DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Uncertainties in dynamic landslide prediction at regional scale“

verfasst von / submitted by

Ekrem Canli, BSc MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, 2018 / Vienna, 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 796 605 452
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Geographie / Geography
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Geogr. Dr. Thomas Glade

Acknowledgements

The last couple of years as a PhD candidate were one hell of a ride. There were times I was endlessly frustrated, but there were also times where I was blown away be all the thrilling experiences I could make. I would like to take the opportunity to express my gratitude to a few people that contributed not only to my research, but my life in general.

Firstly, I would like to say thank you to Thomas Glade for supervising and supporting my PhD thesis. But I would also like to say thank you for giving me the opportunity to fully experience the life as a researcher which is really a once-in-a-lifetime experience I would not have missed for anything.

To all current and past members of the ENGAGE working group I was part in: thank you for all the fun and Pizzas we had! It was good to have you around whenever I needed to vent about some work-related issues.

For programming related issues, I want to show my gratitude to the amazing online communities of stackoverflow, the Cross Validated stackexchange, and the GIS stackexchange. Without those discussions and their collective intelligence, many things in my research probably would not have worked out at all.

It is a matter of great importance to me to say thank you to my family. Without your infinite support during my days as an undergraduate student, I am certain I would not be where I am today. You kept my back free wherever possible which was the greatest aid I could have ever asked for. Thank you!

Verena, thank you for everything. Without your never-ending love and support, I might have thrown in the towel at some point. You kept pushing me in times I loathed even getting out of bed, you helped me getting away from the daily hassles, you motivated me and kept my back free so I could focus on my own work. Thank you so much!

Table of Contents

Table of Contents V
List of Figures VII
List of Tables VIII
Abbreviations IX
German and English Summary XI
1. Background 1
1.1. Thesis outline & scope 1
1.2. Research gap and hypotheses 4
2. The science and philosophy of modeling in geomorphology 9
2.1. Types of models 14
2.1.1. Conceptual models 15
2.1.2. Heuristic methods 16
2.1.3. Empirical-statistical models 17
2.1.4. Deterministic models 19
2.2. Issues in modeling – parametrization, validity and uncertainty 21
3. Physically based landslide modeling and early warning systems — current approaches and challenges 29
3.1. Spatial variability and uncertainties in regional scale landslide modeling 29
3.2. Consideration of spatial and temporal dynamics: rainfall as a crucial component 32
3.3. Current issues in probabilistic landslide modeling and early warning 39
4. Uncertainties in regional scale landslide prediction — a methodological approach 43
4.1. Uncertainties in rainfall information 43
4.2. Uncertainties in geotechnical and hydraulic parametrization at regional scale 45
4.3. Uncertainties in probabilistic modeling 47
5. Discussion of results and hypotheses 51
6. Conclusion and perspectives 59
7. References 69
A. Publications and manuscripts 83
A.1. Canli et al. (2016): Generating web scraped high-quality weather databases for near-real-time derivation of spatial landslide susceptibility 85
A.2. Canli et al. (2017a): Spatially distributed rainfall information and its potential for regional landslide early warning systems 93

A.3. Canli et al. (in prep.): Evaluating landslide dynamics and internal structure — a case study from the Salcher landslide observatory, Austria 121

A.4. Canli et al. (2017b): Probabilistic landslide ensemble prediction systems: Lessons to be learned from hydrology 161

B. Eidesstattliche Erklärung 193

List of Figures

Figure 2.1: A simplified system as a set of objectives together with relationships between the objects and between their attributes (Inkpen 2005) 10

Figure 2.2: Conceptual model of shallow ground-water conditions in hillside soils. The unsaturated zone above the water table has depth d_u. The capillary fringe is between the unsaturated zone and the water table at depth d. The lower boundary, which is treated as impervious in this model, is at depth Z_{max} (Baum et al. 2008) 16

Figure 2.3: Example of a decision tree for classifying landslide susceptibility in which leaves with high landslide susceptibility (more than 5) are emphasized. Variables are listed in the original source in Saito et al. 2009 17

Figure 2.4: Comparison between uncertainty and variation. The reduction in confidence due to unexplained variation, such as uncertainty, decreases through progress in science. Progress in science, however, will not decrease the total variation (Lehmann and Rillig 2014) 26

Figure 3.1: Proposed underground model of the Salcher landslide (Austria) based on all obtained information (inclinometers, drill cores, penetration resistance). Spatial variation in both, horizontal and vertical, direction seem to lack a pronounced spatial organization, which is a challenge for slope scale modeling, yet an even bigger challenge for model parametrization at regional scale (Canli et al. in prep.) 31

Figure 3.2: Results from the automated spatial interpolation: a) Hourly rain gauge data; b) Ordinary Kriging (OK) without filtering; c) OK with filtering; d) IDW interpolation; e) Thiessen polygons. Lines in the b) and c) estimates indicate areas with equal amount 35

Figure 4.1: Flow chart of the proposed methodology showing the automated workflow from obtaining web based rainfall data, through multiple quality assurance filters,

to the application of different interpolation techniques for producing hourly real-time rainfall raster maps (Canli et al. 2017a)..44

Figure 4.2: Probabilistically derived model parameters (Soil depth, Cohesion, Friction angle) based on random sampling from a normally distributed state function (Canli et al. 2017b) 47

Figure 5.1: Probability of failure shown as a proportion of the individual ensemble members that predict an event to occur (FoS < 1.0). Building information was added as an additional layer to express an individual building's exposure to landslides. Building and roads are used from the freely accessible OpenStreeMap database (Canli et al. 2017b; Forest/River/Road data © OpenStreetMap contributors) 56

List of Tables

Table 2.1: Traditionally recommended quantitative methods for landslide susceptibility analysis at different scales (Corominas et al. 2014) 20

Abbreviations

EPS	Ensemble Prediction System
FoS	Factor of Safety
NWP	Numerical Weather Prediction
PoF	Probability of Failure
UNEP	United Nations Environment Programme
WMO	World Meteorological Organization

German and English Summary

Zusammenfassung

Frühwarnsysteme sind im Stande zeitgerecht und effektiv Informationen zu übermitteln, um das Risiko vor einer nahenden Gefahrensituation zu vermeiden oder zu reduzieren. Im letzten Jahrzehnt wurden in der Hydrologie so genannte Ensembleprognosen in die Hochwasservorhersagesysteme implementiert und folgte damit den erzielten Erfolgen in der Wettervorhersage. Dieser probabilistische Ansatz berücksichtigt die inhärente räumliche Variabilität geotechnischer und hydraulischer Parameter sowie deren Unsicherheiten und bringt sie explizit in die Modellergebnisse ein.

Die hier vorgestellte Arbeit befasst sich explizit mit zwei Unsicherheitsaspekten auf regionaler Maßstabsebene in Frühwarnsystemen für Hangrutschungen. Diese betreffen einerseits die Berücksichtigung von Niederschlag als dynamische Komponente und andererseits den Umgang mit räumlicher Variabilität und Unsicherheiten in Parametern für die Modellierung. Ein Ansatz, der überwiegend Anwendung findet, um Niederschlag in Hangrutschungsfrühwarnsystemen zu implementieren, beinhaltet die Verwendung von flächenhaft einheitlichem Niederschlag für ein spezifisches Gebiet basierend auf repräsentativen Niederschlagsmessern. Hier wird eine Alternative vorgestellt, die vorsieht, basierend auf verschiedenen Interpolationsverfahren (deterministisch und geostatistisch), eine räumlich differenzierte Niederschlagsverteilung in Echtzeit zu bestimmen. In einer voll automatisierten Prozesskette werden dazu webbasierte Niederschlagsdaten in mehrere Qualitätschecks geprüft, um eine qualitativ hochwertige Datenbasis zu erlangen.

Für die anschließende Hangrutschungsmodellierung ist das deterministische, physikalisch basierte TRIGRS Modell für eine probabilistische Anwendung modifiziert worden. Um die innewohnenden Unsicherheiten sowie die räumliche Variabilität von Parametern zu adressieren, wurde anstatt einer einzelnen vermeintlich optimalen Parameterkonstellation

basierend auf tatsächlichen Feldmessungen, ein sehr breiter Parameterbereich aus Literaturquellen berücksichtigt. Aus diesem Parameterbereich wurden in einer Zufallsstichprobe mehrere Parameterkonstellationen gezogen, die dann für die jeweiligen Modellläufe herangezogen wurden. Basierend auf einer Vielzahl von gleichermaßen annehmbaren Parameterkonstellationen wurden ebenso viele Modellläufe durchgeführt, die für jede Stunde in einer einzelnen räumlich differenzierten Karte der Hangversagenswahrscheinlichkeit resultierten. Dadurch wird der relative Beitrag jedes einzelnen Modelllaufs, der auf unterschiedlichen, aber gleichermaßen annehmbaren Parametern besteht, berücksichtigt. Dabei deckt die gesamte Spannweite des räumlichen Musters der Hangversagenswahrscheinlichkeit einen Großteil der vorhandenen räumlichen Variabilität und Unsicherheiten ab.

Die Ergebnisse legen nahe, dass für Hangrutschungsmodellierungen auf regionaler Maßstabsebene die Modellparametrisierung basierend auf Literaturquellen ausreichend ist, da a) verschiedene Parameterkonstellationen ähnlich gute Modellergebnisse liefern und damit die Bedeutung einer Modelleichung sowie von teuren und zeitaufwändigen Feldmessungen reduziert wird, und b) die Modellsensitivität der Hangneigung so dominant ist, dass räumliche Unterschiede in der Hangversagenswahrscheinlichkeit mehr durch die räumliche Verteilung des Niederschlags oder der Bodenmächtigkeit beeinflusst werden, als durch geotechnische und hydraulische Parameter. Der hier vorgestellte voll automatisierte Ensembleansatz birgt großes Potential für die zukünftige Ausrichtung von Hangrutschungsfrühwarnsystemen, jedoch sind die Anforderungen an herkömmliche Computerhardware noch zu groß, um die Berechnung stündlicher Hangversagenswahrscheinlichkeiten auf größerer Maßstabsebene in Echtzeit zu bewerkstelligen.

Summary

Early warning aims at providing individuals exposed to a hazard timely and effective information to take action in order to avoid or reduce their risk and prepare for effective response. In the last decade, hydrological modelers have started integrating ensemble prediction systems into their forecasting systems, following on the success of the use of ensembles for weather forecasting. The probabilistic approach acknowledges the presence of unavoidable parameter variability and uncertainty at larger scales and explicitly introduces them into the model results.

The proposed work explicitly addresses two main sources of uncertainties in regional scale landslide early warning. Firstly, how rainfall as a dynamic component is treated and secondly, how spatial variability and uncertainties in geotechnical and hydraulic parameters are considered for regional scale model parametrization. A common approach to introduce rainfall information into landslide early warning system consists of using uniform areal rainfall from representative rain gauges over a specific area. Here, a fully automated process chain is presented that uses web based real-time rain gauge data that is treated with multiple quality checks. This data is then applied to multiple automated interpolation techniques (deterministic and geostatistical) in order to obtain spatially distributed rainfall information. For the landslide prediction, the deterministic, physically based model TRIGRS is modified for a fully automated probabilistic application. In an honest attempt to address parameter variability and uncertainties, broad parameter ranges from literature that are appropriate for the study area are used instead of a presumed best-fit set of parameter values based on actual in situ field data. Out of this parameter range, multiple parameter sets are randomly sampled for iterative model runs. From all parameter sets, which resulted in multiple equally acceptable model realizations, a spatially distributed probability of failure map is derived for each hour. This way, the relative performance of each parameter set is taken into account and depicts the entire model spread with its inherent uncertainties.

Results suggest that for regional scale study areas purely literature based parametrization might be sufficient because a) different parameter sets provide almost equally good results and thus reduces the importance of costly and time consuming field sampling as well as model calibration, and b) slope angle has such a high model sensitivity that in all model runs the predicted areas with the highest slope failure probability are more or less at the same location and differ primarily due to spatially varying soil depth and rainfall. Although the proposed automated landslide ensemble prediction system holds a great potential for the future direction of landslide early warning, computational restraints currently hold back the real-time application for hourly model predictions at regional scale.

1. Background

1.1. Thesis outline & scope

Although substantial effort is put into landslide mitigation methods based on improvements in our understanding of instability mechanisms, landslides are still causing a considerable death toll and major economic losses all over the world (Corominas et al. 2014). Former Secretary-General of the United Nations Ban-Ki Moon called climate change during the 2014 Climate Summit the *defining issue of our time* (Frigg et al. 2015). However, not just climate change, but global change in general will be a critical component in landslide research as it is assumed that its consequences will increase the number of landslides in the future (Crozier 2010, Gariano et al. 2017, Papathoma-Köhle and Glade 2013). Advancements in the past within the field of geotechnical engineering have led to an increasing in situ damage control in many parts of the world, however, landslides triggered by heavy rainstorms still cause substantial losses where protective structures are scarce or where they have not been appropriately designed (Canli et al. 2017a). In this context, landslide risk can be defined as "the expected number of lives lost, persons injured, damage to property and disruption of economic activity due to a particular damaging phenomenon for a given area and reference period" (Varnes 1984, p. 10). In order to manage landslide risk, a multi component analysis is required. The total risk (R) can be expressed as the product of hazard (H), vulnerability (V) and the elements at risk (A) (van Westen et al. 2006):

$$R = \sum \left(H \sum (VA) \right) \quad (1)$$

where:

H Hazard is expressed as the probability of occurrence within a reference period. To distinguish hazard from just susceptibility, is has to contain not just the *spatial probability* of occurrence (based on static environmental factors such as soil depth, friction angle, lithology, etc.), but also a dynamic *temporal probability* (e.g. varying rainfall input, historical records, etc.)

V Vulnerability for a specific type of hazard can be expressed in many ways, such as physical vulnerability (e.g. building shape, construction material, etc.) or social vulnerability (e.g. age-composition of residents, daytime, etc.) and manifests itself in a range from 0 (no loss at all) to 1 (total loss)

A Elements at risk expressed as their total amount or costs (e.g. number of buildings or people in a specific area, cost of buildings)

Consequently, risk implies also the consequences of an event, not just its probability of occurrence. This is important to distinguish at this point, because the scope of this thesis almost exclusively encompasses the hazard component which is, according to van Westen et al. (2006), by far the most complex to determine.

While at first sight landslides are generally considered to be phenomena restricted to the local scale, they can indeed be regarded as a regional phenomenon at specific times (Jaedicke et al. 2014). Contrary, the spatial occurrence of floods, for example, is topographically much more foreseeable and controllable which is far more challenging to asses in distributed landslide prediction due to a landslide's localized nature (Alfieri et al. 2012a). As a consequence, structural protective measures are only feasible were critical infrastructure or persons are directly affected by a potential landslide hazard. However, for covering large areas that are potentially prone to landsliding and where a substantial landslide risk is prevalent, spatial landslide early warning systems (EWS) are indispensable (Glade and Nadim 2014, Thiebes and Glade 2016).

Working in this EWS context requires the observation and timely processing of rainfall events on small spatio-temporal scales, as this is crucial for the successful operation of EWS (Segoni et al. 2009, Thiebes et al. 2013). By far the most common way to implement rainfall into EWS involves the employment of empirical-statistical rainfall thresholds. A certain rainfall threshold is established for a specific area by determining the rainfall amount that triggered landslides in the past (Gariano et al. 2015). With this relationship it is then possible to provide a real-time comparison of current rainfall and the established rainfall threshold to form the basis of landslide warnings (Wieczorek and Guzzetti 1999). However, such empirically derived thresholds rely purely on the relationship between rainfall and landslide occurrence, which reflects quite a strong simplification of the underlying physical processes (Reichenbach et al. 1998, Bogaard and Greco 2018). Most certainly, there is more than just rainfall as the only causative factor involved (Huang et al. 2015). As opposed to those empirical-statistical threshold based approaches, process based approaches are in place that can be used in an early warning context. Those (mostly deterministic) models do not simply establish statistical relationships between the dependent variable and its predictors, but use (physically based) equations to actually represent process interactions. Such process based models are more resembling a white-box approach by describing the underlying physical processes that lead up to the phenomenon being modelled (Corominas et al. 2014). Although computationally very demanding and conceptually challenging to apply at larger scales, physically based models contain "a higher predictive capability and are the most suitable for quantitatively assessing the influence of individual parameters that contribute to shallow landslide initiation" (Corominas et al. 2014, p. 225). Within the scope of this dissertation, only physically based approaches and its associated uncertainties are worked on in a potential early warning context. Those uncertainties primarily address the spatial variability and uncertainties in geotechnical and hydraulic parameters at larger scales as well as the uncertainties introduced by how rainfall as the dynamic component is considered.

With regard to general definitions, this dissertation uses prediction systems and early warning systems synonymously for terminological consistency within the landslide community. However, it is acknowledged that an operative early warning system should additionally consist of a proper dissemination and response strategy as it is suggested by the UNEP (2012). This dissertation also follows the classification scheme of Stähli et al. (2015) for EWS. In this classification scheme, EWS are distinguished between *Alarm*, *Warning* and *Forecasting* systems. Here, the general term *early warning system* is actually referring to warning systems, again for terminological consistency. Warning systems after Stähli et al. (2015) detect significant changes in time-dependent factors before an even occurs. While the initial alert is based on predefined thresholds, the actual alert is only released after expert evaluation (as opposed to alarm systems that immediately release an alarm) and only then (as opposed to forecasting systems that report current modeling results/sensor data in regular intervals). The infinite-slope based landslide modeling approach in this dissertation aims at detecting shallow translational slope failures as this is generally what infinite-slope based models are capable to reproduce and what this dissertation is focusing on. Bell et al. (2014) analyzed 142 landslide entries of an inventory in Lower Austria, which serves as the study area in this dissertation, and estimated a median landslide depth of 1.7 m (mean: 2.2 m). Following the updated *Varnes classification* for landslides of Hungr et al. (2014), shallow translational landslides in this dissertation refer to clay and/or silt planar slides. The modelling scale, the proposed study is embedded in, explicitly targets the regional scale. Although numerical quantification schemes exist to sharply distinguish between different scales (e.g. Corominas et al. 2014 define regional scale in the range between 1:25,000 and 1:250,000), a more qualitative distinction is sufficient for here in a sense that national scale > regional scale > local scale > site-specific scale.

1.2. Research gap and hypotheses

Summarizing the scope of this dissertation, this work primarily deals with uncertainties of regional scale landslide modeling in an early warning context. Uncertainties in this work are linked to a dynamic component, which refers to the real-time assessment of spatial rainfall information on the one hand, and on how deterministic modeling approaches could be improved in a way that spatial parameter variability and uncertainties are addressed and explicitly introduced into the model results on the other hand.

The most common approach in contemporary spatial landslide early warning systems utilizes rainfall information from direct rain gauge measurements. Those rain gauges indicate either a representative amount of rainfall for single landslide locations nearby (e.g. Capparelli and Tiranti 2010) or selected rain gauges indicate representative uniform areal rainfall for an entire region (e.g. Segoni et al. 2015, Rosi et al. 2015). This is quite unfortunate as it is of utmost importance to know exactly the total precipitation accumulated or the rate of precipitation in a given period in order to link on-site rainfall as the triggering event to landslide occurrence (Guzzetti et al. 2007). Thus, using just punctual rain gauge measurements leads inevitably to a situation where the precise amount of a landslide triggering rainfall at a certain location remains mostly unknown. In reality, however, rain gauges with the closest proximity to a landslide location or rain gauges with the supposedly best representation of areal rainfall are selected for determining a landslide-triggering rainfall event (Canli et al. 2017a). Although large efforts are put into establishing appropriate rainfall threshold in early warning applications, an in-depth consideration of the accurate spatial distribution of rainfall is often neglected (Thiebes and Glade 2016). Therefore, parts of this dissertation aim at providing an improved basis for real-time spatio-temporal rainfall data and its potential implementation in a regional landslide EWS. Instead of assuming uniform rainfall over a certain area, different automated interpolation methods are presented. This allows for an approximation of spatially distributed, hourly rainfall predictions in real-time based on rain gauge data (Canli et al. 2017a),

The effect of rainfall on landslide detachment alone, however, is difficult to assess quantitatively without process based modeling approaches, primarily due to the inherent spatial variability in material properties and its associated uncertainties at larger scales (Chae et al. 2017). Although empirical-statistical rainfall thresholds are by far the most common approach in spatial landslide early warning, such thresholds pose a quite considerable simplification between rainfall occurrence and the physical mechanisms leading to landslides by neglecting local environmental conditions and the role of hydraulic processes occurring along slopes (Reichenbach et al. 1998, Bogaard and Greco 2018). However, process based approaches in spatial landslide early warning are almost non-existent. This can be attributed to both, the massive computational power that is required to operate physically based modeling approaches in a timely manner and the difficulty to predict the location of rather small-scale phenomena such as landslides with purely deterministic models with currently available data (Canli et al. 2017b). On the contrary, hydrological sciences have successfully transitioned from threshold based approaches to process based approaches by developing probabilistic models that do not eliminate uncertainty, but explicitly introduce them into the model results to acknowledge the inevitable spatial variety and uncertainties when operating at larger scales (Cloke and Pappenberger 2009). Now that high resolution convective-scale numerical weather predictions (NWP) are available that are particularly suitable for predicting small-scale phenomena such as flash floods and landslides, the next logical step in landslide prediction should be the adaptation of such ensemble prediction systems (EPS). Therefore, this dissertation proposes the application of a probabilistic regional landslide EPS with the aim of investigating the potential of such probabilistic approaches over purely deterministic ones for early warning applications.

Consequently, the identified research gaps are formalized into the following hypotheses and associated research questions respectively:

Hypothesis I: Automated interpolation poses an improvement over selective rain gauge utilization for providing landslide early warning information.

- What data sources are appropriate for being used in real-time applications?
- How to approach automated data quality assurance?
- How do different automated interpolation techniques perform?

Hypothesis II: In situ measurements of geotechnical or hydraulic parameters can be substituted by literature based values for regional scale landslide model parametrization.

- How to approach spatial parameter variability and uncertainties?
- How does a parameter range input affect the spatial prediction pattern?
- What parameters manifest the highest sensitivity?

Hypothesis III: A probabilistic landslide ensemble prediction system is capable of providing timely indication of high resolution landslide exposure at regional scale.

- Does the probability of failure display a realistic image of the most current landslide hazard?
- Does infrastructure data improve the interpretation of probabilistic hazard maps?
- Is it possible to operate landslide ensemble prediction in real-time for the application in regional landslide early warning systems?

2. The science and philosophy of modeling in geomorphology

The philosophical ideas of reductionism have developed and shaped the discipline of physical geography for a long time. Reductionist approaches in different dimensions (ontological, epistemological, and methodological) looked at independent parts of a phenomenon in isolation from each other which culminated in an understanding of the phenomenon as a whole when all parts are combined. This was the predominant paradigm in physical geography to look at natural phenomena until the 1960s when the more holistic and synthetic approach of systems analysis came into play. Systems analysis has developed as the integrative explanatory framework of physical geography and it changed the way of thinking about the physical environment. The rise of systems analysis owes a great deal to the attempt to develop an integrated and all-encompassing framework for all sciences in the twentieth century. According to Inkpen (2005) the existence of such a framework implies:

1.) that all reality is capable of being understood; there are no areas of topics outside of its analytical scope.
2.) All reality can be understood in a common framework using the same sets of terms. This means that understanding in supposedly different subject areas does not require specialist terms of specialist knowledge, but rather translation of these terms to the common terminology of systems analysis.
3.) As there is a common framework, all reality can be expected to behave as predicted by this framework. All reality becomes potentially predictable and, by implication, potentially controllable.

Systems thinking can be represented by a couple of relatively simple ideas (Fig. 2.1). The key components of such a system are the variables or elements, the relationships between the variables or elements, and the bounding of these variables and relationships from the rest of the world (Inkpen 2005). The system itself can be considered as "a set of objectives together with relationships between the objects and between their attributes" (Hall and Fagan 1956, p. 18). Defining variables and relationships imply an ability to define and divide the world into distinct entities and relations. The same applies to the definition and bounding or closure of the system itself which requires a specific view of reality as divisible and understandable by this division. The observer considers the reality of the physical environment from an outside view by defining his own distinct system made of real entities and relations which becomes an entity in itself with its own properties and relations to the rest of the physical environment. This observer's own entity may or may not be the sum of its elements and relations, in any case, however, the observer serves as a passive and objective interpreter of the system outside of the boundaries he has imposed (Inkpen 2005).

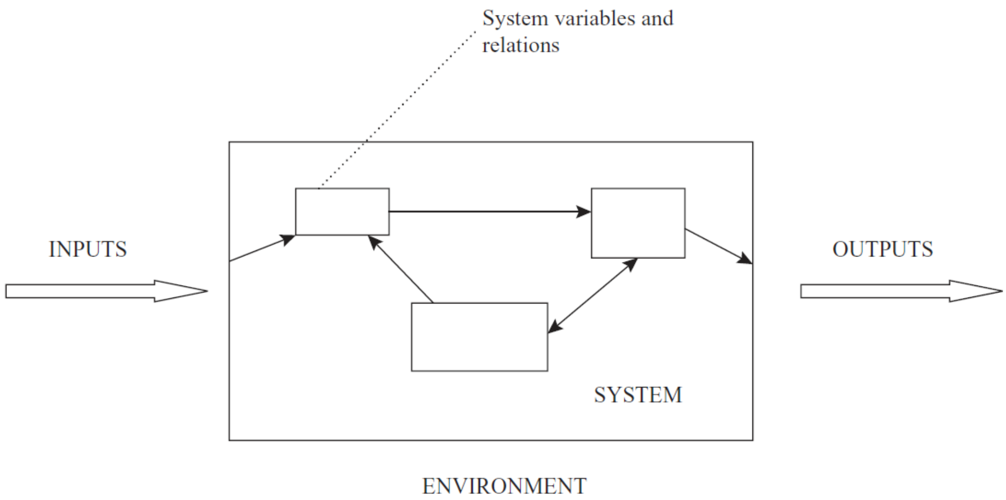


Figure 2.1: A simplified system as a set of objectives together with relationships between the objects and between their attributes (Inkpen 2005)

This omnipresent paradigm in contemporary geomorphology made questions legitimate with respect to processes of change and their rates. The late and seemingly immature development of the subject relative to *hard* sciences condemned geomorphology to a mere descriptive scientific discipline. The theoretical foundation of systems analysis put the identification and understanding of processes in a framework that aimed at systematically answering questions with regard to present and past process rates. The development of a process orientation in physical geography led to changes in the disciplines' own definition. With this new focus, physical geography reached into other disciplines for the necessary theoretical, field and mathematical techniques. The quest for an integrative explanatory framework for physical geography became somewhat superseded by a search for disciplinary homes with their own, existing framework. This problem of discipline definition is a matter that physical geography is still struggling with (Inkpen 2005).

Important to the development of a process basis for physical geography was the development of techniques for quantifying landforms and the landscape. The trend to increasingly complex representations of the physical environment within each subdiscipline further added to the separation of the subject. However, this quantification of the landscape, or rather components of the landscape, ultimately triggered the capability of assessing the contribution of specific processes to the operation of the landscape system. Despite great advances made in technology and data collection, physical geography retained its roots from the past (Inkpen 2005). Physical geography, or rather geomorphology, retained a focus on landscape development that promoted relevant studies of landforms and land forming processes within such features as landslides and, more generally, in environmental management. A plethora of information, the push for relevancy and the focus on process studies – the basic principles of physical geography remained the same: the search for universality, the emphasis on the empirical and a concern with change in the form of equilibrium and process-response. The matter of scale, however, remains as a sticking point for the integration of processes found at different scales and acted as a brake on a purely reductionist view of the scientific endeavor in physical geography (Inkpen 2005).

Modeling, in general, gives us the opportunity to test the reliability of our comprehension of the nature and its processes and phenomena. Its aim is to generalize, put in order and extract all information of interest that are available based on the most current theoretical and experimental knowledge. Modeling also requires finding a middle ground between how we understand and how we represent the complexity of nature. As a result of practical demands, older distributed models, for example, are still being used that do not entirely reflect our current understanding of processes (Semenova and Beven 2015). Models are capable of providing a means of understanding and predicting the operation of systems that are not approachable by experimental methods such as variable control and manipulation. Also, for many practical reasons or due to scale issues (both spatial and temporal), such analyses would be not feasible, hence the requirement for model formulation (Demeritt and Wainwright 2005). Thorn (1988) distinguishes a model from a system in the way how reality is considered. While a model is a fully specified, yet abstract and incomplete, version of reality, a system is viewed as an abstraction that is assumed to exist in reality. A model is an abstraction and a simplification of reality and it is recognized that it does not, nor is intended to, mirror reality. The distinction is generally made to clarify the purpose-led construction of models as opposed to the supposed universal nature of systems. A model is usually created to serve a purpose; it does not, however, need to fully specify reality, nor to be agreed by all. A system may be unknowable in full, but agreement can be achieved that such a set of entities and relationships exist. Consequently, the system can be considered to hold a more universal status whilst the model on the other hand does not (Inkpen 2005).

Although models can be constructed without an explicit underlying philosophy, Beven (2001) considers modeling in the environmental sciences as a form of pragmatic realism. A modeler has in general a clear perceptual model of reality in mind that reflects his qualitative understanding from experience, training and monitoring, including all current constraints that impede model formulation. To add to that, the modeler has to cope with the discrepancy of his own perceptual model of reality and the necessity of building a nomological system with all associated constraints to be able to produce predictions from the model. This also means, however, that there is potential for model rejection, model refinement and for model improvement as the constraints on modeling change (Inkpen and Wilson 2013). Since computational processing became widespread available, study of nature has been strongly driven by this new means of technology. Guzzetti (2005), however, expresses his disappointment in that regard that investigators focus too heavily on applying different tools and methods rather than focusing on the target itself. For the landslide community, this stands more or less valid until this very day. While the degree of sophistication in statistical landslide susceptibility modeling is quite high by now and conceptual errors and biases are actively worked on (e.g. Steger et al. 2016a, 2016b), this is less the case for dynamic threshold or process based approaches. For empirical-statistical threshold based approaches, Bogaard and Greco (2018) and for deterministic modeling approaches Canli et al. (2017b) raised important concerns with respect to deficiencies and challenges in current model applications. Semenova and Beven (2015) raise the question whether this dissatisfaction with current modeling concepts is owed to the current practice of model calibration that allows for a demonstration of success in matching the available data. Klemeš (1986) mentioned already three decades ago in a hydrological context that the current practice of model calibration should not be the be-all-end-all to rigorous model testing and that there is absolutely no guarantee of successfully predicting the future state of a system this way.

In the late 1980s and early 1990s, the paradigm of complex systems research was introduced to geomorphology as an alternative approach to linear explanation of cause and effect. This development was initiated due to field observations where seemingly simple relationships could not be linked to cause and consequence (Temme et al. 2015). This led to the introduction of many new concepts in geomorphology such as *complex response*, *lagged behavior* and *thresholds* (e.g. Knox 1972, Schumm 1973, Thomas and Allison 1993). As most of those concepts do not originate in geomorphology, critical discussion on different concepts and assumptions with respect to complex systems have only accelerated in the last couple of years (e.g. Phillips 2015, Temme et al. 2015, von Elverfeldt et al. 2016).

While there are many definitions on complexity, there are two fundamental properties inherent in complexity theories: a) the system consists of multiple interactive components, and b) these interactions give rise to emergent forms and properties which are not reducible to the sum of the individual components of an observed system (Keiler 2011). This means that cause and effect may not be necessarily related directly as we might think and that a response does not behave in a way as expected. Assumed complex, yet random (stochastic) behavior may result from a simple underlying interaction that is just not (yet) known. Murray and Fonstad (2007) describe unknown scaling interactions as a possible cause for such nonlinear interactions: “Nonlinear interactions often involve multiple feedbacks that lead to surprising and rich, perpetually changing behaviors – behaviors that create themselves, in the sense that ‘events’ do not correspond to changes in the forcing. And simple, local nonlinear interactions provide the basis for the self-organization of global patterns that do not correspond to any forcing template. The related emergent-phenomena perspective points out that analyzing the building blocks of a system – the small-scale processes within a landscape – may not be sufficient to understand the way the larger-scale system works. […] Thus, when nonlinear feedbacks lead to self-organization of large-scale patterns and behaviors, causality extends in both directions through the scales, and the most ‘fundamental’ scale on which to base an analysis may not be the smallest. The extent to which these scale-related phenomena imply that a hierarchy of scales for models and understanding is required in geomorphology is still under vigorous debate” (Murray and Fonstad 2007, p. 173f).

Complexity research in geomorphology is a rather small subfield that aims at introducing tools from non-linear dynamics to explain dynamics and structures of earth surface systems, however, it is not mainstream in any way (Temme et al. 2015). But since commonly applied deterministic modeling approaches have reached their limits of explicability, alternative approaches might see a significant rise in the near future, especially since computational power is getting widely available for considering nonlinear interactions over larger scales (Canli et al. 2017b).

2.1. Types of models

Modeling in physical geography received a huge boost during the so-called *quantitative revolution* of the 1960s and 1970s which resulted in prominent publications such as Chorley and Haggett’s *Models in Geography* (1967) which had a substantial impact on research undertaken by subsequent generations. Since then, different types of models have been established that aim at representing reality (Inkpen and Wilson 2013). The most commonly applied types of models in the field of physical geography can be summarized as conceptual models, heuristic approaches, empirical-statistical models and deterministic models.

2.1.1. Conceptual models

Conceptual models contain a high degree of abstraction and require much knowledge about the underlying processes involved and they reflect the researcher's view of how reality, respectively its variables and relations that were identified as crucial components for the operation of the section of reality that is under investigation, are interconnected (Inkpen and Wilson 2013). Thus, conceptual models reflect the underlying theory about the operation of the physical environment as identified by the researcher. Figure 2.2 depicts a typical conceptual model in physical geography representing shallow ground-water conditions in hillside soils. This figure is a visual representation of the researcher's thought process and how he believes reality to be structured and how the processes in place need to be studied in order to understand the dynamics of the system. Consequently, conceptual models drive the manner in which research is undertaken (Inkpen and Wilson 2013).

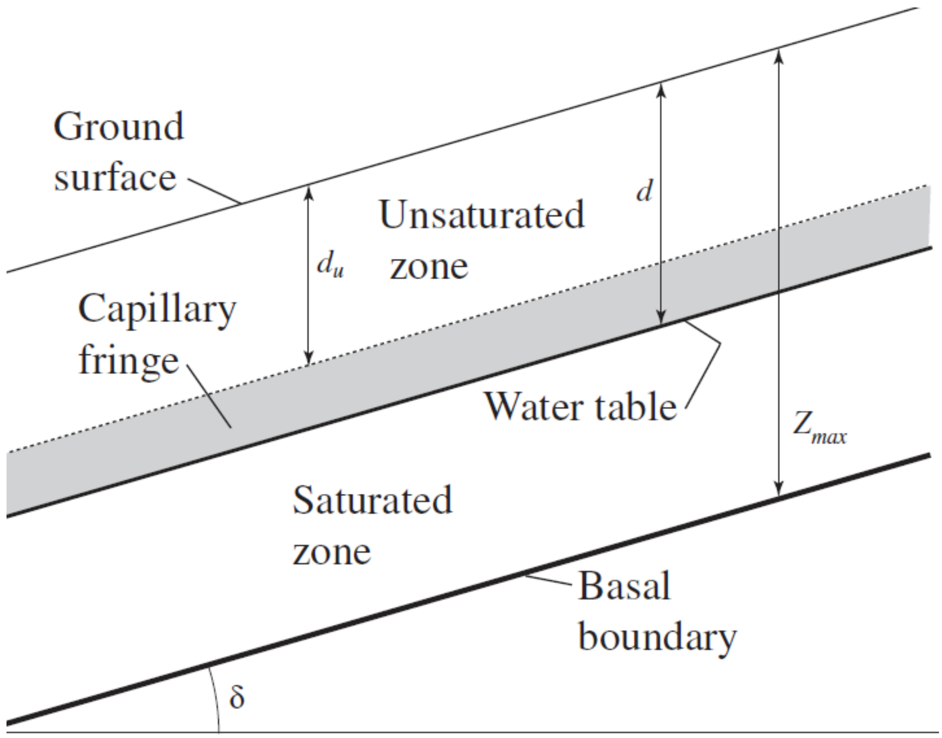


Figure 2.2: Conceptual model of shallow ground-water conditions in hillside soils. The unsaturated zone above the water table has depth d_u. The capillary fringe is between the unsaturated zone and the water table at depth d. The lower boundary, which is treated as impervious in this model, is at depth Z_{max} (Baum et al. 2008)

2.1.2. Heuristic methods

Those subjective approaches are based on expert judgement. A group of specialists assign probabilities to quantify certain process rates, hazard potential, etc. Traditional methods in the domain of heuristic approaches are qualitative or semi-qualitative methods such as geomorphological mapping or index overlay mapping (van Westen et al. 2006). A common and more recent way to systematize heuristic evaluation is based on decision trees (Wong et al. 2005, Corominas et al. 2014). With an increasing number of possible feature characteristics and outcomes, the visual representation of a decision tree spreads out like the branches of a tree (hence its name). Fig. 2.3 shows an example of a decision tree for classifying landslide susceptibility. In general, expert judgement serves as a classifier with respect to variable importance (e.g. average slope angle > 28.7° has a high impact on slope stability). This means that in order to quantify the probabilities of a certain alternative, the branching node probabilities have to be determined. The product of the respective branching node probabilities ultimately results in a particular outcome, such as a slope failure map (Corominas et al. 2014). Although heuristic methods are highly subjective, expert based selection and weighting of variables can indeed serve as a valid alternative to purely automated selection of potentially biased input data, especially when applied over larger areas (van Westen et al. 2006, Steger et al. 2016a).

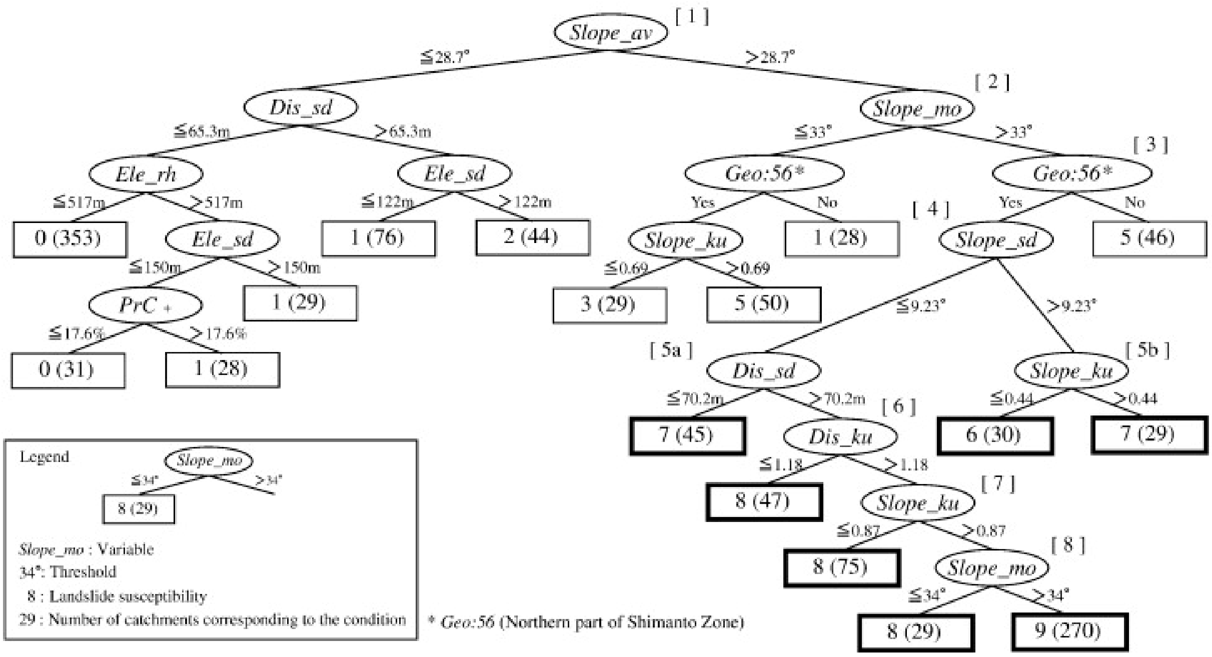


Figure 2.3: Example of a decision tree for classifying landslide susceptibility in which leaves with high landslide susceptibility (more than 5) are emphasized. Variables are listed in the original source in Saito et al. 2009

2.1.3. Empirical-statistical models

"Empirical-statistical models use statistical methods to obtain mathematical expressions that are meant to represent the physical system under study. […] In this way dependent variables are modelled by independent variables and causation is implicit within the model structure" (Inkpen and Wilson 2013, p. 181). Statistical models are highly dependent on the input data and the data has to be available in a suitable format. For statistical landslide susceptibility assessment for example, the dependent variable is usually the landslide initiation location, either as points or polygons, based on a digital elevation model or its derivatives while the independent variables (predictors) trying to explain a landsliding location are commonly distributed maps representing, amongst others, geology, vegetation cover, slope, aspect or distance from rivers. Landslide inventories are often purposefully generated for the use in statistical models (Petschko et al. 2015). Empirical-statistical models are often attributed as simplified *input-output* models that aim at matching input variables to output values through the development of a mathematical expression (Inkpen and Wilson 2013). However, this simplicity comes with a number of drawbacks (van Westen et al. 2006):

- in most cases, only factors that can easily be mapped or derived from a digital elevation model are taken into account
- the generalization of causative factors: landslides are assumed to occur under the same combination of conditions throughout the study area and through time
- different landslide types have different causative factors: a complete and unbiased landslide inventory is almost impossible to come by and in most cases, a proper differentiation of landslide types is lacking

However, in case the underlying data set serving as the modeling basis is good, results of statistical models can perform reasonably well as was demonstrated by rigorous model validation (Steger et al. 2016b). More commonly applied statistical models in landslide research belong to the group of linear models (e.g. logistic regression), while non-linear statistical models (e.g. generalized additive models) or flexible machine learning techniques (e.g. Support Vector Machines, Random Forest) are being increasingly used to increase predictive performance (Micheletti et al. 2014, Goetz et al. 2015, Pham et al. 2016). Although none of these models explicitly state why input and output are intertwined, an appropriate variable selection "does imply a set of processes and a view of the physical system and its operation. Likewise, even if the system and processes cannot be stated accurately, the fact that the model seems to produce predictable results may be sufficient for the model to function adequately in its particular context" (Inkpen and Wilson 2013, p. 181).

2.1.4. Deterministic models

Deterministic models, as opposed to statistical models, do not simply establish statistical relationships between dependent and predictor variables, but use mathematical expressions to actually represent relationships between elements. Working from this basis, "relationships are deduced and the operation of the resultant model can be explained by reference back to these basic principles. The important aspect of this type of modeling is that the relationships must be formalised as mathematical expressions. The behavior of the resultant model is explainable by reference back to the basic principles and their formal relationships. Deterministic modeling relies upon any variable or entity and its relationships being expressed or reduced to a set of basic and fundamental physical principles. This means that it is essential that abstract axioms are linked to real-word entities" (Inkpen and Wilson 2013, p. 180).

Deterministic model output is more concrete and consistent when compared with heuristic and statistical modeling approaches, "given the white-box approach of describing the underlying physical processes leading up to the phenomena being modelled" (Corominas et al. 2014, p. 225). Scale issues are a major concern when applying distributed deterministic models in physical geography when the scale and number of relationships defining the reality under study increases. This ultimately leads to the question whether appropriate laws or relations that are valid for one scale are transferable to another scale since the scale of measurement generally differs significantly from the scale at which the applied model requires 'effective' parameter values to be specified (Beven 1996).

In physical geography, a common distinction of physically based models is made with respect to how the temporal components are treated: models are attributed as either static or dynamic. Static models aim towards the determination of the stimuli that cause, for example, slope instability. Dynamic models consider a temporal component in order to identify cause and effect relationships which makes them especially suitable for simulating future changes under varying initial conditions (van Westen et al. 2012). Based on the used model, physically based modeling approaches are capable both for addressing the spatial and temporal variation of landslide initiation (Horton et al. 2013, Formetta et al. 2016, Zieher et al. 2017) or runout (Hussin et al. 2014, McDougall 2017, Strand et al. 2017). Commonly, their main drawbacks are stated as being computationally very demanding due to the high spatial resolution that is required and a sufficiently high measurement precision of input parameters (van Westen et al. 2012). Therefore, deterministic methods used to be limited to site-specific or local scale applications only (Tab. 2.1). Additionally, the geological and geomorphological conditions should be fairly homogenous and landslides should ideally be rather simple in order to reduce bias and additional uncertainties. As opposed to statistical based approaches, which require a comprehensive landslide inventory that serves as the dependent variable, an incomplete inventory can be sufficient for physically based models as the inventory serves only as a means for model validation and calibration (Corominas et al. 2014).

Table 2.1: Traditionally recommended quantitative methods for landslide susceptibility analysis at different scales (Corominas et al. 2014)

	Quantitative methods	
	Data-driven statistical methods	Deterministic physically based methods
National scale (<1:250,000)	No	No
Regional scale (1:25,000–1:250,000)	Yes	No
Local scale (1:5,000–1:25.000)	Yes	Yes
Site-specific (>1:5,000)	No	Yes

For larger scale applications, infinite-slope based approaches are the most commonly applied family of landslide models as they generally outperform other approaches that try to introduce more complex landslide geometries (Zieher et al. 2017). The most popular static infinite-slope based models for large scale applications are SINMAP (Pack et al. 1998) or SHALSTAB (Dietrich and Montgomery 1998), with regard to dynamic models, TRIGRS (Baum et al. 2008, 2010), STARWARS+PROBSTAB (van Beek 2002) or r.slope.stability (Mergili et al. 2014a and 2014b) should be mentioned. Albeit its drawbacks, physically based models contain “a higher predictive capability and are the most suitable for quantitatively assessing the influence of individual parameters that contribute to shallow landslide initiation” (Corominas et al. 2014, p. 225). Recent advancements in deterministic landslide modeling aim towards a more probabilistic attempt on how to treat spatial parameter variability and uncertainties over larger study areas (Lari et al. 2014, Raia et al. 2014, Salciarini et al. 2017, Canli et al. 2017b).

2.2. Issues in modeling – parametrization, validity and uncertainty

Philosophical questions arise with respect to obtaining data in various ways (Frigg et al. 2015): a) the theory-ladenness of observations. Strictly speaking, instrument obtained data has to be independently tested and confirmed; b) model-filtered raw data and their symbiotic relationship between data and models. Again, those have to be independently tested and confirmed. It is doubtful whether such model-filtered data can be trusted if the models were tested only by the data that they are supposed to be correcting and filtering (*confirmatory circle*); c) the suitability of proxy data in the absence of directly measurable raw data. In geosciences, data is extensively used in the construction of models: models in general contain many observationally derived approximations and heuristics with parametrizations that represent processes incapable of explicitly resolving the spatial or temporal resolution of the model. Consequently, they are replaced by simplified data-driven processes (*data-laden models*) that are partly also physically motivated (Frigg et al. 2015). This data-ladenness is a widely acknowledged phenomenon in modeling environmental processes in which landslide modeling makes no exception.

Model application requires the researcher to make some a priori decisions when deciding upon which model to use for the underlying research question. Model building itself undergoes the same thought process as building a system, namely making the initial decision to what include and exclude from the model and where to draw the boundaries. This decision predetermines the entities and relationships that will be modelled and explained. Any other attributes not considered in the model will be viewed as irrelevant for the model for the sake of simplicity, yet they define and determine the behavior outside of the model's scope. Modeling then requires the parametrization of the defined entities where the choice of these assigned values can substantially affect the operation of the model. The range of input values that are suitable and acceptable for an entity in a model may reflect the experience of the modeler or constraints in the modeling process (Inkpen and Wilson 2013). With regard to empirical-statistical models, dependent variable values are calculated as a mathematical function on independent variable values. Such models can be considered as black-box models since they do not specify anything about how or why inputs are transformed into outputs. This inevitably runs the risk of making erroneous associations between variables whose statistical correlation may either be coincidental or contingent upon some intervening process ignored by the model. Critical realists tend to reject such empirical-statistical models by claiming that they commit the inductivist fallacy of affirming the consequent and failing to explain why the value of a dependent variable necessarily depends on that of an independent one (Demeritt and Wainwright 2005).

On the contrary, deterministic models are generally seeking for a grand unifying theory that tries to find an answer to the question whether processes at higher scales can entirely be reducible to those operating at lower scales. Scale related questions still remain unanswered in most cases and physical geographers tend to be more concerned with practical and computational difficulties of a strictly deductive-deterministic approach to process modeling. According to Demeritt and Wainwright (2005), there are especially those two technical issues that lead to this situation:

a) *Implicit parametrization*: The laws of physics are so abstract that it is required to specify certain boundaries and initial conditions in order to close the gap between the model's underlying theory and its contextual application. In many (if not all) cases, those values to specific certain entities are incompletely known. In general, the process of parametrization links field and other data with the model. As it is usually not possible to measure all parameters directly, implicit parametrization is required that is applicable to the original modeling scale. Often it is the parameters that are just unsuitable rather than the model itself.

b) *Appropriateness of equations*: The identification of suitable equations that are both appropriate and analytically tractable is difficult for environmental processes, especially when changes over time are involved. Additionally, most deterministic models consist of non-linear differential equations (such as the commonly used Richard's equation to represent the movement of water in unsaturated soils) that are difficult to approximate due to the absence of a closed-form analytical solution. Hence, model developers try to find an approximation by means of numerical iteration or finite difference calculation to provide analytically tractable solutions.

This means that values that were initially fixed by reference (e.g. direct measurements or literature values) need alteration or they become optimized (or *fine-tuned*). This optimization based on real-world data is often referred to as *physically based* – with the consequence of limiting the modeling application to the time and place specific domain of optimization. The model then may precisely match the outputs of the empirical data set. For example, an optimized set of soil cohesion and friction angle values may be good for explaining variations in distributed slope stability in a specific location for a particular time period, but they are not as good at explaining variations somewhere else during other time periods. This means that due to parameter tweaking in a model constructed and explained in terms of physical processes, it is entirely unknown whether the physical model is correct and a valid explanation of the empirical data or whether the model is only correct because the parameters have been adjusted to achieve the greatest possible match to the empirical data (Inkpen and Wilson 2013).

This well-established calibration procedure raises the legitimate question whether such a model is transferable to other places and other times. Important in that regard is estimating the validity of a model. Usually this is done by matching the model outcomes to reality which in turn requires the modeler to define a set of criteria against which the outcome properties and the *real* properties can be compared (Inkpen and Wilson 2013). Often, validation and verification are used interchangeably, which is not quite correct from a philosophical point of view. The term validation does not necessarily denote an establishment of truth. Rather, it provides a legitimacy in terms of arguments and methods (Oreskes et al. 1994). When it comes to validating single model outputs based on a best-fit realization (either for statistical or deterministic approaches), there always lies the confirmation bias trap that is omnipresent in landslide research. When comparing a result predicted by a model with observational data and the comparison is unfavorable, the modeler continues to work on the model until a fit is achieved. But the even bigger dilemma awaits if there *is* a match between the model result and observational data because the modeler may be tempted to claim that the model was verified. But again, this would be committing the logical fallacy of *affirming the consequent* (Oreskes et al. 1994).

If a model fails to reproduce observed data, this gives a hint that the model is not yet reliable in some way, but the reverse is rarely the case. Even when using a calibrated model, it is safe to say, at best, that it is empirically adequate. But admitting that calibrated models do need 'additional refinements' suggests that the empirical adequacy of numerical models is forced. So even if a model is consistent with present and past observational data, there is no guarantee that the model will perform in equal measure when predicting the future (Oreskes et al. 1994). But when is it sufficient to attribute a model as a valid representation of reality? Inkpen and Wilson (2013) identified two important issues: "First, how close does the match have to be for the model to be validated? Second, even if a match can be identified how does the modeler know that the match to reality is for the reasons modelled? The matching of model outcomes to observations relies upon there being a clear correspondence or translation from model to some measurable property of reality. Even if a clear and justifiable translation exists, the issue of how 'close' the match in values needs to be before the model is validated needs to be clear. This decision is likely to be driven by the researcher(s) themselves, by the traditions and training in the subject as well as by the potential requirements of the models used" (Inkpen and Wilson 2013, p. 188).

Brown (2004) distinguishes a tripartite division of reality: a) real mechanisms; b) actual events; and c) empirical observations. Those interactions create a feedback loop in which the past persists through the present and into the future to form an *environmental chain of causality*. Within this chain, the observational part always depicts the outcomes of those interactions, but not the causal mechanisms themselves. However, assessing patterns of similarities or difference can indeed provide insights into the real mechanisms involved and also poses the essence of research that goes beyond what pure observation is capable to answer. Due to the non-linear behavior of most environmental processes, a lack of process understanding, multiple (interacting) parameter values, different measurement scales, spatial and temporal heterogeneity or the dependence on the model structure, uncertainties on all ends of environmental modeling are inevitable. This leads environmental processes to be highly dependent upon their contingent conditions as a result of our inability to explain a unique causal world (Brown 2004).

Consequently, uncertainties can either be of epistemic or ontological (aleatory) nature. While the first arises through our lacking knowledge about the nature of the reality and the system under study, the latter occurs as a result of the inherent variability of the reality under investigation (Walker et al. 2003). Lehmann and Rillig (2014) suggest to clearly distinguish between uncertainty and variability that manifests in time and space. While uncertainty is considered as a measure of unexplained variation (i.e. measurement errors, also lack of understanding about cause and effect), spatial and temporal variability in environmental sciences reveal themselves as spatial heterogeneity and will not shrink with scientific progress. Known variation should consequently not be referred to as uncertainty, but explained as variability (Fig. 2.4). Thus, working with environmental systems requires the researched to distinguish between a lack of process understanding and the failure to adequately capture the heterogeneity of responses (Lehmann and Rillig 2014).

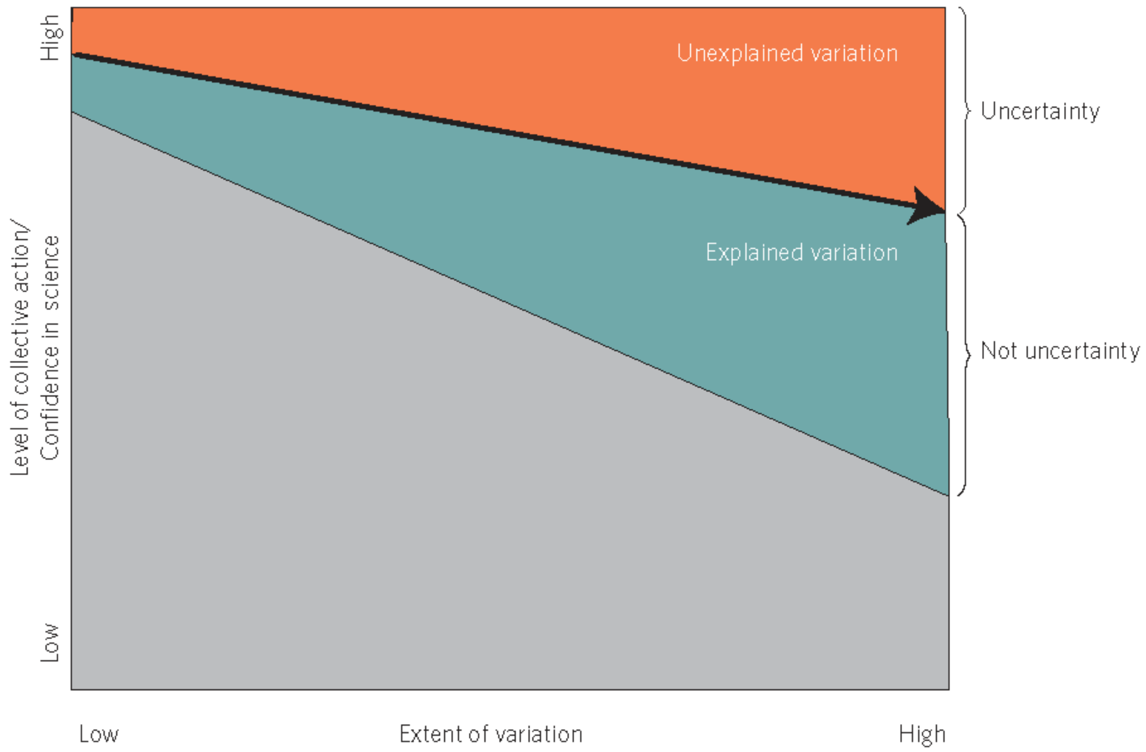


Figure 2.4: Comparison between uncertainty and variation. The reduction in confidence due to unexplained variation, such as uncertainty, decreases through progress in science. Progress in science, however, will not decrease the total variation (Lehmann and Rillig 2014)

Environmental systems reveal a high degree of non-linearity as many of its entities and relations are indeterminate because their causes of change are unknown a priori (e.g. weather predictions based on the same initial conditions may vary drastically the longer the forecasting period). It is generally accepted that our knowledge and understanding of nature is limited, yet deterministic strategies remain to be quite popular in geomorphology. Such strategies “obscure the context of ‘what we know’, as well as ‘how we come to know’ and fail to encourage the transparency of reasoning required for policy-relevant research where even the definitions of an environmental problem may be highly contested” (Brown 2004, p. 368f). Acknowledging the presence of many plausible theories is the core of the equifinality concept that challenges deterministic believe. Equifinality revolves around the rejection of the concept of the optimal model in favor of multiple possibilities for producing acceptable simulators (Beven and Freer 2001). This concept should not come surprising given our understanding of physical theory that there is a plethora of interactions among the components of a system whose resulting representations may be equally acceptable.

