We would like to thank both reviewers for their thoughtful and detailed comments. Here, we have collated our initial, individual responses to both reviewers and added more specific detail as to how we have now altered the manuscript to address each of their points.

Reviewer 1:

1) *The reviewer argues that the method isn't especially unique, and that the sample size used is too large to be considered a 'small initial sub-sample', being approximately half the size of the test sample.*
In this regard, we accept that the manuscript could be clearer in highlighting that the 'small initial sub-sample' we refer to is the ~200 landslides required to undertake the modelling that we describe in the second half of the manuscript. In the final section of the results we have shown that virtually identical membership curves to those derived from the 2006 training landslides can be derived from just 200 randomly selected landslides. Therefore virtually identical hazard models could have been produced from a sample that is only ~5% the size of the test sample. Our motivation for doing this is to identify how soon after the earthquake – in terms of how many landslides are needed to have been mapped - our method could have successfully generated a hazard model. While the use of fuzzy logic is not unique, we argue that our approach to identify the smallest sample required to be successful is innovative and of value.
We have clarified this in the final paragraph of the Introduction (Page 3; Lines 11-24) by more explicitly stating the aims and objectives of the manuscript and more clearly describing the basic approach undertaken.

2) *The reviewer argues that the factors and membership curves identified in this manuscript may not be applicable elsewhere.*
While this may be true, defining a global set of factors is not the aim of this study. Others have already tried such an approach with some success (e.g. Kritikos et al. 2015). Our approach and focus here is to highlight that a very small sample of landslides mapped soon after the earthquake is sufficient to undertake a successful fuzzy logic approach. If this approach is applied in future landslide-triggering earthquakes, different factors and memberships may be needed, and those used in our study may not necessarily be relevant, but this is not our intention.
We have clarified this point by highlighting that others have previously attempted such an undertaking (Page 2; Lines 23-26) but that the aim of this paper is not to generate globally applicable memberships/factors, but instead to rapidly generate location specific memberships/factors (Page 3; Lines 13-16).

3) *The reviewer highlights that other methods may be more applicable than fuzzy logic.*
We agree that other methods are available, each with relative strengths and weaknesses, however our study is not aimed at comparing and contrasting the different methods that are available. We highlight that the use of fuzzy logic builds upon previous work (e.g. Kritikos et al. 2015), the approach performs as well or better than other methods, and critically in a study such as ours where time is of the essence, it is fast to apply. Given that our aim is to quickly apply this in an emergency response, the speed of fuzzy logic is a defining factor in our choice of method.
We highlight this by adding the following to Page 5; Line 2: "…and, importantly for this study, is fast to apply. While other approaches, such as multi-variate statistical analysis, may provide more accurate landslide forecasts, we argue that the marginal gain in forecast accuracy for such approaches is outweighed by the time required to undertake them."

The reviewer asks why we have not tested the use of different values for the fuzzy operator.
Again, we have cited previous work in testing the fuzzy operator value, which identifies 0.9 to be the most appropriate (see: Kritikos et al 2015 & Kritikos and Davies 2016). Lower values result in less over-prediction but more under-prediction, and high values vice versa. With regards to the operator resulting in increasive values, we do not believe this to be the case. The largest hazard value output is 0.93, which occurs in a cell where two input factors have values of 1.
On Page 7; Line 10 we add: "Values less than 0.9 were shown to better forecast landslide non-occurrence, at the expense of landslide occurrence, while values greater than 0.9 achieve better forecasts for landslide occurrence but generally predict high landslide hazard everywhere; 0.9 provides the best compromise (Kritikos et al., 2015)."

4) *The reviewer highlights that the term 'landslide magnitude' has been used differently by other authors, including Malamud, to refer to an event size and that our usage may result in confusion.*
We agree this may lead to confusion, and we address this by referring to 'landslide intensity' throughout the manuscript, which we define as the point density of landslides per unit area, at the first mention of the term.

5) *The reviewer recommends that we explain how the factors selected mechanically cause landslides.*
We feel part of the issue here relates to our use of the term 'mechanically', which we remove to avoid confusion. The intention is to show that each factor has a physical expectation (e.g. ridge tops amplifying shaking) that other studies have shown relate to landslide occurrence.
We add a short description of the physical expectations for each factor in Section 4.2 Data Analysis, Page 8; Lines 13-28.

6) *The reviewer suggests it is important to show how many of the landslides in the training and test samples overlap.*
In this regard, demonstrating the colocation of the two samples is important, and an extra panel in Figure 1 has been added to address this. We also now show the distribution of training and test landslides in regards to slope angle in Fig. 1.
However, we do highlight that the output hazard model is demonstrating the probability of landslides rather than precise locations where landslides will occur. At the scale the model is applied (30 m resolution) and given the variation between contiguous pixels, differences in landslide locations between samples have little effect.

Reviewer 2:

1) *It needs to be clarified in the paper that the developed model is only applicable to the specific region and earthquake…*
Reviewer 1 highlighted a similar issue that the membership curves and factors developed in this manuscript would not necessarily be applicable to other earthquakes or locations. Our response to reviewer 1 (above) highlighted that while others had attempted such an undertaking (e.g. Kritikos et al., 2015; Nowicki et al., 2014) this was not the aim of our study. Given both reviewers highlight this point, a clear and definitive statement as to the precise aims of this study has been added at the end of the introduction (Page 3; Lines 11-24). The aim of this study is to demonstrate a method for rapid post earthquake mapping using fuzzy logic, and thus it is intended that future applications can take the same steps outlined in our study to define their own location and earthquake specific factors and membership functions for rapid response.

2) *…it would be more convincing if the authors can show the similarity or dissimilarity between the training and test datasets, perhaps in the space of the predisposing factors.*
A similar point was also highlighted by reviewer 1 (above), to which we agreed that addressing this issue was necessary. We have therefore added an extra panel in Figure 1 clearly showing distinctions between the training and test datasets and compared the distribution of both datasets in regards to slope angle.

3) *…providing more guidelines and justifications on the choice of membership functions is necessary…*
In the present manuscript, we briefly address the semi-data driven approach to deriving membership functions on Page 5; Lines 25-30. We highlight that the primary fitting method is via linear regression to fit a line with the best 'goodness of fit', but that users may manually alter the function if desired, potentially reducing the goodness of fit.
We add to following description to Page 6; Lines 21-25: "For instance, in circumstances where the best overall goodness of fit is primarily derived from better fitting low influence values rather than high influence values, the output hazard model will be optimistic in its forecasting of landslide hazard. In situations where a conservative forecast is more appropriate, altering the function to better fit high influence values at the expense of overall goodness of fit may be necessary."

4) *Additional information on the landslide non-occurrence data that were used to compute the ROC curves should be provided…*
In this manuscript we have applied the widely accepted approach to deriving ROC curves that have been used in numerous other landslide hazard/susceptibility studies as measures of model success. On Page 6; Lines 22-27 we describe the method used to derive our ROC curves and direct the reader to several appropriate references which describe ROC curve creation in more detail (Metz, 1978; Zweig & Campbell, 1993). We also clearly describe our definition of true and false positive rates, defining false positives (x axis) as being all cells with hazard values above a given threshold within which a landslide is not recorded. True positives (y axis) are defined as all cells above the same hazard threshold within which a landslide has been recorded. The hazard threshold is altered multiple times in order to compute multiple x-y coordinates from which a continuous curve can be plotted. In this regard the false positive rate does include all cells above the hazard threshold with landslide non-occurrence. While we agree with this reviewer that such an approach does have its limitations (particularly around class imbalance) it is beyond the scope of this study to address the potential benefits and limitations of ROC curves and, as previously stated, we note that such an approach to defining ROC curves is common in the landslide modelling community and therefore not at odds with other studies and models.

5) ...*Please provide more explanations on how the map of defined magnitude can benefit the emergency response in addition to providing information on the spatial distribution.*
We add the following explanation to Page 13; Lines 19-22: "While the successful output landslide intensity map (Fig. 7) is coarse in detail, we argue that the output is useful for helping to inform emergency response planning. Overlaying the intensity map with population distributions and critical network data can quickly allow emergency responders to identify regions where landslides are expected to have caused losses and therefore require urgent aid."

6) *The success of the proposed method is highly dependent on the spatial distribution of the initial inventories... This limitation needs to be included in the discussion.*
We have specifically addressed this point in detail in the results section on page 11, lines 15-19 where we describe the variation in goodness of fit for each pre-disposing factor for both the spatially distributed and clustered sub-samples. In fact, it is this observation that leads us to the conclusion that our proposed method is best suited to '*conditions of partial cloud cover where the total visible area of ground... is small but covers a wide area, or where large numbers of landslides are visible...*' (Page 11; Lines 22-23).
We also add a further brief explanation on Page 15; Line 33 – Page 16; Line 9.

Minor Comments:
7) *Page 2, Line 2: Consider adding "for emergency response" after "if an assessment of landsliding is to be useful"...*
Done.

8) *Page 10, Line 18 – 27: Many place names are referred to in the text, but not labelled in Figure 7. Consider adding place names to Figure 7.*
Done.

9) *Page 12, Line 13: "assess" should be replaced by "assesses".*
Done.

10) *Page 13, Line 23: I have trouble understanding the sentence "The model has been shown to be successful despite potential systematic bias in the initial landslide inventories, such as cloud cover above or below specific elevations."... Consider rephrase the sentence. Also consider rewrite "This suggests that if the inventories are systematically biased, the results are unaffected.". In my opinion, the analysis is not enough to conclude that the results will be immune to any systematic bias...*
This sentence refers to the fact that cloud cover post-earthquake predominantly obscured ridgelines and therefore resulted in the systematic bias in the location of landslides in the training inventory. Thus if this bias in landslide locations affected the model, the results should show an inability to forecast test landslides near or at ridgelines. However, we find this is not the case as the model does accurately model test landslides at ridgelines. To address this, we add a description of this systematic bias in the training dataset to section 3 Data (Page 5, Lines 4-6), and reword the sentences suggested by the reviewer for clarification.

11) *Page 14, Line 22: Consider rewrite "this suggests that systematic high fidelity mapping of landslides following an earthquake is not necessary" here and also in the abstract. Although high fidelity landslide mapping takes a lot of time and effort, it is necessary for many applications, such as damage assessment and loss estimation, which require accurate and reliable landslide observations.*
We will add "for informing rapid modelling attempts." here. While we agree that high fidelity landslide mapping has it's uses, such as those suggested by the reviewer, we maintain that it is not necessary to map all landslides in high detail in order to accurately model the spatial extent of landslides post-earthquake, as our modelling has shown.

Rapid post-earthquake modelling of coseismic landslide ~~magnitude~~intensity and distribution for emergency response decision support

Tom R. Robinson[1], Nicholas J. Rosser[1], Alexander L. Densmore[1], Jack G. Williams[1], Mark E. Kincey[1], Jessica Benjamin[1], Heather J.A. Bell[1]

[1]Department of Geography, Durham University, Durham, DH1 3LE, UK

Correspondence to: Tom R. Robinson (tom.robinson@durham.ac.uk)

Abstract. Current methods to identify coseismic landslides immediately after an earthquake using optical imagery are too slow to effectively inform emergency response activities. Issues with cloud cover, data collection and processing, and manual landslide identification mean even the most rapid mapping exercises are often incomplete when the emergency response ends. ~~In t~~This study~~presents~~, we demonstrate how ~~a new,~~ traditional empirical ~~rapid~~ methods for ~~assessing~~modelling the total distribution and relative ~~magnitude~~ intensity (in terms of point density) of coseismic landsliding can be successfully undertaken in the hours and days immediately after an earthquake, allowing the results to effectively inform stakeholders during the response. The method uses fuzzy logic in GIS to quickly assess ~~which~~ and identify the location specific relationships between predisposing factors ~~have influenced~~and landslide occurrence during the earthquake, based on small initial samples of identified landslides. We show that this approach can accurately model both the spatial pattern and the ~~relative magnitude (~~number density~~)~~ of landsliding from the event based on just several hundred mapped landslides, provided they have sufficiently wide spatial coverage, improving upon previous methods. This suggests that systematic high fidelity mapping of landslides following an earthquake is not necessary for informing rapid modelling attempts. Instead, mapping should focus on rapid sampling from the entire affected area to generate results that can inform the model~~ling~~. This method is therefore suited to conditions in which imagery is affected by partial cloud cover, or in which the total number of landslides is so large that mapping requires significant time to complete. The method therefore has the potential to provide a quick assessment of landslide hazard after an earthquake, and may therefore inform emergency operations more effectively compared to current practice.

1 Introduction

Coseismic landslides are one of the most widespread and destructive hazards to result from earthquakes in mountainous environments. Fatalities in earthquakes with landslides have been shown to be up to ten times higher than in comparable earthquakes without landslides (Budimir et al., 2014). Landslides are also a key inhibitor of relief and reconstruction via the

blocking of critical infrastructure and present a chronic hazard, with post-earthquake landslide rates remaining elevated compared to pre-earthquake rates for at least several years (Marc et al., 2015). Rapidly identifying the distribution of landslides following an earthquake is therefore crucial for understanding the total earthquake impacts (Robinson and Davies, 2013), aiding immediate emergency response efforts, including search and rescue (SAR), and assessing the longer-term post-earthquake risks. If an assessment of landsliding is to be useful for emergency response, it needs to be rapid, i.e. generating outputs within the same timeframe (hours to days after the mainshock) as a response is being coordinated, understandable, and communicated to appropriate stakeholders. However, post-earthquake landslide mapping is a difficult and time-consuming task, hindered by issues relating to the collection and processing of appropriate satellite or aerial images, cloud cover, and the slow speeds associated with manually identifying and mapping large numbers of landslides. Following the 2015 M_w 7.8 Gorkha earthquake in Nepal, efforts to rapidly identify and map coseismic landslides using satellite imagery were undertaken by various international groups. However, such issues resulted in most initial mappingthese initial efforts were not being completed until more than one month after the earthquake occurred (e.g., Kargel et al. 2016), and despite most suitable imagery being collected during this time, complete inventories containing polygon data took considerably longer (e.g. Martha et al., 2016; Roback et al., 2017). By comparison, the post-earthquake emergency response began to transition to a recovery phase within days to weeks after the event (Nepal Army, 2016). Consequently, the results of efforts to manually map coseismic landslides can come too late to effectively inform the emergency response to the earthquake.

This issue can potentially be helped through the use of empirically-based models of coseismic landsliding that describe the locations where landslides are more or less likely to have occurred. Several previous attempts have been made to produce landslide models that can be applied rapidly after an earthquake (e.g. Jibson et al., 2000; Godt et al., 2008; Nowicki et al., 2014; Kritikos et al., 2015; Gallen et al., 2016), but with somewhat limited success. These models have attempted to predict the locations where landslides are most likely to have occurred, based either on statistical analysis of the locations of identified landslides from one or more prior earthquakes, or on a simplified Newmark analysis. Approaches using statistical analysis of landslides from previous earthquakes analyse the predisposing factors present at landslide sites and then identify locations with similar combinations of the same predisposing factors. Historically, such approaches are location specific, only using landslides identified from a previous earthquake in the study area to train the model. However, this is not possible in regions without previously compiled landslide inventories. To address this, several have attempteds to develop global relationships between predisposing factors and landslide occurrence have been undertaken (e.g. Nowicki et al., 2014; Kritikos et al., 2015), allowing the model to be location independent and thus rapidly applicableed to the specificany area of interest following an earthquake. However, these models suffer from issues related to the quality and accuracy of input data, availability of training inventories from a sufficient number of different representative environments, and an inability to effectively extrapolate relationships beyond the input data limits.

Approaches using a simplified Newmark analysis adapt the well-established Newmark sliding block model (Newmark, 1965) to predict slope stability using peak ground acceleration (PGA), local slope angle, and estimates of material shear strength properties (Jibson et al., 2000; Gallen et al., 2016). They can therefore be rapidly applied following an earthquake

using only a digital elevation model (DEM) and established ground motion prediction equations. This method is limited however, because estimates of shear strength properties are difficult to obtain and are generally unknown at scales relevant to landsliding (Gallen et al., 2015). Gallen et al. (2016) applied a rapid Newmark analysis model immediately following the 2015 Gorkha earthquake. which estimatedpredicting 2987 landslides above the main slip patch compared to 2214 landslides

5 that they were able to observe. However, their model showed significant discrepancies from the observed spatial pattern of landslides and subsequent mapping (Martha et al., 2016; Roback et al., 2017) has suggested up to ten time s more landslides occurred than Gallen et al's (2016) modelled and observed predicted. This failure is crucial, as the locations where landslides occur dictate the required response as much or perhaps more than the number or volume of landslides. Developin g a method that can accurately model both the spatial distribution and magnitude intensity of landsliding within hours to days of a large

10 earthquake is therefore vital.

