Natural Hazards (2006) 37: 331–372
DOI 10.1007/s11069-005-6141-y


# Predicting Lahar-Inundation Zones: Case Study in West Mount Pinatubo, Philippines

EMMANUEL JOHN M. CARRANZA[1,★] and OFELIA T. CASTRO[2]
[1]*Department of Earth Systems Analysis, International Institute for Geo-Information Science and Earth Observation (ITC), Hengelosestraat 99, P.O. Box 67500, AA Enschede, The Netherlands;* [2]*Mapping Department, National Mapping and Resource Information Authority (NAMRIA), Makati, Philippines*

(Received: 2 May 2003; accepted: 14 April 2005)

**Abstract.** This paper demonstrates techniques for pre-eruption prediction of lahar-inundation zones in areas where a volcano has not erupted within living memory and/or where baseline geological information about past lahars could be scarce or investigations to delimit past lahars might be incomplete. A lahar source (or proximal lahar-inundation) zone is predicted based on ratio of vertical descent to horizontal run-out of eruptive deposits that spawn lahars. Immediate post-eruption distal lahar-inundation zones are predicted based on "pre-eruption" distal lahar-inundation zones and on spatial factors derived from a digital elevation model. Susceptibility to distal lahar-inundation is estimated by weights-of-evidence, by logistic regression and by evidential belief functions. Predictive techniques are applied using a geographic information system and are tested in western part of Pinatubo volcano (Philippines). Predictive maps are compared with a forecast volcanic-hazard map through validation against a field-based volcanic-hazard map. The predictive model of proximal lahar-inundation zone has "true positive" prediction accuracy, "true negative" prediction accuracy, "false positive" prediction error and "false negative" prediction error that are similar to those of the forecast volcanic-hazard map. The predictive models of distal lahar inundation zones have higher "true positive" prediction accuracy and lower "false negative" prediction error than the forecast volcanic-hazard map, although the latter has higher "true negative" prediction accuracy and lower "false positive" prediction error than the former. The results illustrate utility of proposed predictive techniques in providing geo-information could be used, howbeit with caution, for planning to mitigate potential lahar hazards well ahead of an eruption that could generate substantial source materials for lahar formation.

**Key words:** lahars, predictive modeling, weights-of-evidence, logistic regression, evidential belief functions, DEM, GIS

## 1. Introduction

On volcanic landscapes, lahars develop from mixtures of meteoric water and rock debris (Vallance, 2000). Lahars are enormously destructive such that mitigation of lahar hazards is a non-trivial task. In general, hazardous conditions (e.g., volcanic eruptions, heavy rainfall) that favor lahar

★ Author for correspondence: Fax: +31-53-4874336; E-mail: carranza@itc.nl

formation could be recognized and used as warnings that aid mitigation of lahar hazards (e.g., Lecointre *et al.*, 2004). However, time between occurrence of hazardous conditions that favor lahar formation and occurrence of lahars could be insufficient to establish lahar mitigation plans (e.g., Newhall and Punongbayan, 1996). Thus, it is certainly more desirable to develop lahar hazard mitigation plans well ahead of phenomena that favor lahar formation. In such a case, maps depicting probable lahar-inundation zones can be derived through various predictive techniques, which makes use of geological knowledge or information as well as practical experiences gained from previous lahar occurrences.

Knowledge of lahar rheology is essential in creating lahar flow-routing models (e.g., Laenen and Hansen, 1988; Vignaux and Weir, 1990; Macedonio and Pareschi, 1992; Costa, 1997). However, flow-routing models require that input parameters based on lahar rheology should be known in advance to predict behavior of lahars (Pareschi, 1996). In addition, calibration of flow-routing models for efficient interpolation/extrapolation of lahar prone areas requires detailed field data and laboratory analyses (e.g., Canuti *et al.*, 2002). Hence, where data are inadequate to develop robust flow-routing models, alternative methods are needed to predict lahar-inundation zones.

Information about lahar geometry (i.e., cross-sectional areas, planimetric areas, volumes) is useful in predicting lahar-inundation zones (Iverson *et al.*, 1998). Based on scaling and statistical analyses of geometry of 27 lahar paths at nine volcanoes, Iverson *et al.* (1998) derived empirical equations that estimate planimetric and cross-sectional areas of lahar-inundation as functions of lahar volumes. These empirical equations were used by Schilling (1998) to develop LAHARZ – a suite of geographic information system (GIS) programs – for automated mapping of lahar-inundation zones. Utility of LAHARZ for construction of lahar hazard maps has been demonstrated widely in the U.S. (e.g., Scott *et al.*, 1999; Walder *et al.*, 1999; Vallance *et al.*, 2003) and in Central America (e.g., Major *et al.*, 2001; Schilling *et al.*, 2001; Vallance *et al.*, 2001a, 2001b; Canuti *et al.*, 2002). A drawback of LAHARZ modeling is that empirical relationships between planimetric and cross-sectional areas and volumes of lahars were derived specifically at nine volcanoes and that these empirical relationships may not apply accurately to volcanoes in other areas.

In view of certain drawbacks of lahar flow-route and/or LAHARZ modeling mentioned, pre-eruption prediction of lahar-inundation zones pose specific challenges in areas where a volcano has not erupted within living memory. This is because, in such areas, baseline geological information about past lahars could be scarce or investigations to delimit past lahars might be incomplete. The objective of this paper is, thus, to demonstrate techniques to predict lahar-inundation zones prior to an erup-

tion that could generate substantial source materials for lahar formation. These techniques involve estimation, for any location around a volcano, of relative susceptibility to inundation by lahars that (a) form initially within a zone of source materials emplaced on volcano flanks due to an eruption and (b) flow subsequently towards adjacent lowlands. The focus is to predict zones that, in general, could be affected by multiple immediate post-eruption lahars but not to predict inundation by single lahars. Accordingly, predictive maps are validated with a map that depicts cumulative effects of multiple immediate post-eruption lahars.

To predict lahar-inundation zones, we follow some tenets adhered to by previous workers (e.g., Crandell and Mullineaux, 1975; Scott *et al.*, 1995; Iverson *et al.*, 1998). First, knowledge or information of past lahar sources and/or lahar-inundation zones (i.e., geologic precedence) provides basis for predicting future lahar-inundation. Second, lahars follow certain paths and a simplistic expectation is that lahars follow river channels that head on volcano flanks, although complications (e.g., flow avulsion, stream piracy) along lahar paths are not uncommon. Third, there is always uncertainty in predictive modeling of lahar-inundation zones; that is, no one can foretell exactly size of next lahar to descend a given drainage. In following these tenets, however, spatial data concerning lahar rheology or lahar geometry are not used. Of course, significance of such types of data to lahar hazard assessment is not questioned. Rather, it is recognized that, in areas where a volcano has not erupted within living memory, such types of data might be lacking. In our proposed techniques, input spatial data to predictive modeling of lahar-inundation zones were derived from a pre-eruption digital elevation model (DEM), as topographic maps (from which a DEM could be generated) are probably mostly available in situations where a volcano has not erupted within living memory. We demonstrate the predictive techniques in western part of Pinatubo volcano (Figure 1).

## 2. Test Area

Prior to 1991, Pinatubo volcano did not erupt for about 500 years (Newhall *et al.*, 1996). In early April 1991, several geologic indications were recognized that Pinatubo volcano would erupt. These indications prompted Pinatubo Volcano Observatory Team (1991) to prepare and then release, in late May 1991, a forecast volcanic-hazard map (Figure 2). Explosive eruptions in June 1991 resulted in emplacement of $>5$ km$^3$ of pyroclastic-flow deposits (Pierson *et al.*, 1992; Scott *et al.*, 1996) and $\sim$0.2 km$^3$ of tephra-fall deposits (Pierson *et al.*, 1992; Paladio-Melosantos *et al.*, 1996) on the slopes of the volcano. About two-thirds of total estimated volumes of pyroclastic-flows have been deposited in western part of

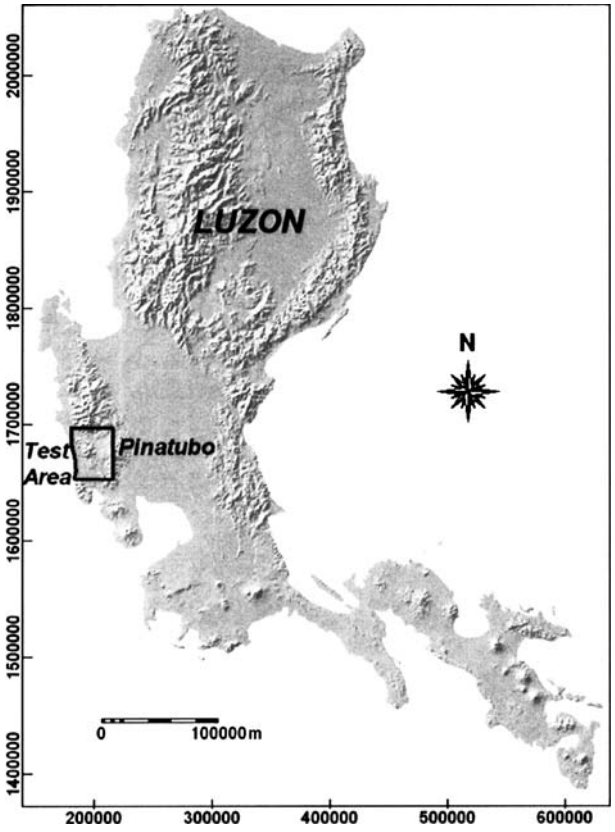


*Figure 1.* Location of test area in Luzon island, Philippines.

Pinatubo volcano (Figure 3; Rodolfo *et al.*, 1996). During and immediately after the eruptions, heavy rainfalls have re-worked the pyroclastic deposits into lahars (Major *et al.*, 1996; Pierson *et al.*, 1996; Rodolfo *et al.*, 1996; Umbal and Rodolfo, 1996).

In western part (as well as in eastern part) of Pinatubo volcano, most lahars originated from pyroclastic-flow deposits while some lahars originated from tephra-fall deposits (Pierson *et al.*, 1992; Rodolfo *et al.*, 1996; Umbal and Rodolfo, 1996). Lahars in 1991 inundated mainly channels of Santo Tomas, Bucao and Maloma river systems (Figure 3). Total volume of 1991 lahars was about $4.6 \times 10^8$ m$^3$ in an aggregate area of about 119 km$^2$ (Rodolfo *et al.*, 1996). These estimates of planimetric area ($B$) and volume ($V$) of 1991 lahars match closely empirical relationship (i.e., $B = 200\ V^{2/3}$) derived by Iverson *et al.* (1998). However, Rodolfo *et al.* (1996) state that estimated volume of 1991 lahars is exclusive of lahars that reached the sea. This suggests that empirical relationships derived by Iverson *et al.* (1998) may not apply accurately to western part of Pinatubo volcano.

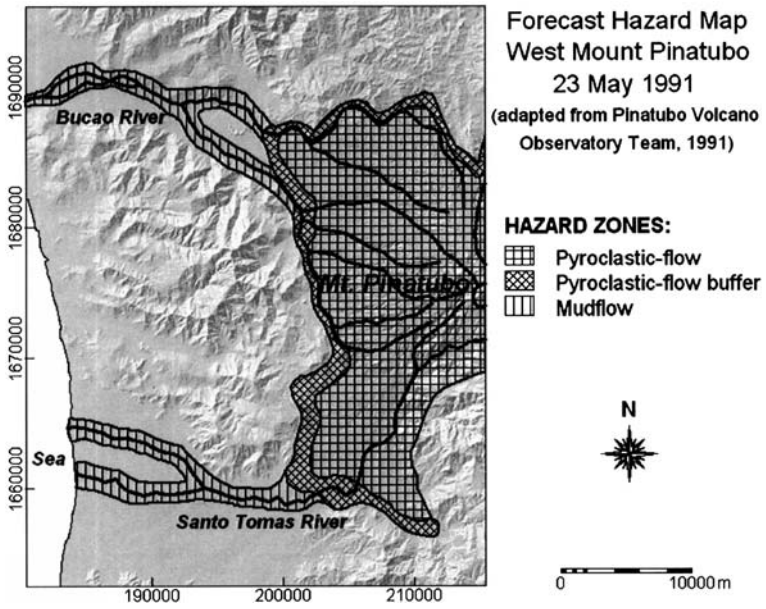


*Figure 2*. Pre-June-1991 volcanic hazard map of western part of Pinatubo volcano (adapted from Pinatubo Volcano Observatory Team, 1991) overlaid on a shaded-relief image of pre-eruption DEM.

## 3. Spatial Datasets

Digital capture of spatial data from analogue (i.e., paper) maps and subsequent spatial data analyses were performed using ILWIS (Integrated Land and Water Information System), a shareware GIS software developed by ITC (International Institute for Geo-Information Science and Earth Observation). In ILWIS, spatial data analysis is carried out in raster mode (i.e., pixel representation of spatial objects).

Spatial data used to predict lahar-inundation zones were derived from a DEM (see Figures 2–4). This DEM was created from pre-eruption 1:50,000 scale analogue topographic maps (with mostly 20-m interval elevation contours) produced by NAMRIA (National Mapping and Resources Information Authority, Philippines) from 1976 air-photos of about 1:44,000 scale. Each topographic map was raster-scanned at 200 dpi resolution, resulting in image pixel size of $\sim$6.3 m. Four corners of each image of topographic map were used as control points in georeferencing to UTM (zone 51) coordinate system, resulting in root mean square error (RMSE) of $< 1$ pixel per image (i.e., horizontal accuracy of $\sim$6 m). Contours in images of topographic maps were digitized (i.e., ''traced'') on-screen as line segments (i.e., vector format) and were assigned respective elevations as spatial attributes. Digitized contours were rasterized using a pixel size of

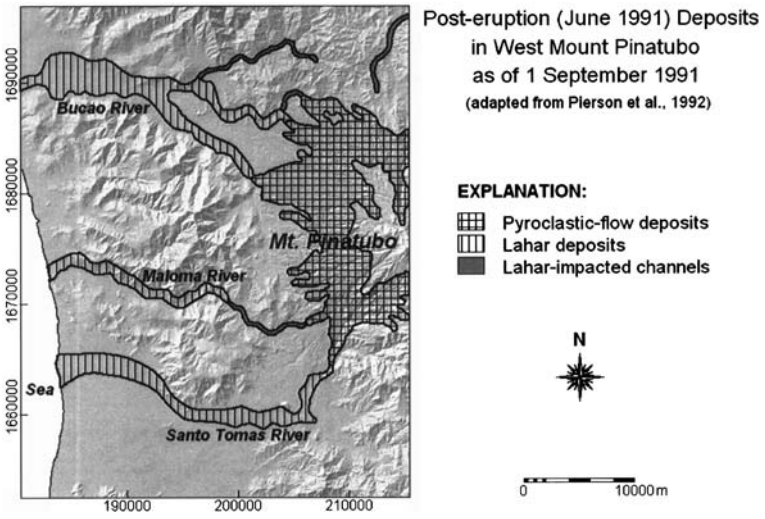


*Figure 3*. Field-based map of pyroclastic-flow deposits of June 1991 eruptions and lahar deposits as of 1 September 1991 in western part of Pinatubo volcano (adapted from Pierson *et al*., 1992) overlaid on a shaded-relief image of pre-eruption DEM.

20 m, which is roughly minimum horizontal distance between contours. The DEM was created by linear interpolation between raster contours (Gorte and Koolhoven, 1990).

