faults) are suitable areas for water infiltration and seepage where different springs and resurgences can upsurge randomly (frequently observed during foundation excavation for infrastructure project, i.e. Beni Haroun dam, RN27 and RN79 maintenance...etc.)

Seismicity

Seismicity in the study area is moderate according to CRAAG (Le Centre de Recherche en Astronomie Astrophysique et Géophysique). Mila basin is located within Zone II (Figure 4.4) and is characterized by moderate seismic activities. However/Moreover, the basin has suffered previously on multiple occasions from various intensive seismic events that vary in terms of magnitude (Figure 4.4). But overall, the upper parts of the basin are previously affected by seismic activities in the past[55] [133].

[54] Can vary depending on factors such as fracturing and karstification states.

[55] Especially in the NW and South-east of the basin.

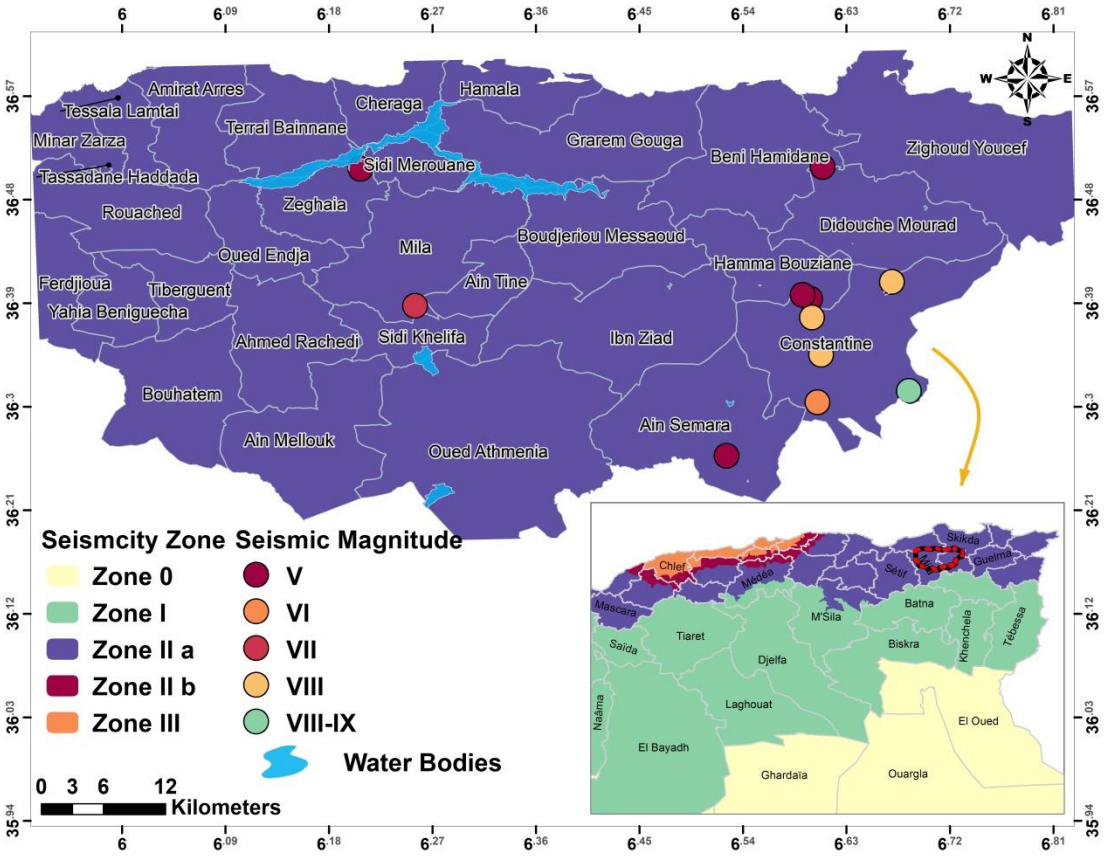


Figure 4.4 Seismic maps with the historic seismic events of the last 50 years of the study area.

(Merghadi, Abderrahmane [138] Edited, after CGS).

Geology

Notably, the study area belongs to a paleogeographic domain known as "domain tellien", which is technically, the oriental segment of a "chaîne" formally known as "chaîne des maghrébides". This chaîne was set up by the "Alpine Orogeny" during the Miocene epoch in north-western Africa[56] [128]. More specifically, a larger Neogene basin known by the "Constantinois basin" encompasses the study area and as stated before[57], several mountainous ranges that belong to different paleogeographic domains surround the study area and constitute the basin bedrock [126].

[56] For that purpose, it's known as "la chaîne alpine d'Afrique du Nord".

[57] See Chapter 04.1.

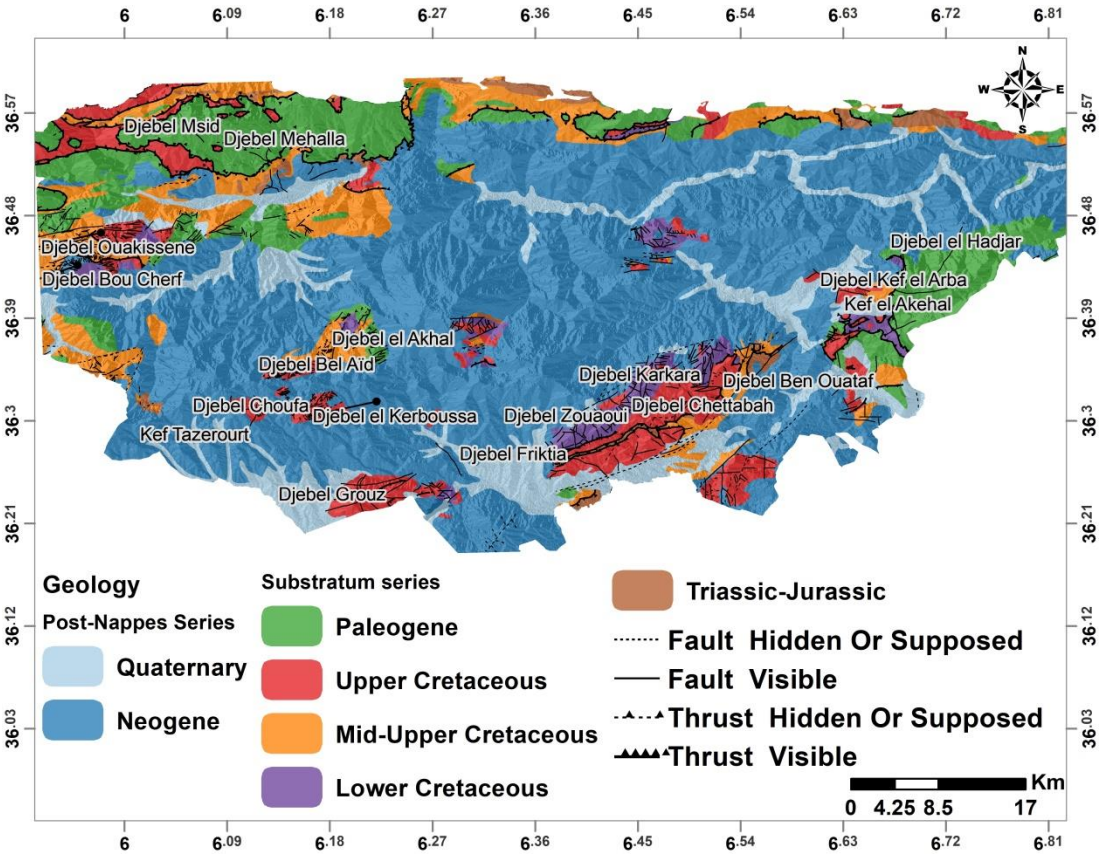


Figure 4.5 The geological map of the study area.

In terms of tectonic activities, the study area shows a tectonic complexity due to some severe conjugation of folds, faults, and thrusts of different ages and styles that are the results of two main tectonic events. The first phase was the "Atalsic phase", which is responsible for forming major recumbent fold structures oriented in a NE-SW direction. The second was the "Alpine phase", which is responsible for breaking and sliding existing formations one over the other to form gigantic thrust faults, which resulted in a thrust belt of structures oriented in N-S direction.

According to Coiffait [126], there exist two general systems of lineaments for the basin based on the orientations of the structures generated by the aforementioned major tectonic events:

- The diagonal system, with NE-SW and NW-SE orientations. This system is dated to the late Eocene. This system is directly responsible for creating and generating some important structures (i.e., folds and Horst-Graben) in the basin. These structures were the source of the detritus materials during the Neogene.

- N-S, E-W lineament system, also known (also known as "Vertical lineament system"). This system belongs to a recent compression phase that is responsible for the current morpho-structure of the study area.

Essentially the local geology of the study area consists of different lithostratigraphic units (Table 4.1 and Figure 4.5) and can be summarized into two groups, called "series" [126]:

- Substratum/Bedrock series, which formulate both the lower base and the bedrock of the basin and consist of Triassic to Paleogene formations.
- Post-nappes series constitute a cover to the bedrock series and consist of Neogene to Quaternary formations. These series, in particular, were slightly affected by recent neotectonic deformations.

Table 4.1 The geological formations present in Mila Basin.

Unit	***Period***	***Epoch***	***Description***
Post-nappes	Quaternary		Alluvium, colluvium, scree, detritus deposits and slopes formations like terraces.
	Neogene		Predominantly detritus composed of clay, marl, limestone, conglomerates, sandstones, sand, lacustral limestone and evaporitic formations.
Substratum	Paleogene	Eocene	Limestone, cherty limestone, and platted marls.
		Paleocene	Opaque to somber marls
	Cretaceous	Upper and Mid-Upper Cretaceous	Marl dominance58.
		Lower Cretaceous	Mainly marly limestone and neritic limestone.
	Jurassic		Mostly thick carbonate formations (dolostone, limestone, and cherty limestone).
	Triassic		Evaporitic and clayey deposits.

[58] Variation are ranging from different horizons of gray marly limestone, alternating marl, and limestone, blueish marl, massive bars of limestone, to alternating marl, cherty limestone, and thin micritic limestone all surmounted by grey marls with conglomerate interbeds.

In general, the study area is predominantly covered by clayey formations. The existing mineralogical units were summarized according to some researcher's previous works (e.g. Zouaoui [137], Chettah [125], and Athmania, Benaissa [123]) into Table 4.2.

Table 4.2 The mineralogical groups existing in Mila Basin.

Group	Mineral	Description
Clay minerals	Illite	Less sensitive toward water content variation (shrink-swell due to water presence). In terms of size and dimension, it varies between 3.96 A° to 7.63 A° and can reach sometimes 8.59 A°.
	Montmorillonite	Highly sensitive to water content variation. Generally, montmorillonite minerals are well structured resulting in highly cohesive soils during the dry season. However, during the wet season with the high amount of water content infiltrate deep soils montmorillonites mineral become vulnerable to water and in the end resulting in solifluction [139].
	Kaolinite	Similar to Illite minerals of being less sensitive toward water content variation. in fact, Kaolinite minerals are counted as the least sensitive toward water content variation in clay minerals family [123].
	Chlorite	Are the most predominant clay minerals in the study area, which generated from the weathering process of the Biotite minerals. Generally speaking, soils with a high percentage of Chlorite (surpass 15%) are counted as susceptible to landsliding (due to high shrink-swell capacity) [123].
	Saponite	Similar to Chlorite, Saponites, and Vermiculites minerals are Highly sensitive to water content variation, but with lower percentages than the remaining clay minerals (i.e. Illite, Montmorillonite, Kaolinite and Chlorite)
	Vermiculite	
Other minerals	Quartz	These minerals are generally found with clay minerals in different percentages that differ depending on the location samples are sampled/drawn from [123, 125].
	Feldspar (Low percentages)	
	Sulfate minerals (e.g. Gypsum, 1 to 3%)	
	Carbonate minerals (e.g. Calcite and Siderite, 10 to 29%)	
	Ferruginous minerals (up to 6%)	

Geomorphology

Geomorphological speaking, Mila basin is composed of multiple terrains of carbonate formations emerging deep within the existing heterogeneous Neogene formations. These structures are the result of the tectonic heritage of the pre-disposition (i.e. pre-sedimentation) of the Neogene detritus formations. Such configuration is responsible for the morphological heterogeneous terrains in terms of the observed spatial morphological entities [137] (Figure 4.6).

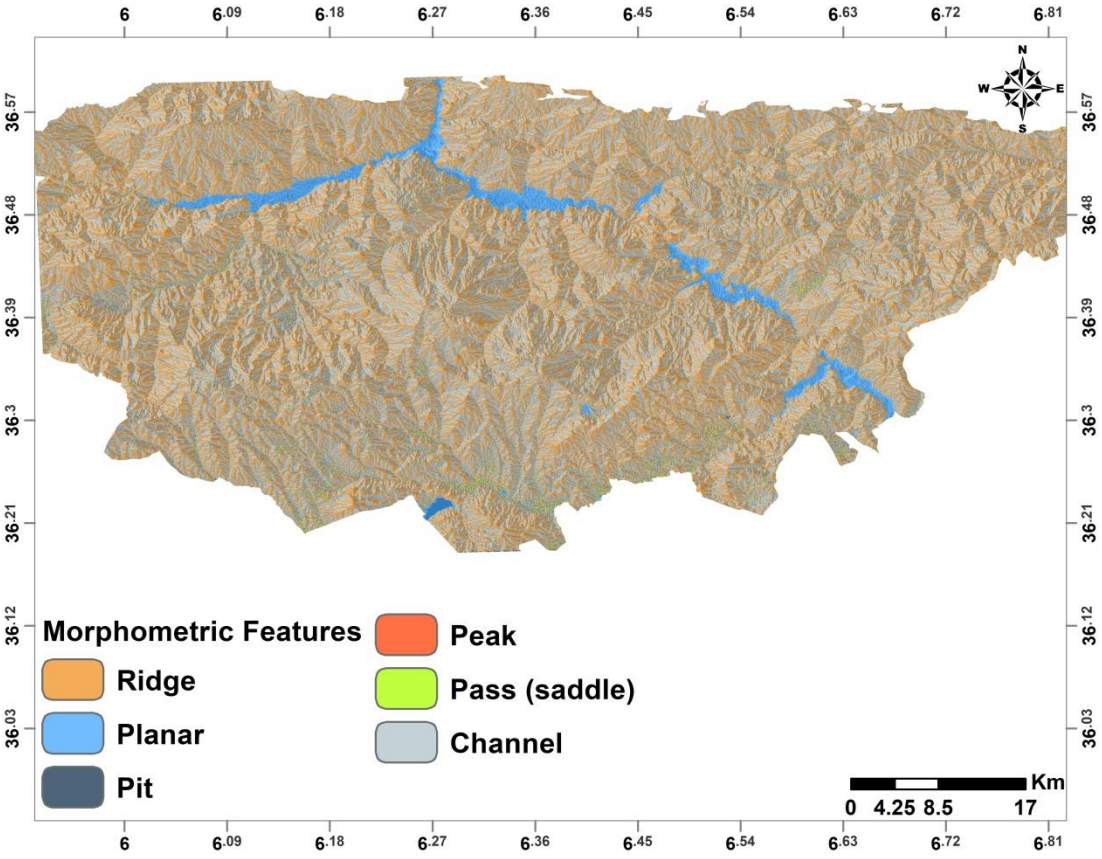


Figure 4.6 The geomorphological map of the study area.

For example, the mountainous landscape of ridges and hills that encompass rugged channels and pits is the most observed landscape. This prominent pattern is generally stretching over large proportions (Table 4.3), especially northern parts of the study area where the terrains elevate rapidly.

Table 4.3 The existing morphometric features present in Mila Basin.

Feature	***Percentage (%)***
Planar	11.581

Pit	0.257
Channel	41.625
Pass (saddle)	0.806
Ridge	45.494
Peak	0.237

According to PDAU (Plan Directeur d'Amenagement et d'Urbanisme)[59] reports, this prominent landscape is essentially characterized by reddish hills and ridges with hummocks and/or undulated terrains of bare to less vegetated lands. During, the wet season, substantial green vegetation covers these terrains, whereas, during the tillage period only the reddish-brown color of the Neogene formations (i.e. clay and marl), is noticeable.

4.2 LANDSLIDES IN MILA BASIN

Landslides are a highly pronounced issue in the study area noticeable by the consistent symptomatic indications of the phenomena through man-made constructions such as roads, pavements, slopes, embankments, powerlines, and water-sewage pipelines. The dramatic variance of landslides in the basin, in terms of spatial repartition and intensity, is a very serious handicap to the urban, local, social and economic development of the basin since 1985. Over the year, the ever-increasing rate and magnitude landslides, increased the number of the element at risk exposed to landslides, especially in the urbanized areas. Different remedial actions and innervations were proposed and highlighted relating instable areas but they were not fully compatible, as they did not consider soils intrinsic properties, landslides characteristics, landslides behaviors and patterns, rending these remedies either insufficient or incompatible (Figure 4.7).

Despite the remedial projects that have been carried out in the recent years, the effects of landslides in term of damages are still persisting and sometimes even worse, as these remedial actually focus on treating the symptoms of landslides rather than the landslides issue itself without consideration of soils intrinsic properties or landslide patterns behaviors.

[59] Published in 2007-2008 for Wilayas of Mila and Constantine.

Figure 4.7 Example of incompatible remedial actions and innervations.

The pile sheet walls were implemented to stop the continuous deformation of the slope mass. However, stabilizing the slope would be appropriate, rather than fixing landslides symptoms (Source: Mila municipality, Location Mila, date: October 2015).

Climatic, geological, geomorphological and human-related characteristics of the basin are in favor of landslides of different shapes, sizes, and types. These predisposition factors, along with the already complex landslide failure mechanism, are complicating this hydro-geomorphological phenomenon, which leads an unexpected evolution of landslides in both spatial and temporal space components. In general, common indicators and evidence of landslides can manifest sometimes in form of scares or continuous slope deformations (e.g. solifluction and creeps) that indicate a deeper and profound dynamics, which needs to be treated with caution and special attention. Usually, disregarding and ignorance about the conditioning and triggering factors of landslides tend to lead to very complex and critical situations. For example, outside or inside urbanized areas in Mila basin (i.e. cities, villages,...etc.), are all experiencing the same symptomatic indicators and evidences such as roads failures, fractures and/or leakage in pipelines [60](i.e. water and sewage), without any concerns to the causes or the potential issues these problems or the

[60] They can be out of service for extended periods of time that can reach couple of weeks.

symptoms may indicate or even influence (e.g. sewage seeping and saturating slopes) (Figure 4.13a-d).

The study area possesses dynamic diversity, in term of landslides that can vary from simple erosions to extremely complex and dangerous types of landslides like, composite landslides (i.e. complex landslides), or even flows and spreads that develop inside gullies and bad-lands. Overall, this diversity is related not only, to the highly dynamic physical proprieties of the basin, but also the human-related factors. These factors are highly contributing to the ever ending complex scenarios by introducing landslide mechanisms that are difficult to predict and generalize (Figure 4.9). The landslide displacement mechanisms observed in the study area can be summarized as the following (for more theoretical details see Chapter 2.1):

- Slope deformation (Solifluction), is frequent in the coherent tender plastic formations[61] that tend to exhibit continuous deformations and suffer from the swelling and shrinking tendencies, but never practically revert to its original state. This is noticeable when runoff-water infiltrates the upper permeable formations, especially in streams (river banks bed) (Figure 4.8).

Figure 4.8 Solifluction of slopes near streams.

Runoff-water infiltrates the upper permeable formations and the slope mass fails gradually under its own weight in downslope movement (Source: Chettah [125], Location: Mila; Date: Unknown).

[61] Mostly clayey and marly formations but sometimes with or without sand.

- Slides[62]:
    - Rotational slides, this is typical for clayey formations that vary from few meters, up to hundreds of meters in length (> 200 m), and can vary in width from 10 ~ 30 meters [139]. For instance, multiple and successive overlapping rotational landslides tend to impregnated slopes with symptomatic indicators and evidence such as scars, that indicate a complex dynamics of conjugated rotational landslides. In most cases, removal of pins and abutments available in slopes feet and toes, exacerbate and increase slopes failure rate (Figure 4.9). This scenario can be accelerated with open ponds and superficial water runoff, which generated mostly from leakages of failed sewer and drainage systems that suffer from leaks and ruptures. In some cases, the leakage can worsen the process by introducing important and significant gullies. Additionally, the scouring of slopes and embankments[63] is noticeable in marly-clayey formations. The overwhelming majority of the landslides are shallow, sometimes successive and deep-seated unless slopes abutments and pins are removed.

Figure 4.9 An example of a landslide where the slope fails due to scouring.

[62] See Figure 4.13 in Chapter 4.3.1 for some examples.

[63] Scouring is the process of cutting slope's feet and/or toes.

(Source: Chettah [125], Location: Mila; Date: Unknown)

-
- Overall, rotational slides often presented in different generations that can be nested and embedded successively one inside the other. As consequence, this will be reflected on the slope surface by generating a chaotic landscape (i.e. embossing, dents, ripple, waves, and creeps) marked by numerous counter-slopes and ridges (Figure 4.10) and small to moderate gullies with fresh notches (cuts) and depletion that favor lateral subsidence's.

Figure 4.10 A chaotic landscape example generated by a successive deep rotational landslide.

Dents, ripple, waves, and creeps marked by numerous counter-slopes, ridges and gullies with fresh notches (cuts) and visible depletion that lateral subsidence's (Location: 185 Log; Date: March 2016, Source: Mila and Constantine municipalities).