This study presents a new approach fordemonstrates how rapid modelling of landslide magnitude intensity (in terms of point density per unit area) and distribution following an earthquake can be successfully undertaken using traditional statistical analysesempirical methods. Instead of attempting to identify global relationships for predisposing factors, this study undertakes a more typical location specific statistical analysis usinges fuzzy logic in Geographic Information Systems (GIS)

15 to rapidly assess relationships for the affected areabased on using small samples of landslides identified by initial mapping efforts soon after the earthquake. The aim is to investigate whether small samples of rapidly mapped landslides these initial mapping efforts can be used to quickly forecast the locations of as-yet unidentified landslides, and what the necessary requirements of these initial mapping efforts are in terms of spatial coverage and number of landslides identifiedrequired. To do this, Wwe first apply the approachmodel to the 2015 Gorkha earthquake using a relatively large training sample of

20 landslides (n = 2006) mapped within 12 days of the event before iteratively reducing the number of training landslides in order to assess the effect on the output landslide forecasts. and discuss the potential speed with which this model could be applied following future earthquakes. The purpose of the study is to demonstrate that meaningful assessments of the magnitude intensity and total distribution of coseismic landslides can be undertaken within a short time after anthe earthquake based on location specific data, allowing the outputs to effectively inform emergency response.

25 **2 Gorkha earthquake application**

The M_w 7.8 Gorkha earthquake occurred on 25 April 2015 with an epicentre ~80 km north-west of Kathmandu (Fig. 1), rupturing a ~150 km section of the Main Himalayan Thrust (MHT) (Avouac et al., 2015; Hubbard et al., 2016). It resulted in intense ground shaking throughout central Nepal leading to large-scale damage to built infrastructure and > 8000 fatalities. The earthquake predominantly affected the Lesser Himalayan region north of Kathmandu, which is characterised by steep

30 slopes and elevations ranging from ~1000 – 3000 m. Martha et al. (2016) and Roback et al. (2017) documented between 15,000 and 25,000 coseismic landslides with a total area of ~90 km^2 and affecting an area > 14,000 km^2 (Fig. 1). The earthquake was followed by a series of powerful aftershocks, including a M_w 7.3 on 12 May 2015, but most of the damage is

believed to have resulted from the mainshock, with just 213 landslides attributed to the 12 May aftershock (Martha et al. 2016). Landsliding from the mainshock led to widespread losses in remote communities, as well as impacts on roads and trails that provide the only access to those communities.

Efforts to rapidly identify and map the resulting landslides led to one of the largest-ever NASA-led satellite image acquisition responses to an earthquake disaster (Kargel et al., 2016). However, persistent cloud cover in the affected region, combined with the time required to download~~ing~~ and georeference~~ing~~ large amounts of imagery, hindered these efforts. By 4 May 2015 (9 days after the mainshock), a joint group from Durham University (DU) and the British Geological Survey (BGS) had mapped the location of just 279 landslides, which increased to 2006 by 7 May, 12 days after the mainshock (Table 1; EwF, 2015~~http://ewf.nerc.ac.uk/2015/05/08/nepal-earthquake-update-on-landslide-hazard-2/~~). A parallel mapping effort was undertaken by a joint group led by the International Centre for Integrated Mountain Development (ICIMOD), NASA, and the University of Arizona (UA), and identified 4312 landslides (Fig. 1) in the period to 2 June 2015 (38 days after the mainshock) (Kargel et al., 2016).

Kargel et al. (2016) argued that this was one of the fastest and broadest emergency remote sensing efforts undertaken by NASA. Importantly, however, the on-the-ground emergency response following the earthquake was decelerated from the end of May, 36 days after the mainshock, and officially ended on 15 June 2015, 51 days after the mainshock (Nepal Army, 2016). Consequently, much of the results of these mapping efforts came too late to inform any humanitarian activities during the emergency response. It is therefore clear that the current approach of manually~~to~~ mapping landslides post-earthquake is not sufficiently fast to effectively inform or support decision making during an emergency response.

In this study, we ~~build a model~~demonstrate an approach that can rapidly assess the potential total distribution of landslides based on an initial, small sample of identified landslides collected soon after the earthquake. To do this, we first model the spatial distribution of landsliding based only on the locations of the 2006 landslides identified by the DU-BGS group up to 7 May 2015. This initial~~e~~ forecast~~model~~ is tested against the 4312 landslides independently mapped by the ICIMOD-NASA-UA group up to 2 June 2015. While this larger dataset is thought to contain some of the 2006 landslides from the DU-BGS group, differences in mapping approaches (e.g. landslide top points versus centroids) between the groups result in minor location differences between the two datasets (Fig. 1). ~~Subsequently~~Following this, we iteratively~~test the effect of~~ reduc~~ing~~e the number of initial training landslides and evaluate the effect on consequent landslide forecasts in order to examine the minimum number of landslides required to achieve a successful~~ly~~ forecast~~model the final mapped spatial distribution and magnitude of landsliding~~.

3 Data

The initial training L~~l~~andslides were manually mapped by the authors in GIS in the immediate aftermath of the earthquake using imagery from a range of sources, including web-hosted high-resolution optical data in Google[TM] Crisis Response (e.g. UK-DMC2 © DMCii, WorldView ©, DigitalGlobe[TM] Inc., SPOT © CNES), imagery accessed via the Disaster Charter,

imagery available from USGS HDDSExplorer, and imagery specifically tasked over regions of interest (e.g. Pleiades © CNES). All data, apart from that which was hosted on web-based GIS platforms, variously required georeferencing, pansharpening, and orthorectification. A database of cloud cover was maintained to monitor where ground had been visible and mapped, and where it remained obscured. Cloud cover predominantly obscured major ridgelines (i.e. watershed boundaries) resulting in the training dataset being systematically biased against landslides that occurred at or near watershed boundaries. All identified landslides were mapped as a single polyline along the long-axis (crest to toe) allowing data on landslide location, extent, and the intersection with infrastructure to be quickly extracted. Point landslide positions at the crest end of the mapped polyline are extracted for this analysis, representing the location of initial failure. Mapping was undertaken at a scale of approximately 1:10,000 and was incrementally posted to the Humanitarian Data Exchange portal (https://data.humdata.org/group/nepal-earthquake), and shared more widely with stakeholders in Nepal in the days and weeks after the earthquake. This mapping effort, and the lessons learned, is described in detail in Williams et al. (forthcoming).

Topographic factor analysis undertaken herein has used the open-source ASTER GDEM V2 (https://asterweb.jpl.nasa.gov/gdem.asp), which has a cell resolution of 30 m, while seismic data was downloaded at the time from the USGS (PGA models and fault plane solutions; http://earthquake.usgs.gov/earthquakes/eventpage/us20002926#shakemap) and Global Centroid Moment Tensor (CMT) catalog (slip vector; http://www.globalcmt.org/).

4 Method and data analysis

Statistical Aapproaches to modelling regional-scale landslide hazard have been summarised in several comparative overviews (e.g., Aleotti and Chowdhury, 1999; Guzzetti et al., 1999; Wang et al., 2005). The present study uses fuzzy logic as the method is computationally simple, can consider highly uncertain data inputs, and has been shown to match or out-perform other approaches (Pradhan, 2010; Bui et al., 2012; Pourghasemi et al., 2012), and, importantly for this study, is fast to apply. While other approaches, such as multi-variate statistical analysis, may provide more accurate landslide forecasts, we argue that the marginal gain in forecast accuracy for such approaches is outweighed by the time required to undertake them.

4.1 Fuzzy logic

Fuzzy logic derives from fuzzy set theory in which an event is assigned a degree of membership, varying between 0 and 1, to a given set, where 1 indicates that the event is entirely related to the set (full membership) and 0 indicates that the event is entirely unrelated to the set (no membership). When utilised for landslide hazard models, this membership value describes how different values of a predisposing factor influence landslide occurrence, with higher values representing greater influence. For instance, steeper slope angles are known to produce higher rates of landsliding compared to shallower slope

angles, meaning the influence of slope angle on landsliding increases with increasing slope angle. Influence (I) is measured

Formatted: Font: Italic

for unique values of each predisposing factor using a kernel density estimate (KDE) by comparing the prevalence of a given

unique value at landslide locations to its prevalence throughout the whole study area:

$$I_x = \frac{Lx_{KDE}}{Sx_{KDE}} / \frac{\sum_{n=1} Ln_{KDE}}{\sum_{n=1} Sn_{KDE}} \tag{1}$$

5   where I_x is the influence of factor value x, Lx_{KDE} is the KDE of value x for landslide locations, Sx_{KDE} is the KDE of value x

throughout the study area, and n is the total number of unique factor values considered. Consequently, if the prevalence of a

value is greater at landslide locations than throughout the whole study area, that value has higher influence on landslide

occurrence. I is then normalised to a 0-1 scale by

$$I_{xnorm} = \frac{I_x - \min(I)}{\max(I) - \min(I)} \tag{2}$$

10   The factors selected for analysis typically have a broad physical expectation in their influence on landsliding, such as

increasing landslide frequency with increasing slope angle, or decreasing landslide frequency with decreasing ground

shaking. These physical expectations are based on the current understanding of landslide processes and previous

observations. Factors that match~~show a clear and definable~~ the broadly expected relationship between influence and factor

values are ~~then~~ carried forward, while factors that ~~show no~~do not match the expected relationship~~, or a relationship that~~

15   ~~cannot be mechanically explained~~ *and* cannot otherwise be explained based on understanding of landslide processes (e.g.

Formatted: Font: Italic

increasing landslide frequency with decreasing slope angle)~~,~~ are removed from the analysis. This assumes that these factors

are not playing an important role in landslide occurrence in the study area.

Factors carried forward undergo a semi data-driven linear regression to find the most appropriate function to describe the

distribution of I_{norm}, using the coefficient of determination (R^2) to assess goodness of fit. A semi data-driven approach is

20   preferred as this allows the user to manually alter the function if necessary (e.g., by using multiple functional forms) to

achieve better goodness of fit for specific influence values at the expense of overall goodness of fit. For instance, in

circumstances where the best overall goodness of fit is primarily derived from better fitting low influence values rather than

high influence values, the output hazard model will be optimistic in its forecasting of landslide hazard. In situations where a

conservative forecast is more appropriate, altering the function to better fit high influence values at the expense of overall

25   goodness of fit may be necessary.~~a conservative approach, it is desirable to achieve greater fit for high influence values at the~~

~~expense of low influence values, as this will produce conservative estimates of landslide hazard.~~ Typical functional forms

considered include linear, exponential, power-law, and polynomial functions. The final function is known as the membership

function, $f(I_{norm})$, which describes how I_{norm} is distributed against all values of a given factor. $f(I_{norm})$ acts as a filter to

convert factor maps (Fig. 2), which show the spatial distribution of unique values within each factor, into influence maps

30   (Fig. 3), which show the spatial variation of I_{norm}.

The resulting influence maps are aggregated on a pixel-by-pixel basis to produce a landslide hazard map that can have values

[0, 1]. This output represents the relative likelihood of landsliding due to the combination of influence from all factors:

values < 0.2 suggest landsliding is very weakly anticipated, while values > 0.8 suggest landsliding is very strongly

anticipated. Landslides are therefore more likely to occur at larger aggregate influence values as landsliding is more strongly anticipated at these locations.

Various approaches to aggregation exist, however the Ffuzzy Gamma approach has been shown to be the most effective for landslide modelling (Bui et al., 2012; Kritikos and Davies, 2015). This approach provides a compromise between the increasing effect of Fuzzy Sum and the decreasing effects of Fuzzy Product (ESRI, 2016). Fuzzy Gamma establishes the combined effect of multiple membership functions for each pixel such that

$$H_{LS} = \left[\prod_{F=1}^{j} f(I_{normF})\right]^{1-\gamma} \cdot \left[1 - \prod_{F=1}^{j}(1 - f(I_{normF}))\right]^{\gamma} \tag{3}$$

where H_{LS} is landslide hazard, I_{normF} is the membership function for factor F, where $F = 1, 2, ..., j$; j is the number of factors to be aggregated; and γ is a user-defined parameter between 0 and 1. The effect of changes in γ has previously been tested, with the optimal value for landslide modelling shown to be ~0.9 (Kritikos and Davies, 2015; Kritikos et al., 2015). Values less than 0.9 were shown to better forecast landslide non-occurrence, at the expense of landslide occurrence, while values greater than 0.9 achieve better forecasts for landslide occurrence but generally predict high landslide hazard everywhere; 0.9 provides the best compromise (Kritikos et al., 2015).

This approach can therefore identify locations in the study area where the combination of predisposing factors results in landsliding being most strongly anticipated, and hence where as-yet unidentified landslides are more likely to have occurred.

4.2 Data analysis

In the present study, landslides mapped by the DU-BGS group prior to 7 May 2015 are used as the initial training dataset (n=2006) and landslides mapped by the ICIMOD-NASA-UA group prior to 2 June 2015 are used as the test dataset (n=4312). The initial training dataset is used to derive the KDEs and membership functions, and the test dataset is used to measure the accuracy and sensitivity of the model via the area under receiver operating characteristic (ROC) curves. The ROC curve is a fundamental tool for evaluating model accuracy and plots the true positive rate of the model against the false positive rate using various cut-off values (Metz, 1978; Zweig and Campbell, 1993). Landslides occurring in cells with H_{LS} values above the cut-off are considered true positives, while cells above the cut-off without landslides are considered false positives. To be considered successful, models must typically achieve area under the curve (AUC) values > 0.7 (Kritikos et al., 2015).

In total, we assess the influence of 12 predisposing factors, each of which has previously been attributed to coseismic landslide occurrence in earthquakes elsewhere (Keefer, 2000; Khazai and Sitar, 2004; Lee et al., 2008; Meunier et al., 2008; Kritikos et al., 2015; Parker et al., 2015) and may therefore influence landsliding in Nepal. They are: (1) slope angle, S; (2) distance from fault plane, D_F; (3) peak ground acceleration (PGA); (4) elevation above sea level (asl), E; (5) distance from rivers, D_R; (6) distance from river confluences, D_C; (7) planform curvature, C_P; (8) normalised ridge-stream distance, (NRSD); (9) hillslope relief, R_H; (10) sub-hillslope relief, R_S; (11) slope aspect relative to the epicentre, Az_{Epi}; and (12) slope aspect relative to the slip vector, Az_{SV}. Distances from fault plane, rivers, and confluences are calculated as the two-

Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic, Subscript
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic, Subscript
Formatted: Font: Italic
Formatted: Font: Italic, Subscript
Formatted: Font: Italic
Formatted: Font: Italic, Subscript
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic, Subscript
Formatted: Font: Italic
Formatted: Font: Italic, Subscript
Formatted: Font: Italic
Formatted: Font: Italic, Subscript
Formatted: Font: Italic
Formatted: Font: Italic, Subscript

dimensional Euclidean (straight line) distance between pixel centroids. For fault plane distance this does not account for depth, assigning ~~landslides~~ pixels directly above the plane with distance = 0. Rivers are defined using a flow accumulation tool with an upstream contributing area threshold of 5 km², and river confluences are defined as the point where two or more rivers intersect. *NRSD* is calculated as the distance between a river and the nearest watershed boundary, where a value of 0 represents a pixel in the river and a value of 1 represents a pixel on the watershed boundary. Relief is calculated as the standard deviation of elevation values within a square window of radius a) 30 cells (hillslope relief); and b) 10 cells (sub-hillslope relief). ~~Slope aspect relative to the epicentre~~ (Az_{Epi}) is derived by combining local slope aspect (*As*) with the Euclidean direction to the epicentre (*Dir*), such that

$$Az_{Epi} = \begin{cases} |(|Dir - As|) - 360|, & |Dir - As| > 180 \\ |Dir - As|, & |Dir - As| \leq 180 \end{cases} \tag{4}$$

Az_{Epi} has values [0, 180], where 0 indicates a hillslope ~~directly~~ facing the epicentre and 180 indicates a hillslope facing ~~directly~~ away from the epicentre. ~~Slope aspect relative to the slip vector~~ (Az_{SV}) is calculated in the same way, by replacing direction to epicentre (*Dir*) in Eq. (4) with slip vector bearing (*SV*).

