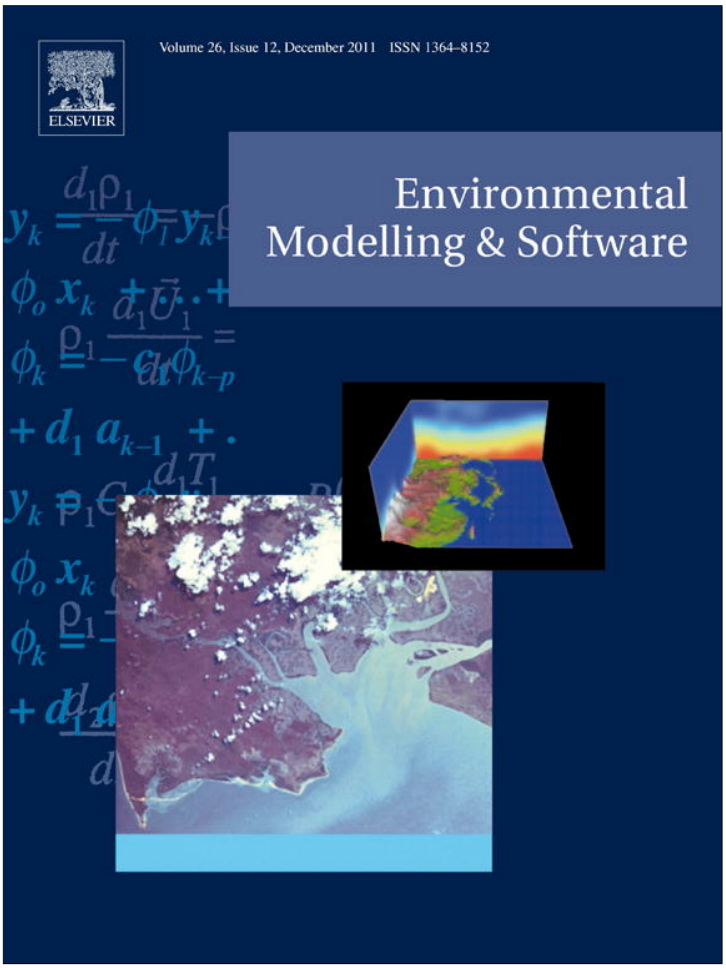
Modeling hydrologic and geomorphic hazards across post–fire landscapes using a self–organizing map approach

Article *in* Environmental Modelling and Software · December 2011

DOI: 10.1016/j.envsoft.2011.07.001

1 author:

Michael James Friedel
GNS Science
101 PUBLICATIONS **365** CITATIONS



Environmental Modelling & Software 26 (2011) 1660–1674

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Modeling hydrologic and geomorphic hazards across post-fire landscapes using a self-organizing map approach

Michael J. Friedel

Crustal Geophysics and Geochemistry Science Center, United States Geological Survey, Denver Federal Center, Box 25046, MS 964, Lakewood, CO 80225, USA

ARTICLE INFO

Article history:
Received 7 August 2010
Received in revised form
11 June 2011
Accepted 3 July 2011
Available online 30 July 2011

Keywords:
Cluster analysis
Cross-validation
Debris flow
Forecast
Flood
Landslide
Runoff
Self-organizing map
Multivariate
Post-fire
Uncertainty

ABSTRACT

Few studies attempt to model the range of possible post-fire hydrologic and geomorphic hazards because of the sparseness of data and the coupled, nonlinear, spatial, and temporal relationships among landscape variables. In this study, a type of unsupervised artificial neural network, called a self-organized map (SOM), is trained using data from 540 burned basins in the western United States. The sparsely populated data set includes variables from independent numerical landscape categories (climate, land surface form, geologic texture, and post-fire condition), independent landscape classes (bedrock geology and state), and dependent initiation processes (runoff, landslide, and runoff and landslide combination) and responses (debris flows, floods, and no events). Pattern analysis of the SOM-based component planes is used to identify and interpret relations among the variables. Application of the Davies–Bouldin criteria following k-means clustering of the SOM neurons identified eight conceptual regional models for focusing future research and empirical model development. A split-sample validation on 60 independent basins (not included in the training) indicates that simultaneous predictions of initiation process and response types are at least 78% accurate. As climate shifts from wet to dry conditions, forecasts across the burned landscape reveal a decreasing trend in the total number of debris flow, flood, and runoff events with considerable variability among individual basins. These findings suggest the SOM may be useful in forecasting real-time post-fire hazards, and long-term post-recovery processes and effects of climate change scenarios.