- Planar slides is explicitly found in streams precisely Oueds and Chaâbats, where materials slide in downward movement inside the riverbed (especially during floods) along successive sub-vertical to parallel slipping-surfaces to the streams hillslopes geometrical plans. This process happen when there is an undermining and sapping of concave streams banks by running water. In some extreme case, landslides can further progress when successive subsidences occur downstream. These types of slides are usually of hundreds of meters in length and 1.5~2 meters in width [139] and mostly simple but

sometimes can evolve into a hybrid of landslides of debris flow during an exceptional rainfall precipitation period and floods.

- Flows, spreads, and Falls, can vary depending on original slopes materials (see Chapter 2.1, from debris flow to mudflow, rock slope spread, sensitive clay spread...etc.), as they are generally sparse and less-spread over the study area and mostly limited in form and size as they scatter around edges of streams, downslopes (Figure 4.11a), or if slopes ridges and peaks are Limestone formations (Figure 4.11b).

(a) Debris Flow (Location: Oued Elkherba; Date: Unknown)

(b) Debris Fall (Location: Mila, Date: Unknown)

Figure 4.11 Landslide examples of spreads and flows available in Mila basin.

(Source: Chettah [125])

According to survey campaigns achieved by local authorities (2003-2017), it was reported that slopes in the study area fails under a conjunction of both predisposition factors (i.e. geology, lithology geomorphology, faults...etc.) and triggering factors (i.e. intense and persistent meteorological events, human activities,...etc.) resulting in landslides of different sizes and types. These Reports suggest that:

- Long and persistent periods of intense to moderate rainfall, are the main culprit in triggering and/or reactivating existent deep-seated landslides due

to the high amount of water infiltrating underground. On the contrary, short and intense to moderate rain storms and/or precipitation, are indirectly affecting slopes stability by an intensive erosive process.

- Rainfall variation turns out to be very significant for this study area as a large amount of precipitation is recorded during the short wet season. Thus, slopes are highly vulnerable during that period of time of the relatively short wet season.
- Human related activities can exacerbate sliding when drainage systems fail or when urban development increases runoff near steep slopes. Additionally, expanding infrastructures such as road and settlement areas are in some cases considered as influencing or triggering factors[64]. Some evidence shows that cutting the gentle slope for building settlement turns out to be a triggering factor for landslides.
- Water in general, has an explicit influence on slopes geotechnical and mechanical properties, especially if absorption and permeability of slopes soils are relatively high.
- Landslide typologies are different and include mostly: (1) translational slides hosted in the shallow deteriorated mantle of the Neogene complex formations and/or stratified Quaternary formations; and (2) Shallow to deep-seated slides hosted in Tertiary formations, wherein the landslides are rainfall-triggered, while sometimes triggered by the erosion-groundwater dynamics.

4.3 GEOSPATIAL DATABASE

Constructing the geospatial database is one of the most crucial steps in successful landslide susceptibility modeling. This process usually consists of constructing an inventory database of landslides from historic events (past and present) and conditioning factors database that control and determine the failure mechanism and the triggering process.

[64] Case of RN27 and RN79.

The geospatial database is prepared as a GIS database that contains all necessary information's regarding geographic location, features, conditioning factors and relative statistics, about the 578 mapped landslides. The conditioning factors and landslide inventory available in the geospatial database are acquired from different sources and they are represented in 2D GIS thematic raster layers, entailing different levels of generalization and different scales. Subsequently, a generalization to common 30 m cells raster resolution is plausible[65] (Fell et al. 2008), considering the fact that the most critical data have reasonably, a small scale that turned out to be too detailed for the research purpose (e.g. Geological map are 1:50000). Therefore, the geospatial database has been recompiled and resampled (either by up-sampling or down-sampling) to an optimal 30 m raster resolution. As a matter of fact, an additional generalization (by aggregating similar classes) took place over such inputs. Finally, assembling of the geospatial database into an input dataset has been prepared by QGis, SagaGIS and R. The resulting database was stored in HDF, JSON and Array formats, which proved to be efficient in absence of a fully integrated standardized solution for data exchange between ML modules and GIS platforms.

4.3.1 Inventory Map

A detailed landslide inventory map has been compiled for the period of January 1985–December 2017 with only slide failure types have been elaborated (Figure 4.12) using mainly:

- Historical records provided publicly by the local agencies (i.e. municipality of Constantine and Mila) with a 531 landslide event.
- Google Earth Pro® software 47 landslide events were detected and mapped (from 2000 to 2017).

On the other hand, the non-landslide samples were easily obtained by random sampling a unique 578 sample site (equal to the total number of landslide samples) from public stability maps available at DUC (Direction d'Urbanisme et Construction) using PAW (Plan d'Amenagement de Wilya) and PDAU.

[65] 30 meters scale is usually acknowledged as optimal scale for regional analysis.

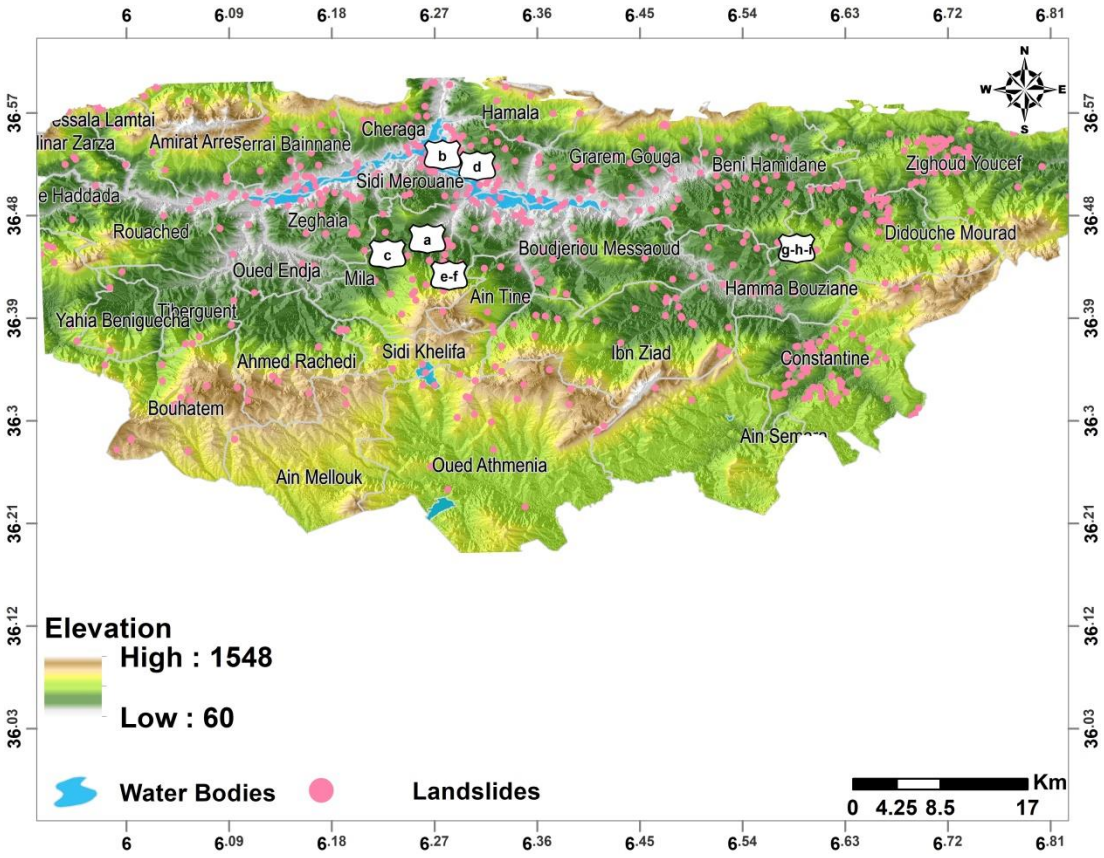


Figure 4.12 The landslide inventory map of the study area.

The validity of the compiled inventory database (landslide and non-landslide samples) has been verified on the field, using extensive field inspections (especially for the landslides that have been mapped with lower certainty) by practicing conventional engineering-geological mapping methodology and low-accuracy navigation device (still sufficient for 30 m inventory). These field inspections were limited to the smaller landslides only, i.e. only those observable on the field. Otherwise, Aerial Photography provided by Google Earth was used to validate and check the mapped sample events. The aforementioned verification methodology implies evidencing of the landslides, but also to find evidence which could support the activity estimation, as well as estimation of the triggering mechanism, landslide depth and type (in order to conform to the adopted classification system). However, these observations and measurements[66] during field inspections are standard, limited in scope, have not been systematically carried and the data have been collected

[66] Measurement of landslide morphology and metrics, depth estimation, tension cracks, object deformations...etc.

randomly[67]. Also, field revisits served the purpose of only ensuring some particular occurrences with adjacent to the available landslide event samples (Figure 4.13). Annual revisits allowed visual monitoring of particular occurrences.

(a) RN 79a, (Type: Deep-Rotational landslide; Date: October 2011)

(b) Sibari (Type: Shallow-Planar landslide; Date: February 2008)

(c) Mila (Type: Deep-Rotational landslide; Date: September 2013)

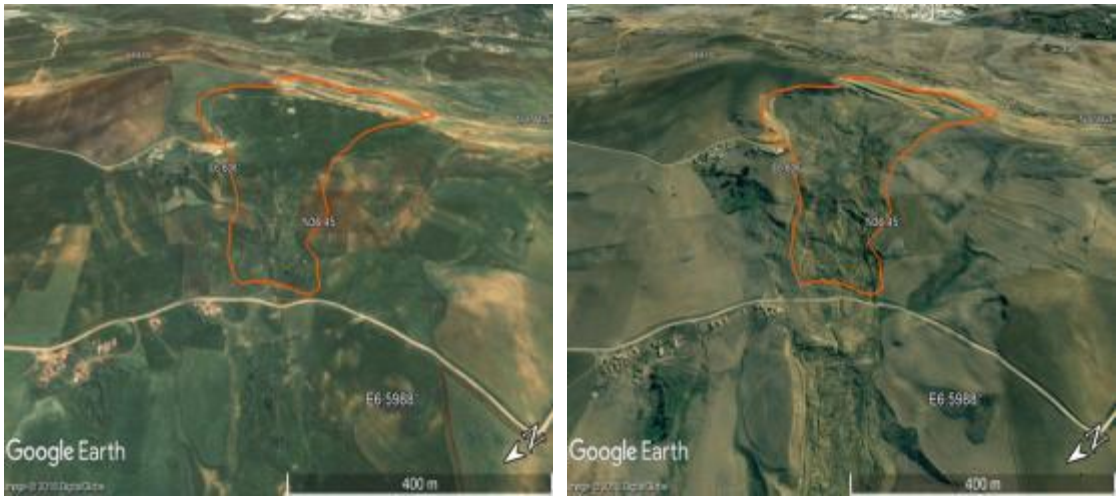
(d) Grarem (Type: Planar landslide; Date: June 2015)

Figure 4.13 Landslide examples used in the landslide inventory.

(Source: Mila and Constantine municipalities, Location: see Figure 4.12).

[67] Especially one provided by local governmental agencies was inconsistent and lack details in most case.

(e,f) Mila (Type: Deep-Rotational landslide; Date: October 2017)



(g, h and i) Didouche Mourad (Type: Deep-Rotational landslide; Date Left: August 2003, Date Right: September 2005, Date Bottom: March 2016).

Figure 4.13 (continued).

The main landslide characteristics were described according to WP/WLI [11] (See Table 2.4) standard recommendations. Common indicators of the active sliding can be found in geomorphological, hydrogeological, botanical evidence, as well as in deformations of the man-made structures. For instance, fresh scars opened tension joints locally filled with water, local pounds and hummocky topography are strong evidence of recent activities in the depletion and accumulation zones. Also, fresh fissures in the buildings or paved roads, disarrangement of the fences and staircases (as most fragile constructions), and tilted tall objects such as poles or trees are further supporting the activity assumption. Information from the members of the local community is also appreciated, especially for the dating of the landslide events, estimating the water table levels in aquifers, estimating the activity rate, estimating the trigger and assessing the damage produced by single or multiple events.

Although slides dominate throughout the study area, several flows spreads, and falls are also present, particularly in the northern part where steeper slopes and narrower valley channels, exist. Since, these types have entirely different phenomenology (geometry, dynamics and mechanism) it is logical to assume that different conditioning factors will have a different effect in each type of movement. Thus, leads to different and separate investigations. However, the emphasis was on susceptibility analysis and the proposed methodology required an analogous type of subject to model. Therefore, in order to remain consistent with the current methodology, the other landslides types (flows, spreads, falls and so forth), have been excluded from the inventory, and only slide type of failure was kept and elaborated. Furthermore, the landslide inventory has been somewhat simplified in order to enhance the statistical representativeness of landslide vs. non-landslide categories. In this context, original landslide classes (WP/WLI classification) have been unified.

4.3.2 Conditioning Factors

In susceptibility analysis, landslide conditioning factors need to be operational, complete, non-uniform, measurable, and non-redundant [140, 141]. However, the selection process of landslide factors is very subjective comes with difficulties (i.e., the study case, scale of the analysis, and data availability, general guidelines for GIS-based studies,…etc.), which explain the variations in landslide susceptibility

modeling studies in term of the conditioning factors used for the analysis. However, despite the fact that there are no clear guidelines about the proper factors to use for such a kind of analysis [141], the conditioning factors (Appendix B) were selected for this case study based on:

- Field survey observations.
- Survey campaign reports achieved by local authorities.
- The most commonly used factors in the literature for landslide susceptibility analysis [e.g. 142, 143, 144].
- Geo-environmental factors of the study area that may directly or indirectly affect landslides and can be used as predisposing factors [33].
- The scale of the analysis and data availability for the case study [145].

In this case study, a total of 16 conditioning factors that describe Mila basin terrain attributes, such as factors regarding geo-morphometric ground surface morphology, hydrological, geotechnical and subsurface, geological and environmental features, as well as some derived synthetic features; were considered as suitable for this thesis.

Geo-morphometric Data

Geo-morphometric or topographic data parameters were generated from the Digital Elevation Model (DEM) of the study area at near 30 meters cell resolution from NASA Shuttle Radar Topography Mission Global 1 arc second (SRTMGL1) mission[68]. The DEM resolution of 30 was chosen for two basic reasons:

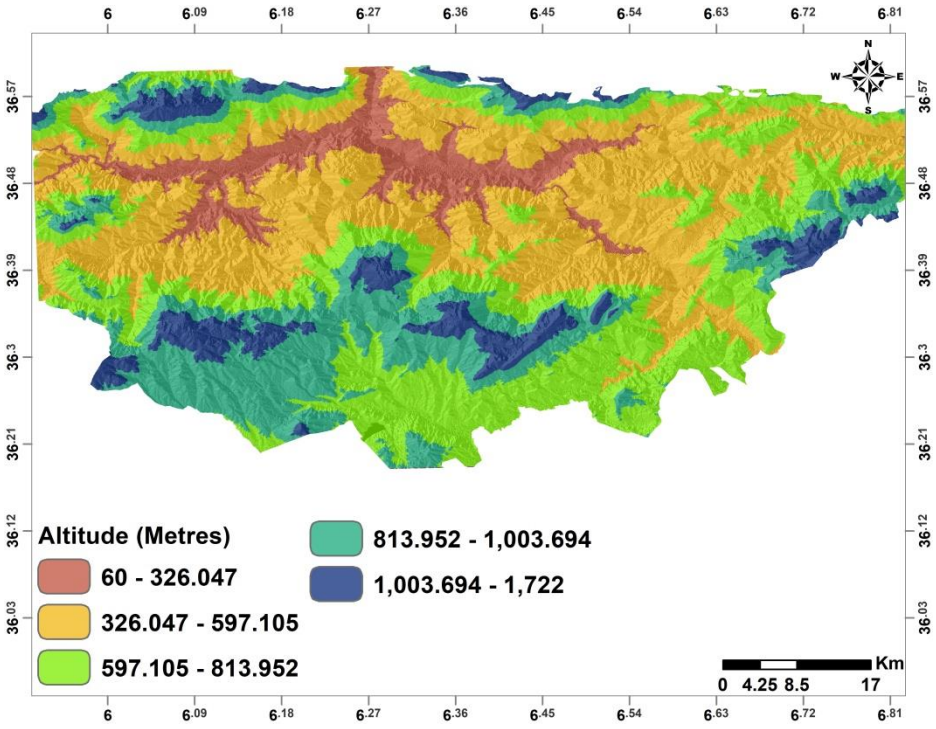
- The sufficiency of 30 meters resolution DEM, considering the overall extent of this case study landslide susceptibility analysis
- The adequate support and compatibility that 30-meter grid size would provide regarding other data sources used for this case study (e.g. geological map at 1:50000 scale). As a result, all geo-morphometric and hydrological DEM derivatives are kept at the same 30 m resolution as the source DEM.

[68] For more details see.

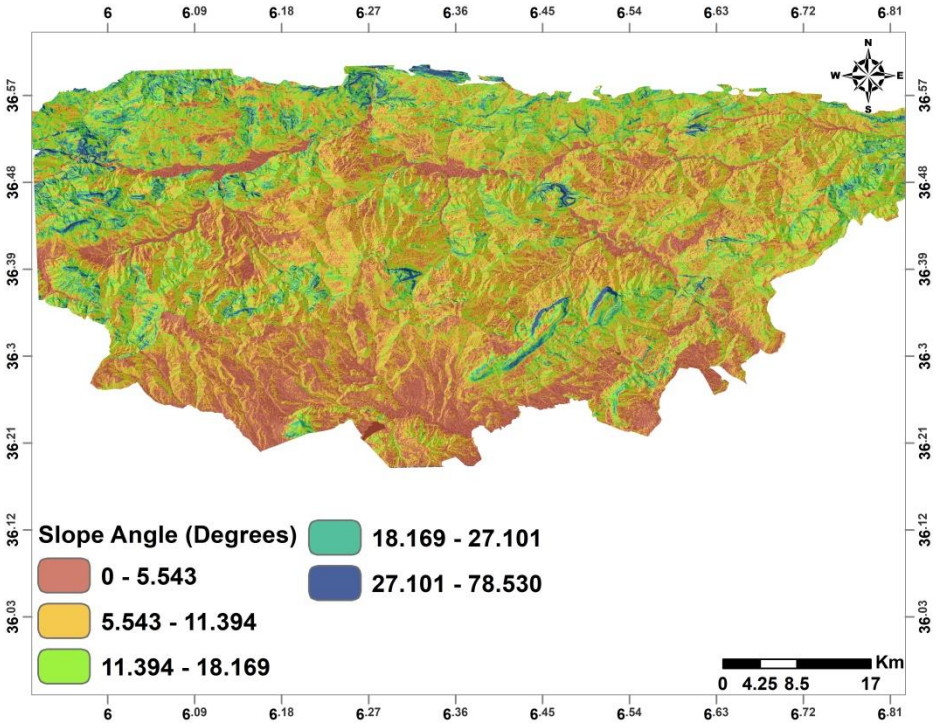
The geomorphometric parameters used in this case study are listed as follows (Figure 4.14a-d):

- Altitude – a float raster (Figure 4.14a), suggesting that the linear increase in potential energy with altitude is associated with higher susceptibility to landslides in the higher elevated grounds. It actually, represents the DEM of the terrain, described earlier.
- Slope angle (Slopes) – a float raster (Figure 4.14b), is highly important due to the fact that slope stability is directly related to landslide phenomenology (i.e. direct physical relationship). If the slope angle is (θ), then the greater (θ), the higher the possibility of slopes instability and vice-versa. However, that depends largely on lithology, rock type and resistance of the lithological units (e.g. in solid rock, the slopes are expected to be stable even with a steep slope angle, while clayey slopes do not need a steep angle to host instability), but in general, steep weathered formation slopes are highly susceptible toward landsliding. From the morphometric point of view, slope angles is considered as DEM first-derivatives and can be computed directly from the DEM by Degree Polynomial (DP) slope algorithms (also called D8 algorithm), referring to a smoothing window of size 9 (3x3 filter). Subsequently, the value of each pixel is defined by the mean of the surrounding 8 pixels [146].
- Slope aspects (Aspects) – a float raster (Figure 4.14c), which refers to the spatial exposure of the ground elements (e.g. azimuth) by controlling the micro-climatic parameters such as exposure to sunlight, wind, rainfall intensity, and the slope material properties. Slope aspects are directly computed from DEM by DP-D8 algorithms in a counter-clockwise fashion (ranges from 0° to 360°), suggesting that susceptibility to landslides accentuate from SW to NW quadrant, since the diurnal solar path influences moisture in slopes and topsoil mantle thickness. Thus, NW slopes are the most inconvenient (with the highest moisture content and the thickest mantle detritus), while SW is the least susceptible.
- Landforms – a float raster (Figure 4.14d), derives a classification for the landscape based on three-part geometric signatures (i.e., slopes, convexity

and surface texture) as the most common form of landslide progression on the slope, as suggested by [147].