The broad physical expectations for each factor in terms of influencing landsliding are based on understanding of landslide processes and previous observations (Keefer, 2000; Khazai and Sitar, 2004; Lee et al., 2008; Meunier et al., 2008; Kritikos et al., 2015; Parker et al., 2015). For slope angle, increasing gradient is expected to yield increasing landslide frequency as the angle between normal stress and gravity increases, producing greater downslope force. Increasing distance from fault plane is expected to produce decreasing landslide frequency due to regional attenuation of ground motions and wave amplitudes. More landslides are expected at higher elevations due to a combination of greater gravitational potential and rock weakening processes such as glacial erosion. Increasing distances from rivers and river confluences are expected to return decreasing landslide frequency as the amount of undercutting of basal hillslopes and downwearing by the rivers also decreases, with river confluences potentially acting as knickpoints where greater erosion rates occur. Planform curvature is used to identify ridges and slope shoulders that can amplify ground shaking and result in higher landslide frequencies (Meunier et al., 2008). *NRSD* is therefore expected to produce a quadratic relationship with landslide frequency, with higher landslide frequencies expected close to rivers (*NRSD* ~ 0) and ridges (*NRSD* ~ 1). Larger relief values may result in larger landslide frequencies as gravitationally-induced shear stresses increase and rock mass strength decreases with the local and regional height of hillslopes (Schmidt and Montgomery, 1995). Slope aspect in relation to epicentre and slip vector is expected to increase landslide frequency on either facing or opposing slopes due to directional patterns of topographic amplification resulting from the incidence angle of seismic waves (Meunier et al., 2008).

With the exception of earthquake-specific parameters (~~fault plane distance~~*D_F*, PGA, *Az_{Epi}* and *Az_{SV}*), each of these factors can be directly derived from a global DEM, such as the ASTER GDEM V2 used herein. This reduces the initial data requirements, helping to facilitate rapid modelling. This is especially important for locations such as Nepal where data for other predisposing factors, such as lithology, structural geology or rainfall, are either absent or difficult (and thus time-consuming) to obtain. PGA, fault plane solutions and CMT data are available from the USGS and Global CMT catalog

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

within a few hours of the earthquake occurring. Consequently, no pre-event data collection is necessary to implement this model, as all necessary data can be gathered post-earthquake from openly available sources.

4.2.1 Peak ground acceleration

PGA is given special consideration in the model, separate to the other predisposing factors. PGA models following an
5 earthquake – such as those produced by the USGS ShakeMap – often change frequently as more acceleration data become available (Fig. 4). Rapid assessment of the influence of PGA on landslide occurrence may therefore give misleading results, as the best PGA model available in the immediate earthquake aftermath may not yet accurately reflect the true pattern. For instance, initial PGA models following the Gorkha earthquake showed that PGA was highest in the region south of Kathmandu. This suggested that the majority of landslides, which occurred north of Kathmandu, occurred in locations with
10 low-to-moderate PGA (Fig. 4). The time required to generate a final (or near-final) PGA model can be on the order of several days or weeks, and the shaking model may therefore be completed too late to be effectively incorporated into the rapid assessments considered herein. Furthermore, there is commonly a high degree of scatter in models that attempt to predict the number, size, and location of coseismic landslides from PGA (e.g. Keefer, 1984; Meunier et al., 2007). Results relying on initial PGA models are therefore likely to have such large uncertainties that they are impractical for informing
15 emergency responders.

To account for difficulties in modelling landslides from PGA, we therefore consider~~model~~ PGA as a threshold factor. This involves setting $I_{norm} = 0.0$ for all values below a given threshold, while values above the threshold are assigned $I_{norm} = 0.5$. This ensures that locations that do not exceed the threshold cannot sustain landslides, while landslide influence in locations above the threshold is dictated by the combination of other factors. From Fig. 4 and Martha et al. (2016) and Keefer (1984)
20 this threshold is suggested to be ~0.08 g.

Importantly, while the location and magnitude of PGA values varied between initial USGS ShakeMap models over the first few days, the area contained within the 0.08 g contour was nearly invariant (Fig. 4). Given that ShakeMap interpolates PGA between seismic stations (Wald et al., 2005), it is easier to rapidly define the footprint of the area experiencing shaking (i.e. PGA > 0.08 g) compared to defining the highest values within this zone, as the number of recording stations increases with
25 distance from epicentre, even in regions with low recording density.

5 Results

5.1 Factor influence modelling

Of the 12 predisposing factors considered, seven show a clear and definable relationship between influence and factor value, and these are carried forward in the modelling process. The remaining factors ~~either~~ showed ~~no clear~~a relationship~~, or the~~
30 ~~relationship demonstrated could not be mechanically defined~~ that did not match the physical expectation *and* could not otherwise be explained, and are therefore removed from the analysis. For instance, *NRSD* ~~appears to~~ shows ~~a~~the expected

relationship ~~in which landslides are more strongly influenced~~ at locations close to rivers but not for locations close to ~~than at~~ ridgelines (Fig. 3i). This is in contrast to the relationship observed for planform curvature (Fig. 3e) ~~as well as observations in the literature that suggest that topographic amplification increases landslide occurrence at ridgelines (Densmore and Hovius, 2000; Meunier et al., 2008; Kritikos et al., 2015)~~. This may be a consequence of ridgelines herein for $NRSD$ being defined using watershed boundaries, which are systematically obstructed from view in the images used to collect the initial training inventory. ~~and thus~~Further, landslides occurring on slope 'shoulders' within a watershed are not accounted for by ~~this factor~~$NRSD$, but are accounted for by planform curvature. This highlights an issue with using $NRSD$ to represent topographic amplification on ridgelines as it cannot account for intra-watershed ridgelines. Similarly, the expected relationship with elevation is not recorded and there is no clear reason why landslides should be strongly influenced at elevations of 2000 – 3000 m asl and only weakly influenced at higher elevations (Fig. 3h). More likely this reflects the fact that higher elevations were predominantly outside of the 0.08 g contour. For these reasons, $NRSD$ and elevation are among the four~~five~~ factors that are not carried forward.

For the factors carried forward, increases in slope angle correlate with increased influence up to slope angles of ~65° (Fig. 3a), at which point the relationship ends due to the small area covered by slope angles > 70° (Fig. 2a). Slopes <~15° show no influence suggesting these slope angles are insufficient to sustain landsliding (i.e. I_{norm} = 0). Increased distances from the fault plane, rivers, and river confluences generally result in decreased influence (Fig. 3b-d), despite a small and unexplained increase at distances of ~3000 m from rivers. Increases in relief at both hillslope and sub-hillslope scales are associated with increased influence up to ~200 m and ~90 m respectively (Fig. 3f-g). Similar to slope angle, the variation above these values likely reflects the small area covered by such values (Fig. 2f-g). Larger planform curvature values also correlate with increased influence, with notably higher influence on laterally convex (ridgeline~~divergent~~) pixels~~noses~~ compared to laterally concave (valleys~~convergent~~) ~~hollows~~pixels (Fig. 3e).

These findings are similar to those of other studies (Khazai and Sitar, 2004; Parker et al., 2015; Kritikos et al., 2015) in that they show coseismic landslides are most strongly encouraged on steep, laterally convex hillslopes close to rivers, and most weakly encouraged on shallow, linear slopes situated far from rivers.

5.2 Landslide hazard model

5.2.1 Model success

Using the membership curves derived in Fig. 3, ~~which~~ showing~~ing~~ the variation in I_{norm} for each factor, we aggregate different combinations of the seven predisposing factors carried forward to find the best performing combination of variables to model the spatial distribution and ~~magnitude~~intensity of the 4312 test landslides. Using all seven predisposing factors results in AUC = 0.838 but the best performance (AUC = 0.870) is achieved by excluding sub-hillslope relief and planform curvature (Fig. 5). The worst performing model (AUC = 0.761) excludes hillslope relief and planform curvature (Fig. 5).

Formatted: Font: Italic

Formatted: Font: Italic
Formatted: Font: Italic

Formatted: Font: Italic

Importantly, all models achieve AUC values > 0.7, suggesting that any combination of these factors is able to model spatial landslide distribution with reasonable accuracy.

The output of the most successful model is shown in Fig. 6 and shows the locations where landslides are strongly-to-very strongly anticipated ($H_{LS} > 0.6$) in red and weakly-to-very weakly anticipated ($H_{LS} < 0.4$) in blue. The maximum modelled $H_{LS} = 0.93$ suggest~~ings~~ landsliding from this earthquake was very strongly anticipated in at least some locations. The total area where landslide occurrence is mildly-to-very strongly anticipated ($H_{LS} > 0.4$) covers ~18,000 km^2, approximately the same size as the total affected area (~14,000 km^2) mapped by Martha et al. (2016) and Roback et al. (2017)~~.~~ Twhile the area where landsliding is very strongly anticipated ($H_{LS} > 0.8$) covers ~85 km^2, similar to the 90 km^2 total landslide area mapped by Martha et al. (2016). This area is concentrated in the major river valleys in Gorkha, Dhading, Rasuwa, and Sindhupalchok districts (Fig. 6), corresponding well with observed landsliding (Fig. 1) and confirming the high AUC value. With a combination of a high maximum H_{LS} value and a large area where landsliding is strongly anticipated, the model suggests that widespread landsliding from this earthquake was highly likely.

5.2.2 Landslide ~~magnitude~~intensity and spatial distribution

In order to be useful for informing emergency responders, it is essential that the final output model explicitly demonstrates both the spatial extent and ~~magnitude~~intensity, either in terms of number, area or volume, of landsliding. To do this, we estimate landslide ~~magnitude~~intensity in terms of number density per unit area as a function of the kernel density of pixels with H_{LS} values exceeding some user-defined threshold (e.g. all cells with $H_{LS} \geq 0.9$). The total spatial extent of landsliding is then represented by the area where kernel density is > 0. This assumes that the ~~magnitude~~intensity of landsliding in any given area is directly linked to the frequency of cells above the corresponding threshold within that area. An area dominated by cells with high H_{LS} is expected to experience a higher ~~magnitude~~intensity of landsliding (i.e. a higher number density) than a comparable area dominated by cells with low H_{LS}.

Taking the most successful hazard model (Fig. 6), we set threshold limits at the 99th, 95th, and 90th percentile of H_{LS} values where the area with values above the threshold accounts for 1%, 5%, and 10% of the total area ($H_{LS} > 0$) respectively, corresponding to $H_{LS} = 0.74$, $H_{LS} = 0.65$ and $H_{LS} = 0.60$ (Fig. 6 inset). The corresponding cells are extracted and the kernel density for each threshold is calculated using a 1 km^2 moving window (Fig. 7). To test how well the training inventory is able to model total landslide ~~magnitude~~intensity and distribution, the results are compared against the kernel density of all 4312 test landslides, also defined using a 1 km^2 moving window. Because this study considers individual pixels, the total number of cells exceeding each threshold is large (1% = 388,694; 5% = 1,927,614; 10% = 3,827,036) and consequently the corresponding kernel densities (100 – 1000 per km^2) are not directly comparable to real landslide number densities (1 – 100 per km^2~~; http://ewf.nerc.ac.uk/2015/06/30/updated-30-june-landslide-inventory-following-25-april-and-12-may-nepal-earthquakes/~~). This occurs because groups of adjacent pixels exceeding the threshold may in reality represent only one or two landslides, but each pixel is considered individually in the kernel density analysis. To account for this, we scale all kernel density models onto a 0-1 scale in which the maximum kernel density = 1 and all other values are scaled

proportionally. The results therefore consider relative landslide ~~magnitude~~intensity (i.e. how locations compare relative to each other) rather than absolute ~~magnitude~~intensity.

By setting a threshold at the 99th percentile ($H_{LS} > 0.74$), the model is able to accurately predict both spatial distribution and relative ~~magnitude~~intensity of landsliding (Fig. 7). Generally, the model performs well, although it over-predicts in Gorkha

5   district near the mainshock epicentre and under-predicts in Sindhupalchok district near the M_w 7.3 aftershock. In particular, it is notable that the model successfully highlights the high relative ~~magnitude~~intensity of landsliding near to the mainshock epicentre, despite no training landslides coming from this region (Fig. 1), as well as closely matching the observed spatial pattern of landsliding. However, it fails to identify landslides in the hills south of Kathmandu, suggesting that landslides here were not influenced in the same way, or by the same factors, as landslides further north. None of the training landslides are

10 located in this region (Fig. 1). Comparatively, setting thresholds at the 95th ($H_{LS} > 0.65$) and 90th ($H_{LS} > 0.60$) percentiles substantially over-estimates both the relative ~~magnitude~~intensity, particularly close to the Nepal-China border, and the total landslide pattern, whilst still underestimating landslide ~~magnitude~~intensity near the M_w 7.3 aftershock (Fig. 7).

5.2.3 Minimum landslide numbers

The current method is trained on 2006 landslides mapped within the first 12 days following the mainshock. We test the

15 impact of reducing the number of landslides involved in training the model in order to estimate how soon after an earthquake this method could be implemented, assuming that the number of landslides mapped immediately after an earthquake may be quite small (Table 1). In reducing the number of training landslides, we also consider the effect of their spatial distribution on the model results by comparing landslide sub-samples randomly distributed across the affected area with sub-samples that are clustered (Fig. 8). These conditions represent mapping efforts undertaken with both partial cloud cover (randomly

20 distributed sub-samples) and more regular cloud cover with intermittent gaps (clustered sub-samples). This allows us to test the effect of both the number and spatial distribution of training landslides ~~and their spatial distribution~~on model performance, as well as~~nd~~ consider ~~to which~~the weather conditions under which this approach may be best suited.

For the randomly distributed sub-samples, we iteratively select 500, 250, and 100 landslides from the original 2006 training landslides. We repeat this five times, producing five sub-samples that mimic the incremental mapping of landslides

25 following an earthquake. For the clustered sub-samples, we split the original training landslides into 10 separate clusters based on landslide proximity to each other. To test the effect of each sub-sample on the model, we calculate the distribution of I_{norm} as before for the seven predisposing factors for which membership functions have been derived. We then evaluate the R^2 fit of each sub-sample to the corresponding membership curves ($f(I_{norm})$) and compare the results with the R² values achieved for the original 2006 training landslides (Fig. 3).

30 For the majority of factors, we achieve R^2 values within 10% of the original using 500 randomly distributed landslides (Fig. 8). For all factors other than distance from fault plane and distance from confluences, we achieve R^2 values within 25% of the original when using 250 randomly distributed landslides, while both relief factors and planform curvature achieve the same result for 100 randomly distributed landslides (Fig. 8). However, for the clustered sub-samples there is little

Formatted: Font: Italic

Formatted: Font: Italic

consistency between factors in regards to the number of landslides and the R² values achieved. For the sub-hillslope relief, distance from rivers, and distance from confluences factors, the largest clusters have the greatest discrepancies in R² values compared to the original, while smaller clusters are able to achieve R² values within 10% of the original (Fig. 8).

This suggests that almost identical estimates of landslide hazard (Fig. 6), ~~magnitude~~intensity and spatial pattern (Fig. 7) could have been made using just several hundred training landslides, provided that these landslides had a sufficiently wide spatial distribution. Had these landslides been clustered then it is unlikely that similar membership curves, and thus hazard models, would have been derived. We therefore suggest that our model is most suited to conditions of partial cloud cover where the total visible area of ground, and therefore mapped landslides, is small but covers a wide area, or where large numbers of landslides are visible and thus the time to complete mapping is dependent on slow manual identification methods.

6 Discussion

6.1 Implications for post-earthquake emergency response decision support

This study has derived a model that can be applied rapidly post-earthquake to aid emergency response. Comparison of modelled relative landslide ~~magnitude~~intensity and pattern (Fig. 7~~e~~) with the location of training landslides (Fig. 1) reveals locations where landslide intensity was expected to be high, but landslides had, at the time, not been mapped. In particular, the model anticipates that landsliding in Gorkha and Dhading districts occurred at similar intensities to Sindhupalchok district despite very few training landslides in the former locations. Likewise, the model correctly identifies that the training inventory under-represents the ~~magnitude~~intensity of landsliding in Rasuwa district.

While the successful output landslide intensity map (Fig. 7) is coarse in detail, we argue that the output is useful for helping to inform emergency response planning. Overlaying the intensity map with population distributions and critical network data can quickly allow emergency responders to identify regions where landslides are expected to have caused losses and therefore require urgent aid.