Research generally follows a working paradigm that should lead to realistic representations of the real processes and characteristics. This idea of identifying a single optimal representation of reality is very distinct in environmental sciences. A major problem arises from the scale discrepancy between sampling and distributed modeling where the use of global parameters undoubtedly leads to errors in predicting local responses at points with unique characteristics (Beven and Freer 2001). By acknowledging that there are many different model structures or many possible parameter sets scattered throughout the parameter space, the range of predicted variables is likely to be larger than linearized solutions would suggest. This equally means acknowledging that there are uncertainties inherent surrounding the area of parameter space around the optimum. As a result, such approaches allow non-linearity to be taken into account (Beven and Freer 2001).

Geomorphological systems can indeed be considered as transient, inheriting remnants of past and present processes. Environmental systems can exhibit certain degrees of chaotic behavior which results in an inability to express the trajectory of their development on the basis of present-day evidence alone. As a consequence, equifinality should not be considered as an indication of a poorly developed methodology, but as something inherent in geomorphological systems (Beven 1996). While it can be of benefit to perform analyses of uncertainty, it needs to be stressed that many uncertainties cannot be quantified or remain difficult to quantify with available information. Uncertainty analysis, however, promotes openness, which implies according to Brown (2004, p. 375):

a) the criteria for evaluating uncertainty are made clear;
b) different ‘informed opinions’ are canvassed when decisions involve large risks or are taken in the ‘public interest’ (i.e. different ‘confidence models’ are proposed alongside different environmental models, and their ‘goodness’ evaluated);
c) the criteria for selecting ‘informed opinions’ are clear;

d) the purpose of assessing uncertainty is clear and
e) expressions of uncertainty are interpretable by different groups of scientists and by non-scientists.

In those cases where decision makers (scientists, practitioners, politicians, etc.) are only interested in reinterpretations or absolute statements of reality, uncertainty analysis might not be the appropriate tool as it is unlikely that decision-making is improved in the short-term. However, if uncertainty analysis is performed as an act of volition and determination to achieve transparency and accountability in scientific research, this can lead to improvements in the quality of data and models and ultimately to our understanding of environmental processes (Brown 2004).

3. Physically based landslide modeling and early warning systems — current approaches and challenges

3.1. Spatial variability and uncertainties in regional scale landslide modeling

Using averaged parameter values from locally measured geotechnical and hydraulic parameters is a common practice for parameterizing physically based landslide models (e.g. Thiebes 2014, Tofani et al. 2017, Zieher et al. 2017). Additionally, databases, published or unpublished technical reports or lookup tables may serve as a source for common parameters (e.g. Schmidt et al. 2008, Kuriakose et al. 2009, Mergili et al. 2014b). In more recent years, the probabilistic treatment of modeling parameters has gained quite some popularity in the landslide community. Probabilistically derived parameters have the potential to consider uncertainties and inherent variability in a way that can be quite beneficial in the absence of a very dense measurement network (Canli et al. 2017b). In general, geotechnical and hydraulic parameters are represented with a univariate distribution consisting of random variables that are based on an underlying probability density function and statistical characteristics (Fan et al. 2016). Common parameters in deterministic landslide modeling that are treated in a probabilistic way are the friction angle or cohesion (e.g. Park et al. 2013, Chen and Zhang 2014, Raia et al. 2014, Salciarini et al. 2017).

While very dense measurement networks at regional scale for assessing required modeling parameters are highly desirable, this goal is hardly achievable in reality. Performing geotechnical and hydraulic measurements at regional scale is difficult, time-consuming and very expensive. Consequently, using such parameters within spatially distributed physically based models is a rather challenging task and in general there is no approach that is universally accepted (Tofani et al. 2017). Even if there is measured data available for one, some or even all parameter values in a model to be able to specify distributions and covariances for the parameter values, some methodological obstacles remain. There is, for example, no guarantee that values measured at one scale will reflect the effective values required in the model to achieve satisfactory predictions of observational data (Beven and Freer 2001).

Reasons that lead to this spatial variability in soil formation processes are manifold. Weathering processes, biological perturbations, atmospheric interactions, etc. are commonly listed processes that lead to spatially varying soil and hydraulic properties (Fan et al. 2016). Yet again, scale matters a lot when considering sampling locations for a regional scale study area. At the slope or catchment scale, variability lacks a pronounced spatial organization. Case study based subsurface exploration, such as Canli et al. (in prep.), clearly demonstrate this randomness in spatial organization (Fig. 3.1). This is less the case at larger scales, where several superimposing factors contribute to spatial variation, such as topography, differences in soil depth, -type and -texture, vegetation characteristics, as well as rainfall patterns. This suggests that the larger the scale, the more soil forming processes manifest a persistent deterministic signature due to the predetermined topography, geology, climate, and other factors (Seyfried and Wilcox 1995, Fan et al. 2016).

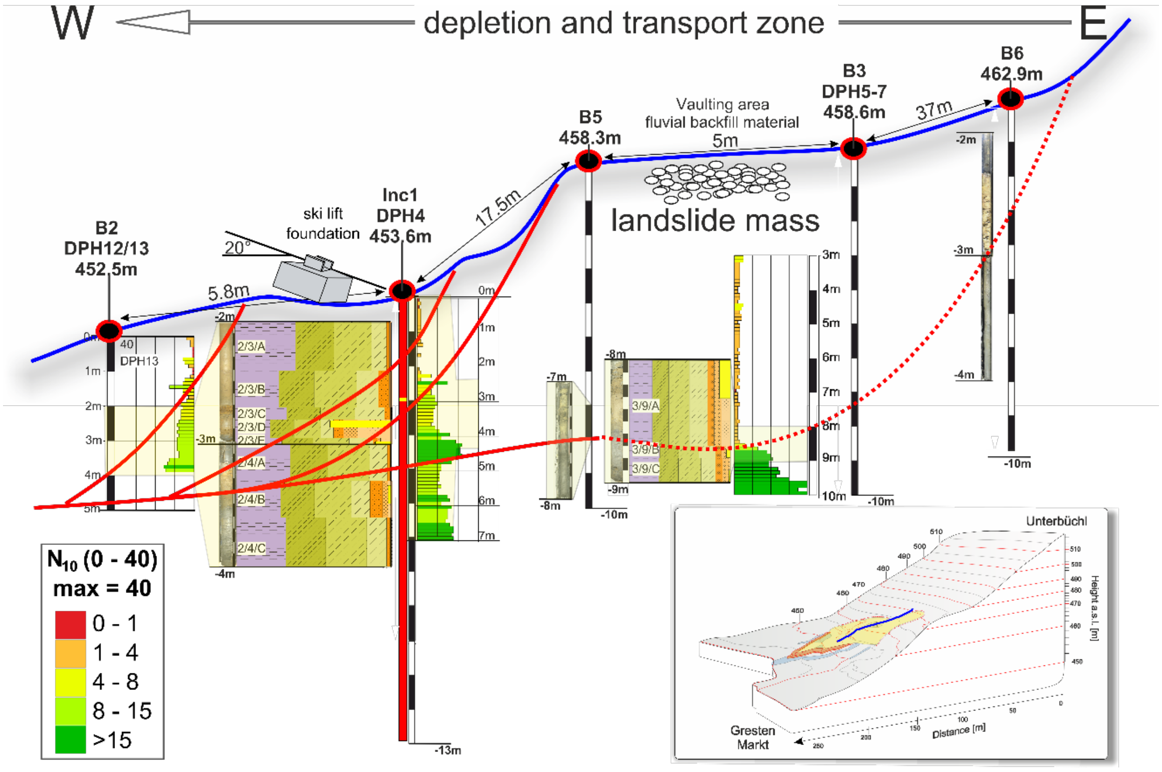


Figure 3.1: Proposed underground model of the Salcher landslide (Austria) based on all obtained information (inclinometers, drill cores, penetration resistance). Spatial variation in both, horizontal and vertical, direction seem to lack a pronounced spatial organization, which is a challenge for slope scale modeling, yet an even bigger challenge for model parametrization at regional scale (Canli et al. in prep.)

To overcome this problem, Neves Seefelder et al. (2016) suggested to apply parameter ranges in physically based modeling applications as their findings yielded results comparable in quality to those derived with best-fit narrow ranges. By acknowledging the fact that at larger scales geotechnical and hydrological parameters are highly variable, uncertain and often poorly understood, narrow parameter rangers or even singular combinations of parameters come with the risk of being highly inaccurate (Neves Seefelder et al. 2016). Canli et al. (2017b) therefore suggest that it might be sufficient to work with literature data for model parametrization alone instead of in situ measured data when working at regional scale.

3.2. Consideration of spatial and temporal dynamics: rainfall as a crucial component

Providing precise and timely rainfall information, no matter whether the approach of the warning system is based on empirical rainfall thresholds or combined hydraulic and slope stability modeling, can be regarded as the most important aspect of any landslide early warning system (Canli et al. 2017a). The spatial variability of real-time rainfall distribution is a crucial aspect to be considered, yet it is insufficiently addressed in early warning applications where the scarcity of rain gauges is a common argument (Chiang and Chang 2009). Common approaches for threshold based early warning systems to determine rainfall for a single landslide site or a specific region are:

- utilization of single rain gauges near a specific landslide site (e.g. Capparelli and Tiranti 2010)
- selection of rain gauges as representative locations for a predefined region (e.g. Segoni et al. 2015, Rosi et al. 2015)

However, using only single point measurements as representative locations is often not really suitable, given the fact that such locations are not only dependent on the distance from the landslide itself, but also from other influencing factors such as elevation, aspect or the wind direction (Aleotti 2004). In a more advanced attempt, Lagomarsino et al. (2013) artificially split their study area in smaller units (*territorial units*; TU) and assigned one representative rain gauge to each TU that indicates areal rainfall for each TU. While this area can be arbitrarily large or small, rainfall is still only considered as uniform across the entire area.

The representation of areal rainfall based on rain gauge measurement is quite common in landslide modeling and early warning (e.g. Segoni et al. 2015, Rosi et al. 2015). However, this approach is rather critical since measurements representing an entire area have been taken from a continuum in space (Oliver and Webster 2014). As a consequence, spatial prediction methods were applied to point measurements to regionalize rainfall in a spatial manner. Historically, spatial prediction was undertaken by purely mathematical interpolation approaches that considered only systematic or deterministic variation, but not any error. Geostatistical prediction, and here it is primarily kriging, is the logical successor that overcomes most of these drawbacks contained in deterministic methods (Webster and Oliver 2001). Geostatistics aims explicitly at correctly portraying spatial variation of spatial random variables such as rainfall (Srivastava 2013).

There is a substantial amount of literature available that aims at comparing different interpolations methods for assessing spatial rainfall distribution, mainly in the fields of hydrology or hydro-meteorology (e.g. Ly et al. 2013, Mair and Fares 2011, Schuurmans et al. 2007, Haberlandt et al. 2007, Goovaerts et al. 2000). Literature generally suggests a separation between deterministic and geostatistical approaches. The most common methods for the deterministic estimation of rainfall are Thiessen polygons and Inverse Distance Weighting (IDW) (Ly et al. 2013). The Thiessen polygon method is one of the earliest and simplest techniques. The targeted region is divided into polygons by perpendicular bisectors between the individual sampling locations. In each polygon, all points are nearer to its enclosed sampling point (the rain gauge) than to any other sampling point (Webster and Oliver 2001). Godt et al. (2006), for example, used this technique to characterize rainfall for shallow landsliding in their study area.

Among the deterministic spatial interpolation techniques, the IDW method is one of the most popular ones. It is based on inverse functions of distance that put a larger weight on unknown locations that are closer to a sampling point that those further away. The advantage of weighting by inverse squared distance is the quick diminishing of the relative weights with increasing distance, making the interpolation sensibly local. The weighting function itself, however, is arbitrary (Webster and Oliver 2001). For modeling rainfall-induced landslides, Chiang and Chang (2009) apply an IDW approach to characterize the spatial rainfall distribution for their study area.

Kriging, on the other hand, is the predominant geostatistical method that links mathematical concepts with geoscientific requirements. Kriging is a generalized least squares regression technique that accounts for the spatial dependence between observations (Schuurmans et al. 2007). Unlike deterministic interpolation techniques, kriging offers a measure of certainty (the kriging variance). In kriging, a weighted sum of the available point observations is calculated in order to estimate the unknown target variable. This is done by minimizing the variance so that the interpolation is biased as little as possible (Ly et al. 2013). This assumption of stationarity in kriging allows to have the same degree of variation from place to place and that the covariance between two observations only depends on the distance between these observations (Oliver and Webster 2014, Ly et al. 2013).

Univariate kriging techniques, such as ordinary kriging, use rain gauge information alone while multivariate kriging techniques, such as kriging with external drift or ordinary cokriging, incorporate additional predictor values (e.g. weather radar information or elevation) to improve the kriging prediction (Goovaerts 1997). Spatial interpolation based on kriging is in general a labor-intensive task as many a priori decisions are necessary for fitting the underlying variogram, on which kriging is based on. Automating attempts in that regard exists, however, when fitting a variogram without supervision, errors might occur. Interpolated rainfall in landslide research is rarely used as an alternative to purely rain gauge based single point measurements. However, specifically for the purpose of implementing interpolated rainfall data into landslide early warning applications, Canli et al. (2017a) proposed an approach to automate the creation of the underlying variogram. Initial modeling parameters were defined and iteratively fitted to the most suitable variogram model. Validation results demonstrated the feasibility of this approach, especially as it is possible to couple the automated interpolation methods with web scraped real-time rainfall data from multiple sources (Canli et al. 2016). Fig. 3.2 shows as an example automated hourly spatial interpolation results from Canli et al. (2017a).

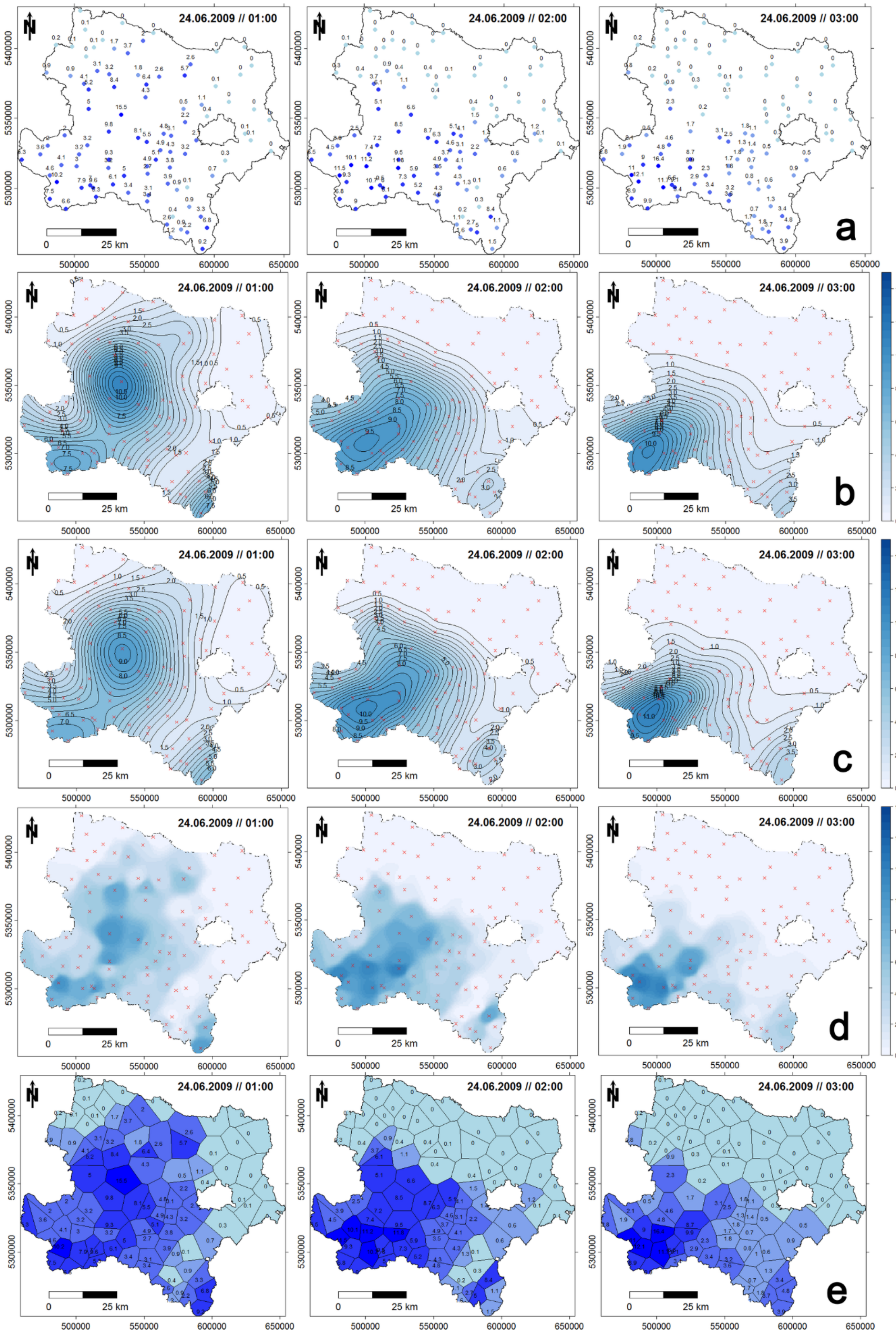


Figure 3.2: Results from the automated spatial interpolation: a) Hourly rain gauge data; b) Ordinary Kriging (OK) without filtering; c) OK with filtering; d) IDW interpolation; e) Thiessen polygons. Lines in the b) and c) estimates indicate areas with equal amounts of rainfall (isohyets). Points in the maps (b) through(d) indicate rain gauge locations

Weather radar data is an attractive alternative, or more a supplement, to determine continuous rainfall fields in near real-time. Due to their high spatial (approx. 1 km) and temporal (10 minutes and less) resolution, Doppler radar technology is highly beneficial for providing spatially distributed rainfall data for landslide studies (Chiang and Chang 2009). The quantification of rainfall estimates using Doppler radar can provide a real-time comparison with rainfall thresholds to form the basis of landslide warnings (Wieczorek and Guzzetti 1999). In case the rain gauge network in a region is not sufficiently dense, radar data is capable of capturing the spatial variation of rainfall fields much better than gauged data (Yang et al. 2004, Segond et al. 2007). However, Doppler radar is only capable of indirectly measuring precipitation, since this technology quantifies rainfall amounts as a magnitude of measured reflectivity from hydrometeors in the atmosphere (Harpold et al. 2017). This requires calibration with actual rain gauge data in order to adjust the high resolution spatial pattern that radar data offers with actual measured rainfall amounts.

The potential of rainfall radar for applications in landslide related research questions remains underexploited as there are only few relevant studies addressing this technology (e.g. Crosta and Frattini 2003, Schmidt et al. 2008, Chiang and Chang 2009, Segoni et al. 2009). Although radar technology has undergone a lot of progress in recent years, the associated uncertainties, and generally low success rates (in terms of correctly predicted landslide occurrences), reduce its applicability within the landslide community (Canli et al. 2017a). Schmidt et al. (2008) and Segoni et al. (2009) concluded that the meteorological uncertainty has the highest influence on slope stability analyses that serves as the basis for physically based landslide early warnings. The hydro-meteorological community is far more involved with the utilization of radar data for the deduction of continuous rainfall fields. However, they share the same concerns (Jasper et al. 2002) or even have taken a step further by implementing numerical weather predictions (Cloke and Pappenberger 2009).

With emerging high-resolution satellite technology, this means of assessing the spatial extent of rainfall could bring huge benefits for dynamic landslide modeling approaches at larger scales (Rossi et al. 2017). This is even more so the case where rain gauge or radar rata is unavailable. Satellite-based rainfall estimates provide synoptic estimates of the spatial distribution of precipitation events (Chappell et al. 2013). Recent satellite data provides those estimates at 0.5 to 3 hours intervals and spatial resolutions between 0.07° and 0.25° (Joyce et al. 2004, Kubota et al. 2007, Huffman et al. 2007, 2010). Until recently, NASA's Tropical Rainfall Measuring Mission (TRMM), which accumulated almost two decades of precipitation data by now, provided the most valuable data archive for global precipitation data (Kirschbaum and Petel 2016). When NASA launched the Global Precipitation Mission (GPM) as a follow-on mission to TRMM in 2014, a huge popularity boost in satellite based precipitation data could be observed (Harpold et al. 2017). GPM provides rainfall and snowfall estimates every three hours. It is equipped with sensors that are far more advanced and that permit better quantification of the physical properties of precipitation particles (Hou et al. 2014). Not many landslide studies have been conducted that incorporate satellite based precipitation data (e.g. Rossi et al. 2012, Kirschbaum et al. 2015, Rossi et al. 2017), and those that exist still rely on TRMM data. In the upcoming years, however, near real-time GPM data with higher spatial resolution holds great potential for the applications in landslide early warning systems (Rossi et al. 2017, Stanley et al. 2017).

All those previously described means of assessing rainfall magnitudes (single location gauged data, interpolated data, radar data, satellite data) share one common ground: they all rely on direct observations. However, for providing timely and effective information that allows individuals exposed to a hazard to act and to avoid or reduce their risk and prepare for effective response, rainfall data as the main trigger for landslides needs to be provided in advance. This inevitably suggests the utilization of numerical weather predictions for such purposes to shift the current paradigm of *warn on detection* towards a *warn on forecast* approach (Stensrud et al. 2009). Flood forecasters in hydrological sciences have adopted such NWP in the last decade into so-called ensemble prediction systems (EPS). The World Meteorological Organization (WMO 2012) defines them as follows: “numerical weather prediction (NWP) systems […] allow us to estimate the uncertainty in a weather forecast as well as the most likely outcome. Instead of running the NWP model once (a deterministic forecast), the model is run many times from very slightly different initial conditions. Often the model physics is also slightly perturbed, and some ensembles use more than one model within the ensemble (multi-model EPS) or the same model but with different combinations of physical parametrization schemes (multi-physics EPS). […] The range of different solutions in the forecast allows us to assess the uncertainty in the forecast, and how confident we should be in a deterministic forecast. […] The EPS is designed to sample the probability distribution function (pdf) of the forecast, and is often used to produce probability forecasts – to assess the probability that certain outcomes will occur" (WMO 2012, p. 1).

EPS approaches started to be viable for smaller scale processes (such as landslides or flash floods) when accurate convective-scale precipitation forecasting was available (as opposed to previously used EPS systems that relied on global or regional rainfall predictions). With spatial resolutions ranging from 1-4 km, convective-scale NWP aim at predicting small-scale atmospheric features such as location and the intensity of thunderstorms (WMO 2012). Besides technical advances, it was mainly the computational challenges that withheld this technology from operational mode and that saw its practical implementation only within this decade (WMO 2012). In, 2012, the German Weather Service (Deutscher Wetterdienst - DWD) started operational mode for their COSMO-DE-EPS with a resolution of 2.8 km (Baldauf et al., 2011, Gebhardt et al., 2011). Similar operational forecasting systems with comparable spatial resolutions have been implemented in the last couple of years, e.g. the AROME model in France (Seity et al., 2011), the MOGREPS-UK model in the UK (Golding et al., 2016) and High Resolution Rapid Refresh (HRRR) model in the USA (Ikeda et al., 2013).

3.3. Current issues in probabilistic landslide modeling and early warning

In the last couple of years, regional scale probabilistic modeling approaches gained increasing popularity. The two most significant explanations for this observation lie in the reduction of computational costs on the one hand, and decreasing confidence in purely deterministic approaches on the other hand (Canli et al. 2017b). Haneberg (2004), Park et al. (2013), Raia et al. (2014), Lee and Park (2016), Zhang et al. (2016) or Salciarini et al. (2017) and others use a probabilistic approach to characterize soil properties at regional scale by randomly selecting variables from a given probability density function.

In two recent studies, Neves Seefelder et al. (2016) and Canli et al. (2017b) propose the application of rather broad parameter ranges for model parametrization instead of best-fit narrow ranges as this is suggested to be a more honest approach in selecting modeling parameters. Canli et al. (2017b) uses this rather broad parameter range in a probabilistic approach to produce a multitude of ensemble members based on hourly rainfall input to express the range of equally possible model iterations. Their case study revealed that broad parameter ranges are indeed feasible for achieving rather narrow ensemble spreads over large areas in a fully automated approach. However, as they were severely lacking computational power to use their result in a real-time scenario for issuing hourly probability of failure maps, it is not yet ready to be used in an early warning context. Schmidt et al. (2008) proposed a coupled regional forecasting system in New Zealand based on multiple process based models (NWP, soil hydrology, slope stability). However, as innovative their research was, it did not find any continuation, probably due to unsatisfying initial results with the rather coarse data back then. Consequently, none of those probabilistic approaches are operated in spatial real-time early warning systems, not even on a prototype basis.

While hydrological sciences have started operational mode of probabilistic ensemble prediction systems (EPS) based on NWP input (Alfieri et al. 2012b, Bartholmes and Todini

2005, Siccardi et al. 2005, Thielen et al. 2009, Vincendon et al. 2011), landslide forecasting is still far from there. Reasons for that shortfall might be found on the conceptual side of model formulation and on the funding situation in landslide research in general (Canli et al. 2017b). With respect to the conceptual difficulties in landslide prediction, Greco and Pagano (2017) distinguish between three stages of a typical predictive system's architecture: I) the predisposing stage, II) the triggering and propagation stage, and III) the collapse stage. While in hydrological applications (II) and (III) are hardly distinguishable from each other, for rainfall-induced landslides this is not necessarily the case. While the predisposing stage (I) is determined by e.g. increasing pore water pressure due to a varying length of rainfall input that worsens the slope stability conditions, the triggering and propagation stage (II) spans from first local slope failures until the formation of associated slip surfaces. The collapse phase (III) ultimately consists of the mobilization of the entire mass leading to the actual failure. However, the time between stages (II) and (III) may vary significantly based on differences in local geomorphology, soil, vegetation, etc. and spans from a couple of minutes (e.g. flow slides in slopes covered with shallow coarse-grained soils) to years (e.g. earth flows in slopes of fine grained soils) (Greco and Pagano 2017). Thus, even the most accurate rainfall predictions might hold significant uncertainties with respect to predicting the spatial and temporal occurrence of landslides given the current approaches of landslide modeling and early warning.

This inability to precisely predict landslide occurrence has therefore consequences on the funding situation of landslide research in general and stands in a stark contrast to hydrological sciences. According to Baum and Godt (2010), losses from landslides are perceived mainly as private and localized economic losses with the result that only few public resources have been allocated to develop appropriate spatial landslide early warning systems. Among the main disaster events, hydrological and meteorological events rank among the costliest ones when comparing global and multi-peril loss databases, while geophysical events take only a small fraction in absolute numbers (Alfieri et al. 2012a, Wirtz et al. 2014). According to Petley (2012), landslide losses are vastly underestimated. Reasons for this observation are manifold: a) major disaster databases, e.g. the NatCatSERVICE from the reinsurance company Munich Re, associate landslides as subordinated hazard types of geophysical (amongst earthquakes) or hydrological hazards (amongst floods or avalanches) (Wirtz et al. 2014); b) landslide databases are inconsistent, incomplete or entirely absent and most of the existing inventories severely lack historical data (Wood et al. 2015, Herrera et al. 2017). As a consequence, much needed initiatives such as the *Hydrological Ensemble Prediction Experiment* (HEPEX) were not established in the landslide community so far. This ongoing bottom-up initiative aims at investigating on how to produce, communicate and use hydrologic ensemble forecasts in a multidisciplinary approach to make use of NWP in flood forecasting (Schaake et al. 2007). This superior position of hydrological forecasting can be primarily attributed to the greater interest international bodies demonstrated towards flood forecasting and thus, the resulting political and financial situation has led to the advancement of ensemble prediction systems in hydrology (Canli et al. 2017b). This is particularly the case with major transboundary flood events that are typically more severe in their consequences, affect larger areas and cause more damage and overall losses (Thielen et al. 2009).

However, transferring over knowledge and past experiences made in hydrological forecasting to the landslide community could significantly change the way how landslide prediction is approached in the near future. An automated landslide EPS framework could open up ways for finetuning input parameters by means of multiple model runs, attributing parameter uncertainties, and, first and foremost, real-time applications with a continuous consideration of antecedent and forecasted rainfall information (Alvioli and Baum 2016). Including measured real-time rainfall magnitudes derived from multiple sources (e.g. rain gauges, radar, satellite) could act as a means of data assimilation to further increase the accuracy of quantitative precipitation estimates and offer a real chance for a shift from the current *warn on detection* to a much needed *warn on forecast* paradigm in landslide early warning (Stensrud et al. 2009, Canli et al. 2017b).

4. Uncertainties in regional scale landslide prediction — a methodological approach

4.1. Uncertainties in rainfall information

Hourly rainfall data, based on rain gauge measurements that is published regularly on different web pages provided by different operators, is used as the basis for creating spatially distributed rainfall raster in near real-time. Hourly data is used to reflect the short-term rainfall intensity that can be considered as the main trigger for rainfall induced landslides in the study area (Lower Austria). Since there were no APIs available (an agreed-on programming interfaces for providing a structure to download and link data), a web scraping service was established. Web scraping mimics the human user interaction with a website by autonomously accessing it, parsing its content to find and extract relevant information and to save those for further use. This automated, time-based scheduling of obtaining hourly rainfall data provides a means of merging multiple data sources into a single database that can be used for storing raw rainfall data.

Multiple automated filters were applied to the raw data to ensure that there is as few errors and uncertainties as possible. This dissertation proposes three filters for quality assurance: a) a *range filter* to ensure physical plausibility; b) a *spatial consistency filter* to ensure there are no suspiciously high or low rain gauges based on the information of neighboring rain gauges within a certain distance; c) an *autocorrelation filter* specifically tailored towards the geostatistical interpolation approach to reduce biases from rain gauges that hold no relevant information and that would adversely influence variogram modeling. Figure 4.1 shows an overview of the automated workflow from obtaining web based rainfall data through the application of different filters to reduce errors and uncertainties in the raw data. The automated spatial interpolation methods applied in Canli et al. (2017a) are the deterministic Inverse Distance Weighting (IDW) method and the Thiessen polygon method, for the geostatistical interpolation an Ordinary Kriging approach was carried out (Fig. 4.1).

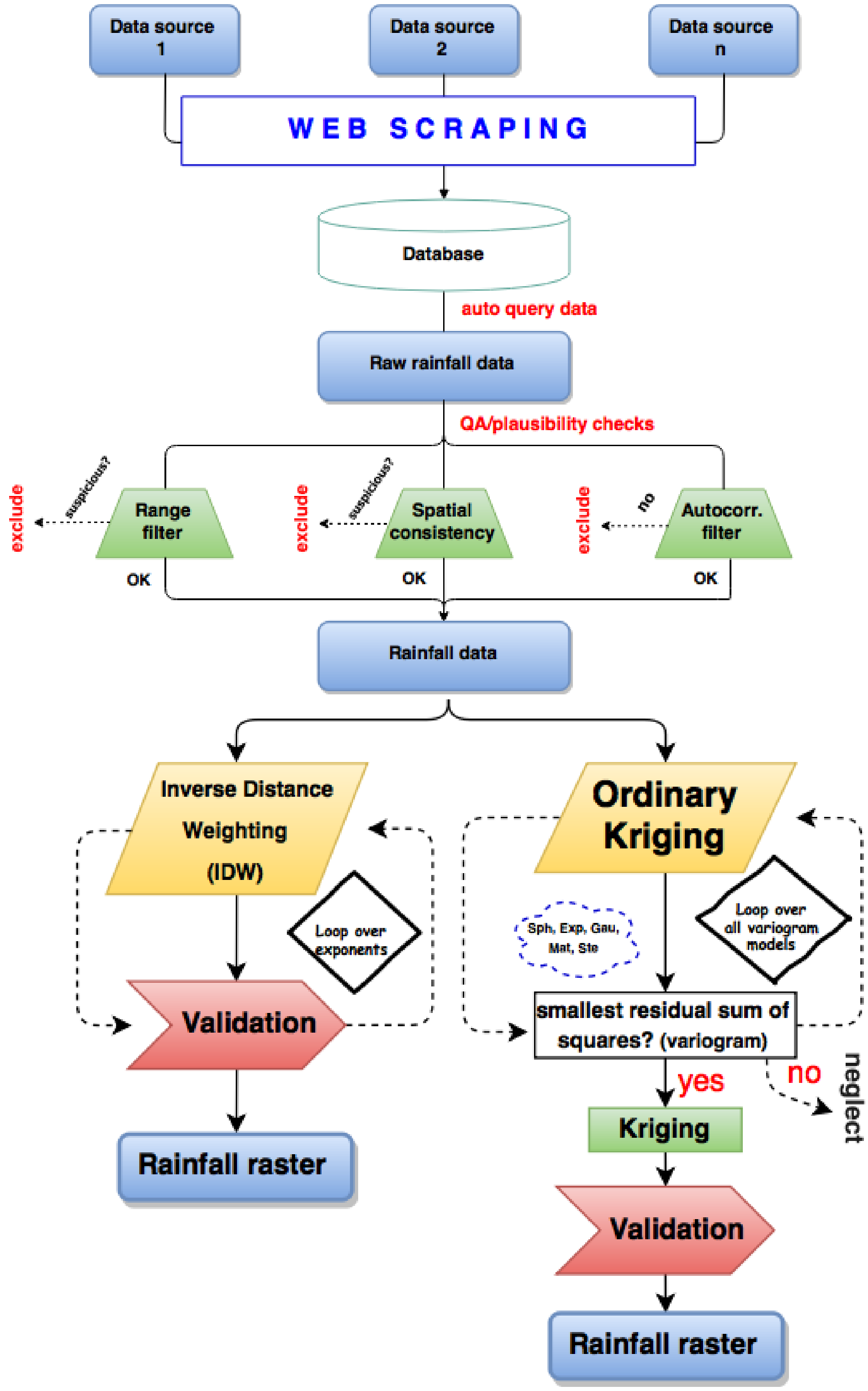


Figure 4.1: Flow chart of the proposed methodology showing the automated workflow from obtaining web based rainfall data, through multiple quality assurance filters, to the application of different interpolation techniques for producing hourly real-time rainfall raster maps (Canli et al. 2017a)

The big advantage of kriging is the consideration of variations in rainfall as a function of distance rather than distance alone in deterministic methods. This means that rain gauges that are in proximity to each other provide data values at unknown sampling locations in between that are quite similar rather than being reduced through the increasing distance from the respective rain gauges. In an automated iterative process, different variogram models are tested to find the best fitting one. A variogram is, in general, a plot of the average squared differences between data pair values and thus a central component in kriging. While creating a suitable variogram is usually a quite labor-intensive task, the proposed approach in this dissertation focusses on a rapid estimation of an appropriate variogram model for real-time applications. Therefore, some a priori decisions based on plausible initial values for the automated processing were made. Details on this automated procedure as well as the more straightforward deterministic interpolation methods are contained in more detail in Canli et al. (2017a). Based on this automated process chain, hourly rainfall was predicted at unsampled locations (on 1-km raster cells) to have a spatially distributed estimate of the most recent rainfall. Kriging additionally comes with the benefit, by being a statistical approach, of calculating the kriging variances. Those serve as an estimation error to reveal the interpolation certainty in a spatially distributed way. Performance comparison between sampling points (rain gauges) and distributed raster map was carried out with a) a leave-one-out cross-validation procedure, and b) by splitting the sampled points randomly into a training and test dataset.

4.2. Uncertainties in geotechnical and hydraulic parametrization at regional scale

Besides rainfall as the dynamic component in a regional rainfall triggered landslide early warning system, process based modelling approaches additionally require a physically based model parametrization. The most common approaches in model parametrization in process based landslide modeling encompass the application of averaged values from field measurements (e.g. Thiebes 2014, Tofani et al. 2017, Zieher et al. 2017) or the utilization of existing data from databases, lookup tables or other published/unpublished data sources (e.g. Schmidt et al. 2008, Kuriakose et al. 2009, Mergili et al. 2014b). Model parametrization at larger scales is by no means a trivial task, mainly due to the lacking spatial comprehension of the spatial organization of involved geotechnical and hydraulic input parameters (Fan et al. 2016). Also, field sampling at more or less representative locations over a large area might be highly biased (what qualifies as a representative location?) or even inappropriate for the modeling itself, as there is no guarantee that measured field values at a single location will reflect the effective values required by the model to achieve satisfactory predictions over a much larger area (Beven and Freer 2001).

Consequently, based on the premise that precise parameters over large areas are essentially unknown or highly uncertain at best, two assumptions were made in an honest attempt to address this lack of spatial comprehension: a) parameters taken from geotechnical literature are sufficient as those values are derived from on a multitude of repeated field and lab measurements and represent *typical* material properties; b) not a singular combination of parameters is used, but the entire parameter range from which parameter sets are randomly sampled in an attempt to cover the entire possible range of material properties.

Therefore, the proposed research suggests a probabilistic approach to derive model parameters based on purely literature based values. In a rather extensive study, Tofani et al. (2017) performed 59 site investigations to parametrize their slope stability model. This is a remarkably large amount of in situ sampled locations and offer a quite unique possibility to determine the underlying probability density function for all measured parameters. Albeit Tofani et al. (2017) reduce the information they use in their modeling attempt to just the median value for each lithological unit, their boxplots suggested normal to lognormal parameter distributions throughout all measured parameters. Wang et al. (2015) argue that this is a common observation and might be a result of the central limit theorem which indicates that lumping data from many different sampling sites tends to yield normal to lognormal distributions. Since the study area in Canli et al. (2017b) is rather large (over 1350 km²), plausible parameter ranges with a normally distributed state function based on geotechnical textbooks to characterize modeling parameters were used as it would be expected that taking many samples over such a large and rather homogenous area would result in quite comparable results. Instead of using a (supposedly) single best-fit value for each parameter (e.g. the median of a sampled value range), a Monte Carlo simulation approach was used to randomly choose multiple parameters sets within a predefined parameter range as the basis for incorporating the inherent parameter variability and uncertainties at larger scales into the model (Fig. 4.2). This way, the subsequent modeling approach is not initialized with a single *best-fit* set of parameters (as it is the case in a purely deterministic model), but run many times from slightly different initial conditions based on the sampled parameter range. This range of different solutions reveals the uncertainty in the model output and how confident we should be in a deterministic forecast (WMO 2012).

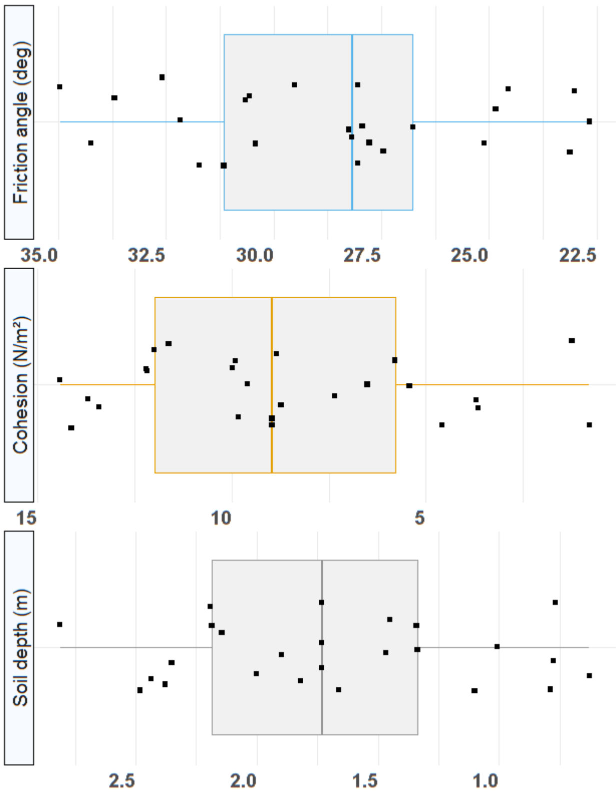


Figure 4.2: Probabilistically derived model parameters (Soil depth, Cohesion, Friction angle) based on random sampling from a normally distributed state function (Canli et al. 2017b)

4.3. Uncertainties in probabilistic modeling

Probabilistic predictions assess the probability that a certain outcome will occur and thus making them particularly desirable (Krzysztofowicz 2001). In the last decade, hydrological models have started integrating ensemble prediction systems (EPS) into their forecasting systems, following on the success of the use of ensembles for weather forecasting (Cloke and Pappenberger 2009). The probabilistic approach acknowledges the presence of unavoidable variability and uncertainty at larger scales and explicitly introduces them into the model results. EPS use ensembles of numerical weather predictions (NWP) to iteratively calculate, for example, a multitude of such probabilistically derived hydrographs for flood events (Cloke and Pappenberger 2009). This results in an expression of the entire model spread with its inherent uncertainties not in absolute terms, but it reveals the relative performance of a model based on different equally probable input parameters. This range of different solutions in the prediction allows for an assessment of uncertainty and how confident modelers and decision makers should be in a prediction (WMO 2012).

This dissertation proposes a fully automated landslide EPS based on different sets of input parameters that are randomly sampled from a broad range of possible parameter values based on geotechnical literature. Due to a lack of NWP data that was not available for this dissertation, spatially distributed rainfall input was considered from hourly geostatistical interpolation (as proposed in Canli et al. 2017a). Strictly speaking, this leaves the rainfall input in a deterministic state (only one rainfall raster per hour), while the probabilistic component is only added through variations in geotechnical and hydraulic parameters. However, the entire model structure is flexible enough to immediately replace the rainfall raster with a multitude of probabilistic NWP raster data sets. Parameters considered in a probabilistic way for the modeling application are soil depth, effective cohesion, effective friction angle and soil saturation (Canli et al. 2017b). As for the modeling itself, the open source, physically based TRIGRS model (transient rainfall infiltration and grid-based regional slope-stability analysis) was used (Baum et al 2008 and 2010). TRIGRS is a quite popular deterministic landslide model that is based on an infinite-slope model approach. Due to its popularity and flexibility, first attempts towards a probabilistic modification were made in the recent past (e.g. TRIGRS_P from Raia et al. (2014) or PG_TRIGRS from Salciarini et al. (2017)). None of those model, however, are operated in an automated way with the purpose of predicting landslides in real-time.

TRIGRS was developed with the aim of modeling the potential occurrences of shallow landslides by incorporating transient pressure response to rainfall and downward infiltration processes (Baum et al. 2008). However, by imposing simplifying assumptions and approximations, the underlying models of ground water flow and slope stability in TRIGRS are subject to limitations. Baum et al. (2008) mention, amongst others, the following restrictions:

- TRIGRS assumes flow in homogeneous, isotropic soil. Additionally, the slope stability model is based on an infinite-slope analysis, which assumes uniform slope, physical properties, thickness, and that pore water pressure is a function of depth and time alone. To reduce errors imposed by abruptly shifting topography and material properties, the study area in this dissertation is limited to a single geological unit (the Rhenodanubian Flyschzone) to keep the subsurface as homogeneous as possible.
- TRIGRS models only one-dimensional vertical infiltration although Baum et al. (2008) acknowledge that during longer storms or periods between storms, lateral flow contributes increasingly to the magnitude and distribution of pore water pressure. This dissertation follows the suggestion to set the initial water table at the ground surface and locking the steady background flux to zero to estimate a *worst-case* scenario, also due to a lack of appropriate initial water conditions.
- TRIGRS does not account for evapotranspiration, which might be quite substantial in the aftermath of a rainfall event. Also, only surface runoff is considered, but not horizontal subsurface flow. Consequently, tracking water conditions and its decay over time is not straightforward if carried out in an automated operational mode.

Being a deterministic model by default, TRIGRS computes a factor of safety (FoS) for each raster cell (10 m spatial resolution). Based on a set of equations, the FoS can be summarized as the ratio of resisting forces (the resisting basal Coulomb friction) and driving forces (the downslope basal driving stress) on the potential failure plane. A FoS ≥ 1.0 indicates stable slope conditions, a FoS < 1.0, on the other hand, slope instability. In the proposed research, TRIGRS was modified in an R and python programming language environment to modify the model to accept probabilistic input in an automated way. This way, multiple model iterations can be calculated (in this case 25), which results in as many equally probable model results based on the different input parameters. Each unstable cell (FoS < 1.0) from all model iterations is tracked and used to calculate the spatially distributed probability of a raster cell to fail. The result is an autonomously generated probability of failure (PoF) map that shows an indication of the most recent slope failure locations. The visualized ensemble spread (the variation in slope failure locations based on probabilistic parameter or rainfall input) gives an indication of the model's precision, and therefore how certain we can be about a prediction, even in the uncalibrated direct model output.

5. Discussion of results and hypotheses

Hypothesis I: Automated interpolation poses an improvement over selective rain gauge utilization for providing landslide early warning information.

For in-depth results and discussion, refer to Canli et al. 2016, Canli et al. 2017a.

This dissertation proposes a fully automated workflow from the hourly, web based collection of rain gauge data to the generation of spatially differentiated rainfall predictions based on deterministic and geostatistical methods. The ultimate goal is to utilize those products in both, threshold based approaches and dynamic physically based modeling approaches to substitute the prevalent practice of using single rain gauge information as a proxy for areal rainfall. The entire methodology proposed in Canli et al. (2017a) was executed purely on an open source basis to make it as easily reproducible as possible. To make use of multiple data sources and with the aim to densify the network of utilized rain gauges, web based hourly rain gauge data was obtained in an automated data workflow.

The results suggested that the Thiessen polygons do not offer any benefit over conventional approaches of selecting single rain gauges as a proxy for areal rainfall due to their arbitrary polygon boundaries that are being unrealistically rough. The IDW method could be a suitable method in case the rain gauge network is sufficiently dense. However, validation results suggested that automated spatial interpolation with kriging yielded the best fit with the available observational data. Additionally, the applied filters further improved the spatial rainfall prediction pattern which resulted in good spatial representations of current rainfall. However, the results also showed that the presence of small-scale, convective heavy rainfall events adversely affect variogram modeling to a rather high degree. This is unfortunate, as many landslide triggering rainfall events originate from such small-scale heavy rainfall events that are based on convection rather than prolonged frontal rainfall. Possible solutions and extension to this approach are the implementation of multivariate kriging methods that use additional predictor variables (such as radar data) or attempt a more probabilistic approach such as conditional simulation that additionally alleviates the smoothing effects of kriging by producing many equally likely scenarios rather than just a best fit scenario.

Therefore, it can be concluded that this dissertation provides a novel approach by applying automated spatial interpolation techniques for producing real-time spatial rainfall patterns from multiple web based sources. Validation results suggested a high spatial agreement with observational data and thus making this approach a possible alternative to purely utilizing single rain gauges as areal rainfall proxy. However, and with regard to the hypothesis, based on the analyses provided in this thesis, it cannot be answered whether the proposed methodology does indeed lead to improved early warning situations as no associated case study was conducted due to a lack of appropriate landslide event data. In order to evaluate whether there is any real benefit, a comparative study in a threshold based or physically based modeling setting is suggested. As a consequence, the results do not support the hypothesis as of yet and need to be further tested in a study area with appropriate landslide event data.

Hypothesis II: In situ measurements of geotechnical or hydraulic parameters can be substituted by literature based values for regional scale landslide model parametrization.

For in-depth results and discussion, refer to Canli et al. 2017b.

Poor spatial comprehension of the spatial organization of the involved geotechnical and hydraulic input parameters makes model parametrization at larger scales a difficult task. Park et al. (2013), Raia et al. (2014), Lee and Park (2016), Zhang et al. (2016) or Salciarini et al. (2017) treat soil parameters in regional scale landslides studies in a probabilistic way to address those inherent parameter uncertainties when it comes to model parametrization. In the probabilistic slope stability modeling approach proposed in this dissertation, each ensemble member was initialized with such probabilistically derived parameters. Results in Canli et al. (2017b) indicate quite significant changes in slope stability across individual members, but also quite high similarities although parameters change drastically between some of the members. For example, a depth of 2.5 m, an effective cohesion of 13.4 Nm^{-2} and an effective friction angle of 35 degree in a singular deterministic output reveals almost the identical spatial distribution of modeled slope failure as a model run with a depth of 2.0 m, an effective cohesion of 5.4 Nm^{-2} and an effective friction angle of 22.7 degree. By using a probabilistic representation that merges the information of all individual ensemble members into a combined representation of slope stability (the probability of failure), the entire range of spatial variability and uncertainty is explicitly introduced into the modeling results. Interestingly, the results of the probability of failure map suggest quite narrow ensemble spreads, which indicates that the different input parameter ranges result in quite similar individual outcomes. This means, that the predicted areas with the highest slope failure probability are consistently modelled more or less at the same locations. Differences in spatial occurrence can thus be considered as some kind of spatial confidence buffer that covers the entire range of used input parameters.

However, the fact that quite broad parameter ranges that are based on textbooks lead to quite similar spatial failure locations, indicates the paramount importance of slope angle as the most sensitive model parameter. This does not come surprising as slope failures are in general associated with higher slope angles (Liao et al. 2011). Also, Neves Seefelder et al. (2016) and Zieher et al. (2017) identified slope angle as the most sensitive modeling parameter in the same model (TRIGRS) as applied in this thesis. This would suggest that no matter how much the parameters within a plausible range vary, it will be consistently the same slope segments that will result in the highest slope failure probabilities. It also suggests that slope failure probability will ultimately only vary based on differences in the most recent spatially distributed dynamic components (e.g. rainfall or soil moisture distribution) or spatially differentiated slope depth maps. As discussed in more detail in Canli et al. (2017b), this raises the question whether model calibration is physically even advisable or if useful conclusions could be drawn from direct model output alone. As a possible explanation, which could also be shown in this dissertation, it can be argued that most models contain multiple combinations of parameter values that provide almost equally good fits to the observed data and that changing the calibration period or the goodness-of-fit-measure results in altered rankings of parameter sets to fit the observations. Consequently, as further pointed out by Beven (1996), there is no single parameter set (or model structure) that serves as the characteristic parameter input for any given area, but there is a certain degree of model equifinality involved when reproducing observations with model predictions. Therefore, given the issues with multiple (interacting) parameter values, measurement scales, spatial and temporal heterogeneity or the dependence on the model structure, there can never be a single set of parameter values for the calibration process that represents an optimum for the study area, but calibration can contribute to the reduction of range in the possible parameter space (Beven 1996, Neves Seefelder et al. 2016, Canli et al. 2017b).

With regard to the hypothesis it can be stated that it indeed appears to be the case that for physically based model applications at regional scale, purely literature based parameter ranges can substitute cost and labor intensive in situ field measurements due to the dominant sensitivity of slope angle and the high degree of equifinality. Thus, the carried-out analysis supports the hypothesis but it is suggested that a comparative study in an area with comprehensive geotechnical and hydraulic data is carried out to empirically affirm this adequacy.

Hypothesis III: A probabilistic landslide ensemble prediction system is capable of providing timely indication of high resolution landslide exposure at regional scale.

For in-depth results and discussion, refer to Canli et al. 2017b.

In ensemble predictions, small perturbations are made to the modeling parameters to be iteratively re-run with those slightly changed starting conditions. If those individual ensemble members are rather similar to each other (small ensemble spread), the prediction confidence is rather high. In case the ensemble spread is large or if they all develop differently, the confidence is much smaller (WMO 2012). As demonstrated for hypothesis II, using even quite large parameter ranges can indeed lead to rather narrow ensemble spreads in the model output as a result of equifinality and the dominant sensitivity of slope angle. To further supplement the resulting probability of failure (hazard) map with additional information to assist decision makers, this thesis suggests the combination with infrastructure data (buildings, roads) towards an exposure map to additionally account for possible consequences (Fig. 5.1).

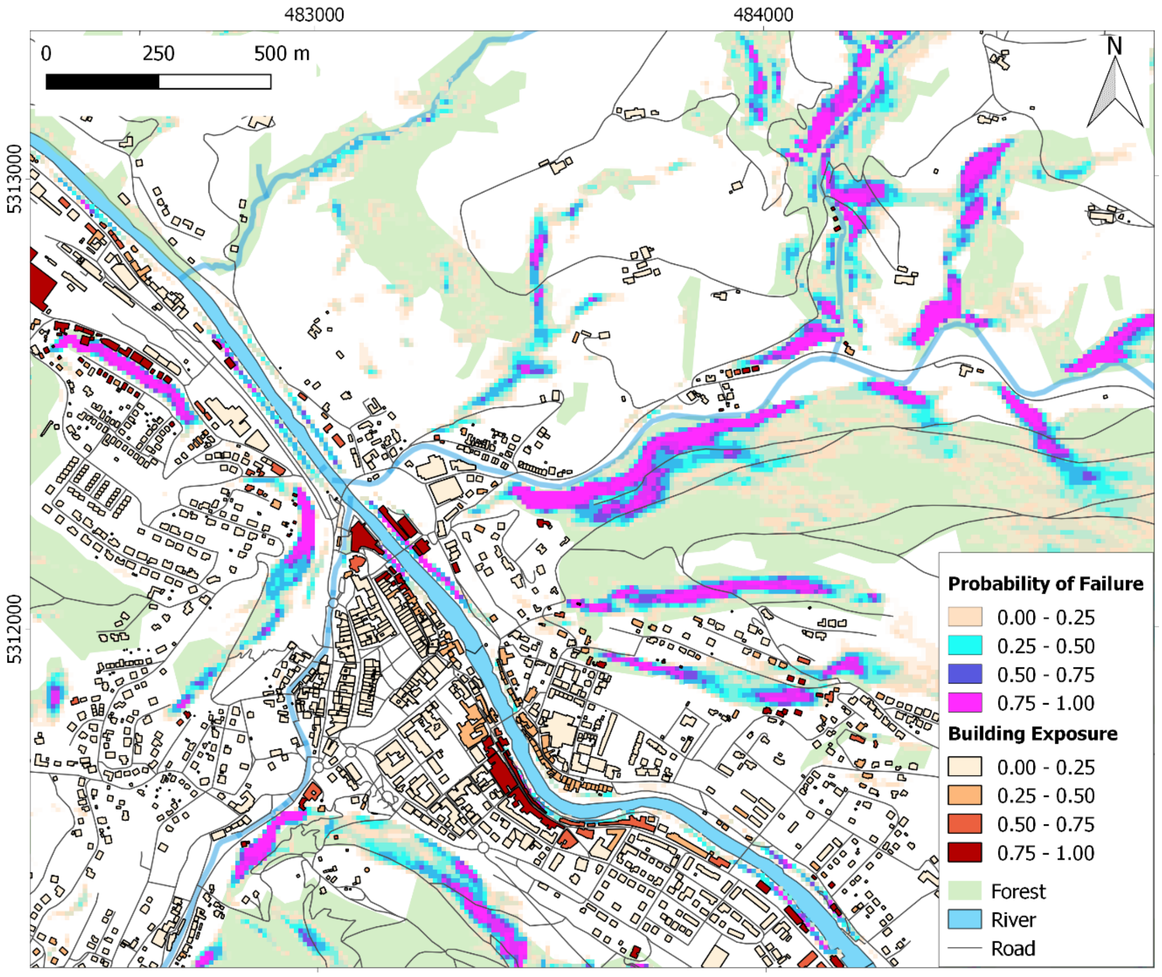


Figure 5.1: Probability of failure shown as a proportion of the individual ensemble members that predict an event to occur (FoS < 1.0). Building information was added as an additional layer to express an individual building's exposure to landslides. Building and roads are used from the freely accessible OpenStreeMap database (Canli et al. 2017b; Forest/River/Road data © OpenStreetMap contributors)

As soon as convective-scale numerical weather predictions are readily available for the implementation in such probabilistic landslide ensemble prediction systems to account for small-scale precipitation input, an operational mode is thinkable from a conceptual point of view. In reality, however, there are some drawbacks. Operating at small spatial scales or even at the scale of individual buildings, as proposed in this dissertation, could suggest a certainty in the modeling results that is simply not achievable. This seems quite disappointing and highly detrimental to what predictive models should be capable of providing: a positionally and temporally accurate mitigation tool. Compared to flood forecasting, where the spatial occurrence of floods is topographically foreseeable and controllable, this is much more difficult in spatial landslide modeling due to the very localized nature of landslide occurrence (Alfieri et al. 2012a). Salciarini et al. (2017) argues that such tools are suitable for a first susceptibility screening of an area prone to landsliding, but less so for single slope/single landslide analyses. As shown in this dissertation, the modelled probability of failure map revealed a high degree of spatial discontinuity in its spatial prediction pattern which undoubtedly puts a decision maker potentially at risk of missing some real landslide occurrences. A more in-depth discussion on the topic of model reliability in landslide prediction and the contribution of model calibration can be found in Canli et al. (2017b). Another drawback in operating landslide ensemble prediction systems in a timely manner is the computational burden involved. The computational time in this thesis to produce an hourly probability of failure map based on 25 individual ensemble members for the entire study area took around 18 hours. Ensembles of numerical weather prediction as dynamic rainfall input would additionally increase computation time quite significantly. Even if the code structure would be optimized to reduce computational time, this is far from what is acceptable in an operational mode. Newer developments in landslide modeling suggest a shift towards parallel computing in order to significantly cut down computation time (Formetta et al. 2016, Mergili et al. 2014a) or even the utilization of high-performance computing (HPC) clusters (Alvioli and Baum 2016).

