DIPARTIMENTO DI SCIENZE DELLA TERRA
Universita` degli Studi di Torino

DOCTORAL SCHOOL OF SCIENCE AND INNOVATIVE TECHNOLOGIES
PhD PROGRAM IN EARTH SCIENCES

CURRENT AND FUTURE GLOBAL DEBRIS FLOW SUSCEPTIBILITY, HUMAN VULNERABILITY, AND THE IMPACT OF SPATIAL DATA UNCERTAINTY ON SUSCEPTIBILITY MODELING

PhD CANDIDATE: **Laurie Jayne Kurilla**

TUTOR: Prof. Giandomenico Fubelli
CO-TUTOR: Prof.ssa Anna Maria Ferrero

EXTERNAL EXAMINER Prof. Gaetano Robustelli
Universita` della Calabria
Department of Earth Sciences

EXTERNAL EXAMINER Prof. Luigi Lombardo
University of Twente
Department of Appled Earth Sciences

PhD CYCLE: XXXIV
PhD COORDINATOR: Prof.ssa Anna Maria Ferrero
ACADEMIC YEARS: 2018/2019, 2019/2020, 2020/2021
SCIENTIFIC DISCIPLINARY SECTOR: GEO/04 Physical geography and geomorphology

Dedication

My family made my dual-continent life easier through their love, acceptance, and support of my life's mission, while enduring many inconveniences with my absence, and not increasing my guilt for pursuing this selfish passion.

Acknowledgements

Throughout the research and writing of this dissertation I have received valuable direct and indirect support and assistance.

I am grateful to the University of Torino Department of Earth Sciences for accepting me into their program and giving me the opportunity to pursue my purpose in life; for the valuable guidance, assistance, and acceptance I received from professors, staff, and PhD students. Grazie mille to my advisors (co-tutors) Prof. Giandomenico Fubelli and Prof. Anna Maria Ferrero. Prof. Fubelli's expertise and guidance were invaluable in supporting my research and methodology questions, and my co-tutor Prof. Ferrero, was always available to respond to my questions and included me in opportune meetings and field excursions.

I would also like to thank Prof. Giuseppe Mandrone (University of Torino/Polytechnical University of Torino) for inviting my participation in professional introductions, and field trips, and for providing a feeling of inclusivity in the Department and Italy.

I wish to thank Dr. Pierluigi Pieruccini (University of Torino) and Dr. Leonardo Disperati (University of Siena) for their availability and guidance for my questions on soils.

I am grateful to my dissertation committee (Prof. Luigi Lombardo, University of Twente, and Prof. Gaetano Robustelli, Universita' della Calabria) for their time and expertise in reviewing and evaluating my dissertation.

In memory of Dr. Joan Petersilia (Stanford University): her excitement, deep understanding, and support for this pursuit of my destiny, was ever present. I think of her every time I use my timesaving, self-organizing bibliography software which she introduced to me as I began this journey. My dear friend inspired me as a human-being and as a professional, and I miss her deeply.

Table of Contents

List of Tables

List of Figures

Abstract

Debris flows, and landslides in general, are worldwide catastrophic phenomena. Due to limitations in funds, time, and data, landslide research is often focused on risk analyses, hazard assessment, and mitigation at a local or basin level.

Debris flows are local phenomena, occurring worldwide, yet less than 1% of the world has been studied for debris flow susceptibility. Debris flow causative and trigger factors are well known and appear to be relatively consistent across a variety of geographies, enhancing the ability to apply past information along with new data sets and processes to the possibility of future events. Sound generalizations about their behavior and predisposing and triggering factors exist at local, basin, and regional levels, and can be upscaled to continental and global scales for global modeling.

Due to forecasted climate changes, debris flow frequency, intensity, and distribution may increase, and rates of return may decrease in the next thirty to fifty years. This eventuality coincides with projections on increased populations and urbanization, leading to the prospect of increased human and economic impact by debris flows. These realizations drive the necessity of global debris flow susceptibility models for today and projected into the future.

This project develops GIS-based "presence only" models utilizing historic debris flow events, and remotely sensed data to model the global distribution of potential debris flow susceptible areas in present day, and the year/decade 2100.

A phased approach, with multiple stages of modeling, was required to reach the ultimate objectives of this research and to address four critical questions: 1) can a single global model adequately represent all continents, 2) what impact does the uncertainty associated with debris flow event locations have on a resulting susceptibility modeling and how to address the issue, 3) how will debris flow susceptibility differ in the future as a result of potential climatic changes,

and 4) looking beyond the traditional debris flow predisposing factors, how do global wildfire probability, and glacier response to climatic changes, amplify future debris flow susceptibility and subsequent impact on humans and economies? These topics are fundamental and necessary to adequately address global debris flow susceptibility, current and future.

The first stage was to determine the efficacy of a single global model, versus individual continental models, in determining global debris flow susceptibility. Global models were developed utilizing five statistical methods (certainty factor, probability density, conditional probability, frequency ratio, and maximum entropy). The best performing global statistical method was used to develop individual continental models. The resulting global and continental models were compared based on model performance testing. The single global model developed with the maximum entropy algorithm was the best performing methodology and was utilized in the subsequent stages of modeling.

The second overarching concern is the impact of data uncertainty. In landside susceptibility modeling, determining those factors and factor classes (environmental variables) with the most impact or significance to slope instability is paramount. The effect of event location uncertainty, on factor and factor class determinations, was analyzed. A model utilizing a subset of the data with the least location uncertainty (<= 5 km) demonstrated the best model performance when verified against test data and model AUCs. A novel scheme was developed to graphically represent the resulting uncertainty to model users and decision-makers, providing them with an ability to choose areas for further study based on either a conservative or more expansive interest.

Based on predicted climate change scenarios (RCP 2.6 and RCP 8.5), the third study develops future debris flow susceptibility models. Utilizing predicted population and urbanization trends by the year 2100, human vulnerability is modeled through the intersect of debris flow susceptibility and population and urbanization.

The fourth and fifth studies employ current and future wildfire probability as a potential augmentation of the debris flow predisposing factors, demonstrating the increased hazard potential; and glacier melting as a potential amplifier to debris flow precipitation triggers.

Global susceptibility models resulting from this research can facilitate the international community in assessing risk, hazard mitigation, and planning by focusing on regions at higher risk. The conservative approach, methodologies and models developed in this research, are a valuable framework for developing an ensemble of global landslide susceptibilities for other landslide types or hazards, and a foundation for further important refinements such as spatial-temporal relationships between susceptibility and seasonal and anomalous precipitation; and more detailed analyses on the impact of debris flows on people and economics, as population, urbanization, and climate warming expand.

Keywords

Climate change, Debris flow, Debris flow susceptibility, GIS, Global perspective, Landslide susceptibility, Maximum Entropy, Presence-only, Uncertainty, Future susceptibility, Debris flow vulnerability

1 Introduction

Debris flows, and landslides in general, are worldwide catastrophic phenomena (Brabb, Colgan, & Best, 1999; Brighenti, Segalini, & Ferrero, 2013; Campbell, 1974; Dowling & Santi, 2013; Froude & Petley, 2018), present on all continents (Dilley, 2005; Gariano & Guzzetti, 2016; Guzzetti et al., 2012; Stanley & Kirschbaum, 2017), however, some continents are underrepresented with regard to focused studies (Broeckx, Vanmaercke, Duchateau, & Poesen, 2018; Gariano & Guzzetti, 2016). Due to limitations in funds, time, data, as well as focused interest in specific locales, landslide research is often necessarily directed at risk analyses, hazard assessments, and mitigation efforts associated with known landslide sites at a local level (Reichenbach, Rossi, Malamud, Mihir, & Guzzetti, 2018).

Debris flows are rapid (often exceeding 10 m/s), gravity-induced mass wasting flows consisting of viscous sediment, water, and air, with bulk densities comparable to those of rock avalanches (~2,000 kg/m^3), which may travel distances of tens of kilometers. Sediment sizes may range from poorly sorted clay to boulders and the flows may entrench and transport vegetation along the way and may reach volumes of 100,000 m^3. They are typically triggered by high intensity rainfall, and often follow preexisting drainage channels and gullies. (Costa, 1984; Hungr, Evans, Bovis, & Hutchinson, 2001; Hungr, Leroueil, & Picarelli, 2013; Iverson, 1997; Varnes, 1978; Varnes & IAEG Commission on Landslides, 1984).

The total land area subject to landslides (of all types) is about 3.7 million square kilometers with an affected population of nearly 300 million (Dilley, 2005) yet, according to Reichenbach et al (Reichenbach et al., 2018) there is "a clear geographical bias in susceptibility study locations with many studies in China, India, Italy, and Turkey, and only a few in Africa, South America, and Oceania". Of 565 projects from 1983 to 2016 reviewed, Reichenbach et al (Reichenbach et al., 2018) found that only 0.03% of the Earth's land area was investigated. "It is estimated that economic loss due to landslides may reach between 1-2% of the gross national product in many developing countries. Evaluating and mitigating the landslide hazard and risk is a major challenge for planners and decision makers in the developing world, as 80% of the reported

fatalities due to landslide is within the developing countries." (Geological Survey of India, 2021). There are many areas on Earth highly susceptible to debris flows which can impact lives and economies, but where it is difficult or impossible to study in the field either due to the dangers of the topographies or the local political climate. Due to worldwide expanding population and urbanization trends, the human and economic impact due to debris flow hazards necessitates broader geographic research. There is a necessity to view debris flow hazard as a global and international problem, and to identify those areas globally which may be highly susceptible.

Much research on debris flows has been conducted at the local level based on direct field surveys, and diverse statistical methodologies, and has resulted in considerable information about predisposing factors, triggers, and return rates associated with debris flows worldwide (Dou et al., 2015; Leoni et al., 2009; L. Lombardo, Fubelli, Amato, & Bonasera, 2016; Park, 2014; Rossi, Guzzetti, Reichenbach, Mondini, & Peruccacci, 2010; Soma, 2018; Van Westen, 2016; Wilford, Sakals, Innes, Sidle, & Bergerud, 2004; Zezere, Pereira, Melo, Oliveira, & Garcia, 2017). Although debris flow behavior and parameters are local phenomena, sound generalizations can be applied to debris flow susceptibility analyses at larger geographic extents based on these criteria (Jacobs et al., 2020; Kirschbaum, Stanley, & Yatheendradas, 2015).

As world population and urbanization grow in number and geographic coverage (Ritchie & Roser, 2018), the need exists to extend focus, research, and modeling to a continental and global scale currently and projected into the future. Localized field surveys to collect event inventories are not a practicable approach in continental and global hazard susceptibility modeling. Thus, debris flow susceptibility analyses at these scales require data-driven, systematic, statistical methodologies which include continental and global coverages of environmental predisposing factors which may influence susceptibility. Susceptibility, in this context, is a qualitative assessment of potential areas of instability with respect to debris flows.

Corominas et al (Corominas et al., 2013) suggest that "a landslide susceptibility map subdivides the terrain into zones with differing likelihoods that landslides of a certain type may occur". A key point is the recognition that likelihoods and susceptibility are associated with a landslide type. Different landslide types will have differing environmental factors, triggers, and different

spatial probabilities (van Westen, van Asch, & Soeters, 2005). van Westen et al (van Westen et al., 2005) consider susceptibilities to be relative indications of spatial probabilities.

Global susceptibility and non-susceptibility models have been developed by other researchers (Farahmand & AghaKouchak, 2013; Froude & Petley, 2018; Hong, Adler, & Huffman, 2007; Jia et al., 2021; Lin, Lin, & Wang, 2017; Nadim, Kjekstad, Peduzzi, Herold, & Jaedicke, 2006; Stanley & Kirschbaum, 2017) with substantial differences in the approaches, and environmental factors from the models developed herein. "Very few studies develop separate statistical models for different landslide types and most merge all active landslides together in one group which is used to generate statistical relations" (van Westen et al., 2005).

Regardless of methodology and environmental factors employed, all of the global landslide maps result in a similarity of results with landslide susceptibility in the same generalized areas across the world and on every continent.

While a landslide-type blind susceptibility model is valuable in identifying hazardous areas in general, landslide-type specific susceptibility models provide insight for specific actionable responses for planning, preventative, and mitigation efforts. An overlay of multiple and distinct landslide-type models will provide a more comprehensive, hazard-specific, and actionable view of global landslide hazards.

1.1 Global versus Continental Modeling

How far-reaching can a susceptibility analysis extend? Can the world be analyzed or modeled as one body for purposes of landslide susceptibility or is this an oversimplification not accounting for the geologic, geomorphic, and tectonic histories and dissimilarities of the continents, and disregard for the latitudinal influences of climatic conditions? This study illuminates the results obtained when performing a 'single' global debris flow susceptibility analysis versus analyses performed on a continent-by-continent basis to determine whether a single global model can adequately represent the continents. The objective of this research is to determine areas of debris flow susceptibility, globally. One can then "drill down" into such areas, regionally and locally, to examine in more detail and determine the potential hazard and risk. Subsequent "drill downs" are beyond the scope of this project.

In this study, susceptibility models were developed for the world, and for each of seven continents (Africa, Asia, Australia, Europe, North America, Oceania, and South America). An analysis was not performed for Antarctica as there are no Antarctic debris flows in the inventories employed.

1.2 Data and Model Uncertainties

Extensive landslide research and modeling has been conducted by the scientific community, with a goal toward assessing hazards and risks, influencing zoning policy, and developing mitigation and response plans. Mapping landslide susceptibility has been more limited, due to a lack of comprehensive historical data which is used in developing heuristic and statistical models (L. Chen et al., 2016; Van Den Eeckhaut & Hervás, 2012).

In a landslide susceptibility analysis, landslide event location accuracy is paramount yet often inaccurately known unless a direct field survey is conducted. Landslide inventories are often constructed based on mapping from aerial imagery, media reports, local governmental agencies, witness accounts, and field work by third party sources (Froude & Petley, 2018; D. Kirschbaum, T. Stanley, & Y. P. Zhou, 2015; Malamud, Turcotte, Guzzetti, & Reichenbach, 2004).

Uncertainties are inherent in all spatial data and at all scales (Chrisman, 1989), however when working at continental and global scales, and in the absence of direct field surveys, the uncertainties are inherently greater (Maffini, Arno, & Bitterlich, 1989; Openshaw, 1989). "Uncertainty exists widely in the natural world, and certainty is conditional and relative" (Shi, 2010). When utilizing methods which overlay and correlate multiple datasets, each with their own uncertainty, their derivative products, such as susceptibility maps, are prone to error propagation of an unknown magnitude (Shi, 2010). It is not a matter of adding more or better data, but rather a "sobering reminder that uncertainty is an irreducible part of sufficiently complex knowledge" (Couclelis, 2003), and thus should be thought of as a natural component of, and addressed in, every project. The presence of data uncertainty does not preclude the use of the data, but rather necessitates a methodology for qualitatively or quantitively characterizing and conveying the level of uncertainty and modeling the associated uncertainty for the benefit of the end-users of the hazard model. "Unlike industrial and other products of material processes,

knowledge products do not carry with them the evidence of their own inadequacy… and the most critical aspects of the quality of its products are often only testable through their indirect and sometimes remote consequences" (Couclelis, 2003). "Uncertainty is an attribute of information and therefore does not mean lack of knowledge" (Bloschl & Montanari, 2009). The aim in hazard modeling is to minimize the societal 'testing' of an inadequately understood susceptibility model.

1.3 The Impact of Climate Change on Future Debris Flow Susceptibility

It is widely accepted that climate changes will affect slope stability and the presence and prevalence of landslides are expected to be exacerbated (Alvioli et al., 2018; Broeckx et al., 2018; S. H. Cannon & DeGraff, 2009; Gariano & Guzzetti, 2016; Geertsema & Clague, 2011; Harrison et al., 2019; C. Huggel, Khabarov, Korup, & Obersteiner, 2012; C. Huggel et al., 2010; IPCC, 2019; D. K. C. Jones, 1993; Keiler, Knight, & Harrison, 2010; Knight & Harrison, 2012; Hanoch Lavee, Imeson, & Sarah, 1998; Li et al., 2018; Moreiras et al., 2021; J. D. Phillips, 2010; Stoffel, Mendlik, Schneuwly-Bollschweiler, & Gobiet, 2013). Many areas of the world are experiencing increases in minimum, mean, and maximum temperatures and more frequent heavy precipitation (Christian Huggel, Clague, & Korup, 2012). Slope instabilities may increase due to increases in temperatures and precipitation over short-term periods, and due to conductive heat transport to the subsurface causing debuttressing effects, with long-term lag-time effects (Christian Huggel et al., 2012). Furthermore, due to both climate change and population and urban growth, it is reasonable to assume the impact of these natural hazards will increase (Alexander, 2004 as cited in (Holec, Bednarik, Liscak, Zilka, & Vitovic, 2018).

In this project, future climate impact on debris flow susceptibility is modeled using the best-case scenario (RCP 2.6) and worst-case scenario (RCP 8.5) ends of the future warming scenario spectrum.

1.4 Human and Economic Impact

At least 14% of total casualties from natural hazards are due to slope failures, and ~49% of natural hazards are landslides (Froude & Petley, 2018). Between 1988-2017 it is estimated that there were more than 56,000 deaths worldwide due to landslides, more than 4.8 million people affected injuriously and/or economically, and about 6 billion euro/year for damages in industrialized countries (Costa, 1984; Froude & Petley, 2018, 2019; Highland & Bobrowsky, 2008; ISPRA, 2020). In Europe, and likely the world, the number of landslides and their societal and economic impact are grossly underestimated (Gunther, Reichenbach, Guzzetti, & Richter, 2007). It is plausible that these numbers have increased considerably due to increasing climate and anthropogenic changes. Yet landslides may be the one natural hazard which is most controllable with respect to human and economic impacts through susceptibility determinations coupled with urban planning policies. "The identification and map portrayal of areas highly susceptible to damaging landslides are the first necessary steps toward loss-reduction" (Zeizel, 1988).

In addition to the expected increases in the geographic extent of debris flow susceptible areas, population and urbanization are also expected to increase in magnitude and geographic extent, leading to an increased likelihood of intersection of debris flow susceptibility and humans. A debris flow susceptibility model, by itself, has marginal value. It is the use of the model that determines its utility and significance. The ultimate use of any hazard susceptibility model is in determining the potential impact on humans and/or economies. When combined with economic, population, and transportation data, a global debris flow susceptibility map provides an international perspective on prioritizing further regional or local investigation and mitigation strategies vis-à-vis those areas with a higher potential risk to humans and economies. To that end, population density is juxtaposed with debris flow susceptibility classifications, resulting in "areas of interest" for human vulnerability.

1.5 Antecedent Wildfire Probability as a Potential Susceptibility Contributor

Frequently debris flows are generated immediately post-fire from runoff-dominated progressive bulking of storm runoff with sediment eroded from hillslopes and channels, and to a lesser extent by infiltration-triggered failure and mobilization. In such areas, slope failures are in response to prolonged periods of rainfall, or prolonged rain fall in combination with rapid snowmelt (S. H. Cannon et al., 2009; Parise & Cannon, 2011). With climate change and increasing temperatures, the likelihood of increased forest fires leads to an increased likelihood in post-fire debris flow frequency, in those burn areas where other debris flow predisposing factors exist (S. H. Cannon & DeGraff, 2009; S. H. Cannon et al., 2009); and rainfall amounts required to induce debris flows decrease (Moody, Shakesby, Robichaud, Cannon, & Martin, 2013; Parise & Cannon, 2011). Given that debris flows may occur during the first rainy season post-fire, one to two years post-fire, and even up to 10 to 30 years (Parise & Cannon, 2011), burned watersheds pose a significant potential threat to humans living in close proximity (De Graff, 2018). Thus, present-day and future wildfire probability models from Moritz et al. (Moritz et al., 2012) are coupled with the debris flow susceptibility models of this project, as a potential debris flow susceptibility amplification.

1.6 Glacier and Permafrost Melting as a Potential Contributor

While there are numerous factors which contribute to landscape instability and landslide susceptibility, such as precipitation, snow melt, temperature, seismic and volcanic activity, and anthropogenic changes to landscape (Gariano & Guzzetti, 2016), precipitation and snow melt are the most significant triggers for debris flows. In mountainous regions, which are most susceptible to slope instabilities, increased temperatures can result in thawing of permafrost and interstitial ice in rock mass cracks, reducing the shear strength, and increasing the frequency and magnitude of rock falls and debris flows (C. Huggel et al., 2010). Glacier-related debris flows may occur during glacier retreat due to the availability and exposure of large quantities of unconsolidated sediments, which are mobilized with warming trends, increased precipitation, snowmelt, and glacial lake outbursts as triggering mechanisms (Marta Chiarle, Iannotti, Mortara, & Deline, 2007). The proximity of glaciers to populations within debris flow susceptible areas is modeled as an additional proxy to population vulnerability.

1.7 Approach

To pursue the above objectives, it is first necessary to discern the variations from continent to continent, to understand the impact of data uncertainty, and to utilize this information to seek the impact on humans. Thus, this research is composed of five studies; a phased approach, where each study builds upon the previous:

Study 1: A single global susceptibility model versus a continental approach to global analysis

Study 2: Impact of spatial data uncertainty in debris flow susceptibility analysis

Study 3: Future debris flow susceptibility and human vulnerability

Study 4: Fire probability as an augmentation to debris flow predisposing factors

Study 5: Future glacial melting as a potential amplifier to debris flow precipitation triggers

A global susceptibility model is not a final result but rather a depiction of areas of interest for further, more detailed investigation. It offers an international focus for understanding the impact on future urbanization and migration trends with the potential to divert such growth into areas less susceptible to such debilitating natural hazards. Landslide susceptibility maps have intrinsic value as a natural hazard risk management tool, and may be meaningful for other applications (Broeckx et al., 2018). The methodologies employed in this project can be appropriately applied to other types of mass wasting.

2 State of the Art

Several researchers have approached the landslide susceptibility problem from a global perspective. There are significant differences between their works and the work performed herein, differences in modeling techniques, predisposing environmental factors considered, with the largest departure in the landslide types being modeled. My work is focused exclusively on debris flows while other researchers use inventories inclusive of all landslide types, or filters on the triggering mechanism, such as rainfall, which in many inventories will include landslide types other than debris flow. Most researchers focus on a smaller select set of environmental factors such as topographic index, precipitation, and soil moisture. A few include ground peak acceleration (GPA) as a proxy for lithologic/slope stability. Many statistical approaches are used. Three of the five research efforts described below use logistic regression. All of these works, regardless of modeling technique, include contrived landslide absence data.

Felsberg et al (Felsberg, Poesen, Bechtold, Vanmaercke, & De Lannoy, 2021) model landslide susceptibility at a coarse spatial resolution of 36 km using predominately satellite soil moisture observations correlated with hydrologically triggered landslide data from the NASA Global Landslide Catalog, the same inventory used herein. Using a hydrologic trigger attribute, as the event filter, will result in the inclusion of a number of different landslide types such as "complex", "creep", "debris flow", "earth flow", "landslide", "lahar", "mudslide", "riverbank collapse", "rock fall", "snow avalanche", "translational slide", and "topple". Felsberg et al use logistic regression which requires presence-absence data; thus, they contrive absences by developing buffered distances around each of the known historical landslide events and define them as "absence" cells.

Similar to Felsberg et al, Kirschbaum and Stanley (Kirschbaum & Stanley, 2018) modeled global susceptibility for landslides based on an attribute of "precipitation-triggered" . In a review of the source database, their 4930 modeled landslide events (which include only those events with a known location accuracy of 10 km or better) could include many landslide types as described above. Five environmental factors (slope, geology, road network, fault zones, and forest loss) are used in concert with satellite-based precipitation estimates to predict potential landslides and generate near real-time hazard warnings.

Lin et al (Lin et al., 2017) use logistic regression with topography (relative relief), extreme precipitation, and lithology environmental predisposing factors, with two landslide "triggers", ground motion, and soil moisture. Their study excludes Greenland and Antarctica. Similar to Felsberg et al (Felsberg et al., 2021) they randomly generate landslide absence cells beyond a buffer zone for each known landslide event. Their presence events include debris slides, rotational slides, slumps, and debris flows.

Using Support Vector Machines (SVM) machine learning, Farahmand et al (Farahmand & AghaKouchak, 2013) develop a "quasi-global" model covering 60^{o} S to 60^{o} N latitude. Using only satellite precipitation data, landuse/landcover, and topographic index with an early version of the NASA GLC inventory for presence data, along with contrived absence cells. They filter the inventory to exclude landslides with slopes and topography index near zero (below 10%), but do not filter on landslide types. Their resulting model provides a landslide event *prediction* rather than landslide *susceptibility* and is geared toward landslide monitoring and potential warning systems.

Hong et al (Hong et al., 2007) use a qualitative and subjective approach in employing six environmental factors (slope, soil type, soil texture, elevation, landcover, and drainage density). Based on results found in literature review, they assign factor weights (from zero to one) accordingly. These weights were then summed over each pixel, with higher sums equating to higher susceptibility. Their model was not trained nor validated against historical or current event data.

Nadim et al (Nadim et al., 2006) developed a global susceptibility model for "slides with rapid mass movement like rock-slides, debris flows, snow avalanches, and rainfall and earthquake induced slides". Their model was based on slope, lithology, soil moisture, precipitation, and seismic conditions Similar to the study herein, they used evapotranspiration data (defined by the aridity factor in my study) as a proxy for antecedent soil moisture. The validation of their global map was limited based on their access to historical event data, and validation was performed only for six countries (Norway, Armenia, Georgia, Nepal, Sri Lanka, and Jamaica). Similar to Hong et al, their determinations of factor influence on susceptibility were subjective, assigning a priori weighted values from "very low" to "very high" to each factor class.

The principal differences in the works described above, and my own, lies in three major categories: 1) their use of 'contrived' absence cells and presence-absence statistical models, 2) the use of a minimal set of environmental factors, and 3) inclusions of landslide types other than debris flow. The research performed herein is decidedly a "presence-only" methodology as discussed in Chapter 7, exclusively for debris flows, using the fourteen environmental predisposing factors most commonly associated with debris flow analyses (Chapter 7).

3 Research Objectives

The overall intent of this project is to develop objective models and processes which can establish debris flow susceptibility at continental and global scales. The outcome of this research are models identifying those areas across the world which have various levels of susceptibility to debris flows based on the empirical factors associated with ~8,000 debris flow events; and then projecting into the future (next 30-50 years) where these susceptible areas may occur, based on climate change and other factors. In this context, susceptibility does not represent any degree of probability, that is, the likelihood or chance of a debris flow occurring, but rather that conditions exist which have been empirically and statistically associated with previous debris flows. Many researchers associate landslide susceptibility with landslide predictability (Guzzetti, Reichenbach, Ardizzone, Cardinali, & Galli, 2006). The use of the term susceptibility in this study, does not imply "predictability". According to the Merriam-Webster Dictionary (*Merriam-Webster.com Dictionary*, Merriam-Webster, https://www.merriam-webster.com/dictionary/predict. Accessed 21 Apr. 2020.), predictable is "to be able to be known, seen, or declared in advance", that is, to say that an event or action will happen in the future. Whereas, susceptible is "to be capable of submitting to an action, process, or operation". The objective is to identify the conditions and processes that promote landscape instability, as it relates to the potential subsequent occurrence of a debris flow and map those locations today and into the future. A priori debris flow events are utilized to train the model by defining "known" susceptibility areas. The model then applied to the entire Earth can identify areas that show "potential". This project is soundly based on the classic premise posited by James Hutton, that the past and present are keys to the future.

Many susceptibility studies have been performed generally at a regional or localized scale. Many studies have been performed on the susceptibility of "landslides" in general. Because conditioning factors differ for the various landslide typologies, this study is specific to debris flows and the conditioning and trigger factors associated with them. The methodologies and processes described herein can be repeated for other landslide types using those conditioning factors specific to those landslide types, resulting in susceptibility maps for those particular landslide typologies. "Following the Uniformitarianism principle, the past and the present are

considered keys to the future, hence it is assumed that future landslides are more likely to occur under the same geologic and geomorphologic conditions that led to past slope instability (Guzzetti et al., 2005; Varnes,1984).”

4 Characterizations, Assumptions, Illuminations

4.1 Debris Flows

According to Iverson (Iverson, 1997) there are several questions in debris flow assessments: (1) where, and (2) when will the debris flow occur, (3) how large will it be, (4) how fast will it travel, and (5) what areas will be impacted downstream. This study attempts to address the question of "where" and "what".

Debris flows are poorly sorted viscous, water-saturated masses with greater than 50% volume of sediment, clasts, and debris. Water content is typically 10-20% by weight. They exhibit high pore pressure and behave as non-newtonian fluids, with both solid and fluid forces at play. Debris flows are characterized by fast, typically unpredicted movements which can exceed 10 m/s with volumes ranging from $10^3 m^3$ to greater than $10^9 m^3$ (Iverson, 1997; Schippa, 2018; Takahashi, 2014; Varnes, 1978; Varnes & IAEG Commission on Landslides, 1984; Whipple & Dunne, 1992).

A clear distinction is not always made by researchers, and those reporting on debris flows, in identifying a debris flow versus an earth-flow, mudflow, mudslide, or lahar (Iverson, 1997) hence landslide inventories typically under-report debris flows, as such.

"It is well known that debris flows are triggered by high-intensity rainstorms, but their spatial distribution is not random. Lithology, altitude, aspect, plant cover, and land use play important roles" (Adrián Lorente, García-Ruiz, Beguería, & Arnáez, 2002), and citations within. Debris flows are best explained by slope inclination class, altitude, soil class, distance to drainage network and surface geology (Grozavu, Pleşcan, Patriche, Mărgărint, & Roşca, 2013). They may occur as hillslope or channelized forms or a combination, dependent on the topography and geology of the area (Nettleton, Martin, Hencher, & Moore, 2005).

4.2 Susceptibility

Susceptibility, as used in the context of this research, represents potentiality. Potential as defined in the Oxford Dictionary "having or showing the capacity to become or develop into something in the future". It does not represent the "likelihood" or "probability".

There are several factors associated with the susceptibility of one site over another for the predisposition of debris flows. Slopes, availability of mobilizable materials, permeability and density of soils, recent denudation, anthropogenic (farming, roads, excavation, mining, etc.), precipitation, and location of tectonics (A. Lorente, Garcia-Ruiz, Beguería, & Arnaez, 2002; Lupiano, Rago, Terranova, & Iovine, 2019; Yang, Han, & Li, 2017). Several of these factors are directly related to climate. Elevation can be a factor but not in all situations when additional factors such as anthropogenic and tectonic factors are prevalent (Cheng, Wang, Zhao, & Zhao, 2016; Lonigro, Gentile, & Polemio, 2015; Skilodimou, Bathrellos, Koskeridou, Soukis, & Rozos, 2018). Presence or absence of former landslides is one of the most important indicators of susceptibility - today (Varnes & IAEG Commission on Landslides, 1984). However due to changing, expanding, climate zones conducive to debris flows, there will most certainly be debris flows in areas that heretofore have not exhibited them.

There is a complex undefined interplay in the predisposing conditions. No single factor can definitively determine the initiation or location of any/all debris flows. Slope angle alone, for example, will not determine a locations susceptibility to debris flows. However, in combination with other predisposing factors such as precipitation, soil, lithology, elevation, landform, a stronger correlation may be made. In addition to current conditions, antecedent conditions and events are important contributory criteria (Crozier, Glade, & Preston, 2013) however due to the temporality of antecedent conditions and the geographic scope of this project, they are not considered.

The spatial unit for debris flow susceptibility utilized throughout this project is the grid cell.

4.3 Assumptions and Illuminations

As is true in any research, there are fundamental assumptions underlying every part of this research and methodology. The following assumptions are the most important with respect to this study.

1. A continental and global model is a generalized model.
2. Models of the real world are necessarily fuzzy.

3. The relationship between debris flows and environmental conditional factors is not necessarily linear.
4. No one debris flow environmental factor is independent of any other environmental factor.
5. Statistical models are “best guesses”, “likelihoods”, “probabilities” based on the limited and fuzzy information available. They are not proof.

5 Data

5.1 Historical Debris Flow Event Data

Two datasets are used for validation and verification throughout this project. A landside inventory was sourced from NASA (D. Kirschbaum, T. Stanley, & Y. Zhou, 2015) for validation data, herein referred to as TRAIN. This inventory contains 11,033 landslides of various types and locational uncertainties. 194 events are classified as debris flow and 2100 as mudslide, herein collectively referred to as "debris flow". "Mudslide" type was included in this study as it is a common misnomer for debris flow. It is noteworthy that landslide classification, itself, may be a source of data uncertainty.

The verification dataset (TEST1) was curated from many agencies across the world, resulting in 5695 debris flow events. The geographic distribution of these events is not as well distributed across the world as would be desired, due to either an inability to locate data sources or inability to obtain inventories from the sources identified. The numeric breakdown of debris flow events by country and continent are presented in Tables 5.1 and 5.2, respectively. The global distribution of TRAIN and TEST1 events are shown in Fig 5.1.

The TRAIN dataset includes a "location accuracy" attribute for each event, which is used in subsequent analyses of the effect of locational uncertainty on debris flow susceptibility. The locational accuracy (or uncertainty) of the global inventory ranges from "exact" to 250 kilometers. The locational accuracies associated with the ninety-nine debris flows in Europe range from an "exact" known location to 50 kilometers (Table 5.3), plus six events which were identified with a location accuracy of "unknown".

Country	Count	% Total	Country	Count	% Total
Afghanistan	3	0.1%	Kyrgyzstan	32	1.4%
Argentina	6	0.3%	Lebanon	1	0.0%
Australia	16	0.7%	Macedonia	2	0.1%
Austria	3	0.1%	Madagascar	2	0.1%
Azerbaijan	1	0.0%	Malaysia	22	1.0%
Bangladesh	9	0.4%	Mexico	38	1.7%
Barbados	1	0.0%	Mongolia	1	0.0%
Belize	1	0.0%	Myanmar [Burma]	4	0.2%
Bhutan	2	0.1%	Namibia	2	0.1%
Bolivia	4	0.2%	Nepal	23	1.0%
Bosnia and Herzegovina	1	0.0%	New Zealand	13	0.6%
Brazil	56	2.4%	Nicaragua	3	0.1%
Bulgaria	1	0.0%	Nigeria	9	0.4%
Burkina Faso	1	0.0%	Norway	8	0.3%
Canada	102	4.4%	Pakistan	8	0.3%
Chile	5	0.2%	Panama	6	0.3%
China	86	3.7%	Papua New Guinea	6	0.3%
Colombia	25	1.1%	Peru	35	1.5%
Costa Rica	8	0.3%	Philippines	32	1.4%
Democratic Republic of the Congo	2	0.1%	Portugal	2	0.1%
Dominica	3	0.1%	Russia	7	0.3%
Dominican Republic	4	0.2%	Rwanda	1	0.0%
Ecuador	9	0.4%	Saudi Arabia	1	0.0%
El Salvador	5	0.2%	Sierra Leone	4	0.2%
Fiji	6	0.3%	Singapore	1	0.0%
France	6	0.3%	Slovenia	1	0.0%
Germany	1	0.0%	South Africa	13	0.6%
Georgia	5	0.2%	South Korea	4	0.2%
Guatemala	18	0.8%	Spain	5	0.2%
Haiti	8	0.3%	Sri Lanka	6	0.3%
Honduras	4	0.2%	Switzerland	11	0.5%
Iceland	1	0.0%	Taiwan	19	0.8%
India	97	4.2%	Tajikistan	17	0.7%
Indonesia	17	0.7%	Thailand	25	1.1%
Iran	1	0.0%	Trinidad and Tobago	14	0.6%
Ireland	7	0.3%	Turkey	3	0.1%
Israel	1	0.0%	U.S. Virgin Islands	1	0.0%
Italy	19	0.8%	Uganda	5	0.2%
Ivory Coast	1	0.0%	Ukraine	3	0.1%
Jamaica	5	0.2%	United Kingdom	23	1.0%
Japan	33	1.4%	United States	1270	55.4%
Kazakhstan	2	0.1%	Venezuela	6	0.3%
Kenya	12	0.5%	Vietnam	7	0.3%

Table 5.1 TRAIN distribution by country

Continent	Approx. Land Area (sq km)	#TRAIN events	Continental TRAIN events as % of Total Events	#TEST1 events	Continental TEST1 events as % of Total Events	TRAIN density (#events/sq km)	TEST1 density (#events/sq km)
Africa	30,001,151	54	2.4	16	0.28	0.0000018	0.0000005
Asia	44,851,729	469	20.4	554	9.7	0.0000105	0.0000124
Australia	7,701,651	16	0.69	1	0.018	0.0000021	0.0000001
Europe	9,898,597	99	4.3	2743	48.1	0.0000100	0.0002771
North America	24,189,365	1473	64.2	2370	41.6	0.0000609	0.0000980
Oceania	429,108	19	0.83	5	0.088	0.0000443	0.0000117
South America	17,757,691	164	7.1	6	0.11	0.0000092	0.0000003
TOTAL	147,144,240	2294		5695		0.0000156	0.0000387

Table 5.2 Continental distribution of TRAIN and TEST1 debris flow events.

Europe			Global		
Uncertainty	Frequency	%	Uncertainty	Frequency	%
Exact	5	5.1%	Exact	174	7.6%
1km	19	19.2%	1km	620	27.0%
5km	31	31.3%	5km	763	33.3%
10km	20	20.2%	10km	277	12.1%
25km	12	12.1%	25km	240	10.5%
50km	6	6.1%	50km	125	5.4%
100km	0	0.0%	100km	9	0.4%
250km	0	0.0%	250km	4	0.2%
unknown	6	6.1%	unknown	82	3.6%
TOTAL	**99**		**TOTAL**	**2294**	

Table 5.3 Locational uncertainty associated with 99 Europe and 2294 global debris flows

There may be sampling bias in this inventory, due to landslide data being commonly collected where the data are easier to gather near cities/towns/villages, where there is known potential danger to facilities and/or people, and at areas more easily accessible for observation or field survey. This bias, if it does exist, may be common among most, if not all, landslide inventories (Reichenbach et al., 2018). An additional component to the sampling bias is a result of the unequal availability of data from every country. A more equitable (representational) geographic distribution of data acquisition across the world would result in more finely tuned continental and global models.

A comprehensive understanding and characterization of the debris flow environments in both TRAIN and TEST1, individually and with respect to one another, was sought. Therefore prior to modeling, numerous evaluations were performed with respect to the frequency of events associated with each predisposing environmental factor, both globally and per continent. These underlying characteristics can be found in Appendix A.

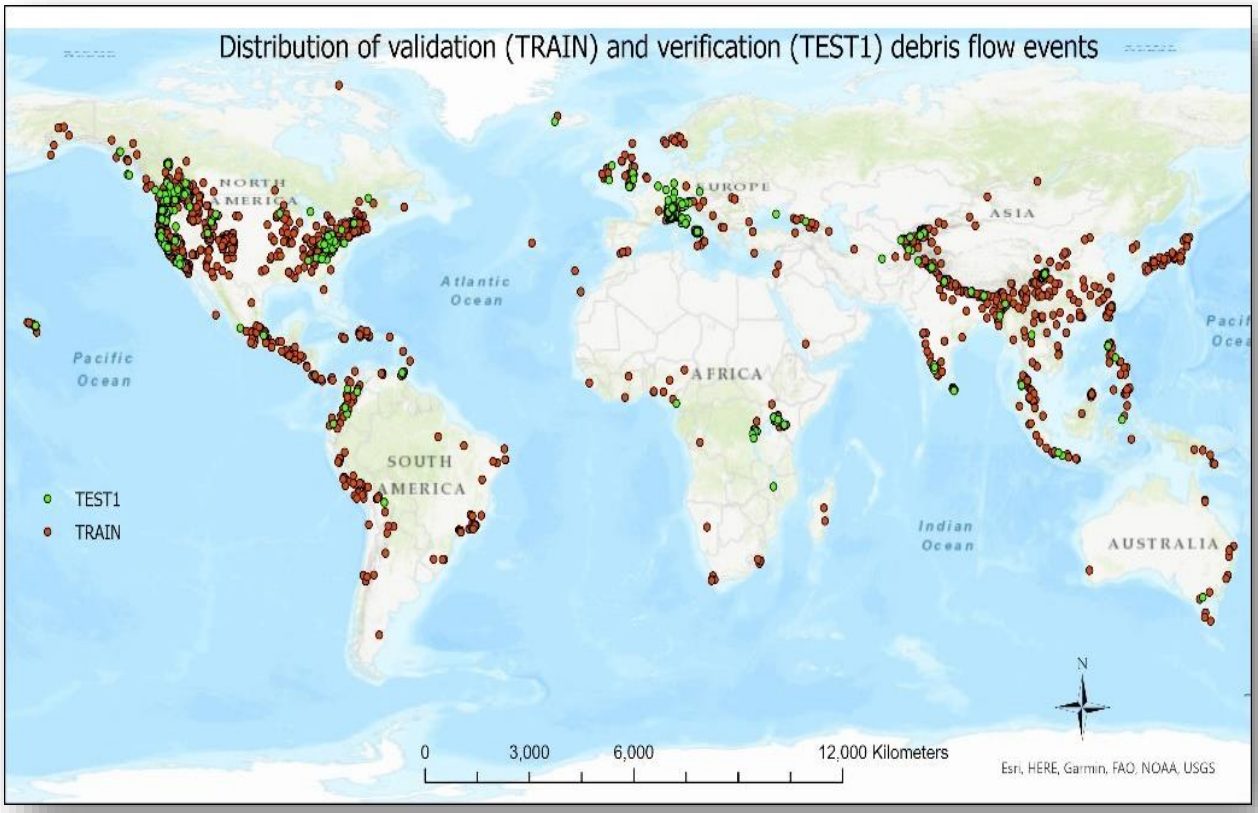


Figure 5.1 Global distribution of validation (TRAIN) in red and verification (TEST1) debris flow events in green

5.2 Predisposing Environmental Data

Environmental variables commonly identified as debris flow predisposing factors across the world (Capitani, Ribolini, & Bini, 2013; N. S. Chen et al., 2010; Corominas et al., 2013; D'Amato Avanzi, Giannecchini, & Puccinelli, 2004; Devkota et al., 2012; Diop, 2012; Dou et al.,

2015; Ferentinou & Chalkias, 2011; Ghosh, Carranza, van Westen, Jetten, & Bhattacharya, 2011; Grozavu & Patriche, 2013; Kirschbaum & Stanley, 2018; Kornejady, Ownegh, & Bahremand, 2017; Adrián Lorente et al., 2002; Meten, PrakashBhandary, & Yatabe, 2015; Nsengiyumva, Luo, Nahayo, Huang, & Cai, 2018) are used in various aspects of the global and continental debris flow susceptibility analyses. They include aridity, climate classification, depth to bedrock, distance to faults, distance to rivers, distance to sea, elevation, fault density, landcover, landform, lithology, precipitation (average monthly), slope, soil drainage, soil thickness, soil type, topsoil clay %, closest water body type, closest active fault type, and closest tectonic boundary type. Factor data information is additionally listed in Appendix B.

While additional factors are relevant, such as slope aspect, and distance to intermittent and perennial rivers, they were difficult to process and summarize at a global level and are therefore not included.

5.2.1 Climate, Aridity, and Precipitation

Climate variables are critical attributes with respect to both predisposing and triggering factors associated with debris flows (C. Huggel et al., 2010; Lonigro et al., 2015; Moreiras et al., 2021; Oakley, 2018; Oakley & Ralph, 2018; Paranunzio et al., 2018). Köppen-Geiger climate classification (Kottek, Grieser, Rudolf, & Rubel, 2006), aridity, and average monthly mean precipitation are used to provide a representation of temperature and precipitation environmental conditions. Antecedent precipitation may also act as a proxy for soil moisture conditions prior to slope destabilization (Nadim et al., 2006). The data are acquired for present day (1980-2016) and future (2071-2100) projection scenarios Representative Concentration Pathway (RCP) 2.6 and RCP 8.5.

5.2.1.1 Climate Classification

Köppen-Geiger (KG) classification was originally developed for biome distributions. Today it is one of the most widely used classification schemes for many different types of studies concerned with climate change impact analyses (Beck et al., 2018). The present day and RCP 8.5 KG data were acquired from Beck et al (Beck et al., 2018) in 1-km raster format. The KG RCP 2.6 1-km data were acquired from Cui, et al (Cui, Liang, Wang, & Liu, 2021).

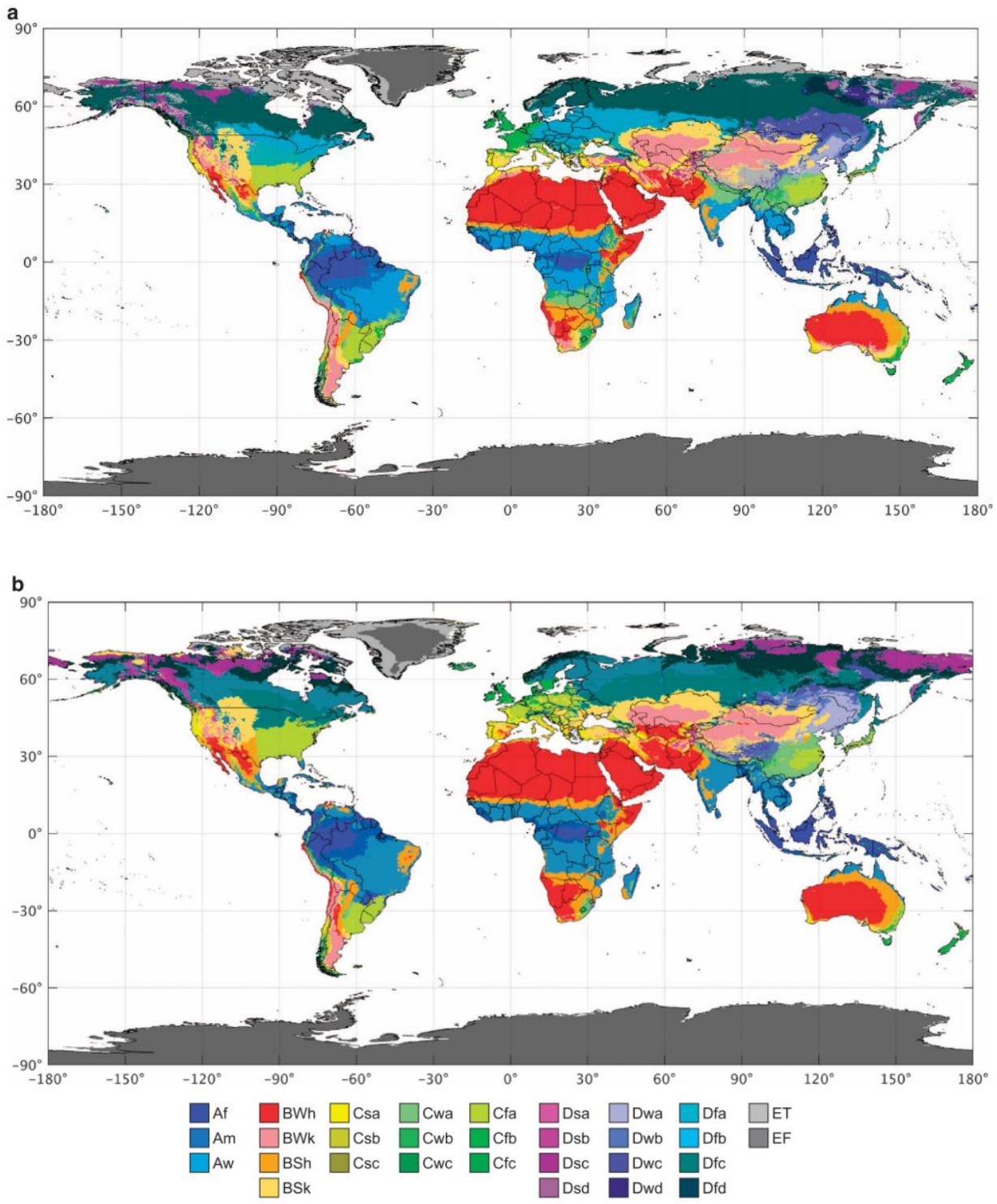


Figure 5.2 (a) Present day (1980-2016) Köppen-Geiger map and (b) future (2071-2100) from Beck et al (Beck et al., 2018)

The scheme uses a three-letter designation, dividing climates into five main climate groups and thirty sub-types. The main climate group is represented by the first letter: A (tropical), B (dry), C (temperate), D (continental), and E (polar). The second letter indicates the seasonal precipitation type, and the third letter indicates the qualitative level of temperature, Table 5.4. Summer is defined as the warmest six-month period (April to September or October to March) and winter is the six-month coolest period.

1st Letter Main climate group	2nd Letter Seasonal precipitation	3rd Letter Qualitative temperature
A - Tropical	f – rainforest m – monsoon w – Savanna, dry winter s – Savanna, dry summer	
B - Arid	W – Desert S - Steppe	h – hot k – cold
C - Temperate	w – Dry winter f – No dry season s – Dry summer	a – Hot summer b – Warm summer c – Cold summer
D – Continental	w – Dry winter f – No dry season s – Dry summer	a (Hot summer) b (Warm summer) c (Cold summer) d (Very cold winter)
E – Polar	T – Tundra F – Eternal frost (ice cap)	

Table 5.4 Köppen-Geiger classification

5.2.1.2 Precipitation

Both current (1980-2017) and RCP 2.6 and RCP 8.5 future (2081-2100) total monthly precipitation values (mm), averaged over twenty years, were obtained from Worldclim.org (Fick & Hijmans, 2017). Data were subsequently summed over the twelve months and divided by 12 to obtain a monthly average of current and future precipitation values. Data are in Geotiff format, with resolution approximately 2 arcminutes (~ 4.5 km at the equator).

5.2.1.3 Aridity

Global aridity index data (Trabucco & Zomer, 2018) for the period 1970-2000 are used as an additional proxy for climate as a debris flow conditioning factor (Fig. 5.3). Aridity is an augmentation to precipitation and climate classification as it is related to the amount of precipitation available after evapo-transpiration. Many debris flow researchers have determined that low intensity, high frequency precipitation over a period of time may be more closely associated with debris flow susceptibility than a single high intensity precipitation event. An aridity index can be used to quantify precipitation availability.

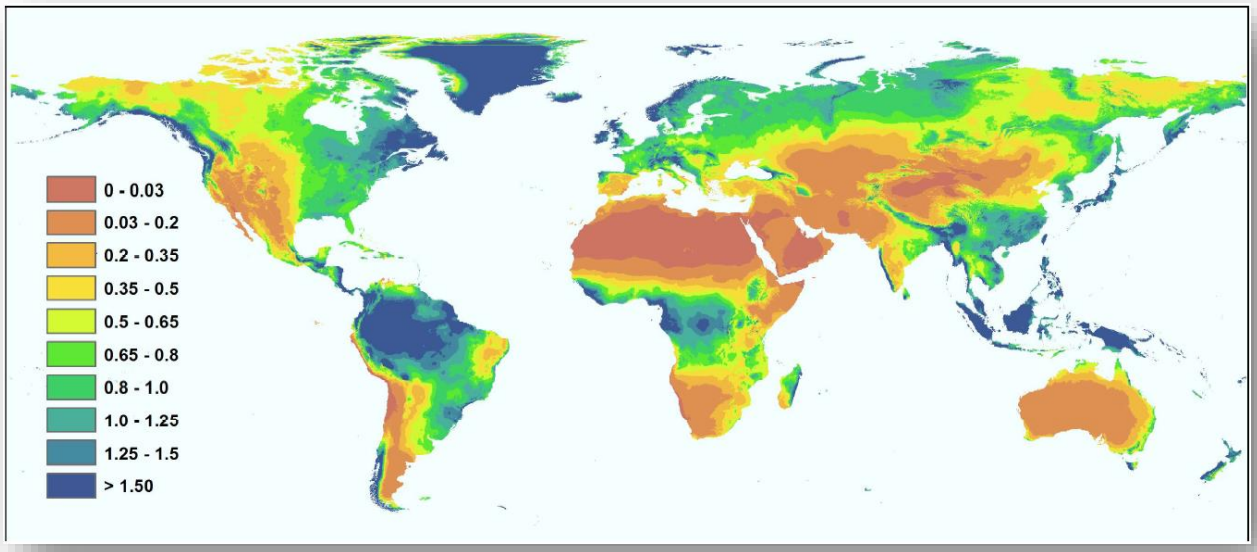


Figure 5.3 Global Aridity Index (1970-2000). Higher values represent more humid conditions, with low (brown/yellow) representing higher aridity

For subsequent analytical processing purposes, this raster dataset with floating point values, was converted to integer format. The INT function of ArcGIS truncates floating point values, therefore a multiplication by 10,000 was performed on each pixel value prior to integer conversion. Table 5.5 presents the original Aridity Index (AI) values, the generalized climate classification scheme, and the calculated values utilized in this project.

Aridity Index Value	Climate Class	Project Derived AI Value
< .03	Hyper Arid	< 300
0.03 – 0.2	Arid	300 - 2000
0.2 – 0.5	Semi-Arid	2000 – 5000
0.5 – 0.65	Dry sub-humid	5000 - 6500
>0.65	Humid	> 6500

Table 5.5 Source data Aridity Index value, classification, and projected derived values

The data have a 30 arcsecond resolution (1 kilometer at the equator, 700 meters at 45N\45S degrees latitude, and 500 meters at 60N\60S degrees latitude), raster format.

5.2.2 Depth to Bedrock

Depth to bedrock provides a view on the amount (depth) of soil column or regolith at the surface as well as a potential indication of ground water and bedrock boundaries (Shangguan, Hengl, Mendes de Jesus, Yuan, & Dai, 2017), all of which may play a role in landscape instability and debris flow susceptibility. Global depth to bedrock data (Pelletier et al., 2016; Shangguan et al., 2017) (Fig. 5.4) was obtained at 30 arcsecond (~1km at the equator) resolution.

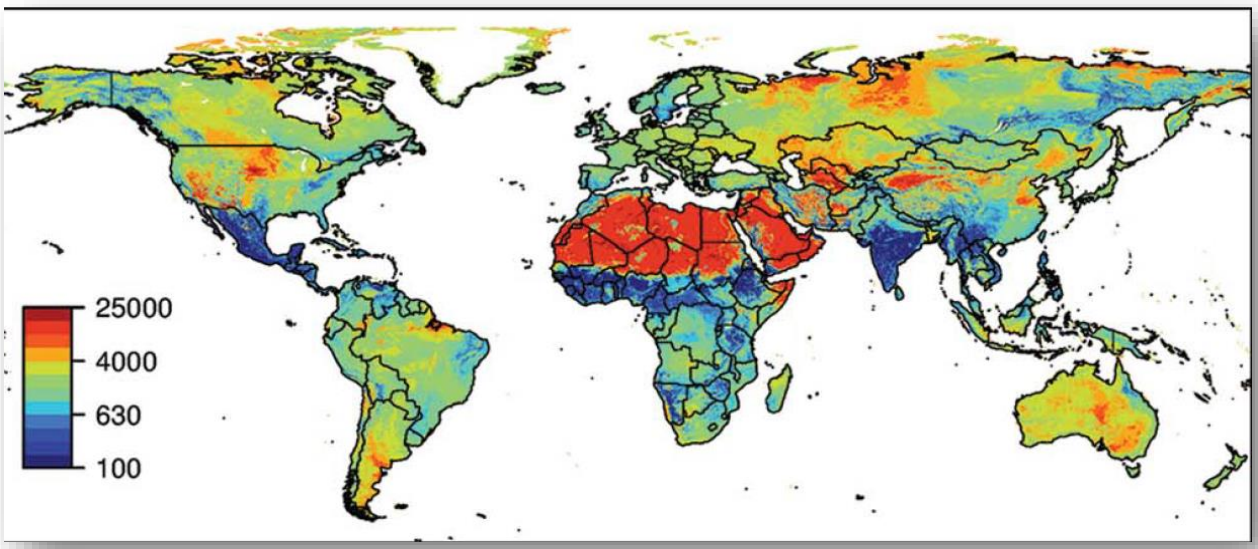


Figure 5.4 Global absolute depth to bedrock (cm) from Shangguan et al **(Shangguan et al., 2017)**

5.2.3 Hydrology

The World Water Bodies polygon dataset of open water rivers, lakes, seas, and oceans of the world, developed by Esri, Inc. (Esri, 2015) was obtained from the University of California Los Angeles (UCLA) for Digital Research and Education (IDRE) geoportal (https://apps.gis.ucla.edu/geodata/dataset/world_water_bodies). The data are provided at 30 arcsecond (about 1km at the equator) resolution. ArcGIS tools "Near", and "Buffer" were used to further process these data to calculate distance to water bodies, and closest water body type for each debris flow event.

5.2.4 Lithology

Lithology plays a key role in landscape instability and debris flow susceptibility. A global lithology raster data set (GLiM) (Fig. 5.5) representing fifteen high level classifications of rock type at the Earth's surface was acquired from Hartmann et al (J. Hartmann & Nils Moosdorf, 2012) at resolution of .5 degrees (about 50 km at the equator).

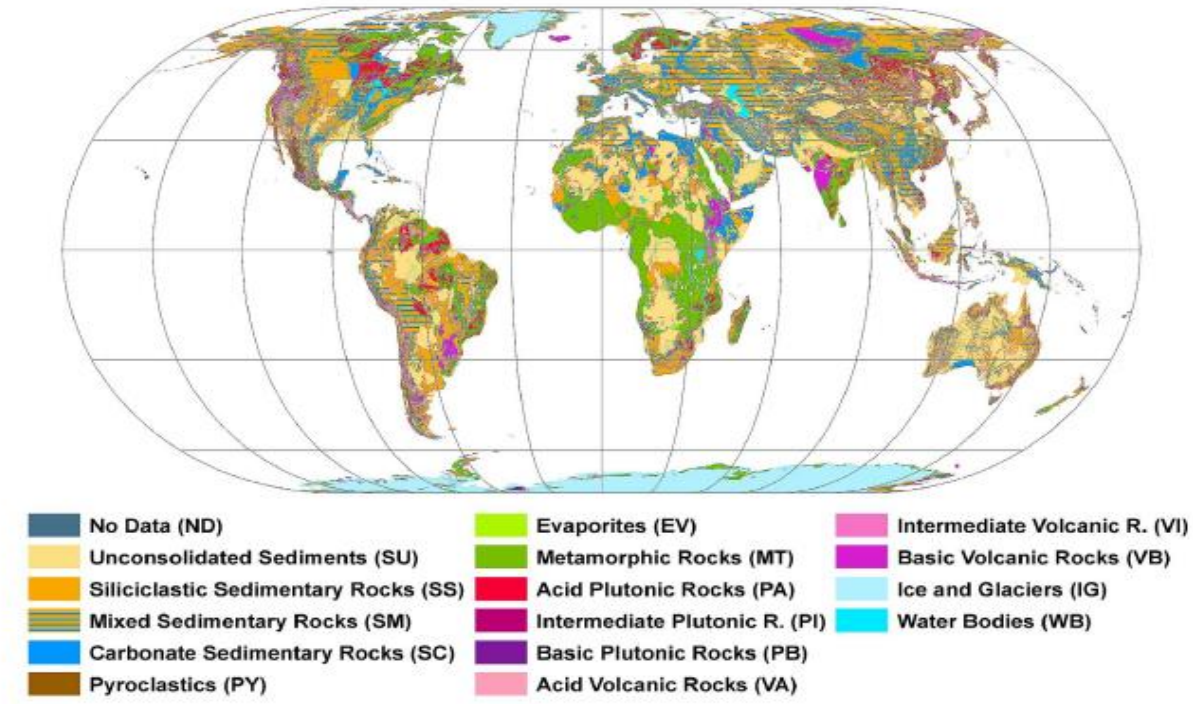


Figure 5.5 Representation of global lithology and basic lithological classes from Hartmann et al (J. Hartmann & Nils Moosdorf, 2012)

5.2.5 Landcover

Landcover is the observed (bio)physical cover of the Earth's surface (Jansen & Gregorio, 2000). Global landcover data (GLCNMO) (Kobayashi et al., 2017) at a resolution of 15 arcseconds (approximately 500 m at the equator) has twenty land cover classes (Table 5.6). For purposes of this study, grouping of certain classifications improves the spatial analysis at a global level, minimizing a "loss of information" when too many classifications occur at too small a scale.

Landcover Code	Grouping Used in this Project	GLCNMO Landcover Class	Label
1	**Forest and Woodland**	Broadleaf Evergreen Forest	Broadleaved Evergreen Closed to Open (100-40%) Trees
2	**Forest and Woodland**	Broadleaf Deciduous Forest	Broadleaved Deciduous Closed to Open (100-40%) Trees
3	**Forest and Woodland**	Needleleaf Evergreen Forest	Needleleaved Evergreen Closed to Open (100-40%) Trees
4	**Forest and Woodland**	Needleleaf Deciduous Forest	Needleleaved Deciduous Closed to Open (100-40%) Trees
5	**Forest and Woodland**	Mixed Forest	Broadleaved Closed to Open Trees and Needleleaved Closed to Open (100-40%) Trees
6	**Forest and Woodland**	Tree Open	Open (40-(20-10)%) Trees (Woodland)
7	**Grassland/Shrubland**	Shrub	Closed to Open Shrubland (Thicket)
8	**Grassland/Shrubland**	Herbaceous	Closed to Open Herbaceous Vegetation, Single Layer
9	**Grassland/Shrubland**	Herbaceous with Sparse Tree / Shrub	Closed to Open Herbaceous Vegetation with Trees and Shrubs
10	**Grassland/Shrubland**	Sparse Vegetation	Sparse Herbaceous Vegetation // Sparse Woody Vegetation
11	**Cropland**	Cropland	Herbaceous Crop(s)
12	**Grassland/Shrubland**	Paddy field	Graminoid Crops
13	**Cropland**	Cropland / Other Vegetation Mosaic	Cultivated and Managed Terrestrial Area(s), and Natural and Semi-Natural Primarily Terrestrial Vegetation //Cult ivated Aquatic or Regularly Flooded Area(s), and Natural and Semi-Natural Primarily Terrestrial Vegetation
14	**Wetlands/Waterbodies**	Mangrove	Closed to Open Woody Vegetation with Water Quality: Saline Water
15	**Wetlands/Waterbodies**	Wetland	Closed to Open Woody Vegetation with Water Quality: Fresh Water //Closed to Open Woody Vegetation with Water Quality: Brackish Water // Closed to Open Herbaceous Vegetation with Water
16	**Bare Ground**	Bare area, consolidated (gravel, rock)	Consolidated Material(s)
17	**Bare Ground**	Bare area, unconsolidated (sand)	Unconsolidated Material(s)
18		Urban	Artificial Surfaces and Associated Area(s)
19		Snow / Ice	Perennial Snow // Perennial Ice
20	**Wetlands/Waterbodies**	Water bodies	Artificial Waterbodies // Natural Waterbodies

Table 5.6 GLCNMO landcover classification from Kobayashi et al (Kobayashi et al., 2017)

5.2.6 Soil

The Digital Soil Map of the World (DSMW) at a scale of 1:5,000,000 was acquired from the Food and Agriculture Organization of the United Nations (FAO) (FAO, 2003). The dataset contains numerous soil characteristics. The characteristics utilized are soil type (dominant soil classification), % clay, and soil drainage.

5.2.7 Tectonics

The GEM Global Active Faults Database (GAF-DB) (Styron & Pagani, 2020) is a database of global active faults containing approximately 13,500 faults, each with associated attributes for geometry, kinematics, and other characteristics. ArcGIS tools "Line Density", "Kernel Density", "Near", and "Buffer" were used to further process these data to calculate fault density, distance to faults, closest active fault type, and closest tectonic boundary type for each debris flow event.

5.2.8 Topography

5.2.8.1 Elevation

Elevation source data is the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) (Danielson, 2011) obtained from the United States Geological Survey (USGS). Three product suites (and resolutions) are offered. The chosen dataset is the "breakline emphasis" which is useful for hydrologic derivatives and modeling over large areas and has a resolution of 7.5 arcseconds (about 250 m at the equator). Topographic breaklines are ridges and streams. The "breakline emphasis" data "maintains the critical topographic features within a landscape by maintaining any minimum or maximum elevation value on a breakline that passes within the data preparation processing window" (Danielson, 2011). Greenland is not covered at this resolution.

The Earth is covered by fifty-nine Geotiff files which were mosaicked in ArcGIS to provide a seamless coverage for this project.

5.2.8.2 Landform

The World Named Landforms dataset obtained from ArcGIS Online (Charles Frye, Sayre, & Soller, 2017) is used in this research because of the robust information provided with their

ecophysical regions. According to Frye et al (Charlie Frye, Sayre, Soller, & Karagulle, 2018) this dataset is based on the work of Richard E. Murphy and E.M. Bridges, where Murphy identified landforms according to erosional and depositional processes, and Bridges used a geomorphological approach. As a result, the ecophysical regions provide a combination of bioclimatic, elevation (qualitative), lithologic, and landform information in the landform descriptions (Sayre, Dangermond, Frye, & Vaughan, 2014). The dataset is a feature layer (polygon) at 250 m resolution.

5.2.8.3 Slope

World Slope GMTED data obtained from ArcGIS Online (*World Slope GMTED*, 2020) is a global raster dataset of slopes calculated from the 7.5 arcsecond (250 m) resolution GMTED2010 elevation data.

5.3 Augmenting Environmental Influences

5.3.1 Wildfire Probability

Current (1971-2000) and future (2070-2099) global wildfire probability data (Moritz et al., 2012) at 0.5 degrees resolution (about 55 km at the equator) were acquired from MoritzFireLab.org. Data available include fire probabilities based on climate with and without vegetation variables. The fire probabilities based on climate-only variables were used in this project as Moritz et al. (Moritz et al., 2012) report that the resulting probabilities were broadly similar between the two datasets.

5.3.2 Glaciers

Glacier outlines, as they were near the beginning of the 21st century, were obtained from the Randolph Glacier Inventory (RGI Consortium, 2017) a "globally complete inventory of glacier outlines". The dataset is provided as a polygon shapefile, at .5 degrees resolution (~ 55 km at the equator). Data attributes of interest include location, area, and slope. ArcGIS tools "Buffer" and "Near" were used to further process these data and associate with debris flow events, areas of susceptibility, and populations.

5.4 Data Associated with Societal Impact Analyses

5.4.1 Population

Current (2020) and future (2100) global population raster data (GAO, 2019; B. Jones & O'Neill, 2016) at 1km resolution, were acquired from CIESIN (*Gridded Population of the World, Version 3 (GPWv3): Population Count Grid*).

5.4.2 Roads

Highway and Primary roads were extracted from the Global Roads Open Access Data Set ((CIESIN), 2013) and further processed using ArcGIS "Line Density" tool. Other (smaller) road classifications were not utilized due to the very high density of such roads and therefore, difficulty in producing meaningful distinctions and associations of roads, road densities, and debris flow susceptibilities. Smaller road classifications are meaningful for more localized impact analyses.

6 Study Area

Debris flow susceptibility analyses are performed for the entire world and for each continent. Although all continents have extensive mountain ranges, vast plains and plateaus, and complex river systems, each continent is unique in the geomorphic expression of these features, their climate, soil, and vegetation (Bridges, 2012), as depicted in Tables 6.1, 6.2, and 6.3.

Continent	Approx. Land Area (sq km)	Mean Elevation (m a.s.l.)	Mountain Belts (% of continent)	Crystalline Shield (% of continent)	Volcanic Plateaus and Plains (% of continent)	Erosional Plains (% of continent)	Depositional Plains (% of continent)
Africa	30,001,151	580	3	**37**	4	30	26
Asia	44,851,729	915	**52**	10	4	12	22
Australia*	8,130,79	305	16	**31**	1	24	28
Europe	9,898,597	300	25	13	1	30	**31**
N. America	24,189,365	610	**35**	25	4	18	18
S. America	17,757,691	550	22	**30**	2	25	21

Table 6.1 Continental geomorphic characteristics (Bridges, 2012) .

*Australia incudes Oceania in this dataset. **Bold red** emphasis indicates major characteristic for that continent

Continent	Forest & Woodland	Grassland & Shrubland	Cropland	Wetlands & Waterbodies	Bare Ground	Urban	Snow & Ice
Africa	**33.4**	31.1	9.1	1.3	25	0.1	0
Asia	**34.6**	27.4	18.3	2.5	16.4	0.4	0.4
Australia	17.9	**72.2**	8.3	1.5	0	0.1	0
Oceania	**66.7**	19.9	7.3	5.4	0.3	0.3	0.1
Europe	**43.8**	11.3	39.4	3.4	.012	1	1
N. America	**37.8**	30.0	14.5	6.2	0.3	0.6	10.6
S. America	**52.1**	26.6	15.4	3.7	1.8	0.3	0.2

Table 6.2 Continental land use/landcover characteristics.

Calculated from global GLCNMO dataset (Kobayashi et al., 2017). Bold red emphasis indicates major characteristic for that continent

Continent	Evaporites (% of continent)	Ice and Glaciers (% of continent)	Metamorphic (% of continent)	Plutonic (% of continent)	Pyroclastic (% of continent)	Sedimentary (% of continent)	Volcanic (% of continent)	Water Bodies (% of continent)
Africa	0.60	0	27.60	1.40	.0002	**65.50**	4.00	0.90
Asia	0.28	.08	6.84	8.95	0.60	**72.90**	8.90	1.26
Australia	0.72	0	2.18	4.56	0.03	**88.79**	3.60	0.11
Europe	0	0.25	9.80	7.80	.05	**77.32**	2.80	2.00
N. America	0.06	7.55	13.19	10.15	1.60	**58.81**	7.28	1.35
Oceania	0	0	10.90	3.90	3.70	**61.10**	19.30	1.10
S. America	0.30	0.10	12.60	10.60	1.40	**65.80**	8.40	0.80

Table 6.3 Continental lithology characteristics.

Calculated in ArcGIS using GLiM (Jens Hartmann & Nils Moosdorf, 2012)data. Bold red emphasis indicates major characteristic for that continent

Global and continental characterization, geographic extent (sq km), and percent coverage was developed for each factor (tectonic type, fault density, soil type, soil thickness, soil drainage, precipitation, aridity, climate classification, elevation, landcover/landuse, and landform), and factor class using, ArcGIS and Excel. These detailed tables and graphs can be found in Appendix C.

For the study on Impact of Spatial Data Uncertainty, Chapter 8, the initial focus is on the continent of Europe to illuminate the resulting discrepancies in predisposing factor classes.

7 Global versus Continental Approach to Susceptibility Analysis

7.1 Data and Methodology

7.1.1 Overview

Debris flow susceptibility models were developed for the world, and for each of seven continents (Africa, Asia, Australia, Europe, North America, Oceania, and South America). An analysis was not performed for Antarctica as there are no Antarctic debris flows in the inventories used.

Probability density (PD), conditional probability(CP), certainty factor (CF), frequency ratio (FR), and maximum entropy (MaxEnt) statistical models were developed and evaluated for best model performance using fourteen environmental factors generally accepted as the most appropriate debris flow predisposing factors (Capitani et al., 2013; N. S. Chen et al., 2010; Corominas et al., 2013; D'Amato Avanzi et al., 2004; Devkota et al., 2012; Diop, 2012; Dou et al., 2015; Ferentinou & Chalkias, 2011; Ghosh et al., 2011; Grozavu et al., 2013). They include aridity, elevation, climate, land use/landcover, landform, fault density, lithology, precipitation (average monthly), soil drainage, soil thickness, soil type, and topsoil % clay. While additional factors are relevant, such as slope aspect, they are difficult to process and summarize at a global level. Climate, average monthly precipitation, and aridity all provide a view on the antecedent weather and precipitation conditions associated with debris flows, as well as areas with potential triggering factors. Aridity (humidity) may act as a proxy for antecedent soil moisture prior to slope destabilization (Nadim et al., 2006). Global models and models for each continent were developed and evaluated against verification data, and results compared.

The first objective was to choose the best statistical approach at a global level. That approach was then used to model each of the seven continents and compare the results with the global model.

The major steps (Fig. 7.1) followed for the whole Earth, and each continent were:

1. Determine best statistical approach at the global level:
    a. All environmental factor layers are overlain in ArcGIS Pro with the historical debris flow event layer (TRAIN).
    b. Debris flow event frequency is determined for each factor and factor class.
    c. Results are input to FR, CF, PD, CP calculations defined in Excel (spreadsheets can be found in Appendix C).
    d. The environmental factors and TRAIN event layers were also input to MaxEnt software.
    e. The "most significant" factor class (individual factor value) is determined, based on the maximum values within each statistical model. The MaxEnt software determines the factor significance values during the susceptibility processing. "Most significant" refers to the factor class with the largest contribution to the statistical model.
    f. A susceptibility map for each statistical model is created in ArcGIS Pro using each factor filtered by its most significant class, for each statistical algorithm, e.g., siliciclastic (ss) is the most significant lithology factor class for the PD algorithm, whereas intermediate plutonic (pi) is the most significant factor class for FR, CP, and CF, Table 7.1. For the MAXENT model, the raster susceptibility map derived by MaxEnt is input to ArcGIS.
    g. For FR, CP, CF, and PD global susceptibility maps, the susceptibility value at each pixel is computed as a sum of the number of factor classes present and then classified into five equal intervals of Very Low, Low, Medium, High, and Very High. For the MAXENT model, susceptibility values at each pixel were computed by the MaxEnt software with susceptibility values ranging from 0 to 1 and classified into the five equal intervals.
    h. TEST1 (verification data) frequencies are computed within each susceptibility classification for each model (FR, CF, PD, CP, and MAXENT). The results are compared across models to determine the best global modeling statistical algorithm.

2. Use the best statistical algorithm to develop continental models
    a. Continental modeling followed the same processes, utilizing environmental factor and debris flow data that fall within each continent's boundaries. These are called "continental models".
    b. A global version of each continent was created by "cutting" the individual continents from the final global susceptibility model. These are called "global continental cut" models.
3. Global versus continental model comparisons were performed in two ways for each continent:
    a. A difference between each of the "global continental cut" models and corresponding "continental" models
    b. A difference between the global model and a mosaic of all continental models
4. All continental susceptibility maps were further processed, classified, and results verified using the same processes as described for global modeling in Step 1.
5. Differencing between the continental and global continental "cut" models was performed on a pixel basis using ArcGIS Pro Cell Statistics tools.
6. Area coverage is then calculated for areas where susceptibility classification in the global model is >, =, and < the continental models.

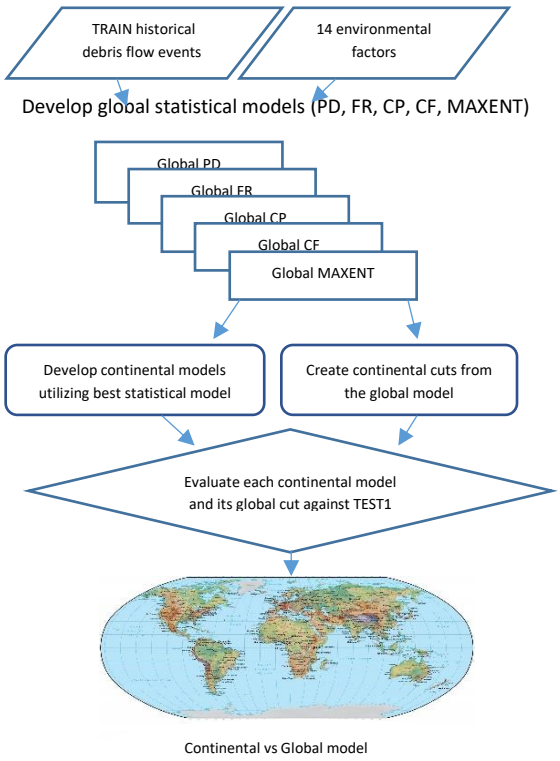


Figure 7.1 Summary overview of the continental versus global modeling process.

7.1.2 Statistical Analyses: Probability Density, Conditional Probability, Certainty Factor, Frequency Ratio

Determination of the most relevant debris flow environmental factors, and factor classes associated with an historical event inventory, is essential to developing susceptibility models (Devkota et al., 2012). Five statistical models were developed to determine the most significant factors and factor classes which would then become input to susceptibility models. Probability Density (PD), Frequency Ratio (FR), Conditional Probability (CP), and Certainty Factor (CF) algorithms were developed in Microsoft Excel (Excel) and a Maximum Entropy model (MAXENT) was developed with MaxEnt software v. 3.4.4 (S. Phillips, Dudik, & Schapire, 2021).

PD, CP, CF, and FR functions were developed for the TRAIN data and environmental predisposing factors using Microsoft Excel (Excel) to determine the most significant factor and

factor classes according to each algorithm, and to determine the best performing algorithm. The maximum entropy and conditional maximum entropy algorithms were also developed in Excel but were ultimately not used in favor of the MaxEnt software. Equations 7.1, 7.2, 7.3, and 7.4 represent the calculations for PD, CP, CF, and FR, respectively. The "most significant" factors and factor classes are those with the largest statistical value as determined by the model algorithm and represents the most significant contribution to the model outcome.

Equation	Statistic	
DF frequency in factor class/Total # DF	Probability Density (PD)	Eq 7.1
DF frequency in factor class/Factor Class area	Conditional Probability (CP)	Eq 7.2
if PPa >= PPs PPa - PPs/PPa (1-PPs) if PPa < PPs PPa - PPs/PPs (1-PPa)	Certainty Factor (CF) (where PPa = Conditional Probability, PPs = Prior Probability)	Eq 7.3
PD/PI	Frequency Ratio (where PI = Predictor Importance = Factor Class Area/Total study area)	Eq 7.4

The spreadsheet calculations and results for lithology and Köppen-Geiger climate environmental predisposing factors (circled in red) and factor classes are shown in Table 7.1. Spreadsheets for the remaining environmental predisposing factors and factor classes can be found in Appendix D. Table 7.2 is the resulting summary of factor contributions per statistical model.

"TRAIN" DATA RESULTS			Earth Total Land Area (sq km)		148,300,000	Total Earth Area	510,000,000	
			KG Total Area (sq km)		147,313,836	Hypsographic Total	140,697,700	
			Total # DF and MS		2294	Total Count of Slope M	7,878,833,570	
			Prior Probability (PPs)		0.00002	Total Count of Soil Thi	142,462,411	
Factor	Factor Class (xx field)	Factor Class Area of Earth Land Surface (sq km) Pr	Factor Class PREDICTOR IMPORTANCE (Pi) Probability Density Class Area/Total Area (Cn/Ht or F3)	Frequency DF & MS per Factor Class	Probability Density DF&MS in Factor Class/TOTAL # DF & MS	FREQUENCY RATIO (FR) Fn/Dn	Conditional Probability (PPa)	CERTAINTY FACTOR (CF) if PPa >= PPs PPa - PPs/PPa (1-PPs) if PPa < PPs
Lithology	null	n/a	n/a	60	0.026	n/a	n/a	n/a
	ev	445,779	0.003	1	0.000	0.15	0.0000022	0.0000
	ig	13,154,821	0.089	0	0.000	0.00	0.0000000	0.0000
	mt	19,570,508	0.132	259	0.113	0.86	0.0000132	-0.1445
	nd	105,611	0.001	0	0.000	0.00	0.0000000	
	pa	8,575,653	0.058	216	0.094	1.63	0.0000252	0.3859
	pb	1,046,408	0.007	11	0.005	0.68	0.0000105	-0.3204
	pi	578,271	0.004	17	0.007	1.90	0.0000294	0.4738
	py	935,649	0.006	15	0.007	1.04	0.0000160	0.0351
	sc	11,765,424	0.079	158	0.069	0.87	0.0000134	-0.1318
	sm	22,026,381	0.149	234	0.102	0.69	0.0000106	-0.3132
	ss	24,507,815	0.165	653	0.285	1.72	0.0000266	0.4195
	su	36,866,217	0.249	421	0.184	0.74	0.0000114	-0.2618
	va	1,535,538	0.010	41	0.018	1.73	0.0000267	0.4207
	vb	5,244,151	0.035	119	0.052	1.47	0.0000227	0.3183
class count:	vi	2,591,440	0.017	74	0.032	1.85	0.0000286	0.4583
15	wb	1,378,972	0.009	15	0.007	0.70	0.0000109	-0.2968
LITHOLOGY SUMS/AVGs		150,328,639	1.014	2294	0.974	1.02		0.0893
Climate Class	NULL	0	0	0	0.000	0	0	0
	Af	7,051,753	0.048	104	0.045	0.95	0.0000147	-0.0466
	Am	5,085,705	0.035	112	0.049	1.41	0.0000220	0.2976
	Aw/As	16,747,775	0.114	104	0.045	0.40	0.0000062	-0.5986
	BSh	7,866,270	0.053	25	0.011	0.20	0.0000032	-0.7945
	Bsk	8,956,705	0.061	161	0.070	1.15	0.0000180	0.1395
	BWk	7,653,999	0.052	38	0.017	0.32	0.0000050	-0.6790
	BWh	21,803,971	0.148	13	0.006	0.04	0.0000006	-0.9615
	Cfa	5,490,516	0.037	225	0.098	2.63	0.0000410	0.6225
	Cfb	2,525,032	0.017	123	0.054	3.13	0.0000487	0.6825
	Cfc	106,362	0.001	3	0.001	1.81	0.0000282	0.4516
	Csa	1,422,928	0.010	183	0.080	8.26	0.0001286	0.8797
	Csb	854,370	0.006	395	0.172	29.69	0.0004623	0.9666
	Csc	7,905	0.000	0	0.000	0.00	0.0000000	
	Cwa	4,199,998	0.029	115	0.050	1.76	0.0000274	0.4351
	Cwb	1,742,172	0.012	82	0.036	3.02	0.0000471	0.6714
	Cwc	6,699	0.000	0	0.000	0.00	0.0000000	
	Dfa	2,011,149	0.014	190	0.083	6.07	0.0000945	0.8363
	Dfb	7,406,505	0.050	198	0.086	1.72	0.0000267	0.4214
	Dfc	15,596,892	0.106	60	0.026	0.25	0.0000038	-0.7513
	Dfd	615,023	0.004	0	0.000	0.00	0.0000000	
	Dsa	255,656	0.002	6	0.003	1.51	0.0000235	0.3409
	Dsb	587,156	0.004	94	0.041	10.28	0.0001601	0.9034
	Dsc	1,600,938	0.011	14	0.006	0.56	0.0000087	-0.4347
	Dsd	58,703	0.000	0	0.000	0.00	0.0000000	
	Dwa	1,207,276	0.008	12	0.005	0.64	0.0000099	-0.3574
	Dwb	1,206,631	0.008	9	0.004	0.48	0.0000075	-0.5178
	Dwc	3,018,367	0.020	6	0.003	0.13	0.0000020	-0.8715
	Dwd	227,722	0.002	0	0.000	0.00	0.0000000	
class count:	ET	7,953,889	0.054	22	0.010	0.18	0.0000028	-0.8212
30	EF	14,045,768	0.095	0	0.000	0.00	0.0000000	
KG SUMS/AVGs		147,313,836	100.0%	2294	100.0%	2.55		0.0271

Table 7.1 Sample probability density, frequency ratio, conditional probability, and certainty factor statistical results for lithology and Köppen-Geiger climate predisposing environment factors

Orange highlights indicate the factor class with highest significance (calculated value) for each statistical model.

FACTOR	PD	CP	CF	FR	MAXENT
Lithology	***.65***	.0000165	.0695	1.07	.001
KG - Climate	.033	.0000397	.027	2.55	.200
Landform	.028	.0000351	-.153	2.27	.138
Elevation	.05	.0000137	-.175	.84	.009
Slope	.017	.0000035	-.777	11.86	***.212***
Land Cover	.12	.0001631	.030	10.54	.033
Soil Thickness	.14	.0000377	.320	2.34	.022
Soil Type	.17	.0000312	.331	3.42	.059
Soil Drainage					.087
Mean Monthly Precipitation	.07	***.0284062***	***.988***	9.18	.140
Aridity Index	.17	.0000125	-.093	.81	.002
Depth to Bedrock					.010
Topsoil Clay %					0
Closest H2O	.33	0.00005	.027	8.48	
Dist2Ocean	.20	.00000035	-.980	.21	
Dist2<>Ocean	.17	.0005964	.967	4.63	
Fault Density	.20	.00490	.833	***34.04***	.088
Closest "Active Fault"	.04	.0013407	.815	.57	
Close Plate Boundary	.14	.0102014	***.994***	1.16	

Table 7.2 Summary of factor contributions to the Probability Density (PD), Conditional Probability (CP), Certainty Factor (CF), Frequency Ratio (FR) and Maximum Entropy (MAXENT) statistical model results.

Most significant factor per statistical method designated with bold italics

7.1.3 Statistical Analyses: Maximum Entropy

Environmental factor layers and debris flow event validation (TRAIN) and verification (TEST1) data were input to MaxEnt v3.4.4 software v. 3.4.4 (S. Phillips & Dudik, 2008; S. Phillips et al., 2021) choosing five replications, logistic output, and jackknife variable importance determinations. The five replications represent five different samplings of the input event data. In all MaxEnt models the mean of the five replications was used as the susceptibility results. For the global analysis, full global extents were input. For continental models, the data were clipped by each continent's boundaries. Full MaxEnt global and continental results are available in Appendix E.

Maximum Entropy, a "presence-only" machine learning algorithm was included in these susceptibility analyses due to the ambiguity of "absence" in this context; and the dependence on

landslide inventories that were not collected through manual field surveys and thus without verified locations. Absence does not necessarily mean that there are or were no debris flows in an area. It means it is not known and/or there are no substantiating data sources, or ability to conduct field surveys, particularly at a continental or global scale. The MaxEnt software evaluates potential distribution from presence-only data by fitting the probability distribution of maximum entropy to the information provided at each observation. MaxEnt is a widely used technique in biological species distribution modeling with recent and growing interest in its use for landslide susceptibility modeling due to its predictive success compared with other methodologies in "presence" only scenarios (Convertino, Troccoli, & Catani, 2013; GÁL, Poszet, & Kerekes, 2018; Kornejady et al., 2017; L. Lombardo et al., 2016; Park, 2014; Yuan et al., 2017) not requiring definitive information as to where events do not occur. The MaxEnt software also determines the relative importance (significance) of each predisposing factor to the susceptibility results.

MaxEnt results include Area Under the Curve (AUC); minimum, maximum, and median "susceptibility"; response curves which show how each environmental factor affects the MaxEnt prediction; analysis of environmental factor contributions ("significance value"); and results of a jackknife test showing environmental variable importance when included, and not included. The resulting data used in this project include the factor significance values and the mean probability output. The MaxEnt susceptibility maps were input to ArcGIS Pro where additional analyses were performed.

The same environmental variables (factors) were used for PD, FR, CP, and CF. A few of these variables were not employed in the MAXENT model and a few of the variables used in the MAXENT model were not utilized in the PD, FR, CP, and CF models. Table 7.3 lists the environmental factors employed in each model. Environmental factors not utilized in PD, FR, CP, and CF is due to the requirement for an area coverage calculation of the underlying factor classes, which is not practicable for such factors (e.g., slope) at continental and global scales, but which is not a MaxEnt software requirement. Variables not employed in the MAXENT model is due to the difficulty of employing distance-based factors (e.g., distance to faults) as "categorical" or "continuous" spatial representations, as required by the MaxEnt software. It is believed that the slight difference in factors is not an impact on the resulting model comparisons and choice,

as the majority of the factors employed in all models are those which are in common use for debris flow susceptibility analyses, as cited earlier. Using only the common subset of factors, in all models, was not chosen as it is believed each model will perform better with the maximum relevant factors employed, while maintaining a parsimonious approach.

Environmental Variable (Factor)	Employed in PD, FR, CP, CF models	Employed in MaxEnt model
Aridity Index	Yes	Yes
Köppen-Geiger Climate Classification	Yes	Yes
Depth to Bedrock	Yes	Yes
Elevation	Yes	Yes
Fault Density	No	Yes
Landform	Yes	Yes
Landuse/Landcover	Yes	Yes
Lithology	Yes	Yes
Precipitation (monthly average mm)	Yes	Yes
Slope (degrees)	No	Yes
Soil Drainage	No	Yes
Soil Thickness	Yes	Yes
Soil Type	Yes	Yes
Topsoil Clay %	No	Yes
Closest Water Body Type	Yes	No
Closest Active Fault Type	Yes	No
Closest Tectonic Boundary Type	Yes	No
Distance to Sea	Yes	No
Distance to Water Body <> Sea	Yes	No
Distance to Faults	Yes	No

Table 7.3 Environmental variables (factors) employed in each statistical model

Environmental variables not used in all statistical models are highlighted in grey cells.

7.1.4 Classification and Verification

The PD, FR, CP, and CF susceptibility models were developed in ArcGIS Pro 2.7 (Esri, 2020) utilizing the factor classes with the highest rankings within each statistical model, as determined in the Excel calculations. The MaxEnt raster susceptibility map, with values from 0 to 1 at each pixel was input directly into ArcGIS. All susceptibility maps were classified into five equal intervals as Very Low, Low, Medium, High, and Very High.

All susceptibility maps (continental and global) were verified against TEST1 events by determining the number of events within each susceptibility classification, along with the number of pixels (proxy for area) for each classification rank. All details can be found in Appendix E.

7.2 Results

7.2.1 Global Results

Frequency ratio, conditional probability, and certainty factor methods yielded the same results for all factors. Therefore, subsequent discussions will refer to "FR_CP_CF" and represent all three models. Probability Density will be referred to as "PD". "MAXENT" represents the Maximum Entropy model.

The FR_CP_CF and PD global models yielded the poorest test results with 23.3% and 19.7%, respectively, of verification events occurring in the Medium to Very High susceptibility classes. MAXENT was the best performing model with 82.9% of the verification events occurring in the Medium to Very High susceptibility classes (Table 7.4). Model AUC cannot be calculated for the FR_CP_CF and PD models with presence-only data, however the AUC for the global MAXENT model is 0.920. The global susceptibility maps for FR_CP_CF, PD, and MAXENT are shown in Figures 7.2, 7.3, and 7.4, respectively.

Based on these results, the global MAXENT model is the best performing statistical model and is used for subsequent comparative analyses with MAXENT continental models to determine the best approach to "global" modeling.

Susceptibility Classification	FR_CP_CF Ratio % TEST1 events	Probability Density % TEST1 events	Maximum Entropy % TEST1 events
Low	76.7%	71.7%	17.1%
Medium	23.3%	7.3%	32.3%
High – Very High	0	12.4%	50.6%

Table 7.4 Susceptibility model results verified with TEST1 events.

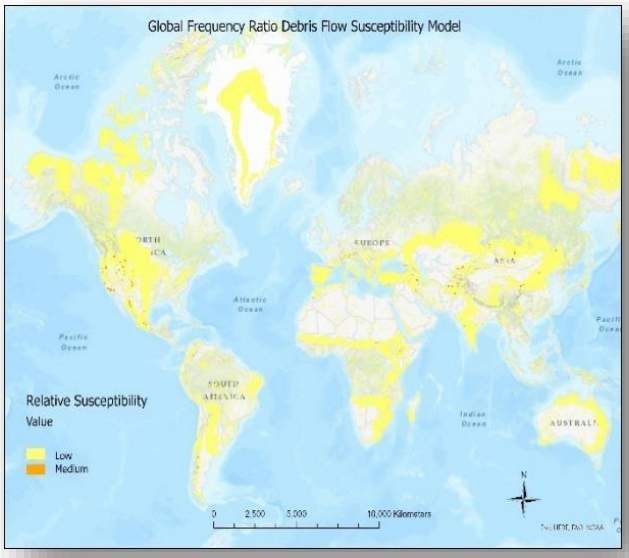


Figure 7.2 Global Frequency Ratio susceptibility model

Figure 7.3 Global Probability Density susceptibility model

Warmer colors represent higher susceptibility.

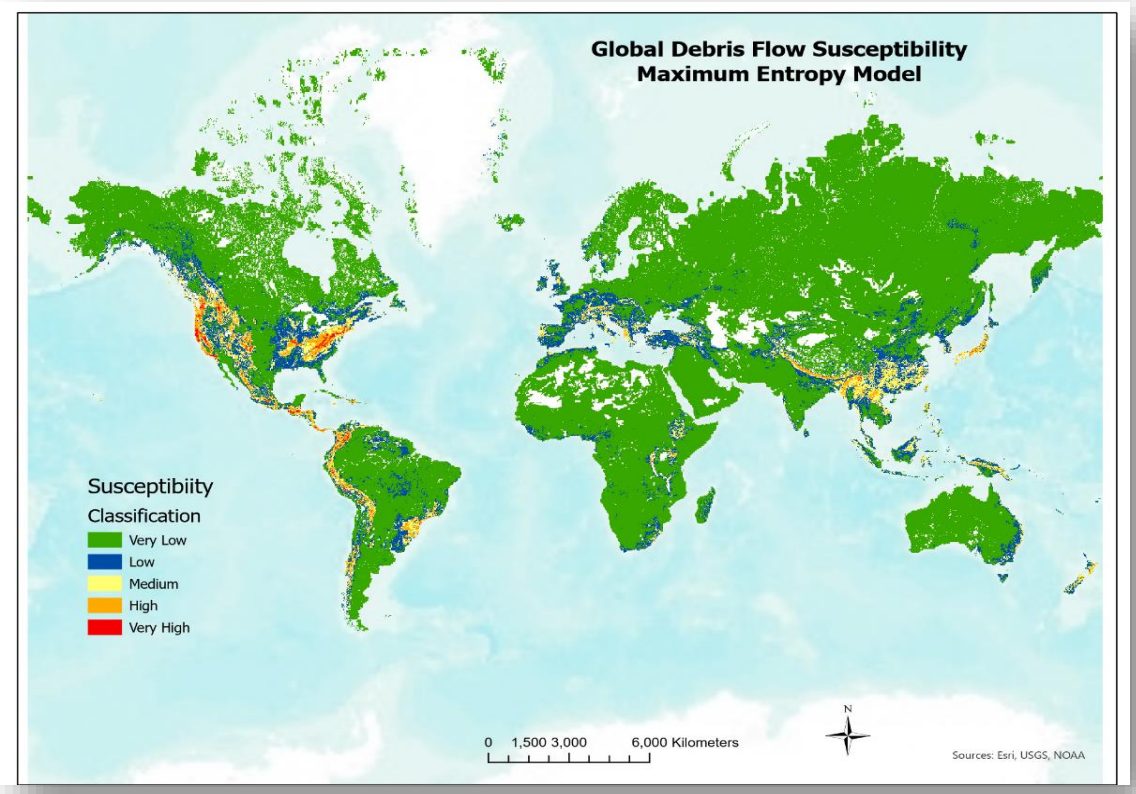


Figure 7.4 Global Maximum Entropy susceptibility model

7.2.2 Continental Results

Each MAXENT continental model was compared with the global MAXENT model utilizing the verification data (TEST1) data. A count was calculated for the Very High susceptibility class and a cumulative count for the Medium, High, and Very High classes for the global and continental models (Table 7.5). Fig. 7.5 is a global susceptibility difference map demonstrating areas where the global susceptibility classification is higher, lower, or equal to a mosaic of the continental models.

Continent	# Test1 Debris Flow Events per continent	# Test1 Debris Flows in Very High Susceptibility Class/Medium to Very High Cumulative (Global Analysis)	% Events in Very High Susceptibility/Medium to Very High Cumulative (Global Analysis)	# Test1 Debris Flows in Very High Susceptibility Class/Medium to Very High Cumulative (Continental Analysis)	% Events in Very High Susceptibility/Medium to Very High Cumulative (Continental Analysis)
Africa	16	5/14	31.3/**87.5**	7/12	**43.8**/75.0
Asia	554	2/478	0.36/**86.3**	4/142	**0.7**/25.6
Australia	1	0/1	0.0/**100.0**	1/1	**100.0/100.0**
Europe	2743	221/2275	**8.1/82.9**	214/1322	7.8/48.2
N. America	2370	797/2350	**33.6/99.2**	566/2338	23.8/98.6
Oceania	5	1/1	**20.0**/20.0	0/3	0/**60.0**
S. America	6	1/4	**16.7/66.7**	0/2	0/33.3

Table 7.5 Percent of verification events (TEST1) and MAXENT classification, by continent.

Bold red emphasis indicates model with highest value for each continent.

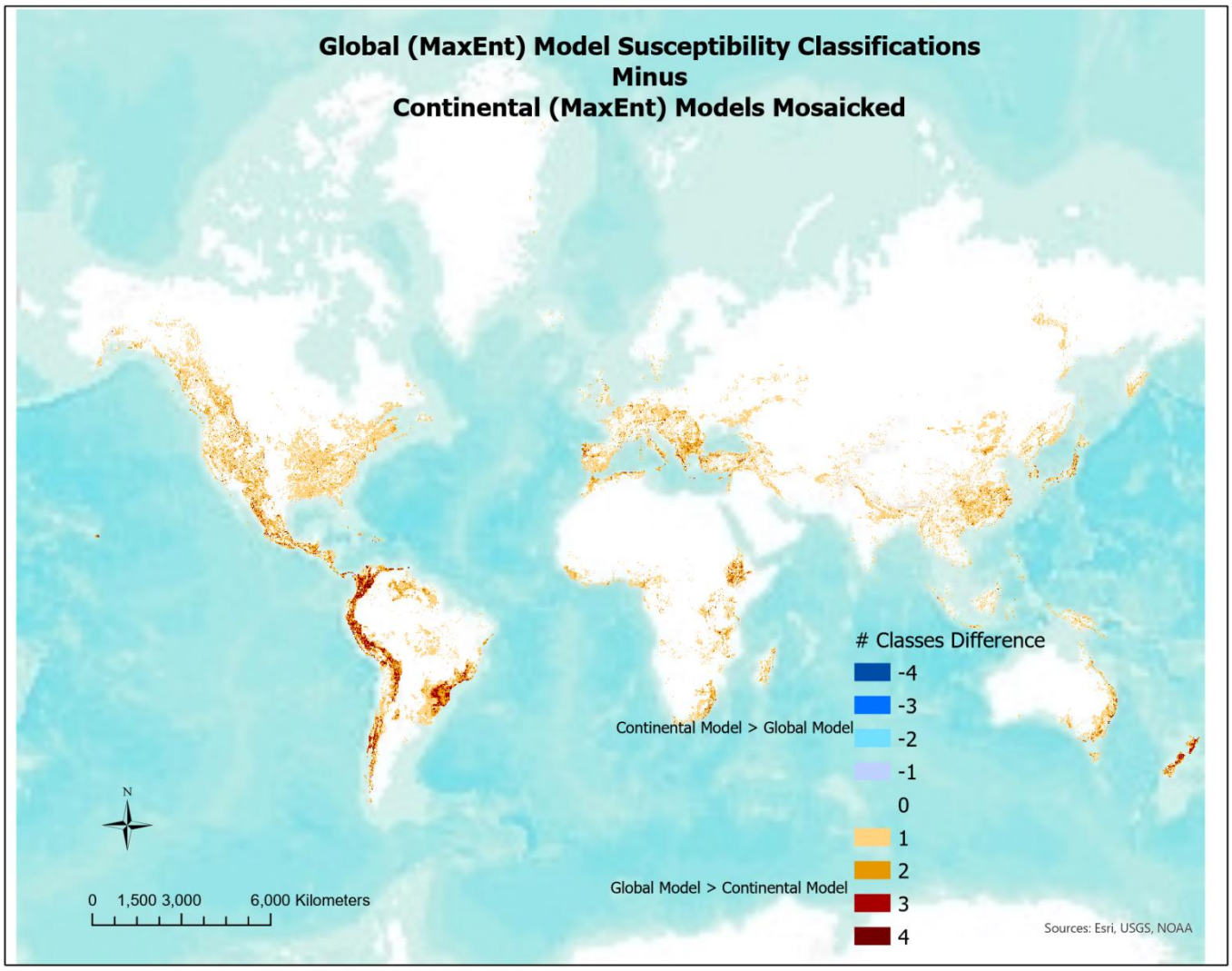


Figure 7.5 Global MaxEnt debris flow susceptibility minus composite (mosaic) of individual continental MaxEnt susceptibility models

Table 7.6 presents a summary of the differences in susceptibility classification between the continental and global models for each continent, by percent area. A difference in susceptibility can be by one or more susceptibility classes. There is significant concurrence (greater than 70% by area) between the susceptibility classifications in the continental and global models for six of the seven continents (Africa, Asia, Australia, Europe, North America, South America). Greater than 95% of the global model is within ± one susceptibility classification of the continental models, and less than 10% of the global model exhibits a greater susceptibility than the continental models (Table 7.7).

Continental Global "Cut" minus Continental Susceptibility	Africa	Asia	Australia	Europe	No America	Oceania	So America
% Contl Susc > Global	1.5%	3.4%	1.3%	6.1%	2.2%	14.0%	0.2%
% Contl Susc = Global	**88.8%**	**86.0%**	**89.8%**	**76.7%**	**73.3%**	**12.5%**	**71.7%**
% Global Susc > Contl	9.7%	10.6%	8.9%	17.2%	24.5%	73.5%	28.1%

Table 7.6 Continental cut from global model minus continental model, susceptibility differences by area (pixel count).

Global Minus Mosaic of All Continentals	Difference in # of Susceptibility Classes	% Area
Contl > Glob	-4	0.02%
Contl > Glob	-3	0.06%
Contl > Glob	-2	0.23%
Contl > Glob	-1	1.21%
Contl = Glob	0	88.78%
Glob > Contl	1	7.77%
Glob > Contl	2	1.68%
Glob > Contl	3	0.25%
Glob > Contl	4	0.01%

Table 7.7 Number of susceptibility classification differences between the global model and mosaic of all continental models

Result for the global model minus individual continental model are presented in the following sections for each continent.

7.2.2.1 *Africa*

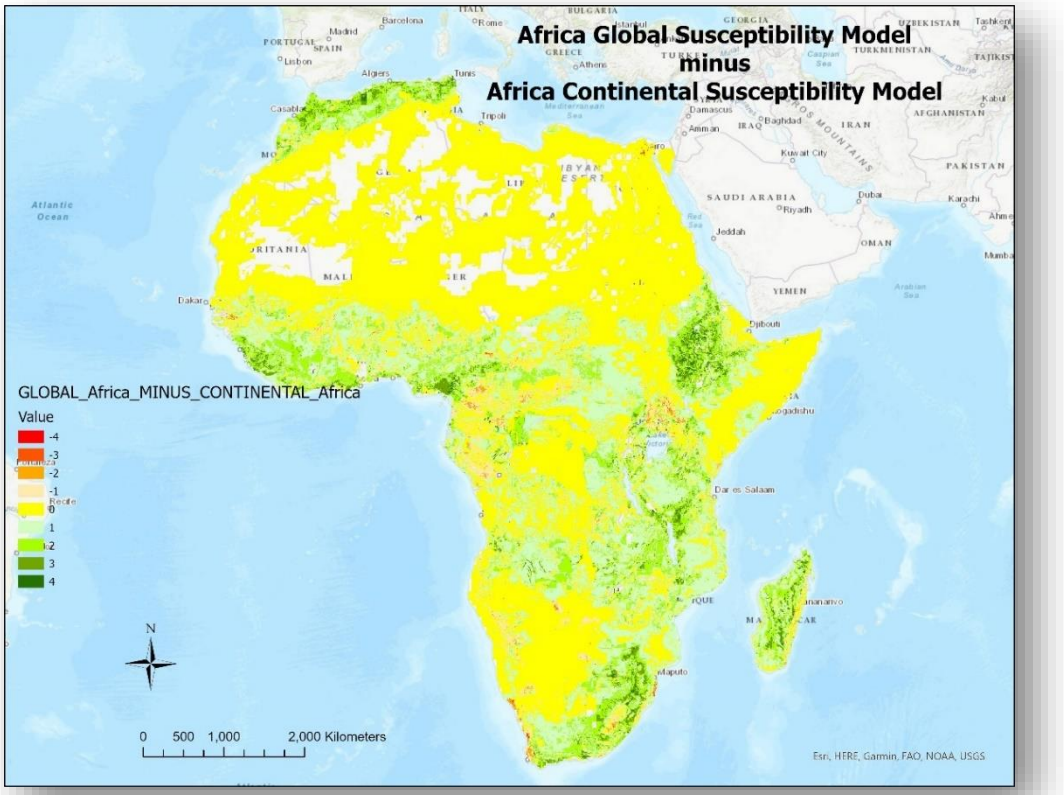


Figure 7.6 Global "cut" Africa susceptibility model minus Continental Africa

Positive numbers represent areas where the global model susceptibility is higher than the continental model, negative numbers represent areas where the continental susceptibility is higher.

Africa			
Comparison	**# Classes Difference**	**% Area**	**Cum %**
Contl > Global	-4	0.02%	
Contl > Global	-3	0.06%	
Contl > Global	-2	0.23%	
Contl > Global	-1	1.21%	1.52%
Contl = Global	**0**	**88.78%**	
Global > Contl	1	7.77%	
Global > Contl	2	1.68%	
Global > Contl	3	0.25%	
Global > Contl	4	0.01%	9.70%

Table 7.8 Global Africa "cut" susceptibility model minus continental Africa model, by percent area coverage.