These results can potentially be used to alert decision makers to the need for response in specific locations before landslides have been confirmed either remotely or on the ground (e.g., in Gorkha and Dhading). In regions badly affected by landslides, ground communications may be severely impacted, meaning the need for response can be overlooked until impacts, which might include landslides, are identified, potentially many days after the event. This study may therefore reduce this risk by enabling decision makers to identify regions likely to have sustained intense landsliding, potentially allowing them to target these areas for response before landslides are conclusively identified. Such an approach may lead to improving response times for badly affected and isolated regions following major earthquakes.

6.2 Potential speed of application

The method presented herein has been developed and applied after the fact and therefore it is necessary to discuss the potential speed with which it could be applied following a future earthquake. Firstly, this study has deliberately focussed on predisposing factors that can be rapidly and directly derived from a DEM, allowing the model to be applied with only minimal additional data requirements. Global DEM datasets at 30 m resolution are freely available from NASA and can be downloaded for the relevant region in a matter of minutes. Likewise, the USGS ShakeMap programme rapidly assesses earthquake parameters within minutes of the earthquake occurring.

We have also shown that it is possible to apply this method from only several hundred landslides (Fig. 8). While the mapping efforts described herein took up to 38 days or more to identify >4000 landslides (Kargel et al., 2016), the joint DU-BGS group had identified 279 landslides within 9 days of the mainshock (Table 1). This was despite the majority of the affected area suffering from cloud cover, suggesting that in events with less cloud cover this time could potential be reduced to just a few days. The proliferation of rapidly available crowdsourced data from unconventional but quickly available sources such as social media may also be used to inform this model (Earle et al., 2010; Bruns and Liang, 2012). Importantly however, the results herein suggest that when cloud free imagery is available for the majority of the affected region, systematically mapping from this imagery in considerable detail, as is currently done, is not required. Instead, mapping could be undertaken in far less detail using a random sample of grid squares covering the entire affected area, with the results then fed into the model described herein. Such an approach may allow landslide mapping and modelling to combine more effectively following an earthquake to inform emergency response operations.

Furthermore, various automated landslide mapping techniques are currently being developed (e.g. Booth et al., 2009; Borghuis et al., 2007; Li et al., 2016), and some methods are even capable of identifying landslides through cloud cover (Kimura and Yamaguchi, 2000). Such techniques can accurately map landslide locations much faster than traditional manual mapping, potentially reducing the time required to map sufficient numbers of landslides. However, whilst automated mapping offers many advantages, at present the range of imagery, the often complex topography, and the poor radiometric distinction between intact slopes and landslides means that manual mapping is often the most reliable approach, particularly in regions without extensive pre-earthquake research.

6.3 Other predisposing factors

There are several other notable factors that have been shown to influence coseismic landsliding elsewhere that have not been considered herein, notably bedrock lithology (Keefer, 2000; Parise and Jibson, 2000; Khazai and Sitar, 2004; Dai et al., 2011), structural geology (Hoek et al., 2005; Selby, 2005; Moore et al., 2009), and rainfall (Iverson, 2000; Dellow and Hancox, 2006). Whilst these factors are certainly important for landslide occurrence, obtaining sufficiently accurate data is often difficult and time-consuming. This is especially true for lower income countries such as Nepal, where the necessary data either may not exist or may not be readily accessible. Further, we have shown that accurate landslide models can be

created without the need for these factors, so long as other important factors, in particular slope angle, are considered (also see Pawluszek and Borkowski (2017) for a comparison of the role of topographic factors and non-topographic factors). Of course, if these data-sets do exist, then they can easily be incorporated into the factor analysis.

6.4 Benefits, limitations, and uncertainties

5 Several other models exist that allow rapid landslide modelling (Jibson et al., 2000; Godt et al., 2008; Nowicki et al., 2014; Kritikos et al., 2015) including the model of Gallen et al. (2016), which was applied following the Gorkha earthquake. The present study provides a useful complement to these existing models, as well as a more general understanding of factors influencing slope failure. One of the primary benefits of our model over others is its ability to accurately model both the spatial distribution and relative ~~magnitude~~intensity of landsliding. Gallen et al. (2016) were able to accurately model the total

10 number of landslides in the Gorkha earthquake but found large discrepancies between their predicted landslide pattern and the observed pattern. Incorrectly modelling the locations of landslides may have detrimental effects on emergency response by focussing responders on areas not requiring immediate response.

Our model also benefits from being trained on landslides known to have occurred during the event. Consequently, the approach allows predisposing factors and the corresponding membership curves~~model is~~ to be tailored to the specific

15 location and earthquake under consideration, rather than relying on global or regional relationships (e.g. Kritikos et al., 2015). No pre-event knowledge of landsliding in the affected area is necessary as the approach allows~~model establishes~~ the influence of various factors to be rapidly assessed post-earthquake. The model has been shown to be successful despite ~~potential~~ systematic bias in the initial landslide inventories, ~~such as~~with cloud cover ~~above or below specific elevations~~predominantly obscuring the major ridgelines. This suggests that ~~if the inventories are systematically biased, the~~

20 ~~results are unaffected.~~ despite the training inventories containing few landslides from major ridgelines, the model forecasts these cells as high hazard, successfully representing the numerous landslides in the test inventory on major ridgelines. This is inferred to be a result of the use of the planform curvature factor, which models high landslide influence in convex cells (i.e. ridgelines), and further highlights the issues with using $NRSD$ to mimic topographic amplification on ridgelines. The lack of pre-event data requirements is also beneficial, as it does not necessitate the collection and storage of global datasets. Where

25 data connections are limited or vulnerable, setting up geodatabases of topography and its derivatives alongside a platform to run models such as the one described prior to an earthquake would be a prudent earthquake preparedness measure.

However, there are important limitations in the model to consider. Firstly, the model is unable to estimate individual landslide area or volume, which becomes of vital importance at the local scale and for assessing post-earthquake impacts like sediment aggradation and flood hazard (e.g., Huang and Fan, 2013). Additional information on the size of landslides would

30 be beneficial, as larger landslides have greater runout and therefore potentially pose a greater hazard. Furthermore, the model only considers relative ~~magnitude~~intensity of landsliding, rather than absolute ~~magnitude~~intensity. Whilst this allows determination of the relative impacts between locations, it does not assess absolute impacts, which may be important. The success of the model is also highly dependent on the spatial distribution of the initial inventories, requiring that the training

landslides have a large spatial distribution with limited clustering. In conditions of widespread cloud cover this will be difficult to achieve as landslides can only be mapped through a few isolated gaps in the cloud, resulting in highly clustered training datasets. Ensuring a wide spatial coverage may also hinder the speed at which this approach can be applied by requiring the collection and processing of numerous satellite images. During the response to the Gorkha earthquake, the majority of time required to manually map landslides was taken up by downloading and processing the necessary images (Williams et al., forthcoming). Consequently, despite this approach requiring just several hundred training landslides, currently image collection and processing speeds are the major components influencing the speed of application. However, this is likely to improve with the increase in more medium resolution satellites, such as Sentinel 2, which can rapidly provide suitable resolution imagery for large areas, enabling modelling methods such as that described herein. Finally, as with all landslide hazard models, the accuracy of this method is reliant on the accuracy and resolution of the input data. Errors in locating landslides can result in incorrect estimates of I_{norm} and therefore the eventual hazard models. Similarly, the model is reliant on the resolution and accuracy of the input DEM from which all non-seismic influence factors are derived.

7 Conclusions

This study has addressed the need for more rapid assessments of coseismic landslide ~~magnitude~~intensity and distribution following a major earthquake. Manual mapping methods can be too slow to effectively inform emergency response operations and are too readily affected by issues with image collection and visibility. The present study has demonstrated an empirical method to model landslide ~~magnitude~~intensity and distribution using fuzzy logic based on initial inventories with small numbers of landslides. We have shown that landslides mapped in the first few days following an earthquake can be successfully used to assess the influence of various predisposing factors, provided that these landslides have sufficiently wide spatial distribution. The influence of these predisposing factors can be used to model the total pattern and relative ~~magnitude~~intensity of landsliding prior to total landslide inventories being completed using a simple threshold analysis. Applied to the 2015 Gorkha earthquake, our approach identifies: slope angle; distance from fault plane, rivers and river confluences; hillslope and sub-hillslope relief; and planform curvature as the key pre-disposing factors influencing landsliding during this event. The model output suggests that landsliding during the event was strongly anticipated in the major river valleys in Gorkha, Dhading, Rasuwa, and Sindhupalchok districts, agreeing with observed patterns of landsliding. Calculating the kernel density of cells in the 99th percentile of modelled hazard values is able to accurately and simultaneously represent both the spatial pattern and relative ~~magnitude~~intensity of landsliding during the event. Assessments of the minimum number of landslides required to achieve these results suggest that just several hundred landslides throughout the affected area need be identified for the method to be successfully applied. Consequently, this suggests that systematic high fidelity mapping of landslides following an earthquake may not be necessary for informing rapid modelling attempts. Instead, mapping should focus on rapid sampling from grid squares covering the entire affected area and using the results to inform the model herein. Doing so could potentially have allowed this model to have been

implemented within hours to days after the Gorkha earthquake occurred, permitting results to be available to emergency responders far earlier than allowed by traditional mapping techniques.

The method herein improves upon other approaches which have been shown to accurately simulate either landslide ~~magnitude~~intensity or distribution, but not both simultaneously. A key strength of our study is the use of open access data sources, allowing the model to be implemented globally without the need for pre-event collection and storage of data. This may therefore present a useful approach for rapidly assessing landslide hazard following an earthquake to effectively inform decision makers during emergency response operations.

Author contribution. TRR conceived and designed the model with contributions from NJR and ALD. JGW, MEK, JB, and HJAB undertook the collection and preparation of imagery and mapped the landslide data. TRR undertook the data analysis and modelling, and prepared the manuscript with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank the various agencies that made satellite imagery freely available through the International Disaster Charter which facilitated mapping following the earthquake. We also thank the members of the World Food Programme (Nepal), United Nations Office for Coordination of Humanitarian Affairs (UN-OCHA), International Committee of the Red Cross/Crescent (ICRC), and the UK Department for International Development (DfID) whose input in helping to define the information and timelines useful for informing emergency response greatly informed the development and approaches of this study. Pippa Whitehouse, Jing Zhu and an anonymous reviewer all provided helpful comments and feedback that helped to improve the manuscript. This study has been supported by the DIFeREns2 (2014-2019) COFUND scheme supported by the European Union's Seventh Framework Programme (grant number 609412).

References

Aleotti, P. and Chowdhury, R.: Landslide hazard assessment: summary review and new perspectives, B Eng Geol Environ, 58(1), 21–44, 1999.

Avouac, J. P., Meng, L., Wei, S., Wang, T. and Ampuero, J. P.: Lower edge of locked Main Himalayan Thrust unzipped by the 2015 Gorkha earthquake, Nat Geosci, 8(9), 708-711, 2015.

Bettinelli, P., Avouac, J. P., Flouzat, M., Jouanne, F., Bollinger, L., Willis, P. and Chitrakar, G. R.: Plate motion of India and interseismic strain in the Nepal Himalaya from GPS and DORIS measurements, J Geodesy, 80(8), 567-589, doi:10.1007/s00190-006-0030-3, 2006.

Booth, A. M., Roering, J. J. and Perron, J. T.: Automated landslide mapping using spectral analysis and high-resolution topographic data: Puget Sound lowlands, Washington, and Portland Hills, Oregon, Geomorphology, 109(3), 132-147, 2009.

Borghuis, A. M., Chang, K. and Lee, H. Y.: Comparison between automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery, Int J Remote Sens, 28(8), 1843-1856, 2007.

Bruns, A. and Liang, Y. E.: Tools and methods for capturing Twitter data during natural disasters, First Monday, 17(4), doi: 10.5210/fm.v17i4.3937, 2012.

Budimir, M. E. A., Atkinson, P. M. and Lewis, H. G.: Earthquake-and-landslide events are associated with more fatalities than earthquakes alone, Nat Hazards, 72(2), 895-914, 2014.

Bui, D. T., Pradhan, B., Lofman, O., Revhaug, I. and Dick, O. B.: Spatial prediction of landslide hazards in Hoa Binh province (Vietnam): a comparative assessment of the efficacy of evidential belief functions and fuzzy logic models, Catena, 96, 28-40, 2012.

Dai, F. C., Xu, C., Yao, X., Xu, L., Tu, X. B. and Gong, Q.M.: Spatial distribution of landslides triggered by the 2008 Ms 8.0 Wenchuan earthquake, China. J Asian Earth Sci, 40, 883–895. doi:10.1016/j.jseaes.2010.04.010, 2011.

Dellow, G. D. and Hancox, G. T.: The influence of rainfall on earthquake-induced landslides in New Zealand, in: Proceedings of Technical Groups, Earthquakes and Urban Development: New Zealand Geotechnical Society 2006 Symposium, Nelson, New Zealand, 355–368, 2006.

Densmore, A. L. and Hovius, N.: Topographic fingerprints of bedrock landslides, Geology, 28(4), 371-374, 2000.

Earle, P., Guy, M., Buckmaster, R., Ostrum, C., Horvath, S. and Vaughan, A.: OMG earthquake! Can Twitter improve earthquake response?, Seismol Res Lett, 81(2), 246-251, 2010.

Earthquakes without Frontiers (EwF): Nepal earthquake: update on landslide hazard (8 May), http://ewf.nerc.ac.uk/2015/05/08/nepal-earthquake-update-on-landslide-hazard-2/, last access: 5 June 2017, 2015.

ESRI: How Fuzzy Overlay works, http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-fuzzy-overlay-works.htm, last access: 6 June 2017, 2016.

Gallen, S. F., Clark, M. K. and Godt, J. W.: Coseismic landslides reveal near-surface rock strength in a high relief, tectonically active setting, Geology, doi: 10.1130/g36080.1, 2015.

Gallen, S. F., Clark, M. K., Godt, J. W., Roback, K. and Niemi, N. A.: Application and evaluation of a rapid response earthquake-triggered landslide model to the 25 April 2015 Mw 7.8 Gorkha earthquake, Nepal, Tectonophysics, doi: 10.1016/j.tecto.2016.10.031, 2016.

Godt, J. W., Baum, R. L., Savage, W. Z., Salciarini, D., Schulz, W. H. and Harp, E. L.: Transient deterministic shallow landslide modeling: requirements for susceptibility and hazard assessments in a GIS framework, Eng Geol, 102, 214–226, doi: 10.1016/j.enggeo.2008.03.019, 2008.

Hoek, E., Carranza-Torres, C. and Corkum, B.: Hoek–Brown Failure Criterion – 2002 Edition, in: Proceedings of NARMS-Tac Conference, Toronto, 267–273, 2002.

Huang, R. and Fan, X.: The landslide story, Nat Geosci, 6(5), 325-326, 2013.

Hubbard, J., Almeida, R., Foster, A., Sapkota, S. N., Bürgi, P. and Tapponnier, P.: Structural segmentation controlled the 2015 Mw 7.8 Gorkha earthquake rupture in Nepal, Geology, 44(8), 639-642, 2016.

Iverson, R. M.: Landslide triggering by rain infiltration, Water Resour Res, 36, 1897–1910, 2000.

Jibson, R. W., Harp, E. L. and Michael, J. A.: A method for producing digital probabilistic seismic landslide hazard maps, Eng Geol, 58, 271–289. doi: 10.1016/S0013-7952(00)00039-9, 2000.

Keefer, D. K.: Landslides caused by earthquakes, Bull Geol Soc Am, 95, 406–421, 1984.

Keefer, D. K.: Statistical analysis of an earthquake-induced landslide distribution—the 1989 Loma Prieta, California event, Eng Geol, 58(3), 231-249, 2000.

Khazai, B. and Sitar, N.: Evaluation of factors controlling earthquake-induced landslides caused by Chi-Chi earthquake and comparison with the Northridge and Loma Prieta events, Eng Geol, 71(1), 79-95, 2004.

Kimura, H. and Yamaguchi, Y.: Detection of landslide areas using satellite radar interferometry, Photogramme Eng Rem S, 66(3), 337-344, 2000.

Kritikos, T. and Davies, T.: Assessment of rainfall-generated shallow landslide/debris-flow susceptibility and runout using a GIS-based approach: application to western Southern Alps of New Zealand, Landslides, 12(6), 1051-1075, 2015.