To conclude this discussion on the feasibility of automated landslide ensemble prediction systems to explicitly introduce uncertainties from geotechnical parameters or from rainfall into the model output, it has to be clearly stated that this is still very much in its infancy. This thesis demonstrated that automated probabilistic landslide prediction is possible with a sufficiently small ensemble spread that indicates a rather high confidence in the spatial prediction pattern. All computational hindrances aside, it is inevitable to apply this model structure in a region with a comprehensive landslide event catalogue to evaluate whether such a high-resolution representation of landslide failure probability is capable of accurately predicting real landslide occurrences. Thus, this dissertation only partly supports the underlying hypothesis: it shows the technical feasibility of automated landslide ensemble predictions, yet it lacks meaningful and rigorous quantitative model validation due to a lack of appropriate landslide event data.

6. Conclusion and perspectives

Decision makers and practitioners in many earth science related fields prefer absolute model outputs. This is especially the case when public safety is at stake and clear thresholds need to be established for liability reasons. However, there is nothing such as an absolute certainty and relying on deterministic models provide an illusion of certainty at best due to a lack of full access to the phenomena of interest, both in time and space (Oreskes et al. 1994). Hence, probabilistic modeling provides an opportunity to increase the reliability and certainty of model outputs by expressing the entire model spread with its inherent uncertainties not in absolute terms, but by showing the relative performance of a model with respect to observational data (Canli et al. 2017b). Since probabilities in decision making are attributed with a lot of concerns, such probabilistic modeling results are not widely accepted yet, however according to Krzysztowicz (2001), this turned out to be unwarranted in a hydrological context. For communicating such probabilistic results, it could be beneficial to use judgmental terms given as a set of likelihood ranges (e.g. virtually certain >99%; very likely >90%; likely >66%; about as likely as not 33% to 66%; unlikely <33%; very unlikely <10%; extremely unlikely <5%; exceptionally unlikely <1%) to express the assessed probability of occurrence (Aven and Renn 2015). In the past decade, the Intergovernmental Panel on Climate Change (IPCC) has brought some tremendous research to light with respect to communicating uncertainty information to provide formal classifications for subjective and objective information (Risbey and Kandlikar 2007, Doyle et al. 2014, Wesselink et al. 2015, Aven and Renn 2015). Based on those IPCC key findings, Lee (2015) reviews attempts to provide a conceptual framework for communicating uncertainty and confidence to decision-makers in landslide risk assessment.

From a modeling point of view, validating deterministic models works on a best-fit realization by assessing the empirical adequacy of a singular model output with its associated observational data (Oreskes et al. 1994). Since this dissertation aims specifically towards the landslide modeling community, the focus lies on describing a potential application of the ROC curve for measuring probabilistic skill as it is by far the most commonly used measure of prediction skill in landslide research, both for rainfall threshold and statistical modeling applications (e.g. Frattini et al. 2010, Petschko et al. 2014, Gariano et al. 2015, Hussin et al. 2016, Steger et al. 2016a, Piciullo et al. 2017, Steger et al. 2017) and physically based ones (e.g. Chen and Zhang 2014, Mergili et al. 2014a, Raia et al. 2014, Formetta et al. 2016, Gioia et al. 2016, Lee and Park 2016). Contingency tables indicate the quality of a forecast system by considering its ability to anticipate correctly the occurrence or non-occurrence of predefined events that are expressed in binary terms, e.g. landslide occurred yes or no (Mason and Graham 1999). The contingency table, basically a two-by-two confusion matrix, holds four possible outcomes, given a certain classifier and instance: if the instance is positive and it is classified as positive, it is counted as a true positive (a hit); if it is classified as negative, it is counted as false negative (a false alarm). If the instance is negative and it is classified as negative, it is counted as true negative (a correct rejection); if it is classified as positive, it is counted as a false positive (a miss) (Fawcett 2006).

For deterministic forecasts, the ROC curve is generated by plotting the hit and false alarm rate for the forecast against the hit and false alarm rates obtained for perpetual warning (equals 1.0) and no-warning (equals 0.0). This means that there is skill only when the hit rate exceeds the false alarm rate. Thus, the ROC curve will ideally lie above the 45° line from the origin if the forecast system is skillful. The closer it is situated to 1.0, the more skillful it is (Mason and Graham 1999). The actual ROC score to compare classifiers can be expressed as the area under the ROC curve (AUC or AUROC), a single scalar value defined as a portion of the area of the unit square, hence creating values between 0 and 1 (Fawcett 2006).

For probabilistic forecasts, a warning can be issued in case the forecast probability for a predefined event exceeds some threshold. For example, if a warning should only be issued when there is at least a 75% confidence that a landslide event will occur (FoS < 1.0), a new contingency table that reflects the occurrences in areas exceeding a 75% probability is constructed. Different warning thresholds can be used for the predefined event, and a set of hit and false-alarm rates can then be determined (which is accordingly used to generate the ROC curve). Consequently, the ROC curve is useful in identifying an optimum warning criterion by indicating the trade-off between misses and false alarms. For a probabilistic system, the greatest value is not necessarily achieved at which the likelihood ratio is maximized. Instead, each decision maker evaluates possible consequences differently and/or has a different cost-loss operating structure, and hence the relative frequencies of hits, false alarms, and misses have to be optimized. In an operational environment, the warning is provided in advance, hence it is not known whether an event is going to occur, but if a warning has been issued. So, there is indeed additional value in knowing the probability of an event occurring, contingent upon the forecast probability (Mason and Graham 1999). Greco and Pagano (2017) suggest calibrating the sensitivity of an EWS based on a cost-benefit analysis that takes several peculiarities into account, such as the uncertainty of the prediction, the cost suffered by the community in case of a false alarm or the costs resulting from a missing alarm with catastrophic event occurrence.

Clearly, efforts put into the validation of probabilistic outcomes in the landslide modeling community are scarce and need substantially more research as of today. In hydrologic sciences on the other hand, some measures to validate probabilistic predictions are in practice. Some are better, some less suitable for distributed model output that is commonly the main form of data representation in landslide modeling and early warning. Thus, it remains to be seen, which skill scores are also feasible for validating probabilistic landslide model predictions. Mason and Graham (1999) and the WMO (2012) mention a few skill scores that are used in validating probabilistic hydrological predictions:

- Brier Score: a root-mean-square error for probability forecasts of a particular event threshold;
- Brier Skill Score: compares the Brier Score of the forecasts with the Brier Score of some reference forecast system;
- Reliability: measures how well forecast probabilities match observed frequencies;

- Receiver (or Relative) Operating Characteristics (ROC): measure how good the forecasts are for decision making
- Relative Operating Levels (ROL): designed to represent the skill of a forecast system from the perspective of the forecasts

The main drawback in this regard is most definitely – not just within this dissertation – the lack of appropriate event data. Complete landslide inventories at regional scale are rarely available – if at all. While statistical landslide susceptibility modelers are paying more and more attention to inventory biases (e.g. Hussin et al. 2016, Steger et al. 2016b, Steger et al. 2017), this does not apply as much to the landslide early warning community (be it the physically based modelers or the rainfall threshold community). On the one hand, this does seem natural since landslide locations serve as the dependent variable in statistical landslide susceptibility mapping that are explained by a set of static preparatory environmental factors (e.g. slope, lithology). Often, landslide inventories are specifically optimized for statistical or machine learning approaches (Petschko et al. 2015). For deterministic modeling approaches, however, the landslide inventory is independent from the modeling itself and serves as a means of validating and calibrating the model. Besides mapping biases (e.g. only reported landslides) or positional uncertainties from remote mapping, the most crucial aspect in dynamic applications is the temporal component. There are only very few event catalogues available that contain all the relevant information: precise location *and* precise timing of landslide initiation, which is crucial for model calibration (Gariano et al. 2015).

Calibration is usually referred to as the process of adjusting model parameters to represent the observation in the model output (landslide initiation at a specific location at a specific time). This implies, however, that the location and the time of landslide initiation is correct. Steger et al. (2016b) found that the only landslide inventory in Austria that occasionally contains temporal landslide information (the Building Ground Registry), exhibits substantial positional biases. The consequences for model calibration therefore are apparent: calibrating a deterministic model to represent optimized parameters at the landslide location might be incorrect, when the location of landslide occurrence is not precise. The same applies to the temporal component, when the landslide initiation time is not exactly known. A common practice to compile an event catalogue retroactively includes incorporating information from newspapers (e.g. Gariano et al. 2015). It seems obvious that this information can only be a rough estimate on where and when a landslide occurred and leads to the following question: how do positionally and temporally erroneous landslide catalogues influence deterministic model output when parameters are calibrated for imprecise landslide observations? Peres et al. (2017) performed such an in-depth analysis to quantify the effects of imprecise identification of triggering rainfall on the assessment and performance of landslide triggering thresholds, Nikolopoulos et al. (2014) analyzed the effect of rain gauge location and density of rainfall networks for the establishment of rainfall thresholds. Both studies concluded that the presence of reporting errors in landslide triggering instants yield thresholds that are significantly underestimated, i.e. lower than the correct ones. Consequently, ubiquitous errors in observed datasets generate further uncertainties in threshold assessment that is of significant magnitude (Peres et al. 2017). Since the landslide inventories with event-based information are the same for physically based approaches, this leads to the assumption that the same issues transfer over accordingly. Hence, probabilistic modeling approaches might alleviate some of those issues in the calibration process by exhibiting the probability of failure for a larger area that could potentially be affected by landsliding (by accommodating spatial uncertainties from larger parameter ranges). Consequently, the likelihood that an inaccurately mapped landslide lies in an area that was predicted to fail in some of the model ensemble members, is higher.

Another model related issue that needs to be considered in upcoming physically based modeling attempts in an early warning context is the application of data assimilation techniques. Data assimilation refers to the blending of multiple sources of dynamic information (for example rainfall or soil moisture data from different sources sampled at different scales) to increase the accuracy of the input data. This has been identified as an increasingly important factor for improving hydrological predictions (Reichle 2008). For dynamic landslide modeling applications, one of the most sensitive calibration parameters, that is usually not readily available at a larger scale, is the steady seepage initial condition. Water flow above the water table (in the unsaturated zone) is dependent on the downward rate of advance of the wetting front, which, in return, depends on antecedent soil moisture conditions. Hence, using a hydrological model with antecedent precipitation and infiltration rates from real-time monitoring could significantly improve slope stability analyses in the long run (Baum et al. 2010). This again gives some indication on the importance of data assimilation by blending multiple sources of information to increase the skill of physically based landslide predictions and to allow for better informed real-world decision making (Liu et al. 2012). Hydrological earth observation is on the verge of a breakthrough in delivering high resolution, accurate soil moisture input on very short time intervals for large regions (McCabe et al. 2017). This has huge potential to overcome the inherent scale incompatibility when using in situ field data that does not necessarily reflect the effective values required by the model itself. Blending those information, together with convection-permitting NWP, into a probabilistic slope stability model could have huge implications on how landslides might be accurately forecasted in the near future.

For actual decision making in landslide early warning situations, however, a combination of different modeling approaches could be beneficial. In the exposure approach presented in this dissertation, every region and every building is treated equally. Since statistical landslide susceptibility approaches have a very long tradition in landslide modeling, some very sophisticated methods have evolved over time that have the potential to supplement probabilistic modeling output. A high statistical likelihood of landslide occurrence means that in those areas it is more likely in the future that landslides occur again (based on available landslide information from the past). Based on a qualitative approach by matching the spatial agreement of statistical susceptibility maps (high spatial likelihood of future occurrence) with real-time probabilistic outcomes (high temporal likelihood of occurrence), this information could serve as an additional layer for decision-making (e.g. low susceptibility + high probability leads to a reduced warning level compared to a spatial match of high susceptibility + high probability). Attempts to combine statistical and physically based landslide susceptibility models were proposed in recent years (e.g. Goetz et al. 2011, Canli et al. 2015, Oliveira et al. 2017). A more simplistic approach could involve the designation of varying safety standards for different regions or objects. Again, for this exposure approach, hospitals or schools for example could be attributed with a higher safety standard that might require action to be taken at lower failure probabilities than a regular building. Similar systems for flood protection and management are in place in the Netherlands where economic analyses were used to differentiate safety standards for different regions (Pilarczyk 2007).

Additionally, and this has to be strongly emphasized, probabilistic forecasts temper the potential for misperception of responsibilities and misattribution of decisions. The task of forecasting, that incorporates solely the principles of science, and the task of decision making, which involves the decision maker's evaluation of consequences, is entirely decoupled (Krzysztowicz 2001). For example, instead of issuing a factor of safety map for a certain area that pinpoints a single estimate, the forecaster may specify a certain probability of failure to be exceeded based on the user's needs. The choice of protection level is thus left entirely to the decision maker, as it should be. There is a long history of discontinued operational landslide early warning systems which can partly be attributed to this mismatch of responsibilities (Baum and Godt 2010).

To conclude this dissertation, five future research topics are proposed:

- **Performance comparison of process based or threshold based landslide EWS with rainfall input from automated interpolation techniques versus uniform areal rainfall based on representative rain gauges.** This dissertation suggests an approach that is capable of predicting rainfall at unsampled locations in real-time from web based data sources. Validation between predicted and observed values resulted in quite satisfactory performance. However, due to a lack of event based landslide data, a performance comparison between the common application of uniform areal rainfall based on representative rain gauges versus the proposed approach of automated rainfall interpolation is pending.

- **Performance comparison of a regional scale process based landslide EWS with parameters derived from in situ sampling versus purely literature based parametrization.** The proposed research indicates an adequacy of purely literature based model parametrization at regional scale. Starting from the premise that selective in situ sampling at representative locations is not only biased, but also potentially unsuitable due to scale discrepancies between sampling resolution and model requirements, the entire parameter range for the respective geological unit was used. First approaches in that direction exist that suggest the application of rather broad parameter ranges (e.g. Neves Seefelder et al. 2016) by acknowledging that best-fit narrow ranges in geotechnical and hydraulic parameters might be off target. Results in this dissertation indeed show that the ensemble spread across all members is rather narrow, and thus suggesting a quite sharp prediction, but it also shows that the geotechnical/hydraulic parameter sensitivity is much lower than the sensitivity of the slope angle. Due to a lack of actual in situ field data from the study area as well as missing event based landslide data, a comparative study in an appropriate study area is suggested. This study should aim at evaluating the performance of literature based parametrization over parametrization with actual field data that was sampled in this specific study area.
- **Development of validation techniques for probabilistic model output in operational landslide early warning.** Model validation with an underlying contingency table considers the ability of a prediction system to correctly differentiate between occurrence or nonoccurrence of predefined events that are expressed in binary terms (e.g. landslide occurred yes/no). Using this contingency table in a ROC plot evaluates the skill of the prediction system. However, for probabilistic predictions systems, warnings can be issued for different predefined exceedance thresholds, and thus multiple contingency tables may exist (e.g. an own contingency table that reflects landslide occurrences in areas

exceeding a 75% probability). Therefore, multiple warning thresholds may be used for the predefined event which results in unique sets of hit and false-alarm rates. By doing so, an optimum warning criterion can be identified by indicating those trade-offs between misses and false alarms. The greatest value in validating such probabilistic systems does not necessarily come from maximizing this likelihood ratio, but by respecting a decision maker's needs. Possible consequences might be evaluated differently or decision makers use different cost-loss operating structures, which results in a requirement for optimized relative frequencies of hits, false alarms, and misses. Such a toolset for validating probabilistic model output based on a decision maker's requirements is missing in landslide research for the most part.

- **Performance comparison of a landslide ensemble prediction system versus common approaches of landslide early warning.** Ensemble prediction systems (EPS) have proven to be quite successful in flood forecasting. While empirical-statistical approaches (e.g. rainfall thresholds) only pose a simplification between the physical mechanisms leading to landslides and rainfall occurrence, process based deterministic approaches use mathematical expressions to represent relationships between elements. Instead of offering a deterministic best-fit model realization that entirely hides predictive uncertainties, probabilistic approaches do not eliminate uncertainty, but they explicitly introduce them into the model results. By considering the proportion of the individual ensemble members that predict slope failure, an estimate of how likely a landslide will occur can be made. This dissertation proposes a fully automated landslide EPS in an early prototype stage. However, due to a lack of numerical weather prediction data as well as event based landslide data, no performance comparison with more established means of landslide early warning could be made. Therefore, such a comparison is proposed in a study area where NWP data is available and that contains a comprehensive landslide event catalogue in order to perform rigorous model validation. Also, performing model code optimization and the application within a high performance computing facility is highly recommended to significantly reduce computation time.

- **Forming up larger interdisciplinary research initiatives for the advancement of landslide ensemble prediction systems.** Based out of a need to aid the World Climate Research Program's Global Water and Energy Cycle Experiment to meet their water-resource applications objectives, HEPEX was launched in March 2004 at a meeting of the European Centre for Medium Range Weather Forecasts (ECMWF). HEPEX stands for *Hydrological Ensemble Prediction Experiment* and is a project specifically designed by hydrologists, meteorologists, and users affiliated with several international organizations. This ongoing initiative aims towards the investigation on how to produce, communicate and use hydrologic ensemble forecasts. HEPEX is an open, participatory project not directly funded by any agency, but rather evolved from a bottom-up initiative by scientists and users who strongly believe that improved forecast techniques arise from interdisciplinary collaboration (Schaake et al. 2007). Although many probabilistic modelling approaches exist in landslide research to address parameter uncertainties and although it is acknowledged by some that such approaches could be the next step forward in landslide prediction (e.g. Alvioli and Baum 2016), no such impactful initiatives have been fostered within the landslide community. This dissertation poses only a small step in that direction, but further cooperation across disciplinary boundaries (e.g. with hydrologists, meteorologists, computer scientists) is envisaged to learn from forecasting experiences made in the last decade and to pursue research towards the improvement of such landslide ensemble prediction systems.

7. References

ALEOTTI P (2004) A warning system for rainfall-induced shallow failures. Engineering Geology 73:247–265. doi: 10.1016/j.enggeo.2004.01.007

ALFIERI L, SALAMON P, PAPPENBERGER F, ET AL (2012a) Operational early warning systems for water-related hazards in Europe. Environmental Science & Policy 21:35–49. doi: 10.1016/j.envsci.2012.01.008

ALFIERI L, THIELEN J, PAPPENBERGER F (2012b) Ensemble hydro-meteorological simulation for flash flood early detection in southern Switzerland. Journal of Hydrology 424–425:143–153. doi: 10.1016/j.jhydrol.2011.12.038

ALVIOLI M, BAUM RL (2016) Parallelization of the TRIGRS model for rainfall-induced landslides using the message passing interface. Environmental Modelling & Software 81:122–135. doi: 10.1016/j.envsoft.2016.04.002

AVEN T, RENN O (2015) An Evaluation of the Treatment of Risk and Uncertainties in the IPCC Reports on Climate Change: An Evaluation of the IPCC Reports on Climate Change. Risk Analysis 35:701–712. doi: 10.1111/risa.12298

BALDAUF M, SEIFERT A, FÖRSTNER J, ET AL (2011) Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities. Monthly Weather Review 139:3887–3905. doi: 10.1175/MWR-D-10-05013.1

BARTHOLMES J, TODINI E (2005) Coupling meteorological and hydrological models for flood forecasting. Hydrology and Earth System Sciences 9:333–346. doi: 10.5194/hess-9-333-2005

BAUM RL, SAVAGE WZ, GODT JW (2008) TRIGRS—A Fortran Program for Transient Rainfall Infiltration and Grid-Based Regional Slope-Stability Analysis, Version 2.0

BAUM RL, GODT JW, SAVAGE WZ (2010) Estimating the timing and location of shallow rainfall-induced landslides using a model for transient, unsaturated infiltration. Journal of Geophysical Research 115. doi: 10.1029/2009JF001321

BAUM RL, GODT JW (2010) Early warning of rainfall-induced shallow landslides and debris flows in the USA. Landslides 7:259–272. doi: 10.1007/s10346-009-0177-0

BELL R, PETSCHKO H, PROSKE H, ET AL (2014) Methodenentwicklung zur Gefährdungsmodellierung von Massenbewegungen in Niederösterreich – MoNOE. Final project report, Vienna, Austria

BEVEN K (1996) Equifinality and Uncertainty in Geomorphological Modelling. In: The Scientific Nature of Geomorphology: Proceedings of the 27th Binghamton Symposium in Geomorphology, Sept. 27-29, 1996. John Wiley & Sons, pp 289–313

BEVEN K (2001) How far can we go in distributed hydrological modelling? Hydrology and Earth System Sciences Discussions 5:1–12. doi: 10.5194/hess-5-1-2001

BEVEN K, FREER J (2001) Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. Journal of hydrology 249:11–29. doi: 10.1016/S0022-1694(01)00421-8

BOGAARD T, GRECO R (2018) Invited perspectives: Hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: proposing hydro-meteorological thresholds. Natural Hazards and Earth System Sciences 18:31–39. doi: 10.5194/nhess-18-31-2018

BROWN JD (2004) Knowledge, uncertainty and physical geography: towards the development of methodologies for questioning belief. Transactions of the Institute of British Geographers 29:367–381. doi: 10.1111/j.0020-2754.2004.00342.x

CANLI E, THIEBES B, PETSCHKO H, GLADE T (2015) Comparing physically-based and statistical landslide susceptibility model outputs - a case study from Lower Austria. EGU General Assembly Conference Abstracts 17. doi: 10.13140/RG.2.2.27952.87042

CANLI E, LOIGGE B, MERGILI M, GLADE T (2016) Generating web scraped high-quality weather databases for near-real-time derivation of spatial landslide susceptibility. In: Aversa S, Cascini L, Picarelli L, Scavia C (eds) Landslides and Engineered Slopes. Experience, Theory and Practice - Proceedings of the 12th International Symposium on Landslides, June 12-19, 2016. CRC Press, Napoli, Italy, pp 545–550

CANLI E, LOIGGE B, GLADE T (2017a) Spatially distributed rainfall information and its potential for regional landslide early warning systems. Natural Hazards. doi: 10.1007/s11069-017-2953-9

CANLI E, MERGILI M, GLADE T (2017b) Probabilistic landslide ensemble prediction systems: Lessons to be learned from hydrology. Natural Hazards and Earth System Sciences Discussions 1–30. doi: 10.5194/nhess-2017-427

CANLI E, ENGELS A, THIEBES B, ET AL (in prep.) Evaluating landslide dynamics and internal structure – a case study from the Salcher landslide observatory, Austria

CAPPARELLI G, TIRANTI D (2010) Application of the MoniFLaIR early warning system for rainfall-induced landslides in Piedmont region (Italy). Landslides 7:401–410. doi: 10.1007/s10346-009-0189-9

CHAE B-G, PARK H-J, CATANI F, ET AL (2017) Landslide prediction, monitoring and early warning: a concise review of state-of-the-art. Geosciences Journal 21:1033–1070. doi: 10.1007/s12303-017-0034-4

CHAPPELL A, RENZULLO LJ, RAUPACH TH, HAYLOCK M (2013) Evaluating geostatistical methods of blending satellite and gauge data to estimate near real-time daily rainfall for Australia. Journal of Hydrology 493:105–114. doi: 10.1016/j.jhydrol.2013.04.024

CHEN HX, ZHANG LM (2014) A physically-based distributed cell model for predicting regional rainfall-induced shallow slope failures. Engineering Geology 176:79–92. doi: 10.1016/j.enggeo.2014.04.011

CHIANG S-H, CHANG K-T (2009) Application of radar data to modeling rainfall-induced landslides. Geomorphology 103:299–309. doi: 10.1016/j.geomorph.2008.06.012

CHORLEY RJ, HAGGETT P (1967) Models in Geography. Methuen, London

CLOKE HL, PAPPENBERGER F (2009) Ensemble flood forecasting: A review. Journal of Hydrology 375:613–626. doi: 10.1016/j.jhydrol.2009.06.005

COROMINAS J, VAN WESTEN C, FRATTINI P, ET AL (2014) Recommendations for the quantitative analysis of landslide risk. Bulletin of Engineering Geology and the Environment 73:209–263. doi: 10.1007/s10064-013-0538-8

CROSTA GB, FRATTINI P (2003) Distributed modelling of shallow landslides triggered by intense rainfall. Natural Hazards and Earth System Science 3:81–93. doi: 10.5194/nhess-3-81-2003

CROZIER MJ (2010) Deciphering the effect of climate change on landslide activity: A review. Geomorphology 124:260–267. doi: 10.1016/j.geomorph.2010.04.009

DEMERITT D, WAINWRIGHT J (2005) Models, Modelling, and Geography. In: Castree N, Rogers A, Sherman DJ (eds) Questioning Geography - Fundamental Debates. Blackwell Pub, Malden, MA, pp 206–225

DIETRICH WE, MONTGOMERY DR (1998) SHALSTAB: a digital terrain model for mapping shallow landslide potential. NCASI (National Council of the Paper Industry for Air and Stream Improvement), Technical Report

DOYLE EEH, MCCLURE J, JOHNSTON DM, PATON D (2014) Communicating likelihoods and probabilities in forecasts of volcanic eruptions. Journal of Volcanology and Geothermal Research 272:1–15. doi: 10.1016/j.jvolgeores.2013.12.006

FAN L, LEHMANN P, OR D (2016) Effects of soil spatial variability at the hillslope and catchment scales on characteristics of rainfall-induced landslides. Water Resources Research 52:1781–1799. doi: 10.1002/2015WR017758

FAWCETT T (2006) An introduction to ROC analysis. Pattern Recognition Letters 27:861–874. doi: 10.1016/j.patrec.2005.10.010

FORMETTA G, CAPPARELLI G, VERSACE P (2016) Evaluating performance of simplified physically based models for shallow landslide susceptibility. Hydrology and Earth System Sciences 20:4585–4603. doi: 10.5194/hess-20-4585-2016

FRATTINI P, CROSTA GB, CARRARA A (2010) Techniques for evaluating the performance of landslide susceptibility models. Engineering Geology 111:62–72. doi: 10.1016/j.enggeo.2009.12.004

FRIGG R, THOMPSON E, WERNDL C (2015) Philosophy of climate science part I: observing climate change. Philosophy Compass 10:953–964. doi: 10.1111/phc3.12294

GARIANO SL, BRUNETTI MT, IOVINE G, ET AL (2015) Calibration and validation of rainfall thresholds for shallow landslide forecasting in Sicily, southern Italy. Geomorphology 228:653–665. doi: 10.1016/j.geomorph.2014.10.019

GARIANO SL, RIANNA G, PETRUCCI O, GUZZETTI F (2017) Assessing future changes in the occurrence of rainfall-induced landslides at a regional scale. Science of The Total Environment 596–597:417–426. doi: 10.1016/j.scitotenv.2017.03.103

GEBHARDT C, THEIS SE, PAULAT M, BEN BOUALLÈGUE Z (2011) Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbations and variation of lateral boundaries. Atmospheric Research 100:168–177. doi: 10.1016/j.atmosres.2010.12.008

GIOIA E, SPERANZA G, FERRETTI M, ET AL (2016) Application of a process-based shallow landslide hazard model over a broad area in Central Italy. Landslides 13:1197–1214. doi: 10.1007/s10346-015-0670-6

GLADE T, NADIM F (2014) Early warning systems for natural hazards and risks. Natural Hazards 70:1669–1671. doi: 10.1007/s11069-013-1000-8

GODT JW, BAUM RL, CHLEBORAD AF (2006) Rainfall characteristics for shallow landsliding in Seattle, Washington, USA. Earth Surface Processes and Landforms 31:97–110. doi: 10.1002/esp.1237

GOETZ JN, GUTHRIE RH, BRENNING A (2011) Integrating physical and empirical landslide susceptibility models using generalized additive models. Geomorphology 129:376–386. doi: 10.1016/j.geomorph.2011.03.001

GOETZ JN, BRENNING A, PETSCHKO H, LEOPOLD P (2015) Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Computers & Geosciences 81:1–11. doi: 10.1016/j.cageo.2015.04.007

GOLDING B, ROBERTS N, LEONCINI G, ET AL (2016) MOGREPS-UK Convection-Permitting Ensemble Products for Surface Water Flood Forecasting: Rationale and First Results. Journal of Hydrometeorology 17:1383–1406. doi: 10.1175/JHM-D-15-0083.1

GOOVAERTS P (1997) Geostatistics for natural resources evaluation. Oxford University Press, New York

GOOVAERTS P (2000) Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology 228:113–129. doi: 10.1016/S0022-1694(00)00144-X

GRECO R, PAGANO L (2017) Basic features of the predictive tools of early warning systems for water-related natural hazards: examples for shallow landslides. Natural Hazards and Earth System Sciences 17:2213–2227. doi: 10.5194/nhess-17-2213-2017

GUZZETTI F (2005) Landslide Hazard and Risk Assessment. Dissertation, University Bonn

GUZZETTI F, PERUCCACCI S, ROSSI M, STARK CP (2007) Rainfall thresholds for the initiation of landslides in central and southern Europe. Meteorology and Atmospheric Physics 98:239–267. doi: 10.1007/s00703-007-0262-7

HABERLANDT U (2007) Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. Journal of Hydrology 332:144–157. doi: 10.1016/j.jhydrol.2006.06.028

HALL AD, FAGAN RE (1956) Definition of systems. In: General Systems. pp 18–28

HANEBERG WC (2004) A rational probabilistic method for spatially distributed landslide hazard assessment. Environmental & Engineering Geoscience 10:27–43. doi: 10.2113/10.1.27

HARPOLD AA, KAPLAN ML, KLOS PZ, ET AL (2017) Rain or snow: hydrologic processes, observations, prediction, and research needs. Hydrology and Earth System Sciences 21:1–22. doi: 10.5194/hess-21-1-2017

HERRERA G, MATEOS RM, GARCÍA-DAVALILLO JC, ET AL (2017) Landslide databases in the Geological Surveys of Europe. Landslides. doi: 10.1007/s10346-017-0902-z

HORTON P, JABOYEDOFF M, RUDAZ B, ZIMMERMANN M (2013) Flow-R, a model for susceptibility mapping of debris flows and other gravitational hazards at a regional scale. Natural Hazards and Earth System Science 13:869–885. doi: 10.5194/nhess-13-869-2013

HOU AY, KAKAR RK, NEECK S, ET AL (2014) The Global Precipitation Measurement Mission. Bulletin of the American Meteorological Society 95:701–722. doi: 10.1175/BAMS-D-13-00164.1

HUANG J, JU NP, LIAO YJ, LIU DD (2015) Determination of rainfall thresholds for shallow landslides by a probabilistic and empirical method. Natural Hazards and Earth System Sciences 15:2715–2723. doi: 10.5194/nhess-15-2715-2015

HUFFMAN GJ, BOLVIN DT, NELKIN EJ, ET Al (2007) The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. Journal of Hydrometeorology 8:38–55. doi: 10.1175/JHM560.1

HUFFMAN GJ, ADLER RF, BOLVIN DT, NELKIN EJ (2010) The TRMM Multi-Satellite Precipitation Analysis (TMPA). In: Gebremichael M, Hossain F (eds) Satellite Rainfall Applications for Surface Hydrology. Springer Netherlands, Dordrecht, pp 3–22

HUNGR O, LEROUEIL S, PICARELLI L (2014) The Varnes classification of landslide types, an update. Landslides 11:167–194. doi: 10.1007/s10346-013-0436-y

HUSSIN HY, CIUREAN R, FRIGERIO S, ET AL (2014) Assessing the Effect of Mitigation Measures on Landslide Hazard Using 2D Numerical Runout Modelling. In: Sassa K, Canuti P, Yin Y (eds) Landslide Science for a Safer Geoenvironment. Springer International Publishing, Cham, pp 679–684

HUSSIN HY, ZUMPANO V, REICHENBACH P, ET AL (2016) Different landslide sampling strategies in a grid-based bi-variate statistical susceptibility model. Geomorphology 253:508–523. doi: 10.1016/j.geomorph.2015.10.030

IKEDA K, STEINER M, PINTO J, ALEXANDER C (2013) Evaluation of Cold-Season Precipitation Forecasts Generated by the Hourly Updating High-Resolution Rapid Refresh Model. Weather and Forecasting 28:921–939. doi: 10.1175/WAF-D-12-00085.1

INKPEN R (2005) Science, Philosophy and Physical Geography, First edition. Routledge, London, New York

INKPEN R, WILSON G (2013) Science, Philosophy and Physical Geography, Second edition. Routledge, London, New York

JAEDICKE C, VAN DEN EECKHAUT M, NADIM F, ET AL (2014) Identification of landslide hazard and risk 'hotspots' in Europe. Bulletin of Engineering Geology and the Environment 73:325–339. doi: 10.1007/s10064-013-0541-0

JASPER K, GURTZ J, LANG H (2002) Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model. Journal of Hydrology 267:40–52. doi: 10.1016/S0022-1694(02)00138-5

JOYCE RJ, JANOWIAK JE, ARKIN PA, XIE P (2004) CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. Journal of Hydrometeorology 5:487–503. doi: 10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

KEILER M (2011) Geomorphology and Complexity – inseparably connected? Zeitschrift für Geomorphologie, Supplementary Issues 55:233–257. doi: 10.1127/0372-8854/2011/0055S3-0060

KIRSCHBAUM D, STANLEY T, SIMMONS J (2015) A dynamic landslide hazard assessment system for Central America and Hispaniola. Natural Hazards and Earth System Science 15:2257–2272. doi: 10.5194/nhess-15-2257-2015

KIRSCHBAUM D, PATEL K (2016) Precipitation Data Key to Food Security and Public Health. Eos 97. doi: 10.1029/2016EO045583

KLEMEŠ V (1986) Operational testing of hydrological simulation models. Hydrological Sciences Journal 31:13–24. doi: 10.1080/02626668609491024

KNOX JC (1972) Valley Alluviation in Southwestern Wisconsin. Annals of the Association of American Geographers 62:401–410. doi: 10.1111/j.1467-8306.1972.tb00872.x

Krzysztofowicz R (2001) The case for probabilistic forecasting in hydrology. Journal of Hydrology 249:2–9. doi: 10.1016/S0022-1694(01)00420-6

Kubota T, Shige S, Hashizume H, et al (2007) Global Precipitation Map Using Satellite-Borne Microwave Radiometers by the GSMaP Project: Production and Validation. IEEE Transactions on Geoscience and Remote Sensing 45:2259–2275. doi: 10.1109/TGRS.2007.895337

Kuriakose SL, van Beek LPH, van Westen CJ (2009) Parameterizing a physically based shallow landslide model in a data poor region. Earth Surface Processes and Landforms 34:867–881. doi: 10.1002/esp.1794

Lagomarsino D, Segoni S, Fanti R, Catani F (2013) Updating and tuning a regional-scale landslide early warning system. Landslides 10:91–97. doi: 10.1007/s10346-012-0376-y

Lari S, Frattini P, Crosta GB (2014) A probabilistic approach for landslide hazard analysis. Engineering Geology 182:3–14. doi: 10.1016/j.enggeo.2014.07.015

Lee JH, Park HJ (2016) Assessment of shallow landslide susceptibility using the transient infiltration flow model and GIS-based probabilistic approach. Landslides 13:885–903. doi: 10.1007/s10346-015-0646-6

Lehmann J, Rillig M (2014) Distinguishing variability from uncertainty. Nature Climate Change 4:153–153. doi: 10.1038/nclimate2133

Liao Z, Hong Y, Kirschbaum D, et al (2011) Evaluation of TRIGRS (transient rainfall infiltration and grid-based regional slope-stability analysis)'s predictive skill for hurricane-triggered landslides: a case study in Macon County, North Carolina. Natural Hazards 58:325–339. doi: 10.1007/s11069-010-9670-y

Liu Y, Weerts AH, Clark M, et al (2012) Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. Hydrology and Earth System Sciences 16:3863–3887. doi: 10.5194/hess-16-3863-2012

Ly S, Charles C, Degré A (2013) Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: a review. Base 17:392–406

Mair A, Fares A (2011) Comparison of Rainfall Interpolation Methods in a Mountainous Region of a Tropical Island. Journal of Hydrologic Engineering 16:371–383. doi: 10.1061/(ASCE)HE.1943-5584.0000330

Mason SJ, Graham NE (1999) Conditional probabilities, relative operating characteristics, and relative operating levels. Weather and Forecasting 14:713–725. doi: 10.1175/1520-0434(1999)014<0713:CPROCA>2.0.CO;2

McCabe MF, Rodell M, Alsdorf DE, et al (2017) The future of Earth observation in hydrology. Hydrology and Earth System Sciences 21:3879–3914. doi: 10.5194/hess-21-3879-2017

McDOUGALL S (2017) 2014 Canadian Geotechnical Colloquium: Landslide runout analysis — current practice and challenges. Canadian Geotechnical Journal 54:605–620. doi: 10.1139/cgj-2016-0104

MERGILI M, MARCHESINI I, ALVIOLI M, ET AL (2014a) A strategy for GIS-based 3-D slope stability modelling over large areas. Geoscientific Model Development 7:2969–2982. doi: 10.5194/gmd-7-2969-2014

MERGILI M, MARCHESINI I, ROSSI M, ET AL (2014b) Spatially distributed three-dimensional slope stability modelling in a raster GIS. Geomorphology 206:178–195. doi: 10.1016/j.geomorph.2013.10.008

MICHELETTI N, FORESTI L, ROBERT S, ET AL (2014) Machine Learning Feature Selection Methods for Landslide Susceptibility Mapping. Mathematical Geosciences 46:33–57. doi: 10.1007/s11004-013-9511-0

MURRAY B, FONSTAD MA (2007) Preface: Complexity (and simplicity) in landscapes. Geomorphology 91:173–177. doi: 10.1016/j.geomorph.2007.07.011

NEVES SEEFELDER C DE L, KOIDE S, MERGILI M (2016) Does parameterization influence the performance of slope stability model results? A case study in Rio de Janeiro, Brazil. Landslides. doi: 10.1007/s10346-016-0783-6

NIKOLOPOULOS EI, CREMA S, MARCHI L, ET AL (2014) Impact of uncertainty in rainfall estimation on the identification of rainfall thresholds for debris flow occurrence. Geomorphology 221:286–297. doi: 10.1016/j.geomorph.2014.06.015

OLIVEIRA SC, ZÊZERE JL, LAJAS S, MELO R (2017) Combination of statistical and physically based methods to assess shallow slide susceptibility at the basin scale. Natural Hazards and Earth System Sciences 17:1091–1109. doi: 10.5194/nhess-17-1091-2017

OLIVER MA, WEBSTER R (2014) A tutorial guide to geostatistics: Computing and modelling variograms and kriging. CATENA 113:56–69. doi: 10.1016/j.catena.2013.09.006

ORESKES N, SHRADER-FRECHETTE K, BELITZ K, OTHERS (1994) Verification, validation, and confirmation of numerical models in the earth sciences. Science 263:641–646. doi: 10.1126/science.263.5147.641

PACK RT, TARBOTON DG, GOODWIN CN (1998) The SINMAP approach to terrain stability mapping. In: 8th Congress of the International Association of Engineering Geology, Vancouver, British Columbia, Canada. p 25

PAPATHOMA-KÖHLE M, GLADE T (2013) The role of vegetation cover change in landslide hazard and risk. In: Renaud FG, Sudmeier-Rieux K, Estrella M (eds) The role of ecosystems in disaster risk reduction. United Nations University Press, Tokyo, New York, Paris, pp 293–320

PARK HJ, LEE JH, WOO I (2013) Assessment of rainfall-induced shallow landslide susceptibility using a GIS-based probabilistic approach. Engineering Geology 161:1–15. doi: 10.1016/j.enggeo.2013.04.011

PERES DJ, CANCELLIERE A, GRECO R, BOGAARD TA (2017) Influence of uncertain identification of triggering rainfall on the assessment of landslide early warning thresholds. Natural Hazards and Earth System Sciences Discussions 1–28. doi: 10.5194/nhess-2017-328

PETLEY D (2012) Global patterns of loss of life from landslides. Geology 40:927–930. doi: 10.1130/G33217.1

PETSCHKO H, BRENNING A, BELL R, ET AL (2014) Assessing the quality of landslide susceptibility maps – case study Lower Austria. Natural Hazards and Earth System Science 14:95–118. doi: 10.5194/nhess-14-95-2014

PETSCHKO H, BELL R, GLADE T (2015) Effectiveness of visually analyzing LiDAR DTM derivatives for earth and debris slide inventory mapping for statistical susceptibility modeling. Landslides 65. doi: 10.1007/s10346-015-0622-1

PHAM BT, PRADHAN B, TIEN BUI D, ET AL (2016) A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India). Environmental Modelling & Software 84:240–250. doi: 10.1016/j.envsoft.2016.07.005

PHILLIPS JD (2015) Badass geomorphology. Earth Surface Processes and Landforms 40:22–33. doi: 10.1002/esp.3682

PICIULLO L, DAHL M-P, DEVOLI G, ET AL (2017) Adapting the EDuMaP method to test the performance of the Norwegian early warning system for weather-induced landslides. Natural Hazards and Earth System Sciences 17:817–831. doi: 10.5194/nhess-17-817-2017

PILARCZYK KW (2007) Flood Protection and Management in the Netherlands. In: Vasiliev OF, van Gelder P, Plate EJ, Bolgov MV (eds) Extreme Hydrological Events: New Concepts for Security. Springer Netherlands, Dordrecht, pp 385–407

RAIA S, ALVIOLI M, ROSSI M, ET AL (2014) Improving predictive power of physically based rainfall-induced shallow landslide models: a probabilistic approach. Geoscientific Model Development 7:495–514. doi: 10.5194/gmd-7-495-2014

REICHENBACH P, CARDINALI M, DE VITA P, GUZZETTI F (1998) Regional hydrological thresholds for landslides and floods in the Tiber River Basin (central Italy). Environmental Geology 35:146–159. doi: 10.1007/s002540050301

REICHLE RH (2008) Data assimilation methods in the Earth sciences. Advances in Water Resources 31:1411–1418. doi: 10.1016/j.advwatres.2008.01.001

RISBEY JS, KANDLIKAR M (2007) Expressions of likelihood and confidence in the IPCC uncertainty assessment process. Climatic Change 85:19–31. doi: 10.1007/s10584-007-9315-7

ROSI A, LAGOMARSINO D, ROSSI G, ET AL (2015) Updating EWS rainfall thresholds for the triggering of landslides. Natural Hazards 78:297–308. doi: 10.1007/s11069-015-1717-7

ROSSI M, KIRSCHBAUM D, LUCIANI S, ET AL (2012) TRMM satellite rainfall estimates for landslide early warning in Italy: preliminary results. In: Hayasaka T, Nakamura K, Im E (eds) Proc. SPIE 8523. Kyoto, Japan, p 85230D

ROSSI M, KIRSCHBAUM D, VALIGI D, ET AL (2017) Comparison of Satellite Rainfall Estimates and Rain Gauge Measurements in Italy, and Impact on Landslide Modeling. Climate 5:90. doi: 10.3390/cli5040090

SAITO H, NAKAYAMA D, MATSUYAMA H (2009) Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: The Akaishi Mountains, Japan. Geomorphology 109:108–121. doi: 10.1016/j.geomorph.2009.02.026

SALCIARINI D, FANELLI G, TAMAGNINI C (2017) A probabilistic model for rainfall—induced shallow landslide prediction at the regional scale. Landslides. doi: 10.1007/s10346-017-0812-0

SCHAAKE JC, HAMILL TM, BUIZZA R, CLARK M (2007) HEPEX: The Hydrological Ensemble Prediction Experiment. Bulletin of the American Meteorological Society 88:1541–1547. doi: 10.1175/BAMS-88-10-1541

SCHMIDT J, TUREK G, CLARK MP, ET AL (2008) Probabilistic forecasting of shallow, rainfall-triggered landslides using real-time numerical weather predictions. Natural Hazards and Earth System Science 8:349–357. doi: 10.5194/nhess-8-349-2008

SCHUMM SA (1973) Geomorphic thresholds and complex response of drainage systems. In: Morisawa M (ed) Fluvial Geomorphology. SUNY Binghamton Publ. in Geomorphology. Colorado State University, Fort Collins, CO, pp 299–310

SCHUURMANS JM, BIERKENS MFP, PEBESMA EJ, UIJLENHOET R (2007) Automatic Prediction of High-Resolution Daily Rainfall Fields for Multiple Extents: The Potential of Operational Radar. Journal of Hydrometeorology 8:1204–1224. doi: 10.1175/2007JHM792.1

SEGOND M-L, WHEATER HS, ONOF C (2007) The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the Lee catchment, UK. Journal of Hydrology 347:116–131. doi: 10.1016/j.jhydrol.2007.09.040

SEGONI S, LEONI L, BENEDETTI AI, ET AL (2009) Towards a definition of a real-time forecasting network for rainfall induced shallow landslides. Natural Hazards and Earth System Science 9:2119–2133. doi: 10.5194/nhess-9-2119-2009

SEGONI S, LAGOMARSINO D, FANTI R, ET AL (2015) Integration of rainfall thresholds and susceptibility maps in the Emilia Romagna (Italy) regional-scale landslide warning system. Landslides 12:773–785. doi: 10.1007/s10346-014-0502-0

SEITY Y, BROUSSEAU P, MALARDEL S, ET AL (2011) The AROME-France Convective-Scale Operational Model. Monthly Weather Review 139:976–991. doi: 10.1175/2010MWR3425.1

SEMENOVA O, BEVEN K (2015) Barriers to progress in distributed hydrological modelling: Invited Commentary. Hydrological Processes 29:2074–2078. doi: 10.1002/hyp.10434

SEYFRIED MS, WILCOX BP (1995) Scale and the Nature of Spatial Variability: Field Examples Having Implications for Hydrologic Modeling. Water Resources Research 31:173–184. doi: 10.1029/94WR02025

SICCARDI F, BONI G, FERRARIS L, RUDARI R (2005) A hydrometeorological approach for probabilistic flood forecast. Journal of Geophysical Research 110. doi: 10.1029/2004JD005314

SRIVASTAVA RM (2013) Geostatistics: A toolkit for data analysis, spatial prediction and risk management in the coal industry. International Journal of Coal Geology 112:2–13. doi: 10.1016/j.coal.2013.01.011

STÄHLI M, SÄTTELE M, HUGGEL C, ET AL (2015) Monitoring and prediction in early warning systems for rapid mass movements. Natural Hazards and Earth System Science 15:905–917. doi: 10.5194/nhess-15-905-2015

STANLEY T, KIRSCHBAUM DB, HUFFMAN GJ, ADLER RF (2017) Approximating Long-Term Statistics Early in the Global Precipitation Measurement Era. Earth Interactions 21:1–10. doi: 10.1175/EI-D-16-0025.1

STEGER S, BRENNING A, BELL R, ET AL (2016a) Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps. Geomorphology 262:8–23. doi: 10.1016/j.geomorph.2016.03.015

STEGER S, BRENNING A, BELL R, GLADE T (2016b) The propagation of inventory-based positional errors into statistical landslide susceptibility models. Natural Hazards and Earth System Sciences 16:2729–2745. doi: 10.5194/nhess-16-2729-2016

STEGER S, BRENNING A, BELL R, GLADE T (2017) The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. Landslides. doi: 10.1007/s10346-017-0820-0

STENSRUD DJ, WICKER LJ, KELLEHER KE, ET AL (2009) Convective-Scale Warn-on-Forecast System: A Vision for 2020. Bulletin of the American Meteorological Society 90:1487–1499. doi: 10.1175/2009BAMS2795.1

STRAND S-A, THAKUR V, L'HEUREUX J-S, ET AL (2017) Runout of Landslides in Sensitive Clays. In: Thakur V, L'Heureux J-S, Locat A (eds) Landslides in Sensitive Clays. Springer International Publishing, Cham, pp 289–300

TEMME A, KEILER M, KARSSENBERG D, LANG A (2015) Complexity and non-linearity in earth surface processes - concepts, methods and applications. Earth Surface Processes and Landforms 40:1270–1274. doi: 10.1002/esp.3712

THIEBES B, BELL R, GLADE T, ET AL (2013) A WebGIS decision-support system for slope stability based on limit-equilibrium modelling. Engineering Geology 158:109–118. doi: 10.1016/j.enggeo.2013.03.004

THIEBES B, BELL R, GLADE T, ET AL (2014) Integration of a limit-equilibrium model into a landslide early warning system. Landslides 11:859–875. doi: 10.1007/s10346-013-0416-2

THIEBES B, GLADE T (2016) Landslide Early Warning Systems—fundamental concepts and innovative applications. In: Aversa S, Cascini L, Picarelli L, Scavia C (eds) Landslides and Engineered Slopes. Experience, Theory and Practice - Proceedings of the 12th International Symposium on Landslides, June 12-19, 2016. CRC Press, Napoli, pp 1903–1911

THIELEN J, BARTHOLMES J, RAMOS M-H, ROO A DE (2009) The European flood alert system–Part 1: concept and development. Hydrology and Earth System Sciences 13:125–140. doi: 10.5194/hess-13-125-2009

THOMAS D, ALLISON RJ (eds) (1993) Landscape sensitivity. Wiley, Chichester, New York

THORN CE (1988) An Introduction to Theoretical Geomorphology. Springer Netherlands, Dordrecht

TOFANI V, BICOCCHI G, ROSSI G, ET AL (2017) Soil characterization for shallow landslides modeling: a case study in the Northern Apennines (Central Italy). Landslides. doi: 10.1007/s10346-017-0809-8

UNEP (2012) Early Warning Systems: A State of the Art Analysis and Future Directions. Division of Early Warning and Assessment (DEWA), United Nations Environment Programme (UNEP), Nairobi, Kenia

VAN BEEK R (2002) Assessment of the influence of changes in land use and climate on landslide activity in a Mediterranean environment. Dissertation, University Utrecht

VAN WESTEN CJ, VAN ASCH T, SOETERS R (2006) Landslide hazard and risk zonation—why is it still so difficult? Bulletin of Engineering Geology and the Environment 65:167–184. doi: 10.1007/s10064-005-0023-0

VAN WESTEN CJ, JAISWAL P, GHOSH S, ET AL (2012) Landslide Inventory, Hazard and Risk Assessment in India. In: Pradhan B, Buchroithner M (eds) Terrigenous Mass Movements. Springer, pp 239–282

VARNES DJ (1984) Landslide hazard zonation: a review of principles and practice. United Nations International, Paris

VINCENDON B, DUCROCQ V, NUISSIER O, VIÉ B (2011) Perturbation of convection-permitting NWP forecasts for flash-flood ensemble forecasting. Natural Hazards and Earth System Science 11:1529–1544. doi: 10.5194/nhess-11-1529-2011

VON ELVERFELDT K, EMBLETON-HAMANN C, SLAYMAKER O (2016) Self-organizing change? On drivers, causes and global environmental change. Geomorphology 253:48–58. doi: 10.1016/j.geomorph.2015.09.026

WALKER WE, HARREMOËS P, ROTMANS J, ET AL (2003) Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. Integrated Assessment 4:5–17. doi: 10.1076/iaij.4.1.5.16466

WANG Y, ZHAO T, CAO Z (2015) Site-specific probability distribution of geotechnical properties. Computers and Geotechnics 70:159–168. doi: 10.1016/j.compgeo.2015.08.002

WEBSTER R, OLIVER MA (2001) Geostatistics for environmental scientists. John Wiley & Sons, Chichester, New York

WESSELINK A, CHALLINOR AJ, WATSON J, ET AL (2015) Equipped to deal with uncertainty in climate and impacts predictions: lessons from internal peer review. Climatic Change 132:1–14. doi: 10.1007/s10584-014-1213-1

WIECZOREK GF, GUZZETTI F (1999) A review of rainfall thresholds for triggering landslides. In: Proc. of the EGS Plinius Conference, Maratea, Italy. pp 407–414

WIRTZ A, KRON W, LÖW P, STEUER M (2014) The need for data: natural disasters and the challenges of database management. Natural Hazards 70:135–157. doi: 10.1007/s11069-012-0312-4

WMO (2012) Guidelines on Ensemble Prediction Systems and Forecasting. Geneva

WONG HN (2005) Landslide risk assessment for individual facilities—state of the art report. In: Hungr O, Fell R, Couture R, Eberhardt E (eds) Landslide risk management: proceedings of the International Conference on Landslide Risk Management, Vancouver, Canada, 31 May - 3 June 2005. Balkema, Leiden, pp 237–296

WOOD JL, HARRISON S, REINHARDT L (2015) Landslide inventories for climate impacts research in the European Alps. Geomorphology 228:398–408. doi: 10.1016/j.geomorph.2014.09.005

YANG D, KOIKE T, TANIZAWA H (2004) Application of a distributed hydrological model and weather radar observations for flood management in the upper Tone River of Japan. Hydrological Processes 18:3119–3132. doi: 10.1002/hyp.5752

ZHANG S, ZHAO L, DELGADO-TELLEZ R, BAO H (2016) A physics-based probabilistic forecasting model for rainfall-induced shallow landslides at regional scale. Natural Hazards and Earth System Sciences Discussions 1–17. doi: 10.5194/nhess-2016-348

ZIEHER T, RUTZINGER M, SCHNEIDER-MUNTAU B, ET AL (2017) Sensitivity analysis and calibration of a dynamic physically based slope stability model. Natural Hazards and Earth System Sciences 17:971–992. doi: 10.5194/nhess-17-971-2017

A. Publications and manuscripts

The following section contains the publications and manuscripts that are part of this article thesis. Each publication is preceded by a co-author declaration that states the individual research contribution of the PhD candidate and all associated co-authors. Each declaration additionally states the title of the publication and where it was published/submitted. The PhD candidate's research contribution is verified by all co-author's signatures.

A.1. Canli et al. (2016): Generating web scraped high-quality weather databases for near-real-time derivation of spatial land-slide susceptibility

Co-author declaration for the following joint paper

This declaration states the research contribution (e.g. research idea and –questions; data compilation, manipulation and modelling, design and preparation of graphics, maps and tables; writing of text) of the candidate, the main supervisor (where he/she is an associate author) and the other authors.
If applicable, the contributions from other PhD candidates who has or intend to include the paper in a thesis should be described. Contributions from master students should be described.

All Authors (please underline corresponding author):
Ekrem Canli, Bernd Loigge, Martin Mergili, Thomas Glade
Title:
Generating web scraped high-quality weather databases for near-real-time derivation of spatial landslide susceptibility
Publication:
Aversa S, Cascini L, Picarelli L, Scavia C (eds) Landslides and Engineered Slopes. CRC Press, Napoli, Italy, pp 545–550 (2016): https://doi.org/10.1201/b21520-59 (published)

Ekrem Canli's independent contribution:
- writing the paper
- research idea and –questions
- data compilation
- data manipulation and modelling
- design and preparation of graphics and maps

Bernd Loigge's independent contribution:
- database design and implementation

Martin Mergili's independent contribution:
- contribution to research idea and –questions
- revision of manuscript

Thomas Glade's independent contribution:
- contribution to research idea and –questions
- revision of manuscript

Has this paper been, or will this paper be part of another doctoral degree thesis?