Published by Elsevier Ltd.

1. Introduction

Peak discharge can increase following a wildland fire because of the commonly exacerbated runoff response. For example, field measurements (Morris and Moses, 1987; DeBano, 2000; Martin and Moody, 2001) and numerical modeling (Beeson et al., 2001; Elliot et al., 2005; Seibert et al., 2010) of unit-area peak discharge reveal increases up to several hundred times over pre-fire rates. Such potentially hazardous peak discharge rates reflect increases in surface runoff that are attributed to reduced rainfall infiltration associated with drying of soil, formation or enhancement of water-repellent (hydrophobic) soils (DeBano, 2000; Robichaud, 2000), decreases in rainfall storage by removal of tree canopy and soil-mantling litter and duff, and increases in source contributing areas (Benavides-Solorio and MacDonald, 2001; Martin and Moody, 2001; Cannon and Gartner, 2005). In addition to changing the local hydrologic response to rainfall, high temperatures associated with burning can cause physical changes in soil that enhance its erodibility (Benavides-Solorio and MacDonald, 2001; Odion and Hanson, 2006.). Over time, the decay of burned plant and tree roots can provide preferential pathways for rainfall infiltration, leading to temporary increases in pore-water pressures (Anderson et al., 2009). Root decay also can reduce the soil cohesion (Uchida et al., 2001), and the combination of increased pore pressures and decreased cohesion can result in landslide failures (Jackson and Roering, 2009).

Rainfall-initiated runoff and landslide failures are the primary processes leading to a post-fire hydrologic and geomorphic response. In steep upper-basin mountain basins, overland and channel flows travel at high velocities (Cannon and Gartner, 2005) resulting in significant erosion (Meyer et al., 2001), sediment transport, and flooding (Gartner et al., 2005; Coe et al., 2008a). Progressive bulking of runoff by sediment eroded from hillslopes and channels can result in flows with broad ranges in sediment concentrations. A debris flow is a spatially continuous rapidly moving mass of water and material composed mainly of coarse debris; typically 20–80% of the particles are greater than 2 mm in diameter (Pierson and Costa, 1987). Hyperconcentrated flows

E-mail address: mfriedel@usgs.gov.

1364-8152/$ – see front matter Published by Elsevier Ltd.
doi:10.1016/j.envsoft.2011.07.001

occupy the boundary between debris and normal stream flows, and they are a mixture of water and sediment with concentrations less than 80% but greater than about 40% by weight (Hutchinson, 1988). In addition to the development of debris flows through progressive sediment bulking, debris flows are observed to mobilize from discrete landslide failures on hillslopes (Meyer and Wells, 1997; Sanchez-Martos et al., 2002), or by a combination of these two processes (Cannon, 2001). Cannon et al. (1998), however, found that considerably more material can be contributed to debris flows from hillslope runoff and channel erosion than from landslide scars.

The observation that a hyperconcentrated flow can develop as floodwater entrains sediment, or conversely, as a debris flow is diluted by water (Wieczorek et al., 1989), underscores the transient and spatial nature of potential post-fire responses to rainfall. Cannon (2001) identified transitory threshold locations within basin channels, where sufficiently eroded material is incorporated (relative to the volume of surface runoff) to generate debris flows that persist down the length of a channel. Understanding the relations between factors governing a transitory threshold is challenging because debris flows are not always generated from all incised channels (Cannon and Gartner, 2005). Likewise, understanding the post-fire related landslide initiation process (Klock and Helvey, 1976; Swanson, 1981; Wondzell and King, 2003) is important but challenging due to a lack of well-controlled data and complex interaction between gradients, pore-water pressures, and physical properties of the near-surface materials (Cannon and Gartner, 2005).