Kritikos, T., Robinson, T. R. and Davies, T.R.: Regional coseismic landslide hazard assessment without historical landslide inventories: A new approach, J Geophys Res-Earth, 120(4), 711-729, 2015.

Lee, C. T., Huang, C. C., Lee, J. F., Pan, K. L., Lin, M. L. and Dong, J. J.: Statistical approach to storm event-induced landslides susceptibility, Nat Hazard Earth Sys, 8(4), 941-960, 2008.

Li, Z., Shi, W., Myint, S. W., Lu, P. and Wang, Q.: Semi-automated landslide inventory mapping from bitemporal aerial photographs using change detection and level set method, Remote Sens Environ, 175, 215-230, 2016.

Marc, O., Hovius, N., Meunier, P., Uchida, T. and Hayashi, S.: Transient changes of landslide rates after earthquakes. Geology, 43(10), 883-886, 2015.

Martha, T. R., Roy, P., Mazumdar, R. Govindharaj, K. B. and Kumar, K. V.: Spatial characteristics of landslides triggered by the 2015 M_w 7.8 (Gorkha) and M_w 7.3 (Dolakha) earthquakes in Nepal. Landslides. doi:10.1007/s10346-016-0763-x, 2016.

Metz, C. E.: Basic principles of ROC analysis, Semin Nucl Med, 8, 283-298, 1978

Meunier, P., Hovius, N. and Haines, A. J.: Regional patterns of earthquake-triggered landslides and their relation to ground motion, Geophys Res Lett, 34, L20408, doi: 10.1029/2007GL031337, 2007.

Meunier, P., Hovius, N. and Haines, A. J.: Topographic site effects and the location of earthquake induced landslides, Earth Planet Sc Lett, 275(3), 221-232, 2008.

Moore, J. R., Gischig, V., Amann, F., Hunziker, M. and Burjanek, J.: Earthquake-triggered rock slope failures: Damage and site effects, in: Proceedings of the 11th International and 2nd North American Symposium on Landslides, Banff, Canada, 3-8 June 2012, 2012.

Nepal Army: The Nepalese Army in the aftermath of the Gorkha earthquake of 2015 (Experiences and Lessons learned), Nepal Army HQ, Kathmandu, 2016.

Newmark, N. M.: Effects of earthquakes on dams and embankments, Géotechnique, 15, 139-159, 1965.

Nowicki, M. A., Wald, D. J., Hamburger, M. W., Hearne, M. and Thompson, E. M.: Development of a globally applicable model for near real-time prediction of seismically induced landslides, Eng Geol, 173, 54–65, doi:10.1016/j.enggeo.2014.02.002, 2016.

Parise, M. and Jibson, R. W.: A seismic landslide susceptibility rating of geologic units based on analysis of characteristics of landslides triggered by the 17 January 1994 Northridge, California earthquake, Eng Geol, 58, 251–270, 2000.

Parker, R. N., Hancox, G. T., Petley, D. N., Massey, C. I., Densmore, A. L. and Rosser, N. J.: Spatial distributions of earthquake-induced landslides and hillslope preconditioning in the northwest South Island, New Zealand, Earth Surface Dynamics, 3(4), 501, 2015.

Pawluszek, K. and Borkowski, A.: Impact of DEM-derived factors and analytical hierarchy process on landslide susceptibility mapping in the region of Rożnów Lake, Poland, Nat Hazards, 86(2), 919-952. doi: 10.1007/s11069-016-2725-y, 2017.

Pourghasemi, H. R., Pradhan, B. and Gokceoglu, C.: Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran, Nat Hazards, 63(2), 965-996, 2012.

Pradhan, B.: Use of GIS-based fuzzy logic relations and its cross application to produce landslide susceptibility maps in three test areas in Malaysia, Environ Earth Sci, 63(2), 329-349, 2011.

Roback, K., Clark, M. K., West, A. J., Zekkos, D., Li, G., Gallen, S. F., Chamlagain, D. and Godt, J.F.: The size, distribution, and mobility of landslides caused by the 2015 Mw 7.8 Gorkha earthquake, Nepal, Geomorphology. doi: 10.1016/j.geomorph.2017.01.030, 2017.

Robinson, T. R. and Davies, T. R. H.: Review article: potential geomorphic consequences of a future great (Mw= 8.0+) Alpine Fault earthquake, South Island, New Zealand. Nat Hazard Earth Sys, 13(9), 2279, 2013.

Schmidt, K. M. and Montgomery, D. R.: Limits to relief. Science, 270, 617-620, 1995.

Selby, M. J.: Hillslope Materials and Processes, Oxford University Press, Oxford, UK, 451 pp, 2005.

Wald, D. J., Worden, B. C., Quitoriano, V. and Pankow, K. L.: ShakeMap manual: technical manual, user's guide, and software guide (No. 12-A1), 2005.

5 Wang, H., Liu, G., Xu, W. and Wang, G.: GIS-based landslide hazard assessment: an overview, Prog Phys Geog, 29(4), 548–567, 2005.

Zweig, M. H. and Campbell, G.: Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine, Clin Chem, 39, 561-577, 1993.

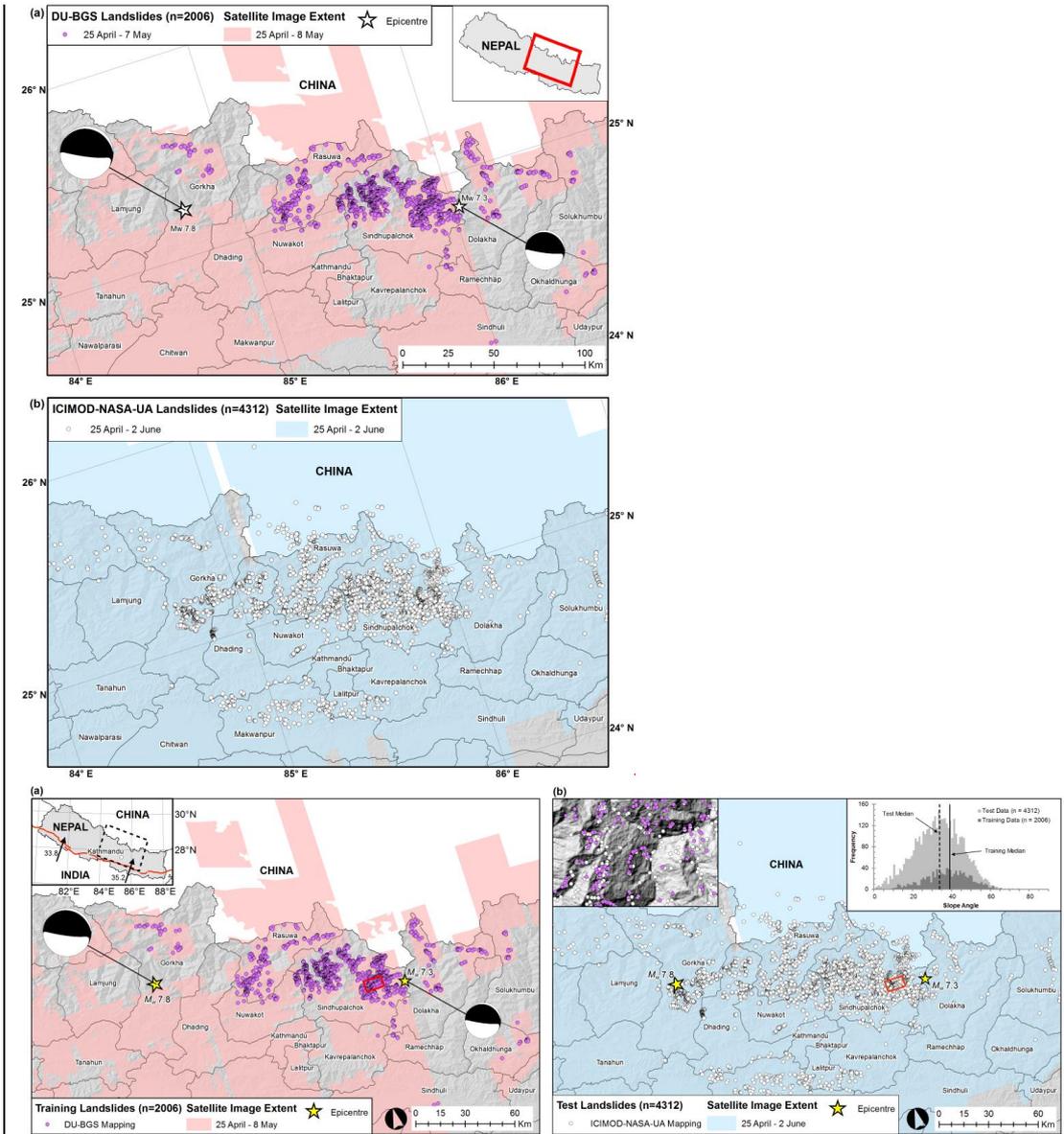



Figure 1: ~~Training and test l~~Landslide inventories from different mapping efforts following the 2015 Gorkha earthquake. (a) Training landslides identified by the joint DU-BGS group using satellite imagery available to 7 May 2015. Inset: study area location within Nepal showing main plate boundary in red and relative plate motion vectors in mm/yr from Bettinelli et al., (2006).

Focal mechanisms are taken from the Global CMT catalog.; **(b) test landslides identified by the joint ICIMOD-NASA-UA group using satellite imagery available to 2 June 2015. Inset: direct comparison between locations of training and test landslides. Graph shows the frequency of landslides in the training and test datasets compared to slope angle: training landslides occur on mean slope angle of 40°; test landslides occur on mean slope angle of 34°. Focal mechanisms are taken from the Global CMT catalog.**

