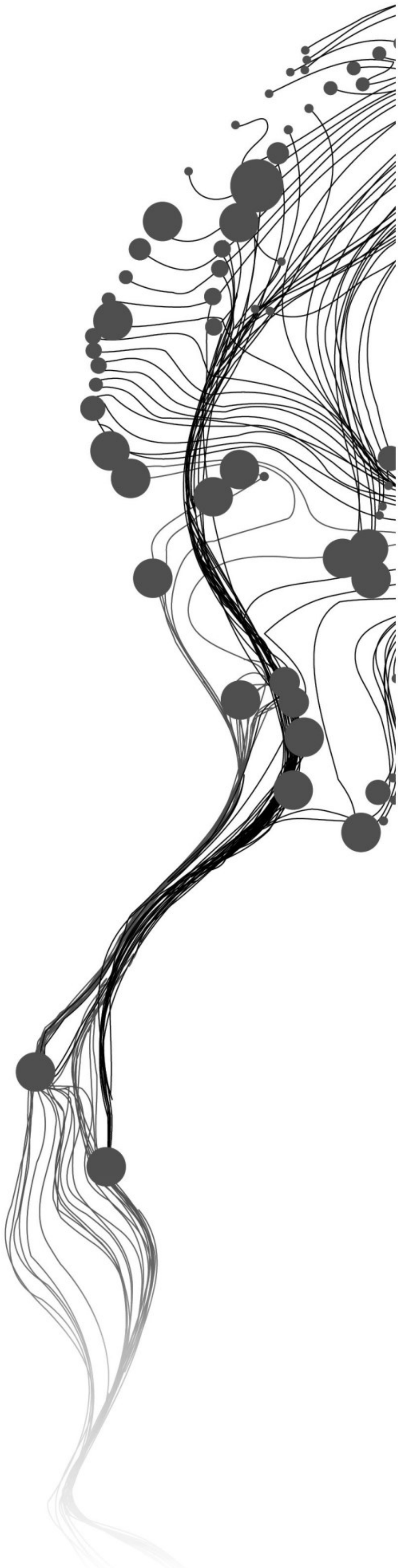


Indigenous Toponyms in Landslide Hazard Mapping for Land Use and Infrastructure Planning

RACHEL GUIMBATAN-FADGYAS

Enschede, The Netherlands, May 2021

Thesis submitted to the Faculty of Geo-information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Spatial Engineering

SUPERVISORS:

First supervisor: dr. Frank O. Ostermann (ITC-GIP)

Second supervisor: dr. Frank B. Osei (ITC-EOS)

ABSTRACT

Reducing disaster risks is a wicked problem that requires integrated knowledge and coordinated action among decision-makers. It acknowledges the combination of indigenous knowledge and scientific knowledge to develop methodologies that improve hazard assessment and reduce community vulnerability. However, few examples exist to operationalize this. Hurdles include knowledge infrastructures that are unable to accommodate different worldviews and knowledge domains. Methods are needed to produce results that are meaningful to the target communities.
The proposed model introduces indigenous toponyms as an interface of co-production between indigenous knowledge and scientific knowledge. It describes how indigenous toponyms can contribute to disaster risk reduction and how the community that provides this type of information can benefit. In a Bayesian approach, indigenous toponyms are used both as data input and explanatory variables for landslide hazard modelling. Translating toponyms into variables for statistical modelling combined qualitative and quantitative methods in the data collection, data processing, and analysis. The workflow was refined as it is evaluated within the context of the study area, situated in the Philippine Cordilleras. First, toponym data obtained in-situ before the research was enriched by structured and unstructured online discussions facilitated by the researcher. Consultations with experts combined with desktop research added details. After which, toponyms were characterized according to their relation with landslide causal factors then regionalized into slope units used to construct models. Using the Deviance Information Criterion, three constructed models were compared for their goodness-of-fit. The selected model was then rendered as a static and dynamic map. The dynamic map version underwent limited testing among actual users as a decision-making tool for land use and infrastructure planning. This mapping output presents a basic tool that the co-producers can improve with updated information and as they prefer. Similar situations may adopt and improve this model.

This research also contributes to indigenous knowledge valorization. As demonstrated, the potential of toponyms as a medium of multidisciplinary collaboration in hazards modelling needs more attention. It opens directions in toponymic research that need further investigation.

Key words
wicked problem, indigenous toponyms, Bayesian, landslide hazards, co-production, disaster risk reduction

ACKNOWLEDGEMENTS

This study topic had been brewing in my mind for a long time and had in part inspired me to return to school where its simplicity can find validation. Little did I know how convoluted it is, covering a range of disciplines and topics which took me back and forth in time between the archives and the present, to groups of people who speak different languages, academic-wise, belief-wise, community-wise. I mulled on various knowledge domains, social and technical, with considerable struggle on the statistical part. Then there's the oscillation between the study area and methods that I could find to fit it. I could not have completed it within the timeframe if not for the consistent supervision of Frank Ostermann, who practically mentored me throughout the research period. I could not have organized my thoughts better. That nudge here and there kept me in the right direction, within boundaries, not losing sight of the main objective to complete a thesis. Thank you, Frank. I also thank Frank Osei for the hints, which made me more confident to change my initial assumptions and to confront data with statistical concepts and their applications.

When I pondered what kind of species this subject matter belongs to and why I had to include extra steps or why I made my study so complicated, the novelty was enough motivation that kept me working. For that, I thank my indigenous heritage and its people who I was only able to reach on social media, but willingly with complete trust, shared the best of their knowledge to me, confident that I would see this to its conclusion. That expectation also gave me some pressure to produce a worthwhile output that will not just gather dust on the shelves.

Thanks to R-INLA, I am less afraid of derivatives, allowing me to see the wonderful world beyond $y=mx+b$ on my regular computer and focus more on the things that count. Throughout this study period, I delved into seemingly unrelated concepts. Thanks to the M-SE programme, the possibilities made me understand what spatial engineering is all about and what a wicked problem demands. Thanks to the WhatsApp SE families for their sane moments. :-)

I thank my husband Gabor for revolving around me, for the countless execuglides to the coffeepot for my caffeine infusement and the spreadsheet and word processing techniques, and many more.

Thank you, God, for creating possibilities.

Table of Contents

List of Figures

List of Tables

Appendices

1. Introduction

This MSc thesis aims to describe how indigenous toponyms can contribute to reduce disaster risks and how the community that provides this type of information can benefit. Through combined production between scientific and indigenous knowledge it covers the entire process of landslide hazard modelling from landslide inventory to map rendering. It is an informed decision-making process where production and its product enable the affected community to make their own risk assessment.

1.1 MOTIVATION AND BACKGROUND

1.1.1 Wicked problem and Indigenous Knowledge

Natural hazards caused over 1.2 million deaths in the last 20 years globally. The economic toll amounts to an annual average of 243 billion dollars (Aon 2020). These defeat sustainable development efforts in less-developed countries, crippling social support systems, making post-disaster recovery an endless struggle, thus exhausting capacities needed to transform economies for the better. It is a wicked problem, compounded by hazard assessment lacking integrated knowledge that is utilized in the best possible way(Weichselgartner and Pigeon 2015). Breaking the vicious cycle requires approaches that understand the root causes of vulnerability and people's abilities to cope and recover from disasters (UN Secretary-General 2016). One approach is to situate this wickedness (Noordegraaf et al. 2019) by using local information to facilitate hazard assessment, in which the necessary experts are identified to collaborate. The strategy is to co-produce hazard assessment models that acknowledges the experiences and knowledge of those with a long history of exposure to natural hazards.

Among these exposed populations are indigenous peoples who possess indigenous knowledge and practices that understand natural hazards (Lambert and Scott 2019). This knowledge resulted from generations of interactions with the environment expressed in cultural meanings and place names (Mark and Turk 2011). Indigenous communities, primarily oral societies, use place names as memory devices where descriptions of the past are stored. These offer geographic information properties and relations that can better assess hazards when combined with scientific knowledge.

1.1.2 Co-knowledge production

The same is also recognized by the Sendai Framework for Disaster Risk Reduction 2015–2030 (UNISDR 2015) which made explicit the use of local and indigenous knowledge in disaster risk reduction . It mentions the use of local and indigenous knowledge (IK) to complement scientific knowledge (SK) in disaster risk assessment that is tailored to localities and context. Before and after the Sendai Framework, participatory methods have been applied to include indigenous communities into scientific studies, yet the transformation of indigenous knowledge into ways that increase community resilience remains underexplored. Despite participatory method awareness, the vast majority of studies on climate research using indigenous knowledge are found to use an extractive model (David-Chavez and Gavin 2018), which limits possibilities of building resilience.

In the integration process, the domination of SK is an issue raised (Battiste 2014; Gasparotto 2016; Mazzocchi 2018; Nakashima, Rou, and Munn 2002), requiring balanced approaches that permit drawing on the best wisdom that these two types of knowledge provide (Kelman, Mercer, and Gaillard 2012). It means co-production of knowledge where IK does not only play a confirmatory role (Alexander et al. 2019; Latulippe and Klenk 2020) which is realized through high community engagement. Indicators of this engagement include transparency of the integration procedure, community authority on analysis, access to findings, and reported outputs (David-Chavez and Gavin 2018; Wheeler and Root-Bernstein 2020). The result is informed decision-making, where studies extend to better outcomes in the community

(Raymond et al. 2010). In informed-decisions, the right to decide on how to save and use IK emphasizes on community engagement not limited only to supplementing science data but to supporting "governance-value" in these systems (Whyte 2018). This further goes to supporting data sovereignty and data governance where the collection, stewardship and dissemination of data is centred on indigenous peoples' rights (David-Chavez and Gavin 2018). Currently, there are limited examples of co-production that adhere to indicators and principles of high community engagement.

Indigenous toponyms are articulations of IK (Heikkilä and Fondahl 2010) that serve as a "medium to reflect on the indigenous ancestral past for guidance on living "right" in the present" (Johnson and Basso 1998). As historical markers of natural hazards that occurred in the distant past, toponyms can provide leads to predictive modelling. Specifically, they can inform on "intensive risks" which is defined by UNISDR (2017; 2015a) as risks with "high-severity, mid to low-frequency disasters," where large concentrations of people and economic activities are exposed to intense hazard events, which are also characterized by underlying risk drivers such as poverty and inequality (UNISDR 2009). The correlation between toponyms and avalanche, landslides, floods, and tsunami was recently explored by the academic community (Dall'Ò 2019; Faccini et al. 2017; Isoda et al. 2019)(Dall'Ò 2019; Faccini et al. 2017; Isoda et al. 2019). While these studies agree that toponyms have informative value, further exploration of their usefulness in disaster risk reduction was not pursued.

1.1.3 Indigenous toponyms and landslides

Indigenous toponyms related to natural hazards abound where these events occur, storing prodigious experiences that serve as references in past spatial decisions to avoid disaster risk. This is evident in old settlement patterns to avoid tsunamis in the coastal regions (Isoda et al. 2019), landslides in the mountainous regions (Dall'Ò 2019), and floods in the riverine and coastal areas (Jones 2016). Land developments may have obscured their importance, but they remain as fixed markers, offering information to build hypotheses for scientific inquiry. However, toponyms that describe ephemeral evidence like those from tsunamis and storm surges approximate the reach of the hazard, and information can be deduced from old settlement patterns or oral literature that may have been embellished through time (Isoda et al. 2019; King, Goff, and Skipper 2007). Their assessment requires additional physical evidence. On the other hand, landslides are downward movements of rocks and soil resulting from natural or human-made actions that leave tangible evidence in the physical landscape. Toponyms related to landslides and associated events describe geomorphology, geology, and deposits, combined with existing oral literature. Regions with a long history of human interaction with landslides offer rich, verifiable, but untapped toponymic information that can help explain landslide occurrence. In societies with existing oral traditions, the narratives are substantiated by what is still observable and vice versa. Scientific investigation validates their importance, which promotes official adoption of their use. For indigenous communities in mountain regions where terrain ruggedness inhibits response during disasters, this is important because official recognition restores confidence in IK. The usage of this information creates an enabling culture where spatial decisions acknowledge the significance of shared local historical observations, thereby promoting awareness of reducing disaster risks.

1.1.4 Usefulness of Landslide Probability Models

Disasters are difficult to resolve due to the lack of understanding of the root causes and risk drivers (Alcántara-Ayala and Oliver-Smith 2019). In areas at risk, stakeholders have to be persuaded first by hazard forecasting from authorities before participating in mitigation measures. Yet, current research on landslide hazard assessment lacks approaches that involve local stakeholders. Although most of these studies refer to their usefulness in land planning and risk reduction, the results do not match their expressed intended use (Hearn and Hart 2019). Prevailing research practice focuses on sophisticated statistical analysis rather than on the relevance of the outputs (Reichenbach et al. 2018). The explanatory variables for these models are often assumed from what is recommended in general literature, derived from geologic surveys and processed from remote sensing. Local knowledge that contextualizes the modelling product is often absent. It may also account for the missing link to land use planning, infrastructure intervention and community awareness, underscoring the need for local input, design and validation of the model. To build this link, toponyms can be utilized as the springboard where the act of eliciting the local peoples' knowledge of these places becomes a process of learning why a name was given to a place and the conditions surrounding it (Perdana and Ostermann 2019). This can provide light on past risk management. For indigenous toponyms, tacit knowledge that the local people have on places linked to landslide hazards can surface. In this process, a collective social exercise memorialized in places (Jenjekwa Vincent 2018) becomes an exercise of awareness that helps shape the usability of resultant mapping. No known methodology exists yet where indigenous toponyms are explored for their usefulness in landslide hazard modelling.

1.1.5 Prior Knowledge Translation

Making room for IK production in academic research that is dominated by science needs a new way of doing things (Latulippe and Klenk 2020). Methods must be based on ethical frameworks where opportunities are provided in which indigenous communities can represent their knowledge and values on their own terms (Hill et al. 2020; Parsons, Fisher, and Nalau 2016). In addition, methods adopted must not undermine IK ontologies where inputs are re-articulated and where outputs lose relevance to the indigenous community. These are concepts of equity and transparency that are consistent with the indigenous rights to self-determination (UNDRIP 2017). From the point of view of science, indigenous knowledge reaches its potential when involved in all steps of the research process and enhanced by an interdisciplinary approach (Bélisle et al. 2018). Therefore the choice of modelling framework must provide this space.

Bayesian models are recognized to be well-suited to meet this methodological challenge. The parallels of the predictive property of IK and Bayesian methods were investigated by Tacher and Golicher (2004) and are now increasingly applied in IK integration with the ecological sciences (Bélisle et al. 2018; Bowles et al. 2020; Gryba 2020; Liedloff et al. 2013; Reid et al. 2021). Fuzzy models are also suggested to capture holism and complex properties of IK in a systematic way (Bélisle et al. 2018; Mackenzie Kierin, Siabato Willington, Reitsma Femke 2017; Sarmiento et al. 2020).

Where indigenous toponyms are assumed to help explain the probability of landslide occurrence over space and time, they are treated as prior knowledge that updates both science and local beliefs. This process of updating is inherently Bayesian, wherein beliefs are also translated into quantifiable form. As a form of indigenous knowledge stored in topographical space, toponyms are used here as local expert knowledge translated into prior distributions for analysis.

1.2 PROBLEM IDENTIFICATION

With the end-view of reducing disaster risks, this study contributes to developing contextualized landslide hazard mapping for land use and infrastructure planning. In particular, indigenous toponyms are used as a co-knowledge production tool and an information source of inputs in a Bayesian model. Thus, the main objective is to devise a co-production process of using indigenous toponym data in landslide hazard modelling.

1.2.1 Research Objectives

The process of using indigenous toponym data into landslide hazard modelling involves four expected outputs in response to the following sub-objectives:

1.To develop a toponymic co-production classification approach that translates toponyms into input variables for landslide hazard modelling

2. To generate, evaluate and select toponymic variables for landslide hazard modelling

3. To implement and select a Bayesian model to use based on their performance

4. To assess the resultant map's usability as a hazard information for land use and infrastructure planning in the study area

1.2.2 Research Questions

The following questions address the respective objective:

Sub-objective 1:

RQ1.1 What are the considerations in representing a toponym as an input variable for landslide hazard modelling?
RQ1.2 How is co-production employed in translating toponyms into model variables?

Sub-objective 2:

RQ2.1 Factoring in the conditions mentioned in RQ1.1, which methods are suited to generate quantitative input variables for modelling from toponyms?
RQ2.2 What probability distribution captures the information provided by the data and toponyms?

Sub-objective 3:
RQ3.1 What are criteria to evaluate and select toponymic variables for landslide hazard modelling?
RQ3.2 Based on which criteria and which process are models selected for their goodness-of-fit?

Sub objective 4:

RQ4.1 Which factors define the usability of the resultant mapping as a piece of base information for land use and infrastructure planning in the study area?
RQ4.2 Based on the factors defined in RQ4.1, what testing method can measure the usability of the resultant map among users in the study area?
RQ4.3 What features in the landslide hazard map needs improvement to make it more usable?

1.2.3 Significance

Limited research has been done where indigenous knowledge and scientific knowledge co-produce to model landslide hazards. There are no known methodologies on how indigenous toponyms translate into prior distributions in Bayesian modelling. The process of integrating indigenous knowledge into state-of-the-art science is an under-researched topic. This research will provide information on how this type of knowledge and its systems integrate with geoscience, which includes associated ethical concerns, that may be helpful for future research in this area. Within the geoscience community, this research can serve as an example of interdisciplinarity and collaboration possibilities that can clarify the tasks of each discipline given a wicked problem. It opens areas for further study. The research also contributes to valorizing indigenous knowledge, which is under-represented in literature and in danger of losing its significance, especially to the communities that possess this.

The research output is a landslide hazard map integrating indigenous toponyms that the concerned local community can use as a reference for their decision-making to reduce disaster risks. The expected result is better risk-governance in the area.

1.3 RESEARCH APPROACH

1.3.1 Overview of analysis workflow

From generating toponymic information to the resultant landslide hazard mapping, co-knowledge production between indigenous knowledge (IK) and scientific knowledge (SK), underpins the general methodology (Figure 1) below.

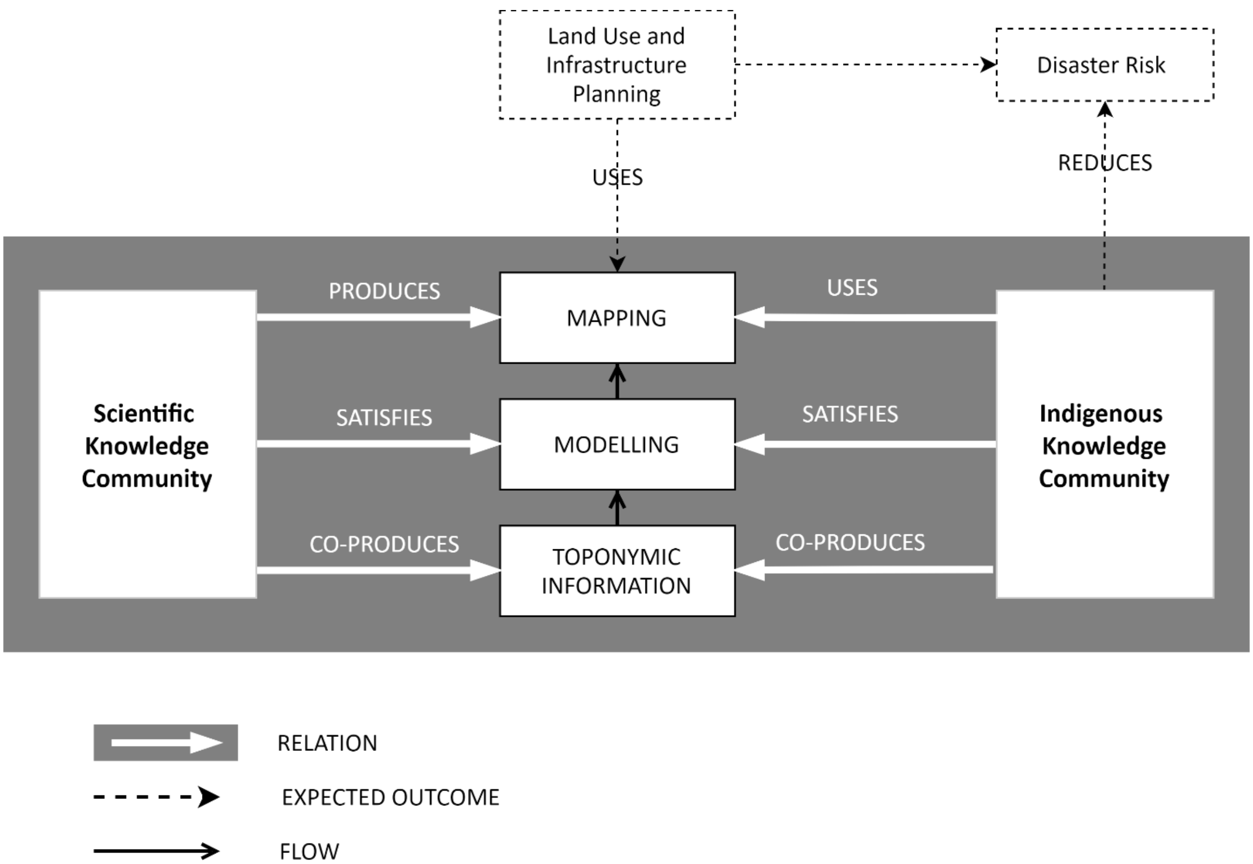


Figure 1.General workflow of co-knowledge production

1.3.2 Research Design

This study combined qualitative and quantitative approaches in data collection, data processing, and analysis following the general sequence of the co-knowledge production workflow shown in Figure 1. This workflow was refined as it is evaluated within the context of a study area. New methods were adopted from knowledge acquired and validated during the research process, which resulted to a proposed model. The study area is a 25-square kilometer composition of contiguous river basins in the Philippines Cordilleras. The region was reported as among the top 5 locations with the largest area at risk of landslides by NASA's LHASA tracking (Parsons and Lister 2018).

Three main characteristics define the study area as a good candidate for the proposed modelling co-production process:

- Richness in indigenous toponyms
- Lived-in by indigenous peoples
- Long history of landslides

The selection of the study area was further influenced by the familiarity of the researcher of the site and local language. Established connections with the target indigenous community made access and verification of data efficient. Another major determinant is the availability of data within the thesis working period and Covid-19 pandemic constraints. The limitations brought about by the pandemic also decided the boundaries of the study area.

1.3.3 Thesis Structure

The thesis document is presented as follows:

Chapter 1: Introduction and problem identification
Chapter 2: Concept Integration
Chapter 3: Design and Implementation
Chapter 4: Modelling Results
Chapter 5: Discussion
Chapter 6: Conclusion discussing insights and possible developments of study

Chapter 2 reviews methods and presents knowledge gaps to explain approaches that are adopted for the design and implementation methods in Chapter 3.

2. Relevant Concepts

This chapter aims to show how different concepts, theories, and models found in literature are integrated relative to the objectives of this study. Five main themes define the concepts reviewed: Co-knowledge production; Landslide hazard modelling; Toponym research approaches; Statistical approach, and; Usefulness. First, it synthesizes the requirements of co-knowledge production, then defines the gaps in conventional landslide modelling. Based on identified gaps and requirements, a discussion on toponym handling and the Bayesian approach follows. These discussions mention state-of-the-art approaches. The last discusses Usefulness measures

2.1 CO-KNOWLEDGE PRODUCTION CONSIDERATIONS

There are few examples of IK-SK co-production available in literature in the past five years that aim at producing concrete results in the form of information or policy action. The list in Table 1 shows co-production where the objective is to generate information for various decision-making processes. The list in Table 2 shows various applications of co-produced knowledge that are directly related to policy decisions.

In the five examples in Table 1, the results provide a better understanding of nature, in the form of enriched data, information equipment, and decision support tools. Co-production methods employed state-of-the-art tools, and joint analyses, involving various skills and disciplines, facilitated by the researchers themselves. In all cases, mixed methods of analysis were employed with specific focus on results that matter to the local community. Some characteristics of the co-production process are outlined as follows:

Early engagement. The concerned indigenous community participate in the study at the earliest possible time wherein protocols specific to the community are defined. It could begin right before the start of the project implementation or as early as the conceptualization stage of the project**.**

Combined data and observations. Different available tools of data acquisition, organization and processing were employed as demonstrated by the combined weights method in South Africa (Nyetanyane and Masinde 2020) and heterogeneous sensing in Swedish Lapland (Roué et al. 2016). Disparate data that are either remote-sensed or "ground-sensed" through mobile applications or sensors are combined in the analysis.

Recognition of complex dimensions. IK has complex dimensions that does not have a one to one correspondence with SK. IK combines different concepts into one to fewer representations, that require appropriate organization, classification or statistical methods that respond to this fuzziness. This is indicated in the use of HREV (Fox et al. 2020) and Shannon's diversity index (Capra et al. 2016) to capture that complexity. Knowledge of a phenomena is unique to the community setting which means that the variables and the methodology adopted derive from the properties of that knowledge.

Table 2 shows that cases aimed at policy decision-making employed co-production methods that facilitate dialogues. IK is visualized with the help of computer-based mapping and GIS tools (Diver 2017). The spatial visualization of IK goals facilitated the formal recognition of this knowledge by the government policy-maker. Continuous reflection through ethically sensitive dialogues ensured that IK diversity is preserved and not subsumed in the dominant knowledge infrastructure (Matuk et al. 2017). In the Brazil case, Matuk et al. (2017) mentioned that an indicator of legitimacy of the resultant knowledge from co-production is its usability among the indigenous community. Another indicator of legitimacy is the community recognition of selected IK resource persons in structured surveys. In this respect, the Delphi technique employed by Wheeler et al (2020) may have to be combined with protocols on selection of representative knowledge resources as well as unstructured elicitation of information.

Table 1. Co-production to generate information

Method	Results	Setting	Author
1. Structured elicitation process and statistical framework to combine indigenous knowledge with survey data 2. Digitizing hand-drawn maps to create spatial polygons 3. Maxent Machine Learning, constraining model parameterization to the IK boundary	Species distribution modelling (SDM) for increased ecological insights in the Martu determination area Ecological insights that improves the impact of research	Australia	(Skroblin et al. 2020)
1. Combined weights of Climate Data, Indigenous Knowledge and Satellite Imagery to determine season onset where IK indicators are elicited through farmer interviews and surveys. 2. Crop monitoring by IK expert using app 3. Crop health prediction through Machine Learning using timeseries algorithms	Optimizing Cropping Decisions by Small-Scale Farmers, uMgungundlovu District Municipality	South Africa	(Nyetanyane and Masinde 2020)
1. Human-relevant environmental variables (HREV) or complex, synthesis variables that when used in conjunction with a host of social variables, assist in informing safe land travel and activities 2. Locating best sites for weather stations	Community-based weather station network in Inuit Nunangat Website where residents can access real-time weather information	Canada	(Fox et al. 2020)
1. Co-design of protocols and joint analyses by herders, climatologists, anthropologists and ethno-biologists 2. Combined observations: Temperature probes Local snow and herding via a smartphone app by Sami herders Cartographic data from Sami on land use patterns Quantitative data about snow metamorphism at the regional scale using remote-sensing and satellite imagery	1. Establishment of Evenk Observatory to monitor climate and environmental change Eurasian reindeer herding peoples 2. Establishment of community-based observatory	Siberia Swedish Lapland	(Roué et al. 2016)
1. Toponym collection from Geoportal, Municipal Library and State Archives, and interviews with local people 2. Database creation and toponym translation 3. Categorization of toponyms 4. Soil field investigation and analysis of several physical–chemical parameters where pedonyms (indigenous knowledge) are matched with prevalent soils (scientific knowledge) 5. Shannon's diversity index to quantify complex variables 6. Statistical comparison between local and scientific knowledge through CCA	Soil resource study of the Sardinian landscape	Italy	(Capra et al. 2015, 2016)

Table 2. Co-production for policy decisions

Method	Results	Location	Author
Delphi technique or structured expert elicitation process	1. Environmental decision-making 2. Identification of participants' experiences of scientists' misconceptions on IK	Arctic	(Wheeler et al. 2020)
1. Workshops through community meetings and fieldtrips 2. “Translation Convergences” through knowledge linkages of land use values and goals visualized through computer-based mapping technology and GIS	Science-policy negotiations 1. Ecosystem-based planning for Xáxli’p Community Forest, 2. Formal recognition of indigenous values by the Ministry of Forests	British Columbia, Canada	(Diver 2017)
Multiple workshops with culturally and ethically sensitive dialogues that include continuous reflection, enabling adaptation and improvisation Method aims at knowledge legitimacy and usability	1. Ethnoecological approach on Amazonian SISA policy (REDD+ program)Integrated Kaxinawá and scientific classifications of the soils on the map	Brazil	(Matuk et al. 2017)

2.3 TOPONYM RESEARCH APPROACHES

This section reviews mapping approaches and the most relevant classification approach that satisfies co-knowledge production.

2.3.1 Toponym collection

Approaches to mapping indigenous toponyms are either related to its protection or usefulness. These activities are expressed in research and rarely for the purpose of placing these indigenous names on the map. The research objectives vary from cultural to historical referencing for various purposes (Cogos, Roué, and Roturier 2017; von Mentz 2017; Morphy et al. 2020).

Toponym collection is also recognized in new disciplines such as ethnopedology that seeks to understand indigenous approaches in soil classification and management (Capra et al. 2015). A new field of inquiry that is starting to employ toponyms is ethnophysiography (Feng and Mark 2017), which studied local conceptualizations of the landscape, and eventually enabling culturally specific geographic information systems (Mark and Turk 2011). Ethnogeology, which was a field identified much earlier (Kamen-Kaye 1975) still has to explore links to toponyms. So far there is no related literature on indigenous toponyms that explain physical mechanisms responsible for landslides.

2.3.2 Participatory approaches

In the mapping of placenames for landslide hazard modelling, the desired information is the point location of the toponym, its etymology, and associations. Since not all indigenous toponyms are recorded in ethnographic atlases, participatory approaches have to be adopted to collect information. Community members are engaged in locating place names and discuss their meanings. Although the concept of participatory mapping is not new in indigenous communities, the mapping of placenames itself is new. Field work is a tried and tested approach of collecting informtion that elicits indepth description. A participatory method through field survey was applied in Indonesia using a collection-verification-publication process (Perdana and Ostermann 2018). In this example, there is emphasis in engaging the local government and the community to improve data collection. The study pointed out that field surveys have limiting factors such as accessibility to location, weather, and density of geographic features in fieldwork areas. The surveyors' familiarity with the place also matters, as it defines navigational tools to use that would enhance their performance. The availability of digital platforms like Open Street Map and Google Earth also offer opportunities for toponym collection.