To date, investigations of post-fire hydrologic and geomorphic hazards typically examine relations and perform modeling between a single initiation process or response and limited number of explanatory variables with no consideration given to prediction uncertainty (Cannon et al., 2003a, 2004; Elliot et al., 2005; Gartner et al., 2008). According to Cannon and Gartner (2005), the susceptibility of a burned basin to various initiation processes (landslides, runoff, and landslide and runoff combination) and responses (debris flows, sediment flows, and flooding) is complicated involving interaction among multiple variables from independent landscape categories (climate, land surface form, geologic texture, and post-fire condition), and independent categorical classes (bedrock geology and location). Also, in some settings the burning of a basin may do little to change existing hillslope processes (Larsen et al., 2006). The inability to describe nonlinear and coupled interaction in this multivariate system contributes to the comparatively poor predictability (with respect to nonuniqueness and uncertainty) associated with recent empirical and numerical modeling efforts discussed by Friedel (2010). Consequently, new tools are needed to improve our understanding of dominant post-fire processes, and predict their responses and quantify uncertainty following rainfall on burned basins.

Given the potentially large number of, and complex interaction among, variables in a burned basin, it is necessary to implement advanced multivariate knowledge extraction and prediction tools. Multivariate methods, such as principal component analysis (Christophersen and Hooper, 1992), factor analysis (Suk and Lee, 1999) and hierarchical clustering (Vesanto and Alhoniemi, 2000), are often used to reduce the dimensionality of data sets for system analysis. As traditionally used, however, these methods reduce complexity assuming linear combinations of the data variables. The usefulness of factor analysis is considered dubious because the factors are not directly observable and results cannot be used in other analytical studies. In hierarchical clustering, the most important attributes defining the branches of a clustering tree are not readily recognizable and important patterns can be lost due to its deterministic nature and high-dimensionality of data. In addition, these methods do not work with disparate and sparsely populated data, and they cannot be used to perform estimation or forecasting. One nonlinear alternative for analysis and modeling of multivariate data is artificial neural networks. Artificial neural networks are sometimes preferable over traditional modeling approaches (Hong and Rosen, 2001) because: (1) they can accommodate the nonlinearities of a system; (2) they can accommodate irregular, sparse, and noisy data; (3) they can be quickly and easily updated; and (4) they can interpret disparate information from multiple and mixed types of variables.

Artificial neural networks are a generalized modeling group that includes supervised and unsupervised methods. Supervised artificial neural networks have been used in predicting the rainfall-response on debris flows (Chang and Chao, 2006; Pak et al., 2009), landslides (Pradhan and Lee, 2010), and flooding (Kalteh et al., 2008). The successful application of supervised training methods, however, is dependent on accurately specifying the weights and output layer of the network prior to its deployment. One alternative that requires no a priori knowledge of underlying relations or designation of an output layer is the self-organizing map (SOM) technique (Kohonen, 2001). This vector quantization technique uses an unsupervised and competitive learning algorithm to identify patterns in the data (Kohonen, 2001). Applications of SOM pattern analysis can be found in ecological (Shanmuganathan et al., 2006), geomorphological (Ehsani and Quiel, 2008), hillslope weathering (Iwashita et al., 2011), groundwater (Hong and Rosen, 2001), and surface-water (Lischeid, 2003; Lu et al., 2007) research. In addition to pattern analysis, the vector basis of a SOM provides the means for estimation and prediction (Wang, 2003). SOM estimation applications can be found in chemical process (Rallo et al., 2002) and surface-water hydrology (Lin and Wu, 2007; Kalteh and Berndtsson, 2007; Kalteh and Hjorth, 2009) research. For a comprehensive review of SOM applications in water resources, the reader is referred to Kalteh and Berndtsson (2007) and Maier et al. (2010).