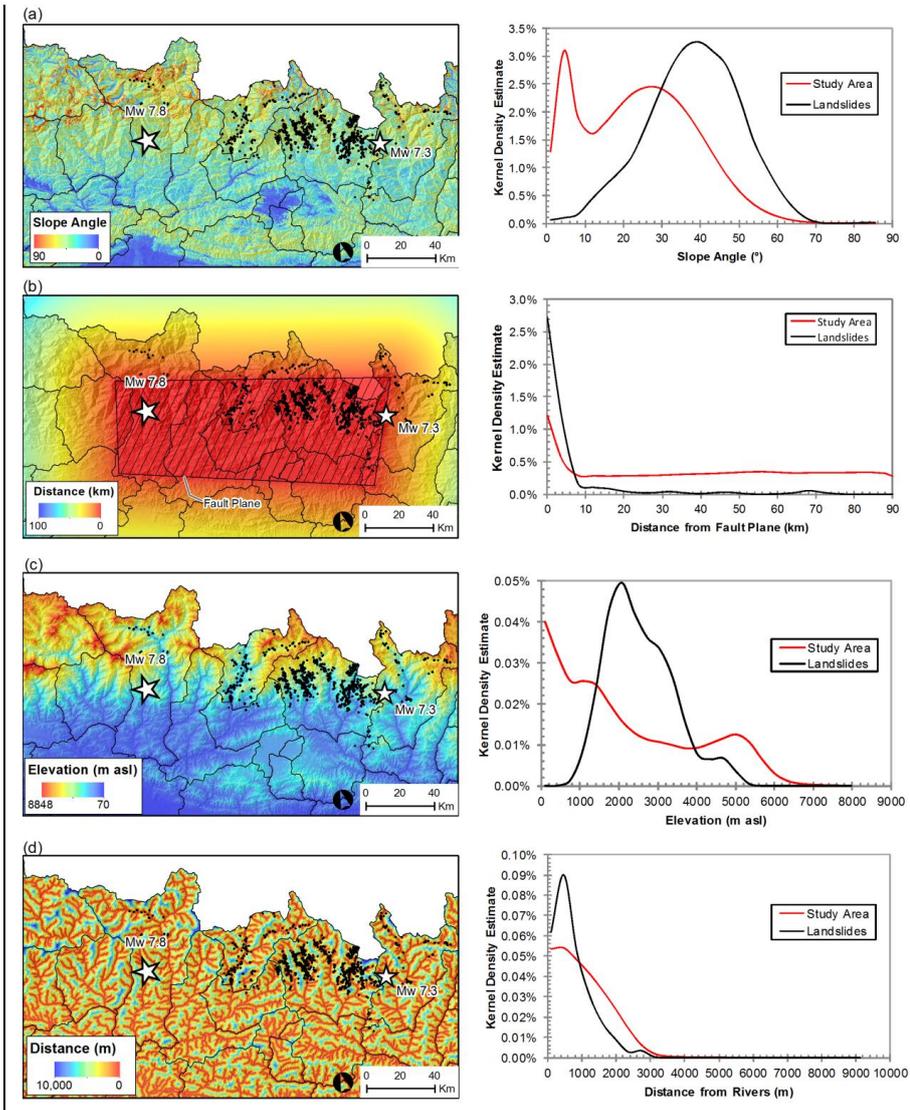
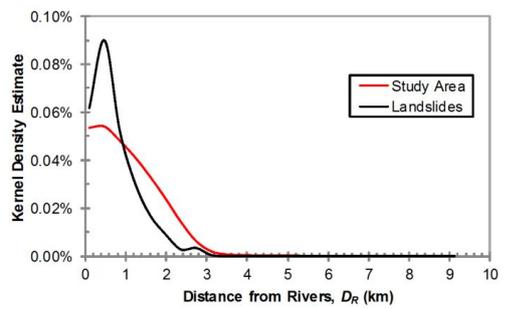
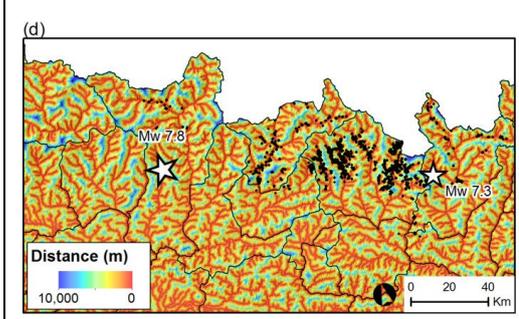
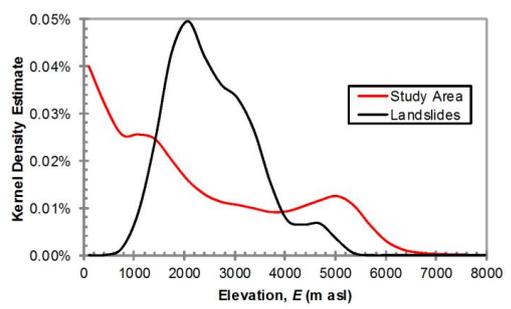
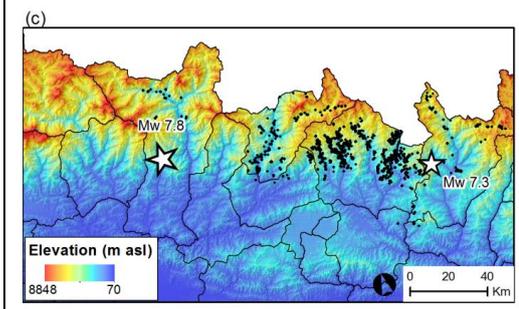
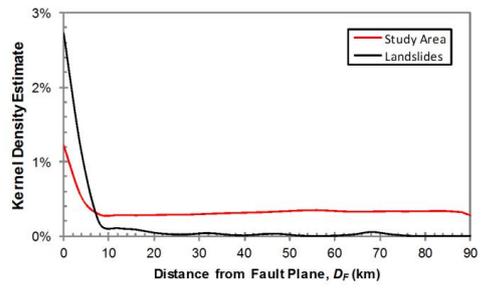
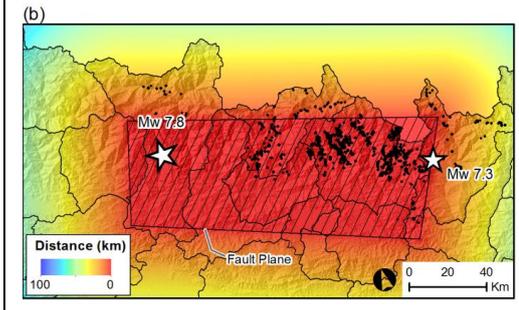
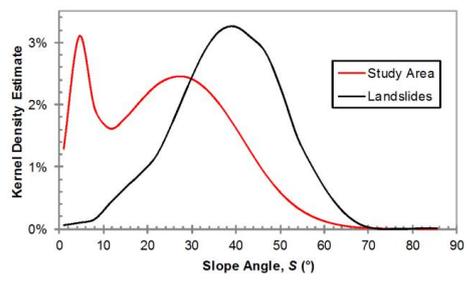
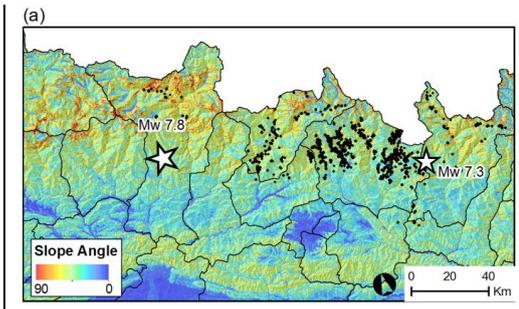
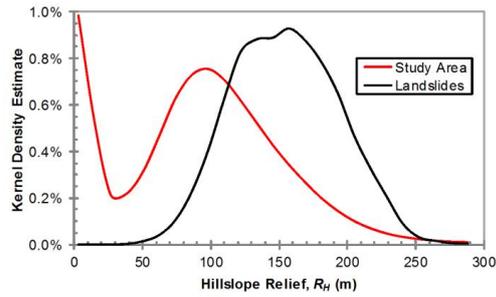
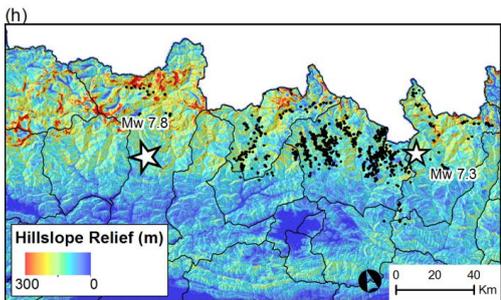
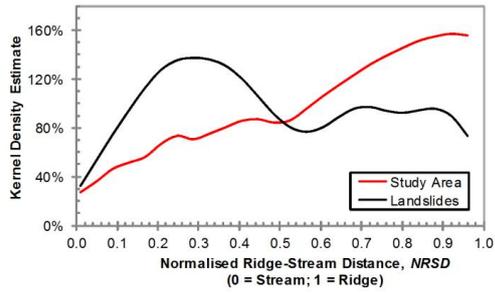
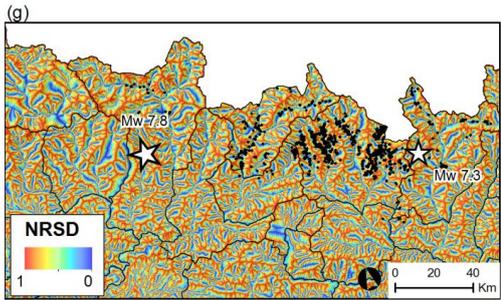
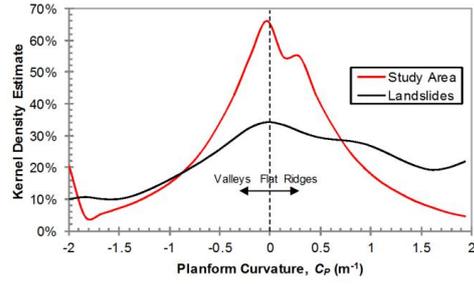
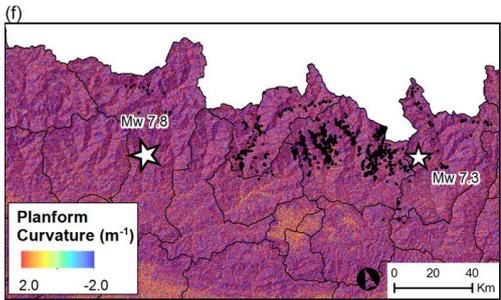
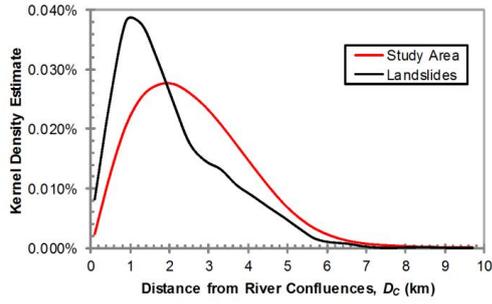
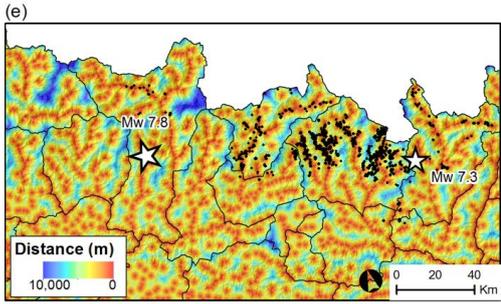
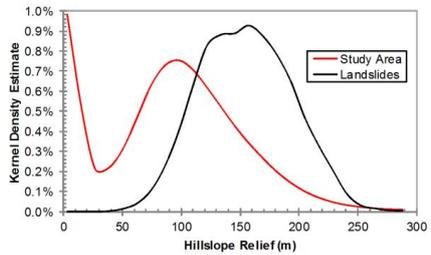
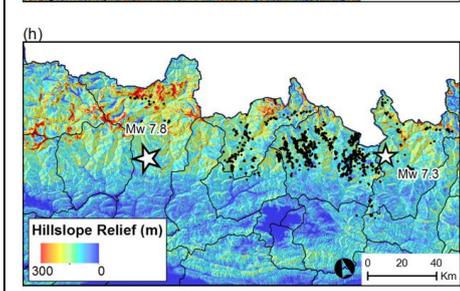
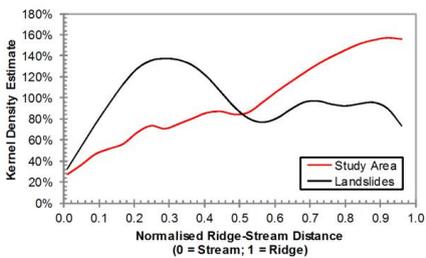
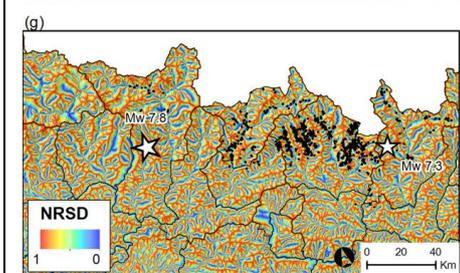
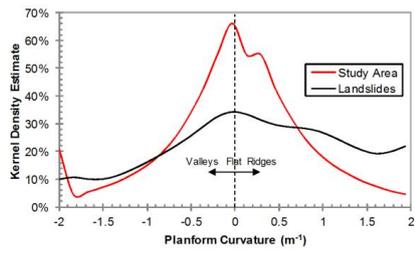
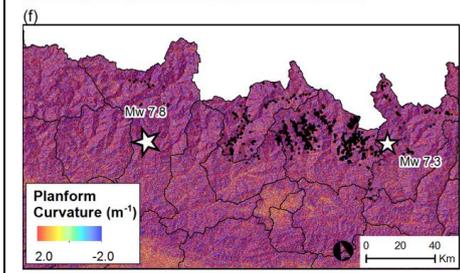
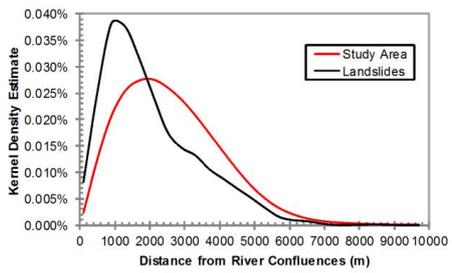
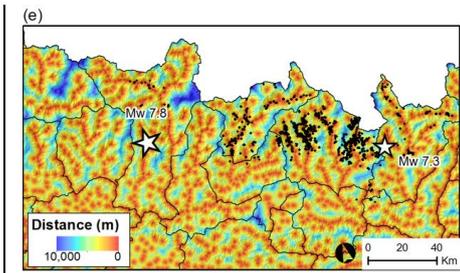
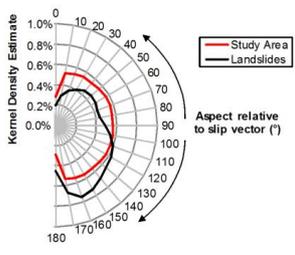
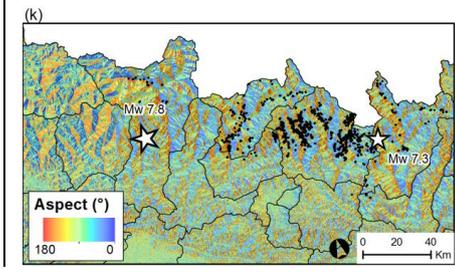
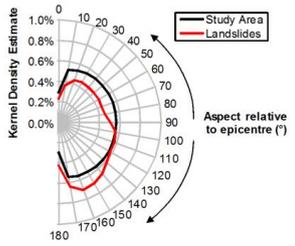
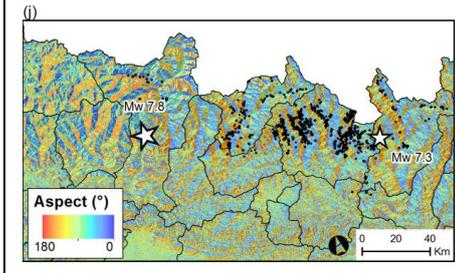
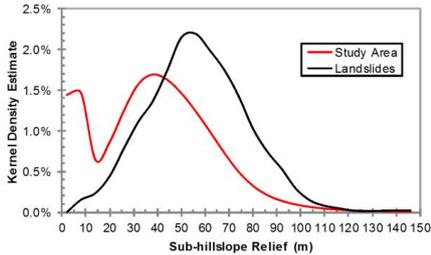
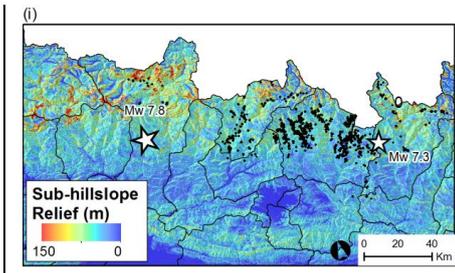
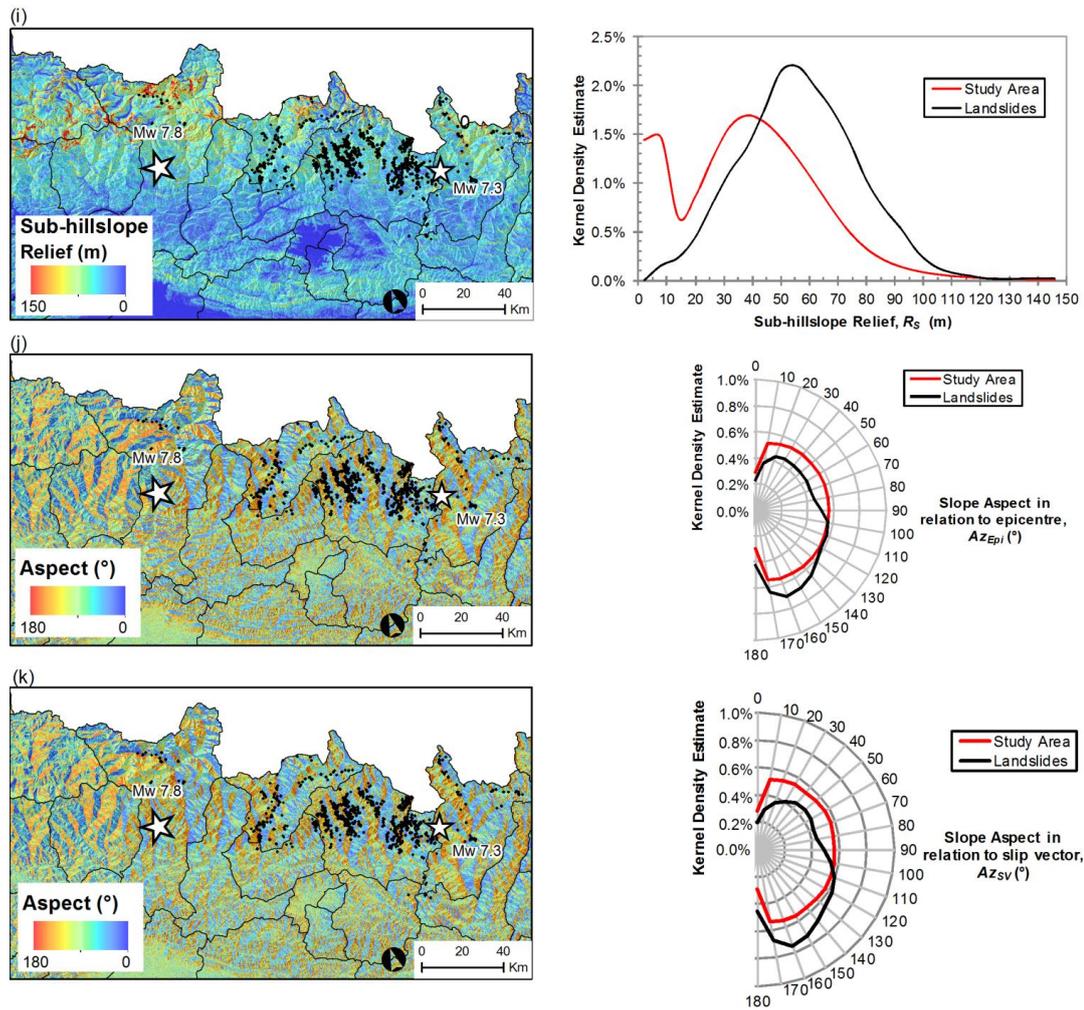



Figure 2: Kernel density estimates of training landslides (black dots; n=2006) compared to the entire study area for different predisposing factors. (a) Slope angle; (b) Euclidean distance from fault plane; (c) elevation; (d) Euclidean distance from rivers; (e) Euclidean distance from river confluences; (f) planform curvature; (g) normalised ridge-stream distance; (h) hillslope relief; (i) sub-hillslope relief; (j) slope aspect in relation to azimuth to epicentre; and (k) slope aspect in relation to slip vector.