2.3.3 Toponym classification

For toponym classification in relation to toponym research, Tent (2015) synthesized different studies and proposed two approaches which he calls "micro" and "macro" levels (Table 3). At the micro level, research is described as qualitative which becomes quantitative at the macro level, where the pattern analysis expresses values in numerical form. Often the qualitative research is followed by a quantitative study in order to find empirical support for hypotheses. Such analyses can reveal much about the following elements: place naming practices and patterns (both temporally and spatially); regional distributions of certain types of toponym, or geographic feature; settlement patterns (both temporally and spatially); the geomorphology of a region (by concentrating on feature types); grammar/syntax of toponyms; linguistic geography such as regional distribution of name types; and, the influence of names on property values.

Table 3. Tent's classification in toponym research

Micro/Qualitative/Intensive	Macro/Quantitative/Extensive
Etymology, meaning, and origin of toponyms	Toponyms of a region and examining patterns of these names
Grassroots-based with basic fields: toponym identification, where sound linguistic knowledge is required; toponym documentation; and, toponym interpretation.	Broader research based on datasets or corpora of toponyms, gazetteers, maps, and atlases. At this stage, placenames function as independent variables which can be tested against dependent variables such as region, toponym type, or feature type.

As shown by Tent, qualitative handling of toponyms requires linguistic competency. Co-production and toponymic organization exercises require ability to communicate in the local language. Local language competency enables recognition of nuances of vocabularies used among different participating communities in the locality.

Another way to classify toponyms is by date of their appearance (Vannieuwenhuyze 2007). The oldest mention of toponyms confirms the name's existence, offering a *terminus ante quem* for the reference to that spatial reality. Therefore, written and oral records that mention the name provide a basis for analyzing linkage to historical events. However, as Vannieuwenhuyze (2007) noted, the date of appearance cannot provide a *terminus post quem* for their existence.

The translation of toponyms into explanatory variables for hazard modelling is not available in literature. An approach on toponym translation into numerical values have so far been explored by Capra et al. (2015) in ethnopedology, in which Shannon entropy index and canonical correspondence analyses were employed.

2.2 LANDSLIDE HAZARD MODELLING

2.2.1 Concepts

Landslide is a slope movement that has several conceptions. The most endorsed is presented by Varnes and the International Association for Engineering Geology (1984) who defined it as "almost all varieties of mass movements on slope including some such as rock falls, topples and debris flow that involve little or no true sliding". The slope movement classification shown in Table 4, is also presented by Varnes (1978)which is widely used today to classify landslides.

Table 4. Classification of slope movements(Varnes 1978)

TYPE OF MOVEMENT			TYPE OF MATERIAL		
			BEDROCK	ENGINEERING SOILS	
				Predominantly coarse material	Predominantly fine material
FALLS			Rock fall	Debris fall	Earth fall
TOPPLES			Rock topple	Debris topple	Earth topple
SLIDES	Rotational	Few units	Rock slump	Debris slump	Earth slump
	Translational	Many units	Rock slide	Debris slide	Earth Slide
LATERAL SPREADS			Rock spread	Debris spread	Earth spread
FLOWS			Rock flow (deep creep)	Debris flow (soil creep)	Earth flow (soil creep)

Abbreviated version of Varnes' classification of slope movements

Landslide susceptibility refers to the spatial probability of landslide occurrence. It predicts "where" landslides are likely to occur (Guzzetti et al. 2005a). According to the USGS, the most important factors determining susceptibility are prior failure, rock or soil strength, and steepness of slope. Varnes and the IAEG (1984) suggested that the definition of natural hazards should also apply to mass movements on a slope, such that landslide hazards would then mean "the probability of occurrence within a specified period and a given area of a potentially damaging phenomenon". Guzzetti et al. (2005) added to the definition by including the concepts of magnitude, geographical location and time recurrence. Magnitude refers to intensity that conditions destructive power. Thus, landslide hazard is susceptibility with a temporal component and intensity that conditions destructive power. With time as a component, hazard also considers the chance that a landslide might travel downslope a given distance (Highland 2008).

Landslide susceptibility models therefore are only concerned with the presence-absence of landslides over space. Landslide hazard modelling goes further than landslide susceptibility modelling by adding temporal, spatial and size probability of events (Corominas and Mavroulli O 2011). Due to scarcity of historical information, few models incorporate time and magnitude.

Assessing landslide hazards needs a collection of landslide occurrence information. These landslide inventories can be prepared through various methods (Guzzetti et al. 2000; Wieczorek 1984) but still pose a challenge because aside from requiring expertise and resources, it is tedious (van Westen, van Asch, and Soeters 2006). The collection of information also requires landslide classification that suffers simplifications, geomorphological deduction, and subjectivity(Guzzetti et al. 2012). Guzetti et al. (2005b) suggest checking against external information on landslide types and processes available for the investigated area.

Both landslide susceptibility mapping and landslide hazard mapping result in zonation or the terrain's subdivision into zones that have a different likelihood of landslide occurrence. The different approaches that lead to landslide hazard zonation are classified into four general approaches (Aleotti and Chowdhury 1999; Guzzetti et al. 1999; Soeters and Westen 1996; van Westen et al. 2006). These are:

- landslide inventory-based probabilistic approach, which interprets data either from remote-sensing, field observation, interviews or historical analysis;
- heuristic approach, which is based on the opinion of geomorphological experts and adopted when landslide data is scarce;
- statistical approach, which uses a landslide inventory map where information on past landslide occurrences are needed to forecast future occurrences using a combination of causal factors that are statistically determined; and
- physically-based modelling approach, which calculates the safety factors and quantitatively produces the stability index using a slope stability model.

These are implemented through various techniques (Shano, Raghuvanshi, and Meten 2020). The best method depends on the scale, the available data, and the characteristics of the study area (Abella et al. 2006).

Landslide susceptibility assessment remains to be dominant in gaining insights into probable slope failures. Statistically-based models also continue to flourish due to advances in technology and increasing availability of data. However, there is concern that the growing number of these models has not made significant changes in terms of quality and usefulness, as discussed in the following section.

2.2.2 Statistically-based models and zonations

The gaps in current practice emphasize the lack of multiplicity and representativeness in landslide temporal inventories, mapping techniques, and model types to increase the quality of information and optimize zonations (Reichenbach et al. 2018). This finding was based on an analysis of 565 peer-reviewed articles on statistically-based landslide susceptibility models. There is also a lack of multiple metrics in evaluating the performance of landslide susceptibility models, which had been mentioned by Rossi et al. (2010)and Steger et al (2016). Aside from this, there is an emphasized need to concentrate on designing new and more reliable methods and indices to evaluate of model quality, thus increasing their credibility and usefulness. The aim is to favour their adoption and use by different stakeholders.

Reichenbach et al. (2018) mentioned the lack of statistical, geomorphological or operational justification for biases regarding the use of modelling tools and landslide information. The observation further includes the lack of careful analysis of available geo-environmental information prior to its use for susceptibility modelling, considering the variables' relevance or lack of relevance.

In general, landslide susceptibility modelling steps start with a landslide inventory followed by selecting causal factors, such as those listed in Table 5, used as independent variables in statistical analysis. Modelling assumes that the factors causing slope-failure in a region are the same as those which will generate landslides in the future. Widely used algorithms make use of either statistical approaches or machine learning techniques. Machine learning algorithms are known to be effective in maximizing predictive performances. But handling analytical tasks, like elaborating parameter uncertainties and effect sizes might require less complex and more transparent algorithms, such as logistic regression or generalized additive models(Goetz et al. 2015; Schmaltz, Steger, and Glade 2017; Steger and Kofler 2019).

Generalized additive models result in continuous probability values in the hazard map that are reclassified into susceptibility zones. Current models mention this output as zonation maps to be used supposedly for land use planning. However, these final zonation maps often have reclassified values that do not define how urban planners and decision-makers could use them (Hearn and Hart 2019).

Table 5. General landslide causal factors (classified by *Süzen and Şener Kaya 2012)*

Environmental	Geotechnical	Topographical	Geological
Anthropogenic parameters	Soil texture	Drainage	Strata-slope interaction
Position within catchment	Soil thickness	Surface roughness	Lineaments/faults
Rainfall	Other geotechnical parameters	Topographic indices	Geology/lithology
Land use/land cover		Elevation	
		Slope aspect	
		Slope length	
		Slope angle	
		Slope curvature	

2.2.3 Spatial units

Landslide hazard assessment uses different spatial units: grid cells, slope units (SUs), or administrative units(Van Den Eeckhaut et al. 2009; Erener and Duzgun 2011). In recent years, the choice of mapping unit for landslide susceptibility zonation has been the subject of study for some researchers. A comparison between grid-cells and SUs yielded findings where the SUs performed better (Ba et al. 2018; Martinello et al. 2020). Another comparison concluded that SU sizes play an important role in the final result (Domènech, Alvioli, and Corominas 2020). An SU has a strong relation with the underlying topography, absent in grid cell–based analyses (Guzzetti et al. 2006).

2.2.4 State-of-the-art on landslide hazard modelling

There are attempts to include temporal probability in landslide hazard models. Techniques were applied that explored available historical benchmarks. For instance, in the absence of historical data on landslides, landslide hazard models have used an indirect approach by analyzing the frequency of rainfall occurrence to derive landslide distribution over time. Fan et al. (2020) and Ha et al. (2020) applied this through simulations in small catchments not exceeding 1 square kilometre. Landslide distribution was also obtained from statistical analysis using Gumbel distribution for a one-time extreme rainfall event (Lee et al. 2020) and Poisson distribution for more frequent rainfall events (Dikshit et al. 2020). Pradhan, Lee, and Kim (2019) applied the same approach on a regional scale using Artificial Neural Network (ANN). Uzielli et al. (2018) also used the triggering probability of rainfall using a Bayesian approach to measure the temporal evolution of landslides. In cases where observed historical data exists but not older than 20 years, machine learning was used, such as Deep Belief Network (DBN) to explore patterns of displacement (Li et al. 2020) and Random Forest (RF) to predict near-future events (Lai and Tsai 2019).

In a similar scenario where the available historical data is not older than eight years, Bayesian-generalized additive models constructed from 3 co-seismic event inventories were used. One model generated predictive realizations over two other inventories (Lombardo and Tanyas 2020). On a global scale, the

National Aeronautics and Space Administration (NASA) offers a near-real-time model called Landslide Hazard Assessment for Situational Awareness (LHASA) which maps susceptibility with 1-kilometre resolution(Kirschbaum and Stanley 2018)In cases where historical data that is older than 50 years were obtained, the Poisson distribution was applied. Fu et al. (2020) used this for community-based modelling in a 34-square kilometre in China. On a 79-square kilometre area in Italy, Lombardo et al. (2020) combined a Poisson model with a Gaussian model using a Bayesian framework that modelled "intensity". Both studies used SUs as spatial units. All models exploited the possibilities of obtainable data, and each has its limitations. Most of it is on account of the accuracy of the data. From these models, there is a correspondence between how far into the future one can predict from how far back in time an event was observed. Records of the distant past have not yet been considered.

2. 4 BAYESIAN APPROACH

Bayesian is a statistical perspective that accommodates one's prior belief in a quantitative analysis. It is based on Bayes' theorem wherein one's prior belief is updated after evidence has been taken into account. This updated belief is called the *posterior probability,* and the object of interest, which can be exploited to make inferences and draw conclusions from. Before seeing evidence, the beliefs held by a modeller about the parameters in a statistical model is called a *prior distribution*, expressed as probability distributions. That belief changes when evidence or new data is obtained. A way of quantifying those belief changes is called conditional probability. The conditional probability distribution given parameters of the data, defined up to a constant is known as the *likelihood function*.

If we have a data set y and model parameters θ, Bayes' rule can be written as:

$$\pi\,(\theta \mid y) = \frac{\pi\,(y \mid \theta)\pi\,(\theta)}{\pi(y)} \qquad (1)$$

where:

$\pi(y|\theta)$ is the *likelihood* of the data y given parameters θ
$\pi(\theta)$ is the *prior distribution* of the parameters and,
$\pi(y)$ is the *marginal likelihood*, which acts as a normalizing constant

$\pi(y)$ is difficult to calculate because it could involve sums and integrals that could be time-consuming, that's why in practice, the *posterior* $\pi\,(\theta|y)$ is the estimated product of the *likelihood* and the *prior distribution.* However, it is needed when computing Bayes factors for model comparison and averaging (Link and Barker 2006). Bayes' rule is written as:

$$\pi(\theta \mid y) \propto \pi(y \mid \theta)\pi(\theta) \qquad (2)$$

In this equation, the posterior can be estimated by re-scaling the product of the likelihood and the prior so that it integrates up to one. The *prior distribution* $\pi(\theta)$ is set by the modeller.

There is an increase in interest in Bayesian statistics due to the recent availability of powerful computational tools. Computational techniques like Markov Chain Monte Carlo (MCMC) makes exact

Bayesian inferences possible even in very complex models. However, the exact computation of marginal likelihoods can be very slow. A fast alternative to this is the integrated nested Laplace approximation (INLA), a method for approximate Bayesian inference introduced by Rue, Martino, and Chopin (2009). This method is used via the R-INLA package. Bayesian modelling in R-INLA allows a free combination of different sorts of modelling approaches with prior information. It does not select model combinations or new arrangements, and it can model information in ways that are deemed realistic. Since it can compute the posterior inferences easily, it allows more time to explore different models. New conceptions of models are therefore analyzed, providing wider usage among those outside the statistical community. In landslide hazard mapping, R-INLA was applied (Lombardo et al. 2020; Lombardo, Opitz, and Huser 2018) using its spatial models for discrete data.

2.6 USEFULNESS AND USABILITY

Nielsen (2012) defined usefulness as the sum of utility and usability, where usability is a quality of attribute that assesses ease of use of user interfaces. Utility defines the functionality of the product. ISO 9241-11 (2018) defines usability further by relating it to the outcome of interacting with a system, product or service. Thus, its failure means the failure of the product to function for users. Modelling efforts and design, should keep the end-users in mind where their spatio-temporal questions find answers in the resultant map(van Elzakker and Ooms 2018). Knowing when a map works depend on user research methods where the kind of people involved determines the evaluation approach (Roth, Ross, and MacEachren 2015). The context of the map use is also a factor to consider. For instance, evaluation depends on whether the mapping would be interactive and accessed via web or desktop application or presented as static information. Whatever the context, the involvement of the end-users from the beginning of the design process ensures that the final product responds to their needs(Gulliksen et al. 2003).

2.7 METHODS ADOPTED

This study combines methods from the preceding related concepts found in literature to develop a contextualized landslide hazard mapping for land use and infrastructure planning. It adopts a Bayesian generalized linear model for landslide assessment. A Bayesian approach is viewed here as appropriate to make inferences from indigenous toponyms. The algorithms of a generalized additive model also satisfy transparency requirements of data transformation gathered from indigenous knowledge.

The collection and analysis of indigenous toponyms follow structured and unstructured co-production. It proceeds from qualitative to quantitative analysis, allowing for iterations to translate indigenous toponyms for modelling. Finally, following the principles of co-production to generate information and aid policy decisions, the output is foreseen as a usable document for its target users. A test guided by ISO 9241-11 is employed to improve the modelling product.

3. Design and Implementation

The methods presented in this chapter derive from both related literature and discoveries about indigenous toponym dimensions within the study area.

Figure 2 further refines Figure 1 by breaking down toponymic information into two process steps that compose the first modelling part. The first part deals with the translation of toponyms into input variables for Bayesian modelling. The second part of modelling deals with statistical modelling. The process is iterative, where adjustments in the model ensue from information updates from the indigenous community when they use the resultant mapping. In the co-production process, the researcher assumed the role of intermediary between scientific knowledge and indigenous knowledge. Unless otherwise indicated, the term "local community" refers to locals engaged by the researcher in social media. These are residents and former residents of the study area who are active in local concerns through their membership in a Facebook discussion group moderated by elected municipal officials.

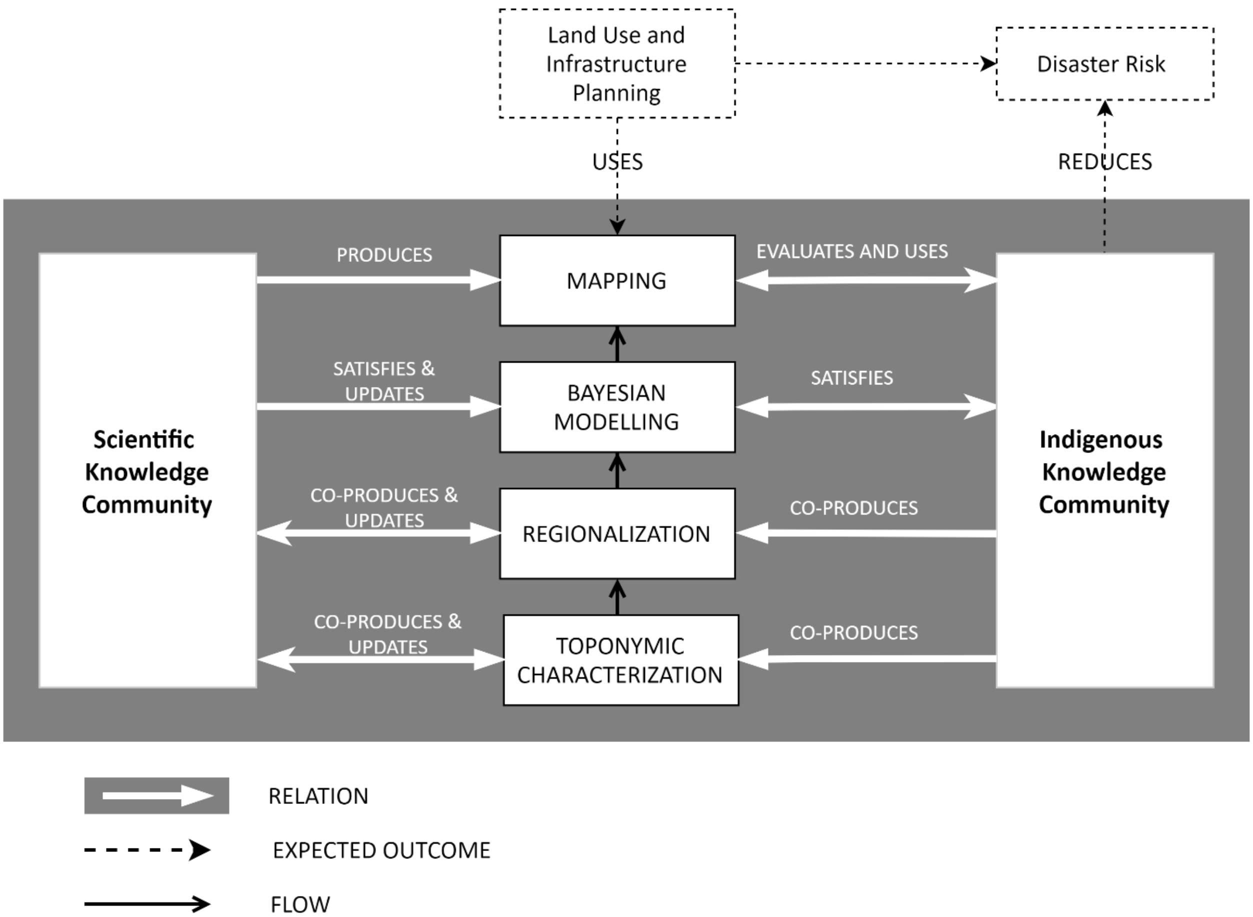


Figure 2. Toponymic co-production landslide hazard modelling process

Toponymic Characterization (3.2) and Regionalization (3.3) discusses the properties of indigenous toponyms in response to RQ1.1. "What are considerations in representing a toponym as an input variable for landslide hazard modelling?" Decisions on the methods adopted required consultations with local community representatives. These are described in Sections 3.2 and 3.3 in response to RQ1.2 "How is co-production employed in toponym translation to variables?"

Section 3.5 discusses the modelling implementation process from selecting significant variables to the selection of the model. Finally, section 3.6 presents how usability of the resultant mapping had been evaluated.

3.1 THE STUDY AREA

The study area (Figure 3) is situated in the largest mass of mountains in the Philippine archipelago known officially as the Cordillera Administrative Region. It is inhabited by different ethnolinguistic communities that share similar mountain-related cultural practices. The region resisted Spain's full sovereign authority in the country's nearly four centuries of Spanish occupation, which accounts for the preservation of indigenous toponyms.

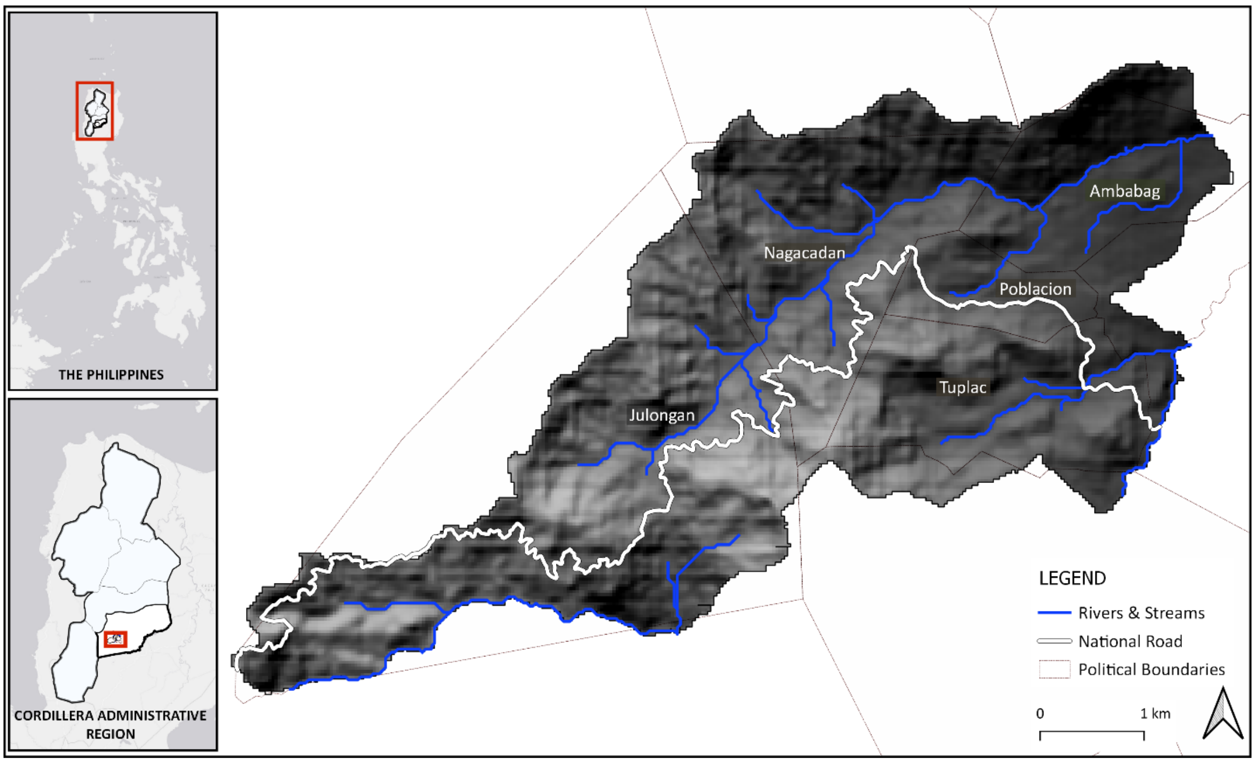


Figure 3. Study area, Philippine Cordilleras

The selected area comprises one drainage basin and three half-basins covering five (5) *barangays*, with the borders of the basins intersecting four (4) other *barangays*. The five barangays include the town centre (Poblacion) and four old villages (Julongan, Nagacadan, Tuplac, and Ambabag), now officially recognized as political-administrative units. *Barangay* is a term for the smallest political unit in the Philippines.

Elevation in this area ranges from 510 to 1626 meters above sea level (masl), with the steepest slopes at 143 per cent. Slopes below 15% comprise 12.6 % of the study area, where the oldest settlements are situated. From this rugged terrain emerged indigenous slope stabilization practices, which are evident in the landscape earning it world heritage recognition (WHC 1995).

Agricultural "water districts" (Barton 1930) partition the basins such that woodlots on the lower reaches of the slope protect and delineate the headwaters on the upper reaches to sustain the rice terraces further down the slopes. Through wet cultivation, east-facing slopes are sealed by moist topsoil all year round to prevent water seepage into the soil foundation that could cause shear failure. Within this system is a drainage network that allows water-saving in the dry season and rapid water exit during heavy precipitation through channels that divert destructive water flow on the slopes. Current land-use practises, however, are undermining this traditional system. Ninety-five per cent (95%) of the 19-km long road that traverses the study area lies on a slope per cent that is 30% and above. A trend of increasing built-up areas is apparent along this road. Protected forest zones where the road cuts through are also undergoing rapid conversion to agriculture. These land-use trends modify natural drainage and account for frequent slope failures that cause annual road blockages and the destruction of settlements in the wet months, from April to January of the following year.

A highly exposed area to hazard is the town centre, *Poblacion*, the newest settlement which sits on a levelled part of the land at the foot of an eroding mountain called *Atade*. In the vernacular language, *Poblacion* is called *Nabagtu*, meaning "higher plateau" but expressed as a verb in the past tense, articulated from the people's perspective in the downhill villages. The population of this centre started in the early 1900s when American occupation established a military barracks, followed by an administrative centre, schools and connecting roads (Barton 1930). The town is the oldest in the province, name *Kiangan*, which derives from *Kiyyangan, an* old abandoned village centre downstream.

The earliest known record of old settlements in the area is found in a 1598 report in an unpublished 1789 manuscript, *Noticias de los infieles igorrotes en lo interior de la Isla de Manila*, by a Dominican missionary (Antolin and Scott 1970). In 1801, foreign entry into the area recorded the population of its old village. However, it was in the early 1900s when more written records mentioned other toponyms in the area. The names of mountains and regions are mentioned in indigenous rituals and myths that invoke deities for protection against calamities. Religious worship focuses on avoiding evil or disaster by giving the gods what they want (Lambrecht 1962). The cosmological conceptions that formulate worship include supernatural geography that mentions some toponyms in the study area. For instance, the eroding mountain, *Atade* is mentioned in invocations because its great deity, *Imbangad,* which means "returned", must hold the rocks from falling (Martin 2021).

3.2 TOPONYMIC CHARACTERIZATION

Figure 4 outlines the steps in toponymic characterization. It is a process of matching toponyms with landslide occurrence based on its meanings and associations in the study area as well as on the information required in modelling. Each step is further illustrated to show how these are further processed and implemented using various tools.

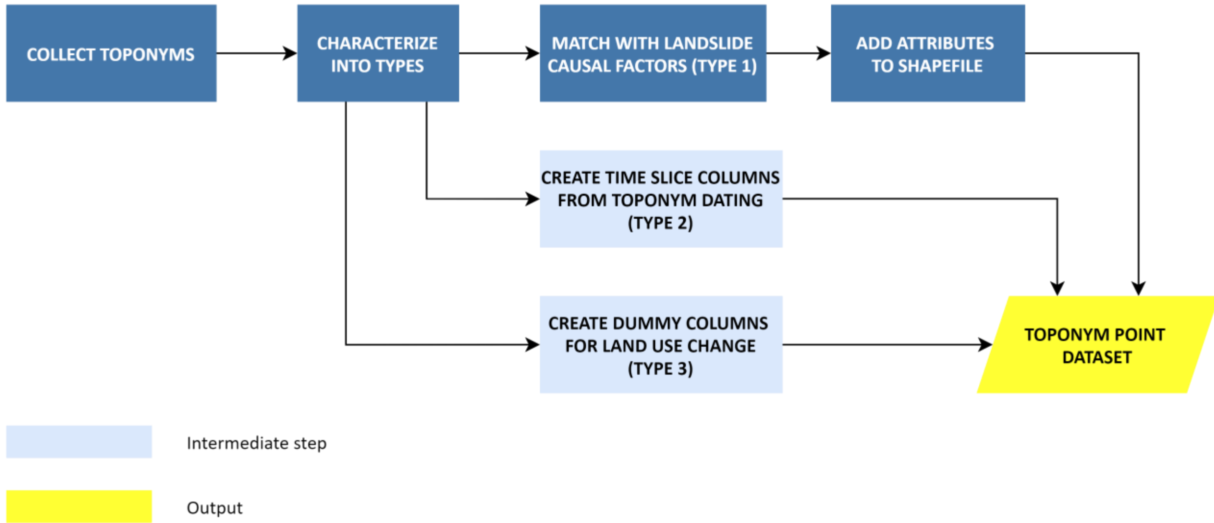


Figure 4. Toponymic characterization steps

3.2.1 Toponym collection and meanings

The output required in this activity is data on the location points of toponyms and associated meanings. These are expressed in a shapefile with corresponding attributes to be used as inputs in the next steps. Figure 5 illustrates the series of activities for this step.

Toponym information may come from existing ethnographic atlas and supplemented by elicited information from the residents. The study area already has an existing toponym dataset collected in situ from three administrative units by the researcher prior to this research. The data was produced from workshops that were organized together with the municipal government, and official representatives of each *barangay*. A representative from the National Commission on Indigenous Peoples (NCIP) was also present to validate the process. The same data set was crossed-checked by an online focus group discussion (FGD) created for this study through a published mapping of toponyms on Google Earth with restricted access to this group. The activity added a few toponyms and also relocated some points to more accurate locations. A resident participant of the FGD relied on sending marks on screenshot images to identify approximate positions.

The toponyms were also presented to a larger FB group which enriched the meanings of toponyms through open-ended discussion threads, which remained active for more than a month. Discussions included morphemes and affixes as well as references to historical events. Toponym meanings were also derived from available dictionaries published by linguists and in documents prepared by priests who resided in the vicinity between the early 1800s and 1960s. Old reports from Spanish friars that date back to the 1500s were used to date the toponyms. Dating the toponyms was done through confirmation with an archaeologist who worked in the area and with community members who confirmed approximates using written records as reference dates. Table 6 summarizes the information sources.

Table 6. Summary of information source

Information	Source	Type of information source
Toponyms	In-situ workshops, 2016	Data set
Map validation and supplementary data	FGD via Facebook discussions, 2020/2021; Google Earth , 2020	Social media + web map tool
Etymology/Associative meaning	Local community via Facebook discussions, 2020/2021	Social media
	Pataueg, 2020	Local language expert
	Summer Institute of Linguistics, 2014; Lambrecht, 1978	Dictionaries
Dating	Antolin and Scott, 1970; United States Philippine Commission (1899-1900) 1904);Roth, 1974	Reports/Notes on ethnohistory
	Acabado, 2021; Martin, 2021	Anthropological archaeologist
		Indigenous culture expert

There are general descriptions of land features that fall between toponyms and geomorphology that are not included in the collection because the local community does not use them as geo references. Instead, these areas become "active" during a natural or human-made event. An example is the formation of natural waterways on concave curvatures known in the vernacular as kulu, which literally mean “scour”. Such feature is considered in classification for certain toponym types described in the next section.