No

Contributions from master students:

No

Do you verify that Ekrem Canli has contributed to this joint paper as described above?

Yes No

If no, specify:

If applicable, first co-author's signature

Yes No

If no, specify:

If applicable, second co-author's signature

Yes No

If no, specify:

If applicable, third co-author's signature

Landslides and Engineered Slopes. Experience, Theory and Practice – Aversa et al. (Eds)
© 2016 Associazione Geotecnica Italiana, Rome, Italy, ISBN 978-1-138-02988-0

Generating web scraped high-quality weather databases for near-real-time derivation of spatial landslide susceptibility

E. Canli, M. Mergili & T. Glade
Department of Geography and Regional Research, University of Vienna, Vienna, Austria

B. Loigge
SPL Tele GmbH & Co. KG, Wolkersdorf, Austria

ABSTRACT: Web scraping is the automated process of gathering data from the internet. This data is stored in a database for further analyses. Scraped data is especially valuable for time dependant research questions, as no own measurement or monitoring devices are necessary to collect real time data. This paper aims at introducing a novel approach for generating large, high quality weather databases from freely available internet weather sources and applying them to near-real-time spatial landslide assessments. We present a first step by automating the procedure from gathering real-time weather data to geostatistical analyses for assessing the spatial pattern of rainfall. The output will be used for the immediate implementation in regional, physical-based landslide susceptibility models.

1 INTRODUCTION

Increasing temperature (e.g. Böhm et al. 2001) and precipitation variability (e.g. Casty et al. 2005) across the European Alps makes it necessary to understand relationships between climate change and landsliding (Crozier 2010). However, it is widely accepted that losses from landsliding are vastly underestimated (Petley 2012). Besides common monitoring and spatial assessment, this increasingly demands reliable and timely landslide warnings. Spatio-temporal predictions of landslides are challenging, mainly due to data constraints. This is even more the case in developing countries, also because the global distribution of fatal landslides clearly shows a strong clustering in such countries. Scarcity of high-quality meteorological data is often referred to as one of the main constraints for performing real-time landslide susceptibility/hazard analysis (e.g. Thiebes et al. 2014). Meteorological data may be expensive or not up-to-date any more soon after it is acquired.

The internet is a great source of freely available, high quality real-time weather data from various operators. By now, web scraping has emerged into a highly valuable technique for extracting information from websites. Hereby, web scraping (or screen scraping) is the process of automatically gathering data from the internet and using those self-generated databases for further analyses (Beran et al. 2009). This technique is of great value for weather data that is released at least hourly to the public, but may be applicable to any other type of continuously released data. Scraping weather data from the internet is furthermore not just limited to sources from officially operated weather stations, but may take advantage of privately operated weather stations whose data is available at online services such as http://openweathermap.org. Additionally, research institutions in developing countries can easily generate their own weather databases without the need of purchasing expensive, ready-made weather data.

In this paper we present a novel approach in landslide research for generating large, high quality weather databases from readily available internet weather sources and applying them to a preliminary step in dynamic landslide susceptibility assessments. The objectives of this paper are covering the following points:

– outlining the technique of web scraping for generating spatially widespread near-real-time weather databases, and
– the instant conversion of each new dataset into a continuous rainfall map based on geostatistical interpolation for subsequent physically based landslide susceptibility analysis.

2 METHODS

2.1 *Web scraping*

In the last few years environmental sciences are witnessing an increasing amount of published informationen on the internet (Vitolo et al. 2013). Therefor Web scraping (often also named Web data extraction/Web harvesting/Web Mining/Screen Scraping), a technique for extracting data and information from websites and storing it in a database, can be used to generate the datasources needed. There is a large variety of different web scraping solutions that range from manual examination and copy-and-paste to fully automated systems. In this paper, two different approaches are disposed for scraping data from the internet. Firstly, the freely available web scraping software Kimonolabs was applied to access data from the Austrian Central Institute for Meteorology and Geodynamics (ZAMG). Secondly, a programming approach was used to retrieve data from the Lower Austrian government (NOEL), which runs its own stations.

The framework used for storing the data is MeteorJS, an open source, cohesive, pure JavaScript framework and development platform. With NodeJS on the server, MongoDB on the backend and a huge variety of freely available packages and libraries, MeteorJS offers an easy and fast way to develop web applications. Figure 1 illustrates the whole framework used in this paper.

2.1.1 *MeteorJS*

MeteorJS was choosen in this work because it allows a rapid prototyping and with a few lines of code a server/client application can be set up.

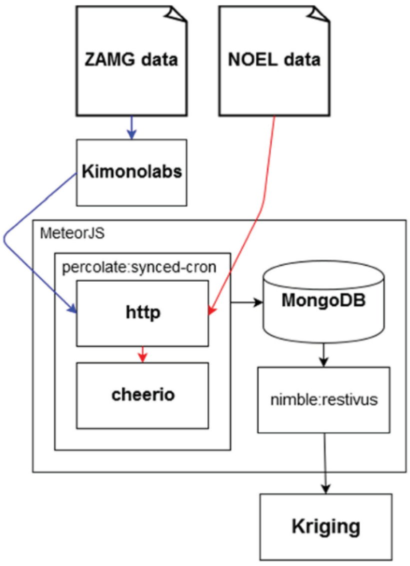


Figure 1. Framework for the automated scraping and storing procedure for weather data from different providers.

MongoDB, classified as a NoSQL database, stores data as JSON–like (Javascript Object Notation) documents making it a fast way to store scraped data. Furthermore, MeteorJS offers a packaging system called AtmosphereJS. For accessing data via APIs (Kimonolabs) and raw HTML content (Cheerio) the http-package was used, which provides an HTTP request API on the client and server. For automatical time-based scheduling the *percolate:synced-cron* package was used, which synchronizes jobs between multiple processes. For data evaluation, the *nimble:restivus*-package was used to create REST APIs and setup CRUD endpoints for the defined MongoDB collections.

Additionally, NoSQL databases provide more flexibility than relational databases because of the changeability of their schema, so the data can be accessed faster in other projects (Vitolo et al. 2013).

2.1.2 *Kimonolabs*

Kimonolabs is a browser based software that enables to glean specific data from any website without coding. It offers an User Interface (UI) that recognizes elements on a webpage that are structurally similar to what is selected. In the background it generates a data model that tries to determine a common pattern. The user is able to edit the queries and combine specific parts via the UI. If there is a need for more complex modifications, the results can be modified by writing javascript functions in the UI. After setting up the scraping mechanism, Kimonolabs creates a REST API which then can be used to scrape the data. Kimonolabs was used to scrape hourly data for 53 weather stations in Lower Austria from the Austrian Central Institute for Meteorology and Geodynamics (ZAMG).

2.1.3 *Cheerio*

Web scraping with Cheerio is another way of accessing the DOM (Document Object Model) of any website. Cheerio is based on jQuery, designed specifically for Node.JS. Via Meteor's *http*-package, the HTML content was loaded. To get cleaner data submodifications on the HTML, the content was preprocessed using the Node.JS module *html-parser2*. Afterwards the weather data was extracted and stored in MongoDB. Using Cheerio, hourly data from the Lower Austrian governmental website (NOEL) was scraped from 63 weather stations.

2.2 *Automated interpolation of scraped rainfall data*

widespread and dense rainfall data are of great importance for the hydrological modelling component in regional slope stability models. Geostatistical methods (such as kriging) are widely applied in spatial interpolation that transfer point

546

measurements to continuous surfaces (Kitsanidis 1997). In this paper we used a kriging approach for automatically generating continuous rainfall maps every hour. Spatial interpolation with geostatistical and Inverse Distance Weighting (IDW) algorithms have shown to outperform interpolation with Thiessen polygons that are used in various hydrological models. Incorporating elevation information from a DEM may aid for multivariate geostatistical analysis, however, Ly et al. (2011) have shown that integrating elevation into kriging did not improve the interpolation accuracy for daily rainfall. For hourly rainfall data, Haberlandt (2007) also mentioned that elevation information plays only a minor role. Goovaerts (2000) concluded that the benefit of incorporating elevation to multivariate techniques may be marginal, if correlations between rainfall and any descriptor (such as elevation) becomes too small which might be the case for rainfall during shorter time steps. In this study we use hourly time steps, therefore we neglect the elevation information as a descriptor variable and apply the ordinary kriging (ORK) method to our dataset.

Kriging aims at estimating the value of a random variable at one or more unsampled points from more or less sparse sample data on a given support (Webster & Oliver 2007). Spatial interpolation can be generalized as:

$$Z_g = \sum_{i=1}^{ns} \lambda_i Zs_i \qquad (1)$$

where Zg is the interpolated value at point g; Zs_i is the observed value at point i; ns is the total number of observed points (rain gauges) and λ_i is the weight contributing to the interpolation. The most crucial step in performing spatial interpolation, thus also in kriging, lies in the calculation of the weights. Values of regionalized variables tend to be related, whereas two points close to each other are more similar than at more widely separated places. The initial step in kriging contains the variogram analysis. The variogram (or theoretical variogram) is a tool for quantifying spatial correlation and is a mathematical expression of the semivariogram (also empirical or experimental variogram) which is computed from the data. The variogram is approximated by a range of different models (such as Spherical, Gaussian, Exponential, etc.) that ensure validity. More in-depth information on geostatistical theories can be found in Goovaerts (1997), Kitanidis (1997) or Webster & Oliver (2007). The behavior of the variogram at greater distances determines whether the function is stationary. For such functions the semivariogram should stabilize around a value, called sill. In case of a stationary function, the length scale at which the sill is obtained determines the scale at which two measurements of the variable become practically uncorrelated (the so called range). In case of non-stationarity, the variogram would keep increasing beyond the separation distance of interest.

The fitting of the semivariogram with a theoretical model is most commonly a manual fitting procedure. This ongoing work presented here ultimately aims at producing automated, unsupervised spatial interpolation of near-real-time rainfall data based on kriging on an hourly basis. Thus, the semivariogram plays an important role as the variogram is selected solely on an automatic model fitting procedure. The coefficients derived from the applied model were then used to determine the weights for the subsequent kriging. In this study we implemented an automated variogram fitting procedure based on a Gaussian model with correlation lengths that are one third of the area's square root range. The sill is automatically determined by the variance of each time slices' rain gauge measurements. To fit the variogram sills to the data, a residual (or restricted) maximum likelihood (REML) estimation was used. The advantage of REML over the more common Ordinary Least Squares (OLS) and Maximum Likelihood (ML) estimation lies in providing unbiased estimates of the variance parameters (thus neglecting nuisance parameters) (Lark & Cullis 2004, Webster et al. 2006). In our ordinary kriging procedure, we consider the semivariogram solely as a measure of spatial correlation that is independent from orientation or direction, thus, anisotropy is not taken into account. A Cross Validation (CV) was performed whereas the goodness of fit was expressed by the related Pearson's r. Additionally, the kriging variance was calculated to take spatial uncertainties into account. We forego the estimation of a Root Mean Square Error (RMSE), although a common performance criterion, as it usually provides little information on the reliability of kriging estimates (Goovaerts 2000, Ly et al. 2011).

The whole interpolation chain is automated with R (open source software for statistical computing) and additional packages for geospatial analysis (R Core Team 2015). The R packages used for kriging are: *sp* (allows R to deal with spatial objects), *gstat* (containing geostatistical tools) and *rgdal* (for handling projection/transformation operations of spatial objects directly in R).

3 RESULTS

3.1 *Web scraping*

The two used approaches in this paper gave slightly different results with respect to reliability of the scraping procedure. On the one hand, Kimonolabs as a free third party scraping tool was easy to imple-

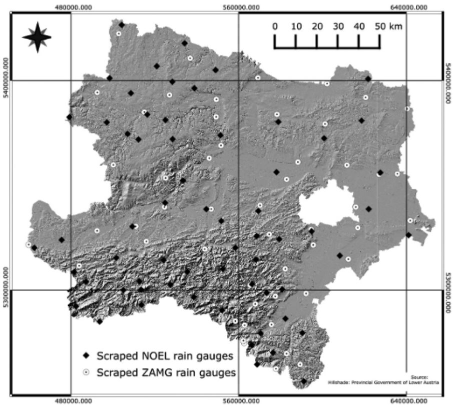


Figure 2. Hillshade of Lower Austria with rain gauges from ZAMG and NOEL that are currently scraped.

ment, but sometimes did not respond properly and therefore produced outage. On the other hand, Cheerio required more development time, as all of the scraping was done by script. Regarding the reliability of both approaches, the extraction of hourly data worked in 72% of cases with Kimobolabs, whereas scraping with Cheerio worked in 96% of cases. Unfortunately, the reasons for those outages with Kimobolabs could not be traced, as the error logged by the system is not sufficiently documented ("500 Server error" which refers to "something went wrong on our end" https://www.kimonolabs.com/apidocs#Limit). The minimal outage using Cheerio can be explained by downtimes of the scraped website. We used 3 scraped datasets of rain gauges to perform the subsequent spatial interpolation: 53 ZAMG stations, 63 NOEL stations and all 116 stations together. Figure 2 shows the distribution of the rain gauges in Lower Austria.

3.2 *Spatial distribution of hourly rainfall*

Following the extreme heat wave in Central Europe in summer 2015, heavy rainfall and hailstorms hit Lower Austria on the evening of July 8, 2015. This storm event brought large amounts of rainfall in the southern part of Lower Austria during the night, whereas the northern part was mostly unaffected. We picked this storm event as it shows a clear spatial differentiation in its rainfall pattern according to the information from the distributed rain gauges. We used 3 consecutive hours (19:00, 20:00 and 21:00) to produce automatically generated rainfall maps.

For every hour, variogram models for all three datasets were generated with an automated procedure (Fig. 3). Our findings are similar to those of Ly et al. (2011), except that we used only a Gaussian model for best fits. As a measure of the dispersion of all observations, the semivariance increased in accordance with the separation distance. This means that rainfall data close to each other are more similar (and their squared difference less significant) than those further apart.

After generating the variograms, kriging was performed with the same 3 datasets for the same 3 consecutive hours. The results are calculated for a 1 × 1 km raster output. Figure 4 shows the maps that resulted from kriging with the related maps of spatial variance contained in the output. The overall picture shows a relatively similar pattern, given the fact that not the distance alone, but weights from surrounding measurements are used to predict values at each gridcell. Thus, the amount of rain gauges and their distribution significantly influence the overall result, although the same automated procedure for creating the variogram was applied to all datasets. The areas with lower

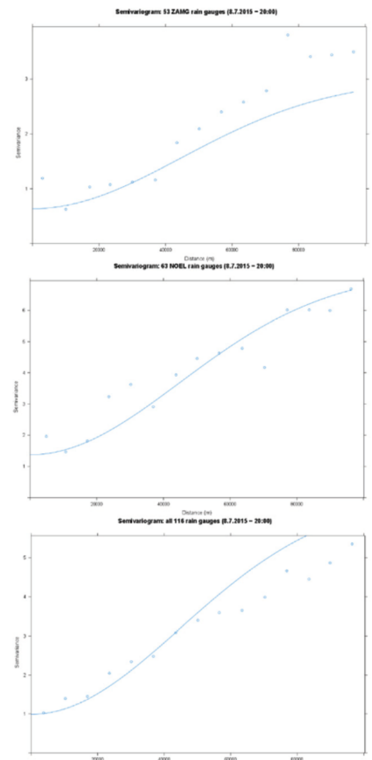


Figure 3. Semivariograms for 20:00 on July 8, 2015 on all 3 datasets with automatically fitted variogram based on a Gaussian model with REML estimates.

548

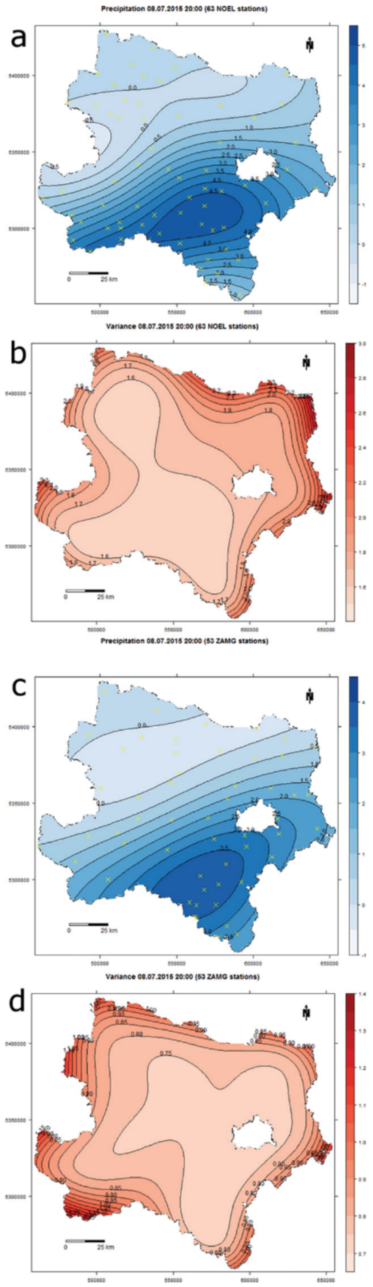


Figure 4. Raster maps (1 × 1 km grid size) produced from automated kriging. Map a and c represent the same timeslice (20:00) for the NOEL (a) and ZAMG (c) dataset. Map b and d show the associated spatial variance.

Table 1. Validation results for all datasets.

Dataset	Rain gauges	Pearson's r
ZAMG 19:00	53	0.63
ZAMG 20:00	53	0.74
ZAMG 21:00	53	0.79
NOEL 19:00	63	0.30
NOEL 20:00	63	0.69
NOEL 21:00	63	0.48
ALL 19:00	116	0.43
ALL 20:00	116	0.74
ALL 21:00	116	0.68

rainfall amounts in the northern part are almost identically calculated, whereas the most evident deviation between the three datasets occurred in the south-eastern part. This may be attributed to a sparsely distributed number of rain gauges in the area in one dataset, but also to orographic effects that cannot be resolved with Ordinary Kriging alone, as this method neglects correlated secondary variables (such as elevation). Also, the variance significantly increases towards the margin of the study area. This can be referred to omitted rain gauges in neighboring states or countries and thus affecting the distance in the semivariogram, as only available gauges in greater distances in a certain direction can be used to calculate the variogram. For the goodness of fit estimation we used a cross validation approach to obtain the linear correlation coefficients between the observed and predicted rainfall. Table 1 gives an overview of the calculated Pearson's r values.

4 CONCLUSIONS AND OUTLOOK

The abundance of data the internet is fed with every day opens up new paths on how geosciences can be pursued in the future. Lots of research done in the field of natural hazards is kind of static and does not incorporate real-time data, which is remarkable, as most natural hazards are often related to immediate events due to external triggering factors. The landslide community does not make an exception, as a majority of modelling approaches deal with retrospect cases. On the contrary, the rainfall threshold community that indeed aims towards early warning of landslide occurrence, relies on the triggering component alone. Only few works so far addressed real-time dynamic components for regional landslide susceptibility and even hazard assessment.

With this paper we aim at taking a preliminary step in combining spatially distributed real-time elements with spatial landslide suscseptibility and hazard assessments. Automated web scraping has proven to be a great technique for extracting information from the internet, but to our best knowledge, this has not been applied in landslide research yet to acquire real-time data that is implemented in dynamic landslide modelling approaches. In this paper we present a novel approach on how to extract real-time meteorological data and produce automated, spatially distributed rainfall maps based on geostatistical methods. The results presented here belong to our first iteration in producing real-time rainfall maps, however, the results are quite promising. We are yet dealing with a couple of obstacles that arise from different sources. The first step in the workflow consists of building a database containing meteorological data. However, some rain gauges produce inconsistent entries that need to be handled beforehand. The more regular and reproducible such inconsistencies are, the easier it is to catch them in the error handling process. Another problem arises from rain gauge malfunctioning. A single rain gauge that is not working during a heavy rainfall event but still shows up on the internet produces a massive bias as the semivariogram gets rather incoherent. This asks for a plausibility check that is yet to be implemented. So a prerequisite for the automated geostatistical analyses is a cleaned up, homogenous dataset. With respect to the scraping procedure itself, we currently advise to set up own scripts to gather all the desired web data in order to increase data availability. However, third party web scraping applications (such as Kimonolabs tested in this work) can become a good alternative, especially for non-programmers and for fast development.

The overall lower correlation coefficients for the 19:00 datasets require further investigation, as other factors clearly contribute to a much higher degree to the predicted amount of rainfall than distance alone (thus the model is underfitted). The automated modelling of the variogram as a function describing the degree of spatial dependence between measurement locations, is common practice, but not as straightforward as a manual fitting procedure. Additionally, the automated fitting for the 19:00 variogram, even when feasible for other times and/or datasets, may not be the best solution in this case. Therefore, a more flexible automating procedure is necessary for iteratively checking for the most suitable model for fitting the variogram.

Ultimately we aim at connecting the rainfall maps immediately into an automated, dynamic slope stability model to assess factor of safety estimations every hour. This workflow will be accompanied by implementing locally monitored hydrological parameters (ground water levels, pore water pressure) that change dynamically in case of a prolonged rainfall event (Bordoni et al. 2015).

REFERENCES

Beran, B., Goodall, J., Valentine, D., Zaslavsky, I., Piasecki, M. 2009. Standardizing Access to Hydrologic Data Repositories through Web Services. In: Dragicevic, S., Dumitru, R., Vlad, T. (eds.), *International Conference on Advanced Geographic Information Systems & Web Services—GEOWS 2009; Proc. intern. symp., Cancun, 1-7 February 2009*. Los Alamitos: IEEE Computer Society.

Böhm, R., Auer, I., Brunetti, M., Maugeri, M., Nanni, T., Schöner, W. 2001. Regional temperature variability in the European Alps: 1760–1998 from homogenized instrumental time series. *International Journal of Climatology* 21: 1779–1801.

Bordoni, M., Meisina, C., Valentino, R., Bittelli, M., Chersich, S. 2015. Site-specific to loca-scale shallow landslides triggering zones assessment using TRIGRS. *Nat. Hazards Earth Syst. Sci.* 15: 1025–1050.

Casty, C., Wanner, H., Luterbacher, J., Esper, J., Böhm, R. 2005. Temperature and precipitation variability in the European Alps since 1500. *International Journal of Climatology* 25: 1855–1880.

Crozier, M.J. 2010. Deciphering the effect of climate change on landslide activity: A review. *Geomorphology* 124: 260–267.

Goovaerts, P. 1997. *Geostatistics for natural resources evaluation*. New York: Oxford University Press.

Goovaerts, P. 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology* 228: 113–129.

Haberlandt, U. 2007. Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *Journal of Hydrology* 332: 144–157.

Kitsanidis, P.K. 1997. *Introduction to Geostatistics—Applications in Hydrogeology*. Cambridge: Cambridge University Press.

Lark, R.M. & Cullis, B.R. 2004. Model-based analysis using REML for inference from systematically sampled data on soil. *European Journal of Soil Science* 55: 799–813.

Ly, S., Charles, C., Degré, A. 2011. Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium. *Hydrology and Earth System Sciences* 15: 2259–2274.

Petley, D. 2012. Global patterns of loss of life from landslides. *Geology* 40: 927–930.

R Core Team. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Thiebes B., Bell R., Glade T., Jäger S., Mayer J., Anderson M., Holcombe L. 2014. Integration of a limit-equilibrium model into a landslide early warning system. *Landslides* 11(5): 859–875.

Vitolo, C., Elkhatib, Y., Reusser, D., Macleod, C.J.A., Buytaert, W. 2015. Web technologies for environmental Big Data. *Environmental Modelling & Software* 63: 185–198.

Webster, R., Welham, S.J., Potts, J.M., Oliver, M.A. 2006. Estimating the spatial scales of regionalized variables by nested sampling, hierarchical analysis of variance and residual maximum likelihood. *Computers & Geosciences* 32: 1320–1333.

Webster, R. & Oliver, M.A. 2007. *Geostatistics for environmental scientists*. Chichester: Wiley.

A.2. Canli et al. (2017a): Spatially distributed rainfall information and its potential for regional landslide early warning systems

Co-author declaration for the following joint paper

This declaration states the research contribution (e.g. research idea and –questions; data compilation, manipulation and modelling, design and preparation of graphics, maps and tables; writing of text) of the candidate, the main supervisor (where he/she is an associate author) and the other authors.
If applicable, the contributions from other PhD candidates who has or intend to include the paper in a thesis should be described. Contributions from master students should be described.

All Authors (please underline corresponding author):
Ekrem Canli, Bernd Loigge, Thomas Glade
Title:
Spatially distributed rainfall information and its potential for regional landslide early warning systems
Journal:
Natural Hazards (2017): https://doi.org/10.1007/s11069-017-2953-9 (published)

Ekrem Canli's independent contribution:
- writing the paper
- research idea and –questions
- data compilation
- data manipulation and modelling
- design and preparation of graphics and maps

Bernd Loigge's independent contribution:
- database design and implementation

Thomas Glade's independent contribution:
- contribution to research idea and –questions
- revision of manuscript

Has this paper been, or will this paper be part of another doctoral degree thesis?
No

Contributions from master students:
No

Do you verify that Ekrem Canli has contributed to this joint paper as described above?

Yes No

If no, specify:

____
If applicable, first co-author's signature

Yes No

If no, specify:

____
If applicable, second co-author's signature

Has this paper been, or will this paper be part of another doctoral degree thesis?
No

Contributions from master students:
No

Do you verify that Ekrem Canli has contributed to this joint paper as described above?

Yes No

If no, specify:

____
If applicable, first co-author's signature

Yes No

If no, specify:

____
If applicable, second co-author's signature

Nat Hazards
DOI 10.1007/s11069-017-2953-9

ORIGINAL PAPER

Spatially distributed rainfall information and its potential for regional landslide early warning systems

Ekrem Canli[1] · Bernd Loigge[2] · Thomas Glade[1]

Received: 3 October 2016 / Accepted: 31 May 2017

Abstract Crucial to most landslide early warning system (EWS) is the precise prediction of rainfall in space and time. Researchers are aware of the importance of the spatial variability of rainfall in landslide studies. Commonly, however, it is neglected by implementing simplified approaches (e.g. representative rain gauges for an entire area). With spatially differentiated rainfall information, real-time comparison with rainfall thresholds or the implementation in process-based approaches might form the basis for improved landslide warnings. This study suggests an automated workflow from the hourly, web-based collection of rain gauge data to the generation of spatially differentiated rainfall predictions based on deterministic and geostatistical methods. With kriging usually being a labour-intensive, manual task, a simplified variogram modelling routine was applied for the automated processing of up-to-date point information data. Validation showed quite satisfactory results, yet it also revealed the drawbacks that are associated with univariate geostatistical interpolation techniques which solely rely on rain gauges (e.g. smoothing of data, difficulties in resolving small-scale, highly intermittent rainfall). In the perspective, the potential use of citizen scientific data is highlighted for the improvement of studies on landslide EWS.

Keywords Rainfall prediction · Web scraping · Geostatistics · Landslides · Early warning system

1 Introduction

Rainfall-induced landslides pose a threat to people and infrastructure around the world (e.g. Guzzetti et al. 1999; Crosta and Frattini 2003; van Westen et al. 2008; Günther et al. 2013). Rising temperature (e.g. Gobiet et al. 2014) and rainfall variability (e.g. Casty et al.

Ekrem Canli
ekrem.canli@univie.ac.at

1 Department of Geography and Regional Research, University of Vienna, Universitaetsstr. 7, 1010 Vienna, Austria

2 SPL Tele GmbH & Co KG, Johann-Galler-Straße 39, 2120 Wolkersdorf, Austria

Published online: 08 June 2017

2005) in the European Alps make it of great importance to increase efforts in dealing with the consequences of global change on landslide susceptibility. Albeit advancements in the past within the field of geotechnical engineering have led to an increasing in situ damage control in many parts of the world, landslides triggered by heavy rainstorms still cause great losses where no protective structures are available or where they have not been appropriately designed. Indeed, fatal landslide-triggering rainstorm events are occurring in many places around the globe. Recent examples include Messina 2009 (Lombardo et al. 2014) or Rio de Janeiro 2010 (Calvello et al. 2015). Besides those landslides occurring in natural, undisturbed conditions, also established geotechnical mitigation measures are not sufficient and fail in many cases. Although landslides are usually restricted to local sites when investigating single events, they can indeed occur in clusters in the aftermath of a storm event and consequently they can be regarded as a regional phenomenon at specific times (Jaedicke et al. 2014). Floods, on the other hand, are not locally restricted either; however, their spatial occurrence is topographically foreseeable and controllable. This situation shows the necessity of landslide early warning systems (EWS) (Glade and Nadim 2014; Thiebes and Glade 2016).

The observation of intense rainfall events on small spatio-temporal scales is crucial for the development of a landslide EWS (Segoni et al. 2009; Thiebes et al. 2013). The predominant approach in implementing rainfall data into landslide EWS is the employment of empirical rainfall thresholds. This requires the precise knowledge of total precipitation accumulated in a given period, or the rate of precipitation in a period, most commonly measured in millimetres/inches per hour (Guzzetti et al. 2007). A certain threshold is then defined for a given spatial extent by determining the rainfall amount that triggered landslides. Hereby, it is important to differentiate the lowest and highest landslide-triggering rainfall thresholds as the lower and upper limit, below which landslides were never reported and above which landslide were always reported (Gariano et al. 2015).

One major challenge within this approach is, how the recorded landslide-triggering rainfall events were selected. In almost all cases, the precise amount of a landslide-triggering rainfall at a certain location remains unknown. In practice, rain gauges with the closest proximity to a landslide location or which provide the best representation of a certain region are selected for determining a landslide-triggering rainfall event. An in-depth consideration of the accurate spatial distribution of rainfall is often neglected for landslide EWS (Thiebes and Glade 2016).

Consequently, the aim of this paper is to provide a useful basis for real-time spatio-temporal rainfall data and to show its potential integration in a regional landslide EWS. Instead of assuming uniform rainfall over a certain area, an automated geo-statistical approach is presented. This allows for an approximation of spatially distributed, hourly rainfall predictions in real time based on gauged rainfall data available on the Internet.

2 Review of current approaches

2.1 Landslide early warning systems

The UNEP (2012) suggests four key elements which are required for any operational EWS: (a) a comprehensive assessment of the risks, (b) the implementation of monitoring and predicting capabilities, (c) a reliable, synthetic and simple dissemination strategy and (d) development of response strategies combined with the need for raising public awareness and education. Besides financial constraints, however, many operational landslide EWS are discontinued due to the fact that only the purely technical components are considered, and the social aspects such as the elements (c) and (d) (UNEP 2012) are not addressed (Baum and Godt 2010). Not many cases of an integrated approach have been published (e.g. Thiebes 2012; Heil et al. 2014).

Many studies describe the calculation of rainfall thresholds best suitable for a given region (e.g. Glade et al. 2000; Guzzetti et al. 2007; Aleotti 2004). The region under consideration for a certain threshold may range from local scale (Keefer et al. 1987) to global scale (Hong et al. 2007). Commonly used thresholds are either attributing rainfall directly or consider additionally the antecedent rainfall. The first one calculates rainfall thresholds for a certain region based on rainfall intensity over a certain time span that caused landslides (e.g. *intensity–duration threshold*), the latter takes the antecedent rainfall before slope failure into account (e.g. *antecedent rainfall threshold*), thus considering indirectly also the soil moisture conditions. Literature reports a landslide EWS in Hong Kong implemented in 1977 probably as the first one of its kind (Chan et al. 2003). The first operational landslide EWS in the USA were implemented in the early 1980s (Baum and Godt 2010). A notable mention is the landslide EWS in the San Francisco bay area in 1985 (Keefer et al. 1987) which is said to be the pioneer of modern landslide EWS in combination with rainfall thresholds (Stähli et al. 2015). Ultimately, every successful landslide EWS is a function of expectations that does not only satisfy the requirements of the developers, but also the decision-makers, end-users and the regulations of the prevalent legal system (Thiebes and Glade 2016).

Based on reviewing over 50 EWS for mass movements in Switzerland, Stähli et al. (2015) suggested the following classification scheme for EWS:

(a) *Alarm systems* are directly coupled to sensors that immediately release an alarming upon exceeding a predefined threshold. The accuracy of the prediction is high and the lead time very short (e.g. rockfall).
(b) *Warning systems* detect significant changes in time-dependent factors before an event occurs (e.g. a rainfall threshold). The initial alert is based on predefined thresholds too but ultimately the alarm is released after an expert evaluation. With an extended lead time, warning systems are mainly used for processes with progressive stages of failure (e.g. rock slides, translational and rotational soil slides).
(c) *Forecasting systems* do not use any predefined thresholds, but the level of danger is evaluated in regular intervals. Experts are issuing warnings based on modelling results or sensor data (e.g. degree of avalanche danger for the next day).

Those thresholds involved in landslide EWS can be derived by different approaches. Alarm systems can be based on direct displacement measurements (e.g. Intrieri et al. 2012), while warning systems and forecasting systems obtain their respective thresholds based on rainfall measures (e.g. Capparelli and Tiranti 2010; Aleotti 2004) or deterministic models (e.g. Montrasio et al. 2014; Chien et al. 2015). The most common types applied in landslide EWS are basic rainfall thresholds. Dealing with thresholds always implies a certain amount of a priori knowledge. This knowledge is not readily available in most cases, especially with respect to the precise date and time of landslide occurrence (van Westen 2006, 2008) and its respective rainfall conditions that ultimately triggered the failure (Wieczorek and Guzzetti 1999).

When dealing with landslide EWS and having the various data constraints in mind, it is evident that it is impossible to cover all early warning situations (Casagli et al. 2010). Although the calculation of rainfall thresholds is the most common approach for assessing regional landslide early warning situations, they only represent a simplification of the physical processes involved (Reichenbach et al. 1998). This means that in most cases there is more than just this one causative factor (rainfall) involved (Huang et al. 2015).

2.2 Observation of rainfall and its spatial representation in landslide EWS

The most important aspect of any landslide EWS is the precise and timely determination of rainfall, no matter whether the approach of the system is based on empirical rainfall thresholds or combined hydrological and slope stability modelling in a landslide forecasting framework. Most researchers are aware of the importance of the spatial variability of rainfall in landslide research. In applications, however, it is commonly not addressed due to the scarcity of available rain gauges (Chiang and Chang 2009). With respect to rainfall thresholds, common approaches to determining the actual rainfall for a single landslide site or a certain region are to utilize single rain gauges near a specific landslide site (e.g. Capparelli and Tiranti 2010) or selected rain gauges as representative locations for a predefined region (e.g. Segoni et al. 2015; Rosi et al. 2015). Aleotti (2004) pointed out that using representative rain gauge locations is dependent on the distance not only from the landslide, but also from other settings such as elevation, aspect or the wind direction. Lagomarsino et al. (2013) present a more advanced approach in their SIGMA warning system by not using a single rainfall threshold for the entire territory of Emilia Romagna (Italy). Instead, they artificially split the region into smaller spatial units, so-called territorial units (TU), which is each characterized by its own threshold. However, each TU is characterized by single rain gauges, thus only offering uniform rainfall for the entire area. With respect to the amount of rain gauges used to characterize aerial rainfall, Guzzetti et al. (2004), for example, use seven rain gauges for an area of ca. 5500 km^2 for their rainfall-induced landslide studies. Similarly, Bathurst et al. (2006) use five for ca. 500 km^2 or Schwab et al. (2008) just one rain gauge for 324 km^2.

Another seemingly very attractive approach to determining continuous rainfall fields is the utilization of radar technology. Radar technology is capable of offering spatially varying rainfall fields at temporal resolutions of 10 min (and less) and spatial resolutions of around 1 km, which is highly desirable for landslide studies (Chiang and Chang 2009). In addition, there is also vertically differentiated information on precipitation values available. The quantification of rainfall estimates using Doppler radar can provide a real-time comparison with rainfall thresholds to form the basis of landslide warnings (Wieczorek and Guzzetti 1999). Radar data may capture the spatial variation of rainfall fields better than gauged data, especially in mountainous regions where rain gauges are sparsely distributed (Yang et al. 2004; Segond et al. 2007). At the same time, however, terrain obstacles may cause (partial) beam blockage (or beam shielding) when radars are operated in mountainous regions. Beam blockage correction offers ways to enhance rainfall estimates in mountainous regions, but sometimes beam occlusion cannot be prevented in a region of complex terrain (Lang et al. 2009; Anagnostou et al. 2010). On the other hand, radar coverage is not available everywhere (Chappell et al. 2013). What also must be kept in mind is that radar is not capable of measuring precipitation directly. Instead, the rainfall magnitudes are derived from the magnitude of measured reflectivity from the hydrometeors in the atmosphere (Harpold et al. 2017). This procedure requires calibration with rain gauge data in order to adjust the high-resolution spatial pattern with actual measured rainfall data. There is an abundance of literature dealing with this merging process (e.g. Collier et al. 1983; Jewell and Gaussiat 2015; Velasco-Forero et al. 2009; Berndt et al. 2014).

Within the landslide community, the potential of radar data remains often underexploited. There are only few studies implementing spatially continuous rainfall data from radar technology for landslide research (e.g. Crosta and Frattini 2003; Chiang and Chang 2009), even fewer for early warning purposes in combination with numerical weather prediction models and radar data (e.g. Schmidt et al. 2008; Segoni et al. 2009). Although the technology has made great advancements in recent years, the associated uncertainties and generally low success rates (in terms of correctly predicted landslide occurrences) reduce its application within the landslide community. Schmidt et al. (2008) proposed a coupled regional forecasting system in New Zealand based on multiple process-based models (weather forecast, soil hydrology, slope stability). Segoni et al. (2009) proposed a similar approach, yet with the same conclusion that the meteorological uncertainty has the highest influence on the final Factor of Safety map that serves as the basis for landslide early warning. The hydro-meteorological community is far more involved with the utilization of radar data for the deduction of continuous rainfall fields. However, they share the same concerns (Jasper et al. 2002) or even have taken a step further by implementing numerical weather predictions (Cloke and Pappenberger 2009).

Another emergent technology for assessing the spatial extent of rainfall is satellite precipitation data, especially in regions without rain gauges and ground radar coverage. Satellite-based rainfall estimates provide synoptic estimates of the spatial distribution of precipitation events (Chappell et al. 2013). This information is provided at 0.5–3-h intervals at spatial resolutions between 0.07° and 0.25° (Joyce et al. 2004; Kubota et al. 2007; Huffman et al. 2007, 2010). Tian and Peters-Lidard (2010) questioned the overall accuracy of satellite rainfall products; however, NASA's Tropical Rainfall Measuring Mission (TRMM) has collected precipitation data for over 17 years by now, which is a valuable data archive for global precipitation data (Kirschbaum and Patel 2016). The popularity of satellite-based precipitation data experienced a considerable rise in February 2014 when NASA launched the Global Precipitation Mission (GPM) as a follow-on mission to TRMM (Harpold et al. 2017). GPM provides rainfall and snowfall estimates every 3 h with sensors that are far more advanced and permit better quantification of the physical properties of precipitation particles (Hou et al. 2014). With respect to landslide research, there are only few studies available that utilize satellite-based precipitation data (e.g. Rossi et al. 2012; Kirschbaum et al. 2015), all of which still rely on TRMM data. It will be interesting to observe how near real-time GPM data with higher spatial resolution will be adopted by the landslide community in terms of early warning applications (Stanley et al. 2017).

What has not been covered in detail for landslide studies so far is the wide topic of spatial interpolation for the generation of continuous rainfall fields, which will be addressed in the next chapter as the focus of this study.

2.3 Spatial interpolation techniques for the generation of continuous rainfall fields

Spatial prediction is almost always based on samples, but in reality the measurements represent a continuum in space from which the samples have been taken (Oliver and Webster 2014). Historically, spatial prediction was undertaken by purely mathematical approaches that considered only systematic or deterministic variation, but not any error.

Geostatistical prediction, namely kriging, is the logical successor that overcomes most of these drawbacks (Webster and Oliver 2001). Geostatistics aim explicitly at correctly portraying spatial variation (Srivastava 2013), or more precisely, kriging has become a term for several related least-squares methods that provide best linear unbiased predictions (BLUP), in which best is meant in the sense of minimum variance (Oliver and Webster 2014). The atmosphere, or any other environmental feature, is the sum of all kinds of physical, chemical or biological interactions. Although physically determined, they still remain more or less a black box due to its complex interactions that are not fully understood, thus making the variation appear to be random (Oliver and Webster 2014). Consequently, many environmental variables, such as rainfall, can be considered as spatial random variables.

There are many studies available that compare different interpolation methods for assessing spatial rainfall distribution, the majority from the hydrological or hydro-meteorological communities (e.g. Ly et al. 2013; Mair and Fares 2011; Schuurmans et al. 2007; Haberlandt 2007; Goovaerts 2000). Most studies focus on monthly or annual rainfall estimates (Ly et al. 2013). There is only a small number of studies available that uses hourly time steps for the spatial interpolation of rainfall (e.g. Haberlandt 2007; Velasco-Forero et al. 2009; Schiemann et al. 2011; Verworn and Haberlandt 2011). Very few studies are available for landslide applications (e.g. Chiang and Chang 2009). The literature suggests a differentiation between deterministic and geostatistical approaches. The most frequently used deterministic methods for estimating rainfall are Thiessen polygons and inverse distance weighting (IDW) (Ly et al. 2013). The Thiessen polygon method (also Voronoi polygons, Dirichlet tessellation) is one of the earliest and simplest techniques. The region sampled is divided into polygons by perpendicular bisectors between the sampling locations. In each polygon, all points are nearer to its enclosed sampling point than to any other sampling point (Webster and Oliver 2001). For example, Godt et al. (2006) used this technique to characterize rainfall for shallow landsliding in Seattle (USA). The IDW method is rather popular among deterministic spatial interpolation techniques. It is based on inverse functions of distance with the result that unknown locations to the sampling point carry larger weight than those further away. The advantage of weighting by inverse squared distance is the quick diminishing of the relative weights with increasing distance, making the interpolation sensibly local. However, the selection of the weighting function is arbitrary; also there is no indication of error (Webster and Oliver 2001). Chiang and Chang (2009), for example, used IDW to characterize the spatial rainfall distribution for modelling rainfall-induced landslides.

With respect to geostatistical approaches, kriging is the predominant method by connecting mathematical concepts with geoscientific requirements. Kriging is a generalized least-squares regression technique that offers accounting for the spatial dependence between observations (Schuurmans et al. 2007). A huge benefit that comes with this technique is the provision of a measure of certainty. In kriging, the unknown target variable is estimated using a weighted sum of the available point observations. The weights of the data are chosen so that the interpolation is unbiased and the variance is minimized (Ly et al. 2013). One important assumption is that the process under consideration is stationary. This allows researchers to assume that there is the same degree of variation from place to place and that the covariance between two observations only depends on the distance between these observations (Oliver and Webster 2014; Ly et al. 2013). For kriging applications, a graphical summary is used to analyse and understand spatial variation: the variogram (or more correctly: the semivariogram, as it depicts half the variance of the difference of the covariance, but for the sake of simplicity it is mostly called variogram).

The variogram plots variation as a function of distance. This means that rain gauges, for example, being in close proximity will have data values that are more similar to each other than rain gauges further away. Thus, the variogram is a plot of the average squared difference between pairs of data values and thus the central part of any subsequent kriging prediction. Just as the variance, the units of the variogram are the square of the units of measurement (Srivastava 2013). The plotted variogram based on the sampled data is called experimental variogram (Oliver and Webster 2014). There are three key characteristics of a variogram: (a) sill: the plateau that is reached by the variogram estimating the variance of the random process; (b) nugget: the y-intercept for the unresolved variation; and (c) range: distance at which the variogram reaches the sill (Srivastava 2013). To make the variogram applicable for the kriging prediction, a curve has to be fitted to the experimental variogram to neglect any inherent point-to-point erratic fluctuations. This fitted curve has a mathematical expression that describes the variance of random processes with changing distance and guarantees non-negative variances in the predictions (Oliver and Webster 2014). Commonly used correlation functions deriving the theoretical variogram are exponential, spherical, Matérn or Gaußian. In practice, there are different types of kriging. In the geoscientific literature, common prediction techniques are (a) ordinary kriging (OK), (b) kriging with external drift (KED), and (c) ordinary cokriging (OCK). Applied to spatial precipitation pattern, OK uses only rain gauge information, while the other two techniques incorporate sampled secondary information (e.g. weather radar information or elevation) to improve the kriging prediction (Goovaerts 1997). There is an abundance of related geostatistical literature available that readers are referred to for more in-depth information (e.g. Isaaks and Srivastava 1989; Cressie 1993; Goovaerts 1997; Webster and Oliver 2001).

3 Materials and methods

3.1 Obtaining real-time weather information from automated web scraping

This study uses web-based, hourly rain gauge data being obtained in an automated data workflow. Hourly time steps were selected due to the fact that longer observation periods would *average* rainfall intensity, which is detrimental for landslide early warning purposes due to the underestimation of peak (maximum) rainfall (Guzzetti et al. 2007). In this study, hourly time steps are considered as a compromise that still allows for creating spatially distributed rainfall information in near real-time and reflects the short-term rainfall intensity requirements for early warning applications, although shorter time steps would be even more desirable. In the past few years, environmental sciences are witnessing an increasing amount of published information on the Internet (Vitolo et al. 2015), which includes high-quality real-time weather data from various data sources. In many scientific fields, data obtained from the Internet are being used in a different way, ranging from simple data extraction tasks to fully automated data processing workflows. This growing demand leads many operators of databases and servers that exhibit a certain volume of traffic and where well-profiled usage expectations are available, to design publicly available application programming interfaces, so-called APIs. Data APIs are agreed-on programming interfaces providing a structure to download and link large chunks of heterogeneous data. Although such web services are the standard and recommended way to enable external access, such APIs are not always available, especially when the potential user base, and consequently the related demand on data, is minor. In such cases, web scraping is an alternative and valuable method for extracting and combining content from the Internet in a systematic way, even in the absence of an API (Glez-Peña et al. 2013).

Web scraping mimics the human user interaction with a website in a systematic way. The web scraping application is accessing as many websites as desired, parses its content to find and extract relevant information and structures it in a way ready to use for subsequent analyses. Although there are some desktop-based web scraping solutions (Glez-Peña et al. 2013 list a few), the most common approach is to use any suitable programming language to achieve maximum flexibility. There are many third-party, open source libraries available for implementing them in the source code for developing own web scraping applications. Figure 1 shows a rough overview of the entire web scraping process on how hourly rainfall data from various data sources is fetched, parsed and stored so that it can be retrieved for the generation of spatially distributed rainfall information. To establish the web scraping application to obtain hourly real-time weather data, the JavaScript based MeteorJS web framework was used in this study. It allows for rapid prototyping and the establishment of a server/client application. The scraped website access is executed by accessing the document object model (DOM) of the HTML content. This is done with Cheerio, which is based on the jQuery library to simplify client-side scripting. Cheerio (https://atmospherejs.com/fermuch/cheerio) itself is specifically designed for the Node.js runtime environment. With Meteor's *http*-package, the HTML content was loaded. To ease parsing, the Node.js module *html-parser2* (https://www.npmjs.com/package/htmlparser2) was used. To provide the extracted data also for later usage and not just real-time applications, the obtained information is stored in Meteor's default database (MongoDB) in a structured way (Fig. 2). For automated, time-based scheduling of the scraping process, the *synced-cron*-package (https://atmospherejs.com/percolate/synced-cron) was used. To access the data directly from the database in the source code of our subsequent analyses, the *restivus*-package (https://atmospherejs.com/nimble/restivus) was used to create RESTful APIs. The MongoDB used can be classified as a NoSQL database that stores data as JSON-like (JavaScript Object Notation) documents which makes it a fast and far more flexible way to store scraped data.

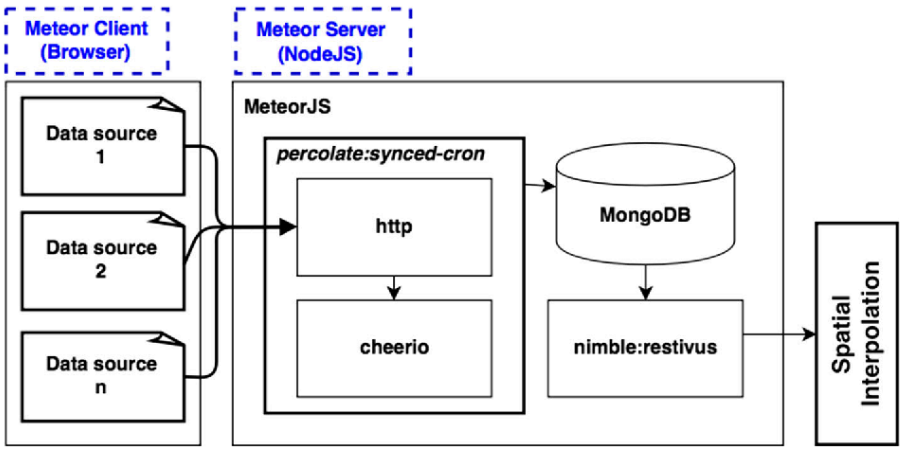


Fig. 1 Flow chart of the implemented web scraping application to fetch, parse and store real-time rainfall data for subsequent analyses

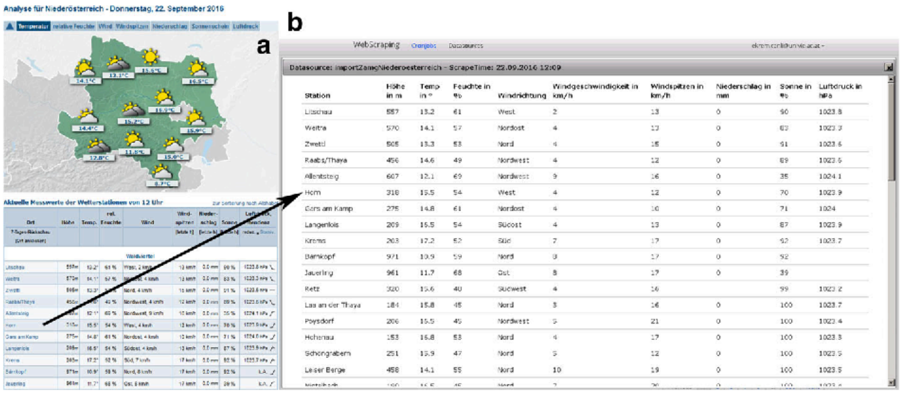


Fig. 2 **a** One of the websites providing hourly weather related data (*Source*: ZAMG); **b** corresponding MongoDB entry from the scraped web data

3.2 Automated spatial interpolation

There are always reliability concerns with respect to web-based data (de Vos et al. 2016). Although no crowdsourced weather data are included in this study at present, there is still the possibility of erroneous or bogus data involved. The quality issues of rain gauge data come from remotely, automated data sampling and the electronic transmission of the data through several ports before being used in an application (Kondragunta and Shrestha 2006). Therefore, real-time rain gauge quality control (QC) is necessary to detect major inconsistencies. Additionally, single-station QC checks alone are not sufficient due to the high spatial variability of rainfall; thus, adjacent stations have to be taken into consideration too. For this purpose, three automated QC/plausibility checks were implemented in this study: (a) a range filter; (b) a spatial consistency filter; and (c) an autocorrelation filter (for the kriging application only). The *range filter* is a simple check performed on a single observation for a given location for a specific time. Rain gauges with hourly rainfall intensities below 0 mm (physically impossible) or above 25 mm (very unlikely in that area) are excluded for subsequent analyses. The *spatial consistency filter* uses a distance matrix to calculate the 95th percentile of all available rain gauge intensities within 20 km (Fig. 3a). This distance was chosen to cover a substantial amount of neighbouring rain gauges and due to pronounced spatial autocorrelation that was shown in the variogram within this distance. If a rain gauge contains an intensity greater than the 95th percentile, it is discarded for subsequent analyses, but only if the intensity is higher than 15 mm. This threshold was chosen due to the fact that predominantly high rainfall intensities are important for landslide early warning purposes while the severe weather centre in Austria (Unwetterzentrale) defines heavy rainfall from 17 mm. Therefore, some tolerance was added to that boundary condition. The spatial consistency filter does not serve purely as a plausibility check, but also as a means to deal with outliers for the geostatistical interpolation, although we cannot know whether an unexpectedly large value is a real outlier (i.e. punctual high-intensity rainfall) or not. Outliers cause serious distortions in a variogram and should be removed if they are suspected to belong to a process other than the one interested (Oliver and Webster 2014). Therefore, we exclude such outliers from the data set to model the spatial relationship between the rain gauges, but use the entire data set

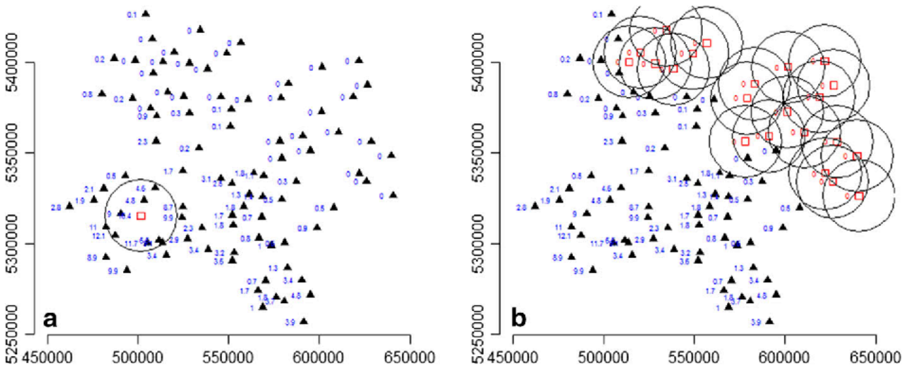


Fig. 3 **a** The spatial consistency filter removes rain gauges from an hourly data set that contain suspiciously high rainfall intensities compared to surrounding rain gauges; **b** the autocorrelation filter removes rain gauges from an hourly data set that contain zero rainfall and have only rain gauges with zero rainfall in their vicinity. Rain gauges that have zero rainfall but contain rain gauges within its vicinity that exhibit rainfall, are kept due to very apparent spatial dependency. The x and y axes refer to projected UTM coordinates. *Triangles* indicate retained rain gauges, *squares* indicate eliminated rain gauges, *large circles* indicate search radius for the distance matrix, and *numbers* refer to rainfall amount (mm h^{-1})

for the kriging prediction. This way, the variogram modelling is not corrupted by outliers, but the kriging surface still accounts for the extreme values.

The last implementation is the *autocorrelation filter*. When examining the distribution of environmental parameters, it often tends to be positively skewed towards the smaller values. This is also the case with hourly rainfall where some regions contain a certain amount of rainfall and other regions do not receive any rainfall at all. So there is a strong clustering of natural zero values in the distribution which makes transformation difficult. The comparison between means of observations is more unreliable because the variances are likely to differ from one set of data to another (Webster and Oliver 2001). Therefore, non-transformed data are used but isolated zero values are treated with the autocorrelation filter with regard to variogram modelling.