Positive numbers represent areas where the global model susceptibility is higher than the continental model by 1,2, 3, or 4 susceptibility class levels, negative numbers represent areas where the continental susceptibility is higher. "0" represents areas where the global and continental model susceptibility classification is equal.

Asia

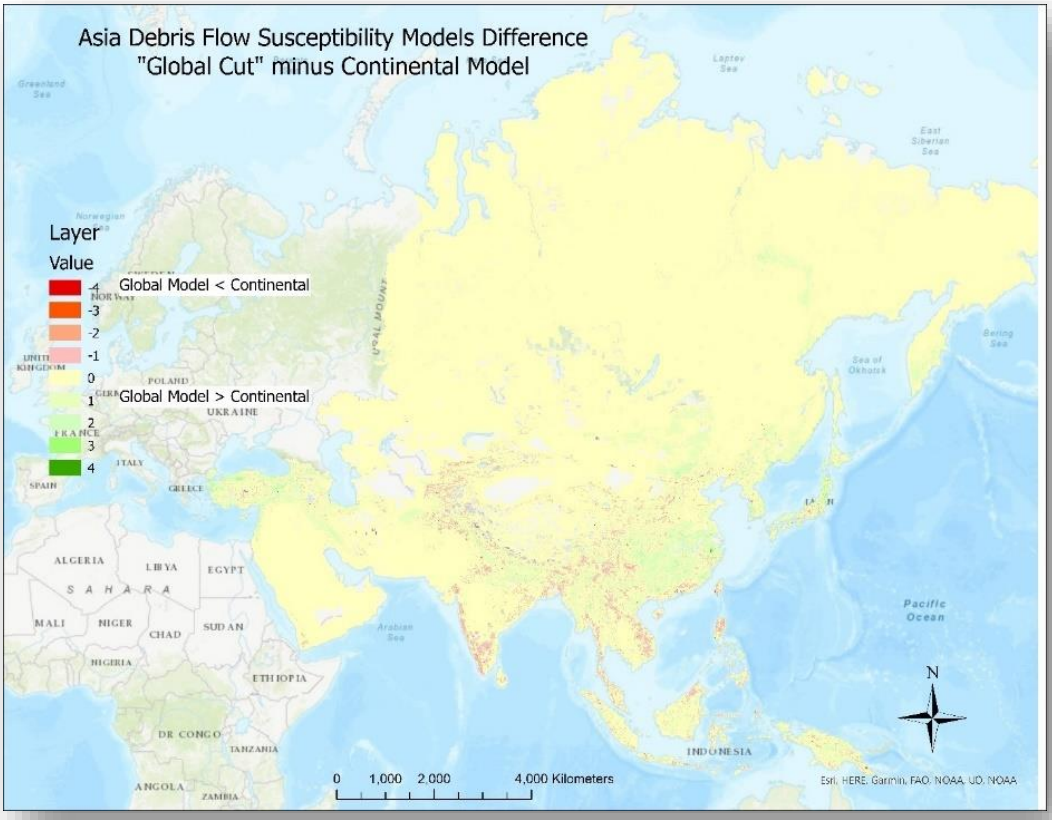


Figure 7.7 Global "cut" Asia susceptibility model minus Continental Asia model

Positive numbers represent areas where the global model susceptibility is higher than the continental model, negative numbers represent areas where the continental susceptibility is higher.

Asia			
Comparison	**# Classes Difference**	**% Area**	**Cum %**
Contl > Global	-4	0.00%	
Contl > Global	-3	0.04%	
Contl > Global	-2	0.26%	
Contl > Global	-1	3.11%	3.42%
Contl = Global	**0**	**85.96%**	
Global > Contl	1	9.24%	
Global > Contl	2	1.27%	
Global > Contl	3	0.11%	
Global > Contl	4	0.00%	10.63%

Table 7.9 Global Asia "cut" susceptibility model minus continental Asia model, by percent area coverage.

Positive numbers represent areas where the global model susceptibility is higher than the continental model by 1,2, 3, or 4 susceptibility class levels, negative numbers represent areas where the continental susceptibility is higher. "0" represents areas where the global and continental model susceptibility classification is equal.

7.2.2.2 *Australia*

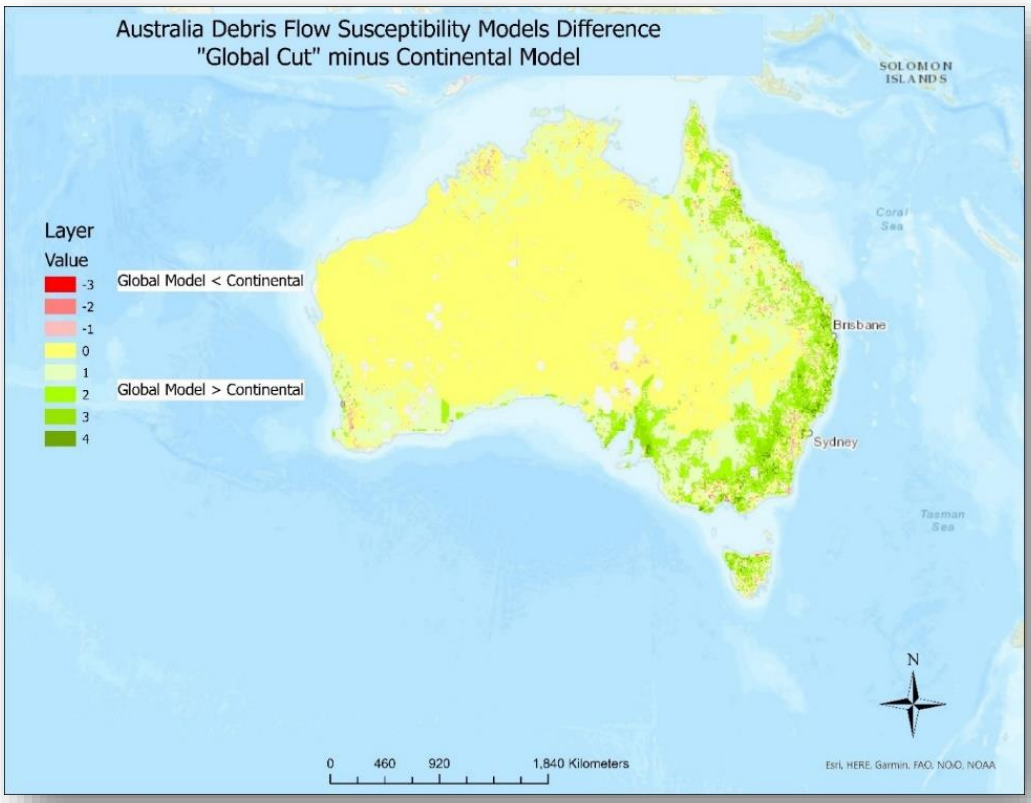


Figure 7.8 Global "cut" Australia susceptibility model minus continental Australia

Positive numbers represent areas where the global model susceptibility is higher than the continental model, negative numbers represent areas where the continental susceptibility is higher.

Australia			
Comparison	**# Classes Difference**	**% Area**	**Cum %**
Contl > Global	-4	0.00%	
Contl > Global	-3	0.04%	
Contl > Global	-2	0.25%	
Contl > Global	-1	0.99%	1.29%
Contl = Global	**0**	**89.80%**	
Global > Contl	1	7.06%	
Global > Contl	2	1.56%	
Global > Contl	3	0.27%	
Global > Contl	4	0.03%	8.91%

Table 7.10 Global Australia "cut" susceptibility model minus continental Australia model, by percent area coverage.

Positive numbers represent areas where the global model susceptibility is higher than the continental model by 1,2, 3, or 4 susceptibility class levels, negative numbers represent areas where the continental susceptibility is higher. "0" represents areas where the global and continental model susceptibility classification is equal.

7.2.2.3 *Europe*

As an expanded comparative analysis example of the process applied to all continents, Figs. 7.9, 7.10, and 7.11 are the MAXENT global, continental, and difference models for Europe, respectively.

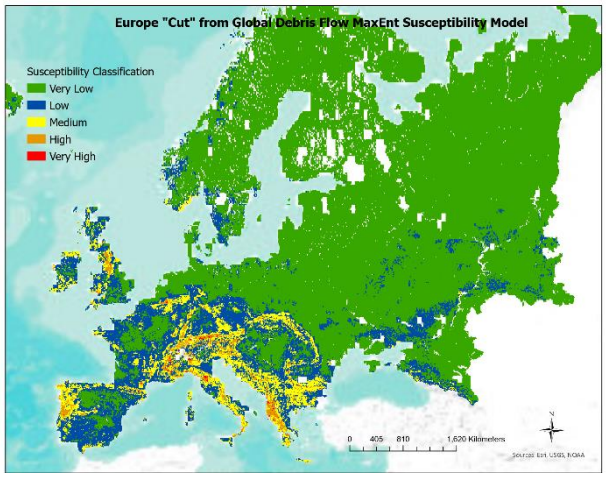


Figure 7.9 Europe global cut MAXENT susceptibility

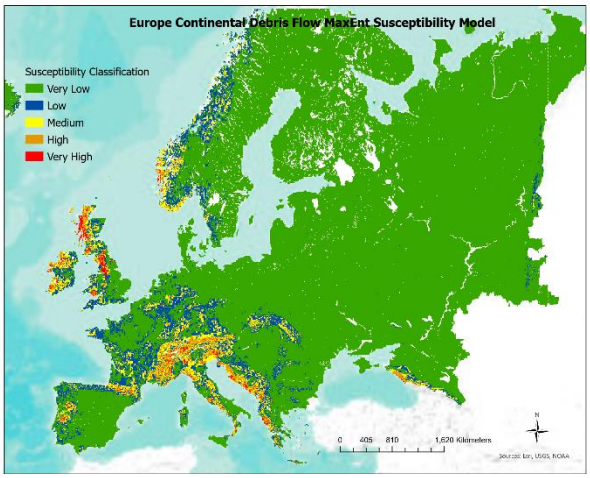


Figure 7.10 Continental MAXENT susceptibility

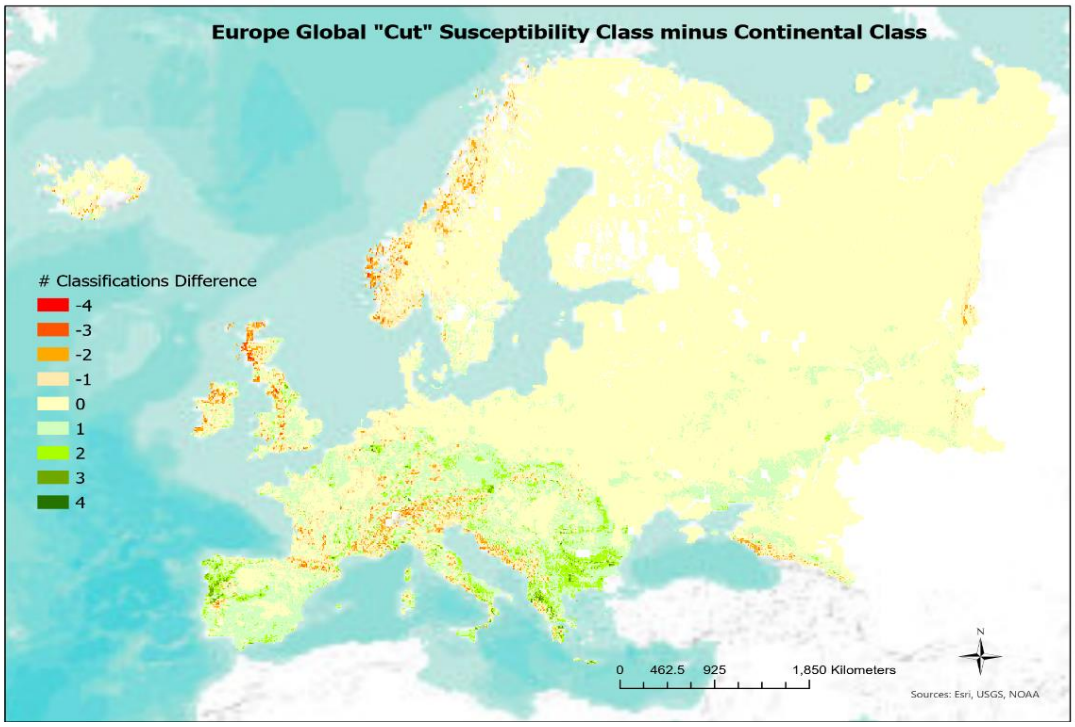


Figure 7.11 Europe debris flow susceptibility difference map. Global cut model minus continental model

Positive numbers represent areas where the global model susceptibility is higher than the continental model, negative numbers represent areas where the continental susceptibility is higher.

Europe			
Comparison	**# Classes Difference**	**% Area**	**Cum %**
Contl > Global	-4	0.01%	
Contl > Global	-3	0.21%	
Contl > Global	-2	1.45%	
Contl > Global	-1	4.43%	6.11%
Contl = Global	**0**	**76.71%**	
Global > Contl	1	14.04%	
Global > Contl	2	2.83%	
Global > Contl	3	0.30%	
Global > Contl	4	0.01%	17.18%

Table 7.11 Global Europe "cut" susceptibility model minus continental Europe model, by percent area coverage.

Positive numbers represent areas where the global model susceptibility is higher than the continental model by 1,2, 3, or 4 susceptibility class levels, negative numbers represent areas where the continental susceptibility is higher. "0" represents areas where the global and continental model susceptibility classification is equal.

7.2.2.4 North America

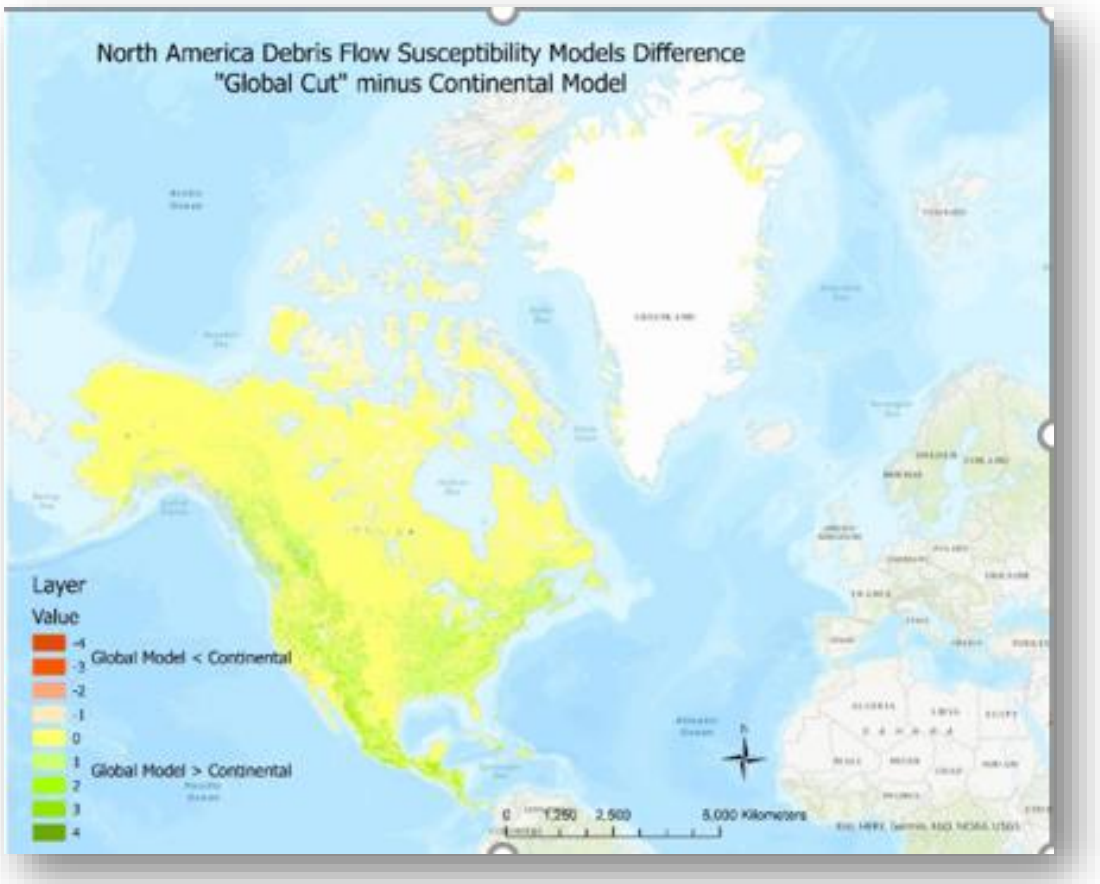


Figure 7.12 North America debris flow susceptibility difference map. Global cut model minus continental model.

Positive numbers represent areas where the global model susceptibility is higher than the continental model, negative numbers represent areas where the continental susceptibility is higher.

North America			
Comparison	**# Classes Difference**	**% Area**	**Cum %**
Contl > Global	-4	0.00%	
Contl > Global	-3	0.01%	
Contl > Global	-2	0.18%	
Contl > Global	-1	2.03%	2.23%
Contl = Global	**0**	**73.28%**	
Global > Contl	1	20.08%	
Global > Contl	2	3.93%	
Global > Contl	3	0.46%	
Global > Contl	4	0.03%	24.49%

Table 7.12 Global North America "cut" susceptibility model minus continental North America model, by percent area coverage.

Positive numbers represent areas where the global model susceptibility is higher than the continental model by 1,2, 3, or 4 susceptibility class levels, negative numbers represent areas where the continental susceptibility is higher. "0" represents areas where the global and continental model susceptibility classification is equal.

7.2.2.5 Oceania

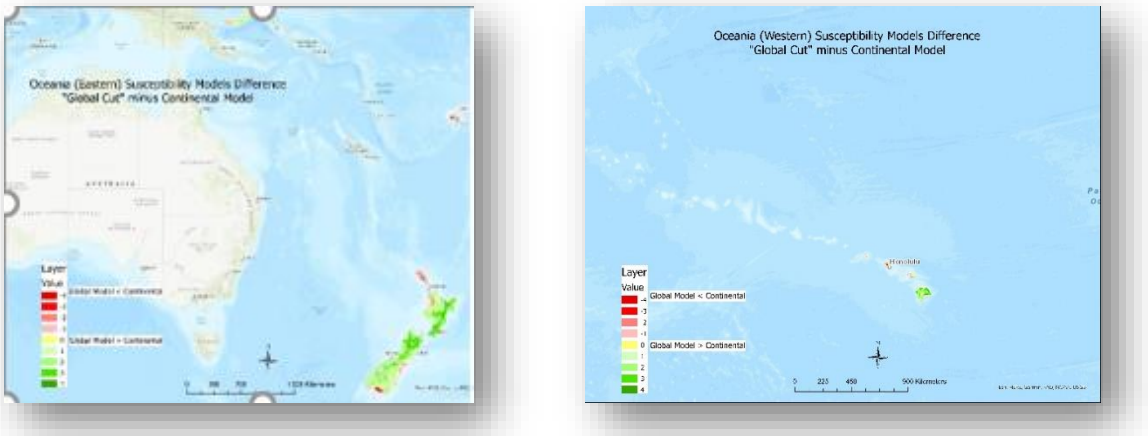


Figure 7.13 Western (left) and eastern hemisphere (right) Oceania debris flow susceptibility difference map. Global cut model minus continental model

Positive numbers represent areas where the global model susceptibility is higher than the continental model, negative numbers represent areas where the continental susceptibility is higher.

Oceania			
Comparison	**# Classes Difference**	**% Area**	**Cum %**
Contl > Global	-4	0.4%	
Contl > Global	-3	1.4%	
Contl > Global	-2	4.5%	
Contl > Global	-1	7.5%	13.9%
Contl = Global	**0**	**12.5%**	
Global > Contl	1	25.3%	
Global > Contl	2	28.1%	
Global > Contl	3	18.4%	
Global > Contl	4	1.6%	73.5%

Table 7.13 Global Oceania "cut" susceptibility model minus continental Oceania model, by percent area coverage.

Positive numbers represent areas where the global model susceptibility is higher than the continental model by 1,2, 3, or 4 susceptibility class levels, negative numbers represent areas where the continental susceptibility is higher. "0" represents areas where the global and continental model susceptibility classification is equal.

7.2.2.6 South America

Figure 7.14 South America debris flow susceptibility difference map. Global cut model minus continental model

Positive numbers represent areas where the global model susceptibility is higher than the continental model, negative numbers represent areas where the continental susceptibility is higher

South America			
Comparison	**# Classes Difference**	**% Area**	**Cum %**
Contl > Global	-4	0.006%	
Contl > Global	-3	0.03%	
Contl > Global	-2	0.06%	
Contl > Global	-1	0.12%	0.22%
Contl = Global	**0**	**71.7%**	
Global > Contl	1	15.4%	
Global > Contl	2	7.5%	
Global > Contl	3	4.6%	
Global > Contl	4	0.6%	28.13%

Table 7.14 Global South America "cut" susceptibility model minus continental South America model, by percent area coverage.

Positive numbers represent areas where the global model susceptibility is higher than the continental model by 1,2, 3, or 4 susceptibility class levels, negative numbers represent areas where the continental susceptibility is higher. "0" represents areas where the global and continental model susceptibility classification is equal.

7.3 Discussion

Several methods were employed to develop a global debris flow susceptibility map. The world was modeled at a single global scale and on a continent-by-continent basis. The results were tested and compared to determine which is the best approach to a debris flow susceptibility map of the world. Of the FR, CP, CF, PD, and MAXENT statistical methods employed, the MAXENT model outperformed PD and FR_CP_CF by 4 and 3.5 times, respectively, as proven with TEST1 verification data. MAXENT was therefore used for subsequent analyses in determining whether a global or continental model provides the better fit.

Evaluating only the Very High susceptibility class, four of the seven continents (Europe, North America, Oceania, South America) demonstrated better performance (greater verified susceptibility results) in the global model than in the individual continental models (Table 7.5). Africa, Asia, and Australia demonstrated better results in the continental model for Very High Susceptibility. However, when evaluating the cumulative results, which may be a better (more generalized) measure, using the Medium (average susceptibility = .45), High (average susceptibility = .64) , and Very High (average susceptibility = .83) susceptibility classes, all continents except Oceania and Australia demonstrated better results in the global model. Greater than 85% of the TEST1 data were correctly classified in Medium to Very High susceptible areas for the six continents (Africa, Asia, Australia, Europe, North America, South America). An average of 54.3% of the TEST1 events were classified in the High and Very High susceptibility classification for all seven continents. Australia exhibited comparable results in both the global and continental models and Oceania exhibited superior results in the continental model for the Medium to Very High cumulative classes. There is only one TEST1 event in Australia (16 TRAIN events) which calls into question results for Australia. The unequal geographic distribution of both TRAIN and TEST1 events have possible implications on this modeling.

As is true in all modeling, subjectivity and biases may be introduced in multiple ways, which may then influence the results. It is acknowledged in this study that there may be a geographic bias in the TRAIN and TEST1 global event inventories based on data collection methodologies and data availability. In comparing two continents (North America and Asia) in the northern

hemisphere with similar overall compositions of climate, and other predisposing factors, the ratio of North America land area to Asia is 0.54, while the TRAIN and TEST1 event density ratios are 5.8 and 7.9, respectively. It is unknown from this research how this may affect the results presented but is an important topic for further study.

Potential explanations for the better or comparable performance of a global model over individual continental models are:

1. Appropriateness of predisposing factors to represent all debris flows regardless of geography: The environmental factors used in this study are those which are generally accepted as the major predisposing factors regardless of region or continent, as cited earlier. Table 7.15 presents the similarity of factor class values associated with the majority of debris flow events for most continents; for example, in eight of the eleven factors: precipitation, aridity, climate, fault density, landcover, landform, lithology, and soil depth.
2. With insufficient geographic distribution and number of debris flow events on each continent, a global statistical model may provide a smoothing effect across all continents, thereby compensating for a statistically smaller number of model training events (Africa, Australia, Oceania), and/or a low density of training events (Africa, Australia, South America), and for what could be an unintended bias in the geographic distribution of the event inventory. A small frequency or density of model training events at the continental level could lead to flawed susceptibility results.
3. These may be valid conclusions only if we can say that debris flow predisposing factors and factor classes are, within reason, universal, at continental and global scales. A follow-up study based on comparisons of accurately defined and assigned debris flow factors and factor classes determined at local and regional scales, based on field surveys or highly accurate remote sensing analyses, can prove beneficial in understanding and improving continental and global modeling.
4. As evidenced in Chapter6 (Tables 6.1, 6.2, and 6.3), there is a stronger similarity than dissimilarity in the geomorphic, land use/landcover, and lithologies (respectively) in the continents.

	AFRICA	ASIA	AUSTRALIA	EUROPE	NORTH AMERICA	OCEANIA	SOUTH AMERICA
Precipitation Avg monthly (mm)	***100-200***	70-80	60-70	***100-200***	***100-200***	***100-200***	***100-200***
Aridity	***Humid***	***Humid***	Dry sub-humid	***Humid***	***Humid***	***Humid***	***Humid***
Climate (KG)	Tropical megathermal (Aws)	***Temporal mesothermal*** (Cwa)	***Temporal mesothermal*** (Cfb)	***Temporal mesotherma l*** (Csa)	***Temporal mesothermal*** (Csb)	***Temporal mesothermal*** (Cfb)	Tropical megathermal (Am)
Distance to Faults (km)	< 200	15-25	15-20	25-50	1-25	15-20	< 15
Fault Density (km/sq km)	***.01-.02***	***.02-.03***	***.02-.03***	***.02-.03***	***.02-.03***	***.01-.02***	***.03-.04***
Landcover	***Forest & Woodland***	***Forest & Woodland***	Urban	***Forest & Woodland***	***Forest & Woodland***	***Forest & Woodland***	Cropland
Landform	Widely spaced ***mountains in alpine system***	***Mountains in alpine system***	Plains in alpine system	***Mountains in alpine system***	Humid ***mountains in alpine system***	Humid ***mountains in alpine system***	Humid ***mountains in alpine system***
Lithology	metamorphic (mt)	***sedimentary*** mixed (sm)	***sedimentary*** siliciclastic (ss)	***sedimentary*** carbonate (sc)	***sedimentary*** siliciclastic (ss)	volcanic basic (vb)	metamorphic (mt)
Soil Type	Nitosol (Nh)	Acrisol (Ao)	Planosol (Ws)	Cambisol (Bd/Be)	Acrisol (Ah)	Acrisol (Ag)	Lithosol (I)
Soil Depth (m)	***1***	***1***	***1***	***1***	***1***	***1***	***1***
Slope (deg)	<=20	<=30	<=10	<=30	<=20	<=20	<=10

Table 7.15 Predominant factor class values associated with the debris flow events (TEST1). Bold italic emphasis indicates those factors which are comparable among the majority of continents.

7.4 Conclusions

Global landslide analyses are possible and meaningful. Conventional wisdom might lead one to believe that a single debris flow susceptibility model for the entire planet may not accurately represent the susceptibility of the individual continents, just as a single model representing susceptibility of any/all landslide types does not accurately represent the susceptibility of individual and distinct landslide types. Thus, a single global debris flow susceptibility model for Earth was developed and compared against debris flow models developed for each continent.

The predominant lithology type (sedimentary) and landuse cover (forest and woodland) is common among all continents, and differences exist from continent to continent with regard to other environmental factors which are known to be associated with debris flows. Mean elevation varies from Asia at 915 m a.s.l. to Europe with 300 m a.s.l. The dominant landcover varies from forest and woodland in Africa, Asia, Oceania, Europe, North and South America to grassland and shrubland in Australia. Mountain belts dominate the geomorphic structures of Asia and North America, crystalline shields dominate Africa, Australia, and South America. Erosional plains dominate Europe. There is sufficient variation to suggest that a single global model may not adequately represent the debris flow susceptibility on all continents. Yet, this research demonstrates that a single global model performs exceptionally well in comparison with individual continental models, as evidenced by six of the seven continents when evaluating on the cumulative results of the Medium, High, and Very High susceptibility classes. This may be due to the fact that debris flow environmental predisposing factors are similar and well-defined in many different types of regions across the world, and although there are dissimilarities in geologic, geomorphologic, hydrologic, tectonic, land use/land cover, climate, and other factors among the continents, there are pockets or regions within each continent with the environmental factors conducive to debris flows. For example, 70% of Australia is composed of arid or semi-arid land, yet the eastern seaboard exhibits High and Very High debris flow susceptibility, and is associated with historic debris flow events.

It is believed that with a larger debris flow event inventory, and an unbiased geographic distribution, these global and individual continental susceptibility models can be further improved.

With respect to the use of the global model versus individual continental models, intended use of debris flow susceptibility maps will dictate whether the error of commission, as is evidenced by the global model, or omission (as seen in the continental models) is better tolerated.

8 Impact of Spatial Data Uncertainty on Debris Flow Susceptibility Analyses

Uncertainties are inherent in all spatial data and at all scales (Chrisman, 1989), however when working at continental and global scales, and in the absence of direct field surveys, the uncertainties are inherently greater (Maffini et al., 1989; Openshaw, 1989). "Uncertainty exists widely in the natural world, and certainty is conditional and relative" (Shi, 2010). When utilizing methods which overlay and correlate multiple datasets, each with their own uncertainty, their derivative products, such as susceptibility maps, are prone to error propagation of an unknown magnitude (Shi, 2010). It is not a matter of adding more or better data, but rather a "sobering reminder that uncertainty is an irreducible part of sufficiently complex knowledge" (Couclelis, 2003), and thus should be thought of as a natural component of, and addressed in, every project. The presence of data uncertainty does not preclude the use of the data, but rather necessitates a methodology for qualitatively or quantitively characterizing and conveying the level of uncertainty and modeling the associated uncertainty for the benefit of the end-users of the hazard model. "Unlike industrial and other products of material processes, knowledge products do not carry with them the evidence of their own inadequacy… and the most critical aspects of the quality of its products are often only testable through their indirect and sometimes remote consequences." (Couclelis, 2003). The aim, herein, is to minimize the societal 'testing' of an inadequately understood debris flow susceptibility model through the awareness of the uncertainty, the potential impact, and introduction of methodologies to develop susceptibility results which include some measure of the awareness.

Although there are many factors and attributes associated with debris flow analyses which are prone to uncertainty, such as debris flow location, type, volume/size, setting, predisposing factors, triggers, etc. (Ardizzone, Cardinali, Carrara, Guzzetti, & Reichenbach, 2002; Carrara, Cardinali, & Guzzetti, 1992; Malamud et al., 2004), for simplicity, and demonstration purposes, in this study, only the uncertainty associated with debris flow event location is investigated.

In a landslide susceptibility analysis, landslide event location accuracy is paramount yet often inaccurately known unless a direct field survey is conducted. Landslide inventories are often constructed based on mapping from aerial imagery, media reports, local governmental agencies,

witness accounts, and field work by third party sources (Froude & Petley, 2018; D. Kirschbaum et al., 2015; Malamud et al., 2004).

In a study of debris flow susceptibility on the European continent, an analysis of the impact between known location and a location accuracy offset for ninety-nine debris flows, demonstrates the impact of uncertainty in appropriately defining predisposing factors associated with debris flows, and consequent analysis for areas of susceptibility.

8.1 Methodology

The impact of event location uncertainty on determination of debris flow predisposing factors and factor classes is investigated for the continent of Europe. The purpose is to investigate and understand the impact and determine susceptibility methodologies to minimize the impact and apply them to the global debris flow susceptibility modeling. The dominant predisposing environmental factors, as determined through Maximum Entropy modeling of Europe, are analyzed with respect to the values found at debris flow event points versus a buffered distance of locational uncertainty around each point. This objective is to illuminate potential discrepancies of predisposing factors associated with debris flow events based on their given point location versus predisposing factors associated with a buffered distance around that point. Distribution of the model validation and verification data for Europe and the 93 debris flow events with known location uncertainty buffers are presented in Figs. 8.1 and 8.2, respectively.

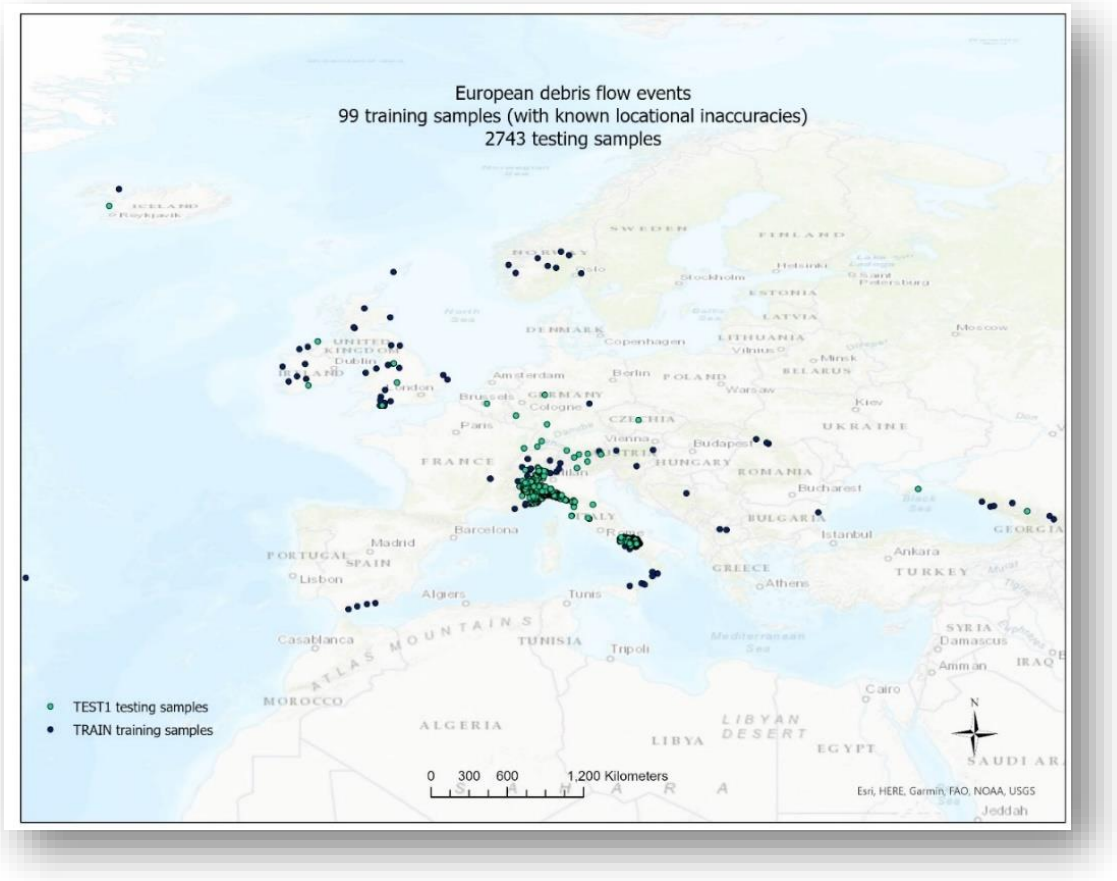


Figure 8.1 Distribution of Europe TRAIN (validation), black dots, and TEST1 (verification) samples, green dots.

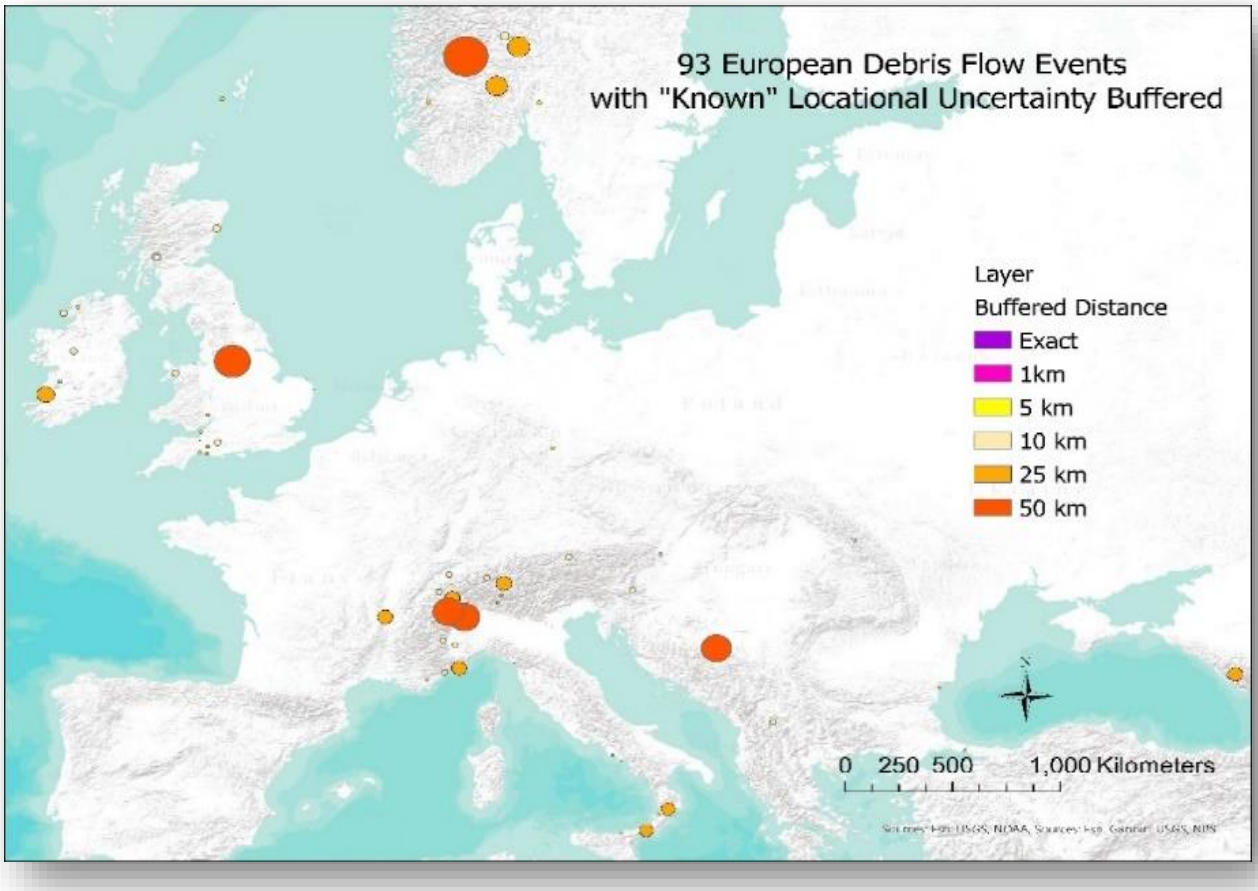


Figure 8.2 93 (of 99) European debris flow events with a buffered "known" locational uncertainty. Some larger buffers overlap and occlude nearby smaller buffers. Base map is from ArcGIS®, the intellectual property of Esri, used herein under license.

Maximum Entropy is used for this uncertainty study based on the modeling performance results in the global versus continental susceptibility analyses documented in Chapter 7.

Six Europe models were developed with MaxEnt software. Each model is analyzed with respect to the AUC, highest contributing predisposing factors, optimal factor classes, susceptibility impact, and comparisons of the susceptibility distributions and susceptibility value differences.

The first model was developed using the original inventory of 99 Europe debris flow events with varying locational uncertainties, as the model training data. This model is subsequently referred to as "Original". The second model utilizes a training dataset created from 93 randomly generated points, one within each of the original 93 locational uncertainty buffers (six of the 99

had an “unknown” uncertainty). This model is referred to as “93 Random”. The third model utilizes only those original inventory event points with known locational uncertainties of 5 km (31 events). This is the “5 km” model. The fourth model, “LTE 5 km” utilizes only those original inventory event points with known locational uncertainties <= 5 km (55 events), including “exact” points, as the training data. The fifth model, “LTE 1 km” utilizes only those original inventory event points with known locational uncertainties <= 1 km (24 events), including “exact” points. The sixth model, “Exact”, utilizes only those points identified as having an “exact” known location (five events). All MaxEnt model runs include a test dataset (TEST1) composed of 2743 debris flow events in Europe, with no location accuracy attribute, collected from numerous sources.

The “Original” model is the model against which all other models are compared, with the purpose of understanding the suitability of such an inventory of events with a broad spectrum of locational uncertainties, and the option for choosing an alternative set of events for the best susceptibility representation.

8.2 Results

AUCs for the “Original”, “93 Random”, “5 km”, “LTE 5 km”, “LTE 1 km”, and “Exact” MaxEnt models are 0.891, 0.857, 0.893, 0.896, 0.921, and 0.93, respectively for the continent of Europe. The “Exact” model, with the highest AUC, is ignored in the subsequent analyses due to the small number of points (four in the United Kingdom and one in Ukraine), and low density (5.05×10^{-7} events per sq km); and hence susceptibility results likely non-representational of the continent.

The “Original” MaxEnt model resulted in precipitation, fault density, and soil type having the highest relative contributions (“gain”) at 42%, 27.6%, and 8.6%, respectively. The gain for all factors included in the model, sum to 100%. Table 8.1 shows the AUC and the factor percent contribution (gain) for the “Original” MaxEnt model, and all factors input to the model.

MaxEnt “gain” addresses the question as to which environmental variables are the most significant to the debris flow events being modeled. The jackknife results, for each variable, were also found to support these three variables as the most significant (Fig. 8.3).

Thus, for purposes of this uncertainty study, the location uncertainty impact is analyzed and demonstrated for only these three factors.

Variable/Model	"Original" (99 points) Percent contribution (gain)
AUC	0.891
precipitation	42.0
fault density	27.6
soil type	8.6
landform	4.9
landcover	4.7
climate	4.5
soil thickness	2.4
drainage	2.3
lithology	2.1
aridity	0.5
elevation	0.2
topsoil % clay	0.1
depth to bedrock	0.1

Table 8.1 AUC and factor percent contribution for Europe "Original" MAXENT susceptibility model.

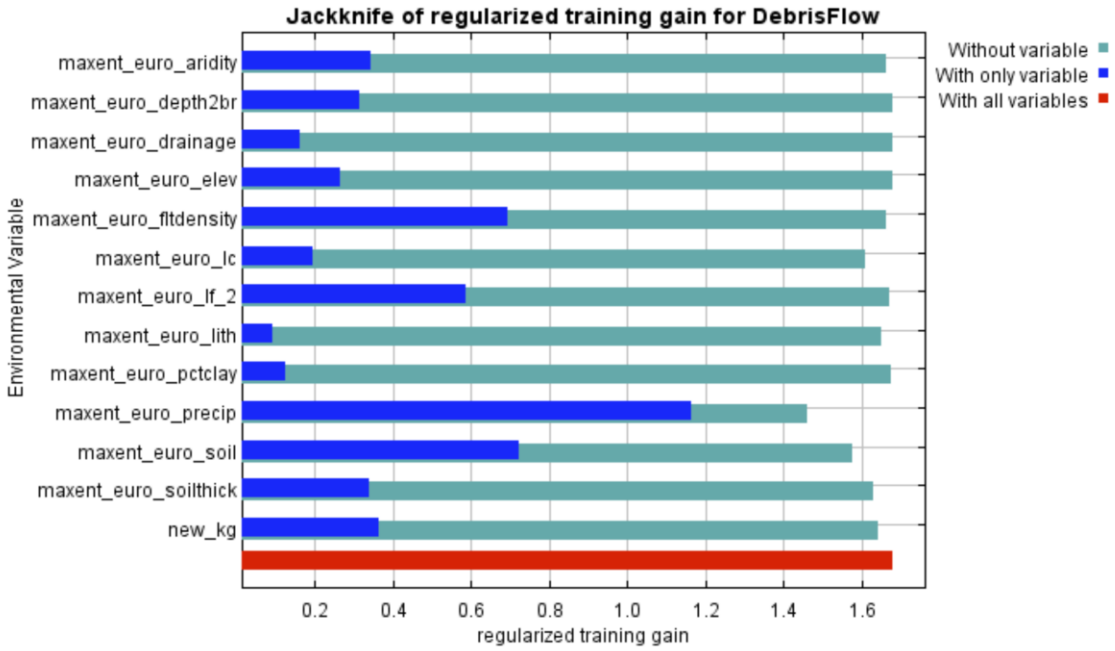


Figure 8.3 Figure 8.3 "Original" MaxEnt jackknife results. "maxent_euro_fltdensity" = the fault density layer, "maxent_euro_precip" = precipitation layer, "maxent_euro_soil" = soil type layer.

8.2.1 Impact of locational uncertainty on precipitation as a predisposing factor - Europe

According to the "Original" MaxEnt model, the monthly average precipitation predisposing factor provided the highest percent contribution (gain) and was the most significant factor in the jackknife test both in most significance as the only variable, and most significant negative impact when removed from the model. In a worldwide study of non-seismic landslide occurrences, Froude and Petley (Froude & Petley, 2018) found a strong correlation between the mean monthly precipitation and landslide events in four of five global regions studied.

Table 8.2 describes the precipitation value range found by an overlay of the 93 "known" locational uncertainty buffered areas and the precipitation factor layer. Two points, events 560 and 6381, each with a 50 km location uncertainty, are selected to view the data and impact at a larger scale. Table 8.3 and Fig. 8.4 depict precipitation values and range of values found within their buffered areas.

Maximum number of different precipitation values within all buffered events	46
Maximum value spread in all buffered events	92 mm
Average number of different precipitation values associated with all buffered events	5.8
Average value spread in all buffered events	13.6 mm

Table 8.2 Range and number of varying precipitation values associated with all 93 debris flow event buffered areas.

Event ID	**560**	**6381**
Precipitation point value	92 mm	73 mm
Locational uncertainty buffer Size	50 km	50 km
Number of different precipitation values within buffered area	32	31
Range of precipitation values within buffered area	58-141 mm	49-94 mm

Table 8.3 Example impact of uncertain location: buffered area values vs point value on associated precipitation factor class determination, for two sample events.

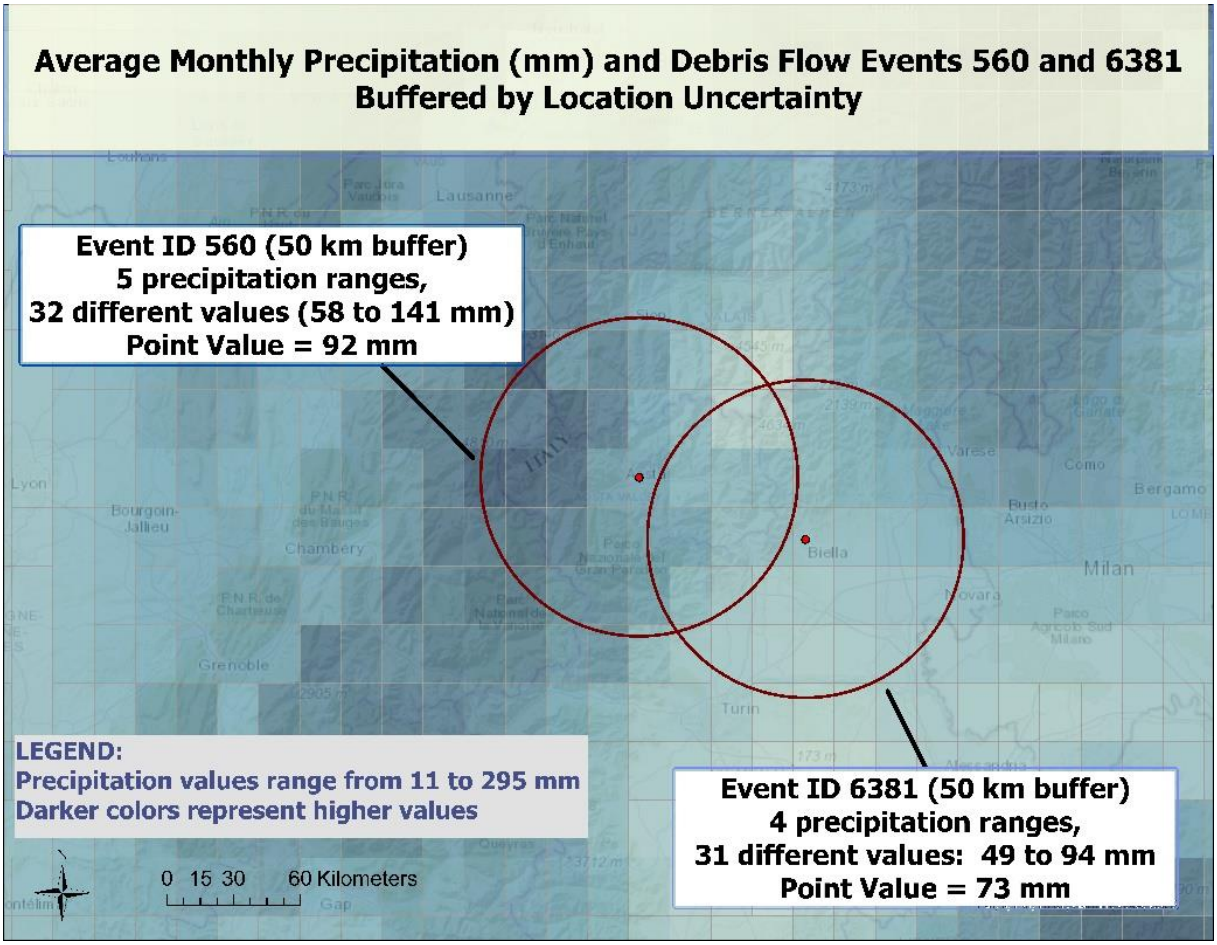


Figure 8.4 Map overlay of monthly average precipitation (mm) raster and debris flow event locational uncertainty buffers, highlighting event 560 and 6381, both with 50 km locational uncertainty.

Each color cell represents an associated precipitation (mm) at that cell. Base map is from ArcGIS®, the intellectual property of Esri, used herein under license. Copyright © Esri.

8.2.2 Impact of locational uncertainty on fault density as a predisposing factor – Europe

According to the "Original" MaxEnt model, the fault density environmental layer provided the second highest percent contribution (gain), at 27.6%, and was the second most significant factor in the jackknife test with most significance as the only variable. . Fig. 8.5(a) depicts the fault density in western Europe with debris flow locational uncertainty buffers. Table 8.4 and Fig. 8.5(b) depict the point fault density value versus the range of values within sample events 560 and 6381.

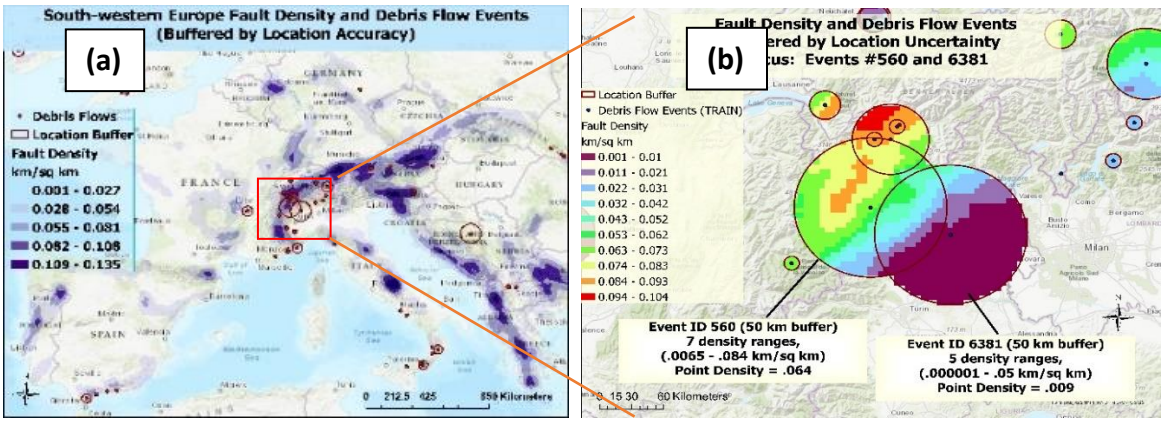


Figure 8.5 5 Fault density (km/sq km) and debris flow event locational uncertainty buffers

(a) Fault densities in western Europe. (b) Fault density (Natural Break (Jenks) classification – 10 breaks) and debris flow event locational uncertainty buffers, highlighting events 560 and 6381, both with 50 km location uncertainties. Base maps and imagery are from ArcGIS®, the intellectual property of Esri, used herein under license. Copyright © Esri.

Event ID	560	6381
Fault density value at point (km/sq km)	.0065 - .0084	.000001 -. 05
Locational uncertainty buffer	50 km	50 km
Number of different Natural Breaks (Jenks) categories within buffer	7	6
Fault density range of values within buffer	.0050 - .0840	0 - .059

Table 8.4 Example impact of uncertain location buffered area vs point values of fault density, for sample events 560 and 6381.

8.2.3 Impact of locational uncertainty on soil as a predisposing factor - Europe

The environmental layer with the third highest percent contribution in the "Original" MaxEnt model is soil type with a gain of 8.6% and tied with fault density as the second most significant factor in the jackknife test with most significance as the only variable. 35% of the 93 buffered events had from two to seven different soil classification types. Fig. 8.6 depicts the soil classifications within each event buffer in a partial view of western Europe, highlighting sample events 560 and 6381. Table 8.5 presents their point value soil classifications versus the range of classification values within their buffered locations.

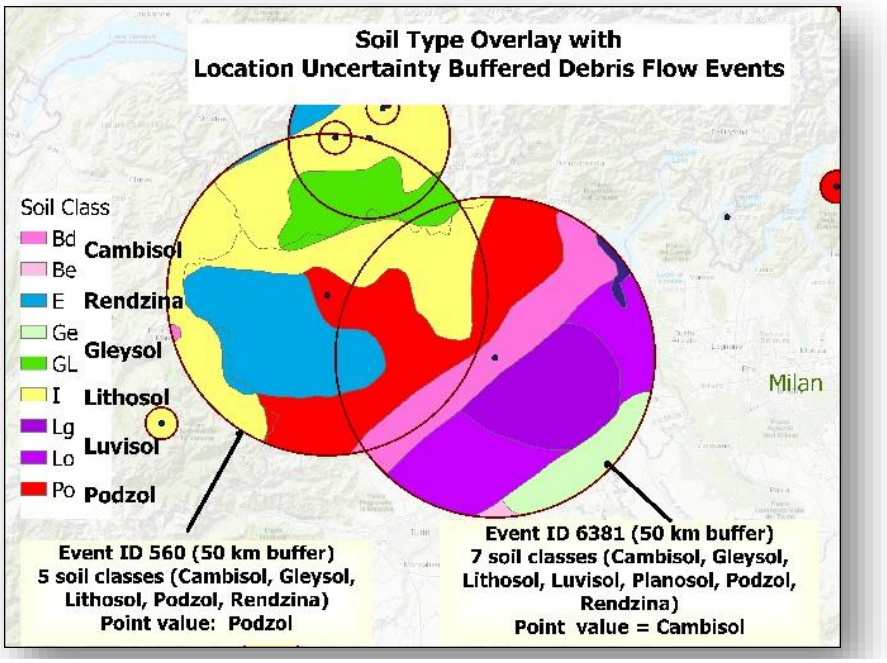


Figure 8.6 Events 560 and 6381 buffered by location uncertainty, with soil classification within each buffer. Base map is from ArcGIS®, the intellectual property of Esri, used herein under license. Copyright © Esri.

Event ID	**560**	**6381**
Soil classification point value	Podzol	Cambisol
Locational uncertainty buffer size	50 km	50 km
Number of different soil classes within buffer	5	7
Soil classes within buffer	Cambisol, Gleysol, Lithosol, **Podzol**, Rendzina	**Cambisol**, Gleysol, Lithosol, Luvisol, Planosol, Podzol, Rendzina

Table 8.5 Uncertain location buffered area vs point values of associated soil factor class, for sample events.

The **bolded soil class** in the list of "Soil classes within buffer" represent the soil class associated with the point location.

8.2.4 Impact of factor uncertainties on susceptibility results - Europe

Susceptibility maps were generated by MaxEnt for the models, "Original", "93 Random", "5 km", "LTE 5 km", and "LTE 1 km" (Fig. 8.7(a),(b),(c),(d), and (e), respectively). The grid cell resolution of the Europe susceptibility maps is 3.36 km. The area coverage for each susceptibility classification for each model is shown in Table 8.6.

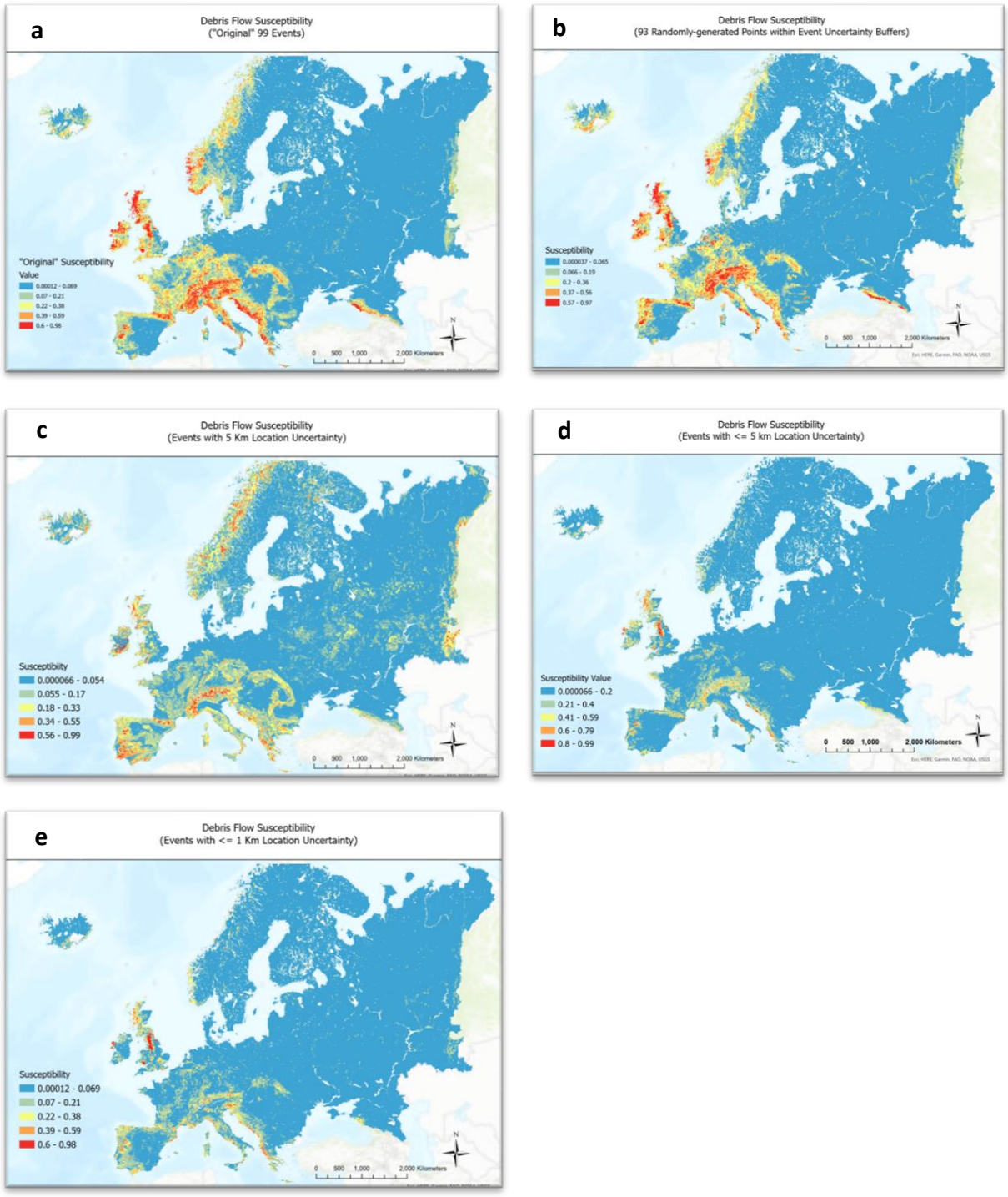


Figure 8.7 MaxEnt debris flow susceptibility maps of Europe. (a) "Original" (b) "93 Random" (c) "5 km" (d) "LTE 5 km" (e) "LTE 1 km". Warmer colors represent higher susceptibilities.

There is close agreement in the AUC and the most significant factor contributions (precipitation, fault density, soil type) in the "Original" and "93 Random" models, and little to no agreement in the remaining models (Table 8.6). Although the "Exact" model has verified locations and the highest AUC, results cannot be used to represent the European continent due to the statistically small number of such events (five) and their localized distribution (four in the United Kingdom and one in Ukraine) (Fig. 8.8), hence the "Exact" model is not further considered.

Variable/Model	"Original" Percent contribution (99 events)	"93 Random" Percent contribution (93 events)	"5 km" Percent contribution (31 events)	"LTE 5 km" Percent contribution (55 events)	"LTE 1 km" Percent contribution (24 events)	"Exact" Percent contribution (5 events)
AUC	0.891	0.893	0.857	0.896	0.921	0.930
precipitation	42	37.4	1.4	30.2	11.9	3
fault density	27.6	29.7	0.5	24.8	10.3	0.1
soil type	8.6	10.5	18.6	13.6	25.8	75.3
landcover	4.7	7.9	17	8.9	18.8	2.3
climate	4.5	2.8	1.5	8.1	11.1	1.4
lithology	2.1	1.6	11.5	4.1	3.1	0.5
soil thickness	2.4	1.8	21	3.3	4	13.3
landform	4.9	3.2	21.6	3.1	10.4	2.2
elevation	0.2	1.6	0	2.1	1.1	1.5
drainage	2.3	2.9	0.2	1.6	1.5	0.2
topsoil % clay	0.1	0.2	1.5	0.1	1.5	0.2
depth to bedrock	0	0.1	5.0	0.1	0.4	0
aridity	0.5	0.3	0	0.1	0.1	0

Table 8.6 AUC and factor contributions for each of the six Europe MaxEnt models.

Figure 8.8 Distribution of the five European debris flow events with a location accuracy of “exact”.

There is little similarity in the predisposing factor classes, identified by the remaining five models, as being associated with debris flow susceptibility , Table 8.7.

Factor /Model	"Original"	"93 Random"	"5 km"	"LTE 5 km"	"LTE 1 km"
precipitation (monthly average mm 1970-2000)	275-300	300-325	260-300	*260-280*	275-300
fault density (km/sq km)	0.02-0.14	0.01-0.02	0 - 0.132	0.13-0.14	0.01-0.14
soil type	Gleysol	Fluvisol, Gleysol	Ferralsol	*Fluvisol, Gleysol*	Gleysol
landcover	urban	urban	Grass/shrubland	urban	sparse vegetation
climate (Köppen-Geiger)	Dfa - hot summer humid continental climate	BSk - semi arid steppe	Csa - Mediterranean hot summer climate and Cfc - subpolar oceanic climate	Csa - Mediterranean hot summer climate and Cfc - subpolar oceanic climate	Csa - Mediterranean hot summer climate
lithology	unconsolidated sedimentary (su) and plutonic basic (pb)	intermediate volcanic (vi)	Sedimentary mixed (sm)	(su) and (vi)	basic plutonic (pb)
soil thickness (m)	0-2.5	0-1	0-1	*0-2*	0-2.5
landform	Plains on sedimentary lithology	Humid plains on sedimentary lithology	Plains in alpine system	*Plains on sedimentary lithology*	Plains in alpine system
elevation (m)	3250-3500	0- 3500	0- 3500	0-500	3250-5000
drainage	"Very poor"	"Imperfectly"	From "Imperfectly" to "Well"	"Imperfectly"	"Moderately well"
topsoil %clay	5-18	20	55-60	*20*	0-22
depth to bedrock (cm)	0	114	0	0	0-2000
aridity (dimensionless)	~1300 (Arid)	~1300 (Arid)	>= 15000 (Humid)	>= 60000 (Humid)	~15000 (Humid)

Table 8.7 Comparison of factor classes with highest significance to susceptibility in four models.

The susceptibility differences of each of the five models are highlighted through pixel difference maps showing the difference between the susceptibility classification and locations of the "Original" model minus "93 Random, "5 km", "LTE 5 km" and "LTE 1 km" models, Fig. 8.9 (a), (b), (c), and (d), respectively. Fig. 8.9 (e) provides a detailed perspective of the large range of susceptibility value differences associated with Events 560 and 6381. Even within a 10 km or 5 km buffer around the points there is significant variability in the susceptibility values. Figure 8.10 and Table 8.8 present the large variance in susceptibility values for Events 560 and 6381 in the "Original" model as viewed within the 50 km location uncertainty buffer, and a glimpse at the significant variance within 10 km and 5 km of their documented event locations. A comparison of the susceptibility classification area coverage for all five Europe models is

presented in Table 8.9. The susceptibility results (0 to 1) are classified using five equal-interval breaks, and classified as Very Low, Low, Moderate, High, and Very High.

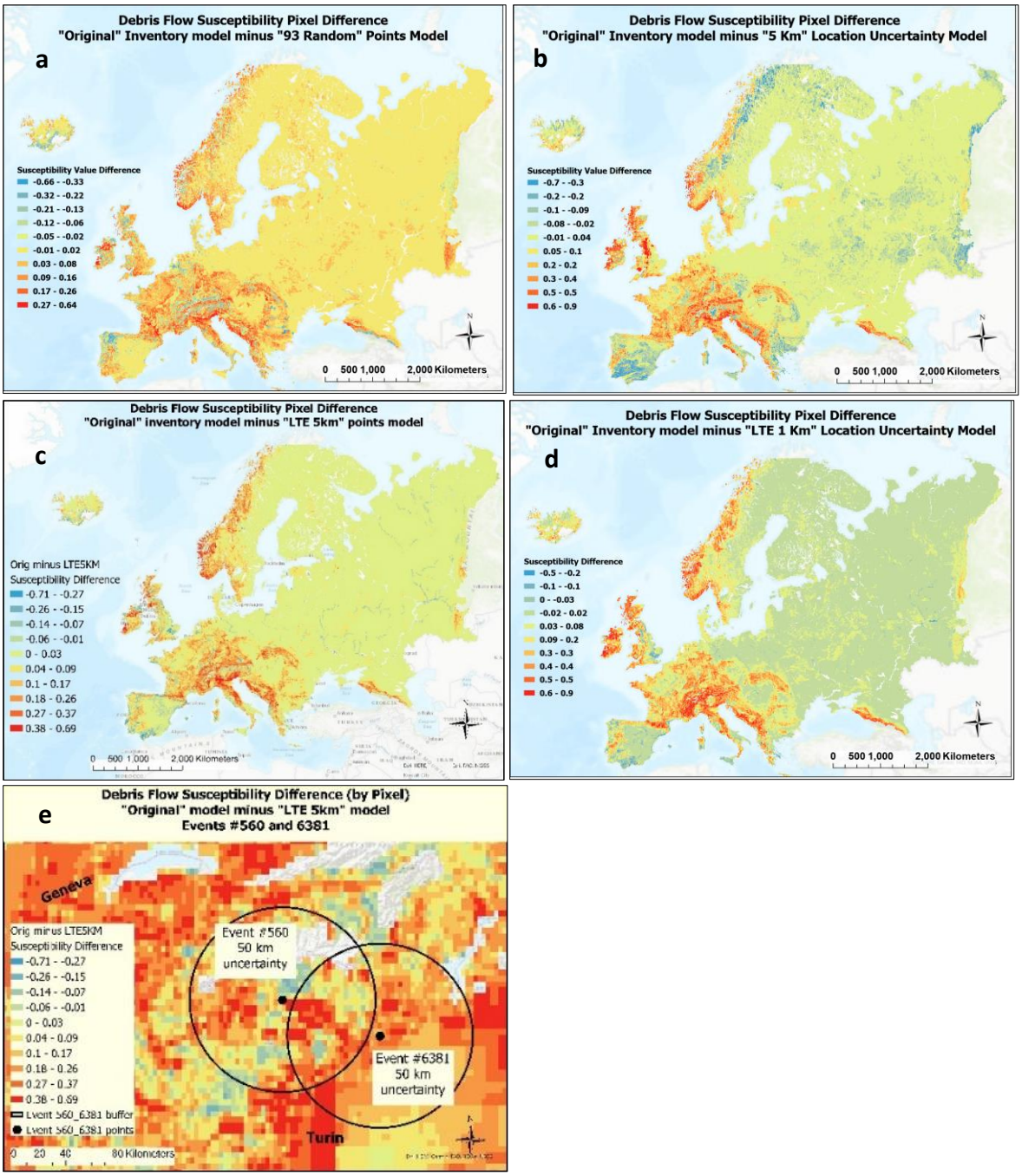


Figure 8.9 Model susceptibility value pixel differences.

(a) "Original" minus "93 Random", (b) "Original" minus "5 km", (c) "Original" minus "LTE 5 km", (d "Original" minus "LTE 1 km", (e) Map 8.9(c) zoomed into area around the Italian Alps for visual enhancement of susceptibility disparities. Warmer colors represent areas of higher susceptibility values in "Original" model. Base maps are from ArcGIS®, the intellectual property of Esri, used herein under license. Copyright © Esri.

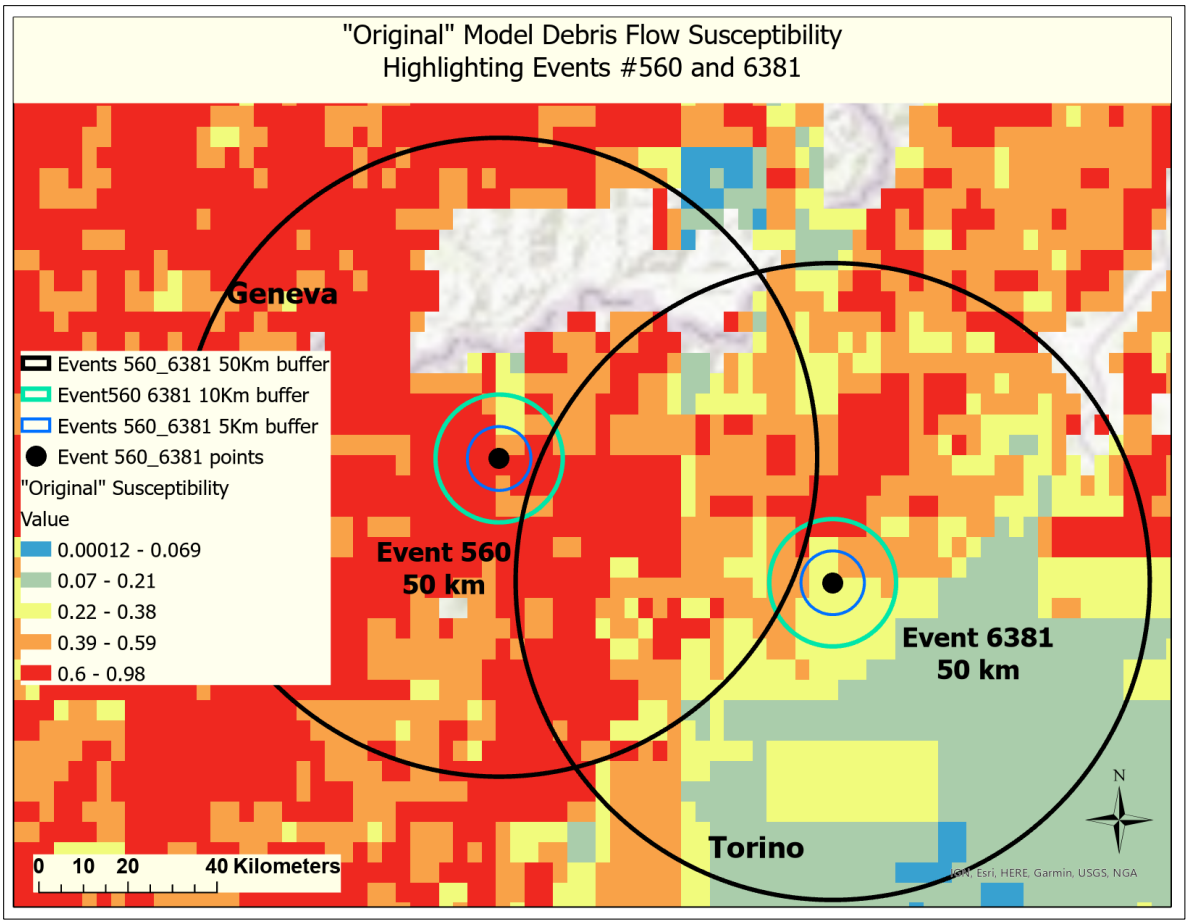


Figure 8.10 "Original" model susceptibility highlighting Events 560 and 6381 with buffers representing their 50 km location uncertainty attribute and at 5 and 10 km.

"Original" Inventory Model Susceptibility Values	Event 560	Event 6381
At Event Point	.62	.35
Lowest Value in 50 km Buffer	.07	.06
Highest Value in 50 km Buffer	.89	.79
Lowest Value in a 10 km Area	.32	.26
Highest Value in a 10 km Buffer	.64	.68
Lowest Value in a 5 km Buffer	.24	.28
Highest Value in a 5 km Buffer	.89	.52

Table 8.8 "Original" model susceptibility values for Events 560 and 6381 within 50 km, 10 km, and 5 km buffers.

Europe MaxEnt Models	"Original"	"5 km"	"LTE 5 km"	"LTE 1 km"	"Exact"
AUC	**0.891**	**0.857**	**0.896**	**0.921**	**0.93**
Susceptibility Value and % Area					
Very Low	84.66%	89.08%	91.29%	96.36%	99.44%
Low	8.54%	8.04%	5.78%	2.63%	0.41%
Medium	4.31%	2.18%	1.99%	0.69%	0.02%
High	1.98%	0.61%	0.74%	0.21%	0.01%
Very High	0.50%	0.10%	0.21%	0.11%	0.12%
Medium to Very High	6.80%	2.89%	2.93%	1.01%	0.15%

Table 8.9 Percent area coverage of susceptibility (based on pixel count) per susceptibility classification for each of five Europe MaxEnt models.

8.2.5 Global Susceptibility Model Comparisons

The global inventory has a location uncertainty from 1 km to 250 km, with the largest percentage (33.3%) having an uncertainty of 5 km (Table 8.10) and only 7.6% with an exact known location.

Global debris flow locational uncertainty		
Uncertainty	Frequency	%
exact	174	7.6%
1 km	620	27.0%
5km	763	33.3%
10km	277	12.1%
25km	240	10.5%
50km	125	5.4%
100km	9	0.4%
250km	4	0.2%
unknown	82	3.6%
Total	**2294**	

Table 8.10 Breakdown of global debris flows and associated locational uncertainty

Global MaxEnt susceptibility models for “Original”, “5 km”, “LTE 5 km”, “LTE 1 km”, and “Exact” are presented in Figs. 8.11 (a), (b), (c), (d), and (e), respectively. The grid cell resolution of the susceptibility maps is 8.3 km. Table 8.11 is a breakdown of the area coverage by susceptibility classification for each model.

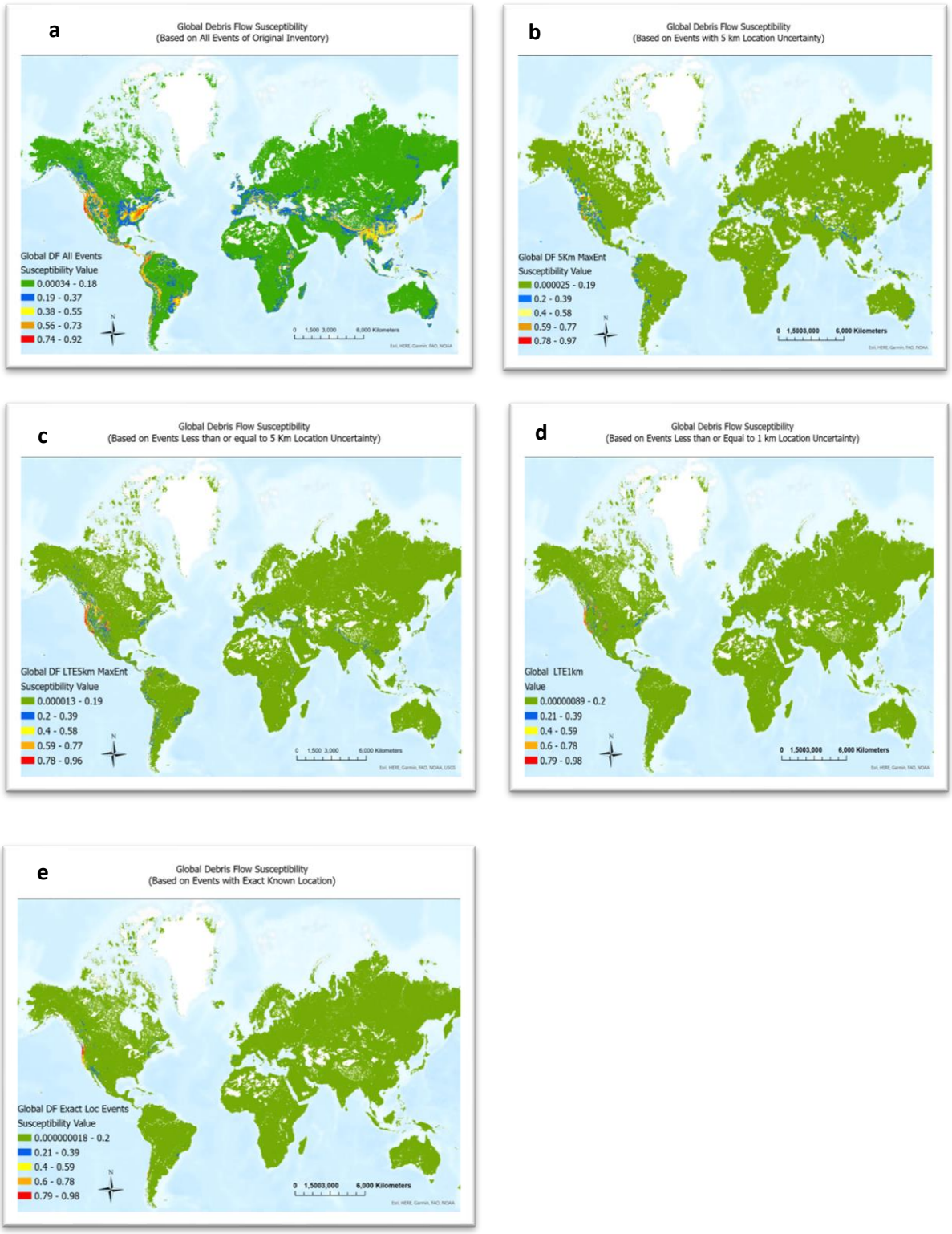


Figure 8.11 Global MaxEnt susceptibility models. (a) "Original", (b) "5 km, (c) "LTE 5 km", (d) "LTE 1 km", and (e) "Exact".

Global Models	"Original"	"5 km"	"LTE 5 km"	"LTE 1 km"	"Exact"
Number of Events	**2294**	**763**	**1557**	**794**	**174**
% Total	100.0%	33.3%	67.9%	34.6%	7.6%
AUC	**0.886**	**0.932**	**0.92**	**0.947**	**0.904**
Very Low	79.00%	96.41%	95.83%	98.34%	99.19%
Low	12.47%	2.57%	2.71%	1.03%	0.47%
Moderate	5.88%	0.75%	0.93%	0.34%	0.19%
High	2.29%	0.23%	0.38%	0.18%	0.09%
Very High	0.36%	0.04%	0.15%	0.10%	0.05%
Moderate to Very High	8.53%	1.01%	1.46%	0.62%	0.34%

Table 8.11 AUC and Area coverage (pixel count) percentage of debris flow susceptibility by model.

As was true with modeling Europe, the "LTE 5 km" global model is the best susceptibility model based on comparisons of AUC, and number and distribution of the model events. "Exact" and "Original" models were discarded based on the small number of "Exact" events and their distribution, and the lack of location fidelity in the "Original" events. "93 Random" is discarded as it was merely a model of fabricated events for demonstration purposes. Figure 8.12 presents the distribution of the 2294 global debris flow events and their location uncertainty values. A difference map of the "Original" global MaxEnt model minus the global "LTE 5 km" model is presented in Fig. 8.13.

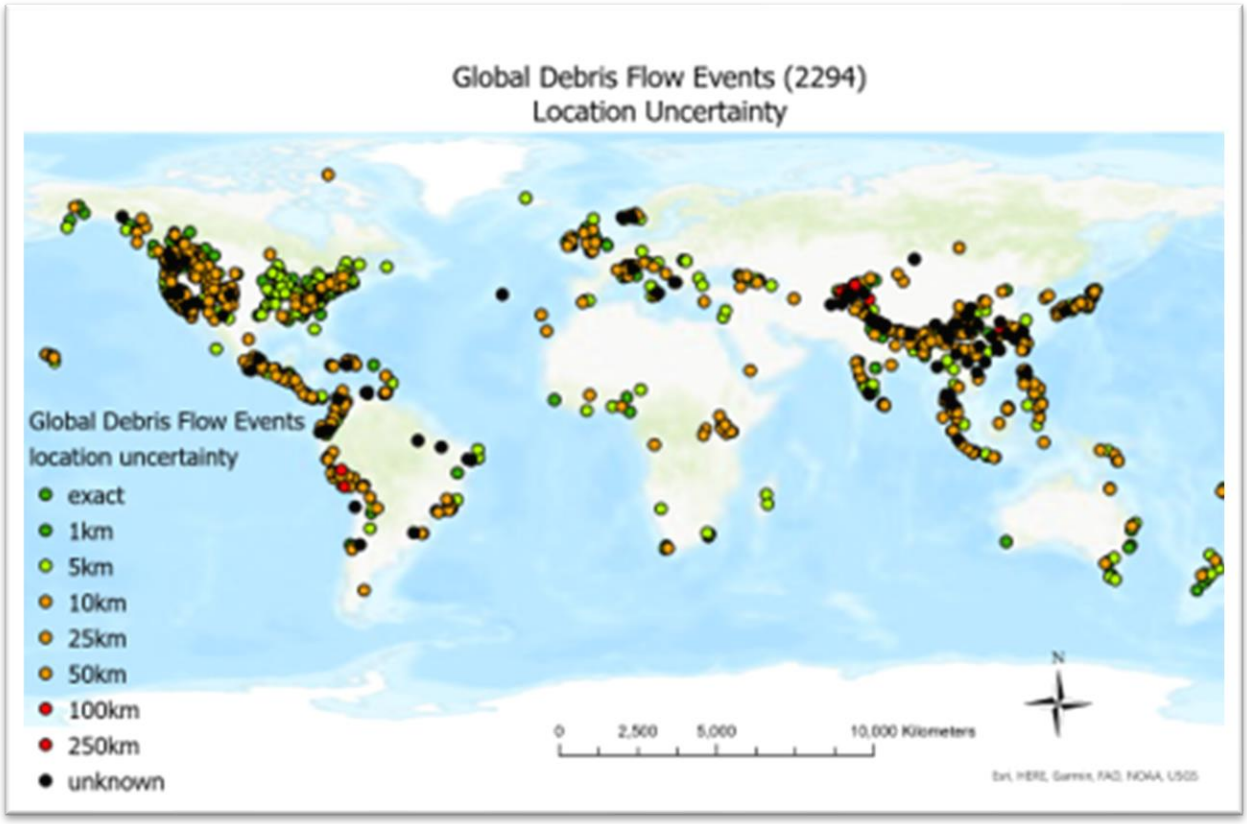


Figure 8.12 Distribution and location uncertainty of 2294 global debris flow events (model training data).

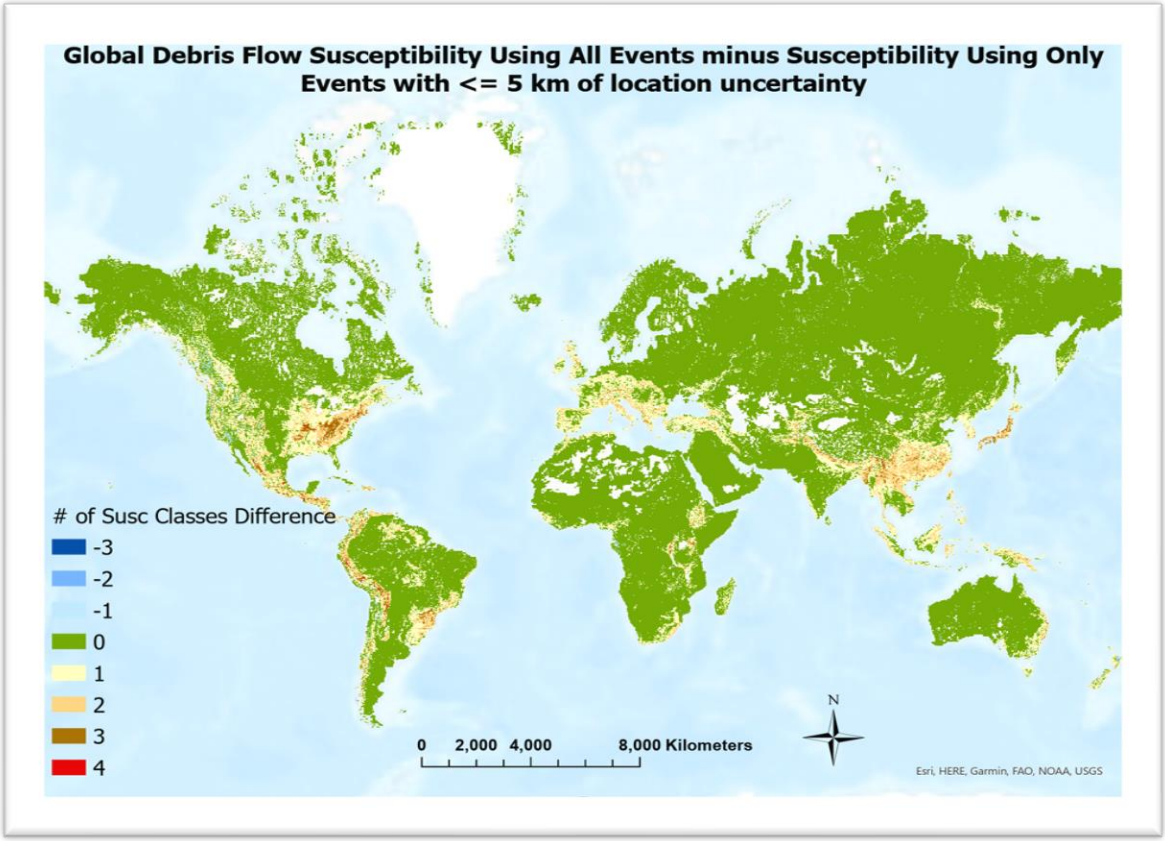


Figure 8.13 Susceptibility difference between "Original" and "LTE 5 km" models.

Warmer colors represent areas where the "Original" model susceptibility classification is greater than the "LTE 5 km" model. Each color is the number of classifications by which the two models differ.

8.3 Discussion

Three general problems are associated with spatial data uncertainty, the first is determining what uncertainties exist, the second is accounting for those uncertainties, and the third is how to represent those uncertainties in our models and/or reasonably adjust the representation of susceptibility results, to account for them.

Identifying specific factor classes (e.g., Cambisol soil class) which have a dominant association with past debris flow events are essential in preparing debris flow susceptibility maps across an area of study (Meten et al., 2015). The impact of locational uncertainty in accurately selecting those predisposing factor classes, which have the most impact on debris flow susceptibility,

varies depending on the scale at which one is performing the analysis, the resolution of the factor itself, the extent of location uncertainty of the debris flow events being used to train a susceptibility model, and the risk associated with an incorrect model (Ardizzone et al., 2002; Carrara et al., 1992). The best models will result from a large number of accurately located and well-distributed debris flow events. However, due to event location uncertainties associated with the landslide inventory, noteworthy differences arose in factor and factor class significance for the environmental layers studied.

The unsuitability of the original debris flow inventory is demonstrated through an inspection of two events each with a 50-km location uncertainty, and the variances in the most significant predisposing factor classes (precipitation, fault density, soil types), as determined through MaxEnt modeling of the Europe continent. The associated precipitation values range from 58 to 141 mm for event #560 and from 49 to 94 mm for event #6381, where the point values were 92 mm and 73 mm, respectively. The fault densities ranged from .0050 to .0840 km/sq km for buffered event #560 and from 0 to .059 km/sq km for event #6381. The point values were .0065 to .0085 and .000001 to .05 km/sq km, respectively. Four soil types (Cambisol, Gleysol, Lithosol, Rendzina) are found within the buffered location of event #560 in addition to Podzol which is identified at the event point location. Six soil types (Gleysol, Lithosol, Luvisol, Planosol, Podzol, Rendzina) are found within the buffered location of event #6381, in addition to Cambisol at the point location. The precipitation point values are almost the mean of the range of values found within the buffered zones for both points; the fault density for event #560 is off by a factor of ten on the high end but the point value for event #6381 is comparable to its buffered range; for both points there is a large discrepancy of soil types between the point and buffered values. Although the impact of location uncertainty (at least at the 50 km buffered examples) on precipitation factor classes is negligible, the impact on soil class is significant. Of the soil classes found in the 50-km buffer there are significant differences which would lead to differing considerations about soils associated with debris flows. For example, Cambisols typically are medium- to fine-grained alluvial, colluvial, aeolian deposits; Lithosols consist of partially weathered rock fragments, usually on steep slopes; Planosols develop on smooth flat uplands with strongly leached upper layers over compacted clay or silt; Podzols are typical of coniferous or boreal forests; Gleysols are wetland soils; Luvisols are characterized by a higher clay content. Each of these soil types tell a different story about debris flow susceptibility.

With substantial factor class variances demonstrated within the event buffers, the resulting susceptibility map is also questionable. The resulting susceptibility map, based on the "Original" inventory of events, exhibits substantial variances. The susceptibility at event #560 is .62 with a range from .07 to .89 within the 50 km buffer. Event #6381 has a point susceptibility of .35 and a range from .06 to .79 within the 50 km buffer.

In general, it is therefore expected that the larger the locational uncertainty, and the finer the granularity of lateral geographic change of the predisposing factors, the larger the uncertainty of the susceptibility results.

Demonstrating the variable and adverse impact of event locational uncertainty on the factor classes, an appropriate approach to susceptibility modeling to accommodate or minimize the impact was pursued. Random points were generated within each of the buffered locations associated with the 93 events having a location uncertainty other than "exact" or "unknown", and a susceptibility model ("93 Random") developed from these points. This model was developed as an additional exercise to evaluate the validity or randomness of a data inventory with events having from zero to 50 km of location uncertainty, and 6% of events with an unknown uncertainty. The AUC values and factor significance values of the "Original" and "93 Random" models are comparable (Table 8.6) and although there is little similarity in the factor class values, this is further evidence that the "Original" inventory of data may not result in a good representative model of debris flow susceptibility and was not further considered in the analysis. It is also likely that the original inventory of global debris flow events may not be suitable for modeling any continent or for a global debris flow susceptibility model.

Thus, four additional MaxEnt debris flow susceptibility models were developed both for Europe and the world, each based on the level of uncertainty of the events. Models with a location uncertainty of 5 km ("5 km"), less than or equal to 5 km ("LTE 5 km"), less than or equal to 1 km ("LTE 1 km"), and a model utilizing only those events with an exact known location ("Exact") were evaluated.

In addition to a concern for the impact on susceptibility modeling by spatial location uncertainty, one must also be concerned with event distribution, and number of events employed. The "exact" Europe and global models were excluded due to the extremely small number of events

and small percentage of the totals (5% and 7.6%, respectively) and their localized distribution. Four of Europe "exact" events are in the United Kingdom and one in Ukraine. The largest percentage of global events with an "exact" known location are within North America (155 events which are 10.5% of the North American events) and no "exact" events are located in Australia. Table 8.12 presents the "exact" known location global events by continent. A small sample size in a large (continental or global) study area may not cover the variability of the environmental factors (Heckmann, Gegg, Gegg, & Becht, 2014). A larger sample size may include more environmental information about the study area, but conversely an overly large sample size may result in overfitting of the model. Based on the number of events (99 for Europe and 2294 for global) it is not believed that these are overly large sample sizes). An exact number of events deemed as appropriate for the best model is not determined herein, rather a reasonableness factor is employed.

CONTINENT	# "Exact" events	# Total Events	% Total
Africa	2	54	3.70%
Asia	7	468	1.50%
Australia	0	16	0.00%
Europe	5	100	5.00%
North America	155	1473	10.52%
Oceania	2	19	10.53%
South America	3	164	1.83%
	174	**2294**	**7.59%**

Table 8.12 Count of global debris flow events with an "exact" known location.

Of the remaining three Europe and global models, "LTE 1 km" has the highest AUC and smallest number of events. The "LTE 5 km" models have the largest percentage of the total inventories (Table 8.13), and either the second or third highest AUC (Table 8.14).

Model	**"5 km"**	**"LTE 5 km"**	**"LTE 1 km"**
Europe # Events (% Total Events)	31 (31%)	55 (55%)	24 (24%)
Global # Events (% Total Events)	763 (33%)	1557 (68%)	794 (35%)

Table 8.13 Number and percentage of total points by model.

Model	"5 km"	"LTE 5 km"	"LTE 1 km"
Europe AUC	.857	.896	.921
Global AUC	.932	.920	.947

Table 8.14 AUC by model.

The "LTE 5 km" model is chosen as the best model for further continental and global analyses herein, due to the larger number, and percentage of total, of well-distributed events (Fig. 8.14 and Fig. 8.15) and a reasonably high AUC.

Figure 8.14 Geographic distribution of Europe events with <= 5 km location uncertainty.

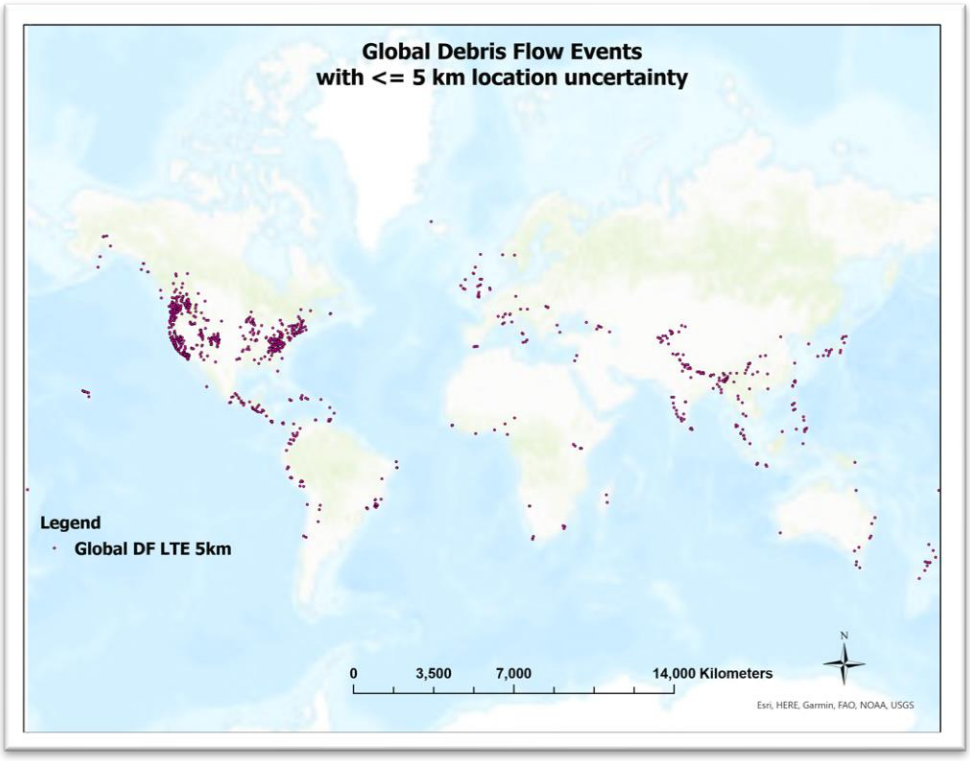


Figure 8.15 Geographic distribution of global events with <= 5 km location uncertainty.

As a further view on susceptibility model choices, and with an attempt to better convey 'uncertainties' to model users, a novel presentation method is developed. Mean and standard deviation maps are developed based on a combination of "Original" and "LTE 5 km" susceptibility maps as input for Europe (Fig. 8.16) and the world (Fig. 8.17). These representations provide the end user with a novel representation of uncertainty associated with the original inventory and the ability to select areas for further detailed study based on areas with a high mean and low standard deviation, from the original, complete set of data ("Original") and the model chosen as a more truthful representation ("LTE 5 km"). An ArcGIS Equal Interval classification scheme was used for susceptibility classification and representation for all models. This scheme divides the low to high values into five equal bins, which are then labeled as "Very Low", "Low", "Medium", "High", and "Very High". Similarly standard deviation and mean maps for the "Original" with the "LTE 1 km" and "Exact" global models are provided in Figures 8.18 and 8.19, respectively.

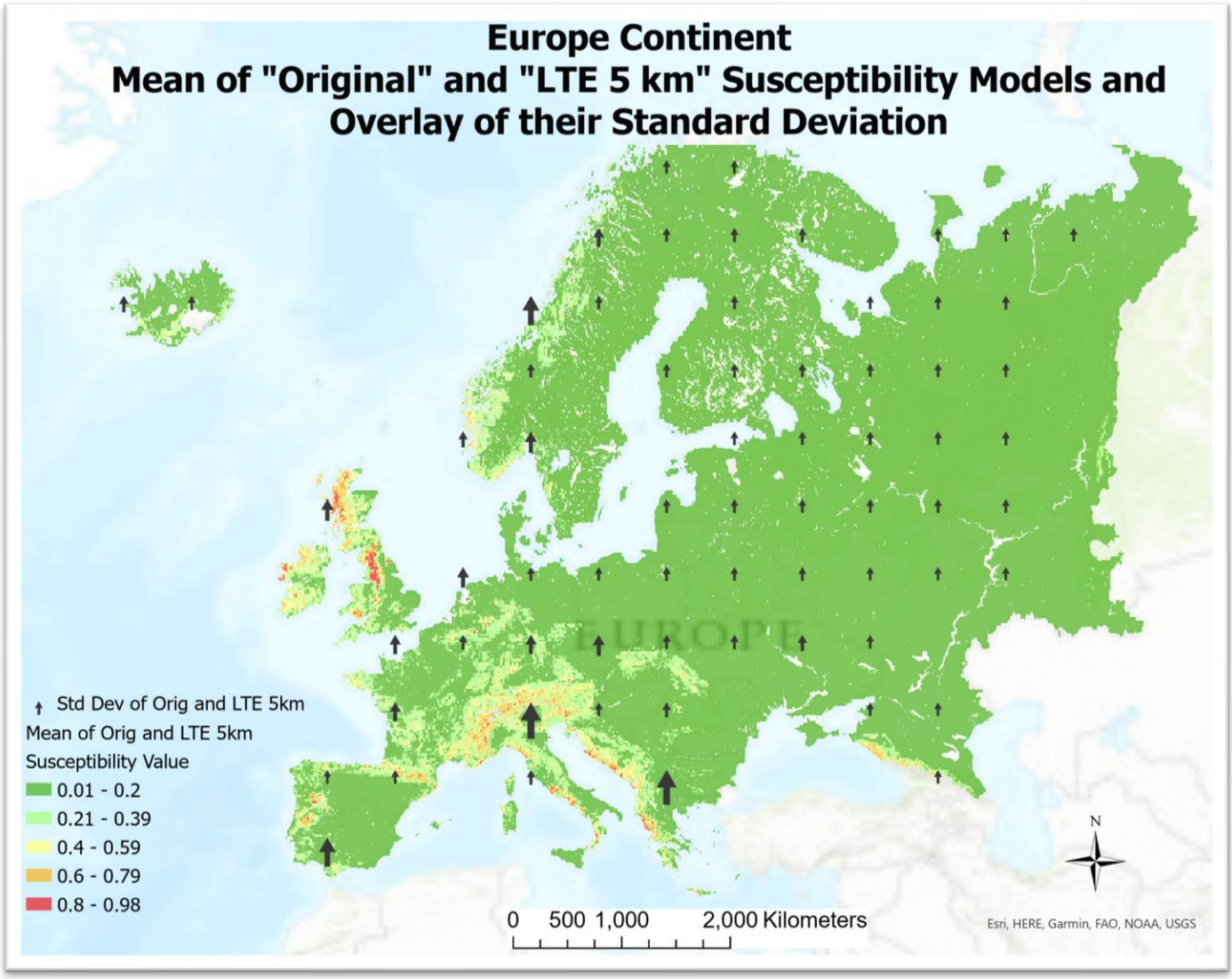


Figure 8.16 Overlay of susceptibility standard deviation and mean for two combined models ("Original" and "LTE 5km").

Standard deviation is represented by a scalar (black arrows), smaller arrows = lower values. Warmers background colors represent higher mean susceptibility.

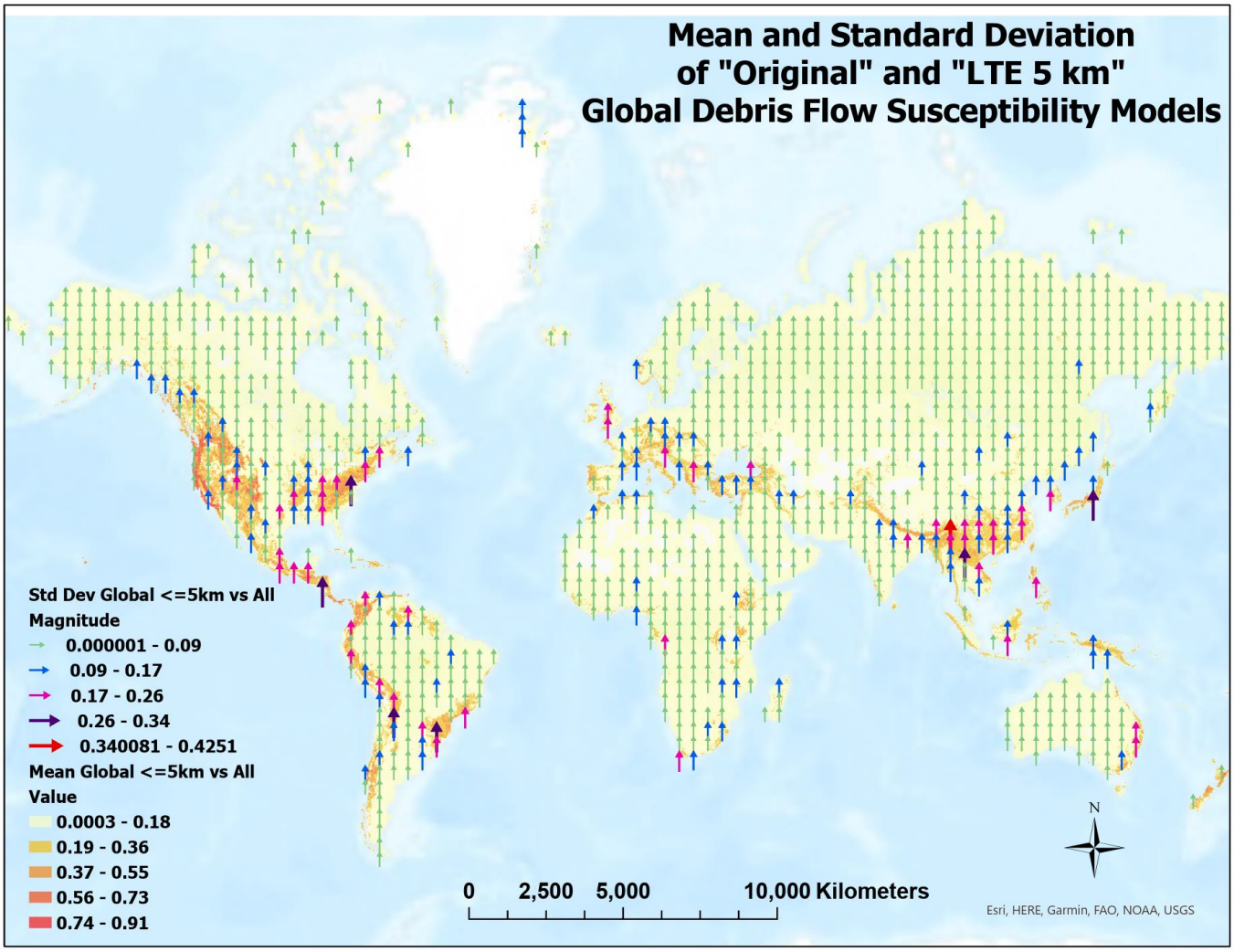


Figure 8.17 Overlay of global susceptibility standard deviation and mean for two combined models ("Original" and "LTE 5 km").

Standard deviation is represented by a scalar (color coded arrows), smaller arrows = lower values. Warmer map colors represent higher mean susceptibility.