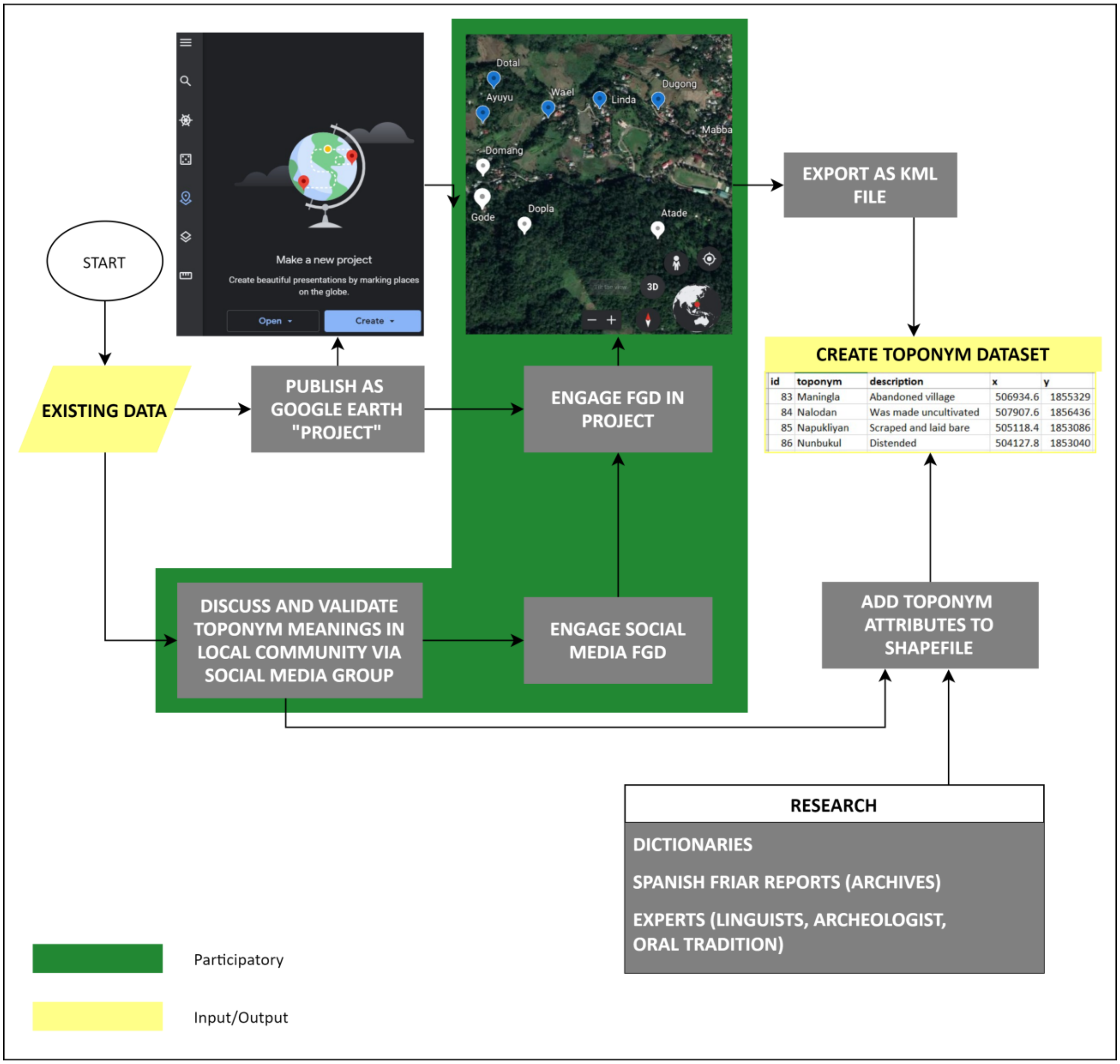


Figure 5. Collect toponyms

3.2.2 Toponym Characterization

From community discussions to online FGDs, there are some toponyms with meanings that also remain unknown. They were classified according to their traditional land use. Etymology and association of toponyms were not always related to a single dimension, category or feature. Thus, some toponyms have multiple meanings and fall into two or even all three classes. The term "characterization" therefore is adopted to distinguish them.

The primary consideration in characterizing toponyms is their relation to landslide occurrence. The second consideration is traditional land use which is reflected in the generic landscape and indicates indigenous knowledge on avoiding and preventing landslides.

Toponyms in the study area are references to locations in the generic landscape (human-made and natural) where human interactions occurred, and events took place. The toponyms feature mountains, vegetation, terrace paddies, old settlement areas and water bodies (springs, rivers, brooks). They also feature morphological, hydrological, and soil attributes that describe the presence of variables that may cause landslides. Locations that do not have toponyms indicate no remarkable feature nor record of significant experience.

In characterizing toponyms related to landslides, two types emerge. The first type is associated with recognized landslide hazard factors. The second type directly describes landslide occurrence or elements of a landslide. Those not related to landslides describe the natural and human-made landscape in terms of land feature, land cover, and land use. This is characterized as the third type. The classification of landslide causal factors is based on the parameters outlined by Suzen and Şener Kaya(2012). Figure 6 illustrates the characterization of toponyms based on their relation to landslides.

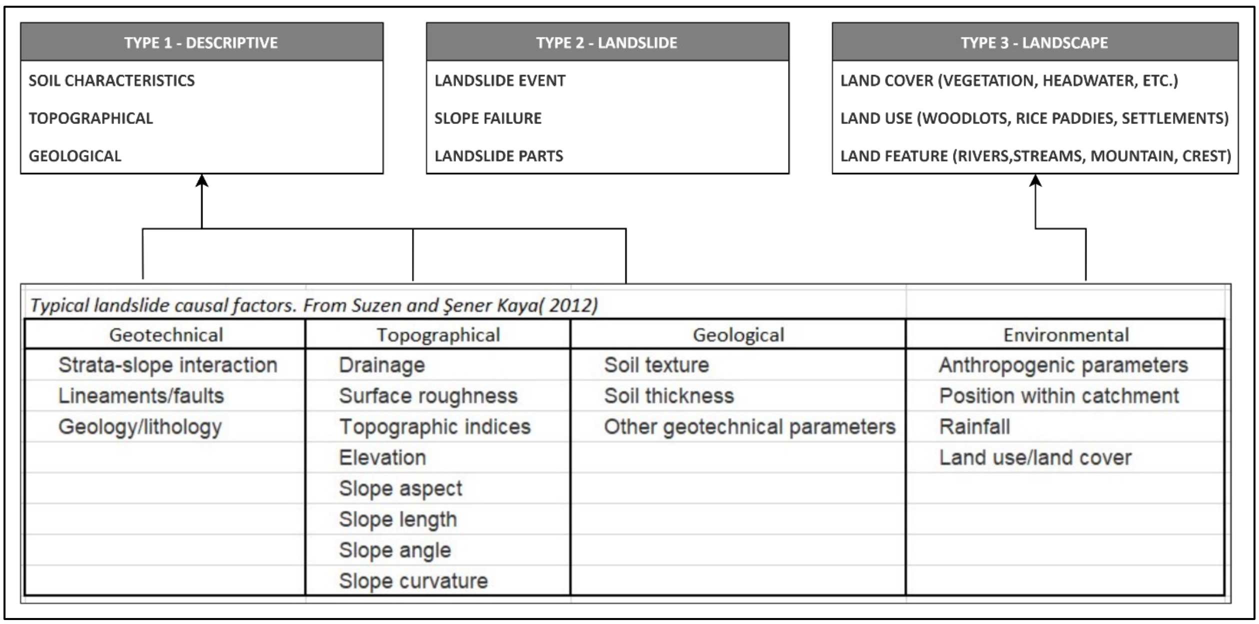


Figure 6. Characterize into types

Each type of toponym is further characterized. Type 1 toponyms indicate the presence of factors that could cause landslides. These causal factors correspond to a list of geological, hydrological, morphological, and other variables that are generally used in landslide susceptibility mapping.

Rapid matching of toponyms with this list of landslide causal factors was first done by the researcher, which narrowed down the list to nine (9) variables. Seven variables that fall under Type 1 toponyms were then used as a checklist for each Type 1 toponym. Appendix A shows a sample of the survey checklist used to generate the scores for each toponym. This checklist was created online through https://www.jotform.com/, shared to the FGD.

The checklist works like a matrix showing how toponyms match with seven possible factors that influence landslide occurrence in the area: Elevation, Slope Steepness, Soil Moisture, Slope Aspect, Lithology, Planar Curvature and, Profile Curvature (Figure 7). The numerical value for Type 1 is the summation of checks for each toponym. The checklist is then presented to the FGD. Decisions are finalized through the larger group, where the checklist questions become part of an open discussion on etymology and associated meanings. The output of this is in the form of added attribute columns to the toponym dataset

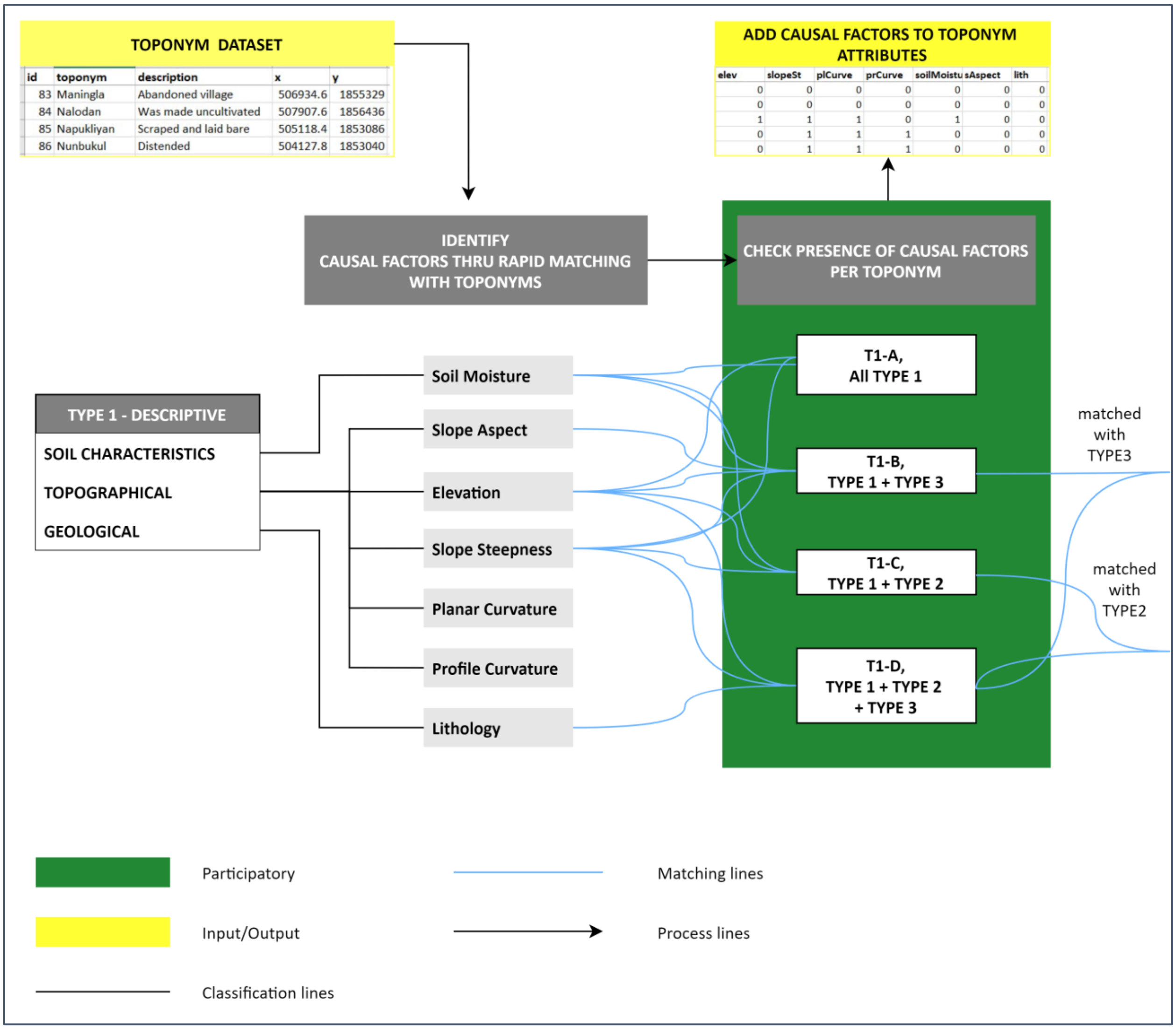


Figure 7.Match with landslide causal factors

The method of matching Type 1 toponyms with a checklist of landslide causal factors captures to some degree the multiple dimensions of indigenous toponyms. From Figure 7, Type 1 toponyms already consist of 4 further categories. This method may still be considered a simplification but necessary to generate a representative numerical value to the toponym.

In the above checklist, soil moisture is often observed on the concave curvatures of the slopes where natural drainage lines are and on the faces of the slopes between these drainage lines to some point when the slope starts to form a convex curve. Thus, planar concave curvatures, known in the language as “kulu” sites, are subsumed under Soil Moisture.

This associative checklist cannot capture other information, such as the dynamic changes mentioned above. Toponyms may therefore be a composite of causal factors and can explain other factors that are not apparent. For example, in the area, the name of one mountain connotes erosion tendencies but scores low on the number of causal factors. Explanatory variables that are available in literature cannot fully capture this effect. Another potential causal factor in the area defined by toponyms is wind direction, which, when combined with the poor status of vegetation cover, results in slope instability. This was not included because of the lack of literature to confirm this.

Toponyms that further match Type 2 characteristics are considered "plus one" landslide counts in the area bounded by the toponym, which will be determined later when connected with spatial units. All toponyms of this type are further classified according to the date they existed, represented as time slices. Figure 8, which is a continuation of Figure 7, illustrates this as an intermediate step where four time slices are created. For example, toponyms that exist in the 16th century (1598) are given a count of 1 event within that time slice. The space in which this happened is assigned when spatial units are defined in Section 3.3. The number of these events are added in the landslide inventory presented in Section 3.4.

Toponyms that match Type 3 characteristics do not indicate the presence of factors that cause landslides. These are toponyms that refer to old settlements (not necessarily settled today), rice terraces, woodlots and old vegetation. All of these describe the traditional landscape. Although vegetation cover may be a variable that can cause landslides, this is not included under Type 1 toponyms because it is its absence that is likely to cause slope failure. Land-use change as a landslide causal factor can be extracted from Type 3 toponyms through dummifying variables. These may be from land-use changes on forest covers such as quarrying, road constructions, and new settlements that alter or block natural drainage lines. Land-use change on forest covers and vegetation-derived toponyms are given a score of 1 to indicate the presence of slope instability. However, this cannot be assigned at this stage because this condition depends on the area covered. The affected area can only be defined during the toponym assignment to slope units described in 3.3.2. On the toponym dataset, attribute columns for land cover and traditional land use are added that classify these land covers/use. Later, dummy variables are added to account for the absence of these descriptions in specific SUs where there is an observed change. Figure 8 illustrates the process of matching toponyms with Type 3 characteristics.

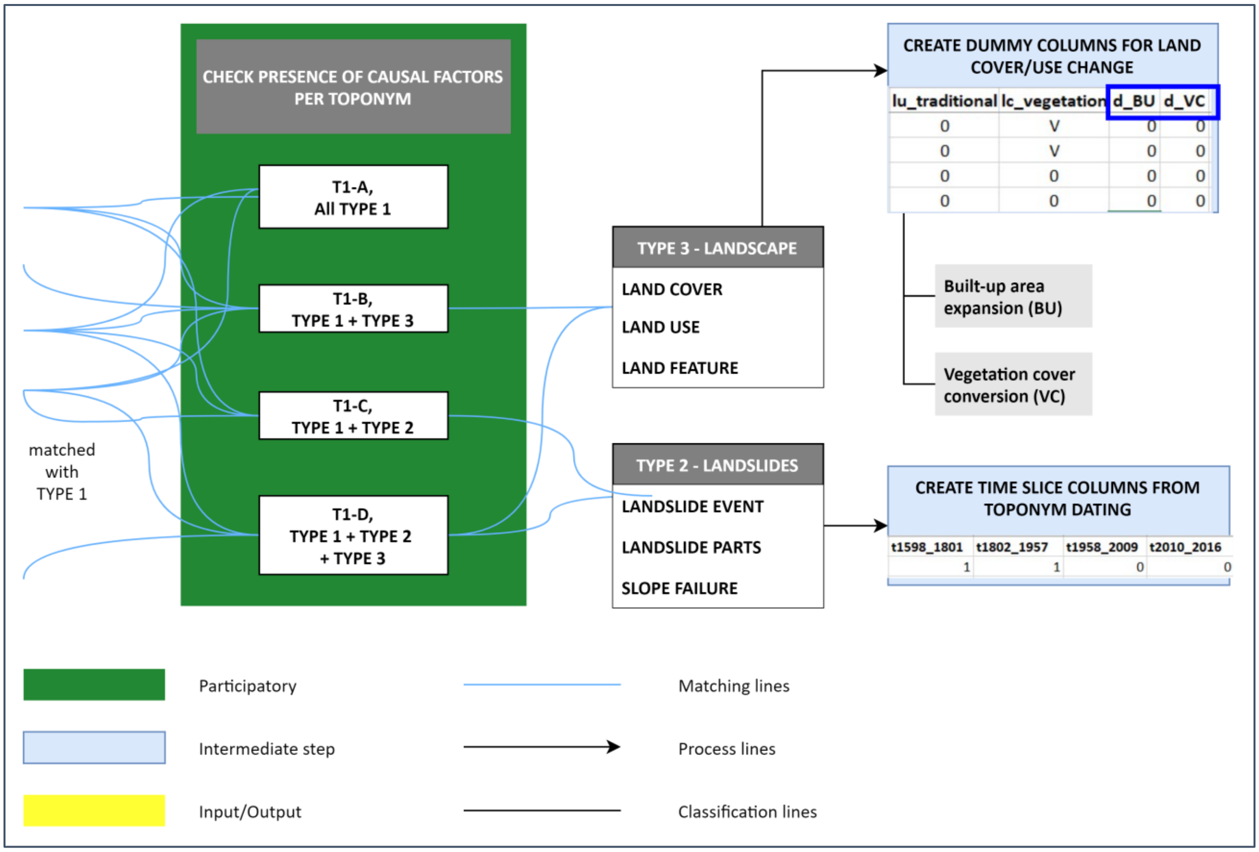


Figure 8. Create columns that match type 2 and type 3 characteristics

3.3 REGIONALIZATION

This section responds to RQ2.1, which states as, “Factoring in the conditions mentioned in RQ1, which methods are suited to generate quantitative input variables for modelling from toponyms?”

Although toponyms are characterized in this model into different types, each can also belong to multiple types. One toponym can represent causal factor features (Type 1), and is counted as a landslide event (Type 2). To prepare toponyms as variables for modelling, these properties must be first referenced in space, then quantified. Figure 9 shows the workflow. The first step is to optimize spatial unit partitioning, followed by toponyms connection to these spatial units. The dummy columns created in Section 3.2 are then populated before these spatial units are grouped according to their properties.

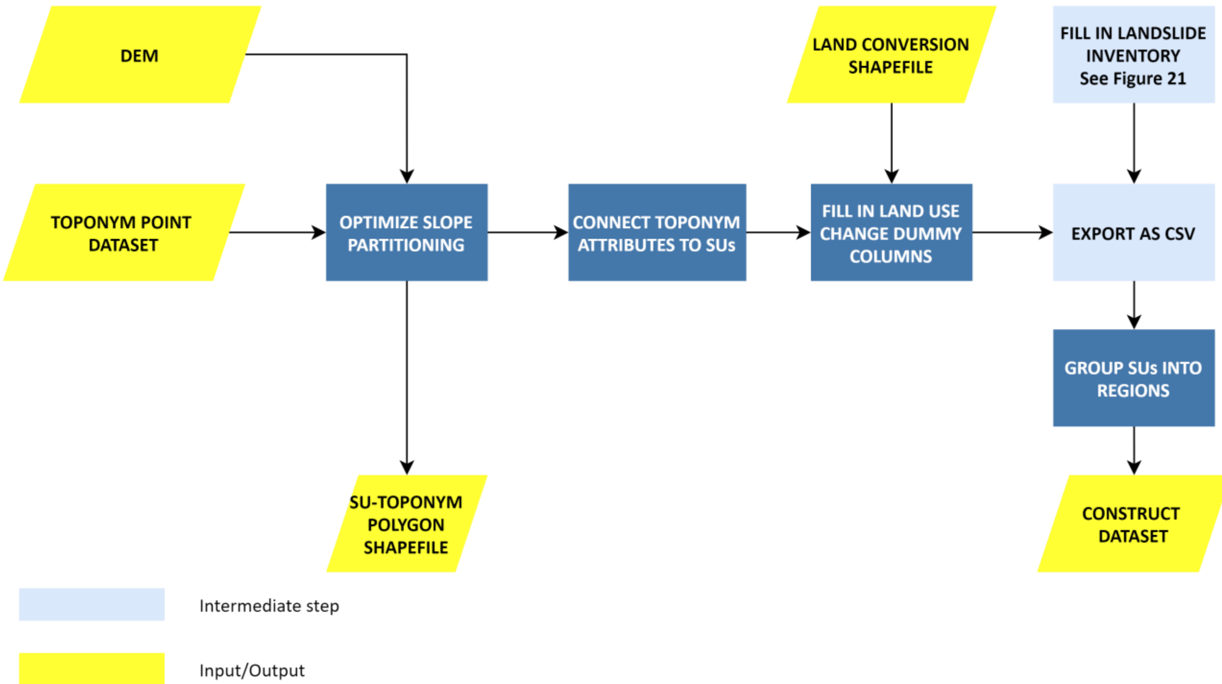


Figure 9. Regionalization steps

3.3.1 Optimizing spatial unit partitioning

The initial step is to decide whether the spatial units to be used should be grid-based or terrain-based. Slope units (SUs) are terrain-based and represent topographical units. An SU is also known as a "half-basin" because ridgelines and valley lines define it (Figure 10). Both landslides and toponyms can adopt different types of mapping units. However, toponyms have ambiguous boundaries. Clearer delineation may only be observed in manmade landforms, such as named plateaus bounded by retaining walls. This is the case in the study area, where settlements are named in various "platform" terminologies and are well-defined. For indigenous placenames, spatial precision is also relative and exclusive to the local community's knowledge. One thing certain about toponym boundaries is that these are not grid-based. Often they adopt natural features such as ridges, rivers and streams (Tsai and Lo 2013) and are close to terrain configurations. Therefore SUs align well with the spatial definitions of indigenous toponyms. For this reason, the proposed model adopts SUs as spatial units.

The purpose of optimizing spatial unit partitioning is to define the minimum size of mapping units that simplifies the model but does not diminish the information content of the two elements for modelling (landslides and toponyms) and their relation. The minimum spatial unit size defines the optimal terrain subdivision by capturing the amount of information detail within each of these elements without these having to be divided further into finer units. In the overall model, larger spatial units mean fewer observation units. Thus, optimizing spatial unit partitioning also means maximizing the amount of information from the fewest observation units. In this study, it is the minimum distance between distinct terrain configurations defined by toponyms that determine the optimal minimum size of spatial units.

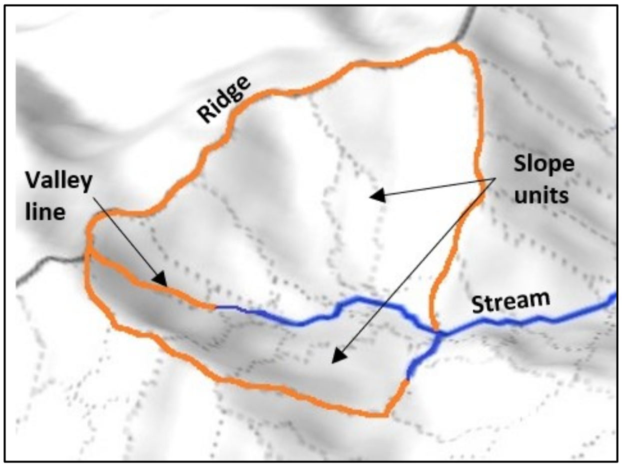


Figure 10. Slope units (SUs)

The “r.watershed” module in GRASS in QGIS extracts “half-basins” based on a threshold value of drainage and divide networks. An algorithm developed by Alvioli et al. (2020) implements an iterative procedure of optimizing SUs by starting the partitions from bigger half-basins, further subdividing these into smaller half-basins. The subdivision of an SU stops then flagged when the half-basin meets the internal homogeneity and size criteria of an SU defined by the modeller. In the work of Alvioli et al. (2016), the determinant for this homogeneity is terrain aspect segmentation.

For the proposed model, the determinant is the toponym meaning descriptive of geomorphology. The iteration procedure starts the partitioning from smaller half-basins where SUs are merged to adjacent SUs that share the longest boundary. Using the “r-watershed” module, smaller half-basins are merged to bigger basins by employing a merging process using the algorithm, “Eliminate selected polygons”. Appendix B displays a detail of this workflow.

Figure 11 illustrates how toponyms define the merging iteration. The number of SUs is reduced iteratively by merging SUs as long as toponyms that describe different geomorphological features continue to be in separate SUs. This is checked with local knowledge, and the minimum distance between toponym points produced in Section 3.2. Finding the closest pair of toponym points that have distinct geomorphological properties can be checked through the “Distance matrix” operation in QGIS which shows the minimum distances between toponym points using the “Summary distance matrix” as output matrix type.

Remnants of lines that are visible in the last iteration can be cleaned through the “v.clean” operation in GRASS. Finding the optimal granularity of SUs may be taken as an optional step here and the decision to use a minimum SU size that is smaller than the optimal size (thereby increasing the number of units) is set by the modeller. In this case, the optimal size is preferred.

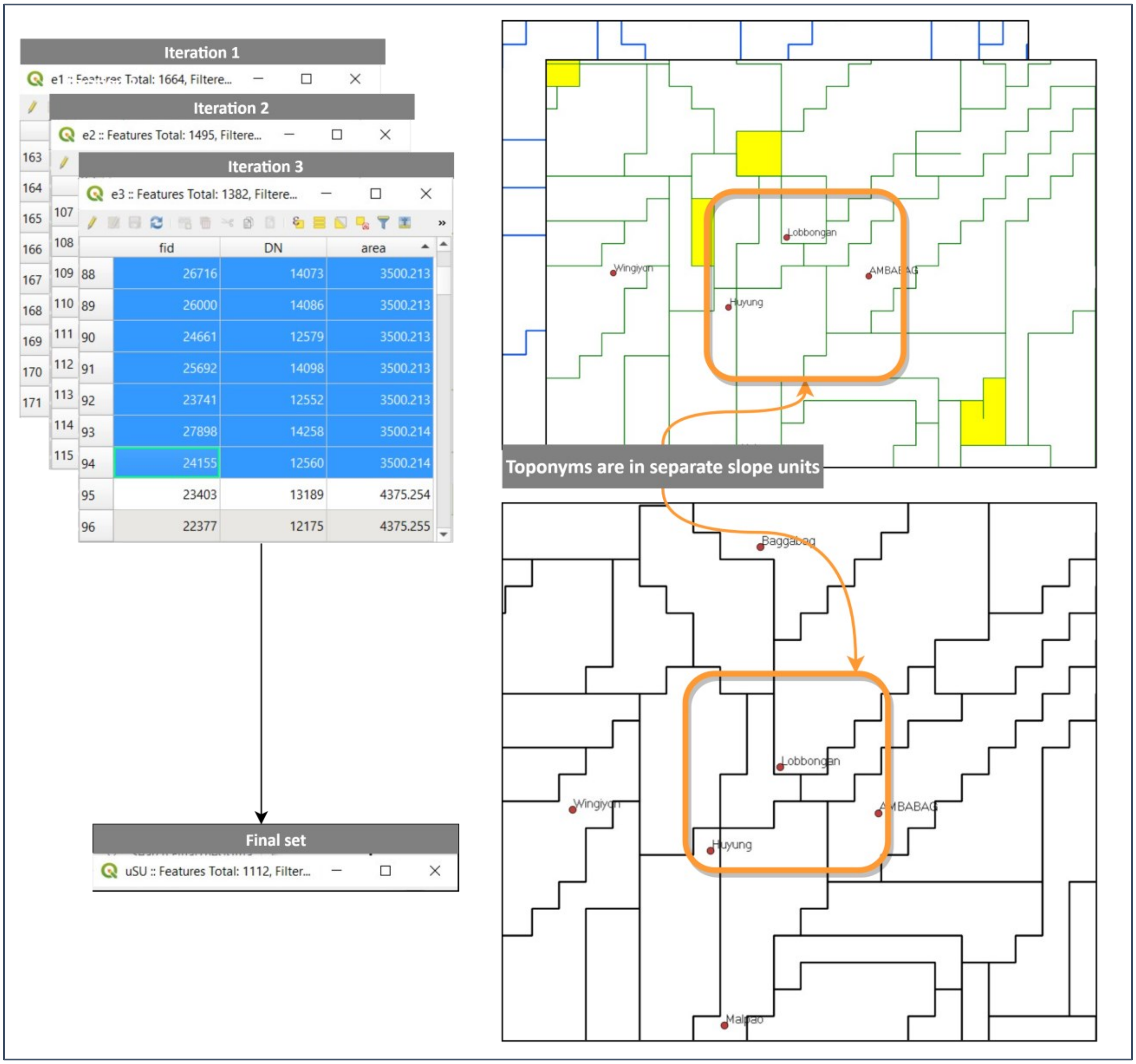


Figure 11. Select polygons to be merged

3.3.2 Connecting toponym attributes with SUs

Inputs for this process are the toponym point dataset and the SU polygon dataset. The output is an SU dataset with toponym attributes exported as comma separated values (CSV) for further processing. The general steps use GIS operations that first connects a point (toponym) to a polygon (slope unit). This polygon then connects to its adjacent neighbouring polygons, which continues *before* these polygons cross the imaginary boundary of the toponym. Polygons that lay across these boundary lines connect to toponyms guided by a set of rules. The imaginary boundary is calculated through Voronoi polygons based on specific toponym points. Figure 12 illustrates the general steps.