Schuurmans et al. (2007) also point out some problems arising with zero rainfall for kriging estimates. Geostatistics is based on the premise of autocorrelation. In an area where rainfall is recorded, there is a certain autocorrelation between the rain gauges. In an area without rainfall, zero values are negligible for the purpose of this study. The boundary between an area containing rainfall and an area without rainfall is, however, still relevant because stations close together are still autocorrelated to a rather high degree. The autocorrelation filter has, consequently, three conditions that must be satisfied: (a) keep all stations that have at least 0.1 mm rainfall; (b) remove all stations that have 0 mm rainfall and only have stations with 0 mm rainfall within distance; and (c) keep all stations that have 0 mm rainfall, but have stations with at least 0.1 mm rainfall within distance (Fig. 3b). As with the spatial consistency filter, the search distance equals 20 km. Additionally, the filtered data set is used for the variogram modelling, but all sample points are used for the kriging prediction. Naturally, this automated filter approach does not work in all situations, especially when the amount of remaining rain gauges is significantly reduced (Webster and Oliver (1992) point out the effect of sample size on variogram estimation). Therefore, a set of data with and without this filter is used in subsequent analyses and omitted on poorer validation results. Large areas with no rainfall that had their rain gauges

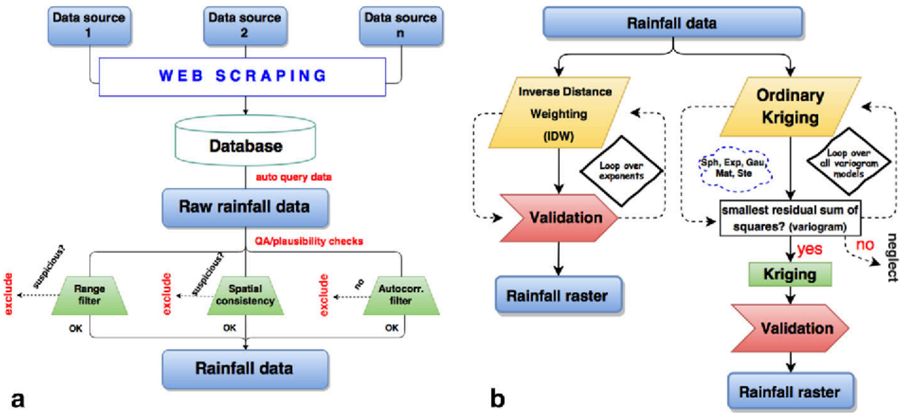


Fig. 4 Flow chart showing the automated workflow from **a** web-based data generation, the application of various filters to the raw rainfall data to eventually exclude single rain gauges, through **b** the interpolation procedures for producing hourly real-time rainfall raster maps

removed this way have a significantly higher variance consequently. Figure 4a shows an overview of all filters that are applied to the raw web-based rainfall data.

3.2.1 Geostatistical methods

Exploratory data analysis was the initial step to assess how representative and consistent the available rain gauges are distributed in the study area. For geostatistical analyses, areas with a high point density provide more reliable estimates at unsampled locations than areas with only a few rain gauges. To assess whether the average distance from an arbitrary point to the next nearest sample point is significantly short, testing for *complete spatial randomness* (CSR) might give an indication (Diggle 2003). The function behind estimating CSR is a nearest neighbour distance distribution function $G(r)$. Testing for CSR covers the horizontal domain. To check whether there are significant differences in distribution with respect to elevation, an overlay sampling of the rain gauge locations was performed on a digital elevation model and tested with a nonparametric KS test. Kriging uses the semi-variances based on a fitted variogram function. Creating a theoretical variogram is usually a labour-intensive task as many choices must be made (such as finding suitable range, sill, nugget values). There are attempts to automate the procedure, yet it carries risks when fitting a variogram without surveillance. Cressie (1985), for example, suggests a specific variogram where the weighted sum of squares between experimental and theoretical variogram becomes a minimum. However, the sensitivity of the variogram estimation on the interpolation itself is often not quite high. On the other hand, if a precise estimation of the error variance is needed (e.g. for uncertainty assessment), a solid variogram estimation is required (Haberlandt 2011). Therefore, the focus of this study lies in the rapid estimation of a suitable variogram for real-time applications.

The initial step for the variogram modelling is importing the georeferenced rain gauge data (filtered and unfiltered) from the web scraping application which happens every hour as soon as up-to-date information is available. To proceed, usually the variogram parameters range, sill and nugget should be estimated by exploratory analysis or expert knowledge. To automate this procedure, the decision is made to assign some plausible initial values. Therefore, the initial range is defined as 0.1 time the diagonal of the input shapefile that defines the boundaries of the study area (the bounding box). The initial sill is calculated as the mean of the maximum and median of the semivariance; the initial nugget is defined as the minimum of the semivariance. The next step uses a loop to iterate over multiple permitted variogram models (Spherical, Matérn incl. M. Stein's parameterization, Gaussian, etc.) to select the model with the smallest residual sum of squares (Oliver and Webster 2014). Additionally, all kappa values are tested for the Matérn model. This procedure with its associated values is based on the *automap*-package in R (Hiemstra et al. 2009). Omnidirectional variograms are created for both filtered and unfiltered input data. In case there is uniformly no rainfall at all or just very little in just a few sampling locations, the variogram is basically a horizontal line, indicating a pure nugget variogram due to the fact that there is no variance between the samples. The prediction in that case would be everywhere the same and thus the mean of the data (Oliver and Webster 2014). It would be difficult to interpret a pure nugget estimate, but the physical justification was given more weight: uniformly zero, or almost zero, rainfall does not matter too much for the underlying research question, which is more interested in larger rainfall intensities for landslide applications. And those are never uniformly distributed.

Here, a univariate approach is applied for spatial prediction, namely ordinary kriging (OK). Adding auxiliary variables as additional predictors might be helpful in some cases, but Ly et al. (2011) have shown, that adding elevation does not improve interpolation accuracy for short time intervals. Also Haberlandt (2007) found that elevation information plays a minor role for hourly rainfall data. When incorporating additional predictors to multivariate approaches, the benefit might be marginal if correlations become too small, as concluded by Goovaerts (2000). Figure 4b shows a flow chart of the automated kriging procedure used in this study. Using OK, hourly rainfall was estimated at unsampled locations on a 1-km square grid across the study area. The 1-km grid spacing used here was set for consistency and subsequent comparison with the radar data that contains the same grid size. Also according to Hengl (2006), who provided some empirical and analytical rules for the selection of suitable grid sizes, the selection of a 1-km grid size is justified based on the amount of rain gauges available. Kriging estimates might take negative values when negative kriging weights are applied. This is undesirable because this can lead to non-physical estimates. Possible solutions to avoid negative estimates are either a posteriori corrections of the kriging weights, as suggested by Deutsch (1996), or simply replacing all negative values with zero (Ly et al. 2013). In this study, the latter approach was used to achieve a physically sound rainfall estimation.

3.2.2 Deterministic methods

With regard to deterministic estimation procedures, the inverse distance weighting (IDW) and Thiessen polygon methods are used. The IDW method gives each rain gauge a weight that is inversely proportional to the distance between that rain gauge and an unknown sample point. Critical user input to this method is a distance parameter (an exponent) that controls the degree of dependence to a rain gauge in closest proximity (Srivastava 2013). A smaller value gives rain gauges further away higher importance, while accordingly larger exponent values assign higher importance to closer rain gauges. IDW is not capable of providing any quantitative indicator of reliability; therefore, an iterative approach was used to test multiple exponent values between 2 and 5, which was considered a plausible range for avoiding biased estimates. The distance exponent with the highest goodness of fit, in terms of the highest coefficient of determination in the validation process, is ultimately used for the final IDW based rainfall estimate (Fig. 4b). Another deterministic method used to compare kriging and IDW estimates is the Thiessen polygon method. This method simply divides the study area into polygons by perpendicular bisectors between the rain gauge locations. Within a polygon, all unknown points are closer to its enclosed rain gauge than to any other rain gauge (Webster and Oliver 2001).

3.3 Radar rainfall estimates

In 1965, Austro Control (the Austrian air navigation services provider) started operational service of its first weather radar at the airport Wien-Schwechat. Since then, many improvements have been made until 2011, when the first dual-polarization ground radar was installed. Those new-generation systems enable the transmission of radio signals with both horizontal and vertical polarization, while the conventional Doppler systems only transmitted and received radio waves with single horizontal polarization (Harpold et al. 2017). Since 2001, radar data in Austria are available with a spatial resolution of $1 \times 1 \times 1$ km and a temporal resolution of 5 min. The operational weather radar network in Austria consists of five stations, each with a range of 224 km covering the entire territory of Austria (Kaltenböck 2012). The used 2d-weather radar composites contain 14 quantization steps given in reflectivity (dBZ). Each radar composite has a spatial resolution of 1×1 km and contains data from a 5-min scan. Consequently, the 5-min rainfall reflectivity was summed by hour and divided by 12 for the hourly average to match the information from the rain gauges. To convert the reflectivity Z to rainfall intensity R (in mm h^{-1}), the empirical Marshall–Palmer equation was used with the relation $Z = 200R^{1.6}$ (Lovejoy et al. 2008). For this study, the radar data remained uncalibrated; thus, the converted rainfall rates from the radar data cannot be directly compared with the rainfall rates from the interpolated rain gauge data, but are used as a qualitative means of validation.

4 Study area

The study area includes the entire federal state of Lower Austria (Niederösterreich) in the north-eastern part of Austria. The size of the study area is 7408 km^2. The mean annual precipitation rates in Lower Austria for the period 2001–2010 show a gradient from lower rates in the northeast (approximately 500 mm) to higher rates in the southwest (approximately 1600–1700 mm) (Petschko et al. 2015). Schweigl and Hervás (2009) and Schwenk (1992) mention exceptional rainfall and/or snow melt as main triggers for landsliding in Lower Austria. A recently compiled landslide inventory for Lower Austria based on 1-m LiDAR DTM derivatives and orthofoto mapping reveals 13,166 landslides (Petschko et al. 2015).

The Cretaceous–Early Tertiary Rhenodanubian Flyschzone (RDF) contains approximately 6300 mapped landslides but contributes only to 14% of the territory of Lower Austria. Although there has been extensive work on statistical landslide susceptibility assessment in Lower Austria (e.g. Petschko et al. 2014; Steger et al. 2015, 2017), there are no published rainfall thresholds or process-based modelling approaches available.

To test the proposed methodology, a rainfall event from June 2009 that triggered many landslides in the southern parts of Lower Austria was selected. In total, 92 rain gauges were available to apply different spatial interpolation techniques for the automated generation of

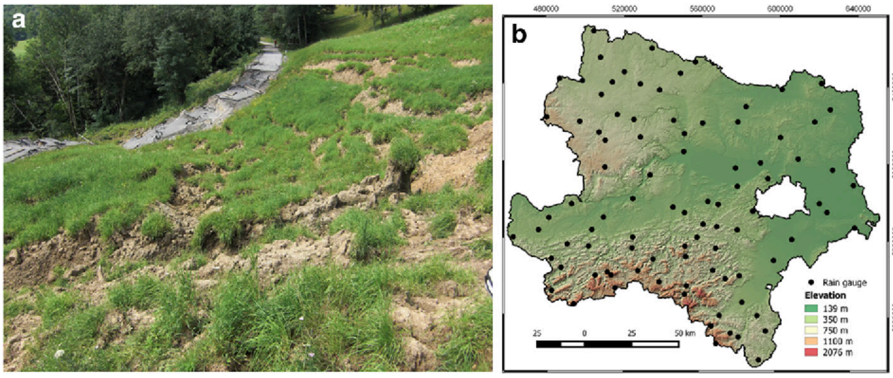


Fig. 5 **a** Destroyed infrastructure near Stössing (Lower Austria) caused by the June 2009 rainfall event (Image: Bertsch); **b** elevation map of Lower Austria; *dots* indicate rain gauge locations used in this study (DEM: CC BY 3.0 AT—Federal state of Lower Austria)

real-time continuous rainfall fields (Fig. 5). All rain gauges used in this study are contained in officially operated, automatic weather station networks from different weather service providers.

5 Results

5.1 Variogram modelling

Exploratory data analysis was performed as a first step to check how the rain gauges are spatially distributed in the study area with respect to spacing and elevation. Figure 6a shows the CSR plot of the empirical function $\hat{G}(r)$ against the theoretical expectations $G(r)$, indicating whether the average distance from an arbitrary point to the next nearest sample point is significantly short. The upper and lower simulation envelopes are calculated based on 100 simulations and indicate significance bands (the number of simulations was chosen arbitrarily but was found justified as no significant changes are expected

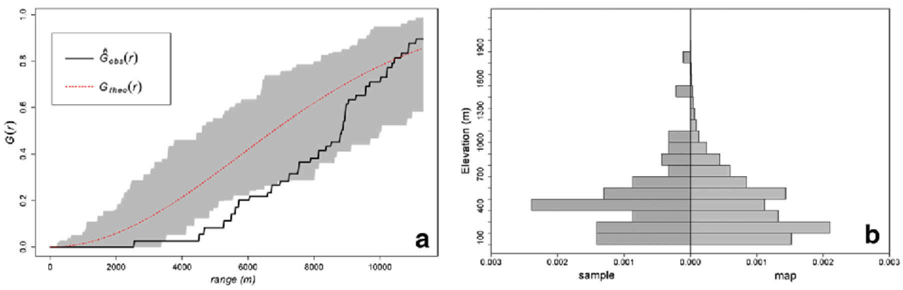


Fig. 6 **a** Testing for complete spatial randomness (CSR) to assess the spatial representativeness of the rain gauge distribution ("sampling design") with respect to geographical space. $\hat{G}(r)$ represents the empirical function (*continuous line*), $G(r)$ the theoretical expectation (*dashed line*) within its significance bands ("envelopes"); **b** relative frequency distribution of rain gauges with respect to overall distribution of elevation in the study area

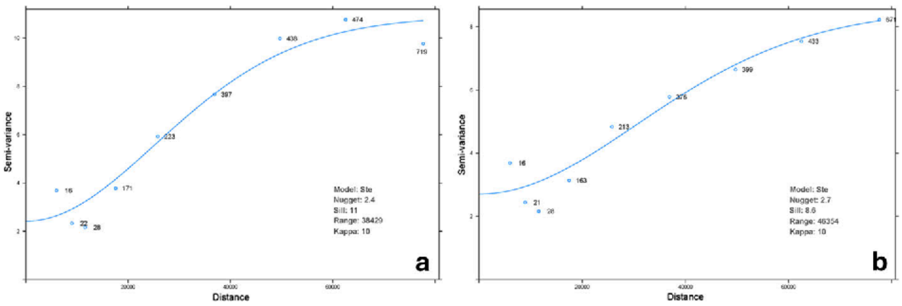


Fig. 7 Exponential variograms for the same time but without data filtering (**a**) and with filtering (**b**). In this case, the filtering leads to a slightly decreased sill which indicates lower variance of the residuals at greater distances

beyond that). Given the fact that there was no sample design involved in the selection of the rain gauge locations and that the average spacing between the rain gauges is around 10 km, the distribution of the sample locations is quite representative with respect to geographical space. With respect to elevation, using a nonparametric KS test showed that the distribution of rain gauges is not significant with respect to the overall distribution of the elevation in the study area (Fig. 6b). Especially higher elevations and elevations in the 400–500 m range are overrepresented by rain gauges.

Variance in every rainfall event differs; thus, also each variogram is different. From the automated fitting procedure, it can be concluded that generally there is quite a large variability in fitted ranges across hourly events. On a multitude of variograms, however, the range is between 30 and 50 km. But there remain many variograms with a very large range leading to spurious autocorrelations, probably caused by large-scale trends extending throughout the study area. From the iterative model fitting procedure, the most common applied model was the Matérn model with Stein's parametrization (resulting from iteratively comparing the smallest residual sum of squares). In some cases, the filtering procedure also leads to a reduction in sill (Fig. 7).

5.2 Real-time rainfall interpolation

To demonstrate the different interpolation techniques described in Methods, three consecutive hours of a frontal rainfall event were selected that caused landslides in the southern parts of Lower Austria. 92 rain gauges served as the basis for the spatial interpolation. Figure 8 shows a composite of those 3 h for all rainfall predictions. All predictions were performed on a 1×1 km grid. All predicted rainfall fields are in good visual accordance; however, the Thiessen polygons (Fig. 8e) are considered to grant no real advantages over the conventional *representative rain gauge* approach to characterizing spatially distributed rainfall for a certain area as the transition between the arbitrary polygon boundaries is unrealistically rough. This is even more extreme for very short time intervals and in mountainous regions where recorded rainfall intensities may vary significantly within short distances. Also in just those 3 h, two problems associated with the kriging technique are apparent. When considering the point information from the rain gauges (Fig. 8a), rainfall in areas with high intensities is much lower for the kriging estimates (Fig. 8b, c). Thus, kriging loses variance by smoothing, yet it gives the best estimates from a statistical point of view. The second issue with kriging is the presence of

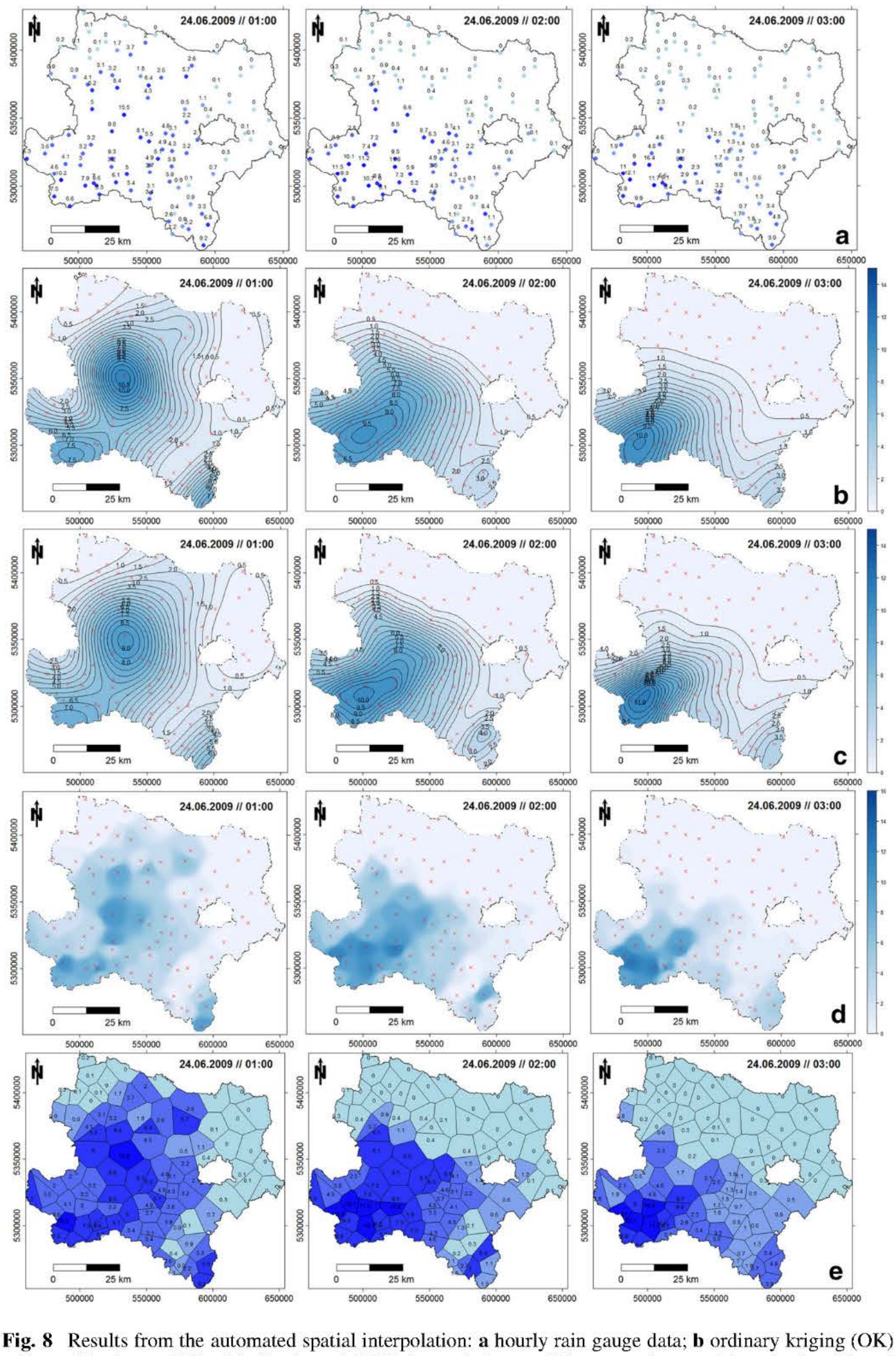


Fig. 8 Results from the automated spatial interpolation: **a** hourly rain gauge data; **b** ordinary kriging (OK) without filtering; **c** OK with filtering; **d** IDW interpolation; **e** Thiessen polygons. Lines in the (**b**) and (**c**) estimates indicate areas with equal amounts of rainfall (isohyets). *Points* in the maps (**b**) through (**d**) indicate rain gauge locations

punctually very high intensities. At 01:00 24 June, 2009 (left column in Fig. 8b), a single high-intensity rain gauge is present that leads to a strange, bullseye-like structure. This rain gauge was removed with the automated filtering (Fig. 8c), still it seems that a spatially limited high-intensity rain cell causes a big influence in this part of the study area that cannot be properly resolved by the automated variogram modelling.

5.3 Performance comparison for hourly interpolation

To estimate the performance of the OK, filtered OK and IDW predictions, two different types of validation were performed. The leave-one-out cross-validation (looCV) removes one rain gauge at a time and recalculates its value from the remaining data. Validation was also performed by splitting the sample size randomly into a training (80%) and a test (20%) subset. The training subset was then used to predict the values of the test subset. For the three consecutive hours shown here, the validation results are presented in Fig. 9a. The problems in the kriging estimates for the local high-intensity rain cell at 01:00 24 June, 2009, are also reflected in the validation results showing much lower coefficients of determination (around 0.65). The next 2 h produced more balanced kriging estimates resulting also in better validation results (around 0.8). For those 2 h, the application of the different filters (range, spatial consistency and autocorrelation filter) also leads to a slight increase in performance. For every rainfall prediction, a standardized residual plot is generated which measures the strength of the difference between observed and predicted values (Fig. 9c). The residuals from the automated rainfall prediction tend to be symmetrically distributed (homoscedastic) and there are no clear patterns in general, which indicates that the automated model prediction is feasible.

Many studies indicate the root-mean-square error (RMSE) as a performance indicator. We refrain from this practice as it does not give justice to the spatially differentiated variances involved in kriging estimates. We found a spatially distributed representation of kriging variances (the *estimation error*) to be a more suitable tool (Fig. 9b). This map also reveals much larger variances near the boundaries of the study area due to the reduced

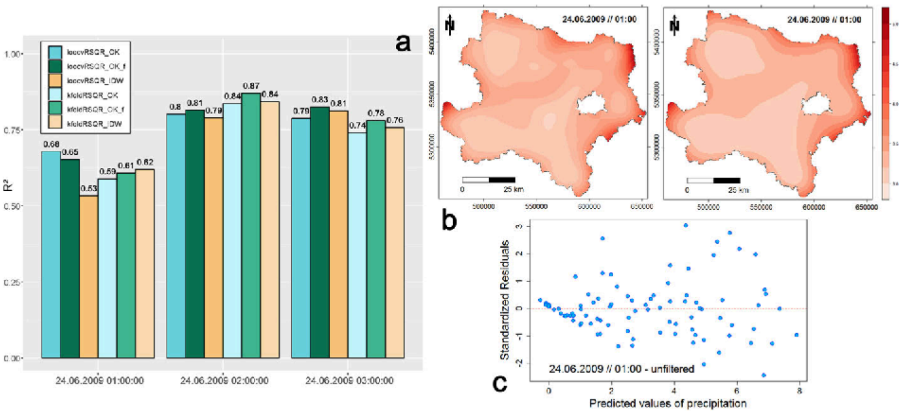


Fig. 9 **a** Validation results (coefficient of determination) for the ordinary kriging (OK) and IDW estimates. (looCV = leave-one-out cross-validation; subsample = sub-sampling into training and test subset; OK_f = OK filtered); **b** kriging variances for an unfiltered (*left*) and filtered (*right*) prediction serving as a spatially distributed error estimation; **c** automatically generated standardized residual plot indicating constant variance

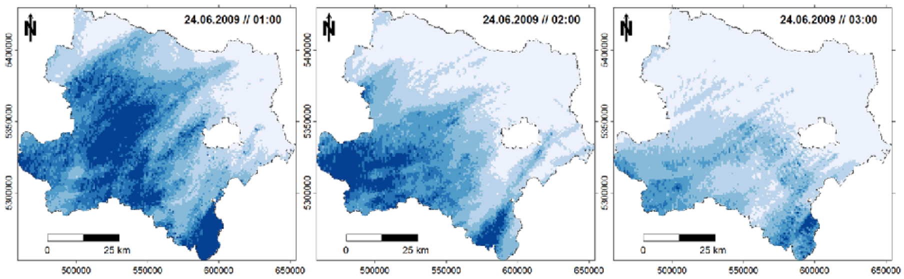


Fig. 10 Uncalibrated radar data for three consecutive hours used for qualitative validation (hence the deliberately omitted legend). The spatial pattern is in good visual accordance with the interpolation rainfall estimates shown in Fig. 8

number of rain gauges in close proximity. Additionally to this quantitative validation, a qualitative validation was performed based on the visual comparison with radar rainfall data. To use the radar imagery for a quantitative validation or as an auxiliary variable for multivariate kriging approaches, it needs to be calibrated first. This would be highly desirable for an additional means of quantitative validation because it provides an independent data set from the interpolation results. Comparing the radar data from Fig. 10 with the interpolated rainfall predictions from Fig. 8, the overall picture shows a rather good match. However, the radar technique is capable of capturing more fine structured, intermittent rainfall fields.

6 Discussion

This study suggests a fully automated workflow from the hourly, web-based collection of rain gauge data to the generation of spatially differentiated rainfall predictions based on deterministic and geostatistical methods. The underlying research question envisages the implementation of those hourly rainfall predictions into a dynamic, combined hydrological and slope stability modelling application with the purpose of estimating landslide failure probabilities in near real time. The entire methodology was implemented solely with open source technology (JavaScript, R, Python, QGIS).

When web data are used, legal and policy issues have to be considered. Legal implications with web data are not always clear in all cases and countries; however, the terms of use should always be respected to limit the permitted data requests within a certain amount of time or, in general, to prevent copyright infringement (Glez-Peña et al. 2013). The usage of meteorological data is better regulated. The Twelfth World Meteorological Congress in Geneva in June 1995 approved a resolution on the international exchange of meteorological data and products (WMO 1996). This resolution stipulates the member states of the World Meteorological Organization the right to use data and products at no costs for non-commercial use. When using web scraping for data generation instead of an API request, a problem might arise in the way how the data are automatically parsed from a website. As soon as the structure of an HTML document is changed, the web scraping application does not work anymore and has to be readjusted, thus requiring constant maintenance. Using these data for geostatistical analyses is usually a very labour-intensive work.

Modelling the variogram is critical for the quality of the kriging estimates; thus, automating this procedure is not so straightforward and required some simplifications in terms of defining initial modelling parameters (i.e. sill, range, nugget). Therefore, a stronger emphasis was put on creating a feasible workflow that produced good quality rainfall predictions for further use. Also worth mentioning is that anisotropy was not considered in this study, although rain gauges recording frontal rainfall are likely to be directionally dependent. However, detecting anisotropy in an automated workflow is not straightforward and could only be approached by iteratively calculating multiple directional variograms and perform kriging predictions for all of them accordingly while comparing their respective validation results. On the other hand, not all rainfall situations are directional (e.g. convective rainfall); therefore, omnidirectional variogram models were used in this study. Additionally, this suggested procedure is only feasible for rainfall and no other forms of precipitation, as only direct rain gauge readings are considered. This is acceptable for this study area due to highest rainfall intensities recorded in the summer months which is relevant for landslide initiation. When passing along those automatically generated rainfall predictions to a landslide modelling application, a solid quality indicator is required as a diagnostic tool beforehand in order to guarantee a good performing model. Or in other words: where to set the threshold between good and bad rainfall predictions. One may use the coefficient of determination or, as suggested by Oliver and Webster (2014), the mean-squared deviation ratio (MSDR) which is the mean of the squared errors divided by the corresponding kriging variances. Possible thresholds would be high coefficients of determination or MSDR values close to 1. However, this has to be extensively tested in the landslide application which one is the most suitable with respect to the underlying research question.

Another major issue commonly reported in other studies (e.g. Schuurmans et al. 2007; Kann et al. 2015) is the presence of small-scale, convective heavy rainfall events, mostly occurring in summer or spatially highly intermittent rainfall (e.g. Chappell et al. 2013). Thus, different types of rainstorms may provide different levels of performance for the proposed methodology. Simple ordinary kriging purely relying on distance-based rain gauge information might not be capable of solving this problem with a limited number of sampling locations alone. In that case, using multivariate kriging approaches that incorporate radar data as additional predictors might be more appropriate. Another interesting approach to mimicking true in situ variability on short-scale changes is the utilization of conditional simulation, or how Srivastava (2013) describes it: the spatial version of Monte Carlo procedures. Kriging usually produces a smoothed surface by losing variance and thus underestimating large values and overestimating small values. Also, kriging estimates produce only a single prediction. Conditional simulation, on the other hand, produces many equally likely scenarios (so-called realizations) by using the same variogram, but at the cost of accuracy. Consequently, this probabilistic procedure produces a more realistic picture of small-scale variations, but should be complemented by kriging estimates when spatial variability and error estimates are crucial for the underlying research question (Srivastava 2013).

With regard to regional landslide EWS, this study offers a direct integration for both threshold-based and process-based approaches. The methodology presented in this study is especially suitable for the implementation in warning systems (following the classification of Stähli et al. 2015) that contain predefined thresholds and are mainly used for processes with progressive stages of failure (e.g. rock slides, translational and rotational soil slides). Instead of using uniformly distributed rainfall for an entire region, spatially differentiated rainfall values can be used for a real-time comparison with previously established rainfall thresholds. Similarly, hourly rainfall predictions can also serve as the time-dependent input in process-based approaches that determine how changes in pore-water conditions alter slope stability conditions. Raia et al. (2014) presented a promising approach where the dynamic, infinite-slope-based model TRIGRS (Baum et al. 2008) was used for the forecasting of rainfall-induced shallow landslides over large regions. In their probabilistic modification of the model, however, they assumed constant rainfall intensity to force slope instability for the entire study area. Salciarini et al. (2017) followed a similar probabilistic approach. It would be interesting to observe how spatially distributed rainfall intensities will behave on slope stability conditions when compared to uniformly distributed rainfall input.

7 Outlook

A major issue in almost all natural sciences is data scarcity. Especially in landslide research proper event data are often lacking. With the rise of Web 2.0 applications, however, there was a large boost in collaborative and fast data acquisition initiatives. Olyazadeh et al. (2016) present a WebGIS Android App for fast data acquisition of landslide hazard, and Klonner et al. (2016) present a review of how volunteered geographical information is collected in natural hazard analysis. Baum et al. (2014) highlight the *Report a landslide* website operated by the USGS for engaging public in the identification of geological hazards. With respect to weather data, there are already much longer lasting initiatives in operation. Such citizen science initiatives have proven to be highly valuable for supplementing primary instrumented rain gauge networks (Harpold et al. 2017). The National Weather Service Cooperative Observer Program (COOP) in the USA was formally created in 1890 for volunteers to take observations. There are many websites that provide open weather data collected by weather enthusiasts that even offer APIs for direct data integration (e.g. http://openweathermap.com/ or http://wunderground.com/ with their personal weather station network). Those privately operated weather stations that are hosted by such online weather networks have explicit terms of service that facilitate API data usage.

Quality issues might be a big concern with such data; however, the benefit of offering a highly densified rain gauge network providing rainfall data in real-time is not to be underestimated and should clearly be addressed in the near future. This paper suggests an automated workflow that enables the quick integration of additional real-time rainfall data from multiple online sources with either an API or web scraping integration. Therefore, very dense rain gauge networks could be established for providing accurate spatially distributed rainfall predictions for the integration in regional landslide EWS. Future work will incorporate this approach in dynamic, grid-based regional slope stability analysis to evaluate in a real-word hindcast situation if and to what degree spatially distributed rainfall information in landslide research contribute to improving regional landslide EWS.

Acknowledgements Open access funding provided by the University of Vienna. The authors thank the Austro Control for providing radar rainfall data and the Provincial government of Lower Austria for their support. The comments and suggestions of two anonymous reviewers have greatly helped us in preparing this manuscript. We gratefully acknowledge this help.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Aleotti P (2004) A warning system for rainfall-induced shallow failures. Eng Geol 73:247–265. doi:10.1016/j.enggeo.2004.01.007

Anagnostou MN, Kalogiros J, Anagnostou EN et al (2010) Performance evaluation of high-resolution rainfall estimation by X-band dual-polarization radar for flash flood applications in mountainous basins. J Hydrol 394:4–16. doi:10.1016/j.jhydrol.2010.06.026

Bathurst JC, Burton A, Clarke BG, Gallart F (2006) Application of the SHETRAN basin-scale, landslide sediment yield model to the Llobregat basin, Spanish Pyrenees. Hydrol Process 20:3119–3138. doi:10.1002/hyp.6151

Baum RL, Godt JW (2010) Early warning of rainfall-induced shallow landslides and debris flows in the USA. Landslides 7:259–272. doi:10.1007/s10346-009-0177-0

Baum RL, Savage WZ, Godt JW (2008) TRIGRS—a Fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis, Version 2.0. US Geological Survey Open-file Report, vol 1159

Baum RL, Highland LM, Lyttle PT, Fee JM, Martinez EM, Wald LA (2014) “Report a Landslide” a website to engage the public in identifying geologic hazards. In: Sassa K, Canuti P, Yin Y (eds) Landslide science for a safer geoenvironment. Springer International Publishing, Cham, pp 95–100

Berndt C, Rabiei E, Haberlandt U (2014) Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios. J Hydrol 508:88–101. doi:10.1016/j.jhydrol.2013.10.028

Calvello M, d’Orsi RN, Piciullo L, Paes N, Magalhaes M, Alvarenga Lacerda W (2015) The Rio de Janeiro early warning system for rainfall-induced landslides: analysis of performance for the years 2010–2013. Int J Disaster Risk Reduct 12:3–15. doi:10.1016/j.ijdrr.2014.10.005

Capparelli G, Tiranti D (2010) Application of the MoniFLaIR early warning system for rainfall-induced landslides in Piedmont region (Italy). Landslides 7:401–410. doi:10.1007/s10346-009-0189-9

Casagli N, Catani F, Del Ventisette C, Luzi G (2010) Monitoring, prediction, and early warning using ground-based radar interferometry. Landslides 7:291–301. doi:10.1007/s10346-010-0215-y

Casty C, Wanner H, Luterbacher J, Esper J, Böhm R (2005) Temperature and precipitation variability in the European Alps since 1500. Int J Climatol 25:1855–1880. doi:10.1002/joc.1216

Chan RKS, Pang PLR, Pun WK (2003) Recent developments in the landslip warning system in Hong Kong. In: Ho KKS, Li KS (eds) Geotechnical engineering—Meeting Society’s Needs: Proceedings of the 14th Southeast Asian geotechnical conference, 10–14 Dec 2001. A.A. Balkema, Rotterdam, pp 219–224

Chappell A, Renzullo LJ, Raupach TH, Haylock M (2013) Evaluating geostatistical methods of blending satellite and gauge data to estimate near real-time daily rainfall for Australia. J Hydrol 493:105–114. doi:10.1016/j.jhydrol.2013.04.024

Chiang S-H, Chang K-T (2009) Application of radar data to modeling rainfall-induced landslides. Geomorphology 103:299–309. doi:10.1016/j.geomorph.2008.06.012

Chien L-K, Hsu C-F, Yin L-C (2015) Warning model for shallow landslides induced by extreme rainfall. Water 7:4362–4384. doi:10.3390/w7084362

Cloke HL, Pappenberger F (2009) Ensemble flood forecasting: a review. J Hydrol 375:613–626. doi:10.1016/j.jhydrol.2009.06.005

Collier CG, Larke PR, May BR (1983) A weather radar correction procedure for real-time estimation of surface rainfall. Q J R Meteorol Soc 109:589–608. doi:10.1002/qj.49710946110

Cressie N (1985) Fitting variogram models by weighted least squares. J Int Assoc Math Geol 17:563–586

Cressie N (1993) Statistics for spatial data, Rev edn. Wiley, New York

Crosta GB, Frattini P (2003) Distributed modelling of shallow landslides triggered by intense rainfall. Nat Hazards Earth Syst Sci 3:81–93. doi:10.5194/nhess-3-81-2003

de Vos L, Leijnse H, Overeem A, Uijlenhoet R (2016) Validation of crowdsourced automatic rain gauge measurements in Amsterdam. In: EGU General Assembly Conference Abstracts, p 11607

Deutsch CV (1996) Correcting for negative weights in ordinary kriging. Comput Geosci 22:765–773

Diggle PJ (2003) Statistical analysis of spatial point patterns, 2nd edn. Arnold, London

Gariano SL, Brunetti MT, Iovine G, Melillo M, Peruccacci S, Terranova O, Vennari C, Guzzetti F (2015) Calibration and validation of rainfall thresholds for shallow landslide forecasting in Sicily, Southern Italy. Geomorphology 228:653–665. doi:10.1016/j.geomorph.2014.10.019

Glade T, Nadim F (2014) Early warning systems for natural hazards and risks. Nat Hazards 70:1669–1671. doi:10.1007/s11069-013-1000-8

Glade T, Crozier M, Smith P (2000) Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical "Antecedent Daily Rainfall Model". Pure Appl Geophys 157:1059–1079. doi:10.1007/s000240050017

Glez-Peña D, Lourenço A, López-Fernández H, Reboiro-Jato M, Fdez-Riverola F (2013) Web scraping technologies in an API world. Brief Bioinform 15:788–797. doi:10.1093/bib/bbt026

Gobiet A, Kotlarski S, Beniston M, Heinrich G, Rajczak J, Stoffel M (2014) 21st century climate change in the European Alps—a review. Sci Total Environ 493:1138–1151. doi:10.1016/j.scitotenv.2013.07.050

Godt JW, Baum RL, Chleborad AF (2006) Rainfall characteristics for shallow landsliding in Seattle, Washington, USA. Earth Surf Proc Landf 31:97–110. doi:10.1002/esp.1237

Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press, New York

Goovaerts P (2000) Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J Hydrol 228:113–129. doi:10.1016/S0022-1694(00)00144-X

Günther A, Reichenbach P, Malet J-P, Van Den Eeckhaut M, Hervás J, Dashwood C, Guzzetti F (2013) Tier-based approaches for landslide susceptibility assessment in Europe. Landslides 10:529–546. doi:10.1007/s10346-012-0349-1

Guzzetti F, Carrara A, Cardinali M, Reichenbach P (1999) Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31:181–216. doi:10.1016/S0169-555X(99)00078-1

Guzzetti F, Cardinali M, Reichenbach P, Cipolla F, Sebastiani C, Galli M, Salvati P (2004) Landslides triggered by the 23 November 2000 rainfall event in the Imperia Province, Western Liguria, Italy. Eng Geol 73:229–245. doi:10.1016/j.enggeo.2004.01.006

Guzzetti F, Peruccacci S, Rossi M, Stark CP (2007) Rainfall thresholds for the initiation of landslides in central and southern Europe. Meteorol Atmos Phys 98:239–267. doi:10.1007/s00703-007-0262-7

Haberlandt U (2007) Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. J Hydrol 332:144–157. doi:10.1016/j.jhydrol.2006.06.028

Haberlandt U (2011) Interpolation of precipitation for flood modelling. In: Schumann AH (ed) Flood risk assessment and management. Springer, Dordrecht, pp 35–52

Harpold AA, Kaplan M, Klos PZ, Link T, McNamara JP, Rajagopal S, Schumer R, Steele CM (2017) Rain or snow: hydrologic processes, observations, prediction, and research needs. Hydrol Earth Syst Sci 21:1–22. doi:10.5194/hess-21-1-2017

Heil B, Petzold I, Romang H, Hess J (2014) The common information platform for natural hazards in Switzerland. Nat Hazards 70:1673–1687. doi:10.1007/s11069-010-9606-6

Hengl T (2006) Finding the right pixel size. Comput Geosci 32:1283–1298. doi:10.1016/j.cageo.2005.11.008

Hiemstra PH, Pebesma EJ, Twenhöfel CJW, Heuvelink GBM (2009) Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network. Comput Geosci 35:1711–1721. doi:10.1016/j.cageo.2008.10.011

Hong Y, Adler RF, Negri A, Huffman GJ (2007) Flood and landslide applications of near real-time satellite rainfall products. Nat Hazards 43:285–294. doi:10.1007/s11069-006-9106-x

Hou AY, Kakar RK, Neeck S, Azarbarzin AA, Kummerow CD, Kojima M, Oki R, Nakamura K, Iguchi T (2014) The global precipitation measurement mission. Bull Am Meteorol Soc 95:701–722. doi:10.1175/BAMS-D-13-00164.1

Huang J, Ju NP, Liao YJ, Liu DD (2015) Determination of rainfall thresholds for shallow landslides by a probabilistic and empirical method. Nat Hazards Earth Syst Sci 15:2715–2723. doi:10.5194/nhess-15-2715-2015

Huffman GJ, Adler RF, Bolvin DT, Gu G, Nelkin EJ, Bowman KP, Hong Y, Stocker EF, Wolff DB (2007) The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J Hydrometeorol 8:38–55. doi:10.1175/JHM560.1

Huffman GJ, Adler RF, Bolvin DT, Nelkin EJ (2010) The TRMM multi-satellite precipitation analysis (TMPA). In: Gebremichael M, Hossain F (eds) Satellite rainfall applications for surface hydrology. Springer, Dordrecht, pp 3–22

Intrieri E, Gigli G, Mugnai F, Fanti R, Casagli N (2012) Design and implementation of a landslide early warning system. Eng Geol 147–148:124–136. doi:10.1016/j.enggeo.2012.07.017

Isaaks EH, Srivastava RM (1989) Applied geostatistics. Oxford University Press, New York

Jaedicke C, Van Den Eeckhaut M, Nadim F, Hervás J, Kalsens B, Vangelsten BV, Smith JT, Tofani V, Ciurean R, Winter MG, Sverdrup-Thygeson K, Syre E, Smebye H (2014) Identification of landslide hazard and risk "hotspots" in Europe. Bull Eng Geol Environ 73:325–339. doi:10.1007/s10064-013-0541-0

Jasper K, Gurtz J, Lang H (2002) Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model. J Hydrol 267(1–2):40–52. doi:10.1016/S0022-1694(02)00138-5

Jewell SA, Gaussiat N (2015) An assessment of kriging-based rain-gauge-radar merging techniques: gauge-radar merging schemes for flood forecasting. Q J R Meteorol Soc 141:2300–2313. doi:10.1002/qj.2522

Joyce RJ, Janowiak JE, Arkin PA, Xie P (2004) CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J Hydrometeorol 5:487–503. doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Kaltenböck R (2012) New generation of dual polarized weather radars in Austria. In: Extended abstracts of the 7th European conference on radar in meteorology and hydrology, 25–29 June 2012, Toulouse, France

Kann A, Meirold-Mautner I, Schmid F, Kirchengast G, Fuchsberger J, Meyer V, Tüchler L, Bica B (2015) Evaluation of high-resolution precipitation analyses using a dense station network. Hydrol Earth Syst Sci 19:1547–1559. doi:10.5194/hess-19-1547-2015

Keefer DK, Wilson RC, Mark RK, Brabb EE, Brown WM III, Ellen SD, Harp EL, Wieczorek GF, Alger CS, Zatkin RS (1987) Real-time landslide warning during heavy rainfall. Science 238:921–925. doi:10.1126/science.238.4829.921

Kirschbaum DB, Patel KG (2016) Precipitation data key to food security and public health. EOS. doi:10.1029/2016EO045583

Kirschbaum DB, Stanley T, Simmons J (2015) A dynamic landslide hazard assessment system for Central America and Hispaniola. Nat Hazards Earth Syst Sci 15:2257–2272. doi:10.5194/nhess-15-2257-2015

Klonner C, Marx S, Usón T, Porto de Albuquerque J, Höfle B (2016) Volunteered geographic information in natural hazard analysis: a systematic literature review of current approaches with a focus on preparedness and mitigation. ISPRS Int J Geo-Inf 5:103. doi:10.3390/ijgi5070103

Kondragunta CR, Shrestha K (2006) Automated real-time operational rain gauge quality-control tools in NWS hydrologic operations. In: Preprints, 20th AMS conference on hydrology, 29 Jan–2 Feb 2006, Atlanta

Kubota T, Shige S, Hashizume H, Aonashi K, Takahashi N, Seto S, Hirose M, Takayubu YN, Ushio T, Nakagawa K, Iwanami K, Kachi M, Okamoto K (2007) Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: production and validation. IEEE Trans Geosci Remote Sens 45:2259–2275. doi:10.1109/TGRS.2007.895337

Lagomarsino D, Segoni S, Fanti R, Catani F (2013) Updating and tuning a regional-scale landslide early warning system. Landslides 10:91–97. doi:10.1007/s10346-012-0376-y

Lang TJ, Nesbitt SW, Carey LD (2009) On the correction of partial beam blockage in polarimetric radar data. J Atmos Ocean Technol 26:943–957. doi:10.1175/2008JTECHA1133.1

Lombardo L, Cama M, Maerker M, Rotigliano E (2014) A test of transferability for landslides susceptibility models under extreme climatic events: application to the Messina 2009 disaster. Nat Hazards 74:1951–1989. doi:10.1007/s11069-014-1285-2

Lovejoy S, Schertzer D, Allaire VC (2008) The remarkable wide range spatial scaling of TRMM precipitation. Atmos Res 90:10–32. doi:10.1016/j.atmosres.2008.02.016

Ly S, Charles C, Degré A (2011) Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium. Hydrol Earth Syst Sci 15:2259–2274. doi:10.5194/hess-15-2259-2011

Ly S, Charles C, Degré A (2013) Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: a review. Base 17:392–406

Mair A, Fares A (2011) Comparison of rainfall interpolation methods in a mountainous region of a tropical island. J Hydrol Eng 16:371–383. doi:10.1061/(ASCE)HE.1943-5584.0000330

Montrasio L, Valentino R, Corina A, Rossi L, Rudari R (2014) A prototype system for space–time assessment of rainfall-induced shallow landslides in Italy. Nat Hazards 74:1263–1290. doi:10.1007/s11069-014-1239-8

Oliver MA, Webster R (2014) A tutorial guide to geostatistics: computing and modelling variograms and kriging. CATENA 113:56–69. doi:10.1016/j.catena.2013.09.006

Olyazadeh R, Sudmeier-Rieux K, Jaboyedoff M, Derron M-H, Devkota S (2016) An offline–online WebGIS android application for fast data acquisition of landslide hazard and risk. Nat Hazards Earth Syst Sci Discuss. doi:10.5194/nhess-2016-267

Petschko H, Brenning A, Bell R, Goetz J, Glade T (2014) Assessing the quality of landslide susceptibility maps—case study Lower Austria. Nat Hazards Earth Syst Sci 14:95–118. doi:10.5194/nhess-14-95-2014

Petschko H, Bell R, Glade T (2015) Effectiveness of visually analyzing LiDAR DTM derivatives for earth and debris slide inventory mapping for statistical susceptibility modeling. Landslides 13:857–872. doi:10.1007/s10346-015-0622-1

Raia S, Alvioli M, Rossi M, Baum RL, Godt JW, Guzzetti F (2014) Improving predictive power of physically based rainfall-induced shallow landslide models: a probabilistic approach. Geosci Model Dev 7:495–514. doi:10.5194/gmd-7-495-2014

Reichenbach P, Cardinali M, De Vita P, Guzzetti F (1998) Regional hydrological thresholds for landslides and floods in the Tiber River Basin (central Italy). Environ Geol 35:146–159. doi:10.1007/s002540050301

Rosi A, Lagomarsino D, Rossi G, Segoni S, Battistini A, Casagli N (2015) Updating EWS rainfall thresholds for the triggering of landslides. Nat Hazards 78:297–308. doi:10.1007/s11069-015-1717-7

Rossi M, Kirschbaum DB, Luciani S, Mondini AC, Guzzetti F (2012) TRMM satellite rainfall estimates for landslide early warning in Italy: preliminary results. In: Hayasaka T, Nakamura K, Im E (eds) Proceedings of SPEE. Kyoto, Japan, p 85230D

Salciarini D, Fanelli G, Tamagnini C (2017) A probabilistic model for rainfall—induced shallow landslide prediction at the regional scale. Landslides. doi:10.1007/s10346-017-0812-0

Schiemann R, Erdin R, Willi M, Frei C, Berenguer M, Sempere-Torres D (2011) Geostatistical radar-raingauge combination with nonparametric correlograms: methodological considerations and application in Switzerland. Hydrol Earth Syst Sci 15:1515–1536. doi:10.5194/hess-15-1515-2011

Schmidt J, Turek G, Clark MP, Uddstrom M, Dymond JR (2008) Probabilistic forecasting of shallow, rainfall-triggered landslides using real-time numerical weather predictions. Nat Hazards Earth Syst Sci 8:349–357

Schuurmans JM, Bierkens MFP, Pebesma EJ, Uijlenhoet R (2007) Automatic prediction of high-resolution daily rainfall fields for multiple extents: the potential of operational radar. J Hydrometeorol 8:1204–1224. doi:10.1175/2007JHM792.1

Schwab M, Rieke-Zapp D, Schneider H, Liniger M, Schlunegger F (2008) Landsliding and sediment flux in the Central Swiss Alps: a photogrammetric study of the Schimbrig landslide, Entlebuch. Geomorphology 97:392–406. doi:10.1016/j.geomorph.2007.08.019

Schweigl J, Hervás J (2009) Landslide mapping in Austria. European Commission Joint Research Centre, Institute for Environment and Sustainability, Italy

Schwenk H (1992) Massenbewegungen in Niederösterreich 1953–1990. Jb Geol B-A 135:597–660

Segond M-L, Wheater HS, Onof C (2007) The significance of spatial rainfall representation for flood runoff estimation: a numerical evaluation based on the Lee catchment, UK. J Hydrol 347:116–131. doi:10.1016/j.jhydrol.2007.09.040

Segoni S, Leoni L, Benedetti AI, Catani F, Righini G, Falorni G, Gabellani S, Rudari R, Silvestro F, Rebora N (2009) Towards a definition of a real-time forecasting network for rainfall induced shallow landslides. Nat Hazards Earth Syst Sci 9:2119–2133. doi:10.5194/nhess-9-2119-2009

Segoni S, Lagomarsino D, Fanti R, Moretti S, Cassagli N (2015) Integration of rainfall thresholds and susceptibility maps in the Emilia Romagna (Italy) regional-scale landslide warning system. Landslides 12:773–785. doi:10.1007/s10346-014-0502-0

Srivastava RM (2013) Geostatistics: a toolkit for data analysis, spatial prediction and risk management in the coal industry. Int J Coal Geol 112:2–13. doi:10.1016/j.coal.2013.01.011

Stähli M, Sättele M, Huggel C, McArdell BW, Lehmann P, Van Herwijnen A, Berne A, Schleiss M, Ferrari A, Kos A, Or D, Springman SM (2015) Monitoring and prediction in early warning systems for rapid mass movements. Nat Hazards Earth Syst Sci 15:905–917. doi:10.5194/nhess-15-905-2015

Stanley T, Kirschbaum DB, Huffman GJ, Adler RF (2017) Approximating long-term statistics early in the global precipitation measurement era. Earth Interact 21:1–10. doi:10.1175/EI-D-16-0025.1

Steger S, Bell R, Petschko H, Glade T (2015) Evaluating the effect of modelling methods and landslide inventories used for statistical susceptibility modelling. In: Lollino G, Giordan D, Crosta G, Corominas J, Azzam R, Wasowski J, Sciarra N (eds) Engineering geology for society and territory, vol 2. Springer International Publishing, Cham, pp 201–204

Steger S, Brenning A, Bell R, Glade T (2017) The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. Landslides. doi:10.1007/s10346-017-0820-0

Thiebes B (2012) Landslide analysis and early warning systems: local and regional case study in the Swabian Alb. Springer, Heidelberg

Thiebes B, Glade T (2016) Landslide early warning systems—fundamental concepts and innovative applications. In: Aversa S, Cascini L, Picarelli L, Scavia C (eds) Landslides and engineered slopes. Experience, theory and practice—proceedings of the 12th international symposium on landslides, 12–19 June 2016. CRC Press, Leiden, pp 1903–1911

Thiebes B, Bell R, Glade T, Jäger S, Mayer J, Anderson M, Holcombe L (2013) Integration of a limit-equilibrium model into a landslide early warning system. Landslides 11:859–875. doi:10.1007/s10346-013-0416-2

Tian Y, Peters-Lidard CD (2010) A global map of uncertainties in satellite-based precipitation measurements: uncertainties in precipitation data. Geophys Res Lett 37:L24407. doi:10.1029/2010GL046008
UNEP (2012) Early warning systems: a state of the art analysis and future directions. Division of Early Warning and Assessment (DEWA), United Nations Environment Programme (UNEP), Nairobi, Kenia
van Westen CJ, Van Asch T, Soeters R (2006) Landslide hazard and risk zonation—why is it still so difficult? Bull Eng Geol Environ 65:167–184. doi:10.1007/s10064-005-0023-0
van Westen CJ, Castellanos E, Kuriakose SL (2008) Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. Eng Geol 102:112–131. doi:10.1016/j.enggeo.2008.03.010
Velasco-Forero CA, Sempere-Torres D, Cassiraga EF, Gómez-Hernández JJ (2009) A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. Adv Water Resour 32:986–1002. doi:10.1016/j.advwatres.2008.10.004
Verworn A, Haberlandt U (2011) Spatial interpolation of hourly rainfall—effect of additional information, variogram inference and storm properties. Hydrol Earth Syst Sci 15:569–584. doi:10.5194/hess-15-569-2011
Vitolo C, Elkhatib Y, Reusser D, Macleod CJ, Buytaert W (2015) Web technologies for environmental big data. Environ Model Softw 63:185–198. doi:10.1016/j.envsoft.2014.10.007
Webster R, Oliver MA (1992) Sample adequately to estimate variograms of soil properties. J Soil Sci 43:177–192. doi:10.1111/j.1365-2389.1992.tb00128.x
Webster R, Oliver MA (2001) Geostatistics for environmental scientists. Wiley, Chichester
Wieczorek GF, Guzzetti F (1999) A review of rainfall thresholds for triggering landslides. In: Proceedings of the EGS Plinius conference, Maratea, Italy, pp 407–414
WMO (1996) Exchanging meteorological data: guidelines on relationships in commercial meteorological activities. WMO policy and practice. Secretariat of the World Meteorological Organization, Geneva
Yang D, Koike T, Tanizawa H (2004) Application of a distributed hydrological model and weather radar observations for flood management in the upper Tone River of Japan. Hydrol Process 18:3119–3132. doi:10.1002/hyp.5752

A.3. Canli et al. (in prep.): Evaluating landslide dynamics and internal structure — a case study from the Salcher landslide observatory, Austria

Co-author declaration for the following joint paper

This declaration states the research contribution (e.g. research idea and –questions; data compilation, manipulation and modelling, design and preparation of graphics, maps and tables; writing of text) of the candidate, the main supervisor (where he/she is an associate author) and the other authors.
If applicable, the contributions from other PhD candidates who has or intend to include the paper in a thesis should be described. Contributions from master students should be described.

All Authors (please underline corresponding author):
Ekrem Canli, Alexander Engels, Benni Thiebes, Bernhard Groiss, Thomas Glade, Joachim Schweigl, Michael Bertagnoli
Title: Evaluating landslide dynamics and internal structure – a case study from the Salcher landslide observatory, Austria
Journal: Bulletin of Engineering Geology and the Environment (in prep.)

Ekrem Canli's independent contribution:
- writing the paper
- further development of research idea and –questions
- extensive field work
- data manipulation and modelling (^{14}C dating, TLS, dynamic probing, percussion drilling, inclinometer measurements)
- design and preparation of graphics and maps

Alexander Engel's independent contribution:
- contribution to research idea and –questions
- extensive field work
- data manipulation and modelling (geomorphological mapping, total station survey, GNSS, dynamic probing, percussion drilling, inclinometer measurements, core sample analysis)
- design and preparation of graphics and maps

Benni Thiebes' independent contribution:
- initial set-up of research idea and –questions
- extensive field work
- revision of graphics & manuscript

Bernhard Groiss' independent contribution:
- extensive field work

Thomas Glade's independent contribution:
- initial set-up of research idea and –questions
- revision of manuscript

Joachim Schweigl's independent contribution:
- support of research idea and –questions

Michael Bertagnoli's independent contribution:
- support of research idea and –questions

Has this paper been, or will this paper be part of another doctoral degree thesis?
No

Contributions from master students:
Yes

If yes, elaborate details:
- contribution to field work

Do you verify that Ekrem Canli has contributed to this joint paper as described above?