In this study, we explore the usefulness of a SOM analysis to help understand the effects of climate variability on hydrologic and geomorphic hazards across the post-fire landscape in western United States (U.S.). The objectives are to: (1) identify dominant post-fire relations among published multivariate data from 540 burned basins; (2) identify conceptual multivariate post-fire regional models from these data for future empirical model development; (3) quantify SOM bias and uncertainty in simultaneous post-fire predictions of initiation processes (runoff, landslide, and runoff with landslide) and responses (none, flooding, and debris flows); and (4) forecast the simultaneous effects of wet and dry climate scenarios on post-fire initiation processes (runoff, landslide, and runoff with landslide) and responses (none, flooding, and debris flows) across the burned landscape. This study uses a SOM approach for modeling which has not previously been done in wildfire studies. It is also the first to attempt simultaneous predictions of multiple dependent variables across a post-fire landscape. This study relies on a data set compiled by Gartner et al. (2005) from field measurements and observations for burned basins in nine states.

2. Method

2.1. Self-organizing map technique

The SOM technique is a type of unsupervised neural network that learns to project, in a nonlinear manner, from a high-dimensional input layer to a low-dimensional discrete lattice of neurons called the output layer (Kohonen, 2001). The algorithm is iterative and after assigning a prototype (weight) vector to each neuron in the output layer it follows these steps:

- Randomly choose a sample data vector from the training set.
- Compute Euclidean distance between data vector and all prototype vectors.
- Identify best matching unit (BMU) as map unit with prototype vector closest to sample data vector.
- Update prototype vectors using a rule-based learning rate and Gaussian neighborhood function centered on the BMU.
- Move BMU and topological neighbors closer to the input data vector.
- Repeat process until a stable map is achieved.

The learning process occurs in two phases: an ordering phase (when large changes are made to the neurons) and tuning phase (when smaller changes are made to units immediately adjacent to the winning neuron). Ultimately, this training process results in an organized map, where the asymptotic local multivariate density of the weights approach that of the training set (Ritter and Schulten, 1986).

Similarities among patterns are mapped into similar weights of neighboring neurons. This infers that the SOM is a clustering algorithm, where there is one cluster associated with each neuron in the map. The winning neuron for each vector therefore denotes to which cluster a particular input vector belongs, and the connection weights of that vector define the center of that cluster. The equations and issues regarding data gathering, normalization, and training are well documented in Vesanto and Alhoniemi (2000) and Kohonen (2001); therefore, no other details are given here.

2.2. Component planes visualization

The component planes visualization technique (Vesanto, 1999) is used to analyze relations among the post-fire landscape variables. Whereas a component plane can be thought of as a slice of the SOM, it actually represents one set of vector component (variable) values in all map units; that is, each component plane portrays the spread of values for the associated variable across the SOM. In that regard they are similar to histograms, the difference being that the same value can be present in multiple places of the map when it belongs to different clusters.

The visualization of multiple component planes allows identification of correlated variables (Vesanto, 1999). Practically speaking, related variables will appear in component planes at coincident locations as red (high values) or blue (low values) colors. In the case of positive correlation, the color distributions (patterns) in component planes are identical among variables meaning that as one variable increases (or decreases) the others do the same. Conversely, a negative correlation among variables appears with the same pattern but opposite color distribution, meaning that as values in one variable increase those in the other variable decrease.

2.3. Estimation and forecasting

The SOM estimates data values based on distances among the available vectors (Fessant and Midenet, 2002; Junninen et al., 2004; Kalteh and Hjorth, 2009). The traditional estimation process is by replacement (called imputation), where the values are taken directly from the prototype vectors of the BMUs. Often times certain data sets will result in biased predictions (Dickson and Giblin, 2007) requiring a modified scheme based on bootstrapping (Breiman, 1996), ensemble average (Rallo et al., 2002), or nearest neighbor (Malek et al., 2008). This study uses an estimation approach based the following iterative scheme:

- An initial SOM is calculated and a first set of replacement values determined.
- The SOM is recalculated and new replacements for the original missing values obtained.
- Repeat the last step until the topographic error stabilizes.