For comparison and for validation of predicted zones that, in general, could be inundated by multiple immediate post-eruption lahars, we used spatial data shown in Figures 2 and 3, respectively. The pre-eruption volcanic hazards map (Figure 2) shows mudflow hazard zones, which probably represent predictions of immediate post-eruption inundation zones by multiple distal lahars. The map of post-eruption deposits (Figure 3) shows zones of lahar deposits and lahar-impacted channels, which represent zones that were inundated by multiple lahars within about two and a half months after the June 1991 eruptions. For example, lahar-inundation zones along the Santo Tomas river system are cumulative effects of at least 30 lahars in August 1991 (Rodolfo *et al*., 1996, p. 995).

The analogue maps for comparison and for validation were each raster-scanned at 1000 dpi resolution resulting in image pixel sizes of $\sim$17 and $\sim$25 m, respectively. Each image of volcanic-hazard map was georeferenced to UTM (zone 51) coordinate system by using at least four control points (e.g., river mouths/junctions, volcano summit) that are conspicuous both in images of volcanic-hazard maps and in georeferenced images of topographic maps. In GIS jargon, procedure to georeference an image (''slave'') by using a georeferenced image (''master'') is called image-to-image registration. The RMSE of georeferencing each image of volcanic-hazard map was $<1$ pixel (i.e., horizontal accuracy of $\sim$17 and $\sim$25 m,

respectively). Pertinent spatial features depicted in images of volcanic-hazard maps were digitized on-screen and then rasterized using a pixel size of 20 m.

It is assumed that image-to-image registration and use of the same pixel size to rasterize spatial data captured digitally from analogue maps ensure proper analysis (i.e., by overlay) of predictive maps against comparison map and with validation map. However, large differences in scale exist among analogue maps from which spatial data were captured. This could result in biased comparison and validation of our predictive maps. However, this problem was alleviated by raster-scanning of maps for comparison and for validation at a pixel resolution 5 times higher than maps for predictive modeling.

## 4. Predictive Modeling

We adopted from previous works (e.g., Iverson *et al.*, 1998; Schilling, 1998; Schilling *et al.*, 2001; Vallance *et al.*, 2001b) the following definitions and assumptions about lahar-inundation zones in terms of proximity to a volcano's vent. Areas proximal to a volcano's vent, which are subject to eruptive phenomena (e.g., pyroclastic flows, etc.), are collectively referred to as proximal-hazard zone and comprise usually source zone of lahars. Areas of lahar-inundation in a source zone can be considered to comprise a ''proximal lahar-inundation'' (hereafter denoted as *PLI*) zone. It is difficult, if not implausible, to predict exactly where *PLI* would occur in a given source zone. Thus, it is assumed that *PLI* zone match roughly with source zone of lahars. As lahars travel away from boundary of source zone, they become ''distal'' lahars and disperse towards adjacent lowlands along and around drainage channels to form distal lahar-inundation (hereafter denoted as *DLI*) zones. Thus, boundary of source zone represent ''end-points'' of *PLI* zone as well as ''start-points'' of *DLI* zones. Based on these definitions/assumptions, *DLI* zones are predicted after delineation of lahar source (or *PLI*) zone.

### 4.1. PREDICTION OF LAHAR SOURCE (OR *PLI*) ZONE

To predict *PLI* zone, we adopt technique of using a ratio of vertical descent ($H$) to horizontal run-out ($L$) of eruptive deposits that spawn lahars (Sheridan and Malin, 1983; Malin and Sheridan, 1982). In test area, pyroclastic-flow deposits are known source materials of most lahars (Newhall *et al.*, 1996). Hayashi and Self (1982) found that (a) pyroclastic-flow deposits have typical $H/L$ ratios of 0.1–0.3 and (b) $H/L$ ratios and volumes of pyroclastic-flow deposits are related inversely. Thus, a $H$ map was calculated using the pre-eruption DEM [i.e., $H$ equals elevation of volcano

summit minus elevation of each pixel] and a $L$ map was calculated depicting distances between each pixel and volcano summit in order to create a $H/L$ map. Assuming a worst-case eruption scenario resulting in voluminous pyroclastic-flow deposits, pixels with $H/L$ values $\geq 0.1$ were considered to represent a source zone (Figure 4). Given a source zone for lahar generation, probability of *PLI* in source zone can be estimated as equal to ratio of area of *PLI* in a source zone to area of source zone. However, as *PLI* zone is assumed to match roughly with source zone of lahars, probability of *PLI* in source zone is estimated simply as equal to 1.

### 4.2. PREDICTION OF *DLI* ZONES

Degree of susceptibility ($S_{DLI}$) to *DLI* can be defined as a function of a number of $X_i$ ($i = 1, 2, \ldots, n$) spatial factors deemed relevant for *DLI*:

$$S_{DLI} = f(X_1, \ldots, X_n). \tag{1}$$

Let map of test area be divided into square units (i.e., pixels). Let maps of $X_i$ spatial factors be partitioned into a number of $C_{ij}$ ($j = 1, 2, \ldots, m$) classes of attributes. For each pixel, $S_{DLI}$ is further defined further as

$$S_{DLI} = f(a_{C_{ij}}), \tag{2}$$

where $a_{Cij}$ represents a map depicting "weights" of $C_{ij}$ attributes of $X_i$ spatial factors with respect to map of past *DLI*. The "weights" relate to degree of spatial coincidence of $C_{ij}$ attributes of $X_i$ spatial factor with past

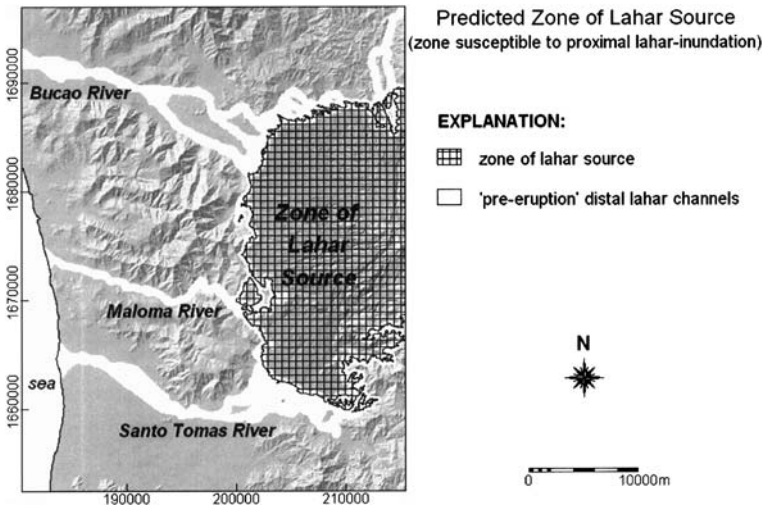


*Figure 4*. Predicted lahar source zone and calculated "pre-eruption" distal lahar-inundation zones overlaid on a shaded-relief image of pre-eruption DEM.

*DLI*. Suppose that aspect, i.e., azimuth of slope (with attributes measured in degrees), controls *DLI*. Suppose also that past distal lahars inundated mostly (or coincided spatially with) east-facing slopes. Thus, with respect to *DLI*, east-facing slopes have more "weight" than other slopes. Values of $a_{Cij}$, therefore, represent measures of spatial associations between *DLI* and $C_{ij}$ attributes of $X_i$ map, and these quantified spatial associations are used to estimate $S_{DLI}$. That is, several maps of $a_{Cij}$, each representing a factor, are integrated (in similar fashion of stacking transparent maps over each other on a light table) in order to delineate areas defined by intersections of high values of $a_{Cij}$ in most map layers. Areas defined by intersections of high values of $a_{Cij}$ in most map layers, thus, have spatial characteristics similar to (in terms of spatial association with) areas of past *DLI*. The concept of predicting *DLI* zones according to definition of $S_{DLI}$ depicted in Equations (1) and (2), thus, involves interpolative/extrapolative delineation of future (or unknown) *DLI* zones based on spatial associations of $C_{ij}$ attributes of $X_i$ spatial factors with past (or known) *DLI* zones.

To estimate values of $a_{Cij}$ for $C_{ij}$ attributes in $X_i$ spatial factor map, it is necessary to have (as geologic precedence) training data concerning past *DLI* zones. Certainly, mapped lahar deposits and lahar-impacted channels shown in Figure 3 cannot be used as training data to predict *DLI* zones before June 1991. Suppose further that baseline data concerning past *DLI* zones, prior to June 1991, are lacking. Based on only a simplistic expectation (as it is also known more commonly) that distal lahars follow drainage channels that head on source zone, we created "pre-eruption" *DLI* zones through an iterative spatial neighborhood algorithm (see Appendix A). This algorithm uses a $3 \times 3$ kernel filter, which moves simultaneously over DEM and map of "start-points" of *DLI* (i.e., pixels along boundary of source zone). Using input map of "start-points" of *DLI*, the iterative algorithm classifies central pixel of $3 \times 3$ kernel as part of a "pre-eruption" *DLI* zone if, and only if, its neighboring pixels (a) form part of a "pre-eruption" *DLI* zone and (b) are topographically higher than central pixel. Each iteration results in a temporary output raster map of "pre-eruption" *DLI* zones, which is used as input to next iteration. Iteration stops when the two spatial neighborhood conditions (in a $3 \times 3$ kernel) are not met anymore in penultimate temporary output map. Pixels classified as parts of a "pre-eruption" *DLI* zone (Figure 4) were then used as training data for estimation of $S_{DLI}$.

Four spatial factors were deemed relevant controls of $S_{DLI}$: (1) proximity to source zone; (2) proximity to drainage lines that head on source zone; (3) elevations outside source zone; and (4) slopes outside source zone. Logic behind first and second spatial factors is that, for any location [topographically] below source zone, it could be expected that $S_{DLI}$ (a) decreases with decreasing proximity to boundary of source zone and (b)

decreases with decreasing proximity to drainage lines that head on source zone. Logic behind third and fourth spatial factors is based on a simplistic expectation that, being saturated with water, lahars would (a) initially travel or flow from locations with higher elevations towards locations with lower elevations and (b) then temporarily or finally deposited in zones with gentle slopes rather than in zones with steep slopes.

Thus, four spatial factor maps were derived from pre-eruption DEM (Figure 5): (1) map of proximity to source zone; (2) map of proximity to drainage lines that head on source zone; (3) map of elevations outside source zone; and (4) map of slopes outside source zone. Map of drainage lines that head on source zone was created through another iterative spatial neighborhood algorithm (see Appendix A). This algorithm also uses a $3 \times 3$ kernel filter, which moves simultaneously over DEM and map of "start-points" of drainage lines (i.e., pixels along boundary of source zone). Using input map of "start-points" of drainage lines, the iterative algorithm classifies central pixel of $3 \times 3$ kernel as part of a drainage line if, and only if, its neighboring pixel(s) (a) form part of drainage line and (b) flow direction(s) from such pixel(s) is toward central pixel. Each iteration results in a temporary output raster map of drainage lines, which is used as input to next iteration. Iteration stops only when the two spatial neighborhood conditions (in a $3 \times 3$ kernel) are not satisfied anymore in penultimate temporary output map. Procedures for creating proximity (distance) map or slope map are basic GIS operations and are not described further.

In order to estimate values of $a_{Cij}$ (Equation (2)) for a spatial factor map of a test area $T$ with $N(T)$ total number of pixels, it is necessary to first determine spatial coincidence between certain pixels in map of training data and certain pixels in a spatial factor map. To do this, attributes of spatial factor maps, which represent continuous variables (Figure 5), were first categorized into classes (or ranges) of attributes because the intent is to estimate values of $a_{Cij}$ for certain classes of attributes rather than for individual attributes. However, choice of intervals with which to classify a spatial factor map of a continuous variable is arbitrary and, unfortunately, there are no rules for this. Each $X_i$ $(i = 1, 2, \ldots, n)$ spatial factor map was classified into $C_{ij}$ $(j = 1, 2, \ldots, m)$ classes using 5-percentile intervals for two reasons. One reason is to make a uniform classification by using equal intervals resulting in equal number of classes in each spatial factor map. The other reason, which is more important, is to create narrow ranges of spatial attributes and thus not a large number of $C_{ij}$ pixels (denoted as $N(C_{ij})$) in $X_i$ spatial factor map in order to optimize estimation of $a_{Cij}$ as it depends on $N(C_{ij})$ (see later below). Unlike the spatial factor maps, map of training data was classified simply into a binary map with $N(DLI)$ and $N(\overline{DLI})$ number of pixels representing areas where "pre-eruption" $DLI$ is present and absent, respectively. Note that $N(\overline{DLI}) = N(T) - N(DLI)$. By

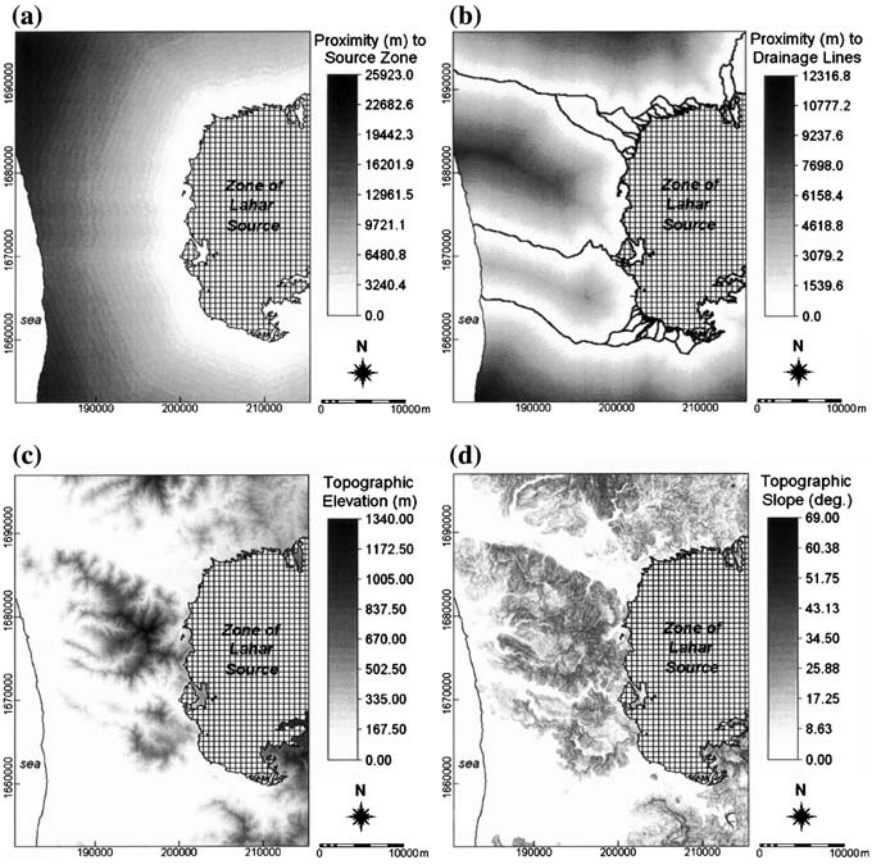


*Figure 5*. Spatial factors used in predictive modeling of distal lahar-inundation zones: (a) proximity to lahar source zone; (b) proximity to drainage lines that head on lahar source zone; (c) elevation outside lahar source zone; (d) slope outside lahar source zone.

overlaying binary map of training data on each multi-class $X_i$ spatial factor map, number of $C_{ij}$ pixels overlapping spatially with $DLI$ pixels [i.e., $N(C_{ij} \bigcap DLI)$] and number of $C_{ij}$ pixels not overlapping spatially with $\overline{DLI}$ pixels [i.e., $N(C_{ij}) - N(C_{ij} \bigcap DLI)$] were determined. Subsequently, numbers of pixels of certain map units and number of pixels in certain spatial overlap conditions were used to estimate $a_{Cij}$.