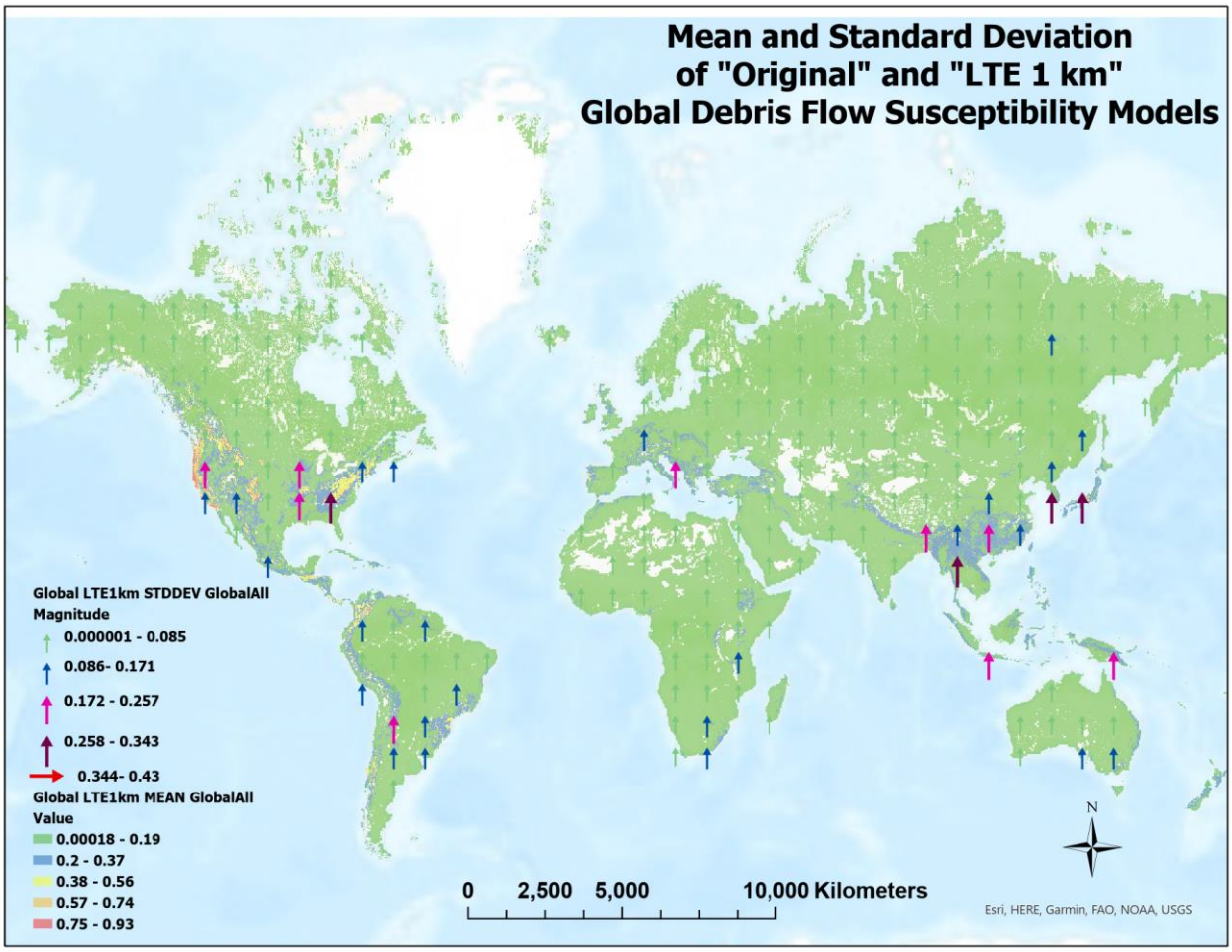


Figure 8.18 Overlay of global susceptibility standard deviation and mean for two combined models ("Original" and "LTE 1 km").

Standard deviation is represented by a scalar (colored arrows), smaller arrows = lower values, etc. Warmer map colors represent higher mean susceptibility

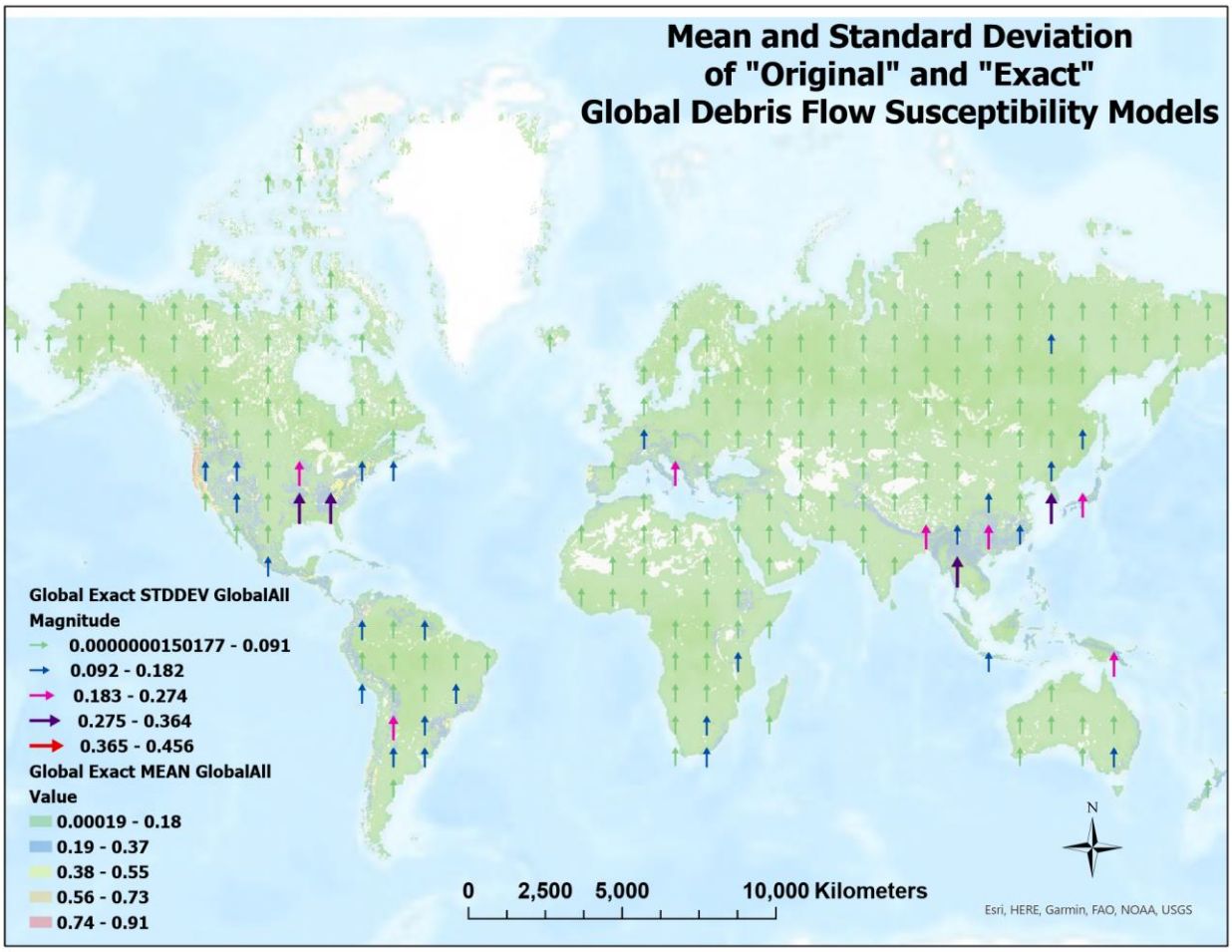


Figure 8.19 Overlay of global susceptibility standard deviation and mean for two combined models ("Original" and "Exact").

Standard deviation is represented by a scalar (colored arrows), smaller arrows = lower values, etc. Warmer map colors represent higher mean susceptibility

8.4 Conclusions

The continent of Europe was used as a case study to understand and highlight the impact of debris flow event location uncertainty with regard to predisposing factors and factor classes and resulting susceptibility zoning. 88.8% of the Europe and global debris flow inventories have a locational uncertainty from 1 to 50 km and 1 to 250 km, respectively, with 6% and 3.6% defined as "unknown", respectively. 82.5% of all global landslide types, in this inventory, have a locational uncertainty from 1 to 250 km, with 4.9% unknown (Table 8.15). It is believed these types of locational uncertainties may be representational of any landslide

inventory whose data is gathered by means other than direct field survey or detailed remote sensing identification tools with a known level of accuracy, and ground control.

Locational Uncertainty Radius (km)	Frequency Distribution (Europe Debris Flows)	% of Total Debris Flow Events	Cum %	Frequency Distribution (World Debris Flows)	% of Total Debris Flow Events	Cum %	Frequency Distribution (World All Landslide Types)	% of Total Landslide Events	Cum %
exact	5	5.1	5.1	174	7.6	7.6	1386	12.6	12.6
1	19	19.2	24.3	620	27.0	34.6	2185	19.8	32.4
5	31	31.3	55.6	763	33.3	67.9	3178	28.8	61.2
10	20	20.2	75.8	277	12.1	79.9	1435	13.0	74.2
25	12	12.1	87.9	240	10.5	90.4	1470	13.3	87.5
50	6	6.1	93.9	125	5.4	95.9	794	7.2	94.7
100	0			9	0.4	96.3	25	0.2	94.9
250	0			4	0.2	96.4	16	0.1	95.1
“unknown”	6	6.1	99.9	82	3.6	100.0	544	4.9	100.0
<null>							2		
Total	**99**			**2294**			**11033**		

Table 8.15 Locational uncertainty frequency distribution of Europe debris flow inventory (left), world debris flow inventory (center), and world inventory of all landslide types (right).

For the Europe continent, five models and three predisposing factors (precipitation, fault density, soil class) were investigated to identify and demonstrate the scope and impact of uncertainty in determining those factor classes most relevant to debris flow susceptibility. First the variances in factor classes within an event’s locational uncertainty buffer were compared to the values associated with the event point location to expose the problem. Then models were developed based on event location uncertainties and evaluated on AUC. The standard deviation and mean of the models based on the original set of events and the model defined on events with less than or equal to a five-kilometer location uncertainty was developed to provide end users with a view of the susceptibility differences, which can facilitate choosing areas of further interest. Using this composite and choosing susceptibility areas with a small standard deviation and high mean, may provide insight to areas which have a higher confidence level.

Utilizing the MaxEnt statistical results of AUC, a model was chosen which would best meet the context, scope, and objectives of this European continent debris flow susceptibility study,

which is to find a set of events and a methodology that can provide the most "truthful" susceptibility results. The "LTE 5 km" model provides a good intermediate result, utilizing fifty-five points, about half of the original data, fairly well distributed across the study area, includes data with a relatively small geographic locational uncertainty from 1 to 5 km, and includes the "exact" known points. This model is an intermediate compromise favoring a slight error of commission over error of omission. Furthermore, by developing susceptibility maps of the mean and standard deviation derived from the "Original" and the "LTE 5 km" models, provides additional tools identifying sites (areas of high susceptibility mean and low standard deviation) for further investigation. Based on the results and conclusions drawn from the Europe models, similar methodologies were applied to the global inventory of debris flow events of varying location uncertainty. The AUC, and number and distribution of events, resulted in the "LTE 5 km" being chosen as the preferred global model.

The cause, nature, and handling of mapping errors (uncertainties) has been the subject of extensive research (Ardizzone et al., 2002; Hunter, 1999; Hunter & Goodchild, 1996; Maffini et al., 1989; Openshaw, 1989; Shi, 2010; Wechsler, 1999; Zufle et al., 2017). Uncertainty and uncertainty handling is context dependent. The combination of uncertainties in data, analytical methods, and the overlay processes, may be additive, multiplicative, and non-linear (Veregin, 1989). If a susceptibility map, at any scale, is to be used to direct further research or mitigation efforts, a confidence level associated with the results is desirable (Soma, 2018). This confidence will be dependent upon many facets of the model, one of which is the reliability of the location of the events used to train the model.

Increasingly, dataset producers are including accuracy (uncertainty) attributions, such as the 'location accuracy' attribute in the NASA landslide inventory (D. Kirschbaum et al., 2015), and the 'epistemic quality' and 'activity confidence' attributes of the GEM Active Faults data (Styron & Pagani, 2020). Utilizing these attributes, when available, provides the researcher insight as to how or if to utilize the uncertainty, based on the context of their project.

Although this project is singularly focused on the impact of landslide inventory locational uncertainty, additional sources of data uncertainty may be inherent in other attributes of the inventory, as well as uncertainties associated both with the spatial assignment and attributes of the environmental factors. The impact of the locational uncertainty will vary according to

the scale of the study area and to the size of the locational uncertainty buffer vis-à-vis the scale of lateral geographic change for a factor class. It would be ideal to create susceptibility maps with an attribute and graphical representation of certainty, but not practicable due to the array, complexity, and relationship of all uncertainties, without modeling across all these variables. According to Goodchild et al 1993 as cited in (Hunter, Goodchild, & Robey, 1994) there are three options for handling uncertainty: (1) omit all reference to it; (2) attach a descriptor to the output; and (3) show samples from the range of possible maps. The latter approach is chosen in this study, that is, to present different models of the susceptibility and allow the user to choose that which is most reasonable for the project at-hand.

In summary, there are several caveats to be noted when working at a continental or global scale. In general, there is no single superior model. The best model is that which best fits the research objective and the error tolerance. While debris flow predisposing factors can be upscaled and generalized for the continental and global scale, the reverse is not valid. The most significant predisposing factors identified at a continental or global scale, through any modeling technique, is a generalized perspective. The predisposing factors identified as most significant are common to debris flows over an extensive geographic scale.

The result of this research demonstrates that there is not a single definitive solution to the uncertainty problem. Including all environmental factors and other data uncertainties (in addition to locational) may help refine the results but is a much more comprehensive and complex model. This research reveals the critical importance of working with data accurately located and applies to landslide analyses of any type. If the location of modeled events is uncertain then any results from such a model will also be uncertain.

9 Future Debris Flow Susceptibility

9.1 Introduction

Climate conditions (precipitation, temperature, aridity) may experience major changes in magnitude and geographic extent in the future. "The frequency and intensity of some extreme weather and climate events have increased as a consequence of global warming and will continue to increase under medium and high emission scenarios (high confidence)" (IPCC, 2019). Climatic variability will be manifested as significantly enhanced rainfall in some areas and marked desertification and enhanced wind erosion in others, operating at scales of tens to hundreds of years (Cherlet et al., 2018; Middleton & Thomas, 1997). The frequency and intensity of droughts are projected to increase in southern Africa, and the Mediterranean region, while frequency and intensity of extreme rainfall events are projected to increase in many other regions across the world.

Both changes in temperature and precipitation affect slope stability and potential debris flow susceptibility (C. Huggel et al., 2012) by increasing sediment availability and triggers resulting from intense and/or prolonged rainfall, rainfall totals, antecedent rainfall, and rain on snowmelt (IPCC, 2019).

Changes in precipitation and temperature levels and patterns are the primary climate changes with a direct impact on soil (types, thickness, composition), geomorphology (slope, hydrography, landform), land use/landcover, and other environmental factors which influence the potential for debris flows. "Magnitude of debris flows could become larger due to larger amounts of sediment delivered to the channels and as a result of the predicted increase in heavy precipitation events." (Stoffel, Tiranti, & Huggel, 2014). These factors can subsequently lead to debris flows in areas of frequent or intense rainfall (Li et al., 2018) (C. Huggel et al., 2012) (Stoffel et al., 2013) (Alvioli et al., 2018) (Gariano & Guzzetti, 2016).

In addition to projecting the future extents of current climates known to be associated with debris flow susceptibility, a more difficult consideration is the potential changes to geomorphologic environmental predisposing factors (Alvioli et al., 2018; Harrison et al., 2019) and therefore geomorphic systems, in the future. Despite the great importance of these climate-induced

geomorphologic changes, they are not considered in this study, thus predisposing factor layers other than climate classification and precipitation, are based on current data.

Researchers have demonstrated the recent impacts of climate change on increased debris flow magnitude and frequency, and change in seasonality in North and South America, Europe, New Zealand, and Asia (Christian Huggel et al., 2012; C. Huggel et al., 2012; C. Huggel et al., 2010; Keiler et al., 2010; Legg, Meigs, Grant, & Kennard, 2014; Moreiras et al., 2021; Stoffel et al., 2014).

Projected climate change impacts on debris flow susceptibility conducted in Asia and Europe indicate that susceptible areas may not increase dramatically, but debris flows may occur more frequently and with a higher magnitude (Alvioli et al., 2018; C. Huggel et al., 2010; Li et al., 2018). This may be due to the projection that annual precipitation may decrease, while the number of heavy rainfall events increase (Donat, Lowry, Alexander, O'Gorman, & Maher, 2016).

The future debris flow susceptibility and vulnerability (Chapter 10) models, developed herein, are based on projected climate and urbanization trends worldwide for the decade 2100. The objective is to identify areas globally with the highest potential for debris flows and future societal vulnerability as indicated by the intersections of high debris flow susceptibility and urbanization. The purpose is not to predict, or provide early warning systems, but rather as a high-level preliminary view of the potential impact to society that can be used to drill-down for more regional and localized research, and planning.

9.2 Model Constraints

There are caveats in future susceptibility modeling. There is not a strict and direct relationship between the climate zones currently associated with debris flow susceptibility, and projected susceptibility based on projected locations of these climate zones. Additional critical factors include changes in precipitation intensity, and seasonality. It is not only the amount of precipitation an area receives, but how anomalous this precipitation is to the region. In general, anomalous variation of rainfall is a key driver of climate change impacts (C. Huggel et al., 2012).

The interactive impacts of various environmental factors on one another, commencing with climate change, are not yet studied, and may have a significant impact on defining future areas of debris flow susceptibility. For example, increased precipitation may affect soil types and thickness, rate of lithologic erosion, slope, geomorphology, etc. Changes in slope may change potential gullying and erosion, landform, etc.

A temporal component is missing from this study, by design. A temporal analysis warrants a separate study due to the complexity of modeling as well as the uncertainty in the source data and the time-periods used in normalizing the data, regression to those time periods when calculating future projections, and uncertainty in the rates of change and overall effect in the longer term projections (Donat et al., 2016).

This study is focused on the future debris flow susceptibility guided exclusively by future precipitation and geographic shifting of associated climate zones; it is an imperfect analysis in the sense that climate classification and precipitation are not the only significant environmental factors which govern debris flow susceptibility. Landuse/land cover can change substantially in the next 50-90 years, due to both climatic and anthropogenic drivers, soils and soil erosion, river incision, and geomorphologic changes due to climate change (Harrison et al., 2019) all of which have an impact on debris flow susceptibility. It is conceivable that the relationship among environmental factors and triggers that are associated with debris flow susceptibility, as known today, could be different in the future considering the possibility of significantly different environmental conditions. In other words, system processes, themselves, may change concurrent with, or as a result of, sustained climate change. Consideration of future changes in other relevant environmental factors is beyond the scope of this project.

Although the models developed herein are relatively simplified vis-à-vis the factors, it is nevertheless a relevant foundational starting point. The impact of the areas defined as potentially susceptible to debris flows may be minimized by reducing the anthropogenic contribution, the only factor over which we have control.

9.3 Methodology

The factors and factor classes employed in the future modeling is derived from the MaxEnt "LTE 5 km" global debris flow susceptibility model which was determined to be the best global model (Chapter 8).

The future models are developed in ArcGIS Pro rather than MaxEnt, as MaxEnt models require historic data as input for modeling. Therefore, for a more equitable comparative analyses between current and future, the "present-day" model is rebuilt in ArcGIS Pro using the same methodology as that employed for the future models.

The present-day "LTE 5 km" Maximum Entropy global debris flow susceptibility model is used to extract those factors with significance values > 1% and their factor classes which exhibit >=90% predicted probability, as suitable debris flow environments. The factors and factor classes, other than climate and precipitation, are used as input to the future models. Representative Concentration Pathways (RCP) 2.6 and 8.5 projected precipitation (Fick & Hijmans, 2017) and Köppen-Geiger climate (Beck et al., 2018) data are used for the future debris flow susceptibility models, extracting a subset based on the climate and precipitation factor classes which exhibited >=90% predicted probability in the "present-day" Maximum Entropy model. Table 9.1 lists the factors and factors classes which met these criteria. Each of the factor layers are weighted based on their "present-day" model Maximum Entropy significance value and summed over each pixel. Each pixel of the resultant susceptibility map is equal to the sum of the weights of all environmental factors present at that pixel, therefore ranging from 0 (no factors present) to 99 (all factors present). Pixels with higher weighted sums represent higher susceptibility. The maps are classified using five equal intervals and qualitative labels of "Very Low", "Low", "Moderate", "High", and "Very High" susceptibility.

Environmental Factor	Factor Weight	Factor Classes Used in Future Model
Slope (deg)	27	10-35
Köppen-Geiger Climate Class	20	Csa, Csb, Dfa
Landform	18	Dry widely spaced mtns in alpine system; dry high tablelands in alpine system
Soil Drainage	9	moderately well
Soil Type	9	Cambisol, Gleysol
Fault Density	6	.02 - 0.14
Land cover	4	Urban, Snow/ice, Needleleaf deciduous, Cropland, Sparse vegetation
Precipitation (Avg mo mm)	2	100-200
Soil Thickness (m)	2	1-3
Lithology	2	pa (acidic intrusive) pb (basic intrusive), pi (intermediate intrusive)

Table 9.1 Predisposing environmental factors, factor classes, and assigned weights for future susceptibility model.

Comparisons are made between future (RCP 2.6 and RCP 8.5) and current models and between the RCP 8.5 and RCP 2.6 future models. The comparisons are developed by subtracting susceptibility values on a pixel-level between two models using ArcGIS Pro "cell statistics" tool.

9.4 Results

The ArcGIS version of current debris flow susceptibility, based on the summed, weighted factor classes from the "LTE 5 km" model, is presented in Fig. 9.1. Figs. 9.2 and 9.3 are the RCP 2.6 and RCP 8.5 future models, respectively.

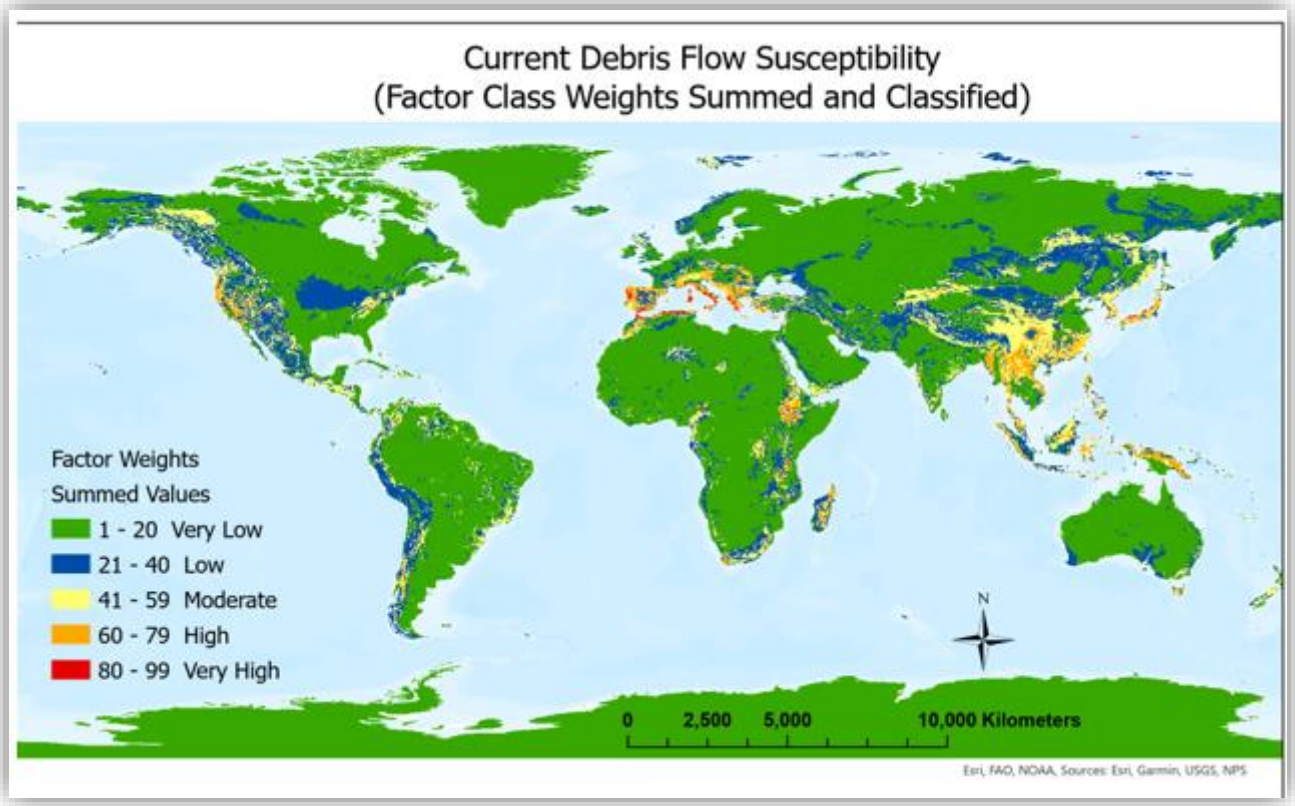


Figure 9.1 ArcGIS version of present day debris flow susceptibility based on summed, weighted MaxEnt factors.

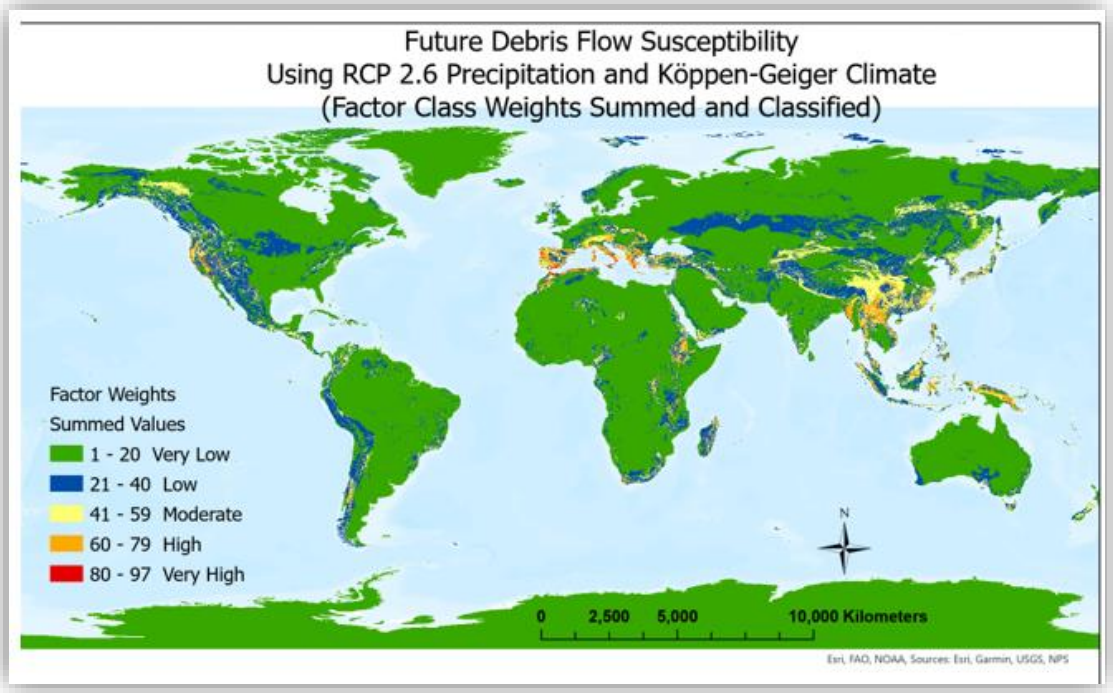


Figure 9.2 Future (RCP 2.6 scenario) debris flow susceptibility based on summed, weighted MaxEnt factors.

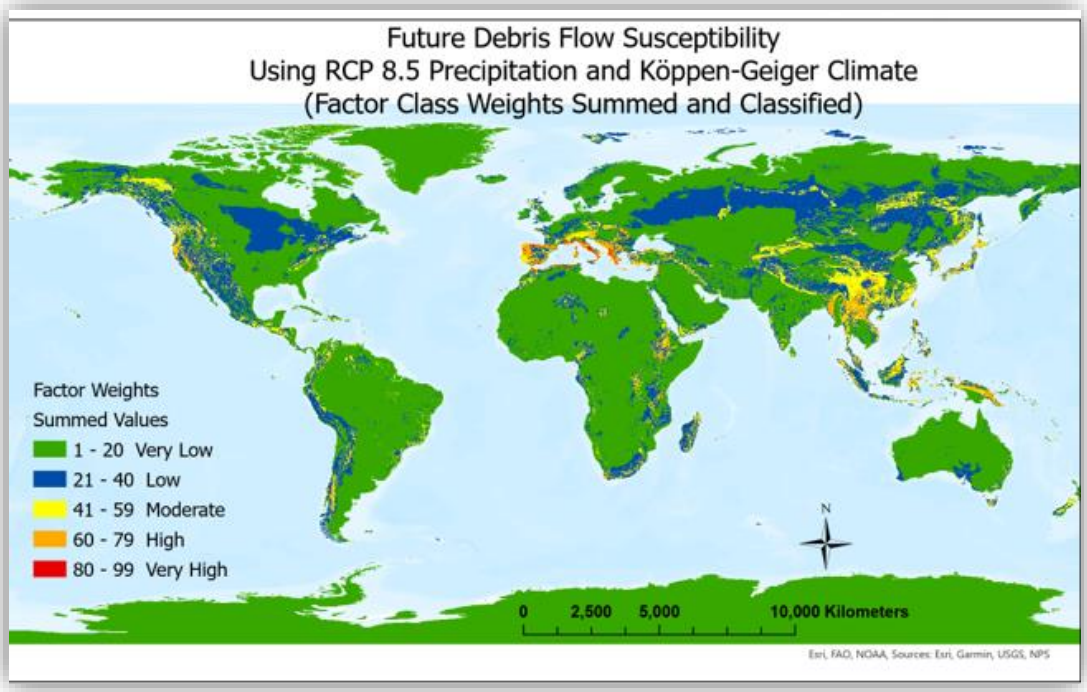


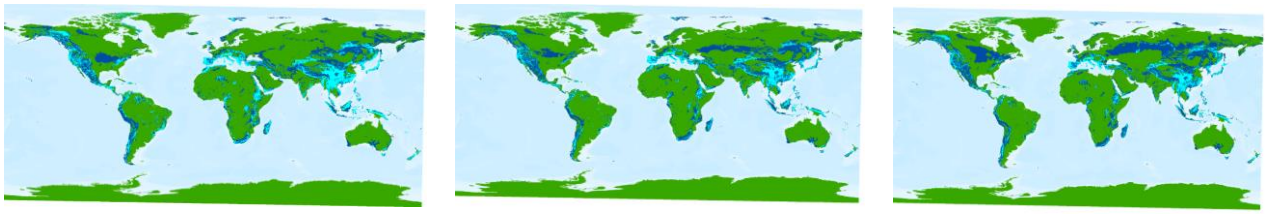
Figure 9.3 Future (RCP 8.5 scenario) debris flow susceptibility based on summed, weighted MaxEnt factors.

The total area of Moderate, High, and Very High debris flow susceptibility zones decreases in both future scenarios as compared to the current model, with slightly more area in the RCP 8.5 than RCP 2.6 model (Table 9.2).

Susceptibility Classification/Model	Current	RCP 2.6	RCP 8.5
Very Low	91.63%	92.47%	90.95%
Low	5.21%	5.24%	6.74%
Moderate	2.54%	1.85%	1.89%
High	0.57%	0.41%	0.38%
Very High	0.05%	0.03%	0.04%
High and Very High	**0.62%**	**0.44%**	**0.42%**
Moderate, High, Very High	**3.16%**	**2.29%**	**2.31%**

Table 9.2 Percent area of debris flow susceptibility classification by model.

There are geographic trends in susceptibility in all models in the Moderate, High, and Very High classifications, but with varying extents (Fig. 9.4). The Moderate to Very High susceptibilities commonly coincide with mountain ranges on all continents, primarily high mountains and scattered mountains as classified by Sayre et al (Sayre et al., 2018) in Fig. 9.5.



Current model RCP 2.6 future scenario RCP 8.5 future scenario

Figure 9.4 Moderate, high, and very high debris flow susceptibility zones (light blue highlights) in current, RCP 2.6, and RCP 8.5 future scenario models.

Mountain environments are vulnerable to mass movements as they provide several of the predisposing environments as well as triggering parameters commonly associated with debris flow susceptibility (Nakileza & Nedala, 2020). They are sources of available sediment, receptors and concentrators of precipitation, and they are composed of slopes of varying degrees, commonly attributed to debris flow susceptibility.

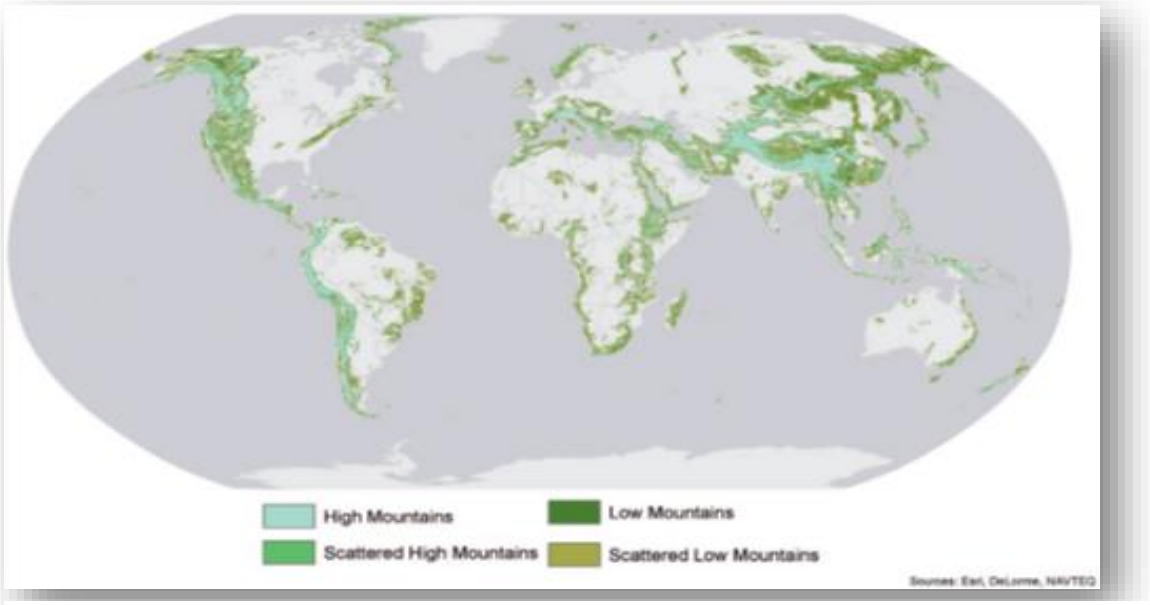


Figure 9.5 Global mountain distribution in 4 classes (high, scattered high, low, scattered low) from Sayre et al.

To discern the geographic shifts/changes in susceptibility, Figs. 9.6, 9.7, and 9.8 model the susceptibility classification differences between RCP 2.6 and present-day; RCP 8.5 and present-day; and RCP 8.5 and RCP 2.6, respectively. Each pixel is evaluated as to whether it increased or decreased in susceptibility classification and by how many classifications.

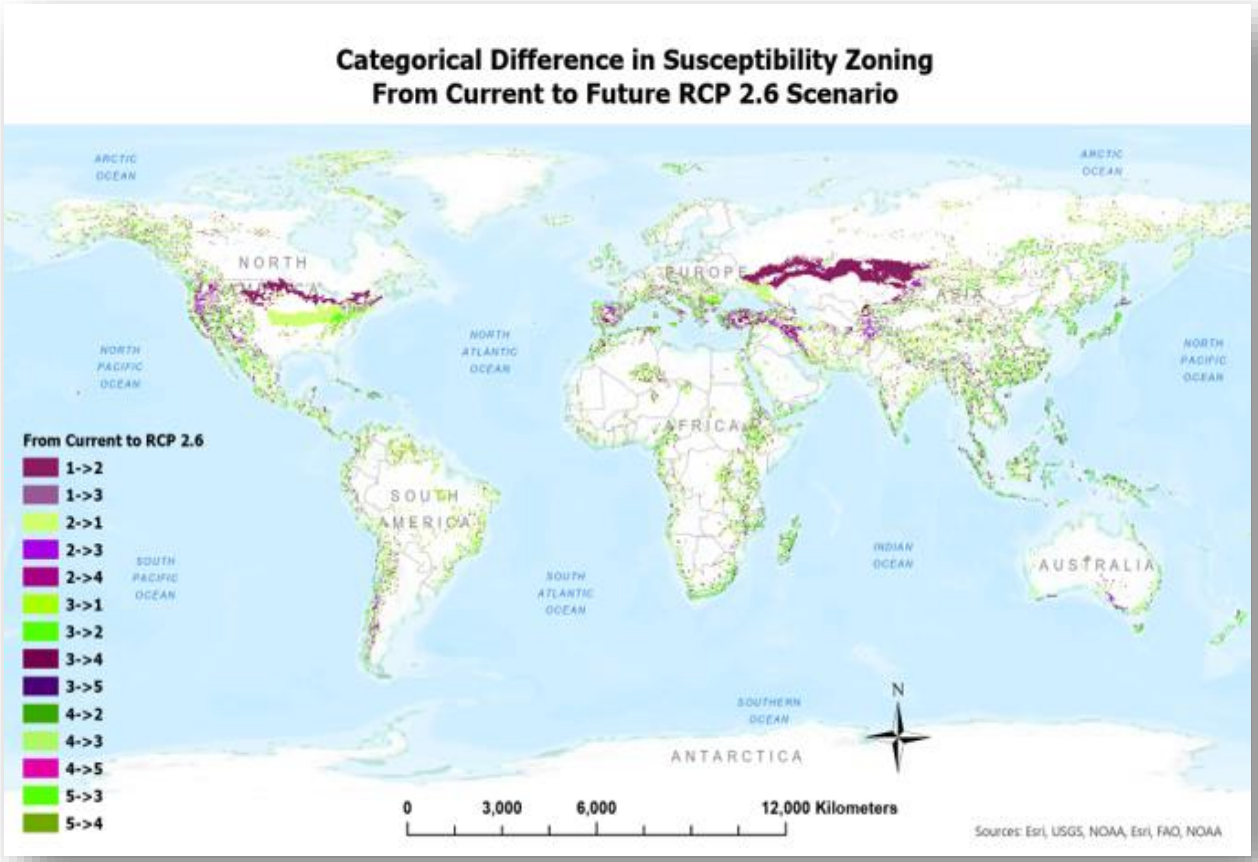


Figure 9.6 Susceptibility zoning classification changes between Current and Future RCP 2.6 models.

Classification mapping is 1 = Very Low, 2 = Low, 3 = Moderate, 4 = High, 5 = Very High.

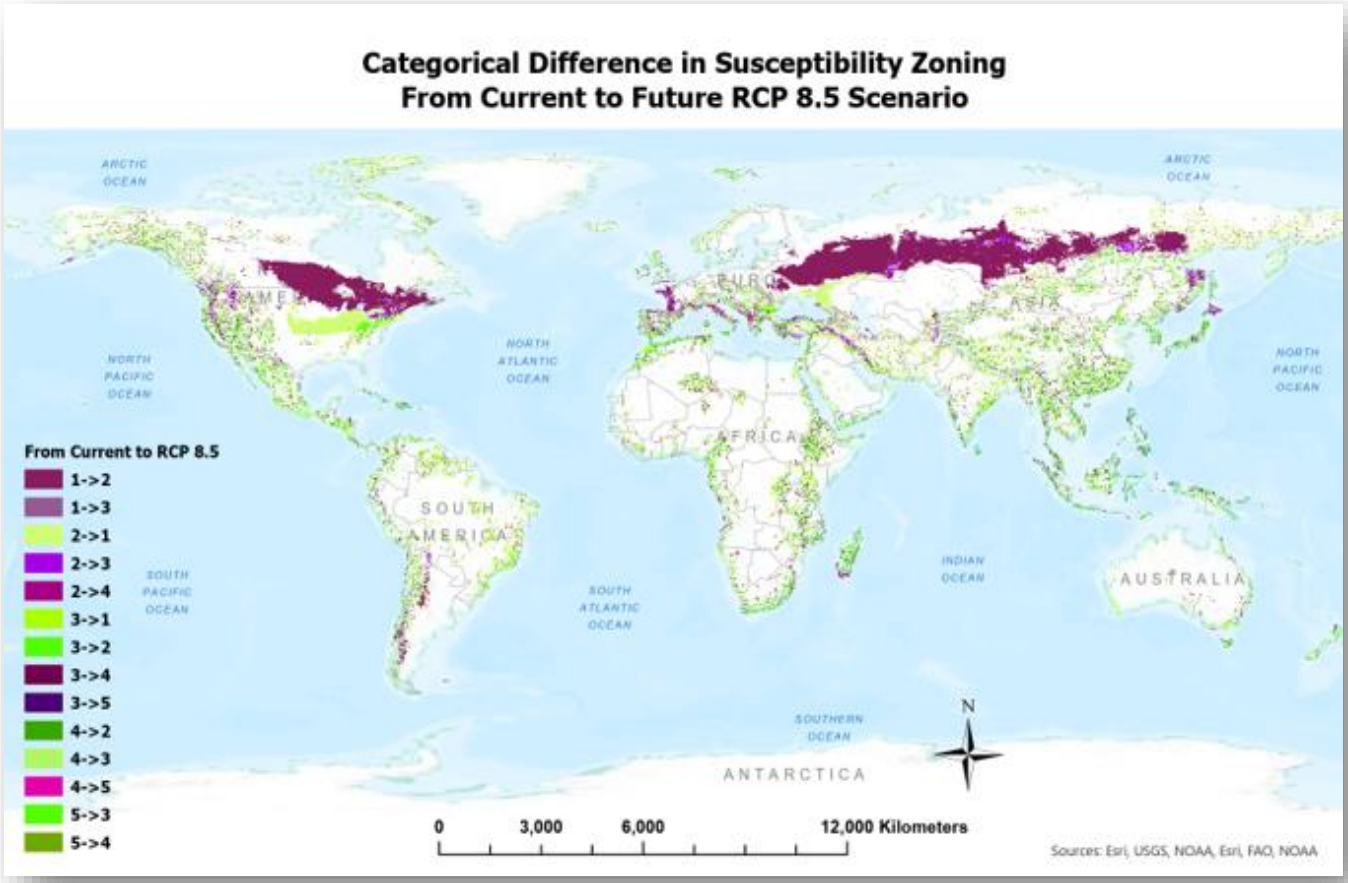


Figure 9.7 Susceptibility zoning classification changes between Current and Future RCP 8.5 models.
Classification mapping is 1 = Very Low, 2 = Low, 3 = Moderate, 4 = High, 5 = Very High.

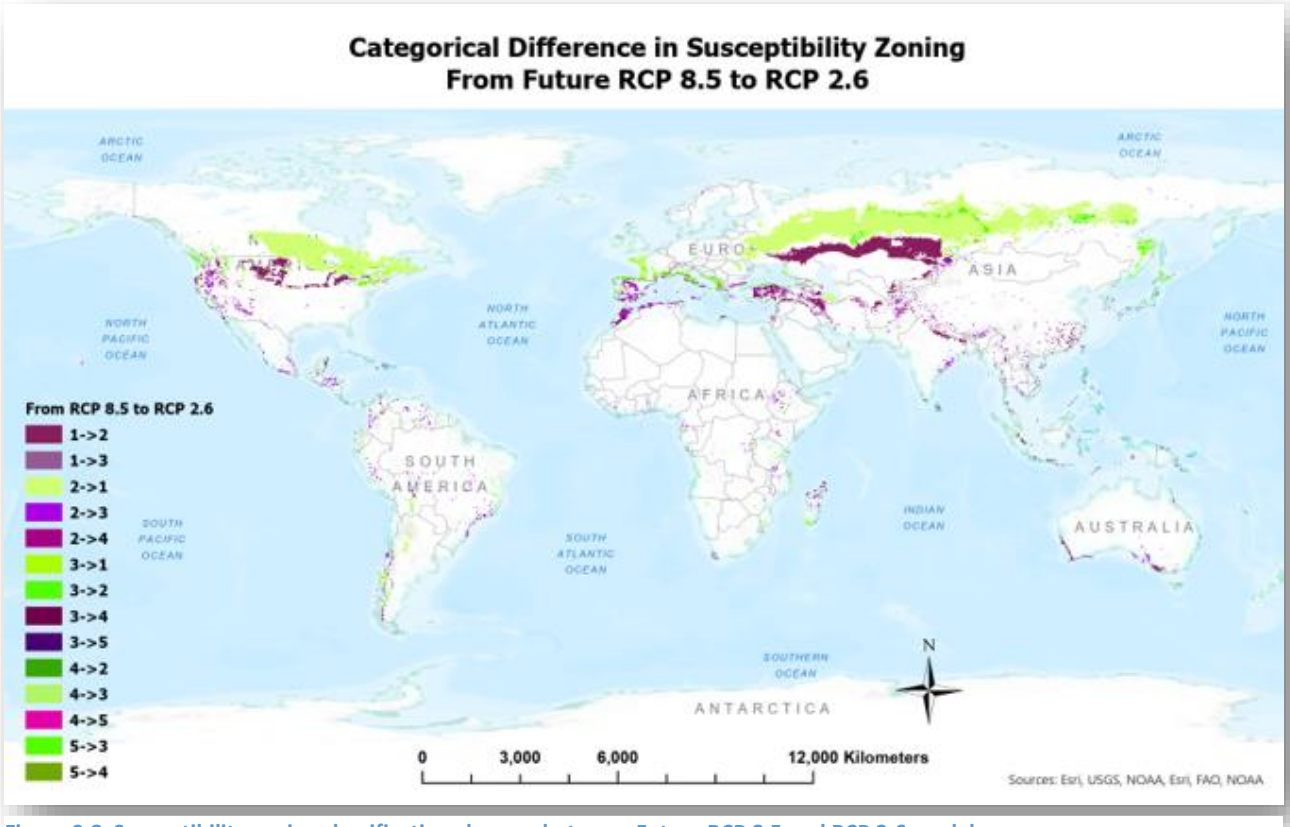


Figure 9.8 Susceptibility zoning classification changes between Future RCP 8.5 and RCP 2.6 models.

Classification mapping is 1 = Very Low, 2 = Low, 3 = Moderate, 4 = High, 5 = Very High.

Table 9.3 presents the changes (increase and decrease) in susceptibility classification in a pairwise comparison between the present-day model and RCP 2.6; present-day model and RCP 8.5; and between RCP 2.6 and RCP 8.5 models, as percent of total land area and percent of changed land area. 1.49% of the total land area (34.17% of total changed land area) exhibit an increase of debris flow susceptibility classification under the future RCP 2.6 scenario, and 3.14% of total land area (51.47% of total changed land area) under RCP 8.5. Conversely, 2.87% of total land area (65.83% of total changed area) exhibits a decrease in susceptibility classification under RCP 2.6 and 2.96% of total land area (48.53% of total changed area) under RCP 8.5.

Susceptibility Change	From Current to RCP 2.6	From Current to RCP 8.5	From RCP 2.6 to RCP 8.5
INCREASE			
% of Total Area	1.49%	3.14%	2.35%
% of Changed Area	34.17%	51.47%	74.47%
DECREASE			
% of Total Area	2.87%	2.96%	0.80%
% of Changed Area	65.83%	48.53%	25.53%
INCREASE TO MODERATE, HIGH, VERY HIGH CLASSIFICATION			
% of Total Area	.47%	.63%	0.29%
% of Changed Area	10.79%	10.29%	9.09%

Table 9.3 Comparison of geographic extent of 'increased' versus "decreased" debris flow susceptibility

Percentage area of any debris flow susceptibility classification to moderate, high, or very high debris flow classification from any other classification.

9.5 Discussion

Although the percent of total land area with an increase in debris flow susceptibility classification under RCP 2.6 and RCP 8.5 scenarios is small (1.49% and 3.14%, respectively), it represents a significant coverage, with 2.2 million km^2 (about the size of Saudi Arabia) under RCP 2.6, and 4.7 million km^2 (about the size of the European Union) under RCP 8.5. However, the global area increase to Moderate, High, Very High debris flow susceptibility under RCP 2.6 and RCP 8.5 scenarios is even smaller with 0.47% (705,000 sq km) and 0.63% (945,000 sq km), respectively. These small projected susceptibility increases belie the potential for more frequent and larger magnitude debris flows (Alvioli et al., 2018; C. Huggel et al., 2010; Li et al., 2018) due to the potential for more numerous heavy rainfall events, in spite of annual precipitation decreases (Donat et al., 2016).

In both future scenarios, the areas of increased susceptibility in the northern hemisphere are coincident with an increase in area extent of Köppen-Geiger (Beck et al., 2018) climate classification Dfa and its projected northward shift in North America and Eurasia. Dfa was found to be one of three most significant climate classifications associated with global debris flow susceptibility in the present-day modeling (Chapter 7). Dfa is a humid continental climate, found between latitudes 30° N and 60° N, characterized by hot summers and cold to severely cold winters. Annual precipitation ranges from 50 to 125 cm, in the form of thunderstorms in summer

and snow in the winter months, higher amounts in the uplands, and no dry season, i.e. precipitation is distributed throughout the year (Beck et al., 2018).

Under RCP 8.5 there is a correlation between areas of increased susceptibility and increased precipitation, that is, where present-day average monthly precipitation is under 100 mm and RCP 8.5 values are projected at 100-200 mm. Significantly impacted areas under both future scenarios, those with Moderate, High, or Very High susceptibility in future models, include the west coasts of North and South America (eastern Cascade Range, Northern Rocky Mountains, and Sierra Nevada of North America, eastern portions of the Andes Mountains in South America); southern Europe/land bounding northern Mediterranean Sea, coastal Turkey, along the Zagros Mountains from Turkey to Iran, southeastern coast of Madagascar; eastern Asia (Himalayan belt, China, Japan, eastern coast of Russia, and mountain ranges of the eastern interior of Asia); and Oceania (New Zealand, and southeastern coast of Australia).

Changes in the geographic extents, of the most significant climate and precipitation factor classes (as determined in the "present-day" model) under RCP 2.6 and RCP 8.5 scenarios, are the main drivers for the geographic changes and level of debris flow susceptibility. Figs. 9.9, 9.10, and 9.11 present the geographic extents of the climate classifications Csa, Csb, Dfa (the three most significant climate factor classes associated with the present-day global debris flow susceptibility model), and average precipitation values of 100 to 200 mm/month for the present-day model, RCP 2.6, and RCP 8.5 models, respectively.

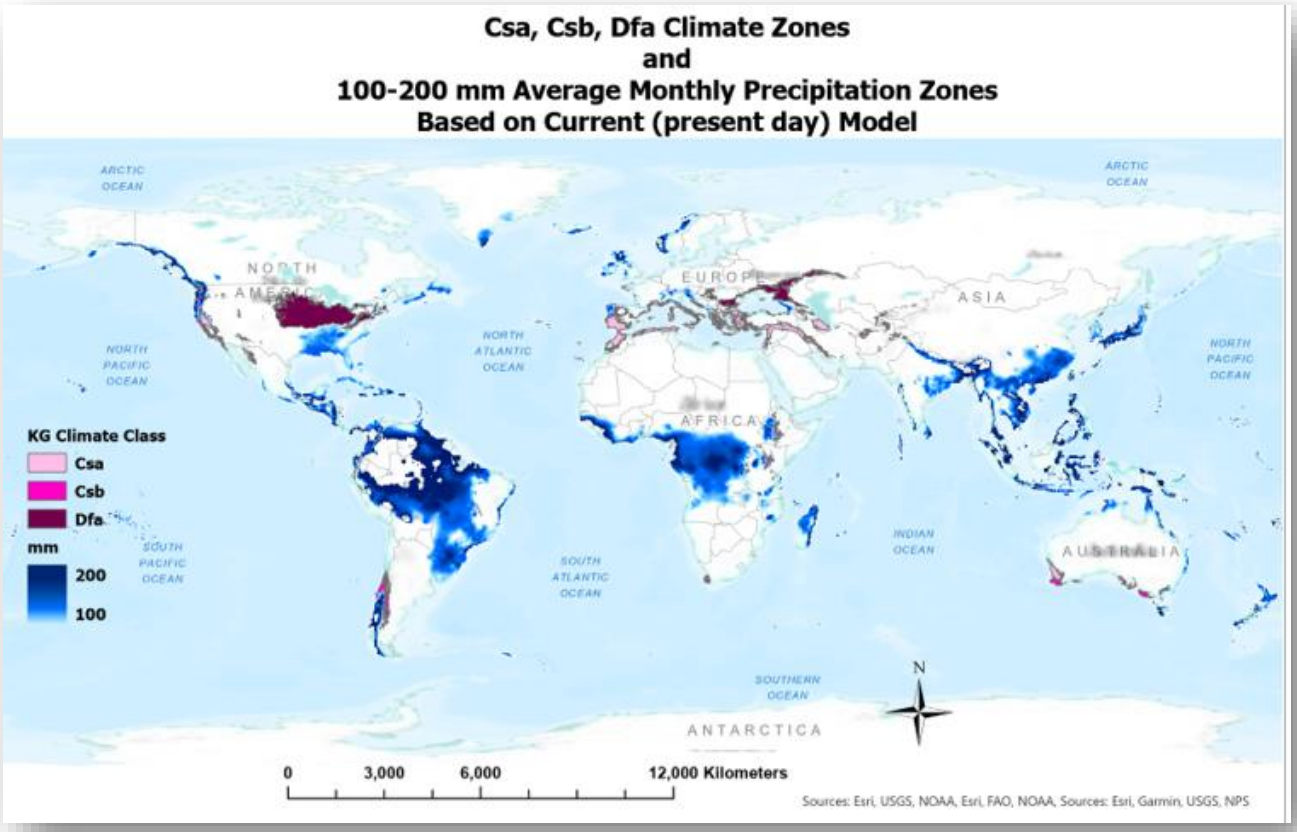


Figure 9.9 Köppen-Geiger climate classification (Csa, Csb, Dfa) and precipitation (100-200 mm) factor class geographic extents in present-day (current) modeling.

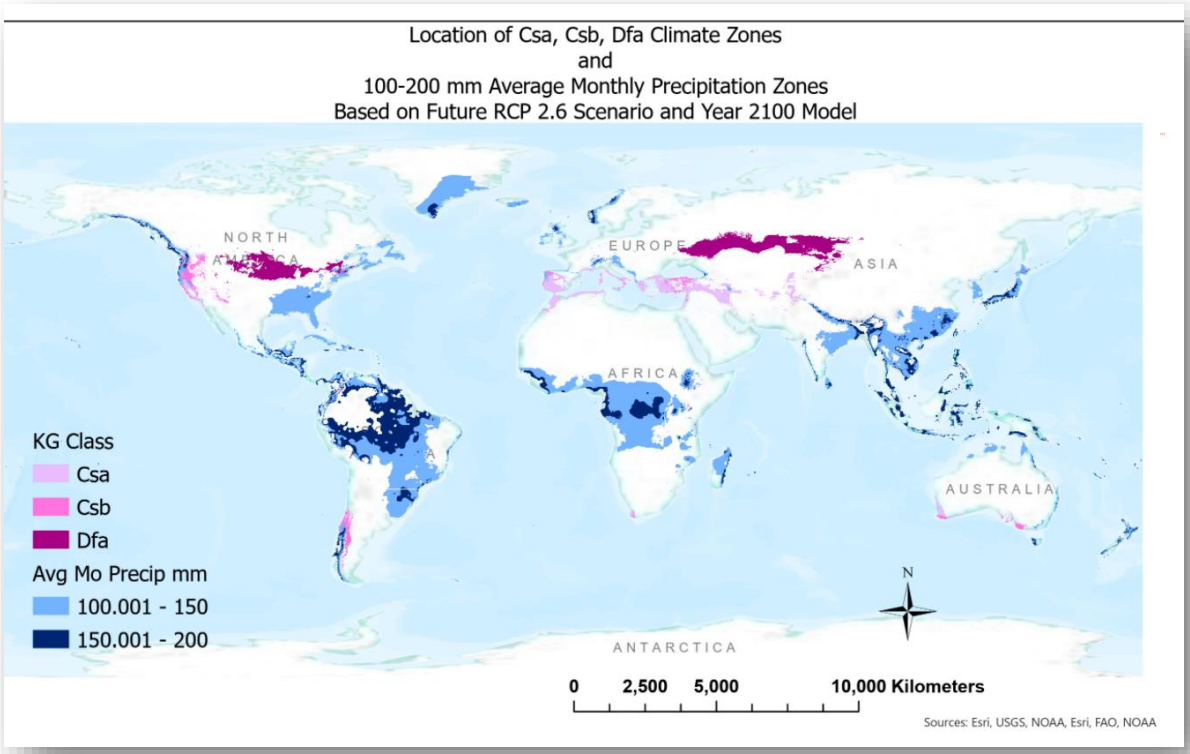


Figure 9.10 Köppen-Geiger climate classification (Csa, Csb, Dfa) and precipitation (100-200 mm) factor class geographic extents in future RCP 2.6 scenario modeling.

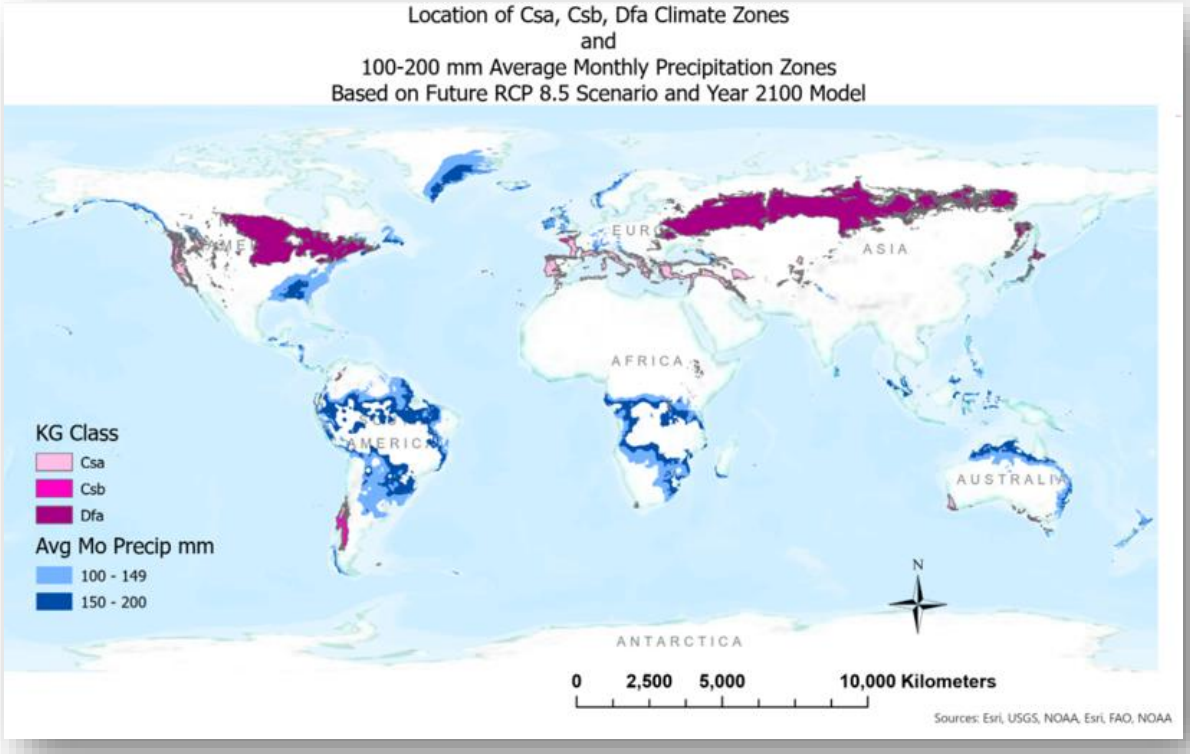


Figure 9.11 Köppen-Geiger climate classification (Csa, Csb, Dfa) and precipitation (100-200 mm) factor class geographic extents in future RCP 8.5 scenario modeling.

While there are numerous areas across the Earth exhibiting a decrease in debris flow susceptibility under scenarios RCP 2.6 and RCP 8.5, a few sizeable areas appear east of the Rocky Mountains to the eastern seaboard of the United States, approximately along 40^{o} N latitude, as well as in eastern Ukraine, western Russia, southern Romania, and northern Bulgaria. The decreased susceptibility in these areas appears to be primarily due to decreasing precipitation in both future scenarios.

9.6 Conclusions

It is acknowledged that there is not a clear, well-defined, linear relationship between climate and debris flows (Froude & Petley, 2018), yet substantial research indicates a causal relationship between climate change and debris flows (Gariano & Guzzetti, 2016), slope evolution and instability directly or indirectly (Borgatti & Soldati, 2010).

Temperature and precipitation climate changes may affect the debris flow seasonality, frequency, magnitude, and locations of susceptibility and may do so over short time periods. Additionally, climate changes will affect other debris flow predisposing factors such as topography or landforms, soil type and thickness, lithology, fracture density, etc. over a longer time span. This unknown interactive impact is a critical component of a more accurate understanding of how and where debris flow susceptibility will change in the future.

Based on climate and precipitation forecasts for RCP 2.6 (best case) and RCP 8.5 (worst case) scenarios, there may be both increases and decreases in debris flow susceptibility in various regions under both scenarios, with an overall 1.49% and 3.14% increase (of total land area) under RCP 2.6 and RCP 8.5, respectively; and a 2.87% and 2.96% decrease in RCP 2.6 and RCP 8.5 scenarios, respectively. Perhaps more importantly is the .47% and .63% change from a lower susceptibility to Moderate, High, or Very High under scenarios RCP 2.6 and RCP 8.5, respectively, as compared to the present-day model. For those areas which increase to the Very High susceptibility class there is an increase of .01% (15,000 km^2) and .02% (30,000 km^2) of total land for RCP 2.6 and RCP 8.5, respectively.

These models were developed with a conservative approach by minimizing the factor classes utilized to those with the highest probability (> 90%) of association, which results in fewer factor

classes (particularly fewer climate and precipitation classes), which may therefore result in smaller areas of Moderate to Very High susceptibility, but which may result in a higher confidence of the results.

10 Human Vulnerability and Debris Flow Susceptibility

10.1 Introduction

There are two ways to view future debris flow susceptibility impacts on humans, one is by looking at future areas of population (population per square kilometer) and the second is to look at projected areas of urbanization (population shift from rural to urban). The latter may be considered a proxy for impact on both humans and economies. Both views are considered in this study. The impact of landslides, be it economic or loss of life, is not uniformly experienced across the world. Proximity of landslides to cities occurs most frequently in countries with lower gross national income, where urbanization frequently extends over unstable slopes and ancient landslides (Cavallo & Noy, 2010; Froude & Petley, 2018; United Nations Office of Disaster Risk Reduction, 2019). The median number of fatalities per event may be as much as 400% greater in developing countries than in advanced countries Dowling and Santi refer to debris flows as a "disaster of social vulnerability" (Dowling & Santi, 2013). Debris flow events are expected to increase in frequency and/or intensity in association with climate change, and human vulnerability will also increase due to increases in population and urbanization, deforestation, and increasing mountain settlements (Froude & Petley, 2018; Haque et al., 2019; United Nations Office of Disaster Risk Reduction, 2019). Although debris flows cannot be prevented their impact may be mitigated through an awareness of the intersection between susceptibility and populations, worldwide, with a recognition of those areas where human loss could be higher. The objective of this work is not to define a quantitative vulnerability index as was done by Papathoma-Kohle et al (Papathoma-Kohle, Schlogl, & Fuchs, 2019), but rather to represent the potential vulnerabilities spatially.

By 2100, a decrease in population in Europe is expected, with increases in the remainder of the world (Table 10.1). Particularly noteworthy is the population of sub-Saharan Africa which is likely to increase by 400%, with West Asia and North Africa doubling (Cherlet et al., 2018).

	Actual (billions)		Projections (billions)		
By Geography	**2015**	**% of Total**	**2030**	**2050**	**2100**
World	7.349		8.501	9.725	11.213
Africa	1.186	16%	1.679	2.478	4.387
Asia	4.393	60%	4.923	5.267	4.889
China	*1.400*	*19%*			
India	*1.300*	*18%*			
Europe	0.738	10%	0.734	0.707	0.646
Latin America	0.634	9%	0.721	0.784	0.721
North America	0.358	5%	0.396	0.433	0.500
Oceania	0.039	1%	0.047	0.057	0.071
By Geomorphology					
High Mountain Regions	0.670	10%		.840	
Low-lying coastal	0.680	10%		1.0	

Table 10.1 World/continental population in 2015 and projections by geographic distribution (Cherlet et al., 2018) and by geomorphic environment (IPCC, 2019b)

Future urbanization projections, used herein, are based on SSP5 (Share Socioeconomic Pathway 5) data at 1-km resolution, developed by Chen et al (G. Chen et al., 2020). SSPs describe alternative trends based on varying criteria. The SSPs consist of five different qualitative narratives describing broad patterns of possible future development at the global level. SSP5 is characterized by fast urbanization and increased migration to cities. The scenario assumes urbanization will reach 92% (of the population) by the end of the 21st century. Jiang et al (Jiang & O'Neill, 2017) predict that the world will continue to urbanize by the end of century under each of the five SSPs with differing outcomes (60%, 79%, and 92%, in SSP3, SSP2, and SSP1/SSP4/SSP5, respectively). It is important to note that Chen et al observe urban land demand decreases in all scenario projections due to post-2030 declining populations, for example in China and Europe. However, they find that urban population declines do not necessarily lead to massive land conversions back to non-urban, but rather abandonment of already built-up areas.

Current and projected population and urbanization data are developed and combined with current ("LTE 5 km") and future debris flow susceptibility models described in Sections 7 and 9, respectively.

10.2 Results and Discussion

Figures 10.1 and 10.2 present the spatial association of areas with >= 74,000 population vis-à-vis future debris flow susceptibility scenarios RCP 2.6 and 8.5, respectively.

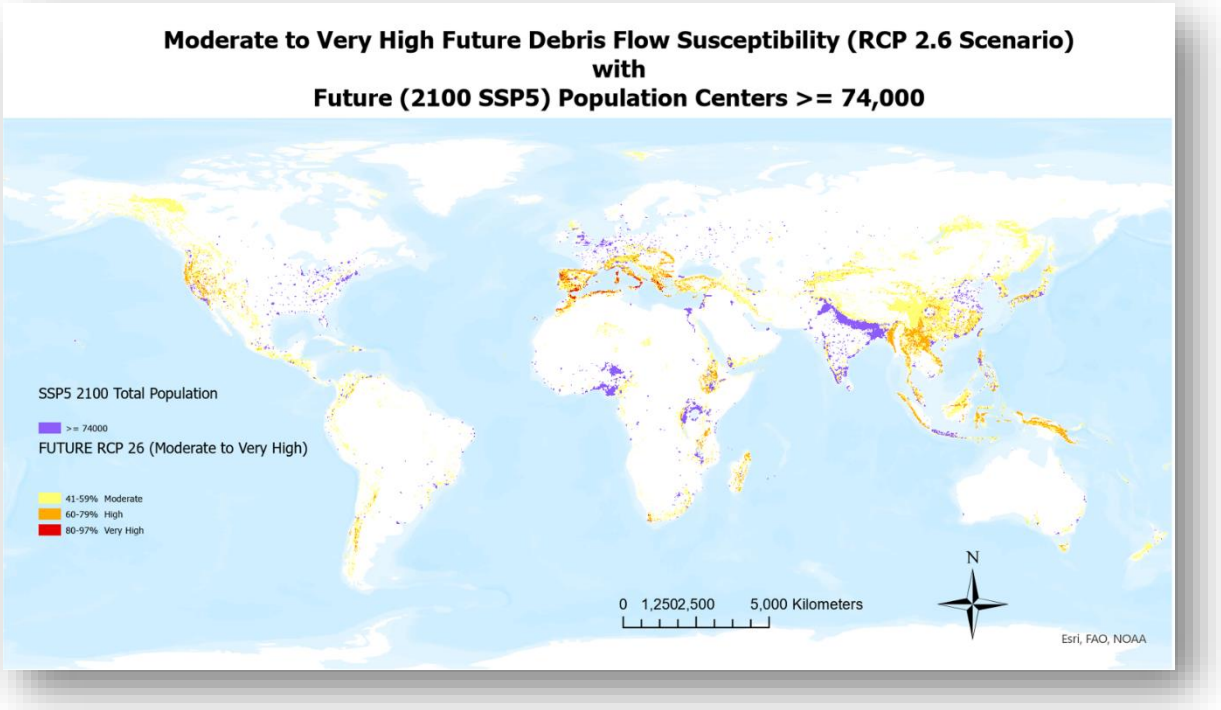


Figure 10.1 Spatial association of centers with >= 74,000 population vis-à-vis future debris flow susceptibility scenario RCP 2.6

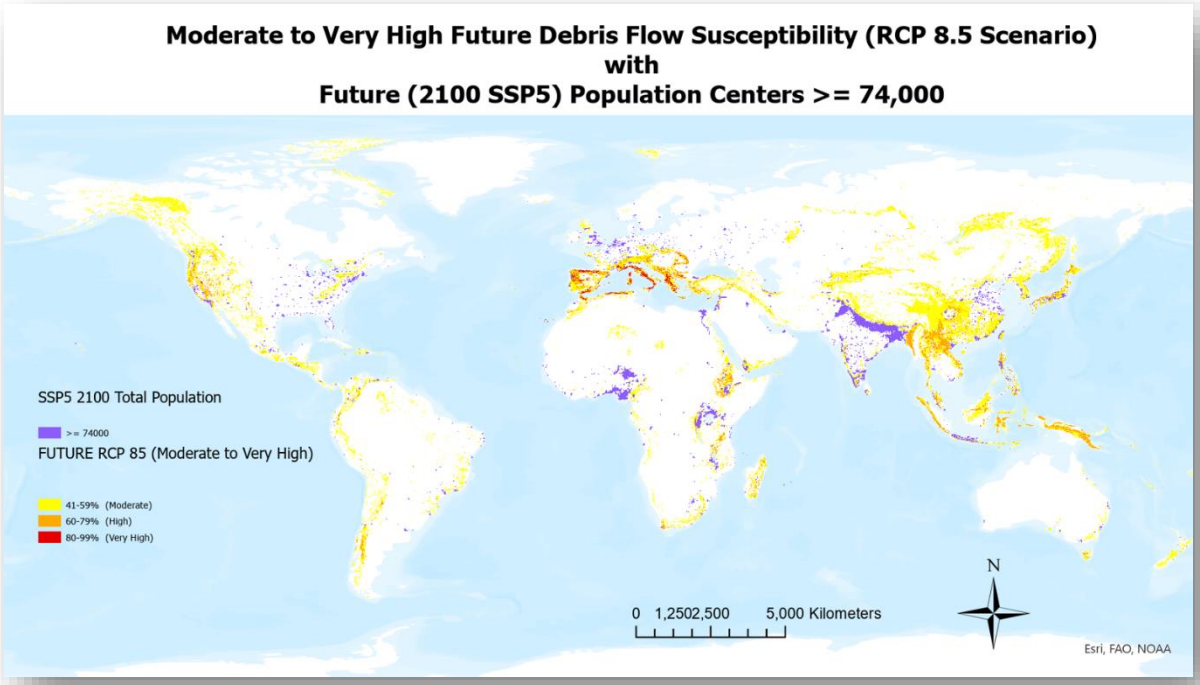


Figure 10.2 Figure 10.2 Spatial association of centers with >= 74,000 population vis-à-vis future debris flow susceptibility scenario 8.5

Under scenario RCP 2.6, 12.8% of the 2100 SSP5 "urban" population will reside in areas classified as Moderate, High, or Very High debris flow susceptibility. Surprisingly, this number decreases to 11.7% under scenario RCP 8.5 (Table 10.2).

Susceptibility Class	% Population RCP 2.6	% Population RCP 8.5
Very Low	61.7%	63.1%
Low	25.5%	25.2%
Moderate	9.3%	8.5%
High	3.2%	2.9%
Very High	0.3%	0.3%
Total Moderate, High, Very High	**12.8%**	**11.7%**

Table 10.2 Future urban population within RCP 2.6 and 8.5 debris flow susceptibility classifications.

Figures 10.3 (a) and (b), 10.4 (a) and (b), 10.5 (a) and (b), 10.6 (a) and (b), 10.7 (a) and (b), and 10.8 (a) and (b) present a continent-by-continent view of the coincidence of areas with a population greater than or equal to 74,000 and Moderate, High, and Very High debris flow

susceptibility for Africa, Asia, Europe, North America, South America, and Australia/Oceania, respectively.

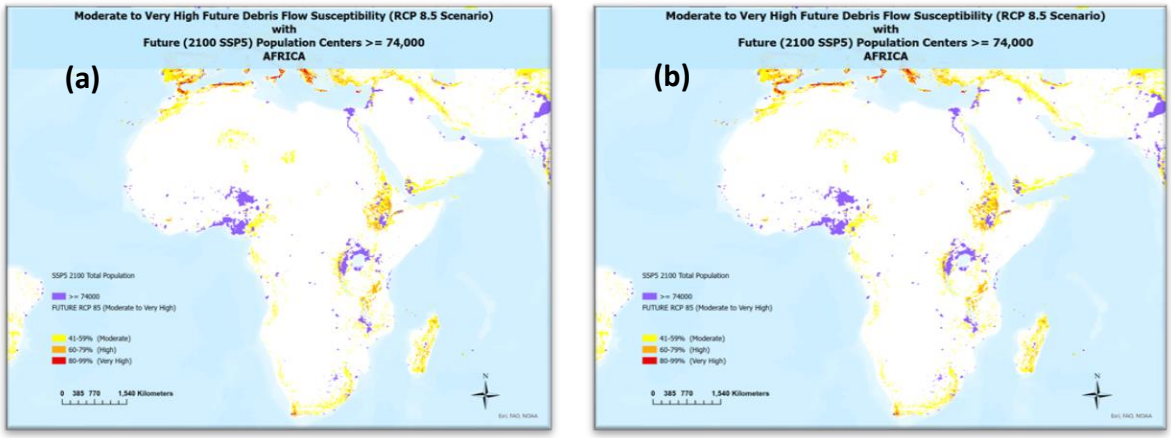


Figure 10.3 Africa SSP5 2100 Population >= 74,000 with Moderate to Very High susceptibility for (a) RCP 2.6 and (b) RCP 8.5

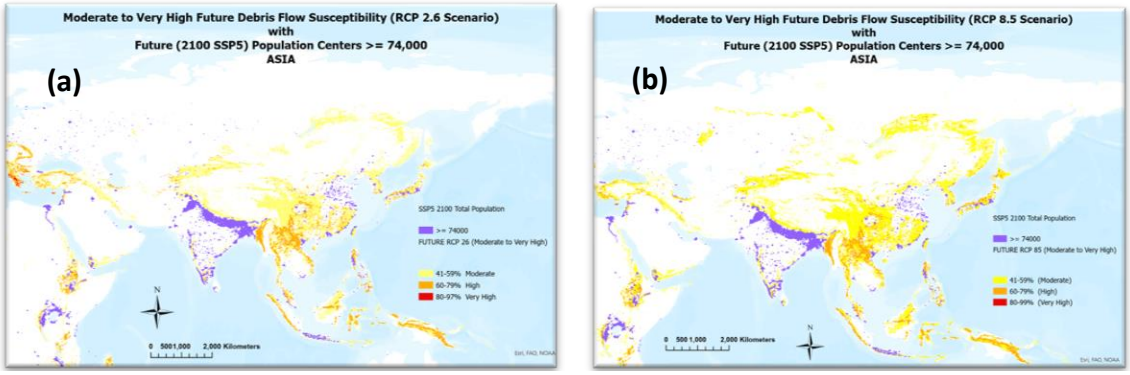


Figure 10.4 Asia SSP5 2100 Population >= 74,000 with Moderate to Very High susceptibility for (a) RCP 2.6 and (b) RCP 8.5

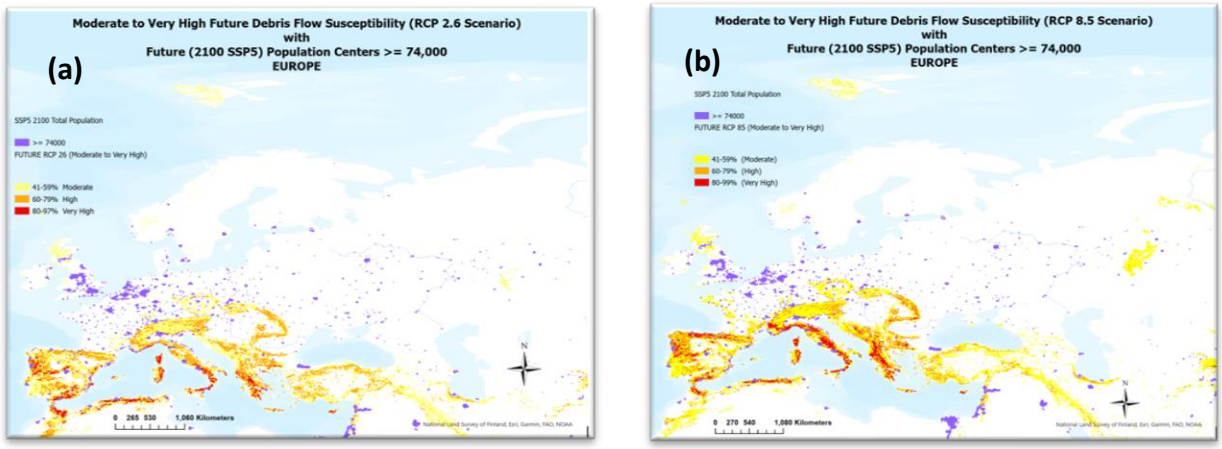


Figure 10.5 Europe SSP5 2100 Population >= 74,000 with Moderate to Very High susceptibility for (a) RCP 2.6 and (b) RCP 8.5

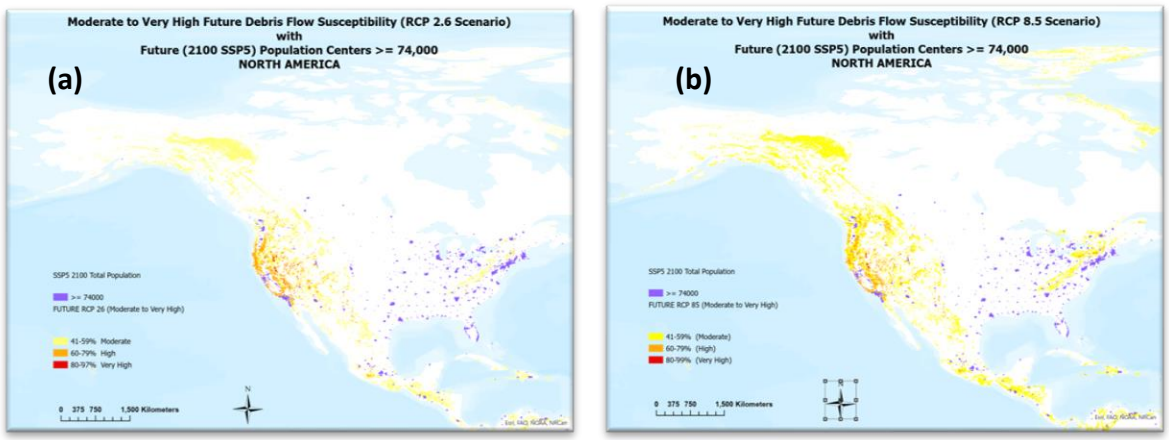


Figure 10.6 North America SSP5 2100 Population >= 74,000 with Moderate to Very High susceptibility for (a) RCP 2.6 and (b) RCP 8.5

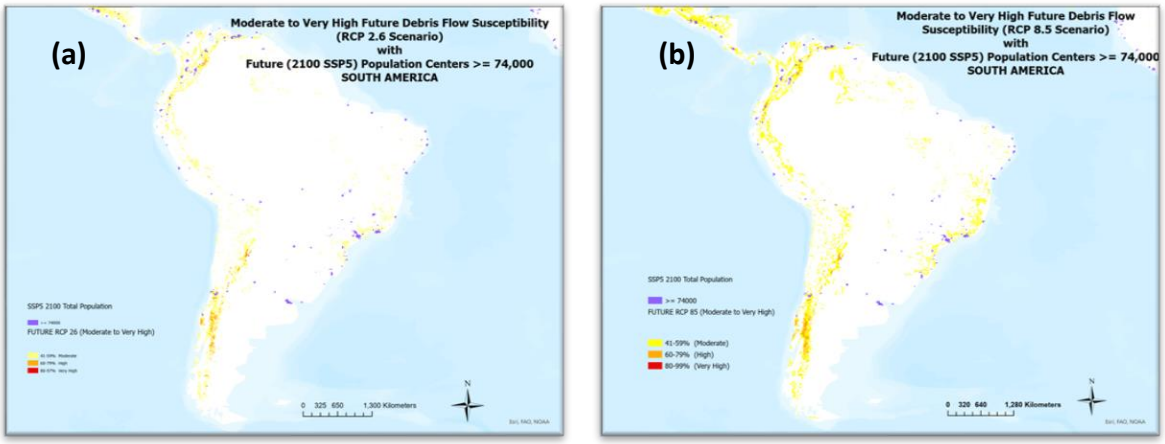


Figure 10.7 South America SSP5 2100 Population >= 74,000 with Moderate to Very High susceptibility for (a) RCP 2.6 and (b) RCP 8.5

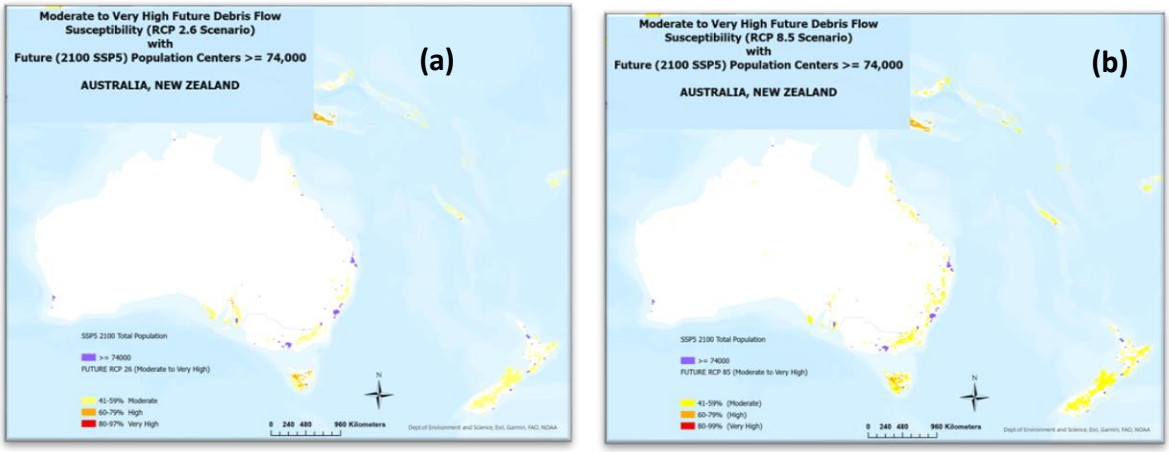


Figure 10.8 Australia, Oceania SSP5 2100 Population >= 74,000 with Moderate to Very High susceptibility for (a) RCP 2.6 and (b) RCP 8.5

Increased population by year 2100 and proximity to debris flow susceptibility may not be the most significant problem for human vulnerability, as evidenced in Figures 10.9 (a) and (b). These maps for a portion of central Europe show a far greater overlap of debris flow susceptibility with urbanization. Jiang et al (Jiang & O'Neill, 2017) project SSP5 to result in the greatest urbanization rate (92%). And although Chen et al (G. Chen et al., 2020) note that urban land demand decreases in all scenario projections due to post-2030 declining populations, they find that urban population declines do not necessarily lead to massive land conversions back to non-urban, but rather abandonment of already developed lands. Such land changes may continue to contribute to slope instabilities.

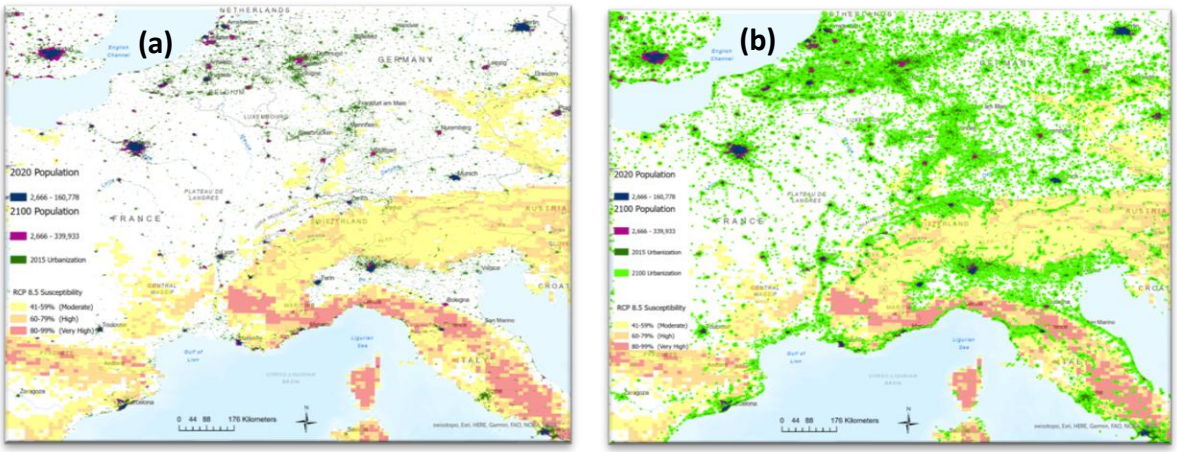


Figure 10.9 Moderate to Very High susceptibility with (a) current and SSP5 2100 population, current urbanization and (b)) current and SSP5 2100 population, current and SSP5 2100 urbanization

10.3 Conclusions

The most notable result of this study, is that population proximity to debris flow susceptibility, alone, may not be the most significant factor in understanding human vulnerability, but rather the proximity of urbanization, where there is a far greater overlap with debris flow susceptibility.

No one SSP is more likely than another to come to fruition, only time will tell based on current and continued practices, changes, and mitigation efforts. However, it is notable that the urbanization rate projected for SSP5 (92%) is the same for three of the five SSPs. The methodologies employed herein are easily modified by replacing SSP5 with any other scenario.

11 Ancillary Factor and Trigger Considerations

11.1 Legacy Factor

Landslide susceptibility is not time-independent (Luigi Lombardo & Tanyas, 2021; Samia et al., 2016), and the past is key to the future (Uniformitarian Principle), therefore "legacy" is introduced as an additional factor to the debris flow susceptibility models. Although there will certainly be increases and decreases in susceptibility and debris flow densities over time as a result of climate change and anthropogenic disturbances, it is nonetheless a useful consideration, particularly in analyses over shorter elapsed time periods, dependent on the location and environmental factors (Samia et al., 2016). While Samia et al use legacy information "without attention for the role of intrinsic attributes (e.g., slope and geology)" in susceptibility analyses, herein legacy data is used to support the role of the intrinsic attributes.

A debris flow density layer, based on the TRAIN global debris flow inventory, was developed as an overlay (Fig 11.1) to the present-day debris flow susceptibility map. The five-classification debris flow density layer is re-classified to integer values for each break with new values from one to five, using ArcGIS Pro. Similarly, the five-classification global susceptibility map is also reclassified to integer values (one to five). The integer values in both layers are multiplied on a pixel-by-pixel basis creating and representing a susceptibility layer augmented by the legacy factor with newly defined susceptibility values from one through twenty-five (Fig. 11.2).

Locally, the true value of a legacy factor may be diminished over time due to the fact that terrain conditions (slope angle, landuse) may change after the occurrence of a landslide (van Westen et al., 2005).

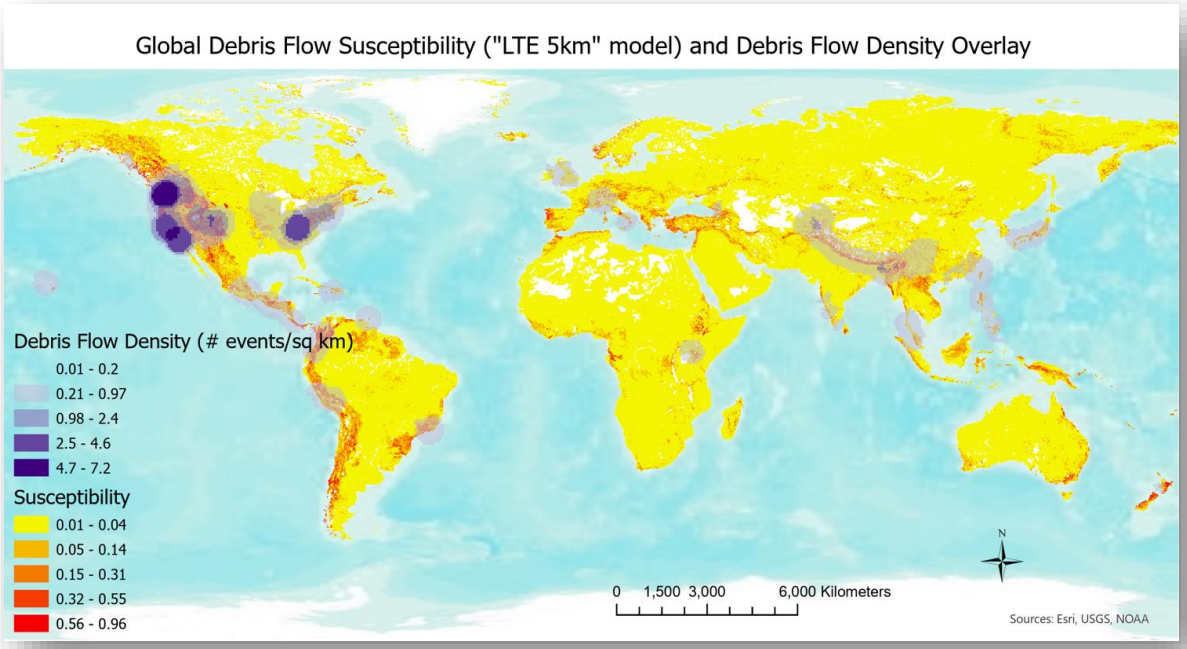


Figure 11.2 Current global debris flow susceptibility using "LTE 5 km" model, with overlay of debris flow density

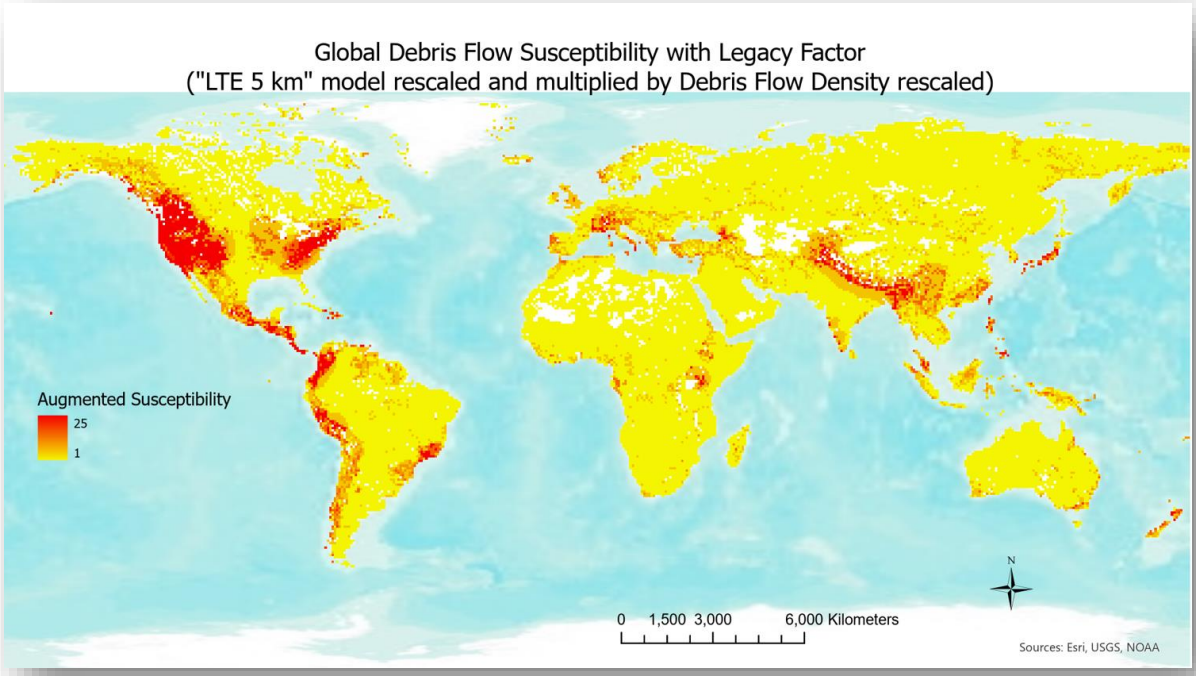


Figure 11.1 Global MaxEnt Debris flow susceptibility with legacy factor.

Classification rescaled values (one through five) multiplied by debris flow density layer rescaled pixel values (one through five), resulting in pixel values one through twenty-five.

11.2 Fire Probability as an Amplifier to Debris Flow Predisposing Factors

Climate and climate change affect the seasonality, return-rate, volume, and location of large, intense wildfires, which subsequently affect the susceptibility of an area to debris flows (S. H. Cannon & DeGraff, 2009).

Wildfires contribute to debris flow susceptibility in numerous ways. During and post-wildfire, stream channels can fill by gravity-driven transport of unconsolidated sediments, post-fire overland flow may increase due to decreased soil infiltration from the formation of a post-wildfire hydrophobic layer (Gabet & Bookter, 2008), removal of vegetation canopy which allows more precipitation to reach the surface , and destruction of root strength, which would otherwise bind the soil (S. Cannon & Reneau, 2000; Neary, Klopatek, DeBano, & Ffolliott, 1999). Vegetative ash can become a significant portion of the fine sediment eroded from hillslopes and contributing to debris flows (S. H. Cannon, Bigio, & Mine, 2001), and decreased roughness of the soil surface (H. Lavee, Kutiel, Segev, & Benyamini, 1995). Such mechanisms may increase susceptibility as well as provide materials and environments for bulked debris flows (Meyer & Wells).

Cannon et al (S. H. Cannon & DeGraff, 2009) found a strong relationship between climatic changes and wildfire frequencies and volume, and post-wildfire debris flow potential in the western United States, with many other parts of the world subject to a similar combination of hazards. They point out the need to identify areas susceptible to post-wildfire debris flows, with potential for triggering rainfall, in order to minimize human risks. Increases in wildfire occurrence are likely attributable to both climate change and urbanization. This highlights the importance of determining the intersection of post-wildfire augmented debris flow susceptibility vis-à-vis population and urbanization.

It is expected that climate change will result in increased desertification in some areas, and increased vegetation, due to increased precipitation, resulting in more burnable materials in other areas. Both are projected to lead to increased risk and severity of wildfires, along with extended fire seasons (IPCC, 2019; Moritz et al., 2012).

Building on the global debris flow susceptibility models (Chapters 7, 8, 9) demonstrating the future geographic extent and impact on growing populations, this study adds wildfire probability as an augmentation. Current and future projected wildfire probability data (Moritz et al., 2012) are used in conjunction with both current and future debris flow susceptibility models to identify areas where wildfire probability may increase, decrease, or have no impact on debris flow susceptibility. Figure 11.3 provides a view of areas which may exhibit wildfire probability changes by differencing the future projected and current wildfire probabilities. The difference model shows a 16.6% increase in wildfire probability, 49% decrease, and 34% of the Earth exhibits no change. Increases are predominantly in mid- to high latitudes of the northern hemisphere, and along the Andes Mountains, southwestern Africa, and central Australia, in the southern hemisphere.

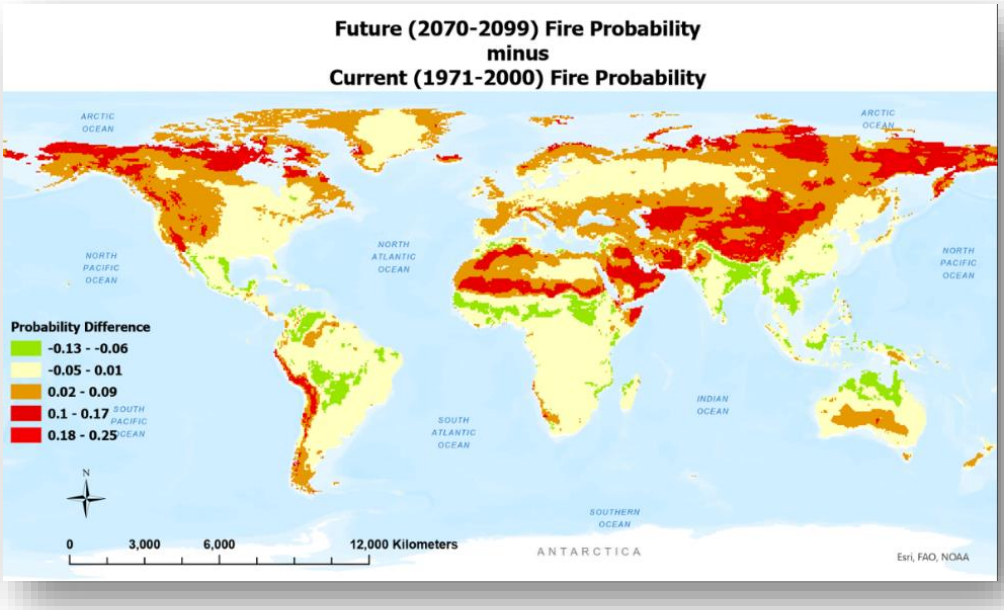


Figure 11.3 Change in current wildfire probabilities and future (2070-2099) projected.

The current ("LTE 5 km" model) debris flow susceptibility (Fig. 11.4 (a)) was summed with current wildfire probability (Fig. 11.4 (b)) on a pixel basis to provide an enhanced debris flow susceptibility (Fig. 11.4 (c)). The IPCC (IPCC, 2019) projected North America, South America, Mediterranean, southern Africa and central Asia as areas that will be increasingly affected by wildfire. These areas exhibit increased debris flow susceptibility when augment with wildfire probability.

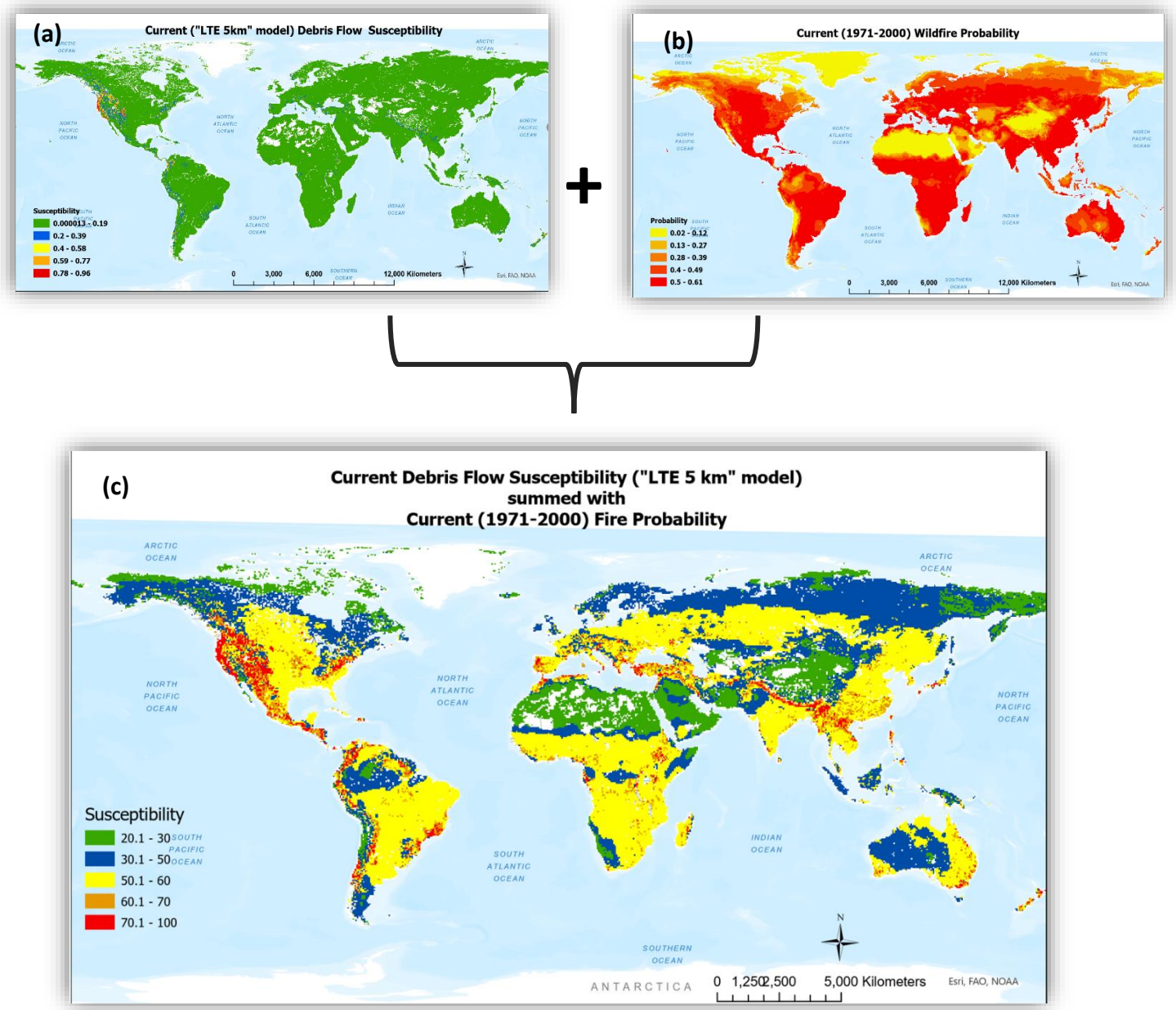


Figure 11.4 (a) Current debris flow susceptibility, (b) current wildfire probability, (c) resulting debris susceptibility when wildfire probability is factored.

Thirty seven percent of the 2294 historical debris flow events (TRAIN) are located in areas with > 47% wildfire probability (Figure 11.5)

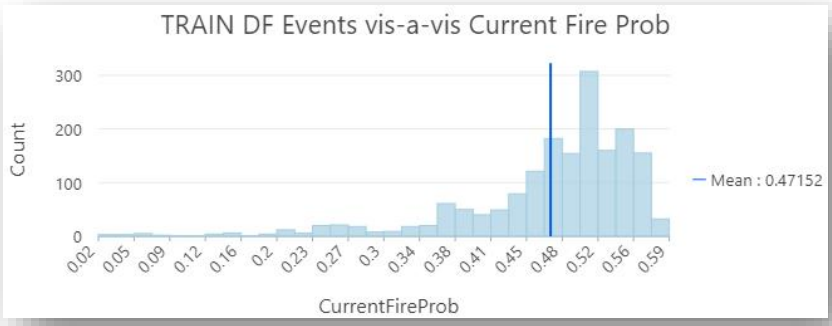
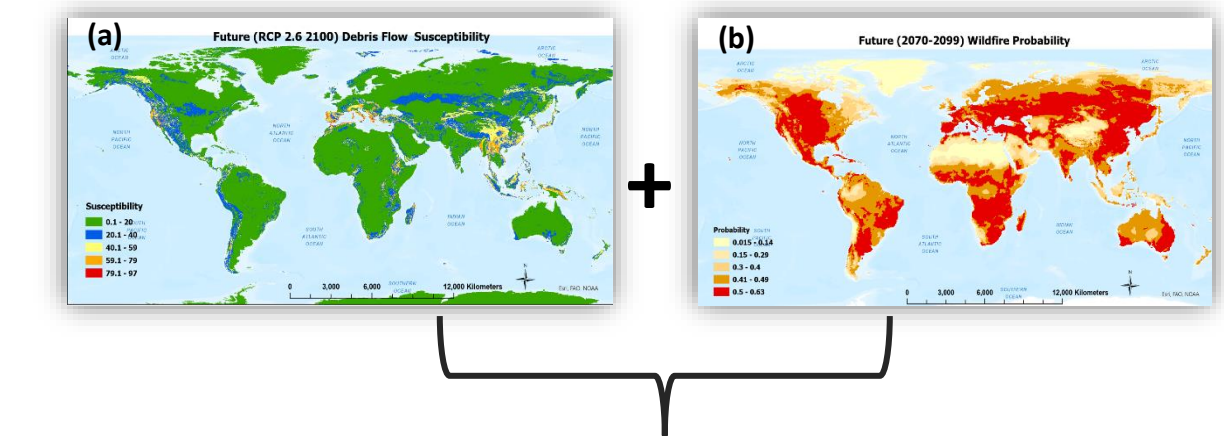


Figure 11.5 Distribution of TRAIN debris flow events and associated wildfire probability.