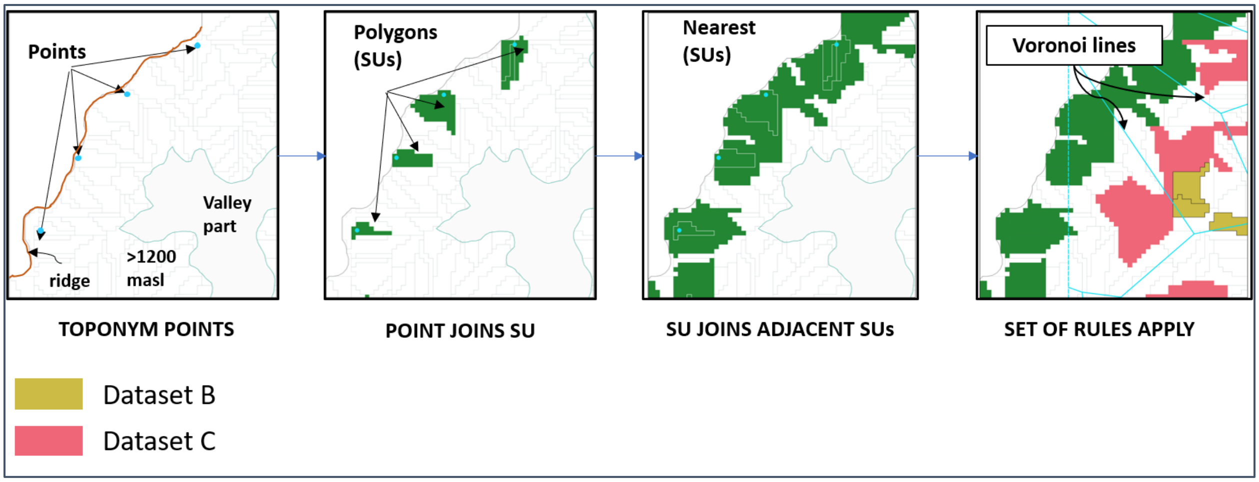


Figure 12. Connecting toponym attributes with SUs

Some important conditions have to be clarified first in the calculation process. First, the toponym points of mountains and those found on the uppers slopes are not centrally located in the area covered by the toponym. In a basin, the central approximation of toponym points more often applies to the valley parts than in the upper slopes due to the denser toponyms in settlement areas. A toponym point in the upper slopes may represent the curved rim of that basin or branch out to another chain of hills. Second, these points are relative to the study area and the residents who mapped these. Thus, a toponym point in the upper slopes may also represent a range of hills with defined ridgelines that extend beyond the study area's boundaries. For the purpose of connecting toponyms with the respective SUs in those areas, more representative points of the same toponym are added in the range of hills that it covers. For headwater toponyms, the representative points are placed on slopes 1200 masl and above. Based on prior knowledge and mapped vegetation, this is an elevation line where it is certain that the area above it is considered headwaters. Below this line is a mixture of land cover(forests) and land use. Mountain/hill toponyms that are found below this elevation are also given more representative points from ridge to river if there are no land use toponyms found on its slopes.

Connecting toponyms starts with SUs from upper slopes, valley parts then on SUs on elevation margins and voronoi lines.

The process is executed in the following steps:

Step 1 From the toponym point dataset, create three separate toponym datasets: Set A) mountains and headwaters; Set B) settlements and agricultural areas and; Set C) remaining toponyms. For the shapefile on headwaters, which consist of a chain of hills, add representative toponym points on the slopes below the ridgeline above 1200 masl.

Step 2 Select slope units that contain these points. Through GIS processing, the "extract by location" operation extracts features from the SU shapefile where the feature contains the toponym points (e.g. mountains and headwater shapefile). This results to an intermediate layer "Extracted (location)" which becomes the input for the operation "Join attributes by location" where the "Joined layer" is the same toponym shapefile. In the "join type" query box, choose "one-to-one". The 2-step process creates a "Joined layer" in the form of an SU shapefile with attributes adopting the toponyms that they contain (Appendix C). This is performed for the three datasets. Alternatively, direct spatial joins and filtering can be performed directly using one dataset but this entails more data cleaning and manual checks.

Step 3 Add surrounding neighbours for sets A and C. The previous step ensured that all toponyms and their attributes are accounted for in the 3 SU datasets. At this stage there are larger gaps between mountain toponyms on the upper reaches of the basin than on the valley parts. The operation "Select by location" picks out slope units that are adjacent to the main SU containing the toponym point extracted in Step 2. These selected SUs are added through the operation "Join attributes by nearest". This creates another new intermediate layer "Joined" for the 2 datasets. It should be noted that there are already some overlaps between the "Joined" C set and the settlement shapefile, "Extracted" B. In this case the vector geoprocessing tool "Difference" is implemented to subtract these polygons from "Joined" C, resulting to "Difference" C. The neighbouring SUs surrounding settlement areas that have now been assigned set B toponyms will be inspected later. Appendix D illustrates these intermediate steps.

Step 4 Merge the 3 datasets. This creates a dataset "Toponym_SU" that only shows SUs that take the properties of the toponyms that they contain (Figure 13). Overlaps between these sets will be manually checked after all gaps are filled.

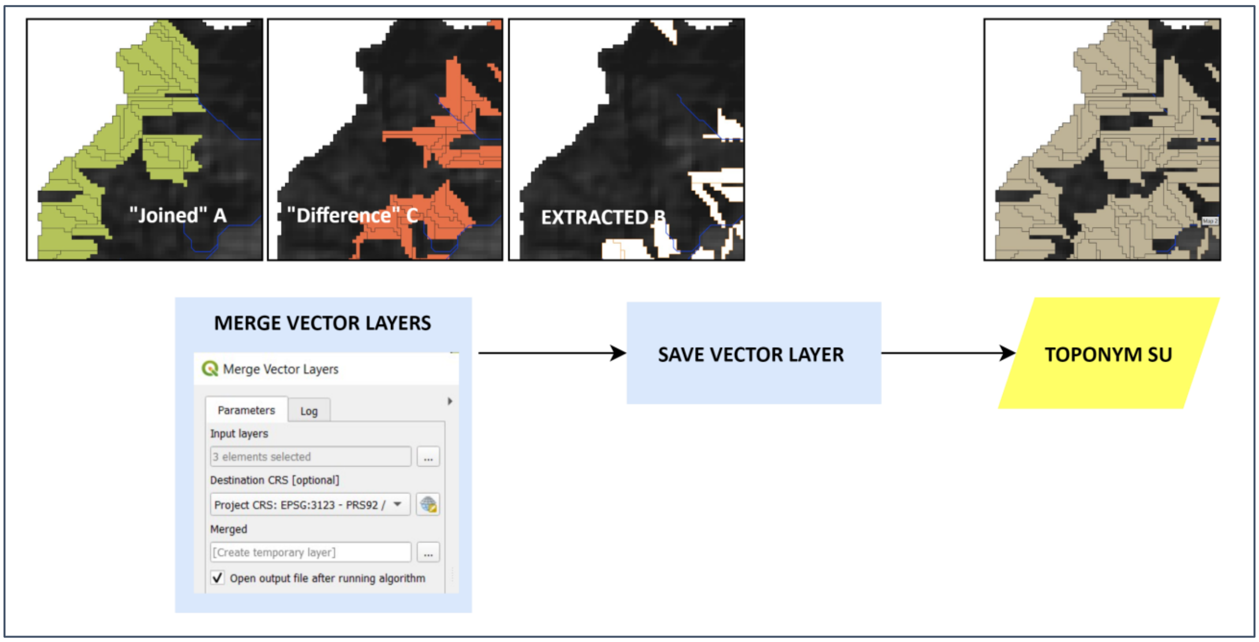


Figure 13. Merging the datasets

Step 5 Connect remaining SUs from the top parts of the basin down to the valley areas. This follows the same process of adding SUs to the nearest toponym set as in Step 3 but this time it will be a manual selection of SUs from the Toponymic SU dataset. To create guides along the slopes, the "Voronoi Polygons" operation is employed using the Set C dataset as input layer. This operation takes the toponym points layer and generates a polygon layer representing approximate boundaries of the toponym points. Figure 14 and 15 show how the voronoi lines are used as guides.

This is a step in the process that needs careful checking because this involves decisions in assigning toponym properties on SUs where they overlap or on SUs that does not have a toponym. In general the land cover toponym from the top takes precedence over land use on the valley. Figure 14 illustrates some rules for Case 1 and Case 2 to serve as guides for this step. The labels show the literal translation of the toponyms.

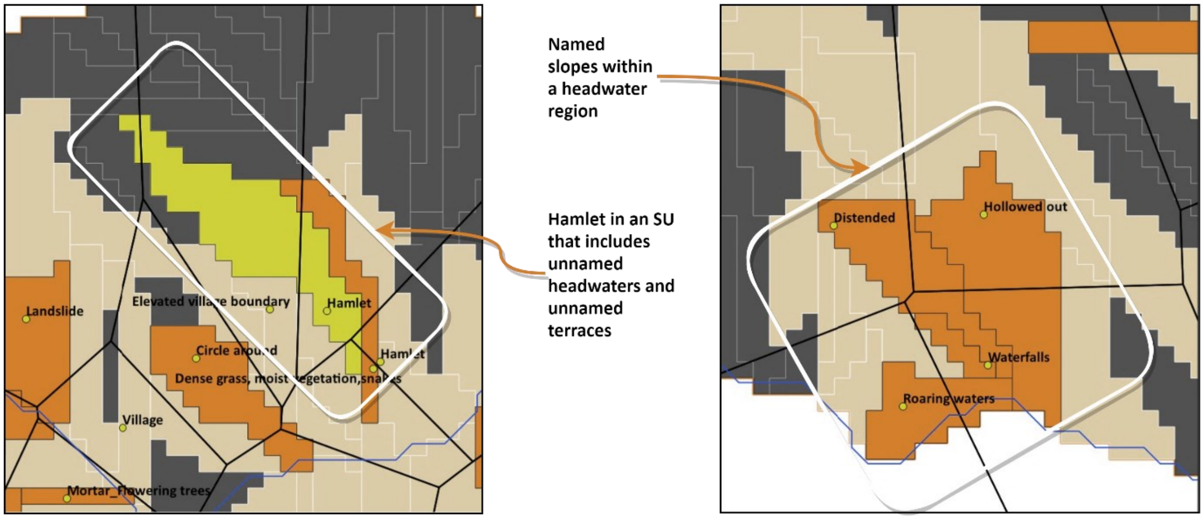
Case 1: Near headwaters. If the SU lies across the voronoi boundary that is transverse to the slope, the SU adopts the headwater name.

Case 2: More than 1 toponym. If there are more than 1 toponym in 1 SU, then the SU is named according to these toponyms. The SU adopts the attributes of all three. If there are attribute values that are conflicting in one SU, the toponym that does not describe the underlying landscape is not included. In the illustrated example, the toponym "abundant beads" was taken out because it does not reflect the underlying landscape after checking this with a satellite basemap.

Figure 14. SU cuts across boundaries

Case 3: Settlement toponym with unnamed land cover and use. If the toponym defined as a settlement area is in an SU that includes traditional agricultural areas and headwaters above, the SU is named as a settlement with headwaters and agricultural area or "H_T_hamlet". In the example in Figure 15, the hamlet toponym is on an SU that cuts across other areas such as an unnamed headwater or forest above and unnamed terraces below it. The forest above and terraces are referred to in the local community as "headwater of " and "terraces of " the hamlet.

Case 4: Toponym within a toponym. If the SU is on a named mountain with a named slope, the SU adopts both headwater/mountain and named slope properties. In the example in Figure 15 the named slopes lie on the lowest slopes of "headwater region". The slope units are therefore renamed "headwaters_hollowed out", "headwaters_distended_hollowed out" (if this is reflected in the underlying terrain), and "headwaters_roaring waters".



CASE 3

CASE 4

Figure 15. Toponyms with other named and unnamed places

Case 5: If there are blank spots or SUs in places with features that are obviously not descriptive of its named neighboring unit, these are labeled as "neighbor_r" to refer to the neighboring SUs. These are described by the locals as "before" this place or "between" two places depending on their reference point. The SUs may adopt the attributes of the nearest toponym within that side or aspect of the slope when landslides occur in this unit.

Step 6 Check overlaps in the final Toponym_SU dataset. The overlaps between those from set A (headwaters/mountain) and set C (thematic descriptions of the slopes) indicate renaming of some slopes (Case 4).

Step 7. Identify SUs that are affected by land use change.

From Step 6, the dataset shows SUs with toponyms and their attributes with the dummy columns for land use change. One column identifies SUs for built-up areas, named "d_BU" which refer to land conversion from toponym-described land cover or land use to roofed structures and pavements. The built-up areas in this case include settlements, paved surfaces, and other industrial activities (e.g. quarrying). The other column identifies SUs where vegetation conversion is evident, named "d_VC". This includes toponyms of mountains and the headwater region. Built-up areas in the form of roads and new settlements are also accounted by superimposing the toponym SUs on existing land use base maps. The two dummy variable columns for vegetation change and built-up areas can now be filled-in to indicate SUs where land use change occurred.

Step 8 Cross-check with community representatives. The process is iterative and may be updated even during usage of the final product.

3.3.3 Summary of Toponym-based landslide causal factors

Table 7 summarizes the landslide causal factors considered as a result of toponym matching and observed changes from toponym-described land use and land cover.

Table 7. Summary of causal factors considered

Causal Factors		Considerations
Topographical	Elevation	Indicated by the toponym
Geomorphology	Slope steepness	Indicated by the toponym
Geomorphology	Planar curvature	Toponyms describe the convex curvature
Geomorphology	Profile curvature	Toponyms describe the convex curvature
Geological	Soil moisture	Toponyms relate this with concave profile and planar curvature and sometimes with slope aspect
Topographical	Slope aspect	Description of south to south-west facing slopes only based on prior knowledge of slope failure during extended heavy precipitation
Geology	Lithology	Associations of toponym
Human Activities	Built-up area	Land use differ from toponym description
Human Activities	Vegetation change	Land use differ from toponym description

3.3.4 Toponymic Regions

The output of the previous process is a dataset that shows the Type 1 attributes of toponyms. These include the checklist of landslide causal factors processed during the toponym characterization stage (3.2.2). SUs are grouped according to the same presence and combination of the causal factors. Henceforth, these are called "predictors".

These regions assume the combination of the predictors present. The grouping of SUs in effect translates toponyms as geospatial regions with composite predictors. The same toponym may be in different regions and different toponyms may belong to one region class. There are cases where one region is equal to one SU. The number and combination of predictors are added as information into the SU dataset. These were then assigned codes to simplify notation.

The regions as a composite representation of the presence of predictors takes the absolute value of their sum as its numerical value. Where there are 3 predictors, the region's value is 3. The assumptions are:

- Since toponyms hold information associated with landslide predictors, the resulting aggregation of predictors combines these effects, where a higher number of predictors means a higher effect.
- The quantity of predictors in each region is proportional to its combined effect, where one predictor equates to a count of 1.
- The effect of predictors is constant for each region.

The kind of predictors present and their interactions within their specific regions should not be ignored and must be preserved in the final analysis. Each region has a different set of predictors that provide information on land use policies and infrastructure interventions. The set of predictors combine inherent characteristics of the terrain and land use activities that require specific and appropriate policy responses.

This information must therefore be rendered in the resultant mapping of the selected model. Figure 16 provides an illustration of information required in the dataset both for modelling and for rendering. The notation *tg* represents "toponymic geospatial regions". In the example, tg10 has three predictors: Elevation (E), Slope steepness (S) and, Lithology (L).

tg	Predictors
0	non-assigned
22	SVc
10	ESL
41	ESLBu
43	ESLBuVc
42	ESLVc
55	SMVc
56	SMBuVc
40	ESMABuVc
21	SBu
25	ESplCprCMALVc
26	ESplCprCMALBuVc
46	SplCMBuVc
45	SplCMVc
51	SplCBu
19	ES
37	ESMBuVc
35	ESMBu
7	ESM

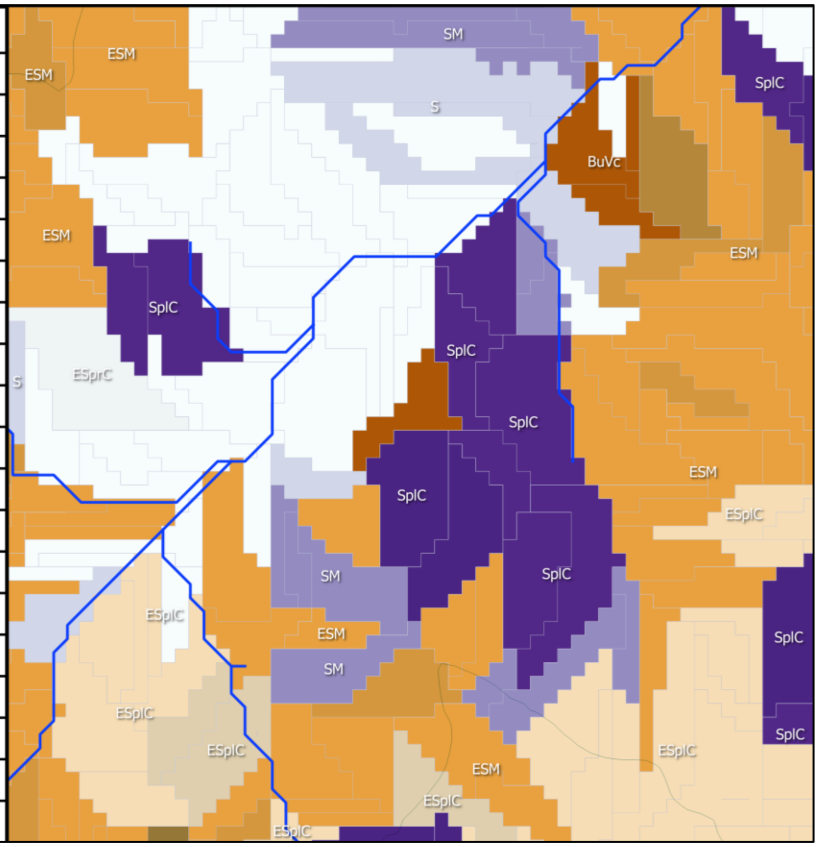


Figure 16. Toponymic geospatial regions, "tg"

3.4 LANDSLIDE INVENTORY

The collection of landslide information is co-produced. Table 8 shows the combined observations of landslide events. The inventory of landslide events from 2009 is a set of points from different sources. Landslide events of 2009 were memorable to the local community, but Google earth images only provided clear images of 2010, which showed old landslide scars of 2009. The images of 2010 were still used to corroborate 2009 recollections. The inventory of landslides before 2009 also comes from recollections of the community, discussed in social media triggered by landslide-related toponym discussion. The inventory from Google earth was exported in Keyhole Markup Language files. These were then converted as shapefiles in QGIS. Each SU covered by the toponym is given the value of "counts" of recalled events plus one count for the toponym. For example, if the local community can recall 2 events in the location covered by a landslide toponym, the SUs that it covers are given a count of two (2) plus one (1). Identification points are described to differentiate multiple events within an SU. Landslide inventory uses Varnes (1978) classification to identify the types of slope movements present in the study area.

Table 8. Combined approaches in landslide inventory

Combined Method	Time	Identification points
Toponyms and community recall on social media	Pre-2009	As counts within the slope units covered
Google Earth images and community recall	2009-2010	Highest position on scar
Google Earth, OSM*,community workshops, GPS Survey	2014-2016	Highest position on scar
Google Earth, OSM*, social media posts, FGD	2017-2020	Highest position on scar
Social media posts and discussions	Early 2021	Approximate point in slope units

*OSM- Open Street Map

3.5 MODEL CONSTRUCTION AND SELECTION

3.5.1 Probability distribution

This section responds to the research question RQ2.2 which is stated as: "What prior distribution captures the information provided by the data and toponyms?".

The analysis here focused on spatial pattern that matches the data. In the Bayesian paradigm, the aim is to estimate the joint posterior distribution. The selected toponymic regions provided spatial information that is assumed to improve the reliability of these estimates. In orer to estimate the posterior marginal distribution, the integrated nested Laplace approximation (INLA) method was employed. This method is implemented in the INLA package available for the R programming language.

The dataset shows multiple temporal landslide events that are partly supplied by indigenous toponyms, which indicate counts of their occurrence. Landslide events are observed within given boundaries in the form of SUs, which presents irregular lattice data. Thus for the landslide occurrences y_i within each SU $i = 1, \dots, n$, the Poisson model is

$$y_i \sim Po(\mu_i)$$
$$\log \mu_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}$$

where μ_i is the mean of the response variable, $x_1, \dots x_p$ are p explanatory variables with effects $\beta_1, \dots, \beta_p$ and β_0 is the intercept representing the overall mean after accounting for the covariate effects.

Spatial models for lattice data are usually defined as random effects with variance-covariance structure that depends on the neighborhood structure of the areas (Gómez-Rubio 2020). Neighbors or areas that share the same boundaries tend to have a similar number of events, which shows spatial autocorrelation. If observed data from neighboring areas exhibit higher correlation than distant areas, this correlation can be accounted for using the class of spatial models called "CAR" models (Conditional Auto-Regressive) introduced by Besag (Besag 1974; Morris 2019). This effect is plausible for landslide hazards, where neighboring SUs have similar surface and subterranean properties. Downward movements also affect surrounding areas. Another important condition to note is that most of the toponymic regions that are used here as explanatory variables are a composite of predictors with various combinations, among which are land use change factors. The type of predictors present has a fixed effect for landslides and their combination may have a random effect. Toponymic regions as explanatory variables however, will be assumed in the modelling to have fixed effects. The underlying structure in the data was confirmed by an over-dispersion test implemented using the dispersion() in the "AER" package. Random effects are therefore added to the model to account for this extra-Poisson variability. The model is expressed as:

$$\log \mu_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + u_i + v_i \qquad (3)$$

where, u_i is a random effect specific to area i to model spatial dependence between the relative effect, and v_i is an unstructured exchangeable component that models uncorrelated noise.

The models considered are expressed as follows:

Model 1: (iid)

$$\log \mu_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + v_i$$

Model 2: (CAR or Besag)

$$\log \mu_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + u_i$$

Model 3: (BYM=iid+Besag)

$$\log \mu_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + u_i + v_i$$

For comparison, the base model (Model 0) will also be presented to show fixed effects. Model 1 assumes that the confirmed over-dispersion or random effects are "independent and identically distributed"(i.i.d). The $i.i.d$ random effects account for any unmeasured explanatory variables and unstructured variability. The assumption is that any data point is independent of any other data point. The other assumption is that the events are identically distributed.

Random effects were also assumed to follow the CAR models specified in Models 2 and 3. The models show how spatial relations present among the data differ from the explanatory variables. CAR modelling is a technique used that borrows information from its neighboring areas. It works in a way such that the probability estimates in a given areal unit is conditional on the level of neighboring values. Relationships between these neighbouring areal units are described by an adjacency matrix of the slope units . In this model, the neighbor relationship is symmetric but not reflexive; if $i \sim j$ then $j \sim i$, but a spatial unit is not its own neighbor.

The spatial relationship between these units are either independent or correlated. The model is called an improper CAR (ICAR) if it assumes complete spatial relationship between these regions. It results to a singular matrix, with some unrealistic consequences (Assuncaõ and Krainski 2009; Lavine and Hodges 2012) which can be overcome by adding a constraint (Cramb et al. 2017; Morris 2019). The BYM model is the result of ICAR and *i.i.d.* essentially combining both structured and unstructured random effects but each cannot be treated individually.

The formulas corresponding to the three models are in the R-INLA package that is used to perform Bayesian analysis in this study. Latent random effects are specified with the f(). The neighborhood structure can be obtained by using the function “poly2nb” from package spdep (Bivand and Wong 2018), which will return an “nb” object.

3.5.2 Selection and analysis of variables

This section responds to the research question RQ3.1 which is stated as: “What are criteria to evaluate and select toponymic variables for landslide hazard modelling?”.

Before the set of toponymic geospatial regions are considered explanatory variables for modelling, this underwent elimination. First, all toponymic geospatial regions (tg) where no observed landslide occurred are taken out from the set, followed by the formulation of the generalized linear model(glm) using all remaining as explanatory variables. The glm describes the relationship between landslide counts and the toponymic geospatial regions. Second, a stepwise selection process in both directions was applied to the glm. Stepwise selection builds up the model step-by-step, each time either adding or subtracting a variable based on the Akaike Information Criterion (AIC). The AIC distinguishes among a set of possible models the best one that fits the data. This was implemented using the function step() in the R “MASS” package. The stepwise process in both directions was selected to arrive at the fewest possible explanatory variables with the least AIC. It is noted that the backward selection process also yielded the same model.

The third step is a multicollinearity diagnosis, simultaneously done with checking how the elimination of variables from the set that passes this test can cause the least loss of information. Identifying variables to eliminate was guided by the result of the stepwise selection process, which shows the variables that can cause minimum information loss when removed. The combination of variables to remove is based on the researcher's observation of variables that may impact the dependent variable (in this case, landslide counts). Multicollinearity test was implemented using the function vif() in the R "car" package, which provided the Variance Inflation Factor (VIF) of each model with a given set of variables. Models with variables that are above 5.0 VIF values are rejected. The model with the set of variables that passed the multicollinearity test and with the least AIC was then used in modelling.

The fourth step of elimination was performed during the Bayesian modelling process by examining the overlap of their 2.5% and 97.5% posterior estimates with zero. Here, Bayesian inference returns the posterior distribution of possible effects presented across a value range. Within this range is the credible interval (CI) containing a particular percentage of probable values. The selection of variables to keep in this experiment made use of the 95% CI due to its wider range. In modelling landslides over an area, the conservative perspective was taken, which considers the widest possibilities of effects given observed data.

The selection of variables utilizing CIs is an iterative process. The models considered, returned posterior distributions where effects are within the zero probability range, which further reduced the set of variables.

3.5.3 Model selection and visualization

This section responds to the research question RQ3.2 which is stated as: "Based on which criteria and which process are models selected for their predictive performance?"

The goodness-of-fit of the models was compared using the Deviance Information Criterion (DIC). A smaller value of DIC indicates a better model fit. Like AIC, the DIC is another criterion that provides an approximation of predictive accuracy but it uses the average log-likelihood over the posterior distribution as a measure of goodness-of-fit. DIC is calculated automatically by R-INLA.

Iteration was also practiced in the selection of the final model where the DIC comparison is combined with the removal of variables that fall within the zero probability range of the CI. Again, this was performed by exploring the data and the impact of variables. The iterative process built the four models by retaining the set of variables that does not increase the DIC value. The model with the lowest DIC is selected for map visualization.

3.6 EVALUATION OF USEFULNESS

3.6.1 Usefulness Factors

This section responds to RQ4.1, stated as follows:

"Which factors define the usefulness of the resultant mapping as a piece of base information for land use and infrastructure planning in the study area?"

Usability is a quality attribute that assesses how easy the user interface (UI) is to use and this includes methods for improving ease-of-use during the design process (Nielsen 2012). An equally important quality attribute according to Nielsen (2012) is "utility" which refers to what users need. If utility is not satisfied then, usability matters little. Likewise, failure to address usability means wasting useful content. In this study, the factors considered in the resultant mapping combines usability as defined by ISO 9241-11 with use requirements and user requirements. First, it must be noted that the resultant mapping here follows from a co-production modelling process in which end-users participated. Thus, this partly satisfies content requirements related to the map purpose.

In the context of land use and infrastructure planning, the utility of the landslide hazard map have to do mainly with its functionality to guide spatial planning interventions in order to reduce disaster risks. The landslide hazard map's basic purpose is to inform its end-users the specific areas in their locality that are at risk of landslides in order for them to make sound spatial planning decisions. This is the use requirement. The use requirement is incorporated in the co-production modelling steps, wherein inputs of end-users have been taken into consideration. As informed participants in the co-production process, they already have an idea of the output, which defines the content and context of use in the output. Therefore, the resultant mapping explored how to optimize content through available devices in the study area.

Another consideration is the usage process of the actual users. Spatial planning interventions in this case refer to land use policies and infrastructure planning that minimize disaster risk. These are tasks assumed by specific personalities in the local government unit or municipality. In the study area, the execution of plans and spatial policies are handled by the municipal planning officers and the municipal disaster risk reduction officer. Before these plans are administered, the decision to implement needs the adoption of the legislative body, after these receive approval of municipal residents. This is the official procedure. The actual users are the municipal planning officers, the municipal disaster risk reduction officer, the legislative body and representative municipal residents. Here, the use and the legal process of using the landslide hazard map in spatial planning interventions have defined the actual map users. Outside of this legal usage process is the general usage of the public in the study area.

The process by which each actual user uses the map interface varies and this is influenced by their individual mapping exposure and group learning dynamics. The latter refers to instances during the adoption of proposed spatial planning interventions where map use is a group workshop. In the approval of land use policies and physical intervention plans, the legislative body inspects this as a group. Hence, evaluation methods take into consideration how maps are used individually and in groups.

3.6.2 Testing methods

This section responds to RQ4.2 "Based on the factors defined in RQ4.1, what testing method can measure the usability of the resultant map among users in the study area?"

Both remote moderated and remote unmoderated methods of usability testing are applied on a dynamic map produced from the selected model. The remote moderated method attempts to assess usability of grouped users by incorporating the dynamics of a group mapping activity, where map use is influenced by reactions of observer-participants. Since this is moderated, some real-time guidance is provided by the researcher. This was conducted as a videoconference with the FGD via MS Teams moderated by the researcher. During the online meeting, an html copy of the map was sent to one user who is assigned to explore the map. This activity was recorded, which captured the mapping activity of a first-time user. The videorecording also showed which and how sections of the panel were checked first, the sequence of panning and zooming, contribution of other participants, and the length of time that specific information is discovered and queried. A simple mapping activity was prepared for this meeting.

The remote unmoderated testing was done through instructions of a mapping exercise and questions with an html copy of the map that were emailed to the municipal disaster risk reduction officer and to the FGD participants after the online meeting. The objective for this assessment is to gather information if the target users and the local public can use the map to identify areas that are affected by landslides. Appendix I and J exhibit the answers of two respondents.