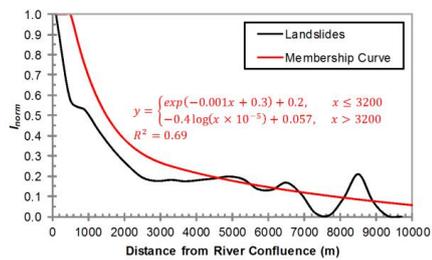
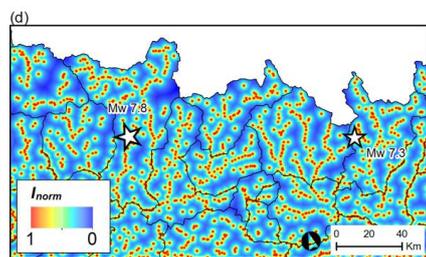
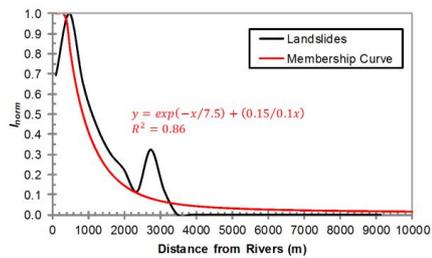
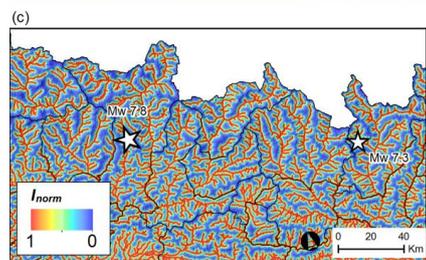
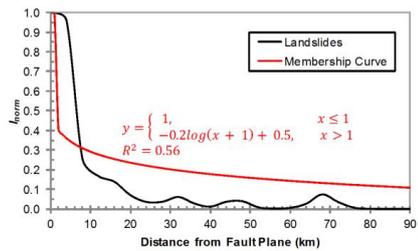
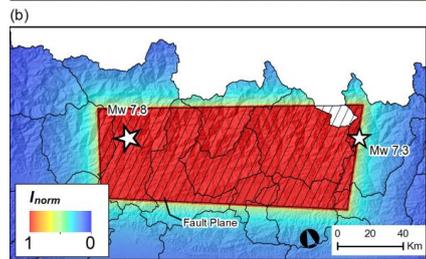
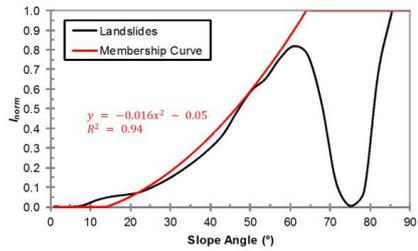
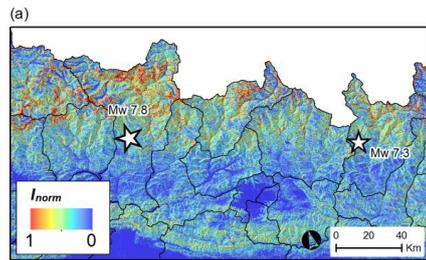
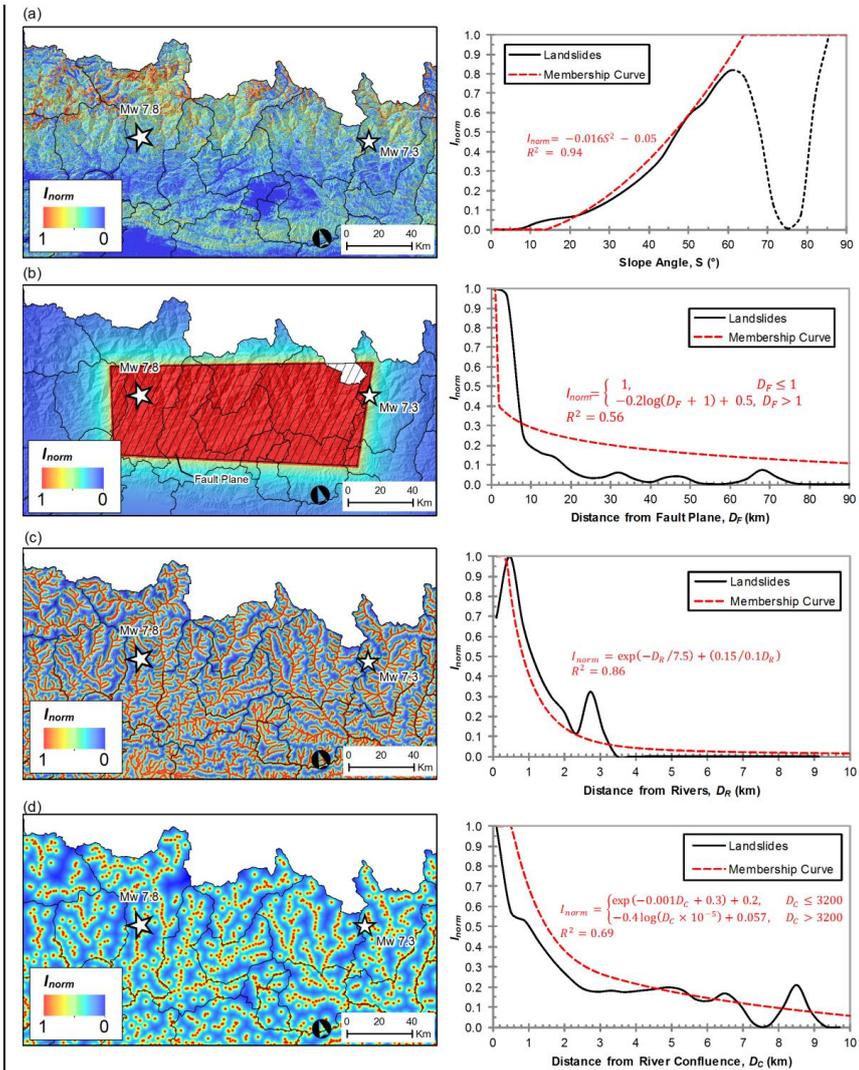
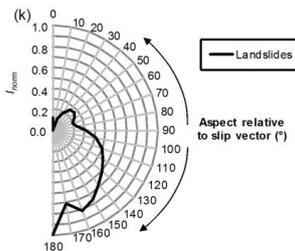
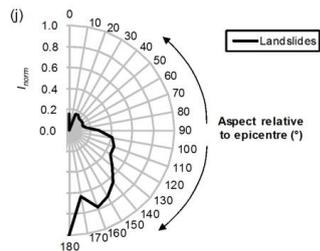
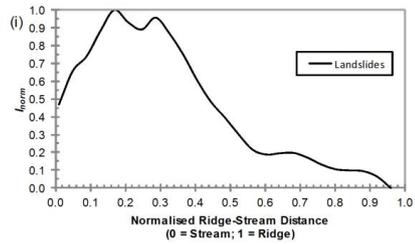
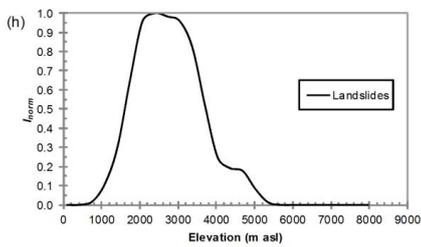
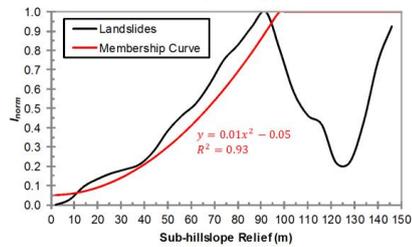
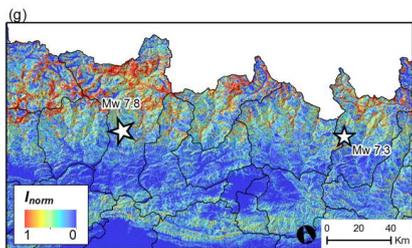
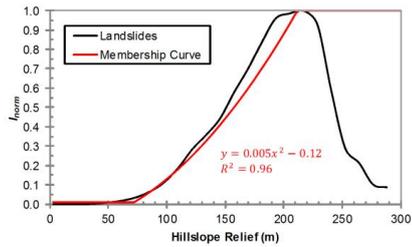
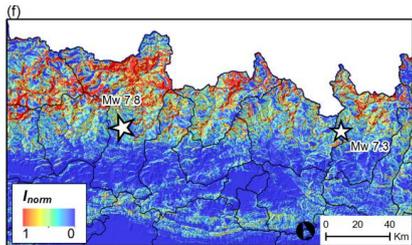
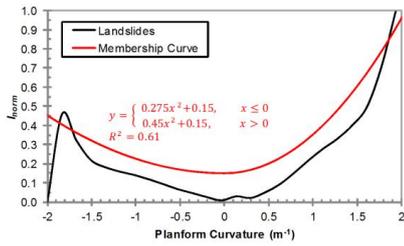
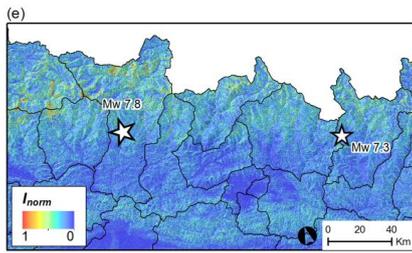
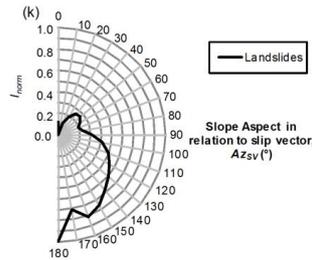
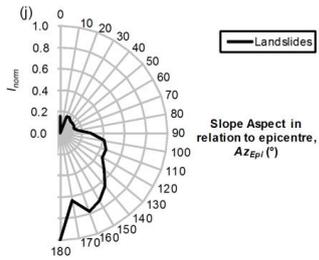
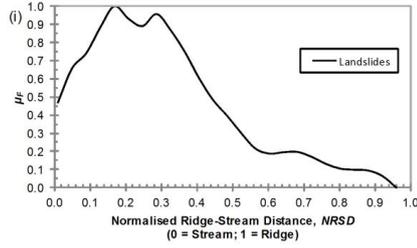
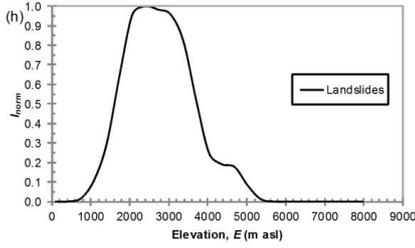
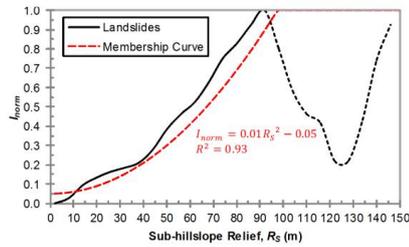
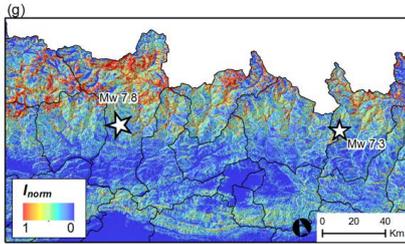
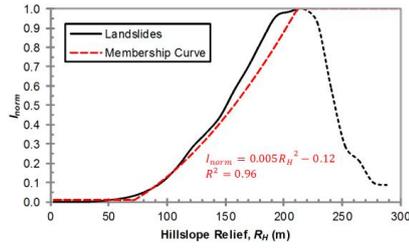
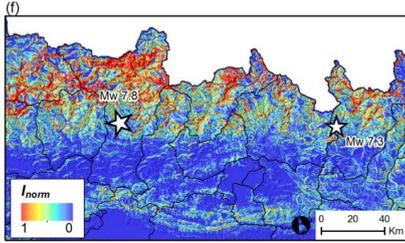
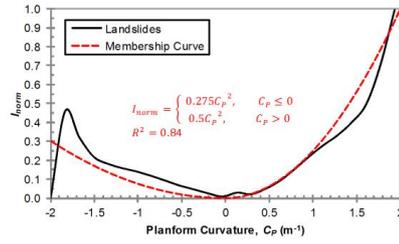
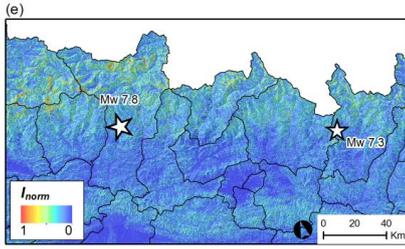



Figure 3: Predisposing factor influence on landslide occurrence. Panels a-g show influence maps, normalised kernel density estimates and membership functions, $f(I_{norm})$, for factors with ~~mechanically definable influence~~ relationships that match the broad physical expectation: (a) Slope angle; (b) distance from fault plane; (c) distance from rivers; (d) distance from river confluences; (e) planform curvature; (f) hillslope relief; and (g) sub-hillslope relief. Short-dashed sections of normalised kernel density estimates are not used in establishing the R² values due to the small number of cells representing these values. Panels h-k show factors ~~with no mechanically discernible~~ with relationships does not match physical expectations and cannot otherwise be explained ~~influence~~: (h) elevation; (i) normalised stream-ridge distance (NSRD); (j) aspect relative to epicentre; and (k) aspect relative to slip vector.

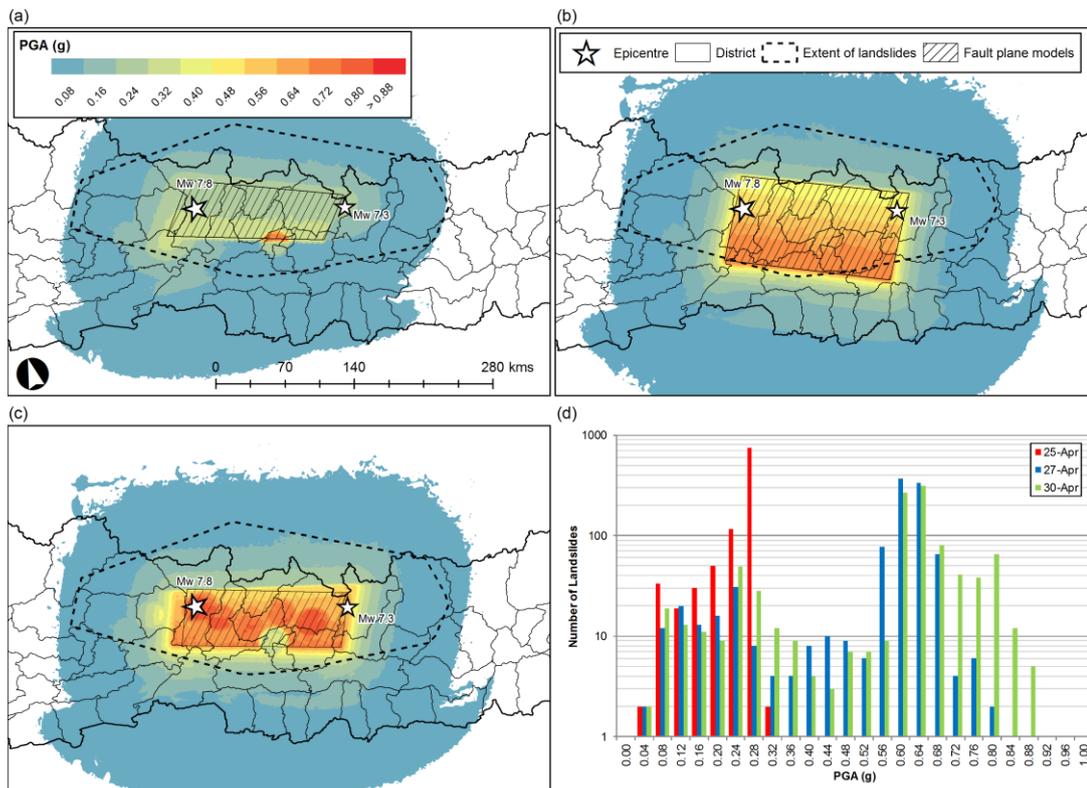



Figure 4: Changes in modelled peak ground acceleration and fault plane following the 2015 Gorkha earthquake. (a) 25 April 2015 (0 days after earthquake); (b) 27 April 2015 (2 days after earthquake); (c) 30 April 2015 (5 days after earthquake); and (d) frequency distribution of training landslides (n=2006) for each PGA model. Data from USGS (http://earthquake.usgs.gov/earthquakes/shakemap/).

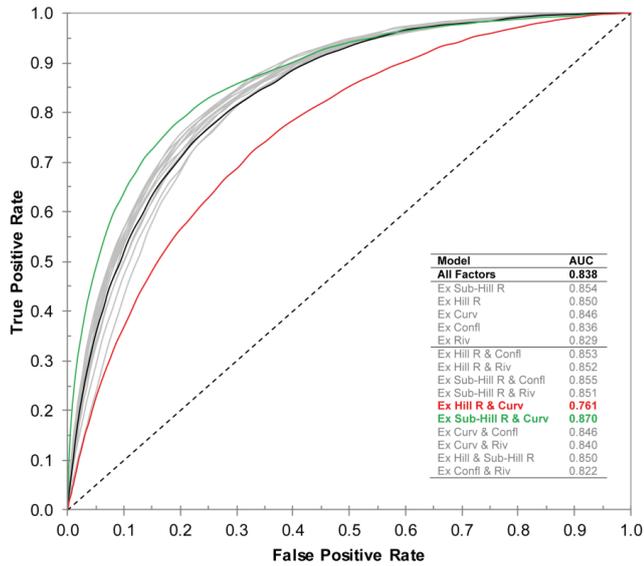
Model	AUC
All Factors	**0.838**
Ex Sub-Hill R	0.854
Ex Hill R	0.850
Ex Curv	0.846
Ex Confl	0.836
Ex Riv	0.829
Ex Hill R & Confl	0.853
Ex Hill R & Riv	0.852
Ex Sub-Hill R & Confl	0.855
Ex Sub-Hill R & Riv	0.851
Ex Hill R & Curv	0.761
Ex Sub-Hill R & Curv	0.870
Ex Curv & Confl	0.846
Ex Curv & Riv	0.840
Ex Hill & Sub-Hill R	0.850
Ex Confl & Riv	0.822

—— All other combinations ——— All Factors
——— Best performing (Ex Sub-hill R & Curv) ——— Worst performing (Ex Hill R & Curv)

Figure 5: ROC Curves for various different combinations of predisposing factors calculated from the test landslides (n=4312) mapped by the joint ICIMOD-NASA-UA group up to 2 June 2015. Black line in plot and black text in legend show the model with all factors; green shows the best performing combination; red shows the worst performing combination; grey shows all other combinations. Ex – Excluding; Sub-Hill R – sub-hillslope relief; Hill R – hillslope relief; Curv – planform curvature; Confl – distance to confluences; Riv – distance to rivers.