Yes No

If no, specify:

If applicable, first co-author's signature

Yes No

If no, specify:

If applicable, second co-author's signature

Yes No

If no, specify:

If applicable, third co-author's signature

Yes No

If no, specify:

If applicable, fourth co-author's signature

Yes No

If no, specify:

If applicable, fifth co-author's signature

Yes No

If no, specify:

If applicable, sixth co-author's signature

<u>Has this paper been, or will this paper be part of another doctoral degree thesis?</u>
No

<u>Contributions from master students:</u>
Yes

If yes, elaborate details:

- contribution to field work

<u>Do you verify that Ekrem Canli has contributed to this joint paper as described above?</u>

Yes No

If no, specify:

____
If applicable, first co-author's signature

Yes No

If no, specify:

____
If applicable, second co-author's signature

Yes No

If no, specify:

____
If applicable, third co-author's signature

Yes No

If no, specify:

____
If applicable, fourth co-author's signature

Yes No

If no, specify:

____
If applicable, fifth co-author's signature

Yes No

If no, specify:

____
If applicable, sixth co-author's signature

Has this paper been, or will this paper be part of another doctoral degree thesis?
No

Contributions from master students:
Yes

If yes, elaborate details:

- contribution to field work

Do you verify that Ekrem Canli has contributed to this joint paper as described above?

Yes No

If no, specify:

If applicable, first co-author's signature

Yes No

If no, specify:

If applicable, second co-author's signature

Yes No

If no, specify:

If applicable, third co-author's signature

Yes No

If no, specify:

If applicable, fourth co-author's signature

Yes No

If no, specify:

If applicable, fifth co-author's signature

Yes No

If no, specify:

If applicable, sixth co-author's signature

Has this paper been, or will this paper be part of another doctoral degree thesis?
No

Contributions from master students:
Yes

If yes, elaborate details:
- contribution to field work

Do you verify that Ekrem Canli has contributed to this joint paper as described above?

Yes No

If no, specify:

If applicable, first co-author's signature

Yes No

If no, specify:

If applicable, second co-author's signature

Yes No

If no, specify:

If applicable, third co-author's signature

Yes No

If no, specify:

If applicable, fourth co-author's signature

Yes No

If no, specify:

If applicable, fifth co-author's signature

Yes No

If no, specify:

If applicable, sixth co-author's signature

Has this paper been, or will this paper be part of another doctoral degree thesis?
No

Contributions from master students:
Yes

If yes, elaborate details:
- contribution to field work

Do you verify that Ekrem Canli has contributed to this joint paper as described above?

Yes No

If no, specify:

____
If applicable, first co-author's signature

Yes No

If no, specify:

____
If applicable, second co-author's signature

Yes No

If no, specify:

____
If applicable, third co-author's signature

Yes No

If no, specify:

____
If applicable, fourth co-author's signature

Yes No

If no, specify:

____
If applicable, fifth co-author's signature

Yes No

If no, specify:

____
If applicable, sixth co-author's signature

Ekrem Canli
Department of Geography and Regional Research
University of Vienna
Universitätsstraße 7/B504
1010 Vienna
Austria
Phone: +43/680/2114451
E-mail: ekrem.canli@univie.ac.at

Evaluating landslide dynamics and internal structure – a case study from the Salcher landslide observatory, Austria

submitted to:

Bulletin of Engineering Geology and the Environment

Ekrem Canli[1], Alexander Engels[1], Benni Thiebes[1], Bernhard Groiss[1], Thomas Glade[1], Joachim Schweigl[2], Michael Bertagnoli[2]

(1) Department of Geography and Regional Research, University of Vienna, Vienna, Austria
(2) Geological Survey, Office of the Lower Austrian Provincial Government, St. Pölten, Austria

ORCID:

Ekrem Canli: 0000-0002-0890-5969

Benni Thiebes: 0000-0002-9067-2120

Thomas Glade: 0000-0001-9676-4416

Evaluating landslide dynamics and internal structure – a case study from the Salcher landslide observatory, Austria

Canli, E.[1], Engels, A.[1], Thiebes, B.[1], Groiss, B.[1], Glade, T.[1], Schweigl, J.[2], Bertagnoli, M.[2]

(1) Department of Geography and Regional Research, University of Vienna, Vienna, Austria (ekrem.canli@univie.ac.at)

(2) Geological Survey, Office of the Lower Austrian Provincial Government, St. Pölten, Austria

Abstract

Rainfall triggered landslides around the globe pose a major threat to both, human life and infrastructure. By now, it is widely accepted that human induced climate change alters temperature and precipitation patterns in some parts of the world. While acknowledging precipitation as the main trigger for landslides, an increased landslide activity can be expected in the near future. This calls for long-term landslide monitoring sites in order to understand kinematic behaviour and triggering conditions. Here, we present a recently established monitoring site on an active landslide in Austria that targets at a decadal persistence. The aim of this study is to characterise the internal structure, assess the current landslide dynamics and to analyse process activity by means of surface and subsurface monitoring installations. Surface methods currently cover terrestrial laserscanning, GNSS, and total station measurements. These reveal actual surface movement rates of several cm per year in the most active part of the landslide. Inclinometer measurements together with results from core drillings and penetrations tests suggest a shear plane in approx. 3 m depth. The combination of different methods within this study provide valuable information for a proper understanding of the landslide structure and its kinematics. As the landslide shows a moderate displacement velocity, it represents an ideal study site for testing new monitoring techniques, developing novel analysis methods, and proposing alert and warning schemes.

Keywords: Landslide monitoring, Rhenodanubian Flyschzone, GNSS, total station, terrestrial laserscanning, inclinometer, core drillings, penetration tests

Introduction

Landslides are natural phenomena that pose a substantial hazard and risk worldwide. Seismic shaking and intense rainstorms are commonly the main triggering agents for landslide occurrences. Strong earthquakes can trigger large numbers of landslides which may cause thousands of fatalities, e.g. during the 2008 Wenchuan earthquake in China (Yin et al. 2009). In addition, rainfall-triggered landslides frequently damage infrastructure and cause significant loss of life. Petley (2012) studied the effects of non-seismic-triggered landslide using a 7-year statistic and approximated an average annual death toll of more than 4,600.

An increased number of landslides is frequently attributed to the future global change, however, large uncertainties remain (Crozier 2010). To better understand the future frequency and magnitude relationship of landslides, and to be able to develop improved alerting and warning capabilities, long-term landslide monitoring systems are required (Thiebes 2012).

A large number of methods have been utilised for landslide monitoring systems (Mikkelsen 1996; Thiebes 2012; Thiebes and Glade 2016). The methods can be grouped into approaches, which analyse the triggering factors, e.g. rainfall and soil-water conditions, or the landslide movement itself. The former include for example piezometers to measure the position of groundwater tables (Massey et al. 2013), TDR (time domain reflectometry) probes to analyse volumetric soil water content (Camek et al. 2010), tensiometers for measuring pore water pressures (Montrasio and Valentino 2007), and also electrical resistivity tomography (ERT) which can give a spatially distributed estimation of soil wetness (Supper et al. 2014; Gance et al. 2016).

Landslide movement monitoring can broadly be distinguished between surface and subsurface methods. Traditionally, total stations (Burghaus et al. 2009; Reyes and Fernández 1996), as well as GNSS-based techniques (Corsini et al. 2012; Yin et al. 2010b) were the primary surface measurements methods. In recent years, terrestrial laserscanning (Abellán et al. 2011; Canli et al. 2015), radar interferometry (Mazzanti et al. 2014; Monserrat et al. 2014; Mulas et al. 2015) and photogrammetry-based methods (Gance et al. 2014; Stumpf et al. 2015; Travelletti et al. 2008) have become popular because they are suitable for displacement measurements over wider areas. Inclinometers, either operated manually or automatic, remain somewhat the gold standard for monitoring of subsurface displacements (Bell and Thiebes 2010; Jongmans et al. 2008; Yin et al. 2010a).

Only few monitoring systems have been actively used for more than a decade. However, these are extremely useful to investigate long-term movement patterns in relation to potential triggers, to be able to test new monitoring equipment, to develop new landslide analysis methods, and to propose novel alerting and warning schemes. Here, we present the characterisation of a complex landslide in the Austrian Prealps as a preparatory step for the installation of a long-term landslide observatory. The aim of the monitoring system is not only to improve the understanding of the landslide under investigation but also to develop and test new methods which then can be implemented on other potentially dangerous landslides, or contribute to already existing systems. We describe the environmental and geomorphological conditions of the landslide, and the results of investigative studies. An outlook highlights the next steps of the implementation of a comprehensive monitoring system, which includes several novel approaches but also reliable traditional methods.

Study Area

The study site is located in the western part of the federal state of Lower Austria in the municipality of Gresten (Scheibbs district in Austria). Gresten is located in a geologically complex area in which three different lithological units are present within a very narrow band of about two kilometres (Fig. 1b). In the alpine SW-NE striking direction, from north to south, the Rhenodanubian Flyschzone is followed by the Gresten Klippen Zone and the Northern Calcareous Alps. The Cretaceous-Early Tertiary Rhenodanubian Flyschzone (RDF) is located at the northern foothills of the East Alps. It is a paleogeographic-tectonic unit as part of the oceanic Penninic zone that was mostly eliminated in the subduction process involved in the Alpine orogeny (Hesse 2011). Flysch materials in the study area are deeply weathered and mainly consist of alterations of pelitic layers (clayey shales, silty shales, marls) and sandstones. The Gresten Unit, mostly termed as the Gresten Klippen Zone (GKZ), is situated in front of the Northern Calcareous Alps. It forms several "Klippen" originating from Jurassic and Lower Cretaceous deposits covered by variegated marls (in German: *Buntmergelserie*) with intercalated sandy limestones (Höck et al. 2005). Being part of the Helvetic system, the GKZ is entirely overthrusted by the main nappe of the RDF. In Lower Austria, both, the RDF and the GKZ as low mountain regions with a highly undulating terrain, are exceptionally prone to landsliding (Gottschling 2006; Petschko et al. 2014). Both units exhibit around five landslides per km^2 in Lower Austria (Petschko et al. 2014).

The study area is located in a warm-temperate, fully humid area characterized by hot summers and cold winters. The mean annual air temperature is 7.0°C, whereas the mean annual rainfall exhibits 1212.9 mm (normal period 1981-2011). However, heavy rainfall events exceeding 100 mm per day may occur. Such events (e.g. September 6, 2007 and June 23, 2009) caused major flooding and triggered landslides in Gresten and other parts Lower Austria. In addition to strong rainfall events, rapid snowmelt has been identified as one of the main triggering factors of landslides in the region (Schwenk 1992; Schweigl and Hervás 2009).

Fig 1 a) The study site is located in the western part of Lower Austria (Digital elevation model (DEM) with overlain orthophoto of the Salcher landslide); b) The Salcher landslide is situated in the municipality of Gresten that is embedded in a highly diverse geological setting between the Flyschzone, the Gresten Klippen Zone, and the Northern Calcareous Alps

The Salcher landslide

The Salcher landslide is situated on a non-forested, east exposed slope on an elevation of ca. 500 m a.s.l. with slope angles between 10° and 20°. The size of the active landslide part is ca. 4000 m^2. The landslide is located centrally in the municipality of Gresten and two roads surround it. Three residential buildings are situated underneath the main sliding direction of the landslide (Fig. 1a).

Historical information

Prior to its initial activation, the Salcher slope was used as a skiing track from the 1950ies onwards (Fig. 2a). In the 1970ies, skiing activities ceased due to repeated skewing of the lift pillars (Irmgard Plank, pers. comm. 2015). Initial significant slope movements were reported in July 1975. Heavy rainfall in the period between June 29 and July 3 in 1975 was assumed to be the triggering cause (internal technical report BD-3120/1-1975 by the Geological Survey of Lower Austria). The weather conditions were also made responsible for the occurrences of 236 other landslides in the region. During this period, precipitation values exceeded the average monthly rainfall amount by more than 200% (Schwenk et al. 1992).

In autumn 1975, remedial measures on the Salcher landslide were carried out on the slope and comprised levelling of the terrain and filling of cracks. Although awareness of the problematic water conditions on the slope existed, countermeasures with respect to the implementation of an upslope drainage system were not realised. Three years later, parts of the landslide were reactivated by a heavy rainfall event on May 31, 1978. For this day, the rainfall record in nearby Randegg exhibited 101.9 mm. Similarly to the rainfall event in 1975, many other landslides and flooding were assigned to this extreme rainfall event (Schwenk 1979). A prominent slope concavity is situated above the currently pronounced scarps (Fig. 1a). The planar and circular area was levelled out in the year 2000 in order to use it as a vaulting area (Hans Plank, pers. comm. 2015). In the course of a rainy period in the first week of August 2006, repeated movements set in. Rainfall measurements indicated 162.5 mm for the period between August 1 and August 7, 2006. The movements caused the formation of tensile cracks, which exhibited 20 m in width and revealed fracture openings of 10 cm. This was accompanied by the creation of fresh minor scarps and hummocks (Fig. 2b).

Fig 2 a) Historical view of the Salcher landslide surface (roughly late 1950s). Photograph: Marcel Mollik; b) Comparison of the Salcher landslide surface in 2007 and 2014

In July 2014, during fieldwork for the current monitoring system, a tree root sample was discovered in a drill core on the landslide in 2.6 m depth. Radiocarbon dating of the wooden sample revealed a calibrated date of 1670 AD (with a conventional ^{14}C age of 265 ± 30 BP). OxCal (Bronk Ramsey 2013) with IntCal13 atmospheric curve (Reimer et al. 2013) were used for calibration. This is interesting for two reasons; a) the oldest obtainable photographs of the Salcher slope (1950s) reveal already a treeless surface, and b) the depth and ambient material of the sample location, which represents a densely bedded, greyish and reducing environment, thus a long time of undisturbed groundwater conditions. This leads to a possible conclusion, that the Salcher slope might have been forested once

4

under relatively stable slope conditions and that the place of discovery was once overrun by the landslide.

Previous monitoring activities

After the reactivation in August 2006, the Geological Survey of Lower Austria set up geodetic piles for total station surveys of the area. Surveys started in April 2007 and were conducted biannually until November 2012. The survey campaigns revealed a total displacement of 0.925 m on the most active part of the Salcher landslide. However, measurements ceased in November 2012 as annual displacement rates between December 2009 and November 2012 totalled only 5 cm, whereas the displacement between April 2007 and December 2009 excelled 4 cm per month on average (unpublished technical report BD1-G-142/001-2007 and GZ-BD5-12476 by the Geological Survey of Lower Austria). The initial survey ended in 2012 because the displacement rates on the landslide were considered as low to very low and no displacement at all was recorded at the elements at risks (three houses and public road).

Jochum et al. (2008) performed a mineralogical characterisation of offsite drill cores together with resistivity measurements and penetration tests. Clay mineral analyses revealed an absence of smectite in the samples. The main mineral content among all samples is weakly crystalized kaolinite. Moreover, illite was determined in all samples. The general occurrence of chlorite in the deepest samples was explained by an early stage of weathering. The interpretation from resistivity measurements pointed towards the existence of an upper active part of the landslide, situated between 0 m and 4 m, and a lower, currently inactive, part between 4 m and 9 m. Based on penetration resistance, slope morphology, and resistivity measurements, a weathering horizon between 9 m and 14 m was suggested.

Methodology

Field measurements conducted at the Salcher landslide consist of surface (GNSS, total station and terrestrial laserscanning) and subsurface methods (dynamic probing heavy, percussion drillings, and inclinometers). Setting up the monitoring site at the Salcher landslide was accompanied by a couple of activities, ranging from preliminary desk study and field reconnaissance to concomitant laboratory and data analysis. Fig. 3 shows the methodological approach of this study.

Fig 3 Methodological approach of this study

Desk study and geomorphological mapping

A better understanding of the historical kinematics was necessary to specify the locations for subsurface investigations and later monitoring. Hence, it was required to assess available

5

professional opinions, damage reports, and findings from previous investigations to get an overview of the kinematic and morphological conditions and changes on the slope. With the intention to identify surface displacement of prominent surface structures from an early stage of fieldwork, GNSS based mapping was initially carried out in 07/2014. Prominent morphological surface structures were surveyed with a Leica GPS 1200. Correction signals were obtained from the Austrian Positioning Service (APOS) that enable measurements with a 3D-uncertainty lower than 1.5 cm.

Total station surveys

The already presented network of geodetic benchmark piles was surveyed via electronic distance measurements. Distance measurements were taken with a Leica TCRP 1201 together with reflecting prisms. The manual of the manufacturer indicates an accuracy of 1 mm. Both, the GNSS and Total station survey data, were visualised with ArcGIS 10.1.

Terrestrial Laserscanning

Two field campaigns (10/2014 and 12/2014) were performed in order to obtain multi-temporal TLS data. Those scans were carried out with a Riegl VZ-6000 long-range terrestrial laserscanner (TLS). The Geological Survey of Lower Austria provided another point cloud of the Salcher landslide dating back to early 2007. Data acquisition consisted of several steps, using a similar approach as Prokop and Panholzer (2009):

1. Location of suitable scan positions that minimize occlusions in the final point clouds.
2. Mounting a GNSS receiver onto the TLS that receives correction signals from the Austrian Positioning Service (APOS). Thus, the point clouds from the respective scan positions were already in the correct coordinate system and coarsely registered. Each scan position was recorded separately by the GNSS so those can act as tie points in the fine registration process afterwards.
3. The scan process itself. This included taking pictures with the integrated, calibrated camera, so the point cloud could be enhanced with RGB information.

The post processing started with the fine registration of each set of point clouds for the 10/2014 and 12/2014 survey. This has been done with the ICP (Iterative Closest Point) algorithm provided by the RiScan Pro software from Riegl. This multi station adjustment (MSA) tool modifies the orientation and position of each scan position in an iterative way in order to calculate the best overall fit. To compare the scan positions, surface data of the scanned objects were used to align the scan positions. Two methods for detecting corresponding plane surface patches were applied: a) manual definition of plane surface patches (from nearby house walls and a street). Hereby, the MSA modifies the scan position by minimising the distance between the defined planes, and b) application of an

6

automatic plane patch filter to detect corresponding points using ICP. Therein, the point cloud is divided in equal sized cubes of a certain size. For each cube, a best-fit plane (least-squares method) is estimated from all points within the cube. After each set of point clouds from a single scan survey is properly registered, another MSA is performed between the three point clouds to match the surfaces from all three available scan surveys (2007, 10/2014, and 12/2014).

For vegetation filtering, we used the RiScan Pro terrain filter that analyses the distance of the points from an estimated ground surface. Based on this analysis, the points are either classified as ground points or non-ground points. After exporting the filtered point cloud for each of the three scan surveys, the point clouds have been used for distance measurements calculations. For this, a DEM of difference (DoD) was created. Therefore, the point clouds have been converted to DEM and hillshades using OPALS (Orientation and Processing of Airborne Laser Scanning data) software (Pfeifer et al. 2014). The DEMs have been created with a cell size of 10 cm. The differences in height (z-axis) were calculated using ArcMap 10.2.

Dynamic Probing Heavy

Dynamic probing heavy (DPH) was performed by using the SRS-15 (German type) penetrometer. The apparatus is shown in Fig. 4a. The aim of this method was to detect changes in mechanical resistance of the subsurface material based on penetration tests (Springman et al. 2009). The device is pneumatically operated with a drop weight of 50 kg. A cone with 43.7 mm in diameter and a dropping height of 500 mm ensured a standardised application according to the European Standard *EN ISO 22476-2* for DPH. The number of blows required for each 10 cm was counted. Subsequent to every advanced meter increment, the rods were rotated to minimize skin friction.

Fig 4 Fieldwork on the Salcher landslide: a-c) Dynamic probing on site DPH1; d-e) Drilling operation with drill core extraction; f) Water outburst of drill core B6; g-h) Inclinometer installation with bentonite grout injection; i) Inclinometer data acquisition

Percussion drilling

Core drillings were carried out with a crawler drill GTR 780V from Geotool. This augering technique is a common method for obtaining only slightly disturbed core samples (Van Den Eeckhaut et al. 2007). The rig can access moderately steep locations with little waterlogging (Fig. 4b). It operates with a standardised weight of 63.5 kg and dropping height of 75 cm. A casing drilling approach was applied to prevent borehole collapses and water inflow. In order to protect the material, drilling progress stopped at a blow-count of 100 for a 10 cm increment. The probe was then extracted from the ground by a vertical-lift hydraulic tube clamp.

Inclinometer measurements

Inclinometer casings obtained from Glötzl were used for field installation (Fig. 4c). The casings consist of flexible acrylonitrile butadiene styrene (ABS) compounds. Each casing is 3 m long and has a diameter of 55 mm. The casings were riveted together to reach the respective final installation depth. The lower end was furnished with a plastic cap. Water tightness of the inclinometer casing was improved by wrapping plastic petrolatum tapes on a polypropylene fleece material base around the connecting elements. Upon reaching the target depth, the drill pipe was removed and the casing was inserted and subsequently filled with water. This way, potentiometric equilibrium with the surrounding was established. Care was taken to orientate the alignment of the leading grooves in direction of the estimated slope movement (Dunnicliff 1993). A tight connection between the casing and the surrounding stratum was achieved by using a bentonite-grout backfill (Bassett 2011). Zero readings were performed after three weeks, which allowed the borehole to settle and the bentonite-grout backfill to harden. Measurements were taken in 50 cm increments using the NMG probe from Glötzl. The probes measuring accuracy varies between 0.01 – 0.1 mm per measurement increment. In order to average errors, the probe was turned by 180° before the second reading. Additionally, surface positions of the inclinometers were determined (Dunnicliff 1993, Stark and Choi 2008). Subsurface displacement was calculated using the software GLNP V4 from Glötzl. Each measurement series was set in relation to the zero reading. Cumulative deformation curves were calculated, which visualize displacement values at corresponding depths.

Core sample analysis

The use of plastic inliners allowed continuous sampling of soil specimens in the lab, which permitted a soil-physical characterisation of the material (Prinz and Strauß 2011). In the laboratory, particle size, natural water content, carbonate content, and consistency were determined. The natural water content was determined according to Austrian standard for gravimetric and volumetric water content (ÖNORM L 1062). The state of consistency was investigated by kneading tests based on the German standard for soil classification in civil engineering (DIN 18196). Particle size analysis was performed by a combined sieving and sedimentation analysis. The samples were split into a coarse soil fraction (> 63 µm) and a fine soil fraction (< 63 µm). Sieve analysis of the coarse fraction was performed in accordance with ÖNORM L 1061–1. Sedimentation analysis of the fine fraction followed ÖNORM L 1061–2. For the sieve analysis, mechanical shakers and test sieves with 2 mm, 630 µm, 200 µm, and 63 µm opening diameter (DIN-ISO 3310/1) were used to derive the weight fractions of gravel, coarse sand, medium sand, and fine sand.

The finer fractions of silt and clay were determined by using a particle size analyser. The respective diameter thresholds were set to 63 µm, 20 µm, 6 µm and 2 µm to establish comparability to the relevant Austrian standard. The SediGraph 5120 and auto sampler MasterTech 052 from Micromeritics were used for analysis. The machine performs X-Ray monitored gravity sedimentation to determine the particle size, which resulted in cumulative finer mass percent versus particle diameter. Prior to analysis, fine soil samples were treated with 0.1% Tetrasodium pyrophosphate, which acts as a deflocculant for clays and prevents coagulation. Additionally, ultrasonic sound was applied before sedimentation analysis to break agglutinated grain matrixes. Granulometric curves were then calculated for each sample. The soil type was determined by plotting the textural composition on the Austrian soil textural triangle.

Results

In several field campaigns between 07/2014 and 02/2015, dynamic probing, core drillings, the installation of inclinometers, terrestrial laserscanning, total station surveys, and inclinometer data acquisition were carried out on the Salcher landslide. Notwithstanding difficult ground conditions and technical limitations with the used drilling rig, six boreholes were drilled, thirteen sites were investigated with DPH, and three inclinometers were installed (Fig. 5). Morphological features and geodetic piles were surveyed and point cloud data acquired during field campaigns.

Fig 5 Overview of surface and subsurface investigation sites

Geomorphological mapping

GNSS-based mapping during 07/2014 was carried out to assess visible scarps to compare them to the measurements from 2007. Heavily waterlogged areas in the perimeter of the lower bulged area was determined. In general, surface morphological observations were consistent with the findings from Jochum et al. (2008), yet with a more pronounced frontal part. Slightly below the slope concavity, several scarps are visible. The missing vegetation and rough surface morphology indicate higher landslide activity in this part. The currently invisible main scarp, as suggested by Jochum et al. (2008), and the lower bulged area delineate the upper and lower boundary of the depletion and transport zone. Altogether, the currently active landslide area was calculated to cover approximately 4000 m². The uppermost visible scarp of the landslide is approximately 110 m long whereas several minor scarps delimit a structure that resembles characteristics of a rotational landslide head. The inclined concrete foundation (a remnant of the skiing lift) on top of the head is turned against the hill by approximately 20°. The steep landslide toe is pointing towards east and stops at approximately 30 m distance to the former lift house.

Since 2007, the toe area remained stagnant in its location, but steepened quite significantly. Apart from some very steep areas, the mostly hummocky landslide surface is densely vegetated with grass. The scarps are mostly free from vegetation, indicating recent activity. The waterlogged areas are covered with patches of *Juncus effusus*.

Terrestrial laserscanning

The ICP algorithm used for registration of point clouds had some problems in defining proper planar surfaces on natural surfaces (e.g. trees). With the automatic plane patch filter, the registration error was in the range of 20-30 mm. After manually defining 10 plane patches distributed across the point cloud (mainly on artificial structures such as house walls or streets that are available at the lower margin of the landslide), the standard deviation of error could be reduced to 10 mm. The RiScan Pro terrain filter removed all points that are non-surface points, however, the dense grass cover was problematic and ground surface information could not be obtained. In both cases (in the 2007 and 12/2014 scan), the grass vegetation remained more or less the ground surface. However, both scans were performed outside the growing season, so that grass conditions on both point clouds can be assumed comparable. For further analyses, we used the filtered point clouds for creating the DEM of difference (DoD).

The DoD between 2007 and 12/2014 revealed an elevation loss along the currently visible scarps up to 75 cm (Fig. 6). Accumulation was identified below the crescent course of the uppermost visible scarp and on the landslide surface that was subject to bulging of slope material. Accumulation was also determined along the toe area of the landslide (up to 85 cm). Nevertheless, these values need to be interpreted with caution, as some uncertainties remain due to possible differences in grass vegetation, registration errors, and DEM surface interpolation. However, the overall pattern is in accordance to the findings of the geomorphological mapping that was done previously and the interpretation of historical photos.

Fig 6 DoD based on point clouds obtained in 2007 and 2014

Total station surveys

Discontinuous total station measurements performed since 2007 showed highly variable movement rates. The most active part of the landslide revealed movement rates up to almost 4 cm per month between 07/2008 and 12/2009, whereas movement rates between 12/2009 and 11/2012 were lower than 0.5 cm per month. A continuation of total station measurements at the persisting geodetic network started in 01/2015. The results indicate substantial movements of PF3 and PF4, which showed displacements of 10.9 cm and 45.7 cm, respectively, corresponding to approximately 0.5 and 2 cm on average per month (Fig. 7). These values are in respect to the previous

10

measurements from 12/2012. The total displacement at the most active part of the landslide surface was captured at PF4 with 1.36 m since 04/2007. The southeast trend of the deformation of PF3 and PF4 is consistent for all measurements that were performed since 2007. PF2 heaved quite significantly until 01/2015, when total station measurements reveal a vertical difference of 5.1 cm in comparison to 2012. Two benchmark points were mounted on the former ski-lift house that showed displacements of 1.4 cm with respect to 2012. Total station measurements were performed assuming PF5 and PF1 to be stable over time. However, benchmark 103 and 104 showed deviations of 1.4 and 1.2 cm in southern direction. Both benchmark points are located on the walls of the buildings below the landslide.

Fig 7 Results from total station measurements

Inclinometer measurements

An anchorage below the suggested upper shear zone was achieved in Inc1, which was installed to a depth of 13 m. The inclinometer neighbours geodetic pile PF4 and corresponds to drill core B2 and the sites of dynamic probing DPH4, DPH12, and DP13 (see Fig. 5 for locations). Inc1 was placed in the upper part of the currently active landslide part. Inc2 neighbours PF3 and was drilled to a final depth of 6.5 m. The site corresponds to drill core B4 and DPH8 and monitored the subsurface conditions near the right flank of the landslide body. Inc3 was drilled to a final depth of 6.5 m; the site corresponds to drill core B1 and DPH3.

Given the average inclination of the three inclinometers, which is below 5.5°, an overall measuring accuracy margin of 0.01 mm to 0.1 mm per measuring step was achieved. In accordance with this benchmark, an error margin of 0.26 to 2.6 mm was attained in Inc1, 0.13 to 1.3 mm in Inc2 and Inc3, respectively (Bell and Thiebes 2010). Inclinometer Inc1 was installed in 06/2014. Manual inclinometer measurements of Inc1 commenced in 07/2014. The reading in 09/2014 already showed 6 mm downslope displacement, at 3 m depth. The reading of Inc1 in 01/2015 revealed substantial deformation at 3 m depth. Within 189 days, Inc1 recorded a cumulative displacement of 38.8 mm in downslope direction (A-axis). The perpendicular movement component, which is measured along the B-axis, revealed little to no deformation (Fig. 8). The final reading reported here was carried out in 02/2015 and confirmed this evaluation. An additional displacement of 0.6 mm lies within the error margin, hence no change to the reading in January could be determined.

Fig 8 Inclinometer measurements (locations shown on Fig 9)

The first zero readings of Inc2 and Inc3 were carried out in 09/2014. The top of the flexible inclinometer casing was subject to minor pulling forces, which occurred when the probe was lifted to the top. The slight bending of the deformation curve in the uppermost meter was caused by this and should therefore be neglected. In 01/2015, Inc2 was assumed to be sheared off, because the blind probe could not be lowered below 3 m depth. The inclinometer was therefore investigated with a borehole camera and found to be intact. It took several attempts to lower the probe down for a reading. The measurement revealed substantial displacement at 3 m depth (Fig. 11a). A maximum lateral displacement of 44.2 mm was detected at 2 m depth. Compared to the downslope component (A) of the movement, negative B values of approximately 2.6 mm indicate a minor trend towards the North. The unfavourable bend in the casing geometry of Inc2 was made responsible for the difficulties with the probe. Likewise Inc2, zero readings of Inc3 were performed in 09/2014. Similarly to Inc2, slight bending of the top of the casing could be seen in the form of minor deflections in downslope direction. Hence, the uppermost measurement was neglected for interpretation. Manual inclinometer measurements in 01/2015 revealed downslope displacements, which were confirmed in the February readings. A deformation of approximately 18.9 mm in positive A-direction and 4.3 mm in positive B-direction could be limited to the zone between 1 m and 2 m depth.

Core sample analysis

Soil-physical properties (particle size, water content, carbonate content, consistency state) of six drill cores were investigated. The location of the drill sites is indicated in Fig. 5. For drill core B1, a total of 17 specimens were extracted, based on transitions in colour and consistency state. According to the Austrian ÖNORM L1050, soil specimens were classified as loam, loamy clay, silty loam, sandy loam, and loamy silt (Fig. 9).

Fig 9 Drill core samples from drilling site B1 with the corresponding soil textures (according to ÖNORM L1050)

The clay fraction varied between 21.5% and 47.9%. Silt takes the largest share in the overall particle size distribution of B1. Whereas the highest values of 66.4% were found in the deepest samples between 4.6 m and 5 m, the lowest values were found within the uppermost meter and range from 29.1% at 0.3 m depth to 34% between 0.7 m and 1 m depth. Sand is present in smaller proportions, ranging from 3.6% to 17.4%. Soil skeleton portion varied between 35.8% and 0.2%, and was highest in the uppermost samples. Fig. 10 summarises particle size, corresponding water content, penetration resistance, and carbonate content, obtained from drill core B1.

Fig 10 Visualisation of particle size, water content, penetration resistance, and carbonate content at corresponding depth in B1

The water content varied strongly between the uppermost 1.5 m and the lowest samples. Highest values and very soft consistency were determined in the uppermost samples and

12

reached 34.4%. The lowest values corresponded to the deepest samples and totalled 7.8%. Apart from elevated levels at approximately 3 m depth, a steady decline towards the bottom of the soil column can be observed. This is in agreement with the constantly increasing blow count of DPH3 at approximately 3 m depth. As seen in B1, the soil samples in B2 were purely cohesive, highest in silt content, and exhibited clay values reaching up to 43.6%. Lowest clay values of 14.1% were determined between 2.8 m and 2.95 m. Within the next 5 cm, however, a transition to loam was determined (Fig. 11). The loamy sample is marked by its stiffness and greyish reduced colour.

Fig 11 Sharp textural transition in 2.95 m depth in B2

Compared to the other drill cores, the water content of B4 was consistently high and exceeded 40% at 0.5 m, 2 m and 3.5 m. The overall cohesive soil samples therefore showed very soft and even slurry consistence. Between 3.45 m and 4 m, as well as just below the surface, clay was generally lower compared to the other drill cores. The synthesis of all drill core analyses is shown in the proposed underground model for the Salcher landslide in the discussion part (Fig. 12)

Discussion and conclusions

Over the period of seven years of total station monitoring, PF3 and PF4 represented by far the most active zones of the landslide. Current total station measurements of the benchmark piles confirm recent slope displacements similar to the most active period in 2009. Particularly PF4 showed surface deformation values of approximately 20 cm per year. However, the long surveying interval of two years does not permit drawing conclusions on a steady movement rate. It is also possible that a sudden displacement of 40 cm occurred in the course of two years. Apart from this, the formation of new tension cracks occurred below the proposed landslide head and indicated the highest activity in this area as well. Based on the findings of the TLS change detection and GNSS measurements, the most active area of the landslide was allocated in the immediate vicinity of PF4. Substantial change in aforementioned areas can also be seen in a comparison of oblique imagery of the slope.

Surface morphological information, together with data from the core analysis, penetration resistance, and inclinometer data, was used to create an underground model of the landslide. The proposed underground model (Fig. 12) suggests that the landslide consists of several interconnected sliding bodies, which resembles the geometry of a larger rotational landslide. The area between the visible scarps and the bulged foot slope exhibits structural characteristics of a translational landslide. The deepest shear zone was determined at approximately 8 m below the surface of the slope concavity. The course of its upper end (dotted line) was drawn based on surface observation and an assumed gradient, which was derived from core analysis. The courses of the minor scarps were derived from local displacement measurements by inclinometers and the structural interpretation of drill core data. The remaining course of the shear zone was constructed based on the interpretation of the surface morphology and blow count values from dynamic probing.

The depth of the shear zone in the lower part of the depletion and transport zone and the bulging zone of accumulation was derived from dynamic probing data and displacement measurements in Inc3.

Fig 12 Proposed underground model of the Salcher landslide based on all obtained information (inclinometers, drill cores, penetration resistance)

Inclinometer-derived movement data demonstrate ongoing movement over the course of the study. Furthermore, core drillings and inclinometers reveal information only at single point locations, limiting their informational value. However, for the time of inquiry, Inc1 and Inc2 underwent a downslope displacement of approximately 4 cm and 2 cm, respectively. What could further be concluded from the inclinometer measurements was that movements in the zone of depletion and transport zone occurred within narrow bands of approximately 1 m thickness in a depth of 3 m in the case of Inc1 and Inc2, and at 2 m depth in Inc3. This is interpreted as the existence of a traceable sliding plane over a longitudinal section.

For the drill cores B1, B2 and B4, a depth related correlation to elevated clay content could be shown. Clay rich soil materials usually undergo a reduction of shear strength over the course of shear processes and residual strength is usually lower due to decreased friction angles. With respect to the composition of the materials in the clay fraction, reference could be made to the findings from Jochum et al. (2008), whose investigations indicated a dominating presence of weakly crystallised kaolinite throughout depth. In conclusion it can be stated, that all of the sampled drill cores showed textural heterogeneity that is interpreted as signs of dynamic behaviour, which cannot be explained by soil diagenetic aspects alone. Movements are assumed to be triggered upon saturation of sensible clay rich layers, which were determined at 2 m depth in drill core B1, and at approximately 3 m depth in drill core B2 and B4. The findings in this study further revealed that the transient parameter water content could not be used to gain straightforward information on the landslide structure. This could be seen in layers, which were underlain by clay rich aquiclude layers. The samples of these layers do not indicate accumulation trends of percolating water above impervious beds. Given the relatively high permeability of the lowest sections of drill core B3, elevated water contents were to be expected above the stiff to hard layer at 8.5 m depth. Despite the use of sealed plastic inliners, evaporation must be taken into account in the two-month timespan between sampling and laboratory analysis of the drill cores. Core loss and resulting empty space in the inliners might have favoured the outgassing of the liquid phase even further. Because little core loss was experienced in drill core B4 and that less time had elapsed between sampling and analysis, the water content in B4 could be used to demonstrate accumulation trends above the depth of 3 m. However, B4 was drilled without a casing, and its drilling was performed several weeks after the drilling of B1-B3. Hence, it cannot be ruled out that water from a perched water table may have biased the results in B4. Likewise, it has to be assumed that different local precipitation conditions and consequently varying soil water conditions were present prior to the sampling campaign of B4.

Moreover, the overall high mixing of material classes, which could be observed in drill core B1, indicates high activity near toe of the landslide. Several clay anomalies did indicate processes, which presumably have rearranged more advanced weathering products on less weathered products. However, no statements with respect to the origin of the clay particles can be made within this study.

Another interesting finding was that the parameter penetration resistance could be used to delineate shear zones at the depth of confirmed movements. Together with drill core interpretation, the assumed shear plane from inclinometer measurements (Fig. 12) could also be traced from dynamic probing and core interpretation alone (Fig. 14).

Fig 13 Proposed scheme of the landslide dimensions solely based on dynamic probing and core analysis

Outlook

The first months of ongoing fieldwork at the Salcher landslide observatory revealed first results of the internal structure and the dynamics of the landslide. However, conclusions towards future slope stability or even triggering conditions cannot be drawn yet. In addition, the discontinuous surface and subsurface measurement intervals of the installed devices aggravate any further predictions. Yet, this preliminary study highlights the importance of long-term monitoring efforts for active landslides in urban areas.

An important issue that could not be addressed in this study is the source of the water for this landslide, which has yet to be determined. It is expected that permanent electrical resistivity tomography (ERT), that has started only recently along the entire length of the slope, will deliver interesting insights into the hydrologic response behaviour. Measurement interval for the ERT is currently at three hours. Likewise, the recently set up weather station on the landslide surface and the implemented piezometers for measuring changes in ground water level are expected to enable investigations on the hydraulic response of the slope in near future. Additionally, TDR

15

probes in different depths are installed for assessing *in situ* soil water content. An innovative device that is currently being worked on is a permanent terrestrial laserscanning (pTLS) system (Canli et al. 2015). Most of the monitoring devices installed contain only point information, whereas TLS data enables spatially widespread information about surface changes over time. However, until now the main constraint of TLS surveys is the temporal resolution. Data is being acquired only very sporadic due to labour costs and time requirements for field campaigns. With this newly developed system, high-resolution point cloud data is obtained once a day and processed in a fully automated way, including data transmission and registration, vegetation filtering and change detection. Movement data then correlated with a newly installed automatic inclinometer and rainfall data (both with 10 min measurement interval). For all automatic measurement devices, a reliable data infrastructure has been implemented by now. Permanent electricity and internet is available on the entire landslide area and data is transferred to a data server in Vienna in (near) real-time. Ultimately, the data gained is used within further analyses including data correlation, threshold analysis, and spatio-temporal slope stability analysis. At present, automatic data post-processing, as well as web-based visualisation of measured data are being developed.

ACKNOWLEDEMENTS

This study was kindly co-financed by the Geological Survey of the Lower Austrian Provincial Government, St. Pölten, Austria. The authors want to thank the municipal authorities of Gresten, especially Irmgard Plank. Additionally, we want to thank the landowner and the neighbouring families for their kind collaboration and their support of our investigations. Robert Peticzka and Christa Hermann are thanked for their aid in the laboratory analyses and for fruitful discussions. We are also grateful to the technical staff, particularly Oliver Szczech and Bamidele Ayoniyi, and to all participating students during fieldwork. Finally, we thank the referees and editor for their valuable remarks that significantly helped in improving the manuscript.

References

Abellán A, Vilaplana JM, Calvet J, et al (2011) Rockfall monitoring by Terrestrial Laser Scanning – case study of the basaltic rock face at Castellfollit de la Roca (Catalonia, Spain). Natural Hazards and Earth System Science 11:829–841. doi: 10.5194/nhess-11-829-2011

Bassett R (2012) A Guide to Field Instrumentation in Geotechnics: Principles, Installation and Reading. CRC Press, Boca Raton

Bell R, Thiebes B (2010) Untergrundbewegungen. In: Bell R, Mayer J, Pohl J, et al (eds) Integrative Frühwarnsysteme für Gravitative Massenbewegungen (ILEWS). Monitoring, Modellierung, Implementierung. Klartext Verlag, Essen, Germany, pp 84–92

Bronk Ramsey C, Scott M, van der Plicht H (2013) Calibration for archaeological and environmental terrestrial samples in the time range 26–50 ka cal BP. Radiocarbon 55:2029–2034. doi: 10.2458/azu_js_rc.55.16935

Burghaus S, Bell R, Kuhlmann H (2009) Improvement of a Terrestric Network for Movement Analysis of a Complex Landslide. In: Proceedings of FIG Working Week 2009, Surveyors Key Role in Accelerated Development, 3. – 8. May 2009. Eilat, Israel

Camek T, Becker R, Öhl S (2010) Bodenfeuchte (TDR, Tensiometer). In: Bell R, Mayer J, Pohl J, et al (eds) Integrative Frühwarnsysteme für Gravitative Massenbewegungen (ILEWS). Monitoring, Modellierung, Implementierung. Klartext Verlag, Essen, Germany, pp 92–100

Canli E, Höfle B, Hämmerle M, et al (2015) Permanent 3D laser scanning system for an active landslide in Gresten (Austria). Geophysical Research Abstracts, 17, EGU2015-2885-2

Corsini A, Iasio C, Mair V (2012) Overview of 2001-08 GPS monitoring at the Corvara landslide and perspectives from 2010-11 use of HR X-band SAR (Dolomites, Italy). In: Eberhardt E, Froese C, Turner K, Leroueil S (ed) Landslides and Engineered Slopes: Protecting Society through Improved Understanding. CRC Press, Banff, Canada, pp 1353–1358

Crozier MJ (2010) Deciphering the effect of climate change on landslide activity: A review. Geomorphology 124:260–267. doi: 10.1016/j.geomorph.2010.04.009

Dunnicliff J (1993) Geotechnical instrumentation for monitoring field performance. John Wiley & Sons, New York

Gance J, Malet J-P, Dewez T, Travelletti J (2014) Target Detection and Tracking of moving objects for characterizing landslide displacements from time-lapse terrestrial optical images. Engineering Geology 172:26–40. doi: 10.1016/j.enggeo.2014.01.003

Gottschling P (2006) Massenbewegungen. In: Wessely G (ed) Geologie von Niederösterreich. Verlag d. Geologischen Bundesanstalt, Vienna, Austria, pp 335–340

Hesse R (2011) Rhenodanubian Flyschzone, Bavarian Alps. In: Carena S, Friedrich AM, Lammerer B (eds) Geological Field Trips in Central Western Europe: Fragile Earth International Conference, Munich, September 2011. Geological Society of America, Boulder, USA, pp 51–73. doi: 10.1130/2011.0022(05)

Höck V, Ślączka A, Gasiński MA, Bąk M (2005) Konradsheim Limestone of the Gresten Klippen Zone (Austria): New insight into its stratigraphic and paleogeographic setting. Geologica Carpathia 56:237. doi: 10.2478/v10096-009-0010-7

Jochum B, Lotter M, Ottner F, Tiefenbach K (2008) Geophysikalische und ingenieurgeologische Methoden zur Untersuchung von durch Massenbewegungen bedingte Bauschäden in Niederösterreich (case study No. BBK-Projekt NC-62/F (2007) and ÜLG-35 (2007)). Geological Survey of Austria and University of Natural Resources and Life Sciences, Vienna

Jongmans D, Renalier F, Kniess U, et al (2008) Characterization of the Avignonet landslide (French Alps) with seismic techniques. In: Chen Z, Zhang J, Li Z, et al (eds) Landslides and Engineered Slopes. From the Past to the Future. CRC Press, Xi'an, China, pp 395–401

Massey CI, Petley DN, McSaveney MJ (2013) Patterns of movement in reactivated landslides. Engineering Geology 159:1–19. doi: 10.1016/j.enggeo.2013.03.011

Mazzanti P, Bozzano F, Cipriani I, Prestininzi A (2015) New insights into the temporal prediction of landslides by a terrestrial SAR interferometry monitoring case study. Landslides 12:55–68. doi: 10.1007/s10346-014-0469-x

Mikkelsen PE (1996) Field instrumentation. In: Turner AK, Schuster RL (eds) Landslides – Investigation and Mitigation: Transportation Research Board Special Report 247. National Academy Press, Washington, USA, pp 278–316

Monserrat O, Crosetto M, Luzi G (2014) A review of ground-based SAR interferometry for deformation measurement. ISPRS Journal of Photogrammetry and Remote Sensing 93:40–48. doi: 10.1016/j.isprsjprs.2014.04.001

Montrasio L, Valentino R (2007) Experimental analysis and modelling of shallow landslides. Landslides 4:291–296. doi: 10.1007/s10346-007-0082-3

Mulas M, Petitta M, Corsini A, et al (2015) Long-term monitoring of a deep-seated, slow-moving landslide by mean of C-band and X-band advanced interferometric products: the Corvara in Badia case study (Dolomites, Italy). ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-7/W3:827–829. doi: 10.5194/isprsarchives-XL-7-W3-827-2015

Petley D (2012) Global patterns of loss of life from landslides. Geology 40:927–930. doi: 10.1130/G33217.1

Petschko H, Brenning A, Bell R, et al (2014) Assessing the quality of landslide susceptibility maps – case study Lower Austria. Natural Hazards and Earth System Science 14:95–118. doi: 10.5194/nhess-14-95-2014

Pfeifer N, Mandlburger G, Otepka J, Karel W (2014) OPALS – A framework for Airborne Laser Scanning data analysis. Computers, Environment and Urban Systems 45:125–136. doi: 10.1016/j.compenvurbsys.2013.11.002

Prinz H, Strauß R (2011) Ingenieurgeologie. Springer, Heidelberg

Reimer PJ, Bard E, Bayliss A, et al (2013) IntCal13 and Marine13 radiocarbon age calibration curves 0-50,000 years cal BP. Radiocarbon 55:1869–1887. doi: 10.2458/azu_js_rc.55.16947

Reyes CA, Fernández LC (1996) Monitoring of surface movements in excavated slopes. In: Senneset K (ed) Landslides. Seventh International Symposium on Landslides. Balkema, Trondheim, Norway, pp 1579–1584

Schweigl J, Hervás J (2009) Landslide mapping in Austria. European Commission Joint Research Centre, Institute for Environment and Sustainability, Italy

Schwenk H (1979) Erhebungsbericht und Gutachten des geologischen Dienstes der Baudirektion (No. BD-G-78156). Geological Survey, Office of the Lower Austrian Provincial Government, St. Pölten

Schwenk H (1992) Massenbewegungen in Niederösterreich 1953-1990. Jahrbuch der Geologischen Bundesanstalt 135:597–660.

Springman S, Kienzler P, Casini F, Askarinejad A (2009) Landslide triggering experiment in a steep forested slope in Switzerland. In: Hamza M, Shahien M, El-Mossallamy Y (eds) Proceedings of the 17th International conference on soil mechanics and geotechnical engineering, ICSMGE 2009. IOS Press, pp 1698–1701

Stark TD, Choi H (2008) Slope inclinometers for landslides. Landslides 5:339–350. doi: 10.1007/s10346-008-0126-3

Stumpf A, Malet J-P, Allemand P, et al (2015) Ground-based multi-view photogrammetry for the monitoring of landslide deformation and erosion. Geomorphology 231:130–145. doi: 10.1016/j.geomorph.2014.10.039

Supper R, Ottowitz D, Jochum B, et al (2014) Geoelectrical monitoring: an innovative method to supplement landslide surveillance and early warning. Near Surface Geophysics 12:133–150. doi: 10.3997/1873-0604.2013060

Thiebes B (2012) Landslide Analysis and Early Warning Systems. Local and Regional Case Study in the Swabian Alb, Germany. Springer, Berlin, Heidelberg

Travelletti J, Oppikofer T, Delacourt C, et al (2008) Monitoring landslide displacements during a controlled rain experiment using a long-range terrestrial laser scanning (TLS). Int Arch Photogramm Remote Sens Spat Inf Sci 37:485–490.