Because the replacement process for missing values is not simply replacement by a prototype vector of the BMU, it is referred to as estimation. The term forecasting is used for the case of estimating future values.

2.4. Multivariate analysis

The principal component analysis (PCA) and k-means cluster analysis techniques are used to evaluate the post-fire SOM relations. The application of these techniques is possible because the SOM preserves nonlinear relations. In this study, PCA (Christophersen and Hooper, 1992) is applied to the component planes for visual identification of correlated variables and their relative significance to variance in the model. As in the traditional application, the variables appearing toward the ends of loading axes are interpreted as those primary contributors to the system variance.

k-means cluster analysis (Vesanto and Alhoniemi, 2000) is used to identify groupings in the SOM output neurons. This technique is considered better than hierarchical clustering because it does not depend on previously found clusters. The k-means algorithm assumes spherical clusters, and it is sensitive to the initialization process. For that reason, the algorithm is run multiple times for each k with different random initializations. The best partitioning for each number of clusters is selected based on the Euclidian distance criterion, and interesting merges are defined using the Davies–Bouldin index (Davies and Bouldin, 1979).

2.5. Model performance

The SOM algorithm is objective, but there is subjectivity when choosing the set of data variables and size of the map thought to affect prediction quality. Moreover, the data variables are sparsely populated, disparate with varying levels of uncertainty in their measurements and observations. For these reasons, the reliability of the SOM as a predictor is evaluated using two approaches: cross-validation and split-sample validation. The basis of cross-validation (Kohavi, 1995) is a leave-one-out strategy. This requires leaving one data value out of the training set while creating a new SOM to estimate that value based on the remaining data. Because a new SOM is created up to 30 times for each value under scrutiny, it forms the basis for the Monte Carlo framework (Rubinstein and Kroese, 2007) from which residuals are used to evaluate error statistics and model bias. In the split-sample validation approach, the original data set are randomly split into two parts: 90% to train the SOM and 10% to evaluate its independent performance.

2.6. Software

The SOM data mining and analysis is carried out using the SiroSOM (CSIRO Exploration and Mining, 2008) graphical user interface (GUI). This GUI provides an interface between data sets and available functions in the freely available SOM Toolbox (Adaptive Informatics Research Center, 2010). A schematic of the SOM data mining attributes is presented in Fig. 1 (after Vesanto and Alhoniemi, 2000).