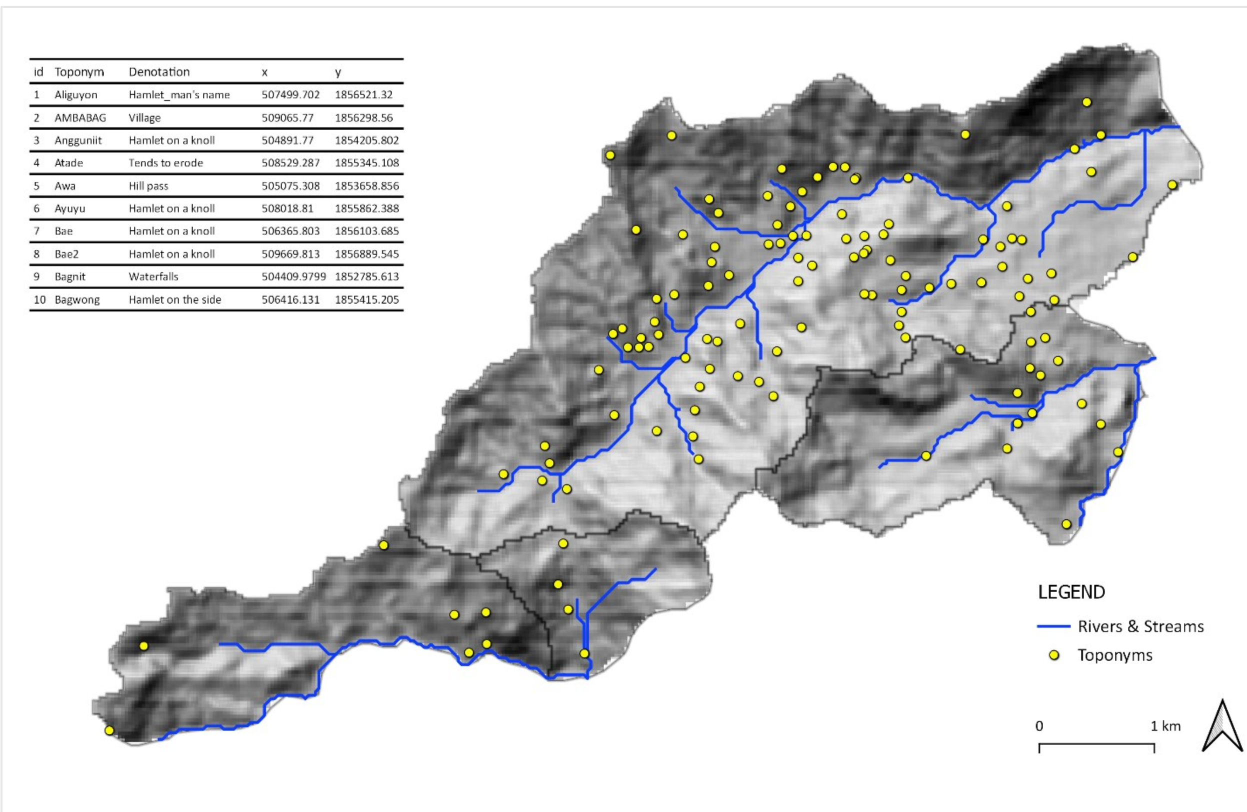
4. Modelling Results

This chapter reports findings from the implementation as described in Chapter 3. Sections 4.1 and 4.2 present the results of the toponym translation process as explanatory variables and as observed data. Section 4.3 presents the statistical modelling results. A key point to note in the results of toponym translation is that the findings of each step partly shaped the procedure and the methods adopted which were described in Chapter 3.

4.1 TOPONYMIC CHARACTERIZATION

4.1.1 Toponym collection and meanings

A total of 123 toponyms was collected from the study area (Figure 17). The southeast half-basin is not a complete list as indicated by more information on toponym points provided by the local community through social media at the end of this research. Some toponyms that connote ownership, such as those featuring land use, are not included. These often relate to human-made boundaries but with no known meanings that describe the landscape. Within the study area, most of these are in the rice terraces. The list of collected toponyms and their longer descriptions are provided in the submitted supplementary material.



id	Toponym	Denotation	x	y
1	Aliguyon	Hamlet_man's name	507499.702	1856521.32
2	AMBABAG	Village	509065.77	1856298.56
3	Angguniit	Hamlet on a knoll	504891.77	1854205.802
4	Atade	Tends to erode	508529.287	1855345.108
5	Awa	Hill pass	505075.308	1853658.856
6	Ayuyu	Hamlet on a knoll	508018.81	1855862.388
7	Bae	Hamlet on a knoll	506365.803	1856103.685
8	Bae2	Hamlet on a knoll	509669.813	1856889.545
9	Bagnit	Waterfalls	504409.9799	1852785.613
10	Bagwong	Hamlet on the side	506416.131	1855415.205

Figure 17. Collected toponyms

4.1.2 Toponymic Characterization

Figure 18 presents 74 toponyms that relate with landslide causal factors elicited through the summary of the survey checklist. Type 1 toponyms denote landscape properties that have local connotations. Eight (8) of these count as landslide occurrence indicated by “T1-C” and “T1-D”.

T1-A TYPE 1

id	Toponym	Denotation	CF
5	Awa	Hill pass	1
13	Bakung	Curve	3
21	Botbotan	Suicide place	3
41	Halipan	Spur	3
47	Imbuliklik	Rock	3
50	JULONGAN	Waterway	3
53	Kahingyan	Slope curve	2
57	Kurba	Spur curve	2
62	Likkod	Slope curve	2
64	Liyang	Cave and water	1
75	Mumbungug	Highly audible	2
86	Nunbukul	Distended	2
87	Nuntuul	Distended	3
98	Pico	Lower slope	3
107	Upla	Spur	4
112	Yukko	Bend	2
117	Wingiyon	Sideway	1
118	Inluplup	Spring	2
119	Wingiyon	Pass - lower slope	1

T1-C TYPE 1 + TYPE 2

id	Toponym	Denotation	CF
31	Domang	Slope	2
32	Dopla	Scarp	3
38	Godde	Landslide	3
39	Gode	Landslide & deposits	3
122	Domang2	Slope	2

T1-D TYPE 1 + TYPE 2 + TYPE 3

id	Toponym	Denotation	CF
4	Atade	High mountain + erosion	3
79	Nabangkawan	Hollowed out	2
85	Napukliyan	Scraped	2

T1-B TYPE 1 + TYPE 3

id	Toponym	Denotation	CF
9	Bagnit	Waterfalls	3
11	Bahakal	Forest	3
16	Bayninan	Knoll hamlet	2
17	Bilid	Hill	3
19	Bilong	Spur hamlet	2
20	Bokyod	Spur hamlet	2
23	Bungubung	Headwater forest	3
26	Dapdap	Terraces	2
27	Daya	Headwater region	3
28	Bolal	Knoll	4
30	Dinilag	Hillside hamlet	3
33	Dinapugan	Spur hamlet	2
40	Gayumhod	Hillside hamlet	1
43	Gitiw	Hillside hamlet	2
44	Golo	Hillside hamlet	1
46	Imbintok	Vegetation	1
48	Huliaban	Spur hamlet	1
49	Indalmogan	Headwater region	4
51	Huyu	Village boundary	3
52	Kadibdib	Windy	2
54	Kappugan	Many hills	2
58	Huyung	Side hamlet	1
59	Lapiddik	Trees	2
60	Lapidik	Trees	2
65	Kabonglahan	Vegetation	3
70	Madannum	Watery soil	3
74	Longnga	Spur hamlet	1
76	Munkilong	Vegetation	3
77	Muyung	Woodland	3
78	Lungngut	Village on a spur	2
80	Nabuluk	Decayed	3
81	Naduntug	Hilly formation	4
83	Maningla	Abandoned village	2
84	Nalodan	Vegetation	2
91	Patkik	Headwater part	3
94	Pa-u	Vegetation	3
97	Patukan	Spur hamlet	3
99	Puloy	Mountain	4
103	Punduntugan	Hill hamlet	2
105	Tangil	Hamlet	3
108	Utu	Waterfalls	2
109	Wa'el	Brook	1
111	TUPLAC	Valley village	3
113	Baggabag	Terraces	2
114	Lobbongan	Flooded ground	1
116	Patuldug	River	1
120	Ambuwaya	Valley village	1
121	Tanibung	Valley village	1

CF = Landslide causal factor

Figure 18. Type 1 characterization of toponyms

The remaining 49 toponyms (Figure 26) show Type 3 toponyms not directly related to landslides. The traditional land use (LU) is classified into settlement and terraces or rice paddies. In general, settlement toponyms indicate their placement along a slope, knolls and hills, or mounds on the valley. Land cover (LC) is classified as either forest or vegetation. The maintenance of forest areas and vegetation is also part of the land management practices. These toponyms either indicate animal habitation or plant species in that area.

		TYPE 3							
id	Toponym	Denotation	LU	LC	id	Toponym	Denotation	LU	LC
1	Aliguyon	Hamlet_man's name	S	0	56	Kikaag	Forest	0	F
2	AMBABAG	Village	S	0	61	Layya	Vegetation	0	V
3	Angguniit	Knoll hamlet	S	0	63	Imbungyaw	Valley hamlet	S	0
6	Ayuyu	Knoll hamlet	S	0	66	Lacdag	Spur hamlet	S	0
7	Bae	Knoll hamlet	S	0	67	Luhong	Trees	0	F
8	Bae2	Knoll hamlet	S	0	68	Linda	Terraces	T	0
10	Bagwong	Side hamlet	S	0	69	Mabbalat	Vegetation	0	V
12	Bahawit	Raised hamlet	S	0	71	Malpao	Vegetation	0	V
14	Balikongkong	Vegetation	0	V	72	Maluhong	Trees	0	F
15	Banaguy	Knoll hamlet	S	0	73	Lohob	Knoll hamlet	S	0
18	Biday	Knoll hamlet	S	0	82	NAGACADAN	Vegetation	S	V
20	Bokyod	Spur hamlet	S	0	88	Olagon	Vegetation	0	V
22	Bumalatuk	Vegetation	0	V	89	Onnop	Vegetation	0	V
24	Bolog	Knoll hamlet	S	0	90	Panniki	Forest	0	F
25	Buyakawan	Vegetation	0	V	92	Motwaon	Terraces	T	0
29	Bunnagan	Hamlet side	S	0	93	NABAGTU	Wide and high plateau	0	0
34	Dotal	Level	S	0	95	Naggawwa	Valley middle	0	0
35	Dotal2	Level	S	0	96	Patugong	Knoll hamlet	S	0
36	Dugung	Village edge hamlet	S	0	100	PINDONGAN	Valley village	S	0
37	Galuwago	Hamlet on a knoll	S	0	101	Pud-awan	Knoll hamlet	S	0
42	Halong	Trees	0	F	102	Pulitang	Knoll hamlet	S	0
45	Ihak	Terraces	T	0	104	Tukyudan	Vegetation	0	V
55	Kibadut	Terraces	T	0	106	Tikma	Knoll hamlet	S	0
					110	Tugawi	Village_abandoned	0	F
					115	Ungbul	Valley hamlet	S	0
					123	Dumanayan	Valley meeting place	S	0

LU = Traditional land use
LC = Land Cover
S = Settlement
F = Forest cover, natural and traditional land management
V = Vegetation
T = Terraces/Rice paddies

Figure 19. Type 3 toponyms

4.2 REGIONALIZATION

4.2.1 Optimizing slope unit partition

The merging of SUs to larger SUs through an iterative elimination process stopped at 6125.4 square meters. To preserve all information from toponyms, any SU should not be larger than this area. This is based on the closest pair of geomorphology-descriptive toponym points, which were 119.5 meters apart. One toponym means "depressed that tends to be flooded", and the other refers to a "water-supplied raised (implied) platform". The final set has 1112 SUs.

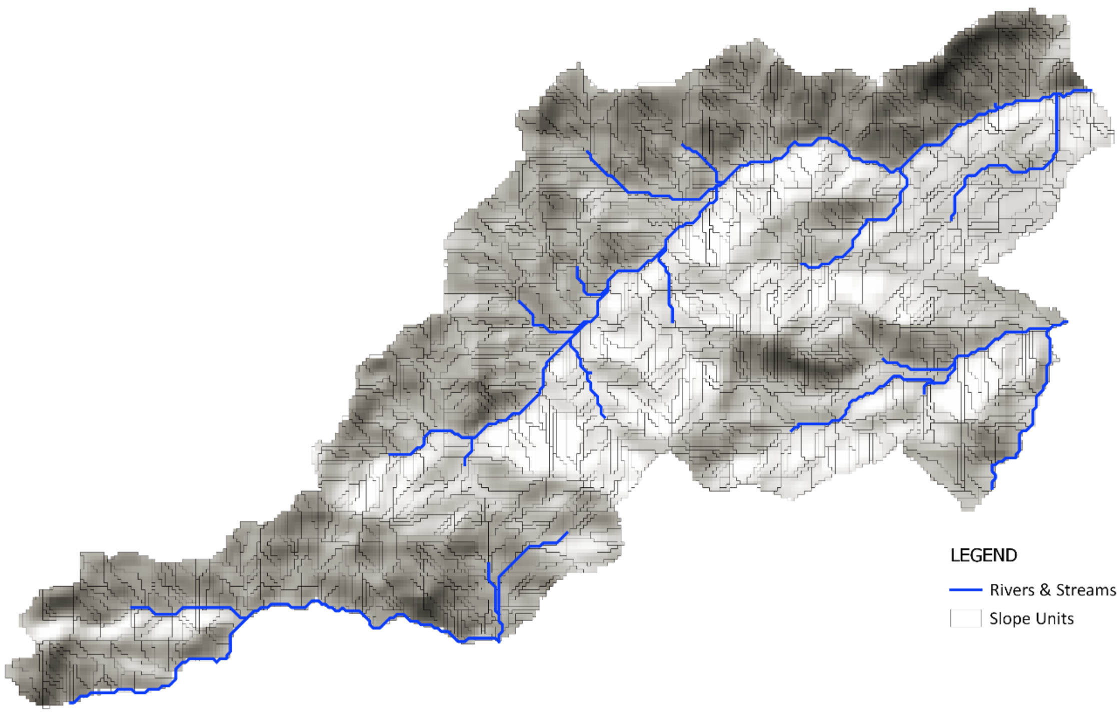


Figure 20. Slope unit partitions

4.2.2 Connecting toponym with SUs

In Figure 21, image A shows SUs connected with landslide causal factors based only on their direct associations with toponyms. The seven (7) causal factors enumerated in Table 7 vary in terms of number and combinations over the area. Image B factors in observed human activities reflected in land use that differs from toponym-described land use or land cover.

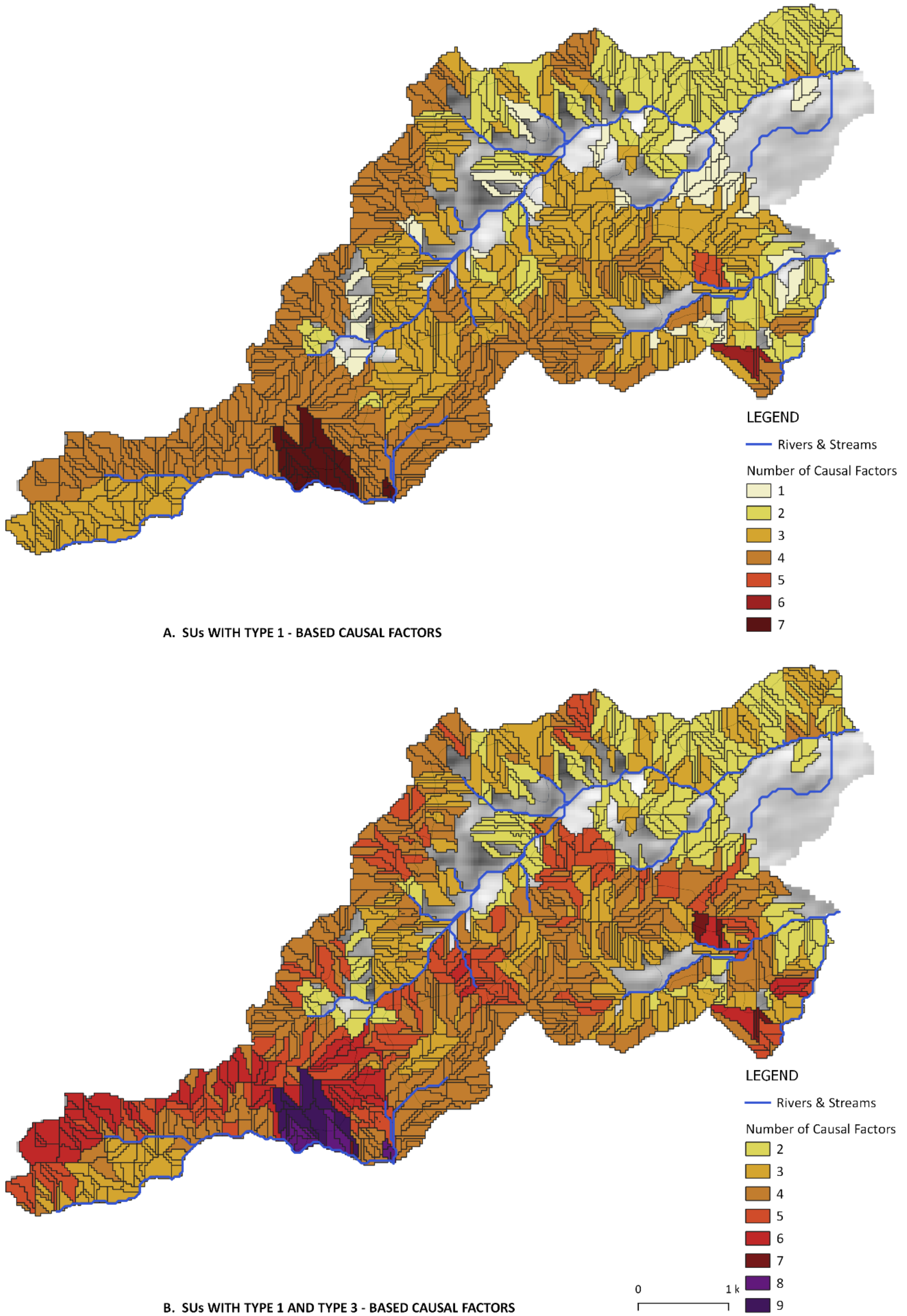


Figure 21. Presence of Toponym-based Landslide Causal Factors

4.2.3 Construction of toponymic regions

Fifty-eight (58) toponymic geospatial regions were formed from the grouping of SUs with the same number and combination of predictors. Figure 29 indicates their placement and the areas that underwent land use change. The areas without predictors are on the valley parts used for settlement and terraced wet cultivation. Figure 22 shows the basic coverage of regions,"tg" and areas that have undergone change due to human activities. Table 9 shows the combination of predictors.

Predictors	Region
BuVc	23
ES	19
ESBuVc	50
ESL	10
ESLBu	41
ESLBuVc	43
ESLVc	42
ESM	7
ESMA	8
ESMABu	27
ESMABu	38
ESMABuVc	40
ESMALBu	6
ESMALBuVc	59
ESMALVc	58
ESMAVc	39
ESMBu	35
ESMBuVc	37
ESMVc	36
ESplCBuVc	9
ESplCL	5
ESplCLBu	33

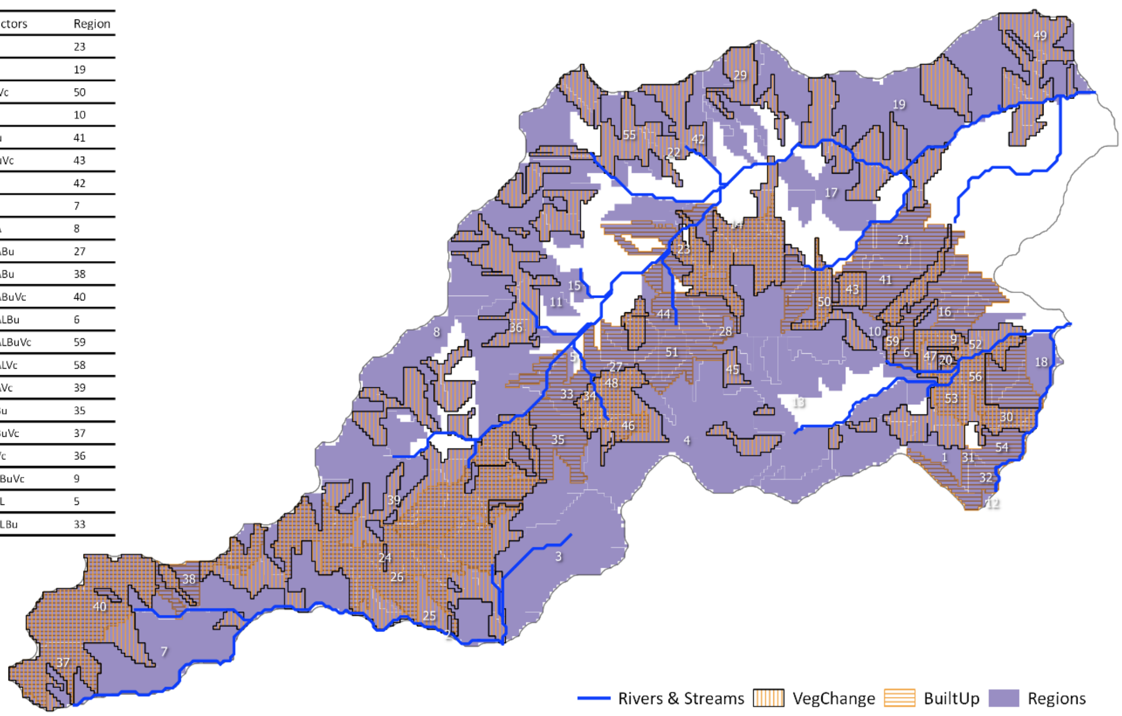


Figure 22. Toponymic regions

Table 9. Combination of predictors

Predictors	tg	SUwLS	SUnLS	nPredictors	Predictors	tg	SUwLS	SUnLS	nPredictors
non-assigned	0	2	192	2	ESplCMBuVc	30	3	1	6
ESplCprCML	1	1	0	6	ESplCprCMLBu	31	1	2	7
ESplCprCMAL	2	1	0	7	SplCprClBu	32	1	4	5
ESplCprC	3	0	33	4	ESplCLBu	33	2	3	5
ESplCM	4	2	69	4	ESplCLBuVc	34	1	0	6
ESplCL	5	0	2	4	ESMBu	35	13	25	4
ESMALBu	6	1	0	6	ESMVc	36	11	16	4
ESM	7	13	91	3	ESMBuVc	37	28	9	5
ESMA	8	17	87	4	ESMABu	38	4	1	5
ESplCBuVc	9	1	0	5	ESMAVc	39	21	25	5
ESL	10	8	0	3	ESMABuVc	40	39	7	6
ESprC	11	0	1	3	ESLBu	41	7	0	4
SplCprCl	12	0	1	4	ESLVc	42	3	0	4
SplCM	13	4	39	3	ESLBuVc	43	4	0	5
SplCprCBuVc	14	2	1	5	SplCMBu	44	3	5	4
SplCprC	15	0	2	3	SplCMVc	45	3	7	4
SMLBu	16	0	2	4	SplCMBuVc	46	5	5	5
SM	17	3	27	2	SAVc	47	0	1	3
SplC	18	2	10	2	SMABuVc	48	1	0	5
ES	19	4	46	2	ESVc	49	6	27	3
SABu	20	1	0	3	ESBuVc	50	1	0	4
SBu	21	0	26	2	SplCBu	51	0	6	3
SVc	22	2	33	2	SplCVc	52	0	2	3
BuVc	23	2	8	2	SplCBuVc	53	0	1	4
ESplCprCMALBu	24	1	0	8	SMBu	54	0	13	3
ESplCprCMALVc	25	3	8	8	SMVc	55	3	12	3
ESplCprCMALBuVc	26	9	0	9	SMBuVc	56	2	2	4
SMABu	27	1	0	4	ESMALVc	57	1	0	6
ESplCMBu	28	0	1	5	ESMALBuVc	58	2	0	7
ESplCMVc	29	5	9	5					

tg – toponymic geospatial region
nPredictors – number of predictors
A – slope aspect
E – elevation
L – lithology
M – slope moisture
plC – planar curvature
SUwLS – Slope units with landslides
SUnLS – Slope units without landslides
prC – profile curvature
S – slope steepness
Bu – built up
Vc – vegetation change

4.2.4 Landslide Inventory

Landslide inventory added current observed data and Type 2 toponym “count” data which totals to 627 landslide events. Of those, 128 events were derived from Type 2 toponyms, expressed as numerical values of the SUs that the Type 2 toponym covers (Figure 23).

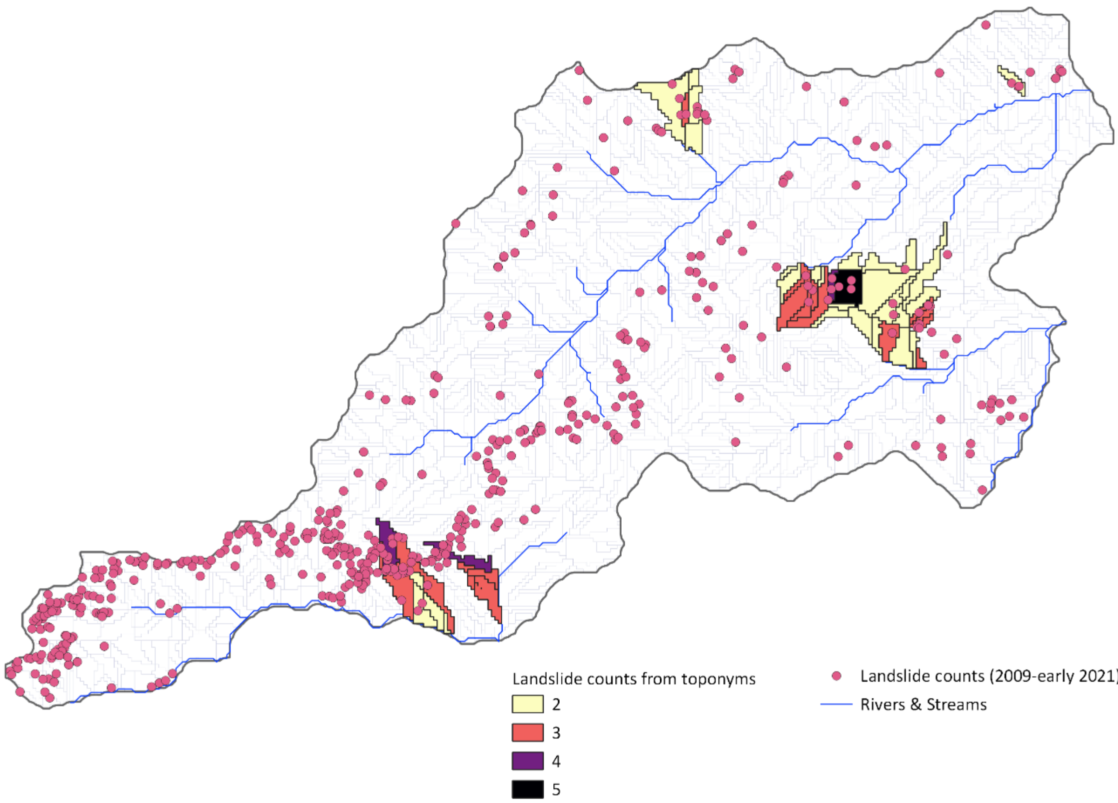


Figure 23. Landslide inventory from toponyms and observed data

4.3 MODEL SELECTION

4.3.1 Selection of Significant Variables

The 58 variables were reduced to 43 using AIC and multi-collinearity diagnosis. Using 95% credible interval, 35 (Table10) are significant with *iid* random effects (Model 1). Model 0 which only accounts for fixed effects has 40 significant variables. However it showed a better DIC value with the set of variables of Model 1. Appendix E and F show which variables were found not significant.

Table 10. Significant variables for Model 0 and Model 1

		mean	**sd**	**0.025quant**	**0.5quant**	**0.975quant**
	(Intercept)	-12.8244	0.172455	-13.1767	-12.8193	-12.5007
1	tg1	0.438282	0.18527	0.067769	0.44036	0.797173
2	tg2	0.663641	0.169025	0.319316	0.66792	0.984536
3	tg6	0.464127	0.197194	0.062417	0.469119	0.838502
4	tg8	0.298461	0.068072	0.162518	0.299239	0.429967
5	tg9	0.618972	0.236632	0.136922	0.624962	1.068219
6	tg10	1.398434	0.141026	1.124503	1.397297	1.67878
7	tg14	0.507888	0.162585	0.171108	0.513896	0.810715
8	tg20	1.11224	0.458817	0.133974	1.139879	1.939834
9	tg24	0.658599	0.134252	0.392618	0.659207	0.921003
10	tg25	0.300938	0.065479	0.166604	0.30293	0.424058
11	tg26	0.477313	0.043223	0.394081	0.476736	0.563928
12	tg27	0.814252	0.344114	0.080551	0.834979	1.434955
13	tg28	0.513425	0.180451	0.132612	0.522692	0.842788
14	tg29	0.40607	0.096205	0.209019	0.408907	0.58721
15	tg30	0.43804	0.120238	0.190767	0.441877	0.663795
16	tg33	0.418766	0.148667	0.105875	0.426077	0.690924
17	tg34	0.510842	0.197195	0.109132	0.515834	0.885217
18	tg35	0.556451	0.077345	0.402745	0.55704	0.706743
19	tg36	0.518198	0.088467	0.340833	0.519419	0.688649
20	tg37	0.656179	0.051586	0.556721	0.655559	0.759233
21	tg38	0.673455	0.114663	0.444311	0.674705	0.895515
22	tg39	0.459883	0.05502	0.3518	0.459878	0.567881
23	tg40	0.594987	0.038909	0.520612	0.594284	0.673411
24	tg41	0.87958	0.115615	0.653287	0.879208	1.107692
25	tg42	0.963913	0.164929	0.638537	0.964206	1.287345
26	tg43	0.733578	0.113034	0.512313	0.733189	0.956749
27	tg44	0.53038	0.145369	0.231804	0.534913	0.803479
28	tg45	0.438338	0.150958	0.123339	0.444755	0.71708
29	tg46	0.554471	0.100405	0.350808	0.556708	0.745675
30	tg49	0.413555	0.136606	0.132049	0.418168	0.668819
31	tg50	0.781059	0.277899	0.225297	0.784177	1.319383
32	tg55	0.421127	0.185995	0.032736	0.429119	0.76426
33	tg56	0.622723	0.173084	0.269642	0.627298	0.950467
34	tg57	0.806518	0.18527	0.436003	0.808597	1.165409
35	tg58	0.529884	0.115569	0.300227	0.530661	0.755068

With CAR effects, tg19 and tg23 which were not significant with *iid* random effects showed significance. The final set has a total of 37 significant variables (Table11). Removing tg27 from Table 11 also resulted to another set of significant variables with CAR effects, but this showed a higher DIC value. Appendix G and H show the iterative selection of significant CAR variables.