Values of $a_{Cij}$ were calculated and then maps of $a_{Cij}$ were integrated by applications of three spatial modeling techniques: (1) weights-of-evidence (or WofE); (2) logistic regression (or LR); and (3) evidential belief functions (or EBF). These spatial modeling techniques (a) quantify spatial associations (i.e., $a_{Cij}$) between training data and a spatial factor (or a set of factors, in case of LR) and (b) estimate, for each location,

values representing $S_{DLI}$ in Equation (2) by pixel-to-pixel integration of $a_{Cij}$ values between map layers. These spatial modeling techniques express values of $a_{Cij}$ and, thus, $S_{DLI}$ in terms of probability (i.e., within unit interval [0,1]).

#### 4.2.1. *Weights-of-Evidence (WofE)*

WofE is based upon a Bayesian probability framework (Bonham-Carter, 1994). Its application here involves a 3-stage process: (1) estimation of prior probability ($P_{\text{Prior}}$) of *DLI*; (2) estimation of WofE of spatial factor with respect to *DLI*; and (3) updating of $P_{\text{Prior}}$ by WofE of spatial factors to estimate posterior probabilities of *DLI*. The "gray-scale" WofE method, which is used with multi-class spatial factor maps (Agterberg and Bonham-Carter, 1990; An, 1992; Goodacre *et al.*, 1993), is adopted here.

Prior probability of *DLI* can be estimated as

$$P_{prior} = \frac{N(DLI)}{N(T)}. \tag{3}$$

WofE of $C_{ij}$, in $X_i$ spatial factor map, with respect to *DLI* pixels in map of training data, is estimated as:

$$W_{Cij} = \log_e\left\{\frac{P(C_{ij} \mid DLI)}{P(C_{ij} \mid \overline{DLI})}\right\} = \log_e \frac{\frac{N(C_{ij} \bigcap DLI)}{N(DLI)}}{\frac{N(C_{ij}) - N(C_{ij} \bigcap DLI)}{N(T) - N(DLI)}}. \tag{4}$$

Estimates of $W_{Cij}$ represent values of $a_{Cij}$ in Equation (2). For spatial overlap conditions in which more *DLI* pixels occur in $C_{ij}$ than would be expected due to chance (i.e., prior probability), $W_{Cij}$ is positive and implies that $C_{ij}$ has positive spatial association with *DLI*. For spatial overlap conditions in which less *DLI* pixels occur in $C_{ij}$ than would be expected due to chance, $W_{Cij}$ is negative and implies that $C_{ij}$ has negative spatial association with *DLI*. Uncertainty (in terms of variance) associated with estimates of $W_{Cij}$ can be calculated as (Bishop *et al.*, 1975):

$$s^2(W_{Cij}) = \frac{1}{N(C_{ij} \bigcap DLI)} + \frac{1}{N(C_{ij}) - N(C_{ij} \bigcap DLI)}. \tag{5}$$

A map of $W_{Cij}$ is created for each $X_i$ spatial factor and all $W_{Cij}$ maps are integrated with $\log_e$ prior odds ($O_{\text{prior}}$), which is related to $P_{\text{prior}}$ (i.e., $O_{\text{prior}} = P_{\text{prior}} / (1 - P_{\text{prior}})$), in order to estimate $\log_e$ posterior odds of *DLI*. Integration of $W_{Cij}$ maps results in so-called "unique-conditions", which are pixels or groups of pixels having same combination of $C_{ij}$ classes of $X_i$ spatial factors. For $k$ $(k = 1, 2, \ldots, p)$ unique-conditions, estimates of $\log_e$

posterior odds are converted to posterior probabilities ($P_{\text{post}}$), which represent values of $S_{DLI}$:

$$P_{\text{post}} = \frac{e^{\sum_{j=1}^{m} W^{k}_{Cij} + \log_e O_{\text{prior}}}}{1 + e^{\sum_{j=1}^{m} W^{k}_{Cij} + \log_e O_{\text{prior}}}}, \tag{6}$$

where the term

$$\sum_{j=1}^{m} W^{k}_{Cij} + \log_e O_{\text{prior}}$$

represents estimates of $\log_e$ posterior odds of $DLI$, and $W^{k}{}_{Cij}$ denotes WofE contributed by $C_{ij}$ class of $X_i$ spatial factor map to $k$th unique-condition.

WofE method assumes that maps of $X_i$ spatial factors are conditionally independent of each other with respect to map of training data. Assumption of conditional independence (CI) was tested using omnibus test or $OT$ (Bonham-Carter, 1994), which involves calculation of ratio number of "pre-eruption" $DLI$ pixels, $N(DLI)$, to number of predicted $DLI$ pixels, $N(DLI)_{\text{pred}}$. The latter is estimated as (Bonham-Carter, 1994):

$$N(DLI)_{\text{pred}} = \sum_{k=1}^{p} P^{k}_{\text{post}} * N(k), \tag{7}$$

where $N(k)$ is number of $k = 1, 2, \ldots, p$ unique-conditions pixels with posterior probabilities $P_{\text{post}}$. Ideally, if maps of $X_i$ spatial factors are conditionally independent of each other with respect to map of training data, then $N(DLI)_{\text{pred}} = N(DLI)$ and $OT = 1$. If $1 \geq OT \geq 0.85$, then assumption of CI is not seriously violated (Bonham-Carter, 1994). Violation of assumption of CI results in either over-estimation or under-estimation of probabilities, which, however, is not a serious concern if prediction is used for relative (e.g., ranking of $S_{DLI}$) and not absolute purposes (Pan and Harris, 2000).

#### 4.2.2. *Logistic Regression (LR)*

One alternative to address problem of lack of CI among maps of spatial factors is to apply LR, whereby classes of spatial factors showing conditional dependence among each other with respect to training data are eliminated automatically in the process of estimating posterior probabilities (Agterberg *et al.*, 1993; Agterberg and Bonham-Carter, 1999). LR is a multivariate technique that considers simultaneously several independent variables to estimate probability of a dichotomous or binary dependent variable (Menard, 1995). We used SPSS v.10 statistical software package in

LR modeling, although the spatial variables were derived using ILWIS software. LR estimates of $S_{DLI}$ in SPSS v.10 were then exported back to ILWIS for visualization and mapping of *DLI*.

Suppose that test area is further divided into $t$ square cells and that $mn$ independent variables (i.e., $C_{ij}$ classes of all $X_i$ spatial factors) are coded for every $t$th cell (i.e., number of pixels of $C_{ij}$ classes is attributed to each $t$th cell). To estimate $S_{DLI}$ in Equation (2), let *DLI* be a binary dependent variable; that is, $DLI = 1$ and $DLI = 0$ if "pre-eruption" *DLI* pixels are, respectively, present and absent in every $t$th cell. In order that estimates of posterior probabilities (i.e., values of $S_{DLI}$ in Equation (2)) lie within unit interval [0,1], the logistic form is postulated; thus,

$$S_{DLI} = \frac{e^{(aC'_{ij}C'_{ij})}}{1 + e^{(aC'_{ij}C'_{ij})}} \tag{8}$$

where $C'_{ij} = (C_{11}, C_{12}, \ldots, C_{mn})$ is $mn$ dimensional vector of $C_{ij}$ classes of $X_i$ spatial factors, and $a_{C'_{ij}}$ is corresponding $mn$ dimensional vector of unknown LR coefficients to be determined for each $C_{ij}$ class.

The term $a_{C'_{ij}}C'_{ij}$ in Equation (8) represents values of $a_{Cij}$ in Equation (2). Cox and Snell (1989) suggest use of maximum likelihood method to determine LR coefficients. A backward stepwise LR was also performed to eliminate classes of spatial factors that do not contribute significantly to LR; a 90% significance level was used as statistical criterion to retain a class of spatial factor in final LR model. The Wald statistic is used to test significance of a LR coefficient associated with a $C_{ij}$ class of $X_i$ spatial factor (Menard, 1995). A goodness-of-fit statistic (*pseudo-*$R^2$) was computed to compare LR model with "pre-eruption" *DLI* model (i.e., training data), defined as (Mark and Ellen, 1995):

$$pseudo - R^2 = 100 \times (1 - [(model - observed)^2/(mean - observed)^2]). \tag{9}$$

where *model* is LR model (i.e., sum of estimated posterior probabilities in all cells), *observed* is "pre-eruption" *DLI* model (i.e., number of cells with $DLI = 1$) and *mean* is average of *observed* probability for all cells. Estimates of *pseudo-*$R^2$ could range from 0 to 100 percent. As estimates of *pseudo-*$R^2$ depends on level of spatial aggregation (i.e., aggregation of pixels into $t$th cells), it is used only for relative comparison of models (Mark and Ellen, 1995).

#### 4.2.3. *Evidential Belief Functions (EBF)*

The Dempster–Shafer theory of evidence provides framework for estimation of EBF (Dempster, 1967; Shafer, 1976), which are integrated accord-

ing to Dempster's (1968) rule of combination. Zadeh (1986) provides a simplification of Dempster–Shafer theory of evidence and shows ability of Dempster's rule of combination to integrate evidences with distinct probability distributions. However, Walley (1987, p. 1460) concluded that "... Dempster's rule should not be used to combine evidence from statistically independent observations ...". Thus, problem of lack of CI among spatial factor maps can be alleviated by application Dempster–Shafer theory of evidential belief. The following discussion for its application here is simplified and informal.

For each $C_{ij}$ class of attributes in $X_i$ spatial factor maps, estimation of EBF is always in relation to a proposition, which in this case is: "This pixel (or location) is susceptible to *DLI*". The EBF to be estimated are *Bel* (degree of belief), *Dis* (degree of disbelief), *Unc* (degree of uncertainty) and *Pls* (degree of plausibility). *Bel* and *Pls* represent, respectively, lower and upper probabilities that a spatial factor supports the proposition. Thus, *Pls* is usually greater but could be equal to *Bel*. Unc is equal to *Pls–Bel* and represents ignorance (or doubt) of one's belief in the proposition given a spatial factor; $Unc = 0$ when $Bel = Pls$. *Dis* is belief that the proposition is false given a spatial factor; it is equal to $1-Pls$. Thus, $Bel + Unc + Dis = 1$. If $Unc = 0$, then $Bel + Dis = 1$, as in probability approach.

Estimates of EBF are usually knowledge-driven, i.e., based on expert knowledge (e.g., Moon, 1989, 1990; An *et al.*, 1992, 1994). Here, due to "lack" of expert knowledge, the following equations were used for data-driven estimation of EBF (Carranza and Hale, 2003).

$$Bel_{C_{ij}} = \frac{W_{C_{ij}DLI}}{\sum_{i=1}^{n} W_{C_{ij}DLI}} \tag{10}$$

where

$$W_{C_{ij}DLI} = \frac{\frac{N(C_{ij} \cap DLI)}{N(C_{ij})}}{\frac{N(DLI) - N(C_{ij} \cap DLI)}{N(T) - N(C_{ij})}}.$$

$$Dis_{C_k} = \frac{W_{C_{ij}\overline{DLI}}}{\sum_{k=1}^{n} W_{C_{ij}\overline{DLI}}}, \tag{11}$$

where

$$W_{C_{ij}\overline{DLI}} = \frac{\frac{N(C_{ij}) - N(C_{ij} \cap DLI)}{N(C_{ij})}}{\frac{N(T) - N(DLI) - [N(C_{ij}) - N(C_{ij} \cap DLI)]}{N(T) - N(C_{ij})}}.$$

$$Unc_{C_{ij}} = 1 - Bel_{C_{ij}} - Dis_{C_{ij}}. \tag{12}$$

$$Pls_{C_{ij}} = Bel_{C_{ij}} + Unc_{C_{ij}} \text{ or } Pls_{C_{ij}} = 1 - Dis_{C_{ij}}. \tag{13}$$

Estimates of $Bel_{C_{ij}}$ are measures of spatial association between map of training data and a spatial factor map and, thus, represent values of $a_{Cij}$ in Equation (2).

According to Dempster's (1968) rule of combination, maps of EBF of spatial factor $X_1$ can be combined with maps of EBF of spatial factor $X_2$ to generate an integrated map of EBF of two spatial factors $X_1$ and $X_2$. Maps of EBF of two spatial factors are combined according to the following equations (adopted from Wright and Bonham-Carter, 1996):

$$Bel_{X_1X_2} = \frac{Bel_{X_1}Bel_{X_2} + Bel_{X_1}Unc_{X_2} + Bel_{X_2}Unc_{X_1}}{\beta.} \tag{14}$$

$$Dis_{X_1X_2} = \frac{Dis_{X_1}Dis_{X_2} + Dis_{X_1}Unc_{X_2} + Dis_{X_2}Unc_{X_1}}{\beta.} \tag{15}$$

$$Unc_{X_1X_2} = \frac{Unc_{X_1}Unc_{X_2}}{\beta,} \tag{16}$$

where $\beta = 1 - Bel_{X_1}Dis_{X_2} + Dis_{X_1}Bel_{X_2}$, which is a normalizing factor that ensures $Bel + Unc + Dis = 1$. Only maps of EBF of two spatial factors can be combined each time; other maps of EBF representing $X_3, \ldots, X_n$ are combined one after another by repeated applications of Equations (14)–(16). Note that Equations (14)–(16) are both commutative and associative, so that any group or order of evidential map combinations does not affect the final output. Final combination of maps of EBF results in integrated degrees of belief, degrees of disbelief, degrees of uncertainty and degrees of plausibility for the proposition based on given spatial factors. As estimates of $Bel_{C_{ij}}$ are here considered to represent values of $a_{C_{ij}}$ in Equation (2), final map of integrated degrees of belief ($Bel_{X_1,\ldots,X_n}$) is considered to represent values of $S_{DLI}$ in Equation (2).

### 4.3. CLASSIFICATION AND VALIDATION OF PREDICTIVE MAPS

Integration of maps depicting values of $a_{C_{ij}}$ by each of above-described techniques results in maps depicting values of $S_{DLI}$ (Equation (2)), which is a continuous variable (i.e., posterior probabilities, degrees of belief) and cumbersome to interpret unless classified. Classification of $S_{DLI}$ maps is convenient because derived posterior probabilities or degrees of belief are useful only in relative but not in absolute terms. That is, it is practically

implausible to know absolutely probability of *DLI*. Predictive maps were created by classification of $S_{DLI}$ values as "very low", "low", "moderate", "high" or "very high". "Very low" $S_{DLI}$ values are those less than estimated prior probability. The other classes are defined by quartile classification of $S_{DLI}$ values equal to or greater than estimated prior probability.

To validate predictive maps, it was convenient to create binary predictive maps, because validation map is also binary. That is, for example, Figure 3 shows areas inundated and not inundated by lahars. Creation of binary predictive map of lahar source (or *PLI*) zone is straightforward, because such zone is essentially binary. To create a binary *DLI* validation map, zones mapped as lahar deposits and as lahar-impacted channels (Figure 3) are considered as mapped *DLI* zones; otherwise, they are considered as mapped non-*DLI* zones. To create a binary *DLI* predictive map, zones classified as having "moderate" to "very high" $S_{DLI}$ values were re-classified as predicted *DLI* zones, whereas zones classified as having "low" to "very low" $S_{DLI}$ values were re-classified as predicted non-*DLI* zones.