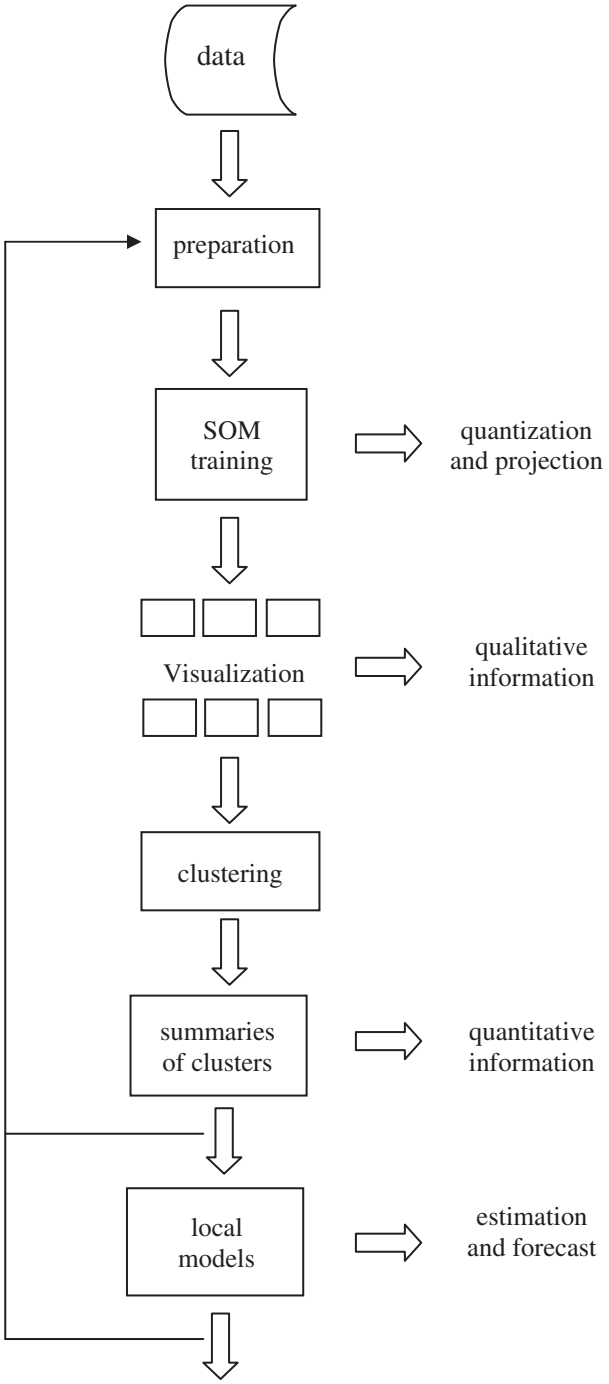


Fig. 1. Schematic of SOM data mining attributes (after Vesanto and Alhoniemi, 2000).

3. Application

3.1. Basin variables

The susceptibility of a landscape to post-fire hydrologic and geomorphic hazards (initiation processes and responses) depends on the interaction of climate with land surface form, geologic texture, and burn severity. The initiation processes, such as runoff, landslides, and landslides plus runoff, occur at locations in a basin; whereas responses, such as debris flows, flooding, no events, occur at a basin outlet. We use measurements and observations, compiled by Gartner et al. (2005), from 600 basins burned by 53 fires in nine

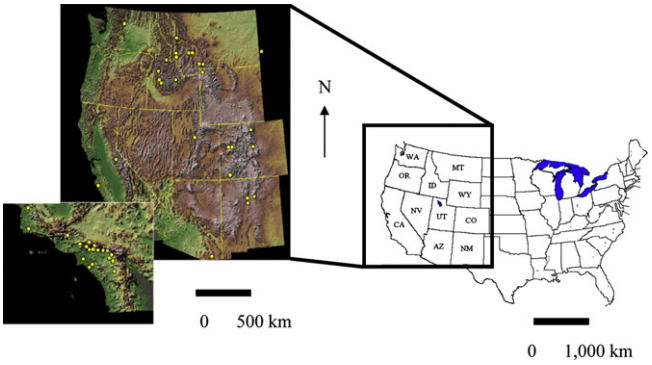


Fig. 2. Map showing locations of burned basin outlets (yellow dots) in the western United States from which data was compiled by Gartner et al. (2005) and used in this study [AZ = Arizona, CA = California, CO = Colorado, ID = Idaho, MT = Montana, NM = New Mexico, UT = Utah, WA = Washington]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Variable notation and data descriptions used in this study.

Type	Category	Acronym	Description
Numerical	Climate	ASI	Average storm intensity in mm per h
		DUR	Rainstorm duration in h
		I10, I15, I30, I60	10-, 15-, 30-, and 60-min per h rainfall intensities
		TS	Total storm rainfall in mm
	Land surface form	AA	Average basin aspect in degrees
		ABG	Average gradient in percent
		BG30, BG50	Basin areas in m^2 with slopes greater or equal to 30 and 50%
		RR	Relief ratio, dimensionless
		RG	Ruggedness, dimensionless
	Geologic texture	AWC	Available water capacity in mm per mm
		CP	Clay content in percentage
		HC	Hydric capacity, dimensionless
		HG	Hydrologic group, dimensionless
		KF	Erodibility factor in tons per unit
		LL	Liquid limit in percent moisture by weight
		OP	Organic matter in percent by weight
		PM	Hydraulic permeability in cm per h
		ST	Soil thickness, m
		SS	Soil slope, dimensionless
	Post-burn condition	ABL, ABM, ABH, ABMH, ABT	Burn severities characterized as area burned low, medium, high, medium plus high, and total in m^2
		PSK, PSO, PMN, PMD	Burned soil diameter skewness, sorting, mean, median statistics
Binary	Response	DF	Debris flow
		FL	Flood (water and hyperconcentrated)
		NR	No response
	Initiation process	LS	Landslide
		RO	Runoff
		ROLS	Runoff plus landslide
		UN	Unknown
	State	AZ	Arizona
		CA	California
		CO	Colorado
		ID	Idaho
		MT	Montana
		NM	New Mexico
		SD	South Dakota
		UT	Utah
		WA	Washington
	Geology	Grn	Igneous–granite
		Met	Metamorphic
		Sed	Sedimentary
		Vol	Igneous–volcanic