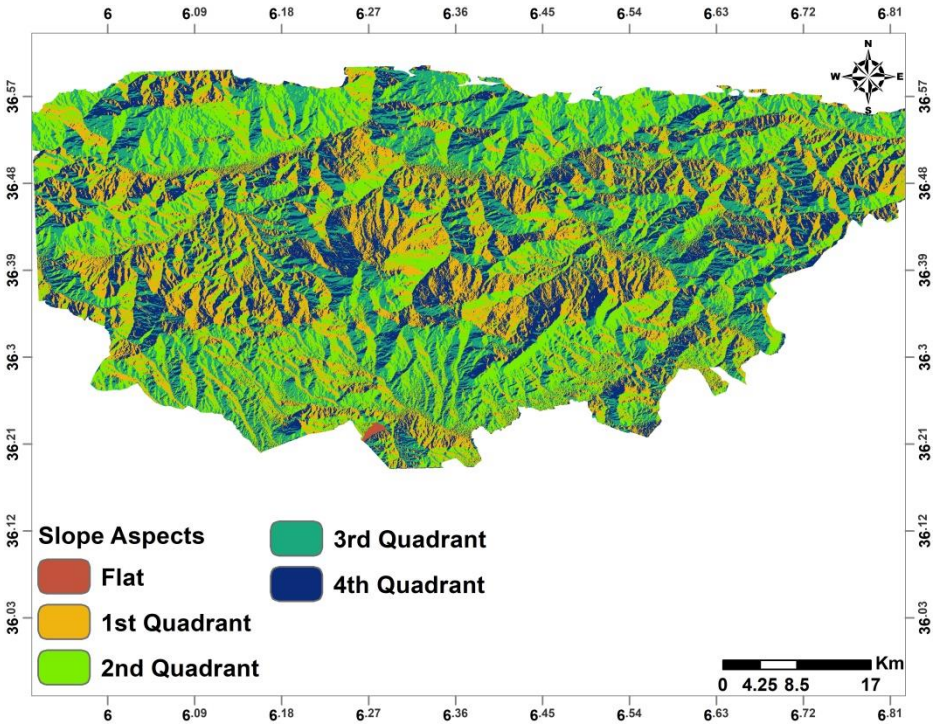


(a) Altitude

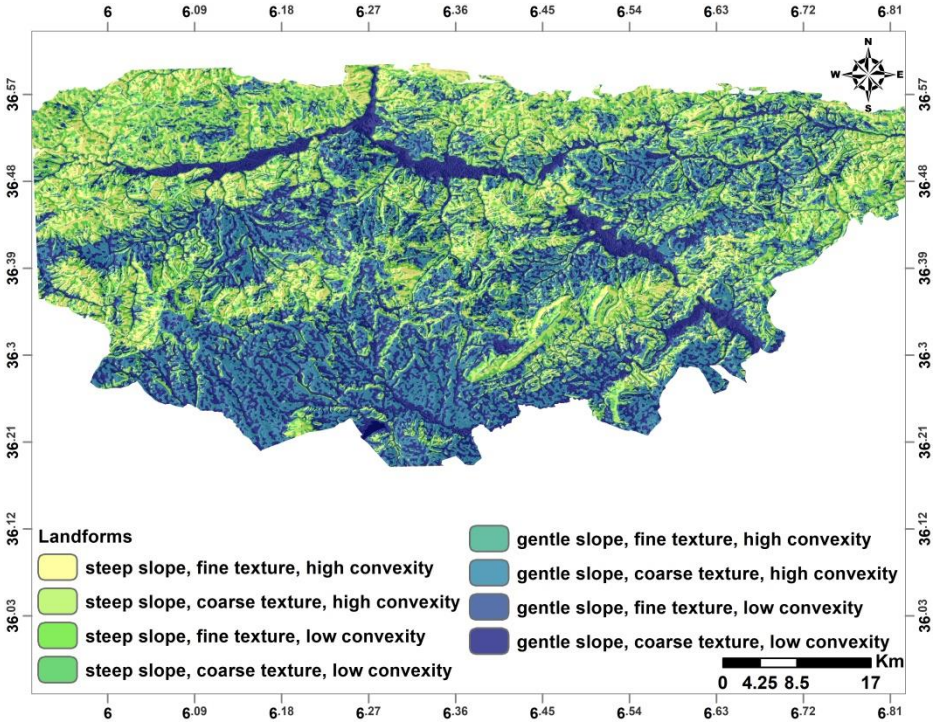


(b) Slope Angle

Figure 4.14 Geo-morphometric conditioning factors



(c) Slope Aspects



(d) Landforms

Figure 4.14 (continued).

Hydrological Data

Water, in general, is undoubtedly playing a primordial role in the triggering process of landslides and decrease slopes stability by effects such as:

- Increasing or decreasing the shear strength, cohesion, permeability and the overall mass of the slope.
- Weathering of slopes materials,
- Eroding of slopes footing,
- Saturating slopes.

These effects, influence slopes stability balance and can either increase or decrease landslides depending on water presence. Thus, the following parameters were used to express the hydrological effect on the landslide susceptibility in the study area:

- Rainfall – is a float raster (Figure 4.15a), that was generated from the Annual Mean of Precipitation (AMP) for the period of 1985 to 2017, using the Inverse Euclidean Distance Weighted (IDW) method. Rainfall is one of the triggering factors for landslides and considered and one of the most dangerous factors on slopes stabilities (especially upslope), because slides often occur following persistent periods of intense to moderate rainfall, where a high amount of water runoff water infiltrate and saturate formations and soils located beneath slopes (especially steep slopes, and can also introduce groundwater levels fluctuations depending of the amount of water infiltrated). On the other hand, short and intense to moderate rainstorms and precipitations, less amount of water infiltrated deep underground but affects slope stability indirectly by an intensive erosive process generated by the high amount of dissipated water on the surface ground.
- Topographic Wetness Index (TWI) – is float raster (Figure 4.15b), and represents a morphometric parameter[69] that pinpointing the effect of local topography on certain locations and the size of the saturated source area of

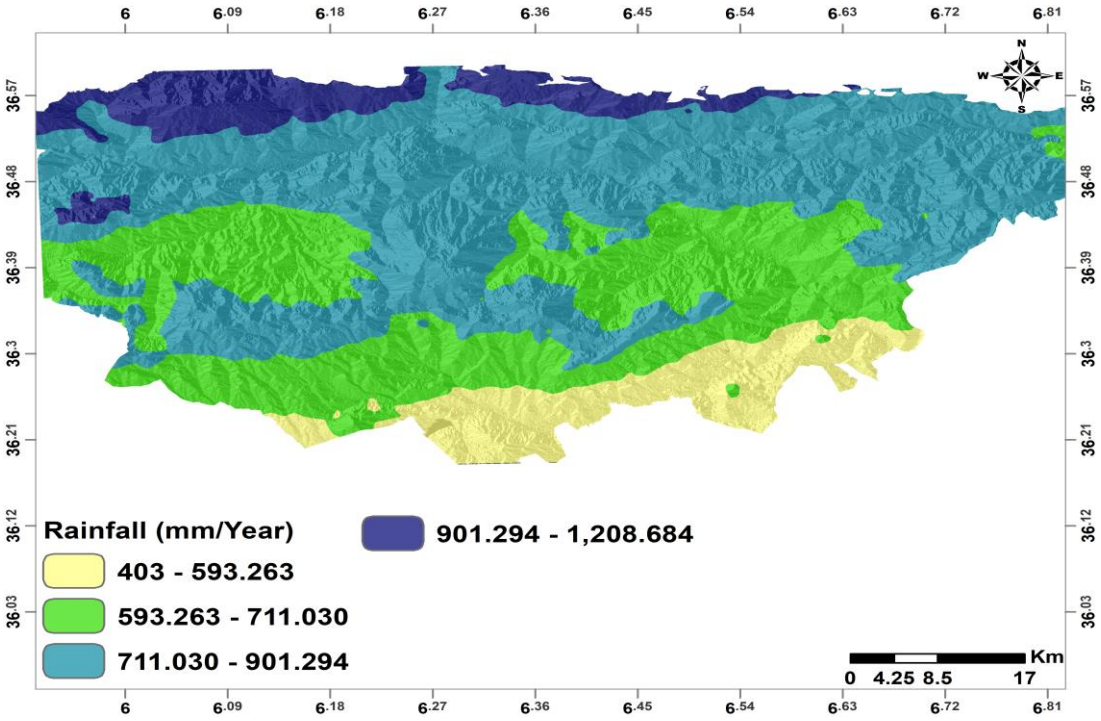
[69] From morphometric point of view, TWI is DP/D8 second order DEM-based derivative.

run-off generation by defining terrain retention (i.e. moisture distribution), which is correlated with the hydrogeological conditions, that influence surface run-off and infiltration [148]. Therefore, by expressing water retention distribution throughout the study area, TWI influence slope stability by fact that effective stress decrease in saturated slopes, and thus areas with higher TWI values as relatively more prone to instabilities. According to Beven and Kirkby [149] and Moore, Grayson [150], TWI can be calculated by the topographic water retention potential given by a relation of the upslope drainage area and slope gradient, using Equation (4.1):

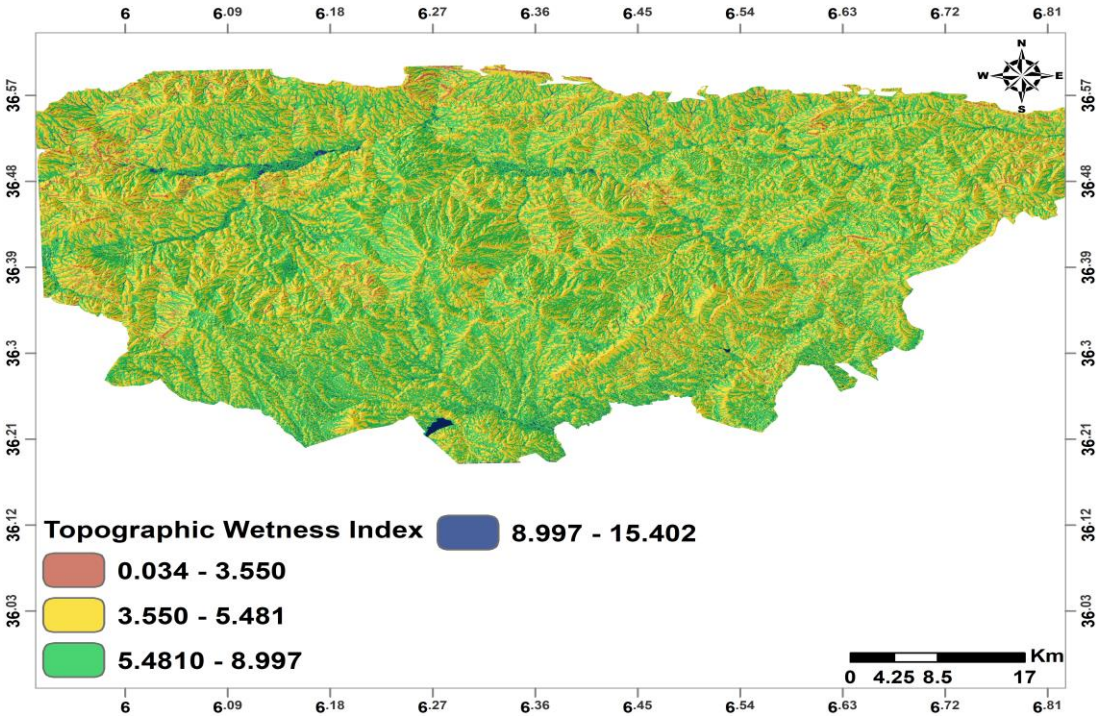
$$TWI = Ln(\frac{A_s}{tan\beta}) \quad (4.1)$$

- Where: A_s is the specific catchment area (m^2/m) and β is the local slope in degree (°).
- Distance to Hydrographic Network (WDist) – is a float-buffer raster (Figure 4.15c), which introduces the influence of linear erosion on the slope stability, since deformation and failure processes develop regressively upslope under the vertical and lateral influence of the linear erosion. In narrow upper sections of the valleys, vertical erosion dominates, steepening the slopes and destabilizing rock masses. On the other hand, lower sections tend to develop lateral erosion, widening the valley bed, once again pushing slopes off the balance. The foregoing discussion suggests that areas closer to the streamlines are more affected than remote ones, thus buffering out the landslide susceptibility toward the ridges of local watersheds. Distance to the hydrographic network was computed from vectorized hydrographic streams network using IDW in SagaGIS.

Precipitation data and hydrographic networks were provided by ANRH (L'Agence Nationale des Ressources Hydrauliques) and ONM (Office National de Meteo).

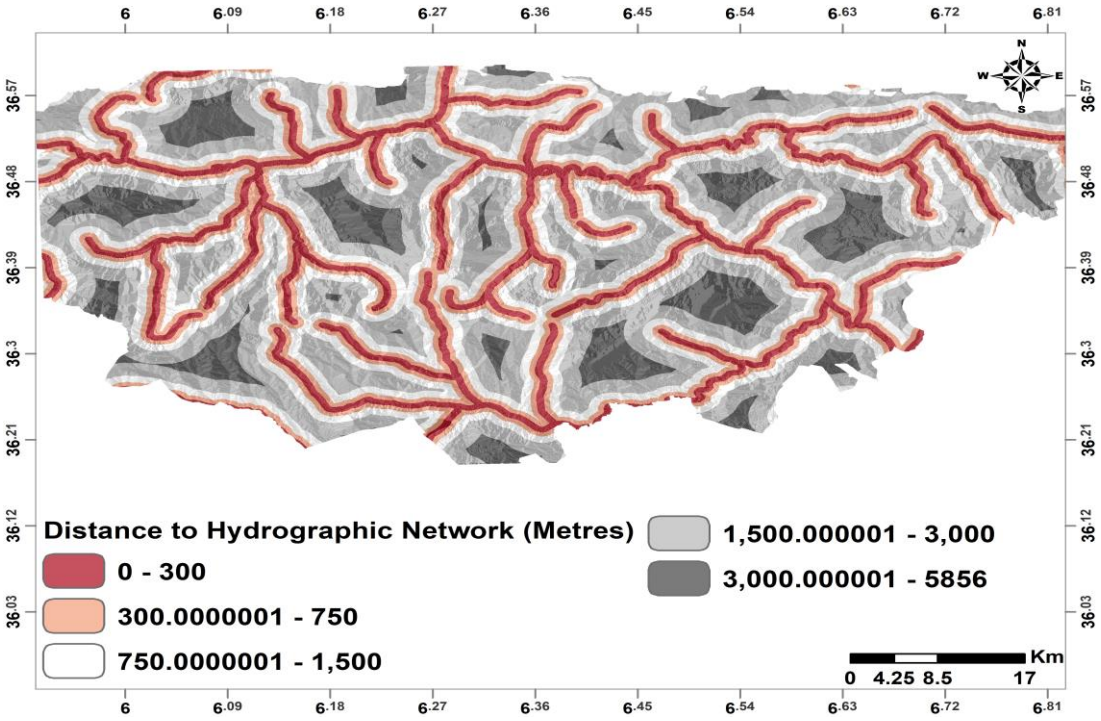


(a) Rainfall



(b) Topographic Wetness Index.

Figure 4.15 Hydrological conditioning factors.



(c) Distance to Hydrographic Network

Figure 4.15 (continued).

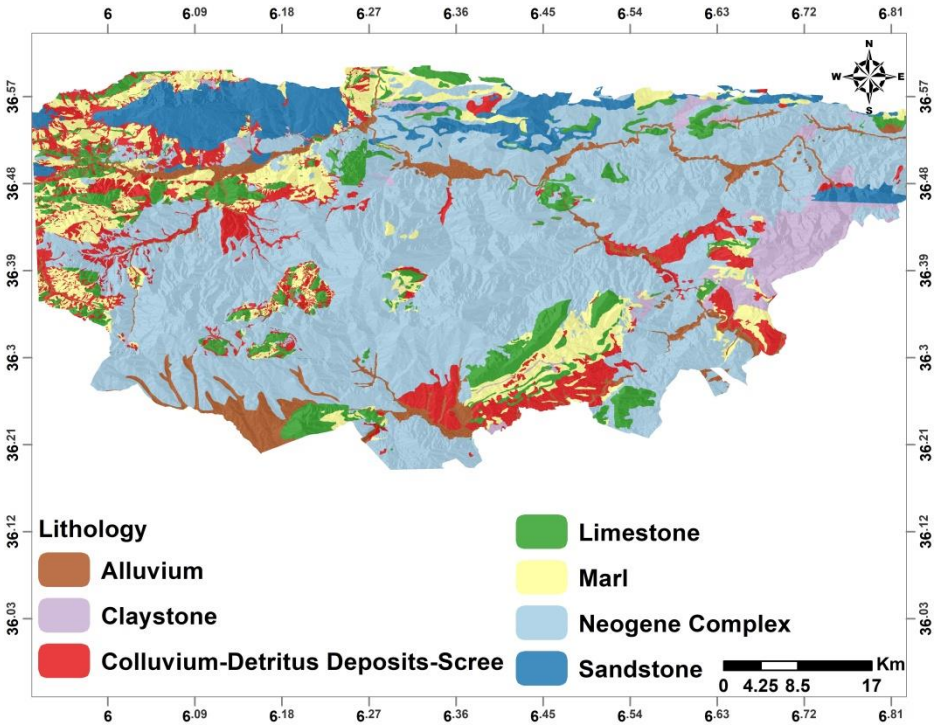
Geological Data

Geological data for the study area were compiled from a total of seven hard-copy maps covering the study area and each is a 1:50000 scale provided by ASGA (L'Agence du Service Géologique de l'Algérie). These maps were further simplified to meet the requirements of this case study[70]. Therefore, the generalization to a raster map with a 30 m resolution was justifiable. The map was also used to derive the synthetic data such as the Euclidean distance buffer to geological structures (Figure 4.16a-c).

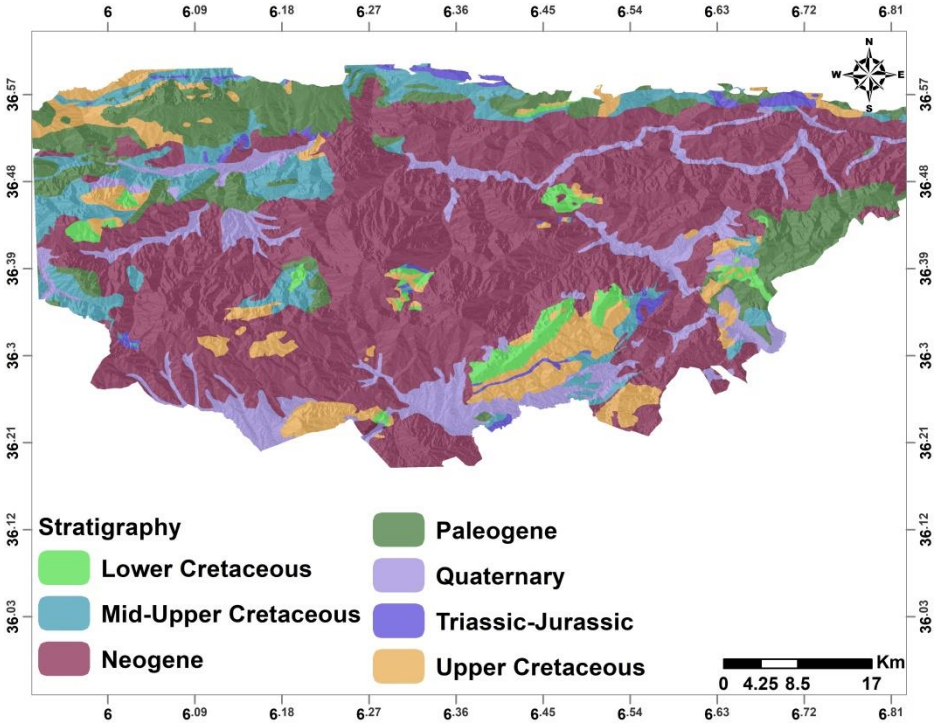
- Lithology – a discrete (i.e. categorical) raster (Figure 4.16a), that represents the outcropping lithology derived after the geological maps, as mentioned above. The map depicts 7 lithological units namely, Alluvium, Claystone, Colluvium-Detritus Deposits-Scree, Limestone, Marl, Neogene Complex, and Sandstone that are different in their physical and mechanical behaviors. Thus, differently prone to instabilities,

[70] In order to enhance the statistical representativeness.

- Stratigraphy – a discrete raster (Figure 4.16b), that represent the stratigraphy of the outcropping lithology derived after geological map, as mentioned above. The map depicts 7 chronostratigraphic units namely, Quaternary, Neogene, Paleogene, Upper Cretaceous, Upper-Mid Cretaceous, and Lower Cretaceous and Triassic-Jurassic.
- Distance to Faults (FDist) – is a float-buffer raster (Figure 4.16c), and represent the distance from the available geological structures such as faults and joints in the geological map using IDW in SagaGIS. Since faults and joints, were considered as zones of weakest shear resistance (limited only to a residual shear resistance) and also affected by the infiltrated water and fill material, it is logical to assume that instabilities are more prone in the areas closer to these structures. In more seismically active areas such parameters could be much more appreciated since the shear resistance faces further effects, related to the fault dynamics.