To validate a binary predictive map, it is overlaid on a binary validation map in order to determine number of pixels in overlap zones between the units of the two binary maps (Figure 6). "True positive" prediction accuracy is estimated as proportion (expressed as percentage) of number of pixels in overlap between predicted *DLI* zones and mapped *DLI* zones to total number of pixels in mapped *DLI* zones. "True negative' prediction accuracy is estimated as proportion (expressed as percentage) of 4 number of pixels in overlap between predicted non-*DLI* zones and mapped non-*DLI* zones to total number of pixels in mapped non-*DLI* zones, "False negative" prediction error is estimated as proportion (expressed as percentage) of number of pixels in overlap between predicted non-*DLI* zones and mapped *DLI* zones to total number of pixels in mapped *DLI* zones, "False positive" prediction error is estimated as proportion (expressed as percentage) of number of pixels in overlap between predicted *DLI* zones and mapped non-*DLI* zones to total number of pixels in mapped non-*DLI* zones,

| | | Validation field-based map | |
|---|---|---|---|
| | | Mapped hazard zones | Mapped non-hazard zones |
| Predictive map | Predicted hazard zones | **A** | **B** |
| | Predicted non-hazard zones | **C** | **D** |

'True positive' prediction accuracy = **A**÷(**A**+**C**)× 100
'True negative' prediction accuracy = **D**÷(**B**+**D**)× 100
'False positive' (Type II) prediction error = **B**÷(**B**+**D**)× 100
'False negative' (Type I) prediction error = **C**÷(**A**+**C**)× 100

**A**, **B**, **C** and **D** are number of pixels

*Figure 6*. Error matrix for estimation of prediction accuracy and prediction error.

''False negative' prediction error and ''false positive'' prediction error are, respectively, called Type I error and Type II error in statistical hypothesis testing. Type I prediction error represents under-estimation of lahar-inundation hazard, whereas Type II prediction error represents over-estimation of lahar-inundation hazard. It is, of course, desirable to have minimal Type I and Type II prediction errors. It is, however, more desirable to have Type I prediction error lower than Type II prediction error, because it is conceivable that gross under-estimation of lahar-inundation hazard, rather than gross over-estimation of lahar-inundation hazard, could potentially result in higher damage to lives and properties.

## 5. Results

### 5.1. PREDICTED LAHAR SOURCE (OR *PLI*) ZONE

Prior to describing and validating results of predicting a lahar source (or *PLI*) zone, it is imperative to describe pertinent spatial information about test area, pre-eruption forecasted hazards (Figure 2) and mapped eruptive deposits (Figure 3). Total number of ($20 \times 20$ m) pixels in test area is 3,693,032 ($\approx$1,477 km$^2$). Forecasted zone of pyroclastic-flow deposits (Figure 2) contains 808,729 pixels; whereas mapped pyroclastic-flow deposits (Figure 3) consist of 388,552 pixels. Number of pixels in overlap between forecasted *PLI* zone and mapped *PLI* zone is 375,638, whereas number of pixels in overlap between forecasted non-*PLI* zone and mapped non-*PLI* zone is 2,871,389 (Table I). The forecasted *PLI* zone has 96.7% ''true positive'' prediction accuracy and 86.9% ''true negative'' prediction accuracy. In addition, the forecasted *PLI* zone has 3.3 Type I prediction error and 13.1% Type II prediction error. Thus, forecasted *PLI* zone of pyroclastic-flow deposits has very high ''true positive'' prediction accuracy, high ''true negative'' prediction accuracy, very low Type I error, and low Type II error.

Predicted lahar source (or *PLI*) zone (Figure 4) contains 822,471 pixels. Number of pixels in overlap between mapped *PLI* zone and predicted *PLI* zone is 378,224, whereas number of pixels in 6 overlap between mapped non-*PLI* zone and predicted non-*PLI* zone is 2,860,223 (Table I). The predicted *PLI* zone has 97.3% ''true positive'' prediction accuracy and 86.6% ''true negative'' prediction accuracy. In addition, predicted *PLI* zone has 2.5% Type I prediction error and 13.4% Type II prediction error. Thus, predicted *PLI* zone has very high ''true positive'' accuracy, high ''true negative'' accuracy, very low Type 1 error, and low Type II error. These show that the forecasted *PLI* zone and the predicted *PLI* zone have similar ''true negative'' prediction accuracy and Type II error, although the latter has slightly higher ''true positive'' prediction accuracy and slightly lower Type I prediction error than the former.

*Table I.* Error matrices for validation of forecasted *PLI* zone and predicted *PLI* zone against field-based map of *PLI* zone.

| | | Field-based map of *PLI* zone (Figure 3) | |
|---|---|---|---|
| | | *PLI* zone | Non-*PLI* zone |
| Forecast map of *PLI* zone | *PLI* zone | 375,638 | 433,091 |
| (Figure 2) | Non-*PLI* zone | 12,914 | 2,871,389 |
| Predictive map of *PLI* zone | *PLI* zone | 378,224 | 444,247 |
| (Figure 4) | Non-*PLI* zone | 10,328 | 3,304,480 |

Values in table are number of pixels.

### 5.2. PREDICTED *DLI* ZONES

Procedures for predicting *DLI* zones were undertaken in area outside predicted *PLI* zone (see Figure 4 or 5), which contains 2,871,150 pixels (i.e., $N(T)$). Area outside predicted *PLI* zone is also used as reference area for comparison between forecasted *DLI* zones and predicted *DLI* zones through validation against mapped *DLI* zones. Number of "pre-eruption" *DLI* pixels (i.e., $N(L)$) of training data is 153,475. Thus, estimated prior probability of *DLI* (Equation (3)) in every pixel is about 0.053.

Before describing and validating results of predicting *DLI* zones, it is imperative to describe pertinent spatial information about (a) pre-eruption forecasted *DLI* zones (i.e., mudflows; Figure 2) and (b) mapped *DLI* zones (i.e., lahar deposits and lahar-impacted channels; Figure 3). Number of pixels of forecasted *DLI* zones is 235,281, whereas number of pixels of mapped *DLI* zones is 299,348. Number of pixels in overlap between forecasted *DLI* zones and mapped *DLI* zones is 137,477, whereas number of pixels in overlap between forecasted non-*DLI* zones and mapped non-*DLI* zones is 2,474,048 (Table V). The forecasted *DLI* zones have 45.9% "true positive" prediction accuracy and 96.2% "true negative" prediction accuracy, In addition, the forecasted *DLI* zones have 54.1% Type I prediction error and 3.8% Type II prediction error. Thus, forecasted *DLI* zones have low "true positive" accuracy, very high "true negative" prediction accuracy, high Type I prediction error and very low Type II prediction error.

#### 5.2.1. *WofE Model*

For a multi-class spatial factor map representing a continuous variable (e.g., distance, slope, elevation), WofE can be calculated via either a cumulative increasing approach or a cumulative decreasing approach (Raines *et al.*, 2000). Cumulative increasing approach was used to calculate WofE

for proximity to lahar source (or *PLI*) zone and for proximity to drainage lines that head on lahar source (or *PLI)* zone. This is because one would expect that distal lahars travel away (i.e., increasing distance) from lahar source (or *PLI*) zone or from drainage lines that head on lahar source (or *PLI*) zone. Cumulative decreasing approach was used to calculate WofE for elevations and slopes, as distal lahars, being saturated with water, generally tend to flow from locations of higher elevations towards locations of lower elevations (thus, decreasing elevation) and then deposited from locations of steeper slopes towards locations of gentler slopes (thus, decreasing slope).

Table II shows variations in values of $W_{C_{ij}}$ for classes of attributes of spatial factors with respect to "pre-eruption" *DLI* zones. Proximity to lahar source (or *PLI*) zone and proximity to drainage lines that head on lahar source (or *PLI*) zone have positive spatial associations with "pre-eruption" *DLI* zones. Values of $W_{C_{ij}}$ indicate that areas within about 4 km of lahar source (or *PLI*) zone and areas within about 2 km of drainage lines that head on lahar source (or *PLI*) zone are highly susceptible to *DLI*. Elevations and slopes have negative spatial associations with "pre-eruption" *DLI* zones. This latter quantitative characterization is, in general, consistent with how water-saturated materials follow a transportation and deposition process on [topographic] surfaces. Values of $W_{C_{ij}}$ indicate that areas with elevations $< 200$ m and areas with slopes $< 7°$ are susceptible to *DLI*.

In posterior probability map resulting from integration of $W_{C_{ij}}$ maps, number of pixels in predicted *DLI* zones is 209,963 (i.e., $N(L)_{\text{pred}}$ in Equation (7)). Thus, $OT$ is 0.73, which indicates lack of CI of among $C_{ij}$ classes of $X_i$ spatial factor maps with respect to "pre-eruption" *DLI* zones. This is expected as spatial factor maps were all derived from DEM. However, estimated posterior probabilities are interpreted only in relative terms and were thus used to rank (i.e., classify) $S_{DLI}$ values. Figure 7 shows classified map of $S_{DLI}$ values based on WofE modeling. Zones of "very low" $S_{DLI}$ occupy about 69% of area outside zone of lahar source. Each of other zones occupies about 7.7% of area outside lahar source (or *PLI*) zone. (Reclassified binary predictive map of *DLI* zones based on WofE modeling is not shown here because of space.)

Number of pixels in overlap between mapped *DLI* zones and predicted *DLI* zones is 235,085, whereas number of pixels in overlap between mapped non-*DLI* zones and predicted non-*DLI* zones is 2,155,135 (Table V). The WofE model has 78.5% "true positive" prediction accuracy and 83.8% "true negative" prediction accuracy. In addition, for the WofE model has 21.5% Type I prediction error and 16.2% Type II prediction error. Thus, WofE model has fairly high "true positive" prediction accuracy, high "true negative" prediction accuracy, fairly low Type I error and low Type II error. These show that the WofE model of predicted *DLI* zones is

*Table II.* Estimates of WofE (and their standard deviations) for classes of spatial factors; (A) proximity to lahar source zone; (B) proximity to drainage lines that head on lahar source zone; (C) elevations outside lahar source zone; (D) slopes outside lahar source zone.

| (A) Proximity to lahar source zone | | | | | (B) Proximity to drainage that head on lahar source zone | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Ascending 5-percentile no. of classes (km) | Cumulative no. of class pixels | Cumulative no. of distal lahar pixels | $W_{C_{ij}}$ | $s(W_{c_{ij}})$ | Ascending 5-percentile no. of classes (km) | Cumulative no. of class pixels | Cumulative no. of distal lahar pixels | $W_{C_{ij}}$ | $s(W_{c_{ij}})$ |
| 0.00–0.55 | 143,391 | 37,570 | 1.838 | 0.005 | 0.00–0.10 | 142,651 | 80,746 | 3.140 | 0.004 |
| 0.55–1.44 | 286,782 | 63,395 | 1.614 | 0.004 | 0.10–0.25 | 286,161 | 119,121 | 2.536 | 0.003 |
| 1.44–2.36 | 430,126 | 77,442 | 1.358 | 0.004 | 0.25–0.43 | 427,403 | 141,224 | 2.168 | 0.003 |
| 2.36–3.29 | 574,038 | 84,467 | 1.117 | 0.003 | 0.43–0.67 | 572,905 | 150,414 | 1.841 | 0.003 |
| 3.29–4.20 | 716,899 | 89,262 | 0.924 | 0.003 | 0.67–0.93 | 717,328 | 52,980 | 1.569 | 0.003 |
| 4.20–5.09 | 860,838 | 91,637 | 0.746 | 0.003 | 0.93–1.21 | 860,654 | 153,325 | 1.345 | 0.003 |
| 5.09–5.94 | 1,004,513 | 93,698 | 0.600 | 0.003 | 1.21–1.51 | 1,003,782 | 153,466 | 1.162 | 0.003 |
| 5.94–6.76 | 1,148,198 | 94,827 | 0.466 | 0.003 | 1.51–1.84 | 1,147,767 | 153,475 | 1.006 | 0.003 |
| 6.76–7.57 | 1,291,195 | 95,876 | 0.351 | 0.003 | 1.84–2.19 | 1,290,368 | 153,475 | 0.871 | 0.003 |
| 7.57–8.44 | 1,435,295 | 97,985 | 0.260 | 0.003 | 2.19–2.54 | 1,434,662 | 153,475 | 0.752 | 0.003 |
| 8.44–9.40 | 1,578,893 | 99,940 | 0.179 | 0.003 | 2.54–2.92 | 1,578,450 | 153,475 | 0.646 | 0.003 |
| 9.40–10.43 | 1,722,150 | 106,464 | 0.154 | 0.003 | 2.92–3.30 | 1,721,687 | 153,475 | 0.550 | 0.003 |
| 10.43–11.53 | 1,865,701 | 117,130 | 0.171 | 0.003 | 3.30–3.70 | 1,865,637 | 153,475 | 0.462 | 0.003 |
| 11.53–12.68 | 2,009,599 | 124,276 | 0.155 | 0.003 | 3.70–4.17 | 2,00,9554 | 153,475 | 0.381 | 0.003 |
| 12.68–13.89 | 2,153,065 | 127,570 | 0.109 | 0.003 | 4.17–4.69 | 2,151,700 | 153,475 | 0.308 | 0.003 |
| 13.89–15.12 | 2,296,540 | 129,893 | 0.060 | 0.003 | 4.69–5.25 | 2,296,448 | 153,475 | 0.238 | 0.003 |
| 15.12–16.38 | 2,440,138 | 132,452 | 0.016 | 0.003 | 5.25–5.89 | 2,439,676 | 153,475 | 0.173 | 0.003 |
| 16.38–17.83 | 2,583,832 | 138,946 | 0.006 | 0.003 | 5.89–6.57 | 2,583,333 | 153,475 | 0.112 | 0.003 |
| 17.83–19.99 | 2,727,247 | 144,000 | −0.013 | 0.003 | 6.57–7.48 | 2,727,373 | 153,475 | 0.054 | 0.003 |
| 19.99–25.93 | 2,871,150 | 153,475 | 0.000 | 0.003 | 7.48–12.32 | 2,871,150 | 153,475 | 0.000 | 0.003 |