western states (Fig. 2). The spatial coverage differs from other landscape frameworks (Omernik, 1987; Wolock et al., 2004), because it is discontinuous and effectively occurs as a random sampling of basin sizes (1 km^2 to 10,000 km^2) burned by wildland fires over the period: 1914–2008. The notation and descriptions of variables used in this study are presented in Table 1.

The climate variables differ from Winter (2001) and Wolock et al. (2004), in which they use the mean annual precipitation (rainstorms plus snowmelt) minus the potential evapotranspiration. In our study basins, the annual evapotranspiration is essentially constant and less than about 20 cm (Hanson, 1991). For that reason, we focus on characteristics of storms that are known to trigger hydrologic and geomorphic hazards in burned basins (Cannon et al., 2003b; Coe et al., 2008b). Such rainfall descriptors include the 10-min (I10), 15-min (I15), 30-min (I30) and 60-min (I60) peak rainfall intensities; average storm rainfall intensity (ASI), storm duration (DUR), and total storm rainfall amount (TS).

The land surface form, geologic texture, and burn severity variables describe characteristics that can enhance or mitigate the effect of climate on post-fire hazards. In this study, the land surface variables include terrain measures, such as average aspect ratio (AA), average basin gradient (ABG), area of slopes greater than 30 (BG30) and 50° (BG50), ruggedness (RG), and relief ratio (RR). A geographic information system is used to extract these measures from a 10-m and 30-m digital elevation model (U.S. Geological Survey, 2009). The geologic texture variables include soil and geologic properties affecting the surface runoff, infiltration and ground-water flow, as well as other soil properties potentially associated with erosion and slope failure. The State Soil Geographic database (U.S. Department of Agriculture, 1994) is used to identify soil characteristics that include available water capacity (AWC), percent clay content (CP), hydric capacity (HC), hydrologic group (HG), erodibility factor (KF), liquid limit (LL), percent organic matter (OP), permeability (PM), slope (SS), and thickness (ST).

The changes in soil properties affecting post-fire hazards are often related to the burn intensity (severity). According to Odion and Hanson (2006), there are three general burn severity classes: low, medium, and high. The low burn severity (ABL) is characterized by an area with partly burned duff and debris, absent to low hydrophobicity, infiltration and erosion potential are not significantly changed, and no sediment delivery. The medium burn severity (ABM) is characterized by an area with consumed duff, moderate hydrophobicity, low soil erosion and delivery. The high burn severity (ABH) is characterized by an area with consumed duff and shrubs (no stumps), moderate to high hydrophobicity; moderate to high soil erosion and sediment delivery. Other burn severity variables that we use are the moderate plus high (ABMH) and total (ABT) area burned (Martin and Moody, 2001) and grain size statistics. The computed statistics of burned soil particle diameters include median (PMD), mean (PMN), sorting (PMS), and skewness (PSK).

In addition to numerical variables, two independent and two dependent categorical variable types are incorporated into the data set. This is possible by converting the categorical variables from present and absent to a binary equivalent using 1 and 0. The independent categorical variables are the uppermost rock types, such as granite (Grn), metamorphic (Met), sedimentary (Sed), and volcanic (Vol); and states, such as Arizona (AZ), California (CA), Colorado (CO), Idaho (ID), Montana (MT), New Mexico (NM), South Dakota (SD), Utah (UT), and Washington (WA). The dependent variables are basin initiation processes, such as runoff (RO), landslides (LS), or combined runoff and landslides (ROLS); and responses, such as debris flows (DF), water and sediment-laden floods (FL), and no events (NR).