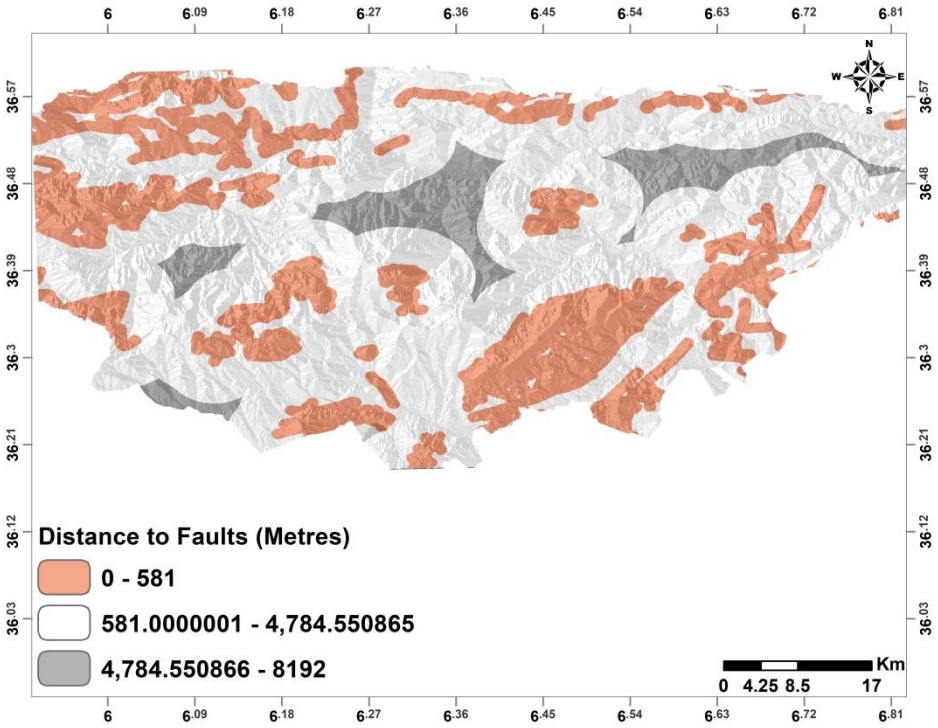


(a) Lithology



(b) Stratigraphy

Figure 4.16 Geological conditioning factors.



(c) Distance to Faults

Figure 4.16 (continued).

Geotechnical Data

Geotechnical data were directly from obtained Mila and Constantine local agencies, i.e. municipalities, at 30 meters resolution. This data vary in term of the overall influence on landslide occurrences, but certainly, introduce relative interpretation on the geotechnical context of the landslides distribution patterns at the study area.

- Soil Textures (Texture) – a discrete raster (Figure 4.17a), that represents the available soil units[71] available in the study area according to the relative proportion of sand, silt, and clay content. The soil textures types were assigned according to USDA54 classification[72] and six soil units that are different in their physical, mechanical and geotechnical behavior were obtained, i.e. Sandy Clay, Clay Loam, Silty Clay Loam, and Sandy Clay

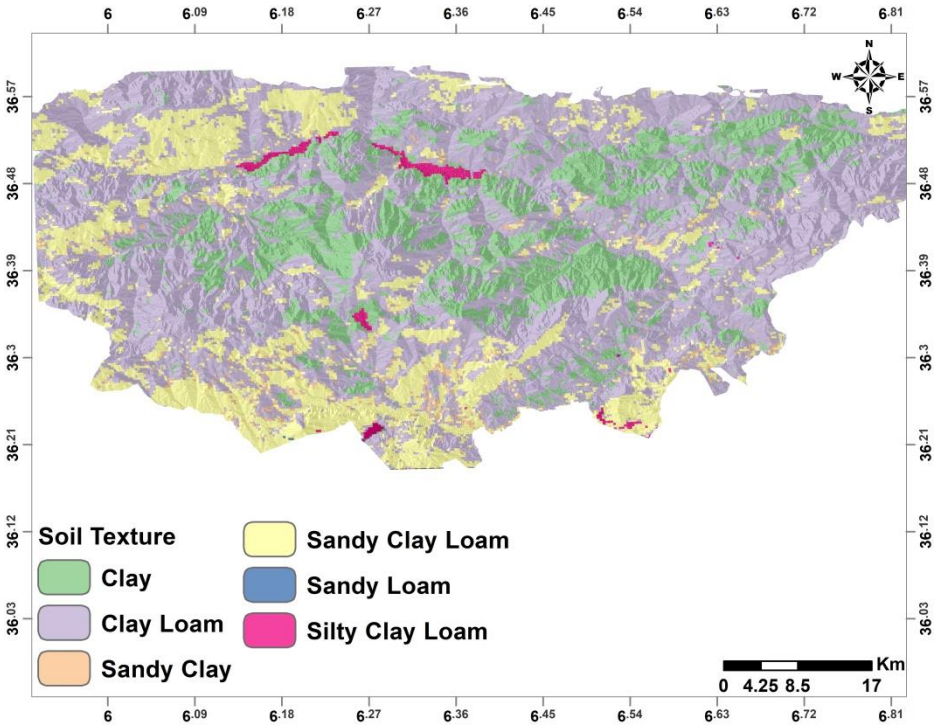
[71] From geotechnical engineering standpoint.

[72] Rely on the relative proportion of clay, sand, and silt. For more details check the following website.

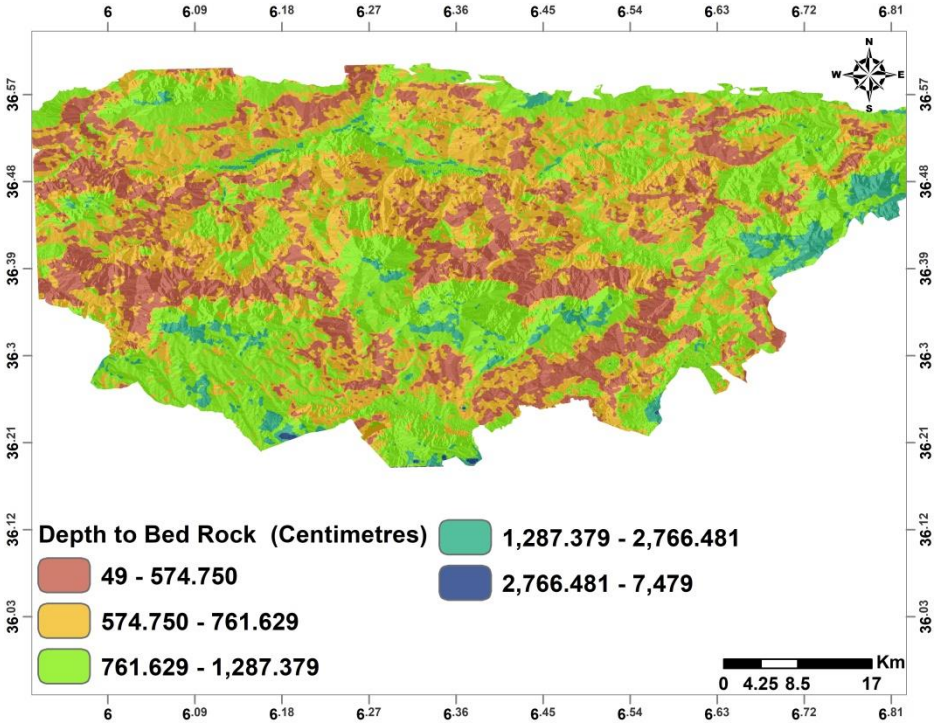
Loam. In general, soils with a high percentage of clay form very stable aggregates resistant to detachment, but so sensitive to water. On the other hand, lighter soils like sandy soils or coarse loams are easy to detach as they have low organic matter content, resulting in their inability to form very stable aggregates (Das and Agarwal 2002). Hence, soils with a high content of sand and clay, steeper slopes, and intensive rainfall, which constitutes the most dominant factors of the landslide, cause severe damage to the land (Patanakanog 2001).

- Depth to Bedrock (DepthBR) – is a float raster (Figure 4.17b), that forms one of most the important factors for assessing the stability of the soil and landslide susceptibility of the land. With the increase in soil depth to bedrock, the tendencies of the soil to absorb moisture also increase. Thus, reducing the runoff rate. Hence, shallow soil is considered to be more unstable and prone to landslide than the deep soil.
- Bulk Density (BDensity) – is float raster (Figure 4.17c), that in general have tight relationship with soil properties, such as soil textures and depth to bedrock, especially in areas that are affected by landslides where this parameter can help in understanding the interaction between the soil structure and the geotechnical behavior of the slopes[73].

[73] One example is saturated hydraulic conductivity, which depends on water leaching due to the macro porosity, which, in turn, is related to soil texture, particle arrangement (structure) and bulk density. The free passage of water is crucial to reduce runoff (Gomes et al., 2011).

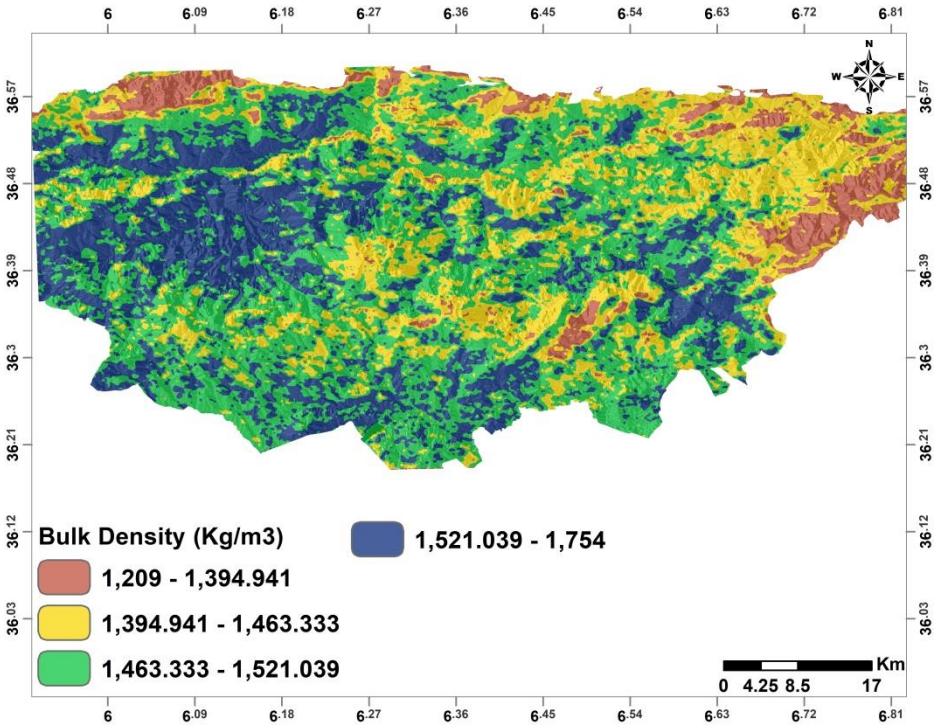


(a) Soil Texture



(b) Depth to Bedrock.

Figure 4.17 Geotechnical conditioning factors.



(c) Bulk Density

Figure 4.17 (continued).

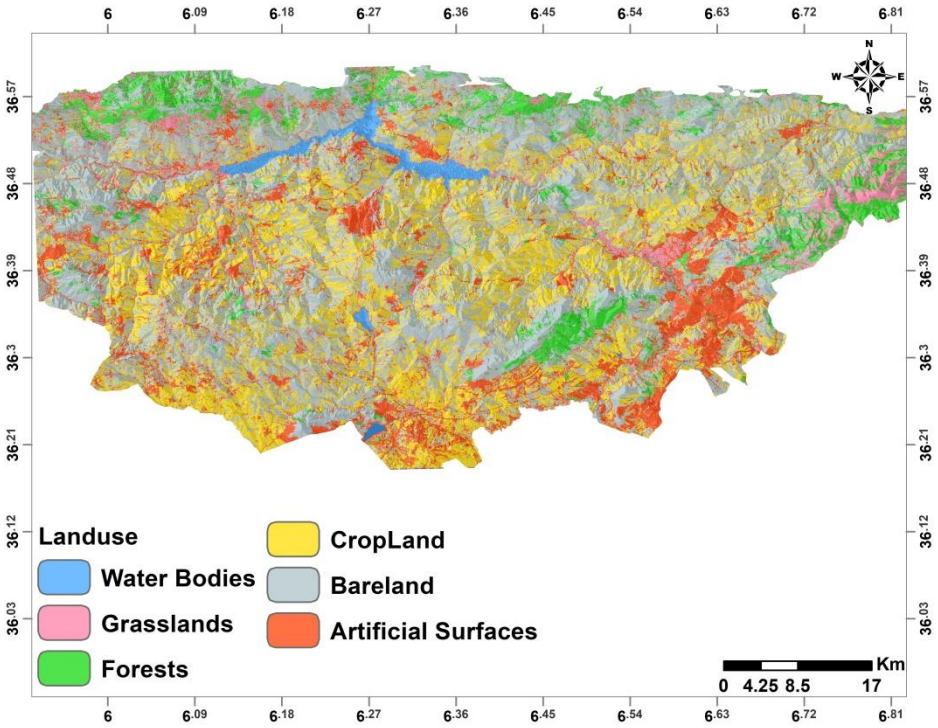
Environmental Data

Environmental information's particularly regarding parameters that may influence landslides slopes stabilities and landslide distribution. These informations, besides natural environment-oriented information, are human-related one way or another, and it's rarely incorporated in landslide susceptibility analysis, due to various reasons (e.g. the lack the data). The following parameters were used to express the environmental effect on landslide susceptibility in the study area:

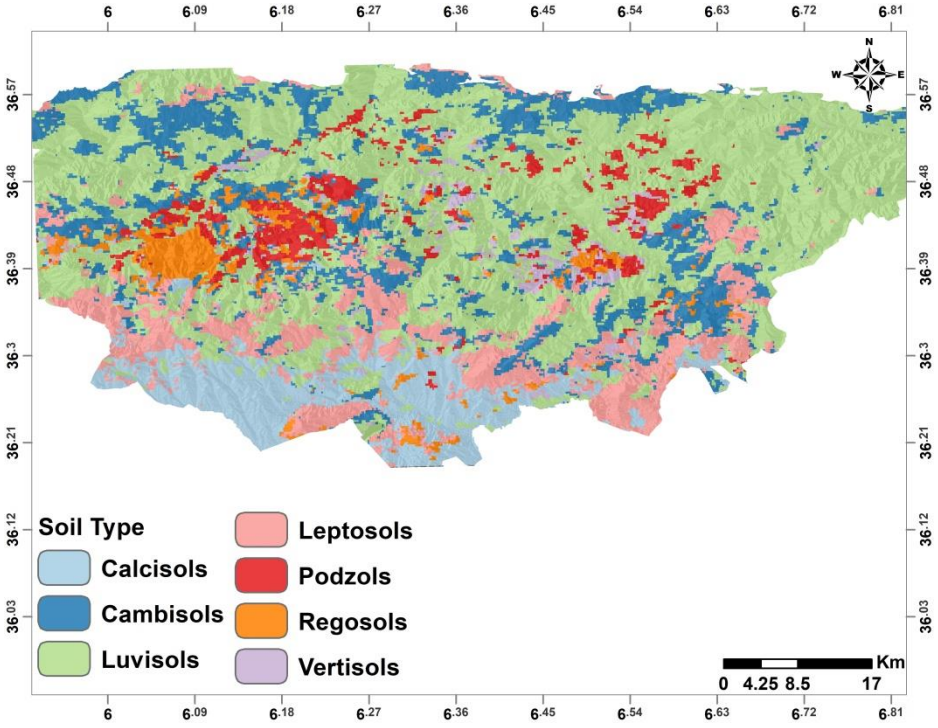
- Landuse (Landuse) – a discrete raster (Figure 4.18a), which considered one of the most influential parameters on landslides occurrence. Theoretically, barren land and shifting cultivation are more prone to landslides than other landuse units. It could happen because there is no deep root that can hold the soil. Contrarily, forest areas tend to decrease the landslide occurrences due to the natural anchorage provided by the tree roots. Landuse classes are categorized into Artificial Surfaces, Forests, Grasslands, CropLand and Bareland.

- Soil types (Soils) – a discrete raster (Figure 4.18b), represent the available soil units[74] available in the study area. The map depicts 6 soil units, i.e. Calcisols, Cambisols, Luvisols, Leptosols, Podzols, Regosols, Vertisols. These units are drastically different in their water retention and root cohesion behaviors. Thus differently prone to erosion, which directly and/or indirectly instabilities.
- Distance to Roads (RDist) – is a float-buffer raster (Figure 4.18c), and represent the distance from the roads network using IDW in SagaGIS. Human-induced factors may raise the probability of landslide occurrences. Cutting the toe of a steep slope and filling along the road are the common human activities on the hilly areas which increase the susceptible area to a landslide. It is convinced; when the many landslide events were nearby cutting road areas (e.g. RN 79a and RN27). Therefore, the best way to contain the effect of road factors in landslide study is by making a buffer on the upslope part.

[74] From agriculture-engineering standpoint.

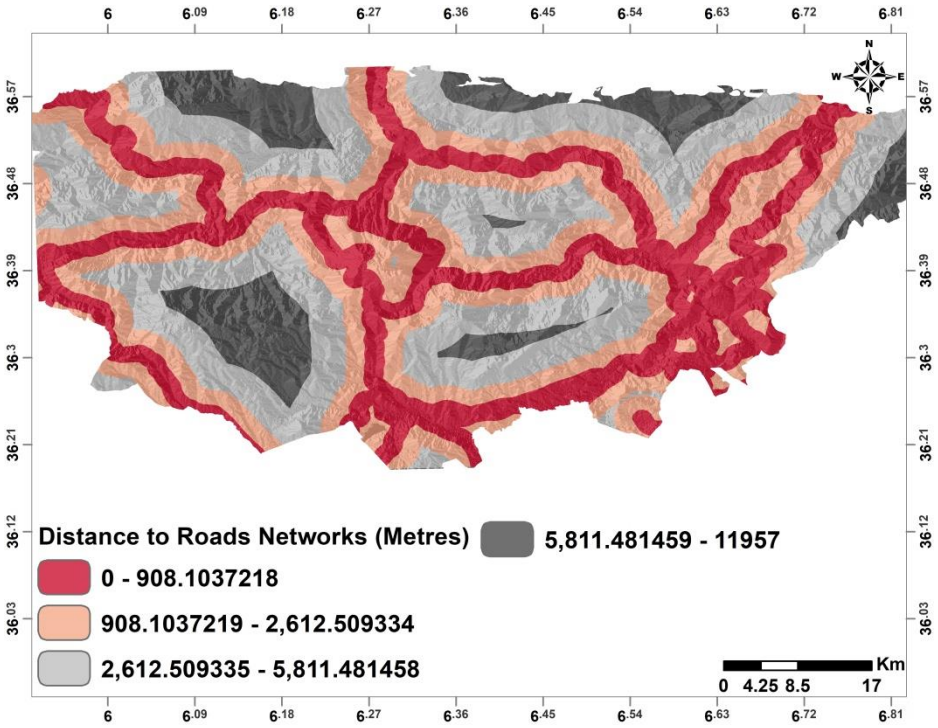


(a) Landuse



(b) Soil Type

Figure 4.18 Environmental conditioning factors.



(c) Distance to Roads

Figure 4.18 (continued).

4.4 DATA SUMMARY

Detailed descriptions of all used factors in the geospatial database, their class's breaks, and categories are given in Appendix B. A general statistics for the conditioning factors were expressed in terms of frequency statistics using overall terrain data such as the percentage of the overall area extent covered by each class and landslide percentage in each distinct class. It is common practice, for ranging (to reclassify it into different intervals) numeric conditioning factors and quantifying nominal conditioning factors. In order to avoid subjective quantification, the reclassification process (the class intervals and the total number of classes) of the continuous factors (altitude, slopes, rainfall, and so forth) into different intervals, was performed automatically using the Geometrical Intervals[75] reclassification method due to the non-uniform distribution of the data in these factors. On the other hand, the categorical or nominal factors (i.e. Lithology, Stratigraphy, and so forth) remain

[75] This method, rely on data distribution, standard deviation and the supported mean value in order to determine class breaks.

intact due to the fact that classes are unique and predefined by the nature of the factor itself. Thus, they do not require any reclassification.

Chapter 5: Results and Discussions

In this Chapter, a special focus was given to reporting and discussing the main results obtained from the analysis of the implemented landslide susceptibility models (Chapter 3.2.2). The following Sections contain results and outcomes of the landslide susceptibility analysis modeling according to the implemented research workflow (Chapter 3.6). The suitability of each given model and the generated landslide susceptibility maps are also discussed.

5.1 RESULTS

Herein, this section will focus on reporting the results of all of the proposed models (See Chapter 3.2.2) in the same order following of subsection reported in the workflow (See Chapter 3.6).

5.1.1 Analyzing and Optimizing Landslide Conditioning Factors

In a comparative study, constructing the necessary conditioning factors does not necessarily imply that it is suitable for use as an input dataset for models. In fact, it is crucial to check the integrity of the input dataset by performing some sort of analysis (i.e., Pearson's correlation coefficient analysis, and multicollinearity detection) before conducting the modeling, mainly to ensure:

- The non-independence among conditioning factors to the landslide inventory.
- Determine the suitability of the underlying assumption behind choosing the factors.

In this research, PCC and VIF analyses were performed against 16 conditioning factors by taking into account the aforementioned criteria. Values in the PCC correlogram (Figure 5.1) are lower than the critical threshold of (0.7). The highest PCC recorded was between TWI and the Slope angles pair at 0.54. In fact, a high correlation is expected between the generated variables and the source variables (i.e., TWI, Slopes, and Altitude that were derived from the DEM). On the other hand, the VIF results (Figure 5.2), show that all factors should be used since the highest value is less than the theoretical critical value of 5 [35-38].