*Table II.* Continued

| (C) Elevations outside lahar source zone | | | | | (D) Slopes outside lahar source zone | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Descending 5-percentile no. of classes (m) | Cumulative no. of class pixels | Cumulative no. of distal lahar pixels | $W_{C_{ij}}$ | $s(W_{c_{ij}})$ | Descending 5- percentile no. of classes (degrees) | Cumulative no. of class pixels | Cumulative no. of distal lahar pixels | $W_{C_{ij}}$ | $s(W_{c_{ij}})$ |
| 690–1340 | 145,489 | 879 | −2.229 | 0.034 | 31.51–69.19 | 143,444 | 545 | −2.695 | 0.043 |
| 555–690 | 287,344 | 1,306 | −2.515 | 0.028 | 27.54–31.51 | 287,049 | 939 | −2.845 | 0.033 |
| 480–555 | 435,287 | 1,646 | −2.700 | 0.025 | 24.64–27.54 | 430,271 | 1,343 | −2.892 | 0.027 |
| 420–480 | 596,557 | 2,270 | −2.694 | 0.021 | 22.24–24.64 | 574,253 | 1,813 | −2.881 | 0.023 |
| 380–420 | 730,536 | 3,180 | −2.559 | 0.018 | 20.14–22.24 | 717,418 | 2,271 | −2.878 | 0.021 |
| 340–380 | 865,570 | 4,259 | −2.435 | 0.015 | 18.17–20.14 | 861,927 | 2,793 | −2.855 | 0.019 |
| 300–340 | 1,005,048 | 7,124 | −2.068 | 0.012 | 16.33–18.17 | 1,004,645 | 3,441 | −2.799 | 0.017 |
| 255–300 | 1,149,107 | 10,845 | −1.780 | 0.010 | 14.57–16.33 | 1,148,360 | 4,361 | −2.696 | 0.015 |
| 214–255 | 1,292,162 | 15,543 | −1.534 | 0.008 | 12.84–14.57 | 1,291,840 | 5,256 | −2.626 | 0.014 |
| 180–214 | 1,439,170 | 25,290 | −1.150 | 0.006 | 11.04–12.84 | 1,435,342 | 6,501 | −2.519 | 0.012 |
| 140–180 | 1,600,390 | 36,583 | −0.881 | 0.005 | 9.03–11.04 | 1,579,244 | 8,945 | −2.294 | 0.011 |
| 115–140 | 1,722,805 | 56,073 | −0.518 | 0.004 | 6.54–9.03 | 1,722,379 | 13,848 | −1.941 | 0.008 |
| 88–115 | 1,866,688 | 78,757 | −0.248 | 0.004 | 3.69–6.54 | 1,865,894 | 24,207 | −1.458 | 0.006 |
| 60–88 | 2,013,672 | 91,611 | −0.170 | 0.003 | 1.85–3.69 | 2,009,716 | 49,142 | −0.812 | 0.005 |
| 42–60 | 2,153,436 | 99,262 | −0.156 | 0.003 | 1.01–1.85 | 2,155,085 | 77,381 | −0.416 | 0.004 |
| 27–42 | 2,296,953 | 105,935 | −0.155 | 0.003 | 0.58–1.01 | 2,302,209 | 93,462 | −0.289 | 0.003 |
| 18–27 | 2,440,416 | 109,778 | −0.181 | 0.003 | 0.37–0.58 | 2,433,317 | 104,971 | −0.225 | 0.003 |
| 13–18 | 2,584,125 | 114,958 | −0.193 | 0.003 | 0.26–0.37 | 2,568,745 | 112,624 | −0.208 | 0.003 |
| 7–13 | 2,727,580 | 127,964 | −0.137 | 0.003 | 0.17–0.26 | 2,717,602 | 127,901 | −0.134 | 0.003 |
| 0–7 | 2,871,150 | 153,475 | 0.000 | 0.003 | 0.00–0.17 | 2,871,150 | 153,475 | 0.000 | 0.003 |

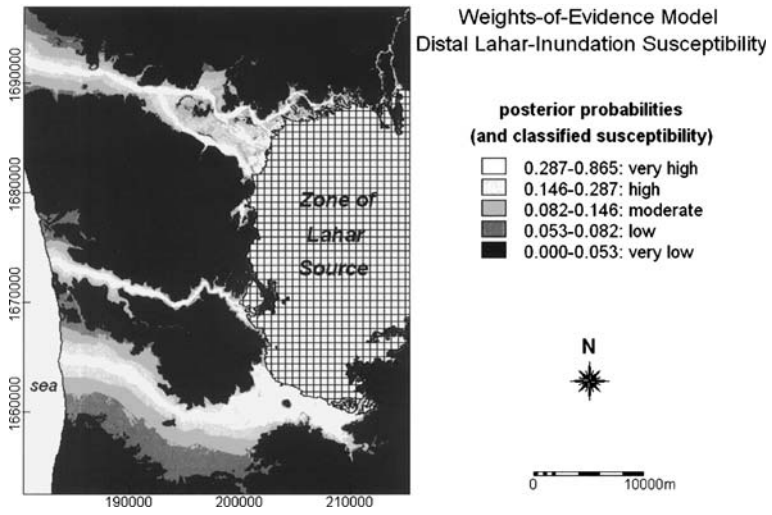


*Figure 7.* Probabilistic map of distal lahar-inundation zones derived by WofE modeling.

superior to the forecasted *DLI* zones in terms of "true positive" prediction accuracy and Type I prediction error, although it is inferior to the forecasted *DLI* zones in terms of "true negative" prediction accuracy and Type II prediction error.

### 5.2.2. *LR Model*

Test area was further divided into 500-m square cells (i.e., each cell contains 625 20 m $\times$ 20 m pixels) to estimate posterior probabilities of *DLI* according to Equation (8). Table III shows regression coefficients of classes of spatial factors remaining in final model achieved by backward stepwise LR.

Classes of proximity to lahar source (or *PLI*) zone in final LR model are those $<6$ km. Classes of proximity to drainage lines that head on lahar source (or *PLI*) zone in final LR model are those $<2$ km. These indicate that pixels $<6$ km of lahar source (or *PLI*) zone and pixels $<2$ km of drainage lines that head on lahar source (or *PLI*) zone lack CI among each other with respect to "pre-eruption" *DLI* zones. Exclusion of $>6$ km classes of proximity to lahar source (or *PLI*) zone and inclusion of $>2$ km classes of proximity to drainage lines that head on lahar source (or *PLI*) zone in final LR model indicate the following. Proximity to drainage lines that head on lahar source (or *PLI*) zone, rather than proximity to lahar source (or *PLI*) zone, is a statistically significant predictor of *DLI* zones. Coefficients for classes of proximity to lahar source (or *PLI*) zone in final LR model are negative, which indicates negative spatial association between pixels of "pre-eruption" *PLI* zones and pixels at distances $<6$ km of

*Table III*. LR coefficients of classes of spatial factors and their of indices of statistical significance obtained by backward regression modeling.

| Independent variable class* | $a_{C'_{ij}}$ | Asymptotic standard error | Wald-statistic | Significance level ($\alpha$) |
|---|---|---|---|---|
| PS-035 | −0.006 | 0.002 | 9.562 | 0.002 |
| PS-040 | −0.011 | 0.004 | 9.379 | 0.002 |
| PS-045 | 0.009 | 0.002 | 13.645 | 0.000 |
| PS-050 | −0.006 | 0.003 | 6.211 | 0.013 |
| PS-055 | −0.004 | 0.002 | 6.114 | 0.013 |
| PS-075 | 0.003 | 0.002 | 4.497 | 0.034 |
| PS-080 | −0.006 | 0.002 | 10.887 | 0.001 |
| PS-085 | −0.008 | 0.002 | 12.481 | 0.000 |
| PS-100 | −0.005 | 0.002 | 6.885 | 0.009 |
| PD-005 | 0.066 | 0.009 | 50.688 | 0.000 |
| PD-010 | 0.024 | 0.005 | 22.899 | 0.000 |
| PD-015 | 0.023 | 0.003 | 41.877 | 0.000 |
| PD-020 | 0.010 | 0.002 | 17.195 | 0.000 |
| PD-025 | 0.018 | 0.003 | 46.472 | 0.000 |
| PD-035 | 0.021 | 0.004 | 22.153 | 0.000 |
| PD-040 | −0.165 | 0.069 | 5.669 | 0.017 |
| EL-010 | −0.002 | 0.001 | 2.893 | 0.089 |
| EL-040 | 0.005 | 0.002 | 9.958 | 0.002 |
| EL-050 | −0.003 | 0.001 | 4.937 | 0.026 |
| EL-070 | −0.012 | 0.004 | 9.415 | 0.002 |
| EL-100 | −0.008 | 0.004 | 4.707 | 0.030 |
| SL-005 | −0.008 | 0.002 | 14.181 | 0.000 |
| SL-015 | −0.009 | 0.003 | 12.599 | 0.000 |
| SL-020 | −0.009 | 0.002 | 14.522 | 0.000 |
| SL-025 | −0.012 | 0.002 | 25.215 | 0.000 |
| SL-030 | −0.012 | 0.002 | 25.246 | 0.000 |
| SL-035 | −0.017 | 0.003 | 350.001 | 0.000 |
| SL-040 | −0.010 | 0.004 | 70.099 | 0.008 |
| SL-045 | −0.032 | 0.007 | 19.203 | 0.000 |
| SL-060 | −0.069 | 0.013 | 27.996 | 0.000 |
| SL-075 | −0.051 | 0.012 | 18.181 | 0.000 |
| SL-090 | −0.053 | 0.011 | 22.773 | 0.000 |

*PS = proximity to lahar source zone; PD = proximity to drainage lines that head on lahar source zone; EL = elevations outside lahar source zone; SL = slopes outside lahar source zone; numbers indicate upper bounds of 5-percentile classes (see Tables II and/or IV).

lahar source (or *PLI*) zone. This suggests that areas $<6$ km of lahar source (or *PLI*) zone are less susceptible to *DLI* than areas $<6$ km of lahar source (or *PLI*) zone. Coefficients for classes of proximity to drainage lines that head on lahar source (or *PLI*) zone in final LR model are positive, which indicates positive spatial association between pixels of "pre-eruption" *DLI* zones and pixels at distances $<2$ km of drainage lines that head on lahar source (or *PLI*) zone. This suggests that areas $<2$ km of drainage lines that head on lahar source (or *PLI*) zone are more susceptible to *DLI* than areas $>2$ km of drainage lines that head on lahar source (or *PLI*) zone.

Only five out of 20 elevation classes remain in final LR model; these elevation classes pertain to almost full range of elevations outside lahar source (or *PLI*) zone and their coefficients are negative. These suggest that (a) elevation is a minor factor of *DLI* or (b) *DLI* could occur in any elevation. Eleven out of 20 slope classes remain in final model; these slope classes are mostly $<9°$ and their coefficients are negative. These suggest that (a) slope is a non-trivial factor of *DLI* and (b) gentler slopes are generally more susceptible to *DLI* than steeper slopes.

Final LR model has *pseudo*-$R^2$ of $>99\%$, which indicate a very satisfactory goodness-of-fit with "pre-eruption" *DLI* zones. Figure 8 shows classified map of $S_{DLI}$ values based on LR modeling. Zones of "very low" $S_{DLI}$ values occupy about 79% of area outside zone of lahar source. Each of other zones occupies about 5.2% of area outside lahar source (or *PLI*) zone. (Re-classified binary predictive map of *DLI* zones based on LR modeling is not shown here because of space.)

Number of pixels in overlap between mapped *DLI* zones and predicted *DLI* zones is 168,447, whereas number of pixels in overlap between mapped non-*DLI* zones and predicted non-*DLI* zones is 2,289,600 (Table V). The LR model has 56.3% "true positive" prediction accuracy and 89.0% "true negative" prediction accuracy. In addition, the LR model has 43.7% Type I prediction error and 11.0% Type II prediction error. Thus, the LR model has low "true positive" prediction accuracy, high "true negative" prediction accuracy, high Type I prediction error, and low Type II prediction error. These show, however, that the LR model of predicted *DLI* zones is superior to the forecasted *DLI* zones in terms of "true positive" prediction accuracy and Type I prediction error, although it is inferior to the forecasted *DLI* zones in terms of "true negative" prediction accuracy and Type II prediction error.

### 5.2.3. *Evidential Belief Model*

Table IV shows variations of values of *Bel*, *Dis* and *Unc* for classes of spatial factors with respect to "pre-eruption" *DLI* zones. Proximity classes

$<4$ km of lahar source (or *PLI*) zone have higher *Bel* and lower *Dis* and *Unc* than most proximity classes $>4$ km of lahar source (or *PLI*) zone. Proximity classes $<1.5$ km of drainage lines that head on lahar source (or *PLI*) zone have higher values of *Bel* and lower values of *Dis* and *Unc* than proximity classes $>1.5$ km of drainage lines that head on lahar source (or *PLI*) zone. These imply that areas proximal to lahar source (or *PLI*) zone and to drainage lines that head on lahar source (or *PLI*) zone (a) have positive spatial associations with "pre-eruption" *DLI* zones and (b) are more susceptible to *DLI* than areas that are distal. Values of *Bel* and *Unc* generally increase and decrease, respectively, with decreasing elevation and decreasing slope. This suggests that (a) higher elevations and steeper slopes have negative spatial associations with *DLI* zones and (b) lower elevations and gentler slopes are more susceptible to *DLI*, respectively, than higher elevation and steeper slopes. Overall, the maximum values of *Bel* and minimum values of (*Unc* for each spatial factor indicate that the proximity variables are more important spatial factors of *DLI* than the topographic variables).

Classified map of $S_{DLI}$ values based on integrated values of *Bel* is displayed in Figure 9. Zones of "very low" susceptibility occupy about 28% of area outside lahar source (or *PLI*) zone. Each of other zones occupies about 18% of area outside lahar source (or *PLI*) zone. (Re-classified binary predictive map of *DLI* zones based on EBF modeling is not shown here because of space.)

Number of pixels in overlap between mapped *DLI* zones and predicted *DLI* zones is 279,680, whereas number of pixels in overlap between mapped non-*DLI* zones and predicted non-*DLI* zones is 1,308,194 (Table V). The EBF model has "true positive" prediction accuracy of 93.4% and "true negative" prediction accuracy of 50.9%. In addition, the EBF model has Type I prediction error of 6.6% and Type II error of 49.1%. Thus, the EBF model has very high "true positive" accuracy, low "true negative" accuracy, very low Type I error, and high Type II error. These show that the EBF model is superior to the forecast hazard map in terms of "true positive" prediction accuracy and Type I prediction error, although it is inferior to the forecast hazard map in terms of "true negative" prediction accuracy and Type II prediction error.

## 6. Discussion

### 6.1. ON PREDICTION OF LAHAR SOURCE (OR *PLI*) ZONE

Our work on predicting lahar source (or *PLI*) zone follows the "*H*/*L*" technique developed by previous workers (e.g., Malin and Sheridan, 1982; Sheridan and Malin, 1983). Thus, in this regard, we present nothing new except

*Table IV*. Estimates of EBF for classes of spatial factors: (A) proximity to lahar source zone; (B) proximity to drainage lines that head on lahar source zone; (C) elevations outside lahar source zone; (D) slopes outside lahar source zone.