Table 11. Significant variables from Model 2 and 3

		mean	sd	0.025quant	0.5quant	0.975quant
	(Intercept)	-13.014	0.259108	-13.5406	-13.0082	-12.521
1	tg1	0.55525	0.186219	0.184846	0.55666	0.917631
2	tg2	0.643804	0.190871	0.259082	0.647089	1.010306
3	tg6	0.510208	0.182175	0.135182	0.516147	0.851973
4	tg8	0.310826	0.091945	0.126726	0.311996	0.48831
5	tg9	0.710903	0.199357	0.295641	0.719074	1.080183
6	tg10	1.42814	0.156904	1.123256	1.426936	1.73963
7	tg14	0.579378	0.163089	0.242161	0.585213	0.883572
8	tg19	0.624624	0.260217	0.110707	0.624993	1.136336
9	tg20	1.357521	0.509424	0.290325	1.380565	2.296482
10	tg23	0.756381	0.343218	0.025274	0.776827	1.374907
11	tg24	0.578025	0.136241	0.307817	0.578788	0.843821
12	tg25	0.297079	0.090034	0.114058	0.299235	0.46808
13	tg26	0.452747	0.056349	0.341912	0.452733	0.563568
14	tg27	0.789831	0.355772	0.034562	0.809932	1.434744
15	tg28	0.506715	0.185389	0.115065	0.516419	0.844477
16	tg29	0.483292	0.106354	0.268821	0.485198	0.687076
17	tg30	0.445951	0.131696	0.174798	0.450301	0.692475
18	tg33	0.444261	0.166763	0.099784	0.45014	0.755487
19	tg34	0.477725	0.181848	0.103066	0.483776	0.818449
20	tg35	0.561956	0.086765	0.390695	0.562224	0.731569
21	tg36	0.494336	0.097235	0.299777	0.495576	0.681871
22	tg37	0.57825	0.061391	0.459113	0.577767	0.700014
23	tg38	0.558086	0.140236	0.275874	0.560401	0.8273
24	tg39	0.470576	0.070516	0.330991	0.470883	0.608324
25	tg40	0.522887	0.058859	0.405851	0.523351	0.637357
26	tg41	0.91168	0.132899	0.652819	0.910855	1.17494
27	tg42	0.937493	0.15626	0.626998	0.938599	1.241628
28	tg43	0.80797	0.113202	0.587279	0.807314	1.0321
29	tg44	0.596021	0.148803	0.293297	0.599695	0.8783
30	tg45	0.532524	0.147179	0.227923	0.537908	0.806439
31	tg46	0.595903	0.108636	0.378265	0.597327	0.805501
32	tg49	0.571232	0.176251	0.224547	0.571016	0.919034
33	tg50	0.56243	0.209665	0.130844	0.569295	0.955424
34	tg55	0.522274	0.187367	0.137649	0.52811	0.873902
35	tg56	0.722098	0.176031	0.365435	0.725862	1.057796
36	tg57	0.895617	0.196377	0.505061	0.897062	1.277944
37	tg58	0.595931	0.113179	0.370788	0.596799	0.816093

4.3.2 Goodness of fit

The DIC values and effective number of parameters for the three competing models are shown in Table 12. Model 3 showed the least DIC value; however, this formulation only slightly improves model fit over the other CAR model. It differs from Model 2 by less than 2 DIC. Between the CAR models, Model 3 is less complex, indicated by the effective number of parameters. This will therefore be selected as the best model. Notable too is how the random effects are accounted for between Model 1 and the CAR models. The CAR models consider more toponymic priors than the iid model and show significant differences in the DIC value.

Table 12. DIC values

Models	Effective number of parameters	DIC
Model 0 (Fixed effects)	35.82	1687.128
Model 1 (Fixed effects + iid random effects)	212.03	1481.020
Model 2 (Fixed effects + proper CAR effects)	194.85	**1429.984**
Model 3 (Fixed effects + iCAR and iid effects)	**185.85**	1431.157

Figure 24 shows a summary of the posterior distribution of all the estimated coefficients that appeared to be significant for Model 3. The estimated coefficients for Model 0 and Model 1 are also plotted to highlight the difference. By absolute value, the *mean tg10,tg20,tg41,tg42*, and *tg57* variables gave the strongest contribution to the models.

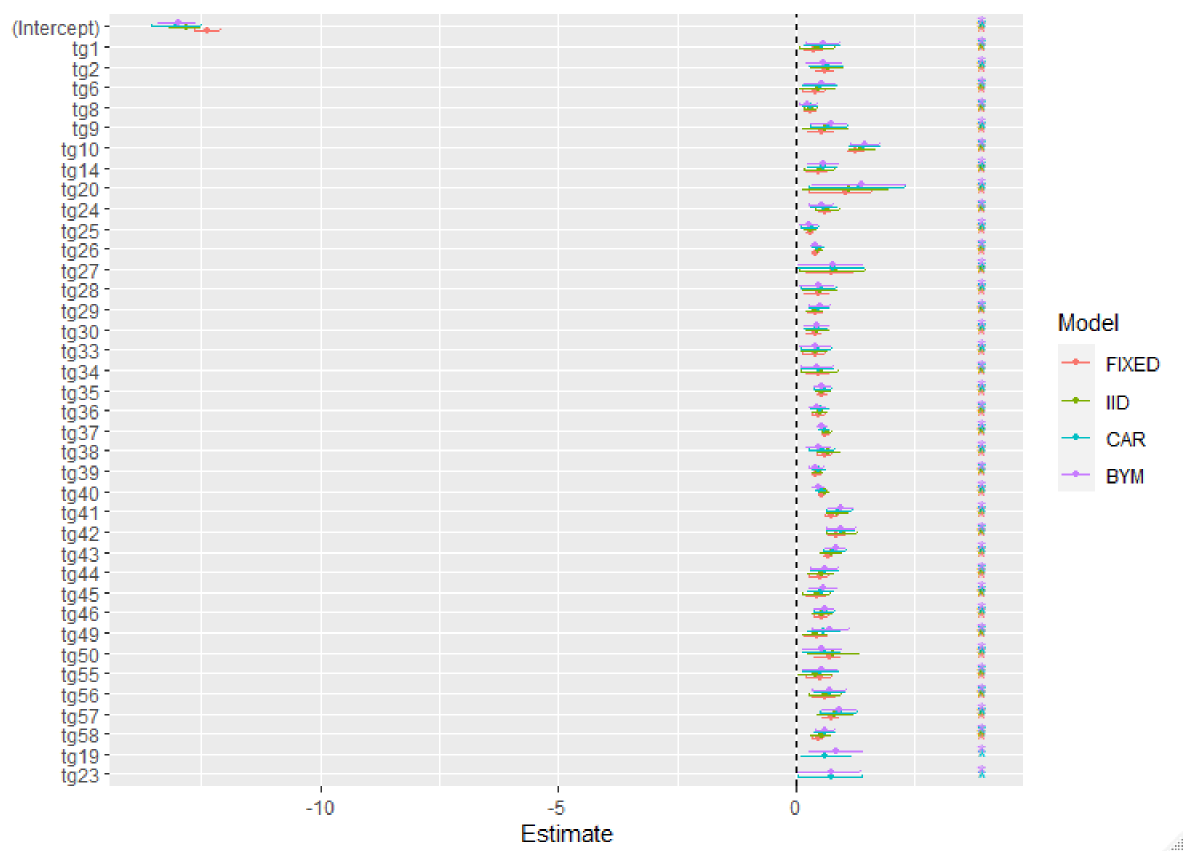


Figure 24. Estimated intercept and coefficient effects

Figure 25 shows the fixed and mixed effects of the four models. At this scale a small difference between the fixed and the mixed models can be seen.

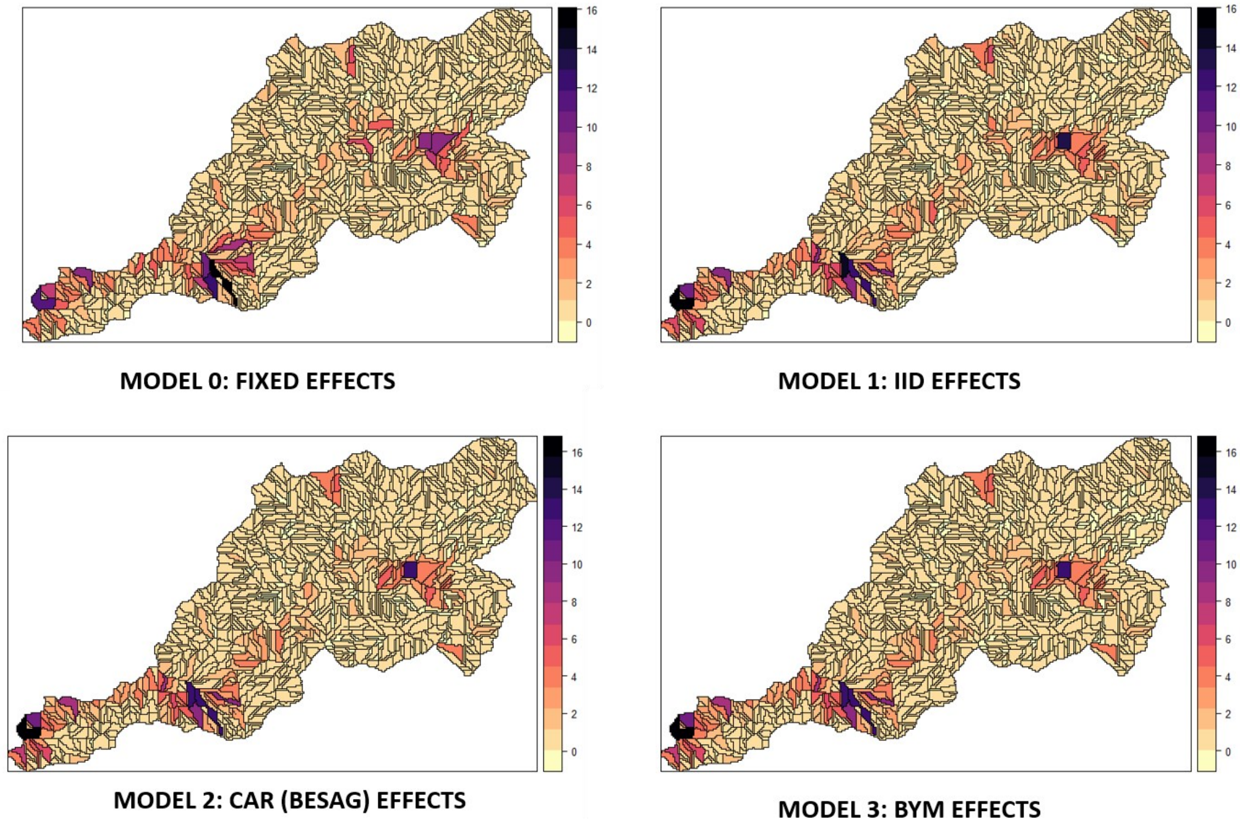


Figure 25. Fixed and mixed effects

Figure 26 shows the plotted posterior mean of Model 3 presented as "Intensity" values. The upper and lower limits are presented in Table 13. This is the resultant map to be rendered for users in the study area.

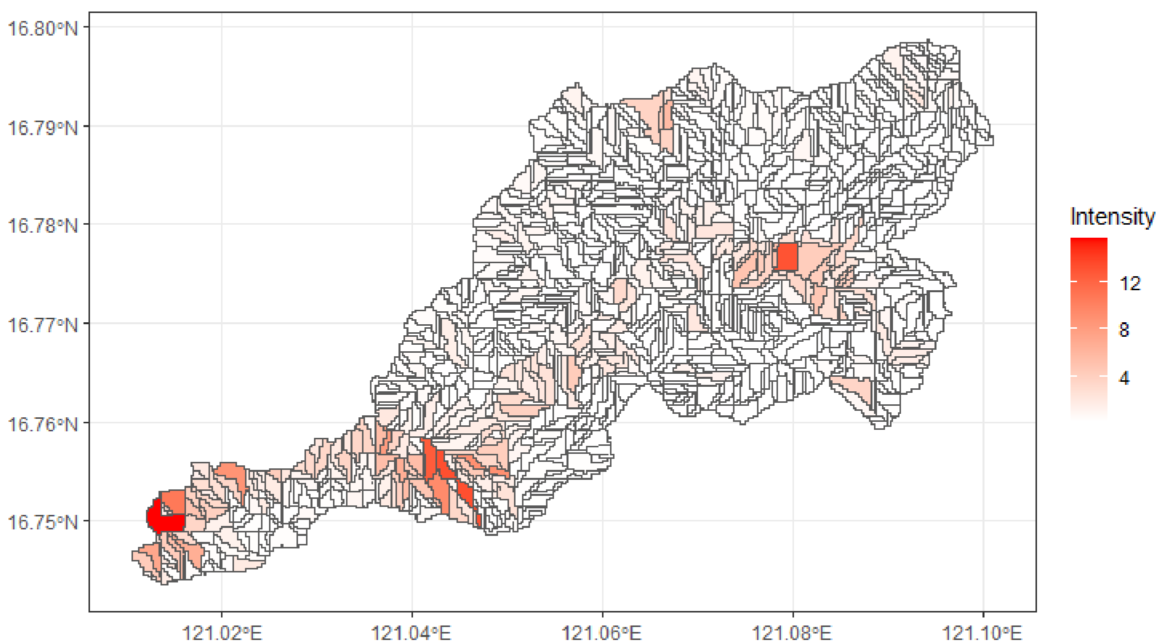


Figure 26. Model 3 posterior mean as Intensity values

Table 13. Upper, lower and relative limits of Model 3

RL	LL	UL
Min. : 0.007518	Min. :0.000325	Min. : 0.03133
1st Qu.: 0.044070	1st Qu.:0.004508	1st Qu.: 0.17106
Median : 0.104101	Median :0.011820	Median : 0.38181
Mean : 0.605143	Mean :0.189557	Mean : 1.42929
3rd Qu.: 0.416444	3rd Qu.:0.062170	3rd Qu.: 1.34442
Max. :15.714723	Max. :9.468316	Max. :23.56907

4.4 MODEL SUMMARY

Model 3 was also rendered as a dynamic map. This is available in the supplementary material provided.

From preprocessing to model rendering, a number of softwares and packages were used. Table 14 shows a summary of the modelling activities and citations. Computers that have standard operating systems for which QGIS and R-INLA are available can run the models. The recommended memory for each is 1 Gb.

Table 14. Citation and software used

Software/Package	Purpose	Citation/url
QGIS, SAGA, GRASS	Pre-processing	https://www.qgis.org
R-INLA	Analysis, Modelling	Bivand, Gómez-Rubio, and Rue 2015
SPDE	Analysis	Bakka et al. 2018
R-MASS	Selection of variables	Venables and Ripley 2002
R-CAR	Selection of variables/VIF test	Fox and Weisberg 2019
R-AER	Modelling, Over-dispersion test	Kleiber and Zeileis 2008
R-Leaflet	Dynamic Map	Graul 2016
Tidyverse/ggplot2	Plotting	Wickham 2016

4.6 USAGE TEST

This section partly responds to RQ4.3 “What features in the landslide hazard map needs improvement to make it more usable?”.

From the unsupervised exercise sent to eight actual users and stakeholders, only two responded, but with very similar positive comments. One is a municipal planner and the other an environmental planner. Both are familiar with mapping and the purpose of the landslide hazard map. The non-response of the others is taken here as a response in itself that needs follow-up. Two participants joined the supervised exercise where one was assigned as the map navigator. This user is not a professional mapper but often uses the Google mapping platform. The first activities observed were panning and zooming, searching for river lines and roads, aided by comments from the other participant. Notable is the length of time spent searching for basic landmarks that were not apparent in the initial moments of checking the map. Instructions of the researcher-facilitator aided awareness of the base map to use. The preferred base map is the ESRI world imagery; however, the data is available to a limited resolution in the study area. The dynamic map used in this exercise was from initial modelling outputs where one participant observed the negative "intensity" values of landslides. This was an input error that was corrected in the final map. The exercise also led to some corrections in the toponym properties that were adopted by some slope units.

Features to improve in the resultant mapping is obtained from observations in the video recorded user test and the answers in the survey questions. One suggestion is to have more satellite imagery choices. The recorded user test also revealed that the colour transparency of the landslide map layer needs adjustment to reveal features of the base map used.

5. Discussion

5.1 TOPONYMIC CHARACTERIZATION

5.1.1 Toponym collection and meanings

A significant part of the toponyms collected was previously mapped by the researcher before the research started through direct in-person FGDs officially organized for each administrative unit, which provided a good base for discussion in social media where this was presented for verification. If there were no initial map to begin with, the information would have been incomplete given the limited time to map this under pandemic constraints. However, the advantage of social media cannot be ignored. In this case, it has enriched the data with narratives from residents of the study area who are knowledgeable about the topic but are not available in person or would not have been invited as official representatives in FGD workshops on site.

5.1.2 Toponymic Characterization

The method developed here focused on toponym relation with landslides that requires a basic understanding of the elements (landslide inventory and causal factors) in landslide susceptibility modelling. It also requires a basic understanding of the local landscape pattern and the language. Toponym characterization is a process of deciphering its meaning and translating it into a quantifiable form that involves several decision-making steps where local knowledge is paramount. Thus, more prolonged engagement with the locals would have led to richer and more contextual information on the semantics. During the latter part of the research, new inputs from the locals that relate the abundance of vegetation to sun exposure or slope aspect (a landslide causal factor) were not incorporated in the computation because there was not enough time for a complete review of vegetation properties. From cross-checking with the slope aspect derived from DEM, there is a correlation between special vegetation-derived toponyms and slope aspect and landslide occurrence.

While this information is covered in the model, it is not exhaustive, which means that there may be other vegetation-derived toponyms that correspond with slope aspect that were not included. Exhaustive extraction of information from vegetation-derived toponyms may have to include experts in botany. Another toponym expressed information that may need consideration is how wind direction can affect landslide occurrence. Literature on the correlation of wind direction and landslide occurrence is scarce; hence this was not factored in. Although these predictors were not defined and therefore not counted, the correlating toponymic variables cover unexplained effects. The model is also designed to be iterative, where usage of the mapping product opens opportunities to enrich the basic information used.

Characterization also identified landslide causal factors to consider in the locality. Toponyms only identify their presence in general terms but provide leads for further studies. Lithology, for instance, is only associated with descriptions of loose rocks in specific areas that provide information for geologic mapping.

5.2 REGIONALIZATION

5.2.1 Optimizing slope units

Optimizing slope units using toponym meanings was a preliminary step to establish the minimum area of slope units that do not reduce the information content of toponyms. The attention was on identifying and ensuring that the closest pair of toponyms with distinct morphological properties are in separate slope

units. Because it is based on toponym semantics, it is subject to other interpretations, which introduces subjectivity.
In this case, the determinants are toponyms describing geomorphology, which define discrete land features. This could be taken as a simplification of an optimization process but offers a simple method of determining when information may be lost from a model element.

The optimized partition of SUs also removed zeros within finer partitions. Excessive zeros would have required either a different distribution or models that specifically address zero inflation. Fewer and larger-sized slope units, however produced coarser rendering in the final mapping. As realized later in the selected model, the gradation of values from a "hot spot" (very high intensity) to surrounding slope units is based on this granularity, where the intensity values depend on the nearness of neighbours, regardless of the distance length.

5.2.2 Connecting toponym with SUs

The imaginary boundary between toponyms which served as a guide to determine which slope units (SUs) should connect to specific toponyms, employed Voronoi polygons from a QGIS geometry tool operation. Ideally, this imaginary line should have been set by the study area residents who can quickly sketch where a toponym description ends or which areas cover the line where toponyms meet. However, the purpose of this model is not to find discrete boundaries of toponyms. It is instead designed to identify polygons or SUs that contain toponym properties through points or possible boundaries. The model's proposed set of rules that apply to slope units containing these boundaries works to combine toponym properties in these areas rather than making distinctions between toponyms. In this model, SUs, where toponyms meet, provide information on all possible effects of present toponyms. Any incorrect toponym placement is checked and updated when users of the resultant map offer better approximations of toponyms that each SU adopts.

An SU is a half-basin that may cover the line where a landslide related toponym meets a settlement toponym. Typically, this is from the middle slope of a hill down to the valley part. To predict the downward path of a landslide and the distance that it can reach on the inhabited valley parts, the SUs that capture these toponym combinations provide rich information of both probabilities and areas at risk.

5.2.3 Construction of toponymic regions

The toponymic regions represent different composites of predictors, which assume categorical values or the maximum effect of predictors. The method also assumes that the effect of the composite of predictors is constant over the surface area. It takes a conservative view which is preferable in hazards modelling. However, the predictors present in toponymic regions are based on interpretations of toponym meanings that may miss other variables. Regions, however, cover this because they already represent the presence of predictors regardless of the number of predictors.

The application of toponymic regions in modelling filters out areas where predictors are assumed to be absent. In effect, this eliminates noise which could happen when predictors cover areas that they do not describe. In traditional landslide susceptibility models, predictors such as elevation and slope steepness are treated as continuous variables covering the entire area, possibly introducing noise when applied in the study area. For instance, slope steepness would have included active rice terraces where landslides do not occur.

An option that was not studied is whether it makes a difference if these predictors assume their statistical units for these regions. For example, elevation as a predictor would instead assume a standardized value from a range of 500 to 1500 masl rather than a categorical value of 1. A comparative study would reveal if this approach is better than the proposed model.

5.2.4 Landslide Inventory

This inventory focuses on flows defined by Varnes (1978), which are precipitation-triggered events and frequent occurrences in the study area. From an 11-year count, 499 events in the study area were observed compared to the events recalled and toponym-deducted information that dates back to about 400 years. It is not clear, too, if all these toponym-based events are triggered by precipitation. The 11-year observation shows frequent landslide events that may differ from Type 2 toponyms. Type 2 toponyms describe landslide events or parts left by a landslide. They are named as such because they are indelible. Based on the local community's recollection of the last time that landslides occurred in areas covered by Type 2 toponym, the erosion of *Atade* was "powerful", as described by eyewitnesses with reference to the details of what they were doing that day. There are considerable gaps in the inventory, but these records of events carry weight in inferring probabilities of recurrence for these non-frequent, episodic events. More time dedicated to research would have allowed exploration in this direction.

5.3 MODELLING

5.3.1 Goodness of fit

Model 3 has the best value for goodness-of-fit compared to all considered models. Parsimony was the criterion that prevailed in the selection of the best model. In this case, it is the model where random effects were accounted for by the convolution of unstructured and structured effects. It is the best explanation at this point. In general, the use of toponymic priors is more compatible with the CAR models.Further exploration of other models may better represent the problem. Also, information criteria aside from DIC may be considered. An expansion of the study area within the same linguistic group can confirm the strong correlation of some variables. This may yield information on a list of toponymic variables that can be used in heuristic landslide assessments in similar areas where data are hard to obtain.

The model suits this study area where landslides are not apparent on the slope units adjacent to rivers. However, its application in other areas must be checked because the smoothening of random effects does not distinguish terrain and assumes as if the entire area is a flat surface. In areas where landslide events are observed to be intense beside rivers, the smoothening could extend to the slope units across the river. The range at which autocorrelation fades in this space needs to be checked by prior knowledge. Local knowledge input is essential.

5.3.2 Toponymic priors

Bayesian modelling does not limit how prior beliefs are incorporated, nor does it discriminate against any belief, including biases. In this proposed model, the bias on using toponyms as explanatory variables is laid out for scrutiny, confronted with evidence and known landslide hazard predictors. The use of toponymic variables is based on the prior belief that they have value. On the opposite end of this is a prior belief that they have zero value. For as long as both beliefs acknowledge mounting evidence, both will eventually converge on the right answer. The difference lies in how long it would take for both to arrive at that point, which is crucial in disaster risk preparedness, where evidence is in the form of a likely disastrous event.

Toponyms by themselves embody observed history, marking events and terrain properties. In this model, some are treated as observed data, and some are treated as explanatory variables. Those considered "observed" data from time periods that are not recent, add evidence otherwise unavailable. Those treated as explanatory variables are assessed using this "observed" data plus current observations. Toponyms that are characterized as both Type 1 (causal factors) and Type 2 (landslide count) are a convolution of both. In the model, it would appear as if it is a circular reference when the same toponym is used as the explanatory and dependent variable. However, the reference to the same toponym does not mean a reference to the same dimension. These dimensions of the toponym cannot be treated individually as different members.

5.3.3 A partial solution

Toponyms offer opportunities to explore a deeper insight of the problem because it provides a conversation interface between scientific and indigenous knowledge. In this model, the conversation continues in the mapping output, which serves as a platform to highlight issues about underlying risk drivers which were previously unacknowledged. This is only a partial solution to a wicked problem. But the information needed to define a local problem has improved and the modelling product is potentially useful.

The proposed model highlights the needed participation of different disciplines. Knowledge, expertise and skills are required to handle the conjunction of substantive and statistical areas. One researcher is not enough. This study was conducted from the perspective of a researcher who is an architect and an urban planner with a personal bias for indigenous knowledge. Some knowledge areas are outlined in this study and need finer research. Other perspectives and insights are also needed to improve the model.

5.6 USAGE TEST

If the usage test were done in person with potential users, an in-depth discussion of usefulness would have ensued. Tests from public users would also provide a broader perspective. These tests would have revealed more specific cartographic features to improve as well as information that needs highlighting. However, the current resultant mapping already offers a starting point for further improvements and refinements. The positive responses of two planners who are themselves government employees indicate that the hazards map can be used as a tool for local spatial planning and policy decisions. The timeline of this study does not bound usage testing.

6. Conclusion and Future Developments

5.1 CONCLUSION

In this study, indigenous toponyms were translated as input variables for landslide hazard modelling. The process explored the dimensions of toponyms, then designed their translation as inputs to a set of Bayesian models analyzed for their goodness-of-fit. The selected model underwent limited usage testing to assess its usefulness among actual users. The process involved matching information content from toponyms with landslide hazard assessment requirements. The procedure is iterative, where methods adopted drew from what is found in literature and from data gathered from the study area through a series of consultations and engagement of local community representatives.

The following summarizes answers to the research questions posed in Chapter 1.

RQ1.1 What are the considerations in representing a toponym as an input variable for landslide hazard modelling?

Relation to landslides is the primary consideration in translating a toponym as a variable for landslide hazard modelling. This governs the translation process where other considerations also arise, such as:

- the multiple dimensions of toponyms which characterize them as explanatory variables (causal factors) and observed variables (landslide events),
- the choice of the spatial unit that is common to both toponyms and landslide assessment, and
- the context of the causal factor in the study area, which specifies it as a local variable (e.g. planar curvature only refers to convex curvature, slope aspect refers to the south to southwest-facing faces).

Toponym interpretations are relative to the study area where prior knowledge of the researcher and the local community is also considered throughout this translation process.

RQ1.2 How is co-production employed in translating toponyms into model variables?

First, the purpose of the toponym collection was announced to a large discussion group in social media. Second, a structured (focus group) group dedicated to toponym collection and landslide hazard mapping was organized. Third, unstructured discussions in the large group were facilitated to enrich the existing data. The small group focused on the map and placement of toponyms during the toponym collection stage, whereas the large group served as a validating group, where toponym meanings were discussed at length. From these discussions, the relation between toponyms and landslides were clarified. The FGD played a part in the survey checklist, testing of the dynamic map, and updating incorrect connections between toponyms and SUs.

RQ2.1 Factoring in the answers to RQ1.1, which methods are suited to generate quantitative input variables for modelling from toponyms?

Two general methods are designed to suit the given toponymic information: Systematic characterization and regionalization. Toponyms undergo systematic characterization to match their properties with the elements of landslide hazard modelling and regionalization to translate these properties into georeferenced quantified variables. From their given meanings and associations, the toponyms were characterized into three types: Type 1) landslide causal factors; Type 2) landslide events, and; Type 3) land cover/use, where the difference with present-day land cover/use counts as a landslide causal factor. These causal factors are spatially referenced through geoprocessing operations by joining the respective toponym with slope unit (SU) partitions of the study area. An SU is a spatial unit that is deemed suited to both toponyms and landslides. The SUs carrying the same number and combination of causal factors are grouped into "toponymic geospatial regions' or *tg*, which are then used as explanatory variables. The value of a *tg* is the sum of causal factors (referred to in the region as the number of predictors) present. Toponyms characterized as Type 2 or landslide events are added as "counts" of the observed variable.

RQ2.1 What probability distribution captures the information provided by the data and toponyms?

The data and toponyms provided information on the count of landslide events for a given area. The Poisson probability distribution applies.

RQ3.1 What are the criteria to evaluate and select toponymic variables for landslide hazard modelling?

The selection of variables for modelling used a combination of Akaike Information Criterion (AIC) and multicollinearity diagnosis to select toponymic variables. The selection of significant variables used a 95% Credible Interval (CI).

RQ3.2 Based on which criteria and which process are models selected for their goodness-of-fit?

The Deviance Information Criterion (DIC) which measures fit and complexity was used in model selection. Among competing models, the model with the lowest DIC and the least number of parameters was selected.

RQ4.1 Which factors define the usefulness of the resultant mapping as a piece of base information for land use and infrastructure planning in the study area?

The factors that define the usefulness of the resultant mapping are:

- its utility or functionality to guide spatial planning interventions in order to reduce disaster risks,
- the usage process of the actual users (official) and public users in the study area, and
- usability or ease of use.

The process by which each actual user uses the map interface depends on their individual mapping exposure and group learning dynamics (e.g. during a legislative session to adopt a proposed spatial planning intervention).

RQ4.2 Based on the factors defined in RQ4.1, what testing method can measure the usefulness of the resultant map among users in the study area?

Remote moderated, and remote unmoderated testing methods can measure the usefulness of the resultant map. The remote moderated method attempts to assess the usability of grouped users by incorporating the dynamics of a map interface use, where reactions of observer-participants could influence map navigation. The remote unmoderated testing is more focused on the overall usefulness of the resultant mapping, where mapping activity is done by individual users guided by instructions and questions to extract information from the given map.

For this study, both methods tested the resultant dynamic map. The moderated testing method was conducted through a videoconference, where map navigation of a first-time user in the presence of another user was recorded for analysis.

RQ4.3 What features in the landslide hazard map needs improvement to make it more usable?

From the limited tests, satellite base maps that are available in the study area were preferred. In addition, findings from the video recording of map navigation showed that the transparency of the landslide probability map needs to be improved.
This will reveal details of the selected base map, which allows the user to find landmarks easily.

The conceptualization of indigenous toponym-based co-production landslide hazard modelling follows a process that begins and ends with the local community. The methodology is anchored on the principles of co-production and the modelling gaps described in Chapter 2, which emphasize usefulness. As a motivation and a guide in designing the modelling steps, usefulness also fulfils an ethical responsibility of this research. The inclusion of an evaluation of map usefulness for the target users is an extra step here as it is seen as good practice in local community engagement. In co-production modelling, the information elicited from the local community returns to them in a processed form that they can use. Underpinning the entire process is informed-decision making, which results from modelling steps that are reasonable and explicable. These become digestible to the local community, who contributed to its production.