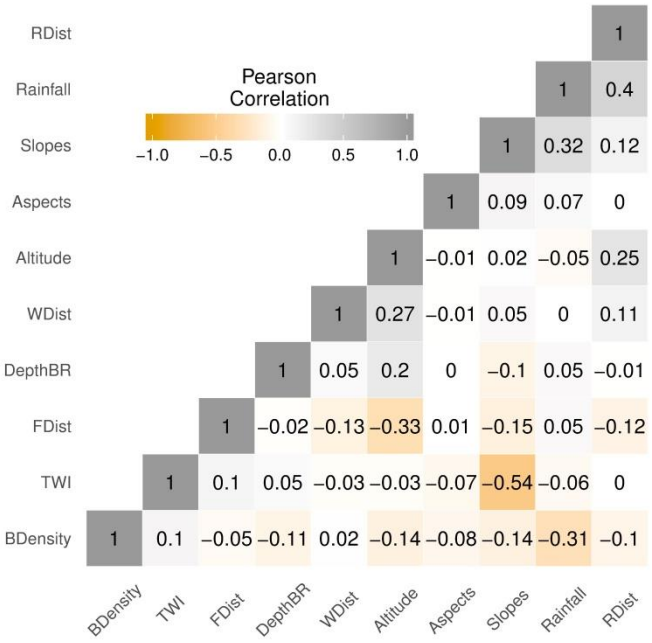


Figure 5.1 Correlogram based on Pearson correlation matrix of numerical conditioning factors.

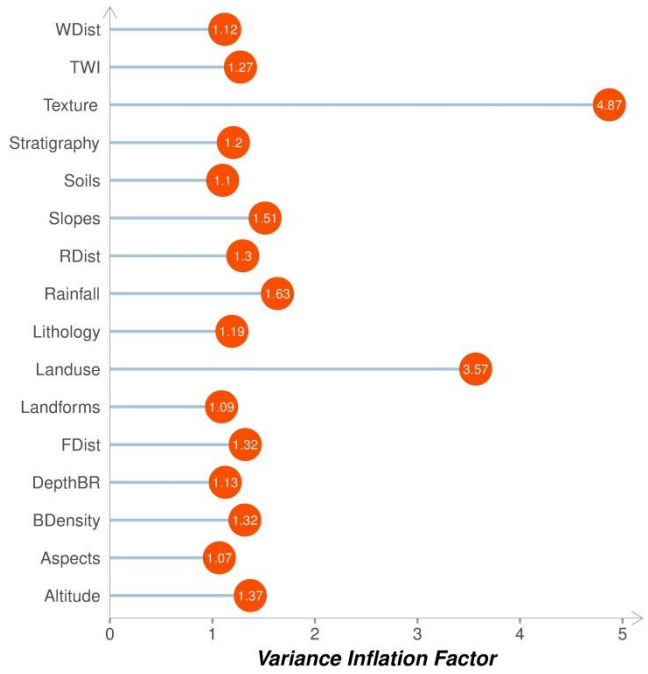


Figure 5.2 Variance inflation factor analysis results in landslide conditioning factors.

5.1.2 Model Training

During tuning, hyperparameters need to be carefully optimized, so as much accuracy models can be achieved, the model selection will be reliable. In general, the tuning process must be a formal and quantified part of model evaluation. Yet, in

most cases, personal experience and intuition heavily intervene by influencing the process in ways that are hard to quantify or describe [87]. For that reason, there has been no intervention in the optimization process, as further model optimization was limited to SMBO, which is one of the best techniques for optimizing numerous hyperparameters at once, for the purpose of objectifying the simulation as much as possible.

In the midst of the training process, the optimal hyperparameters are carefully optimized by SMBO for all models[76] according to Table 5.1 using the following procedures:

- Set a single objective function for each learner using "smoof" [151] with AUC to maximize it as a single performance criterion.
- Use "lhs" package [152] to set an initial design grid that covers the supplied search space of each model parameter by drawing a Latin Hypercube Sample Design (LHS) using a Column wise Pairwise (CP) algorithm to generate an optimal design with respect to the S optimality criterion [153].
- During every single iteration, a new point is being proposed through LCB infill optimization of the estimated standard error. This error is usually obtained by a surrogate model that is either kriging-based for a purely numeric space or random forest for a mixed search space.
- Select and return the optimum values of the desired hyperparameters based on the highest AUC (Table 5.1).

Table 5.1 The optimum parameters obtained by the tuning process.

Model	***Hyperparameter***	***Optimal Value***
GBM	Shrinkage	0.020
	n.trees	570
	interaction.depth	8
NNET	Size	29
	Decay	0.809

[76] Of all the implemented models in this case study, i.e. GBM, LR, NNET, RF and SVM, only LR is straightforward and doesn't require tuning.

RF	Replace	FALSE
	sample.fraction	0.953
	num.trees	1012
	mtry	5
SVM	kernel	radial
	cost	28.382
	gamma	2−8.398
	degree	N/A

According to the tuning results reported in Table 5.1, it turned out these values were the optimal hyperparameters (parameter of choice) for the final susceptibility model. In fact, these parameters achieved the best and highest performance (in terms of AUC) for each respective model. For example, GBM achieved the best performance by (0.02), (570), and (8) as Shrinkage, n.trees and interaction.depth, respectively. While, NNET achieved optimal AUC with (29) and (0.809) as Size and Decay, respectively. On the other hand, (FALSE), (0.953), (1012) and (5) as Replace, sample.fraction, num.trees, and mtry, respectively, achieved the best AUC possible for RF, but ($2^{8.382}$) and ($2^{-8.398}$) as cost and gamma concerning SVM.

5.1.3 Model Evaluation and Comparison

Given the optimal hyperparameters sets (see Table 3.2 and Table 5.1), were used to train each respective model. Afterward, the predictive performance capabilities and the quality of the resulting models were evaluated using the input dataset based on performance indicator metrics like AUC, ACC, and the Kappa index.

The Overall performance results (Figure 5.3 and Table 5.2), show that all the models have a "substantial agreement" between the observed and the predicted landslides expressed in term of a kappa index ranging between 0.5605 and 0.6405. The AUC and ACC values range from 0.8575 to 0.8967, and 0.7803 to 0.8203, respectively, indicating that all the models have "very good" predictive capabilities. In particular, the ensembles models that benefit from a divide-and-conquer approach such as RF and GBM yielded significantly better results than traditional methods like NNET, SVM, and LR. In fact, GBM was the highest-ranked model in terms of the performance of the AUC, ACC, and Kappa index with values of 0.8967, 0.8203, and

0.6405, respectively (Table 5.2). RF held the second-highest ranked model with performances similar to GBM with values of 0.8957, 0.8178, and 0.6356 for AUC, ACC, and kappa, respectively. NNET, on the other hand, was able to achieve the highest performance after the ensemble tree models, followed up by SVM. In contrast, the LR performance was consistently lower than the rest of the models in every metric, with values of 0.8575, 0.7803, and 0.5605 for AUC, ACC, and kappa, respectively.

Table 5.2 The overall performances of the trained landslide models.

Metrics	***Model***				
	GBM	*LR*	*NNET*	*RF*	*SVM*
Acc	0.820	0.780	0.809	0.817	0.802
Kappa Index	0.640	0.560	0.619	0.635	0.605

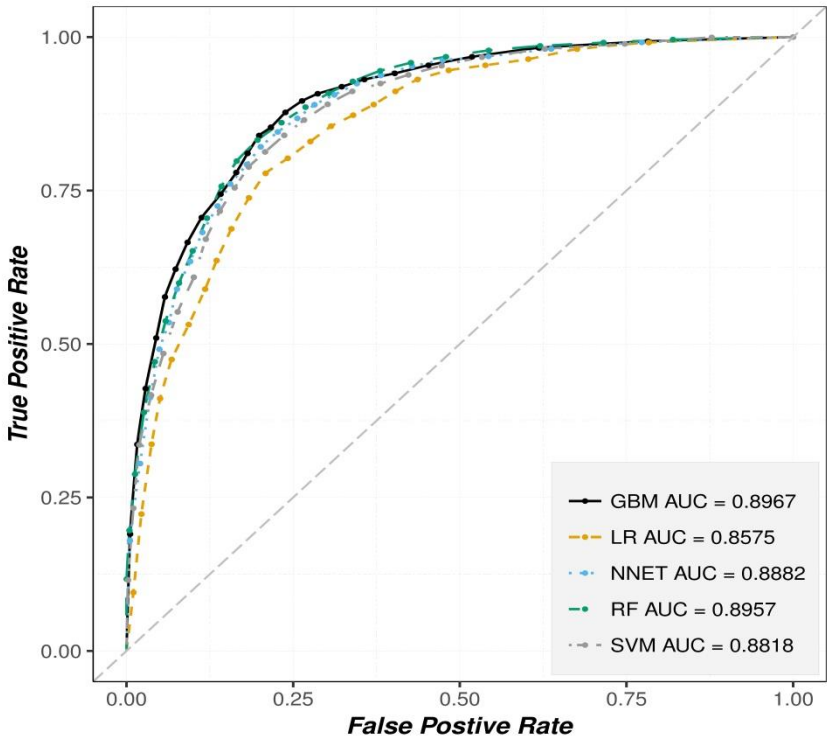


Figure 5.3 The stacked receiver operating characteristic (ROC) curves of the implemented models.

Despite the brief quantitative report concerning the overall performance results, presented in Table 5.2 and Figure 5.3, some additional pair-wise statistics have been calculated for better understanding the differences in the performance capabilities of each model against each other. In fact, in order to determine if the differences in performance between the five landslides susceptibility models are statistically

significant, a systematic pairwise comparison using the Wilcoxon signed-rank test at the 5% significance level was conducted (Table 5.3).

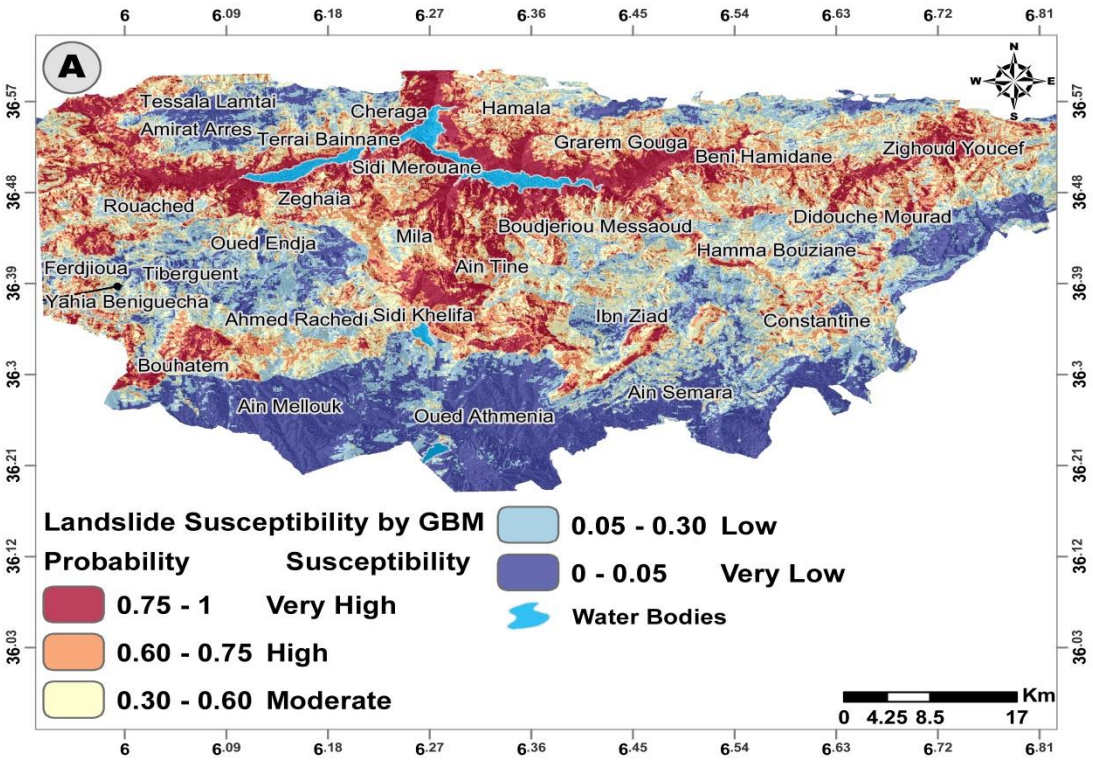
The results show that there is a systematic difference in the performance results between each pair of models except for the GBM and RF pair, where the difference in performance was found to be statistically insignificant (that is, $p.value \geq 0.05$ and $-1.96 \leq z.value \leq +1.96$, so, the null hypothesis was accepted). Overall, this plausible result rather goes in its favor of both GBM and RF as the best models for the data at hand in this study.

Table 5.3. The pairwise comparison of the five landslide susceptibility models using the Wilcoxon signed-rank test.

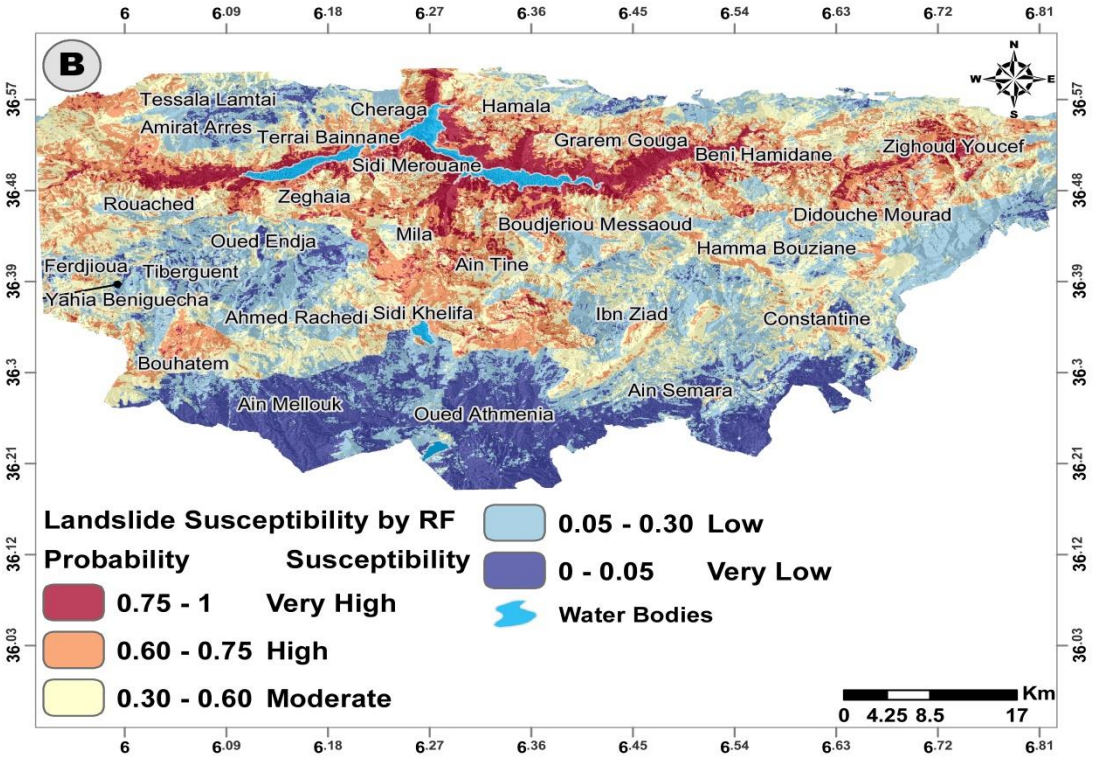
No.	*Pairwise comparison*	$z.value$	$p.value$	*Significance*
1	GBM vs. RF	-0.579	0.562	No
2	GBM vs. LR	6.111	0.000	Yes
3	GBM vs. NNET	3.606	0.001	Yes
4	GBM vs. SVM	5.266	0.000	Yes
5	RF vs. LR	6.149	0.000	Yes
6	RF vs. NNET	2.905	0.004	Yes
7	RF vs. SVM	4.025	0.000	Yes
8	SVM vs.LR	5.589	0.000	Yes
9	SVM vs. NNET	-3.223	0.001	Yes
10	NNET vs. LR	5.995	0.000	Yes

5.1.4 Generating Landslide Susceptibility Map

Once the final models were evaluated and validated, the tuned models were used to successfully predict and generate landslide occurrence in the study area in the form of probability grids ranging from 0 to 1, then they were reclassified into five susceptibility classes (Table 3.5). The implemented models successfully generated susceptibility maps that can be acknowledged as plausible as they overall produce a fine and smooth prediction surfaces that correspond very apparent with the spatial trends of the actual landslides that indicate that the dispersion is very limited (Figure 5.4). Thus, indicate that no potential post-processing (e.g. majority filtering, neighboring smoothing...etc.) is required.

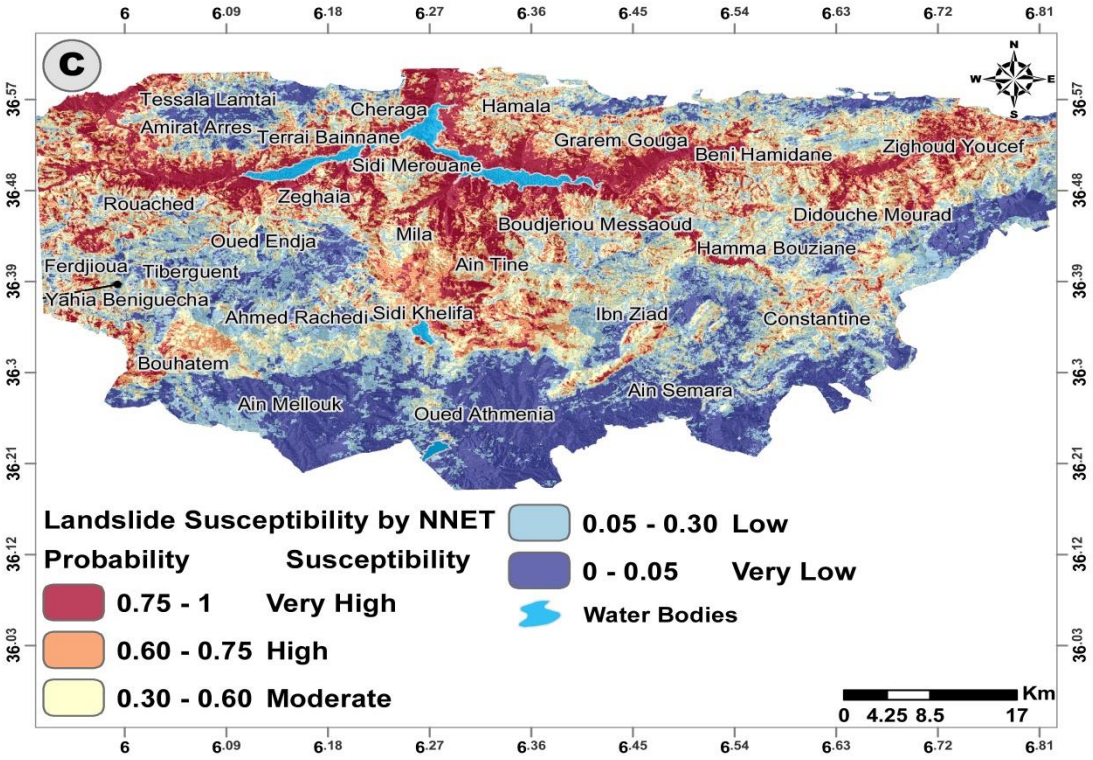


(a) GBM

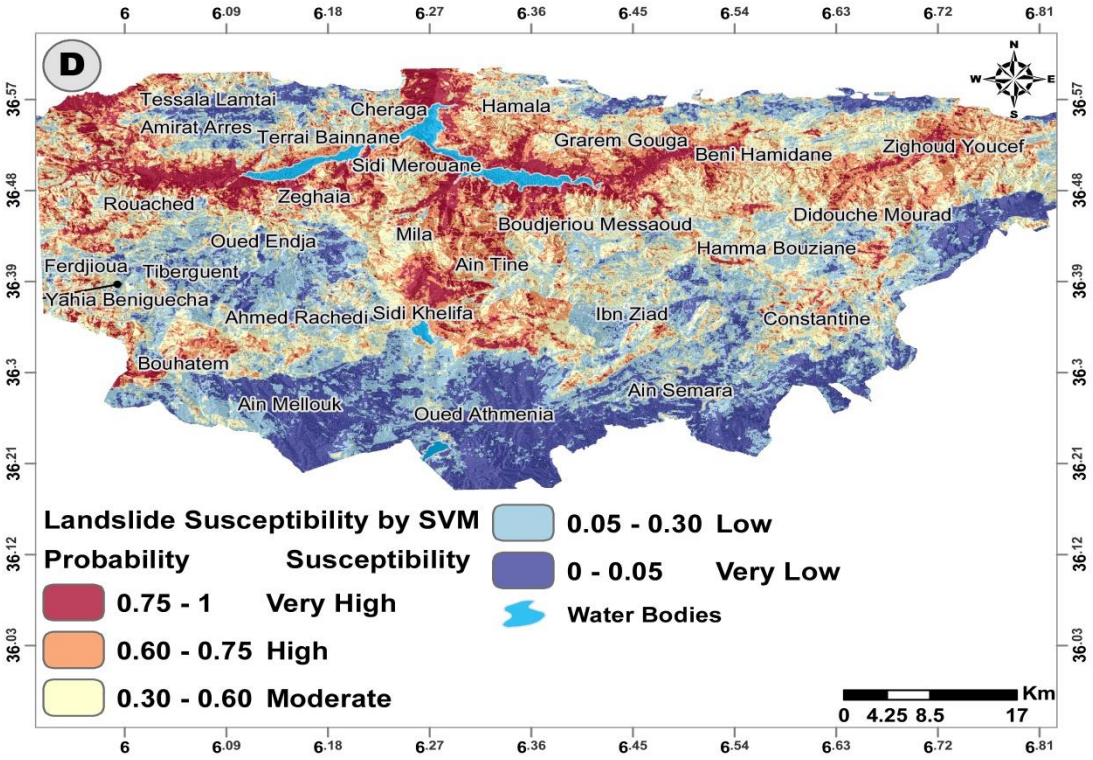


(b) RF

Figure 5.4 The generated landslide susceptibility maps.