The future debris flow models (RCP 2.6 and RCP 8.5) were also summed with future (2070-2099) wildfire probability, creating wildfire-augmented future debris susceptibility models (Figures 11.6 and 11.7, respectively).



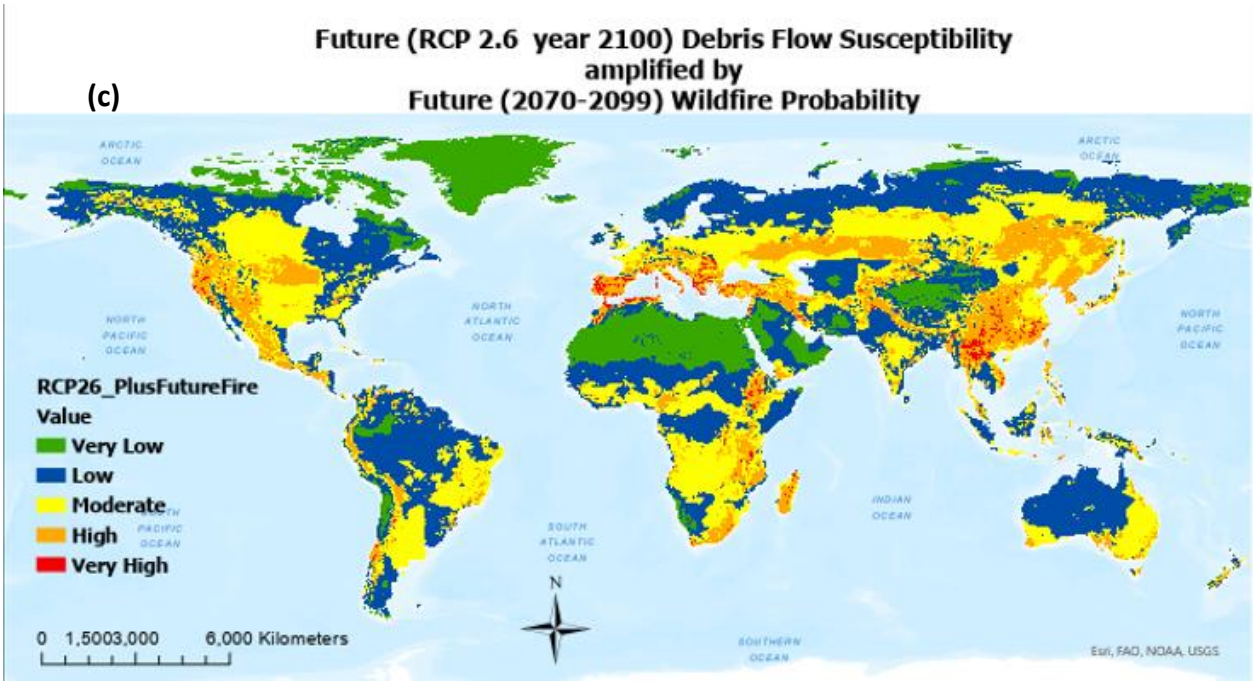


Figure 11.6 (a) Future (2100, RCP 2.6) debris flow susceptibility, (b) future (2070-2099) wildfire probability, (c) resulting debris susceptibility when wildfire probability is factored.

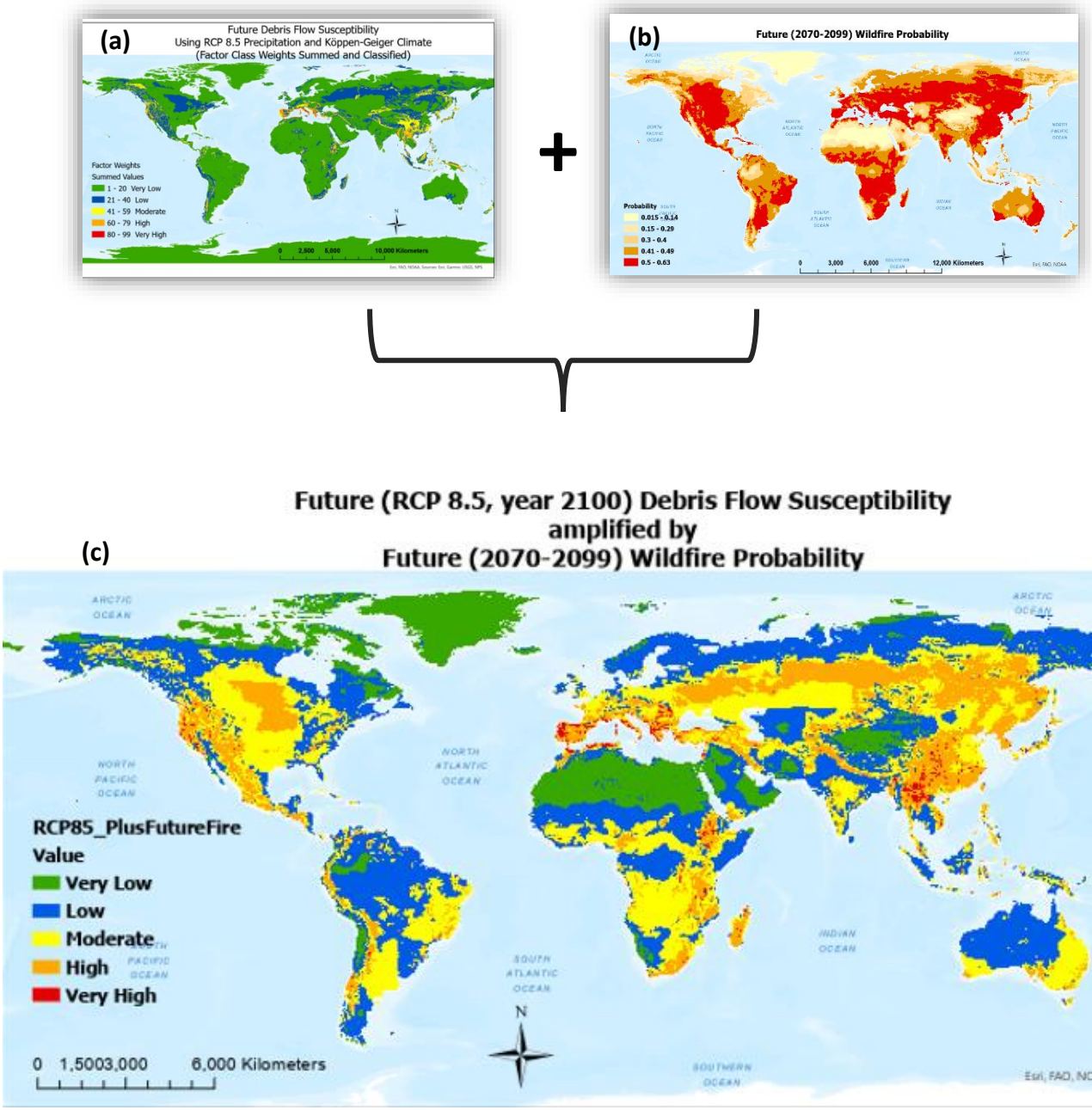


Figure 11.7 (a) Future (2100, RCP 8.5) debris flow susceptibility, (b) future (2070-2099) wildfire probability, (c) resulting debris susceptibility when wildfire probability is factored.

Table 11.1 is a summation of the debris flow susceptibility change after adding wildfire probability. There is a significant increase in areas classified as Moderate to Very High in all models.

	% Change Due to Fire		
Susceptibility Class	**Current**	**RCP 2.6**	**RCP 8.5**
Very Low	-62.1%	-75.8%	-74.2%
Low	27.0%	33.3%	30.5%
Moderate	23.6%	26.9%	25.5%
High	10.3%	10.8%	17.3%
Very High	1.2%	4.8%	0.9%
Sum Moderate to Very High	**35.1%**	**42.5%**	**43.8%**

Table 11.1 Percent change in land associated with each susceptibility classification, after adding wildfire probability, for the current, RCP 2.6, and RCP 8.5 debris flow susceptibility models.

11.3 Future Glacial Melting as a Potential Amplifier to Debris Flow Triggers

It is believed that climate change has driven the increase in large slope failures in the European Alps and other high mountainous regions such as Alaska, and New Zealand, in the past three or four decades, and therefore may pose a significant increase in these hazards in the future in areas with predisposing factors and future warming trends (C. Huggel et al., 2010).

"Debris-flow volumes in high-mountain areas will depend chiefly on the stability and/or movement rates of permafrost bodies, and destabilized rock glaciers could lead to debris flows without historic precedents in the future" (Stoffel et al., 2014). Glacier retreats leading to debuttressing of hillslopes, increased amounts of unconsolidated, unvegetated, erodible, mobilizable material, and glacial lakes with unstable water reservoirs in high mountain areas have been demonstrated in North America and Europe and increasing frequency of debris flows at the margins of glaciers during twentieth century glacier retreat. (M. Chiarle, Geertsema, Mortara, & Clague, 2011; Marta Chiarle et al., 2007; Christian Huggel et al., 2012; Moreiras et al., 2021).

Glacier contribution to precipitation in triggering a debris flow may be associated with the presence of buried glacier ice in the debris, infiltrated water which induces high water pressures linking water pockets in the glaciers, and a sudden emptying of ice-marginal lakes, the latter which may occur in the absence of precipitation, all of which may also occur with high air temperatures, contributing to melting of the glaciers (Marta Chiarle et al., 2007). Glacier-related

debris flows occur in high elevation mountainous regions which may intersect with villages and/or recreational areas (e.g., skiing, camping, hiking).

Because glacial melt, as well as precipitation, can act as a trigger to debris flows, glaciers in slopes of 20 to 35 degrees were factored in as a potential modifier to future debris flows. Distances were calculated from these glaciers, in high debris flow susceptibility environments, to future urban and population centers.

Although glaciers at any degree of slope may provide increased meltwaters to surrounding areas as a result of global warming, a conservative approach is followed by including only those glaciers on slopes which were found to be most commonly associated with debris flows. Of the latest version (RGI6.0) of the Randolph Glacier Inventory (RGI Consortium, 2017) with 214,429 glaciers worldwide, a subset of glaciers with slope >=20 and <=35 degrees was created with a 5 km buffer. Temporal coverage of the inventory is 2000 to 2010, horizontal resolution is approximately 15-30 m. The continental distribution of the subset with 126,372 glaciers is shown in Table 11.2.

The use of these data does not imply future glacier melt, only that should melt occur in these areas they may result in or contribute to debris flows.

The majority of the glacier subset are in Asia (51.3%) and North America (32.9%). Thirteen of the North American glaciers are within 10 km of population centers, and none are within 5 km.

CONTINENT	# Glaciers	% Events	Glacial Area Total (sq km)	% Area
Africa	19	0.02%	3	0.003%
Antarctica	535	0.4%	804	0.8%
Asia	**64,861**	**51.3%**	**52,339**	**53.1%**
Europe	4,676	3.7%	2,424	2.5%
North America	**41,588**	**32.9%**	**32,492**	**33.0%**
Oceania	2,291	1.8%	709	0.7%
South America	12,402	9.8%	9,768	9.9%
TOTAL	**126,372**		98,538	

Table 11.2 126,372 glaciers with slope 20-35 degrees, by continent

The juxtaposition of glaciers with current, future RCP 2.6 and RCP 8.5 are shown in Figs. 11.8, 11.9, and 11.10, respectively.

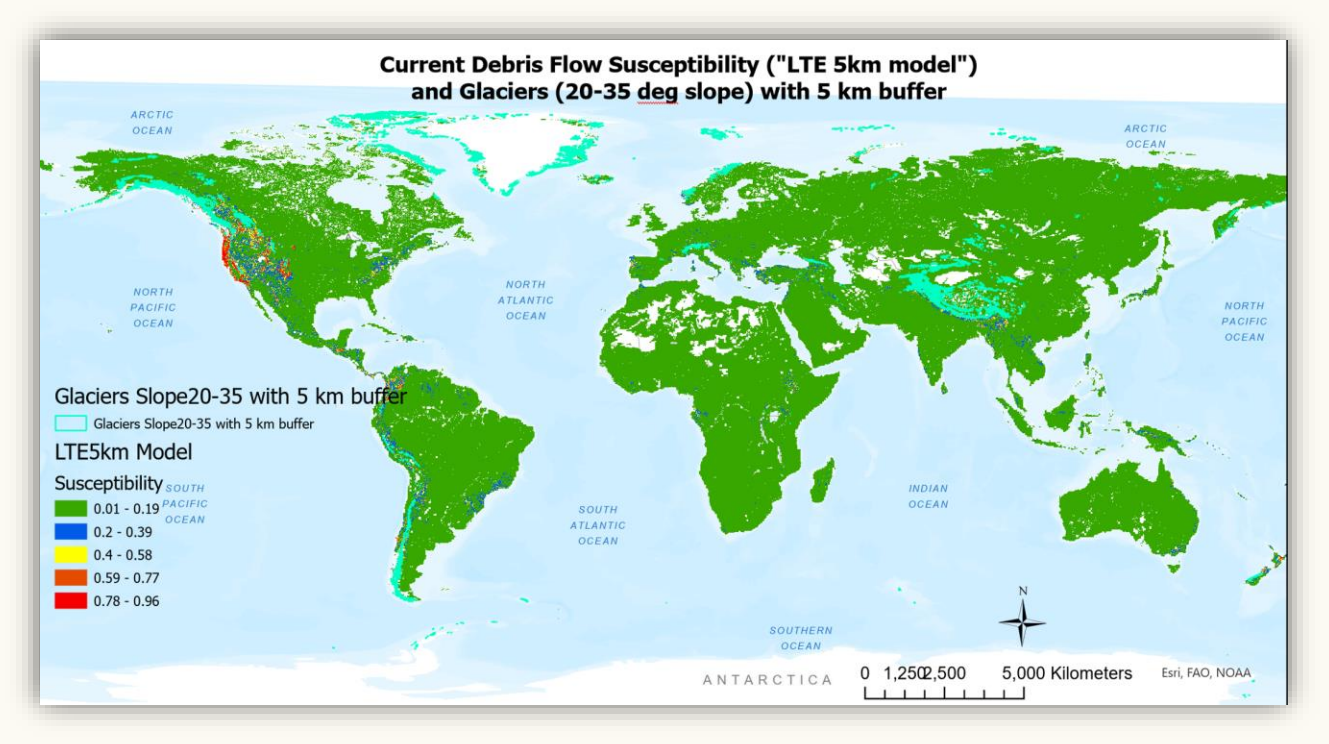


Figure 11.8 Current ("LTE 5km") debris flow susceptibility and buffered glaciers (with 20-35 degree slope)

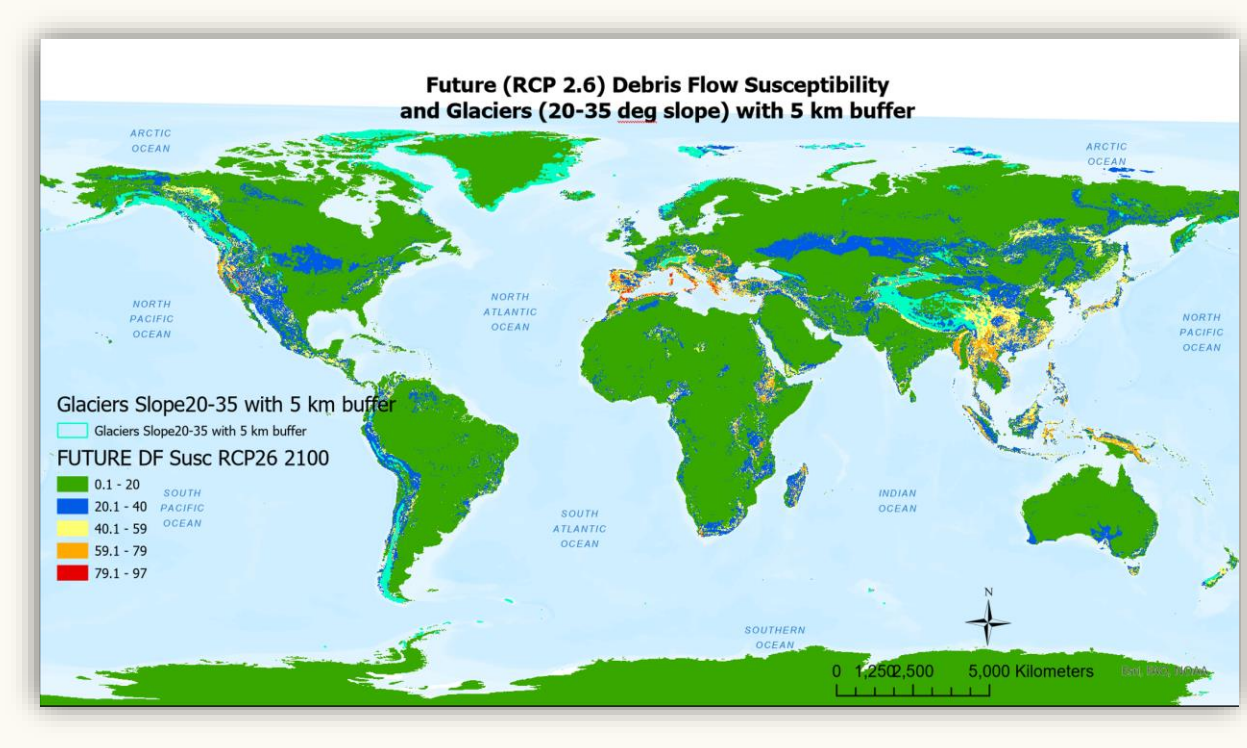


Figure 11.10 Future (RCP2.6 2100) debris flow susceptibility and buffered glaciers (with 20-35 degree slope)

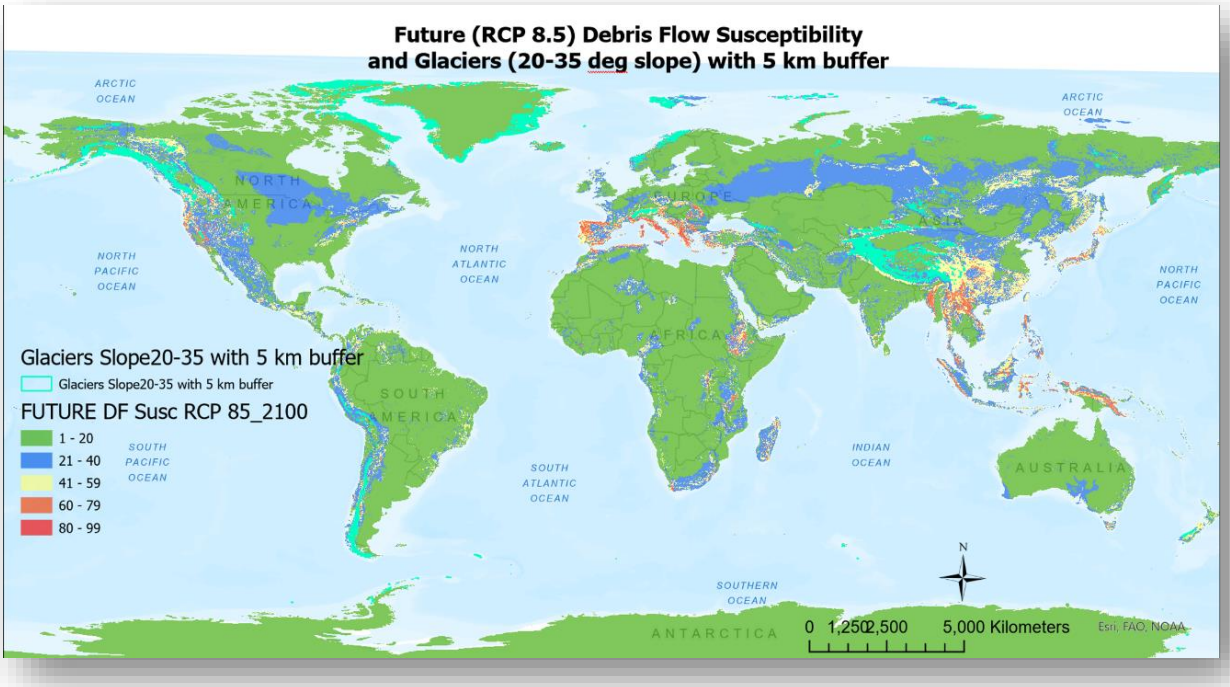


Figure 11.9 Future (RCP28.5 2100) debris flow susceptibility and buffered glaciers (with 20-35 degree slope).

The subset of glaciers was buffered by 5 km, as a representative boundary of potential influence to surrounding areas. This is a subjective choice, based loosely and anecdotally on findings in literature, since there are no existing studies (to my knowledge) with regard to the potential distance which a melting or retreating glacier may impact. The model is easily adjusted to use larger buffers.

The percent of glaciers within a Moderate, High, or Very High debris flow susceptibility area increases substantially from the current model to the RCP 2.6 and RCP 8.5 models, without wildfire probability (Table 11.3), however, the reverse is true when wildfire probability is included in the susceptibility models. As was found in Chapter 11.2, when debris flow susceptibility was augmented with wildfire probability, only the current debris flow model showed a substantial increase in area (68.2%) covered by the Moderate to Very High susceptibility classifications, whereas the RCP 2.6 and RCP 8.5 models saw a significant decrease in areal coverage (19.3% and 22.1%, respectively). These results are due to the decrease in wildfire global area coverage by 49% in the Moritz et al (Krawchuk, Moritz, Parisien, Van Dorn, & Hayhoe, 2009) 1971-2000 to the projected 2070-2099 models.

In both RCP 2.6 and RCP 8.5 scenarios, debris flow susceptibility with and without wildfire probability, the number of glaciers coincident with Moderate to Very High susceptibility areas, is significant.

Areas most at risk are the northwestern coasts of North America and Norway, western coast of South America, European Alps, and the Himalayan belt, under all modeled scenarios.

	Without Fire Probability			With Fire Probability		
Susc Class/Model	**Current**	**RCP 2.6**	**RCP 8.5**	**Current**	**RCP 2.6**	**RCP 8.5**
Very Low	85.4%	15.8%	9.1%	30.8%	16.8%	16.5%
Low	10.7%	63.6%	55.4%	29.5%	54.4%	57.3%
Moderate	3.4%	20.3%	33.3%	22.7%	27.6%	25.0%
High	0.5%	0.3%	2.3%	12.6%	1.2%	1.2%
Very High	0.01%	0	0	4.40%	0	0
Sum Moderate, High, Very High	**3.9%**	**20.6%**	**35.5%**	**39.6%**	**28.8%**	**26.2%**

Table 11.3 Percent of glaciers within debris susceptibility classification, by current and future models, with and without wildfire probability

12 Conclusion

12.1 Summary

A conservative approach was taken in all aspects of this research and with an attempt to reduce subjectivity. The ultimate objective was to develop current and future global debris flow susceptibility models and identify associated human vulnerability. There are several considerations which must be addressed prior to the final objectives.

The first question was "can the world be analyzed or modeled as one body for purposes of debris flow susceptibility or is this an oversimplification not accounting for the geologic, geomorphic, and tectonic histories and dissimilarities of the continents, and disregard for the latitudinal influences of climatic conditions?" This research demonstrates (Chapter 7) that a single global model performs exceptionally well in comparison with individual continental models, as evidenced by six of the seven continents when evaluating against a verification dataset and their presence in the Medium, High, and Very High susceptibility classes. There is significant concurrence (greater than 70% by area) between the susceptibility classifications in the continental and global models for Africa, Asia, Australia, Europe, North America, and South America. Greater than 95% of the single global model is within ± one susceptibility classification difference of the continental models, and less than 10% of the global model exhibits a greater susceptibility than the continental models.

Upon determining that a single global model, developed with Maximum Entropy, was the optimal approach, it was then necessary to look more deeply at the contributions to the model to further improve and refine the inputs and the resulting presentation.

When working at any scale, particularly global, it is often required to work with historical event inventories, rather than perform direct detailed field surveys of events. Due to the nature of the data collection methodologies and sources, historical inventories are prone to inaccuracies, uncertainties, imprecisions of various types and associated with various attributes of the inventory. Location is a critical attribute, and uncertainties associated with location will affect the determination of the most important environmental predisposing factors and factor classes

and therefore the susceptibility results. The global debris flow inventory used in this research has location uncertainties ranging from 1 km to 250 km. A number of methods were used to study the impact of these locational deviations and it was determined that a susceptibility model developed with a subset of the inventory (events with a location accuracy of <= 5 km) provides more accurate results, and certainly a more conservative representation. It is this model which is promoted as the present-day global debris flow susceptibility model.

With projected climate changes, and the known changes to debris flow susceptibility due to climate changes, over the past several decades, future global debris flow susceptibility was modeled based on climate and precipitation forecasts using a best-case scenario (RCP 2.6) and a worst-case scenario (RCP 8.5) for the decade 2100. Future models were analyzed with respect to one another and the present-day ("LTE 5 km") model. Results show that under both future scenarios there will be an increase in land susceptible to debris flows, and areas which will see an increase in susceptibility classification (i.e., from a lower susceptibility to Moderate, High, or Very High), and the locations of various susceptibilities may change. For those areas which increase to the Very High susceptibility class there is an increase of .01% (51,000 km^2) and .02% (102,000 km^2) of total land for RCP 2.6 and RCP 8.5, about the size of Costa Rica, and Iceland, respectively.

The potential human vulnerability to current and future debris flow susceptibility was studied using both population and urbanization data. The most notable result when studying the vulnerability, is that proximity to debris flow susceptibility by population data alone, may not be the most significant factor in understanding human vulnerability but rather urbanization, where there is a far greater overlap with debris flow susceptibility.

This research goes beyond the commonly employed landslide susceptibility modeling methods and inputs. The present-day and future susceptibility models developed herein were further studied and augmented with debris flow legacy, wildfire probability, and glacier melting/retreat data to determine whether and where there is an amplification, lessening, or no impact to susceptibility. As may be expected, all three were found to be critical indicators and factors associated with debris flow susceptibility which should continue to be considered in all susceptibility modeling efforts.

While local and regional hazard studies will always be essential, we must reframe our thinking and research to include studies at much broader scales. All hazard problems are global problems. We cannot nibble away at global problems, but rather must understand where problems exist, what problem areas may begin to coalesce, and where and when these problems may begin to encroach on human populations or vice versa. The more we know about our problems at the global scale the more information we will have to better understand and address them locally and regionally, as well.

Continental and global debris flow susceptibility analyses are an important foundation and first cut for an international perspective on the impact of debris flows on people and economies, as population, urbanization, and climate warming expand. With a much larger debris flow event inventory, and an unbiased geographic distribution, results can be further improved and refined in global and individual continental susceptibility models.

The present-day and future global debris flow models can be improved as more predisposing factors become available at global scales and higher resolution. The methods and outcomes of this work are an important foundation and framework for more in-depth studies with regard to vulnerability and hazard, as well as being used for other landslide typologies, and other types of hazards.

Future refinements to the models developed herein, and important follow-on research of interest to this author, include spatial-temporal modeling by evaluating the seasonal periodicity of high monthly mean precipitation values compared with the dates of the historical events, and debris flows identified as triggered by anomalous precipitation, evaluated with recorded anomalous high-intensity precipitation data. With relevant changes, the methodology followed herein will be applied to other landslide types such as co-seismic and rock avalanches, resulting in a suite of global models specific to each landslide type.

A summary of the graphics and tables for key models and findings resulting from this research are:

1. Present-day debris flow susceptibility model based on the events with <= 5 km location uncertainty (Fig. 12.1),
2. Present-day model augmented with wildfire probability (Fig. 12.2),
3. Present-day model overlay with glaciers (Fig. 12.3),
4. RCP 2.6 debris flow susceptibility model (Fig. 12.4),
5. RCP 2.6 model augmented with future wildfire probability (Fig. 12.5),
6. RCP 2.6 model overlay with glaciers (Fig. 12.6),
7. RCP 8.5 debris flow susceptibility model (Fig. 12.7),
8. RCP 8.5 model augment with future wildfire probability (Fig. 12.8),
9. RCP 8.5 model overlay with glaciers (Fig. 12.9)

Summary of key findings:

1. Percent of land by current, RCP 2.6, and RCP 8.5 models, by susceptibility classification (Table 12.1)
2. Susceptibility changes from current to future models (Table 12.2)
3. Susceptibility changes due to wildfire probability, by model Table 12.3)
4. Future populations within susceptibility classifications, by future models (Table 12.4)
5. Glacier distributions vis-à-vis susceptibility classifications current and future, with and without fire probability (Table 12.5)

12.2 Key Models

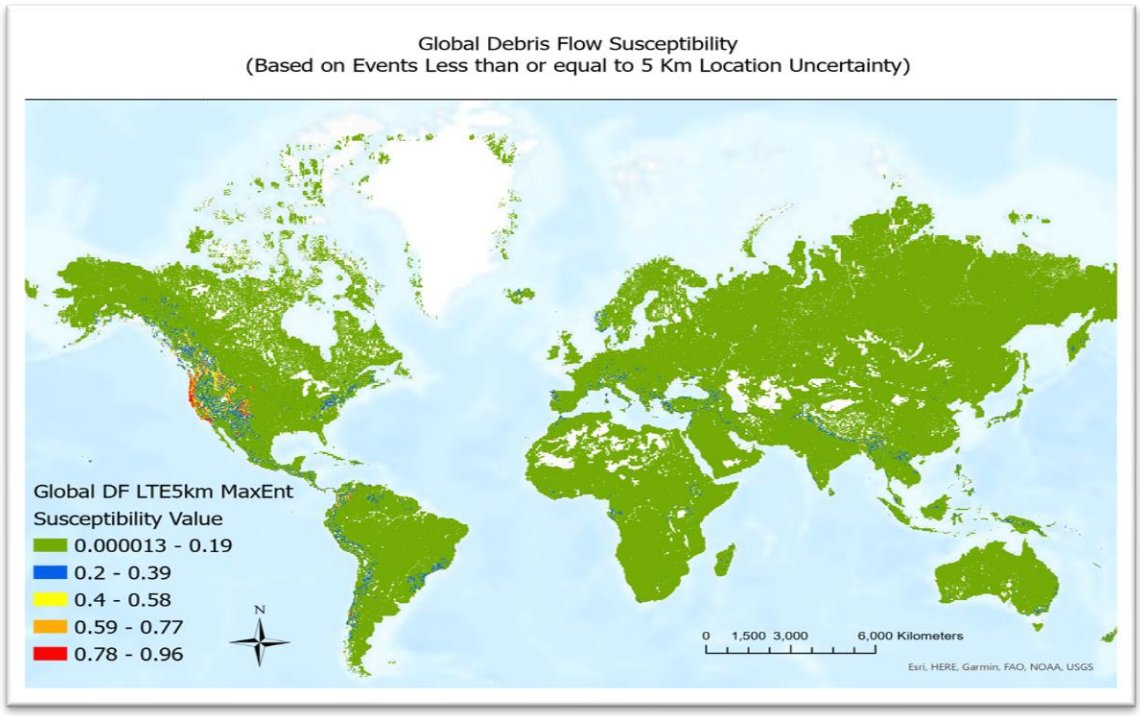


Figure 12.1 1. Present-day debris flow susceptibility model based on the events with <= 5 km location uncertainty

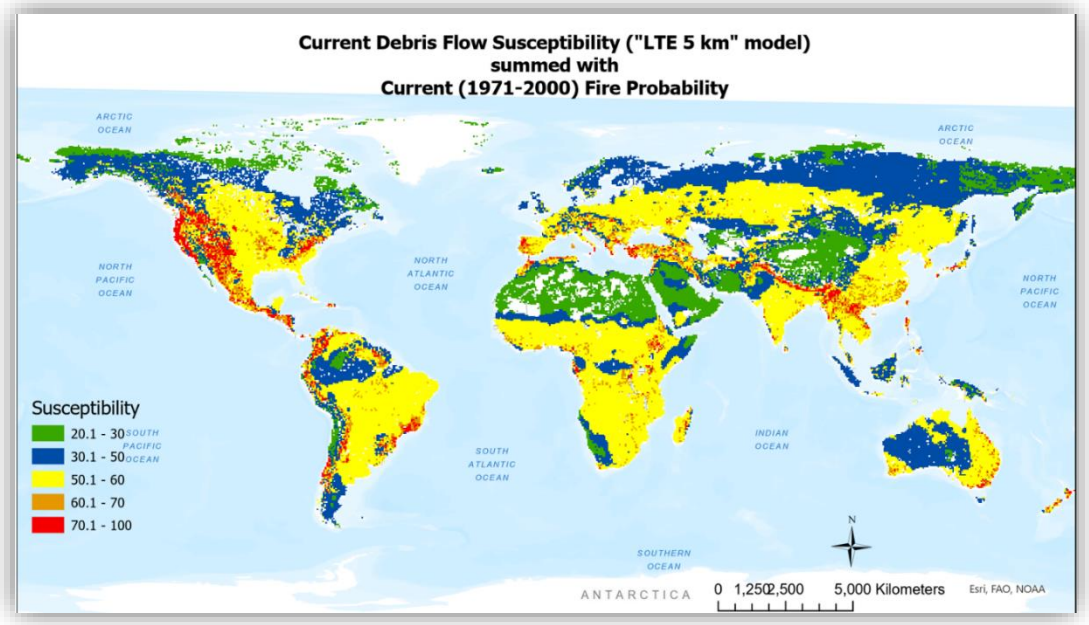


Figure 12.2 2. Present-day model augmented with wildfire probability

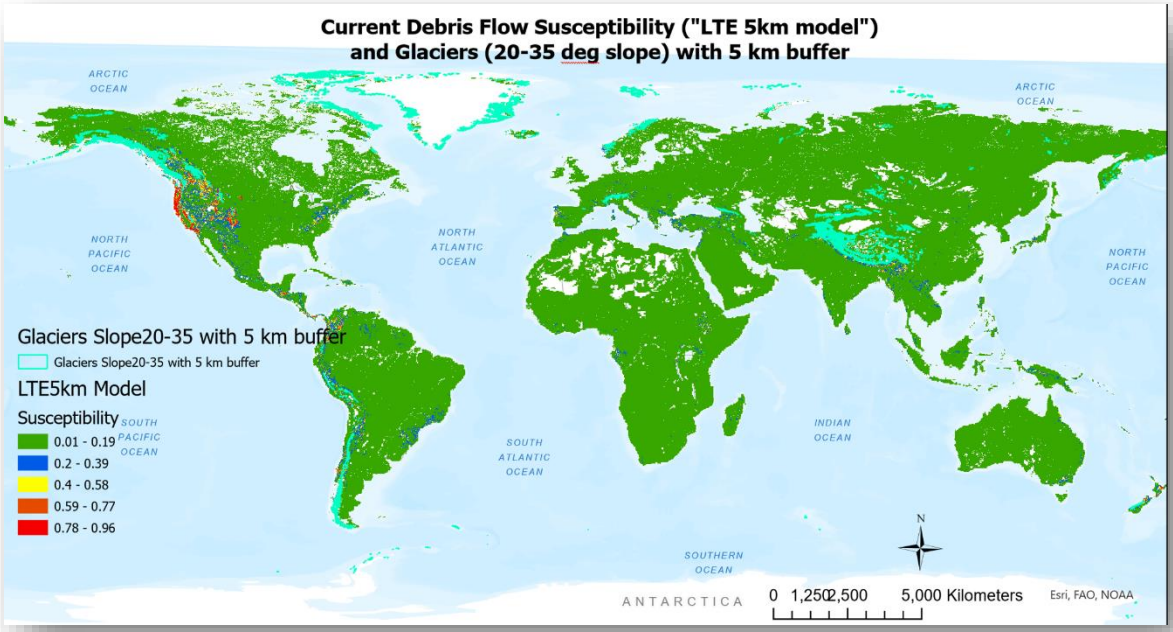


Figure 12.3 Present-day model overlay with glaciers

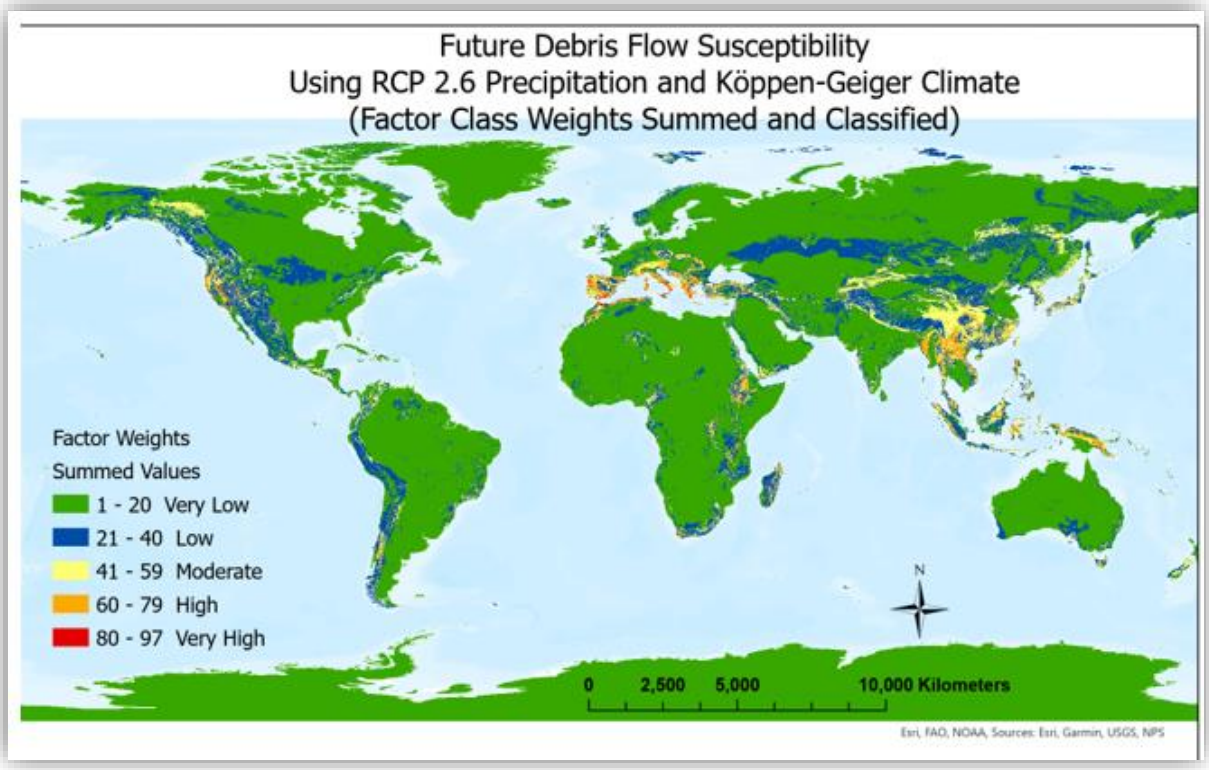


Figure 12.4 RCP 2.6 debris flow susceptibility model

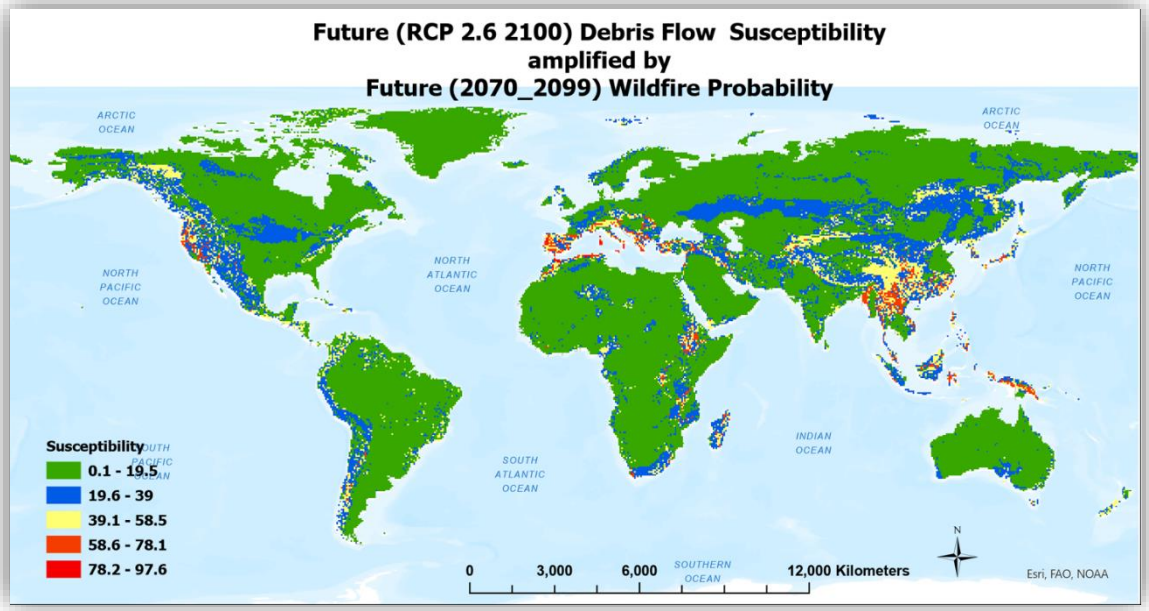


Figure 12.6 RCP 2.6 model augmented with future wildfire probability

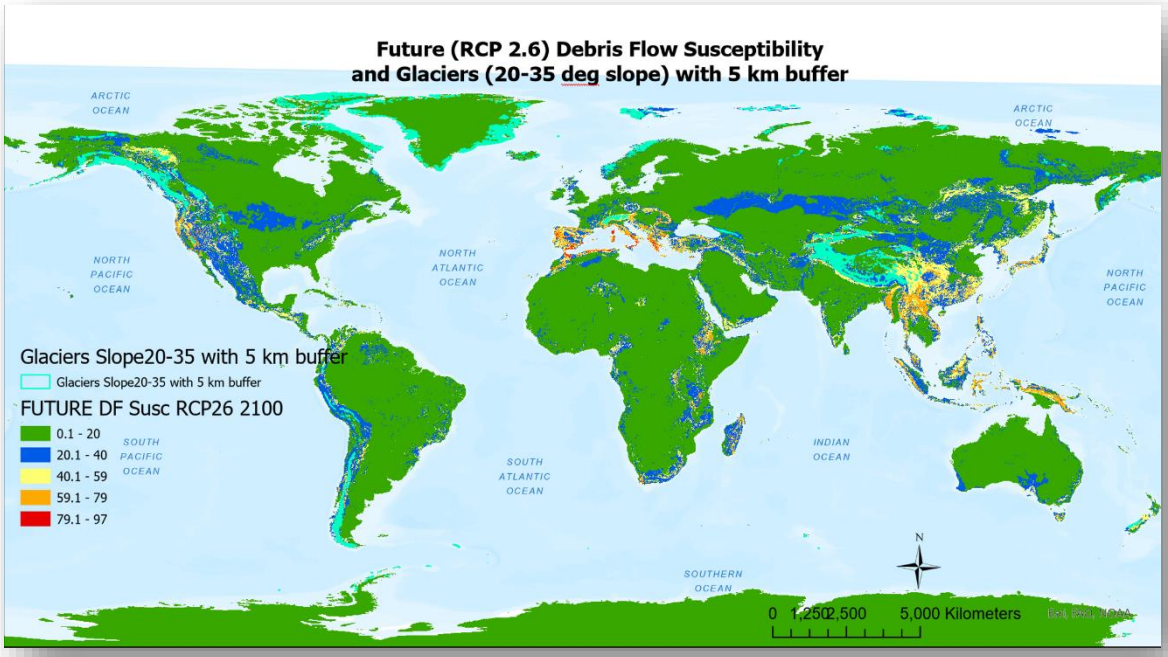


Figure 12.5 RCP 2.6 model overlay with glaciers

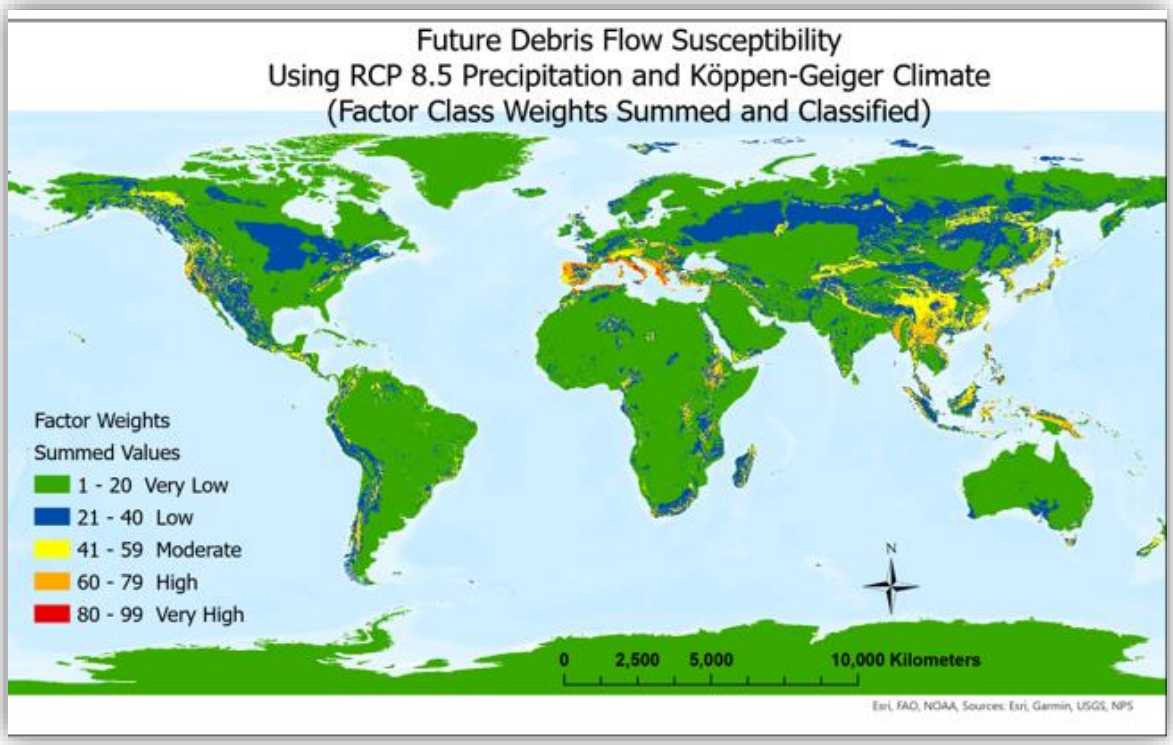


Figure 12.7 RCP 8.5 debris flow susceptibility model

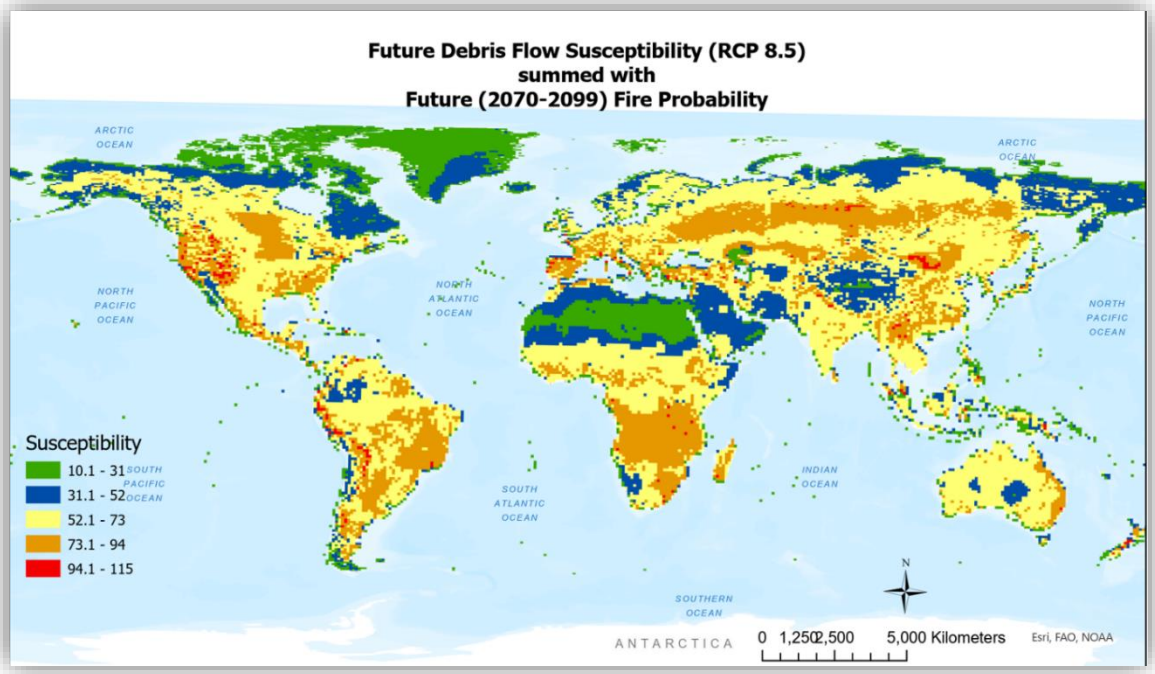


Figure 12.8 RCP 8.5 model augment with future wildfire probability

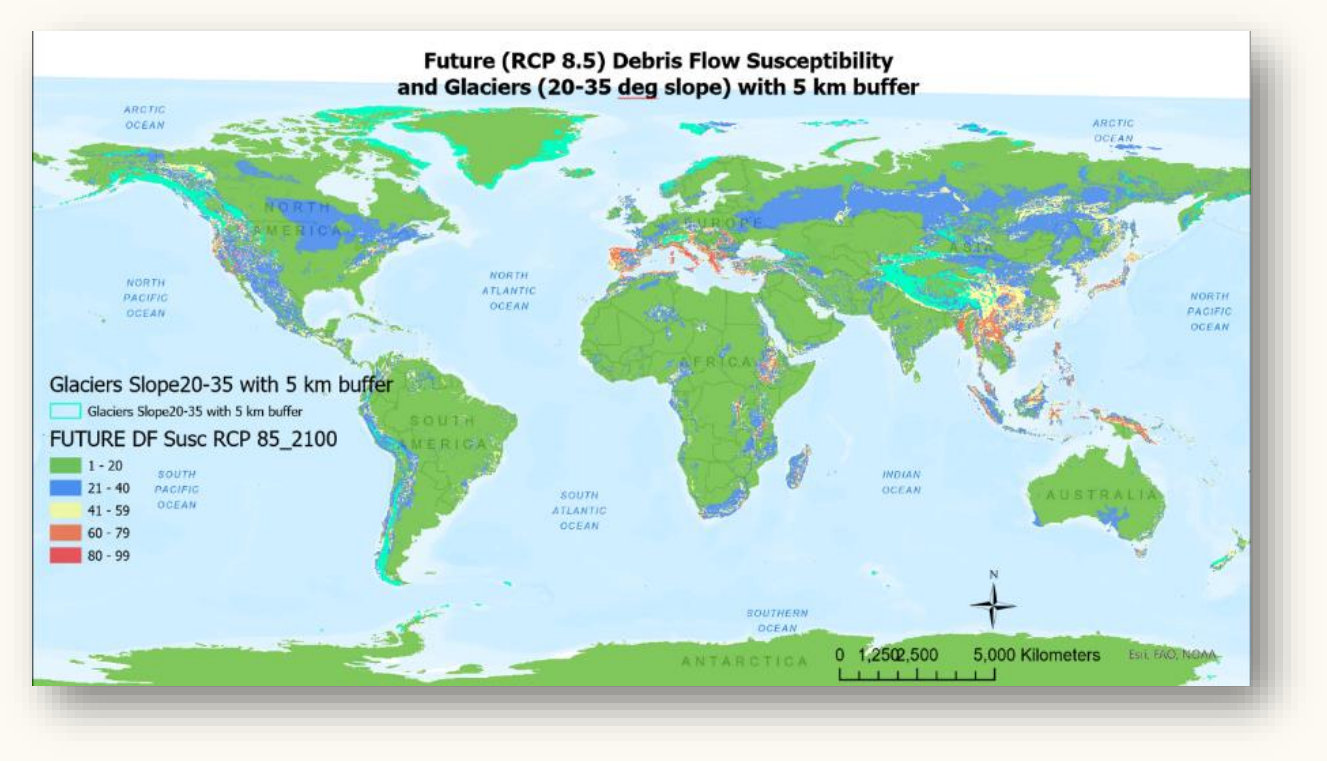


Figure 12.9 RCP 8.5 model overlay with glaciers

12.3 Key Findings

Susceptibility Classification/Model	Current	RCP 2.6	RCP 8.5
Very Low	91.63%	92.47%	90.95%
Low	5.21%	5.24%	6.74%
Moderate	2.54%	1.85%	1.89%
High	0.57%	0.41%	0.38%
Very High	0.05%	0.03%	0.04%
High and Very High	**0.62%**	**0.44%**	**0.42%**
Moderate, High, Very High	**3.16%**	**2.29%**	**2.31%**

Table 12.1 Percent of land by current, RCP 2.6, and RCP 8.5 models, by susceptibility classification

Susceptibility Change	From Current to RCP 2.6	From Current to RCP 8.5	From RCP 2.6 to RCP 8.5
INCREASE			
% of Total Area	1.49%	3.14%	2.35%
% of Changed Area	34.17%	51.47%	74.47%
DECREASE			
% of Total Area	2.87%	2.96%	0.80%
% of Changed Area	65.83%	48.53%	25.53%
INCREASE TO MODERATE, HIGH, VERY HIGH CLASSIFICATION			
% of Total Area	.47%	.63%	0.29%
% of Changed Area	10.79%	10.29%	9.09%

Table 12.2 2. Susceptibility changes from current to future models

	% Change Due to Fire		
Susceptibility Class	**Current**	**RCP 2.6**	**RCP 8.5**
Very Low	-82.6%	-19.3%	-22.1%
Low	14.5%	12.9%	15.8%
Moderate	56.0%	5.0%	4.9%
High	10.1%	1.3%	1.2%
Very High	2.1%	0.1%	0.1%
Sum Moderate to Very High	**68.2%**	**6.4%**	**6.2%**

Table 12.3 3. Susceptibility changes due to wildfire probability, by model

Susceptibility Class	% Population RCP 2.6	% Population RCP 8.5
Very Low	61.7%	63.1%
Low	25.5%	25.2%
Moderate	9.3%	8.5%
High	3.2%	2.9%
Very High	0.3%	0.3%
Total Moderate, High, Very High	**12.8%**	**11.7%**

Table 12.4 4. Future populations within susceptibility classifications, by future models

% glaciers within	Without Fire Probability			With Fire Probability		
Susc Class/Model	**Current**	**RCP 2.6**	**RCP 8.5**	**Current**	**RCP 2.6**	**RCP 8.5**
Very Low	85.4%	15.8%	9.1%	30.8%	16.8%	16.5%
Low	10.7%	63.6%	55.4%	29.5%	54.4%	57.3%
Moderate	3.4%	20.3%	33.3%	22.7%	27.6%	25.0%
High	0.5%	0.3%	2.3%	12.6%	1.2%	1.2%
Very High	0.01%	0	0	4.40%	0	0
Sum Moderate, High, Very High	**3.9%**	**20.6%**	**35.5%**	**39.6%**	**28.8%**	**26.2%**

Table 12.5 Glacier distributions vis-à-vis susceptibility classifications current and future, with and without fire probability

13 References

(CIESIN), C. f. I. E. S. I. N. (2013). *Global Roads Open Access Data Set, Version 1 (gROADSv1)* [geodatabase]. Retrieved from: http://sedac.ciesin.columbia.edu/data/set/groads-global-roadsopen-access-v1

Alvioli, M., Melillo, M., Guzzetti, F., Rossi, M., Palazzi, E., von Hardenberg, J., . . . Peruccacci, S. (2018). Implications of climate change on landslide hazard in Central Italy. *Sci Total Environ, 630*, 1528-1543. doi:10.1016/j.scitotenv.2018.02.315

Ardizzone, F., Cardinali, M., Carrara, A., Guzzetti, F., & Reichenbach, P. (2002). Impact of mapping errors on the reliability of landslide hazard maps. *Natural Hazards and Earth System Sciences, 2*(1-2), 3-14. doi:DOI 10.5194/nhess-2-3-2002

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Koppen-Geiger climate classification maps at 1-km resolution. *Scientific Data, 5*, 180214. doi:10.1038/sdata.2018.214

Bloschl, G., & Montanari, A. (2009). Climate change impactsâ" throwing the dice? *Hydrological Processes*, n/a-n/a. doi:10.1002/hyp.7574

Borgatti, L., & Soldati, M. (2010). Landslides as a geomorphological proxy for climate change: A record from the Dolomites (northern Italy). *Geomorphology, 120*(1-2), 56-64. doi:10.1016/j.geomorph.2009.09.015

Brabb, E. E., Colgan, J. P., & Best, T. C. (1999). *Map showing inventory and regional susceptibility for Holocene debris flows and related fast-moving landslides in the conterminous United States*.

Bridges, E. M. (2012). *World Geomorphology*. Cambridge: Cambridge University Press.

Brighenti, R., Segalini, A., & Ferrero, A. M. (2013). Debris flow hazard mitigation: A simplified analytical model for the design of flexible barriers. *Computers and Geotechnics, 54*, 1-15. doi:10.1016/j.compgeo.2013.05.010

Broeckx, J., Vanmaercke, M., Duchateau, R., & Poesen, J. (2018). A data-based landslide susceptibility map of Africa. *Earth-Science Reviews, 185*, 102-121. doi:10.1016/j.earscirev.2018.05.002

Campbell, R. H. (1974). Debris flows originating from soil slips during rainstorms in Southern California. *Journal of Engineering Geology, 7*, 339-349.

Cannon, S., & Reneau, S. (2000). Conditions for generation of fire-related debris flows, Capulin Canyon, New Mexico. *Earth Surface Processes and Landforms, 25*. doi:10.1002/1096-9837(200009)25:103.0.CO;2-H

Cannon, S. H., Bigio, E. R., & Mine, E. (2001). A process for fire-related debris flow initiation, Cerro Grande fire, New Mexico. *Hydrological Processes, 15*, 3011-3023.

Cannon, S. H., & DeGraff, J. V. (2009). *Cascading consequences of climate change and expanding population on the threat of wildfire and post fire debris-flow hazards, Western U.S. (Invited).* Paper presented at the American Geophysical Union.

Cannon, S. H., Gartner, J. E., Rupert, M. G., Michael, J. A., Rea, A. H., & Parrett, C. (2009). Predicting the probability and volume of postwildfire debris flows in the intermountain western United States. *Geological Society of America Bulletin, 122*(1-2), 127-144. doi:10.1130/b26459.1

Capitani, M., Ribolini, A., & Bini, M. (2013). The slope aspect: A predisposing factor for landsliding? *Comptes Rendus Geoscience, 345*(11-12), 427-438. doi:10.1016/j.crte.2013.11.002

Carrara, A., Cardinali, M., & Guzzetti, F. (1992). Uncertainty in assessing landslide hazard and risk. *ITC Journal, 2*, 172-183.

Cavallo, E., & Noy, I. (2010). *The Economics of Natural Disasters - A Survey*. IDB Working Paper Series No. IDB-WP-124. Working Paper. Inter-American Development Bank. Retrieved from http://www.iadb.org/res/publications/pubfiles/

Chen, G., Li, X., Liu, X., Chen, Y., Liang, X., Leng, J., . . . Huang, K. (2020). Global projections of future urban land expansion under shared socioeconomic pathways. *Nat Commun, 11*(1), 537. doi:10.1038/s41467-020-14386-x

Chen, L., van Westen, C. J., Hussin, H., Ciurean, R. L., Turkington, T., Chavarro-Rincon, D., & Shrestha, D. P. (2016). Integrating expert opinion with modelling for quantitative multi-hazard risk assessment in the Eastern Italian Alps. *Geomorphology, 273*, 150-167. doi:10.1016/j.geomorph.2016.07.041

Chen, N. S., Zhou, W., Yang, C. L., Hu, G. S., Gao, Y. C., & Han, D. (2010). The processes and mechanism of failure and debris flow initiation for gravel soil with different clay content. *Geomorphology, 121*(3-4), 222-230. doi:10.1016/j.geomorph.2010.04.017

Cheng, W., Wang, N., Zhao, M., & Zhao, S. (2016). Relative tectonics and debris flow hazards in the Beijing mountain area from DEM-derived geomorphic indices and drainage analysis. *Geomorphology, 257*, 134-142. doi:10.1016/j.geomorph.2016.01.003

Cherlet, M., Hutchinson, C., Reynolds, J., Hill, J., Sommer, S., & von Maltitz, G. (2018). World Atlas of Desertification - Third Edition. In (3rd ed., pp. 248). Luxembourg: Office of the European Union.

Chiarle, M., Geertsema, M., Mortara, G., & Clague, J. J. (2011). Impacts f climate change on debris flow occurrence in the cordillera of western Canada and the European Alps. *Italian Journal of Engineering Geology and Environment*.

Chiarle, M., Iannotti, S., Mortara, G., & Deline, P. (2007). Recent debris flow occurrences associated with glaciers in the Alps. *Global and Planetary Change, 56*(1-2), 123-136. doi:10.1016/j.gloplacha.2006.07.003

Chrisman, N. (1989). Modeling error in overlaid categorical maps. In M. Goodchild & S. Gopal (Eds.), *The Accuracy of Spatial Databases* (pp. 21-43). Pennsylvania: Taylor & Francis.

Convertino, M., Troccoli, A., & Catani, F. (2013). Detecting fingerprints of landslide drivers: A MaxEnt model. *Journnal of Geophysical Research: Earth Surface, 118*, 1367-1386. doi:10.1002/jgrf.20099,

Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J. P., Fotopoulou, S., . . . Smith, J. T. (2013). Recommendations for the quantitative analysis of landslide risk. *Bulletin of Engineering Geology and the Environment*. doi:10.1007/s10064-013-0538-8

Costa, J. E. (1984). Physical Geomorphology of Debris Flows. In J. E. Costa & P. J. Fleisher (Eds.), *Developments and Applications of Geomorphology* (pp. 268-317). Berlin, Heidelberg: Springer Berlin Heidelberg.

Couclelis, H. (2003). The Certainty of Uncertainty: GIS and the Limits of Geographic Knowledge. *Transactions in GIS, 7*(2), 165-175. doi:10.1111/1467-9671.00138

Crozier, M. J., Glade, T., & Preston, N. (2013). Antecedent Conditions. In P. Bobrowsky (Ed.), *Encyclopedia of Natural Hazards* (pp. 10-13): Springer.

Cui, D., Liang, S., Wang, D., & Liu, Z. (2021). A 1-km global dataset of historical (1979-2017) and future (2020-2100) Koppen-Geiger climate classification and bioclimatic variables. *Earth Science Data*. doi:10.5194/essd-2021-53

D'Amato Avanzi, G., Giannecchini, R., & Puccinelli, A. (2004). The influence of the geological and geomorphological settings on shallow landslides. An example in a temperate climate environment: the June 19, 1996 event in northwestern Tuscany (Italy). *Engineering Geology, 73*(3-4), 215-228. doi:10.1016/j.enggeo.2004.01.005

Danielson, J. J., Gesch, D. B. (2011). *Global multi-resolution Terrain Elevation Data 2010 (GMTED2010)*. Reston, Virginia: U.S. Geological Survey Retrieved from https://pubs.usgs.gov/of/2011/1073/pdf/of2011-1073.pdf

De Graff, J. V. (2018). A rationale for effective post-fire debris flow mitigation within forested terrain. *Geoenvironmental Disasters, 5*(1). doi:10.1186/s40677-018-0099-z

Devkota, K. C., Regmi, A. D., Pourghasemi, H. R., Yoshida, K., Pradhan, B., Ryu, I. C., . . . Althuwaynee, O. F. (2012). Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya. *Natural Hazards, 65*(1), 135-165. doi:10.1007/s11069-012-0347-6

Dilley, M., Chen, R.S., Uwe, D., Lerner-Lam, A., Arnold, M. , Agwe, J., Buys, P., Kjekstad, O., Lyon, B., Yetman, G. (2005). *Natural Disaster Hotspots, A Global Risk Analysis* (Vol. 5). Washington, D.C.: The World Bank.

Diop, S. (2012). *An Overview of Landslide Occurrence, Inventorization and Susceptibility Mapping in South Africa*. Paper presented at the UR Forum Mapping Global Risk.

Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A., & Maher, N. (2016). More extreme precipitation in the world's dry and wet regions. *Nature Climate Change, 6*(5), 508-513. doi:10.1038/nclimate2941

Dou, J., Tien Bui, D., Yunus, A. P., Jia, K., Song, X., Revhaug, I., . . . Zhu, Z. (2015). Optimization of Causative Factors for Landslide Susceptibility Evaluation Using Remote Sensing and GIS Data in Parts of Niigata, Japan. *PLoS ONE, 10*(7), e0133262. doi:10.1371/journal.pone.0133262

Dowling, C. A., & Santi, P. M. (2013). Debris flows and their toll on human life: a global analysis of debris-flow fatalities from 1950 to 2011. *Natural Hazards, 71*(1), 203-227. doi:10.1007/s11069-013-0907-4

Esri, I. (2015). *World Water Bodies* [Polygon shapefile].

Esri, I. (2020). ArcGIS Pro (Version 2.7): Esri, Inc. Retrieved from https://esri.com/en-us/arcgis/products/arcgis-pro/overview

FAO, F. a. A. O. o. t. U. N. (2003). *The Digital Soil Map of the World*.

Farahmand, A., & AghaKouchak, A. (2013). A satellite-based global landslide model. *Natural Hazards and Earth System Sciences, 13*(5), 1259-1267. doi:10.5194/nhess-13-1259-2013

Felsberg, A., Poesen, J., Bechtold, M., Vanmaercke, M., & De Lannoy, G. J. M. (2021). Estimating global landslide susceptibility and its uncertainty through ensemble modelling. *Natural Hazards and Earth System Science*. doi:10.5194/nhess-2021-360

Ferentinou, M., & Chalkias, C. (2011). *Mapping Mass Movement Susceptibility across Greece with GIS, ANN and Statistical Methods.* Paper presented at the Second World Landslide Forum, Rome, Italy.

Fick, S. E., & Hijmans, R. J. (2017). *WorldClim 2.1: new 1km spatial resolution climate surfaces for global land areas.* [10 minutes (~340 km2) Geotiff]. Monthly Precipitation. Retrieved from: https://worldclim.org/data/worldclim21.html

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences, 18*(8), 2161-2181. doi:10.5194/nhess-18-2161-2018

Froude, M. J., & Petley, D. N. (2019). *Guide to the Global Fatal Landslide Database on ArcGIS Online*

Frye, C., Sayre, R., Soller, D., & Karagulle, D. (2018). *World Named Landforms*.

Frye, C., Sayre, R. J., & Soller, D. R. (2017). *World Named Landforms*.

Gabet, E. J., & Bookter, A. (2008). A morphometric analysis of gullies scoured by post-fire progressively bulked debris flows in southwest Montana, USA. *Geomorphology, 96*(3-4), 298-309. doi:10.1016/j.geomorph.2007.03.016

GÁL, A., Poszet, S. L., & Kerekes, A. H. (2018). Landslide susceptibility assessment using the maximum entropy model in a sector of the Cluj–Napoca Municipality, Romania. *Revista de Geomorfologie, 20*(1), 130-146. doi:10.21094/rg.2018.039

GAO, J. (2019). *Downscaling Global Spatial Population Projections from 1/8-degree to 1-km Grid Cells*. Retrieved from: https://search.earthdata.nasa.gov/search?p=C1399902925-SEDAC&m=-0.0703125!0!2!1!0!0%2C2&q=population&ok=population

Gariano, S. L., & Guzzetti, F. (2016). Landslides in a changing climate. *Earth-Science Reviews, 162*, 227-252. doi:10.1016/j.earscirev.2016.08.011

Geertsema, M., & Clague, J. J. (2011). Impacts of climate change on debris flow occurrence in the Cordillera of western Canada and the European Alps. *Italian Journal of Engineering Geology and Environment, 3*. doi:10.4408/IJEGE.2011-03.B-006

Geological Survey of India. (2021). Landslide Hazard. Retrieved from https://www.gsi.gov.in/webcenter/portal/OCBIS/pageGeoInfo/pageLANDSLIDEHAZRD?_afrLoop=35938692434541212&_adf.ctrl-state=13osiqwfwl_59#!%40%40%3F_afrLoop%3D35938692434541212%26_adf.ctrl-state%3D13osiqwfwl_63

Ghosh, S., Carranza, E. J. M., van Westen, C. J., Jetten, V. G., & Bhattacharya, D. N. (2011). Selecting and weighting spatial predictors for empirical modeling of landslide susceptibility in the Darjeeling Himalayas (India). *Geomorphology, 131*(1-2), 35-56. doi:10.1016/j.geomorph.2011.04.019

Gridded Population of the World, Version 3 (GPWv3): Population Count Grid.

Grozavu, A., & Patriche, C. V. (2013). Landslide Susceptibility Assessment: GIS Application to a Complex Mountainus Environment. doi:10.1007/978-3-642-12725-0_4

Grozavu, A., Pleşcan, S., Patriche, C. V., Mărgărint, M. C., & Roşca, B. (2013). Landslide Susceptibility Assessment: GIS Application to a Complex Mountainous Environment. In J. Kozak, K. Ostapowics, A. Bytnerowicz, & B. Wyzga (Eds.), *The Carpathians: Integrating Nature and Society Towards Sustainability* (pp. 31-44). Heidelberg: Springer Berlin.

Gunther, A., Reichenbach, P., Guzzetti, F., & Richter, A. (2007). *Criteria for the identification of landslide risk areas in Europe: the Tier 1 approach*. Retrieved from Luxembourg: https://publications.jrc.ec.europa.eu

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K.-T. (2012). Landslide inventory maps: New tools for an old problem. *Earth-Science Reviews, 112*, 42-66.

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., & Galli, M. (2006). Estimating the quality of landslide susceptibility models. *Geomorphology, 81*(1-2), 166-184. doi:10.1016/j.geomorph.2006.04.007

Haque, U., da Silva, P. F., Devoli, G., Pilz, J., Zhao, B., Khaloua, A., . . . Glass, G. E. (2019). The human cost of global warming: Deadly landslides and their triggers (1995-2014). *Sci Total Environ, 682*, 673-684. doi:10.1016/j.scitotenv.2019.03.415

Harrison, S., Mighall, T., Stainforth, D. A., Allen, P., Macklin, M., Anderson, E., . . . Shannon, S. (2019). Uncertainty in geomorphological responses to climate change. *Climatic Change, 156*(1-2), 69-86. doi:10.1007/s10584-019-02520-8

Hartmann, J., & Moosdorf, N. (2012). *Global Lithological Map Database v1.0 (gridded to 0.5° spatial resolution)*. World Lithologic Map.

Hartmann, J., & Moosdorf, N. (2012). The new global lihological map database GLiM: A representation of rock properties at the Earth surface. *Geochemistry, Geophysics, Geosystems, 13*(12). doi:10.1029/2012GC004370

Hartmann, J., & Moosdorf, N. (2012). The new global lithological map database GLiM: A representation of rock properties at the Earth surface. *Geochemistry, Geophysics, Geosystems, 13*(12). doi:10.1029/2012gc004370

Heckmann, T., Gegg, K., Gegg, A., & Becht, M. (2014). Sample size matters: investigating the effect of sample size on a logistic regression susceptibility model for debris flows. *Natural Hazards and Earth System Sciences, 14*(2), 259-278. doi:10.5194/nhess-14-259-2014

Highland, L. M., & Bobrowsky, P. (2008). *The Landslide Handbook - A Guide to Understanding Landslides*. U.S. Department of the Interior

U.S. Geological Survey

Holec, J., Bednarik, M., Liscak, P., Zilka, A., & Vitovic, L. (2018). Assessment of debris flow susceptibility using bivariate and multivariate statistical analyses and verification based on catastrophic events from 2014 in the Krivanska Fatra Mountains, Slovakia. *Acta Geologica Slovaca, 10*, 1-19.

Hong, Y., Adler, R., & Huffman, G. (2007). Use of satellite remote sensing data in the mapping of global landslide susceptibility. *Natural Hazards, 43*, 245-256. doi:10.1007/s11069-006-9104-z.

Huggel, C., Clague, J. J., & Korup, O. (2012). Is climate change responsible for changing landslide activity in high mountains? *Earth Surface Processes and Landforms, 37*(1), 77-91. doi:10.1002/esp.2223

Huggel, C., Khabarov, M., Korup, O., & Obersteiner, M. (2012). Physical impacts of climate change on landslide occurrence and related adaptation. In J. J. Clague & D. Stead (Eds.), *Landslides: Types, Mechanisms, and Modeling*. Cambridge: Cambridge University Press.

Huggel, C., Salzmann, N., Allen, S., Caplan-Auerbach, J., Fischer, L., Haeberli, W., . . . Wessels, R. (2010). Recent and future warm extreme events and high-mountain slope stability. *Philos Trans A Math Phys Eng Sci, 368*(1919), 2435-2459. doi:10.1098/rsta.2010.0078

Hungr, O., Evans, S., Bovis, M., & Hutchinson, J. N. (2001). *Review of the classification of landslides of the flow type* (Vol. 7).

Hungr, O., Leroueil, S., & Picarelli, L. (2013). The Varnes classification of landslide types, an update. *Landslides, 11*(2), 167-194. doi:10.1007/s10346-013-0436-y

Hunter, G. J. (1999). Managing uncertainty in GIS. *Geographical Information Systems, 2*, 633-641.

Hunter, G. J., & Goodchild, M. (1996). Communicating uncertainty in spatial databases. *Transactions in GIS, 1*(1), 13-24. doi:doi:10.1111/j.1467-9671.1996.tb00030.x

Hunter, G. J., Goodchild, M., & Robey, M. (1994). *A Toolbox for Assessing Uncertainty in Spatial Databases*. Paper presented at the 22nd Annual Conference of the Australasian Urban and Regional Information Systems Association, Inc., Sydney.

IPCC. (2019). *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*. Retrieved from

ISPRA, I. S. p. l. P. e. l. R. A. (2020). World Landslides Forum: Each year, landslide cause damage to 6 billion Euros [Press release]

Iverson, R. M. (1997). The Physics of Debris Flows. *American Geophysical Union, 35*(3), 245-296.

Jacobs, L., Kervyn, M., Reichenbach, P., Rossi, M., Marchesini, I., Alvioli, M., & Dewitte, O. (2020). Regional susceptibility assessments with heterogeneous landslide information: Slope unit- vs. pixel-based approach. *Geomorphology, 356*. doi:10.1016/j.geomorph.2020.107084

Jansen, L. J. M., & Gregorio, A. (2000). *Land Cover Classification System (LCCS): Classification Concepts and User Manual*. Rome: Food and Agriculture Organization of the United Nations.

Jia, G., Alvioli, M., Gariano, S. L., Marchesini, I., Guzzetti, F., & Tang, Q. (2021). A global landslide non-susceptibility map. *Geomorphology, 389*. doi:10.1016/j.geomorph.2021.107804

Jiang, L., & O'Neill, B. C. (2017). Global urbanization projections for the Shared Socioeconomic Pathways. *Global Environmental Change, 42*, 193-199. doi:10.1016/j.gloenvcha.2015.03.008

Jones, B., & O'Neill, B. C. (2016). Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. *Environmental Research Letters, 11*(8). doi:10.1088/1748-9326/11/8/084003

Jones, D. K. C. (1993). Global warming and geomorphology. *The Geographical Journal, 159*, 124-130. doi:0016-7398/93/0002-0124

Keiler, M., Knight, J., & Harrison, S. (2010). Climate change and geomorphological hazards in the eastern European Alps. *Philos Trans A Math Phys Eng Sci, 368*(1919), 2461-2479. doi:10.1098/rsta.2010.0047

Kirschbaum, D., & Stanley, T. (2018). Satellite-Based Assessment of Rainfall-Triggered Landslide Hazard for Situational Awareness. *Earth's Future, 6*(3), 505-523. doi:10.1002/2017ef000715

Kirschbaum, D., Stanley, T., & Yatheendradas, S. (2015). Modeling landslide susceptibility over large regions with fuzzy overlay. *Landslides, 13*(3), 485-496. doi:10.1007/s10346-015-0577-2

Kirschbaum, D., Stanley, T., & Zhou, Y. (2015). Spatial and temporal analysis of a global landslide catalog. *Geomorphology, 249*, 4-15. doi:10.1016/j.geomorph.2015.03.016

Kirschbaum, D., Stanley, T., & Zhou, Y. P. (2015). Spatial and temporal analysis of a global landslide catalog. *Geomorphology, 249*, 4-15. doi:10.1016/j.geomorph.2015.03.016

Knight, J., & Harrison, S. (2012). Evaluating the impacts of global warming on geomorphological systems. *Ambio, 41*(2), 206-210. doi:10.1007/s13280-011-0178-9

Kobayashi, T., Tateishi, R., Alsaaideh, B., Sharma, R. C., Wakaizumi, T., Miyamoto, D., . . . Maitiniyazi, A. (2017). *Production of Global Land Cover Data - GLCNMO2013* [15 ArcSecond (500 m), raster]. Landuse/Landcover.

Kornejady, A., Ownegh, M., & Bahremand, A. (2017). Landslide susceptibility assessment using maximum entropy model with two different data sampling methods. *Catena, 152*, 144-162. doi:10.1016/j.catena.2017.01.010

Kottek, M. J., Grieser, C. B., Rudolf, B., & Rubel, F. (2006). *World Map of the Koppen-Geiger climate classification updated.*

Krawchuk, M. A., Moritz, M. A., Parisien, M. A., Van Dorn, J., & Hayhoe, K. (2009). Global pyrogeography: the current and future distribution of wildfire. *PLoS ONE, 4*(4), e5102. doi:10.1371/journal.pone.0005102

Lavee, H., Imeson, A. C., & Sarah, P. (1998). The impact of climate change on geomorphology and desertification along a Mediterranean-arid transect *Land Degradation and Development, 9*, 407-422. doi:10.1002/(SICI)1099-145x(199809/10)9:53.0.CO

Lavee, H., Kutiel, P., Segev, M., & Benyamini, Y. (1995). Effect of surface roughness on runff and erosion in a Mediterranean ecosystem: the role of fire. *Geomorphology, 11*, 227-234.

Legg, N. T., Meigs, A. J., Grant, G. E., & Kennard, P. (2014). Debris flow initiation in proglacial gullies on Mount Rainier, Washington. *Geomorphology, 226*, 249-260. doi:10.1016/j.geomorph.2014.08.003

Leoni, G., Barchiesi, F., Catallo, F., Dramis, F., Fubelli, G., Lucifora, S., . . . Puglisi, C. (2009). GIS Methodology to Assess Landslide Susceptibility: Application to a River Catchment of Central Italy. *Journal of Maps, 5*(1), 87-93. doi:10.4113/jom.2009.1041

Li, M., Tian, C.-s., Wang, Y.-k., Liu, Q., Lu, Y.-f., & Shan, W. (2018). Impacts of future climate change (2030-2059) on debris flow hazard: A case study in the Upper Minjiang River basin, China. *Journal of Mountain Science, 15*(8), 1836-1850. doi:10.1007/s11629-017-4787-z

Lin, L., Lin, Q., & Wang, Y. (2017). Landslide susceptibility mapping on a global scale using the method of logistic regression. *Natural Hazards and Earth System Sciences, 17*(8), 1411-1424. doi:10.5194/nhess-17-1411-2017

Lombardo, L., Fubelli, G., Amato, G., & Bonasera, M. (2016). Presence-only approach to assess landslide triggering-thickness susceptibility: a test for the Mili catchment (north-eastern Sicily, Italy). *Natural Hazards: Journal of the International Society for the Prevention and Mitigation of Natural Hazards, 84*(1), 565-588. Retrieved from https://EconPapers.repec.org/RePEc:spr:nathaz:v:84:y:2016:i:1:d:10.1007_s11069-016-2443-5

Lombardo, L., & Tanyas, H. (2021). From scenario-based seismic hazard to scenario-based landslide hazard: fast-forwarding to the future via statistical simulations. *Stochastic Environmental Research and Risk Assessment*. doi:10.1007/s00477-021-02020-1

Lonigro, T., Gentile, F., & Polemio, M. (2015). The influence of climate variability and land use variations on the occurrence of landslide events (Subappennino Dauno, Southern Italy). *Rendiconti Online della Società Geologica Italiana, 35*, 192-195. doi:10.3301/rol.2015.98

Lorente, A., Garcia-Ruiz, J., Beguería, S., & Arnaez, J. (2002). Factors explaining the spatial distribution of hillslope debris flows. *Mountain Research and Development, 22*, 32-39. doi:10.1659/02764741(2002)022[0032:FETSDO]2.0.CO;2

Lorente, A., García-Ruiz, J. M., Beguería, S., & Arnáez, J. (2002). Factors Explaining the Spatial Distribution of Hillslope Debris Flows. *Mountain Research and Development, 22*(1), 32-39. doi:10.1659/0276-4741(2002)022[0032:Fetsdo]2.0.Co;2

Lupiano, V., Rago, V., Terranova, O. G., & Iovine, G. (2019). Landslide inventory and main geomorphological features affecting slope stability in the Picentino river basin (Campania, southern Italy). *Journal of Maps*, 1-11. doi:10.1080/17445647.2018.1563836

Maffini, G., Arno, M., & Bitterlich, W. (1989). Observations and comments on the generation and treatment of error in digital GIS data. In M. Goodchild & S. Gopal (Eds.), *Accuracy of Spatial Databases* (pp. 55-68). Philadelphia: Taylor and Francis.

Malamud, B. D., Turcotte, D. L., Guzzetti, F., & Reichenbach, P. (2004). Landslide inventories and their statistical properties. *Earth Surface Processes and Landforms, 29*(6), 687-711. doi:10.1002/esp.1064

Meten, M., PrakashBhandary, N., & Yatabe, R. (2015). Effect of Landslide Factor Combinations on the Prediction Accuracy of Landslide Susceptibility Maps in the Blue Nile Gorge of Central Ethiopia. *Geoenvironmental Disasters, 2*(1). doi:10.1186/s40677-015-0016-7

Meyer, G. A., & Wells, S. G. Fire-related sedimentation events on alluvial fans, Yellowstone National Park, U.S.A. *Journal of Sedimentary Research, 67 (5)*, 776-791.

Middleton, N., & Thomas, D. S. (1997). World Atlas of Desertification - Second Edition. In (2nd ed.). London: United Nations Environment Program.

Moody, J. A., Shakesby, R. A., Robichaud, P. R., Cannon, S. H., & Martin, D. A. (2013). Current research issues related to post-wildfire runoff and erosion processes. *Earth-Science Reviews, 122*, 10-37. doi:10.1016/j.earscirev.2013.03.004

Moreiras, S. M., Sepúlveda, S. A., Correas-González, M., Lauro, C., Vergara, I., Jeanneret, P., . . . Lara, M. (2021). Debris Flows Occurrence in the Semiarid Central Andes under Climate Change Scenario. *Geosciences, 11*(2). doi:10.3390/geosciences11020043

Moritz, M. A., Parisien, M.-A., Batllori, E., Krawchuk, M. A., Van Dorn, J., Ganz, D. J., & Hayhoe, K. (2012). Climate change and disruptions to global fire activity. *Ecosphere, 3*(6). doi:10.1890/es11-00345.1

Nadim, F., Kjekstad, O., Peduzzi, P., Herold, C., & Jaedicke, C. (2006). Global landslide and avalanche hotspots. *Landslides, 3*(2), 159-173. doi:10.1007/s10346-006-0036-1

Nakileza, B. R., & Nedala, S. (2020). Topographic influence on landslides characteristics and implication for risk management in upper Manafwa catchment, Mt Elgon Uganda. *Geoenvironmental Disasters, 7*(1). doi:10.1186/s40677-020-00160-0

Neary, D. G., Klopatek, C. C., DeBano, L. F., & Ffolliott, P. F. (1999). Fire effects on belowground sustainability: a review and synthesis. *Forest Ecology and Management, 122*, 51-71.

Nettleton, I. M., Martin, S., Hencher, S., & Moore, R. (2005). Debris Flow Types and Mechanisms. In *Scottish Road Network Landslide Study* Scotland.

Nsengiyumva, J. B., Luo, G., Nahayo, L., Huang, X., & Cai, P. (2018). Landslide Susceptibility Assessment Using Spatial Multi-Criteria Evaluation Model in Rwanda. *Int J Environ Res Public Health, 15*(2). doi:10.3390/ijerph15020243

Oakley, N. S. (2018). *Characteristics of extreme precipitation in southern California and applications to post-fire debris flows, shallow landslides, and water resources.* (PhD). University of Nevada, Reno,

Oakley, N. S., & Ralph, F. M. (2018). *Meterological Conditions Associated with the Deadly 9 January 2018 Debris Flow on the Thomas Fire Burn Area Impacing Montecito, CA: A Preliminary Analysis*. Retrieved from

Openshaw, S. (1989). Learning to live with errors in spatial databases. In M. Goodchild & S. Gopal (Eds.), *Accuracy of Spatial Databases* (pp. 263-276). Philadelphia: Taylor and Francis.

Papathoma-Kohle, M., Schlogl, M., & Fuchs, S. (2019). Vulnerability indicators for natural hazards: an innovative selection and weighting approach. *Sci Rep, 9*(1), 15026. doi:10.1038/s41598-019-50257-2

Paranunzio, R., Chiarle, M., Laio, F., Nigrelli, G., Turconi, L., & Luino, F. (2018). New insights in the relation between climate and slope failures at high-elevation sites. *Theoretical and Applied Climatology, 137*(3-4), 1765-1784. doi:10.1007/s00704-018-2673-4

Parise, M., & Cannon, S. H. (2011). Wildfire impacts on the processes that generate debris flows in burned watersheds. *Natural Hazards, 61*(1), 217-227. doi:10.1007/s11069-011-9769-9

Park, N.-W. (2014). Using maximum entropy modeling for landslide susceptibility mapping with multiple geoenvironmental data sets. *Environmental Earth Sciences, 73*(3), 937-949. doi:10.1007/s12665-014-3442-z

Pelletier, J. D., Broxton, P. D., Hazenberg, P., Zeng, X., Troch, P. A., Niu, G. Y., . . . Gochis, D. (2016). A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling. *Journal of Advances in Modeling Earth Systems, 8*(1), 41-65. doi:10.1002/2015ms000526

Phillips, J. D. (2010). Amplifiers, filters and geomorphic responses to climate change in Kentucky rivers. *Climate Change, 103*, 571-595. doi:10.1007/s10584-009-0775-z

Phillips, S., & Dudik, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography, 31*, 161-175. doi:10.1111/j.2007.0906-7590.05203.x

Phillips, S., Dudik, M., & Schapire, R. E. (2021). Maxent software for modeling species niches and distributions (Version 3.4.1). Retrieved from http://biodiversityinformatics.amnh.org/open_source/maxent/

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. *Earth-Science Reviews, 180*, 60-91. doi:10.1016/j.earscirev.2018.03.001

RGI Consortium. (2017). *Randolph Glacier Inventory - A Dataset of Global Glacier Outlines: Version 6.0*. Retrieved from Boulder, Colorado, USA:

Ritchie, H., & Roser, M. (Producer). (2018). Urbanization. Retrieved from https://ourworldindata.org/urbanization

Rossi, M., Guzzetti, F., Reichenbach, P., Mondini, A. C., & Peruccacci, S. (2010). Optimal landslide susceptibility zonation based on multiple forecasts. *Geomorphology, 114*(3), 129-142. doi:10.1016/j.geomorph.2009.06.020

Samia, J., Temme, A., Bregt, A., Wallinga, J., Guzzetti, F., Ardizzone, F., & Rossi, M. (2016). Do landslides follow landslides? Insights in path dependency from a multi-temporal landslide inventory. *Landslides, 14*(2), 547-558. doi:10.1007/s10346-016-0739-x

Sayre, R., Dangermond, J., Frye, C., & Vaughan, R. (Cartographer). (2014). A New Map of Global Ecological Land Units - An Ecophysiographic Stratification Approach

Sayre, R., Frye, C., Karagulle, D., Krauer, J., Breyer, S., Aniello, P., . . . Cress, J. (2018). A New High-Resolution Map of World Mountains and an Online Tool for Visualizing and Comparing

Characterizations of Global Mountain Distributions. *Mountain Research and Development, 38*(3), 240-249. doi:10.1659/mrd-journal-d-17-00107.1

Schippa, L. (2018). The Effects of Sediment Size and Concentration on the Rheological Behavior of Debris Flows. In G. K. a. A. C. Mitropoulos (Ed.), *Granularity in Materials Science*: IntechOpen.

Shangguan, W., Hengl, T., Mendes de Jesus, J., Yuan, H., & Dai, Y. (2017). Mapping the global depth to bedrock for land surface modeling. *Journal of Advances in Modeling Earth Systems, 9*(1), 65-88. doi:10.1002/2016ms000686

Shi, W. (2010). *Principles of Modeling Uncertainties in Spatial Data and Spatial Analyses*. Florida: CRC Press.

Skilodimou, H., Bathrellos, G., Koskeridou, E., Soukis, K., & Rozos, D. (2018). Physical and Anthropogenic Factors Related to Landslide Activity in the Northern Peloponnese, Greece. *Land, 7*(3). doi:10.3390/land7030085

Soma, A. (2018). Landslide susceptibility Map using Certainty Factor for Mitigation in Mountainous Area of Ujung-Loe Watershed South Sulawesi Indonesia.

Stanley, T., & Kirschbaum, D. B. (2017). A heuristic approach to global landslide susceptibility mapping. *Natural Hazards, 87*(1), 145-164. doi:10.1007/s11069-017-2757-y

Stoffel, M., Mendlik, T., Schneuwly-Bollschweiler, M., & Gobiet, A. (2013). Possible impacts of climate change on debris-flow activity in the Swiss Alps. *Climatic Change, 122*(1-2), 141-155. doi:10.1007/s10584-013-0993-z

Stoffel, M., Tiranti, D., & Huggel, C. (2014). Climate change impacts on mass movements--case studies from the European Alps. *Sci Total Environ, 493*, 1255-1266. doi:10.1016/j.scitotenv.2014.02.102

Styron, R., & Pagani, M. (2020). The GEM Global Active Faults Database. *Earthquake Spectra, 36*(1_suppl), 160-180. doi:DOI: 10.1177/8755293020944182

Takahashi, T. (2014). *Debris Flow: Mechanics, Prediction and Countermeasures* (2nd ed.). London, UK: Taylor & Francis Group.

Trabucco, A., & Zomer, R. (2018). *Global Aridity Index and Potential Evapo-Transpiration (ETO) Climate Database v2.* . Retrieved from: https://cgiarcsi.community

United Nations Office of Disaster Risk Reduction. (2019). *Global Assessment Report on Disaster Risk Reduction*. Retrieved from Geneva, Switzerland:

Van Den Eeckhaut, M., & Hervás, J. (2012). *Landslidie inventories in Europe and policy recommendations for otheir interoperability and harmonization* (EUR 25666 EN). Retrieved from

Van Westen, C. J. (2016). National Scale Landslide Susceptibility Assessment for Saint Lucia. In *Carribean Handbook onRisk Information Management*.

van Westen, C. J., van Asch, T. W. J., & Soeters, R. (2005). Landslide hazard and risk zonation—why is it still so difficult? *Bulletin of Engineering Geology and the Environment, 65*(2), 167-184. doi:10.1007/s10064-005-0023-0

Varnes, D. J. (1978). *Slope movement types and processes* (Special Report 176). Retrieved from Washington , D.C. :

Varnes, D. J., & IAEG Commission on Landslides. (1984). Landslide hazard zonation: a review of principles and practice. In (pp. 63). Paris, France: UNESCO.

Veregin, H. (1989). Error modeling for the map overlay operation. In M. Goodchild & S. Gopal (Eds.), *The Accuracy of Spatial Databases* (pp. 3-18). Pennsylvania: Taylor & Francis.

Wechsler, S. P. (1999). *Digital Elevation Model (DEM) Uncertainty: Evaluation and Effect on Topographic Parameters*. Paper presented at the ESRI User Conference, San Diego, CA USA.

Whipple, K. X., & Dunne, T. (1992). The influence of debris-flow rheology on fan morphology, Owens Valley, California. *Geological Society of America Bulletin, 104*, 887-900.

Wilford, D. J., Sakals, M. E., Innes, J. L., Sidle, R. C., & Bergerud, W. A. (2004). Recognition of debris flow, debris flood and flood hazard through watershed morphometrics. *Landslides, 1*(1), 61-66. doi:10.1007/s10346-003-0002-0

World Slope GMTED. (2020). Retrieved from: https://www.arcgis.com/home/item.html?id=1f34944490cd43c59d379773d4b7f052

Yang, D., Han, Z., & Li, Y. (2017). Spatial Distribution Features of Debris Flows at Active Fault Zone Along Ya-Lu Highway, China. *The Open Civil Engineering Journal, 11*(1), 563-571. doi:10.2174/1874149501711010563

Yuan, S., Huang, G., Xiong, H., Gong, Q., Wang, J., & Chen, J. (2017). Maximum Entropy-Based Model of High-Threat Landslide Disaster Distribution in Zhaoqing, China. *Journal of Risk Analysis and Crisis Response, 7*(3), 108-126.

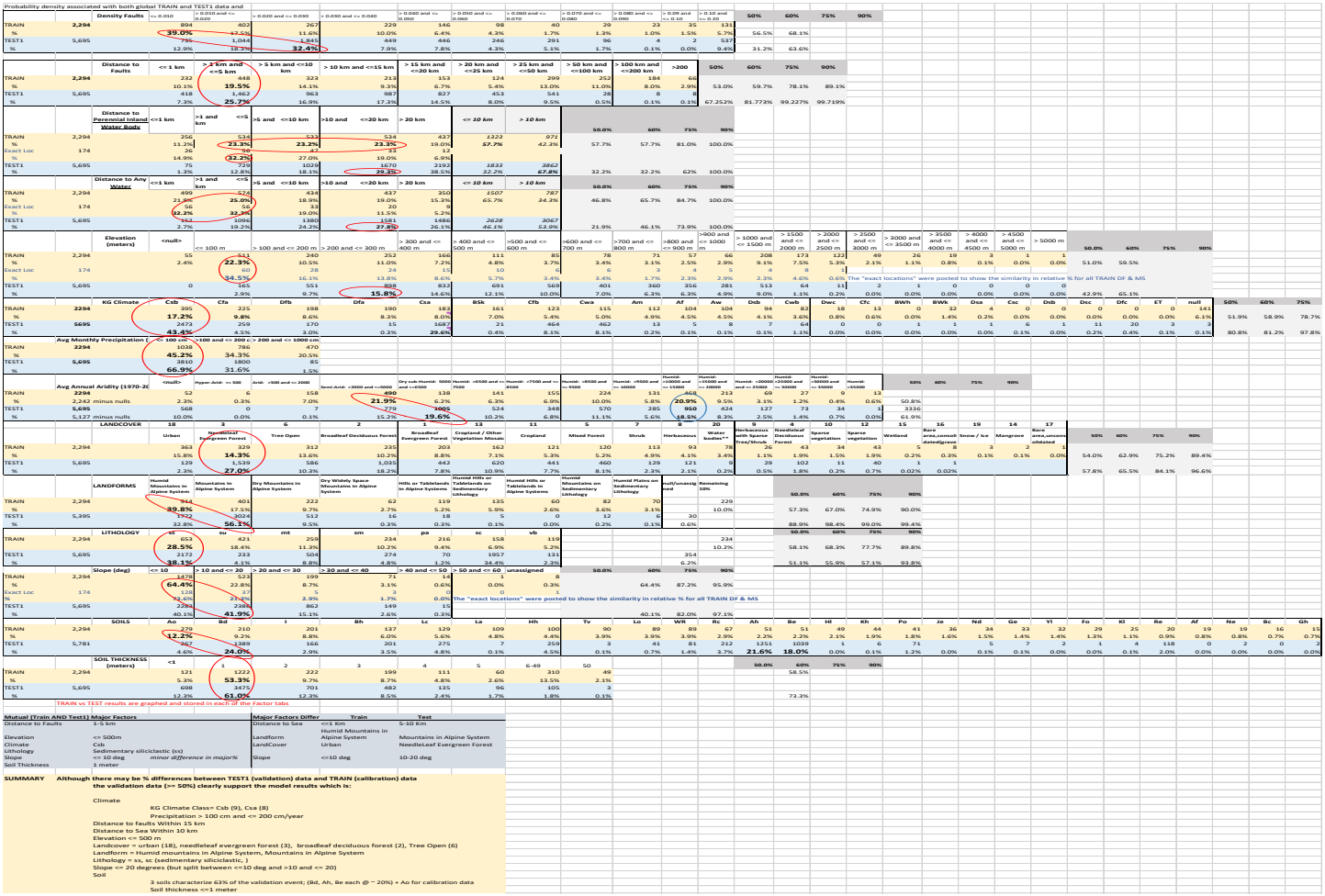
Zeizel, A. J. (1988). *Colorado Landslide Hazard Mitigation Plan*. Colorado: Colorado Geological Survey

Zezere, J. L., Pereira, S., Melo, R., Oliveira, S. C., & Garcia, R. A. C. (2017). Mapping landslide susceptibility using data-driven methods. *Sci Total Environ, 589*, 250-267. doi:10.1016/j.scitotenv.2017.02.188

Zufle, A., Trajcevski, G., Pfoser, D., Renz, M., Rice, M. T., Leslie, T., . . . Emrich, T. (2017). *Handling Uncertainty in Geo-Spatial Data*. Paper presented at the 2017 IEEE 33rd International Conference on Data Engineering (ICDE).