| (A) Proximity to lahar source zone | | | | | | (B) Proximity to drainage that head on lahar source zone | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5-percentile classes (km) | Number of class pixels | Number of distal lahar pixels | *Bel* | *Dis* | *Unc* | 5-percentile classes (km) | Number of class pixels | Number of distal lahar pixels | *Bel* | *Dis* | *Unc* |
| 0.00–0.55 | 143,391 | 37,570 | 0.284 | 0.039 | 0.677 | 0.00–0.10 | 142,651 | 80,746 | 0.655 | 0.022 | 0.323 |
| 0.55–1.44 | 143,391 | 25,825 | 0.177 | 0.043 | 0.780 | 0.10–0.25 | 143,510 | 38,375 | 0.196 | 0.038 | 0.766 |
| 1.44–2.36 | 143,344 | 140,47 | 0.088 | 0.048 | 0.864 | 0.25–0.43 | 141,242 | 22,103 | 0.100 | 0.044 | 0.856 |
| 2.36–3.29 | 143,912 | 7,025 | 0.042 | 0.050 | 0.908 | 0.43–0.67 | 145,502 | 9,190 | 0.037 | 0.049 | 0.914 |
| 3.29–4.20 | 142,861 | 4,795 | 0.028 | 0.051 | 0.921 | 0.67–0.93 | 144,423 | 2,566 | 0.010 | 0.052 | 0.938 |
| 4.20–5.09 | 143,939 | 2,375 | 0.014 | 0.052 | 0.934 | 0.93–1.21 | 143,326 | 345 | 0.001 | 0.053 | 0.946 |
| 5.09–5.94 | 143,675 | 2,061 | 0.012 | 0.052 | 0.936 | 1.21–1.51 | 143,128 | 141 | 0.001 | 0.053 | 0.946 |
| 5.94–6.76 | 143,685 | 1,129 | 0.006 | 0.053 | 0.941 | 1.51–1.84 | 143,985 | 9 | 0.000 | 0.053 | 0.947 |
| 6.76–7.57 | 142,997 | 1,049 | 0.006 | 0.053 | 0.941 | 1.84–2.19 | 142,601 | 0 | 0.000 | 0.053 | 0.947 |
| 7.57–8.44 | 144,100 | 2,109 | 0.012 | 0.052 | 0.936 | 2.19–2.54 | 144,294 | 0 | 0.000 | 0.053 | 0.947 |
| 8.44–9.40 | 143,598 | 1,955 | 0.011 | 0.052 | 0.937 | 2.54–2.92 | 143,788 | 0 | 0.000 | 0.053 | 0.947 |
| 9.40–10.43 | 143,257 | 6,524 | 0.039 | 0.050 | 0.911 | 2.92–3.30 | 143,237 | 0 | 0.000 | 0.053 | 0.947 |
| 10.43–11.53 | 143,551 | 10,666 | 0.065 | 0.049 | 0.886 | 3.30–3.70 | 143,950 | 0 | 0.000 | 0.053 | 0.947 |
| 11.53–12.68 | 143,898 | 7,146 | 0.043 | 0.050 | 0.907 | 3.70–4.17 | 143,917 | 0 | 0.000 | 0.053 | 0.947 |
| 12.68–13.89 | 143,466 | 3,294 | 0.019 | 0.052 | 0.929 | 4.17–4.69 | 142,146 | 0 | 0.000 | 0.053 | 0.947 |
| 13.89–15.12 | 143,475 | 2,323 | 0.013 | 0.052 | 0.935 | 4.69–5.25 | 144,748 | 0 | 0.000 | 0.053 | 0.947 |
| 15.12–16.38 | 143,598 | 2,559 | 0.015 | 0.052 | 0.933 | 5.25–5.89 | 143,228 | 0 | 0.000 | 0.053 | 0.947 |
| 16.38–17.83 | 143,694 | 6,494 | 0.039 | 0.050 | 0.911 | 5.89–6.57 | 143,657 | 0 | 0.000 | 0.053 | 0.947 |
| 17.83–19.99 | 143,415 | 5,054 | 0.030 | 0.051 | 0.919 | 6.57–7.48 | 144,040 | 0 | 0.000 | 0.053 | 0.947 |
| 19.99–25.93 | 143,903 | 9,475 | 0.057 | 0.049 | 0.894 | 7.48–12.32 | 143,777 | 0 | 0.000 | 0.053 | 0.947 |

*Table IV*. Continued

| (C) Elevations outside lahar source zone | | | | | | (D) slopes outside lahar source zone | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5-percentile classes (degrees) | Number of class pixels | Number of distal lahar pixels | *Bel* | *Dis* | *Unc* | 5-percentile classes (m) | Number of class pixels | Number of distal lahar pixels | *Bel* | *Dis* | *Unc* |
| 690–1340 | 145,489 | 879 | 0.005 | 0.053 | 0.942 | 31.51–69.19 | 143,444 | 545 | 0.003 | 0.053 | 0.944 |
| 555–690 | 141,855 | 427 | 0.003 | 0.053 | 0.944 | 27.54–31.51 | 143,605 | 394 | 0.002 | 0.053 | 0.945 |
| 480–555 | 147,943 | 340 | 0.002 | 0.053 | 0.945 | 24.64–27.54 | 143,222 | 404 | 0.002 | 0.053 | 0.945 |
| 420–480 | 161,270 | 624 | 0.003 | 0.053 | 0.944 | 22.24–24.64 | 143,982 | 470 | 0.003 | 0.053 | 0.944 |
| 380–420 | 133,979 | 910 | 0.006 | 0.053 | 0.941 | 20.14–22.24 | 143,165 | 458 | 0.003 | 0.053 | 0.944 |
| 340–380 | 135,034 | 1,079 | 0.007 | 0.053 | 0.940 | 18.17–20.14 | 144,509 | 522 | 0.003 | 0.053 | 0.944 |
| 300–340 | 1,39,478 | 2,865 | 0.017 | 0.052 | 0.931 | 16.33–18.17 | 142,718 | 648 | 0.004 | 0.053 | 0.943 |
| 255–300 | 144,059 | 3,721 | 0.022 | 0.052 | 0.926 | 14.57–16.33 | 143,715 | 920 | 0.005 | 0.053 | 0.942 |
| 214–255 | 143,055 | 4,698 | 0.028 | 0.051 | 0.921 | 12.84–14.57 | 143,480 | 895 | 0.005 | 0.053 | 0.942 |
| 180–214 | 147,008 | 9,747 | 0.059 | 0.049 | 0.892 | 11.04–12.84 | 143,502 | 1,245 | 0.007 | 0.052 | 0.941 |
| 140–180 | 161,220 | 11,293 | 0.062 | 0.049 | 0.889 | 9.03–11.04 | 143,902 | 2,444 | 0.014 | 0.052 | 0.934 |
| 115–140 | 122,415 | 19,490 | 0.153 | 0.044 | 0.803 | 6.54–9.03 | 143,135 | 4,903 | 0.029 | 0.051 | 0.920 |
| 88–115 | 143,883 | 22,684 | 0.154 | 0.044 | 0.802 | 3.69–6.54 | 143,515 | 10,359 | 0.064 | 0.049 | 0.887 |
| 60–88 | 146,984 | 12,854 | 0.079 | 0.048 | 0.873 | 1.85–3.69 | 143,822 | 24,935 | 0.171 | 0.043 | 0.786 |
| 42–60 | 139,764 | 7,651 | 0.048 | 0.050 | 0.902 | 1.01–1.85 | 145,369 | 28,239 | 0.197 | 0.042 | 0.761 |
| 27–42 | 143,517 | 6,673 | 0.040 | 0.050 | 0.910 | 0.58–1.01 | 147,124 | 16,081 | 0.101 | 0.047 | 0.852 |
| 18–27 | 143,463 | 3,843 | 0.023 | 0.052 | 0.925 | 0.37–0.58 | 131,108 | 11,509 | 0.079 | 0.048 | 0.873 |
| 13–18 | 143,709 | 5,180 | 0.031 | 0.051 | 0.918 | 0.26–0.37 | 135,428 | 7,653 | 0.049 | 0.050 | 0.901 |
| 7–13 | 143,455 | 13,006 | 0.082 | 0.048 | 0.870 | 0.17–0.26 | 148,857 | 15,277 | 0.094 | 0.047 | 0.859 |
| 0–7 | 143,570 | 25,511 | 0.177 | 0.043 | 0.780 | 0.00–0.17 | 153,548 | 25,574 | 0.165 | 0.044 | 0.791 |

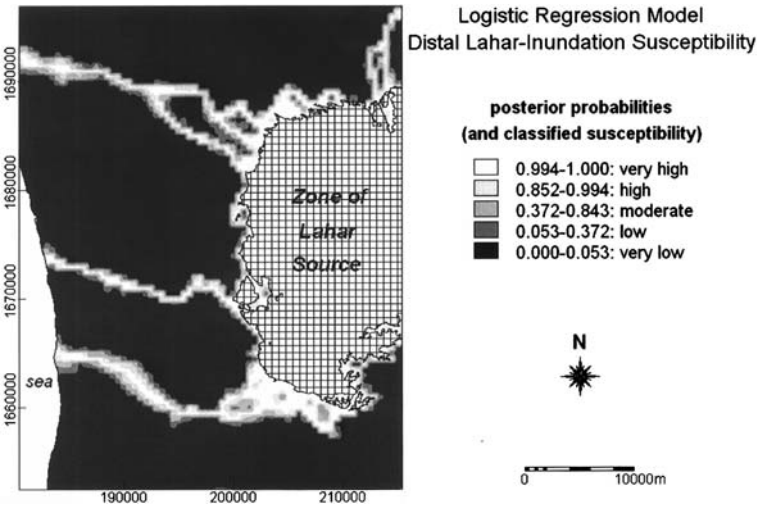


*Figure 8*. Probabilistic map of distal lahar-inundation zones derived by LR modeling.

to demonstrate further this technique in the test area and to use predicted boundary of predicted *PLI* zone in predicting *DLI* zones. With respect to a field-based map of pyroclastic-flow deposits (Figure 3), which are known sources of most lahars in test area, predictive map of *PLI* zone has very high "true positive" prediction accuracy, high "true negative" prediction accuracy, very low Type I prediction error and low Type II prediction error. These results further demonstrate utility of "$H/L$" technique to predict *PLI* zone. However, it is emphasized here that delineation of *PLI* zone requires baseline geological information to determine "$H/L$" ratios appropriate for type of eruptive deposits that spawn lahars in certain volcanoes (e.g., Waitt *et al*., 1995; Waythornas and Miller, 1999; Scott *et al*., 2001).

### 6.2. ON PREDICTION OF *DLI* ZONES

Our methodology to predict *DLI* zones is similar but different to LAHARZ modeling (Iverson *et al*., 1998; Schilling, 1998). The main similarities lean on, aside from thematic objective, three aspects: (a) non-application of physics; (b) application of DEM; and (c) application of GIS. As these similarities are obvious, the focus here is to explain the differences. The main differences also lean on three aspects: (a) application of DEM; (b) type of input data to predictive modeling; and (c) type of output of predictive modeling. In LAHARZ modeling, a DEM is used to generate outputs. In our predictive modeling, a DEM is used to derive spatial data to be used as inputs. In LAHARZ modeling, input data consist of hypothetical volumes of lahars. In our predictive modeling, input data consist of a map of training data and maps of topographic factors relevant for

*Table V*. Error matrices for validation of forecast map of *DLI* zones and binary predictive models of *DLI* zones against field-based map of *DLI* zones.

| | | Field-based map of *DLI* zones (Figure 3) | |
|---|---|---|---|
| | | *DLI* zones | non-*DLI* zones |
| Forecast map of *DLI* zones (Figure 2) | *DLI* zones | 137,477 | 100,804 |
| | Non-*DLI* zones | 161,871 | 2,470,998 |
| WofE model of *DLI* zones | *DLI* zones | 235,085 | 416,667 |
| | Non-*DLI* zones | 64,263 | 2,155,135 |
| LR model of *DLI* zones | *DLI* zones | 168,447 | 282,202 |
| | Non-*DLI* zones | 130,901 | 2,289,600 |
| EBF model of *DLI* zones | *DLI* zones | 279,680 | 1,263,608 |
| | Non-*DLI* zones | 19,668 | 1,308,194 |

Values in table are number of pixels.

*DLI*. In LAHARZ modeling, outputs are planimetric and cross-sectional areas of *DLI* zones. In our predictive modeling, outputs are probabilities of *DLI* zones.

It is type of output in prediction of *DLI* zones that motivated us to demonstrate applications of probabilistic spatial modeling techniques described above. This is because Iverson *et al*. (1998, p. 980) stressed desirability to provide quantitative and probabilistic analysis of lahar-inundation zones,

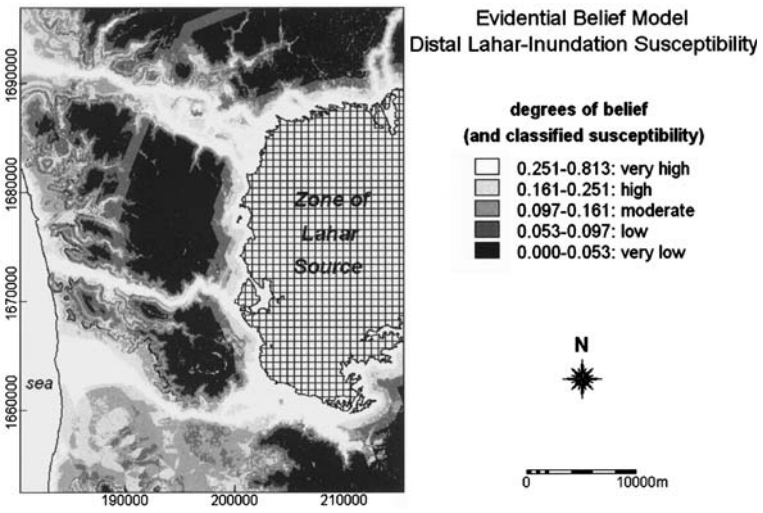


*Figure 9*. Probabilistic map of distal lahar-inundation zones derived by EBF modeling.

The probabilistic techniques used here were developed for and are commonly applied to mineral potential mapping (e.g., Bonham-Carter *et al.*, 1988, 1989; Moon, 1989, 1990; Agterberg and Bonham-Carter, 1990, 1999; Agterberg, 1992; An, 1992; An *et al.*, 1992, 1994; Agterberg *et al.*, 1993; Bonham-Carter, 1994; Carranza and Hale, 2000, 2001, 2003; Carranza, 2002). Similar spatial modeling techniques have also been applied to natural hazard zonation studies (e.g., Mark and Ellen, 1995; Binaghi *et al.*, 1998; Connor *et al.*, 2001; Rupert *et al.*, 2003) but not specifically to predictive mapping of *DLI* zones. It is pointed out here, however, that estimates of $S_{DLI}$ values by applications of techniques presented do not pertain to distal lahar recurrence probabilities, as these are, according to Iverson *et al.* (1998, p. 980), assessed most logically based on lahar volumes. Nonetheless, estimates of $S_{DLI}$ values by applications of techniques presented enable mapping and ranking of zones with varying degrees of $S_{DLI}$.

### 6.3. ON ACCURACY AND ERROR OF PREDICTIVE MODELS OF *DLI* ZONES

With respect to the field-based hazard map (Figure 3) and in terms of predicting *DLI* zones, the three predictive models have much higher ''true positive' prediction accuracy and much lower Type I prediction error than the forecast hazard map (Figure 2). However, with respect to the field-based hazard map (Figure 3) and in terms of predicting non-*DLI* zones, the three predictive model have lower ''true negative'' prediction accuracy and higher Type II prediction error than the forecast hazard map. These results indicate that the forecast map under-estimated the *DLI* zones, whereas the three predictive models over-estimated the *DLI* zones. The results further suggest that, if later field-based maps of *DLI* zones were used here for validation (i.e., Figure 3), then prediction accuracy (both ''true positive'' and ''true'' negative) and prediction error (both Type I and Type II) of the three predictive models would probably be even higher and lower, respectively, than the accuracy and error estimated here. However, validation of the predictive models of *DLI* zones with an immediate post-eruption field-based map of *DLI* zones is considered appropriate than validation with field-based maps of *DLI* zones much later after the eruption because of the following reason. The predictive models of *DLI* zones are based on a model of pre-eruption topographic surface, which was probably still much similar to the immediate post-eruption topographic surface over which distal lahars flowed and were deposited.

The predictive model with highest ''true positive'' prediction accuracy and lowest Type I prediction error is the EBF model, followed by the WofE model, and then by the LR model. This is because the EBF model resulted in highest number of *DLI* pixels, the WofE model resulted in second highest number of *DLI* pixels, and the LR model resulted in lowest

number of *DLI* pixels. The predictive model with highest "true negative" prediction accuracy and lowest Type II prediction error is the LR model, followed by the WofE model, and then by the EBF model, This is because the LR model resulted in highest number of non-*DLI* pixels, the WofE model resulted in second highest number of non-*DLI* pixels, and the EBF model resulted in lowest number of non-*DLI* pixels. These results suggest the following. The LR model could represent a "poor-case" *DLI* hazard scenario; that is, highest under-estimation of *DLI* zones and lowest over-estimation of non-*DLI* zones. The WofE model could represent a "worse-case" *DLI* hazard scenario; that is, moderate under-estimation of *DLI* zones and moderate over-estimation of non-*DLI* zones. The EBF model could represent a "worst-case" *DLI*. hazard scenario; that is, lowest under-estimation of *DLI* zones and highest over-estimation of non-*DLI* zones. Thus, with the three proposed techniques for predictive mapping of *DLI* zones, different *DLI* hazard scenarios could be depicted and then provided as inputs to lahar hazard mitigation planning.