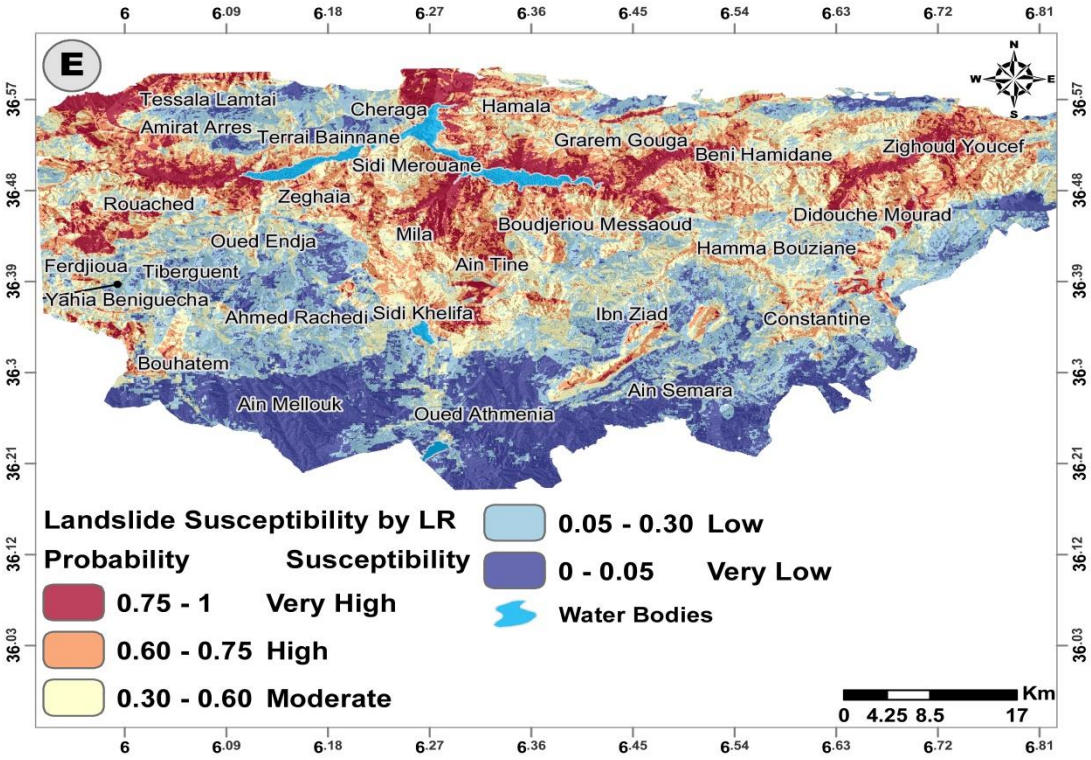


(c) NNET



(d) SVM

Figure 5.4 (continued).



(e) LR

Figure 5.4 (continued).

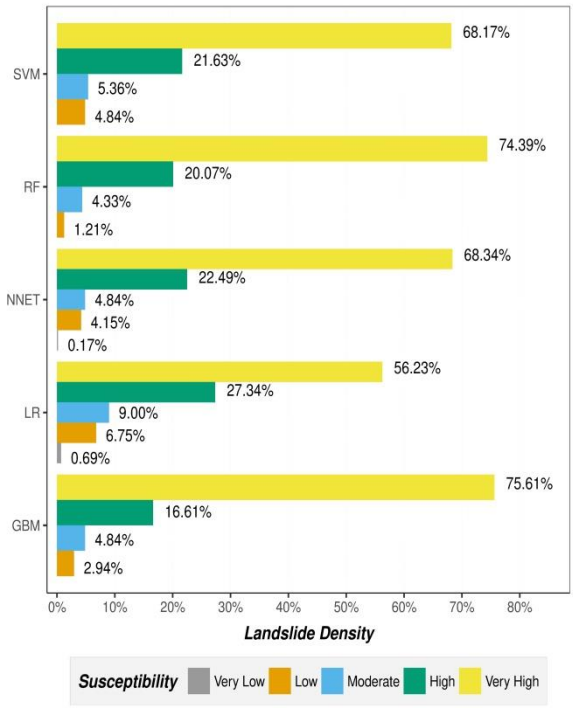
In the case of a landslide susceptibility assessment, the models usually evaluated by probabilistic performance metrics such as AUC, ACC, and Kappa index, but this actually is not enough. Models with close or even similar performance results (for example, GBM and RF have no statistical significance in the performance difference in this case study) and they do not necessarily generate similar predictive output surfaces. The spatial predictive output surface is critical for assessing the quality of landslide susceptibility models. Overall, by performing a sufficiency analysis on the predictive output surface in the form of summary statistics (that is, landslide density distribution and the area extent covered by each susceptibility class), it is possible to gain an insight into the model's quality by:

- The spatial predictive output surface details
- The results of the landslide distribution analysis.

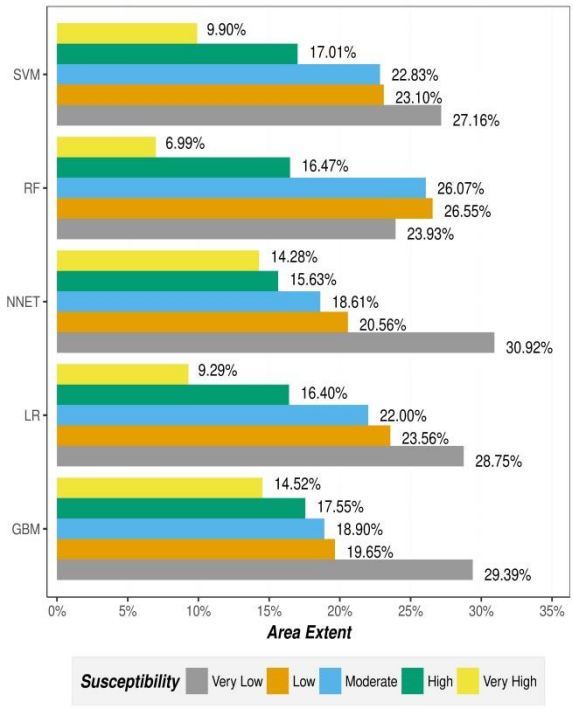
In fact, Once the final models are evaluated and benchmarked against the outer testing instance, the next phase is to use them successfully to predict the study area,

this step resulted in a positive class (landslide occurrence) probability grids, which reclassified according to Table 3.5 into five susceptibility classes toward landsliding (Figure 5.4). Then, by overlapping the landslide inventory and the reclassified susceptibility maps (Figure 5.5), a sufficiency analysis summary statistic was obtained in the form of a landslide density distribution (Figure 5.5a) and the total area extent covered by each susceptibility class (Figure 5.5b). The results are satisfying because they fulfill two spatial conditions: (1) the landslide pixels should be located at the very high and high susceptible classes and (2) the extent of the areas covered by the very high and high susceptible classes should be as small as possible. All the reclassified models show an increase in the landslide density ratio distribution in positive trend when moving from low to high susceptible classes, with GBM scoring the best results[77] of approximately 75.61% and 14.52% for landslide density occurrences and the area extent covered by the highest susceptibility class (that is, "very high"). RF scored 74.39% and 6.99% followed by NNET with 68.34% and 14.28%, SVM with 68.17% and 9.90%, and LR with 56.23% and 9.29%.

[77] For Landslide susceptibility assessment only "Very High" class has been regarded as the base for assessing the models.



(a) Landslide density distribution by susceptibility zones



(b) The total area extent covered by susceptibility zones

Figure 5.5 The sufficiency analysis of the susceptibility maps.

A positive indicator of the classification capability of the generated models is that they do not show any landslide events in the "Very Low" susceptibility class (that is, if the landslide density is null, then the class is absent) or they only show a very small percentage (<0.70% of the total landslide events) (Figure 5.5a). In general, the "Very Low" and "Low" susceptibility classes are grouped pixels with low probabilities toward landslides, which mean that those pixels have higher confidence probability toward stability. Therefore, having a lower percentage (or even better, an absence) of the lower susceptibility classes indicates higher confidence in the misclassification error (equal to 1−ACC) achieved by those models. Further, they indicate that the misclassification error achieved was near the classification threshold (for binary equal-proportions, the classifications threshold is 0.5) and not at the extremes. These results have revealed similar, but somewhat better insight into the overall performance capabilities of each susceptibility model (i.e. for this case study) compared to the results of probabilistic performance metrics such as AUC, ACC and Kappa index. However, despite the fact that majority of the actual landslide instances fall into the "Very High" susceptibility class, the "somehow" dominance of the "Very low" class (Figure 5.5a) in term of Area extent covered by each class is very disturbing as it shows that study area is (at maximum) 31% mostly stable, while similarly the most adverse/opposite class, i.e. "Very High" zones occupy at least about 7% of the total area which high considering the scale of the analysis. Thus, reflecting the engrave danger the study area is facing in terms of the landslide.

Results of both landslide densities (Figure 5.5a) and the area covered by each susceptibility class (Figure 5.5b) within each landslide susceptibility model were satisfying with similar outcomes to model evaluation and comparison, yet fulfill the two spatial conditions mentioned earlier. All models show an increase in landslide density ratio when moving from low to high susceptible classes, with RF scoring the best results of approximately (90.83%) and (6.74%) for landslide density occurrences in the higher susceptible classes (both "High" and "Very High") and the area extent covered by only the "very high" susceptibility class, followed up by GBM with (89.44%) and (15.66%), SVM with (87.89%) and (12%), NNET with (86.33%) and (14.13%), and LR with (82.53%) and (11.36%).

5.2 DISCUSSIONS

The most effective way to reduce casualties and economic losses resulting from landslides are landslide risk planning and management; therefore, high-quality landslide susceptibility maps are an important tool [154]. However, it is still a challenge to produce high accuracy landslide susceptibility maps at a regional scale due to the complex nature of landslides and it is widely recognized that the prediction quality of landslides is dependent on the algorithm used. Thus, although various methodologies for producing landslide susceptibility maps have been developed, and yet the prediction accuracy of these methods is still debated [116]. On the other hand, it usually practical to experiment landslide prediction simulation that approximate predict future landslide pattern. Therefore, in the present study, five classifications algorithms (GBM, LR, NNET, RF, and SVM) were investigated and compared for landslide susceptibility mapping at Mila Basin.

The results obtained in this study (see Figure 5.3 and Table 5.2) show that all the implemented models achieved high performance[78] (AUC > 0.85, Acc > 78% and kappa > 0.56). However, two ensemble trees models (GBM and RF) yielded the highest prediction results compared to the others. This better performance is confirmed to be statistically significant with the used Wilcoxon signed-rank test. This finding is in agreement with the results from recent studies i.e., in ([155-158]) that reported that the ensemble models outperform single ML models. In contrast to GBM and RF, LR consistently yields the lowest results compared to the other implemented models. This finding is in line with the literature where LR achieves the worst, if not the poorest, performance of all models [100, 102, 122, 144, 159].

The better-fit and higher performance of GBM and RF compared to LR, NNET, and SVM in this research is due to the divide-and-conquer approach that the ensemble technique implements in both models (i.e., benefiting from aggregating weak learners to solve the issue). In fact, the main causes of error in the landslide modeling at the basin scale in this study is due to noise and the uncertainty that existed in the landslide conditioning factor maps (which were collected from various sources and scales). It is still difficult to eliminate noise and uncertainty, though

[78] It could be speculated that the balanced methodology used in this research, i.e. SMBO can also resulted in high performance.

several fuzzy modeling approaches have been proposed. However, ensemble learning, RF, and GBM, which use random sampling with replacement strategy, could minimize these due to their diversity and stability [160], which are two key issues of ensemble learning. Thus, both RF and GBM are capable models that work well over noise and uncertainty environments [161] such as landslide modeling, and therefore, they are very pleasant, robust, and better than the other models in this study for predicting the future landslide pattern.

Generally, GBM models offer similar or even better performance results than RF, but the large number of sensitive parameters and the tendency to easily over-fit makes it difficult to implement it right out the box compared to RF, which is easier to implement and less prone to both over-fitting and outliers. Additionally, some studies [e.g. 162] have found that GBM performs exceptionally well when the dimensionality is low (≈4000 predictors). Above that, RF has the best overall performance. Notably, the results obtained by SVM for typical binary landslide susceptibility problems are very satisfying. Even if it is lower than GBM, RF, and NNET, it is still relevant compared to the results produced by similar studies [e.g 102, 143, 144, 163, 164]. NNET, on the other hand, unsurprisingly outperforms SVM and LR, but fails to capture the underlying model of the input data like RF and GBM, simply because neural networks need a large number of observations. However, in the case of landslides, the observation events are scarce and very hard to obtain[79] . On top of samples being scarce, the most recent landslide susceptibility studies [100, 144, 165] do not benefit from the full potential of NNET by implementing NNET models with vanilla "Backpropagation" or one of its variances for the weight adjustments. In fact, Back-propagation based NNET are extremely slow to converge, which leads to a long execution time and a heavy computational load, not to mention both a large number of parameters to tune in and the special input data preparation required. Unlike Back-propagation NNETs, the implemented feed forward BFGS NNET are faster to converge with fewer hyperparameters to tune in and provided arguably better results than similar studies that implemented NNET [100, 144, 165].

[79] Obviously, the modeling performance increase with the increase of the overall number of landslide samples but, such strategy is not practical in life real situations that require not only experts and agencies commitments but also decent financial budget.

Theoretically speaking, the spatial pattern generalization which is provided by the ML algorithm and models, has led to a significant accuracy, but it is also apparent that the optimization strategy (SMBO) along with resampling strategy plays a crucial role. The implemented workflow methods explore all possible instances for each model and fine-tune it to the maximum, which eliminates bias and subjectivity and focuses on yielding results for any spatially correlated phenomenon, not just landslide distribution. For these reasons the expectations were not too optimistic, which has been proven right. The models rather served as a demonstration of the predictive capabilities[80] of the implemented models in landslide modeling in particularly in susceptibility assessment, which have proven to provide valuable results for some later decision-making process.

In the end, in order to achieve a statistically meaningful procedure, it is convenient to comment on the strengths and weaknesses of each model as it is widely accepted that no single or particular model can be depicted as the most suitable for all case scenarios. For example, the LR model is simple, fast, easy to implement, and is only able to capture the linear relationship between the conditioning factors and the landslide susceptibility. The merit of LR is that it does not compulsorily require normal distribution data. Additionally, both continuous and discrete data types can be used as an input for the LR model combined with the fact that LR models, in general, don't require heavy computational budget (CPU, Memory and Time) compared to the other models, but that is probably not of great importance for the task of predictive landslide mapping, outside the disaster management framework (where very quick but plausible solutions are needed). However, landslides are complex phenomena with non-linear mechanisms. SVMs are useful non-linear classifiers whose goal is not only to correctly classify landslide instances but also to keep the distance between instances and keep the separation of the hyperplane at a maximum. This makes SVM models appealing for susceptibility evaluation considering the number of hyperparameters to tune in. In addition, since the solution for the SVM separating hyperplane is found from the convex quadratic programming optimization problem, it is guaranteed that the solution is globally optimal. Therefore, the SVM is a good replacement for Artificial Neural Networks

[80] If implemented with appropriate optimization and resampling strategy.

which are usually stuck at local optima and are very difficult to train. However, if those hyperparameters are inappropriately set, SVM will often lead to unsatisfactory results. NNET models are very effective for simulating non-linear complex phenomena with multiple conditioning factors (preferably continuous input dataset). However, being a black box model and a large number of samples required to obtain a reliable model are the only downsides to this kind of model. Ensemble tree models (GBM and RF) offer excellent performance with decent interpretability and a moderate number of hyperparameters to tune in but require a considerable time budget (they require a lot of time to converge, especially if used on large-scale analyses). Though some studies (such as in Vorpahl, Elsenbeer [166]) highly recommend RF and GBM due to the outstanding performance, they suggest that a rather fast and simple model, such as LR would be much better than advanced ML models.

5.3 SUMMARY

As a summary for this chapter we conclude the following:

- The implemented models turned out to be relatively accurate, as expected from ML-based models.
- Models were able to produce relatively reliable prognosis, with a slight underestimation, while still being reasonably simple and GIS-integrated.
- The model can be characterized as underestimating in terms of landslide instances, but yet following logical trends of landslide occurrence.
- The comparison among the five implemented models in terms of the results demonstrates that the implemented models are able to provide very pleasant and robust results with GBM and RF being optimal for predicting the future landslide pattern.
- In some cases, the Occam's razor directly applies, so that the simplest solution – the simplest modeling method can provide an optimal solution. It would provide the optimal balance between the quality and complexity of the model [6, 18].

Apart from these points, the experiment design was valid (selection of the splits, optimization of the parameters, preprocessing of the inputs were apparently correct) as shown in the workflow.

Chapter 6: Main Achievements

Since the previous chapters have been rather voluminous and the information turned abundant and very detailed, some essential achievements and their relation to the initial research objectives are to be clarified in the following paragraphs. The objectives (structured as in Chapter 1.3) have been compiled by the following achievements:

This chapter will focus on explaining briefly the essential achievements fulfilled in this research and their relation to the initial research objectives (see Chapter 1.3):

1. Address the shortage in literature for Mila basin in term of landslide susceptibility mapping through investigating, implementing, assessing and comparing prediction capability of advanced statistical-based models such as Machine Learning methods and algorithms

   For this case study, a special highlight to the unique features, caveats, advantages, and drawbacks of the statistical approach in general and Machine learning in specific was outlined with specific attention to landslide susceptibility paradigm. In-depth analysis, assessment, and comparison are performed for each model and technique used in this research. Therefore, it could be said that Objective 1 has been appreciated consistently throughout this thesis.

2. The production of useful landslide susceptibility mapping and assessment frameworks with a reproducible and unbiased optimization process and exploit the possibility of automating the process of landslide susceptibility mapping or landslide mapping by taking advantage of low-cost data resources available at the local agencies and open source community.

Open source solutions, such as SagaGIS, GDAL, R, and others, have been fully exploited in the processing and modeling of the landslide susceptibility mapping and assessment framework. SMBO in particular was implemented for objectively automating the tuning process of the implemented model. In addition, the data that have been used were obtained for free from local agencies. These data turned sufficient for conducting the proposed methodology and utterly rounds-up Objective 2 of this thesis.

3. Standardizing the procedure regarding landslide assessment in the study area (i.e. acquisition, scaling, pre-processing, optimization, and evaluation procedures) by preparing custom and reproducible algorithms for specifically the purpose of landslide assessment in the study area using GIS.

   Although each case study may be different than others and even the same case study may vary over time thus exactly the input dataset may not be the same. Additionally, the used data are sometimes available in different time series mishandled, and sometimes the quality (i.e. resolution) of these images might not always satisfy the requirements. Therefore, standardizing and homogenizing the process for landslide assessment by a standard procedure[81] will ensure that all models will underwent the same processing procedure, and increase not only the objectiveness but also the reproducibility of the results and thus gaining confidence in using the results safely as demanded local agencies. These procedures are clearly explained (see Chapter 3: Chapter 3 and Chapter 2.3), which leads to a conclusion that the Objective 3 (see Chapter 1.3), has been fully perceived throughout the thesis. Thus, the Objective 2 was partially fulfilled.

[81] Providing the source code for the custom algorithms is one of the most efficient ways for reproducible experiment.

4. Implementing a variety of known models and techniques that rely on statistical modeling approaches, but also experimenting with the state-of-the-art techniques, advanced methods and unprecedented solutions for landslide assessment using GIS.

   Mila basin has been mostly and extensively elaborated in terms of classic and in-situ landslide studies, but no similar investigation performed over this area before. Thus, a different gamut of ML-based models and techniques have been intentionally implemented and elaborated with the GIS paradigm[82], which would supplement the next investigations, conducted by other practitioners. The resulting models are to present transient relative values over the area, pinpointing landslide-endangered zones and safe zones. In this sense, the fulfillment of Objective 4 has been asserted.

5. Evaluating the model's performance and the results obtained using the most appropriate procedures and methods, in favor of gaining qualitative and quantitative descriptors evaluations of the model's performance using GIS in combination with statistical tools.

[82] Keep in mind, typical details about the landslides phenomenology such as the triggering mechanism, distribution…etc., are out of the scope of this research, as they are widely presented and discussed in other authors work.

The evaluation of the individual models for this case study has been always given by several performance indicator metrics, such as ACC, κ-index, ROC curves and AUC. Nevertheless, the evaluation of the modeling performance predictive capability has been addressed by the sufficiency analysis of the generated susceptibility maps. An appropriate method, for model comparison (see Chapter 3) was based not only on pure model performance (e.g. ACC, AUC and so forth.) but also statistical significance between each pair of models. These evaluation procedures are counted as the most appropriate procedure for model evaluation since it will allow for qualitative and quantitative evaluation of the model. Thus, the Objective 1 was partially fulfilled and Objective 5 has been practically fulfilled.

6. Address the issues of availability, visualization and publishing of the detailed results in the form of reproducible, reliable, generic landslide susceptibility map per each model using GIS, and web-GIS and estimating their applicability for better environmental management and for reducing the victims and damages caused by future landslide occurrences.

   Visualization of the most of the models has been given by separate maps (see Chapter 3 and Chapter 2.3), while some of the insignificant results have not been visualized on purpose, the predictive models have been additionally featured as interactive web-maps and made publicly available (Figure 5.4). Their applicability is left for discussion of those who find them appealing or useful for their particular needs (planning, modeling, mapping, managing…etc.). Therefore, the final objective of this thesis, Objective 6, has also been completed.