APPENDIX A TRAIN and TEST1 Environmental Conditions

A.1 TRAIN and TEST1 Comparisons - Overview



Probability density associated with both global TRAIN and TEST1 data and

		Density Faults	<= 0.010	> 0.010 and <= 0.020	> 0.020 and <= 0.030	> 0.030 and <= 0.040	> 0.040 and <= 0.050	> 0.050 and <= 0.060	> 0.060 and <= 0.070	> 0.070 and <= 0.080	> 0.080 and <= 0.090	> 0.09 and <= 0.10	> 0.10 and <= 0.20	50%	60%	75%	90%
TRAIN	2,294		894	402	267	229	146	98	40	29	23	35	141				
%			39.0%	17.5%	11.6%	10.0%	6.4%	4.3%	1.7%	1.3%	1.0%	1.5%	5.7%	56.5%	68.1%		
TEST1	5,695		736	1,044	1,845	449	446	246	291	96	4	2	537				
%			12.9%	18.3%	32.4%	7.9%	7.8%	4.3%	5.1%	1.7%	0.1%	0.0%	9.4%	31.2%	63.6%		

		Distance to Faults	<= 1 km	> 1 km and <=5 km	> 5 km and <=10 km	> 10 km and <=15 km	> 15 km and <=20 km	> 20 km and <=25 km	> 25 km and <=50 km	> 50 km and <=100 km	> 100 km and <=200 km	>200	50%	60%	75%	90%
TRAIN	2,294		232	448	323	213	153	124	299	252	184	66					
%			10.1%	19.5%	14.1%	9.3%	6.7%	5.4%	13.0%	11.0%	8.0%	2.9%	53.0%	59.7%	78.1%	89.1%	
TEST1	5,695		418	1,462	963	987	827	453	541	28	8	8					
%			7.3%	25.7%	16.9%	17.3%	14.5%	8.0%	9.5%	0.5%	0.1%	0.1%	67.252%	81.773%	99.227%	99.719%	

		Distance to Perennial Inland Water Body	<=1 km	>1 and <=5 km	>5 and <=10 km	>10 and <=20 km	> 20 km	<= 10 km	> 10 km	50.0%	60%	75%	90%
TRAIN	2,294		256	534	533	534	437	1223	971				
%			11.2%	23.3%	23.2%	23.3%	19.0%	27.7%	42.3%	57.7%	57.7%	81.0%	100.0%
Exact Loc	174		26	94	47	23	12						
%			14.9%	32.2%	27.0%	19.0%	6.9%						
TEST1	5,695		75	729	1029	1670	2192	1833	3862				
%			1.3%	12.8%	18.1%	29.3%	38.5%	32.2%	67.8%	32.2%	32.2%	62%	100.0%

		Distance to Any Water	<=1 km	>1 and <=5 km	>5 and <=10 km	>10 and <=20 km	> 20 km	<= 10 km	> 10 km	50.0%	60%	75%	90%
TRAIN	2,294		499	574	434	437	350	1507	787				
%			21.8%	25.0%	18.9%	19.0%	15.3%	65.7%	34.3%	46.8%	65.7%	84.7%	100.0%
Exact Loc	174		56	56	33	20	9						
%			32.2%	32.2%	19.0%	11.5%	5.2%						
TEST1	5,695		562	1096	1380	1581	1486	2628	3067				
%			2.7%	19.2%	24.2%	27.8%	26.1%	46.1%	53.9%	21.9%	46.1%	73.9%	100.0%

		Elevation (meters)	<null>	<= 100 m	> 100 and <= 200 m	> 200 and <= 300 m	> 300 and <= 400 m	> 400 and <= 500 m	>500 and <= 600 m	>600 and <= 700 m	>700 and <= 800 m	>800 and <= 900 m	>900 and <= 1000 m	> 1000 and <= 1500 m	> 1500 and <= 2000 m	> 2000 and <= 2500 m	> 2500 and <= 3000 m	> 3000 and <= 3500 m	> 3500 and <= 4000 m	> 4000 and <= 4500 m	> 4500 and <= 5000 m	> 5000 m	50.0%	60%	75%	90%
TRAIN	2,294		55	511	240	252	166	111	85	78	71	57	66	208	173	122	49	26	19	3	1	1				
%			2.4%	22.3%	10.5%	11.0%	7.2%	4.8%	3.7%	3.4%	3.1%	2.5%	2.9%	9.1%	7.5%	5.3%	2.1%	1.1%	0.8%	0.1%	0.0%	0.0%	51.0%	59.5%		
Exact Loc	174			60	28	24	15	10	6	6	3	4	5	4	8	1										
%				34.5%	16.1%	13.8%	8.6%	5.7%	3.4%	3.4%	1.7%	2.3%	2.9%	2.3%	4.6%	0.6%	The "exact locations" were posted to show the similarity in relative % for all TRAIN DF & MS								
TEST1	5,695		0	155	511	808	833	691	569	401	360	356	281	513	64	11	2	1	0	0	0	0				
%				2.9%	9.7%	15.8%	14.6%	12.1%	10.0%	7.0%	6.3%	6.3%	4.9%	9.0%	1.1%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	42.9%	65.1%		

		KG Climate	Csb	Cfa	Dfb	Dfa	Csa	Bsk	Cfb	Cwa	Am	Af	Aw	Dsb	Cwb	Dwc	Cfc	BWh	BWk	Dsa	Csc	Dsb	Dfc	ET	0	null	50%	60%	75%
TRAIN	2294		395	225	198	190	183	161	123	115	112	104	104	94	82	18	13	0	32	4	0	0	0	0	141			
%			17.2%	9.8%	8.6%	8.3%	8.0%	7.0%	5.4%	5.0%	4.9%	4.5%	4.5%	4.1%	3.6%	0.8%	0.6%	0.0%	1.4%	0.2%	0.0%	0.0%	0.0%	0.0%	6.1%	51.9%	58.9%	78.7%
TEST1	5695		2473	259	170	15	1687	21	464	462	13	5	8	7	64	0	0	1	1	1	6	1	11	20	3	64	3		
%			43.4%	4.5%	3.0%	0.3%	29.6%	0.4%	8.1%	8.1%	0.2%	0.1%	0.1%	0.1%	1.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.2%	0.4%	0.1%	0.1%	80.8%	81.2%	97.8%

		Avg Monthly Precipitation (	<= 100 cm	>100 and <= 200 c	> 200 and <= 1000 cm																								
TRAIN	2294		1038	786	470																								
%			45.2%	34.3%	20.5%																								
TEST1	5,695		3810	1800	85																								
%			66.9%	31.6%	1.5%																								

		Avg Annual Aridity (1970-2(	<null>	Hyper-Arid: <= 300	Arid: >300 and <= 2000	Semi-Arid: >2000 and <=5000	Dry sub-Humid: 5000 and <=8500	Humid: >8500 and <= 7500	Humid: >7500 and <= 8500	Humid: >8500 and <= 9500	Humid: >9500 and <= 10000	Humid: >10000 and <= 15000	Humid: >15000 and <= 20000	Humid: >20000 and <= 25000	Humid: >25000 and <= 30000	Humid: >30000 and <= 35000	Humid: >35000	50%	60%	75%	90%
TRAIN	2294		52	6	158	490	138	141	155	224	131	468	213	69	27	9	13				
%	2,242 minus nulls		2.3%	0.3%	7.0%	21.9%	6.2%	6.3%	6.9%	10.0%	5.8%	20.9%	9.5%	3.1%	1.2%	0.4%	0.6%	50.8%			
TEST1	5,695		568	0	7	279	1005	524	348	570	285	950	424	127	73	34	1	3336			
%	5,127 minus nulls		10.0%	0.0%	0.1%	15.2%	19.6%	10.2%	6.8%	11.1%	5.6%	18.5%	8.3%	2.5%	1.4%	0.7%	0.0%	61.9%			

		LANDCOVER	18	3	6	2	1	13	11	5	7	8	20	9	4	10	12	15	16	19	14	17	50%	60%	75%	90%
			Urban	Needleleaf Evergreen Forest	Tree Open	Broadleaf Deciduous Forest	Broadleaf Evergreen Forest	Cropland / Other Vegetation Mosaic	Cropland	Mixed Forest	Shrub	Herbaceous	Water bodies**	Herbaceous with Sparse Tree/Shrub	Mixed/deal Deciduous Forest	Sparse vegetation	Sparse vegetation	Wetland	Bare area,consolidated	Snow / Ice	Mangrove	Bare area,unconsolidated				
TRAIN	2,294		363	329	312	235	203	162	121	120	113	93	78	26	43	34	43	5	8	3	2	1				
%			15.8%	14.3%	13.6%	10.2%	8.8%	7.1%	5.3%	5.2%	4.9%	4.1%	3.4%	1.1%	1.9%	1.5%	1.9%	0.2%	0.3%	0.1%	0.1%	0.0%	54.0%	62.9%	75.2%	89.4%
TEST1	5,695		129	1,539	586	1,035	442	620	441	460	129	121	9	29	102	11	40	1	1							
%			2.3%	27.0%	10.3%	18.2%	7.8%	10.9%	7.7%	8.1%	2.3%	2.1%	0.2%	0.5%	1.8%	0.2%	0.7%	0.02%	0.02%				57.8%	65.5%	84.1%	96.6%

		LANDFORMS	Humid Mountains in Alpine System	Mountains in Alpine System	Dry Mountains in Alpine System	Dry Widely Space Mountains in Alpine System	Hills or Tablelands in Alpine Systems	Humid Hills or Tablelands on Sedimentary Lithology	Humid Hills or Tablelands in Alpine Systems	Humid Mountains on Sedimentary Lithology	Humid Plains on Sedimentary Lithology	null/unassigned	Remaining 10%	50.0%	60%	75%	90%
TRAIN	2,294		914	401	222	62	119	135	60	82	70		229				
%			39.8%	17.5%	9.7%	2.7%	5.2%	5.9%	2.6%	3.6%	3.1%		10.0%	57.3%	67.0%	74.9%	90.0%
TEST1	5,395		1922	2024	512	16	18	5	0	12	6	30					
%			32.8%	56.1%	9.5%	0.3%	0.3%	0.1%	0.0%	0.2%	0.1%	0.6%		88.9%	98.4%	99.0%	99.4%

		LITHOLOGY	ss	su	mt	sm	pa	sc	vb	50.0%	60%	75%	90%
TRAIN	2,294		653	421	259	234	216	158	119				
%			28.5%	18.4%	11.3%	10.2%	9.4%	6.9%	5.2%				
TEST1	5,695		2172	233	504	274	70	1957	131				
%			38.1%	4.1%	8.8%	4.8%	1.2%	34.4%	2.3%				

		Slope (deg)	<= 10	> 10 and <= 20	> 20 and <= 30	> 30 and <= 40	> 40 and <= 50	> 50 and <= 60	unassigned	50.0%	60%	75%	90%
TRAIN	2,294		1478	523	199	71	14	1	8				
%			64.4%	22.8%	8.7%	3.1%	0.6%	0.0%	0.3%	64.4%	87.2%	95.9%	
Exact Loc	174		126	37	5	3	0	0	1				
%			72.4%	21.3%	2.9%	1.7%	0.0%	The "exact locations" were posted to show the similarity in relative % for all TRAIN DF & MS					
TEST1	5,695		2284	2386	862	149	13						
%			40.1%	41.9%	15.1%	2.6%	0.3%			40.1%	82.0%	97.1%	

		SOILS	Ao	Bd	I	Bh	Lc	La	Hh	Tv	Lo	WR	Rc	Ah	Be	Hl	Kh	Po	Je	Nd	Ge	Yl	Fo	Kl	Re	Af	Ne	Bc	Gh
TRAIN	2,294		279	210	201	137	129	109	100	90	89	89	67	51	51	49	44	41	36	34	33	32	29	25	20	19	19	16	15
%			12.2%	9.2%	8.8%	6.0%	5.6%	4.8%	4.4%	3.9%	3.9%	3.9%	2.9%	2.2%	2.2%	2.1%	1.9%	1.8%	1.6%	1.5%	1.4%	1.4%	1.3%	1.1%	0.9%	0.8%	0.8%	0.7%	0.7%
TEST1	5,781		967	1389	166	201	275	7	259	3	41	81	212	1251	1039	1	6	71		5	7	2	1	4	118	0	2	0	2
%			4.6%	24.0%	2.9%	3.5%	4.8%	0.1%	4.5%	0.1%	0.7%	1.4%	3.7%	21.6%	18.0%	0.0%	0.1%	1.2%	0.0%	0.1%	0.1%	0.0%	0.0%	0.1%	2.0%	0.0%	0.0%	0.0%	0.0%

		SOIL THICKNESS (meters)	<1	1	2	3	4	5	6-49	50	50.0%	60%	75%	90%
TRAIN	2,294		121	1222	222	199	111	60	310	49				
%			5.3%	53.3%	9.7%	8.7%	4.8%	2.6%	13.5%	2.1%		58.5%		
TEST1	5,695		698	3475	701	482	135	96	105	2				
%			12.3%	61.0%	12.3%	8.5%	2.4%	1.7%	1.8%	0.1%		73.3%		

TRAIN vs TEST results are graphed and stored in each of the Factor tabs

Mutual (Train AND Test1)	Major Factors
Distance to Faults	1-5 km
Elevation	<= 500m
Climate	Csb
Lithology	Sedimentary siliciclastic (ss)
Slope	<= 10 deg
Soil Thickness	1 meter

minor difference in major%

Major Factors Differ	Train	Test
Distance to Sea	<=1 Km	5-10 Km
Landform	Humid Mountains in Alpine System	Mountains in Alpine System
LandCover	Urban	NeedleLeaf Evergreen Forest
Slope	<=10 deg	10-20 deg

SUMMARY **Although there may be % differences between TEST1 (validation) data and TRAIN (calibration) data the validation data (>= 50%) clearly support the model results which is:**

Climate
KG Climate Class= Csb (9), Csa (8)
Precipitation > 100 cm and <= 200 cm/year
Distance to faults Within 15 km
Distance to Sea Within 10 km
Elevation <= 500 m
Landcover = urban (18), needleleaf evergreen forest (3), broadleaf deciduous forest (2), Tree Open (6)
Landform = Humid mountains in Alpine System, Mountains in Alpine System
Lithology = ss, sc (sedimentary siliciclastic,)
Slope <= 20 degrees (but split between <=10 deg and >10 and <= 20)
Soil
3 soils characterize 63% of the validation event; (Bd, Ah, Be each @ ~ 20%) + Ao for calibration data
Soil thickness <=1 meter

A.1.1 TRAIN and TEST1 Tectonic Comparisons

Slip Type	"TRAIN" FREQUENCY	"TRAIN" Probability Density	"TEST1" FREQUENCY	"TEST1" Probability Density
<null>	25	1.1%	2	0.0%
Anticline	36	1.6%	5	0.1%
Dextral	280	12.2%	1295	22.7%
Dextral-Normal	9	0.4%	405	7.1%
Dextral-Reverse	13	0.6%	428	7.5%
Dextral_Transform	10	0.4%	1	0.0%
Normal	579	25.2%	990	17.4%
Normal-Dextral	3	0.1%	0	0.0%
Normal-Sinistral	1	0.0%	0	0.0%
Reverse	**1035**	**45.1%**	**1700**	29.9%
Reverse-Dextral	13	0.6%	646	11.3%
Reverse-Sinistral	3	0.1%	0	0.0%
Reverse-Strike-Slip	17	0.7%	1	0.0%
Sinistral	140	6.1%	132	2.3%
Sinistral-Normal	7	0.3%	3	0.1%
Sinistral-Reverse	41	1.8%	80	1.4%
Sinistral_Transform	26	1.1%	0	0.0%
Spreading_Ridge	14	0.6%	4	0.1%
Strike-Slip	10	0.4%	0	0.0%
Subduction_Thrust	32	1.4%	3	0.1%
TOTAL	**2294**	**100.0%**	**5695**	**100.0%**

Frequency Analysis TEST vs TRAIN - DISTANCE to TECTONICS FACTOR

DISTANCE to "ACTIVE FAULTS"	TRAIN DATA		TEST1 DATA	
	Probability Density	Frequency	Probability Density	Frequency
<= 1 km	7.8%	179	1.5%	88
> 1 km and <=5 km	**13.0%**	**298**	14.2%	810
> 5 km and <=10 km	8.3%	191	15.2%	867
> 10 km and <=15 km	7.4%	170	14.6%	833
> 15 km and <=20 km	4.2%	96	11.4%	649
> 20 km and <=25 km	2.7%	62	7.5%	428
> 25 km and <=50 km	9.5%	217	**21.6%**	**1232**
Top 50% prediction	**52.9%**	**1213**	**86.2%**	**4907**
> 50 km and <=100 km	9.7%	223	12.5%	712
> 100 km and <=200 km	7.1%	163	0.2%	14
> 200 km	30.3%	695	1.1%	62
90% Totals	**100.0%**	**2294**	**100.0%**	**5695**

NEAREST...	"TRAIN" Data	"TRAIN"	"TEST1" Data	"TEST1"	
PLATE BOUNDARY (STEP) TYPE	**FREQUENCY**	Prob Density	FREQUENCY	Prob Density	
CCB	521	22.7%	243	4.3%	Continental Convergent Boundary
CRB	160	7.0%	544	9.6%	Continental Rift Boundary
CTF	569	24.8%	1715	30.1%	Continental Transform Fault
OCB	17	0.7%	1	0.0%	Oceanic Convergent Boundary
OSR	58	2.5%	10	0.2%	Oceanic Spreading Ridge
OTF	369	16.1%	33	0.6%	Oceanic Transform Fault
SUB	600	26.2%	3149	55.3%	Subduction (see comment)
TOTAL	**2294**	**100.0%**	**5695**	**100.0%**	

A.1.2 TRAIN and TEST1 Landcover/Land Use Comparisons

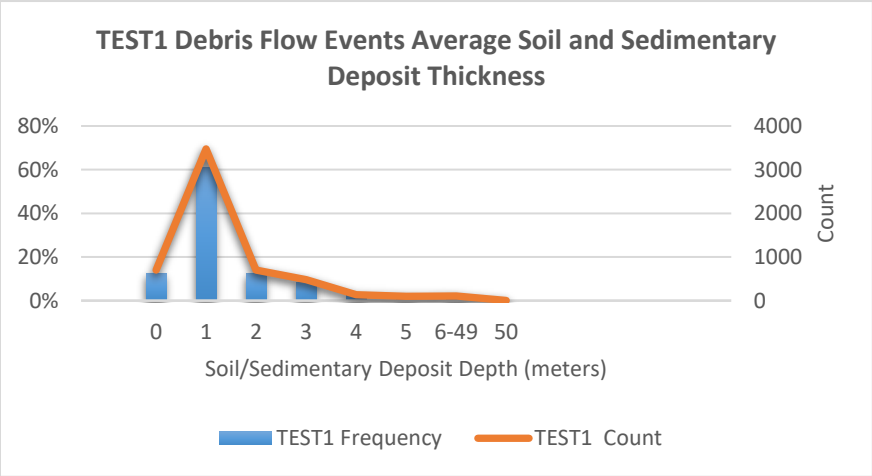
Landcover Code	LANDCOVER	TRAIN DATA			TEST1 DATA		
		DF & MS			Debris Flow		
GLCNMO Lan	LandCover Description	Freq	TRAIN Count		TEST1 Count	Frequency	
18	Urban	15.8%	363		129	2.3%	
3	Needleleaf Evergreen Forest	14.3%	329		1539	**27.0%**	
6	Tree Open	13.6%	312		586	10.3%	
2	Broadleaf Deciduous Forest	10.2%	235	53.9%	1035	**18.2%**	57.8%
1	Broadleaf Evergreen Forest	8.8%	203		442	7.8%	
13	Cropland / Other Vegetation Mosaic	7.1%	162		620	**10.9%**	
11	Cropland	5.3%	121		441	7.7%	
5	Mixed Forest	5.2%	120		460	8.1%	
7	Shrub	4.9%	113		129	2.3%	
8	Herbaceous	4.1%	93	89.3%	121	2.1%	96.7%
9	Herbaceous with Sparse Tree/Shrub	1.1%	26		29	0.5%	
4	Needleleaf Deciduous Forest	1.9%	43		102	1.8%	
10	Sparse vegetation	1.5%	34		11	0.2%	
12	Sparse vegetation	1.9%	43		40	0.7%	
15	Wetland	0.2%	5		1	0.0%	
16	Bare area,consolidated(gravel,rock)	0.3%	8		1	0.0%	
19	Snow / Ice	0.1%	3		0	0.0%	
14	Mangrove	0.1%	2		0	0.0%	
17	Bare area,unconsolidated (sand)	0.0%	1		0	0.0%	
20	**Water bodies****	3.4%	78		9	0.2%	
	90% Totals	**100.0%**	**2294**		**5695**	**100%**	
	Number of Training Data Events	2294					
	Number of TEST1 Data Events	5695					

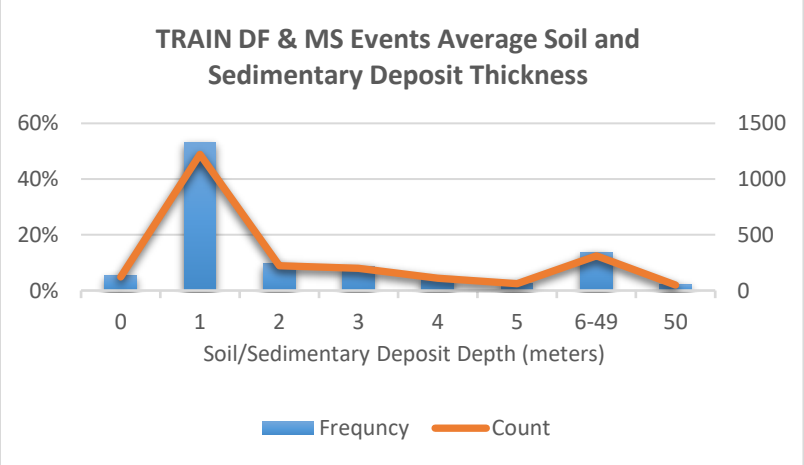
A.1.3 TRAIN and TEST1 Soil Type Comparisons

		TRAIN DATA		TEST1	
Soil Code	**SOILS**	Train		Test1	
		Freq	**Count**		
Ao & Ah	**Acrisols**	**14.4%**	**330**	**26.7%**	1518
Bd & Bh &	**Cambisols**	**17.3%**	**398**	**46.2%**	2629
I	Lithosol	**8.8%**	**201**	**2.9%**	166
Lc & La	Luvisol	**10.4%**	**238**	**5.0%**	282
Hh	Phaeozem	**4.4%**	100	4.5%	259
Tv		3.9%	90	0.1%	3
Lo		3.9%	89	0.7%	41
WR		3.9%	89	1.4%	81
Rc & Re	Regosols	3.8%	87	5.8%	330
Hl		2.1%	49	0.0%	1
Kh		1.9%	44	0.1%	6
<null)		1.9%	43	0.1%	4
Po	703	1.8%	41	1.2%	71
Je		1.6%	36	0.0%	
Nd		1.5%	34	0.1%	5
Ge		1.4%	33	0.1%	7
Yl		1.4%	32	0.0%	2
Fo		1.3%	29	0.0%	1
Kl		1.1%	25	0.1%	4
Af		0.8%	19	0.0%	
Ne		0.8%	19	0.0%	2
Bc		0.7%	16	0.0%	
Gh		0.7%	15	0.0%	2

A.1.4 TRAIN and TEST1 Soil Thickness Comparisons

Thickness (m)	TEST1 Frequency	TEST1 Count	TRAIN Frequency	TRAIN Count
0	12.3%	698	5.3%	121
1	61.0%	3475	53.3%	1222
2	12.3%	701	9.7%	222
3	8.5%	482	8.7%	199
4	2.4%	135	4.8%	111
5	1.7%	96	2.6%	60
6-49	1.8%	105	13.5%	310
50	0.05%	3	2.1%	49



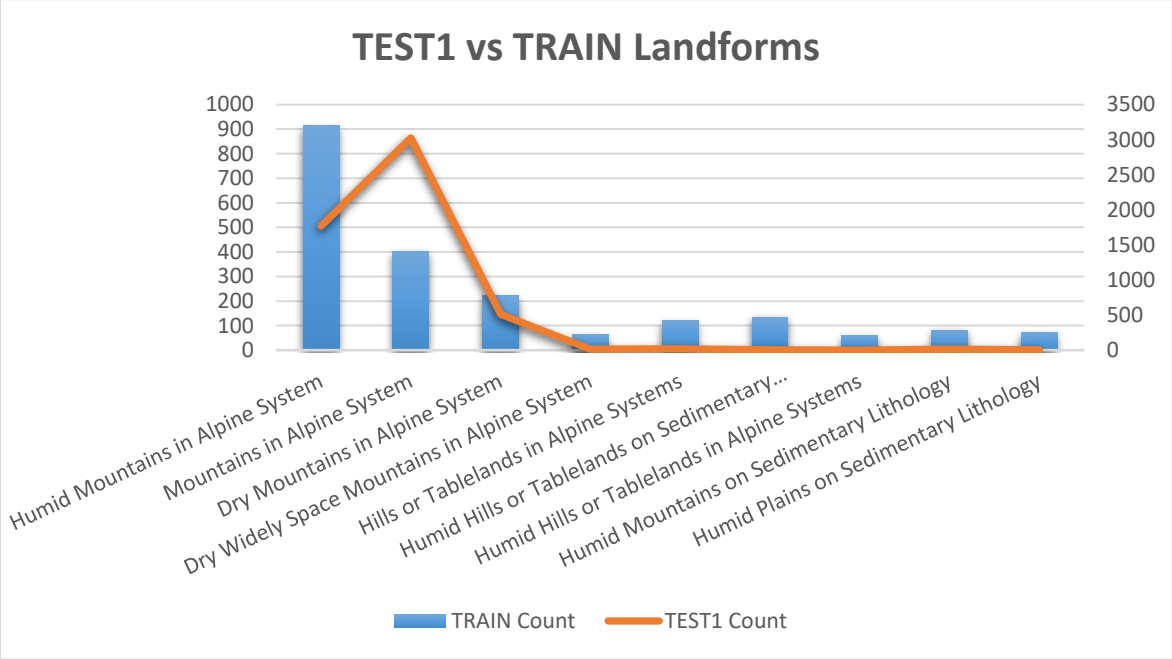


A.1.5 TRAIN and TEST1 Slope Comparisons

	TRAIN DATA		TRAIN DATA 'EXACT' LOC KNOWN		TEST1 DATA	
Slope	**Freq**	**Count**	**Count**	**Freq**	**Freq**	**Count**
<= 10 deg	**64.4%**	1478	128	73.6%	40.1%	2283
> 10 and <= 20 deg	22.8%	523	37	21.3%	**41.9%**	2386
> 20 and <= 30 deg	8.7%	199	5	2.9%	15.1%	862
> 30 and <= 40 deg	3.1%	71	3	1.7%	2.6%	149
> 40 and <= 50 deg	0.6%	14	0	0.0%	0.3%	15
> 50 and <= 60 deg	0.0%	1	0	0.0%		
unassigned	0.3%	8	1	0.6%		
TOTALS	100.0%	2294	174	100%	100.0%	5695

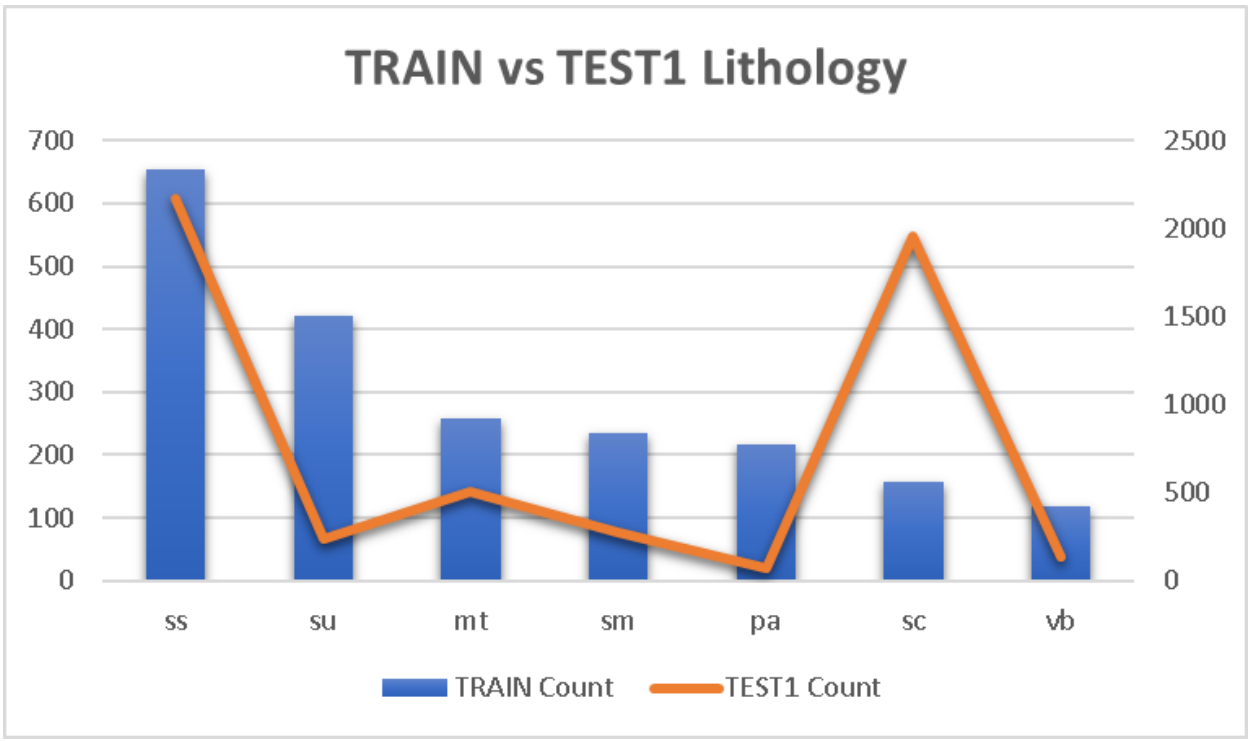
A.1.6 TRAIN and TEST1 Landform Comparisons

Frequency Analysis TEST vs TRAIN - LANDFORM FACTOR								
Sq km earth:	510,100,000							
Sq km land:	148,000,000		TRAIN DATA			TEST1		
LANDFORM (Ecophysical Region - Murphy)	Area Sq km	% Land	Freq	TRAIN Count	Prob. Density (En*Dn)	Freq	TEST1 Count	Prob. Density (Hn*Dn)
Humid Mountains in Alpine System	11,592,742	7.8%	**39.8%**	914	3.1%	32.7%	1772	2.6%
Mountains in Alpine System	20,598,908	13.9%	17.5%	401	2.4%	55.7%	3024	4.4%
Dry Mountains in Alpine System	2,030,136	1.4%	9.7%	222	0.1%	9.4%	512	0.7%
Dry Widely Space Mountains in Alpine System	2,006,612	1.4%	2.7%	62	0.0%	0.3%	16	0.0%
Hills or Tablelands in Alpine Systems	15,173,714	10.3%	5.2%	119	0.5%	0.3%	18	0.0%
Humid Hills or Tablelands on Sedimentary Lithology	3,499,916	2.4%	5.9%	135	0.1%	0.1%	5	0.0%
Humid Hills or Tablelands in Alpine Systems	10,420,469	7.0%	2.6%	60	0.2%	0.0%	0	0.0%
Humid Mountains on Sedimentary Lithology	205,585	0.1%	3.6%	82	0.0%	0.2%	12	0.0%
Humid Plains on Sedimentary Lithology	15,280,805	10.3%	3.1%	70	0.3%	0.1%	6	0.0%



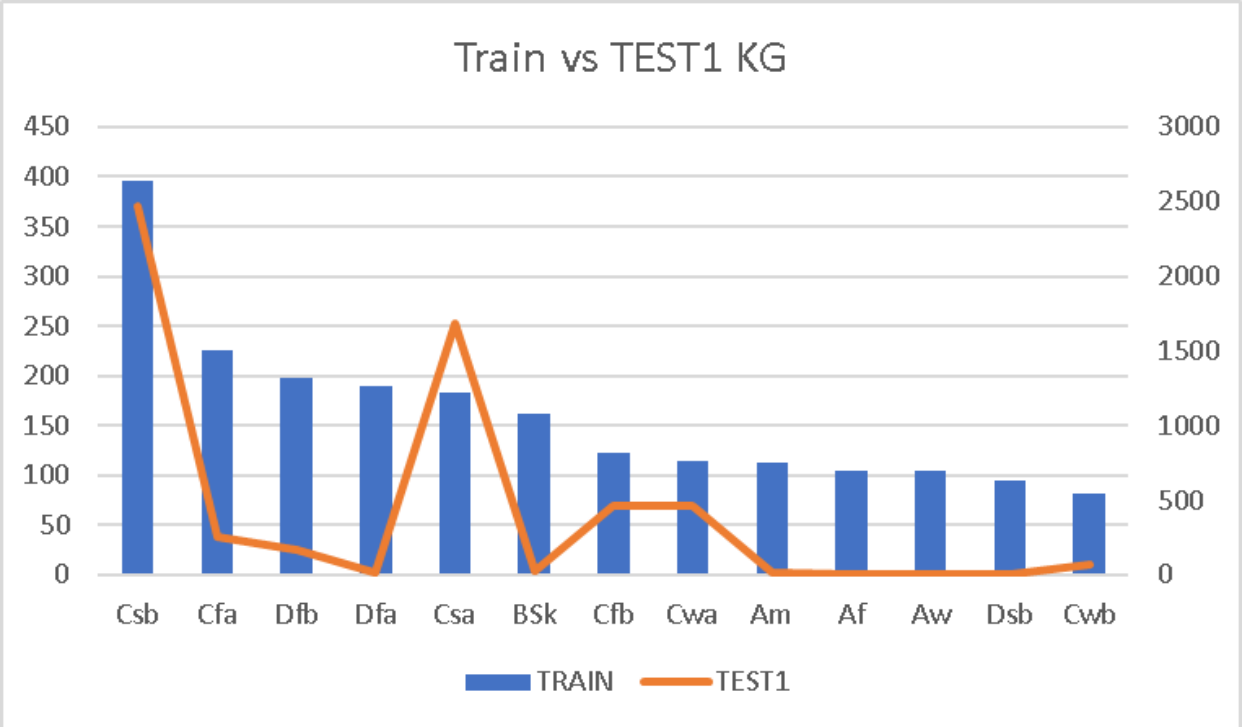
A.1.7 TRAIN and TEST1 Lithology Comparisons

			TRAIN DATA			TEST1 DATA		
LITHOLOGY	Area Sq Km	% Earth Land	Freq	'RAIN Coun	Prob Density (En*Dn)	Freq	'EST1 Coun	Prob Density (Hn*Dn)
ss	24,507,815	16.6%	28.5%	653	4.7%	38.1%	2172	6.3%
su	36,866,217	24.9%	18.4%	421	4.6%	4.1%	233	1.0%
mt	19,570,508	13.2%	11.3%	259	1.5%	8.8%	504	1.2%
sm	22,026,381	14.9%	10.2%	234	1.5%	4.8%	274	0.7%
pa	8,575,653	5.8%	9.4%	216	0.5%	1.2%	70	0.1%
sc	11,765,424	7.9%	6.9%	158	0.5%	34.4%	1957	2.7%
vb	5,244,151	3.5%	5.2%	119	0.2%	2.3%	131	0.1%



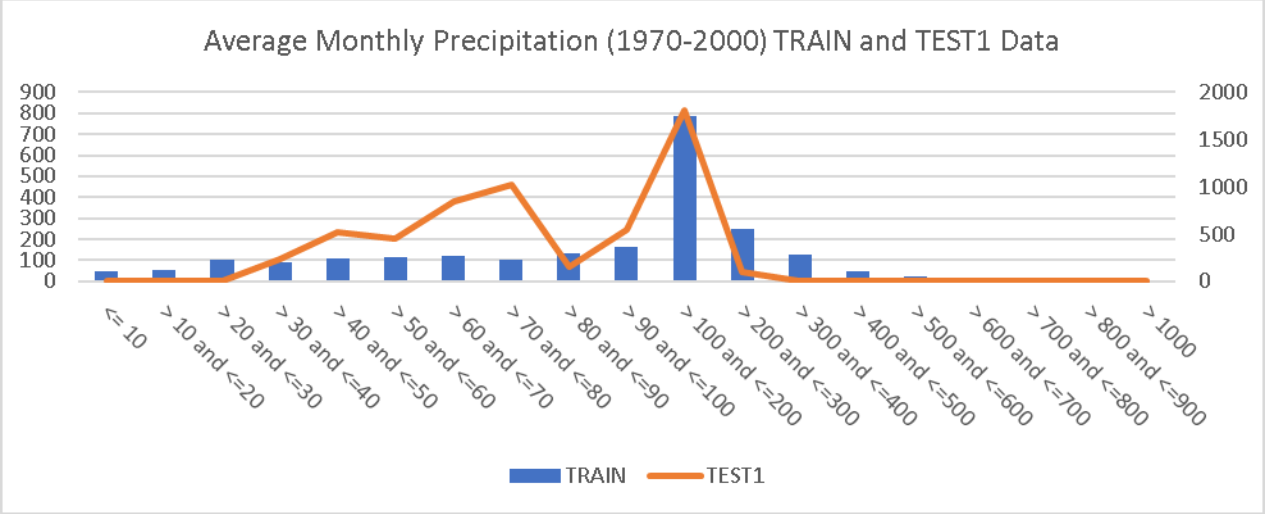
A.1.8 TRAIN and TEST1 Köppen-Geiger (KG) Climate Classification Comparison

			Sq Km	% World Land	TRAIN DATA		Prob. Density (En*Fn)	TEST1 DATA		Prob. Density (Jn*En)
	KG Code	KG CLASS			Freq	Count		Freque	Count	
	9	Csb	854,370	0.6%	**17.2%**	395	0.1%	**43.4%**	2473	0.3%
Top 50%	14	Cfa	5,490,516	3.7%	9.8%	225	0.1%	4.5%	259	0.2%
Top 50%	26	Dfb	7,406,505	5.0%	8.6%	198	0.0%	3.0%	170	0.1%
Top 50%	25	Dfa	2,011,149	1.4%	8.3%	190	0.0%	0.3%	15	0.0%
	8	Csa	1,422,928	1.0%	8.0%	183	0.0%	29.6%	1687	0.3%
	7	BSk			7.0%	161	0.04%	0.4%	21	
	15	Cfb			5.4%	123	0.03%	8.1%	464	
Top 90%	11	Cwa			5.0%	115	0.03%	8.1%	462	
Top 90%	2	Am			4.9%	112	0.03%	0.2%	13	
Top 90%	1	Af			4.5%	104	0.03%	0.1%	5	
Top 90%	3	Aw			4.5%	104	0.03%	0.1%	8	
	18	Dsb			4.1%	94	0.02%	0.1%	7	
	12	Cwb			3.6%	82	0.02%	1.1%	64	

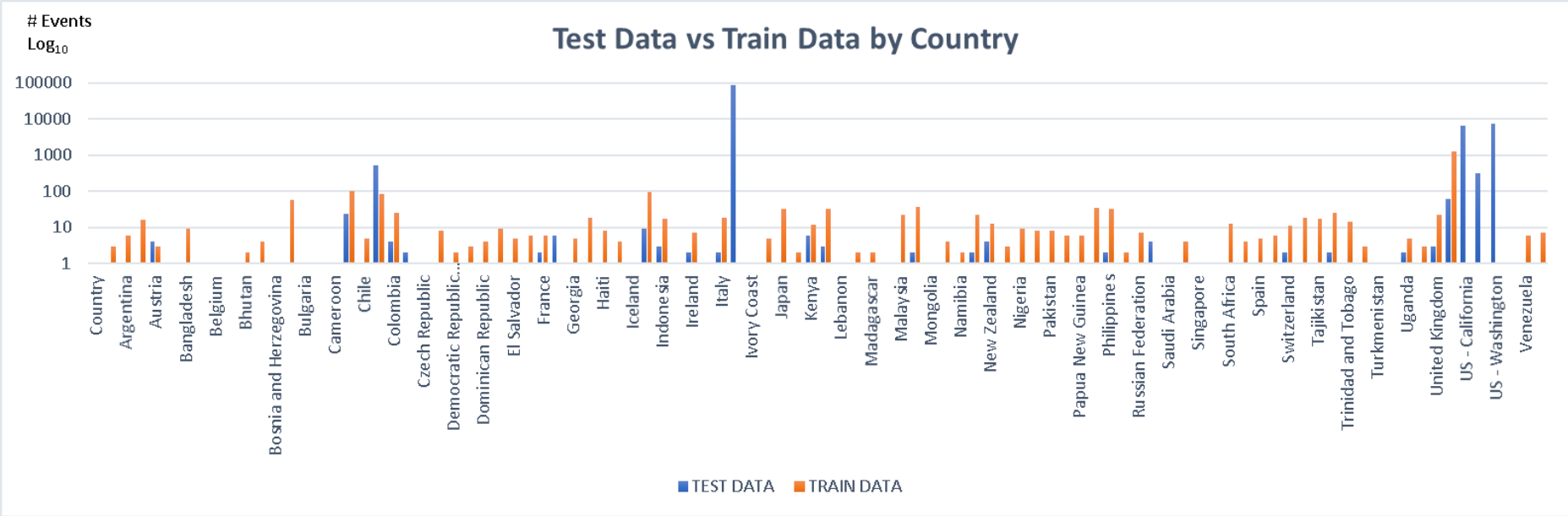


A.1.9 TRAIN and TEST1 Average Monthly Precipitation Comparison

Avg Monthly Precip (mm)	TRAIN	TEST1
<= 10	45	0
> 10 and <=20	54	1
> 20 and <=30	103	11
> 30 and <=40	90	234
> 40 and <=50	109	517
> 50 and <=60	114	458
> 60 and <=70	118	845
> 70 and <=80	105	1020
> 80 and <=90	134	154
> 90 and <=100	166	551
> 100 and <=200	**786**	**1810**
> 200 and <=300	251	94
> 300 and <=400	129	**0**
> 400 and <=500	47	0
> 500 and <=600	21	0
> 600 and <=700	9	0
> 700 and <=800	5	0
> 800 and <=900	4	0
> 1000	4	0
	2294	**5695**



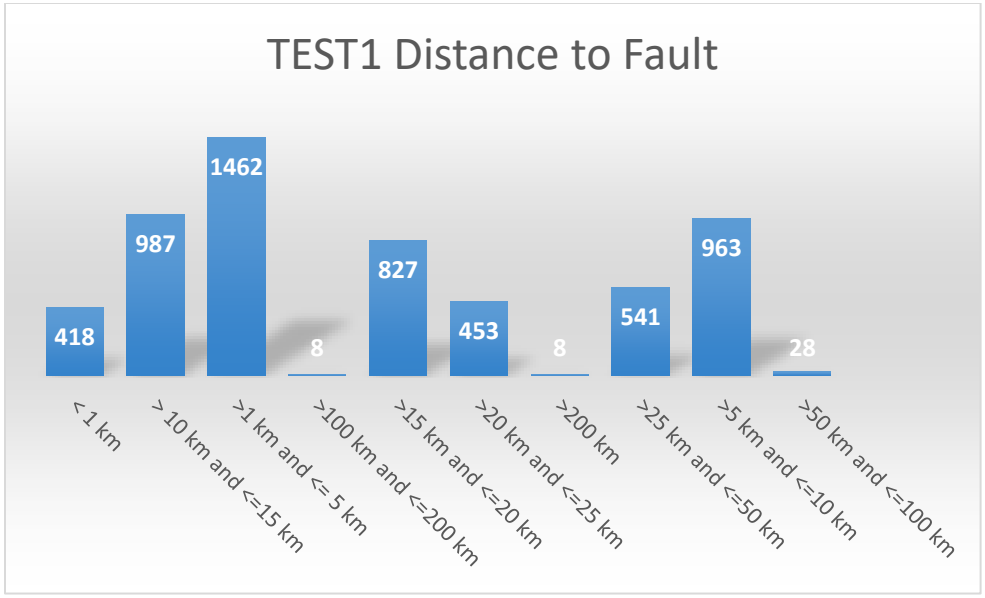
A.1.10 TRAIN and TEST1 Distribution by Country Comparison



A.2 TEST1 Environmental Conditions

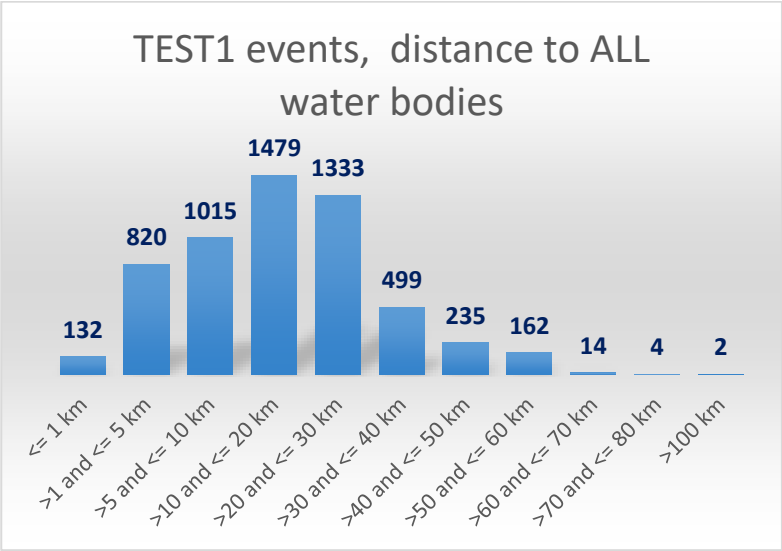
A.2.1 TEST1 Distance to Faults

TEST1 Distance to Fault	Count	Frequency
< 1 km	418	7.3%
> 10 km and <=15 km	987	17.3%
>1 km and <= 5 km	1462	25.7%
>100 km and <=200 km	8	0.1%
>15 km and <=20 km	827	14.5%
>20 km and <=25 km	453	8.0%
>200 km	8	0.1%
>25 km and <=50 km	541	9.5%
>5 km and <=10 km	963	16.9%
>50 km and <=100 km	28	0.5%
TOTAL	**5695**	

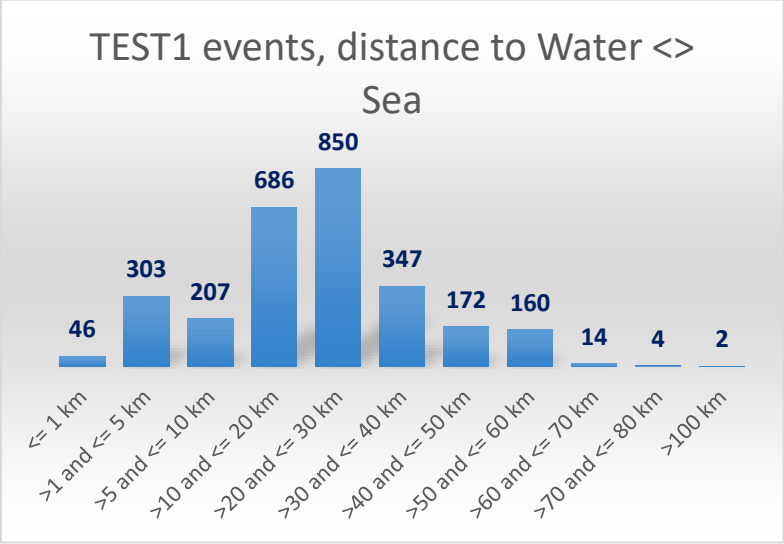

TEST1 Distance to Fault
418
987
1462
8
827
453
8
541
963
28
< 1 km
> 10 km and <=15 km
>1 km and <= 5 km
>100 km and <=200 km
>15 km and <=20 km
>20 km and <=25 km
>200 km
>25 km and <=50 km
>5 km and <=10 km
>50 km and <=100 km

A.3.2 TEST1 Distance to Water Bodies

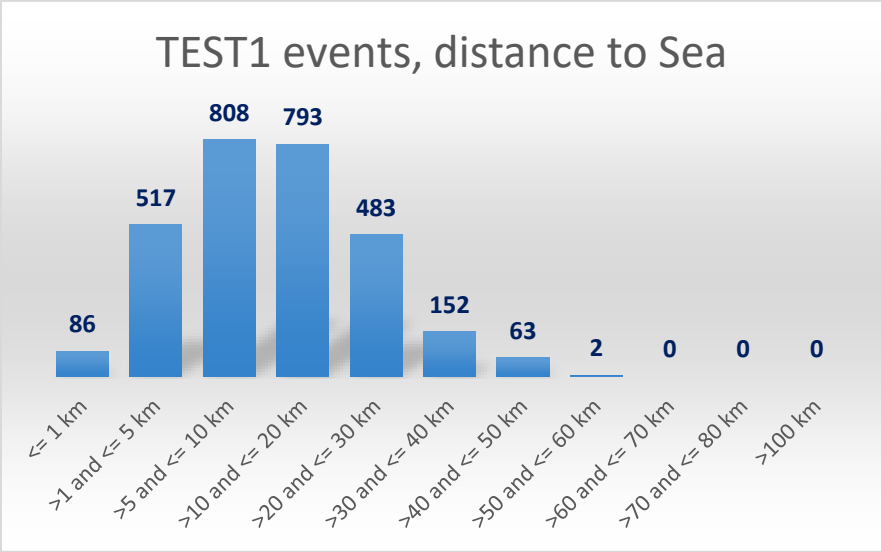
All Water Bodies			
D2H20Bin	Frequency	%	Cum%
<= 1 km	132	2.3%	2.3%
>1 and <= 5 km	820	14.4%	16.7%
>5 and <= 10 km	1015	17.8%	34.5%
>10 and <= 20 km	**1479**	**26.0%**	60.5%
>20 and <= 30 km	**1333**	**23.4%**	83.9%
>30 and <= 40 km	499	8.8%	92.7%
>40 and <= 50 km	235	4.1%	96.8%
>50 and <= 60 km	162	2.8%	99.6%
>60 and <= 70 km	14	0.2%	99.9%
>70 and <= 80 km	4	0.1%	100.0%
>100 km	2	0.0%	100.0%
TOTAL	**5695**		



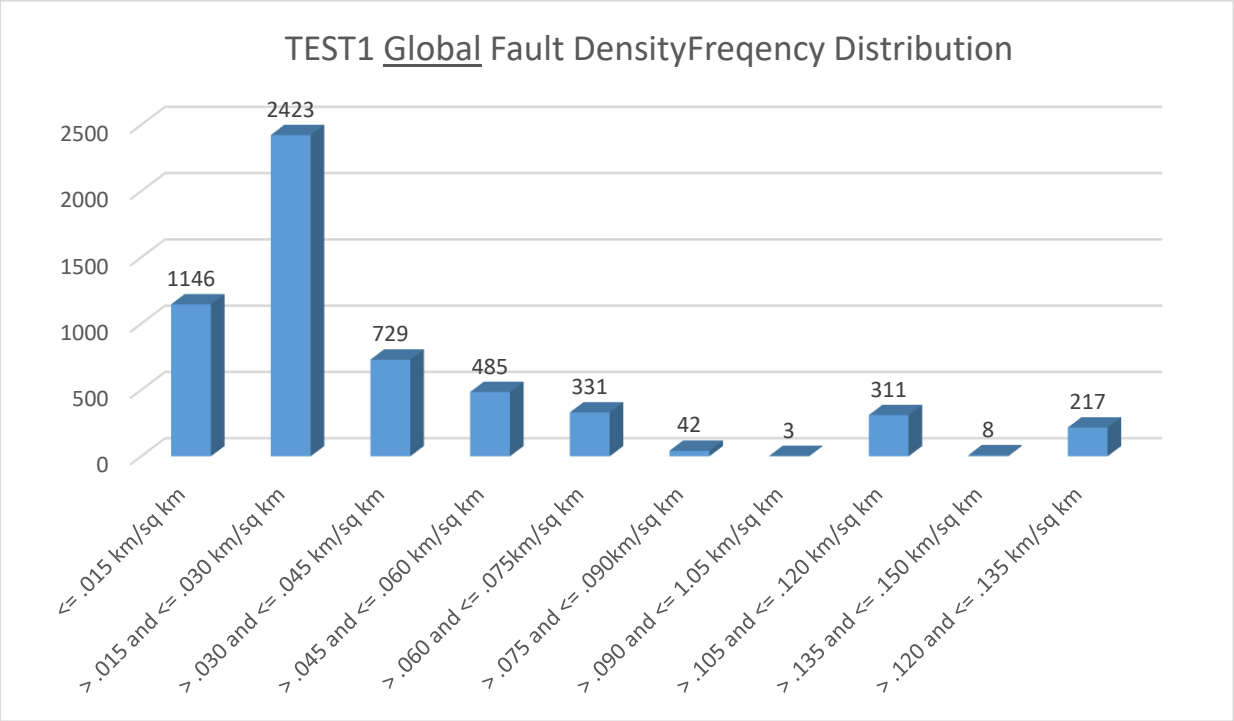
Water <> SEA (WATERFID <> 89111384)			
D2H20Bin	Frequency	%	Cum %
<= 1 km	46	1.6%	1.6%
>1 and <= 5 km	303	10.9%	12.5%
>5 and <= 10 km	207	7.4%	19.9%
>10 and <= 20 km	**686**	**24.6%**	44.5%
>20 and <= 30 km	**850**	**30.5%**	75.0%
>30 and <= 40 km	347	12.4%	87.4%
>40 and <= 50 km	172	6.2%	93.6%
>50 and <= 60 km	160	5.7%	99.3%
>60 and <= 70 km	14	0.5%	99.8%
>70 and <= 80 km	4	0.1%	99.9%
>100 km	2	0.1%	100.0%
	2791		



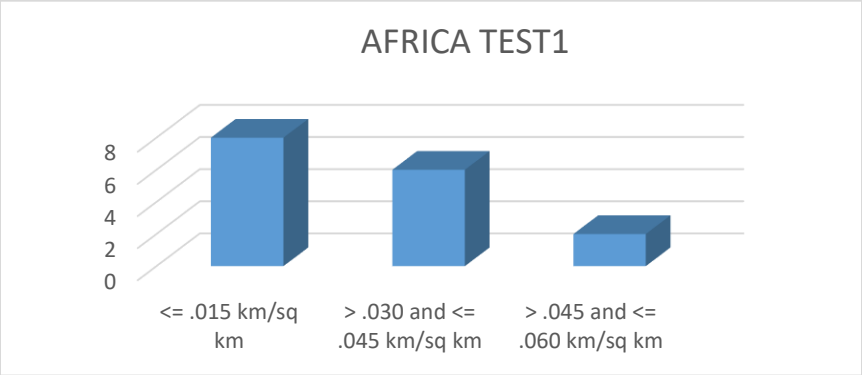
Water = SEA (WATERFID = 89111384)			
D2H20Bin	Frequency	%	Cum %
<= 1 km	86	3.0%	3.0%
>1 and <= 5 km	517	17.8%	20.8%
>5 and <= 10 km	**808**	**27.8%**	48.6%
>10 and <= 20 km	**793**	**27.3%**	75.9%
>20 and <= 30 km	483	16.6%	92.5%
>30 and <= 40 km	152	5.2%	97.8%
>40 and <= 50 km	63	2.2%	99.9%
>50 and <= 60 km	2	0.1%	100.0%
>60 and <= 70 km	0	0.0%	100.0%
>70 and <= 80 km	0	0.0%	100.0%
>100 km	0	0.0%	100.0%
	2904		

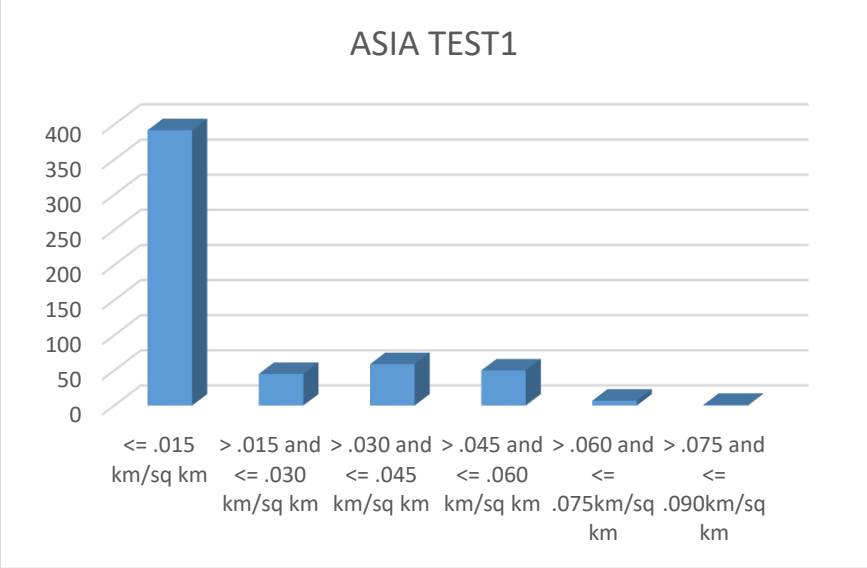


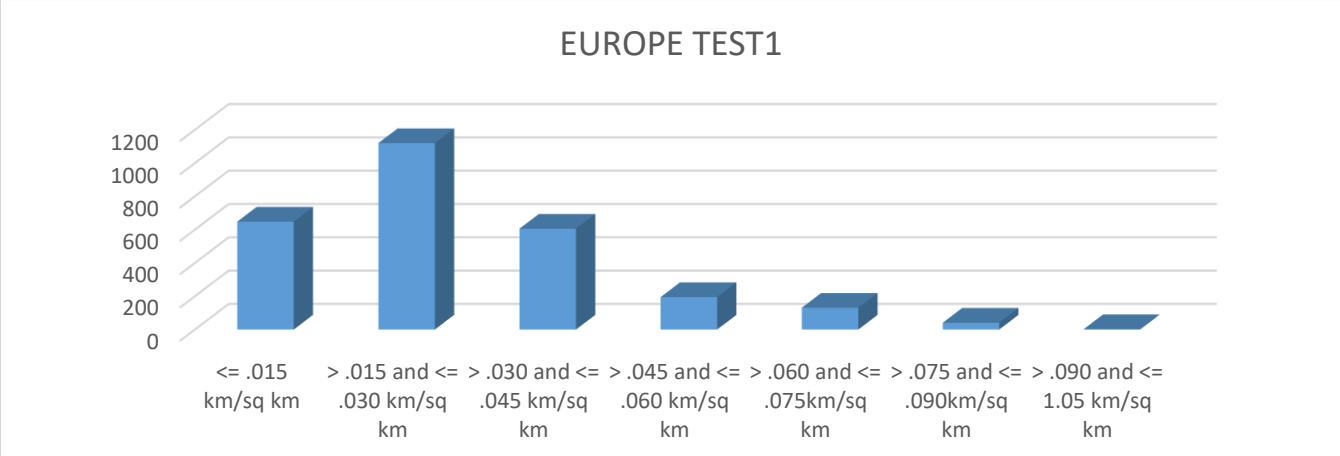
A.3.3 TEST1 Global Fault Density

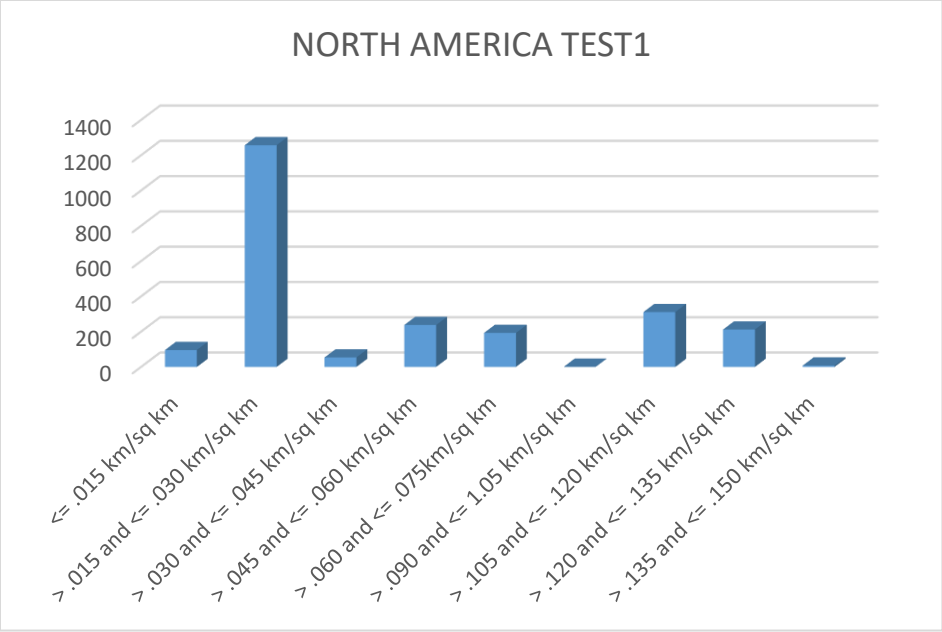


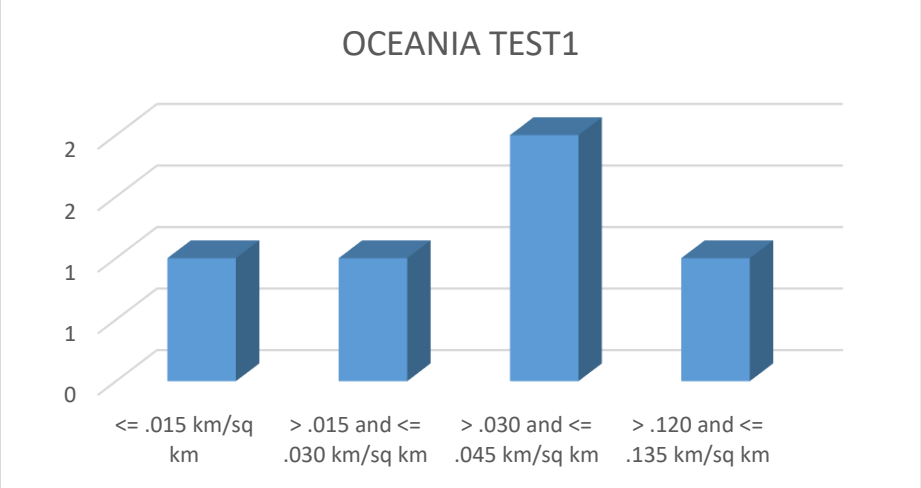
A.3.4 TEST1 Fault Density by Continent

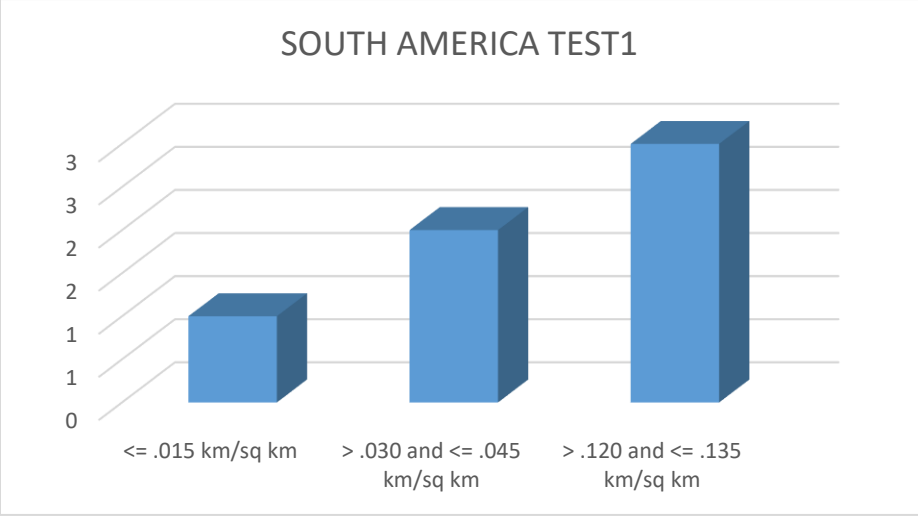







NORTH AMERICA TEST1
1400
1200
1000
800
600
400
200
0
<= .015 km/sq km
> .015 and <= .030 km/sq km
> .030 and <= .045 km/sq km
> .045 and <= .060 km/sq km
> .060 and <= .075km/sq km
> .090 and <= 1.05 km/sq km
> .105 and <= .120 km/sq km
> .120 and <= .135 km/sq km
> .135 and <= .150 km/sq km


OCEANIA TEST1
2
2
1
1
0
<= .015 km/sq km
> .015 and <= .030 km/sq km
> .030 and <= .045 km/sq km
> .120 and <= .135 km/sq km


SOUTH AMERICA TEST1
3
3
2
2
1
1
0
<= .015 km/sq km
> .030 and <= .045 km/sq km
> .120 and <= .135 km/sq km

CONTINENT	Qualitative Density	Frequency	% of Continent TEST1 Events
Africa	Very Low	10	62.5%
Africa	Low	6	37.5%
Asia	Very Low	446	80.5%
Asia	Low	104	18.8%
Asia	Moderate	4	0.7%
Australia	Very Low	1	100.0%
Europe	Very Low	1957	71.3%
Europe	Low	662	24.1%
Europe	Moderate	124	4.5%
North America	Very Low	1386	58.5%
North America	Low	327	13.8%
North America	Moderate	126	5.3%
North America	High	527	22.2%
North America	Very High	4	0.2%
Oceania	Very Low	2	40.0%
Oceania	Low	2	40.0%
Oceania	High	1	20.0%
South America	Very Low	2	33.3%
South America	Low	1	16.7%
South America	High	3	50.0%

Appendix B. Data Sources

B.1 Landslide Inventories

Source	(D. Kirschbaum et al., 2015)
Description	Database of 11,033 landslides, various landslide types, subset used for "TRAIN"
Resolution	not applicable
Coverage	global
Format	Point .shp file and database

B.2 Lithology

Source	(Jörg Hartmann & Nils Moosdorf, 2012)
Description	Representation of rock properties at the Earth surface
Resolution	0.5°
Coverage	global
Format	Raster grid

B.3 Topography

Source	(Danielson, 2011) U.S.G.S.
Description	Elevation model
Resolution	7.5 arcseconds (~250 meters)
Coverage	global
Format	Raster grid

B.4 Climate classification

Source	(Beck et al., 2018)
Description	Köppen-Geiger
Resolution	0.0083° resolution (approximately 1 km at the equator),
Coverage	global
Format	Raster grid

B.5 Precipitation

Source	(Fick & Hijmans, 2017)
Description	Monthly precipitation, temperature, (and more) aggregated across a temporal range of 1970 – 2000; also for 2081-2100
Resolution	1 km
Coverage	global
Format	Raster grid

B.6 Aridity

Source	(Trabucco & Zomer, 2018)
Description	Aridity and evapo-transpiration data for 1970-2000
Resolution	30 arcseconds
Coverage	global
Format	Raster grid

B.7 Soils

Source	(FAO, 2003) https://www.fao.org/land-water/land/land-governance/land-resources-planning-toolbox/category/details/en/c/1026564/
Description	Digitized version of the FAO-UNESCO Soil Map of the World produced in paper version at scale 1:5 million. 4931 mapping units consisting of soil associations, which are mixtures of different soil types, classified according to the FAO-UNESCO Legend
Resolution	1:5 000 000
Coverage	global
Format	polygon .shp file

B.8 Land use/Land cover

Source	(Kobayashi et al., 2017)
Description	Global Land Cover by National Mapping Organizations (GLCNMO). Version 3 (gm_lc_v3_2_2)
Resolution	500 m, 15’
Coverage	global
Format	Raster grid

B.9 Landforms

Source	(Sayre et al., 2014), (Charlie Frye et al., 2018)
Description	Landforms characterized with the climate regime, landform, geology, and land cover
Resolution	250 m
Coverage	global
Format	polygon .shp file

Appendix C –Global and Continental Environmental Data Coverages

C.1 Climate Classification Current/Future Global Coverage

							SORTED descending order on absolute value						
uture Scenario RCP 8.5			**1988-2016 Observed**		**Change in Sq Km**	**Abs**(Change in Sq Km	Abs(Change in Sq Km)		Future - Current				
KG Class	**Classification Sq Km**	**% Total**	**Classification Sq Km**	**% Total**	**Future - Current**	**Future - Current**	**Future - Current**		**Future**	**Current**	**% Change**	**Abs % Change**	
NonLand	362,751,940.43	71.12%	362,751,785.41	71.12%	155.02	155.02	Dfc	9,076,202.22	4.43%	10.59%	-6.16%	**6.16%**	Cold, no dry season, cold summer
Af	6,361,672.68	4.32%	7,051,752.77	4.79%	-690,080.08	690,080.08	Dfa	5,984,792.43	5.43%	1.37%	4.06%	**4.06%**	Cold, no dry season, hot summer
Am	5,552,329.63	3.77%	5,085,705.19	3.45%	466,624.44	466,624.44	ET	4,871,063.98	2.09%	5.40%	-3.31%	**3.31%**	Polar, tundra
Aw	20,216,665.35	13.72%	16,747,774.70	11.37%	3,468,890.65	3,468,890.65	BWh	3,603,680.75	17.25%	14.80%	2.45%	**2.45%**	Arid, desert, hot
BWh	25,407,651.81	17.25%	21,803,971.06	14.80%	3,603,680.75	3,603,680.75	Aw	3,468,890.65	13.72%	11.37%	2.35%	**2.35%**	Tropical, savannah
BWk	6,020,748.59	4.09%	7,653,999.43	5.20%	-1,633,250.84	1,633,250.84	BSh	2,839,928.64	7.27%	5.34%	1.93%	**1.93%**	Arid, steppe, hot
BSh	10,706,198.45	7.27%	7,866,269.80	5.34%	2,839,928.64	2,839,928.64	Cfa	1,867,704.85	4.99%	3.73%	1.27%	**1.27%**	Temperate, no dry season, hot summer
BSk	9,289,476.52	6.31%	8,956,705.06	6.08%	332,771.46	332,771.46	Dsc	1,777,334.18	2.29%	1.09%	1.21%	**1.21%**	Cold, dry summer, cold summer
Csa	1,968,415.66	1.34%	1,422,928.25	0.97%	545,487.41	545,487.41	BWk	1,633,250.84	4.09%	5.20%	-1.11%	**1.11%**	Arid, desert, cold
Csb	555,629.36	0.38%	854,369.60	0.58%	-298,740.24	298,740.24	Dwc	1,558,537.41	0.99%	2.05%	-1.06%	**1.06%**	Cold, dry winter, cold summer
Csc	6,261.21	0.00%	7,905.39	0.01%	-1,644.18	1,644.18	Cwa	1,280,208.18	1.98%	2.85%	-0.87%	**0.87%**	Temperate, dry winter, hot summer
Cwa	2,919,790.08	1.98%	4,199,998.25	2.85%	-1,280,208.18	1,280,208.18	Cwb	1,035,759.99	0.48%	1.18%	-0.70%	**0.70%**	Temperate, dry winter, warm summer
Cwb	706,411.72	0.48%	1,742,171.71	1.18%	-1,035,759.99	1,035,759.99	Dwa	883,263.54	1.42%	0.82%	0.60%	**0.60%**	Cold, dry winter, hot summer
Cwc	2,046.54	0.00%	6,698.84	0.00%	-4,652.30	4,652.30	EF	772,228.04	9.01%	9.53%	-0.52%	**0.52%**	Polar, frost
Cfa	7,358,221.29	4.99%	5,490,516.44	3.73%	1,867,704.85	1,867,704.85	Af	690,080.08	4.32%	4.79%	-0.47%	**0.47%**	Tropical, rainforest
Cfb	1,967,741.29	1.34%	2,525,032.43	1.71%	-557,291.14	557,291.14	Dfd	615,016.17	0.00%	0.42%	-0.42%	**0.42%**	Cold, no dry season, very cold winter
Cfc	139,557.87	0.09%	106,362.40	0.07%	33,195.48	33,195.48	Dfb	611,064.43	5.44%	5.03%	0.41%	**0.41%**	Cold, no dry season, warm summer
Dsa	254,186.00	0.17%	255,656.19	0.17%	-1,470.19	1,470.19	Cfb	557,291.14	1.34%	1.71%	-0.38%	**0.38%**	Temperate, no dry season, warm summ
Dsb	749,099.82	0.51%	587,156.44	0.40%	161,943.38	161,943.38	Csa	545,487.41	1.34%	0.97%	0.37%	**0.37%**	Temperate, dry summer, hot summer
Dsc	3,378,272.03	2.29%	1,600,937.85	1.09%	1,777,334.18	1,777,334.18	Am	466,624.44	3.77%	3.45%	0.32%	**0.32%**	Tropical, monsoon
Dsd	21.32	0.00%	58,703.30	0.04%	-58,681.98	58,681.98	BSk	332,771.46	6.31%	6.08%	0.23%	**0.23%**	Arid, steppe, cold
Dwa	2,090,540.03	1.42%	1,207,276.49	0.82%	883,263.54	883,263.54	Csb	298,740.24	0.38%	0.58%	-0.20%	**0.20%**	Temperate, dry summer, warm summer
Dwb	1,312,075.34	0.89%	1,206,630.77	0.82%	105,444.56	105,444.56	Dwd	227,454.28	0.00%	0.15%	-0.15%	**0.15%**	Cold, dry winter, very cold winter
Dwc	1,459,829.37	0.99%	3,018,366.78	2.05%	-1,558,537.41	1,558,537.41	Dsb	161,943.38	0.51%	0.40%	0.11%	**0.11%**	Cold, dry summer, warm summer
Dwd	267.69	0.00%	227,721.97	0.15%	-227,454.28	227,454.28	Dwb	105,444.56	0.89%	0.82%	0.07%	**0.07%**	Cold, dry winter, warm summer
Dfa	7,995,941.31	5.43%	2,011,148.88	1.37%	5,984,792.43	5,984,792.43	Dsd	58,681.98	0.00%	0.04%	-0.04%	**0.04%**	Cold, dry summer, very cold winter
Dfb	8,017,569.22	5.44%	7,406,504.79	5.03%	611,064.43	611,064.43	Cfc	33,195.48	0.09%	0.07%	0.02%	**0.02%**	Temperate, no dry season, cold summer
Dfc	6,520,689.87	4.43%	15,596,892.09	10.59%	-9,076,202.22	9,076,202.22	Cwc	4,652.30	0.00%	0.00%	0.00%	**0.00%**	Temperate, dry winter, cold summer
Dfd	7.04	0.00%	615,023.21	0.42%	-615,016.17	615,016.17	Csc	1,644.18	0.00%	0.01%	0.00%	**0.00%**	Temperate, dry summer, cold summer
ET	3,082,824.69	2.09%	7,953,888.66	5.40%	-4,871,063.98	4,871,063.98	Dsa	1,470.19	0.17%	0.17%	0.00%	**0.00%**	Cold, dry summer, hot summer
EF	13,273,539.52	9.01%	14,045,767.56	9.53%	-772,228.04	772,228.04	NonLand	155.02					

C.2 Landcover Global Coverage

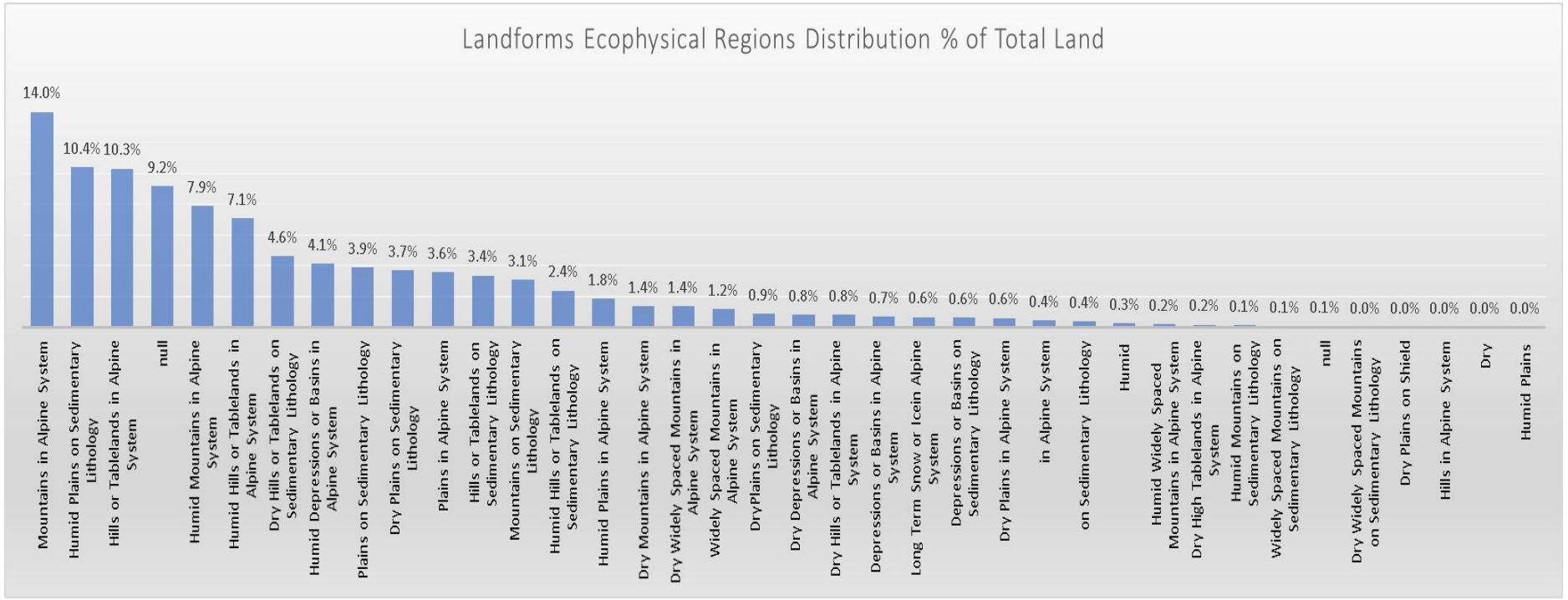
GLCNMO Landuse/Landcover		
Description	**sq km**	**% Earth**
Broadleaf Evergreen Forest	*9,686,507*	*1.90%*
Broadleaf Deciduous Forest	*6,636,292*	*1.30%*
Needleleaf Evergreen Forest	*4,010,189*	*0.79%*
Needleleaf Deciduous Forest	*5,625,814*	*1.10%*
Mixed Forest	*4,996,624*	*0.98%*
Tree Open	*13,430,207*	*2.63%*
FOREST/WOODLAND	**44,385,633**	**8.70%**
Shrub	*8,346,598*	*1.64%*
Herbaceous	*14,410,931*	*2.83%*
Herbaceous with Sparse Tree/Shrub	*2,559,545*	*0.50%*
Sparse vegetation	*10,336,721*	*2.03%*
GRASSLAND/SHRUBLAND	**35,653,794**	**6.99%**
Cropland	*10,626,938*	*2.08%*
Paddy field	*1,676,802*	*0.33%*
Cropland / Other Vegetation Mosaic	*7,111,474*	*1.39%*
CROPLAND	**19,415,214**	**3.81%**
Mangrove	**64,428**	**0.01%**
Wetland	**1,206,413**	**0.24%**
Bare area, consolidated (gravel, rock)	*5,106,962*	*1.00%*
Bare area, unconsolidated (sand)	*6,003,915*	*1.18%*
BARE GROUND	**12,381,717**	**2.43%**
Urban	**420,736**	**0.08%**
Snow / Ice	**54,430,306**	**10.67%**
Water bodies	**343,384,598**	**67.32%**

C.3 Aridity Global and Continental Coverage

	Aridity Classification Distribution							
	GLOBAL	**Africa**	**Asia**	**Australia**	**Europe**	**No America**	**Oceania**	**So America**
Hyper-Arid	6.4%	28.2%	4.5%	0.0%	0.0%	0.1%	0.0%	1.7%
Arid	15.6%	21.6%	19.0%	65.2%	1.0%	5.7%	0.0%	9.1%
Semi-Arid	19.6%	19.8%	21.4%	25.2%	12.5%	21.6%	1.1%	12.9%
Dry Sub-Humid	12.0%	8.2%	14.7%	4.7%	12.3%	14.1%	4.4%	7.6%
Humid	46.3%	22.2%	40.4%	4.9%	74.3%	58.5%	94.5%	68.7%

C.4 Landform Global Coverage

Murphy Bridges Ecophysical Region	% Total	Area (km2)
Mountains in Alpine System	14.0%	20,598,913
Humid Plains on Sedimentary Lithology	10.4%	15,280,803
Hills or Tablelands in Alpine System	10.3%	15,173,709
null	9.2%	13,477,087
Humid Mountains in Alpine System	7.9%	11,592,738
Humid Hills or Tablelands in Alpine System	7.1%	10,420,467
Dry Hills or Tablelands on Sedimentary Lithology	4.6%	6,791,995
Humid Depressions or Basins in Alpine System	4.1%	6,040,808
Plains on Sedimentary Lithology	3.9%	5,700,728
Dry Plains on Sedimentary Lithology	3.7%	5,414,018
Plains in Alpine System	3.6%	5,269,229
Hills or Tablelands on Sedimentary Lithology	3.4%	4,938,424
Mountains on Sedimentary Lithology	3.1%	4,520,414
Humid Hills or Tablelands on Sedimentary Lithology	2.4%	3,499,916
Humid Plains in Alpine System	1.8%	2,698,149
Dry Mountains in Alpine System	1.4%	2,030,136
Dry Widely Spaced Mountains in Alpine System	1.4%	2,006,615
Widely Spaced Mountains in Alpine System	1.2%	1,708,549
DryPlains on Sedimentary Lithology	0.9%	1,254,926
Dry Depressions or Basins in Alpine System	0.8%	1,223,009
Dry Hills or Tablelands in Alpine System	0.8%	1,161,698
Depressions or Basins in Alpine System	0.7%	1,023,209
Long Term Snow or Icein Alpine System	0.6%	955,745
Depressions or Basins on Sedimentary Lithology	0.6%	939,951
Dry Plains in Alpine System	0.6%	867,898
in Alpine System	0.4%	658,119
on Sedimentary Lithology	0.4%	548,906
Humid	0.3%	392,736
Humid Widely Spaced Mountains in Alpine System	0.2%	262,228
Dry High Tablelands in Alpine System	0.2%	232,286
Humid Mountains on Sedimentary Lithology	0.1%	205,585
Widely Spaced Mountains on Sedimentary Lithology	0.1%	121,110
null	0.1%	120,793
Dry Widely Spaced Mountains on Sedimentary Lithology	0.0%	33,782
Dry Plains on Shield	0.0%	20,317
Hills in Alpine System	0.0%	17,828
Dry	0.0%	6,136
Humid Plains	0.0%	21
Total	**100.0%**	**147,208,983**


Landforms Ecophysical Regions Distribution % of Total Land
14.0%
10.4%
10.3%
9.2%
7.9%
7.1%
4.6%
4.1%
3.9%
3.7%
3.6%
3.4%
3.1%
2.4%
1.8%
1.4%
1.4%
1.2%
0.9%
0.8%
0.8%
0.7%
0.6%
0.6%
0.6%
0.4%
0.4%
0.3%
0.2%
0.2%
0.1%
0.1%
0.1%
0.0%
0.0%
0.0%
0.0%
0.0%
Mountains in Alpine System
Humid Plains on Sedimentary Lithology
Hills or Tablelands in Alpine System
null
Humid Mountains in Alpine System
Humid Hills or Tablelands in Alpine System
Dry Hills or Tablelands on Sedimentary Lithology
Humid Depressions or Basins in Alpine System
Plains on Sedimentary Lithology
Dry Plains on Sedimentary Lithology
Plains in Alpine System
Hills or Tablelands on Sedimentary Lithology
Mountains on Sedimentary Lithology
Humid Hills or Tablelands on Sedimentary Lithology
Humid Plains in Alpine System
Dry Mountains in Alpine System
Dry Widely Spaced Mountains in Alpine System
Widely Spaced Mountains in Alpine System
DryPlains on Sedimentary Lithology
Dry Depressions or Basins in Alpine System
Dry Hills or Tablelands in Alpine System
Depressions or Basins in Alpine System
Long Term Snow or Icein Alpine System
Depressions or Basins on Sedimentary Lithology
Dry Plains in Alpine System
in Alpine System
on Sedimentary Lithology
Humid
Humid Widely Spaced Mountains in Alpine System
Dry High Tablelands in Alpine System
Humid Mountains on Sedimentary Lithology
Widely Spaced Mountains on Sedimentary Lithology
null
Dry Widely Spaced Mountains on Sedimentary Lithology
Dry Plains on Shield
Hills in Alpine System
Dry
Humid Plains

C.5 Elevation Global and Continental Coverage

from "Continental Hypsography" Harrison 1983 except Oceania which was calculated in ArcGIS														
ELEVATION (m)	**AFRICA (km^2)**	**%**	**ASIA (km^2)**	**%**	**N AMERICA (km^2)**		**S AMERICA (km^2)**	**%**	**EUROPE (km^2)**	**%**	**AUSTRALIA (km^2)**	**%**	**OCEANIA (km^2)**	**%**
0-1000	23,381,600	80.6%	29,498,900	72.3%	16,878,000	73.1%	14,557,600	86.9%	8,790,000	96.7%	7,634,500	96.6%	375,998	88.2%
1000-1500	4,617,000	15.9%	4,838,300	11.9%	2,645,300	11.5%	484,600	2.9%	186,300	2.0%	108,200	1.4%	35,330	8.3%
1500-2000	823,000	2.8%	2,301,300	5.6%	2,213,800	9.6%	488,500	2.9%	87,500	1.0%	49,200	0.6%	11,518	2.7%
2000-2500	135,100	0.5%	741,500	1.8%	717,000	3.1%	173,300	1.0%	25,000	0.3%	24,600	0.3%	2,158	0.5%
2500 - 3000	54,800	0.2%	698,200	1.7%	628,700	2.7%	308,500	1.8%	2,300	0.0%	61,500	0.8%	593	0.1%
3000-3500	3,100	0.0%	290,100	0.7%	15,100	0.1%	173,900	1.0%	0	0.0%	12,300	0.2%	298	0.1%
3500-4000	0	0.0%	319,600	0.8%	0	0.0%	186,700	1.1%	0	0.0%	12,300	0.2%	200	0.0%
4000-5000	0	0.0%	1,561,200	3.8%	0	0.0%	383,400	2.3%	0	0.0%	0	0.0%	17	0.0%
5000-6750	0	0.0%	569,700	1.4%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%

C.6 Fault Density Global and Continental Coverage

from ArcGIS and miscellaneous tools									
Qualitative Label	**Value Range (km/sq km)**	**Global %**	**Africa %**	**Asia %**	**Australia %**	**Europe %**	**North America %**	**Oceania %**	**South America %**
Very Low	<= 0.033885	99.2%	97.3%	96.4%	98.3%	94.0%	99.1%	59.4%	95.6%
Low	> 0.033885 and <= 0.067771	0.7%	2.3%	3.4%	1.6%	4.8%	0.8%	23.7%	4.1%
Moderate	> 0.067771 and <= 0.101656	0.1%	0.4%	0.3%	0.1%	1.1%	0.1%	12.7%	0.3%
High	> 0.101656 and <= 0.135542	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	3.7%	0.0%
Very High	> 0.135542	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%

C.7 Tectonics Global and Continental Coverage

Four of seven continents (Asia, Australia, Oceania, South America) have Reverse faults as major percentage in fault length (km). Three of seven continents (Africa, Europe, North America) exhibit Normal faults as the majority by length (km). Only Oceania switches dominant slip type when evaluating based on frequency versus length. Globally, Reverse faults dominate by km length. Globally, Normal faults dominate by frequency. However, this is dependent on the level of data uncertainty in the source tectonic maps.

C.7.1 Tectonics Global Coverage

Slip Type	GLOBAL			
	SLIP TYPE LINEAR TOTAL (km)	% Total (by km)	Slip Type Count	% Total by Count
unknown	*15,259*	*3.2%*	*1,275*	*11.2%*
Anticline	19,043	4.0%	312	2.7%
Blind Thrust	35	0.0%	1	0.0%
Dextral	83,691	17.8%	1,168	10.3%
Dextral-Normal	9,309	2.0%	202	1.8%
Dextral-Oblique	47	0.0%	1	0.0%
Dextral-Reverse	12,158	2.6%	360	3.2%
Dextral_Transform	22	0.0%	1	0.0%
Normal	87,488	18.6%	2,966	26.1%
Normal-Dextral	880	0.2%	26	0.2%
Normal-Sinistral	635	0.1%	15	0.1%
Normal-Strike-Slip	40	0.0%	2	0.0%
Reverse	132,711	28.2%	2,635	23.2%
Reverse-Dextral	3,424	0.7%	77	0.7%
Reverse-Sinistral	1,303	0.3%	24	0.2%
Reverse-Strike-Slip	2,796	0.6%	396	3.5%
Sinistral	75,613	16.1%	946	8.3%
Sinistral-Normal	7,966	1.7%	158	1.4%
Sinistral-Reverse	11,220	2.4%	172	1.5%
Sinistral_Transform	135	0.0%	6	0.1%
Spreading_Ridge	366	0.1%	13	0.1%
Strike-Slip	5,106	1.1%	568	5.0%
Subduction_Thrust	975	0.2%	15	0.1%
Syncline	718	0.2%	9	0.1%
TOTALS	**470,940**		**11,348**	

C.7.2 Tectonics Africa Coverage

Slip Type	AFRICA				
	Count	Slip Type Count % By Continent	Linear Total (km)	SlipType (km)% Continent	SlipType (km)% Global
unknown	*51*	*8.9%*	*4,138*	*10.81%*	*0.88%*
Anticline	4	0.7%	167	0.44%	0.04%
Blind Thrust	1	0.2%	35	0.09%	0.01%
Dextral	7	1.2%	228	0.60%	0.05%
Dextral-Normal	3	0.5%	72	0.19%	0.02%
Dextral-Oblique	0	0	0	0	0
Dextral-Reverse	3	0.5%	117	0.31%	0.02%
Dextral_Transform	0	0	0	0	0
Normal	416	72.5%	26,524	69.29%	5.63%
Normal-Dextral	0	0.0%	0	0	0
Normal-Sinistral	2	0.3%	41	0.11%	0.01%
Normal-Strike-Slip	0	0	0	0	0
Reverse	70	12.2%	5,915	15.45%	1.26%
Reverse-Dextral	1	0.2%	36	0.09%	0.01%
Reverse-Sinistral	1	0.2%	100	0.26%	0.02%
Reverse-Strike-Slip	0	0	0	0	0
Sinistral	8	1.4%	510	1.33%	0.11%
Sinistral-Normal	1	0.2%	28	0.07%	0.01%
Sinistral-Reverse	5	0.9%	302	0.79%	0.06%
Sinistral_Transform	0	0	0	0	0
Spreading_Ridge	0	0	0	0	0
Strike-Slip	1	0.2%	65	0.17%	0.01%
Subduction_Thrust	0	0	0	0	0
Syncline	0	0	0	0	0
TOTALS	**574**		**38,277**	**1**	**8.1%**

C.7.3 Tectonics Asia Coverage

Slip Type	ASIA				
	Count	Slip Type Count % By Continent	Linear Total (km)	SlipType (km)% Continent	SlipType (km)% Global
unknown	*18*	*0.53%*	*379*	*0.15%*	*0.08%*
Anticline	308	9.05%	18,876	7.43%	4.01%
Blind Thrust	0	0	0	0	0
Dextral	745	21.89%	59,637	23.49%	12.66%
Dextral-Normal	94	2.76%	5,031	1.98%	1.07%
Dextral-Oblique	1	0.03%	47	0.02%	0.01%
Dextral-Reverse	65	1.91%	5,276	2.08%	1.12%
Dextral_Transform	0	0	0	0	0
Normal	439	12.90%	22,258	8.77%	4.73%
Normal-Dextral	1	0.03%	58	0.02%	0.01%
Normal-Sinistral	0	0	0	0	0
Normal-Strike-Slip	0	0	0	0	0
Reverse	810	23.80%	67,401	26.55%	14.31%
Reverse-Dextral	10	0.29%	710	0.28%	0.15%
Reverse-Sinistral	6	0.18%	403	0.16%	0.09%
Reverse-Strike-Slip	7	0.21%	348	0.14%	0.07%
Sinistral	698	20.51%	58,807	23.16%	12.49%
Sinistral-Normal	97	2.85%	5,281	2.08%	1.12%
Sinistral-Reverse	83	2.44%	7,367	2.90%	1.56%
Sinistral_Transform	0	0	0	0	0
Spreading_Ridge	0	0	0	0	0
Strike-Slip	6	0.18%	691	0.27%	0.15%
Subduction_Thrust	6	0.18%	617	0.24%	0.13%
Syncline	9	0.26%	718	0.28%	0.15%
TOTALS	**3403**		**253,906**	**1**	**53.9%**

C.7.4 Tectonics Australia Coverage

Slip Type	AUSTRALIA				
	Count	Slip Type Count % By Continent	Linear Total (km)	SlipType (km)% Continent	SlipType (km)% Global
unknown	*859*	*50.71%*	*9,073*	*30.4%*	*1.9%*
Anticline	0		0	0	0
Blind Thrust	0		0	0	0
Dextral	0		0	0	0
Dextral-Normal	0		0	0	0
Dextral-Oblique	0		0	0	0
Dextral-Reverse	170	10.0%	222	0.7%	0.0%
Dextral_Transform	0		0	0	0
Normal	7		254	0.9%	0.1%
Normal-Dextral	0		0	0	0
Normal-Sinistral	0		0	0	0
Normal-Strike-Slip	0		0	0	0
Reverse	625	36.9%	19,182	64.3%	4.1%
Reverse-Dextral	0		0	0	0
Reverse-Sinistral	0		0	0	0
Reverse-Strike-Slip	0		0	0	0
Sinistral	0		0	0	0
Sinistral-Normal	0		0	0	0
Sinistral-Reverse	10	0.6%	208	0.7%	0.0%
Sinistral_Transform	0		0	0	0
Spreading_Ridge	0		0	0	0
Strike-Slip	23	1.4%	892	3.0%	0.2%
Subduction_Thrust	0		0	0	0
Syncline	0		0	0	0
TOTALS	**1694**		**29,832**	**1**	**6.3%**

C.7.5 Tectonics Europe Coverage

Slip Type	EUROPE				
	Count	Slip Type Count % By Continent	Linear Total (km)	SlipType (km)% Continent	SlipType (km)% Global
unknown	*1*	*0.001366*	*80*	*0.2%*	*0.02%*
Anticline	0	0	0	0	0
Blind Thrust	0	0	0	0	0
Dextral	48	6.6%	3,806	9.6%	0.8%
Dextral-Normal	41	5.6%	1,534	3.9%	0.3%
Dextral-Oblique	0	0	0	0	0
Dextral-Reverse	56	7.7%	2,917	7.4%	0.6%
Dextral_Transform	1	0.1%	22	0.1%	0.0%
Normal	300	41.0%	12,779	32.3%	2.7%
Normal-Dextral	0	0	0	0	0
Normal-Sinistral	4	0.5%	120	0.3%	0.0%
Normal-Strike-Slip	0	0	0	0	0
Reverse	116	15.8%	7,380	18.7%	1.6%
Reverse-Dextral	0	0	0	0	0
Reverse-Sinistral	2	0.3%	54	0.1%	0.0%
Reverse-Strike-Slip	0	0	0	0	0
Sinistral	80	10.9%	7,035	17.8%	1.5%
Sinistral-Normal	40	5.5%	1,940	4.9%	0.4%
Sinistral-Reverse	24	3.3%	1,118	2.8%	0.2%
Sinistral_Transform	3	0.4%	103	0.3%	0.0%
Spreading_Ridge	11	1.5%	351	0.9%	0.1%
Strike-Slip	0	0	0	0	0
Subduction_Thrust	5	0.7%	280	0.7%	0.1%
Syncline	0	0	0	0	0
TOTALS	**732**		**39,519**	**1**	**8.4%**

C.7.6 Tectonics North America Coverage

Slip Type	NORTH AMERICA				
	Count	Slip Type Count % By Continent	Linear Total (km)	SlipType (km)% Continent	SlipType (km)% Global
unknown	*185*	*6.0%*	*1,015*	*0*	*0*
Anticline	0		0	0	0
Blind Thrust	0		0	0	0
Dextral	175	5.7%	8,272	16.92%	1.76%
Dextral-Normal	17	0.6%	595	1.22%	0.13%
Dextral-Oblique	0		0	0	0
Dextral-Reverse	7	0.2%	398	0.82%	0.08%
Dextral_Transform	0		0	0	0
Normal	1071	34.8%	18,653	38.16%	3.96%
Normal-Dextral	8	0.3%	255	0.52%	0.05%
Normal-Sinistral	5	0.2%	343	0.70%	0.07%
Normal-Strike-Slip	0		0	0	0
Reverse	590	19.2%	7,806	15.97%	1.66%
Reverse-Dextral	6	0.2%	364	0.74%	0.08%
Reverse-Sinistral	2	0.1%	70	0.14%	0.01%
Reverse-Strike-Slip	374	12.2%	2,078	4.25%	0.44%
Sinistral	93	3.0%	5,019	10.27%	1.07%
Sinistral-Normal	6	0.2%	269	0.55%	0.06%
Sinistral-Reverse	15	0.5%	597	1.22%	0.13%
Sinistral_Transform	3	0.1%	32	0.06%	0.01%
Spreading_Ridge	2	0.1%	16	0.03%	0.003%
Strike-Slip	516	16.8%	3,081	6.30%	0.65%
Subduction_Thrust	1	0.0%	13	0.03%	0.00%
Syncline	0		0	0	0
TOTALS	**3076**		**48,876**	**1**	**10.4%**

C.7.7 Tectonics Oceania Coverage

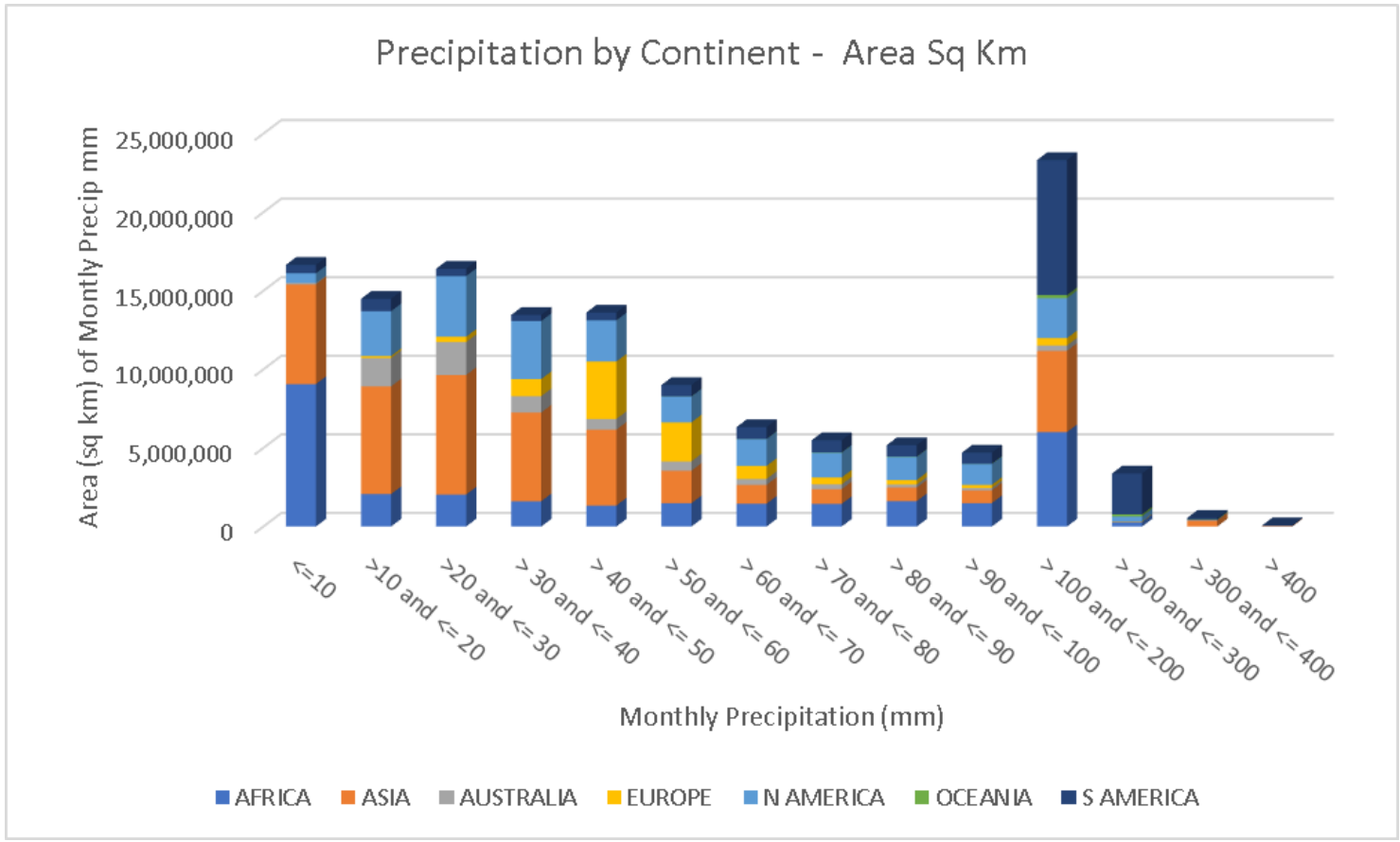
Slip Type	OCEANIA				
	Count	Slip Type Count % By Continent	Linear Total (km)	SlipType (km)% Continent	SlipType (km)% Global
unknown	*156*	*15.8%*	*363*	*3.2%*	*0.1%*
Anticline	0		0	0	0
Blind Thrust	0		0	0	0
Dextral	25	2.5%	784	7.0%	0.2%
Dextral-Normal	18	1.8%	726	6.5%	0.2%
Dextral-Oblique	0		0	0	0
Dextral-Reverse	45	4.6%	2,298	20.5%	0.5%
Dextral_Transform	0		0	0	0
Normal	587	59.5%	1,825	16.3%	0.4%
Normal-Dextral	13	1.3%	358	3.2%	0.1%
Normal-Sinistral	0		0	0	0
Normal-Strike-Slip	1	0.1%	20	0.2%	0.0%
Reverse	71	7.2%	2,512	22.4%	0.5%
Reverse-Dextral	53	5.4%	1,926	17.2%	0.4%
Reverse-Sinistral	8	0.8%	236	2.1%	0.1%
Reverse-Strike-Slip	2	0.2%	42	0.4%	0.0%
Sinistral	0		0	0	0
Sinistral-Normal	0		0	0	0
Sinistral-Reverse	1	0.1%	38	0.3%	0.0%
Sinistral_Transform	0		0	0	0
Spreading_Ridge	0		0	0	0
Strike-Slip	4	0.4%	12	0.1%	0.0%
Subduction_Thrust	2	0.2%	59	0.5%	0.0%
Syncline	0		0	0	0
TOTALS	**986**		**11,199**	**1**	**2.4%**

C.7.8 Tectonics South America Coverage

Slip Type	SOUTH AMERICA				
	Count	Slip Type Count % By Continent	Linear Total (km)	SlipType (km)% Continent	SlipType (km)% Global
unknown	*5*	*0.6%*	*211*	*0.4%*	*0.0%*
Anticline	0		0	0	0
Blind Thrust	0		0	0	0
Dextral	168	19.0%	10,965	22.2%	2.3%
Dextral-Normal	29	3.3%	1,352	2.7%	0.3%
Dextral-Oblique	0		0	0	0
Dextral-Reverse	14	1.6%	930	1.9%	0.2%
Dextral_Transform	0		0	0	0
Normal	146	16.5%	5,195	10.5%	1.1%
Normal-Dextral	4	0.5%	208	0.4%	0.0%
Normal-Sinistral	4	0.5%	131	0.3%	0.0%
Normal-Strike-Slip	1	0.1%	20	0.0%	0.0%
Reverse	353	40.0%	22,514	45.6%	4.8%
Reverse-Dextral	7	0.8%	389	0.8%	0.1%
Reverse-Sinistral	5	0.6%	440	0.9%	0.1%
Reverse-Strike-Slip	13	1.5%	328	0.7%	0.1%
Sinistral	67	7.6%	4,242	8.6%	0.9%
Sinistral-Normal	14	1.6%	448	0.9%	0.1%
Sinistral-Reverse	34	3.9%	1,590	3.2%	0.3%
Sinistral_Transform	0		0	0	0
Spreading_Ridge	0		0	0	0
Strike-Slip	18	2.0%	364	0.7%	0.1%
Subduction_Thrust	1	0.1%	5	0	0
Syncline	0		0	0	0
TOTALS	**883**		**49,332**	**1**	**10.5%**

C.8 Precipitation

PRECIPITATION - Average Monthly by Continent														
	Area Sq Km													
PRECIPITATION (mm)	**AFRICA**	**%**	**ASIA**	**%**	**AUSTRALIA**	**%**	**EUROPE**	**%**	**N AMERICA**	**%**	**OCEANIA**		**S AMERICA**	**%**
<=10	9,066,803	30.3%	6,344,900	14.8%	92,048	1.2%	0		610,948	2.5%	0		537,617	0
>10 and <= 20	2,066,967	6.9%	6,850,466	16.0%	1,805,885	23.5%	123,715		2,854,980	11.8%	0		788,207	4.5%
>20 and <= 30	2,035,205	6.8%	7,606,165	17.8%	2,094,027	27.2%	341,539		3,861,280	15.9%	0		470,412	2.7%
> 30 and <= 40	1,609,586	5.4%	5,644,396	13.2%	1,042,781	13.6%	1,087,101	10.9%	3,692,472	15.2%	337	0.1%	386,998	2.2%
> 40 and <= 50	1,338,048	4.5%	4,825,631	11.3%	679,843	8.8%	3,651,505	36.7%	2,623,367	10.8%	2,490	0.6%	497,066	2.8%
> 50 and <= 60	1,487,293	5.0%	2,065,903	4.8%	596,058	7.8%	2,479,932	25.0%	1,644,290	6.8%	9,318	2.2%	723,257	4.1%
> 60 and <= 70	1,454,380	4.9%	1,196,688	2.8%	378,328	4.9%	**831,297**	**8.4%**	1,702,339	7.0%	11,991	2.8%	750,900	4.2%
> 70 and <= 80	1,440,546	4.8%	934,733	2.2%	304,349	4.0%	439,571	4.4%	1,578,428	6.5%	14,454	3.4%	778,317	4.4%
> 80 and <= 90	1,616,140	5.4%	866,756	2.0%	189,120	2.5%	291,208	2.9%	1,467,088	6.0%	17,998	4.2%	730,348	4.1%
> 90 and <= 100	1,492,886	5.0%	806,579	1.9%	149,785	1.9%	204,398	2.1%	1,321,350	5.4%	20,416	4.8%	703,828	4.0%
> 100 and <= 200	6,010,443	20.1%	5,166,234	12.1%	345,681	4.5%	465,634	4.7%	2,541,883	10.5%	194,421	45.4%	8,593,374	48.5%
> 200 and <= 300	279,528	0.9%	10,510	0.0%	7,890	0.1%	21,916	0.2%	327,755	1.4%	129,577	30.2%	2,601,303	14.7%
> 300 and <= 400	6,709	0.0%	386,911	0.9%	653	0.0%	0	0.0%	26,983	0.1%	26,986	6.3%	99,929	0.6%
> 400	0	0.0%	34,965	0.1%	0	0.0%	0	0.0%	416	0.0%	522	0.1%	41,776	0.2%


Precipitation by Continent - Area Sq Km
Area (sq km) of Montly Precip mm
25,000,000
20,000,000
15,000,000
10,000,000
5,000,000
0
<=10
>10 and <= 20
>20 and <= 30
> 30 and <= 40
> 40 and <= 50
> 50 and <= 60
> 60 and <= 70
> 70 and <= 80
> 80 and <= 90
> 90 and <= 100
> 100 and <= 200
> 200 and <= 300
> 300 and <= 400
> 400
Monthly Precipitation (mm)
AFRICA
ASIA
AUSTRALIA
EUROPE
N AMERICA
OCEANIA
S AMERICA

C.9 Elevation

ELEVATION	area square kilometers						
meters	AFRICA	ASIA	N AMERICA	S AMERICA	EUROPE	AUSTRALIA	OCEANIA
0-1000	23,381,600	29,498,900	16,878,000	14,557,600	8,790,000	7,634,500	375,998
1000-1500	4,617,000	4,838,300	2,645,300	484,600	186,300	108,200	35,330
1500-2000	823,000	2,301,300	2,213,800	488,500	87,500	49,200	11,518
2000-2500	135,100	741,500	717,000	173,300	25,000	24,600	2,158
2500 - 3000	54,800	698,200	628,700	308,500	2,300	61,500	593
3000-3500	3,100	290,100	15,100	173,900	0	12,300	298
3500-4000	0	319,600	0	186,700	0	12,300	200
4000-5000	0	1,561,200	0	383,400	0	0	17
5000-6750	0	569,700	0	0	0	0	0
TOTALS	**29,014,600**	**40,818,800**	**23,097,900**	**16,756,500**	**9,091,100**	**7,902,600**	**426,112**

C.10 Lithology

Lithology by Continent							
	AFRICA	ASIA	AUSTRALIA	EUROPE	No AMERICA	OCEANIA	So AMERICA
Type	Area Sq Km	Area Sq Km	Area Sq Km	Area Sq Km	Area Sq Km	Area Sq Km	Area Sq Km
plutonic - acidic	373,293	3,316,140	286,835	716,338	2,060,871	7,662	1,717,469
plutonic - basic	62,061	574,243	47,072	67,426	189,384	5,438	36,911
plutonic - intermediate	27,009	155,075	16,321	32,049	217,095	3,116	126,847
sedimentary - carbonate	3,077,931	4,410,360	291,071	1,507,074	2,189,922	42,996	275,248
sedimentary - mixed	1,510,617	10,723,231	686,052	3,452,679	3,724,199	90,565	2,181,832
sedimentary - siliciclastic	5,329,738	6,852,629	1,022,339	1,277,254	5,262,107	31,738	4,555,071
sedimentary - unconsolidated	11,451,336	10,986,007	4,818,516	1,876,390	3,114,967	85,462	4,674,884
volcanic-acidic	29,883	664,456	71,429	89,539	388,304	11,048	281,951
volcanic - basic	1,062,051	1,982,073	153,265	167,489	960,664	52,416	751,347
volcanic-intermediate	205,647	1,356,158	52,351	37,935	419,073	15,676	451,993
volcanic - pyroclastic	496	269,988	2,261	4,841	388,852	15,293	254,970
evaporite	200,971	127,639	55,565	0	15,719	0	45,386
ice/glacier	0	34,182	0	26,490	1,833,729	0	17,207
metamorphic	8,982,688	3,092,788	167,228	1,027,947	3,205,990	44,716	2,240,437
no data	4,877	78,692	74	748	15,306	0	3,030
water body	280,966	568,490	8,109	210,297	328,431	4,438	136,966
TOTAL	**32,599,565**	**45,192,150**	**7,678,488**	**10,494,496**	**24,314,613**	**410,564**	**17,751,547**