There is a significant departure from the usual modelling process when indigenous toponyms are used to improve usefulness. This is indicated in the treatment of using landslide causal factors. The significance of each variable is not rendered in the resultant map. In the proposed process, toponyms are translated and used both as data and as explanatory variables after these are matched with landslide causal factors. In the resultant mapping, landslide causal factors (elevation, slope, curvature etc.) are still presented together with the history of landslide occurrence in specific areas. Current models do not present this. The model works so that regions of implementation in the final zonation map reflect landslide causal factors that need attention. This gives pieces of explicit information that aid decisions on land use and infrastructure planning.

5.2 FUTURE RESEARCH WORK

This study opened the subject of using indigenous toponyms in landslide hazard modelling, touching on areas that need further investigation and topics for further exploration, not limited to geospatial studies.

For modelling, these are the list of recommendations for further research:

- Further case studies within the region that share the same language family validate the significance of toponyms in landslide hazard assessment in the area. A sufficient number of samples and analysis is expected to fine-tune the identification of significant toponyms. The foreseen result is a "better" toponymic model that can be adopted in a heuristic approach to assessing landslides in areas in the region where landslide data are scarce.
- Estimate the temporal exceedance probability of landslides using the *terminus post ante quem* of toponyms

- Explore the combination of landslide causal factors that are specific to place and setting. In particular, investigate the effects of wind direction in forested high slopes and the exacerbating effect of land use within these areas.
- Develop a module that facilitates the joining of slope units with the nearest toponym that is constrained within half-basins
- Apply other models that account for the random effects that are not covered by toponyms
- Apply other methods of analysis using the current landslide data and toponym collection.
- Test toponymic approaches to optimize slope unit partitioning
- Explore application of toponyms to predict other hazards
- Compare the proposed mode with other model constructions using the standardized statistical values values of the same predictors

In the area of co-production, the following lists some recommendations:

- Design a platform that allows citizens to contribute local knowledge of toponym meanings and associations
- Explore toponymic information for modelling with other disciplines such as linguists, ethnobotanists, ethnopedologist, archaeologists, etc.

References

Abella, Enrique A. Castellanos, Cees J. Van Westen, Geoinformation Science, and Earth Observation. n.d. "Qualitative Landslide Hazard and Risk Assessment Using Multicriteria Analysis ; a Case Study from San Antonio Del Sur ,."

Acabado, Stephen (personal communication, February 5, 2021) commented on *terminus ante quem* of toponym.

Alcántara-Ayala, Irasema, and Anthony Oliver-Smith. 2019. "Early Warning Systems: Lost in Translation or Late by Definition? A FORIN Approach." *International Journal of Disaster Risk Science* 10(3):317–31.

Aleotti, P., and R. Chowdhury. 1999. "Landslide Hazard Assessment: Summary Review and New Perspectives." *Bulletin of Engineering Geology and the Environment* 58(1):21–44.

Alexander, Steven M., Jennifer F. Provencher, Dominique A. Henri, Jessica J. Taylor, and Steven J. Cooke. 2019. "Bridging Indigenous and Science-Based Knowledge in Coastal-Marine Research, Monitoring, and Management in Canada: A Systematic Map Protocol." *Environmental Evidence* 8(1):1–7.

Antolin, Francisco, and William Henry Scott. 1970. "Notices of the Pagan Igorots in 1789." *Asian Folklore Studies* 29(2):177.

Aon. 2020. *Weather, Climate & Catastrophe Insight*. Chicago,Illinois.

Assuncaõ, Renato, and Elias Krainski. 2009. "Neighborhood Dependence in Bayesian Spatial Models."

Ba, Qianqian, Yumin Chen, Susu Deng, Jiaxin Yang, and Huifang Li. 2018. "A Comparison of Slope Units and Grid Cells as Mapping Units for Landslide Susceptibility Assessment." *Earth Science Informatics* 11.

Bakka, Haakon, Håvard Rue, Geir-Arne Fuglstad, Andrea Riebler, David Bolin, Elias Krainski, Daniel Simpson, and Finn Lindgren. 2018. "Spatial Modelling with R-INLA: A Review."

Battiste, Marie. 2014. *Research Ethics for Protecting Indigenous Knowledge and Heritage: Institutional and Researcher Responsibilities. In Denzin, N. K., Lincoln, Y. S., & Smith, L. T. Handbook of Critical and Indigenous Methodologies (Pp. 497-510).* Thousand Oaks, CA: SAGE Publications, Inc.

Bélisle, Annie Claude, Hugo Asselin, Patrice Leblanc, and Sylvie Gauthier. 2018. "Local Knowledge in Ecological Modeling." *Ecology and Society* 23(2).

Besag, Julian. 1974. "Spatial Interaction and the Statistical Analysis of Lattice Systems Author (s): Julian Besag Source : Journal of the Royal Statistical Society . Series B (Methodological), Vol . 36 , No . 2 Published by : Blackwell Publishing for the Royal Statistical." *Journal of Royal Statistical Society* 36(2):192–236.

Bivand, Roger, Virgilio Gómez-Rubio, and Håvard Rue. 2015. "Spatial Data Analysis with R-INLA with Some Extensions." *Journal of Statistical Software, Articles* 63(20):1–31.

Bowles, Ella, Kia Marin, Pamela MacLeod, and Dylan J. Fraser. 2020. "A Three-Pronged Approach That Leans on Indigenous Knowledge for Northern Fish Monitoring and Conservation." *Evolutionary Applications* (September 2020):653–57.

Capra, Gian Franco, Antonio Ganga, Andrea Buondonno, Eleonora Grilli, Carla Gaviano, and Sergio Vacca. 2015. "Ethnopedology in the Study of Toponyms Connected to the Indigenous Knowledge on Soil Resource." *PLoS ONE* 10(3).

Capra, Gian Franco, Antonio Ganga, Peter Filzmoser, Carla Gaviano, and Sergio Vacca. 2016. "Combining Place Names and Scientific Knowledge on Soil Resources through an Integrated Ethnopedological Approach." *Catena* 142:89–101.

Cogos, Sarah, Marie Roué, and Samuel Roturier. 2017. "Sami Place Names and Maps: Transmitting Knowledge of a Cultural Landscape in Contemporary Contexts." *Arctic, Antarctic, and Alpine Research* 49(1).

Corominas, Jordi, and Mavroulli O. 2011. *Guidelines for Landslide Susceptibility, Hazard and Risk Zoning 2 SUMMARY.*

Cramb, SM, EW Duncan, PD Baade, and KL Mengersen. 2017. “Investigation of Bayesian Spatial Models.” (9):1–109.

Dall’Ò, Elisabetta. 2019. “Historicizing Vulnerability: Place-Names, Risk and Memory in the Mont Blanc Area.” *AIMS Geosciences* 5(3):493–508.

David-Chavez, Dominique M., and Michael C. Gavin. 2018. “A Global Assessment of Indigenous Community Engagement in Climate Research.” *Environmental Research Letters* 13(12).

Dikshit, Abhirup, Raju Sarkar, Biswajeet Pradhan, Ratiranjan Jena, Dowchu Drukpa, and Abdullah M. Alamri. 2020. “Temporal Probability Assessment and Its Use in Landslide Susceptibility Mapping for Eastern Bhutan.” *Water (Switzerland)* 12(1).

Diver, Sibyl. 2017. “Negotiating Indigenous Knowledge at the Science Policy-Interface: Insights from the Xáxli’p Community Forest.” *Environmental Science and Policy* 73:1–11.

Domènech, Guillem, Massimiliano Alvioli, and Jordi Corominas. 2020. “Preparing First-Time Slope Failures Hazard Maps: From Pixel-Based to Slope Unit-Based.” *Landslides* 17(2):249–65.

Van Den Eeckhaut, M., P. Reichenbach, F. Guzzetti, M. Rossi, and J. Poesen. 2009. “Combined Landslide Inventory and Susceptibility Assessment Based on Different Mapping Units: An Example from the Flemish Ardennes, Belgium.” *Natural Hazards and Earth System Sciences* 9(2):507–21.

van Elzakker, Corné P. J. M., and Kristien Ooms. 2018. “Understanding Map Uses and Users.” Pp. 55–67 in *A. J. Kent, & P. Vujakovic (Eds.)The Routledge Handbook of Mapping and Cartography*. Routledge.

Erener, Arzu, and Sebnem Duzgun. 2011. “Landslide Susceptibility Assessment: What Are the Effects of Mapping Unit and Mapping Method?” *Environmental Earth Sciences - ENVIRON EARTH SCI* 66:1–19.

Faccini, Francesco, Robert A. Hearn, Sergio Pedemonte, and Alessandro Sacchini. 2017. “Geo-Hydrological Hazard Underlined by Toponymy of Italian Military Geographycal Ins:Tute Maps at 1:25,000 Scale. Contribu:Ons to the Knowledge of Floodplain and Hillslope Dinamycs.”

Fan, Linfeng, Peter Lehmann, Chunmiao Zheng, and Dani Or. 2020. “Rainfall Intensity Temporal Patterns Affect Shallow Landslide Triggering and Hazard Evolution.” *Geophysical Research Letters* 47(1).

Feng, Chen Chieh, and David M. Mark. 2017. “Cross-Linguistic Research on Landscape Categories Using GEOnet Names Server Data: A Case Study for Indonesia and Malaysia.” *Professional Geographer* 69(4).

Fox, John, and Sanford Weisberg. 2019. *An {R} Companion to Applied Regression*. Third. Thousand Oaks {CA}: Sage.

Fox, Shari, Esa Qillaq, Ilkoo Angutikjuak, Dennis Joseph Tigullaraq, Robert Kautuk, Henry Huntington, Glen E. Liston, and Kelly Elder. 2020. “Connecting Understandings of Weather and Climate: Steps towards Co-Production of Knowledge and Collaborative Environmental Management in Inuit Nunangat.” *Arctic Science* 6(3):267–78.

Fu, Sheng, Lixia Chen, Tsehaie Woldai, Kunlong Yin, Lei Gui, Deying Li, Juan Du, Chao Zhou, Yong Xu, and Zhipeng Lian. 2020. “Landslide Hazard Probability and Risk Assessment at the Community Level: A Case of Western Hubei, China.” *Natural Hazards and Earth System Sciences* 20(2).

Gasparotto, Melissa. 2016. “Digital Colonization and Virtual Indigeneity: Indigenous Knowledge and Algorithm Bias.”

Goetz, J. N., A. Brenning, H. Petschko, and P. Leopold. 2015. “Evaluating Machine Learning and Statistical Prediction Techniques for Landslide Susceptibility Modeling.” *Computers and Geosciences* 81:1–11.

Graul, Christian. 2016. “LeafletR: Interactive Web-Maps Based on the Leaflet JavaScript Library.”

Gryba, Rowenna. 2020. “How Indigenous Knowledge Is Helping Us Fine-Tune Statistical Models.” *UBC Science*. Retrieved May 9, 2021 (https://focus.science.ubc.ca/stats-660805dd930a).

Gulliksen, Jan, Bengt Göransson, Inger Boivie, Stefan Blomkvist, Jenny Persson, and Åsa Cajander. 2003. “Key Principles for User-Centred Systems Design.” *Behaviour and Information Technology* 22(6):397–409.

Guzzetti, F., M. Cardinali, P. Reichenbach, and A. Carrara. 2000. "Comparing Landslide Maps: A Case Study in the Upper Tiber River Basin, Central Italy." *Environmental Management* 25(3):247–63.

Guzzetti, Fausto, Alberto Carrara, Mauro Cardinali, and Paola Reichenbach. 1999. "Landslide Hazard Evaluation: A Review of Current Techniques and Their Application in a Multi-Scale Study, Central Italy." *Geomorphology* 31(1):181–216.

Guzzetti, Fausto, Alessandro Cesare Mondini, Mauro Cardinali, Federica Fiorucci, Michele Santangelo, and Kang-Tsung Chang. 2012. "Landslide Inventory Maps: New Tools for an Old Problem." *Earth-Science Reviews* 112(1):42–66.

Guzzetti, Fausto, Paola Reichenbach, Francesca Ardizzone, Mauro Cardinali, and Mirco Galli. 2006. "Estimating the Quality of Landslide Susceptibility Models." *Geomorphology* 81(1):166–84.

Guzzetti, Fausto, Paola Reichenbach, Mauro Cardinali, Mirco Galli, and Francesca Ardizzone. 2005a. "Probabilistic Landslide Hazard Assessment at the Basin Scale." *Geomorphology* 72(1–4):272–99.

Guzzetti, Fausto, Paola Reichenbach, Mauro Cardinali, Mirco Galli, and Francesca Ardizzone. 2005b. "Probabilistic Landslide Hazard Assessment at the Basin Scale." *Geomorphology* 72(1):272–99.

Ha, Nguyen Duc, Takahiro Sayama, Kyoji Sassa, Kaoru Takara, Ryosuke Uzuoka, Khang Dang, and Tien Van Pham. 2020. "A Coupled Hydrological-Geotechnical Framework for Forecasting Shallow Landslide Hazard—a Case Study in Halong City, Vietnam." *Landslides* 17(7).

Hearn, G. J., and A. B. Hart. 2019. "Landslide Susceptibility Mapping: A Practitioner's View." *Bulletin of Engineering Geology and the Environment* 78(8):5811–26.

Heikkilä, Karen, and Gail Fondahl. 2010. "Indigenous Toponyms as Pedagogical Tools: Reflections from Research with Tl'azt'en Nation, British Columbia." *Fennia : International Journal of Geography* 188(1):105–22.

Highland, Lynn M. 2008. "Introduction The Landslide Handbook-A Guide to Understanding Landslides." *The Landslide Handbook - A Guide to Understanding Landslides* 4–42.

Hill, Rosemary, Fiona J. Walsh, Jocelyn Davies, Ashley Sparrow, Meg Mooney, Russell M. Wise, and Maria Tengö. 2020. "Knowledge Co-Production for Indigenous Adaptation Pathways: Transform Post-Colonial Articulation Complexes to Empower Local Decision-Making." *Global Environmental Change* 65(August):102161.

Isoda, Yuzuru, Akio Muranaka, Go Tanibata, Kazumasa Hanaoka, Junzo Ohmura, and Akihiro Tsukamoto. 2019. "Strengths of Exaggerated Tsunami-Originated Placenames: Disaster Subculture in Sanriku Coast, Japan." *ISPRS International Journal of Geo-Information* 8(10).

Jenjekwa Vincent. 2018. "A TOPONYMIC PERSPECTIVE ON ZIMBABWE'S POST-2000 LAND REFORM PROGRAMME (THIRD CHIMURENGA)."

Johnson, Sarah, and Keith H. Basso. 1998. "Wisdom Sits in Places: Landscape and Language among the Western Apache." *The Journal of American Folklore* 111(442):444.

Jones, Richard. 2016. "Responding to Modern Flooding: Old English Place-Names as a Repository of Traditional Ecological Knowledge." *Journal of Ecological Anthropology* 18(1).

Kamen-Kaye, Maurice. 1975. "Ethnogeology?" (March):13210.

Kelman, Ilan, Jessica Mercer, and Jc Gaillard. 2012. "Indigenous Knowledge and Disaster Risk Reduction." *Geography* 97(1):12–21.

King, Darren N. T., James Goff, and Apanui Skipper. 2007. "Māori Environmental Knowledge and Natural Hazards in Aotearoa-new Zealand." *Journal of the Royal Society of New Zealand* 37(2):59–73.

Kirschbaum, Dalia, and Thomas Stanley. 2018. "Satellite-Based Assessment of Rainfall-Triggered Landslide Hazard for Situational Awareness." *Earth's Future* 6(3):505–23.

Kleiber, Christian, and Achim Zeileis. 2008. *Applied Econometrics with {R}*. New York: Springer-Verlag.

Lai, Jhe Syuan, and Fuan Tsai. 2019. "Improving GIS-Based Landslide Susceptibility Assessments with Multi-Temporal Remote Sensing and Machine Learning." *Sensors (Switzerland)* 19(17).

Lambert, Simon J., and John C. Scott. 2019. "International Disaster Risk Reduction Strategies and Indigenous Peoples." *International Indigenous Policy Journal* 10(2).

Lambrecht, Francis. 1978. *Ifugao-English Dictionary*. Baguio City: Catholic Vicar Apostolic of the

Mountain Province.
Latulippe, Nicole, and Nicole Klenk. 2020. "Making Room and Moving over: Knowledge Co-Production, Indigenous Knowledge Sovereignty and the Politics of Global Environmental Change Decision-Making." *Current Opinion in Environmental Sustainability* 42:7–14.
Lavine, Michael L., and James S. Hodges. 2012. "On Rigorous Specification of ICAR Models." *American Statistician* 66(1):42–49.
Lee, Jung Hyun, Hanbeen Kim, Hyuck Jin Park, and Jun Haeng Heo. 2020. "Temporal Prediction Modeling for Rainfall-Induced Shallow Landslide Hazards Using Extreme Value Distribution." *Landslides*.
Levy Tacher, Samuel Israel, and J. D. Duncan Golicher. 2004. "How Predictive Is Traditional Ecological Knowledge? The Case of the Lacandon Maya Fallow Enrichment System." *Interciencia* 29(9):496.
Li, Huajin, Qiang Xu, Yusen He, Xuanmei Fan, and Songlin Li. 2020. "Modeling and Predicting Reservoir Landslide Displacement with Deep Belief Network and EWMA Control Charts: A Case Study in Three Gorges Reservoir." *Landslides* 17(3):693–707.
Liedloff, A. C., E. L. Woodward, G. A. Harrington, and S. Jackson. 2013. "Integrating Indigenous Ecological and Scientific Hydro-Geological Knowledge Using a Bayesian Network in the Context of Water Resource Development." *Journal of Hydrology* 499.
Link, William A., and Richard J. Barker. 2006. "Model Weights and the Foundations of Multimodel Inference." *Ecology* 87(10):2626–35.
Lombardo, Luigi, Thomas Opitz, Francesca Ardizzone, Fausto Guzzetti, and Raphaël Huser. 2020. "Space-Time Landslide Predictive Modelling." *Earth-Science Reviews* 209.
Lombardo, Luigi, Thomas Opitz, and Raphael Huser. 2018. "Numerical Recipes for Landslide Spatial Prediction by Using R-INLA: A Step-by-Step Tutorial." P. To be assigned in.
Lombardo, Luigi, and Hakan Tanyas. 2020. "Chrono-Validation of near-Real-Time Landslide Susceptibility Models via Plug-in Statistical Simulations." *Engineering Geology* 278.
Mackenzie Kierin, Siabato Willington, Reitsma Femke, Claramunt Christophe. 2017. "Spatio-Temporal Visualisation and Data Exploration of Traditional Ecological Knowledge/Indigenous Knowledge." *Conservat Soc* 15:41–58.
Mark, David, and Andrew Turk. 2011. "Ethnophysiography."
Martin, Marlon (personal communication, February 4, 2021) quoted indigenous invocations.
Martinello, Chiara, Chiara Cappadonia, Christian Conoscenti, Valerio Agnesi, and Edoardo Rotigliano. 2020. "Optimal Slope Units Partitioning in Landslide Susceptibility Mapping." *Journal of Maps* 0(0):1–11.
Matuk, Fernanda Ayaviri, Carlos Ernesto Gonçalves Reynaud Schaefer, Felipe Nogueira Bello Simas, Thiago Torres Costa Pereira, Davi Feital Gjorup, and France Maria Gontijo Coelho. 2017. "Ethnopedology of a Quilombola Community in Minas Gerais: Soils, Landscape, and Land Evaluation." *Revista Brasileira de Ciencia Do Solo* 41.
Mazzocchi, Fulvio. 2018. "Why 'Integrating' Western Science and Indigenous Knowledge Is Not an Easy Task: What Lessons Could Be Learned for the Future of Knowledge?" *Journal of Futures Studies* 22(3):19–34.
von Mentz, B. 2017. "Toponyms and Chronology: Notes on a New Approach to the Study of the Past." *Historia Mexicana* 67(1):7–59.
Morphy, Frances, Howard Morphy, Patrick Faulkner, and Marcus Barber. 2020. "Toponyms from 3000 Years Ago? Implications for the History and Structure of the Yolŋu Social Formation in North-East Arnhem Land." *Archaeology in Oceania* 55(3):153–67.
Morris, Mitzi. 2019. "Spatial Models in Stan: Intrinsic Auto-Regressive Models for Areal Data." Retrieved May 27, 2021 (https://mc-stan.org/users/documentation/case-studies/icar_stan.html).
Nakashima, Douglas, Marie Rou, and Ted Munn. 2002. *Indigenous Knowledge , Peoples and Sustainable Practice Indigenous Knowledge , Peoples and Sustainable*. Vol. 5. Wiley.
Nielsen, Jakob. 2012. "Usability 101: Introduction to Usability." *Nielsen Norman Group*. Retrieved May 22, 2021 (https://www.nngroup.com/articles/usability-101-introduction-to-

usability/#:~:text=Usability and utility are equally,it's not what you want.&text=Definition of Usability %3D how easy,of Useful %3D usability %2B utility.).

Noordegraaf, Mirko, Scott Douglas, Karin Geuijen, and Martijn Van Der Steen. 2019. "Weaknesses of Wickedness: A Critical Perspective on Wickedness Theory." *Policy and Society* 38(2):278–97.

Nyetanyane, John, and Muthoni Masinde. 2020. "Integration of Indigenous Knowledge, Climate Data, Satellite Imagery and Machine Learning to Optimize Cropping Decisions by Small-Scale Farmers. a Case Study of UMgungundlovu District Municipality, South Africa." Pp. 3–19 in *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*. Vol. 321 LNICST. Springer.

Pataueg, Josephine (personal communication, January 31, 2021-February 3, 2021) discussed morphemes of the local language.

Parsons, Meg, Karen Fisher, and Johanna Nalau. 2016. "Alternative Approaches to Co-Design: Insights from Indigenous/Academic Research Collaborations." *Current Opinion in Environmental Sustainability* 20:99–105.

Parsons, Sarah, and Kristine Lister. 2018. "Italy, Austria And China Top The List Of Countries At High Risk Of Landslides Right Now." *RESOURCEWATCH*. Retrieved May 7, 2021 (https://blog.resourcewatch.org/2018/08/27/italy-austria-and-china-top-the-list-of-countries-at-high-risk-of-landslides-right-now/).

Perdana, Aji Putra, and Frank O. Ostermann. 2019. "Eliciting Knowledge on Technical and Legal Aspects of Participatory Toponym Handling." *ISPRS International Journal of Geo-Information* 8(11):500.

Pradhan, Ananta Man Singh, Seung Rae Lee, and Yun Tae Kim. 2019. "A Shallow Slide Prediction Model Combining Rainfall Threshold Warnings and Shallow Slide Susceptibility in Busan, Korea." *Landslides* 16(3):647–59.

Raymond, Christopher M., Ioan Fazey, Mark S. Reed, Lindsay C. Stringer, Guy M. Robinson, and Anna C. Evely. 2010. "Integrating Local and Scientific Knowledge for Environmental Management." *Journal of Environmental Management* 91(8):1766–77.

Reichenbach, Paola, Mauro Rossi, Bruce D. Malamud, Monika Mihir, and Fausto Guzzetti. 2018. "A Review of Statistically-Based Landslide Susceptibility Models." *Earth-Science Reviews* 180:60–91.

Reid, Andrea J., Lauren E. Eckert, John-Francis Lane, Nathan Young, Scott G. Hinch, Chris T. Darimont, Steven J. Cooke, Natalie C. Ban, and Albert Marshall. 2021. "'Two-Eyed Seeing': An Indigenous Framework to Transform Fisheries Research and Management." *Fish and Fisheries* 22(2):243–61.

Rossi, Mauro, Fausto Guzzetti, Reichenbach Paola, Alessandro Mondini, and Silvia Peruccacci. 2010. "Optimal Landslide Susceptibility Zonation Based on Multiple Forecasts." *Geomorphology* 114:129–42.

Roth, Dennis. 1974. "Notes on the Ethnohistory of Northern Luzon." *Ethnohistory* 21(4):371–78.

Roth, Robert E., Kevin S. Ross, and Alan M. MacEachren. 2015. "User-Centered Design for Interactive Maps: A Case Study in Crime Analysis." *ISPRS International Journal of Geo-Information* 4(1):262–301.

Roué, Marie, Samuel Roturier, AlexandrA Lavrillier, Semen Gabyshev, and Douglas Nakashima. 2016. *Coproduction of Indigenous and Scientific Knowledge as a Response to Global Change in the Arctic: Case Studies from Eurasian Reindeer Herding Peoples*. Fairbanks.

Rue, H., S. Martino, and N. Chopin. 2009. "Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with Discussion)." *Journal of the Royal Statistical Society, Series B* 71(2):319–92.

Sarmiento, Ivan, Sergio Paredes-Solís, David Loutfi, Anna Dion, Anne Cockcroft, and Neil Andersson. 2020. "Fuzzy Cognitive Mapping and Soft Models of Indigenous Knowledge on Maternal Health in Guerrero, Mexico." *BMC Medical Research Methodology* 20(1):125.

Schmaltz, Elmar M., Stefan Steger, and Thomas Glade. 2017. "The Influence of Forest Cover on Landslide Occurrence Explored with Spatio-Temporal Information." *Geomorphology* 290:250–64.

Shano, Leulalem, Tarun Kumar Raghuvanshi, and Matebie Meten. 2020. "Landslide Susceptibility

Evaluation and Hazard Zonation Techniques – a Review." *Geoenvironmental Disasters* 7(1).

Skroblin, Anja, Tracy Carboon, Gladys Bidu, Nganjapayi Chapman, Minyawu Miller, Karnu Taylor, Waka Taylor, Edward T. Game, and Brendan A. Wintle. 2020. "Including Indigenous Knowledge in Species Distribution Modeling for Increased Ecological Insights." *Conservation Biology* 35(2):587–97.

Soeters, R., and C. J. Westen. 1996. "Slope Instability Recognition, Analysis and Zonation." *In: Turner, A.K., Schuster, R.L. (Eds.), Landslide: Investigations and Mitigation. Special Report, Vol. 247. Transportation Research Board, National Research Council, National Academy Press, Washington, D.C.* 129–17.

Standardization, International Organization for. 2018. "Ergonomics of Human-System Interaction — Part 11: Usability: Definitions and Concepts." Retrieved (https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en).

Steger, S., A. Brenning, R. Bell, and T. Glade. 2016. "The Propagation of Inventory-Based Positional Errors into Statistical Landslide Susceptibility Models." *Natural Hazards and Earth System Sciences* 16(12):2729–45.

Steger, Stefan, and Christian Kofler. 2019. "Statistical Modeling of Landslides." Pp. 519–46 in *Spatial Modeling in GIS and R for Earth and Environmental Sciences*. Elsevier.

Summer Institute of Linguistics. 2014. "Tuwali Ifugao Dictionary and Grammar Sketch." Retrieved (https://www.webonary.org/tuwali-ifugao/).

Süzen, Mehmet Lütfi, and Basak Şener Kaya. 2012. "Evaluation of Environmental Parameters in Logistic Regression Models for Landslide Susceptibility Mapping." *International Journal of Digital Earth* 5(4):338–55.

Tent, Jan. 2015. "Approaches to Research in Toponymy." *Names* 63(2):65–74.

Tsai, Bor-Wen, and Yung-Ching Lo. 2013. *THE SPATIAL KNOWLEDGE OF INDIGENOUS PEOPLE IN MOUNTAINOUS ENVIRONMENTS: A CASE STUDY OF THREE TAIWANESE INDIGENOUS TRIBES*.

TSMC. 2017. "Annual Report."

UN Secretary-General. 2016. *Report of the Open-Ended Intergovernmental Expert Working Group on Indicators and Terminology Relating to Disaster Risk Reduction*. Vol. 21184.

UNDRIP. 2017. "The 2007 United Nations Declaration on the Rights of Indigenous Peoples." *Austrian Review of International and European Law Online* 17(1):137–265.

UNISDR. 2009. *Chapter 1. The Global Challenge: Disaster Risk, Poverty and Climate Change*.

UNISDR. 2015. *Sendai Frameworkfor Disaster Risk Reduction2015 - 2030*.

United States Philippine Commission (1899-1900). 1904. *Report of the Philippine Commission to the Secretary of War*.

Uzielli, Marco, Guido Rianna, Fabio Ciervo, Paola Mercogliano, and Unni K. Eidsvig. 2018. "Temporal Evolution of Flow-like Landslide Hazard for a Road Infrastructure in the Municipality of Nocera Inferiore (Southern Italy) under the Effect of Climate Change." *Natural Hazards and Earth System Sciences* 18(11).

Vannieuwenhuyze, Bram. 2007. "The Study and Classification of Medieval Urban Toponymy: The Case of Late Medieval Brussels (13th-16th Centuries)." *Onoma: Bibliographical and Information Bulletin* 42:189–211.

Varnes, D. J. 1978. "Slope Movement Types and Processes." Pp. 11–33 in *Special Report 176: Landslides: Analysis and Control*, edited by R. L. Schuster and R. J. Krizek. Washington D. C.: Transportation and Road Research Board, National Academy of Science,.

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

Weichselgartner, Juergen, and Patrick Pigeon. 2015. "The Role of Knowledge in Disaster Risk Reduction." *International Journal of Disaster Risk Science* 6(2):107–16.

van Westen, C. J., T. W. J. van Asch, and R. Soeters. 2006. "Landslide Hazard and Risk Zonation—Why Is It Still so Difficult?" *Bulletin of Engineering Geology and the Environment* 65(2):167–84.

WHC. 1995. *Convention Concerning The Protection Of The World Cultural And Natural Heritage WHC-*

95/CONF.203/16. Berlin.

Wheeler, Helen C., Finn Danielsen, Maryann Fidel, Vera Hausner, Tim Horstkotte, Noor Johnson, Olivia Lee, Nibedita Mukherjee, Amy Amos, Heather Ashthorn, Øystein Ballari, Carolina Behe, Kaitlin Breton-Honeyman, Gunn-Britt Retter, Victoria Buschman, Pâviârak Jakobsen, Frank Johnson, Bjarne Lyberth, Jennifer A. Parrott, Mikhail Pogodaev, Rodion Sulyandziga, and Nikita Vronski. 2020. "The Need for Transformative Changes in the Use of Indigenous Knowledge along with Science for Environmental Decision-making in the Arctic." *People and Nature* 2(3):544–56.