Chapter 7: Conclusions

This thesis rounds-off, summing up a detailed methodological for framework proposal for mapping landslide susceptibility using non-conventional approaches such as GIS and statistics using efficient and advanced modeling techniques and methods. These have been tailored specifically according to underlined research motifs and objectives, which have been consistently followed. The thesis focus on Mila Basin on which the proposed methodology has been fully employed, tested and discussed. It outputs a handful of different interpretable models, which have to depend on the underlined goals, motifs and objectives of the research, limitations, drawbacks, benefits, caveats and different practical relevancies that need to be taken into account and consideration, are highlighted.

7.1 BENEFITS AND DRAWBACKS

Although this research had been conducted and gained some result, some benefits and drawbacks have to be mentioned as follow:

- An obvious relationship between the complexity of the implemented models and their GIS integration, in which the difficulty of implementing such a model would be increased proportionally to the complexity of the model. Thus, rending the simple models fairly used and implemented. Yet, this limits to a certain degree the overall benefit from advanced models due to the fact that either manual handling, or an additional programming effort and skills for data manipulation outside the GIS environment. However, despite the fact that Statistical based-models such as ML tend to provide excellent results, they generally[83] computationally expensive and unsuitable for quick predictions. Obviously, this wouldn't be an issue in the modern era of the outstanding advancement in computer science, software solutions, parallel and cloud computing, which should maximize the performance and shortens the processing time. Therefore, It plausible to assume that complex models nowadays are eventually becoming

[83] Some exceptions exist. For example, models like LR.

obsolete with the easy deploying systems, but then even more complex models will take over with new demands and new challenges posed to the hardware and software solutions.

- In this particular research, it has been inferred that:
  - Linux as an operating system is more computational and programming friendly[84] compared to the Microsoft Windows as it only user-friendly:
  - ArcGIS is the most robust GIS platform, but fails to follow up the module development as fast as its open-source counterparts (GDAL, SagaGIS, QGIS...etc.);
  - R is very customizable and very flexible, plus it is practically GIS-integrated, but not too user-friendly and not so robust for handling large datasets like C and C++ and therefore some under-the-hood optimization needs to be performed.

Overall, a combination of various solutions is still necessary, but holistic solutions are perceivable and R is one solid example of it.

- Another critical issue that has been discussed briefly (in Chapter 2.3 and Chapter 3.2) and can be considered a drawback is the model's evaluation. It is very hard to evaluate the predictions in a landslide susceptibility scenario[85] due to the fact that only present (and past) landslides can be available. Future landslides on the other hand, are foreseeable, and therefore cannot be taking into account. Moreover, it is obvious that all performance indicators are not necessarily correct as it may be deceiving, because of the predictive nature of the model should not be suppressed by the strict performance metrics. Some of the performance evaluation methods (i.e. non-parametric tests or sufficiency analysis) are taking into account the difference in performance or the micro differences in each particular landslide susceptibility class. These approaches could re-endorse

[84] Tools for compiling and developing optimized libraries are very easy in Linux compared to Windows.

[85] Unlike hazard risk assessment scenarios where the prognosis relates to the specified time series

the model which has been underestimated. It is probably the most objective evaluation method, thus far.

- It turn out that there is an important benefit from up-scaling to be discussed, tilling of the area into several sub-areas is not beneficial as it is important to mention that ML-based models would be affected and compromised by such solution. Experiences drawn from this research, suggest that the area with $1\sim 4.5 * 10^6$ points (pixels) is a fair upper limit for the size for the study area, while the lower limit could be 100000. These limits apply only to particular circumstances[86].
- From the implemented models, only the two ensemble tree models (RF and GBM) were proven the most suitable models for this case study when comparing them to the remaining models (NNET, SVM, and LR), as they significantly outperformed the rest of the models based on the excellent performance results achieved. Despite that, the remaining three models are considered viable options, as they are adequately capable of satisfactory performance compared to similar studies.
- The achieved results demonstrate that there is a significant difference between the implemented models. Even if the obtained results are underlined with a clear objective of comparing and assessing those models, finding the most suitable model for the case study was very challenging as it does not depend solely on the performance results, but also on the high level of uncertainty behind landslide modeling and the limitation and caveats that come with each model.
- There still some difficulties and uncertainties behind landslides modeling and producing accurate results, due to the fact that the modeling processes is heavily depended on the tuning details, the used approach and the supplied data [4, 39-41]. Despite the invested efforts in hazard modeling and landslide susceptibility specifically, an absence of solid agreement about the suitability of a given technique or method for landslide-prone areas prediction is still present [31]. Therefore, assess and compare the

[86] Depends on hardware solutions and software capabilities.

prediction capabilities of advanced ML methods for landslide susceptibility should be carried out.

- ML-based inherent one critical disadvantage of from the statistical approach which is requiring a significant number of conditioning factors to obtain reliable results.
- Finally, the data are sparse and very hard to come by, especially when local agencies don't collaborate on projects and work using unified standards (different software, different formats, resolution…etc.) making the process of obtaining and normalizing the data very difficult. For this reason, only data from local municipalities are used for the conditioning factors involved in building the susceptibility maps. However, it is desirable to once again underline that high quality of input data can guarantee a plausible result, even by using the simplest modeling solutions, while on the other hand, no model, no matter how sophisticated cannot help if the input data are poor in quality.
- Difficulties also arise from purely technical causes, such as the lack of independent, long-lasting, institutionalized landslide agencies on a national level, which would focus on all the aspects of landslide problematic, including their assessment and provide the research continuity. At present, individual projects at universities or institutes are treating this problem, but only during the project lifetime. At best, there are cases where multi-scaled and nation-wide researches are involved, but most commonly landslide assessment is disconnected into separate case-studies and focused on very specific project objectives, rather than revealing of the fundamental breakthroughs in landslide knowledge [5, 31].

7.2 APPLICABILITY

The results obtained in this research highlight the effectiveness of all Five ML technique classifiers, especially ensemble tree models such as the GBM and RF algorithms for the assessment of landslide susceptibility. However, depending on the purpose each model can find its purpose at some level of the assessment that can vary from the preliminary to the detailed research stage. Yet, the most beneficiary

that can substantially benefit from each model is the detailed landslide mapping analysis.

Despite the overwhelming advantages of such models in landslide susceptibility framework, they are not for the purpose of replacing conventional mapping but rather supplement it in different stages of the landslide hazard map development[87] as they are in the current stat a semi-products of landslide assessment[88] . On top of that, finding the most suitable model for the case study is very challenging as it does not depend solely on the performance results, but also on the high level of uncertainty behind landslide modeling and the limitation and caveats that come with each model.

Summing up, the obtained landslide susceptibility maps by the implemented models can successfully pinpoint the critical areas and guide the practitioners towards more efficient mapping rending the resultant susceptibility maps as preliminary planning framework for planners or as a technical framework for countermeasures and regulatory policies by decision-makers to minimize the damages introduced by either existing or future landslides by the Mila and Constantine municipalities. Thus, each model output (depending on purpose) can easily find their purpose in regional, small scale planning, urban planning, strategic planning, but also some preliminary insurance analysis, planning of detailed research or sampling, updating the inventories, tracking changes and so forth.

7.3 RECOMMENDATIONS AND FURTHER NOTICES

Based on this research, some recommendations have been suggested as much more needs to be done to achieve reliable semi-automated landslide mapping and landslide susceptibility assessment for future studies which adapting this research and for mitigation plans:

First, there are only 16 (sixteen) factors used in building the landslide susceptibility maps. Introducing more richness to the input data pool by i.e. several

[87] Especially in the preliminary stage (i.e. in the early stage of research planning).

[88] Intermediate models which are used by the experts to compile a final map.

factors related to influence landslide occurrences[89] that can be added into the model from new resources of inputs is one of the milestones for further model refinement. Expert-based inputs could significantly contribute to more accurate analysis, especially if they are focusing on the terrain features[90]. For instance, a landslide is area-based, thus having this kind of information on each event it opens the possibility of generating additional synthetic inputs, such as statistical parameters (variances, means, standard deviations, etc.) of other inputs. This would offer the whole new source of relations between the landslide occurrence and the input data. Similarly, inputting geological domains as quasi-homogeneous areas in terms of stratigraphy and lithology has been proven useful in landslide assessment. Unfortunately, such inputs require serious additional engagements of experts, local agencies and resources which can turn insurmountable problems (e.g. generating of geological quasi-homogeneous domains require extensive RS and field techniques and qualified experts to generate it, although there are some trends toward creating simple domains automatically).

Moreover, multi-temporal inputs[91] are very desirable but unfortunately, they are rarely available if not impossible to come by, they actually help in overhauling the susceptibility assessment to a hazard or risk framework. Such an integrated approach does sound optimal, and with the present development of RS systems, it is realistic to expect that in a couple of decades from now it will be much easier to model landslide hazard and risk. Another idea for more precise modeling is either introducing more landslide events or includes more precise landslide events such as landslide source areas in the inventory, i.e. to discern between the source and accumulation areas of the landslide body at inventory level, and to train the model only over the areas which have suffered the conditions leading to failure.

Furthermore, the implemented models in this research (i.e. LR, GBM, NNET, RF, and SVM) are counted as most advanced techniques which theirs rely on statistical-based approach for landslide susceptibility assessment, produced very

[89] Such as deformation formations based on InSAR space born imagery or distance to settlement also can be used as the causal factors.

[90] Geological or Engineering-geology.

[91] Historical repositories, especially on dating slope displacements, aerial and satellite images in monitoring context, terrestrial monitoring techniques, such as surveying, LiDAR and Radar scanning.

accurate results due to the fact that their generalization power is getting fully exploited (with the current research workflow). It will be better to not hesitate with exploring and challenging other advanced ML algorithms and techniques if proven to be effective. For example, experimenting with ensembles of classifiers in the form of classifiers chains by combining different techniques in the same manner as GBM and RF would theoretically produce more robust and readily post-processed models should be expected therein.

All these comments are proposing the ideas for improvements in the susceptibility or spatial landslide prediction. However, the susceptibility map shows that many regions of the study area are in high susceptibility. It is too difficult to make communities staying out of these areas. Slope stabilization methods should be implemented to reduce the possibility of landslide occurrences. Structural mitigation activities such as sub drains, retaining wall, gabion…etc.; are assumed can diminish the mass movement. Biotechnical mitigation such as planting the deep root vegetation can also be an effective way of slope stabilization. In addition to structural mitigation, non-structural mitigation can also reduce the effect of the mass movement. Increasing the awareness of local communities to mitigate the mass movement is believed as the key to reducing the occurrences and the effects of the landslides.

In the end, assuming that at one point, the most optimal solution for susceptibility framework will be reached, it would then be an entirely new challenge to deal with the hazard and risk frameworks, which is the author's remote objective, from the current stand-point.

Bibliography

1. Van Westen, C.J., *Application of geographic information systems to deterministic landslide hazard zonation.* Boletin de Vias, 1994. **21**(79): p. 11-141.

2. Soeters, R. and C.J. Van Westen, *Landslides: Investigation and mitigation. Chapter 8-Slope instability recognition, analysis, and zonation.* Transportation research board special report, 1996(247).

3. Guzzetti, F., et al., *Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy.* Geomorphology, 1999. **31**(1): p. 181-216.

4. Tien Bui, D., et al., *Landslide susceptibility assessment in the Hoa Binh province of Vietnam: A comparison of the Levenberg–Marquardt and Bayesian regularized neural networks.* Geomorphology, 2012. **171-172**: p. 12-29.

5. Van Westen, C.J., T.W.J. Van Asch, and R. Soeters, *Landslide hazard and risk zonation—why is it still so difficult?* Bulletin of Engineering geology and the Environment, 2006. **65**(2): p. 167-184.

6. Brenning, A., *Improved spatial analysis and prediction of landslide susceptibility: Practical recommendations.* Landslides and Engineered Slopes, Protecting Society through Improved Understanding, edited by: Eberhardt, E., Froese, C., Turner, AK, and Leroueil, S., Taylor & Francis, Banff, Alberta, Canada, 2012: p. 789-795.

7. Fell, R., et al., *Guidelines for landslide susceptibility, hazard and risk zoning for land use planning.* Engineering Geology, 2008. **102**(3): p. 85-98.

8. Gerath, R., et al., *Guidelines for legislated landslide assessments for proposed residential development in British Columbia.* Association of Professional Engineers and Geoscientists of British Columbia, Vancouver, BC Google Scholar, 2006.

9. Highland, L.M. and P. Bobrowsky, *The Landslide Handbook - A Guide to Understanding Landslides*, in *Circular*. 2008, US Geological Survey.

10. Varnes, D.J., *Landslide hazard zonation: a review of principles and practice*. Natural hazards. 1984: UNESCO.

11. WP/WLI, *A suggested method for describing the rate of movement of a landslide.* Bulletin of the International Association of Engineering Geology - Bulletin de l'Association Internationale de Géologie de l'Ingénieur, 1995. **52**(1): p. 75-78.

12. Cruden, D.M. and D.J. Varnes, *Landslide Types and Processes*. Vol. 247. 1996. 36-57.

13. Cruden, D. and R. Couture, *More Comprehensive Characterization of Landslides: Review and additions*. 2010.

14. Bell, F.G., *ENGINEERING GEOLOGY | Problematic Rocks*. Encyclopedia of Geology. 2005: Elsevier. 543-554.

15. Cruden, D., *The Working Classification of Landslides: material matters*. 2019.

16. Cruden, D. and H. Lan, *Using the Working Classification of Landslides to Assess the Danger from a Natural Slope*. 2014.

17. Lee, E.M., *Landslide risk management: Key issues from a British perspective*, in *Landslide Risk Assessment*. 2018, Routledge. p. 227-237.

18. Lee, E.M. and D.K.C. Jones, *Landslide Risk Assessment*. Landslide risk assessment. 2004: Thomas Telford Ltd.

19. WP/WLI, *A suggested method for describing the activity of a landslide.* Bulletin of the International Association of Engineering Geology - Bulletin de l'Association Internationale de Géologie de l'Ingénieur, 1993. **47**(1): p. 53-57.

20. Varnes, D.J., *Slope movement types and processes.* Special report, 1978. **176**: p. 11-33.

21. Guzzetti, F., et al., *Landslide inventory maps: New tools for an old problem.* Earth-Science Reviews, 2012. **112**(1): p. 42-66.

22. Chacón, J., et al., *Engineering geology maps: landslides and geographical information systems.* Bulletin of Engineering Geology and the Environment, 2006. **65**(4): p. 341-411.

23. Castellanos Abella, E.A. and C.J. van Westen, *Landslide hazard assessment using the heuristic model.* 2001.

24. Van Westen, C.J., T.W.J. Van Asch, and R. Soeters, *Landslide hazard and risk zonation—why is it still so difficult?* Bulletin of Engineering Geology and the Environment, 2005. **65**(2): p. 167-184.

25. Montgomery, D.R. and W.E. Dietrich, *A physically based model for the topographic control on shallow landsliding.* Water Resources Research, 1994. **30**(4): p. 1153-1171.

26. Van Westen, C.J., *Application of geographic information systems to landslide hazard zonation.* 1993.

27. Lary, D.J., et al., *Machine learning in geosciences and remote sensing.* Geoscience Frontiers, 2016. **7**(1): p. 3-10.

28. Kanevski, M., A. Pozdnoukhov, and V. Timonin, *Machine Learning for Spatial Environmental Data*. 2009: EFPL Press.

29. Mitchell, T.M., *Machine learning. WCB*. 1997, McGraw-Hill Boston, MA:.

30. Witten, I.H., E. Frank, and M.A. Hall, *Embedded Machine Learning*, in *Data Mining: Practical Machine Learning Tools and Techniques*. 2011, Elsevier. p. 531-538.

31. Carrara, A. and R.J. Pike, *GIS technology and models for assessing landslide hazard and risk.* Geomorphology, 2008. **94**(3-4): p. 257-260.

32. Bonham-Carter, G.F., *Geographic Information Systems for geoscientists-modeling with GIS*. Vol. 13. 1994: Elsevier BV. 398-398.

33. Van Westen, C.J., E. Castellanos, and S.L. Kuriakose, *Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview.* Engineering Geology, 2008. **102**(3): p. 112-131.

34. Mondini, A.C., et al., *Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images.* Remote Sensing of Environment, 2011. **115**(7): p. 1743-1757.

35. Kennedy, P., *A Guide to Econometrics*. 5th ed. 2003: The MIT Press. 500-500.

36. Marquaridt, D.W., *Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation.* Technometrics, 1970. **12**(3): p. 591-612.

37. Hair, J.A.R.T.R. and W. Black, *Multivariate data analysis*. 1998.

38. Neter, J., W. Wasserman, and M.H. Kutner, *Applied linear regression models*. 2 ed. 1989: Richard D Irwin. 667-667.

39. Yilmaz, I., *Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat—Turkey).* Computers & Geosciences, 2009. **35**(6): p. 1125-1138.

40. Tien Bui, D., et al., *Landslide susceptibility assessment in vietnam using support vector machines, decision tree, and Naive Bayes Models.* Mathematical Problems in Engineering, 2012.

41. Pradhan, B., et al., *Land subsidence susceptibility mapping at Kinta Valley (Malaysia) using the evidential belief function model in GIS.* Natural Hazards, 2014. **73**(2): p. 1019-1042.

42. Breiman, L., *Random forests.* Machine learning, 2001. **45**(1): p. 5-32.

43. Breiman, L., et al., *Classification and regression trees.* 1984: p. 368-368.

44. Breiman, L. and A. Cutler. *Random Forests*. 2004 30 August 2017]; Available from: https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm.

45. Efron, B., *The Jackknife, the Bootstrap and Other Resampling Plans*. CBMS-NSF Regional Conference Series in Applied Mathematics. 1982: Society for Industrial and Applied Mathematics. 96.

46. Efron, B. and R.J. Tibshirani, *An Introduction to the Bootstrap*. 1993: Springer US.

47. Ferreira, A.J. and M.A.T. Figueiredo, *Boosting Algorithms: A Review of Methods, Theory, and Applications*, in *Ensemble Machine Learning: Methods and Applications*, C. Zhang and Y. Ma, Editors. 2012, Springer US: Boston, MA. p. 35-85.

48. Trigila, A., et al., *Comparison of Logistic Regression and Random Forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy).* Geomorphology, 2015. **249**(Supplement C): p. 119-136.

49. Breiman, L., M. Last, and J. Rice, *Random Forests: Finding Quasars*, in *Statistical Challenges in Astronomy*, E.D. Feigelson and G.J. Babu, Editors. 2003, Springer New York: New York, NY. p. 243-254.

50. Valiant, L.G., *A Theory of the Learnable.* Commun. ACM, 1984. **27**(11): p. 1134-1142.

51. Freund, Y. and R.E. Schapire, *Experiments with a New Boosting Algorithm*, in *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*. 1996, Morgan Kaufmann Publishers Inc.: Bari, Italy. p. 148-156.

52. Freund, Y. and R.E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting.* Journal of Computer and System Sciences, 1997. **55**(1): p. 119-139.

53. Friedman, J., R. Tibshirani, and T. Hastie, *Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors).* The Annals of Statistics, 2000. **28**(2): p. 337-407.

54. Friedman, J.H., *Greedy function approximation: a gradient boosting machine.* Annals of statistics, 2001. **29**(5): p. 1189-1232.

55. Natekin, A. and A. Knoll, *Gradient boosting machines, a tutorial.* Frontiers in Neurorobotics, 2013. **7**.

56. Hastie, T., R. Tibshirani, and J. Friedman, *Boosting and Additive Trees*, in *The Elements of Statistical Learning*. 2009, Springer. p. 79-113.

57. Meir, R. and G. Rätsch, *An introduction to boosting and leveraging*, in *Advanced lectures on machine learning*. 2003, Springer-Verlag New York, Inc. p. 118-183.

58. Schapire, R.E., *The Boosting Approach to Machine Learning: An Overview*, in *Nonlinear Estimation and Classification*, D.D. Denison, et al., Editors. 2003, Springer New York: New York, NY. p. 149-171.

59. Breiman, L., *Arcing classifier (with discussion and a rejoinder by the author).* Ann. Statist., 1998. **26**(3): p. 801-849.

60. Mason, L., et al., *Boosting Algorithms as Gradient Descent in Function Space*. 1999.

61. Mason, L., et al., *Boosting Algorithms As Gradient Descent*, in *Proceedings of the 12th International Conference on Neural Information Processing Systems*. 1999, MIT Press: Denver, CO. p. 512-518.

62. Nelder, J.A. and R.J. Baker, *Generalized Linear Models*. 2004, John Wiley & Sons, Inc.

63. Paola, J.D. and R.A. Schowengerdt, *A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery.* International Journal of Remote Sensing, 1995. **16**(16): p. 3033-3058.

64. Møller, M.F., *A scaled conjugate gradient algorithm for fast supervised learning.* Neural Networks, 1993. **6**(4): p. 525-533.

65. Ripley, B.D., *Pattern recognition and neural networks*. 2007: Cambridge university press. 403-403.

66. Cristianini, N. and B. Schölkopf, *Support Vector Machines and Kernel Methods: The New Generation of Learning Machines.* Artificial Intelligence Magazine, 2002.

67. Yao, X. and F.C. Dai, *Support vector machine modeling of landslide susceptibility using a GIS: A case study.* The Geological Society of London, IAEG, 2006.

68. Ballabio, C. and S. Sterlacchini, *Support Vector Machines for Landslide Susceptibility Mapping: The Staffora River Basin Case Study, Italy.* Mathematical Geosciences, 2012. **44**(1): p. 47-70.