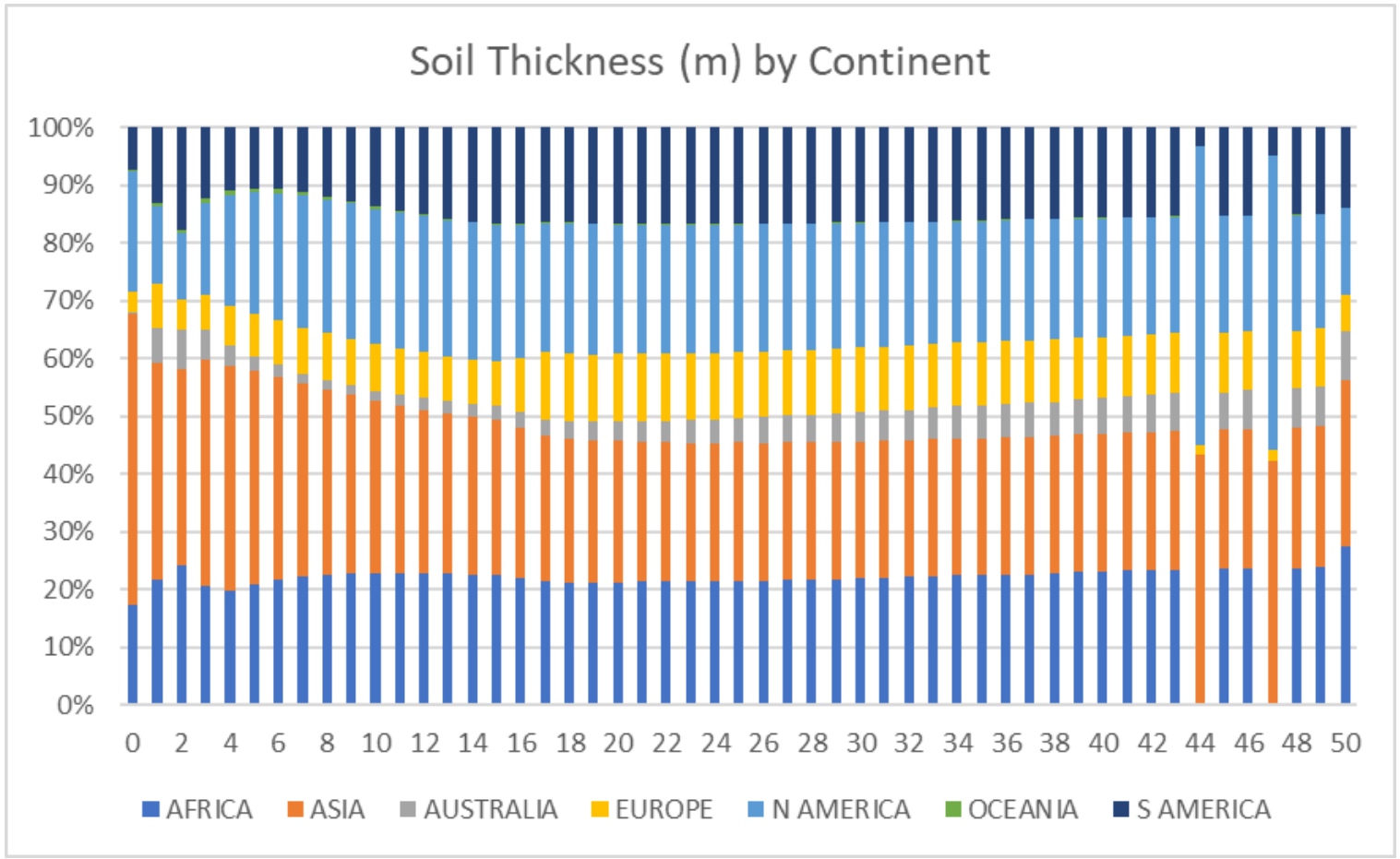
C.11 Soils

C.11.1 Soil Type

	AFRICA	ASIA	AUSTRALIA	EUROPE	No AMERICA	OCEANIA	So AMERICA
SOIL TYPE	**Area (sq km)**	**Area (sq km)**	**Area (sq km)**	**Area (sq km)**	**Area (sq km)**	**Area (sq km)**	**Area (sq km)**
ACRISOLS	856,919	3,963,149	132,005	1,679	1,437,185	32,251	1,938,123
CAMBISOLS	1,054,624	4,391,356	48,678	1,564,837	1,772,861	211,195	347,408
CHERNOZEMS	0	851,327	0	1,006,993	404,003	0	0
PODZOLUVISOLS	1,643,824	2,113,467		1,510,630	52,292	0	0
RENDZINAS	27,699	69,999	0	193,594	151,060	13,208	0
FERRALSOLS	4,203,830	150,413	112,472	0	6,824	16,340	6,117,678
GLEYSOLS	586,558	3,280,335	1,571	193,260	3,597,646	2,040	668,299
PHAEOZEMS	6,691	199,287	30,359	92,491	691,886	1,074	495,653
LITHOSOLS	3,450,274	13,735,854	611,995	703,905	1,941,058	37,546	2,004,454
FLUVISOLS	693,788	1,440,179	0	375,620	135,805	8,090	613,597
KASTAZNOZEMS	25,965	1,813,779	16,596	535,113	1,934,216	0	491,074
LUVISOLS	2,558,189	1,394,243	605,364	1,157,510	2,278,849	17,915	1,155,597
GREYZEMS	0	194,467	0	55,307	52,057	0	0
NITOSOLS	1,141,582	453,172	103,058	0	142,294	143	282,302
HISTOSOLS	16,867	1,120,912	649	282,035	1,031,973	1,463	38,133
PODZOLS	26,701	217,992	80,367	1,838,721	2,453,913	3,523	15,996
ARENOSOLS	4,017,573	744,027	1,269,080	26,276	0	0	784,375
REGOSOLS	1,925,662	930,976	689,115	188,951	3,087,498	12,308	272,880
SOLONETZ	124,470	408,653	379,573	61,695	127,723	0	205,084
ANDOSOLS	47,168	262,329	0	41,258	390,852	62,118	213,805
RANKERS	0	0	0	29,819	0	0	25,888
VERTSOLS	983,145	812,272	868,449	37,382	258,108	399	152,839
PLANOSOLS	438,067	389,395	484,864	199,835	811,478	3,590	710,781
XEROSOLS	1,306,259	1,633,673	725,417	62,429	377,502	0	447,001
YERMOSOLS	4,687,105	3,632,589	1,408,091	0	1,014,523	0	730,328
SOLONCHAKS	160,318	918,552	114,934	15,338	3,505	0	100,616
TOTALS	**29,983,278**	**45,122,397**	**7,682,637**	**10,174,678**	**24,155,111**	**423,203**	**17,811,911**

C.11.2 Soil Thickness

SOIL	AFRICA		ASIA		AUSTRALIA		EUROPE		N AMERICA		OCEANIA		S AMERICA	
THICKNESS (m)	Area SqKm	%	Area SqKm	%	Area SqKm	%	Area SqKm	%	Area SqKm	%	Area SqKm	%	Area SqKm	%
<= 3 m	15,927,869	53.7%	28,502,777	64.8%	4,175,264	54.5%	5,213,407	54.7%	10,387,748	50.7%	319,655	77.4%	9,575,692	54.6%
>3 and <=10	1,514,578	5.1%	2,409,490	5.5%	161,098	2.1%	532,755	5.6%	1,525,066	7.4%	42,123	10.2%	809,340	4.6%
>10 and <= 30	2,737,026	9.2%	3,174,267	7.2%	417,049	5.4%	1,294,069	13.6%	2,825,456	13.8%	23,864	5.8%	2,043,676	11.7%
>30 and <=50	9,486,047	32.0%	9,895,688	22.5%	2,912,552	38.0%	2,496,226	26.2%	5,767,840	28.1%	27,458	6.6%	5,109,649	29.1%



C.11.3 Soil Drainage

Global Soil Drainage		
Classification	**Pixel Count**	**% World**
Very Poor	443,811	2.6%
Poor	2,352,556	14.0%
Imperfectly	3,464,833	20.6%
Moderately Well	8,126,252	48.2%
Well	1,050,419	6.2%
Somewhat Excessive	1,420,806	8.4%
Excessive	0	0.0%

Appendix D Probability Density, Frequency Ratio, Conditional Probability, and Certainty Factor Statistical Model Calculations

D.1 Precipitation, Aridity, Closest Water Body, Distance to Water Bodies

Factor	Factor Class (xx field)	Factor Class Area of Earth Land Surface (sq km) Pr	Factor Class PREDICTOR IMPORTANCE (Pi) Probability Density Class Area/Total Area (Cn/H1 or F3)	Frequency DF & MS per Factor Class	Probability Density DF&MS in Factor Class/TOTAL # DF & MS	FREQUENCY RATIO (FR) Fn/Dn	Conditional Probability (PPa)	CERTAINTY FACTOR (CF) if PPa >= PPs PPa - PPs/PPa (1-PPs) if PPa < PPs
Avg Mthly Precip (mm)		Occurrence						
	<=10	225,342	30.4%	45	2.0%	0.06	0.0001997	0.9226
	>10 and <= 20	105,710	14.3%	54	2.4%	0.165	0.0005108	0.9697
	>20 and <= 30	98,352	13.3%	103	4.5%	0.34	0.0010473	0.9852
	>30 and <= 40	75,339	10.2%	90	3.9%	0.39	0.0011946	0.9871
	>40 and <= 50	71,603	9.7%	109	4.8%	0.49	0.0015223	0.9899
	>50 and <= 60	42,060	5.7%	114	5.0%	0.88	0.0027104	0.9943
	>60 and <= 70	27,090	3.7%	118	5.1%	1.41	0.0043559	0.9965
	>70 and <= 80	21,728	2.9%	105	4.6%	1.56	0.0048325	0.9968
	>80 and <= 90	19,949	2.7%	134	5.8%	2.17	0.0067171	0.9977
	>90 and <= 100	17,685	2.4%	166	7.2%	3.03	0.0093865	0.9984
	>100 and <= 200	14,273	1.9%	786	34.3%	17.80	0.0550690	0.9997
Count:	>200 and <= 300	19,754	2.7%	251	10.9%	4.11	0.0127063	0.9988
14	>300 and <= 400	2,129	0.3%	129	5.6%	19.58	0.0605918	0.9998
	> 400	380	0.1%	90	3.9%	76.54	0.2368421	1.0000
PRECIP SUMS/AVGs		**741,394**	**100.0%**	**2294**	**100.0%**	**9.18**		**0.9883**
Aridity Index (x10,000)			from EU Comm World Atlas of Desertification					
	Hyper-Arid	9,491,200	6.4%	6	0.3%	0.04	0.0000006	-0.9591
	Semi-Arid	20,168,800	13.6%	493	21.5%	1.58	0.0000244	0.3672
	Arid	17,054,500	11.5%	161	7.0%	0.61	0.0000094	-0.3897
	Dry Sub-Humid	8,453,100	5.7%	141	6.1%	1.08	0.0000167	0.0726
Count:	Humid	62,879,200	42.4%	1493	65.1%	1.53	0.0000237	0.3485
6	Cold	30,253,200	20.4%	0	0.0%	0.00	0.0000000	
Aridity SUMS/AVGs		**148,300,000**	**100.0%**	**2294**	**100.0%**	**0.81**		**-0.0934**
Closest H2O Body		**sq km**						
	Intermittant Stream	172,696	0.05%	3	0.1%	2.77	0.00002	0.1095434
Count:	Perennial Stream	13,050,066	3.57%	1839	80.2%	22.47	0.00014	0.8902439
3	Ocean/Sea	352,547,751	96.38%	452	19.7%	0.20	0.00000	-0.9171176
Closest H2O SUMS/AVGs		**365,770,513**	**100.0%**	**2294**	**100.0%**	**8.48**		**0.0275565**
Distance to Ocean/Sea		Ocean/All Water	96.3%					
	<= 1 km	352,237,004	96.3%	256	41.8%	0.43	0.00000073	-0.95
	>1 and <= 5 km	352,237,004	96.3%	148	24.1%	0.25	0.00000042	-0.97
	>5 and <= 10 km	352,237,004	96.3%	65	10.6%	0.11	0.00000018	-0.99
Count=	>10 and <= 20 km	352,237,004	96.3%	76	12.4%	0.13	0.00000022	-0.99
5	> 20 km	352,237,004	96.3%	68	**11.1%**	0.12	0.00000019	-0.99
Ocean/Sea SUMs/AVGs				**613**	**100.0%**	**0.21**		**-0.98**
Distance to Water <> Sea			**3.6%**					
	<= 1 km	469,802	3.6%	102	6.1%	1.69	0.00022	**0.93**
	>1 and <= 5 km	469,802	3.6%	237	14.1%	3.92	0.00050	**0.97**
	>5 and <= 10 km	469,802	3.6%	264	15.7%	4.36	0.00056	0.97
	>10 and <= 20 km	469,802	3.6%	386	23.0%	6.38	0.00082	0.98
Count=	>20 and <= 30 km	469,802	3.6%	217	12.9%	3.59	0.00046	0.97
6	>30 km	469,802	3.6%	475	28.3%	7.85	0.00101	0.98
Water <> Sea SUMs/AVGs				**1681**	**100.0%**	**4.63**		**0.9672**

D.2 Elevation, Slope, Landcover/Land Use, Soil Thickness, Soil Type

Factor	Factor Class (xx field)	Factor Class Area of Earth Land Surface (sq km) Pr	Factor Class PREDICTOR IMPORTANCE (Pi) Probability Density Class Area/Total Area (Cn/H1 or F3)	Frequency DF & MS per Factor Class	Probability Density DF&MS in Factor Class/TOTAL # DF & MS	FREQUENCY RATIO (FR) Fn/Dn	Conditional Probability (PPa)	CERTAINTY FACTOR (CF) if PPa >= PPs PPa - PPs/PPa (1-PPs) if PPa < PPs
Elevation	<null>			55				
	<= 100 m	20,578,000	14.6%	511	22.3%	1.52	0.0000248	0.3771
	>100 and <= 200 m	18,332,800	13.0%	240	10.5%	0.80	0.0000131	-0.1537
	>200 and <= 300 m	15,785,500	11.2%	252	11.0%	0.98	0.0000160	0.0310
	>300 and <= 400 m	12,673,000	9.0%	166	7.2%	0.80	0.0000131	-0.1532
	>400 and <= 500 m	10,151,200	7.2%	111	4.8%	0.67	0.0000109	-0.2931
	>500 and <= 600 m	7,577,900	5.4%	85	3.7%	0.69	0.0000112	-0.2749
	>600 and <= 700 m	5,621,800	4.0%	78	3.4%	0.85	0.0000139	-0.1031
	>700 and <= 800 m	5,025,600	3.6%	71	3.1%	0.87	0.0000141	-0.0867
	>800 and <= 900 m	4,114,600	2.9%	57	2.5%	0.85	0.0000139	-0.1044
	>900 and <= 1000 m	4,312,600	3.1%	66	2.9%	0.94	0.0000153	-0.0106
	>1000 and <= 1500 r	14,087,000	10.0%	208	9.1%	0.91	0.0000148	-0.0455
	>1500 and <= 2000 r	7,686,500	5.5%	173	7.5%	1.38	0.0000225	0.3127
	>2000 and <= 2500 r	3,290,800	2.3%	122	5.3%	2.27	0.0000371	0.5828
	>2500 and <= 3000 r	4,612,500	3.3%	49	2.1%	0.65	0.0000106	-0.3132
	>3000 and <= 3500 r	1,830,400	1.3%	26	1.1%	0.87	0.0000142	-0.0817
	> 3500 and <= 4000	1,755,600	1.2%	19	0.8%	0.66	0.0000108	-0.3004
	> 4000 and <= 4500	1,267,000	0.9%	3	0.1%	0.15	0.0000024	-0.8469
class count:	> 4500 and <= 5000	712,600	0.5%	1	0.0%	0.09	0.0000014	-0.9093
19	> 5000 m	1,282,300	0.9%	1	0.0%	0.05	0.0000008	-0.9496
ELEV SUMS/AVGs		**140,697,700**	**100.0%**	**2239**	**97.6%**	**0.84**		**-0.1749**
Slope (deg)		**Count of Occurrence**						
	<null>			8				
	<= 10	7,565,225,120	96.0%	1478	64.4%	0.67	0.0000002	-0.9874
	>10 & <= 20	219,810,048	2.8%	523	22.8%	8.17	0.0000024	-0.8462
	>20 & <= 30	75,012,212	1.0%	199	8.7%	9.11	0.0000027	-0.8285
	>30 & <= 40	16,660,481	0.2%	71	3.1%	14.64	0.0000043	-0.7245
Count:	>40 & <= 50	1,854,600	0.0%	14	0.6%	25.93	0.0000075	-0.5120
6	> 50	271,109	0.0%	1	0.0%	12.67	0.0000037	-0.7615
SLOPE SUMS/AVGs		**7,878,833,570**	**100.0%**	**2286**	**99.7%**	**11.86**		**-0.7767**
Land Cover Aggregated	per Latham et al							
	Urban	296,600	0.2%	363	15.8%	79.12	0.0012239	0.9874
	Forest & Woodland	43,600,200	29.4%	1242	54.1%	1.84	0.0000285	0.4570
	Cropland	23,283,100	15.7%	283	12.3%	0.79	0.0000122	-0.2142
	Grassland/Shrubland	44,490,000	30.0%	309	13.5%	0.45	0.0000069	-0.5510
	Bare Ground	19,723,900	13.3%	9	0.39%	0.03	0.0000005	-0.9705
	Mangrove			2	0.09%			
Count:	Wetlands/Waterboc	2,521,100	1.7%	83	3.62%	2.13	0.0000329	0.5302
8	Snow / Ice	14,385,100	9.7%	3	0.13%	0.01	0.0000002	-0.9865
Totals/Avgs		**148,300,000**	**100.0%**	**2294**	**99.87%**	**10.54**		**0.0298**
Soil Thickness	**meters**	**Occurrence**						
	0	8,602,978	6.0%	121	5.3%	0.87	0.0000141	-0.0807
	1	62,728,049	44.0%	1222	53.3%	1.21	0.0000195	0.2060
	2	7,530,111	5.3%	222	9.7%	1.83	0.0000295	0.4753
	3	2,514,686	1.8%	199	8.7%	4.91	0.0000791	0.8045
	4	1,604,108	1.1%	111	4.8%	4.30	0.0000692	0.7765
Count:	5	1,288,984	0.9%	60	2.6%	2.89	0.0000465	0.6677
7	6-50	58,193,495	40.8%	359	15.6%	0.38	0.0000062	-0.6012
Soil Thickness SUMS/AVGs		**142,462,411**	**100.0%**	**2294**	**100.0%**	**2.34**		**0.3197**
Soil Type		Approx						
	Acrisols (Ao, Ah)	11,864,000	8.0%	330	24.4%	3.05	0.0000278	0.4439
	Cambisols (Bd, Bh, B	17,796,000	12.0%	398	29.4%	2.45	0.0000224	0.3083
	Lithosol (I)	14,830,000	10.0%	201	14.8%	1.48	0.0000136	-0.1238
	Luvisol (Lc, La)	7,415,000	5.0%	238	17.6%	3.52	0.0000321	0.5181
	Phaeozem (Hh)	1,927,900	1.3%	100	7.4%	5.68	0.0000519	0.7018
Count:	Regosols (Rc, Re)	2,966,000	2.0%	87	6.4%	3.21	0.0000293	0.4727
6	Other (many minor)	91,501,100	61.7%	940	69.4%	1.13	0.0000103	-0.3359
SOILS SUMS/AVGs		**56,798,900**	**38.3%**	**1354**	**100.0%**	**3.42**		**0.3308**

D.3 Landform

Factor	Factor Class (xx field)	Factor Class Area of Earth Land Surface (sq km) Pr	Factor Class PREDICTOR IMPORTANCE (Pi) Probability Density Class Area/Total Area (Cn/H1 or F3)	Frequency DF & MS per Factor Class	Probability Density DF&MS in Factor Class/TOTAL # DF & MS	FREQUENCY RATIO (FR) Fn/Dn	Conditional Probability (PPa)	CERTAINTY FACTOR (CF) if PPa >= PPs PPa - PPs/PPa (1-PPs) if PPa < PPs
Landform	null	13,597,881	0.09	22	0.010	0.10	0.0000016	-0.8954
	Depressions or Basins ir	1,023,209	0.01	1	0.00044	0.06	0.0000010	-0.9368
	Depressions or Basins on Sedimentary Lithology	939,951	0.01	0	0.00000	0.00	0.0000000	
	Dry	6,136	0.00004	2	0.00087	21.07	0.0003259	0.9526
	Dry Depressions or Basins in Alpine System	1,223,009	0.01	13	0.00567	0.69	0.0000106	-0.3128
	Dry High Tablelandsin Alpine System	232,286	0.00	30	0.01308	8.35	0.0001292	0.8802
	Dry Hills or Tablelands on Sedimentary Lithology	6,791,995	0.05	8	0.00349	0.08	0.0000012	-0.9239
	Dry Hills or Tablelandsin Alpine System	1,161,698	0.01	4	0.00174	0.22	0.0000034	-0.7774
	Dry Mountains in Alpine System	2,030,136	0.01	222	0.097	7.07	0.0001094	0.8586
	Dry Plains in Alpine System	867,898	0.01	3	0.001	0.22	0.0000035	-0.7765
	Dry Plains on Sedimentary Lithology	5,414,018	0.04	10	0.004	0.12	0.0000018	-0.8806
	Dry Plains on Shield	20,317	0.00	0	0.000	0.00	0.0000000	
	Dry Widely Spaced Mountains in Alpine	2,006,612	0.01	62	0.027	2.00	0.0000309	0.4994
	Dry Widely Spaced Mou	33,782	0.00	1	0.000	1.91	0.0000296	0.4774
	DryPlains on Sedimentary Lithology	1,254,926	0.01	0	0.000	0.00	0.0000000	
	Hills in Alpine System	17,828	0.00	0	0.000	0.00	0.0000000	
	Hills or Tablelands on Se	4,938,424	0.03	1	0.000	0.01	0.0000002	-0.9869
	Hills or Tablelandsin Alpine System	15,173,709	0.10	119	0.052	0.51	0.0000078	-0.4930
	Humid	392,736	0.00	0	0.000	0.00	0.0000000	
	Humid Depressions or Basins in Alpine System	6,040,808	0.04	19	0.008	0.20	0.0000031	-0.7967
	Humid Hills or Tablelands on Sedimentary Lithology	3,499,916	0.02	135	0.059	2.49	0.0000386	0.5990
	Humid Hills or Tablelandsin Alpine	10,420,469	0.07	60	0.026	0.37	0.0000058	-0.6278
	Humid Mountains in Alpine System	11,592,742	0.08	914	0.398	5.10	0.0000788	0.8038
	Humid Mountains on Sedimentary Lithology	205,585	0.00	82	0.036	25.79	0.0003989	0.9612
	Humid Plains	21	0.00	0	0.000	0.00	0.0000000	
	Humid Plains in Alpine System	2,698,149	0.02	21	0.009	0.50	0.0000078	-0.4968
	Humid Plains on Sedimentary Lithology	15,280,805	0.10	70	0.031	0.30	0.0000046	-0.7039
	Humid Widely Spaced Mountains in Alpine	262,228	0.00	7	0.003	1.73	0.0000267	0.4205
	in Alpine System	658,119	0.00	0	0.000	0.00	0.0000000	
	Long Term Snow or Icein Alpine System	955,745	0.01	0	0.000	0.00	0.0000000	
	Mountains in Alpine	20,598,908	0.14	401	0.175	1.26	0.0000195	0.2054
	Mountains on Sedimen	4,520,414	0.03	1	0.000	0.01	0.0000002	-0.9857
	on Sedimentary Lithology	548,906	0.00	0	0.000	0.00	0.0000000	
	Plains in Alpine System	5,269,229	0.04	44	0.019	0.54	0.0000084	-0.4602
	Plains on Sedimentary Lithology	5,700,728	0.04	27	0.012	0.31	0.0000047	-0.6938
class count:	Widely Spaced Mountains in Alpine System	1,708,549	0.01	15	0.007	0.57	0.0000088	-0.4324
36	Widely Spaced Mountains on Sedimentary Lithology	121,110	0.001	0	0.000	0.00	0.0000000	
LANDFORM SUMS/AVGs		**147,208,983**	**99.3%**	**2,294**	**100.00%**	**2.27**		**-0.1534**

D.4 Closest "Active Faults" Type, Closest Plate Boundary Type, Fault Density

Factor	Factor Class (xx field)	Factor Class Area of Earth Land Surface (sq km) Pr	Factor Class PREDICTOR IMPORTANCE (Pi) Probability Density Class Area/Total Area (Cn/H1 or F3)	Frequency DF & MS per Factor Class	Probability Density DF&MS in Factor Class/TOTAL # DF & MS	FREQUENCY RATIO (FR) Fn/Dn	Conditional Probability (PPa)	CERTAINTY FACTOR (CF) if PPa >= PPs PPa - PPs/PPa (1-PPs) if PPa < PPs
Closest "Active Faults" Type		**Linear km**						
	<null>	26,373	2.7%	25	1.1%	0.401	0.0009479	0.9837
	Anticline	23,829	2.5%	36	1.6%	0.639	0.0015108	0.9898
	Blind Thrust	42	0.0%	0	0.0%	0.000	0.0000000	
	Dextral	120,425	12.4%	280	12.2%	0.984	0.0023251	0.9934
	Dextral-Normal	14,924	1.5%	9	0.4%	0.255	0.0006031	0.9744
	Dextral-Oblique	67	0.0%	0	0.0%	0.000	0.0000000	
	Dextral-Reverse	17,508	1.8%	13	0.6%	0.314	0.0007425	0.9792
	Dextral_Transform	33,024	3.4%	10	0.4%	0.128	0.0003028	0.9489
	Normal	120,819	12.4%	579	25.2%	2.028	0.0047923	0.9968
	Normal-Dextral	1,390	0.1%	3	0.1%	0.913	0.0021576	0.9928
	Normal-Sinistral	889	0.1%	1	0.0%	0.476	0.0011243	0.9863
	Normal-Strike-Slip	45	0.0%	0	0.0%	0.000	0.0000000	
	Reverse	198,473	20.4%	1035	45.1%	2.207	0.0052148	0.9970
	Reverse-Dextral	7,012	0.7%	13	0.6%	0.785	0.0018540	0.9917
	Reverse-Sinistral	2,422	0.2%	3	0.1%	0.524	0.0012386	0.9875
	Reverse-Strike-Slip	6,261	0.6%	17	0.7%	1.149	0.0027153	0.9943
	Sinistral	100,131	10.3%	140	6.1%	0.592	0.0013982	0.9890
	Sinistral-Normal	11,269	1.2%	7	0.3%	0.263	0.0006212	0.9751
	Sinistral-Reverse	15,215	1.6%	41	1.8%	1.141	0.0026947	0.9943
	Sinistral_Transform	38,436	4.0%	26	1.1%	0.286	0.0006764	0.9771
	Spreading_Ridge	110,108	11.3%	14	0.6%	0.054	0.0001271	0.8784
	Strike-Slip	11,943	1.2%	10	0.4%	0.354	0.0008373	0.9815
Count=	Subduction_Thrust	109,460	11.3%	32	1.4%	0.124	0.0002923	0.9471
24	Syncline	849	0.1%		0.0%	0.000	0.0000000	
Active Faults SUMs/AVGs		**970,914**	**100.0%**	**2294**	**100.0%**	**0.57**		**0.8149**
Closest "Plate Boundary" Type		Linear km						
	CCB	23,003	8.8%	521	22.7%	2.57	0.0226488	0.9993
	CRB	27,471	10.5%	160	7.0%	0.66	0.0058244	0.9974
	CTF	26,174	10.0%	569	24.8%	2.47	0.0217388	0.9993
	OCB	17,447	6.7%	17	0.7%	0.11	0.0009744	0.9841
	OSR	67,335	25.8%	58	2.5%	0.10	0.0008614	0.9821
Count=	OTF	47,894	18.4%	369	16.1%	0.88	0.0077045	0.9980
7	SUB	51,470	19.7%	600	26.2%	1.33	0.0116573	0.9987
Plate Boundary SUMs/AVGs		**260,795**	**100.0%**	**2294**	**100.0%**	**1.16**		**0.9941**
FaultDensity km/sq km		using 5km cell, 25 km search						
0	0.009691	49,706,457	94.86%	1044	45.5%	0.48		
0.009691	0.03015	2,003,088	3.82%	492	21.4%	5.61		
0.03015	0.05384	491,268	0.94%	300	13.1%	13.95		
0.05384	0.087221	157,207	0.30%	263	11.5%	38.21		
0.087221	0.274585	39,788	0.08%	195	8.5%	111.94		

Appendix E Maximum Entropy Results

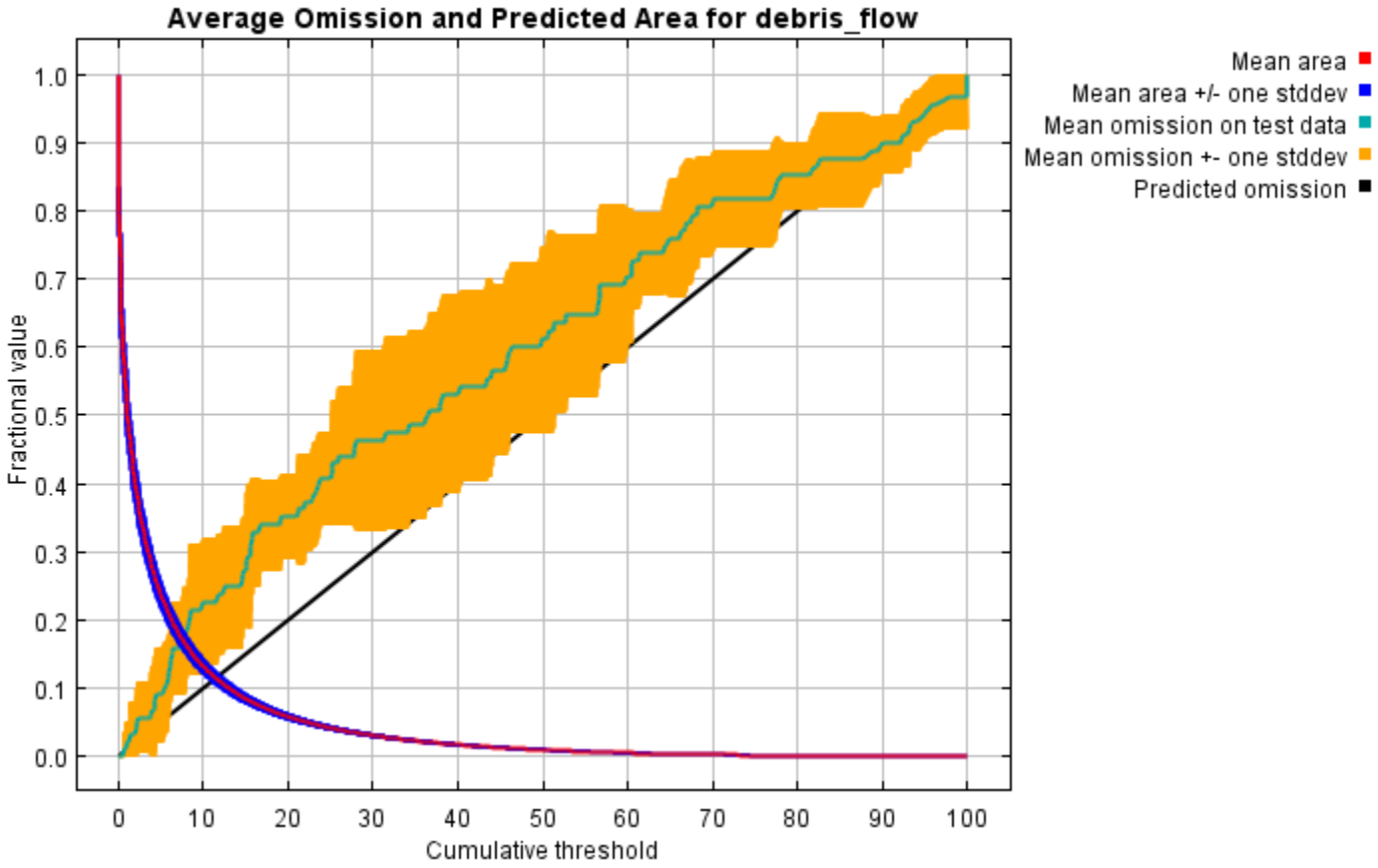
E.1 Global

Replicated maxent model for debris_flow “LTE 5 km”

This page summarizes the results of 5-fold cross-validation for debris_flow, created Sat Sep 25 11:48:47 CEST 2021 using Maxent version 3.4.4. The individual models are here: [0] [1] [2] [3] [4]

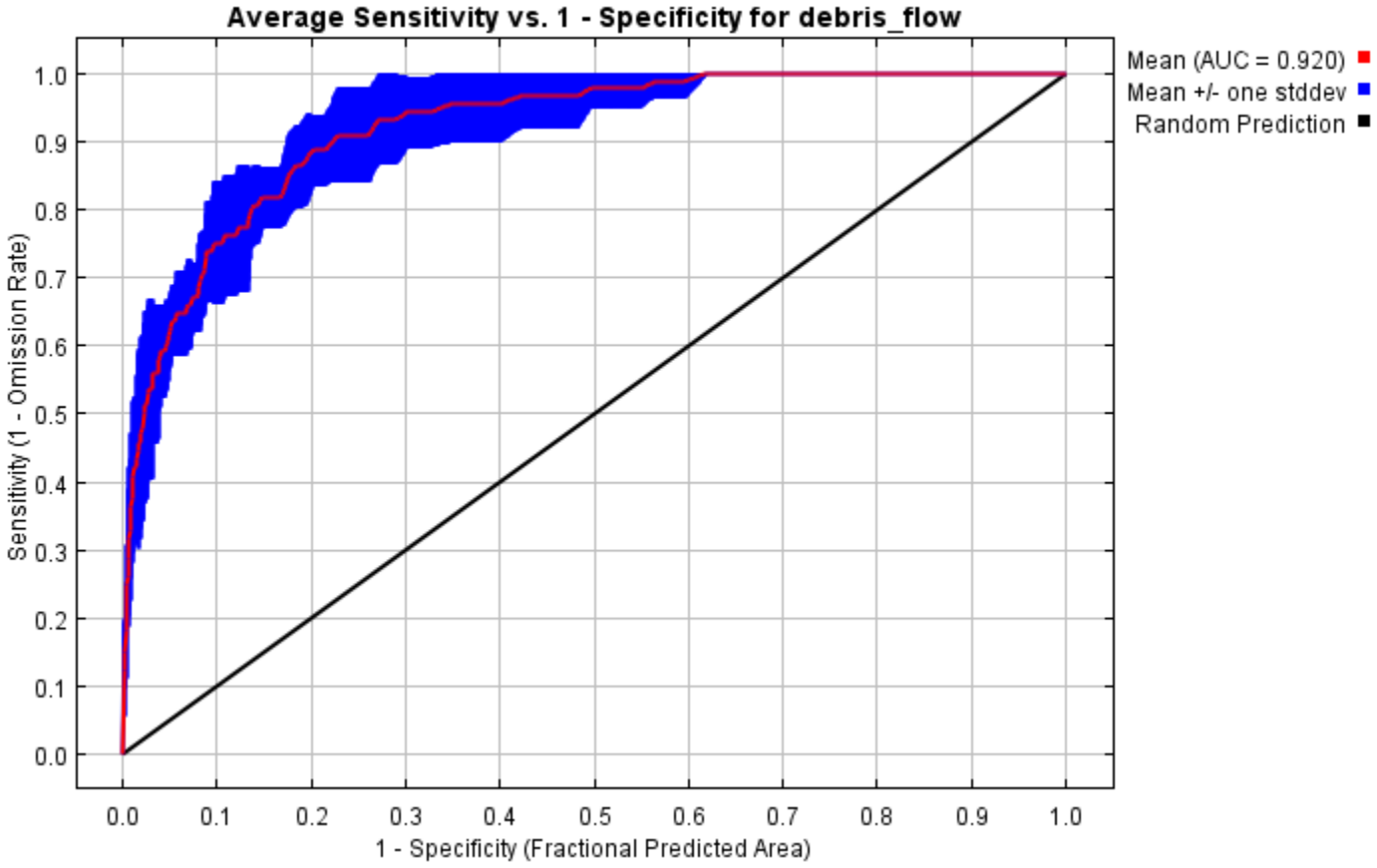
Analysis of omission/commission

The following picture shows the test omission rate and predicted area as a function of the cumulative threshold, averaged over the replicate runs. The omission rate should be close to the predicted omission, because of the definition of the cumulative threshold.



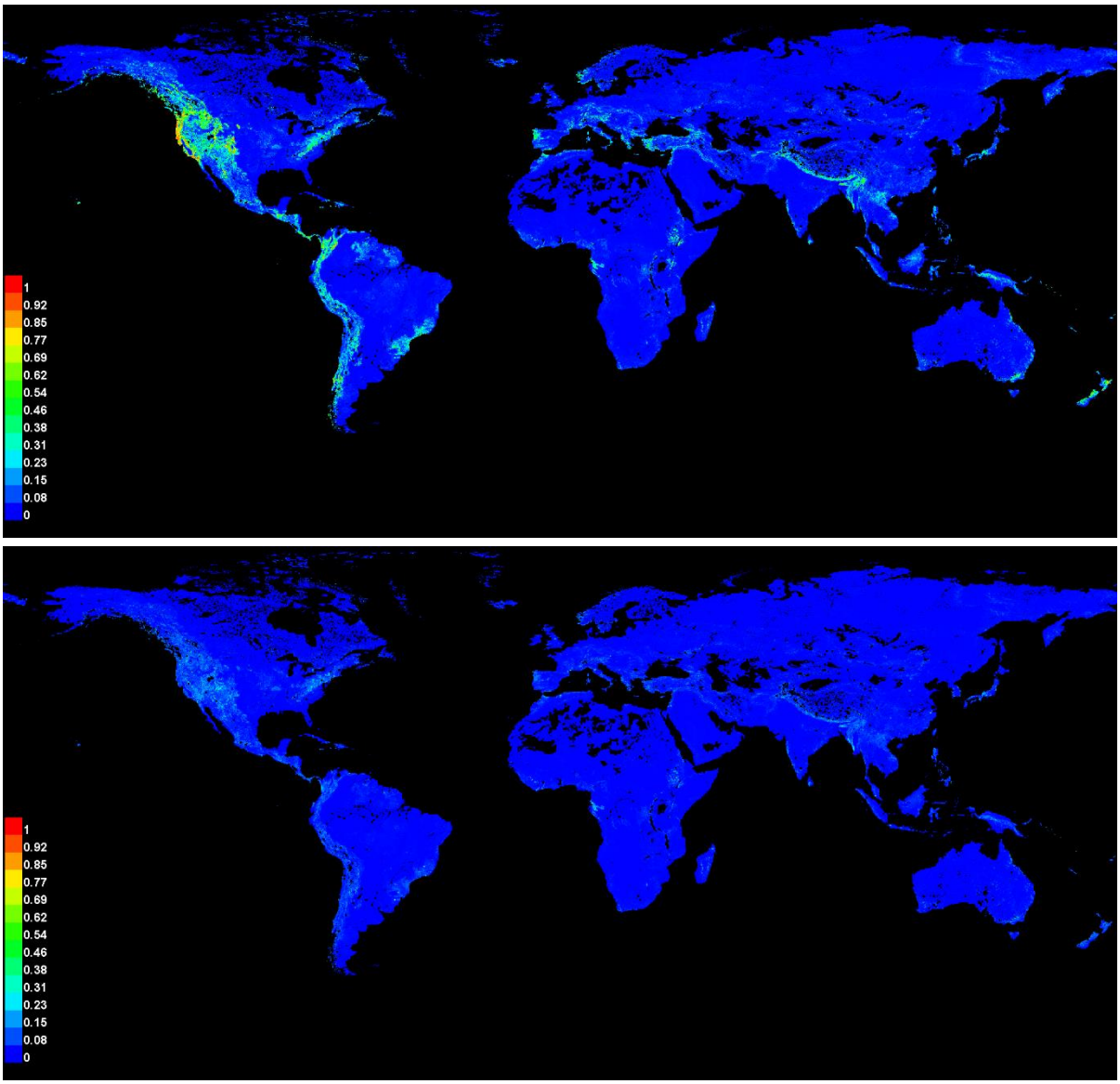
The next picture is the receiver operating characteristic (ROC) curve for the same

data, again averaged over the replicate runs. Note that the specificity is defined using predicted area, rather than true commission (see the paper by Phillips, Anderson and Schapire cited on the help page for discussion of what this means). The average test AUC for the replicate runs is 0.920, and the standard deviation is 0.022.



Pictures of the model

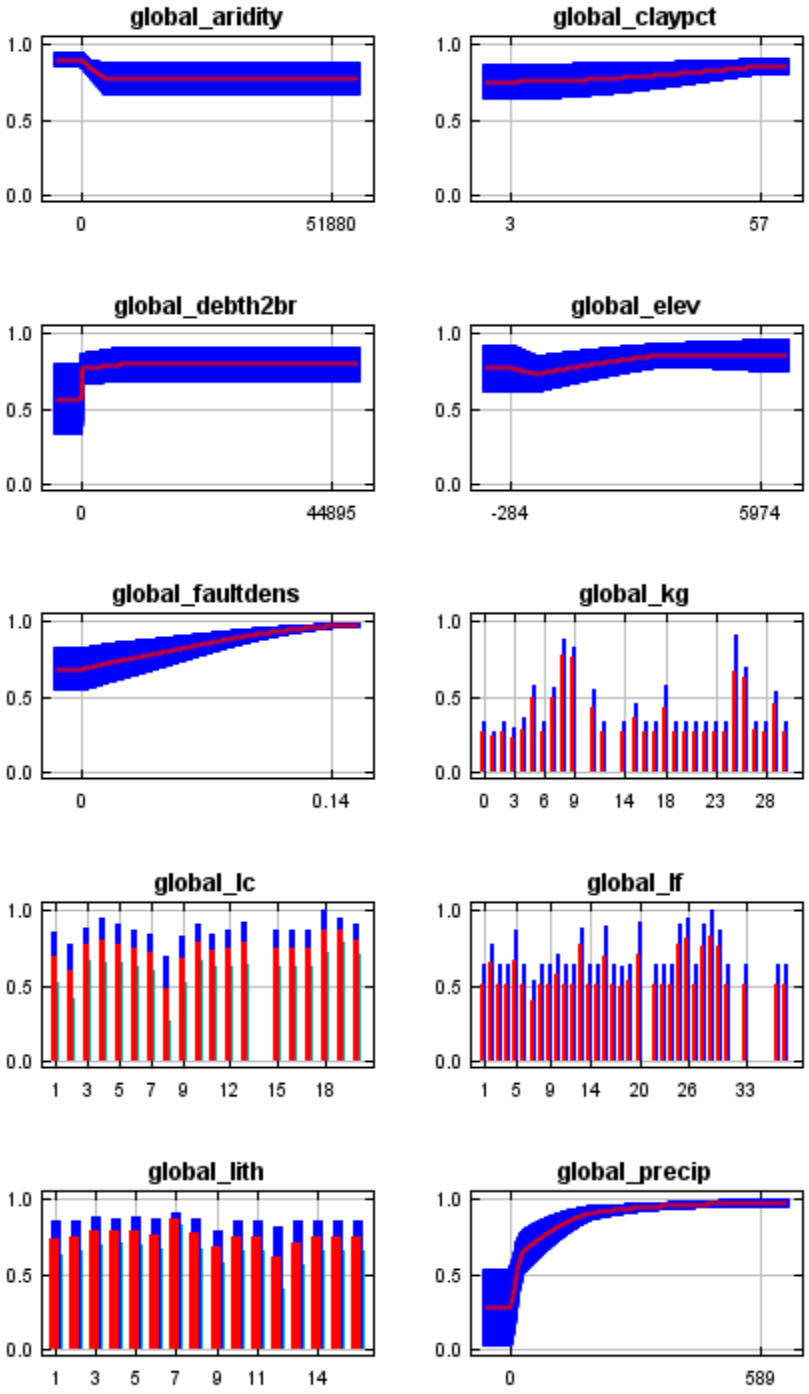
The following two pictures show the point-wise mean and standard deviation of the 5 output grids. Other available summary grids are min, max and median.

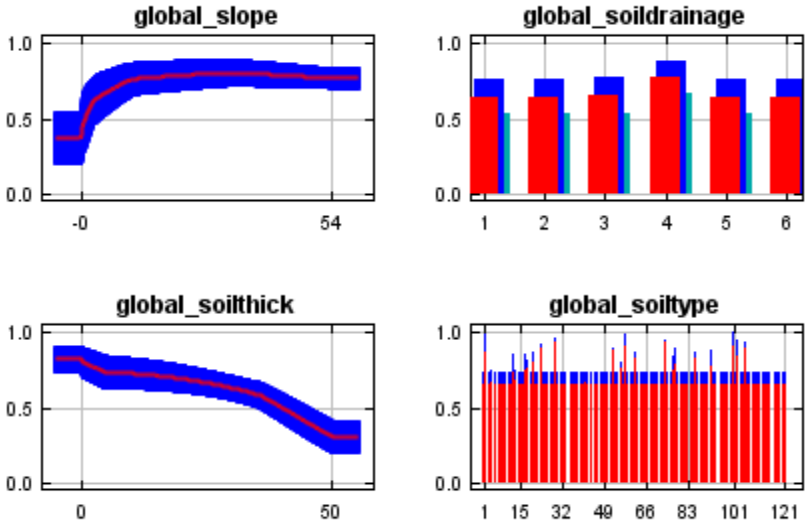


Response curves

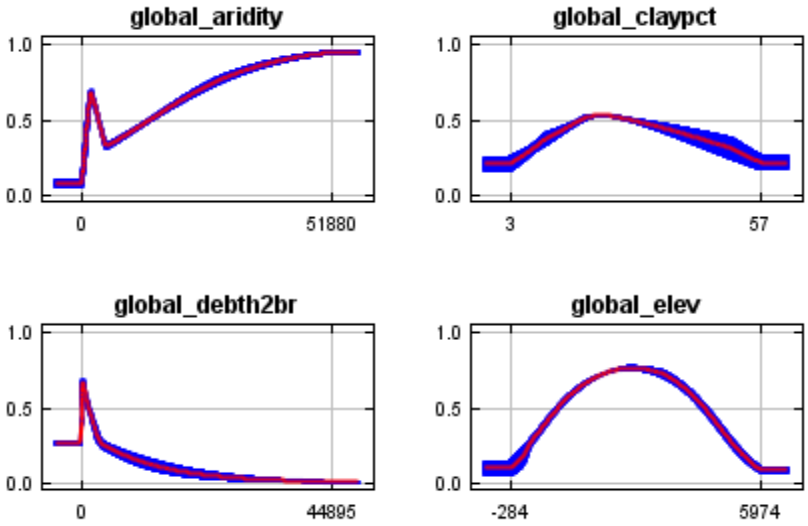
These curves show how each environmental variable affects the Maxent prediction. The curves show how the predicted probability of presence changes as each environmental variable is varied, keeping all other environmental variables at their average sample value. Click on a response curve to see a larger version. Note that the curves can be hard to interpret if you have strongly correlated variables, as the model may depend on the correlations in ways that are not evident in the curves. In other words, the curves show the marginal effect of changing exactly one variable, whereas the model may take advantage of sets of variables changing together. The curves show the mean response of the 5 replicate Maxent runs (red) and and the mean +/- one

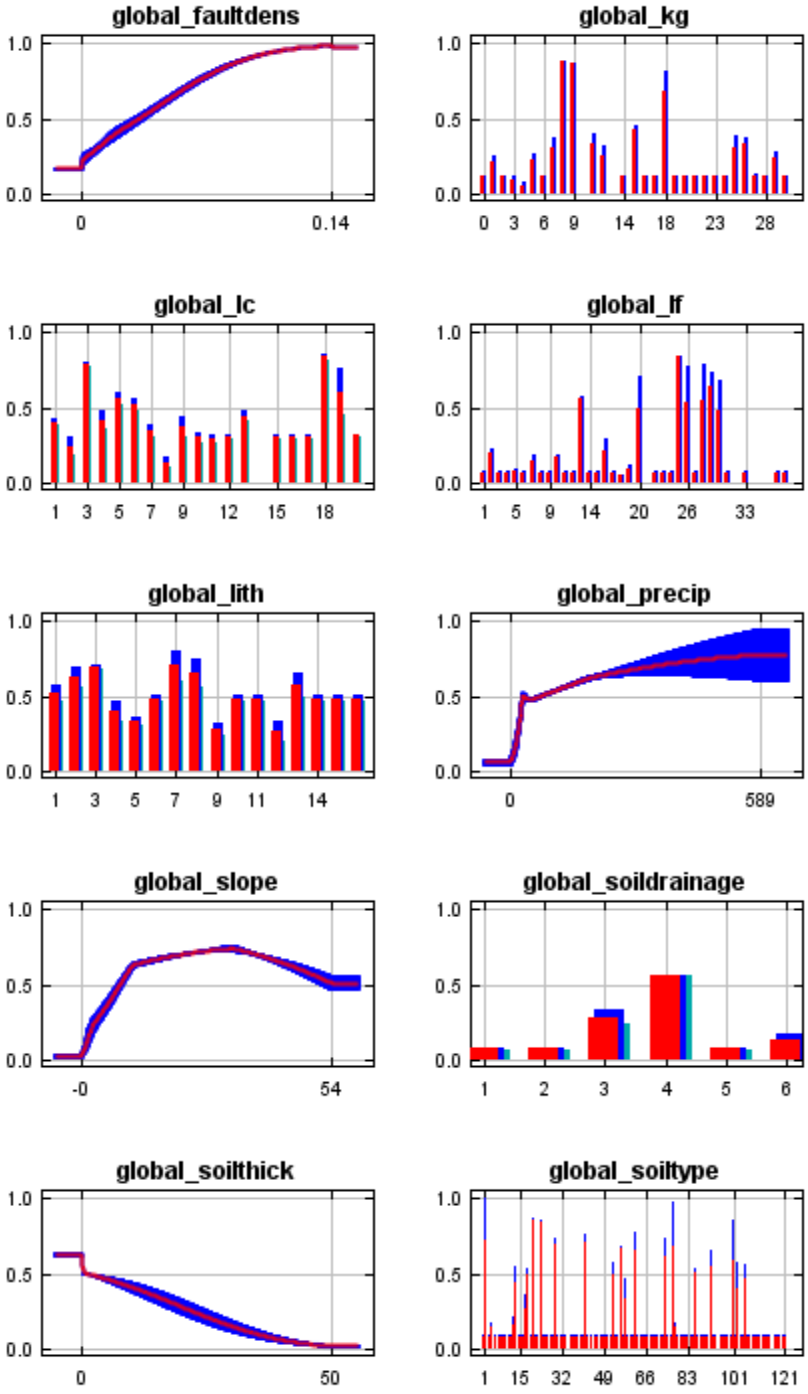
standard deviation (blue, two shades for categorical variables).





In contrast to the above marginal response curves, each of the following curves represents a different model, namely, a Maxent model created using only the corresponding variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other variables. They may be easier to interpret if there are strong correlations between variables.





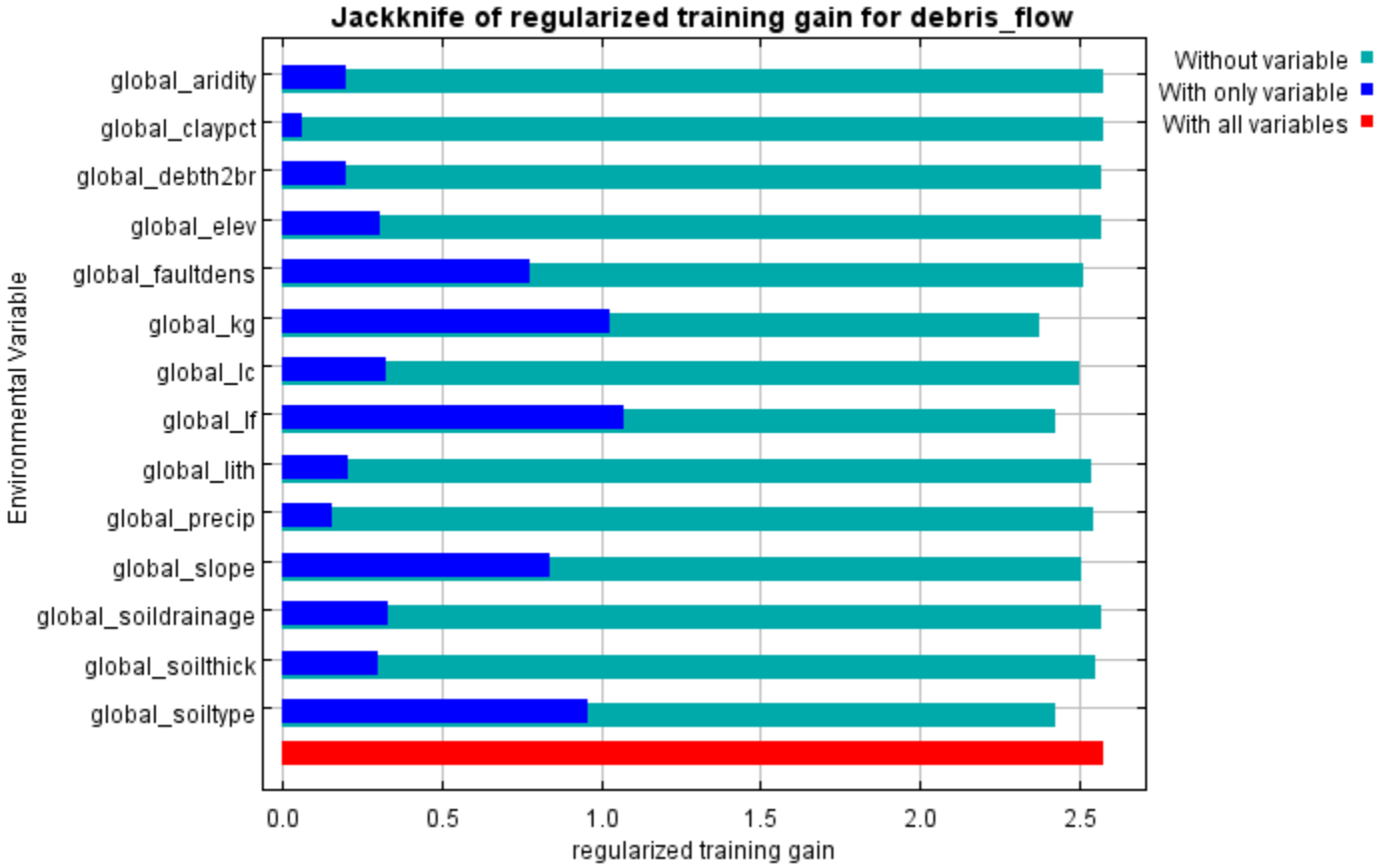
Analysis of variable contributions

The following table gives estimates of relative contributions of the environmental variables to the Maxent model. To determine the first estimate, in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the

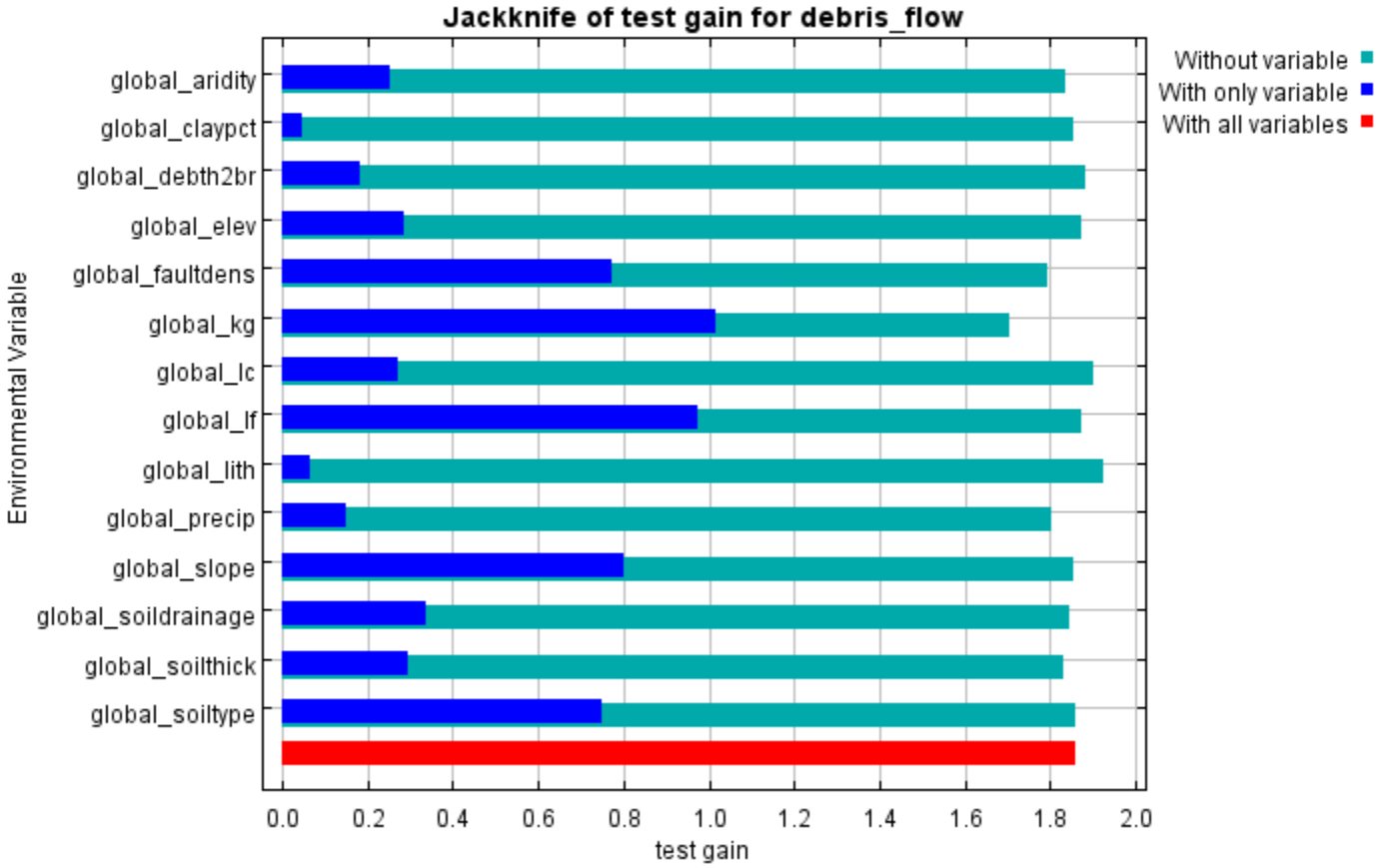
corresponding variable, or subtracted from it if the change to the absolute value of lambda is negative. For the second estimate, for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is reevaluated on the permuted data, and the resulting drop in training AUC is shown in the table, normalized to percentages. As with the variable jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated. Values shown are averages over replicate runs.

Variable	Percent contribution	Permutation importance
global_slope	26.6	8.4
global_kg	20.2	21.8
global_lf	18.3	11.3
global_soiltype	9.4	10.5
global_soildrainage	9	1.8
global_faultdens	5.6	4
global_lc	4.3	7.9
global_lith	2.2	2.5
global_soilthick	1.9	8.6
global_precip	1.6	18.6
global_aridity	0.4	1.7
global_elev	0.3	1
global_debth2br	0.2	1.6
global_claypct	0	0.2

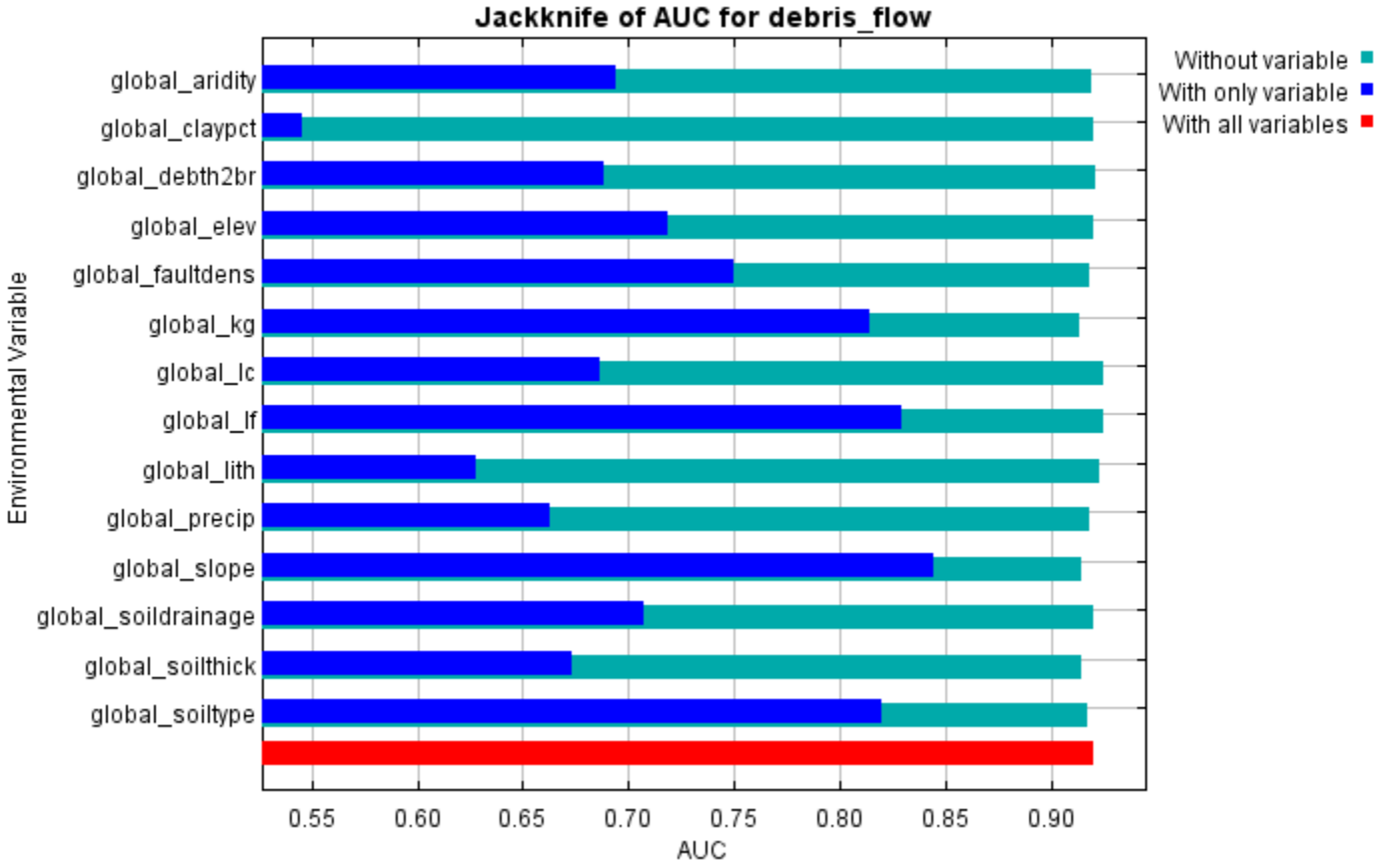
The following picture shows the results of the jackknife test of variable importance. The environmental variable with highest gain when used in isolation is global_lf, which therefore appears to have the most useful information by itself. The environmental variable that decreases the gain the most when it is omitted is global_kg, which therefore appears to have the most information that isn't present in the other variables. Values shown are averages over replicate runs.



The next picture shows the same jackknife test, using test gain instead of training gain. Note that conclusions about which variables are most important can change, now that we're looking at test data.



Lastly, we have the same jackknife test, using AUC on test data.



Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E debris_flow responsecurves jackknife outputformat=logistic "outputdirectory=C:_PhD_PhD PROJECT\MaxEntropy_Stats\MaxEntResults\GlobalAnalysis\Global_DF_TRAIN_LTE5km_TEST1" "samplesfile=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EventData_samples\TRAIN_W

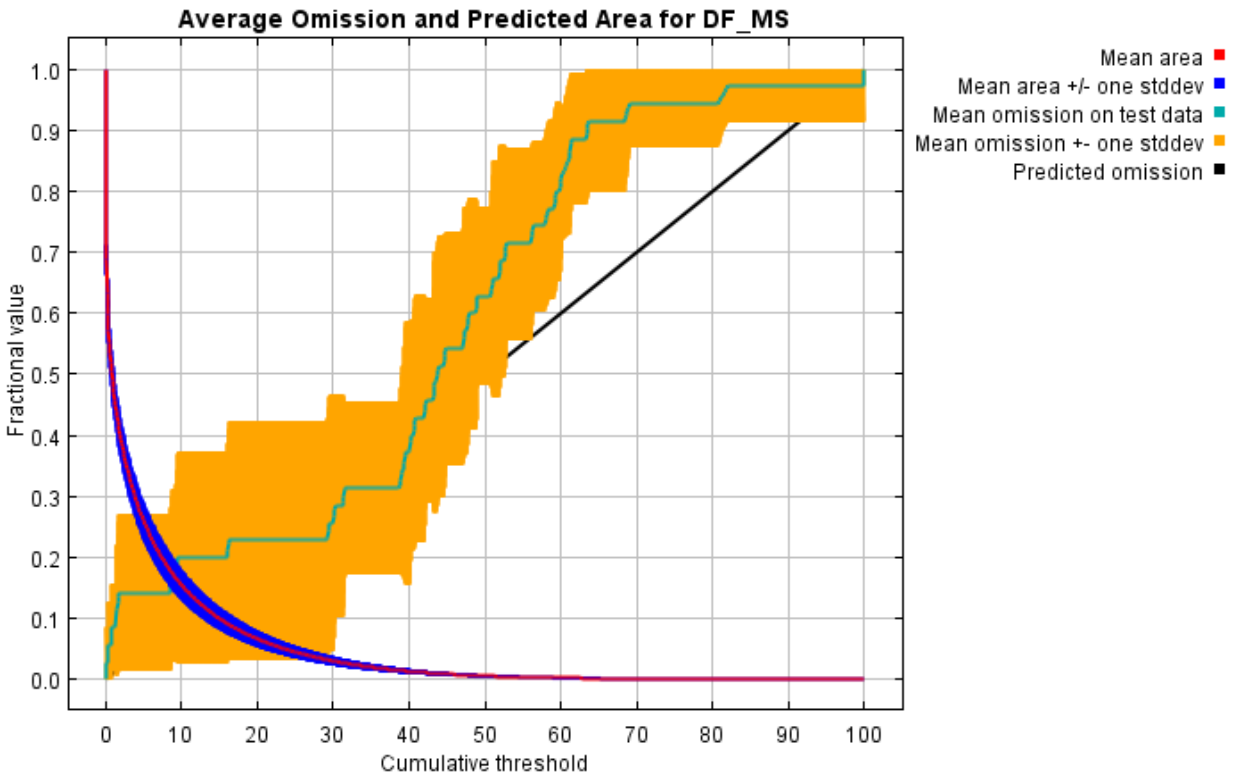
E.2 Africa

Replicated maxent model for DF_MS

This page summarizes the results of 5-fold cross-validation for DF_MS, created Tue Jun 01 19:15:31 CEST 2021 using Maxent version 3.4.4. The individual models are here: [0] [1] [2] [3] [4]

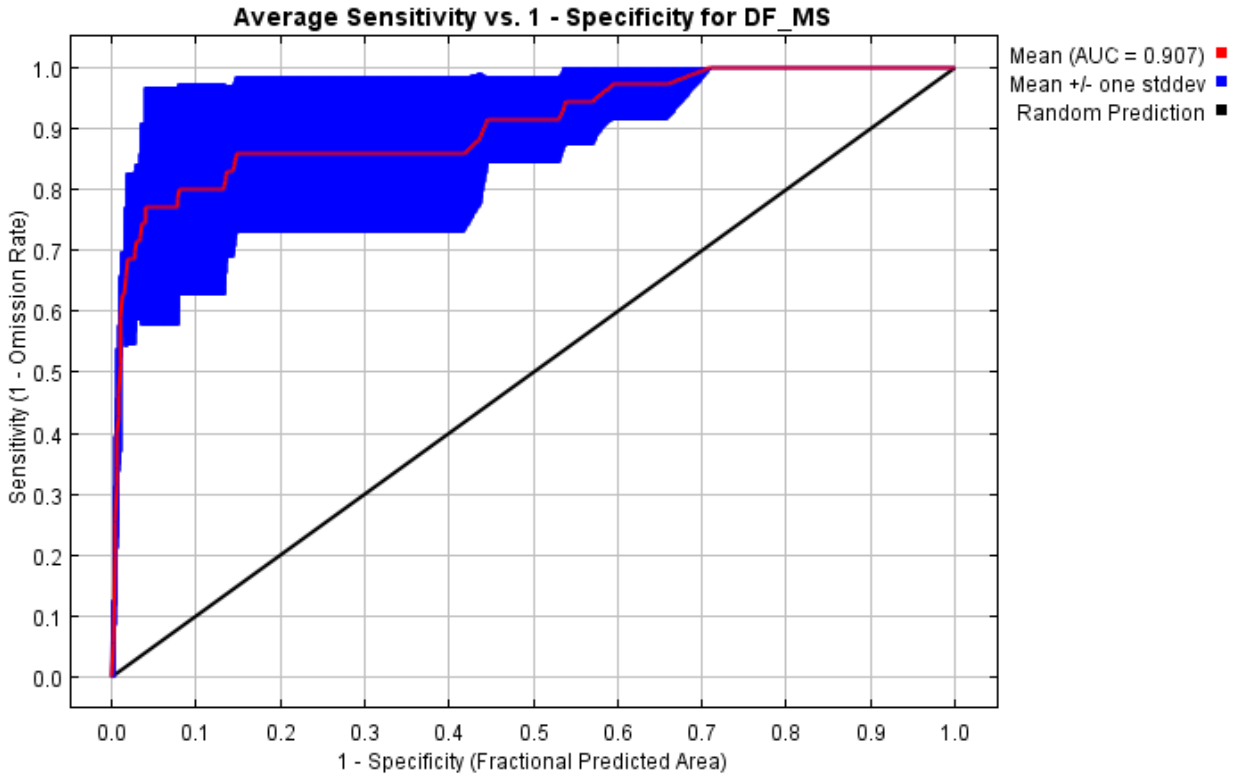
Analysis of omission/commission

The following picture shows the test omission rate and predicted area as a function of the cumulative threshold, averaged over the replicate runs. The omission rate should be close to the predicted omission, because of the definition of the cumulative threshold.



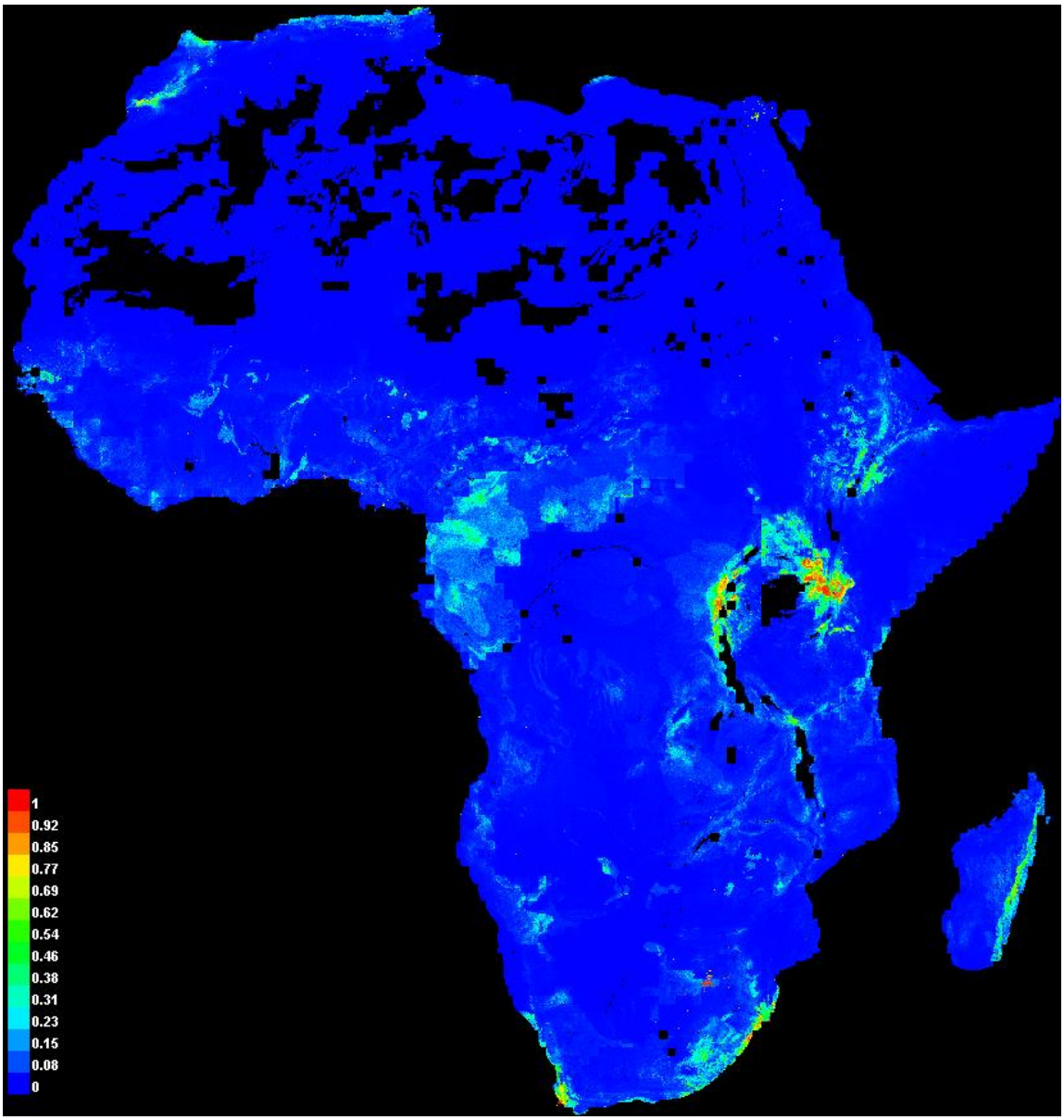
The next picture is the receiver operating characteristic (ROC) curve for the same data, again averaged over the replicate runs. Note that the specificity is defined using predicted area, rather than true commission (see the paper by Phillips, Anderson and Schapire cited on the help page for discussion of what this means). The average test AUC for the replicate runs is 0.907, and the standard deviation is

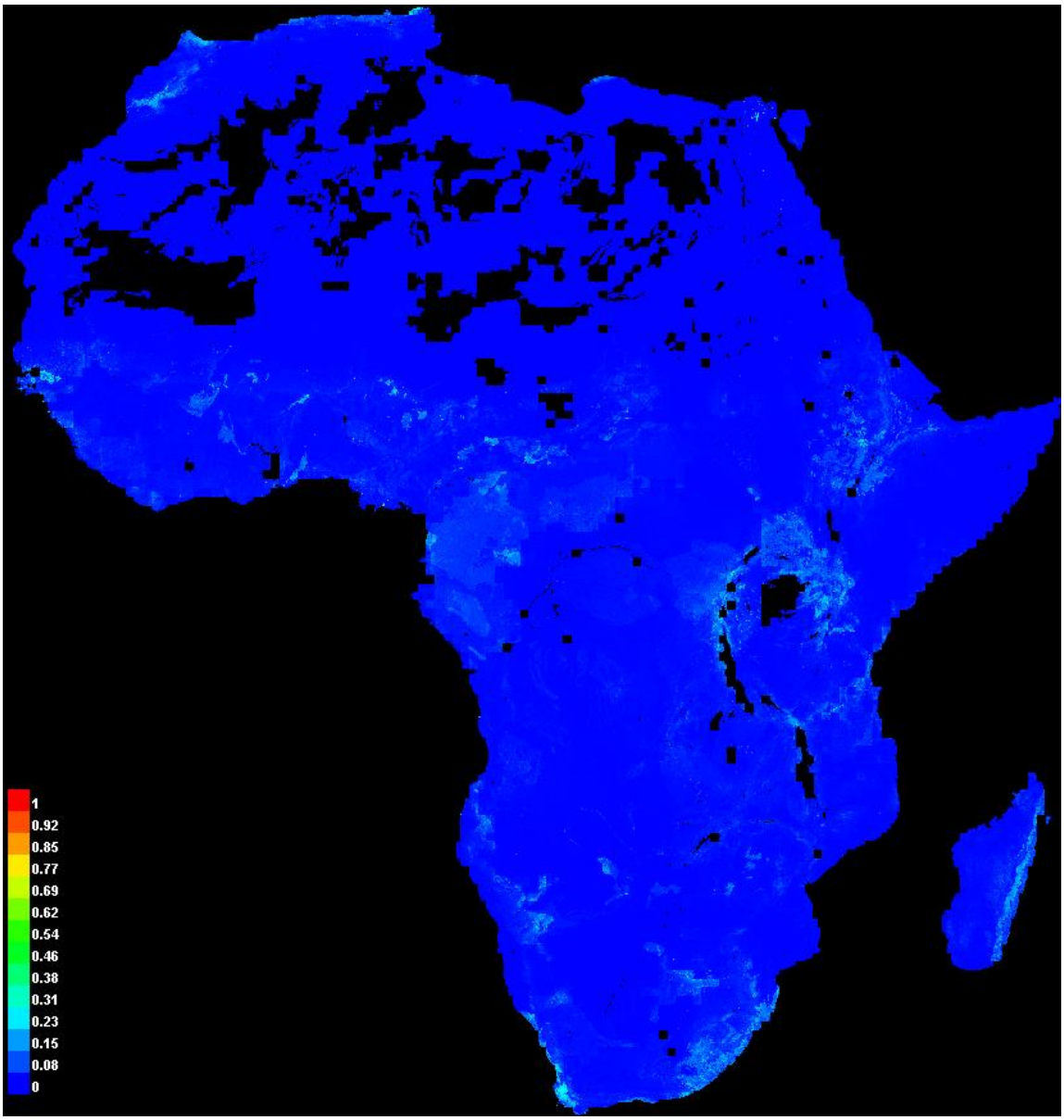
0.067.



Pictures of the model

The following two pictures show the point-wise mean and standard deviation of the 5 output grids. Other available summary grids are min, max and median.

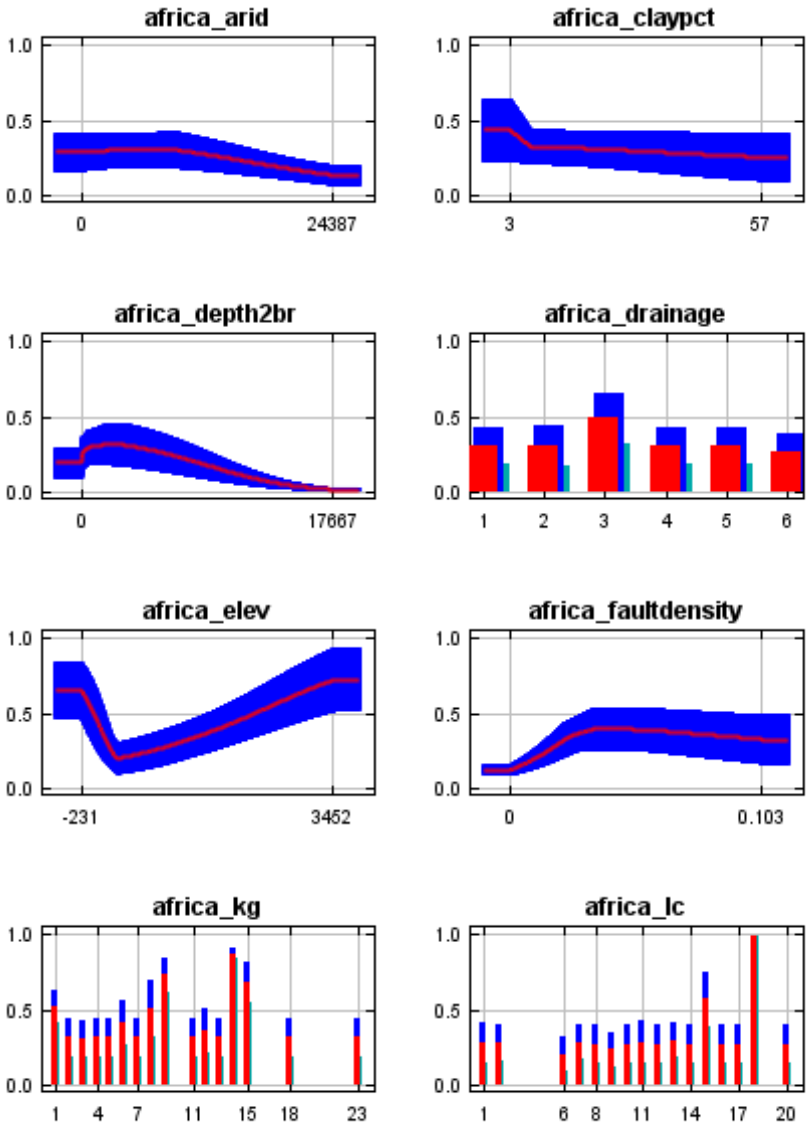

1
0.92
0.85
0.77
0.69
0.62
0.54
0.46
0.38
0.31
0.23
0.15
0.08
0

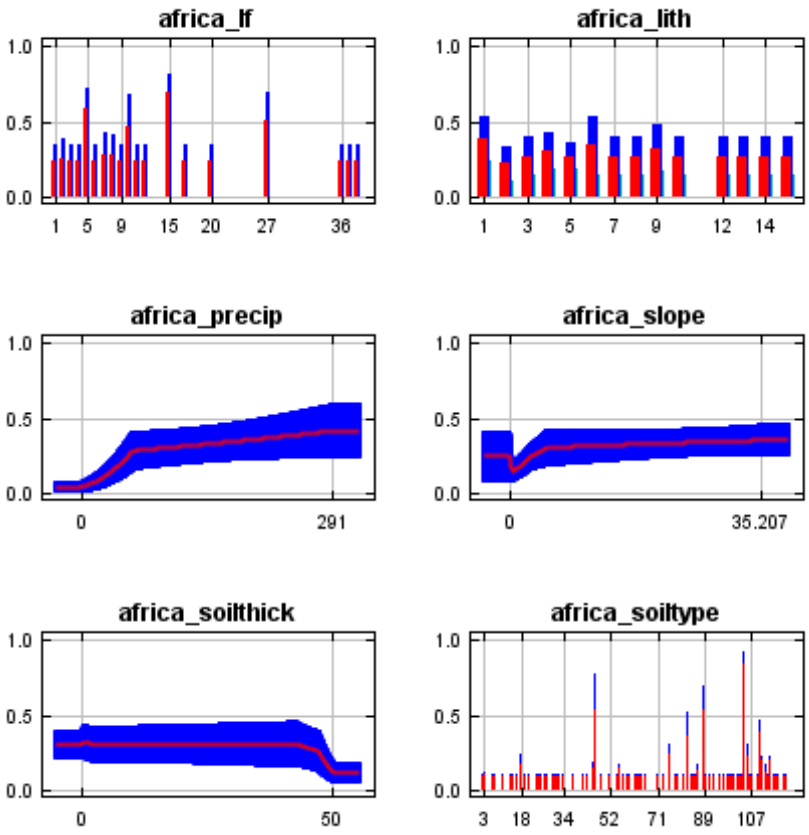


Response curves

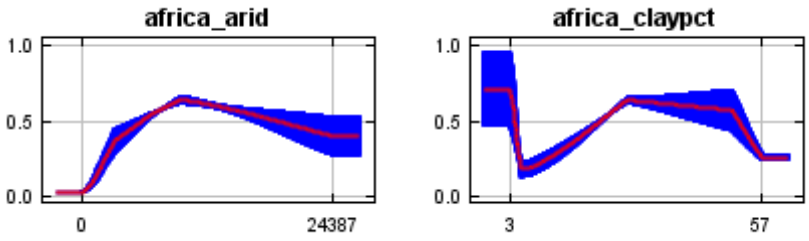
These curves show how each environmental variable affects the Maxent prediction. The curves show how the predicted probability of presence changes as each environmental variable is varied, keeping all other environmental variables at their average sample value. Click on a response curve to see a larger version. Note that the curves can be hard to interpret if you have strongly correlated variables, as the

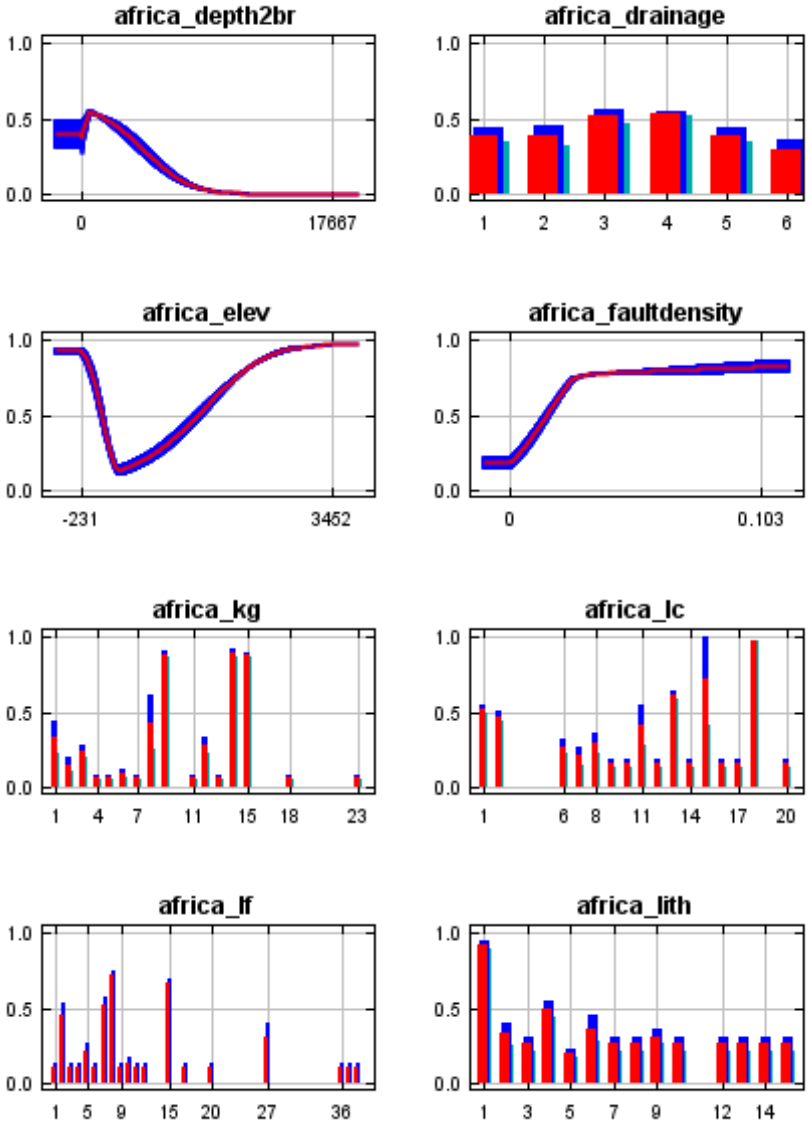
model may depend on the correlations in ways that are not evident in the curves. In other words, the curves show the marginal effect of changing exactly one variable, whereas the model may take advantage of sets of variables changing together. The curves show the mean response of the 5 replicate Maxent runs (red) and and the mean +/- one standard deviation (blue, two shades for categorical variables).

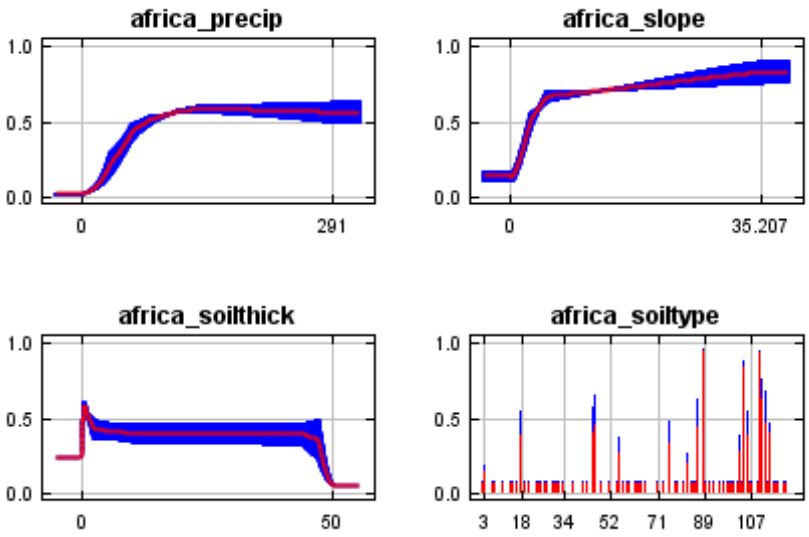




In contrast to the above marginal response curves, each of the following curves represents a different model, namely, a Maxent model created using only the corresponding variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other variables. They may be easier to interpret if there are strong correlations between variables.




africa_depth2br
1.0
0.5
0.0
0
17667
africa_drainage
1.0
0.5
0.0
1
2
3
4
5
6
africa_elev
1.0
0.5
0.0
-231
3452
africa_faultdensity
1.0
0.5
0.0
0
0.103
africa_kg
1.0
0.5
0.0
1
4
7
11
15
18
23
africa_lc
1.0
0.5
0.0
1
6
8
11
14
17
20
africa_lf
1.0
0.5
0.0
1
5
9
15
20
27
36
africa_lith
1.0
0.5
0.0
1
3
5
7
9
12
14



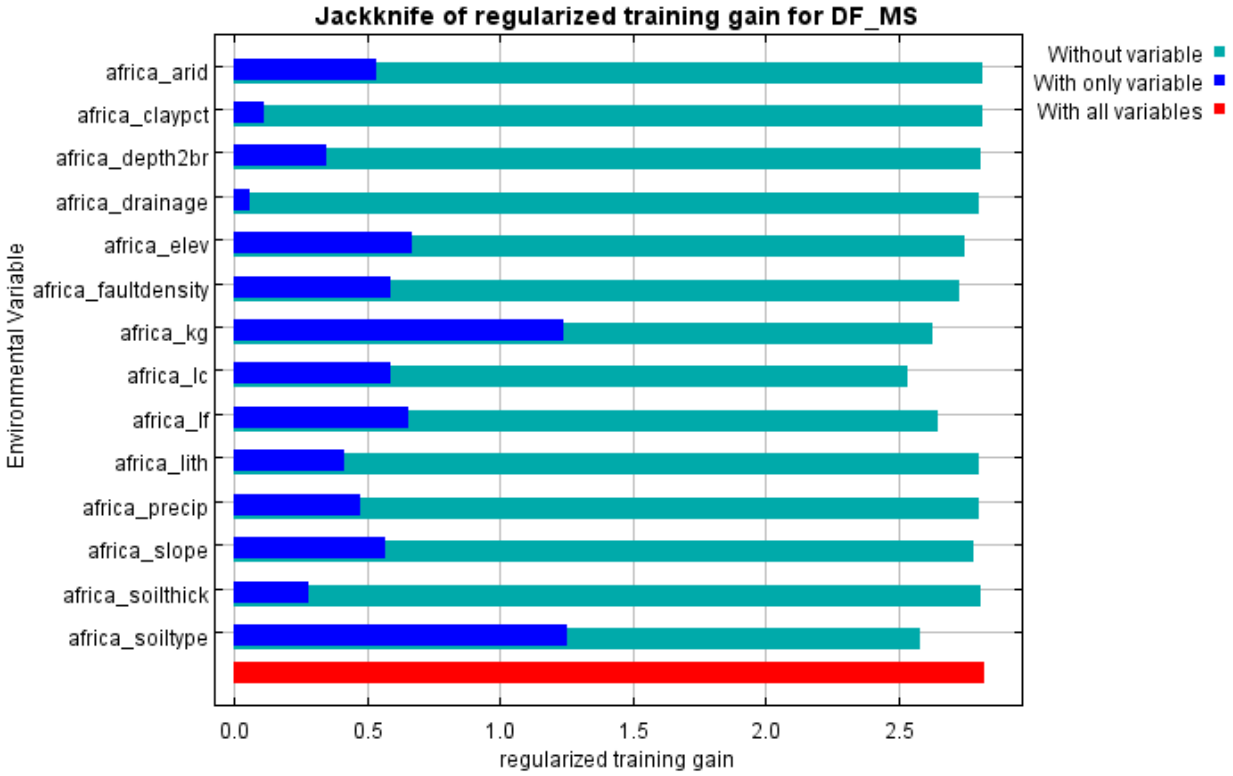
Analysis of variable contributions

The following table gives estimates of relative contributions of the environmental variables to the Maxent model. To determine the first estimate, in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the corresponding variable, or subtracted from it if the change to the absolute value of lambda is negative. For the second estimate, for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is reevaluated on the permuted data, and the resulting drop in training AUC is shown in the table, normalized to percentages. As with the variable jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated. Values shown are averages over replicate runs.

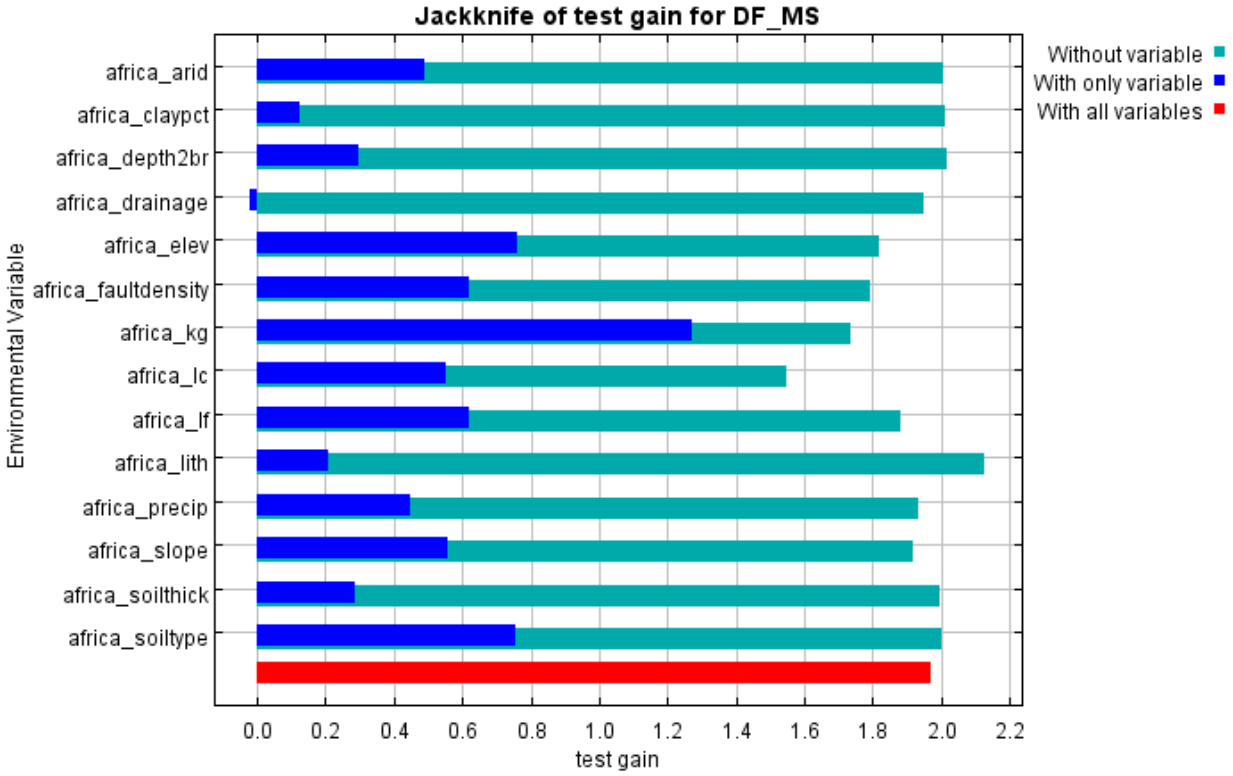
Variable	Percent contribution	Permutation importance
africa_kg	19.3	8.3
africa_soiltype	18.1	18.1

africa_faultdensity	14.1	3
africa_lc	13.2	29.6
africa_arid	12.9	2
africa_lf	5.9	11.7
africa_slope	5.6	3.4
africa_elev	4.5	4.6
africa_lith	2.3	5.1
africa_soilthick	1.5	3.4
africa_precip	1.4	5.6
africa_drainage	0.8	1.7
africa_depth2br	0.3	1.5
africa_claypct	0.2	2

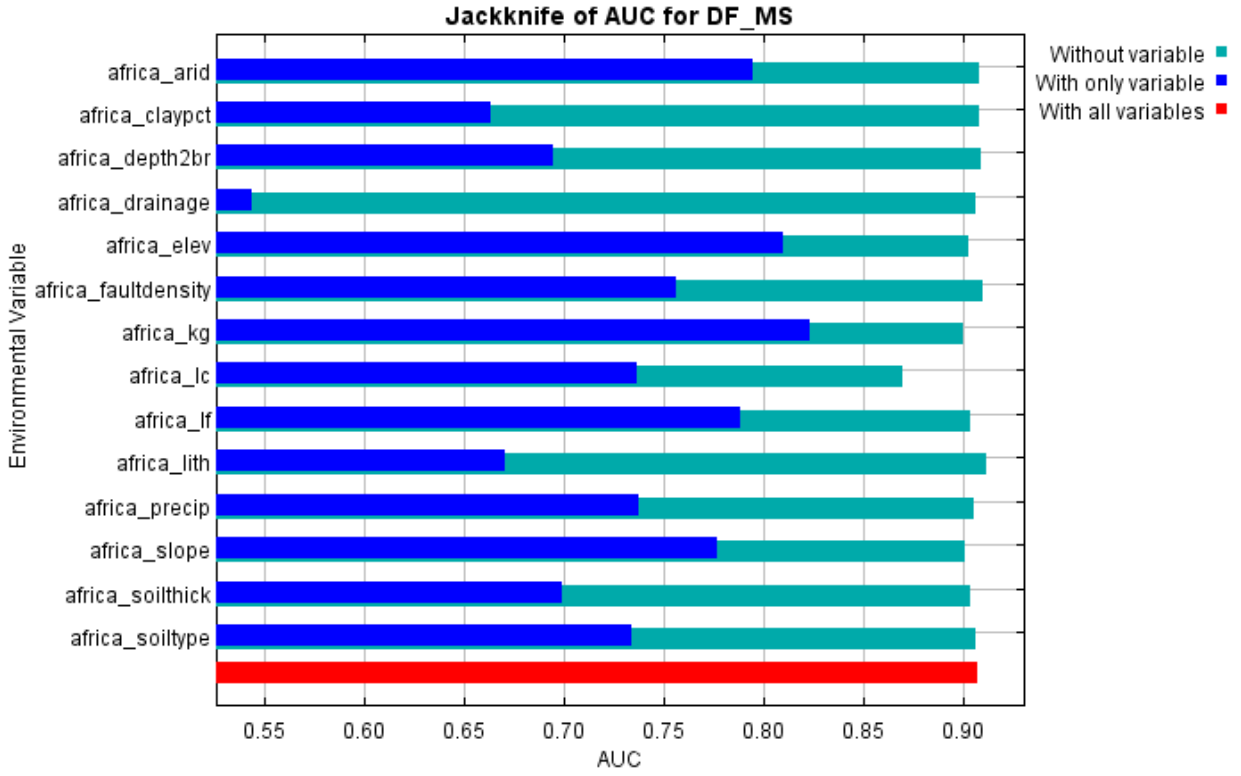
The following picture shows the results of the jackknife test of variable importance. The environmental variable with highest gain when used in isolation is africa_soiltype, which therefore appears to have the most useful information by itself. The environmental variable that decreases the gain the most when it is omitted is africa_lc, which therefore appears to have the most information that isn't present in the other variables. Values shown are averages over replicate runs.



The next picture shows the same jackknife test, using test gain instead of training gain. Note that conclusions about which variables are most important can change, now that we're looking at test data.



Lastly, we have the same jackknife test, using AUC on test data.



Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E DF_MS responsecurves jackknife outputformat=logistic "outputdirectory=C:_PhD_PhD PROJECT\MaxEntropy_Stats\MaxEntResults\ContinentalAnalysis\Africa" "samplesfile=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EventData_samples\Africa_TRAIN.csv" "environmentallayers=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EnvLayers\Africa" randomseed replicates=5 -t africa_drainage -t africa_kg -t africa_lc -t africa_lf -t africa_lith -t africa_soiltype

E.3 Asia

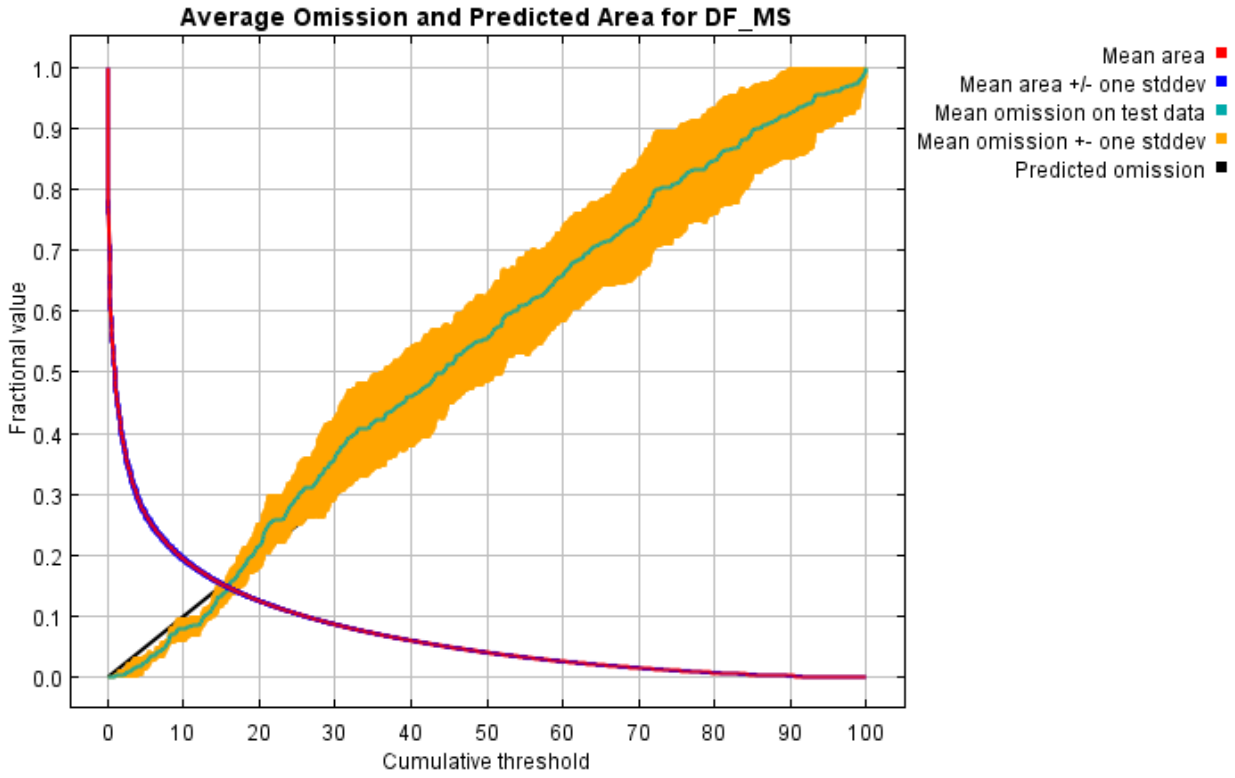
Replicated maxent model for DF_MS

This page summarizes the results of 5-fold cross-validation for DF_MS, created Sun

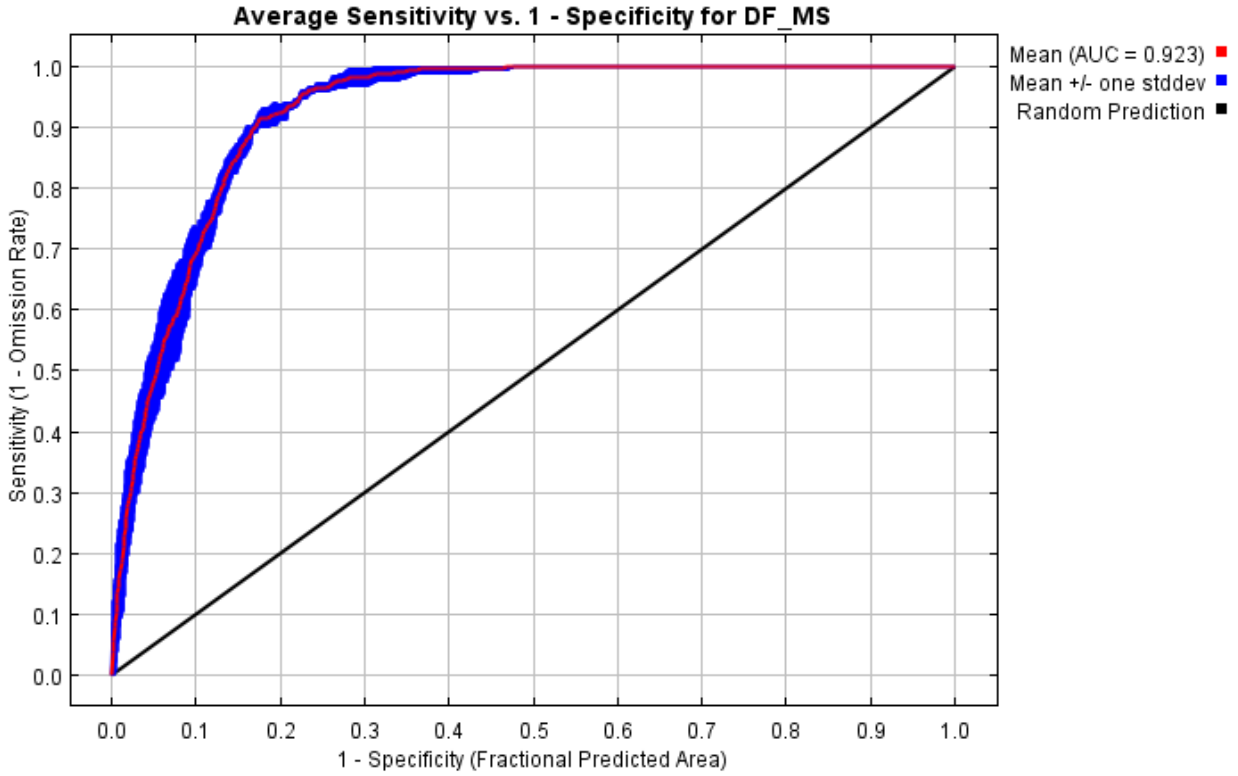
Jun 13 15:42:17 CEST 2021 using Maxent version 3.4.4. The individual models are here: [0] [1] [2] [3] [4]

Analysis of omission/commission

The following picture shows the test omission rate and predicted area as a function of the cumulative threshold, averaged over the replicate runs. The omission rate should be close to the predicted omission, because of the definition of the cumulative threshold.