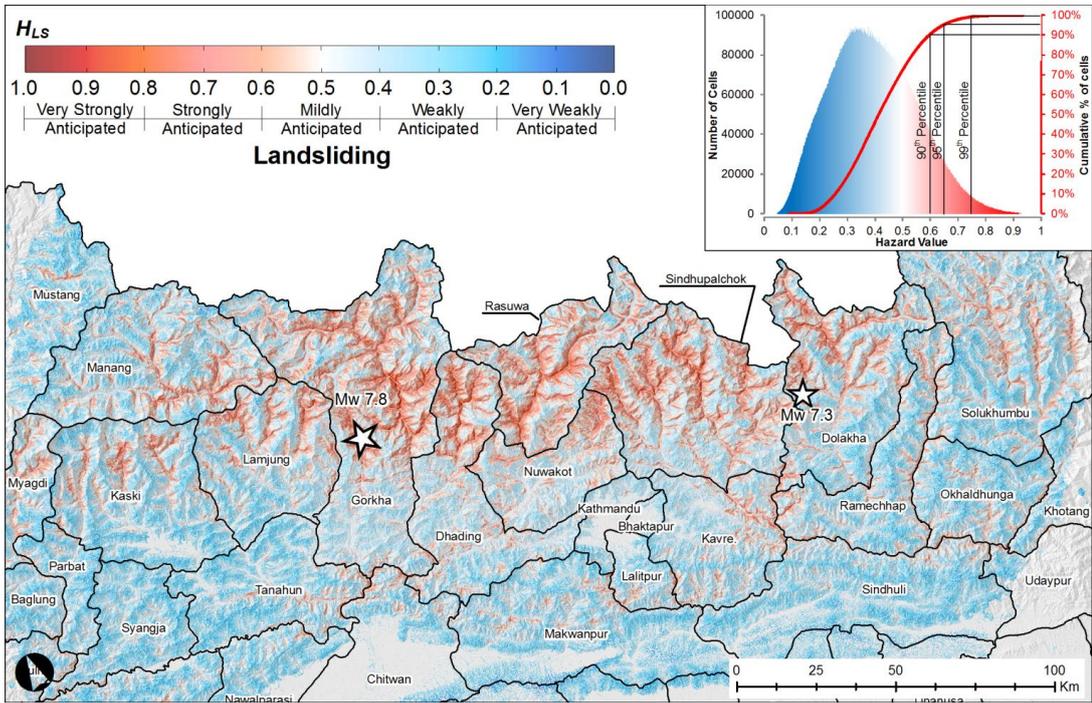



Figure 6: Landslide hazard resulting from the most successful factor combination (Fig. 5) and distribution of hazard values (inset) showing hazard values corresponding to the 99th, 95th, and 90th percentile of hazard area.

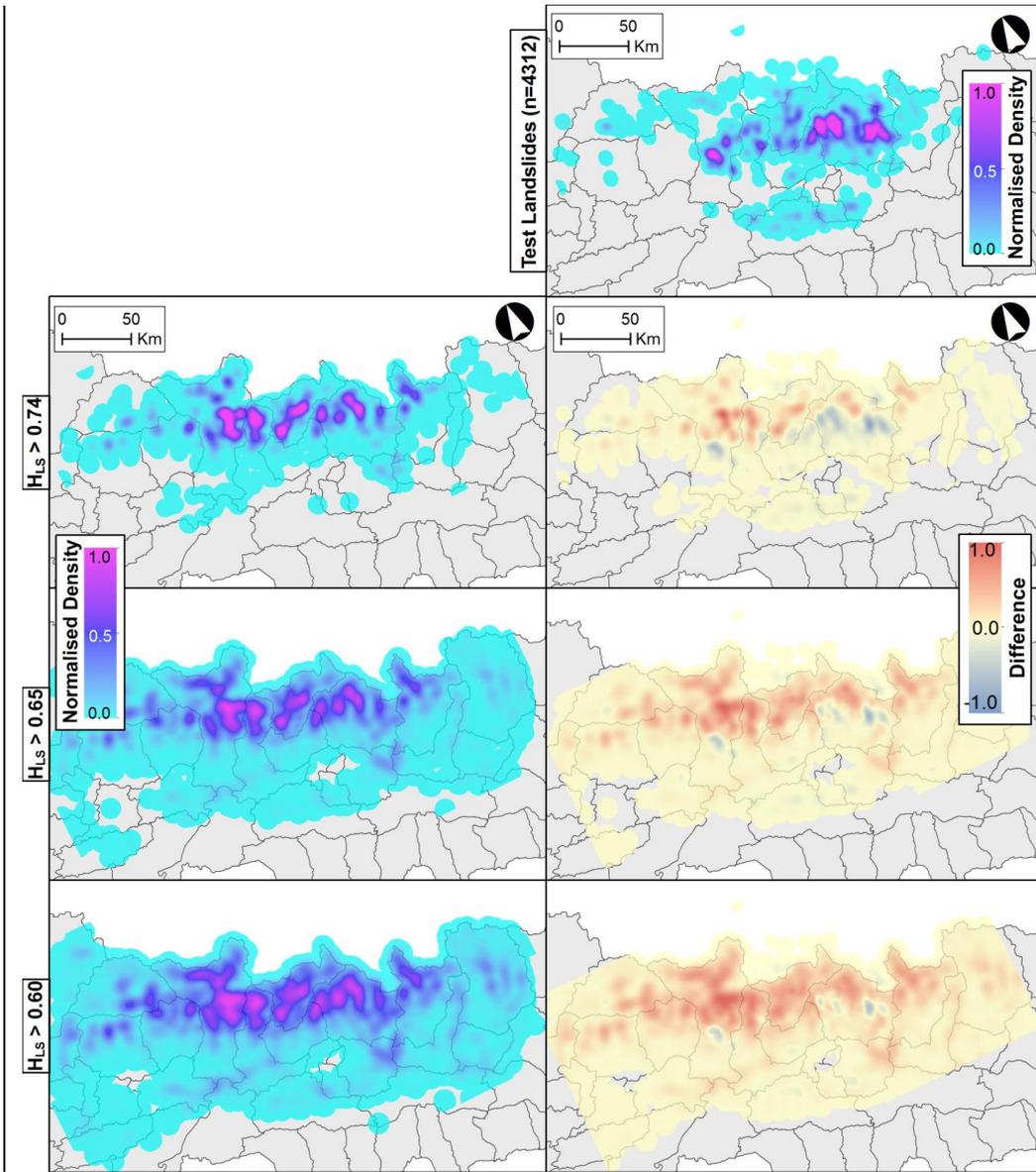
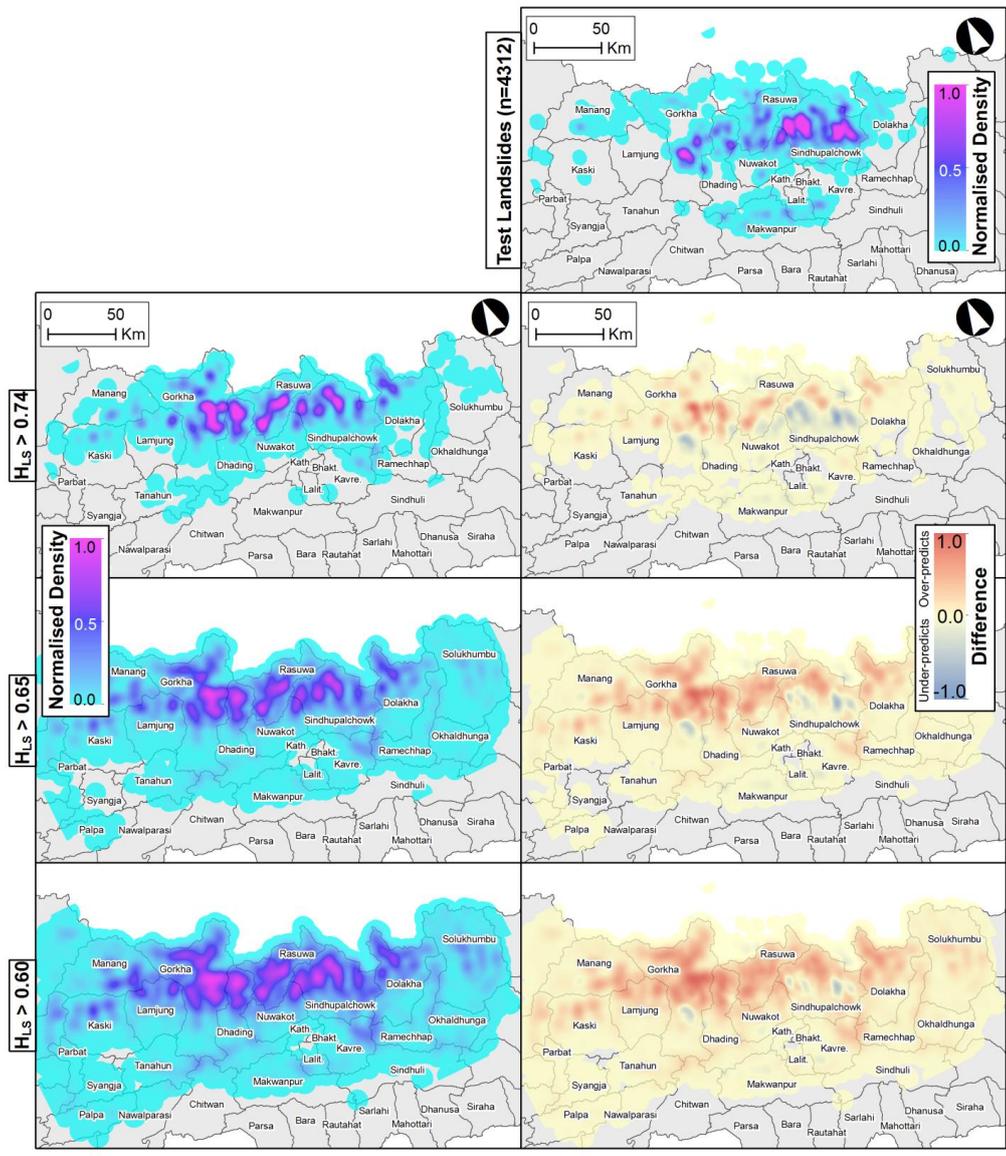



Figure 7: Normalised landslide ~~magnitude~~intensity (number density per unit area) and spatial pattern calculated as kernel density of pixels exceeding various H_{LS} values compared to test landslides for cells with $H_{LS} > 0.74$ (99th percentile), $H_{LS} > 0.65$ (95th percentile), $H_{LS} > 0.60$ (90th percentile).

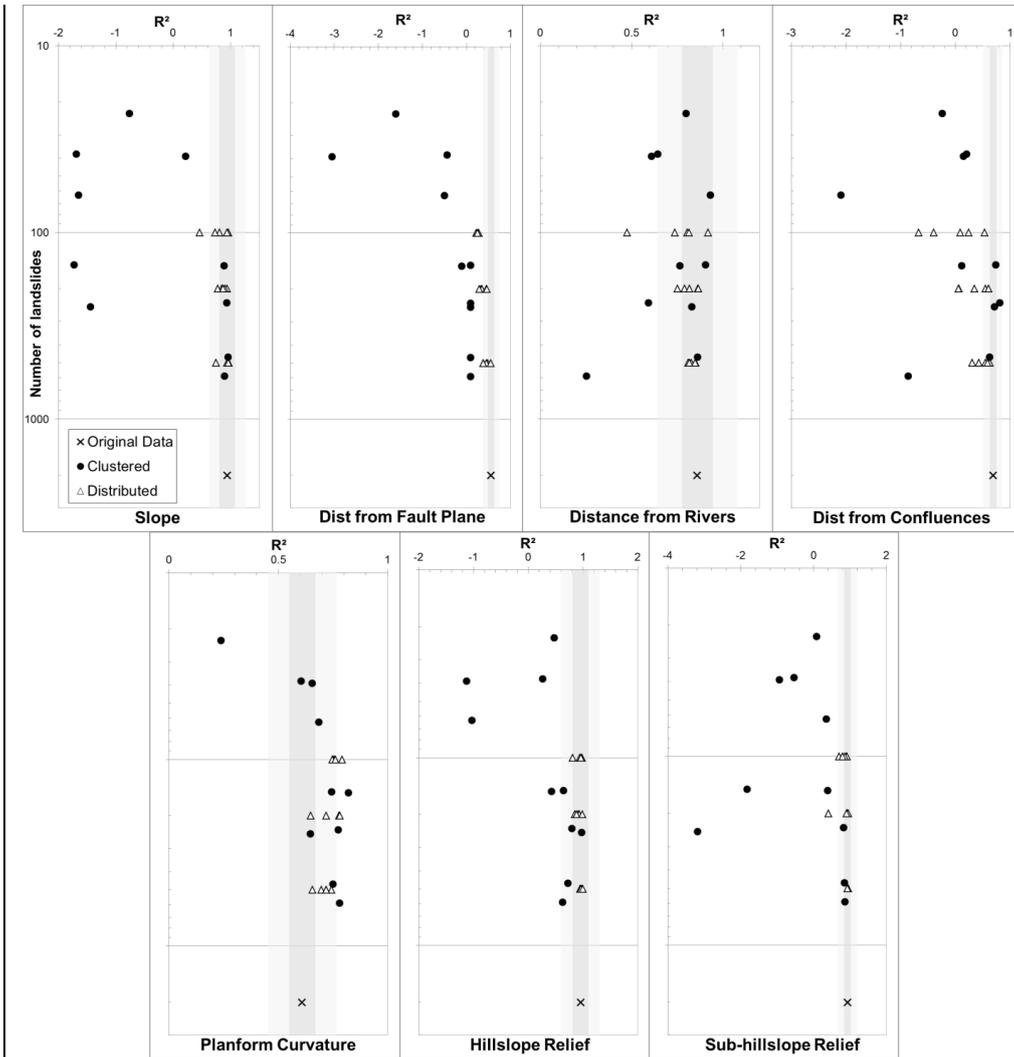
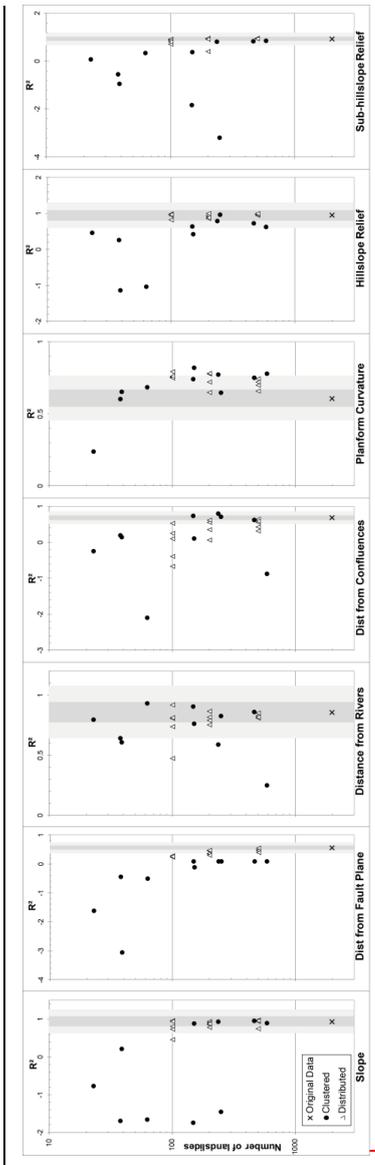



Figure 8: Effect of using smaller sub-sampled landslide inventories, either clustered or distributed across the entire affected area, on R^2 fit of membership curves for different predisposing factors (Fig. 3). Dark grey box shows values within 10% of original R^2 value; Light grey box shows values within 25% of original R^2 value.

Table 1. Total number of landslides identified by date by the joint DU-BGS group following the 2015 Gorkha earthquake.

Date (dd.mm.yy)	Number of days since mainshock (inclusive)	Number of landslides identified
25.04.15	0	0
04.05.15	9	279
07.05.15	12	2006
21.05.15	26	3600
19.06.15	55	5600