It was not, however, the intention to show which of three spatial modeling techniques would provide the best predictive *DLI* model, because it cannot be guaranteed that the performance of each of these techniques in the test area would be the same as in other areas. Our intention here was to show potential alternative techniques that could be useful in providing probabilistic analysis of *DLI* zones, In this regard, we discuss further ramifications of these techniques to prediction of *DLI* zones.

### 6.4. ON RAMIFICATIONS OF PROBABILISTIC MAPPING OF *DLI* ZONES

The techniques presented here to predict *DLI* zones are data-driven and thus sensitive to areas where they are applied, because spatial associations between training data and spatial factors vary from one area to another. Consequently, outputs of probabilistic techniques used here depend on (a) quality and quantity of training data and (b) quality and quantity of spatial factor maps.

Ideally, training data to be used in predicting *DLI* zones, using the techniques presented here, should consist of field-based maps of past *DLI* zones, because spatial accuracy of such maps are known, or verifiable. However, such type of baseline data might be lacking in areas where a volcano has not erupted within living memory. Thus, here, we used training data of "pre-eruption" *DLI* zones, which were calculated using a DEM. However, we only recommend using same type of training data as surrogate if baseline geological information concerning past *DLI* zones are lacking. Nonetheless, the fair to very good validation results for the predictive *DLI* models suggest that type of training data used here is applicable to volcanic areas where little or no geological information is available. Alter-

natively, the results indicate possibility of using *DLI* zones delineated by LAHARZ modeling as training data in the probabilistic techniques presented here. Whether real or surrogate training data are used, quantified spatial associations between training data and spatial factor maps could be statistically non-significant (i.e., meaningless) if the former is insufficient. The number of pixels of training data used here (Figure 4) is about half as small as number of mapped lahar deposits and lahar-impacted channels (Figure 3). However, estimates of $W_{C_{ij}}$ in WofE modeling and $a_{C_{ij}}$ coefficients in LR modeling are statistically significant (Tables II and III), which indicate that quantity of training data used here was sufficient to obtain statistically significant results.

Quality, in terms of spatial accuracy, of factor maps certainly influences outputs of techniques presented although we did not evaluate quantitatively this aspect here as this relates to accuracy of pre-eruption DEM (see further below). It is, therefore, emphasized here that estimates of $S_{DLI}$ values (Equation (2)) derived by applications of techniques presented are pseudo-probabilities because these were calculated by using (a) training data that are not real *DLI* zones and (b) some spatial factor data that do not carry sufficient information about training data. From the results (Tables II–IV), it is shown that elevation and slope do not carry sufficient information about (i.e., they have negative spatial associations with) training data. In applying these techniques to mineral potential mapping (see references cited above), maps of geological factors that show negative spatial associations with mineral occurrences are practically not used to delineate potentially-mineralized; i.e., negative spatial associations means negative evidence. Here, in spite of their negative spatial associations with the training data, elevation and slope were still used because, in western parts of test area, *DLI* is influenced by elevation and slope rather than by proximity to lahar source zone. That is, in western parts of test area, *DLI* zones are mostly lowlands and floodplains that are distal to boundary of lahar source zone. However, in a volcanic terrain one could expect that good correlation exists between elevation and slope, such that using both of them is tantamount to using redundant spatial factor maps and that using only one of them would be sufficient to obtain meaningful results. At this point, we digress to issue of quantity of spatial factor maps but revisit subsequently issue of redundant spatial factor maps.

In GIS-based predictive modeling, it is certainly desirable to make use of several layers of spatial data depicting individual factors that influence occurrence of geographic phenomena of interest. However, in applications of the spatial modeling techniques presented, large quantity (in terms of number of map layers) of spatial factor data could also be undesirable. This is because, as number of map layers of spatial factor data increases, problem of lack of CI among map layers with respect to training data

increases. The problem of CI, however, can be solved by combining maps of spatial factors that show lack of CI with respect to training data. There are a number of ways to do this and either principal components analysis or factor analysis is robust (Davis, 1973). Combining maps of spatial factors showing lack of CI with respect to training data should be combined according to a theme. Thus, going back to issue of redundancy of elevation and slope in this study, these two spatial factors may be combined into a ''landscape'' factor. However, when a new spatial factor map is derived by combining maps of spatial factors that lack CI with respect to training data, posterior probabilities and degrees of belief have to be re-estimated. Concerning problem of lack of CI among spatial factor maps with respect to training data, the results show that modeling with either LR or EBF is advantageous over WofE modeling. Alternatively, there are other probabilistic techniques that are not seriously affected by lack of CI among spatial factor maps with respect to training data. One such robust probabilistic technique is application of neural networks, which is also commonly used in mineral exploration (e.g., Harris *et al*., 2003; Porwal *et al*., 2003) but is becoming useful to volcanic-hazard studies (e.g., Bertucco *et al*., 1999; De Falco *et al*., 2002). Application of neural networks to predictive delineation of *DLI* zones is yet to be demonstrated; however, its utility would depend on sufficient training data about not only known past *DLI* zones but also about zones ''never'' inundated or have very low probability of being inundated by lahars. For this reason, we have not applied neural work to the present case study.

### 6.5. ON LIMITATIONS OF PRE-ERUPTION OEM

It has been shown by Stevens *et al*. (2002) that predictive modeling of volcanic flows such as lahars is sensitive to spatial accuracy of a DEM. However, the fair to very good validation results of the predictive models created indicate that spatial (horizontal and vertical) accuracy of DEM used here is adequate for mapping of zones that could, in general, be inundated by multiple lahars. Nonetheless, we recognize that DEMs of high spatial and temporal accuracy is crucial in predicting lahar-inundation zones. On one hand, changes in topography due to erosion and/or deposition of volcanic materials can alter lahar flow paths along drainage lines that head on lahar source zone. On the other hand, structures built to mitigate lahar hazards or other cultural features that impede drainage could also alter lahar flow paths. These imply that predictive models of lahar-inundation zones based on spatial data derived from a DEM have to be updated whenever recent DEMs become available. Modern methods of topographic data acquisition (e.g., laser altimetry, radar interferometry) are now available to create spatially and temporally accurate DEMs.

## 7. Conclusions

(1) A DEM is a potentially valuable source of spatial data that can be used in predicting lahar-inundation zones in areas where a volcano has not erupted within living memory and where baseline geological information is lacking. However, predictive models of lahar-inundation zones based on the techniques presented have to be up-to-dated whenever DEMs of higher spatial and temporal accuracy become available.
(2) Potential zone of lahar source (or *PLI* zone) can be determined by using a ratio of vertical descent to horizontal run-out that is typical of eruptive deposits that spawn lahars. By this technique, predictive map of *PLI* zone has very high "true positive" prediction accuracy, high "true negative" prediction accuracy, very low Type I prediction error and low Type II prediction error, which are similar to those of the forecast volcanic-hazard map.
(3) Zones susceptible to *DLI* can be determined by using spatial data derived from a DEM. Spatial training data derived from a DEM consist of "pre-eruption" *DLI* zones. Spatial factor data derived from a DEM or a DEM-derived product include; (a) proximity to lahar source zone, (b) proximity to drainage lines that head on lahar source zone, (e) elevation and (d) slope.
(4) The probabilistic techniques used to predict zones susceptible to *DLI* involve modeling by weights-of-evidence, by logistic regression and by evidential belief functions. These techniques concern quantification of spatial associations between training data and spatial factors in order to interpolate/extrapolate from layers of spatial factors unknown (or future) *DLI* zones. Performance of these techniques used depend on (a) study area, (b) quality and quantity of training data and (c) quality and quantity of spatial factor data. The output of these probabilistic techniques are *pseudo*-probabilities of *DLI*, which nonetheless enable mapping and ranking of zones that could, in general, be susceptible to *DLI*.
(5) The predictive models of *DLI* zones have higher "true positive" prediction accuracy and lower Type I prediction error but lower "true negative" and higher Type II prediction error as compared to forecast map of mudflow zones. The lower "true negative" prediction accuracy and higher Type II prediction error of the predictive models as compared to the forecast model, however, do not negate usefulness of the proposed techniques. This is because over-estimation of lahar-inundation hazard, rather than under-estimation of lahar-inundation hazard, could be potentially beneficial in hazard mitigation planning to save lives and properties.

(6) Results of the study further indicate that the techniques presented to predict *DLI* zones could be useful to predict *DLI* zones in situations where [mapped] potential lahar sources have not yet been re-worked to lahars. Results of the study suggest that the techniques presented to predict *DLI* zones are applicable to other volcanic areas that are as prone to *DLI* as the environs of Pinatubo volcano. However, whether or not the techniques described to predict *DLI* zones are applicable to volcanic areas not as prone to *DLI* as environs of Pinatubo volcano needs further testing.
(7) Results of study indicate that predictive lahar-inundation maps derived from applications of presented predictive techniques can be considered, howbeit interpreted with caution, in decision-making to mitigate lahar hazards in areas where a volcano has not erupted within living memory and where baseline geological information is lacking.

**Acknowledgments**

Parts of this work are related to O.T. Castro's M.Sc., research in Geoinformatics at ITC with funding from the Netherlands Fellowship Programme. We thank three anonymous reviewers for their critical comments, which helped greatly to improve an earlier version of the manuscript.

**Appendix A: Spatial Neighborhood Calculations**

Spatial neighborhood calculations in ILWIS uses a $3 \times 3$ kernel filter, which is moved over a raster map. Pixels in this filter are coded as follows:

| | | |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 6 |
| 7 | 8 | 9 |

Result of a calculation is stored in central pixel of 3×3 kernel in output map. If a neighborhood operation is performed on pixels on top or bottom line or on very first or last column of a raster map, new neighbors are created by duplicating this boundary line or column. General syntax of iterative spatial neighborhood calculation is: *Map2* = MAPITERP (*Map1*#(*iterexpr*)); where *Map2* is output raster map, *Map1* is input raster

map, # is neighborhood operator indicating all neighbors are to be used and *iterexpr* is iteration expression until pixels do not change anymore in output map.

Calculation of "pre-eruption" distal lahar-inundation zones and drainage lines that head on boundary of lahar source zone are performed once a DEM without pits, i.e., pixels with values lower than all its surrounding pixels, is obtained. Spatial neighborhood expression for mapping pits in a DEM is $Pit = \text{NBMINP(DEM\#)} = 5$, where *Pit* is output map, NBMINP is neighborhood operator that returns in output map position of neighbor with smallest value. The term (DEM#) = 5 means that, if on a DEM, central pixel of 3× 3 kernel filter has smallest value, then central pixel is considered a pit. Result of this operation is a Boolean map in which pits are labeled "True", otherwise "False". Pits in a DEM are filled automatically (see http://www.itc.nl/ilwis/general_info/new_in_32.asp).

## Calculation of "Pre-eruption" Distal Lahar-Inundation Zones from Given "Start-Points"

Calculation uses as inputs a *start* map (in which pixels representing "start-points" of distal lahar-inundation are assigned a value of 1; otherwise, a value of 0). Syntax for calculating map of distal lahar-inundation channels is

$$Map2 = \text{MAPITERP}(iff(start, start, \text{NBMAX}(start\#, (\text{DEM}\#\text{DEM})))).$$

This expression means that, if (denoted by *iff*) central pixel, in a $3 \times 3$ kernel, over *start* map has a value of 1, it is unchanged. Otherwise, neighbor pixels of central pixel with value equals 1 (denoted by operator NBMAX) is used in the calculation provided that this neighbor pixels has an elevation greater than elevation of central pixel (denoted by expression DEM# > DEM). This means that, if a neighbor pixel belongs to a distal lahar-inundation zone, then central pixel will only be labeled a part of the distal lahar-inundation zone if neighbor pixel is topographically higher than central pixel. This expression can also be used to model lava flows, debris flows and pollution dispersion from a given source.

## Calculation of Drainage Lines from Given "Start-Points"

The first operation for this procedure to create a map of flow directions, calculated by the expression:

$$Flowdir = 10 - \text{NBMINP(DEM\#)}.$$

The result is a raster map with values ranging from 1 to 9. For example, a value of 1 means that flow direction towards central pixel is from northwest (see coding above). With a *Flowdir* map and the same *start* map as described above, drainage lines can be tracked from given "start-points" by the calculation expression:

$$Map2 = \mathrm{MAPITERP}(iff(start, start, \mathrm{NBMAX8}(\#, (Flowdir\# = \mathrm{NBPOS})))).$$

This expression means that, if (denoted by *iff*) central pixel, in a $3 \times 3$ kernel, over start map has a value of 1, it is unchanged. Otherwise, if, in $3 \times 3$ kernel, value of a neighbor pixel in *start* map excluding central pixel (denoted by NBMAX8) equals neighbor position in *Flowdir* map (denoted by *Flowdir#* = NBPOS, meaning that flow from that pixel is towards central pixel), then value of central pixel in *start* map is replaced with value of that neighbor. If more than one neighbor pixels in *start* map satisfy the condition, then value of neighbor corresponding to pixel in *Flowdir* map with highest value is used (denoted by NBMAX8(*start#*,(*Flowdir#* = NBPOS))).

## References

Agterberg, F. P.: 1992, Combining indicator patterns in weights of evidence modeling for resource evaluation, *Nonrenew. Resour.* **1**, 39–50.

Agterberg, F. P. and Bonham-Carter, G. F.: 1990, Deriving weights of evidence from geoscience contour maps for the prediction of discrete events. In: *Proceedings 22nd APCOM (Applications of Computers and Mathematics in the Mineral Industries) Symposium*, Berlin, Germany, September 17–21, 1990, Technical University of Berlin, pp. 381–395.

Agterberg, F. P. and Bonham-Carter, G. F.: 1999, Logistic regression and weights of evidence modeling in mineral exploration. In: *Proceedings 28th APCOM (Applications of Computers and Mathematics in the Mineral Industries) Symposium*, Golden, CO, USA, 20–22 October 1999, Colorado School of Mines, pp. 483–490.

Agterberg, F. P., Bonham-Carter, G. F., Cheng, Q., and Wright, D. F.: 1993. Weights of evidence modeling and weighted logistic regression for mineral potential mapping. In J. C. Davis & U. C. Herzfeld (Eds.), *Computers in Geology – 25 Years of Progress* (pp. 13–32). New York: Oxford University Press.

An, P.: 1992, *Spatial Reasoning and Integration Techniques for Geophysical and Geological Exploration Data*, Ph.D. thesis, University of Manitoba, Canada, 280 pp.

An, P., Moon, W. M., and Bonham-Carter, G. F.: 1992, On knowledge-based approach on integrating remote sensing, geophysical and geological information. In: *Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS)*, 1992, pp. 34–38.

An, P., Moon, W. M., and Bonham-Carter, G. F.: 1994, An object-oriented knowledge representation structure for exploration data integration, *Nonrenew. Resour.* **3**, 132–145.

Bertucco, L., Coltelli, M., Nunnari, G., and Occhipinti, L.: 1999, Cellular neural network for real-time monitoring of volcanic activity, *Comput. Geosci.* **25**, 101–117.

Binaghi, E., Luzi, L., Madella, P., Pergalani, F., and Rampini, A.: 1998, Slope instability zonation: A comparison between certainty factor and fuzzy Dempster–Shafer approaches, *Nat. Hazards* **17**, 77–97.