69. Brown, M.P.S., et al., *Knowledge-based analysis of microarray gene expression data by using support vector machines.* Proceedings of the National Academy of Sciences, 2000. **97**(1): p. 262-267.

70. Chen, W., et al., *Landslide susceptibility mapping based on GIS and support vector machine models for the Qianyang County, China.* Environmental Earth Sciences, 2016. **75**(6): p. 1-13.

71. Colkesen, I., E.K. Sahin, and T. Kavzoglu, *Susceptibility mapping of shallow landslides using kernel-based Gaussian process, support vector machines and logistic regression.* Journal of African Earth Sciences, 2016. **118**(Supplement C): p. 53-64.

72. Decoste, D. and B. Schölkopf, *Training Invariant Support Vector Machines.* Machine Learning, 2002. **46**(1): p. 161-190.

73. Guo, Q., M. Kelly, and C.H. Graham, *Support vector machines for predicting distribution of Sudden Oak Death in California.* Ecological Modelling, 2005. **182**(1): p. 75-90.

74. Huang, C., L.S. Davis, and J.R.G. Townshend, *An assessment of support vector machines for land cover classification.* International Journal of Remote Sensing, 2002. **23**(4): p. 725-749.

75. Joachims, T., *Text categorization with Support Vector Machines: Learning with many relevant features BT - Machine Learning: ECML-98: 10th European Conference on Machine Learning Chemnitz, Germany, April 21–23, 1998 Proceedings*, C. Nédellec and C. Rouveirol, Editors. 1998, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 137-142.

76. Xu, C., et al., *GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China.* Geomorphology, 2012. **145**(Supplement C): p. 70-80.

77. Hastie, T., R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. 2 ed. Vol. 1. 2009: Springer. 745.

78. Burges, C.J.C., *A Tutorial on Support Vector Machines for Pattern Recognition.* Data Mining and Knowledge Discovery, 1998. **2**(2): p. 121-167.

79. Burges, C.J.C. and B. Schölkopf. *Improving the accuracy and speed of support vector machines*. 1997.

80. Schölkopf, B., C. Burges, and A.J. Smola, *Advances in kernel methods: support vector learning. 1999*. 1999, The MIT Press.

81. Scholkopf, B., et al., *Input space versus feature space in kernel-based methods.* IEEE Transactions on Neural Networks, 1999. **10**(5): p. 1000-1017.

82. Schölkopf, B. and A.J. Smola, *A Short Introduction to Learning with Kernels*, in *Advanced Lectures on Machine Learning: Machine Learning Summer School 2002 Canberra, Australia, February 11–22, 2002 Revised Lectures*, S. Mendelson and A.J. Smola, Editors. 2003, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 41-64.

83. Schölkopf, B., A.J. Smola, and F. Bach, *Learning with kernels: support vector machines, regularization, optimization, and beyond*. 2002: MIT press.

84. Shawe-Taylor, J., et al., *Structural risk minimization over data-dependent hierarchies.* IEEE transactions on Information Theory, 1998. **44**(5): p. 1926-1940.

85. Molinaro, A.M., R. Simon, and R.M. Pfeiffer, *Prediction error estimation: a comparison of resampling methods.* Bioinformatics, 2005. **21**(15): p. 3301-3307.

86. Stone, M., *Cross-Validation and Multinomial Prediction.* Biometrika, 1974. **61**(3): p. 509.

87. Bergstra, J., D. Yamins, and D.D. Cox, *Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures*, in *Proceedings of the 30th International Conference on Machine Learning*, D. Sanjoy and M. David, Editors. 2013, PMLR: Atlanta, GA, USA. p. 115-123.

88. Bischl, B., et al., *mlrMBO: A Toolbox for Model-Based Optimization of Expensive Black-Box Functions*, in *useR*. 2016: Stanford, California.

89. Bergstra, J., et al. *Algorithms for Hyper-Parameter Optimization*. 2011.

90. Hutter, F., H.H. Hoos, and K. Leyton-Brown, *Sequential Model-Based Optimization for General Algorithm Configuration BT - Learning and Intelligent Optimization: 5th International Conference, LION 5, Rome, Italy, January 17-21, 2011. Selected Papers*, C.A.C. Coello, Editor. 2011, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 507-523.

91. Snoek, J., H. Larochelle, and R.P. Adams. *Practical bayesian optimization of machine learning algorithms*. 2012. Curran Associates Inc.

92. Thornton, C., et al. *Auto-WEKA: combined selection and hyperparameter optimization of classification algorithms*. 2013. ACM Press.

93. López-Ibáñez, M., et al., *The irace package: Iterated racing for automatic algorithm configuration.* Operations Research Perspectives, 2016. **3**: p. 43-58.

94. B. Mockus, J. and L. Mockus, *Bayesian approach to global optimization and application to multiobjective and constrained problems*. Vol. 70. 1991. 157-172.

95. Frean, M. and P. Boyle, *Using Gaussian Processes to Optimize Expensive Functions*. 2008. 258-267.

96. Cox, D. and S. John, *SDO: A Statistical Method for Global Optimization*. 1997.

97. Fawcett, T., *An introduction to ROC analysis.* Pattern Recognition Letters, 2006. **27**(8): p. 861-874.

98. Hanley, J.A. and B.J. McNeil, *The meaning and use of the area under a receiver operating characteristic (ROC) curve.* Radiology, 1982. **143**(1): p. 29-36.

99. Negnevitsky, M., *Artificial intelligence: a guide to intelligent systems*. 2005: Pearson Education.

100. Tsangaratos, P. and I. Ilia, *Comparison of a logistic regression and Naïve Bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size.* Catena, 2016. **145**(Supplement C): p. 164-179.

101. Kantardzic, M., *Data Mining: Concepts, Models, Methods, and Algorithms: Second Edition*. 2011.

102. Tien Bui, D., et al., *Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree.* Landslides, 2016. **13**(2): p. 361-378.

103. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. **33**(1): p. 159-74.

104. Dormann, C.F., et al., *Collinearity: a review of methods to deal with it and a simulation study evaluating their performance.* Ecography, 2012. **36**(1): p. 27-46.

105. Booth, G.D., M.J. Niccolucci, and E.G. Schuster, *Identifying proxy sets in multiple linear regression: an aid to better coefficient interpretation.* Research paper INT (USA), 1994.

106. Ridgeway, G., *gbm: Generalized Boosted Regression Models*. 2017.

107. R Core Development Team, *R: A Language and Environment for Statistical Computing*. 2017, R Foundation for Ftatistical Computing: Vienna, Austria.

108. Venables, W.N. and B.D. Ripley, *Modern Applied Statistics with S*. Fourth ed. 2002, New York: Springer.

109. Wright, M.N. and A. Ziegler, *ranger: A fast implementation of random forests for high dimensional data in C++ and R.* Journal of Statistical Software, 2015. **77**(1): p. 1-17.

110. Meyer, D., et al., *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*. 2017.

111. Kertész, C., *Rigidity-Based Surface Recognition for a Domestic Legged Robot.* IEEE Robotics and Automation Letters, 2016. **1**(1): p. 309-315.

112. Oshiro, T.M., P.S. Perez, and J.A. Baranauskas. *How Many Trees in a Random Forest?* in *Machine Learning and Data Mining in Pattern Recognition*. 2012. Berlin, Heidelberg: Springer Berlin Heidelberg.

113. Kavzoĝlu, T., *An investigation of the design and use of feed-forward artificial neural networks in the classification of remotely sensed images.* 2001.

114. Hecht, N., *IEEE First Annual International Conference on Neural Networks San Diego, California June 21-24, 1987.* IEEE Expert, 1987. **2**(2): p. 14-14.

115. Ripley, B.D., *Statistical aspects of neural networks*. 1993, Springer US: Boston, MA. p. 40-123.

116. Wang, C., *A theory of generalization in learning machines with neural network applications*. 1994, University of Pennsylvania. p. 305.

117. Aldrich, C., J.S.J. Van Deventer, and M.A. Reuter, *The application of neural nets in the metallurgical industry.* Minerals Engineering, 1994. **7**(5): p. 793-809.

118. Kaastra, I. and M. Boyd, *Designing a neural network for forecasting financial and economic time series.* Neurocomputing, 1996. **10**(3): p. 215-236.

119. Kanellopoulos, I. and G.G. Wilkinson, *Strategies and best practice for neural network image classification.* International Journal of Remote Sensing, 1997. **18**(4): p. 711-725.

120. Can, T., et al., *Susceptibility assessments of shallow earthflows triggered by heavy rainfall at three catchments by logistic regression analyses.* Geomorphology, 2005. **72**(1-4): p. 250-271.

121. Youssef, A.M., M. Al-Kathery, and B. Pradhan, *Landslide susceptibility mapping at Al-Hasher area, Jizan (Saudi Arabia) using GIS-based frequency ratio and index of entropy models.* Geosciences Journal, 2015. **19**(1): p. 113-134.

122. Youssef, A.M., et al., *Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and*

comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. Landslides, 2015. **13**(5): p. 839-856.

123. Athmania, D., A. Benaissa, and M. Bouassida, *Propriétés minéralogiques des argiles gonflantes de la wilaya de Mila*. 2009.

124. Athmania, D., et al., *Clay and marl formation susceptibility in Mila province, Algeria.* Geotechnical and Geological Engineering, 2010. **28**(6): p. 805-813.

125. Chettah, W., *Investigation des propriétés minéralogiques et géomécaniques des terrains en mouvement dans la ville de Mila «Nord-Est d'Algérie»*, in *Sciences de la terre et de l'univers*. 2009, Université Hadj Lakhdar: Université Hadj Lakhdar. p. 175.

126. Coiffait, P.-E., *Un bassin post-nappes dans son cadre structural: l'exemple du bassin de Constantine (Algérie Nord-Orientale)*. 1992, Universite de Nancy. p. 405-405.

127. Côte, M., *Les régions bioclimatiques de l'est algérien*. 1974. 6-6.

128. Delga, M.D., *Mise au point sur la structure du Nord-Est de la Berbérie*. 1969.

129. Durand Delga, M., *Mise au point sur la structure du Nord-Est de la Berbérie. Publ. Serv. Carte géol. Algérie, NS.* Bull. soc. Géol. fr.,(7), xiii, 1969: p. 328-337.

130. Labiod, F., *Mouvements de masses et instabilités des terrains dans le bassin versant de l'Oued Mila caractérisations et enjeux socio économiques.* 2009.

131. Mebarki, A., *Le Bassin du Kébir-Rhumel: Hydrologie de surface et aménagement des ressources en eau.* 1982: p. 304-304.

132. Mebarki, A. and C. Thomas, *Analyse des relations entre écoulements superficiels et souterrains à partir des hydrogrammes des cours d'eau. Application au bassin du Kébir-Rhumel dans le Constantinois (Algérie).* Revue Hydrologie continentale, ORSTOM, Paris, 1988. **3**(2): p. 89-103.

133. Mounira, A., *Problèmes géologiques et géotechniques dans le bassin de Mila: Leur impact sur les ouvrages d'art*. 2002, Université Larbi Tebessi de Tébessa.

134. Remmache, I., *Potentiel en substances utiles non métalliques (gypse et sel gemme) du bassin de Mila (Algérie nord orientale).* 2006.

135. Rullan-Perchirin, F., *Recherches sur l'érosion dans quelques bassins du Constantinois (Algérie)*. 1985, Université Panthéon-Sorbonne (Paris): Paris, France. p. 356-356.

136. Rullan-Perchirin, F. and A. Rullan, *Les mouvements de masse dans le bassin-versant du Rhumel constantinois: essai méthodologique.* Travaux de l'Institut de Géographie de Reims, 1987. **69**(1): p. 151-171.

137. Zouaoui, S., *Etude géologique et géotechnique des glissements de terrains dans le bassin néogéne de mila:glissement de sibari*, in *Sciences de la terre et de l'univers*. 2008, Université Hadj Lakhdar: Université Hadj Lakhdar. p. 150.

138. Merghadi, A., B. Abderrahmane, and D. Tien Bui, *Landslide Susceptibility Assessment at Mila Basin (Algeria): A Comparative Assessment of Prediction Capability of Advanced Machine Learning Methods.* ISPRS International Journal of Geo-Information, 2018. **7**(7): p. 268.

139. Amireche, H., *L'eau, le substrat, la tectonique et l'anthropisation dans les phénomènes érosifs du tell nord-Constantine*. 2001, Univ. Mentouri Constantine. p. 266.

140. Ayalew, L. and H. Yamagishi, *The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan.* Geomorphology, 2005. **65**(1): p. 15-31.

141. Ayalew, L., et al., *Landslides in Sado Island of Japan: Part II. GIS-based susceptibility mapping with comparisons of results from two methods and verifications.* Engineering Geology, 2005. **81**(4): p. 432-445.

142. Chen, W., et al., *A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility.* Catena, 2017. **151**: p. 147-160.

143. Pham, B.T., et al., *A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India).* Environmental Modelling & Software, 2016. **84**: p. 240-250.

144. Tien Bui, D., et al., *Spatial prediction of rainfall-induced landslides for the Lao Cai area (Vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization.* Landslides, 2017. **14**(2): p. 447-458.

145. Yalcin, A., *GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): Comparisons of results and confirmations.* Catena, 2008. **72**(1): p. 1-12.

146. Olaya, V., *A gentle introduction to SAGA GIS.* The SAGA User Group eV, Gottingen, Germany, 2004. **208**.

147. Iwahashi, J. and R.J. Pike, *Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature.* Geomorphology, 2007. **86**(3-4): p. 409-440.

148. Conforti, M., et al., *Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy).* CATENA, 2014. **113**: p. 236-250.

149. Beven, K.J. and M.J. Kirkby, *A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant.* Hydrological Sciences Bulletin, 1979. **24**(1): p. 43-69.

150. Moore, I.D., R.B. Grayson, and A.R. Ladson, *Digital terrain modelling: A review of hydrological, geomorphological, and biological applications.* Hydrological Processes, 1991. **5**(1): p. 3-30.

151. Bossek, J., *smoof: Single- and Multi-Objective Optimization Test Functions.* The R Journal, 2017.

152. Carnell, R., *lhs: Latin Hypercube Samples*. 2016.

153. Stocki, R., *A method to improve design reliability using optimal Latin hypercube sampling.* Computer Assisted Mechanics and Engineering Sciences, 2005. **12**(4): p. 393-393.

154. Klimeš, J., et al., *Challenges for landslide hazard and risk management in 'low-risk'regions, Czech Republic—landslide occurrences and related costs (IPL project no. 197).* Landslides, 2017. **14**(2): p. 771-780.

155. Pham, B.T., D. Tien Bui, and I. Prakash, *Bagging based Support Vector Machines for spatial prediction of landslides.* Environmental Earth Sciences, 2018. **77**(4): p. 146.

156. Dang, V.-H., et al., *Enhancing the accuracy of rainfall-induced landslide prediction along mountain roads with a GIS-based random forest classifier.* Bulletin of Engineering Geology and the Environment, 2018.

157. Chen, W., et al., *GIS-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method.* CATENA, 2018. **164**: p. 135-149.

158. Hong, H., et al., *Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China).* CATENA, 2018. **163**: p. 399-413.

159. Conoscenti, C., et al., *Assessment of susceptibility to earth-flow landslide using logistic regression and multivariate adaptive regression splines: A case of the Belice River basin (western Sicily, Italy).* Geomorphology, 2015. **242**(Supplement C): p. 49-64.

160. Dai, Q., R. Ye, and Z. Liu, *Considering diversity and accuracy simultaneously for ensemble pruning.* Applied Soft Computing, 2017. **58**: p. 75-91.

161. Brillante, L., et al., *Investigating the use of gradient boosting machine, random forest and their ensemble to predict skin flavonoid content from berry physical–mechanical characteristics in wine grapes.* Computers and Electronics in Agriculture, 2015. **117**: p. 186-193.

162. Caruana, R., N. Karampatziakis, and A. Yessenalina, *An empirical evaluation of supervised learning in high dimensions*, in *Proceedings of the 25th international conference on Machine learning*. 2008, ACM: Helsinki, Finland. p. 96-103.

163. Goetz, J.N., et al., *Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling.* Computers & Geosciences, 2015. **81**: p. 1-11.

164. Dou, J., et al., *Shallow and Deep-Seated Landslide Differentiation Using Support Vector Machines: A Case Study of the Chuetsu Area, Japan.* Terrestrial, Atmospheric and Oceanic Sciences, 2015. **26**(02): p. 13.

165. Dou, J., et al., *An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan.* Natural Hazards, 2015. **78**(3): p. 1749-1776.

166. Vorpahl, P., et al., *How can statistical models help to determine driving factors of landslides?* Ecological Modelling, 2012. **239**(Supplement C): p. 27-39.

Appendices

Appendix A

This appendix is featuring Objective 6 (see Chapter 1.3).

- All scripts source codes used in this experiment are available on-line in Github (https://github.com/aminevsaziz/lsm_in_Mila_basin) or (https://github.com/aminevsaziz/lsm_in_Mila_basin).
- A reproducible container is available for this research repository is available on-line on (https://github.com/aminevsaziz/lsm_in_Mila_basin)
-

Appendix B

Table 7.1 The spatial relationship between the landslide conditioning factors and landslides.

Conditioning factors	***Class***	***Class Percentage (%)***	***Landslide Percentage (%)***
Altitude (m)	60 - 326.047	8.79	19.55
	326.047 - 597.105	36.06	48.79
	597.105 - 813.952	28.97	18.51
	813.952 - 1,003.694	18.64	7.79
	1,003.694 - 1,722	7.56	5.36
Slope angles (Slopes) (°)	0 - 5.543	26.67	21.11
	5.543 - 11.394	39.88	37.89
	11.394 - 18.16987664	23.33	28.37
	18.169 - 27.101	8.30	10.90
	27.101 - 78.530	1.83	1.73
Slope Aspects (Aspects)	Flat	0.76	1.04
	1st Quadrant (0° to 90°)	23.71	26.30
	2nd Quadrant(90° to 180°)	28.20	25.26
	3rd Quadrant(180° to 270°)	22.59	21.45
	4th Quadrant(270° to 360°)	24.75	25.95
Landforms	Steep slope, fine texture, high convexity	13.06	16.09
	Steep slope, coarse texture, high convexity	16.18	16.26
	Steep slope, fine texture, low convexity	5.85	6.92
	Steep slope, coarse texture, low convexity	10.67	13.67
	Gentle slope, fine texture, high convexity	3.17	4.15
	Gentle slope, coarse texture, high convexity	24.29	11.59
	Gentle slope, fine texture, low convexity	2.22	1.73
	Gentle slope, coarse texture, low convexity	24.56	29.58
Rainfall (mm/Year)	403 - 593.263	8.52	3.98
	593.263 - 711.030	50.81	21.28
	711.030 - 901.294	31.08	67.65
	901.294 - 1,208.684	9.60	7.09
Topographic Wetness Index (TWI)	0.034 - 3.550	3.35	5.19
	3.550 - 5.481	50.11	48.10
	5.481 - 8.997	44.91	45.16

	8.997 - 15.402	1.63	1.56
Distance to Hydrographic Network (m) (WDist)	0 - 300	10.79	18.86
	300 - 750	34.42	21.63
	750 - 1,500	25.68	27.16
	1,500 - 3,000	17.02	26.12
	3,000 - 5856	12.09	6.23
Lithology	Alluvium	5.33	7.44
	Claystone	4.06	2.94
	Colluvium-Detritus Deposits-Scree	9.44	7.61
	Limestone	8.22	7.61
	Marl	9.54	7.79
	Neogene Complex	56.05	59.86
	Sandstone	7.36	6.75
Stratigraphy	Quaternary	10.52	14.01
	Neogene	56.84	61.42
	Paleogene	12.38	6.40
	Upper Cretaceous	7.57	4.50
	Upper-Mid Cretaceous	8.61	9.86
	Lower Cretaceous	2.62	2.77
	Triassic-Jurassic	1.45	1.04
Distance to Faults (m) (FDist)	0 - 581	30.17	22.32
	581 - 4,784.550	61.89	61.25
	4,784.550 - 8192	7.94	16.44
Soil Texture (Texture)	Clay	19.01	25.43
	Sandy Clay	1.69	1.38
	Clay Loam	59.08	60.03
	Silty Clay Loam	0.80	1.73
	Sandy Clay Loam	19.42	11.42
Depth to Bedrock (cm) (DepthBR)	49 - 574.750	22.04	19.03
	574.7502397 - 761.629	33.81	34.43
	761.6293378 - 1,287.379	39.62	42.91
	1,287.379578 - 2,766.481	4.46	3.63
	2,766.481936 - 7,479	0.07	0.00
Bulk Density (Kg/m3) (Bdensity)	1,209 - 1,394.941	6.92	2.42
	1,394.941 - 1,463.333	25.29	32.01
	1,463.333 - 1,521.039	41.07	40.83
	1,521.039 - 1,754	26.72	24.74

Landuse	Water Bodies	1.47	3.98
	Artificial Surfaces	13.96	17.99
	Forests	7.56	6.92
	Grasslands	4.75	6.23
	CropLand	26.13	23.70
	Bareland	46.12	41.18
Soil type	Calcisols	9.80	1.21
	Cambisols	15.50	19.72
	Luvisols	50.17	58.65
	Leptosols	13.08	9.00
	Podzols	5.22	5.88
	Regosols	3.78	2.25
	Vertisols	2.46	3.29
Distance to Roads networks (m) (RDist)	0 - 908.103	25.70	40.31
	908.103 - 2,612.509	30.19	28.89
	2,612.509 - 5,811.481	32.09	24.74
	5,811.481 - 11957	12.02	6.06