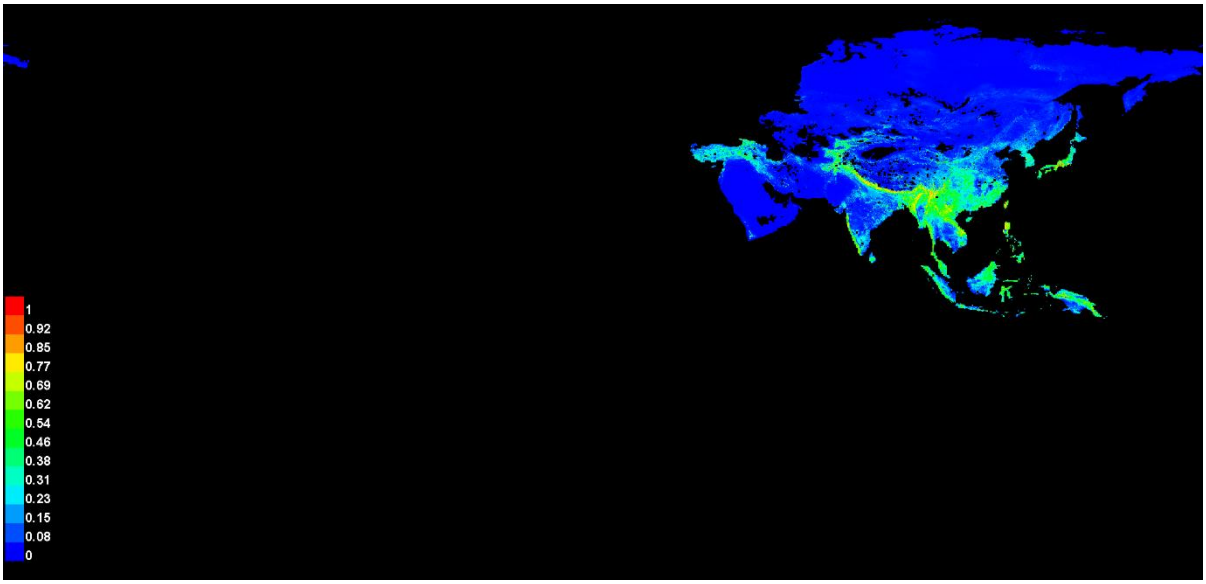


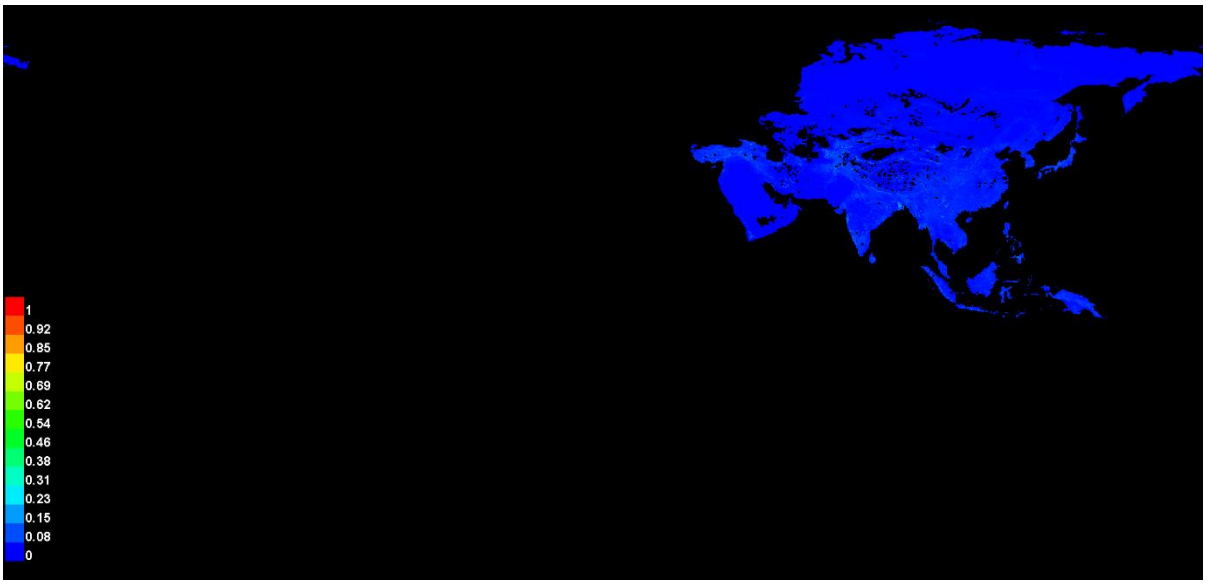
The next picture is the receiver operating characteristic (ROC) curve for the same data, again averaged over the replicate runs. Note that the specificity is defined using predicted area, rather than true commission (see the paper by Phillips, Anderson and Schapire cited on the help page for discussion of what this means). The average test AUC for the replicate runs is 0.923, and the standard deviation is 0.008.



Pictures of the model

The following two pictures show the point-wise mean and standard deviation of the 5 output grids. Other available summary grids are min, max and median.

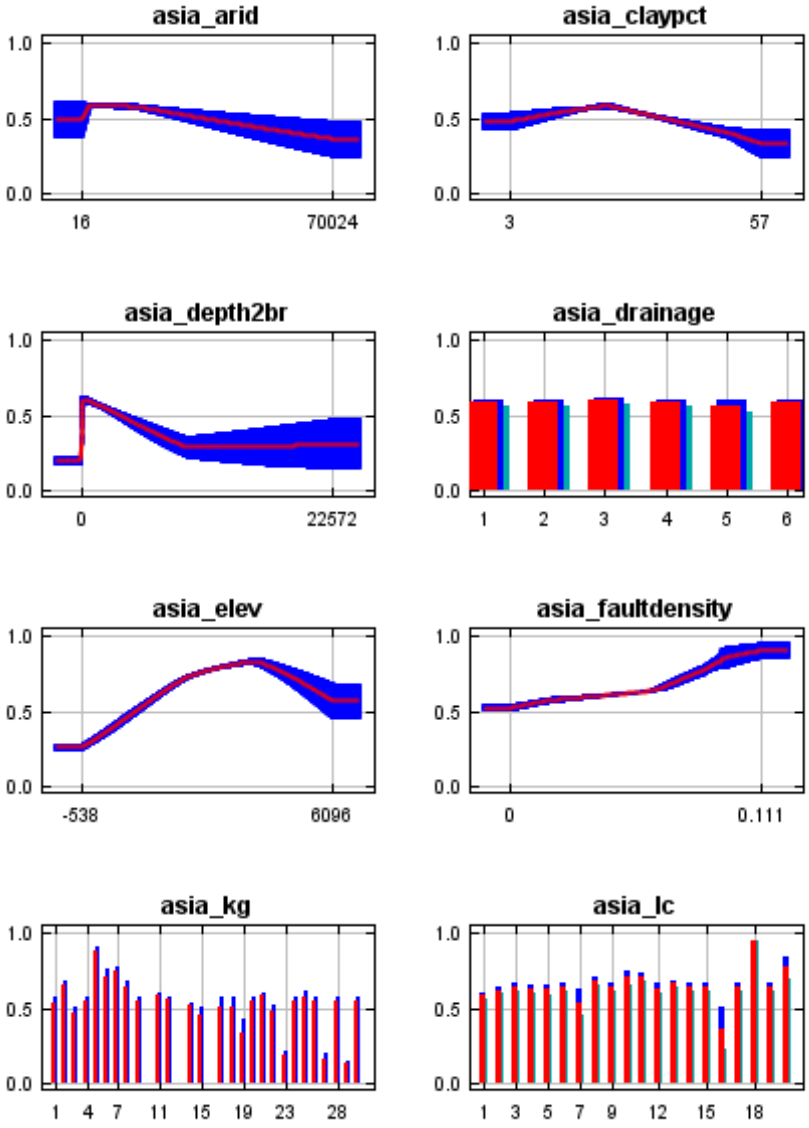


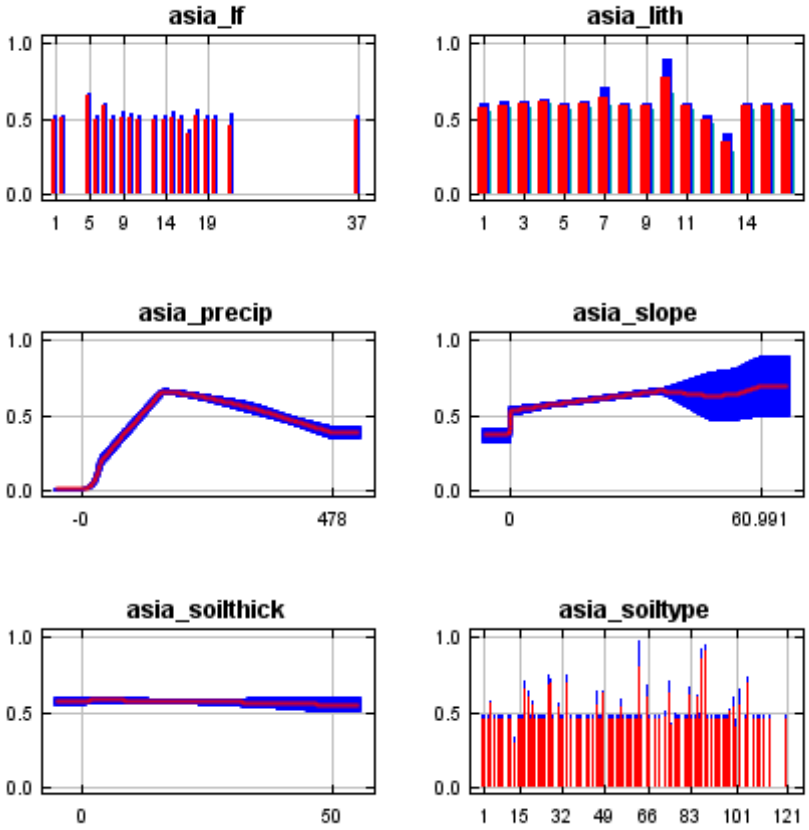


Response curves

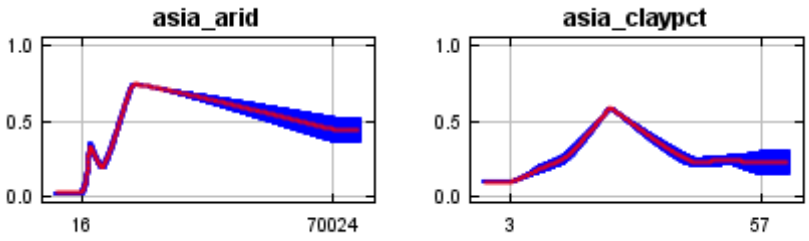
These curves show how each environmental variable affects the Maxent prediction. The curves show how the predicted probability of presence changes as each environmental variable is varied, keeping all other environmental variables at their average sample value. Click on a response curve to see a larger version. Note that the curves can be hard to interpret if you have strongly correlated variables, as the model may depend on the correlations in ways that are not evident in the curves. In

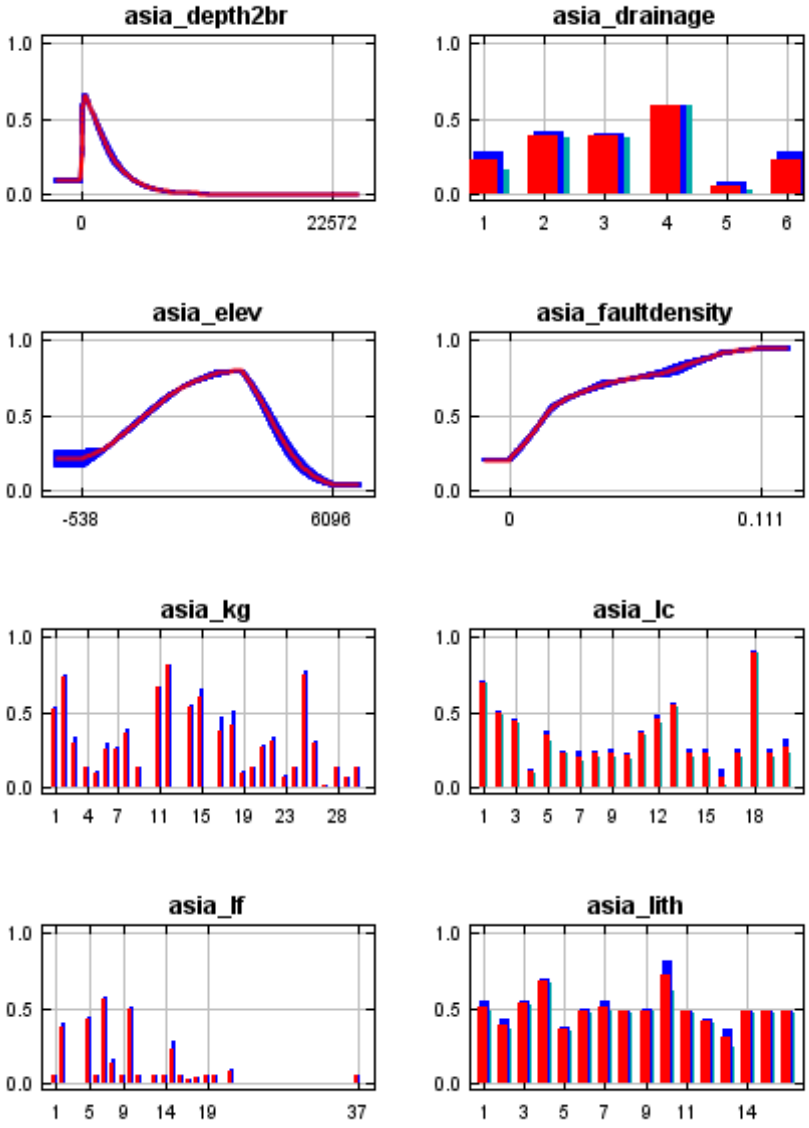
other words, the curves show the marginal effect of changing exactly one variable, whereas the model may take advantage of sets of variables changing together. The curves show the mean response of the 5 replicate Maxent runs (red) and and the mean +/- one standard deviation (blue, two shades for categorical variables).

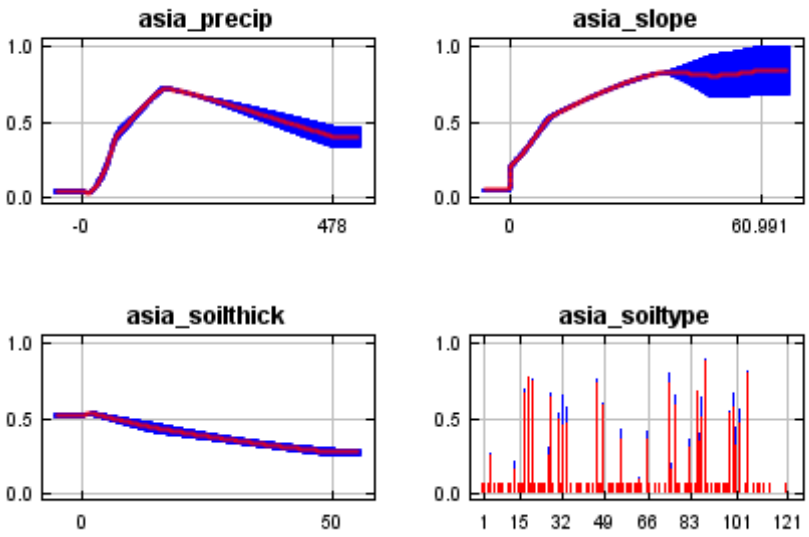




In contrast to the above marginal response curves, each of the following curves represents a different model, namely, a Maxent model created using only the corresponding variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other variables. They may be easier to interpret if there are strong correlations between variables.




asia_depth2br
1.0
0.5
0.0
0
22572
asia_drainage
1.0
0.5
0.0
1 2 3 4 5 6
asia_elev
1.0
0.5
0.0
-538
6096
asia_faultdensity
1.0
0.5
0.0
0
0.111
asia_kg
1.0
0.5
0.0
1 4 7 11 15 19 23 28
asia_lc
1.0
0.5
0.0
1 3 5 7 9 12 15 18
asia_lf
1.0
0.5
0.0
1 5 9 14 19 37
asia_lith
1.0
0.5
0.0
1 3 5 7 9 11 14



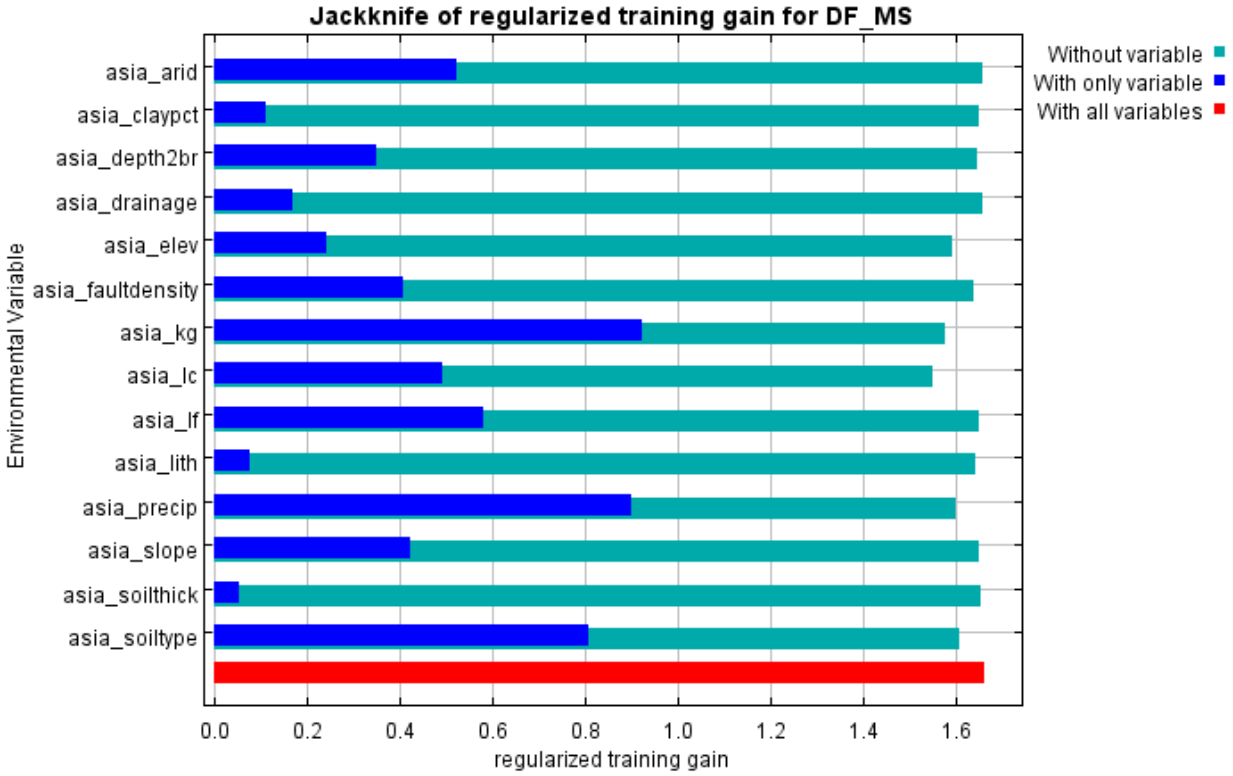
Analysis of variable contributions

The following table gives estimates of relative contributions of the environmental variables to the Maxent model. To determine the first estimate, in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the corresponding variable, or subtracted from it if the change to the absolute value of lambda is negative. For the second estimate, for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is reevaluated on the permuted data, and the resulting drop in training AUC is shown in the table, normalized to percentages. As with the variable jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated. Values shown are averages over replicate runs.

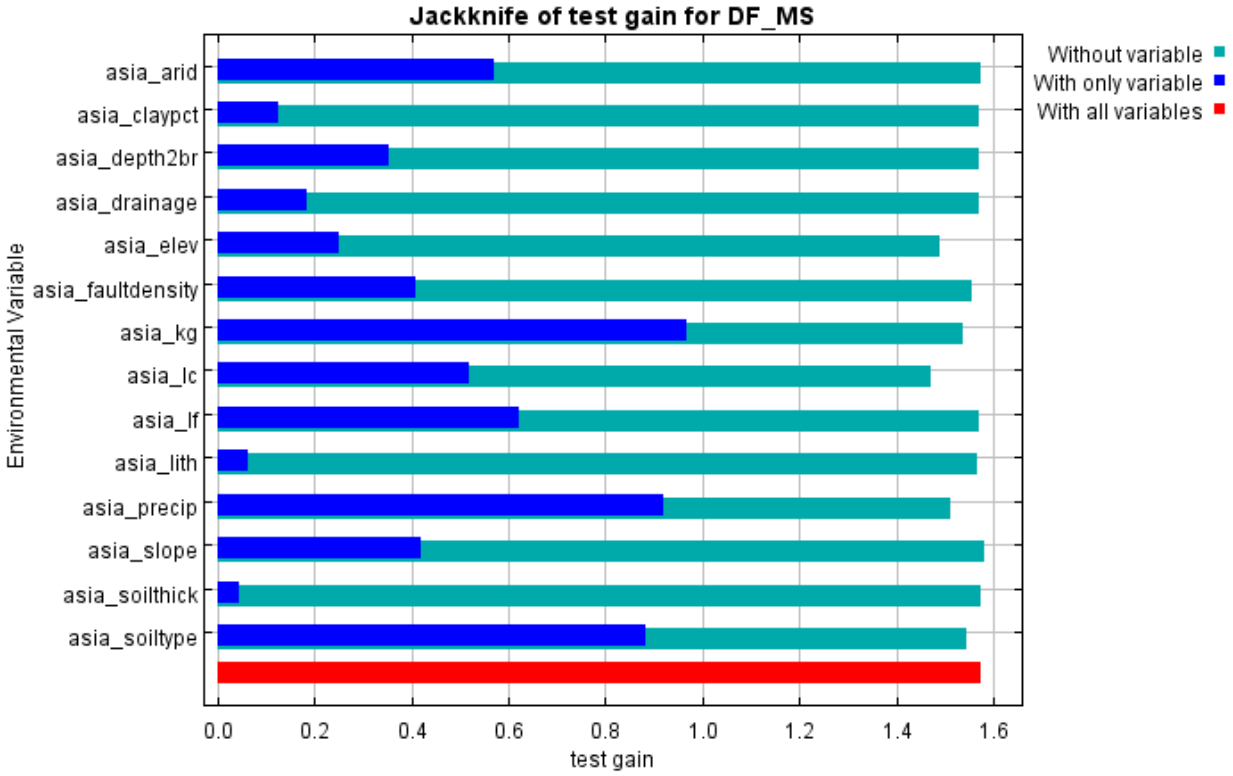
Variable	Percent contribution	Permutation importance
asia_precip	46.6	55
asia_kg	15.6	19.2

asia_lc	8.1	5
asia_faultdensity	8	1
asia_slope	5.9	0.7
asia_elev	4.6	8.7
asia_soiltype	4.6	3.9
asia_lf	2.7	1.3
asia_claypct	1.5	0.8
asia_lith	1.2	1
asia_depth2br	0.6	2.3
asia_soilthick	0.4	0.6
asia_drainage	0.1	0.1
asia_arid	0.1	0.2

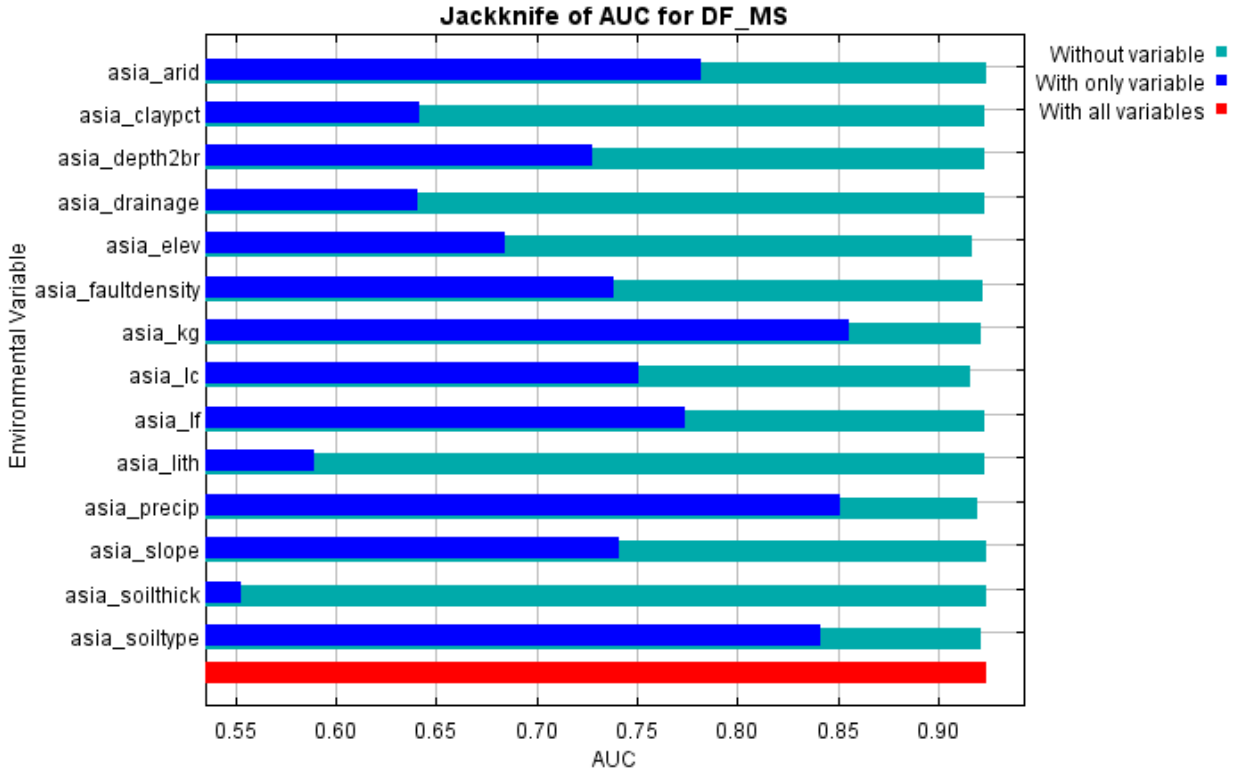
The following picture shows the results of the jackknife test of variable importance. The environmental variable with highest gain when used in isolation is asia_kg, which therefore appears to have the most useful information by itself. The environmental variable that decreases the gain the most when it is omitted is asia_lc, which therefore appears to have the most information that isn't present in the other variables. Values shown are averages over replicate runs.



The next picture shows the same jackknife test, using test gain instead of training gain. Note that conclusions about which variables are most important can change, now that we're looking at test data.



Lastly, we have the same jackknife test, using AUC on test data.



Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E DF_MS responsecurves jackknife outputformat=logistic "outputdirectory=C:_PhD_PhD PROJECT\MaxEntropy_Stats\MaxEntResults\ContinentalAnalysis\Asia" "samplesfile=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EventData_samples\Asia_TRAIN.csv" "environmentallayers=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EnvLayers\Asia\Asia_Redo" randomseed noaskoverwrite replicates=5 -t asia_drainage -t asia_kg -t asia_lc -t asia_lf -t asia_lith -t asia_soiltype

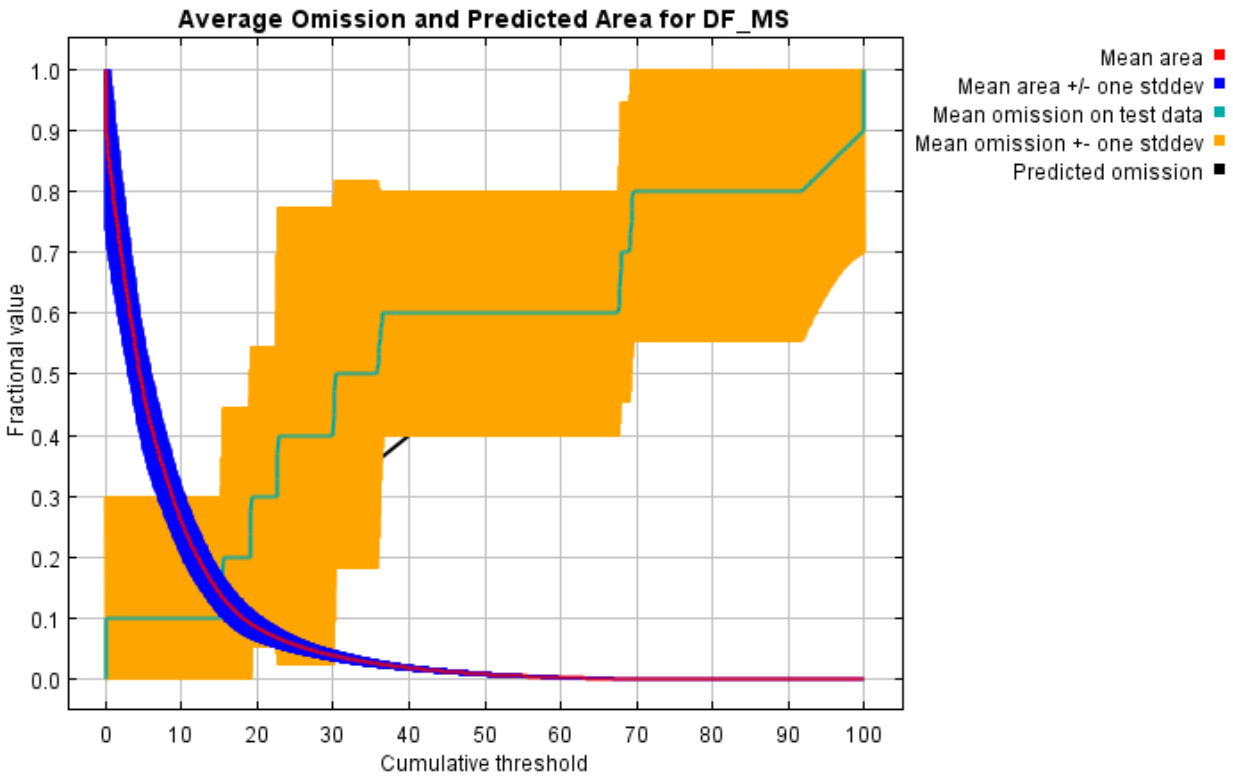
E.4 Australia

Replicated maxent model for DF_MS

This page summarizes the results of 5-fold cross-validation for DF_MS, created Mon Jun 14 17:33:38 CEST 2021 using Maxent version 3.4.4. The individual models are here: [0] [1] [2] [3] [4]

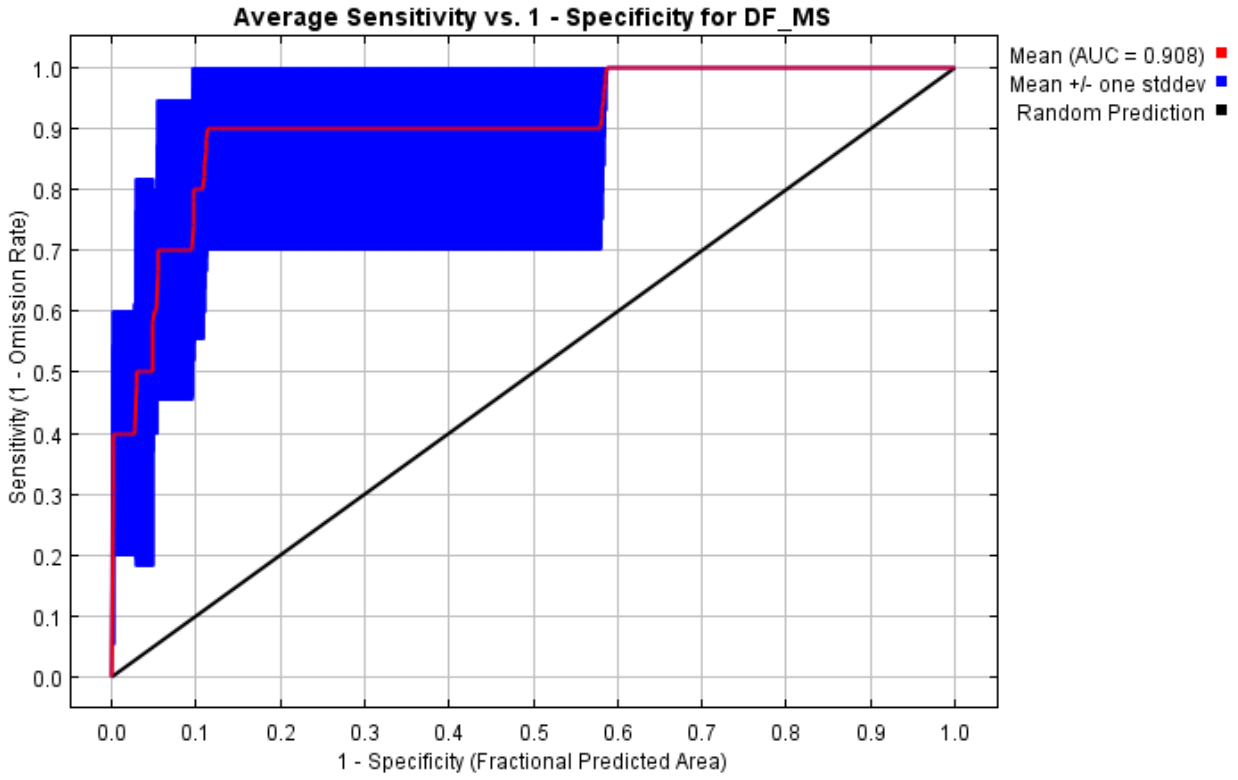
Analysis of omission/commission

The following picture shows the test omission rate and predicted area as a function of the cumulative threshold, averaged over the replicate runs. The omission rate should be close to the predicted omission, because of the definition of the cumulative threshold.



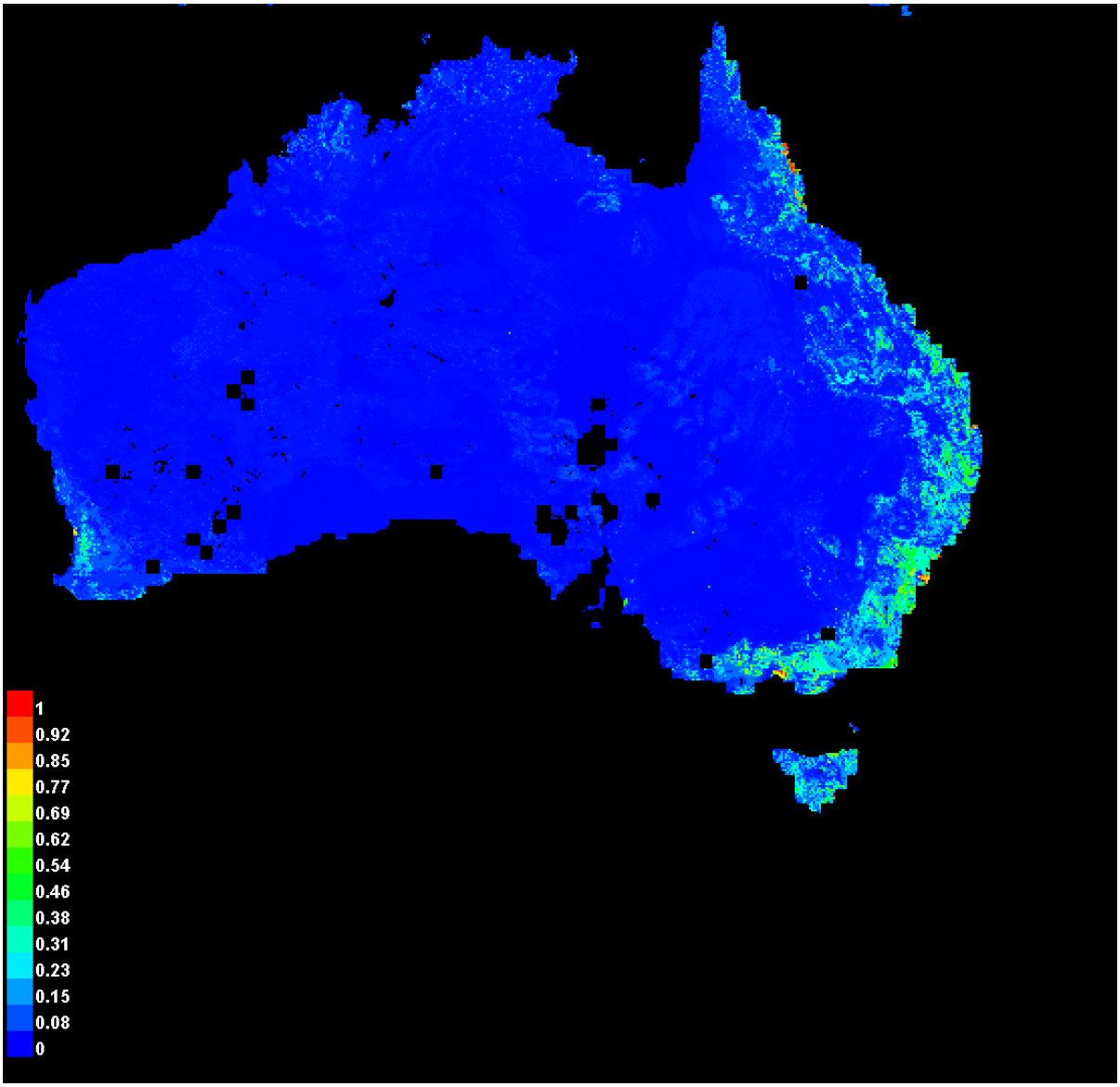
The next picture is the receiver operating characteristic (ROC) curve for the same data, again averaged over the replicate runs. Note that the specificity is defined using predicted area, rather than true commission (see the paper by Phillips, Anderson and Schapire cited on the help page for discussion of what this means). The average test AUC for the replicate runs is 0.908, and the standard deviation is

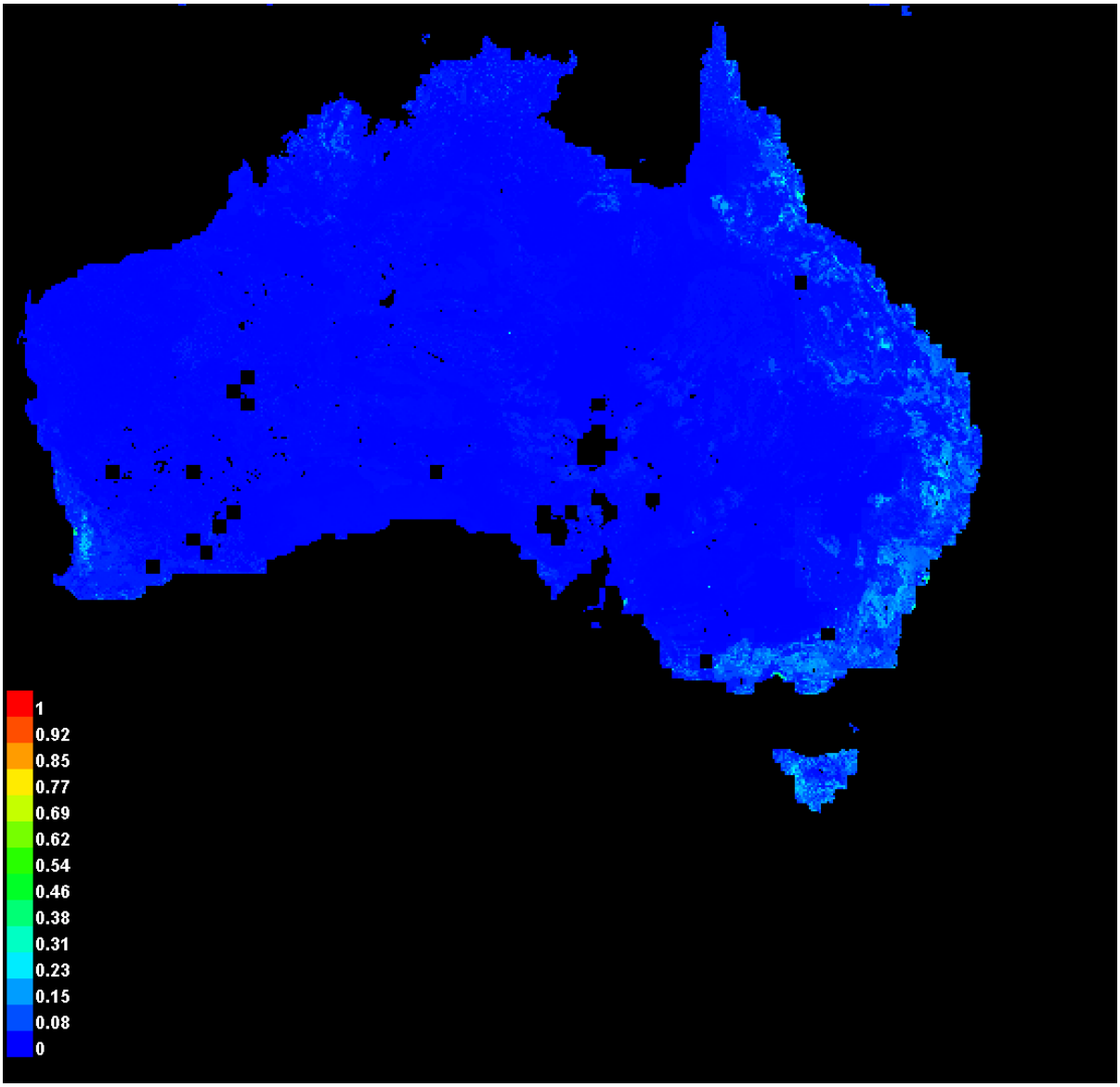
0.101.



Pictures of the model

The following two pictures show the point-wise mean and standard deviation of the 5 output grids. Other available summary grids are min, max and median.

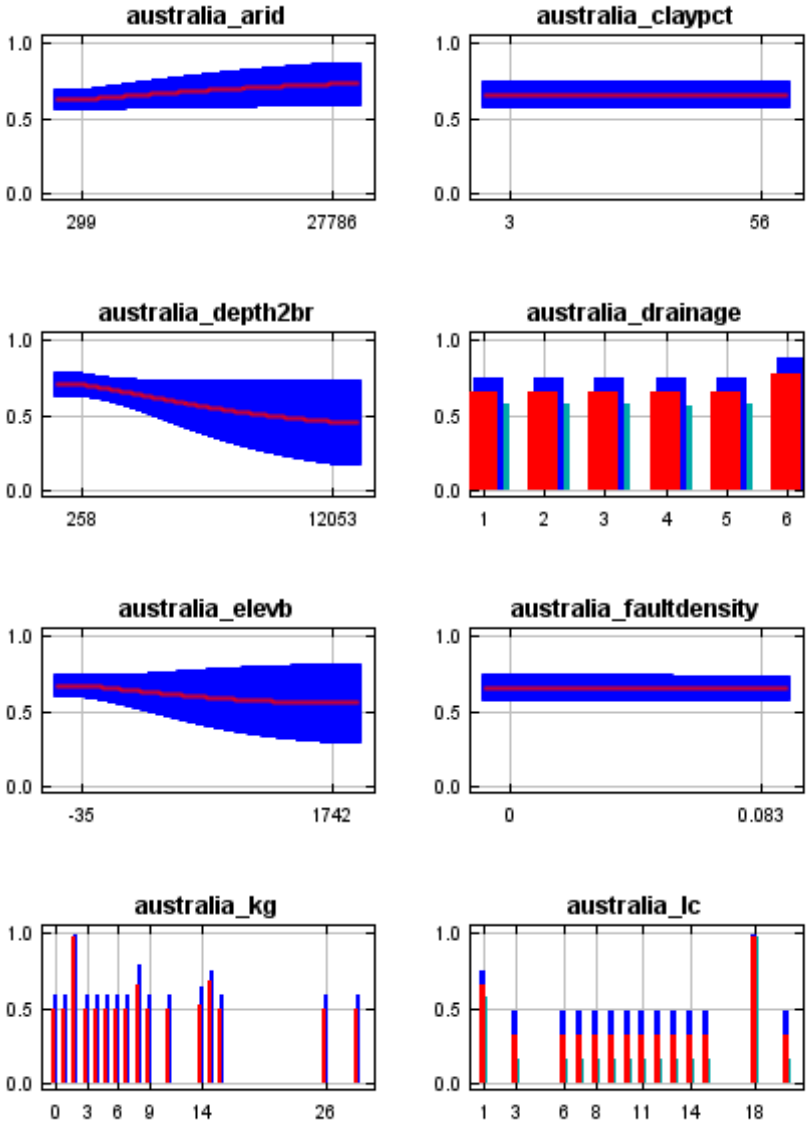

1
0.92
0.85
0.77
0.69
0.62
0.54
0.46
0.38
0.31
0.23
0.15
0.08
0

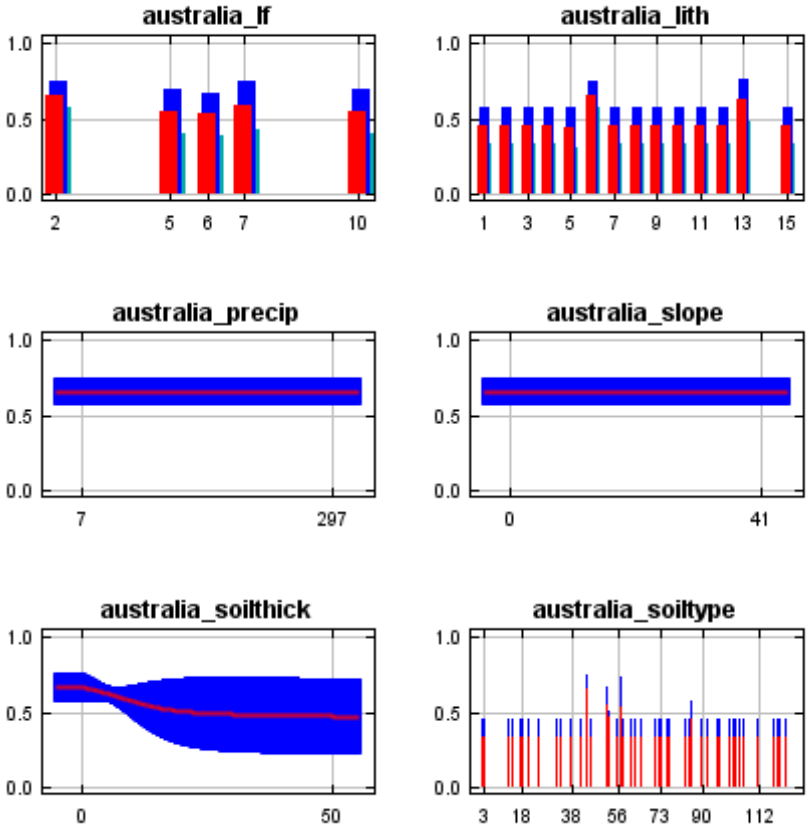


Response curves

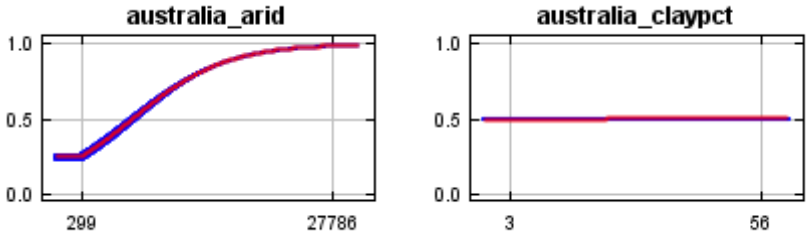
These curves show how each environmental variable affects the Maxent prediction. The curves show how the predicted probability of presence changes as each environmental variable is varied, keeping all other environmental variables at their average sample value. Click on a response curve to see a larger version. Note that the curves can be hard to interpret if you have strongly correlated variables, as the model may depend on the correlations in ways that are not evident in the curves. In

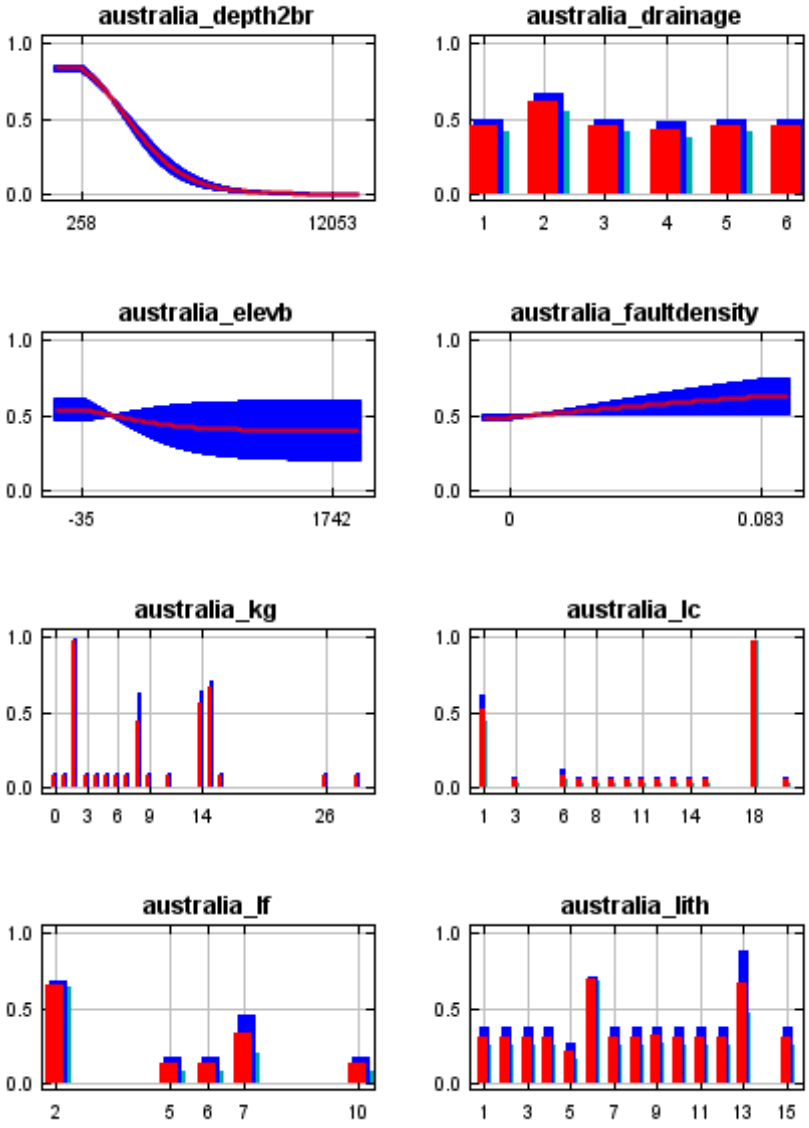
other words, the curves show the marginal effect of changing exactly one variable, whereas the model may take advantage of sets of variables changing together. The curves show the mean response of the 5 replicate Maxent runs (red) and and the mean +/- one standard deviation (blue, two shades for categorical variables).

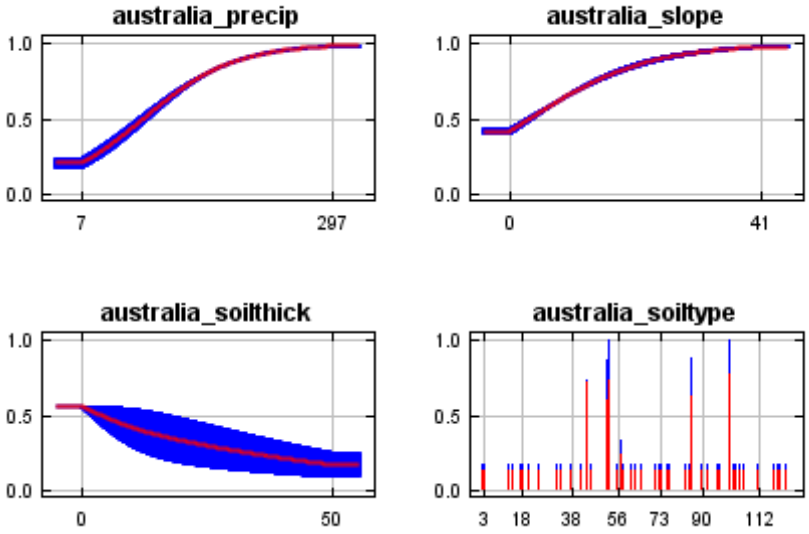




In contrast to the above marginal response curves, each of the following curves represents a different model, namely, a Maxent model created using only the corresponding variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other variables. They may be easier to interpret if there are strong correlations between variables.




australia_depth2br
1.0
0.5
0.0
258
12053
australia_drainage
1.0
0.5
0.0
1 2 3 4 5 6
australia_elevb
1.0
0.5
0.0
-35
1742
australia_faultdensity
1.0
0.5
0.0
0
0.083
australia_kg
1.0
0.5
0.0
0 3 6 9 14 26
australia_lc
1.0
0.5
0.0
1 3 6 8 11 14 18
australia_lf
1.0
0.5
0.0
2 5 6 7 10
australia_lith
1.0
0.5
0.0
1 3 5 7 9 11 13 15



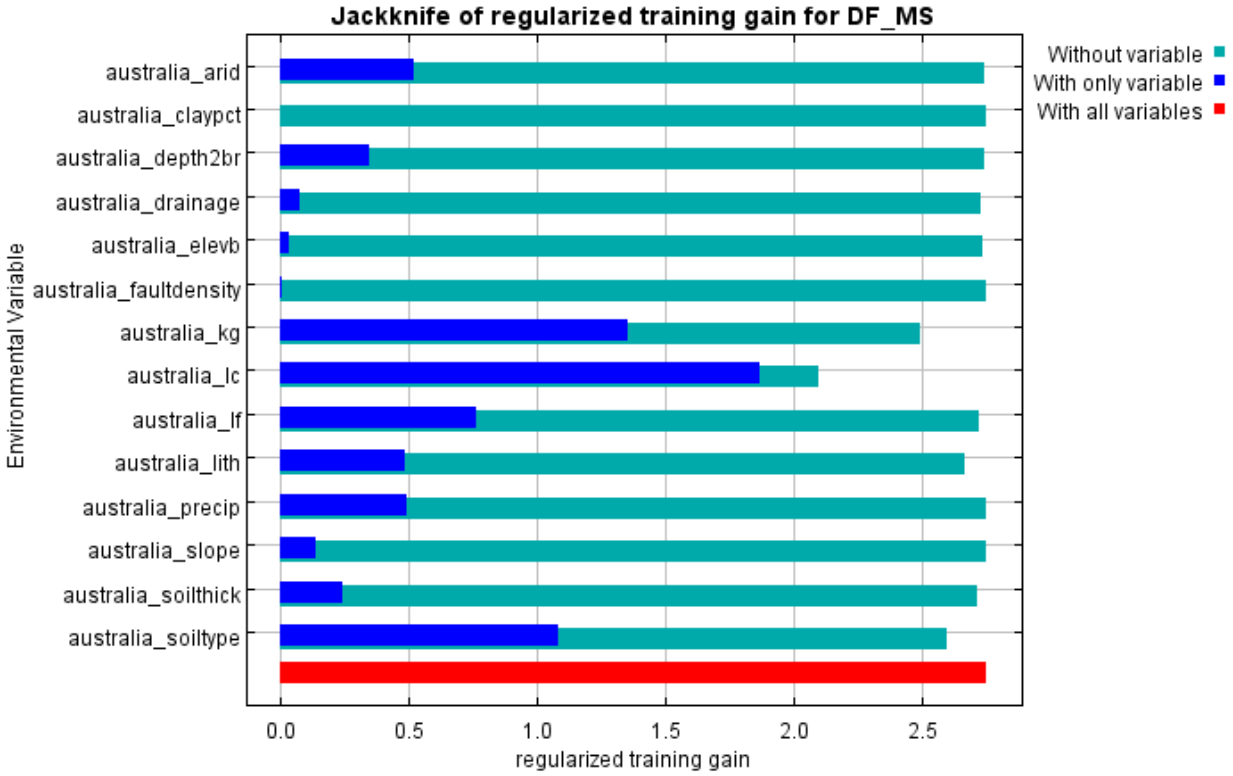
Analysis of variable contributions

The following table gives estimates of relative contributions of the environmental variables to the Maxent model. To determine the first estimate, in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the corresponding variable, or subtracted from it if the change to the absolute value of lambda is negative. For the second estimate, for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is reevaluated on the permuted data, and the resulting drop in training AUC is shown in the table, normalized to percentages. As with the variable jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated. Values shown are averages over replicate runs.

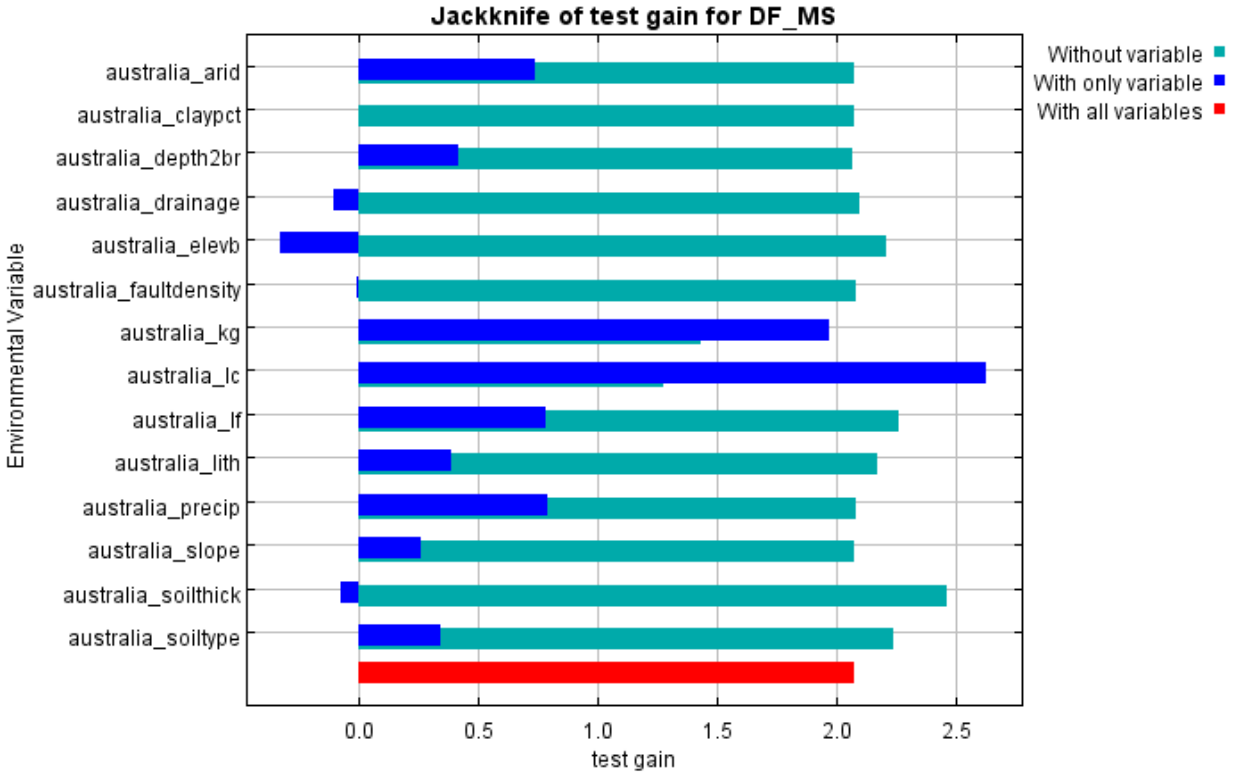
Variable	Percent contribution	Permutation importance
australia_lc	46	31.8
australia_kg	16.9	6.8

australia_soiltype	13.1	21.8
australia_lf	12.7	7.9
australia_lith	7.5	4.4
australia_soilthick	2.5	18
australia_drainage	0.8	4.9
australia_depth2br	0.2	3.6
australia_elevb	0.1	0.8
australia_arid	0.1	0
australia_faultdensity	0	0
australia_claypct	0	0
australia_slope	0	0
australia_precip	0	0

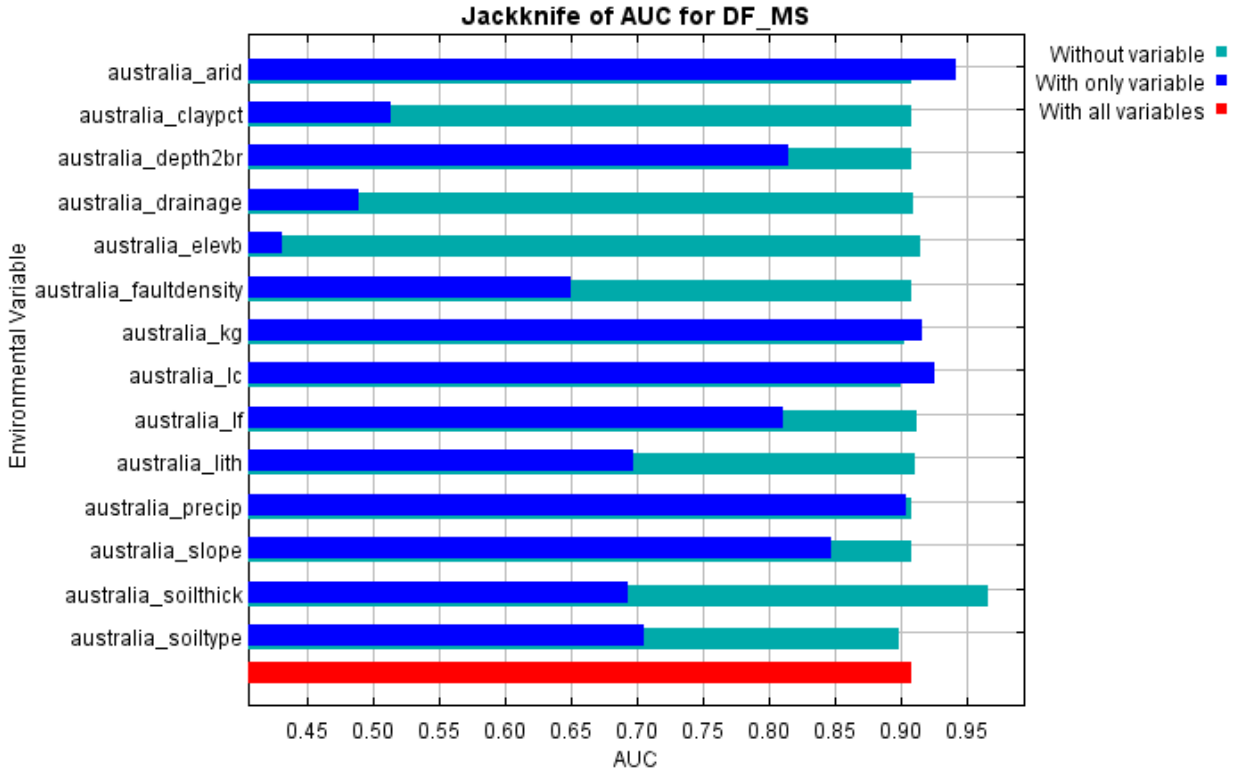
The following picture shows the results of the jackknife test of variable importance. The environmental variable with highest gain when used in isolation is australia_lc, which therefore appears to have the most useful information by itself. The environmental variable that decreases the gain the most when it is omitted is australia_lc, which therefore appears to have the most information that isn't present in the other variables. Values shown are averages over replicate runs.



The next picture shows the same jackknife test, using test gain instead of training gain. Note that conclusions about which variables are most important can change, now that we're looking at test data.



Lastly, we have the same jackknife test, using AUC on test data.



Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E DF_MS responsecurves jackknife outputformat=logistic "outputdirectory=C:_PhD_PhD PROJECT\MaxEntropy_Stats\MaxEntResults\ContinentalAnalysis\Australia" "samplesfile=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EventData_samples\Australia_TRAIN.csv" "environmentallayers=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EnvLayers\Australia" randomseed noaskoverwrite replicates=5 -t australia_drainage -t australia_kg -t australia_lc -t australia_lf -t australia_lith -t australia_soiltype

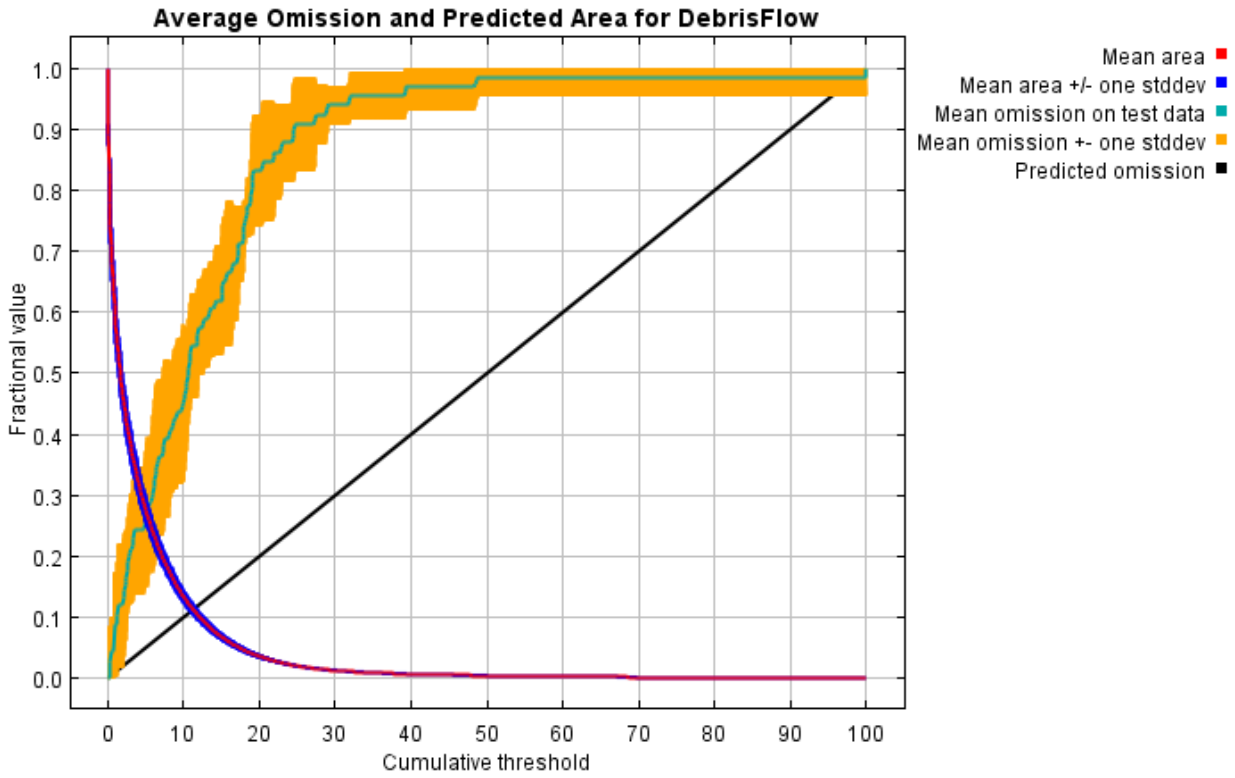
E.5 Europe

Replicated maxent model for DebrisFlow

This page summarizes the results of 5-fold cross-validation for DebrisFlow, created Mon Aug 02 14:57:26 PDT 2021 using Maxent version 3.4.4. The individual models are here: [0] [1] [2] [3] [4]

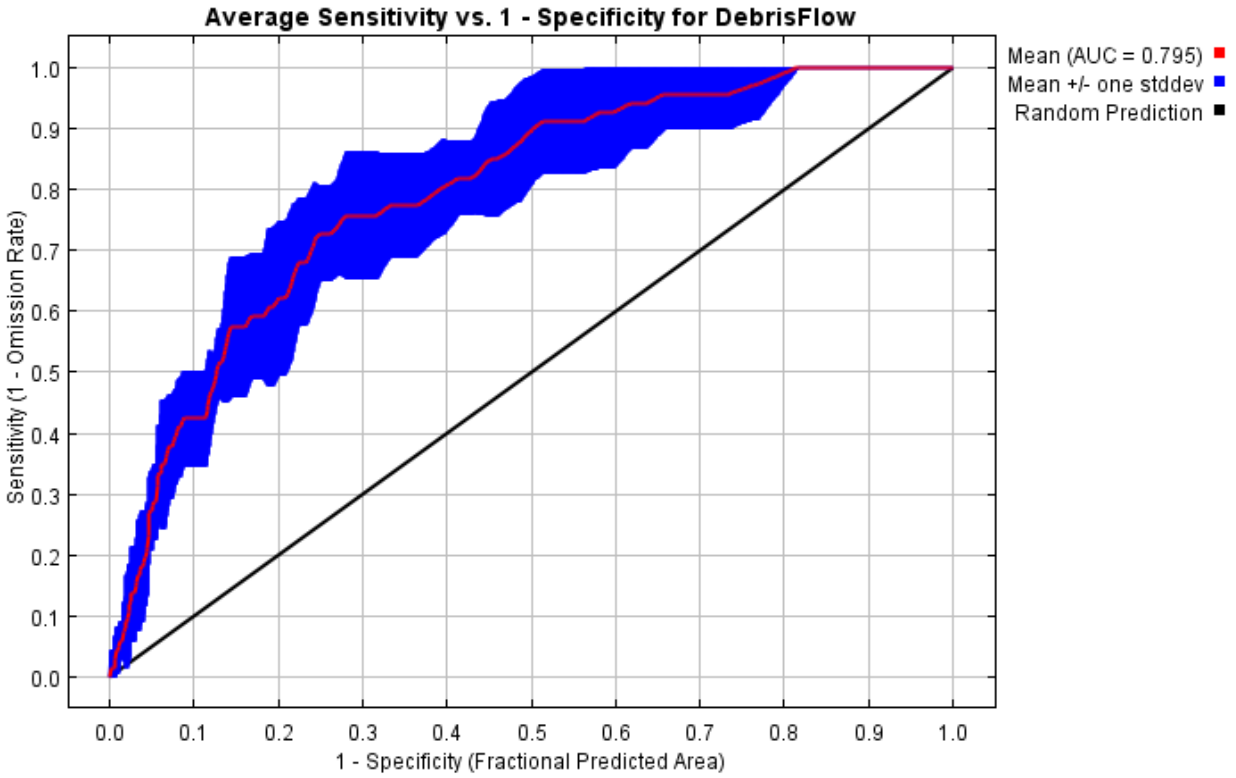
Analysis of omission/commission

The following picture shows the test omission rate and predicted area as a function of the cumulative threshold, averaged over the replicate runs. The omission rate should be close to the predicted omission, because of the definition of the cumulative threshold.



The next picture is the receiver operating characteristic (ROC) curve for the same

data, again averaged over the replicate runs. Note that the specificity is defined using predicted area, rather than true commission (see the paper by Phillips, Anderson and Schapire cited on the help page for discussion of what this means). The average test AUC for the replicate runs is 0.795, and the standard deviation is 0.049.



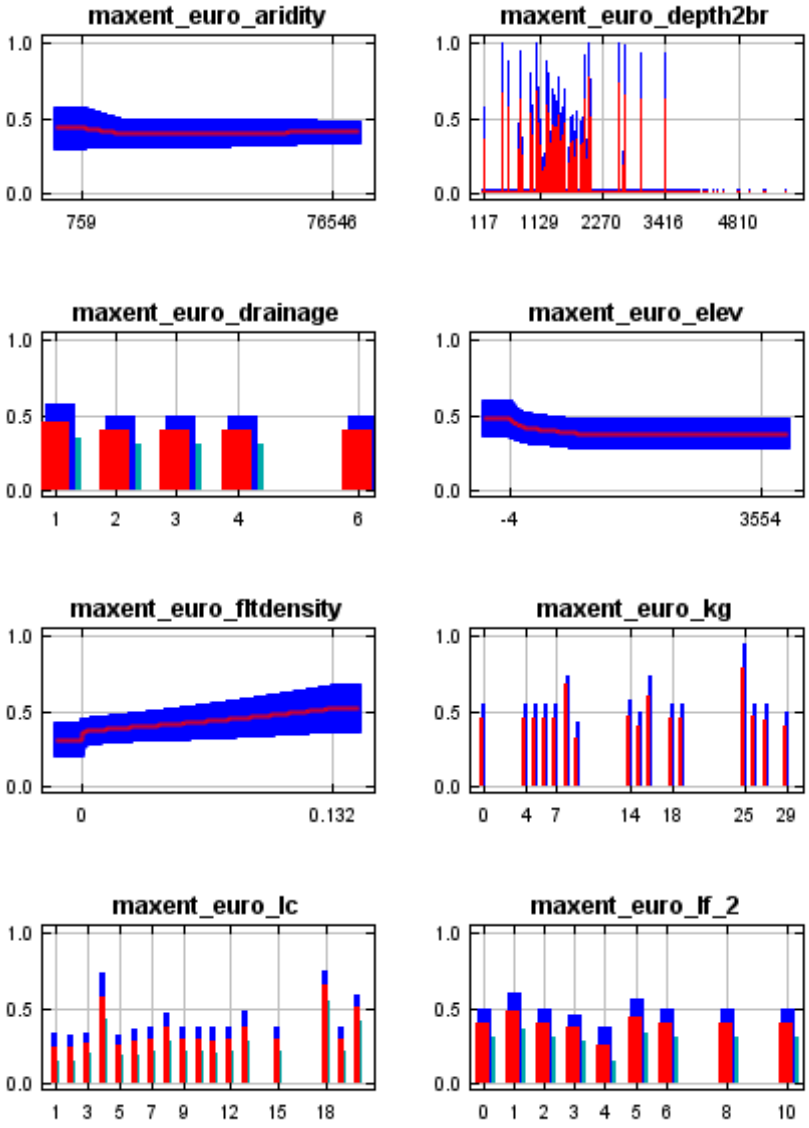
Pictures of the model

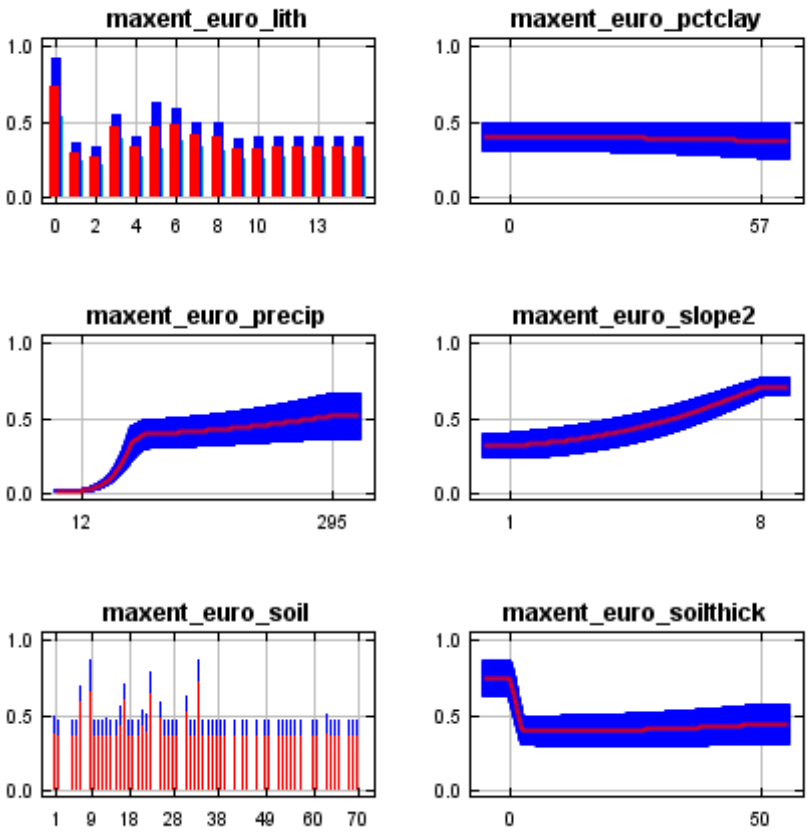
The following two pictures show the point-wise mean and standard deviation of the 5 output grids. Other available summary grids are min, max and median.

Response curves

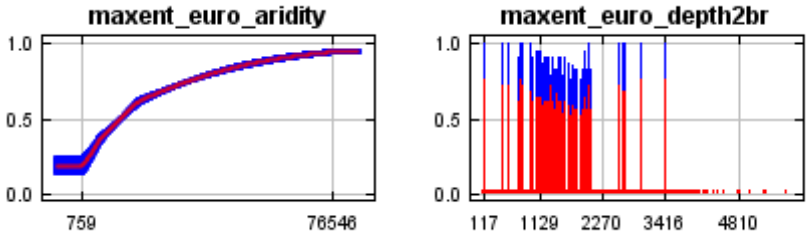
These curves show how each environmental variable affects the Maxent prediction. The curves show how the predicted probability of presence changes as each environmental variable is varied, keeping all other environmental variables at their average sample value. Click on a response curve to see a larger version. Note that the curves can be hard to interpret if you have strongly correlated variables, as the model may depend on the correlations in ways that are not evident in the curves. In

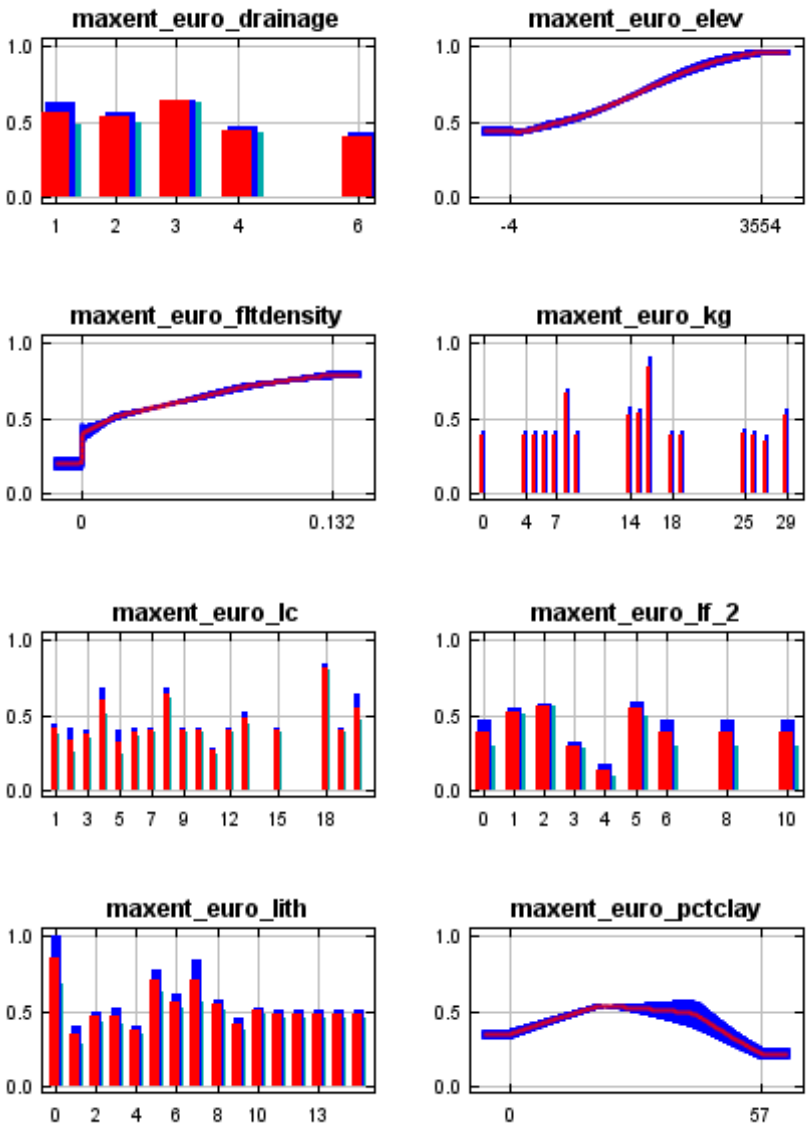
other words, the curves show the marginal effect of changing exactly one variable, whereas the model may take advantage of sets of variables changing together. The curves show the mean response of the 5 replicate Maxent runs (red) and and the mean +/- one standard deviation (blue, two shades for categorical variables).

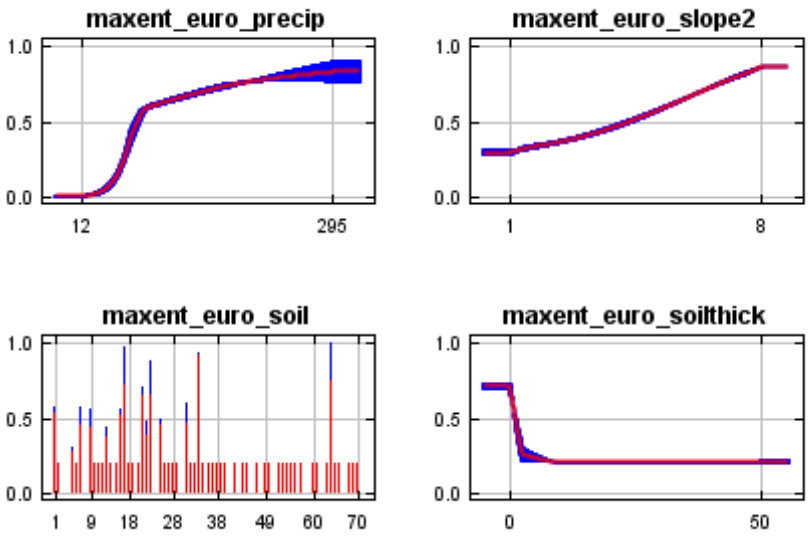




In contrast to the above marginal response curves, each of the following curves represents a different model, namely, a Maxent model created using only the corresponding variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other variables. They may be easier to interpret if there are strong correlations between variables.




maxent_euro_drainage
1.0
0.5
0.0
1
2
3
4
6
maxent_euro_elev
1.0
0.5
0.0
-4
3554
maxent_euro_fltdensity
1.0
0.5
0.0
0
0.132
maxent_euro_kg
1.0
0.5
0.0
0
4
7
14
18
25
29
maxent_euro_lc
1.0
0.5
0.0
1
3
5
7
9
12
15
18
maxent_euro_lf_2
1.0
0.5
0.0
0
1
2
3
4
5
6
8
10
maxent_euro_lith
1.0
0.5
0.0
0
2
4
6
8
10
13
maxent_euro_pctclay
1.0
0.5
0.0
0
57



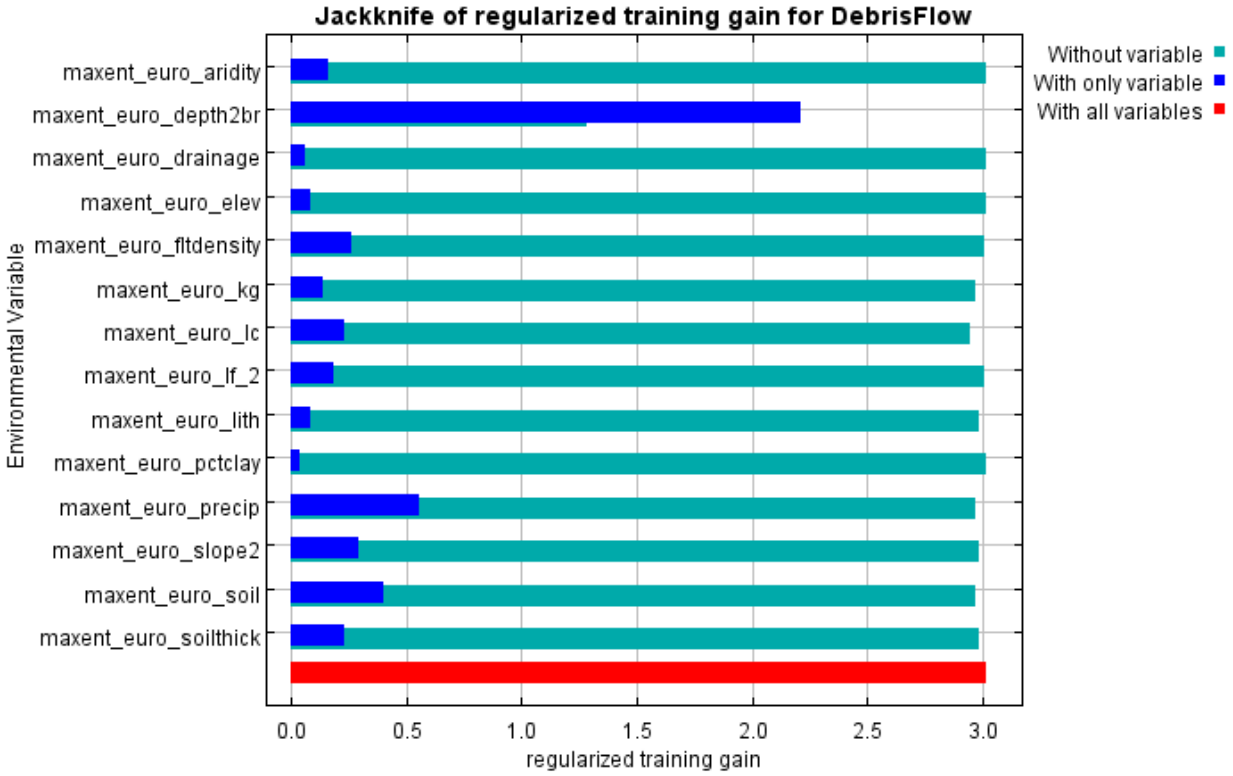
Analysis of variable contributions

The following table gives estimates of relative contributions of the environmental variables to the Maxent model. To determine the first estimate, in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the corresponding variable, or subtracted from it if the change to the absolute value of lambda is negative. For the second estimate, for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is reevaluated on the permuted data, and the resulting drop in training AUC is shown in the table, normalized to percentages. As with the variable jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated. Values shown are averages over replicate runs.

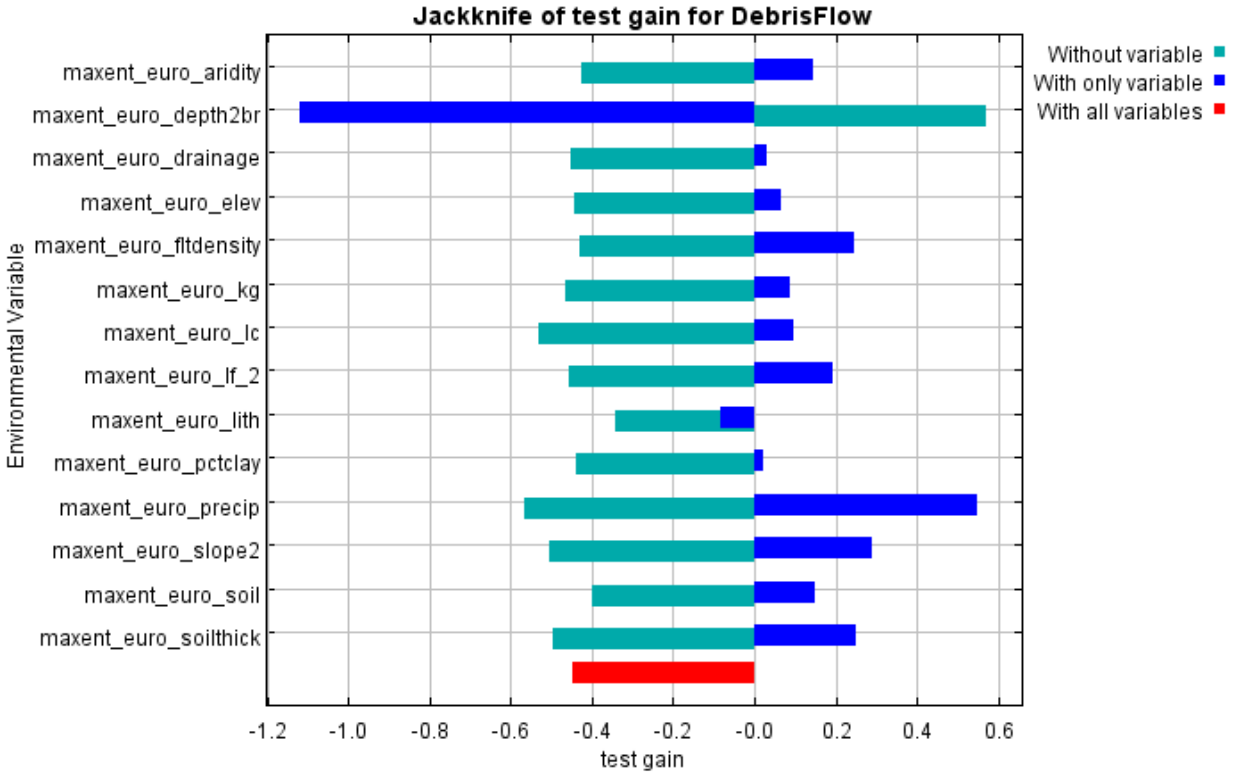
Variable	Percent contribution	Permutation importance
maxent_euro_depth2br	64.4	93.5
maxent_euro_precip	11.1	1.7

maxent_euro_lc	5	1
maxent_euro_fltdensity	4.7	0.1
maxent_euro_slope2	4.6	0.5
maxent_euro_soil	4.5	0.7
maxent_euro_kg	2	0.8
maxent_euro_lith	1.8	0.7
maxent_euro_soilthick	1	0.6
maxent_euro_lf_2	0.5	0.3
maxent_euro_drainage	0.1	0
maxent_euro_elev	0.1	0.1
maxent_euro_aridity	0	0
maxent_euro_pctclay	0	0

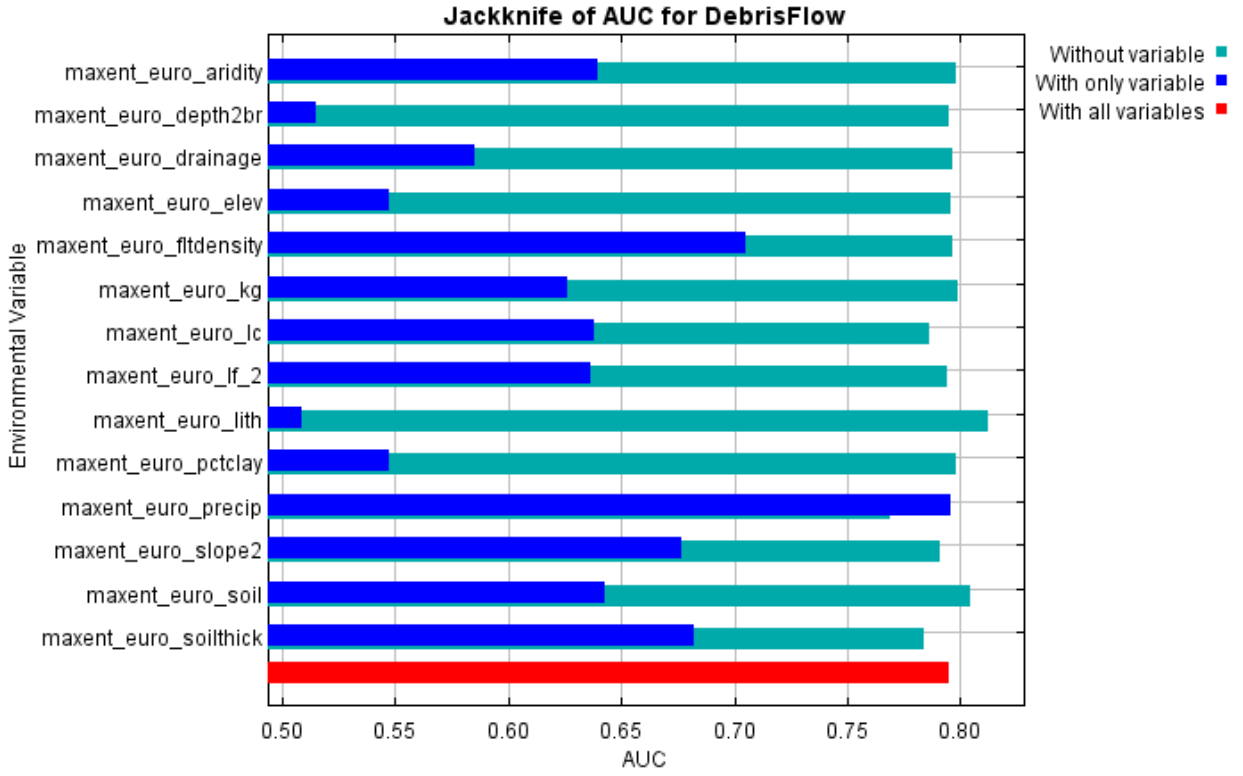
The following picture shows the results of the jackknife test of variable importance. The environmental variable with highest gain when used in isolation is maxent_euro_depth2br, which therefore appears to have the most useful information by itself. The environmental variable that decreases the gain the most when it is omitted is maxent_euro_depth2br, which therefore appears to have the most information that isn't present in the other variables. Values shown are averages over replicate runs.



The next picture shows the same jackknife test, using test gain instead of training gain. Note that conclusions about which variables are most important can change, now that we're looking at test data.



Lastly, we have the same jackknife test, using AUC on test data.



Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E DebrisFlow responsecurves jackknife outputformat=logistic "outputdirectory=C:_PhD_PhD PROJECT\MaxEntropy_Stats\MaxEntResults\ContinentalAnalysis\Europe\NEWEST_RESULTS_TrainWithTestWithSlope" "samplesfile=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EventData_samples\TRAIN_Europe_For_MaxEnt2.csv" "environmentallayers=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EnvLayers\Europe" randomseed "testsamplesfile=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EventData_samples\Test1_Europe.csv" replicates=5 -t maxent_euro_depth2br -t maxent_euro_drainage -t maxent_euro_kg -t maxent_euro_lc -t maxent_euro_lf_2 -t maxent_euro_lith -t maxent_euro_soil

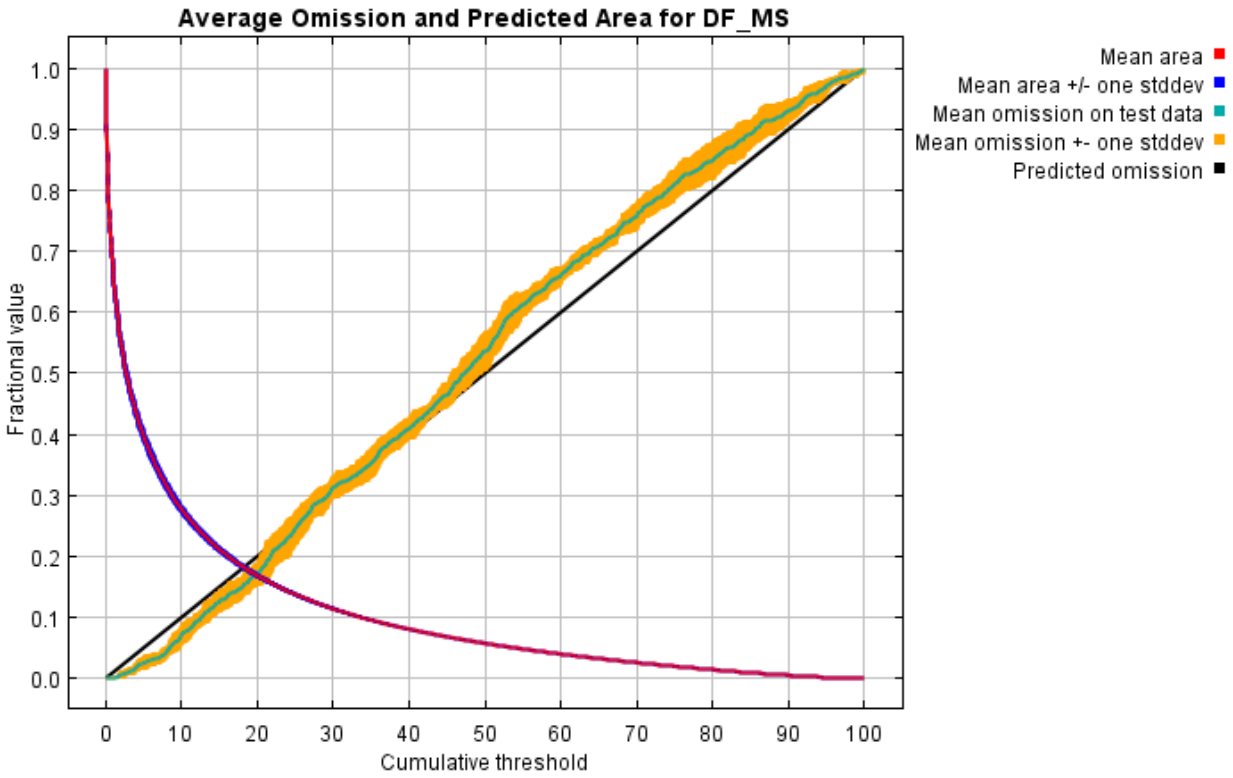
E.6 North America

Replicated maxent model for DF_MS

This page summarizes the results of 5-fold cross-validation for DF_MS, created Fri Jun 11 13:42:08 CEST 2021 using Maxent version 3.4.4. The individual models are here: [0] [1] [2] [3] [4]

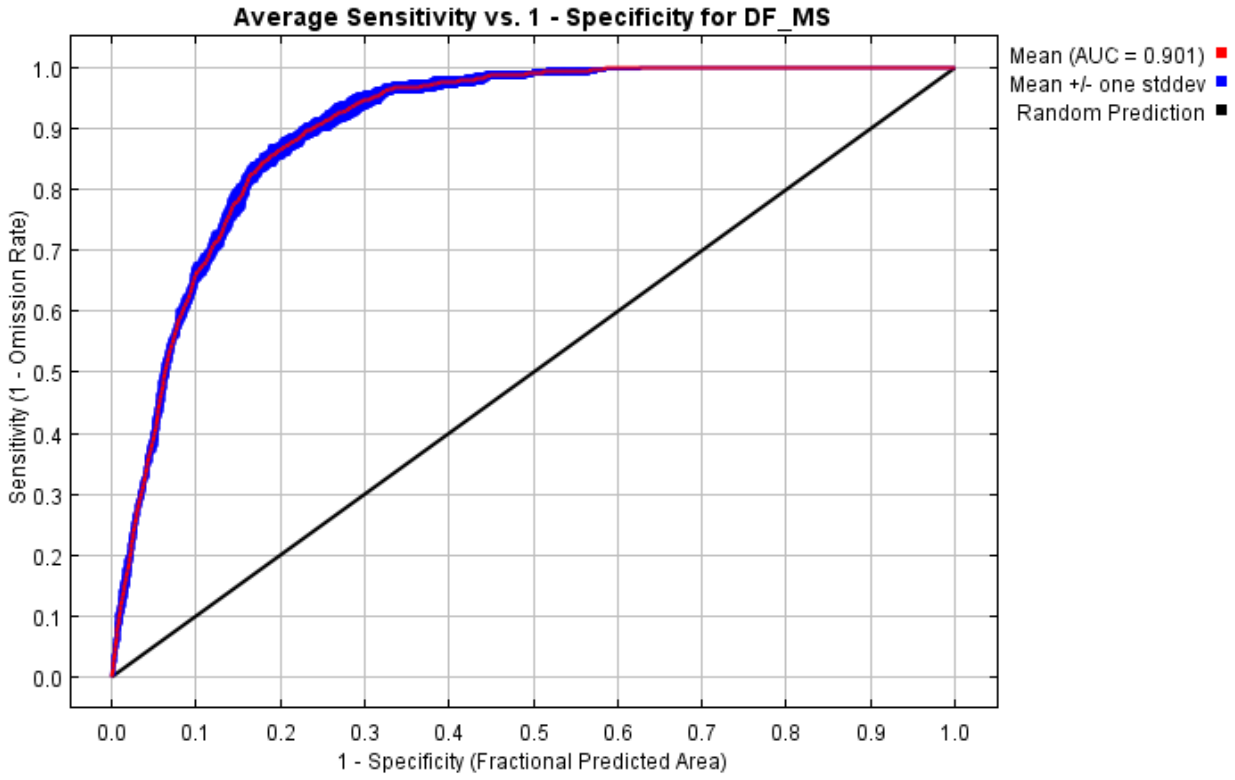
Analysis of omission/commission

The following picture shows the test omission rate and predicted area as a function of the cumulative threshold, averaged over the replicate runs. The omission rate should be close to the predicted omission, because of the definition of the cumulative threshold.



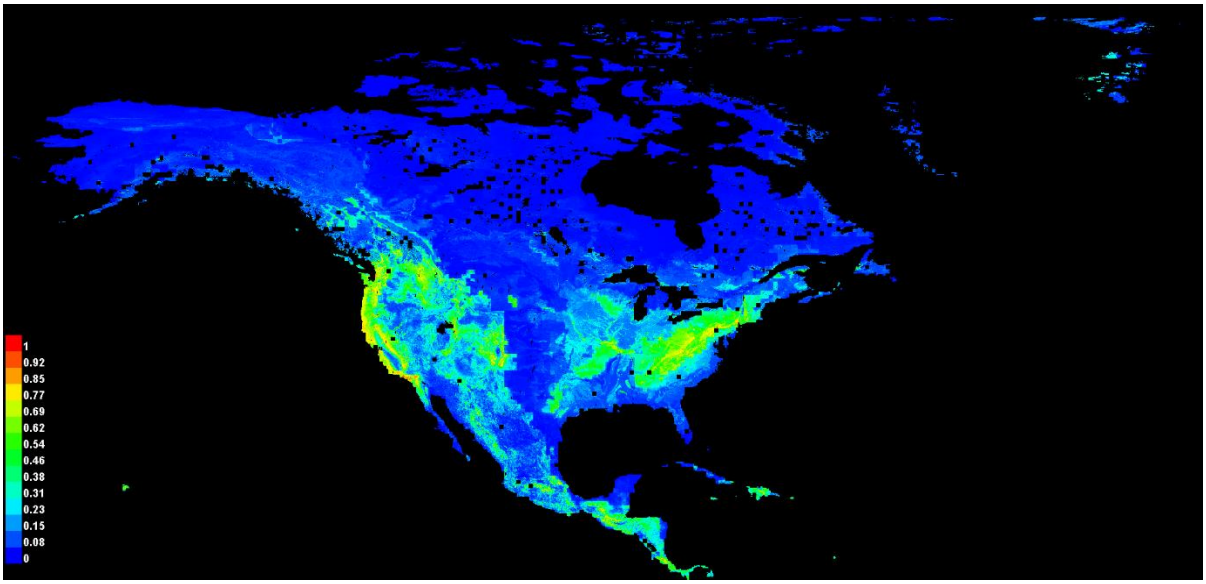
The next picture is the receiver operating characteristic (ROC) curve for the same data, again averaged over the replicate runs. Note that the specificity is defined

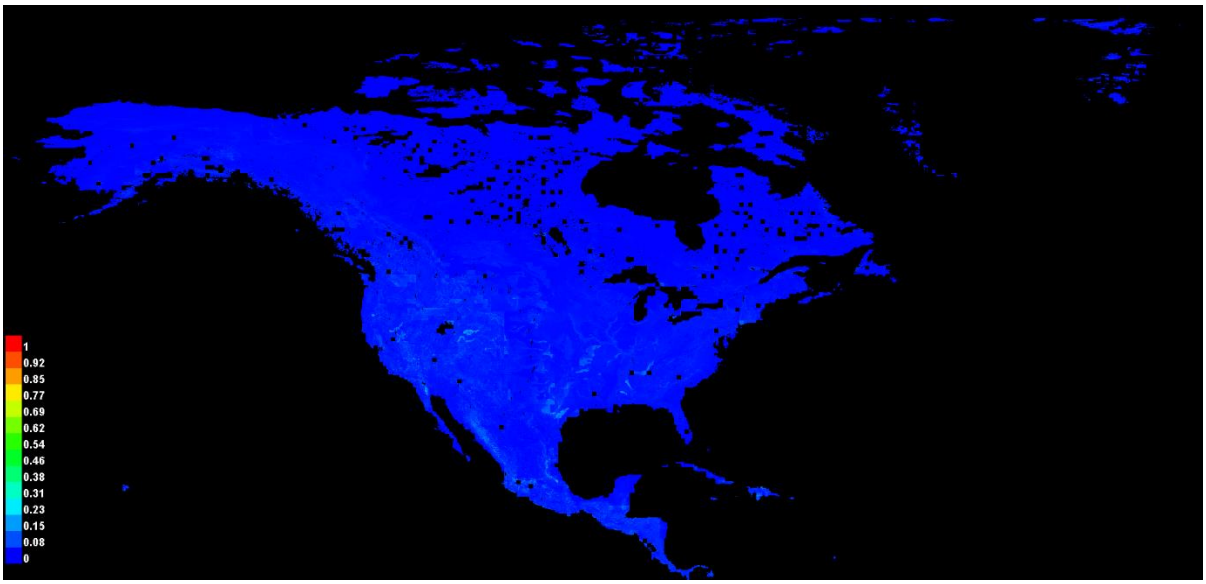
using predicted area, rather than true commission (see the paper by Phillips, Anderson and Schapire cited on the help page for discussion of what this means). The average test AUC for the replicate runs is 0.901, and the standard deviation is 0.005.