Van Den Eeckhaut M, Verstraeten G, Poesen J (2007) Morphology and internal structure of a dormant landslide in a hilly area: The Collinabos landslide (Belgium). Geomorphology 89:258–273. doi: 10.1016/j.geomorph.2006.12.005

Yin Y, Wang F, Sun P (2009) Landslide hazards triggered by the 2008 Wenchuan earthquake, Sichuan, China. Landslides 6:139–152. doi: 10.1007/s10346-009-0148-5

Yin Y, Wang H, Gao Y, Li X (2010a) Real-time monitoring and early warning of landslides at relocated Wushan Town, the Three Gorges Reservoir, China. Landslides 7:339–349. doi: 10.1007/s10346-010-0220-1

Yin Y, Zheng W, Liu Y, et al (2010b) Integration of GPS with InSAR to monitoring of the Jiaju landslide in Sichuan, China. Landslides 7:359–365. doi: 10.1007/s10346-010-0225-9

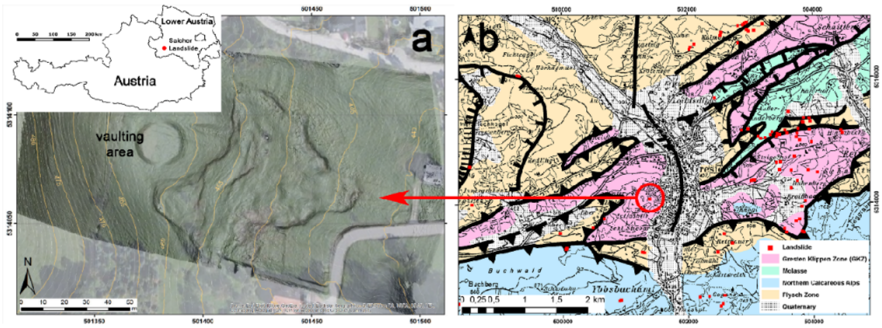


Fig 2 a) The study site is located in the western part of Lower Austria (Digital elevation model (DEM) with overlain orthophoto of the Salcher landslide); b) The Salcher landslide is situated in the municipality of Gresten that is embedded in a highly diverse geological setting between the Flyschzone, the Gresten Klippen Zone, and the Northern Calcareous Alps

21

Fig 2 a) Historical view of the Salcher landslide surface (roughly late 1950s). Photograph: Marcel Mollik; b) Comparison of the Salcher landslide surface in 2007 and 2014

22

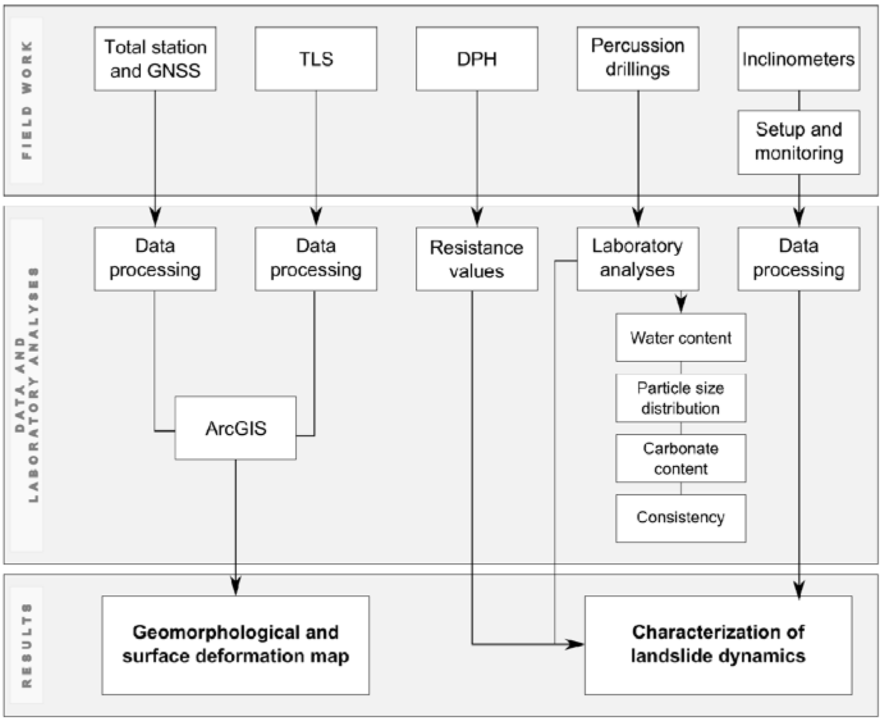
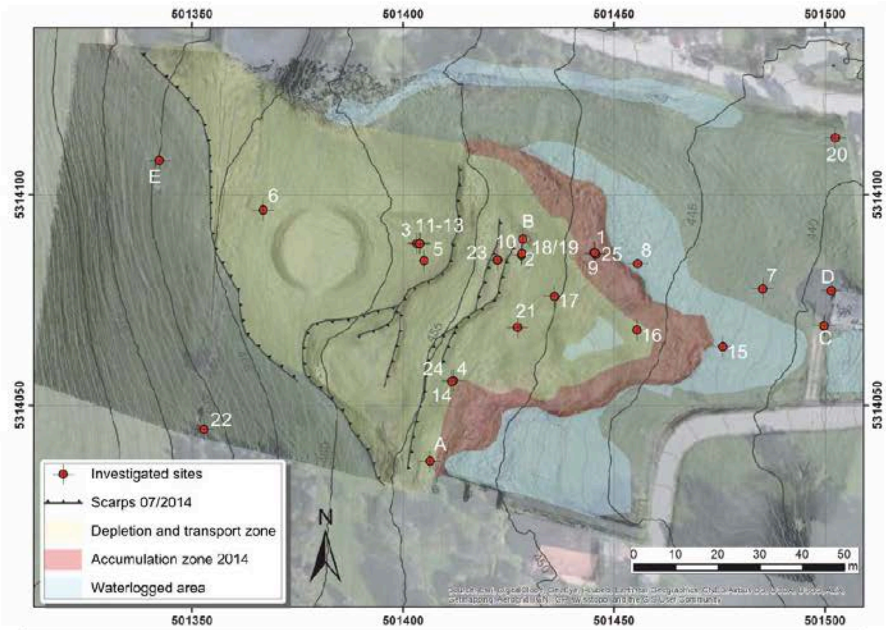


Fig 3 Methodological approach of this study

23

Fig 4 Fieldwork on the Salcher landslide: a-c) Dynamic probing on site DPH1; d-e) Drilling operation with drill core extraction; f) Water outburst of drill core B6; g-h) Inclinometer installation with bentonite grout injection; i) Inclinometer data acquisition

24



Core drillings	Dynamic probing	TLS	Inclinometers	Total station
1 B1	7 DPH1	20 TLS	23 Inc1	A PF3
2 B2	8 DPH 2	21 TLS	24 Inc 2	B PF4
3 B3	9 DPH 3	22 TLS	25 Inc3	C 1oo
4 B4	10 DPH 4			D 102
5 B5	11-13 DPH5-7			E PF2
6 B6	14 DPH8			
	15 DPH9			
	16 DPH10			
	17 DPH11			
	18 DPH12			
	19 DPH13			

Fig 5 Overview of surface and subsurface investigation sites

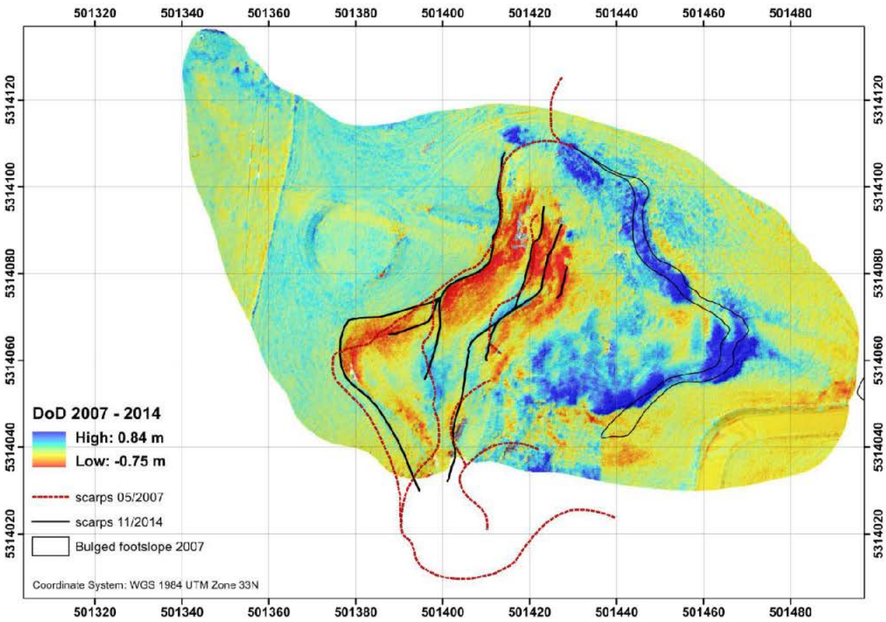


Fig 6 DoD based on point clouds obtained in 2007 and 2014

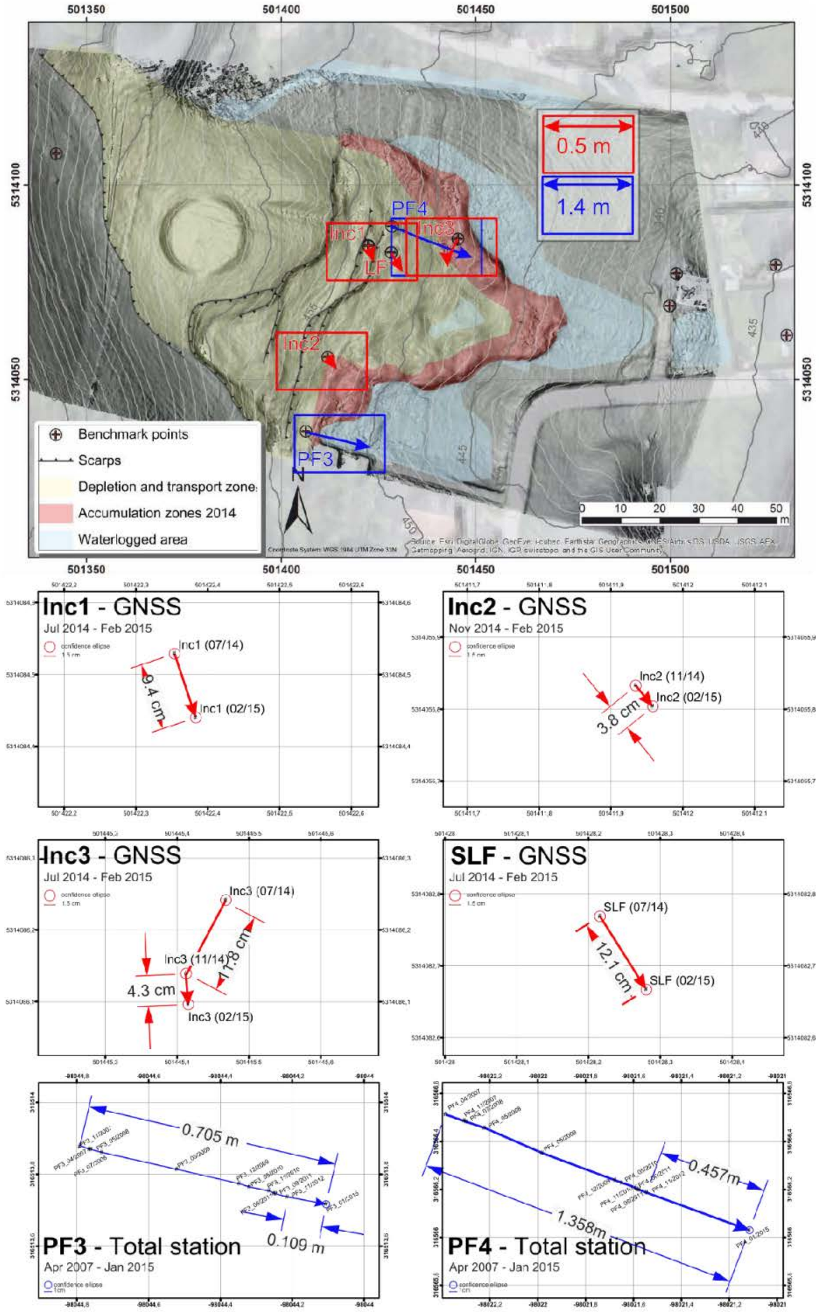


Fig 7 Results from total station measurements

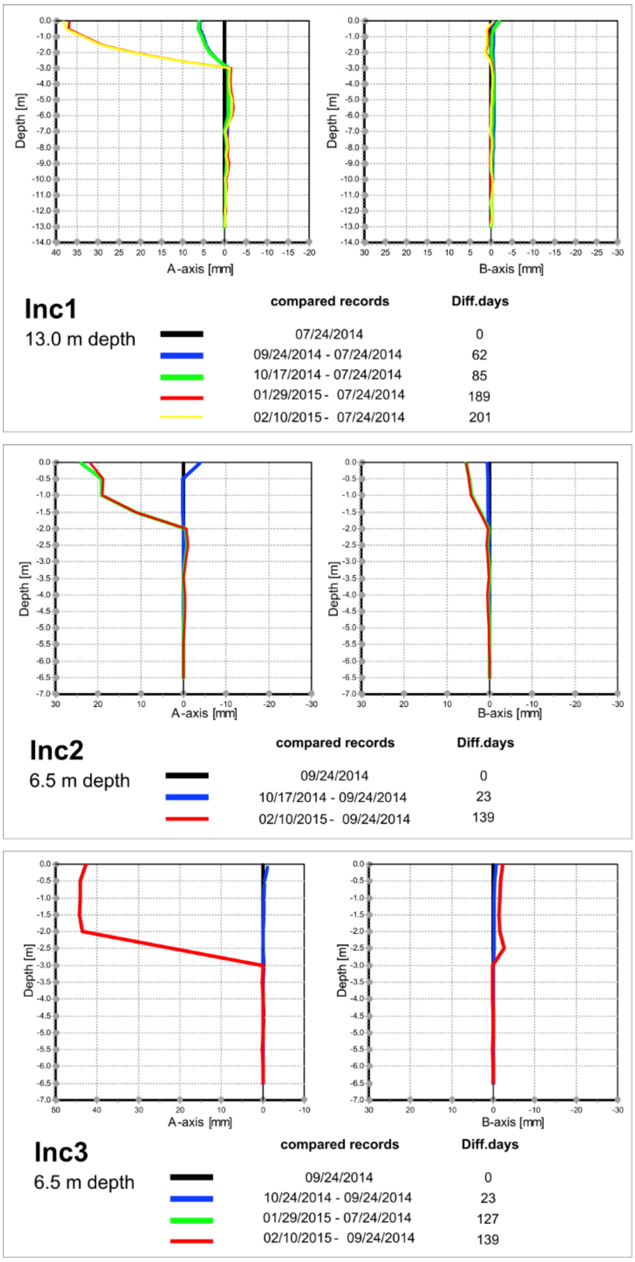


Fig 8 Inclinometer measurements (locations shown on Fig 9)

28

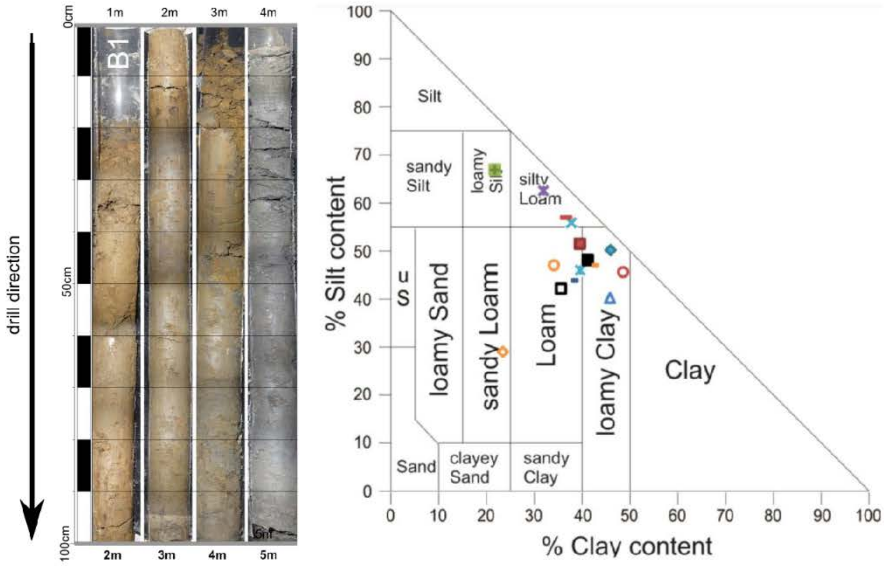


Fig 9 Drill core samples from drilling site B1 with the corresponding soil textures (according to ÖNORM L1050)

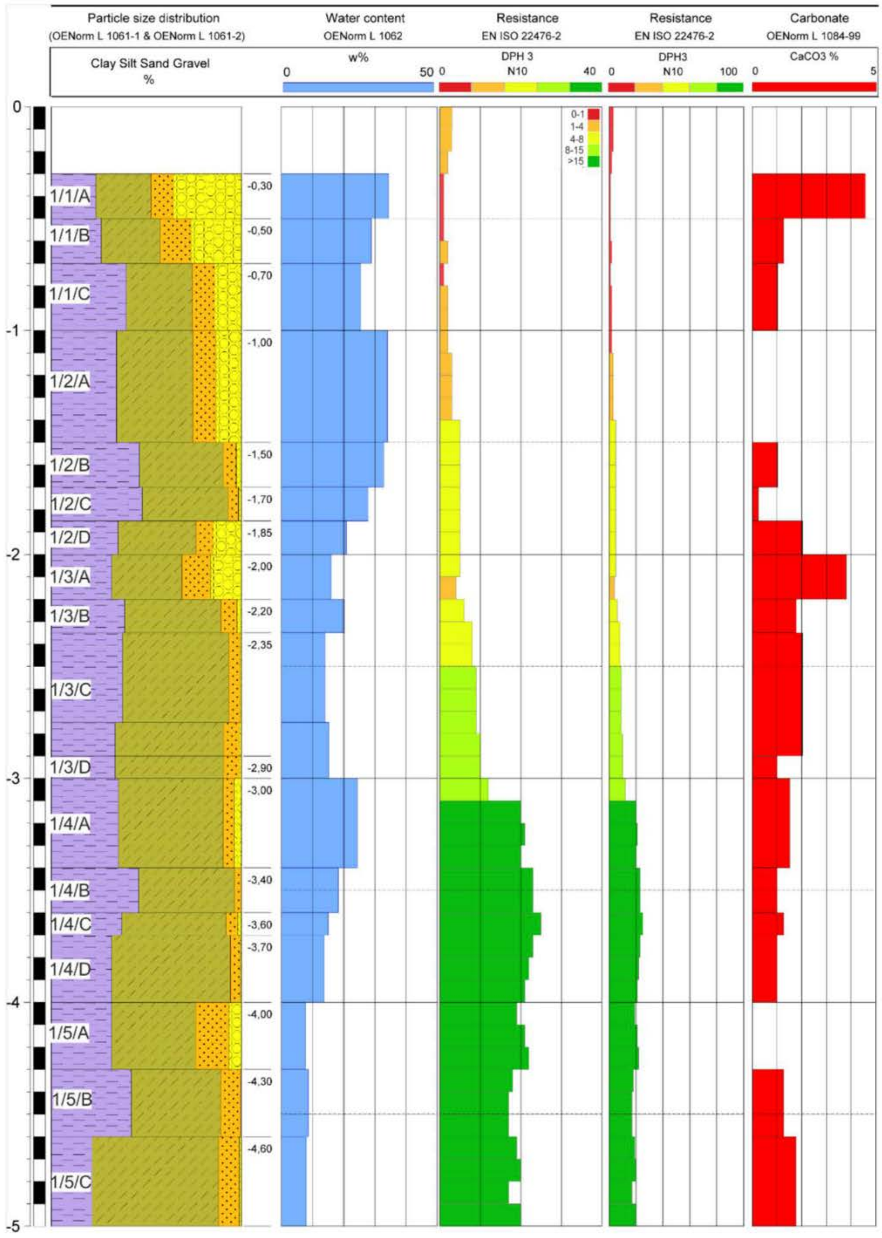


Fig 10 Visualisation of particle size, water content, penetration resistance, and carbonate content at corresponding depth in B1

30

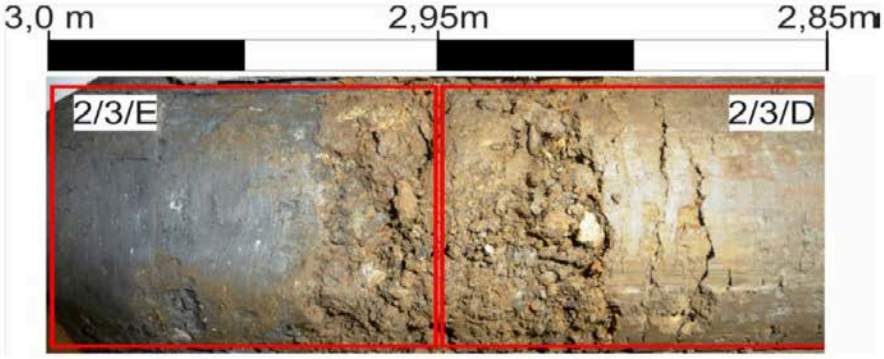


Fig 11 Sharp textural transition in 2.95 m depth in B2

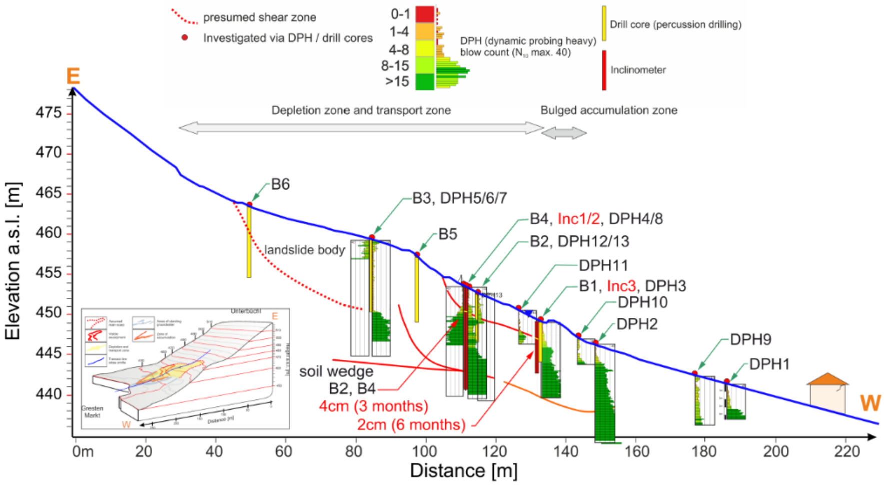


Fig 12 Proposed underground model of the Salcher landslide based on all obtained information (inclinometers, drill cores, penetration resistance)

32

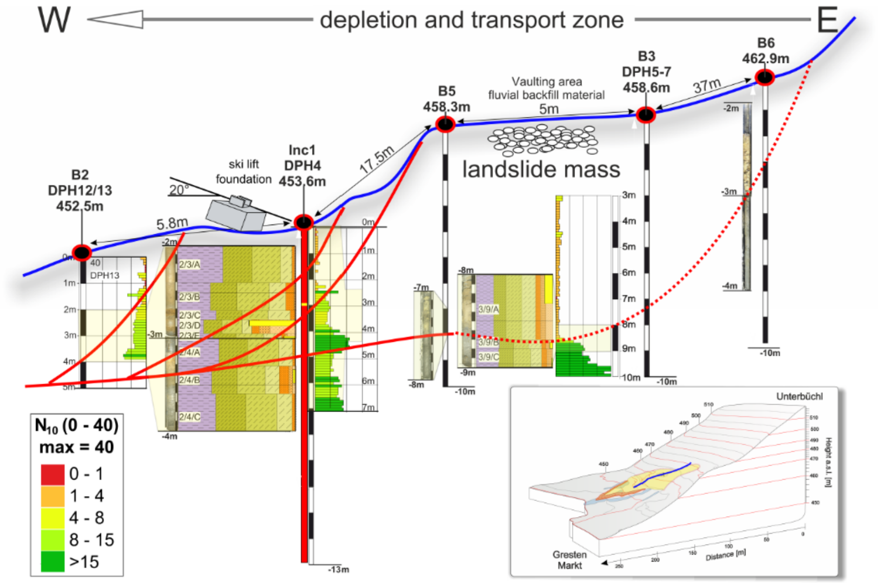


Fig 13 Proposed scheme of the landslide dimensions solely based on dynamic probing and core analysis

33

A.4. Canli et al. (2017b): Probabilistic landslide ensemble prediction systems: Lessons to be learned from hydrology

Co-author declaration for the following joint paper

This declaration states the research contribution (e.g. research idea and –questions; data compilation, manipulation and modelling, design and preparation of graphics, maps and tables; writing of text) of the candidate, the main supervisor (where he/she is an associate author) and the other authors.
If applicable, the contributions from other PhD candidates who has or intend to include the paper in a thesis should be described. Contributions from master students should be described.

All Authors (please underline corresponding author):
Ekrem Canli, Martin Mergili, Thomas Glade
Title:
Probabilistic landslide ensemble prediction systems: lessons to be learned from hydrology
Journal:
Natural Hazards and Earth System Sciences: https://doi.org/10.5194/nhess-2017-427 (in open review)

Ekrem Canli's independent contribution:
- writing the paper
- research idea and –questions
- data compilation
- data manipulation and modelling
- design and preparation of graphics and maps

Martin Mergili's independent contribution:
- contribution to research idea and –questions
- revision of manuscript

Thomas Glade's independent contribution:
- contribution to research idea and –questions
- revision of manuscript

Has this paper been, or will this paper be part of another doctoral degree thesis?
No

Contributions from master students:
No

Do you verify that Ekrem Canli has contributed to this joint paper as described above?

Yes No

If no, specify:

If applicable, first co-author's signature

Yes No

If no, specify:

If applicable, second co-author's signature

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences
Open Access
Discussions

Probabilistic landslide ensemble prediction systems: Lessons to be learned from hydrology

Ekrem Canli[1], Martin Mergili[1,2], Thomas Glade[1]

[1] Department of Geography and Regional Research, University of Vienna, Universitätsstr. 7, 1010 Vienna, Austria

[2] Institute of Applied Geology, University of Natural Resources and Life Sciences, Gregor-Mendel-Straße 33, 1180 Vienna, Austria

Correspondence to: Ekrem Canli (ekrem.canli@univie.ac.at)

Abstract. Landslide early warning has a long tradition in landslide research. Early warning can be defined as the provision of timely and effective information that allows individuals exposed to a hazard to take action to avoid or reduce their risk and prepare for effective response. In the last decade, hydrological forecasting started operational mode of so called ensemble prediction systems (EPS) following on the success of the use of ensembles for weather forecasting. Those probabilistic approaches acknowledge the presence of unavoidable variability and uncertainty at larger scales and explicitly introduce them into the model results. Now that convective-scale numerical weather predictions and high-performance computing are getting more common, landslide early warning should attempt to learn from past experiences made in the hydrological forecasting community. This paper reviews and summarizes concepts of ensemble prediction in hydrology and how ties to landslide research could improve landslide forecasting. Three future research directions were identified: 1.) evaluation of how and to what degree probabilistic landslide forecasting improves predictive skill; 2.) adaptation and development of methods for validating and calibrating probabilistic landslide models; 3.) application of data assimilation methods to increase the quality of physical parametrization and increased forecasting accuracy.

Keywords: ensemble prediction systems, probabilistic forecasting, landslide early warning

1. Introduction

Landslide prediction at regional scale is a hot topic within the scientific community as the time-varying aspects of landslide susceptibilities, hazards and even risks are crucial for emergency response planning and protecting public safety (Baum et al., 2010, Glade and Crozier, 2015). Further, the number of landslides is assumed to increase due to global change (Crozier, 2010, Gariano et al., 2017, Papathoma-Köhle and Glade, 2013). This calls for an increased demand in early warning procedures with the aim of issuing timely warnings of an upcoming hazardous event to temporarily reduce the exposure of vulnerable persons or infrastructure (Thiebes and Glade, 2016). In this paper, we use prediction systems synonymously with early warning systems for terminological consistency within the landslide community although we acknowledge that early warning should also cover dissemination and response strategies (UNEP, 2012). Warnings can be considered as calls for the public to take protective action, and the time scale of a warning depends on the associated weather event (Stensrud et al., 2009). For natural hazard

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences Discussions

types with a rapid onset, such as shallow landslides and flash floods, rainfall can be considered as the main triggering mechanism. Rainfall nowcasting, or short-term rainfall forecasting, is an important and well-established tool for numerical weather predictions (NWP) in meteorological and hydrological applications that offer rainfall predictions several hours ahead. While it is evident that processes with potentially very short response time require more effort for timely early warning than just real-time measurement of rainfall, real forecasting initiatives are scarce especially in the landslide community (Tiranti et al., 2017).

The reasons for the rare application of NWP products within the landslide early warning community are manifold. One reason might be the complexity of single landslide detachments: the same landslide triggering event does not necessarily cause other landslides as the time between propagation stage and the collapse phase may vary significantly based on differences in local conditions of topography, materials such as soil, regolith and rock, vegetation, etc. and spans from minutes (e.g. flow slides on slopes covered with shallow coarse-grained soils) to years (e.g. earth flows in slopes of fine grained soils) (Greco and Pagano, 2017). Based on empirical-statistical relationships between landslide occurrence and its associated rainfall event, rainfall thresholds within a certain confidence interval aim at accounting for those differences in slope failure behavior (Glade, 2000). Guzzetti et al. (2007) give an overview of rainfall and climate variables used in the literature for the definition of rainfall thresholds for the initiation of landslides, however, such empirical-statistical approaches only pose a simplification between rainfall occurrence and the physical mechanisms leading to landslides, neglecting local environmental conditions and the role of the hydrological processes occurring along slopes (Reichenbach et al, 1998, Bogaard and Greco, 2017). Attempts to relate landslide-triggering thresholds to weather and other physically based characteristics can be very challenging given the quality of currently available data (Peres et al., 2017). Another reason for the negligence of physically based forecasting initiatives used to be the lacking spatial resolution and computational power for considering such convective-scale phenomena which are of particular interest for modelling small scale related phenomena with a rapid onset such as shallow landslides and flash floods. This became, however, increasingly less of an issue. Convective-scale NWP with spatial resolutions of 1 to 4 km issued in very short time intervals are already available in many parts of the world. The hydrological community has recently adopted to those advancements by implementing such convective-permitting models into operational flood prediction systems (Hapuarachchi et al., 2011, Liu et al., 2012, Yu et al., 2015).

This paper reviews and summarizes concepts of ensemble prediction systems (EPS) in hydrology and how those can be translated to be applicable also in process-based landslide early warning systems. A strong emphasis is put on how to deal with spatial uncertainties by demonstrating the benefits of probabilistic model application which does not eliminate uncertainty, but it explicitly introduces in into the model results. In a case study, we highlight possible spatially distributed physically based landslide early warning products for decision makers and point out specific challenges that landslide research has to face in the upcoming years. The aims of this paper are:

a) to critically evaluate the current state of physically based landslide early warning, its limitations and possible ties to hydrological forecasting;

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences
Open Access
Discussions

b) on this basis, to foster cooperation across disciplinary boundaries to bring together scientists from different fields to pursue research based on forecasting experiences gained in the last couple of years.

2. Probabilistic forecasting in hydrology and ties to landslide research

When considering ensemble prediction systems (EPS), one should clarify what is expressed with the term *ensemble* and why EPS should be used at all since it is virtually unused in the landslide community. In the *Guidelines on Ensemble Prediction Systems and Forecasting* issued by the World Meteorological Organization (WMO, 2012), EPS are defined as "numerical weather prediction (NWP) systems that allow us to estimate the uncertainty in a weather forecast as well as the most likely outcome. Instead of running the NWP model once (a deterministic forecast), the model is run many times from very slightly different initial conditions. Often the model physics is also slightly perturbed, and some ensembles use more than one model within the ensemble (multi-model EPS) or the same model but with different combinations of physical parametrization schemes (multi-physics EPS). […] The range of different solutions in the forecast allows us to assess the uncertainty in the forecast, and how confident we should be in a deterministic forecast. […] The EPS is designed to sample the probability distribution function (pdf) of the forecast, and is often used to produce probability forecasts – to assess the probability that certain outcomes will occur" (WMO, 2012, p. 1).

Krzysztofowicz (2001) argues that forecasts should be stated in probabilistic, rather than deterministic, terms and that this "has been argued from common sense and decision-theoretic perspectives for almost a century" (Krzysztofowicz, 2001, p. 2). But still, by the new millennium, most operational hydrological forecasting systems relied on deterministic forecasts and there was a too strong emphasis on finding the *best* estimates rather than quantifying the predictive uncertainty (Krzysztofowicz, 2001). However, those times have been overcome a decade later (Cloke and Pappenberger, 2009). From a scientific and historical perspective, landslide prediction has very strong roots in empirical-statistical threshold based approaches (Wieczorek and Glade, 2005, Guzzetti et al., 2007). This stands valid until today, since most operational landslide early warning systems rely purely on the relationship between rainfall and landslide occurrence, thus representing only a simplification of the underlying physical processes. Baum and Godt (2010), Alfieri et al. (2012a), Thiebes (2012) and Thiebes and Glade (2016) give an overview of present and past operational landslide early warning systems (EWS). Bogaard and Greco (2017) critically analyze the role of rainfall thresholds for shallow landslides and debris flows from a hydro-meteorological point of view.

One reason why landslide forecasting is seemingly more challenging can be attributed to the spatial and temporal predictability of landslide processes. The spatial occurrence of floods is topographically foreseeable and controllable which is much more difficult to assess for landslides in distributed modelling due to their very localized nature (Alfieri et al., 2012a, Canli et al., 2017). Also, the prediction domain in flooding, which is usually streamflow, is rather straightforward to observe and to be measured accurately over a long time. In the past 15 years, a mindset of adapting probabilistic concepts to account for inherent uncertainties has taken over in the hydrologic community and the move towards ensemble prediction systems (EPS) in flood

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences
Discussions

forecasting represents the state of the art in forecasting science, following on the success of the use of ensembles for weather forecasting (Buzzia et al., 2005, Cloke and Pappenberger, 2009).

Unfortunately, initiatives such as the *Hydrological Ensemble Prediction Experiment* (HEPEX) were not fostered in the landslide community to date. The general aims of this ongoing bottom-up initiative are to investigate how to produce, communicate and use hydrologic ensemble forecasts in a multidisciplinary approach (Schaake et al., 2007). One reason for the absence of such cooperative efforts might be the political, and therefore also financial, situation that led to the advancement of ensemble predictions in hydrology. Many international bodies demonstrated their interest in EPS which led to this superior position of hydrological prediction. This is even more so the case when taking into account transboundary floods that are typically more severe in their magnitude, affect larger areas and cause more damage and overall losses (Thielen et al., 2009). Beven (1996) argues that the importance of water resources management led to considerably higher efforts by both researchers and government agencies in hydrological data collection.

Losses from landslides are perceived as mainly private and localized economic losses and thus, only few public resources have been allocated to develop sound spatial landslide early warning systems (Baum and Godt, 2010). As a result, spatial operational landslide early warning systems are scarce and many of them never surmounted their prototype status. Consequently, long monitoring time series, which are indispensable for sound and reliable early warning systems (such as available e.g. for floods, storms, etc.), are commonly not available. Additionally, methodological issues or inadequate monitoring together with insufficient warning criteria significantly reduce the ability of existing systems to issue effective warnings (Baum and Godt, 2010). When looking at the raw numbers, hydrological events rank among the main disaster events together with meteorological events when comparing events in global and multi-peril loss databases, while geophysical events take only a small fraction in absolute numbers (Alfieri et al., 2012a, Wirtz et al., 2014). However, it is widely accepted that landslide losses are vastly underestimated (Petley, 2012). There are several reasons for this observation: a) major disaster databases, e.g. the NatCatSERVICE from the reinsurance company Munich Re, associate landslides as subordinated hazard types of geophysical (amongst earthquakes) or hydrological hazards (amongst floods or avalanches) (Wirtz et al., 2014); b) landslide databases are inconsistent, incomplete or entirely absent and most of the existing inventories severely lack historical data (Wood et al., 2015).

3. Benefits and types of probabilistic approaches

Generally speaking, in an ensemble forecast small changes (perturbations) are made to the model parameters and then the model is re-run with these slightly perturbed starting conditions. If the different model realizations (*ensemble members*) are similar to each other, the forecasting confidence is rather high. Contrary, if they all develop differently, the confidence is much lower (WMO, 2012). By considering the proportion of the ensemble members that predict a storm or a landslide, we can make an estimate of how likely the storm or landslide occurs.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

The term *ensemble prediction* for environmental applications was coined in the field of meteorology, thus describing the application of numerical weather prediction systems, but it is used in different ways in neighboring disciplines. The atmospheric component is consistently described as weather ensemble input, yet the same applies to how observations of the land surface are incorporated into distributed forecasting models. In the data assimilation stage, ensembles of plausible land surface state observations (initial streamflow, soil moisture, snowpack, etc.) are created. Using multiple feasible parameter sets for each model or for each model run will realistically increase the spread of possible outcomes, yet it is more objective in terms of considered input parameters that were not directly observed (Schaake et al., 2007). Thus, the term ensemble prediction may be used in any instance of multi-parametric or multi-model data input that is used for forecasting the target variable.

In landslide research, there are a few attempts that explicitly address ensemble techniques as a means of overcoming limitations from purely deterministic approaches or by increasing the predictive performance of statistically based susceptibility mapping. None of them, however, incorporate ensemble techniques in real-time applications. Pradhan et al. (2017) used an ensemble approach to evaluate the output of a physically based model for a statistical machine learning model in varying hydrological conditions. Their ensemble model is based on a maximum entropy model that creates and combines multiple models to improve modeling results. However, their distributed output does not predict when or exactly where landslide will occur, but yields a classified map with information where landslide occurrence can be expected over the long-term. Thus, their presented ensemble approach indicates landslide susceptibility that may be applicable for regional/spatial planning. While the term *ensemble* is by no means used a lot in landslide studies, it seems that it is predominantly used by the statistical landslide susceptibility modeling community (e.g. Lee and Oh, 2012, Althuwaynee et al., 2014a, Althuwaynee et al., 2014b). It is, however, not used in any way to address uncertainties in a forecasting model (Bartholmes and Todini, 2005, Vincendon et al., 2011). In a very promising approach, Chen et al. (2016) couple a deterministic model with probabilistically treated geotechnical parameters with rainfall input from an operational multi-scale and multi-member NWP system (GRAPES) to forecast spatial landslide occurrences with their ensemble prediction model (GRAPES-Landslide).

While there are not many landslide studies using or at least addressing ensemble techniques, there has been quite some work done on probabilistic landslide hazard analysis in the recent past. Lari et al. (2014) propose a probabilistic approach expressing hazard as a function of landslide destructive power where landslide intensity (in terms of displacement rate) is considered rather than their magnitude. Haneberg (2004), Park et al. (2013), Raia et al. (2014), Lee and Park (2016) and Zhang et al. (2016) treat soil properties at regional scale applications in a probabilistic way by randomly selecting variables from a given probability density function, mostly by means of Monte Carlo (MC) simulation. Salciarini et al. (2017) tried to enhance those approaches by considering geostatistical methods to provide the spatial distribution of soil properties and by using the Point Estimate Method (PEM) as a computationally more efficient method compared to MC simulation. But still, none of those probabilistic approaches are operated in spatial real-time early warning systems, not even on a prototype basis. The research of Schmidt et al. (2008) represents a remarkable exception: they proposed a coupled regional forecasting system in New

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences Discussions

Zealand based on multiple process-based models (NWP, soil hydrology, slope stability). Unfortunately, a continuation of this research was not further pursued.

In general, it is possible to distinguish between three types of EPS: global, regional and convective-scale EPS. They each address different spatial and temporal scales in the forecast. For rainfall-induced landslide applications, the latter is the most appealing; thus, we will focus on this one alone. Convective-scale NWP, with model grid sizes of 1–4 km, can attempt to predict details such as the location and intensity of thunderstorms (WMO, 2012). Therefore, those systems reduce the effect of highly intermittent rainfall events that cause serious issues with small-scale rainfall events when applying geostatistical rainfall interpolation techniques (Canli et al., 2017). Convective-scale NWP models are likely to better resolve the intensity and spatial scale of local precipitation, especially in convective precipitation when topographic forcing is involved. Therefore, they are particularly valuable for predicting small scale phenomena, such as flash floods or landslides. However, the major drawback of convective-scale EPS is the immense cost of running (WMO, 2012).

In the past 15 years, many experimental and operational mesoscale EPS have been developed, yet very few with regard to convection-permitting EPS. In, 2012, the German Weather Service (Deutscher Wetterdienst - DWD) started operational mode for their COSMO-DE-EPS with a resolution of 2.8 km (Baldauf et al., 2011, Gebhardt et al., 2011). Similar operational forecasting systems have been implemented in the last couple of years by the weather services of France using their 2.5 km AROME model (Seity et al., 2011), the UK with their 2.2 km MOGREPS-UK model (Golding et al., 2016) and the USA using the 3 km High Resolution Rapid Refresh (HRRR) model (Ikeda et al., 2013).

4. The hydrological equivalent of rainfall-induced shallow landslides: the case of flash floods

One major difference between flood and landslide early warning is the available lead time. While the lead time in larger river basins is sufficiently long to prevent any hazardous situations from river flooding, shallow landslides, in the case of first time failures, generally occur suddenly and spatially unforeseeable in a specific area susceptible to landsliding. As opposing to regular floods, however, flash floods can indeed be considered as an appropriate counterpart to rainfall-induced shallow landslide occurrence. Flash floods are, similar to shallow landslides, characterized by the superior importance of small-scale extreme precipitation events and their rapid onset, which leaves only little response time. it is therefore appropriate to examine how flash flood forecasting is performed and how it is applicable to landslide forecasting. What makes landslide forecasting particularly challenging is the evolutionary sequence of the process.

Greco and Pagano (2017) distinguish between three stages of a typical predictive system's architecture: I) the predisposing stage, II) the triggering and propagation stage, and III) the collapse stage. While in hydrological applications (II) and (III) are hardly distinguishable from each other, for rainfall-induced landslides this is not necessarily the case. While the predisposing stage (I) is determined by e.g. increasing pore water pressure due to a varying length of rainfall input that worsens the slope stability conditions, the triggering and propagation stage (II) spans from first local slope failures until the formation of associated slip surfaces. The collapse phase (III) ultimately consists of the mobilization of the entire mass leading to the actual

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

failure. However, the time between stages (II) and (III) may vary significantly based on differences in local geomorphology, soil, vegetation, etc. and spans from a couple of minutes (e.g. flow slides in slopes covered with shallow coarse-grained soils) to years (e.g. earth flows in slopes of fine grained soils) (Greco and Pagano, 2017). Even when spatially distributed process-based landslide predictions are performed in relatively homogeneous regions, this time offset still prevails and makes landslide modelling in any context a challenging task. Therefore, warnings should generally be issued during indications of stage (II) since the lead time of stage (III) might be too short given the rapid kinematic characterization of the post-failure behavior, as recent disastrous examples in Italy have shown (Greco and Pagano, 2017).

Hydrological forecasting systems relying only on rainfall observations do not allow for a sufficiently long lead time for warnings. Extending this forecasting lead time further than the watershed response times requires the use of quantitative precipitation forecasts (QPF) from numerical weather predictions (NWP) (Vincendon et al., 2011). Additionally, models to represent hydrologic and hydraulic processes within a catchment to determine how rainfall-runoff accumulates is required (Hapuarachchi et al., 2011). With regard to producing quantitative precipitation estimates (QPE) in real-time, research has gone into blending multiple sources of information (radar, satellite and gauged data) to increase the accuracy of QPEs. This process is generally referred to as *data assimilation* and is considered as increasingly important for improving hydrological predictions (Reichle, 2008).

For predicting flash floods, however, longer lead times are necessary and thus high resolution QPFs with 1–6-hour lead times are generated. In recent years, the spatial (<5 km) and temporal (<1 h) resolutions of NWP model rainfall forecasts have significantly improved, while the combination of such NWP model forecasts with blends of the advected patterns of recent radar, satellite and gauged rainfall data additionally increased the accuracy of nowcasting products (Hapuarachchi et al., 2011). Based on those high-resolution NWP model forecasts, probabilistic ensemble prediction systems have aided in exploring and quantifying uncertainties. Numerous studies have used those probabilistic precipitation forecasts to drive hydrological models (Vincendon et al., 2011, Bartholmes and Todini, 2005, Siccardi et al., 2005, Thielen et al., 2009). The application of such convective-permitting ensemble NWP is computationally very demanding and still in its infancy with respect to flash flood prediction (Alfieri et al., 2012b). However, a further reduction of the spatial uncertainties of high-resolution rainfall fields is highly desirable, given the fact that rainfall is still considered as the most uncertain parameter in hydrological forecasting systems (Hapuarachchi et al., 2011, Alfieri et al., 2012b).

5. Many sizes fit all: the concept of equifinality

The concept of equifinality is deeply rooted in the hydrological community. It expresses an acceptance that many sets of parameters may provide equally acceptable forecasts (Beven, 1996, Beven and Freer, 2001, Collier, 2007). The concept of equifinality revolves around the rejection of the concept of *the* optimal model in favor of multiple possibilities for producing acceptable simulators (Beven and Freer, 2001). This concept is based on the understanding of physical theory and relates to the plethora of interactions among the components of a system whose resulting representations may be equally acceptable.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Research generally follows a working paradigm that should lead to realistic representations of the real processes and characteristics. This idea of identifying a single optimal representation of reality is very distinct in environmental sciences. A major problem arises from the scale discrepancy between sampling and distributed modeling where the use of global parameters undoubtedly leads to errors in predicting local responses at points with unique characteristics (Beven and Freer, 2001). By acknowledging that there are many different model structures or many possible parameter sets scattered throughout the parameter space, the range of predicted variables is likely to be larger than linearized solutions would suggest. This equally means acknowledging that there are uncertainties inherent surrounding the area of parameter space around the *optimum*. As a result, such approaches allow nonlinearity to be considered for predictions (Beven and Freer, 2001).

Geomorphological systems can indeed be considered as transient, inheriting remnants of past and present processes. Environmental systems can exhibit certain degrees of chaotic behavior which results in an inability to express the trajectory of their development based on present-day evidence alone. Therefore, equifinality should not be considered as an indication of a poorly developed methodology, but as something inherent in geomorphological systems (Beven, 1996). However, it should most certainly *not* serve as a loophole for an inadequate methodology or model setup! A practical consequence of this equifinality may lead to a more robust approach to testing the viability of different model setups with the aim to reject some, but to retain many of the offered solutions (Beven, 1996). Similarities and differences in model results should ultimately lead to an improved process understanding and, hence, predictive models with a higher sensitivity and specificity.

6. Calibration and validation of probabilistic forecasts

A model is an abstraction and a simplification of reality, hence the need for assessing its validity. Model validation provides a legitimacy in terms of arguments and methods (Oreskes et al., 1994). However, model validation is difficult when the most interesting events are rare, which is generally the case for flash floods or landslides. Also, calibration might be difficult for certain variables, or where suitable observations are not available. The WMO (2012) suggests that direct model output (DMO) from ensembles, although not ideal, still provide valuable information (WMO, 2012). The probabilistic forecasts with a DMO might not be as sharp (e.g. larger ensemble spread), but they still offer an estimate of the uncertainties and thus pose an advantage over purely deterministic forecasts. But even where measurements of modeling parameters are available, it has often shown that those parameters cannot be assumed constant in space or time, which makes calibration even more difficult. Additionally, the scale of measurement generally differs significantly from the scale at which the applied model requires "effective" parameter values to be specified (Beven, 1996).

Deterministic models for landslide prediction synthesize the interaction between hydrology, topography, vegetation and soil mechanics in order to physically understand and predict the location and timing that trigger landslides. These models usually contain a hydraulic and a slope stability component with different degrees of simplification (Formetta et al., 2016). In most cases, the target variable is the slope safety factor (FoS), which is useful as it enables decision makers to take actions when if

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

falls short of a certain threshold (e.g. FoS < 1.0). Also, when talking about the probability of an event occurring, this event must be defined:

- What is the threshold value to be exceeded?
- What is the exact time or time period to which the forecast refers?
- What is the exact location or area to which the forecast applies?
- Which uncertainties are considered and what is their role in the modelling process?

With regard to those questions and as a starting point, the FoS is a suitable variable for probabilistic forecasting. Yet it has two major flaws: a) it is only a ratio of resisting forces to driving forces that is commonly not directly measured in the field and cannot be directly monitored, and b) landslide events are rare and (unlike streams for example) their future location of occurrence remains unknown until they occur. This makes landslide calibration a really challenging task. And there are limitations of model calibration in the case of rare events. Commonly, calibration will improve the reliability of forecasts (i.e. the match of the target variable or forecast probabilities to frequency of observations of the event) but reduce the resolution of the forecast (the ability to discriminate whether an event will occur or not). Consequently, calibration will improve forecasts of common events, but will also lead to the underprediction of more extreme events. The WMO (2012) argues that this is the case for rare events, since the statistical distributions are trained to the more common occurrences. For rare events, hence, calibration cannot be expected to provide significant improvement over the raw forecasts.

Besides model calibration, validation is an important part within forecasting. Validation unfortunately comes with a rather strong emphasis on either-or-situations. In practice, few (if any) models are entirely confirmed by observational data, and few are completely refuted (Oreskes et al., 1994). On top of that, for most models there may be multiple combinations of parameter values that provide almost equally good fits to the observed data. Thus, changing the calibration period or the goodness-of-fit measure results in an altered ranking of parameter sets to fit the observations. Consequently, there is no single parameter set (or model structure) that serves as the *characteristic* parameter input for any given area, but there is a certain degree of model equifinality involved when reproducing observations with model predictions (Beven, 1996). Therefore, given the issues with multiple (interacting) parameter values, measurement scales, spatial and temporal heterogeneity or the dependence on the model structure, there can never be a single set of parameter values for the calibration process that represents an optimum for the study area, but calibration can contribute to the reduction of range in the possible parameter space.

As a result, this is a field where probabilistic model output really shines, as it expresses the entire model spread with its inherent uncertainties not in absolute terms, but shows the relative performance of a model with respect to observational data. Many decision makers and practitioners in all kind of earth science related fields still favor absolute model output, especially in areas where public policy and public safety is at stake. Unfortunately, certainty is an illusion and ultimately the reason for modeling: the lack of full access, either in time or space, to the phenomena of interest (Oreskes et al., 1994). In practice, there are many measures that attempt to validate probabilistic forecasts. Some are better, some less suitable for distributed model output that is commonly the main form of data representation in landslide early warning. Without going into detail in this paper, we

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences Discussions Open Access EGU

highlight the work of Mason and Graham (1999) and the WMO (2012) that mention a few skill scores suitable for probabilistic outcomes.

7. Case Study

In a simplified ensemble modelling approach applied to a larger study area in Austria (approx. 1366 km²), this specific case study aims to investigate a) how equifinality influences modelling outcome with purely literature based geotechnical parametrization, b) which ways of visual representation are viable for presenting probabilistic data, and c) how infrastructure data can further supplement early warning procedures in an exposure context.

7.1 Study Area

The Rhenodanubian Flyschzone (RDF) in the federal state of Lower Austria stretches over approx. 130 km in a SW–NE striking direction. The study area is limited to this geological zone in order to keep the subsurface as homogeneous as possible (Fig. 1). The Cretaceous–early Tertiary RDF is located in the northern foothills of the East Alps, in between the Molasse basin to the North and the Northern Calcareous Alps to the South. The RDF is a paleogeographic-tectonic unit as part of the oceanic Penninic zone that was to a large part eliminated in the subduction process involved in the Alpine orogeny (Hesse, 2011). Flysch materials in the RDF are typically deeply weathered and mainly consist of alterations of pelitic layers (clayey shales, silty shales, marls) and sandstones. Physiographically, the RDF can be characterized as a low mountain region with a highly undulating terrain. It is exceptionally prone to landsliding, exhibiting around five landslides per km2 (Petschko et al., 2014). Heavy rainfall events (exceeding 100 mm per day) as well as rapid snowmelt are considered to be the main triggering factors for slope failure in the region (Schwenk, 1992, Schweigl and Hervás, 2009).

Figure 1: (A) Location of the Rhenodanubian Flyschzone in Lower Austria (DEM: CC BY 3.0 AT–Federal state of Lower Austria); (B) Typical earth slide in Lower Austria after a heavy rainfall event in May 2014 (Picture: K. Gokesch).

7.2 Modeling Approach

7.2.1 TRIGRS

Physically based models used to be attributed to local scale applications (e.g. Corominas et al., 2014, van Westen et al., 2008) because of their computational requirements and data constraints. This has clearly shifted in the last couple of years and by now, physically based models can be quite commonly found to evaluate rainfall-induced landslide susceptibility at the regional scale. The majority is infinite-slope model based with only a few necessary input parameters to be suitable at a regional scale. Increasing the physical basis of a model comes at the cost of introducing even more parameters, while the available data for calibration does not increase at the same time and could lead to problematic overparameterization (Beven, 1996). Even the simplest infinite-slope stability models generally require more parametrization than can be justified by available data.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

However, there are some general features of hillslope hydrology that are relevant to slope instability that can be considered to a certain degree by infinite-slope models: vertical infiltration, dependence of infiltration on initial soil moisture conditions, varying time scales for infiltration and lateral flow (Baum et al., 2010). As a result, TRIGRS (*transient rainfall infiltration and grid-based regional slope-stability analysis*; refer to Baum et al. (2008) for details), which we use in this case study, offers a good trade-off between model complexity and flexibility while we acknowledge the availability of other dynamic, physically based models that were applied at a regional scale, such as STARWARS/PROBSTAB (Kuriakose et al., 2009) or r.slope.stability (Mergili et al., 2014a). Raia et al. (2014) with their TRIGRS_P model and Salciarini et al. (2017) with their PG_TRIGRS model have already attempted a probabilistic TRIGRS derivative in the recent past that gave us the confidence to use TRIGRS in an automated probabilistic approach.

TRIGRS was specifically developed for modeling the potential occurrences of shallow landslides by incorporating transient pressure response to rainfall and downward infiltration processes (Baum et al., 2008). Initial soil conditions are assumed either saturated or tension-saturated. TRIGRS computes transient pore-pressure changes to find analytical solutions to partial differential equations, representing one-dimensional vertical flow in isotropic, homogeneous materials due to rainfall infiltration from rainfall events with durations ranging from hours to a few days. It uses a generalized version of Iverson's (2000) infiltration model solution to the boundary problem posed by Richard's equation. This solution assesses the effects of transient rainfall on the timing and location of landslides by modeling the pore water pressure of a steady component and a transient component (Liao et al., 2011). However, the model is limited by its distributed one-dimensional modeling approach with noninteracting grid cells and its simplified soil-water characteristic curve (Baum et al., 2010). The entire theoretical basis together with all model related assumptions and equations can be found in Baum et al. (2008, 2010). TRIGRS computes a factor of safety (FoS) for each grid cell based on an infinite-slope model. It allows for the implementation of spatially varying raster input (e.g. rainfall, property zones, soil depth, infiltration, etc.) to account for horizontal heterogeneity. The FoS can generally be referred to as the ratio of resisting forces (the resisting basal Coulomb friction) over driving forces (the gravitationally induced downslope basal driving stress) on the potential failure surface, with a FoS < 1.0 indicating slope instability and a FoS ≥ 1 slope stability respectively.

7.2.2 Model Setup

The probabilistic modeling setup is realized entirely in an open source framework. This was done not only to make it as easily reproducible as possible, but also because it offered the largest flexibility. TRIGRS, which itself is open source, is operated by providing input text files that contain many lines. Those input files are used to specify the numerical values of the input parameters, the location of the input raster files in the filesystem, and all other relevant grids to be considered (e.g. spatially distributed rainfall maps, different property zones to subdivide the study area in homogenous regions, spatially distributed soil depth maps, etc.). We used python programming language in a script for all string formatting procedures that receives its data from an initialization file. That python script is also used for parsing the raw input into variables usable for TRIGRS. User provided arguments in this initialization file hold the number of property zones needed, the rainfall duration pattern, number

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences Discussions

of timesteps and all variables that are used for the probabilistic treatment of parameters, such as min/max values for soil depth, effective cohesion and effective friction angle as well as the number of model runs. The most recent rainfall input can be automatically imported by predefined naming conventions.

We used the GDAL package (GDAL Development Team, 2017) for reading and writing raster files and the NumPy package (van der Walt et al., 2011) for all raster calculations in the python script. Based on the number of predefined model runs, for each run a single deterministic output is generated based on the selected input parameters derived randomly from a normal distribution. We computed 25 model runs for each hour which resulted in 25 equally probable model results based on the different input parameters. After the initial deterministic model run, a new file is updated after each iteration that is used as the probability of failure (PoF) output. It tracks for each raster cell the initial value of the deterministic factor of safety output, and in case a cell holds a FoS < 1.0 (unstable cell), the corresponding PoF raster cell receives this information by diving the count of unstable raster cells by the number of model runs in order to calculate a probability value for this raster cell to fail at this location given the different input parameters. All used variables, deterministic model outputs (the FoS maps) and the probabilistic model output (the PoF map) are parsed through to R (R Core Team, 2017). In R, all piped arguments from the python script are used for producing ready-to-use maps (packages: rgdal (Bivand et al., 2017), sp (Pebesma and Bivand, 2005)) or to visualize performance measures such as ROC plots (package: ROCR (Sing et al., 2005)). The entire procedure from importing raw data to producing usable maps is fully automated within an executable file that may be initiated every hour. This open code structure is flexible enough to enable the direct implementation of the most recent available data (rainfall data, soil moisture data, etc.) with minimal effort and thus makes it a useful tool in considering data assimilation techniques.

7.2.3 Parametrization

Model parametrization over large areas is a difficult task given the poor spatial comprehension of the spatial organization of involved geotechnical and hydraulic input parameters. Tofani et al. (2017) performed 59 site investigations to parametrize their distributed slope stability model. This amount of in situ soil samplings with associated lab measurements is exceptional and a great source to determine the prescribed probability density function of all measured parameters, especially since all measurements from all sampling sites were published. Although Tofani et al. (2017) ultimately used the median value for each lithological class, the boxplots suggested normal to lognormal parameter distributions. This is a common observation and might be a result of the central limit theorem, which indicates that lumping data from many different sources (i.e. different in situ soil sampling sites in this case) tends to result in a normal or lognormal distribution (Wang et al., 2015). This gives us confidence to use plausible parameter ranges with a normally distributed state function based on geotechnical textbooks to characterize soils in our study area.

In accordance to the generalized likelihood uncertainty estimation (GLUE) methodology proposed by Beven and Binley (1992), we use a simple Monte Carlo simulation of multiple randomly chosen parameter sets within a predefined parameter range and within a single model structure as the basis for incorporating the inherent parameter uncertainties. Parameters that are considered in a probabilistic way are soil depth, effective cohesion and the effective friction angle (Fig. 2). We assume

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

fully saturated conditions (θ = 40%) and slope-parallel groundwater flow for the sake of simplicity and given the absence of appropriate initial water conditions. Using all this information, it is now possible to have a spatially distributed probabilistic assessment of the FoS, expressed as the probability of failure (PoF). As TRIGRS is capable of calculating the increase in pore water pressure within the soil, the result is a distributed representation of the decrease in shear strength until slope failure (FoS < 1.0) is reached at a certain depth.

Figure 2: Probabilistically derived modeling parameters based on random sampling from a normally distributed state function. Jittering dots (to prevent overplotting) indicate individual samples within a plausible parameter range.

The raster cell size of the DEM to derive all model relevant topographical parameters used in this case study, is 10 meters. This cell size allows for a sufficiently high representation surface topography without losing too much information through surface aggregation and smoothing. For the rainfall input, three hourly timesteps were applied with spatially distributed rainfall raster maps representing hourly rainfall based on automated geostatistical interpolation (the methodology is described in detail in Canli et al. (2017)). Using interpolated rainfall input is sufficient as a proof on concept for this case study, but this can be immediately exchanged for any other raster input, such as numerical weather predictions, in a real-time application. The selection of hourly rainfall input as well as the decision to choose a three-hour timeframe to force the model was made arbitrarily as for the study area there are no published information available on the hydrological response of landslides to rainfall. The spatial resolution of 1 km for the rainfall input was resampled to match the cell size of the DEM, which is a prerequisite of TRIGRS.

7.3 Results

Fig. 3 shows the results for 24 model iterations for the same time based on spatially distributed, hourly rainfall input over the last three hours. Each ensemble member was initialized with probabilistically derived parameters that are displayed on each map. The WMO (2012) describes this form of EPS representation *postage stamp map* that shows each individual ensemble member which allows the forecaster to view the scenarios in each member forecast. The results indicate quite significant changes across individual members, but also quite high similarities although parameters change drastically between some of the members. For example, a depth of 2.5 m, an effective cohesion of 13.4 Nm^{-2} and an effective friction angle of 35 degree in one of the deterministic outputs reveal almost the identical FoS distribution with a depth of 2.0 m, an effective cohesion of 5.4 Nm^{-2} and an effective friction angle of 22.7 degree.

Figure 3: Postage stamp map for 24 model iterations for the same time. Each ensemble member was initialized with altered parameters within a plausible range to account for variability and spatial uncertainty. Factor of Safety (FoS) values < 1 indicate slope instability.

By using a probabilistic representation, this variability and uncertainty is accounted for. Here, the probability is estimated as a proportion of the ensemble members that predict an event to occur (FoS < 1.0) at a specific raster cell. For example, a

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences Discussions

probability between 0.75 and 1.0 means that a specific raster cell, under varying input parameters, indicates slope failure in 75% to 100% of all model runs for this specific time.

To provide additional information, which supports different actors responsible to manage landslide hazards, the PoF is underlain with accurately mapped building polygons and roads for a direct exposure visualization of the elements at risk towards landslides (Fig. 4). Buildings and roads are imported from the freely accessible OpenStreetMap (OSM) database. OSM covers almost the entirety of existing buildings in Austria and is based off official Austrian administrative data, which stands under an open government data (OGD) license. Building exposure is a result of a simple spatial join that assigns each building the highest PoF value within 25 m. This value, while arbitrarily chosen, further accounts for spatial uncertainties since TRIGRS models only the location of actual landslide initiation. High building exposure along the river is a modeling artifact introduced by the steep retaining wall and the associated sudden and steep decline in slope angle. Results of the PoF map suggest quite a narrow ensemble spread, which means that the different input parameters indicate an expression of equifinality. This can be considered as some kind of *spatial confidence buffer* that gives some reliance that under varying rainfall forcing the location of possible slope failure is modelled quite consistently at the more or less same location.

Figure 4: Probability of Failure depicted as a proportion of the ensemble members that predict an event to occur (FoS < 1.0). Building exposure to current slope failure predictions adds an additional information layer for decision makers. Buildings and roads are imported from the freely accessible OpenStreetMap (OSM) database (© OpenStreetMap contributors).

8 Discussion

Since landslides generally tend to occur in steeper slopes (Liao et al., 2011), this spatial confidence buffer modelled in the probabilistic approach presented here could partially alleviate two issues: a) reduce the influence of positionally imprecise landslide inventory data in the calibration process since a larger slope proportion reveals instability; b) reduce the false alarm ratio since landslide locations are more likely to be situated within a certain slope failure probability segment (as would be the case in Liao et al., 2011 for example). In this case study, we can only perform some kind of qualitative validation for the following reasons: a) for Lower Austria a very comprehensive and spatially accurate landslide inventory based on high-resolution airborne LiDAR based DEM mapping exists (Petschko et al., 2015), however, it does not contain any temporal information; b) the Building Ground Registry (BGR) is the most comprehensive source of reported damage causing landslides in Austria, however, its spatial and temporal accuracy is insufficient for physically based model calibration and validation. Qualitative validation by visual comparison (Fig. 5) indicate, for this specific time and under the given rainfall input, that there is an agreement between some of the landslide initiation points and areas of high failure probability.