Bishop, M. M., Feinberg, S. E., and Holland, P. W.: 1975, *Discrete Multivariate Analysis: Theory and Practice*, MIT Press, Cambridge, MA, pp. 587.

Bonham-Carter, G. F.: 1994, *Geographic Information Systems for Geoscientists: Modelling with GIS*, Pergamon, Ontario, pp. 398.

Bonham-Carter, G. F., Agterberg, F. P., and Wright, D. F.: 1988, Integration of geological datasets for gold exploration in Nova Scotia, *Photogrammet. Eng. Remote Sens.* **54**, 1585–1592.

Bonham-Carter, G. F., Agterberg, F. P., and Wright, D. E.: 1989, Weights of Evidence Modelling: A New Approach to Mapping Mineral Potential. In: F.P. Agterberg and G.F. Bonham-Carter (eds.), *Statistical Applications in the Earth Sciences*, Geological Survey of Canada, Paper 89-9, pp. 171–483.

Canuti, P., Casagli, N., Catani, F., and Falorni, G.: 2002, Modeling of the Guagua Pichincha volcano (Ecuador) lahars, *Phys. Chem. Earth* **27**, 1587–4599.

Carranza, E. J. M.: 2002, *Geologically-Constrained Mineral Potential Mapping: Examples from the Philippines*, Ph.D. thesis, Delft University of Technology, The Netherlands, ITC Publication No. 86, International Institute for Geo-Information Science and Earth Observation (ITC), Enschede, 480 pp.

Carranza, E. J. M. and Hale, M.: 2000, Geologically-constrained probabilistic mapping of gold potential, Baguio district, Philippines, *Nat. Resour. Res.* **9**, 237–253.

Carranza, E. J. M. and Hale, M.: 2001, Logistic regression for geologically-constrained mapping of gold potential, Baguio district, Philippines, *Explorat. Mining Geol.* **10**, 165–175.

Carranza, E. J. M. and Hale, M.: 2003, Evidential belief functions for data-driven geologically-constrained mapping of gold potential, Baguio district, Philippines, *Ore Geol. Rev.* **22**, 117–132.

Connor, C. B., Hill, B. E., Winfrey, B., Franklin, N., and La Femina, P. C.: 2001, Estimation of volcanic hazards from tephra fallout, *Nat. Hazards Rev.* **2**, 33–42.

Costa, J. E.: 1997, Hydraulic modeling for lahar hazards at Cascades Volcanoes, *Environ. Eng. Geosci.* **3**, 21–30.

Cox, D. R. and Snell, E. J.: 1989, *Analysis of Binary Data*, 2Chapman & Hall, London, pp. 236.

Crandell, D. R. and Mullineaux, D. R.: 1975, Technique and rationale of volcanic-inundations appraisals in the Cascade Range, northwestern United States, *Environ. Geol.* **1**, 23–32.

Davis, J. C.: 1973, *Statistics and Data Analysis in Geology*, 2John Wiley and Sons, Singapore, pp. 646.

De Falco, I., Giodano, A., Luongo, G., Mazzarella, A., and Tarantino, E.: 2002, The eruptive activity of Vesuvius and its neural architecture, *J. Volcanol. Geotherm. Res.* **113**, 111–118.

Dempster, A. P.: 1967, Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Stat.* **38**, 325–339.

Dempster, A. P.: 1968, A generalization of Bayesian inference, *J. Roy. Stat. Soc. B* **30**, 205–247.

Goodacre, A. K., Bonham-Carter, G. F., Agterberg, F. P., and Wright, D. F.: 1993, A statistical analysis of the spatial association of seismicity with drainage patterns and magnetic anomalies in western Quebec, *Tectonophysics* **217**, 285–305.

Gorte, B. G. H. and Koolhoven, W.: 1990, Interpolation between isolines based on the Borgefors distance transform, *ITC J.* **1900-3**, 245–247.

Harris, D., Zurcher, L., Stanley, M., Marlow, J., and Pan, G.: 2003, A comparative analysis of favorability mappings by weights of evidence, probabilistic neural networks, discriminant analysis, and logistic regression, *Nat. Resour. Res.* **12**, 241–255.

Hayashi, J. N. and Self, S.: 1992, A comparison of pyroelastic flow and debris avalanche mobility, *J. Geophys. Res.* **97**, 9063–9071.

Iverson, R. M., Schilling, S. P., and Vallance, J. W.: 1998, Objective delineation of lahar-inundation zones, *Geol. Soc. Am. Bull.* **110**, 972–984.

Laenen, A. and Hansen, R. P.: 1988, Simulation of three lahars in the Mount St. Helens area, Washington, using a one-dimensional unsteady-state streamflow model, *U.S. Geological Survey Water-Resources Investigations Report* 88-4004.

Lecointre, J., Hogdson, K., Neall, V., and Cronin, S.: 2004, Lahar-triggering mechanisms and hazard at Ruapehu volcano, New Zealand, *Nat. Hazards* **31**, 85–109.

Macedonio, G. and Paresehi, M. T.: 1992, Numerical simulation of some lahars from Mount St. Helens, *J. Volcanol. Geotherm. Res.* **54**, 65–80.

Major, J. J., Janda, R. J., and Daag, A. S.: 1996, Watershed disturbance and lahars on the east side of Pinatubo volcano during the mid-June 1991 eruptions. In C. G. Newhall & R. S. Punongbayan (Eds.), *Fire and Mud, Eruption and Lahars of Mt. Philippines* (pp. 895–919). Pinatubo: University of Washington Press.

Major, J. J., Schilling, S. P., Pullinger, C. R., Escobar, C. D., and Howell, M. M.: 2001, Volcano-hazard zonation for San Vicente Volcano, El Salvador, *U.S. Geological Survey Open File Report* 01-367.

Malin, M. C. and Sheridan, M. F.: 1982, Computer-assisted mapping of pyroclastic surges, *Science* **217**, 637–640.

Mark, R. K. and Ellen, S. D.: 1995, Statistical and simulation models for mapping debris-flow hazard. In A. Carrara & F. Guzzetti (Eds.), *Geographical Information Systems in Assessing Natural Hazards* (pp. 93–106). Dordrecht: Kluwer Academic Publishers.

Menard, S..: 1995, *Applied Logistic Regression Analysis, Series: Qualitative Applications in the Social Sciences, No. 106*, Sage Publications, Thousand Oaks, CA, pp. 98.

Moon, W. M.: 1989, Integration of remote sensing and geological/geophysical data using Dempster–Shafer approach. In: *Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS)*, 1989, pp. 838–841.

Moon, W. M.: 1990, Integration of geophysical and geological data using evidential belief function, *IEEE Trans. Geosci. Remote Sens.* **28**, 711–720.

Newhall, C. G. and Punongbayan, R. S.: 1996, The narrow margin of successful of volcanic-risk mitigation. In R. Scarpa & R. I. Tilling (Eds.), *Monitoring and Mitigation of Volcano Hazards* (pp. 807–838). Berlin: Springer-Verlag.

Newhall, C. G., Daag, A. S., Delfin, F. G. Jr., Hoblitt, R. P., McGeehin, J., Pallister, J. S., Regalado, M. T. M., Rubin, M., Tubianosa, B. S., Tamayo, R. A. Jr., and Umbal, J. V.: 1996, Eruptive history of Pinatubo volcano. In C. G. Newhall & R. S. Punongbayan (Eds.), *Fire and Mud, Eruption and Lahars of Mt. Philippines* (pp. 165–195). Pinatubo: University of Washington Press.

Paladio-Melosantos, M. L. O., Solidum, R. U., Scott, W. E., Quiambao, R. B., Umbal, J. V., Rodolfo, K. S., Tubianosa, B. S., Delos Reyes, P. J., Alonso, R. A., and Ruelo, H. B.: 1996, Tephra falls of the 1991 eruptions of Pinatubo volcano. In C. G. Newhall & R. S. Punongbayan (Eds.), *Fire and Mud, Eruption and Lahars of Mt Pinatubo Philippines* (pp. 513–535). Seattle, WA: University of Washington Press.

Pan, G. C. and Harris, D. P.: 2000, *Information Synthesis for Mineral Exploration*, Oxford University Press Inc, New York.

Pareschi, M. T.: 1996, Physical modeling of eruptive phenomena lahars. In R. Scarpa & R. I. Tilling (Eds.), *Monitoring and Mitigation of Volcano Hazards* (pp. 463–489). Berlin: Springer-Verlag.

Pierson, T. C., Daag, A. S., Delos Reyes, P. J., Regalado, M. T. M., Solidum, R. U., and Tubianosa, B. S.: 1996, Flow and deposition of posteruption hot lahars on the east side of Pinatubo volcano, July–October 1991. In C. G. Newhall & L. S. Punongbayan (Eds.), *Fire and Mud, Eruption and Lahars of Mt. Philippines* (pp. 921–950). Pinatubo: University of Washington Press.

Pierson, T. C., Janda, R. J., Umbal, J. V., and Daag, A. S.: 1992, Immediate and long-term hazards from lahars and excess sedimentation in rivers draining Pinatubo volcano, Philippines, *U.S. Geological Survey Water Resources Investigations Report* 92-4039.

Pinatubo Volcano Observatory Team: 1991, Lessons from a major eruption: Mt. Pinatubo, Philippines, *Eos, Trans., Am. Geophys. Union* **72**, 545, 552–553, 555.

Porwal, A., Carranza, E. J. M., and Hale, M.: 2003, Artificial neural networks for mineral potential mapping, *Nat. Resour. Res.* **12**, 155–171.

Raines, G. L., Bonham-Carter, G. F., and Kemp, L.: 2000, Predictive probabilistic modeling using ArcView GIS, *ArcUser* April–June 2000, 45–48, http//esri.com/news/arcuser/0400/files/wofe.pdf.

Rodolfo, K. S., Umbal, J. V., Alonso, R. A., Remotigue, C. T., Paladio-Melosantos, M. L., Salvador, J. H. G., Evangelista, D., and Miller, Y.: 1996, Two years of lahars on the western flank of Pinatubo volcano: Initiation, flow processes, deposits, and attendant geomorphic and hydraulic changes. In C. G. Newhall & R. S. Punongbayan (Eds.), *Fire and Mud, Eruption and Lahars of Mt, Pinatubo, Philippines* (pp. 989–1013). Seattle, WA: University of Washington Press.

Rupert, M. G., Cannon, S. H., and Gartner, J. E.: 2003, Using logistic regression to predict the probability of debris flow occurring on areas recently burned by wildland fires, *U.S. Geological Survey Open-File Report* 03-500.

Schilling, S. P.: 1998, LAHARZ: GIS programs for automatic delineation of lahar hazard zones, *U.S. Geological Survey Open-File Report* 98-638.

Schilling, S. P., Vallance, J. W., Mactías, O., and Howell, M. M.: 2001, Lahar hazards at Agua Volcano, Guatemala, *U.S. Geological Survey Open-File Report* 01-432.

Scott, K. M., Vallance, J. W. and Pringle, P. T.: 1995, Sedimentology, behavior, and hazards of debris flows at Mount Rainier, Washington, *U.S. Geological Survey Professional Paper* 1547.

Scott, K. M., Macíias, J. L., Naranjo, J. A., Rodríguez, S., and McGeehin, J. P.: 2001, Catastrophic debris from landslides in volcanic terrains, mobility hazard assessment, and mitigation strategies, *U.S. Geological Survey Professional Paper* 1630.

Scott, W. E., Hoblitt, R. P., Torres, R. C., Self, S., Martinez, M. L., and Nillos, T. Jr.: 1996, Pyroclastic flows of the June 15, 1991, climactic eruption of Pinatubo volcano. In C. G. Newhall & R. S. Punongbayan (Eds.), *Fire and Mud, Eruption and Lahars of Mt, Pinatubo, Philippines* (pp. 545–570). Seattle, WA: University of Washington Press.

Scott, W. E., Iverson, R. M., Schilling, S. P., and Fisher, B. J.: 1999, Volcano hazards in the Three Sisters Region, Oregon, *U.S. Geological Survey Open-File Report* 99-437.

Shafer, G..: 1976, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, pp. 297.

Sheridan, M. F. and Malin, M. C.: 1983, Application of computer-assisted mapping to volcanic hazard evaluation of surge eruptions: Vulcano, Lipari, and Vesuvius, *J. Volcanol. Geotherm. Res.* **17**, 467–488.

Stevens, N. S., Manville, V., and Heron, D. W.: 2002, The sensitivity of a volcanic flow model to digital elevation model accuracy: Experiments with digitised map contours and

interferometric SAR at Ruapehu and Taranaki volcanoes, New Zealand, *J. Volcanol. Geotherm. Res.* **119**, 89–105.

Umbal, J. V. and Rodolfo, K. S.: 1996, The 1991 lahars of southwestern Pinatubo volcano and evolution of the lahar-dammed Mapanuepe lake. In C. G. Newhall & R. S. Punongbayan (Eds.), *Fire and Mud, Eruption and Lahars of Mt, Pinatubo, Philippines* (pp. 951–970). Seattle, WA: University of Washington Press.

Vallance, J. W.: 2000, Lahars. In H. Sigurdsson, B. Houghton, S. McNutt, H. Rymer & J. Stix (Eds.), *Encyclopedia of Volcanoes* (pp. 601–616). San Diego, CA: Academic Press.

Vallance, J. W., Cunico, M. L., and Schilling, S. P.: 2003, Debris-flow hazards caused by hydrologic events at Mount Rainer, Washington, *U.S. Geological Survey Open-File Report* 03-368.

Vallance, J. W., Schilling, S. P., and Devoli, G.: 2001a, Lahar hazards at Mombacho volcano, Nicaragua, *U.S. Geological Survey Open-File Report* 01-455.

Vallance, J. W., Schilling, S. P., Matías, O., Rose, W.E., and Howell, M. M., and Devoli, G.: 2001b, Volcano hazards at Fuego and Acatenango, Guatemala, *U.S. Geological Survey Open-File Report* 01-431.

Vignaux, M., and Weir, G. J.: 1990, A general model for Mt. Ruapehu lahars, *Bull. Volcanol.* **52**, 381–390.

Walder, J. S., Gardner, C. A., Conrey, R. M., Fisher, B. J., and Schilling, S. P.: 1999, Volcano hazards in the Mount Jefferson region, Oregon, *U.S. Geological Survey Open-File Report* 99-024.

Waitt, R. B., Mastin, G. L., and Begét, J. E.: 1995, Volcanic-hazard zonation for Glacier Peak volcano, Washington, *U.S. Geological Survey Open-File Report* 95-499.

Walley, P.: 1987, Belief function representations of statistical evidence, *Ann. Stat.* **15**, 1439–1465.

Waythomas, C. F. and Miller, T. P.: 1999, Preliminary volcano-hazard assessment for Iliamna volcano, Alaska, *U.S. Geological Survey Open-File Report* 99-373.

Wright, D. F. and Bonham-Carter, G. F.: 1996, VHMS favourability mapping with GIS-based integration models, Chisel Lake–Anderson Lake area. In: G. F. Bonham-Carter, A. G. Galley, and G. E. M. Hall (eds.), EXTECH I: *A Multidisciplinary Approach to Massive Sulphide Research in the Rusty Lake–Snow Lake Greenstone Belts, Manitoba*, Geological Survey of Canada, Bulletin 426, pp. 339–376, 387–401.

Zadeh, L. A.: 1986, A simple view of the Dempster–Shafer theory of evidence and its implication for the rule of combination, *AI Mag.* **7**, 85–90.