Pictures of the model

The following two pictures show the point-wise mean and standard deviation of the 5 output grids. Other available summary grids are min, max and median.

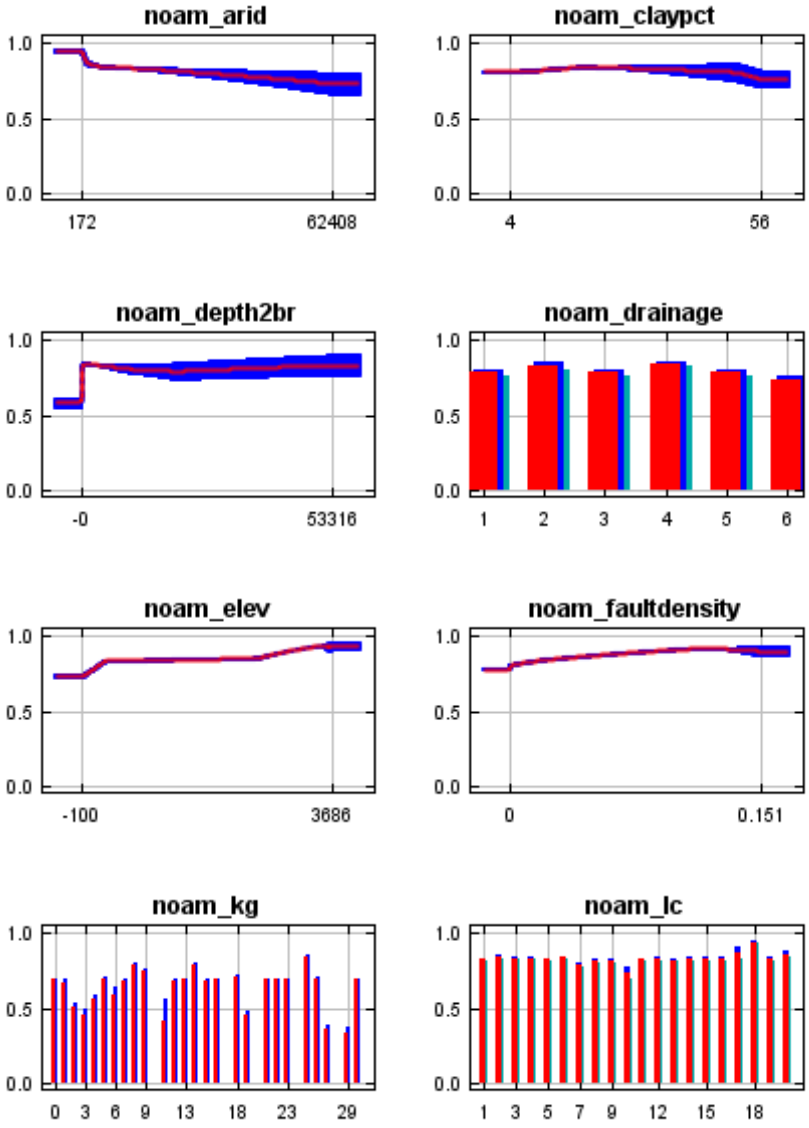


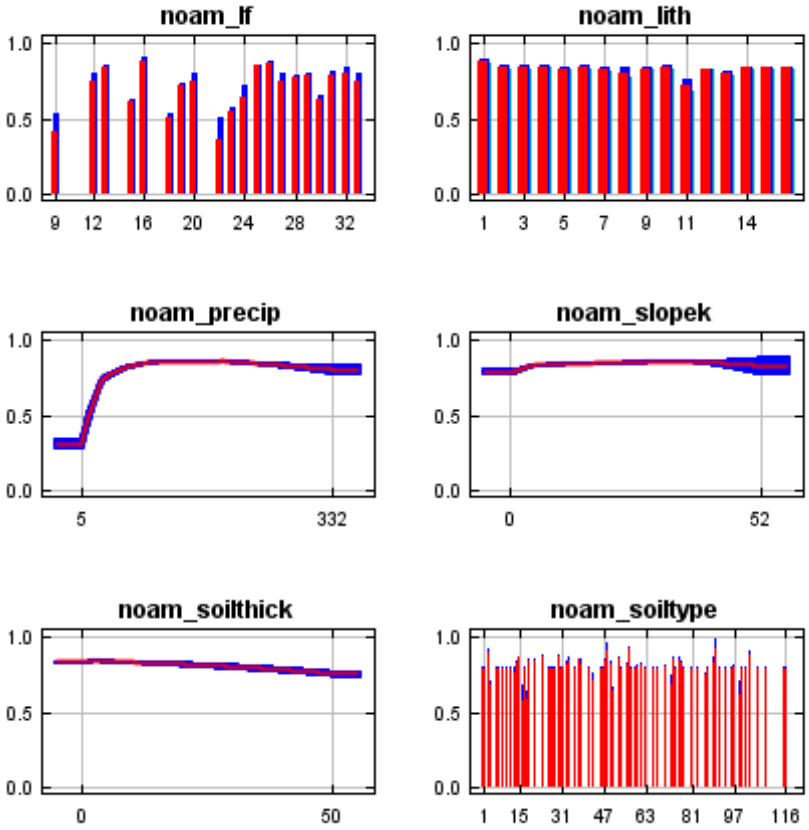


Response curves

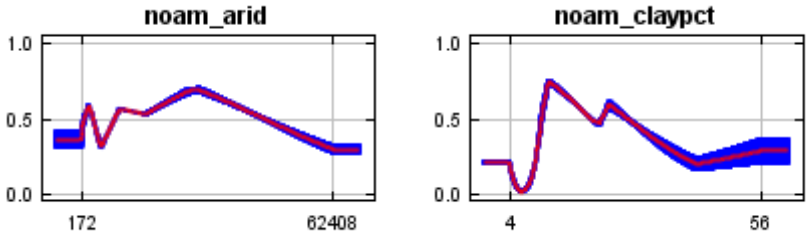
These curves show how each environmental variable affects the Maxent prediction. The curves show how the predicted probability of presence changes as each environmental variable is varied, keeping all other environmental variables at their average sample value. Click on a response curve to see a larger version. Note that the curves can be hard to interpret if you have strongly correlated variables, as the model may depend on the correlations in ways that are not evident in the curves. In

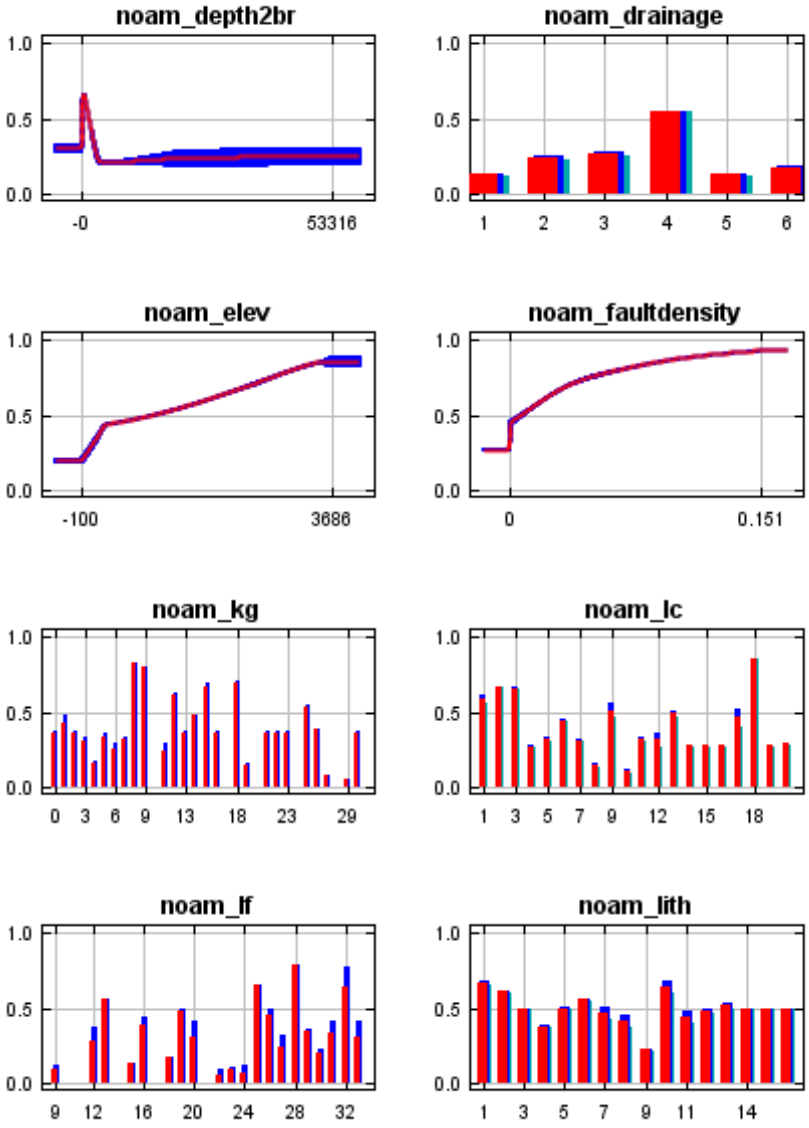
other words, the curves show the marginal effect of changing exactly one variable, whereas the model may take advantage of sets of variables changing together. The curves show the mean response of the 5 replicate Maxent runs (red) and and the mean +/- one standard deviation (blue, two shades for categorical variables).

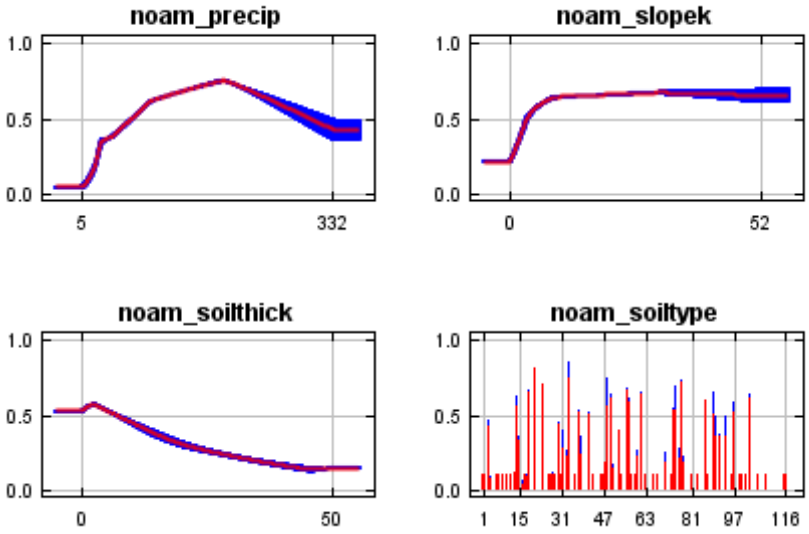




In contrast to the above marginal response curves, each of the following curves represents a different model, namely, a Maxent model created using only the corresponding variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other variables. They may be easier to interpret if there are strong correlations between variables.




noam_depth2br
1.0
0.5
0.0
-0
53316
noam_drainage
1.0
0.5
0.0
1 2 3 4 5 6
noam_elev
1.0
0.5
0.0
-100
3686
noam_faultdensity
1.0
0.5
0.0
0
0.151
noam_kg
1.0
0.5
0.0
0 3 6 9 13 18 23 29
noam_lc
1.0
0.5
0.0
1 3 5 7 9 12 15 18
noam_lf
1.0
0.5
0.0
9 12 16 20 24 28 32
noam_lith
1.0
0.5
0.0
1 3 5 7 9 11 14



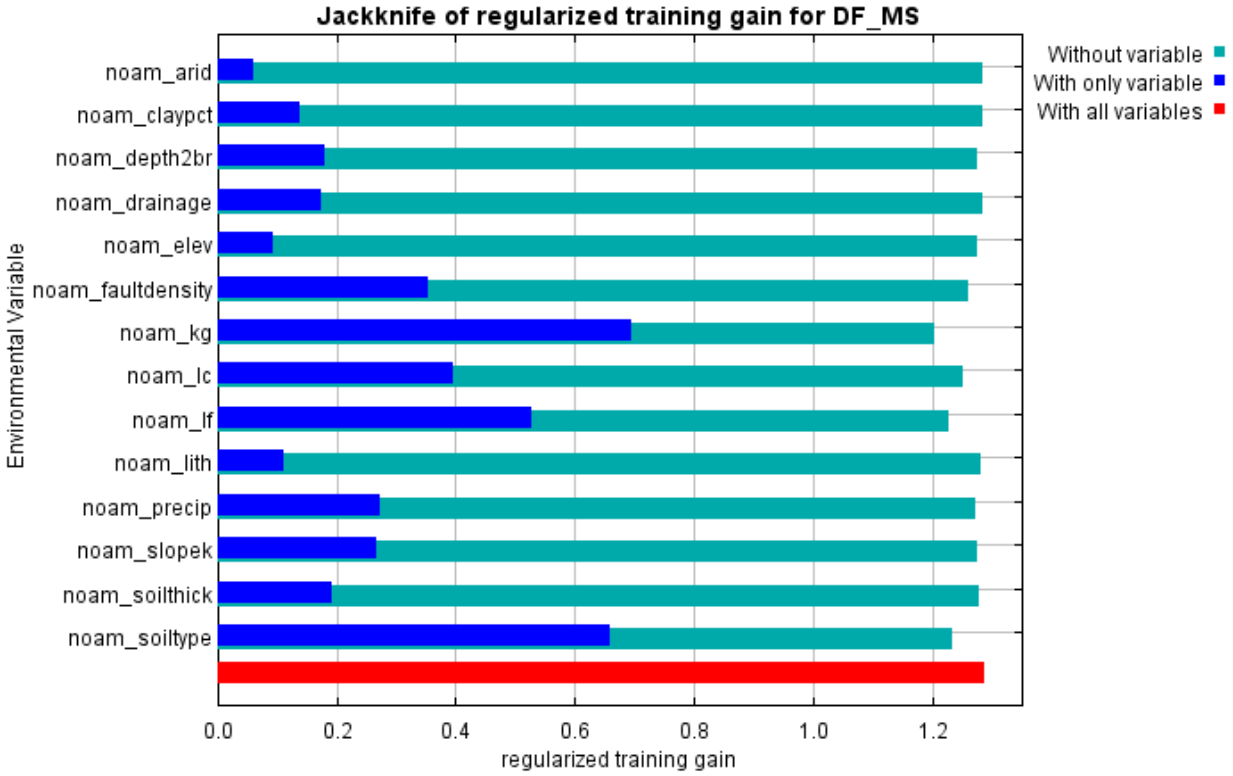
Analysis of variable contributions

The following table gives estimates of relative contributions of the environmental variables to the Maxent model. To determine the first estimate, in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the corresponding variable, or subtracted from it if the change to the absolute value of lambda is negative. For the second estimate, for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is reevaluated on the permuted data, and the resulting drop in training AUC is shown in the table, normalized to percentages. As with the variable jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated. Values shown are averages over replicate runs.

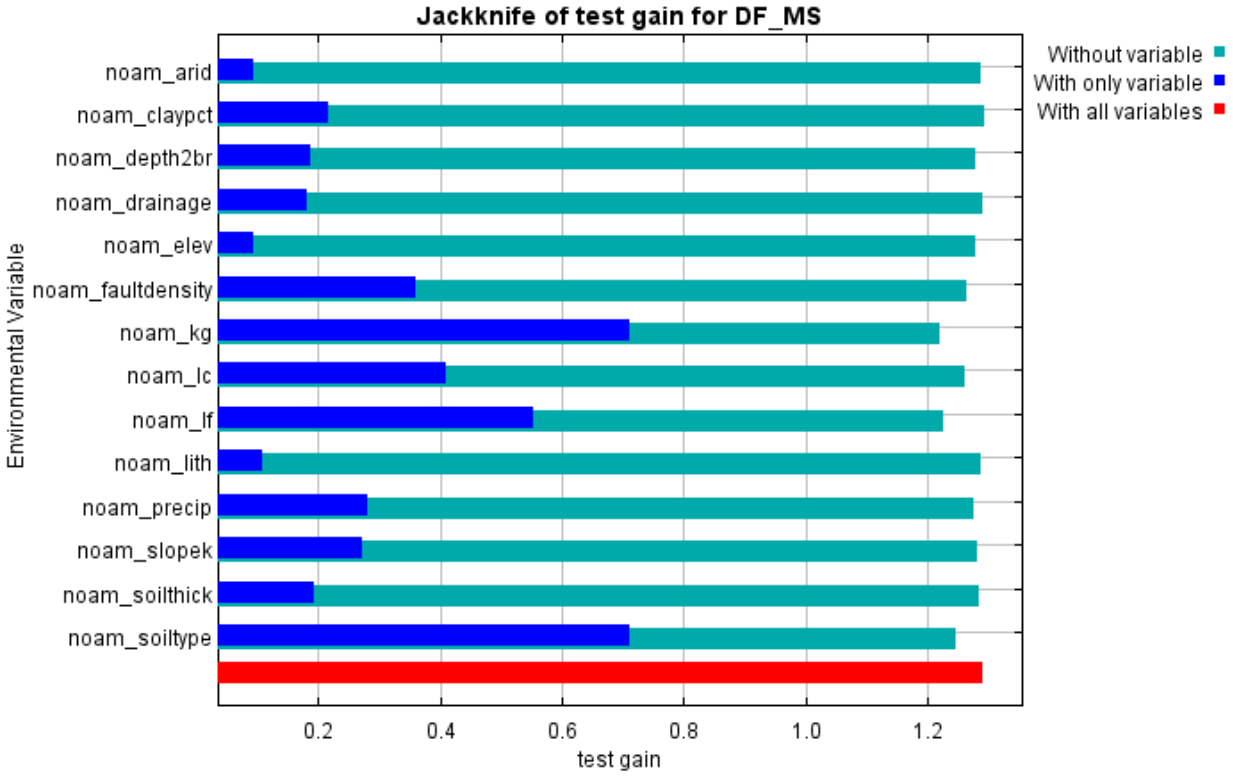
Variable	Percent contribution	Permutation importance
noam_kg	31.7	30.3
noam_faultdensity	14.3	4

noam_lf	10	23.9
noam_soiltype	9.2	10.5
noam_slopek	7.3	2
noam_precip	6.7	15.9
noam_lc	5.9	4
noam_drainage	5.7	1.8
noam_soilthick	5.2	1.3
noam_depth2br	1.7	2.1
noam_lith	1	0.9
noam_elev	0.8	1.4
noam_claypct	0.4	0.3
noam_arid	0.2	1.7

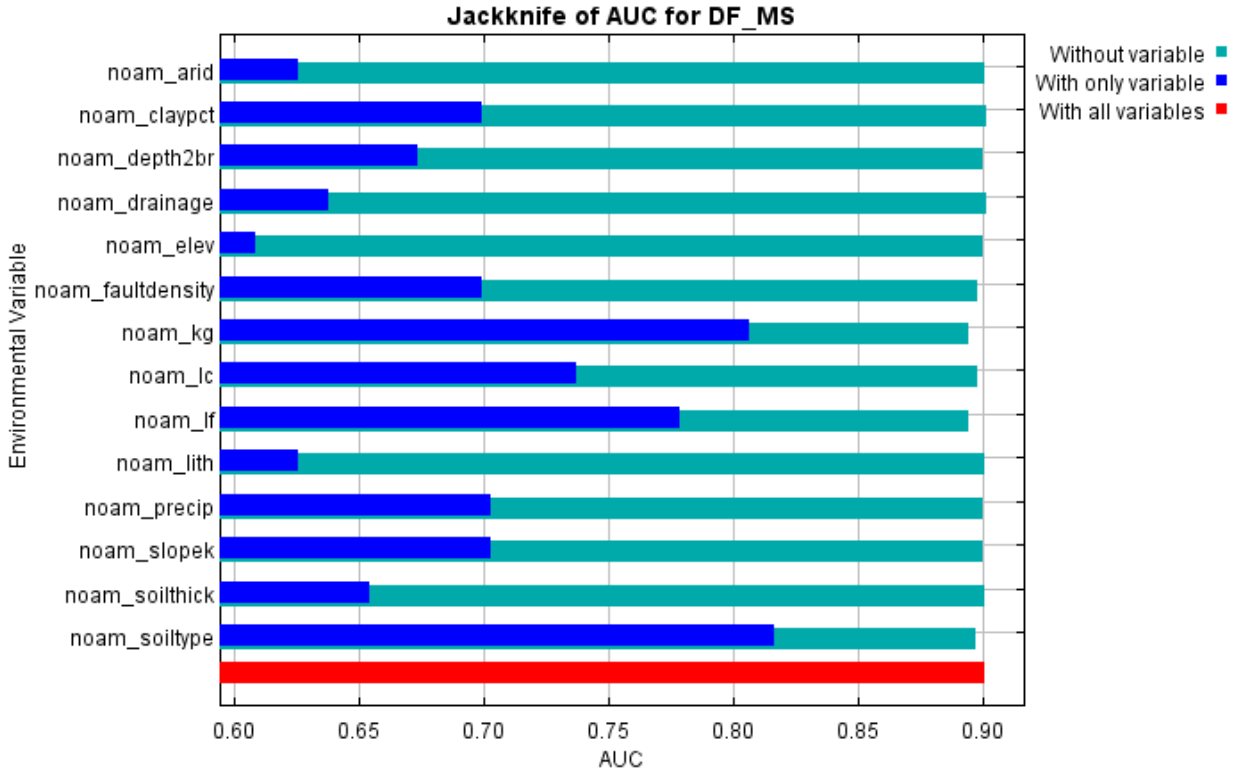
The following picture shows the results of the jackknife test of variable importance. The environmental variable with highest gain when used in isolation is noam_kg, which therefore appears to have the most useful information by itself. The environmental variable that decreases the gain the most when it is omitted is noam_kg, which therefore appears to have the most information that isn't present in the other variables. Values shown are averages over replicate runs.



The next picture shows the same jackknife test, using test gain instead of training gain. Note that conclusions about which variables are most important can change, now that we're looking at test data.



Lastly, we have the same jackknife test, using AUC on test data.



Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E DF_MS responsecurves jackknife outputformat=logistic "outputdirectory=C:_PhD_PhD PROJECT\MaxEntropy_Stats\MaxEntResults\ContinentalAnalysis\NoAmerica2" "samplesfile=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EventData_samples\No_America_TRAIN.csv" "environmentallayers=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EnvLayers\No_America" randomseed replicates=5 -N noam_slopeb -N noam_slopec -N noam_sloped -N noam_slopee -N noam_slopef -N noam_slopeg -N noam_slopei -t noam_drainage -t noam_kg -t noam_lc -t noam_lf -t noam_lith -t noam_soiltype

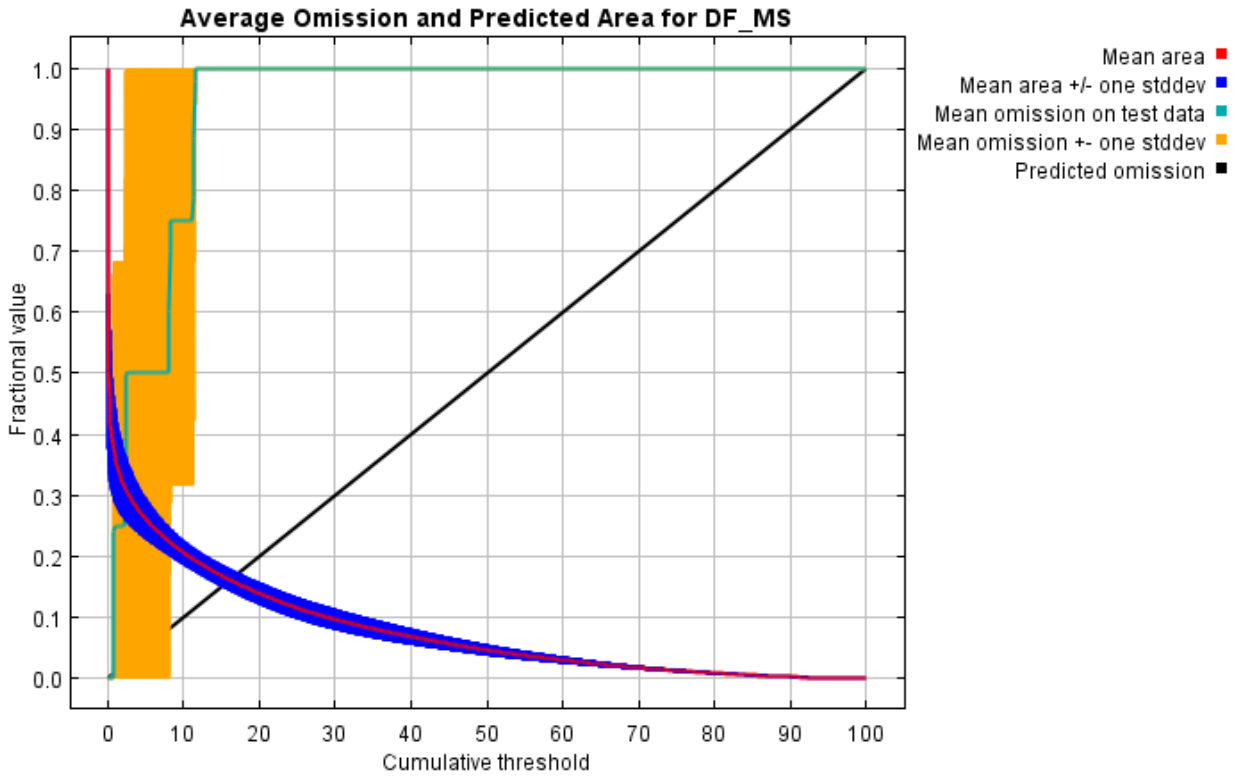
E.7 Oceania

Replicated maxent model for DF_MS

This page summarizes the results of 4-fold cross-validation for DF_MS, created Wed Jun 02 13:55:18 CEST 2021 using Maxent version 3.4.4. The individual models are here: [0] [1] [2] [3]

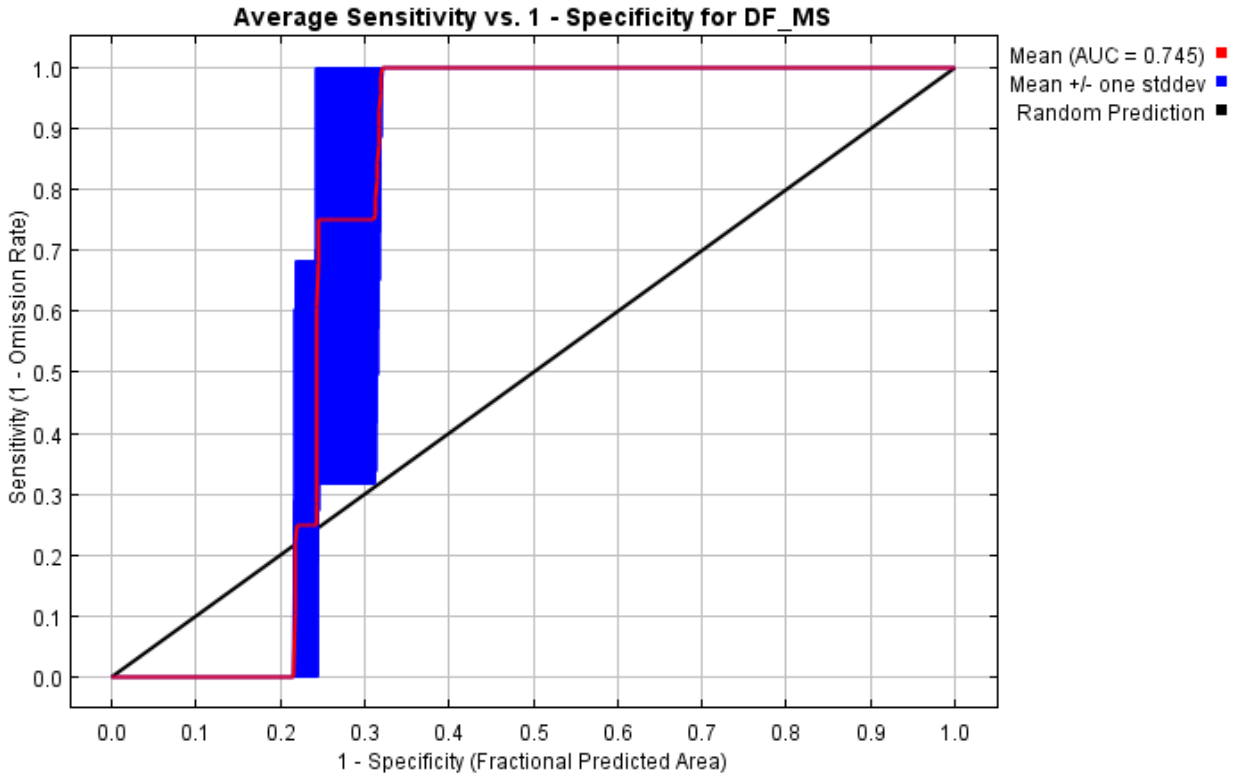
Analysis of omission/commission

The following picture shows the test omission rate and predicted area as a function of the cumulative threshold, averaged over the replicate runs. The omission rate should be close to the predicted omission, because of the definition of the cumulative threshold.



The next picture is the receiver operating characteristic (ROC) curve for the same

data, again averaged over the replicate runs. Note that the specificity is defined using predicted area, rather than true commission (see the paper by Phillips, Anderson and Schapire cited on the help page for discussion of what this means). The average test AUC for the replicate runs is 0.745, and the standard deviation is 0.037.

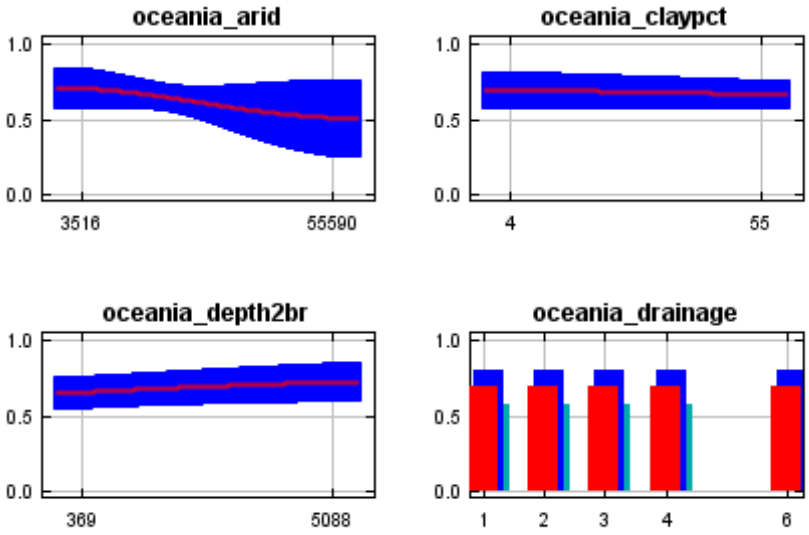


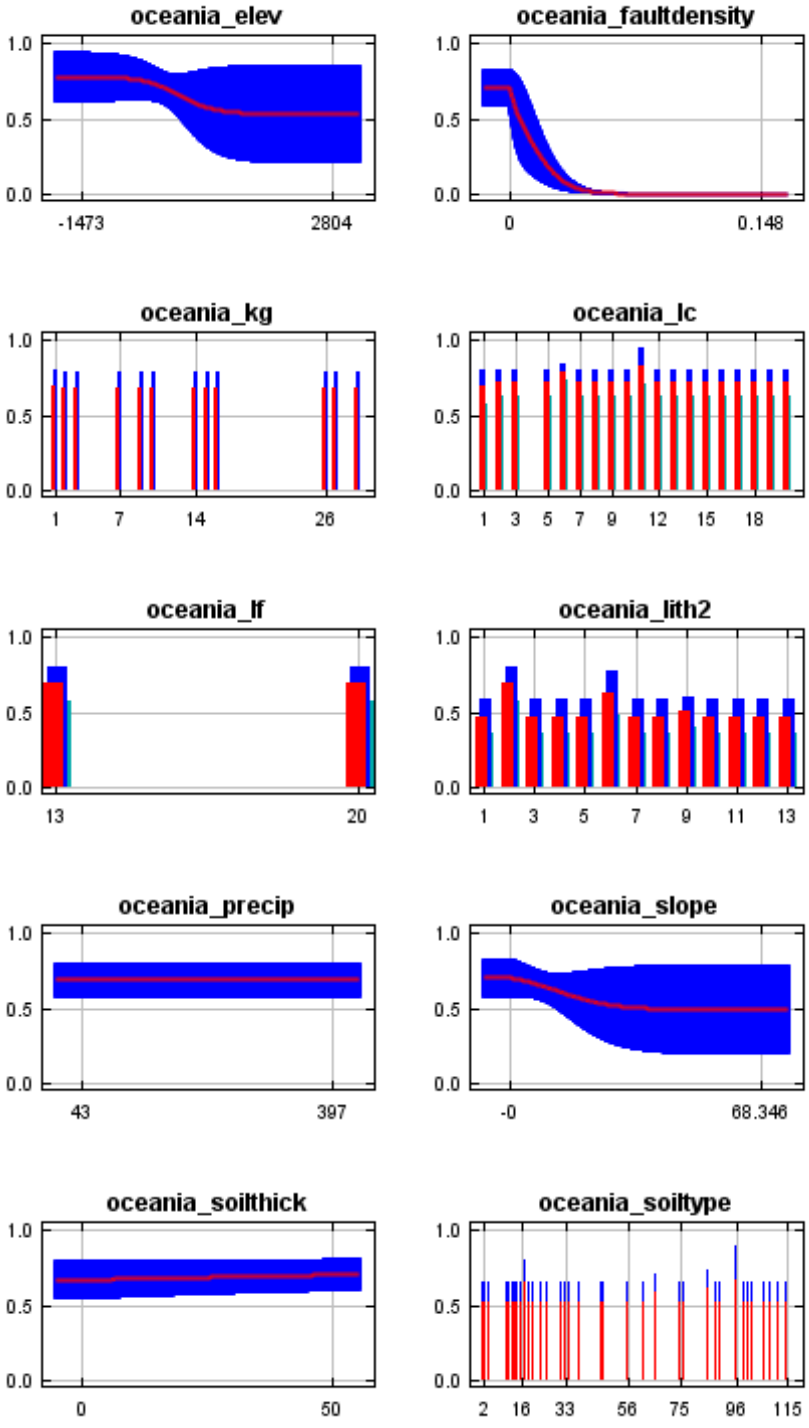
Pictures of the model

The following two pictures show the point-wise mean and standard deviation of the 4 output grids. Other available summary grids are min, max and median.

Response curves

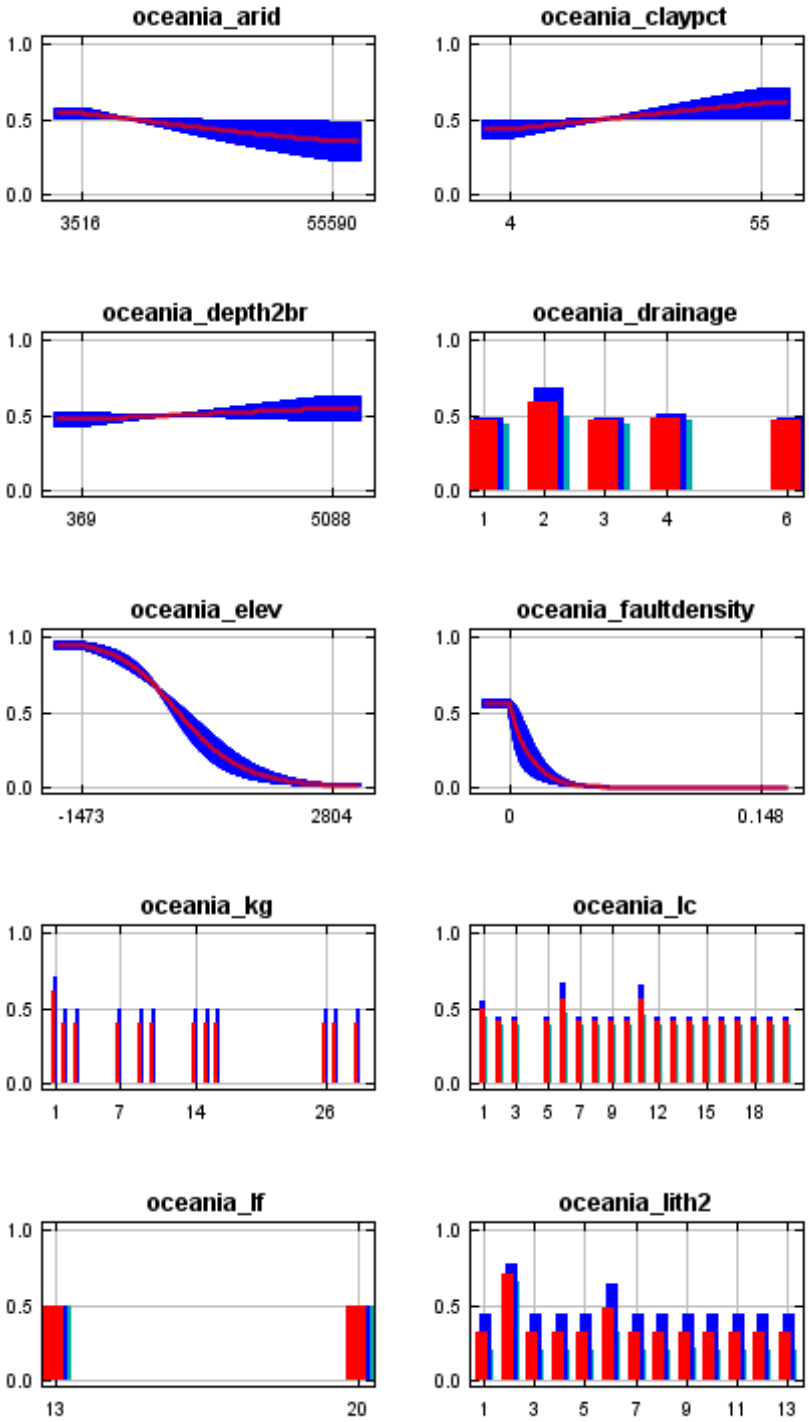
These curves show how each environmental variable affects the Maxent prediction. The curves show how the predicted probability of presence changes as each environmental variable is varied, keeping all other environmental variables at their average sample value. Click on a response curve to see a larger version. Note that the curves can be hard to interpret if you have strongly correlated variables, as the model may depend on the correlations in ways that are not evident in the curves. In other words, the curves show the marginal effect of changing exactly one variable, whereas the model may take advantage of sets of variables changing together. The curves show the mean response of the 4 replicate Maxent runs (red) and and the mean +/- one standard deviation (blue, two shades for categorical variables).

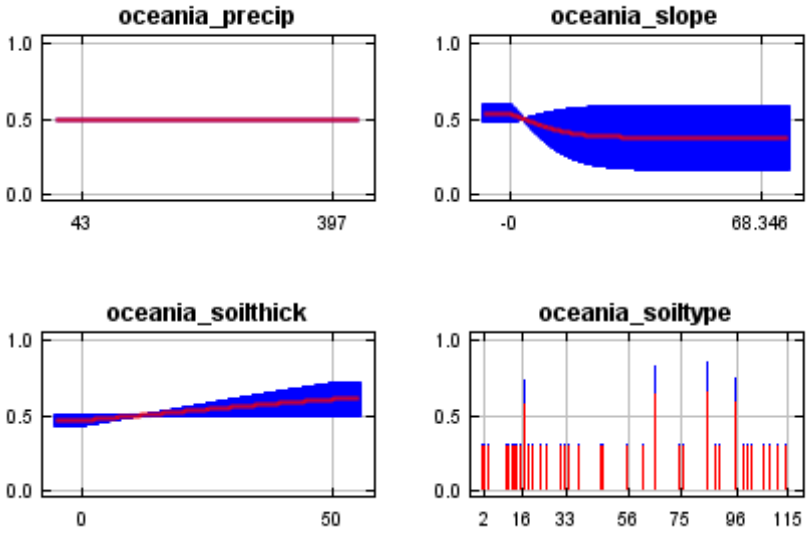




In contrast to the above marginal response curves, each of the following curves represents a different model, namely, a Maxent model created using only the corresponding variable. These plots reflect the dependence of predicted suitability

both on the selected variable and on dependencies induced by correlations between the selected variable and other variables. They may be easier to interpret if there are strong correlations between variables.





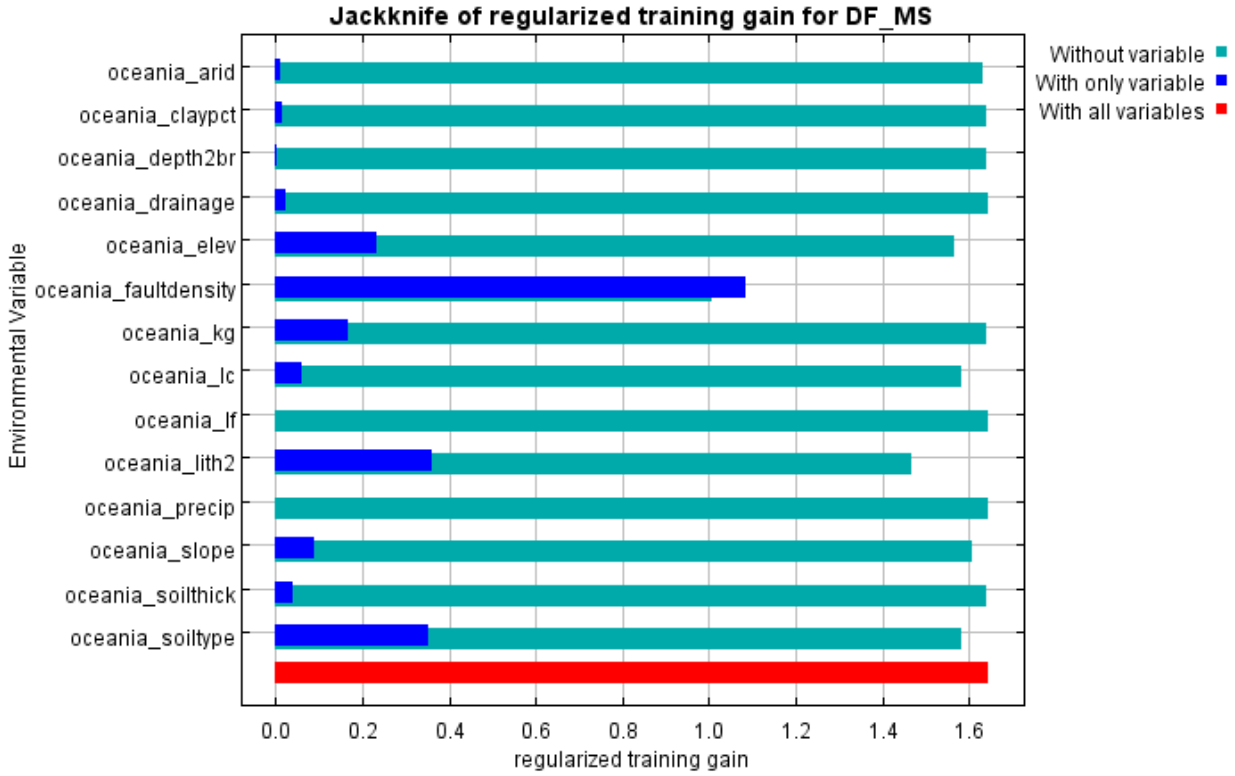
Analysis of variable contributions

The following table gives estimates of relative contributions of the environmental variables to the Maxent model. To determine the first estimate, in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the corresponding variable, or subtracted from it if the change to the absolute value of lambda is negative. For the second estimate, for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is reevaluated on the permuted data, and the resulting drop in training AUC is shown in the table, normalized to percentages. As with the variable jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated. Values shown are averages over replicate runs.

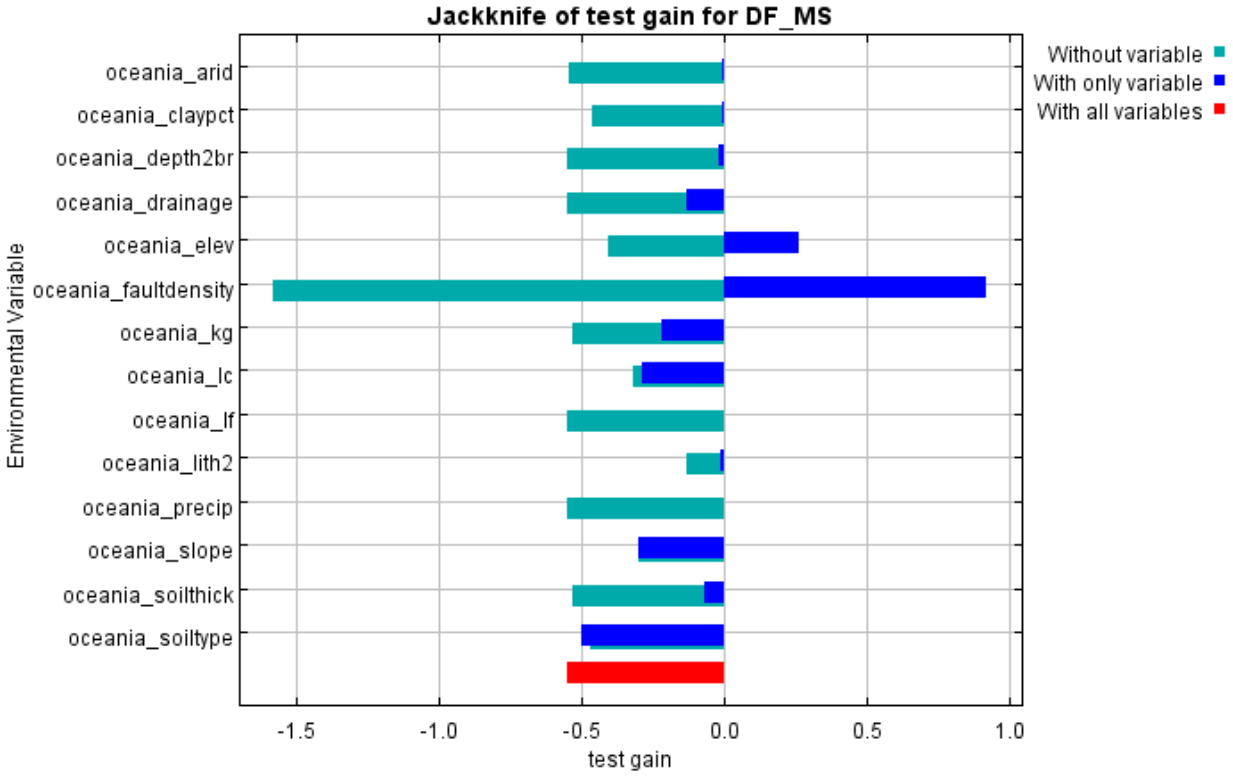
Variable	Percent contribution	Permutation importance
oceania_faultdensity	59.9	77.1
oceania_lith2	16.8	2.1

oceania_soiltype	8.1	3.9
oceania_elev	5.1	6.1
oceania_lc	4.2	1.8
oceania_slope	2.9	8.6
oceania_kg	1.3	0
oceania_soilthick	1.2	0.2
oceania_arid	0.3	0
oceania_depth2br	0.1	0
oceania_claypct	0.1	0.1
oceania_lf	0	0
oceania_drainage	0	0
oceania_precip	0	0

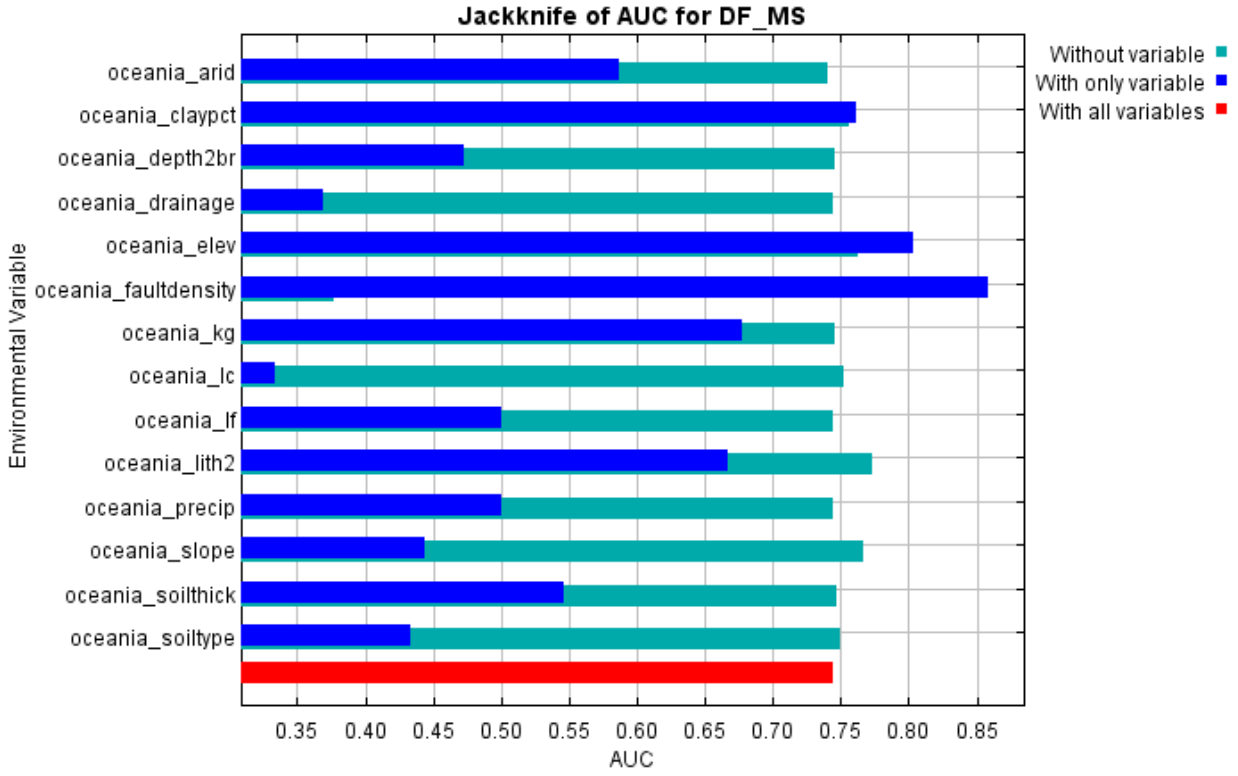
The following picture shows the results of the jackknife test of variable importance. The environmental variable with highest gain when used in isolation is oceania_faultdensity, which therefore appears to have the most useful information by itself. The environmental variable that decreases the gain the most when it is omitted is oceania_faultdensity, which therefore appears to have the most information that isn't present in the other variables. Values shown are averages over replicate runs.



The next picture shows the same jackknife test, using test gain instead of training gain. Note that conclusions about which variables are most important can change, now that we're looking at test data.



Lastly, we have the same jackknife test, using AUC on test data.



Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E DF_MS responsecurves jackknife outputformat=logistic "outputdirectory=C:_PhD_PhD PROJECT\MaxEntropy_Stats\MaxEntResults\ContinentalAnalysis\Oceania" "samplesfile=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EventData_samples\Oceania_TRAIN.csv" "environmentallayers=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EnvLayers\Oceania" randomseed replicates=5 writeplotdata -t oceania_drainage -t oceania_kg -t oceania_lc -t oceania_lf -t oceania_lith2 -t oceania_soiltype

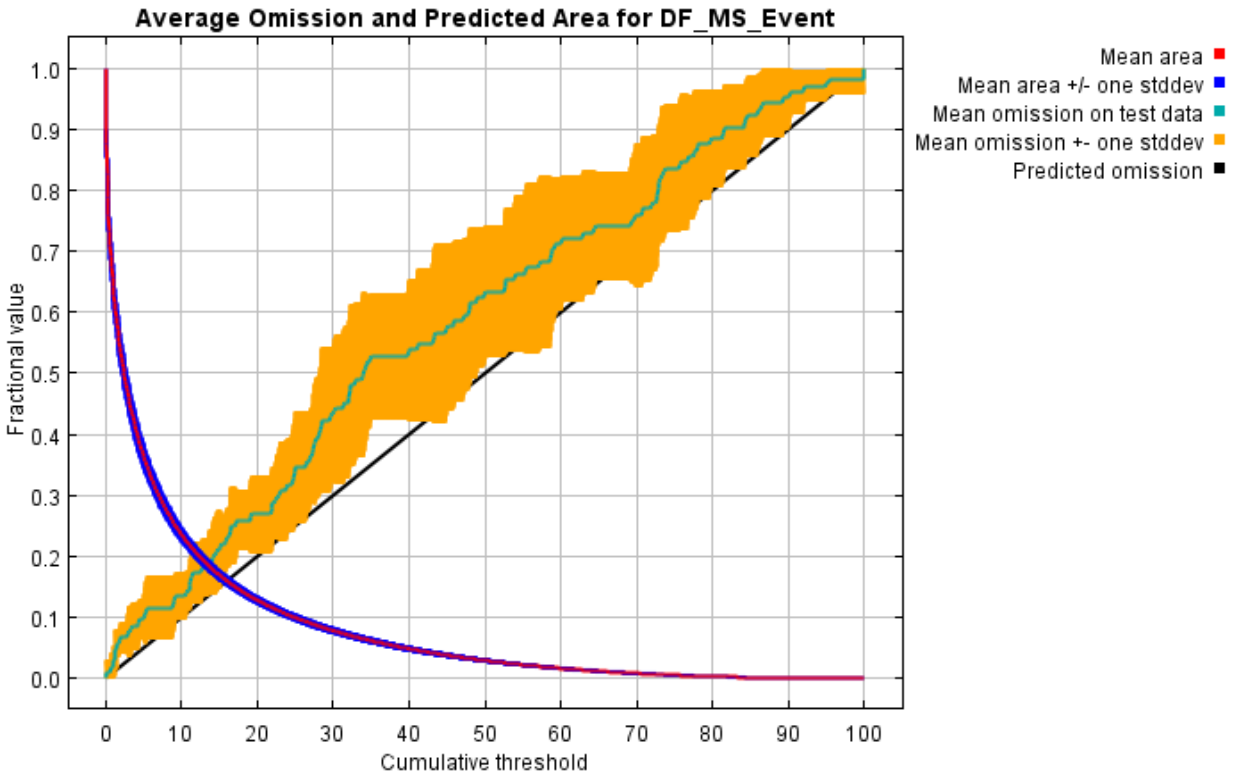
E.8 South America

Replicated maxent model for DF_MS_Event

This page summarizes the results of 5-fold cross-validation for DF_MS_Event, created Sat Jun 12 12:44:39 CEST 2021 using Maxent version 3.4.4. The individual models are here: [0] [1] [2] [3] [4]

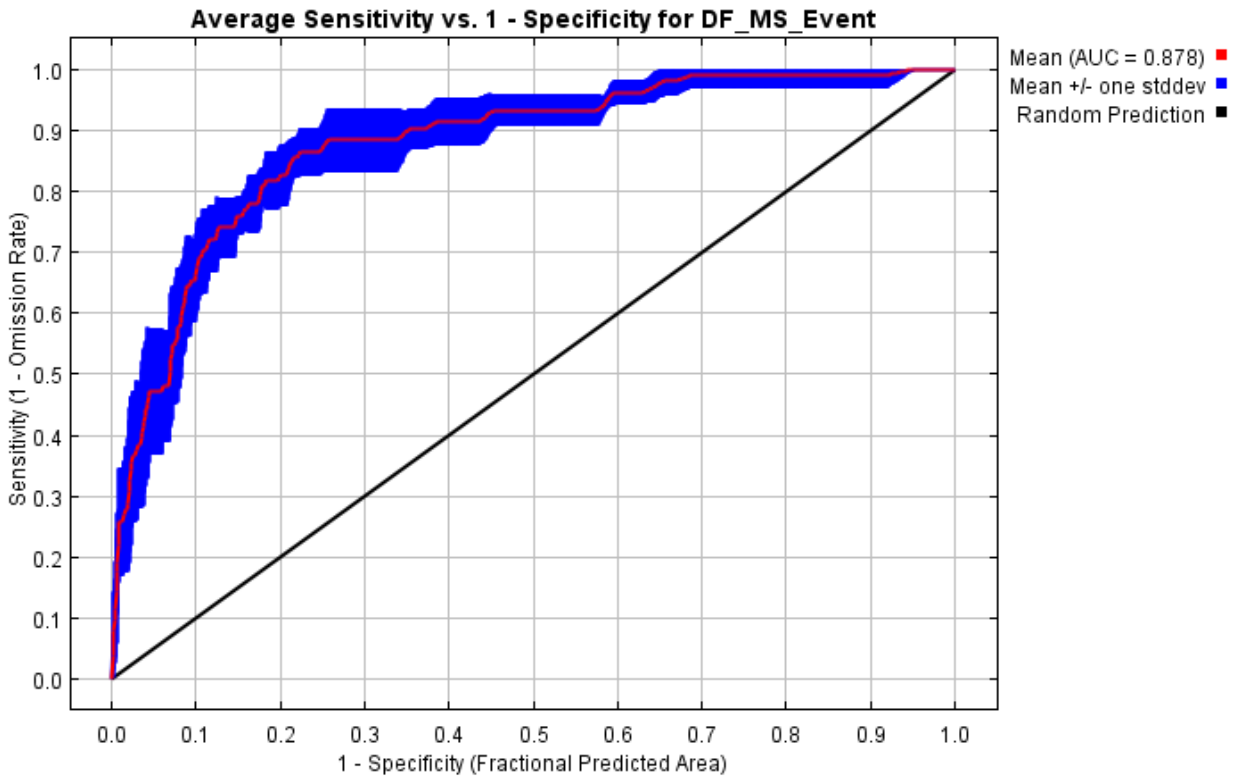
Analysis of omission/commission

The following picture shows the test omission rate and predicted area as a function of the cumulative threshold, averaged over the replicate runs. The omission rate should be close to the predicted omission, because of the definition of the cumulative threshold.



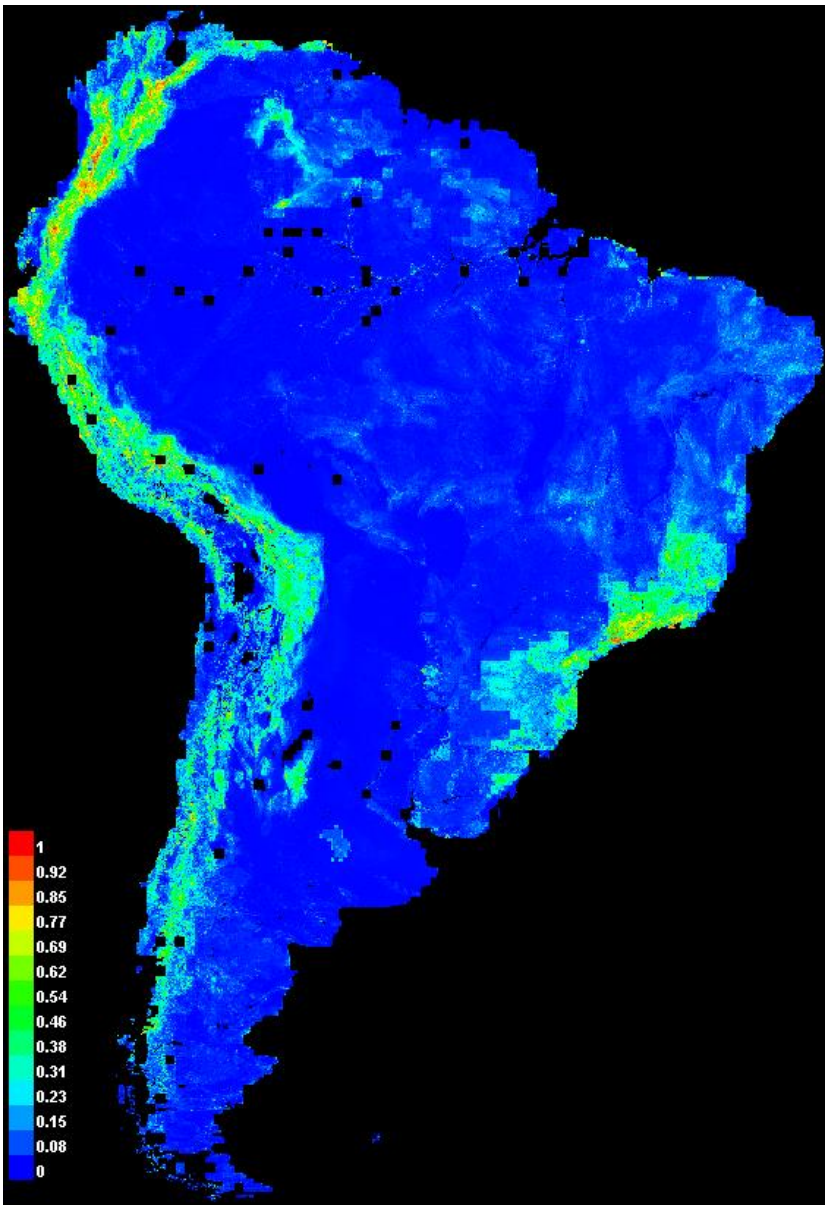
The next picture is the receiver operating characteristic (ROC) curve for the same

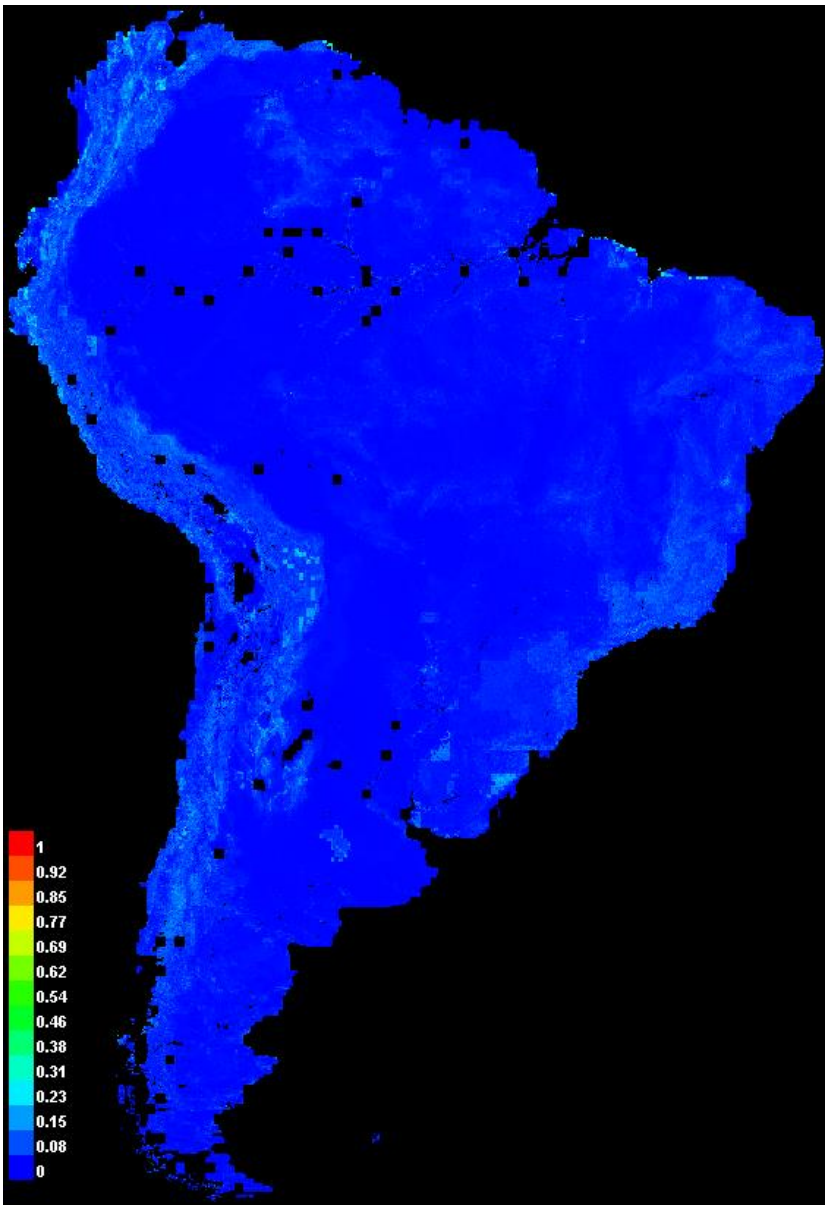
data, again averaged over the replicate runs. Note that the specificity is defined using predicted area, rather than true commission (see the paper by Phillips, Anderson and Schapire cited on the help page for discussion of what this means). The average test AUC for the replicate runs is 0.878, and the standard deviation is 0.021.



Pictures of the model

The following two pictures show the point-wise mean and standard deviation of the 5 output grids. Other available summary grids are min, max and median.

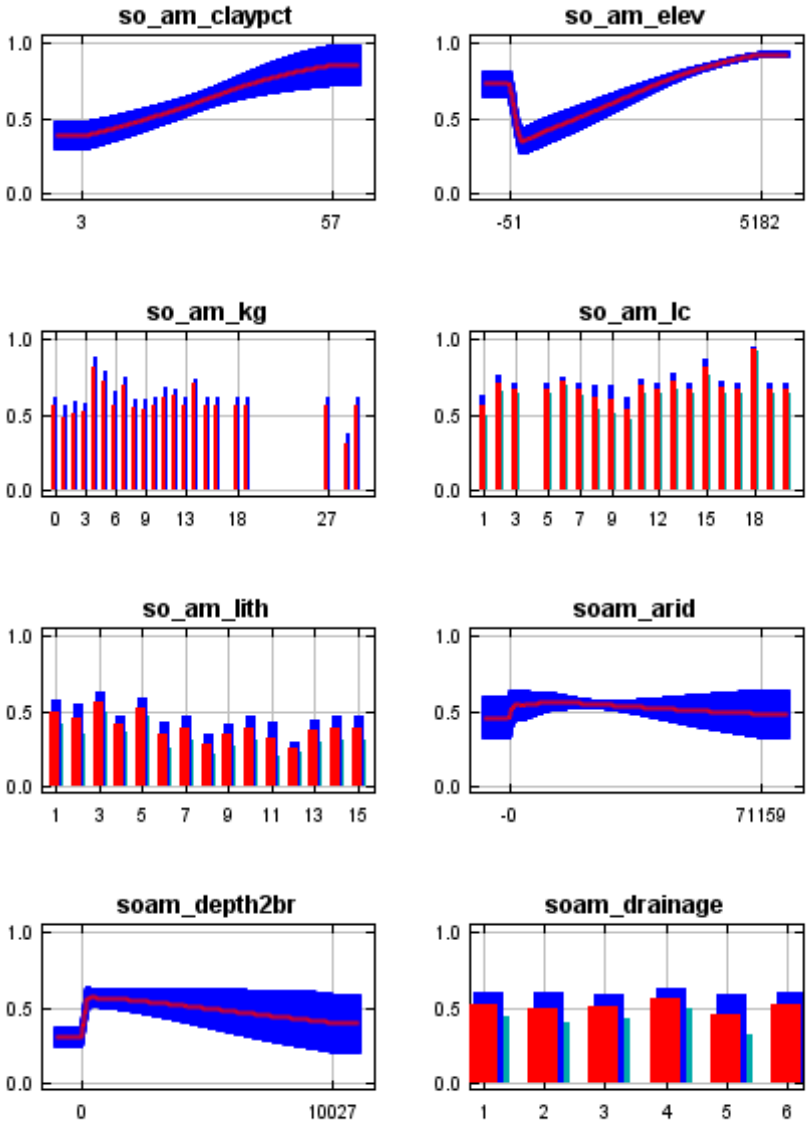

1
0.92
0.85
0.77
0.69
0.62
0.54
0.46
0.38
0.31
0.23
0.15
0.08
0

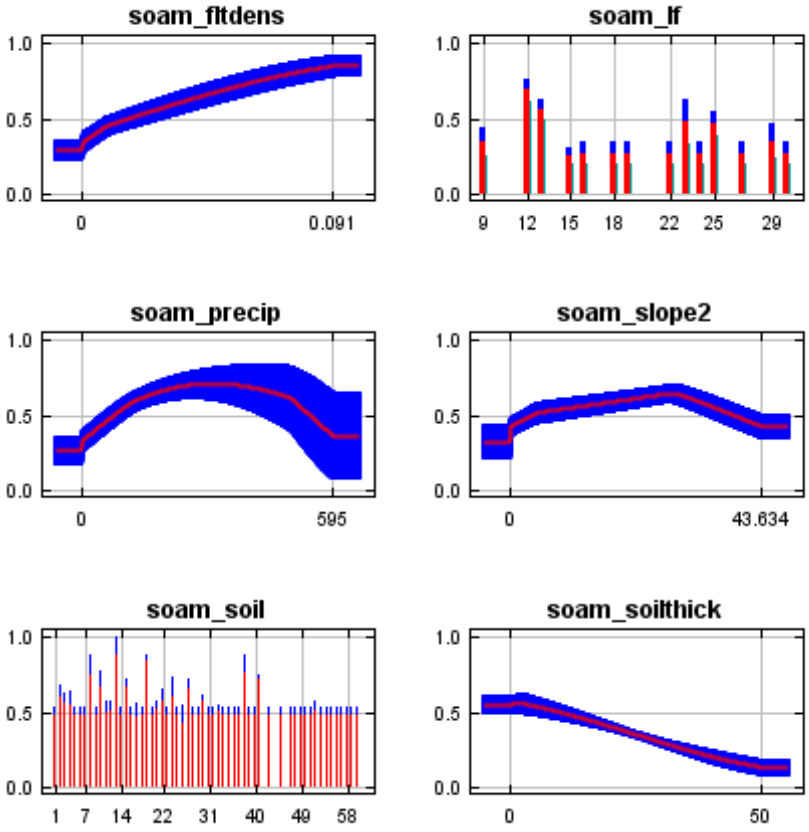


Response curves

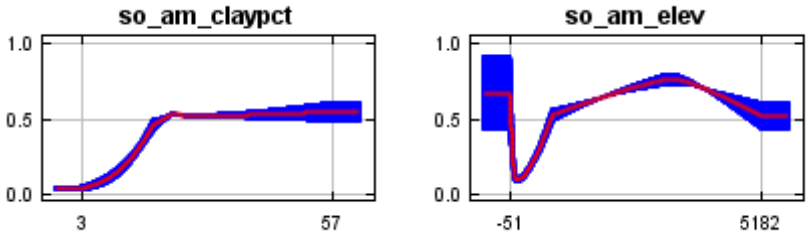
These curves show how each environmental variable affects the Maxent prediction. The curves show how the predicted probability of presence changes as each environmental variable is varied, keeping all other environmental variables at their average sample value. Click on a response curve to see a larger version. Note that the curves can be hard to interpret if you have strongly correlated variables, as the

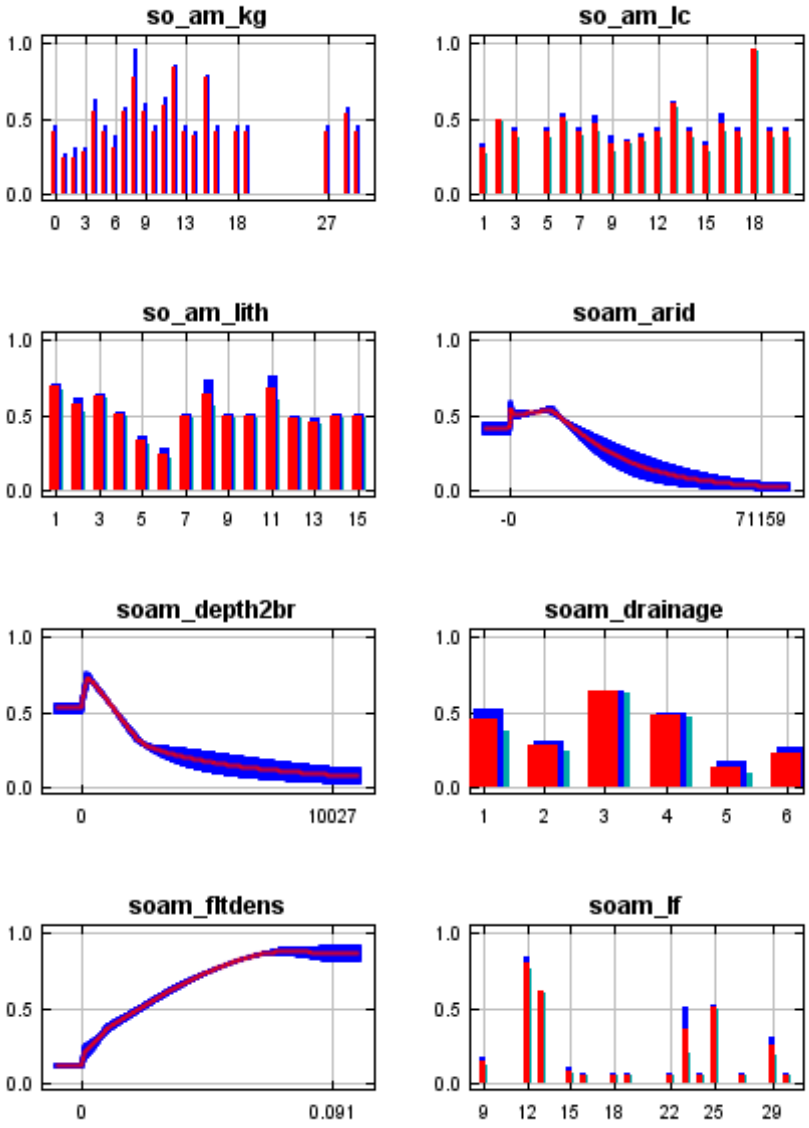
model may depend on the correlations in ways that are not evident in the curves. In other words, the curves show the marginal effect of changing exactly one variable, whereas the model may take advantage of sets of variables changing together. The curves show the mean response of the 5 replicate Maxent runs (red) and and the mean +/- one standard deviation (blue, two shades for categorical variables).

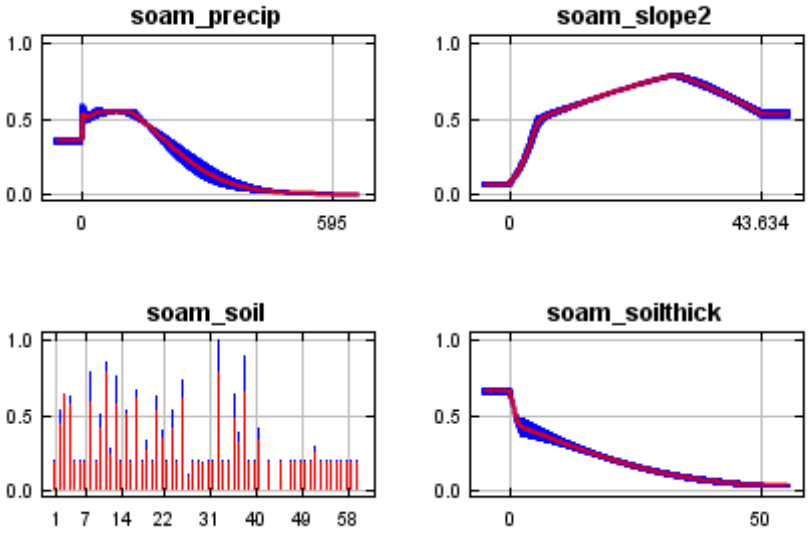




In contrast to the above marginal response curves, each of the following curves represents a different model, namely, a Maxent model created using only the corresponding variable. These plots reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by correlations between the selected variable and other variables. They may be easier to interpret if there are strong correlations between variables.




so_am_kg
1.0
0.5
0.0
0 3 6 9 13 18 27
so_am_lc
1.0
0.5
0.0
1 3 5 7 9 12 15 18
so_am_lith
1.0
0.5
0.0
1 3 5 7 9 11 13 15
soam_arid
1.0
0.5
0.0
-0 71159
soam_depth2br
1.0
0.5
0.0
0 10027
soam_drainage
1.0
0.5
0.0
1 2 3 4 5 6
soam_fltdens
1.0
0.5
0.0
0 0.091
soam_lf
1.0
0.5
0.0
9 12 15 18 22 25 29



Analysis of variable contributions

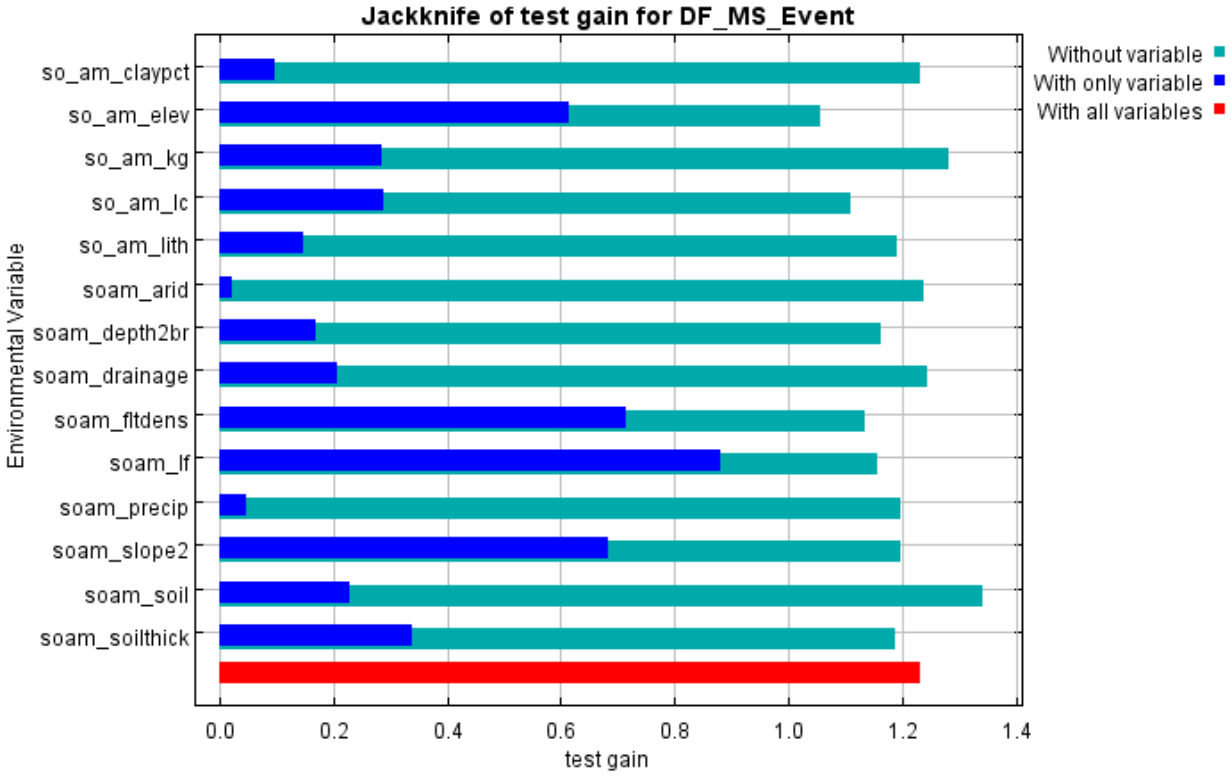
The following table gives estimates of relative contributions of the environmental variables to the Maxent model. To determine the first estimate, in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the corresponding variable, or subtracted from it if the change to the absolute value of lambda is negative. For the second estimate, for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is reevaluated on the permuted data, and the resulting drop in training AUC is shown in the table, normalized to percentages. As with the variable jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated. Values shown are averages over replicate runs.

Variable	Percent contribution	Permutation importance
soam_fltdens	25.5	8.3
soam_lf	19.4	9.7

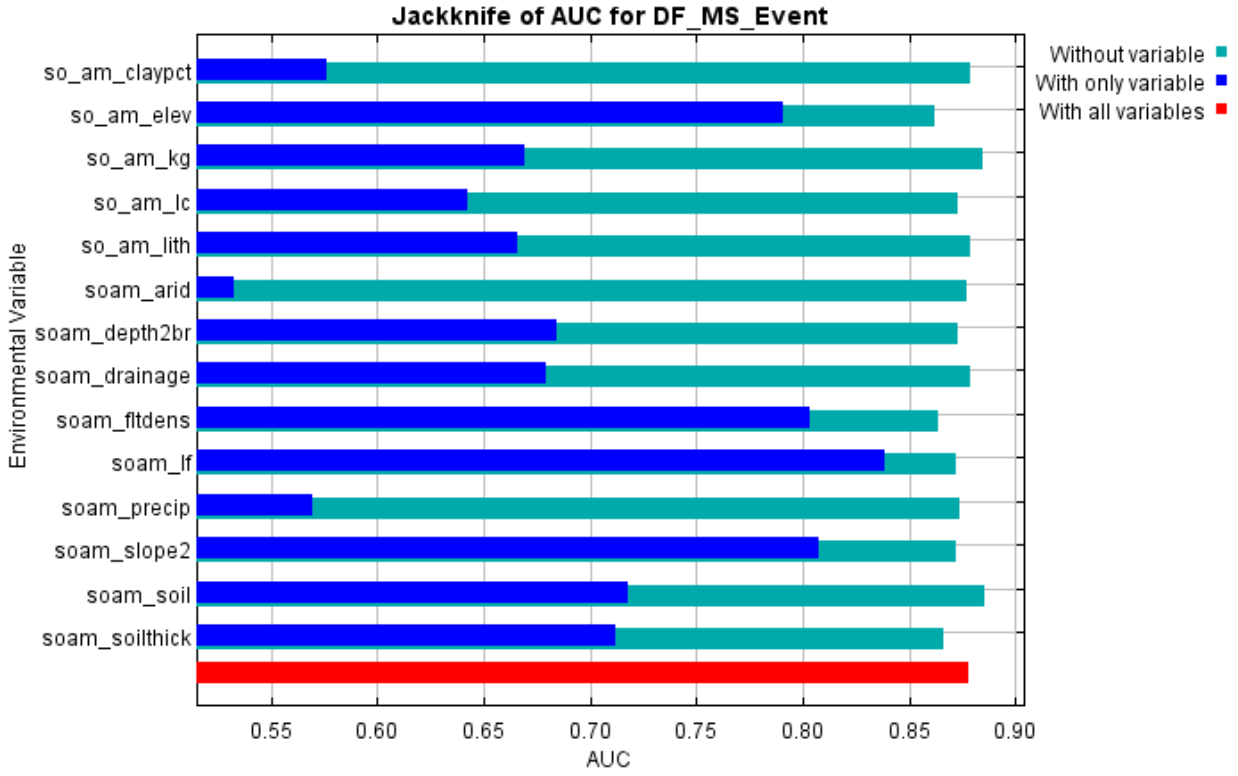
so_am_lc	11.9	5.3
soam_slope2	11.3	4.5
so_am_elev	7.9	16.6
so_am_kg	6.4	7.4
soam_soilthick	4.8	17.6
soam_soil	4.5	7.6
so_am_lith	4.2	4.1
soam_depth2br	1.4	7.8
soam_precip	1.4	6.8
so_am_claypct	0.9	2.5
soam_arid	0.3	1.2
soam_drainage	0.2	0.6

The following picture shows the results of the jackknife test of variable importance. The environmental variable with highest gain when used in isolation is soam_lf, which therefore appears to have the most useful information by itself. The environmental variable that decreases the gain the most when it is omitted is so_am_elev, which therefore appears to have the most information that isn't present in the other variables. Values shown are averages over replicate runs.

The next picture shows the same jackknife test, using test gain instead of training gain. Note that conclusions about which variables are most important can change, now that we're looking at test data.



Lastly, we have the same jackknife test, using AUC on test data.



Command line to repeat this species model: java density.MaxEnt nowarnings noprefixes -E "" -E DF_MS_Event responsecurves jackknife outputformat=logistic "outputdirectory=C:_PhD_PhD PROJECT\MaxEntropy_Stats\MaxEntResults\ContinentalAnalysis\SoAmerica" "samplesfile=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EventData_samples\So_America_TRAIN.csv" "environmentallayers=C:_PhD_PhD PROJECT\MaxEntropy_Stats\EnvLayers\So_America" randomseed noaskoverwrite replicates=5 -t so_am_kg -t so_am_lc -t so_am_lith -t soam_drainage -t soam_lf -t soam_soil

APPENDIX F Susceptibility Analyses

F.1 Comparison of global susceptibility models with verification data (TEST1)

FREQUENCY RATIO						PROBABILITY DENSITY					MAXIMUM ENTROPY				
	# Factors Present	#TEST1 Events	% TEST1 of Total	TRAIN	TRAIN %	# Factors Present	#TEST1 Events	% TEST1 of Total	TRAIN	TRAIN %	MaxEnt 5 Equal Interval Susc Classes	#TEST1 Events	% TEST1 of Total	TRAIN	TRAIN %
Low	0	**3013**	52.9%	1344	58.6%	0	**4024**	70.7%	1411	61.5%	0 - Very Low	**1638**	28.8%	512	22.3%
Low	1	**1353**	23.8%	646	28.2%	1	**62**	1.1%	7	0.3%	1 - Very Low	**42**	0.7%	94	4.1%
Medium	2	**791**	13.9%	169	7.4%	2	**26**	0.5%	33	1.4%	2 - Low	**821**	14.4%	257	11.2%
Medium	3	**479**	8.4%	103	4.5%	3	**388**	6.8%	149	6.5%	3 - Medium	**1581**	27.8%	448	19.5%
Medium	4	**59**	1.0%	32	1.4%	4	**217**	3.8%	218	9.5%	4 - High	**1413**	24.8%	710	31.0%
High	5	**0**				5	**39**	0.7%	283	12.3%	5 - Very High	**200**	3.5%	273	11.9%
High	6	**0**				6	**234**	4.1%	133	5.8%	N/A	**0**	0		
High	7	**0**				7	**273**	4.8%	37	1.6%	N/A	**0**	0		
Very High	8	**0**				8	**96**	1.7%	10	0.4%	N/A	**0**	0		
Very High	9	**0**				9	**336**	5.9%	13	0.6%	N/A	**0**	0		
Very High	10	**0**				10	**0**	0			N/A	**0**	0		
		5695		2294			**5695**		2294			**5695**		2294	
SUMMARY	No TEST1 events found in high or very high susceptibility					Only 4.5% of TEST1 events found in high or v ery high susceptibility						3194	**28.3% of TEST1 in high or very high susceptibili**		
	Medium	23.3%			13.3%	6.8% in medium		11.1%					**27.8% in medium**		
	Med to V.H	23.3%		TRAIN:	13.3%	11.3% medium to very		28.3%	TRAIN:	38.2%			**56% of TEST1 m**	TRAIN:	62.4%
		All Low	76.7%					71.7%					43.9%		
		All Medium	23.3%					11.1%					27.8%		
		All High	0					9.6%					24.8%		
		All Very High	0					7.6%					3.5%		

F.2 MaxEnt global vs continental susceptibility results with Equal Interval Classification

Max Ent Model, Equal Interval susceptbility classfication										
EQUAL INTERVAL CLASSIFICATION										
Africa Global Model						Africa Continental Model				
Class	# Events	% Events	# Pixels	% Pixels	Prob Density	# Events	% Events	# Pixels	% Pixels	Prob Density
Very Low	3	18.8%	291,767	89.2%	0.2	4	25.0%	319,542	97.3%	0.3
Low	2	12.5%	26,438	8.1%	1.5	2	12.5%	5,998	1.8%	6.8
Medium	4	25.0%	6,857	2.1%	11.9	3	18.8%	1,564	0.5%	39.4
High	7	**43.8%**	1,947	0.6%	**73.5**	0	0.0%	770	0.2%	0.0
Very High	0	0.0%	92	0.0%	0.0	7	**43.8%**	569	0.2%	**252.5**
	16		327,101			16		328,443		
Medium to Very High =	**11**	**68.8%**				**10**	**62.5%**			
EQUAL INTERVAL CLASSIFICATION										
Asia Global Model						Asia Continental Model				
Very Low	17	3.1%	559,660	79.9%	0.04	13	2.3%	607,695	85.5%	0.03
Low	61	11.0%	80,998	11.6%	0.95	399	**72.0%**	59,733	8.4%	**8.57**
Medium	341	**61.6%**	42,969	6.1%	10.03	130	23.5%	29,827	4.2%	5.59
High	133	24.0%	15,395	2.2%	**10.92**	8	1.4%	11,629	1.6%	0.88
Very High	2	0.4%	1,185	0.2%	2.13	4	0.7%	1,528	0.2%	3.36
	554		700,207			554		710,412		
Medium to Very High =	**476**	**85.9%**				**142**	**25.6%**			
EQUAL INTERVAL CLASSIFICATION										
Oceania Global Model						Oceania Continental Model				
Class	# Events	% Events	# Pixels	% Pixels	Prob Density	# Events	% Events	# Pixels	% Pixels	Prob Density
Very Low	1	20.0%	485	12.1%	1.6	1	20.0%	3118	75.7%	0.3
Low	1	20.0%	1433	**35.9%**	0.6	1	20.0%	500	12.1%	1.6
Medium	1	20.0%	1249	31.3%	0.6	1	20.0%	297	7.2%	2.8
High	2	**40.0%**	754	18.9%	**2.1**	2	40.0%	173	4.2%	**9.5**
Very High	0	0.0%	73	1.8%	0.0	0	**0.0%**	31	0.8%	**0.0**
	5		3,994			5		4,119		
Medium to Very High =	**3**	**60.0%**				**3**	**60.0%**			
EQUAL INTERVAL CLASSIFICATION										
Europe Global Model						Europe Continental Model				
Class	# Events	% Events	# Pixels	% Pixels	Prob Density	# Events	% Events	# Pixels	% Pixels	Prob Density
Very Low	8	0.3%	133150	74.0%	0.0	917	33.4%	1,300,672	84.7%	0.4
Low	466	17.0%	31428	17.5%	1.0	632	23.0%	131,231	8.5%	2.7
Medium	1282	46.7%	12468	6.9%	6.7	559	20.4%	66,283	4.3%	4.7
High	761	**27.7%**	2786	1.5%	17.9	393	14.3%	30,444	2.0%	**7.2**
Very High	226	8.2%	220	0.1%	**67.4**	242	**8.8%**	7,743	0.5%	**17.5**
	2743		180,052			2743		1,536,373		
Medium to Very High =	**2269**	**82.7%**				**1194**	**43.5%**			
EQUAL INTERVAL CLASSIFICATION										
No America Global Model						No America Continental Model				
Class	# Events	% Events	# Pixels	% Pixels	Prob Density	# Events	% Events	# Pixels	% Pixels	Prob Density
Very Low	4	0.2%	393,707	73.7%	0.0	12	0.5%	447,486	83.7%	0.0
Low	14	0.6%	75,182	14.1%	0.0	18	0.8%	55,502	10.4%	0.1
Medium	30	1.3%	39,363	7.4%	0.2	140	5.9%	19,037	3.6%	1.7
High	1510	**63.7%**	21,371	4.0%	15.9	1532	**64.6%**	11,056	2.1%	**31.3**
Very High	812	34.3%	4,871	0.9%	**37.6**	668	28.2%	1,500	0.3%	**100.5**
	2370		534,494			2370		534,581		
Medium to Very High =	**2352**	**99.2%**				**2340**	**98.7%**			
EQUAL INTERVAL CLASSIFICATION										
So America Global Model						So America Continental Model				
Class	# Events	% Events	# Pixels	% Pixels	Prob Density	# Events	% Events	# Pixels	% Pixels	Prob Density
Very Low	1	16.7%	150,648	71.7%	0.2	6	100.0%	211665	99.5%	1.0
Low	1	16.7%	32,557	15.5%	1.1	0	0.0%	573	0.3%	0.0
Medium	3	**50.0%**	15,809	7.5%	**6.6**	0	0.0%	193	0.1%	0.0
High	1	16.7%	9,670	4.6%	3.6	0	0.0%	235	0.1%	**0.0**
Very High	0	0.0%	1,325	0.6%	0.0	0	**0.0%**	69	0.0%	**0.0**
	6		210,009			6		212,735		
Medium to Very High =	**4**	**66.7%**				**0**	**0.0%**			

F.3 MaxEnt global vs continental susceptibility results with Jenks Natural Breaks Classification. Detailed and summarized

Max Ent Model, Jenks Natural Breaks susceptbility classfication										
JENKS CLASSIFICATION										
Africa Global Model						Africa Continental Model				
Class	# Events	% Events	# Pixels	% Pixels	Prob Densi	# Events	% Events	# Pixels	% Pixels	Prob Density
Very Low	0	0.0%	173,334	53.0%	0.0	1	6.3%	274,279	83.5%	0.07
Low	1	6.3%	88,871	27.2%	0.2	0	0.0%	38,382	11.7%	0.00
Medium	1	6.3%	43,785	13.4%	0.5	5	31.3%	10,992	3.3%	9.34
High	3	18.8%	15,332	4.7%	4.0	3	18.8%	3,396	1.0%	18.13
Very High	11	**68.8%**	5,779	1.8%	**38.9**	7	**43.8%**	1,402	0.4%	**102.49**
	16		327,101			16		328,451		
Medium to Very High =										
	15	**93.8%**				**15**	**93.8%**			
JENKS CLASSIFICATION										
Asia Global Model						Asia Continental Model				
Very Low	3	0.5%	416,749	59.5%	0.0	1	0.2%	522,945	73.6%	0.002
Low	17	3.1%	137,891	19.7%	0.2	11	2.0%	83,070	11.7%	0.170
Medium	39	7.0%	71,968	10.3%	0.7	343	**61.9%**	51,373	7.2%	**8.562**
High	180	32.5%	46,681	6.7%	4.9	185	33.4%	34,813	4.9%	6.814
Very High	315	**56.9%**	26,918	3.8%	**14.8**	14	2.5%	18,211	2.6%	0.986
	554		700,207			554		710,412		
Medium to Very High =										
	534	**96.4%**				**542**	**97.8%**			
JENKS CLASSIFICATION										
Oceania Global Model						Oceania Continental Model				
Class	# Events	% Events	# Pixels	% Pixels	Prob Densi	# Events	% Events	# Pixels	% Pixels	Prob Density
Very Low	1	20.0%	737	18.5%	1.1	1	20.0%	2751	66.8%	0.30
Low	1	20.0%	943	23.6%	0.8	0	0.0%	495	12.0%	0.00
Medium	2	40.0%	914	22.9%	**1.7**	1	20.0%	391	9.5%	2.11
High	0	0.0%	884	22.1%	0.0	1	20.0%	293	7.1%	2.81
Very High	1	**20.0%**	516	12.9%	1.5	2	**40.0%**	189	4.6%	**8.72**
	5		3,994			5		4,119		
Medium to Very High =										
	1	**60.0%**				**3**	**80.0%**			
JENKS CLASSIFICATION										
Europe Global Model						Europe Continental Model				
Class	# Events	% Events	# Pixels	% Pixels	Prob Densi	# Events	% Events	# Pixels	% Pixels	Prob Density
Very Low	1	0.0%	110654	53.1%	0.0	486	17.7%	1,129,009	73.5%	0.24
Low	13	0.5%	46712	22.4%	0.0	549	20.0%	182,209	11.9%	1.69
Medium	237	8.6%	27278	13.1%	0.7	563	20.5%	115,145	7.5%	2.74
High	1212	44.2%	16867	8.1%	5.5	612	**22.3%**	72,629	4.7%	4.72
Very High	1280	**46.7%**	6809	3.3%	**14.3**	533	19.4%	37,381	2.4%	**7.99**
	2743		208,320			2743		1,536,373		
Medium to Very High =										
	2492	**99.5%**				**1145**	**62.3%**			
JENKS CLASSIFICATION										
No America Global Model						No America Continental Model				
Class	# Events	% Events	# Pixels	% Pixels	Prob Densi	# Events	% Events	# Pixels	% Pixels	Prob Density
Very Low	1	0.0%	304,459	57.0%	0.0	5	0.2%	137647	148.3%	0.00
Low	6	0.3%	110,044	20.6%	0.0	7	0.3%	37812	53.6%	0.01
Medium	13	0.5%	57,897	10.8%	0.1	13	0.5%	15560	28.2%	0.02
High	35	1.5%	38,953	7.3%	0.2	23	1.0%	8675	19.0%	0.05
Very High	2315	**97.7%**	23,141	4.3%	**22.6**	2322	**98.0%**	5587	11.3%	**8.69**
	2370		534,494			2370		205,281		
Medium to Very High =										
	2363	**99.7%**				**2358**	**99.5%**			
JENKS CLASSIFICATION										
So America Global Model						So America Continental Model				
Class	# Events	% Events	# Pixels	% Pixels	Prob Densi	# Events	% Events	# Pixels	% Pixels	Prob Density
Very Low	0	0.0%	95,905	45.4%	0.0	1	16.7%	177825	83.6%	0.20
Low	1	16.7%	59,705	28.3%	0.6	1	16.7%	28000	13.2%	1.27
Medium	1	16.7%	26,229	12.4%	1.3	4	66.7%	5906	2.8%	**24.01**
High	0	0.0%	16,771	7.9%	0.0	0	**0.0%**	619	0.3%	0.00
Very High	4	**66.7%**	12,613	6.0%	**11.2**	0	0.0%	385	0.2%	0.00
	6		211,223			6		212,735		
Medium to Very High =										
	5	**83.3%**				**4**	**66.7%**			

SUMMARIZED DATA

Max Ent Model, Equal Interval susceptbility classfication				Max Ent Model, Jenks Natural Breaks classfication			
EQUAL INTERVAL CLASSIFICATION				JENKS CLASSIFICATION			
	# Events	% Events	# Events	% Events	# Events	% Events	# Events	% Events
Africa Global Model			Africa Continental Model		Africa Global Model			Africa Continental Model	
Medium to Very High =	11	68.8%	10	62.5%	**Medium to Very High =**	**15**	**93.8%**	**15**	**93.8%**
Asia Global Model			Asia Continental Model		Asia Global Model			Asia Continental Model	
Medium to Very High =	476	85.9%	142	25.6%	**Medium to Very High =**	**534**	**96.4%**	**542**	**97.8%**
Oceania Global Model			Oceania Continental Model		Oceania Global Model			Oceania Continental Model	
Medium to Very High =	3	60.0%	3	60.0%	**Medium to Very High =**	**1**	**60.0%**	**3**	**80.0%**
Europe Global Model			Europe Continental Model		Europe Global Model			Europe Continental Model	
Medium to Very High =	2269	82.7%	1194	43.5%	**Medium to Very High =**	**2492**	**99.5%**	**1145**	**62.3%**
No America Global Model			No America Continental Model		No America Global Model			No America Continental Model	
Medium to Very High =	2352	99.2%	2340	98.7%	**Medium to Very High =**	**2363**	**99.7%**	**2358**	**99.5%**
So America Global Model			So America Continental Model		So America Global Model			So America Continental Model	
Medium to Very High =	4	66.7%	0	0.0%	**Medium to Very High =**	**4**	**83.3%**	**4**	**66.7%**

Appendix G Global MaxEnt Susceptibility Model Backward Analyses

	Köppen-Geiger Climate					
KG Class	# Pixels	% Total	Grouped			
Water	4	0.0%				
Af	4,499	6.3%				
Am	4,356	6.1%				
Aw	1,079	1.5%	**13.9%**	**A - Tropical/megathermal**		
BWh	4	0.0%				
BWk	694	1.0%				
BSh	85	0.1%				
BSk	4,754	6.6%	**7.7%**	**B - Dry (desert and semi-arid)**		
Csa	2,169	3.0%				
Csb	4,297	6.0%				
Csc	44	0.1%				
Cwa	**9,378**	**13.1%**				
Cwb	5,484	7.7%				
Cwc	17	0.0%				
Cfa	**10,332**	**14.4%**				
Cfb	5,571	7.8%				
Cfc	381	0.5%	**52.6%**	**C - Continental/mesothermal**		
Dsa	74	0.1%				
Dsb	2,535	3.5%				
Dsc	698	1.0%				
Dwa	70	0.1%				
Dwb	669	0.9%				
Dfa	**6,152**	**8.6%**				
Dfb	**6,020**	**8.4%**				
Dfc	1,715	2.4%	**25.1%**	**D - Continental/microthermal**		
ET	473	0.7%				
EF	29	0.04%	**0.7%**	**E - Polar**		

Precipitation SUMMARY	
mm	% area (pixels)
<100	43.70%
>=100 and <200	**46.80%**
>=200 and <300	8.50%
>=300 and <400	0.80%
>=400	0.20%

Soil Drainage		
Class	**#Pixels**	**% Total**
Very Poor	32	0.05%
Poor	4,334	6.3%
Imperfectly	9,735	14.1%
Moderately Well	**53,672**	**77.7%**
Well	86	0.1%
Somewhat Excessive	1,187	1.7%

Soil Thickness		
meters	# Pixels	Grouped %
<=3	**1,604**	**87.0%**
>3 and <=50	1,505	13.0%

Clay Pct	# Pixels	% Total
<= 15%	5026	7.3%
16 - 25%	61614	89.5%
26-57%	2238	3.2%

Soil Type		
Grouped	**# Pixels**	**%**
Acrisols	18,707	26.1%
Cambisols	13,231	18.5%
Luvisols	9,830	13.7%
SubTotal	**41,768**	58.3%
22 other types	29,845	41.7%

Murphy-Bridges Ecophysiographic Region	
% Area	EcoPhysReg
1.2%	Plains in Alpine System
0.5%	Plains on Sedimentary Lithology
23.7%	**Mountains in Alpine System**
0.8%	Widely Spaced Mountains in Alpine System
7.2%	Hills or Tablelandsin Alpine System
37.2%	**Humid Mountains in Alpine System**
1.2%	Humid Hills or Tablelandsin Alpine System
0.4%	Dry Depressions or Basins in Alpine System
1.2%	Humid Plains on Sedimentary Lithology
6.7%	Humid Hills or Tablelands on Sedimentary Lithology
1.2%	
0.9%	Humid Plains in Alpine System
9.6%	Dry Mountains in Alpine System
0.8%	Dry High Tablelandsin Alpine System
3.8%	Humid Mountains on Sedimentary Lithology
3.5%	Dry Widely Spaced Mountains in Alpine System

Lithology		
% Total		Grouped
0.0%	**evaporites**	0.0%
0.01%	**ice & glaciers**	0.01%
12.3%	**metamorphic**	12.3%
0.1%	**no data**	0.1%
8.4%	acidic	
1.4%	basic	
1.4%	intermediate	
	Plutonic	11.2%
1.2%	**pyroclastic**	1.2%
9.6%	carbonate	
15.7%	mixed	
26.5%	siliciclastic	
9.4%	unconsolidated	
	Sedimentary	**61.2%**
3.0%	acidic	
6.7%	basic	
4.0%	intermediate	
	volcanic	13.6%
0.4%	water bodies	0.4%

Elevation (m)	% area (pixels)
>0 and <500	30.7%
>=500 and <1000	21.3%
>=1000 and < 2500	**37.4%**
>=2500 and <=5791	10.6%

Slope Deg	# Pixels	% Total
<20 deg	2,676,759	77.2%
20-30 deg	535,795	15.4%
31-40 deg	196,111	5.7%
> 40 deg	59,507	1.7%

LandCover/LandUse	
% Total	Grouped
68.5%	Forest and Woodland
13.0%	Grassland/Shrubland
16.8%	Cropland
1.2%	Urban
0.0%	Snow/Ice
0.3%	Wetlands

Fault Density		
km/sq km	% area (pixels)	#Pixels
.000001-.0302862	**71.2%**	**268,263**
.0302863-.0605723	21.5%	81,000
.0605724-.0908585	5.7%	21,620
.0908586-.121145	1.2%	4,672
.121146-.151431	0.3%	977