For personnel responsible to manage landslides in a given region, however, this situation would be quite challenging in order to take appropriate action. The probabilistic approach depicts spatial variability and uncertainty much better than any purely deterministic result, yet there are still many unaccounted uncertainties involved with respect to actual slope failure prediction. Thus, a map representation of slope failure probability at such high spatial resolution could suggest a certainty that simply is

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

not achievable in landslide modeling. It has to be stressed that this probabilistic approach does not eliminate uncertainty, but it explicitly introduces it into the model results. This is quite detrimental to the ultimate goal of predictive modeling: to be a positionally and temporally accurate mitigation tool. Salciarini et al. (2017) points out that such a tool can be suitable for a first susceptibility screening of an area prone to landsliding, but not for single slope/single landslide analyses. Since such a map reveals a high degree of spatial discontinuity in its spatial prediction pattern, this undoubtedly puts the forecaster at risk of missing some real landsliding occurrences. This raises the question whether putting high efforts into probabilistic landslide forecasting is warranted compared to a combination of statistical susceptibility maps with an early warning approach including empirical-based rainfall thresholds (see conclusions *Challenge 2: Rare events and model averaging*). Kirschbaum et al. (2012) present such a nowcasting attempt at a regional and global scale by using remotely sensed precipitation data.

Figure 5: Probability of failure map detail for a specific time under prevailing rainfall conditions. Known historic landslide initiation points (ellipses) partly overlap with current slope stability conditions. However, high spatial resolution, and therefore a high degree of spatial discontinuity, poses a risk for missing many real landslide events in an early warning situation.

This spatial confidence buffer that indicates a rather narrow ensemble spread is an equifinal result of the main predetermining factor: slope angle. Neves Seefelder et al. (2016) and Zieher et al. (2017) identified slope angle as one of the most sensitive modeling parameter in TRIGRS, which is not surprising since slope failures are in general associated with higher slope angles (Liao et al., 2011). Therefore, no matter what the geotechnical or hydraulic input parameters are, it will be always the same slope segments that will result the highest slope failure probability. Slope failure probability will ultimately vary only based on the dynamic component (here: rainfall) or if a spatially distributed soil depth map is provided. The ensemble members in Fig. 3 indicate very similar results under greatly varying input parameters because of equifinality. This raises the question if model calibration is physically advisable or if we could draw useful conclusions from the direct model output alone (see conclusions *Challenge 1: Parameter uncertainties at regional scale modeling*).

Deterministic forecasts suppress information and judgement about uncertainty. They generally pretend to be absolute based on an optimal set of input parameters. Empirical approaches, such as the commonly used rainfall thresholds in landslide early warning applications, started to incorporate estimates of uncertainty only recently by defining rainfall thresholds at different exceedance probabilities (e.g. Melillo et al., 2016, Piciullo et al., 2017), yet they rely on very good landslide event catalogues and thus purely on past reportings, which adds a tremendously large source of error (Peres et al., 2017). Gariano et al. (2015) found that an underestimation of only 1% in the number of considered landslides can result in a significant decrease in the performance of a threshold based landslide EWS. Additionally, rainfall thresholds represent a simplification of the underlying physical processes by establishing purely a relationship between rainfall and landslide occurrence (Bogaard and Greco, 2017). Both, deterministic and empirical approaches may create the illusion of certainty in a user's mind, which can easily lead to wrong conclusions. Krzysztofowicz (2001) mentions a notable event in the spring of 1997, where a falsely issued deterministic forecast on the Red River in Grand Forks, North Dakota, led to evacuations and left a devastated city. After the event a City Council Member in Grand Forks stated (p. 3): "…the National Weather Service continued to predict that the river's crest at

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences Discussions

Grand Forks would be 49 ft… If someone had told us that these estimates were not an exact science,… we may have been better prepared."

Based on these words, the presumption that hiding the predictive uncertainty behind the façade of a precise estimate serves better the public need is wrong and careless. Concerns about the acceptance of probabilities in decision making turned out to be unwarranted (Krzysztofowicz, 2001). Based on our observations we found that many published landslide studies dealing with physically based hindcasting applications rely too strong on purely number based validation outputs. Deterministic results are taken as given when the modellers achieve *satisfactory* results based on the model validation, without defining what the criteria are for this satisfaction or when this state of satisfaction is reached. Beven (1996) argues that this is generally owed to relativism, when there is a need to adopt less stringent criteria of acceptability or to acknowledge that it is not possible to predict all the observations all the time (with common arguments ranging from scale issues, spatial heterogeneity, uncertainty in model structure or process understanding, etc.). In all other cases, probabilistic approaches should be prioritized since they allow not only for the incorporation of parametric uncertainties, but also facilitate the geomorphic plausibility control in the absence of proper calibration/validation data. However, narrowing down uncertainties is a good first step, but not the be-all and end-all of ensemble approaches. It is the differences that matter between model predictions and determining and unpicking those differences should be the ultimate goal of ensemble approaches which requires high quality data (Challinor et al., 2014). The scarcity of such high-quality data in landslide research is well known. The potential of local-scale studies to draw conclusions for a larger scale (e.g. Bordoni et al., 2015) remains to be a very important field of study in the near future. In this regard, data assimilation might be a key factor for producing accurate model predictions while reducing those inherent uncertainties. Data assimilation can be referred to as (real-time) parameter updating with observations of flow, soil moisture, groundwater, displacement or rainfall (continuously measured through e.g. radar, rain gauges, etc.) and appropriate uncertainty modeling to correct model predictions (Collier, 2007, Reichle, 2008). Liu et al. (2012) give an in-depth review on the current state of data assimilation applications in both, hydrologic research and operational practices that are in many parts valid for landslide prediction too. While there are a few adaptive systems in landslide early warning based on empirical thresholds (e.g. the SIGMA early warning system in Italy (Martelloni et al., 2012, Segoni et al., 2017)), there are none that use physically based predictions with blends of most recent QPEs or other independent observations. For extreme events, this might be key if the probability of extreme floods or landslides occurring is continuously and objectively evaluated and updated in real-time, especially when it comes to assimilating new observations from multiple sources across a range of spatiotemporal scales (Liu et al., 2012).

9 Conclusions

We would like to conclude this paper by raising awareness for a couple of technical and conceptual challenges the landslide forecasting community has to face in the near future. Since physically based, probabilistic landslide forecasting is still in its

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences
Discussions

infancy, we refrain from addressing challenges in operational practices that are currently discussed in hydrological forecasting (e.g. Pagano et al., 2014), but are of equal importance for the operational use of landslide forecasting nonetheless.

Challenge 1: Parameter uncertainties at regional scale modeling

Current practices for geotechnical parametrization in physically based landslide modeling include the application of averaged values from in situ measurements (e.g. Thiebes, 2014, Tofani et al., 2017, Zieher et al., 2017) or using values from existing databases, lookup tables or other published/unpublished sources (e.g. Schmidt et al., 2008, Kuriakose et al., 2009, Mergili et al., 2014b). In the landslide research community, probabilistic treatment of input parameters for regional model application has seen a rise only in the last couple of years. Probabilistic approaches allow for a more thorough consideration of uncertainties and inherent variability of model specific parameters. Spatially varying parameters (both geotechnical and hydraulic) are usually represented as univariate distributions of random variables based on an underlying probability density function and statistical characteristics (Fan et al., 2016). Friction angle and cohesion are commonly considered as such varying variables that are treated in a probabilistic way for model parametrization (e.g. Park et al., 2013, Chen and Zhang, 2014, Raia et al., 2014, Salciarini et al., 2017). Interestingly, in hydrological streamflow prediction the parameter uncertainty of the hydraulic model is often neglected in favor of a deterministic parameter input. This is explained by the superior proportion of total estimation uncertainty introduced by the weather predictions alone, which blurs the streamflow variability that the meteorological input data cannot explain (Alfieri et al., 2012b).

Measuring geotechnical and hydrological parameters for large areas is difficult, time-consuming, and expensive. Therefore, applying spatially distributed physically based models with spatially variable geotechnical parameters is not straightforward and it is impossible to find an approach that is universally accepted (Tofani et al., 2017). Even if there is a sufficiently large amount of measured values available for one, some or even all parameter values in a model up to the point that it is possible to specify distributions and covariances for the parameter values, there remain some methodological obstacles. For example, there is no guarantee that values measured at one scale will reflect the effective values required in the model to achieve satisfactory predictions of observed variables (Beven and Freer, 2001). At larger scales (e.g. > 1:25,000), there are several factors that cause spatial variation of, for example, soil water content, topography, differences in soil depth, -type and -texture, vegetation characteristics, as well as rainfall patterns. Additionally, spatially varying soil and hydraulic properties are influenced by interrelated soil formation processes (such as weathering processes, biological perturbations, atmospheric interactions) (Fan et al., 2016), and thus making selective in situ soil sampling a tricky task when performed at a larger scale.

Small scale (e.g. < 1:10,000) variability usually lacks a spatial organization, hence its representation as stochastic process. The larger the scale, however, the more soil forming processes manifest a persistent deterministic signature due to the predetermined geology, topography, climate, etc. (Seyfried and Wilcox, 1995, Fan et al., 2016). Neves Seefelder et al. (2016) suggest applying rather broad ranges of parameters for physically based approaches to be on the "safe side" as they yield results comparable in quality to those derived with best-fit narrow ranges. By acknowledging the fact that geotechnical and hydrological parameters – when applied on a larger scale – are highly variable, uncertain and often poorly known, narrow

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

parameter ranges or even singular combinations of parameters come with the risk of being off target (Neves Seefelder et al., 2016). This basically implies that, when working at a regional scale and beyond, an actual parametrization with in situ measured samples might not be necessary at all when using literature values instead. This could mean enormous savings in time and money spent, yet this clearly needs further research to evaluate whether there is and to what degree the benefits of actual sampled in situ data are compared to just utilizing literature values in broad ranges when modeling at larger scales.

Challenge 2: Rare events and model averaging

Like flood events, landslides types with a rapid onset can indeed be considered as an extreme event. Hereby, extreme does not necessarily refer to huge displaced landslide volumes – also small landslides might be considered as extreme in terms of potential consequences. While it is possible to continuously monitor and forecast regular streamflow, extreme events are scarce which makes model calibration and, consequently, forecasting a real challenge. We argue that this is even more so the case for landslides since there are no directly observable target variables to be monitored at a regional scale. Landslide models can only be calibrated on a case by case basis. Shallow landslides are one of the most common landslide types (van Asch et al., 1999). While they occur quite in abundance when looking at their spatial distribution, they are typically low-frequency events. And most of them do occur in so called 'low-risk' environments as defined by Klimeš et al. (2017): low annual frequency of landslides; the majority of the landslides are of small size and are low impact events. Due to the scarcity of such extreme events, Collier (2007) argues that such events may lie outside of what model calibration if capable of providing for forecasting approaches. Commonly, calibration will improve the reliability of forecasts (i.e. the match of the target variable or forecast probabilities to frequency of observations of the event) but reduce the resolution of the forecast (the ability to discriminate whether an event will occur or not). Consequently, calibration will improve forecasts of common events, but reduces the probability of forecasting more extreme events.

The WMO (2012) argues that this is the case when events are rare, since the statistical distributions are trained to the more common events. For rare events, hence, calibration cannot be expected to provide significant improvement over the raw forecasts. Therefore, it is very difficult to *validate* a model for future use, as it can be only continually evaluated in the light of the most recent data (Oreskes et al., 1994, Challinor et al., 2014). And landslides are per se extreme events with no *common* events attributed to them as they only occur under exceptional circumstances given the environmental interactions involved.

This raises the question if any averaged model output, and that is by definition every model output based on model calibration from past events, will ever be able to precisely forecast extreme events at the regional scale. The sensitivity of the model had to be lowered in a way that much larger areas of slope failure need to be forecasted to catch a few real extreme events at the cost of significantly raising the number of false alerts. This is especially the case when engineering conservatism comes into play in decision making, thus leaving probabilistic forecasting attempts in a nonsuperior state over purely deterministic approaches. This is a known issue (e.g. Baum et al., 2010) in a way that FoS computations usually are more likely to identify areas prone to slope failure during a given rainfall event rather than predicting exact locations of specific landslides. A term

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

such as landslide susceptibility forecasting seems more appropriate in that case. Our results in the Flyschzone of Lower Austria seem to point in that direction so far. This is definitely an issue that needs far more in-depth research in the future.

What else has to be kept in mind are the technical specifications of the modeling approach for slope stability analysis at a regional scale. The most commonly applied modeling approach relies on the infinite-slope stability model which reduces the landslide geometry to a slope-parallel layer of infinite length and width. Modeling approaches that try to introduce more complex landslide geometries in a GIS environment are generally outperformed by the infinite-slope stability model (Zieher et al., 2017). Consequently, parameters representing the landslide geometry assumed by the model (i.e. slope angle and depth) are highly sensitive (Zieher et al., 2017). This means that the underlying model itself already performs some sort of averaging too since the precise landslide geometry cannot be adequately resolved in the infinite-slope stability model.

Challenge 3: Computational burden

In literature, physically based approaches for modeling rainfall-induced shallow landslides were suggested to be applied to smaller scale study areas while statistical based approaches were recommended for larger scale susceptibility assessments (e.g. van Westen et al., 2006, Corominas et al., 2014). One reason usually mentioned is the poor comprehension of the spatial organization of the geotechnical and hydraulic input parameters (e.g. Tofani et al., 2017, Park et al., 2013). However, as outlined above, it does not make too much difference whether the underlying study area is 50 km² or 5000 km² investigated at a scale of 1:1,000 or 1:25,000 — the model is still influenced by errors or uncertainties from the input parameters to the same degree given the fact how input parameters are derived. Therefore, one major drawback used to be the computational costs involved when modeling physically based at a regional scale. Because as soon as computational power was available at reasonable costs, the area size, associated with a high-resolution DEM, steadily increased over time for physically based applications and currently exceeding thousands of square kilometers (e.g. Tofani et al., 2017, Alvioli and Baum, 2016).

Recent landslide model development is aiming towards featuring multithreading and parallelization. Since high resolution DEM are available in many parts of the world, the computational demands increased significantly, especially when applied in a dynamic/time-dependent modeling framework. Parallelization has great potential in grid-based landslide modeling, especially for the time-consuming hydraulic model components, for several reasons: in case of TRIGRS, for example, which is a coupled slope stability and hydraulic model, only excess water from infiltration is directed to the neighboring cells which makes it the only variable that relies on explicit neighborhood relations. This needs to be done only once, however. Vertical groundwater flow and one-dimensional slope stability in a two-dimensional array of noninteracting columns can subsequently be computed independently for each cell, which is a prime example for parallelization purposes (Baum et al., 2010, Alvioli and Baum, 2016). Besides TRIGRS v2.1, which received its parallel implementation by Alvioli and Baum (2016) only recently, other models for physically based landslide applications are using a parallelized module: NewAge-JGrass (Formetta et al., 2016) or r.slope.stability (Mergili et al., 2014a).

In our case study, the computational time for one model iteration is about 45 minutes, which is far too long for computing a large set of different ensemble members in an operational real-time application. We did not yet use the parallel implementation

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

of TRIGRS on our regular commodity machine (3.40GHz quad-processor equipped with 32GB RAM), but Alvioli and Baum (2016) reported that parallel computation on a multi-core node already led to a significant speedup compared to a single-node local machine. When applied to a HPC (High-Performance Computing) cluster, they were even able to reduce running time from one day to one hour. This allows the exploration of new possibilities in how landslide forecasting can be approached in the future. While HPC applications are common in meteorological (Bauer et al., 2015) and hydrological forecasting (Shi et al., 2015), this is a field clearly underexploited in the field of landslide forecasting. This opens up possibilities to encompass fine tuning of input parameters by means of multiple model runs, probabilistic applications and, first and foremost, real-time applications with a continuous consideration of antecedent and forecasted rainfall information (Alvioli and Baum, 2016).

Acknowledgements

The authors thank the Provincial government of Lower Austria for their support.

Compliance with ethical standards.
The authors declare that they have no conflict of interest.

References

Alfieri, L., Salamon, P., Pappenberger, F., Wetterhall, F. and Thielen, J.: Operational early warning systems for water-related hazards in Europe, Environmental Science & Policy, 21, 35–49, doi:10.1016/j.envsci.2012.01.008, 2012a.

Alfieri, L., Thielen, J. and Pappenberger, F.: Ensemble hydro-meteorological simulation for flash flood early detection in southern Switzerland, Journal of Hydrology, 424–425, 143–153, doi:10.1016/j.jhydrol.2011.12.038, 2012b.

Althuwaynee, O. F., Pradhan, B., Park, H.-J. and Lee, J. H.: A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping, CATENA, 114, 21–36, doi:10.1016/j.catena.2013.10.011, 2014a.

Althuwaynee, O. F., Pradhan, B., Park, H.-J. and Lee, J. H.: A novel ensemble decision tree-based CHi-squared Automatic Interaction Detection (CHAID) and multivariate logistic regression models in landslide susceptibility mapping, Landslides, 11(6), 1063–1078, doi:10.1007/s10346-014-0466-0, 2014b.

Alvioli, M. and Baum, R. L.: Parallelization of the TRIGRS model for rainfall-induced landslides using the message passing interface, Environmental Modelling & Software, 81, 122–135, doi:10.1016/j.envsoft.2016.04.002, 2016.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M. and Reinhardt, T.: Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities, Monthly Weather Review, 139(12), 3887–3905, doi:10.1175/MWR-D-10-05013.1, 2011.

Bartholmes, J. and Todini, E.: Coupling meteorological and hydrological models for flood forecasting, Hydrology and Earth System Sciences, 9(4), 333–346, doi:10.5194/hess-9-333-2005, 2005.

Bauer, P., Thorpe, A. and Brunet, G.: The quiet revolution of numerical weather prediction, Nature, 525(7567), 47–55, doi:10.1038/nature14956, 2015.

Baum, R. L., Savage, W. Z. and Godt, J. W.: TRIGRS - A Fortran Program for Transient Rainfall Infiltration and Grid-Based Regional Slope-Stability Analysis, Version 2.0, 2008.

Baum, R. L. and Godt, J. W.: Early warning of rainfall-induced shallow landslides and debris flows in the USA, Landslides, 7(3), 259–272, doi:10.1007/s10346-009-0177-0, 2010.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Baum, R. L., Godt, J. W. and Savage, W. Z.: Estimating the timing and location of shallow rainfall-induced landslides using a model for transient, unsaturated infiltration, Journal of Geophysical Research, 115(F3), doi:10.1029/2009JF001321, 2010.

Beven, K. and Binley, A.: The future of distributed models: model calibration and uncertainty prediction, Hydrological processes, 6(3), 279–298, doi:10.1002/hyp.3360060305, 1992.

Beven, K.: Equifinality and Uncertainty in Geomorphological Modelling, in: The Scientific Nature of Geomorphology: Proceedings of the 27th Binghamton Symposium in Geomorphology, held 27-29 September, 1996, John Wiley & Sons, 27, 289–313, 1996.

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, Journal of hydrology, 249(1), 11–29, doi:10.1016/S0022-1694(01)00421-8, 2001.

Bivand, R., Keitt, T., Rowlingson, B., Pebesma, E., Sumner, M., Hijmans, R. and Rouault, E.: rgdal: Geospatial Data Abstraction Library. Available at: https://cran.r-project.org/package=rgdal, last access: 30 November 2017.

Bogaard, T. and Greco, R.: Invited perspectives. A hydrological look to precipitation intensity duration thresholds for landslide initiation: proposing hydro-meteorological thresholds, Natural Hazards and Earth System Sciences Discussions, 1–17, doi:10.5194/nhess-2017-241, 2017.

Bordoni, M., Meisina, C., Valentino, R., Bittelli, M. and Chersich, S.: Site-specific to local-scale shallow landslides triggering zones assessment using TRIGRS, Natural Hazards and Earth System Science, 15(5), 1025–1050, doi:10.5194/nhess-15-1025-2015, 2015.

Canli, E., Loigge, B. and Glade, T.: Spatially distributed rainfall information and its potential for regional landslide early warning systems, Natural Hazards, doi:10.1007/s11069-017-2953-9, 2017.

Challinor, A., Martre, P., Asseng, S., Thornton, P. and Ewert, F.: Making the most of climate impacts ensembles, Nature Climate Change, 4(2), 77–80, doi:10.1038/nclimate2117, 2014.

Chen, H. X. and Zhang, L. M.: A physically-based distributed cell model for predicting regional rainfall-induced shallow slope failures, Engineering Geology, 176, 79–92, doi:10.1016/j.enggeo.2014.04.011, 2014.

Chen, Y., Chen, D., Li, Z., Wu, Y. and Huang, J.: An ensemble prediction model for rainfall-induced landslides and its preliminary application, Chinese Journal of Atmospheric Sciences, 40(3), 515–527, doi:10.3878/j.issn.1006-9895.1503.15120, 2016.

Cloke, H. L. and Pappenberger, F.: Ensemble flood forecasting: A review, Journal of Hydrology, 375(3–4), 613–626, doi:10.1016/j.jhydrol.2009.06.005, 2009.

Collier, C. G.: Flash flood forecasting: What are the limits of predictability?, Quarterly Journal of the Royal Meteorological Society, 133(622), 3–23, doi:10.1002/qj.29, 2007.

Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J.-P., Fotopoulou, S., Catani, F., van ven Eeckhaut, M., Mavrouli, O., Agliardi, F., Pitilakis, K., Winter, M. G., Pastor, M., Ferlisi, S., Tofani, V., Hervás, J. and Smith, J. T.: Recommendations for the quantitative analysis of landslide risk, Bulletin of Engineering Geology and the Environment, 73(2), 209–263, doi:10.1007/s10064-013-0538-8, 2014.

Crozier, M. J.: Deciphering the effect of climate change on landslide activity: A review, Geomorphology, 124(3–4), 260–267, doi:10.1016/j.geomorph.2010.04.009, 2010.

Fan, L., Lehmann, P. and Or, D.: Effects of soil spatial variability at the hillslope and catchment scales on characteristics of rainfall-induced landslides, Water Resources Research, 52(3), 1781–1799, doi:10.1002/2015WR017758, 2016.

Formetta, G., Capparelli, G. and Versace, P.: Evaluating performance of simplified physically based models for shallow landslide susceptibility, Hydrology and Earth System Sciences, 20(11), 4585–4603, doi:10.5194/hess-20-4585-2016, 2016.

Gariano, S. L., Brunetti, M. T., Iovine, G., Melillo, M., Peruccacci, S., Terranova, O., Vennari, C. and Guzzetti, F.: Calibration and validation of rainfall thresholds for shallow landslide forecasting in Sicily, southern Italy, Geomorphology, 228, 653–665, doi:10.1016/j.geomorph.2014.10.019, 2015.

Gariano, S. L., Rianna, G., Petrucci, O. and Guzzetti, F.: Assessing future changes in the occurrence of rainfall-induced landslides at a regional scale, Science of The Total Environment, 596–597, 417–426, doi:10.1016/j.scitotenv.2017.03.103, 2017.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences Discussions

GDAL Development Team: GDAL - Geospatial Data Abstraction Library, Version 2.1.3, Open Source Geospatial Foundation. Available at: http://www.gdal.org, last access: 30 November 2017.

Gebhardt, C., Theis, S. E., Paulat, M. and Ben Bouallègue, Z.: Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbations and variation of lateral boundaries, Atmospheric Research, 100(2–3), 168–177, doi:10.1016/j.atmosres.2010.12.008, 2011.

Glade, T.: Modelling landslide-triggering rainfall thresholds at a range of complexities, edited by E. Bromhead, N. Dixon and M.-L. Ibsen: Proceedings of the 8th International Symposium on Landslides.- 26.-30. June 2000, Cardiff, 2, Thomas Telford, 633–640, 2000.

Glade, T. and Crozier, M.J.: A review of scale dependency in landslide hazard and risk analysis, edited by T. Glade, M. Anderson and M. Crozier: Landslide hazard and risk, Wiley, Chichester 75–138, 2015.

Golding, B., Roberts, N., Leoncini, G., Mylne, K. and Swinbank, R.: MOGREPS-UK Convection-Permitting Ensemble Products for Surface Water Flood Forecasting: Rationale and First Results, Journal of Hydrometeorology, 17(5), 1383–1406, doi:10.1175/JHM-D-15-0083.1, 2016.

Greco, R. and Pagano, L.: Basic features of the predictive tools of early warning systems for water-related natural hazards: examples for shallow landslides, Natural Hazards and Earth System Sciences Discussions, 1–31, doi:10.5194/nhess-2017-269, 2017.

Guzzetti, F., Peruccacci, S., Rossi, M. and Stark, C. P.: Rainfall thresholds for the initiation of landslides in central and southern Europe, Meteorology and Atmospheric Physics, 98(3–4), 239–267, doi:10.1007/s00703-007-0262-7, 2007.

Haneberg, W. C.: A rational probabilistic method for spatially distributed landslide hazard assessment, Environmental & Engineering Geoscience, 10(1), 27–43, doi:10.2113/10.1.27, 2004.

Hapuarachchi, H. A. P., Wang, Q. J. and Pagano, T. C.: A review of advances in flash flood forecasting, Hydrological Processes, 25(18), 2771–2784, doi:10.1002/hyp.8040, 2011.

Hesse, R.: Rhenodanubian Flyschzone, Bavarian Alps, in Geological Field Trips in Central Western Europe: Fragile Earth International Conference, Munich, September 2011, edited by S. Carena, A. M. Friedrich, and B. Lammerer, Geological Society of America, Boulder, USA, 51–73, 2011.

Ikeda, K., Steiner, M., Pinto, J. and Alexander, C.: Evaluation of Cold-Season Precipitation Forecasts Generated by the Hourly Updating High-Resolution Rapid Refresh Model, Weather and Forecasting, 28(4), 921–939, doi:10.1175/WAF-D-12-00085.1, 2013.

Iverson, R. M.: Landslide triggering by rain infiltration, Water Resources Research, 36(7), 1897–1910, doi:10.1029/2000WR900090, 2000.

Kirschbaum, D. B., Adler, R., Hong, Y., Kumar, S., Peters-Lidard, C. and Lerner-Lam, A.: Advances in landslide nowcasting: evaluation of a global and regional modeling approach, Environmental Earth Sciences, 66(6), 1683–1696, doi:10.1007/s12665-011-0990-3, 2012.

Klimeš, J., Stemberk, J., Blahut, J., Krejčí, V., Krejčí, O., Hartvich, F. and Kycl, P.: Challenges for landslide hazard and risk management in "low-risk" regions, Czech Republic–landslide occurrences and related costs (IPL project no. 197), Landslides, 14(2), 771–780, doi:10.1007/s10346-017-0798-7, 2017.

Krzysztofowicz, R.: The case for probabilistic forecasting in hydrology, Journal of Hydrology, 249(1–4), 2–9, doi:10.1016/S0022-1694(01)00420-6, 2001.

Kuriakose, S. L., van Beek, L. P. H. and van Westen, C. J.: Parameterizing a physically based shallow landslide model in a data poor region, Earth Surface Processes and Landforms, 34(6), 867–881, doi:10.1002/esp.1794, 2009.

Lari, S., Frattini, P. and Crosta, G. B.: A probabilistic approach for landslide hazard analysis, Engineering Geology, 182, 3–14, doi:10.1016/j.enggeo.2014.07.015, 2014.

Lee, S. and Oh, H.-J.: Ensemble-Based Landslide Susceptibility Maps in Jinbu Area, Korea, in Terrigenous Mass Movements, edited by B. Pradhan and M. Buchroithner, Springer Berlin Heidelberg, 193–220, 2012.

Lee, J. H. and Park, H. J.: Assessment of shallow landslide susceptibility using the transient infiltration flow model and GIS-based probabilistic approach, Landslides, 13(5), 885–903, doi:10.1007/s10346-015-0646-6, 2016.

Liao, Z., Hong, Y., Kirschbaum, D., Adler, R. F., Gourley, J. J. and Wooten, R.: Evaluation of TRIGRS (transient rainfall infiltration and grid-based regional slope-stability analysis)'s predictive skill for hurricane-triggered landslides: a case study in Macon County, North Carolina, Natural Hazards, 58(1), 325–339, doi:10.1007/s11069-010-9670-y, 2011.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., Seo, D.-J., Schwanenberg, D., Smith, P., van Dijk, A. I. J. M., van Velzen, N., He, M., Lee, H., Noh, S. J., Rakovec, O. and Restrepo, P.: Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities, Hydrology and Earth System Sciences, 16(10), 3863–3887, doi:10.5194/hess-16-3863-2012, 2012.

Martelloni, G., Segoni, S., Fanti, R. and Catani, F.: Rainfall thresholds for the forecasting of landslide occurrence at regional scale, Landslides, 9(4), 485–495, doi:10.1007/s10346-011-0308-2, 2012.

Mason, S. J. and Graham, N. E.: Conditional probabilities, relative operating characteristics, and relative operating levels, Weather and Forecasting, 14(5), 713–725, 1999.

Melillo, M., Brunetti, M. T., Peruccacci, S., Gariano, S. L. and Guzzetti, F.: Rainfall thresholds for the possible landslide occurrence in Sicily (Southern Italy) based on the automatic reconstruction of rainfall events, Landslides, 13(1), 165–172, doi:10.1007/s10346-015-0630-1, 2016.

Mergili, M., Marchesini, I., Alvioli, M., Metz, M., Schneider-Muntau, B., Rossi, M. and Guzzetti, F.: A strategy for GIS-based 3-D slope stability modelling over large areas, Geoscientific Model Development, 7(6), 2969–2982, doi:10.5194/gmd-7-2969-2014, 2014a.

Mergili, M., Marchesini, I., Rossi, M., Guzzetti, F. and Fellin, W.: Spatially distributed three-dimensional slope stability modelling in a raster GIS, Geomorphology, 206, 178–195, doi:10.1016/j.geomorph.2013.10.008, 2014b.

Neves Seefelder, C., Koide, S. and Mergili, M.: Does parameterization influence the performance of slope stability model results? A case study in Rio de Janeiro, Brazil, Landslides, doi:10.1007/s10346-016-0783-6, 2016.

Oreskes, N., Shrader-Frechette, K., Belitz, K. and others: Verification, validation, and confirmation of numerical models in the earth sciences, Science, 263(5147), 641–646, doi:10.1126/science.263.5147.641, 1994.

Pagano, T. C., Wood, A. W., Ramos, M.-H., Cloke, H. L., Pappenberger, F., Clark, M. P., Cranston, M., Kavetski, D., Mathevet, T., Sorooshian, S. and Verkade, J. S.: Challenges of Operational River Forecasting, Journal of Hydrometeorology, 15(4), 1692–1707, doi:10.1175/JHM-D-13-0188.1, 2014.

Papathoma-Köhle M. and Glade T.: The role of vegetation cover change for landslide hazard and risk, edited by G. Renaud, K. Sudmeier-Rieux and M. Estrella: The Role of Ecosystems in Disaster Risk Reduction. UNU-Press, Tokyo, 293–320, 2013.

Park, H. J., Lee, J. H. and Woo, I.: Assessment of rainfall-induced shallow landslide susceptibility using a GIS-based probabilistic approach, Engineering Geology, 161, 1–15, doi:10.1016/j.enggeo.2013.04.011, 2013.

Pebesma, E. J. and Bivand, R. S.: Classes and methods for spatial data in R, R News, 5(2), 9–13, 2005.

Peres, D. J., Cancelliere, A., Greco, R. and Bogaard, T. A.: Influence of uncertain identification of triggering rainfall on the assessment of landslide early warning thresholds, Natural Hazards and Earth System Sciences Discussions, 1–28, doi:10.5194/nhess-2017-328, 2017.

Petley, D.: Global patterns of loss of life from landslides, Geology, 40(10), 927–930, doi:10.1130/G33217.1, 2012.

Petschko, H., Brenning, A., Bell, R., Goetz, J. and Glade, T.: Assessing the quality of landslide susceptibility maps – case study Lower Austria, Natural Hazards and Earth System Science, 14(1), 95–118, doi:10.5194/nhess-14-95-2014, 2014.

Petschko, H., Bell, R. and Glade, T.: Effectiveness of visually analyzing LiDAR DTM derivatives for earth and debris slide inventory mapping for statistical susceptibility modeling, Landslides, 65, doi:10.1007/s10346-015-0622-1, 2015.

Piciullo, L., Gariano, S. L., Melillo, M., Brunetti, M. T., Peruccacci, S., Guzzetti, F. and Calvello, M.: Definition and performance of a threshold-based regional early warning model for rainfall-induced landslides, Landslides, 14(3), 995–1008, doi:10.1007/s10346-016-0750-2, 2017.

Pradhan, A. M. S., Kang, H.-S., Lee, J.-S. and Kim, Y.-T.: An ensemble landslide hazard model incorporating rainfall threshold for Mt. Umyeon, South Korea, Bulletin of Engineering Geology and the Environment, doi:10.1007/s10064-017-1055-y, 2017.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org, last access: 30 November 2017.

Raia, S., Alvioli, M., Rossi, M., Baum, R. L., Godt, J. W. and Guzzetti, F.: Improving predictive power of physically based rainfall-induced shallow landslide models: a probabilistic approach, Geoscientific Model Development, 7(2), 495–514, doi:10.5194/gmd-7-495-2014, 2014.

Reichenbach, P., Cardinali, M., De Vita, P. and Guzzetti, F.: Regional hydrological thresholds for landslides and floods in the Tiber River Basin (central Italy), Environmental Geology, 35(2–3), 146–159, doi:10.1007/s002540050301, 1998.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences Discussions

Reichle, R. H.: Data assimilation methods in the Earth sciences, Advances in Water Resources, 31(11), 1411–1418, doi:10.1016/j.advwatres.2008.01.001, 2008.

Salciarini, D., Fanelli, G. and Tamagnini, C.: A probabilistic model for rainfall–induced shallow landslide prediction at the regional scale, Landslides, doi:10.1007/s10346-017-0812-0, 2017.

Schaake, J. C., Hamill, T. M., Buizza, R. and Clark, M.: HEPEX: The Hydrological Ensemble Prediction Experiment, Bulletin of the American Meteorological Society, 88(10), 1541–1547, doi:10.1175/BAMS-88-10-1541, 2007.

Schmidt, J., Turek, G., Clark, M. P., Uddstrom, M. and Dymond, J. R.: Probabilistic forecasting of shallow, rainfall-triggered landslides using real-time numerical weather predictions, Natural Hazards and Earth System Science, 8(2), 349–357, 2008.

Schweigl, J. and Hervás, J.: Landslide mapping in Austria, European Commission Joint Research Centre, Institute for Environment and Sustainability, Italy. Available at: https://esdac.jrc.ec.europa.eu/ESDB_Archive/eusoils_docs/Images/EUR23785EN.pdf, last access: 30 November 2017, 2009.

Schwenk, H.: Massenbewegungen in Niederösterreich 1953-1990, Jahrbuch der Geologischen Bundesanstalt, 135(2), 597–660, 1992.

Segoni, S., Rosi, A., Lagomarsino, D., Fanti, R. and Casagli, N.: Brief communication: Using punctual soil moisture estimates to improve the performances of a regional scale landslide early warning system, Natural Hazards and Earth System Sciences Discussions, 1–9, doi:10.5194/nhess-2017-361, 2017.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C. and Masson, V.: The AROME-France Convective-Scale Operational Model, Monthly Weather Review, 139(3), 976–991, doi:10.1175/2010MWR3425.1, 2011.

Seyfried, M. S. and Wilcox, B. P.: Scale and the Nature of Spatial Variability: Field Examples Having Implications for Hydrologic Modeling, Water Resources Research, 31(1), 173–184, doi:10.1029/94WR02025, 1995.

Shi, H., Li, T., Liu, R., Chen, J., Li, J., Zhang, A. and Wang, G.: A service-oriented architecture for ensemble flood forecast from numerical weather prediction, Journal of Hydrology, 527, 933–942, doi:10.1016/j.jhydrol.2015.05.056, 2015.

Sing, T., Sander, O., Beerenwinkel, N. and Lengauer, T.: ROCR: visualizing classifier performance in R, Bioinformatics, 21(20), 7881, doi:10.1093/bioinformatics/bti623, 2005.

Stensrud, D. J., Wicker, L. J., Kelleher, K. E., Xue, M., Foster, M. P., Schaefer, J. T., Schneider, R. S., Benjamin, S. G., Weygandt, S. S., Ferree, J. T. and Tuell, J. P.: Convective-Scale Warn-on-Forecast System: A Vision for 2020, Bulletin of the American Meteorological Society, 90(10), 1487–1499, doi:10.1175/2009BAMS2795.1, 2009.

Thiebes, B.: Landslide Analysis and Early Warning Systems. Local and Regional Case Study in the Swabian Alb, Germany, Springer Berlin Heidelberg, Germany, 2012.

Thiebes, B., Bell, R., Glade, T., Jäger, S., Mayer, J., Anderson, M. and Holcombe, L.: Integration of a limit-equilibrium model into a landslide early warning system, Landslides, 11(5), 859–875, doi:10.1007/s10346-013-0416-2, 2014.

Thiebes, B. and Glade, T.: Landslide Early Warning Systems–fundamental concepts and innovative applications, in Landslides and Engineered Slopes. Experience, Theory and Practice - Proceedings of the 12th International Symposium on Landslides, June 12-19, 2016, 3, edited by S. Aversa, L. Cascini, L. Picarelli, and C. Scavia, CRC Press, Napoli, 1903–1911, 2016.

Thielen, J., Bartholmes, J., Ramos, M.-H. and Roo, A. de: The European flood alert system–Part 1: concept and development, Hydrology and Earth System Sciences, 13(2), 125–140, doi:10.5194/hess-13-125-2009, 2009.

Tiranti, D., Devoli, G., Cremonini, R., Sund, M. and Boje, S.: Regional landslide forecasting in Piemonte (Italy) and in Norway: experiences from 2013 late spring, Natural Hazards and Earth System Sciences Discussions, 1–30, doi:10.5194/nhess-2017-411, 2017.

Tofani, V., Bicocchi, G., Rossi, G., Segoni, S., D'Ambrosio, M., Casagli, N. and Catani, F.: Soil characterization for shallow landslides modeling: a case study in the Northern Apennines (Central Italy), Landslides, doi:10.1007/s10346-017-0809-8, 2017.

UNEP: Early Warning Systems: A State of the Art Analysis and Future Directions, Division of Early Warning and Assessment (DEWA), United Nations Environment Programme (UNEP), Nairobi, Kenia. Available at: http://na.unep.net/siouxfalls/publications/early_warning.pdf, last access: 30 November 2017, 63 pp., 2012.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Vincendon, B., Ducrocq, V., Nuissier, O. and Vié, B.: Perturbation of convection-permitting NWP forecasts for flash-flood ensemble forecasting, Natural Hazards and Earth System Science, 11(5), 1529–1544, doi:10.5194/nhess-11-1529-2011, 2011.

van Asch, T. W., Buma, J. and van Beek, L. P. H.: A view on some hydrological triggering systems in landslides, Geomorphology, 30(1), 25–32, doi:10.1016/S0169-555X(99)00042-2, 1999.

van der Walt, S., Colbert, S. C. and Varoquaux, G.: The NumPy array: a structure for efficient numerical computation, Computing in Science & Engineering, 13(2), 22–30, doi:10.1109/MCSE.2011.37, 2011.

van Westen, C. J., van Asch, T. and Soeters, R.: 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, 2006.

van Westen, C. J., Castellanos, E. and Kuriakose, S. L.: Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview, Engineering Geology, 102(3–4), 112–131, doi:10.1016/j.enggeo.2008.03.010, 2008.

Wang, Y., Zhao, T. and Cao, Z.: Site-specific probability distribution of geotechnical properties, Computers and Geotechnics, 70, 159–168, doi:10.1016/j.compgeo.2015.08.002, 2015.

Wieczorek, G. F. and Glade, T.: Climatic factors influencing occurrence of debris flows, in Debris-flow Hazards and Related Phenomena, edited by M. Jakob and O. Hungr, Springer Berlin Heidelberg, 325–362, 2005.

Wirtz, A., Kron, W., Löw, P. and Steuer, M.: The need for data: natural disasters and the challenges of database management, Natural Hazards, 70(1), 135–157, doi:10.1007/s11069-012-0312-4, 2014.

WMO: Guidelines on Ensemble Prediction Systems and Forecasting, World Meteorological Organization, WMO-No. 1091, Geneva. Available at: http://www.wmo.int/pages/prog/www/Documents/1091_en.pdf, last access: 30 November 2017, 23 pp., 2012.

Wood, J. L., Harrison, S. and Reinhardt, L.: Landslide inventories for climate impacts research in the European Alps, Geomorphology, 228, 398–408, doi:10.1016/j.geomorph.2014.09.005, 2015.

Yu, W., Nakakita, E., Kim, S. and Yamaguchi, K.: Improvement of rainfall and flood forecasts by blending ensemble NWP rainfall with radar prediction considering orographic rainfall, Journal of Hydrology, 531, 494–507, doi:10.1016/j.jhydrol.2015.04.055, 2015.

Zhang, S., Zhao, L., Delgado-Tellez, R. and Bao, H.: A physics-based probabilistic forecasting model for rainfall-induced shallow landslides at regional scale, Natural Hazards and Earth System Sciences Discussions, 1–17, doi:10.5194/nhess-2016-348, 2016.

Zieher, T., Rutzinger, M., Schneider-Muntau, B., Perzl, F., Leidinger, D., Formayer, H. and Geitner, C.: Sensitivity analysis and calibration of a dynamic physically based slope stability model, Natural Hazards and Earth System Sciences, 17(6), 971–992, doi:10.5194/nhess-17-971-2017, 2017.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences
Discussions

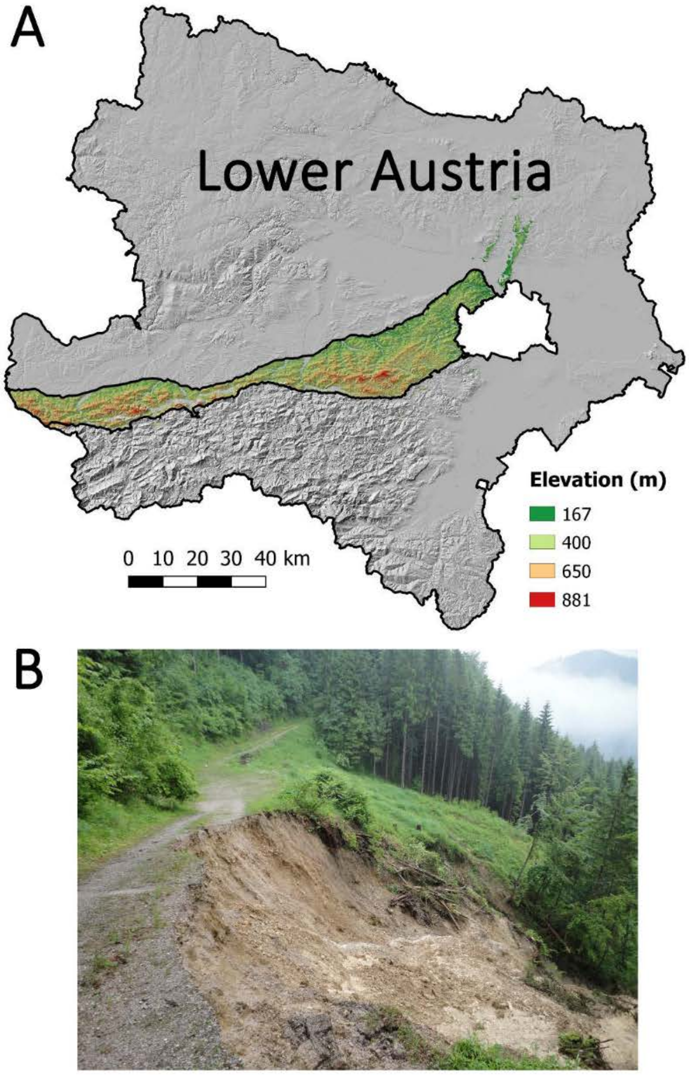


Figure 1: (A) Location of the Rhenodanubian Flyschzone in Lower Austria (DEM: CC BY 3.0 AT–Federal state of Lower Austria); (B) Typical earth slide in Lower Austria after a heavy rainfall event in May 2014 (Picture: K. Gokesch).

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

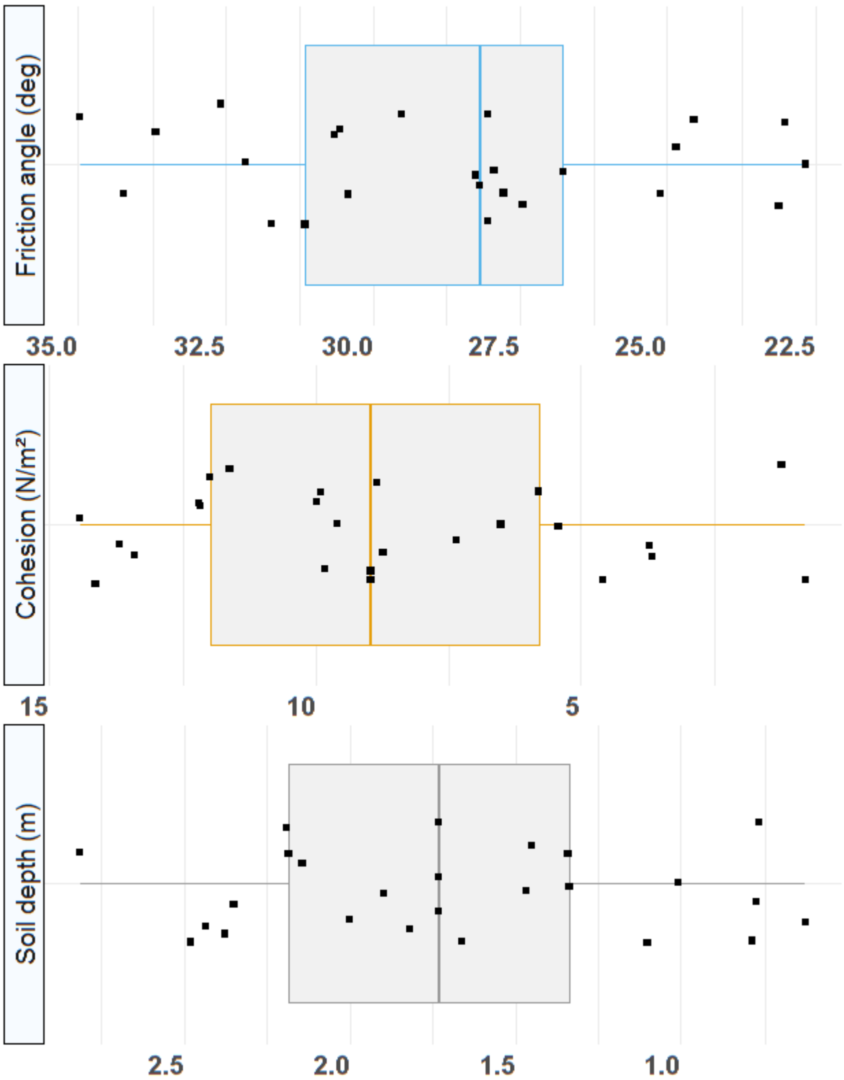


Figure 2: Probabilistically derived modeling parameters based on random sampling from a normally distributed state function. Jittering dots (to prevent overplotting) indicate individual samples within a plausible parameter range.

27

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences Discussions

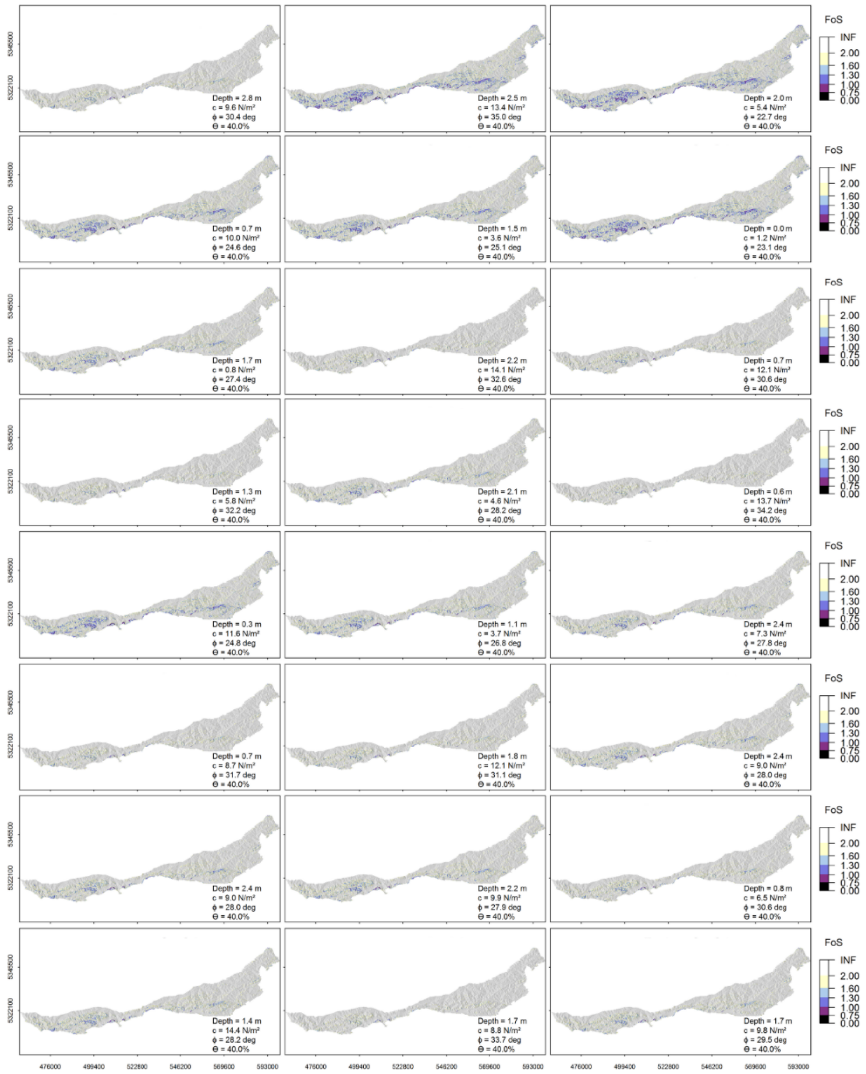


Figure 3: Postage stamp map for 24 model iterations for the same time. Each ensemble member was initialized with altered parameters within a plausible range to account for variability and spatial uncertainty. Factor of Safety (FoS) values < 1 indicate slope instability.

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences
Open Access
EGU
Discussions

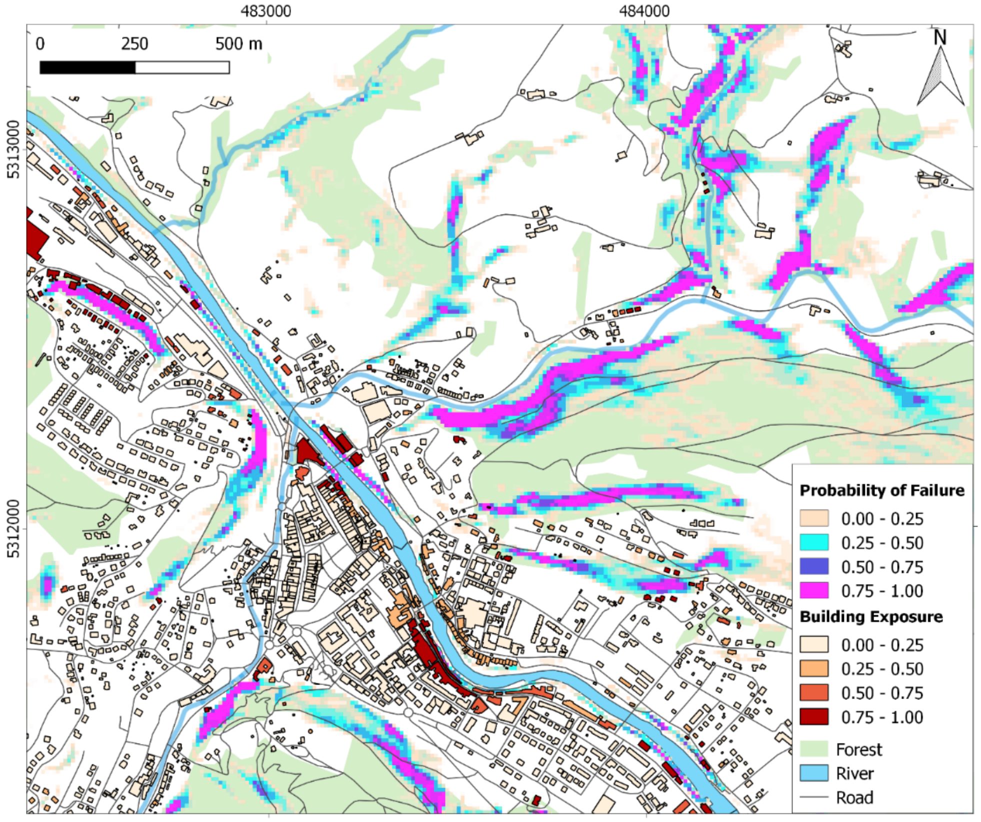


Figure 4: Probability of Failure depicted as a proportion of the ensemble members that predict an event to occur (FoS < 1.0). Building exposure to current slope failure predictions adds an additional information layer for decision makers. Buildings and roads are imported from the freely accessible OpenStreetMap (OSM) database (© OpenStreetMap contributors).

Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-427
Manuscript under review for journal Nat. Hazards Earth Syst. Sci.
Discussion started: 12 December 2017

Natural Hazards and Earth System Sciences
Discussions
Open Access

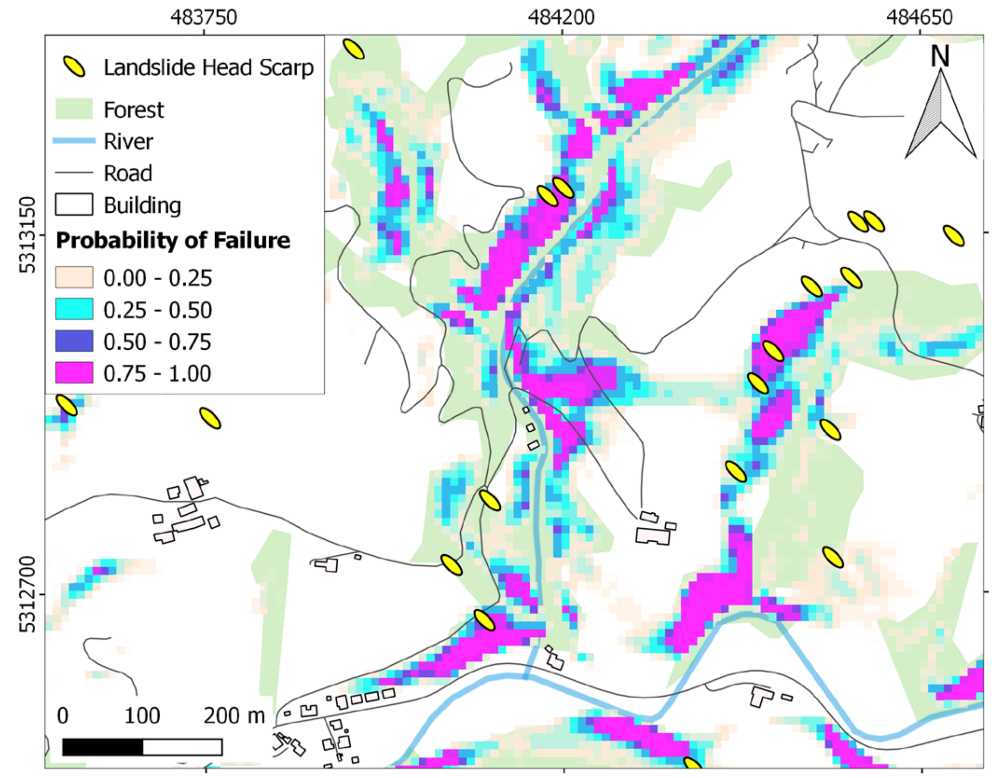


Figure 5: Probability of failure map detail for a specific time under prevailing rainfall conditions. Known historic landslide initiation points (ellipses) partly overlap with current slope stability conditions. However, high spatial resolution, and therefore a high degree of spatial discontinuity, poses a risk for missing many real landslide events in an early warning situation.

B. Eidesstattliche Erklärung

Hiermit erkläre ich, Ekrem Canli, geboren am 07.05.1985 in Horn, die vorliegende Dissertation selbständig angefertigt zu haben. Aus fremden Quellen direkt oder indirekt übernommene Informationen oder Gedanken sind als solche kenntlich gemacht.

Die Arbeit wurde bisher weder in gleicher noch in ähnlicher Form einer anderen Prüfungsbehörde vorgelegt oder veröffentlicht.

Wien, Jänner 2018