Wheeler, Helen C., and Meredith Root-Bernstein. 2020. "Informing Decision-Making with Indigenous and Local Knowledge and Science." *Journal of Applied Ecology* 57(9):1634–43.

Whyte, Kyle Powys. 2018. "What Do Indigenous Knowledges Do for Indigenous Peoples." in *Traditional Ecological Knowledge: Learning from Indigenous Practices for Environmental Sustainability*, edited by M. Nelson and D. Shilling. Cambridge: Cambridge University Press.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wieczorek, Gerald f. 1984. "Preparing a Detailed Landslide-Inventory Map for Hazard Evaluation and Reduction." *Environmental and Engineering Geoscience* xxi(3):337–42.

Appendices

APPENDIX A. SAMPLE SURVEY CHECKLIST

Place Name Association 1

Please tick on the boxes that suits the meaning of the place

AMBABAG

- ☐ Is the name exclusive for places in high elevations or can it also be used for places in lower elevations like places near rivers?
- ☐ Does the name denote placement on a slope only? Can it be used on the plains?
- ☐ Does the name denote wetness, moisture, or nearness to a wet area?
- ☐ Does the name denote geographic direction? Does it refer to placement with respect to sun movement/directions?
- ☐ Does the name indicate the presence of rocks?
- ☐ Does the name indicate convex curvature on the planar dimension? (It bulges/curves out)
- ☐ Does the name indicate profile curvature? Is it curving outwards from the inside?

What type of land cover, land use or landform is it?

- ☐ Old settlement/village
- ☐ Forest/Headwater
- ☐ Vegetation (grass and shrubs)
- ☐ Agroforest (Muyung or Swidden)
- ☐ Rice terraces or paddies
- ☐ Mountain/Hill
- ☐ River/stream/pond

BAGNIT

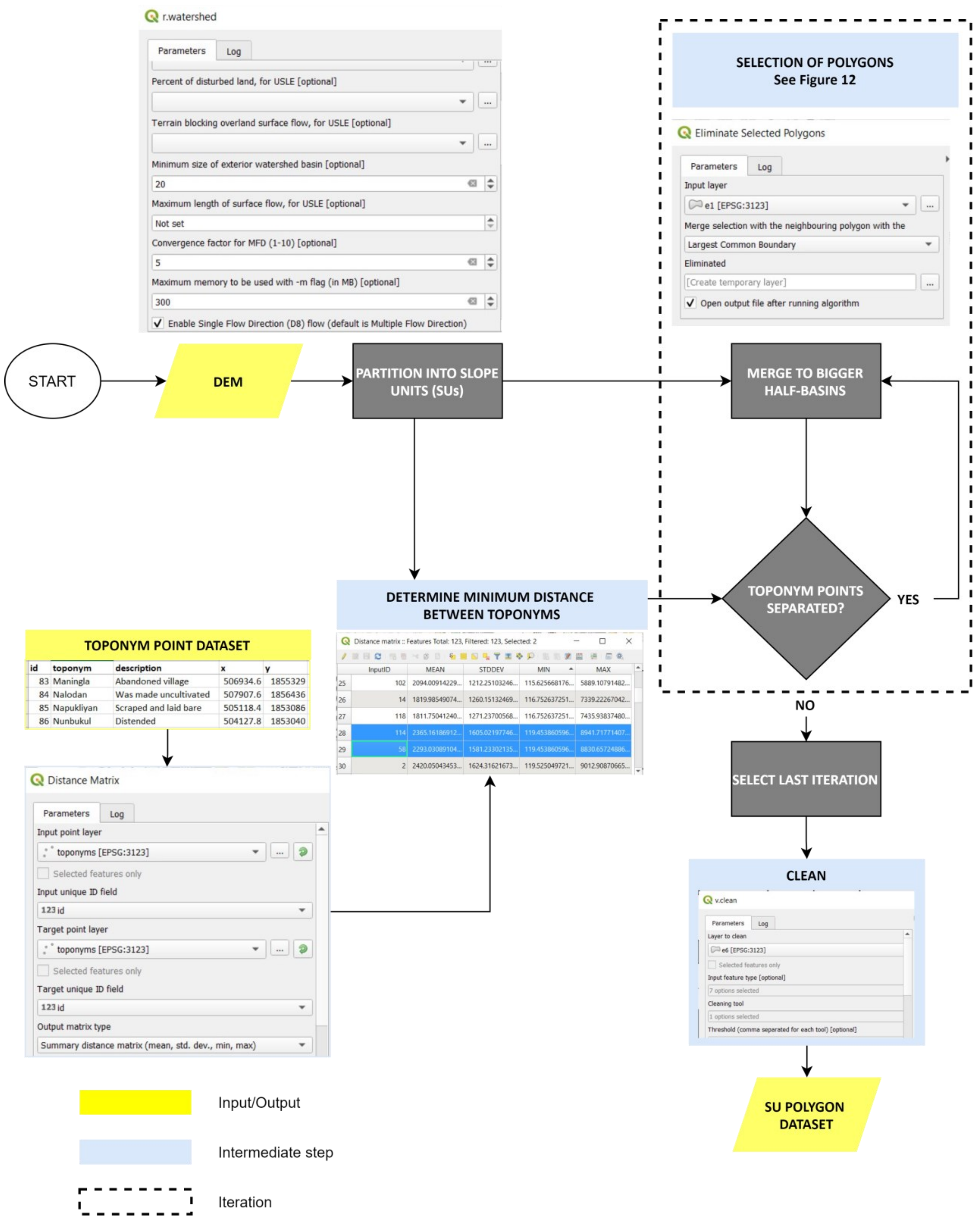
- ☐ Is the name exclusive for places in high elevations or can it also be used for places in lower elevations like places near rivers?
- ☐ Does the name denote placement on a slope only? Can it be used on the plains?
- ☐ Does the name denote wetness, moisture, or nearness to a wet area?
- ☐ Does the name denote geographic direction? Does it refer to placement with respect to sun movement/directions?
- ☐ Does the name indicate the presence of rocks?
- ☐ Does the name indicate convex curvature on the planar dimension? (It bulges/curves out)
- ☐ Doesthe name indicate profile curvature? Is it curving outwards from the inside?

What type of land cover, land use or landform is it?

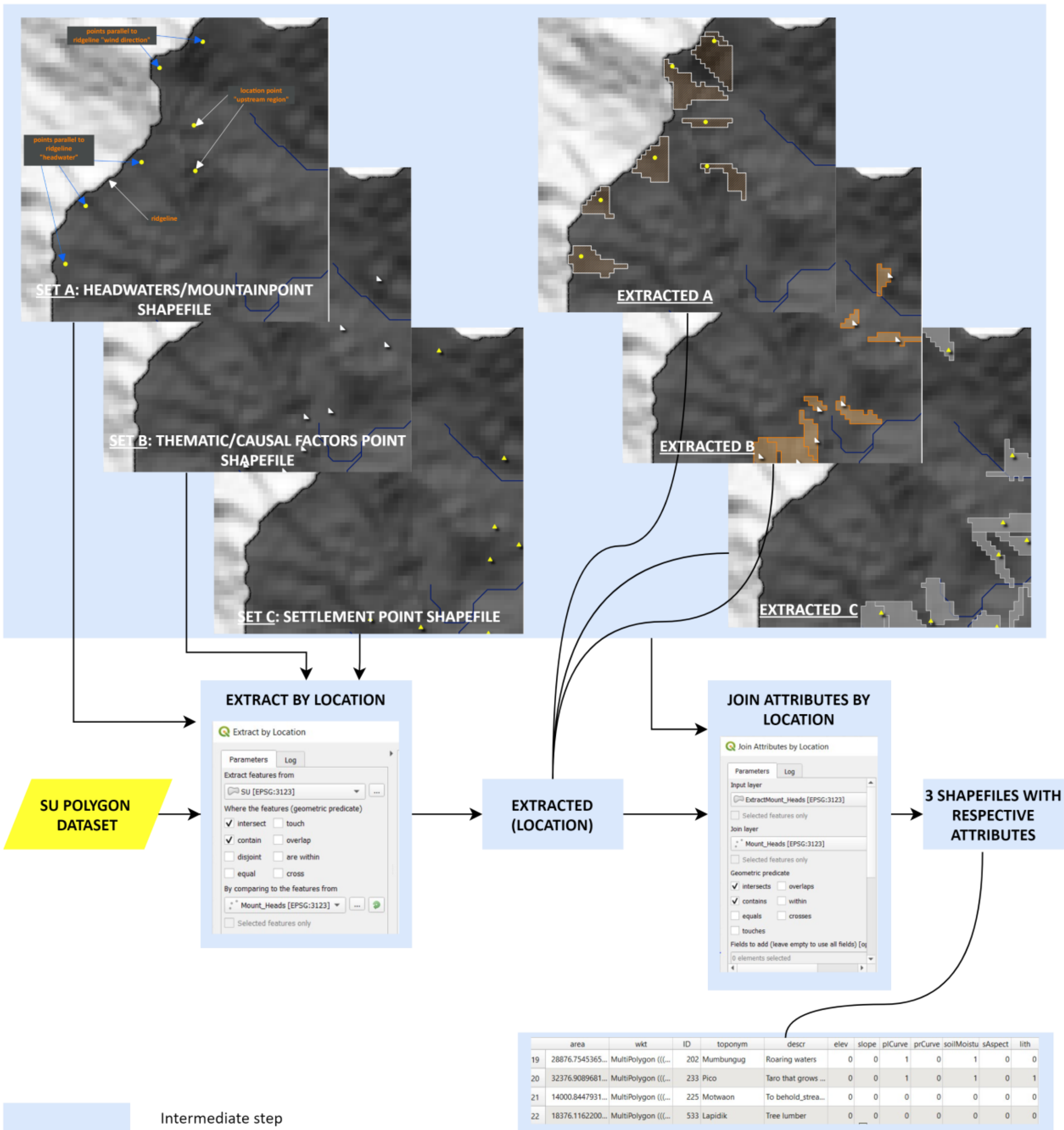
- ☐ Old settlement/village
- ☐ Vegetation (forest)
- ☐ Vegetation (grass and shrubs)
- ☐ Agroforest (Muyung or Swidden)
- ☐ Rice terraces or paddies
- ☐ Mountain/Hill
- ☐ River/stream/pond

Next

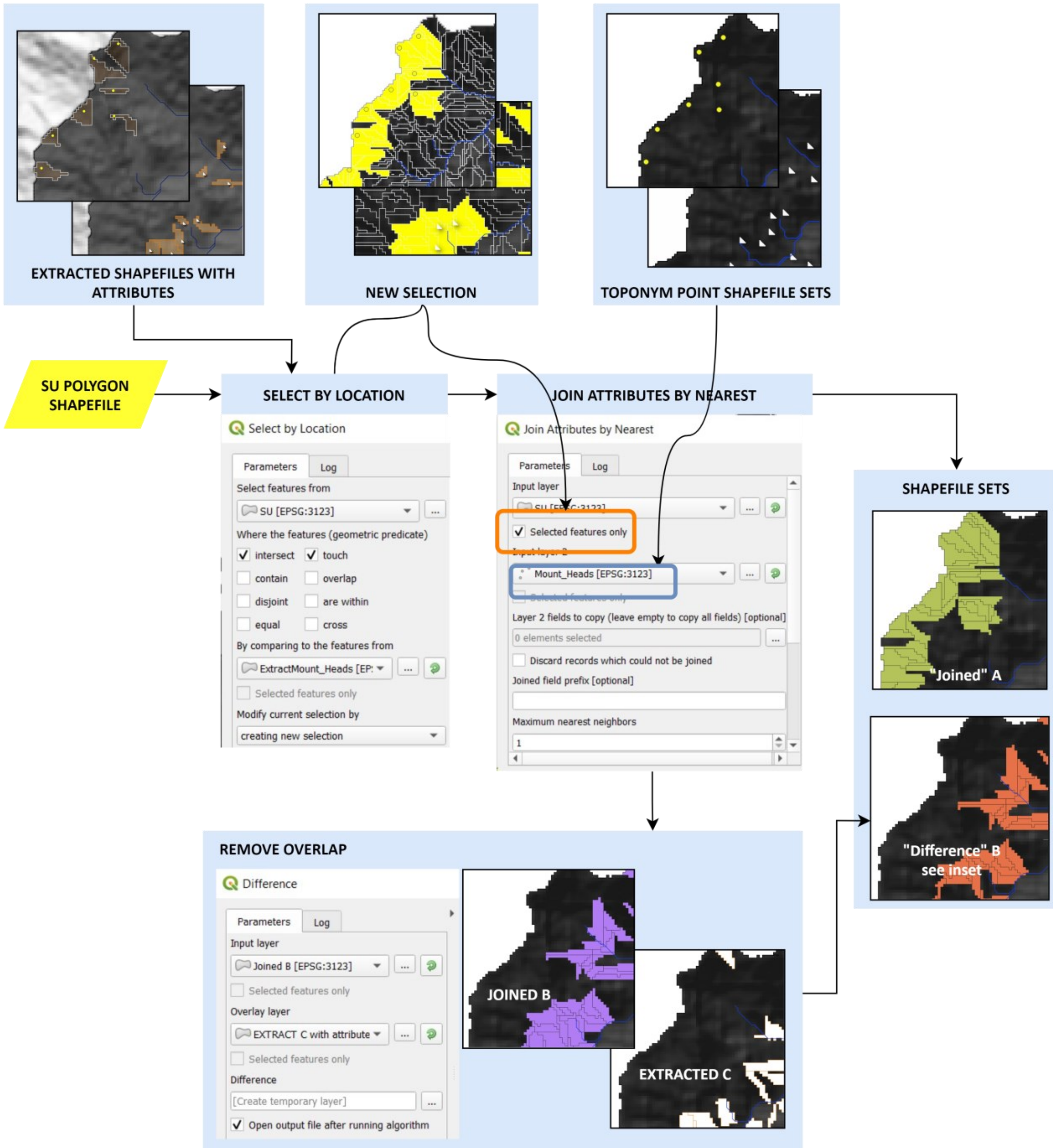
APPENDIX B. OPTIMIZING SLOPE UNITS WORKFLOW DETAIL



APPENDIX C. TWO-STEP PROCESS OF JOINING ATTRIBUTES



APPENDIX D. ADDING NEIGHBOURING SLOPE UNITS TO SET A AND SET B



APPENDIX E. SELECTION OF SIGNIFICANT VARIABLES MODEL 0 AND MODEL 1

	mean	sd	0.025quant	0.5quant	0.975quant
(Intercept)	-12.9917	0.18982	-13.3799	-12.9859	-12.6358
tg1	0.466708	0.184696	0.097224	0.468825	0.824368
tg2	0.688038	0.168565	0.344525	0.692357	1.007905
tg6	0.49259	0.196657	0.091828	0.497628	0.865766
tg8	0.342162	0.070986	0.201195	0.342699	0.480021
tg9	0.653126	0.235988	0.172214	0.659172	1.100935
tg10	1.455169	0.142821	1.177829	1.454007	1.739065
tg14	0.542353	0.16276	0.20527	0.548351	0.845527
tg19	0.423098	0.208342	-0.00737	0.430541	0.811028
tg20	1.169348	0.457907	0.192674	1.197085	1.994864
tg23	0.782822	0.344151	0.050026	0.803192	1.403503
tg24	0.679906	0.133806	0.414734	0.680539	0.941372
tg25	0.32281	0.066027	0.187585	0.324746	0.447156
tg26	0.496092	0.043893	0.411568	0.495508	0.584022
tg27	0.857087	0.343431	0.124582	0.877888	1.476231
tg28	0.554143	0.180796	0.172832	0.56334	0.884335
tg29	0.440905	0.097277	0.242069	0.443646	0.624409
tg30	0.466658	0.120541	0.218869	0.470474	0.693054
tg33	0.453479	0.149087	0.139872	0.46075	0.726537
tg34	0.539305	0.196658	0.138543	0.544343	0.912481
tg35	0.600144	0.079749	0.442223	0.600563	0.755577
tg36	0.561415	0.090543	0.380422	0.562484	0.736326
tg37	0.69035	0.053876	0.586706	0.689636	0.798157
tg38	0.707863	0.115259	0.477647	0.709085	0.931144
tg39	0.494351	0.057223	0.382361	0.494213	0.606994
tg40	0.623355	0.041021	0.545063	0.622574	0.706138
tg41	0.922166	0.116787	0.693671	0.92177	1.152606
tg42	1.006662	0.165236	0.680682	1.006962	1.330658
tg43	0.767727	0.11352	0.545524	0.767339	0.991824
tg44	0.573819	0.146255	0.273781	0.578268	0.848876
tg45	0.48202	0.15184	0.165575	0.488337	0.762702
tg46	0.589128	0.101343	0.38383	0.591292	0.782357
tg48	0.505557	0.274743	-0.08044	0.522197	1.000872
tg49	0.472179	0.138974	0.186782	0.476552	0.732806
tg50	0.823697	0.277039	0.269479	0.826872	1.360176
tg55	0.479974	0.187331	0.089463	0.487789	0.82612
tg56	0.666145	0.173485	0.312436	0.670677	0.994769
tg57	0.834945	0.184697	0.465459	0.837061	1.192605
tg58	0.554199	0.11551	0.324635	0.554988	0.779229

ITERATION 1

APPENDIX F. SELECTION OF SIGNIFICANT VARIABLES MODEL 0 AND MODEL 1

	mean	sd	0.025quant	0.5quant	0.975quant
(Intercept)	-12.8701	0.176279	-13.2304	-12.8649	-12.5394
tg1	0.446002	0.18528	0.075479	0.448076	0.804922
tg2	0.670258	0.169033	0.325925	0.674533	0.991178
tg6	0.471847	0.197204	0.070127	0.476834	0.846251
tg8	0.310083	0.068654	0.173139	0.310812	0.442862
tg9	0.628234	0.236644	0.146174	0.63422	1.077517
tg10	1.413847	0.141492	1.139066	1.412692	1.695153
tg14	0.517167	0.162693	0.180212	0.523162	0.820246
tg20	1.127674	0.458833	0.149389	1.155307	1.955317
tg23	0.718791	0.342908	-0.01182	0.739274	1.33677
tg24	0.664389	0.13426	0.398401	0.664994	0.926815
tg25	0.306761	0.065612	0.172213	0.308738	0.430185
tg26	0.482443	0.043393	0.398897	0.481861	0.569404
tg27	0.825829	0.344127	0.092114	0.846552	1.446569
tg28	0.524327	0.180577	0.143327	0.533568	0.854
tg29	0.415369	0.096449	0.217918	0.418181	0.597059
tg30	0.445762	0.120353	0.198301	0.449588	0.671778
tg33	0.428057	0.148801	0.114944	0.435352	0.700517
tg34	0.518561	0.197204	0.116842	0.523549	0.892966
tg35	0.568087	0.077846	0.413513	0.568637	0.719461
tg36	0.529791	0.08891	0.351659	0.530977	0.701206
tg37	0.665416	0.052085	0.565054	0.664772	0.769513
tg38	0.682735	0.114848	0.45327	0.683971	0.905193
tg39	0.469147	0.055484	0.360247	0.469112	0.578131
tg40	0.602674	0.03937	0.527449	0.601952	0.682054
tg41	0.891143	0.11593	0.664282	0.890756	1.119903
tg42	0.975498	0.165105	0.649813	0.975779	1.299298
tg43	0.74284	0.113216	0.521251	0.742442	0.966391
tg44	0.542004	0.145603	0.243048	0.546511	0.815637
tg45	0.449973	0.151182	0.134607	0.456362	0.729216
tg46	0.56376	0.100635	0.359701	0.565976	0.755473
tg49	0.42909	0.137092	0.14679	0.43365	0.685471
tg50	0.792638	0.277915	0.236863	0.795749	1.331006
tg55	0.436687	0.186316	0.047782	0.444637	0.780549
tg56	0.634353	0.173252	0.281009	0.638908	0.962484
tg57	0.814238	0.185281	0.443714	0.816312	1.173158
tg58	0.536495	0.115634	0.306726	0.537266	0.761819

ITERATION 2

APPENDIX G. SELECTION OF SIGNIFICANT VARIABLES MODEL 2 AND MODEL 3

	mean	sd	0.025quant	0.5quant	0.975quant
(Intercept)	-13.3537	0.243717	-13.8555	-13.3453	-12.8997
tg1	0.71178	0.19322	0.331766	0.711767	1.091623
tg2	0.6316	0.189426	0.249647	0.63492	0.995147
tg6	0.608575	0.180887	0.236631	0.614312	0.948406
tg8	0.320453	0.104368	0.115828	0.320308	0.525689
tg9	0.82604	0.200384	0.409643	0.833884	1.198165
tg10	1.593948	0.167301	1.27101	1.591995	1.928018
tg13	0.376056	0.218467	-0.07659	0.384259	0.781991
tg14	0.638765	0.164883	0.298931	0.644302	0.947341
tg17	0.676041	0.304787	0.036774	0.690267	1.2355
tg18	0.852853	0.420405	-0.04828	0.880281	1.604374
tg19	1.014462	0.301826	0.42127	1.014487	1.606977
tg20	1.572006	0.507943	0.508823	1.594597	2.50944
tg23	0.906617	0.348918	0.165646	0.926466	1.537912
tg24	0.567582	0.133673	0.301957	0.568529	0.827801
tg25	0.291301	0.09518	0.099839	0.292809	0.474239
tg26	0.444024	0.060854	0.326244	0.443402	0.56529
tg27	0.901072	0.353308	0.150862	0.921073	1.541592
tg28	0.545141	0.187094	0.150479	0.55471	0.886558
tg29	0.576393	0.112514	0.351637	0.57764	0.794061
tg30	0.536581	0.135201	0.259133	0.54075	0.790516
tg31	0.467281	0.190368	0.059082	0.479764	0.807543
tg32	0.467405	0.290278	-0.14822	0.483546	0.994158
tg33	0.487073	0.169027	0.139502	0.492499	0.804132
tg34	0.530214	0.179336	0.160137	0.536386	0.865525
tg35	0.630204	0.092984	0.448388	0.629905	0.813517
tg36	0.525061	0.101853	0.323059	0.525728	0.723188
tg37	0.605959	0.064245	0.482407	0.605083	0.734593
tg38	0.53499	0.142185	0.250255	0.536817	0.809381
tg39	0.477665	0.080217	0.322247	0.476938	0.637131
tg40	0.512223	0.06419	0.388472	0.51143	0.640532
tg41	1.034788	0.142036	0.759892	1.033359	1.317724
tg42	1.054466	0.15883	0.740436	1.055057	1.364955
tg43	0.910086	0.118907	0.679966	0.908871	1.14704
tg44	0.720412	0.15422	0.408955	0.723343	1.015458
tg45	0.677277	0.154186	0.36167	0.681751	0.967809
tg46	0.688619	0.113393	0.463265	0.689464	0.909113
tg48	0.534765	0.252431	-0.01023	0.553021	0.981148
tg49	0.835425	0.205323	0.433529	0.834901	1.239867
tg50	0.641913	0.20722	0.214691	0.648953	1.029317
tg55	0.675865	0.197029	0.27574	0.680527	1.049917
tg56	0.822595	0.1811	0.457452	0.825801	1.169775
tg57	1.00514	0.195405	0.617641	1.006179	1.386631
tg58	0.680548	0.115003	0.452947	0.681028	0.905294

ITERATION 1

,APPENDIX H. SELECTION OF SIGNIFICANT VARIABLES MODEL 2 AND MODEL 3,

	mean	sd	0.025quant	0.5quant	0.975quant
(Intercept)	-13.1129	0.213562	-13.5503	-13.1063	-12.7129
tg1	0.617176	0.184788	0.251066	0.618091	0.977992
tg2	0.608717	0.187028	0.231118	0.612176	0.967171
tg6	0.549691	0.177552	0.183178	0.555815	0.881732
tg8	0.284139	0.101329	0.084881	0.284201	0.482857
tg9	0.754462	0.196577	0.344345	0.762755	1.118015
tg10	1.492756	0.158933	1.184504	1.491358	1.808737
tg14	0.593705	0.162112	0.258515	0.599516	0.895961
tg17	0.578521	0.299275	-0.05139	0.593327	1.12573
tg19	0.908818	0.292236	0.332143	0.909625	1.480474
tg20	1.440978	0.500219	0.39068	1.464491	2.360881
tg23	0.79341	0.342647	0.063676	0.813769	1.410975
tg24	0.547873	0.131748	0.2858	0.548903	0.804093
tg25	0.272568	0.093613	0.084066	0.274121	0.452308
tg26	0.425916	0.059387	0.310833	0.425352	0.544138
tg27	0.790556	0.347696	0.049644	0.811379	1.417513
tg28	0.496554	0.184155	0.10715	0.506347	0.831609
tg29	0.519888	0.108078	0.302704	0.521553	0.727693
tg30	0.472094	0.132096	0.200174	0.476452	0.719308
tg31	0.377433	0.18283	-0.0172	0.390622	0.70054
tg33	0.438073	0.166212	0.095089	0.443821	0.748556
tg34	0.470229	0.176184	0.105184	0.476824	0.797926
tg35	0.567877	0.087778	0.395071	0.567994	0.739866
tg36	0.484348	0.098635	0.287849	0.4853	0.675401
tg37	0.563777	0.060416	0.446805	0.563201	0.683972
tg38	0.506866	0.139865	0.22638	0.508805	0.776397
tg39	0.44722	0.077653	0.296337	0.446653	0.601161
tg40	0.485705	0.061878	0.366041	0.485054	0.609046
tg41	0.969187	0.13647	0.704127	0.968114	1.240148
tg42	0.981699	0.153436	0.677004	0.982734	1.280425
tg43	0.856008	0.114592	0.63351	0.855069	1.0837
tg44	0.632394	0.148979	0.329909	0.635854	0.91559
tg45	0.596551	0.148125	0.290806	0.6017	0.873057
tg46	0.623319	0.108778	0.405863	0.624586	0.833602
tg49	0.766277	0.199374	0.374629	0.76623	1.157811
tg50	0.581703	0.203751	0.160226	0.589112	0.961088
tg55	0.597316	0.190959	0.20783	0.602487	0.957934
tg56	0.744091	0.175529	0.388775	0.747734	1.079072
tg57	0.93752	0.191111	0.557389	0.938943	1.309558
tg58	0.628999	0.111526	0.407458	0.62975	0.846202

ITERATION 2

APPENDIX I. ANSWERS TO USAGE TEST EXERCISES, RESPONDENT 1

This is a usage test of a dynamic map on landslide hazard probability. Kindly answer the following questions while exploring the attached "webM2.html" dynamic map. Map coverage: Julongan, Nagacadan, Poblacion, Ambabag, Tuplac drainage basins.

<u>Code of Landslide Causal Factors:</u>
E = Elevation, S=slope, plC=planar curvature, prC=profile curvature
L=Lithology, A=slope aspect, M=moisture, Bu=built-up, Vc=Vegetation cover change(land conversion)

The area is partitioned into slope units that define the movement of slope failure.

Interview questions:

1. Please identify at least 3 villages/settlements that are exposed to landslide occurrence:
   a. Poblacion area down to Mabbalat-Dumanayan-Malpao including Kiangan Central School at the foot of Mount Atade. It has a history of landslides in the past.
   b. Gode-Domang – There are now increasing settlements along the road going up to Patukan. It also has a history of slides as the name denotes. "Gode" means slide
   c. Bilong – Also has increasing settlements and history of slides.

   Note: although Indalmogan is a high risk community, there are no known settlement in that area as far as I know.

2. Give two (2) sites with a **very high** value of landslide probability:

   a. Mount Atade
   b. Indalmogan

Write "Y" if Yes, "N" if No for the statements below. For "N" answers, you may answer which part is difficult.

I can gauge how exposed my house is to landslide occurrence. <u>**Yes**</u>
I can understand the map and its elements - <u>**Yes**</u>
This map can help identify alternative safe routes, or where to construct them.- <u>**Yes**</u>
I can identify which factors cause landslide occurrence. <u>**Yes**</u>
I can easily identify which areas are not safe from landslides. <u>**Yes**</u>
I am able to identify possible infrastructure interventions to minimize landslide effects in specific areas. - <u>**Yes**</u>
This map can help me suggest zoning policies to minimize exposure to hazards (examples: forest protection policy, drain cleaning). <u>**Yes**</u>
I can identify which settlements are likely to be affected by landslides. <u>**Yes**</u>

3. Given a choice between a static map and this dynamic map, what do you prefer to use? A. Static map B. Dynamic map C. <u>**Both**</u>

4. What other information would you like to see on this map?
None.

5. Will you recommend this map to others in your locality for disaster risk reduction? <u>**Yes**</u> **Why?** <u>**To save lives and property by clearly understanding what is indicated in the map on what would likely to happen in any disaster situation and to improve on zoning policies that would minimize exposures to hazards.**</u>

APPENDIX J. ANSWERS TO USAGE TEST EXERCISES, RESPONDENT 2

Interview questions:

1. Please identify at least 3 villages/settlements that are exposed to landslide occurrence:
- Kadibdib (Code 40) with intensity 15.65
- Indalmogan-Nabangkawan (code 26) with intensity 13.61
- Atade (code 43) with intensity 12.99

2. Give two (2) sites with a **very high** value of landslide probability:
Indalmogan-Napukliyan (code 40) and
Gode-Domang (code 35)

Write "Y" if Yes, "N" if No for the statements below. For "N" answers, you may answer which part is difficult.

ð I can gauge how exposed my house is to landslide occurrence. (if you do not have a house there, choose a building) "Y"
ð I can understand the map. "Y"
ð This map can help identify alternative safe routes, or where to construct them. "Y"
ð I can identify which factors cause landslide occurrence. "Y"
ð I can easily identify which areas are not safe from landslides. "Y"
ð I am able to identify possible infrastructure interventions to minimize landslide effects in specific areas. "N"
ð This map can help me suggest zoning policies to minimize exposure to hazards (examples: forest protection policy, drain cleaning). "Y"
ð I can identify which settlements are likely to be affected by landslides. "Y"

3. Given a choice between a static map and this dynamic map, what do you prefer to use? A. Static map B. Dynamic map C. Both Answer=C

4. What other information would you like to see on this map? 3-D map sana, para Makita ang elevations and slopes parang google map. Sana yung creeks/streams or drainage ay color coded din to reflect intensity.

Translation: 3-D map would be cool, so that elevations and slopes are visible like Google Map. The creeks/streams/drainage could have been color coded, too to reflect intensity

5. Will you recommend this map to others in your locality for disaster risk reduction? Why? Yes, I will recommend this map for them to include in their plans like MDRRM risk reduction plan, CLUP, ADSDPP among others.

MDRRM – Municipal Disaster Risk Reduction Plan
CLUP – Comprehensive Land Use Plan
ADSDPP – Ancestral Domain Sustainable Development and Protection Plan