The complete data set comprises 22,533 possible samples (present and missing). Of these samples, there are 14,523 data values present reflecting independent variables and another 2436 dependent variables. A comprehensive summary of available and missing data for variables in climate, lands surface form, geologic texture, and post-burn condition categories is summarized in Table 2. The respective variables with fewest and most data values are the hydric capacity (2% present; 98% missing) and total area burned at all severities (97% present; 3% missing).

3.3. SOM preparation and training criteria

Application of the SOM to the post-fire landscape data is done using a fixed number of neurons and topological relations. Specifically, the selected grid shape is a toroid (wraps from top to bottom and side to side) with hexagonal neurons whose size is established based on the ratio between eigenvalues of the training data (Adaptive Informatics Research Center, 2010). Using this approach results in 20 rows (along its length) and 13 columns (along its width) for a total of 260 neurons. Prior to deployment, the training data (except soil grain size statistics) are log transformed and randomly assigned as an initial set of map vectors. One reason for the transformation is to linearize data so that one value does not dominate the Euclidean distance metric used in the mapping process (Vesanto and Alhoniemi, 2000). Training of the map is conducted using default values in a sequential two-step training algorithm: ordering and tuning (CSIRO Exploration and Mining, 2008). The ordering phase has a training length of 20 map units together with an initial and final Gaussian neighborhood radius of 24 map units (one-fourth of the maximum) and 6 map units (one-fourth of initial). The tuning phase has a length of 400 map units (at least four times the first phase) together with an initial and final neighborhood radius of 6 map units and 1 map unit. In both instances, a learning function (inversely proportion to time) is used to ensure all input samples have approximately equal influence on the training result. The respective standard learning rates that we use during the initial and fine-tuning phases are 0.5 and 0.05.

Table 2
Percentage of actual and missing data for each variable in the sparsely populated post-fire data set [*Climate* (blue = low, red = high): ASI = Average storm intensity in mm per h; DUR = rainstorm duration in h; I10, I15, I30, I60 are the 10-, 15-, 30-, and 60-min per h rainfall intensities; TS = total storm rainfall in mm. *Land surface form* (blue = low, red = high): AA = average basin aspect in degrees, ABG = average gradient in percent, BG30 and BG50 = basin areas in m^2 with slopes greater or equal to 30 and 50%, RR = relief ration, and RG = ruggedness. *Geologic texture* (blue = low, red = high): AWC = Available water capacity in mm per mm, CP = clay content in percentage, HC = hydric capacity, HG = hydrologic group, KF = erodibility factor in tons per unit, LL = liquid limit in percent moisture by weight, OP = organic matter in percent by weight, PM = hydraulic permeability in inches per h, ST = soil thickness, and SS = soil slope. *Post-burn condition* (blue = low, red = high): ABL, ABM, ABH, ABMH, ABT are the burn severities characterized as area burned low, medium, high, medium plus high, and total in m^2; PSK, PSO, PMN, PMD = burned soil diameter skewness, sorting, mean, median statistics.].

Category	Status	Variables						
Climate		ASI	DUR	I10	I15	I30	I60	TS
	Present, %	78	78	77	68	75	73	90
	Missing, %	22	22	23	32	25	27	10
Land surface form		AA	ABG	BG30	BG50	RG	RR	
	Present, %	80	83	83	82	82	53	
	Missing, %	20	17	17	18	18	47	
Geologic texture		AWC	CP	HC	HG	KF	LL	
	Present, %	63	62	2	63	58	62	
	Missing, %	37	38	98	37	42	38	
		OP	PM	SS	ST			
	Present, %	63	63	63	63			
	Missing, %	37	37	37	37			
Post-burn condition		ABL	ABM	ABH	ABMN	ABT		
	Present, %	60	63	53	70	97		
	Missing, %	40	37	47	30	3		
		PMD	PMN	PSO	PSK			
	Present	48	48	35	35			
	Missing	52	52	65	65			

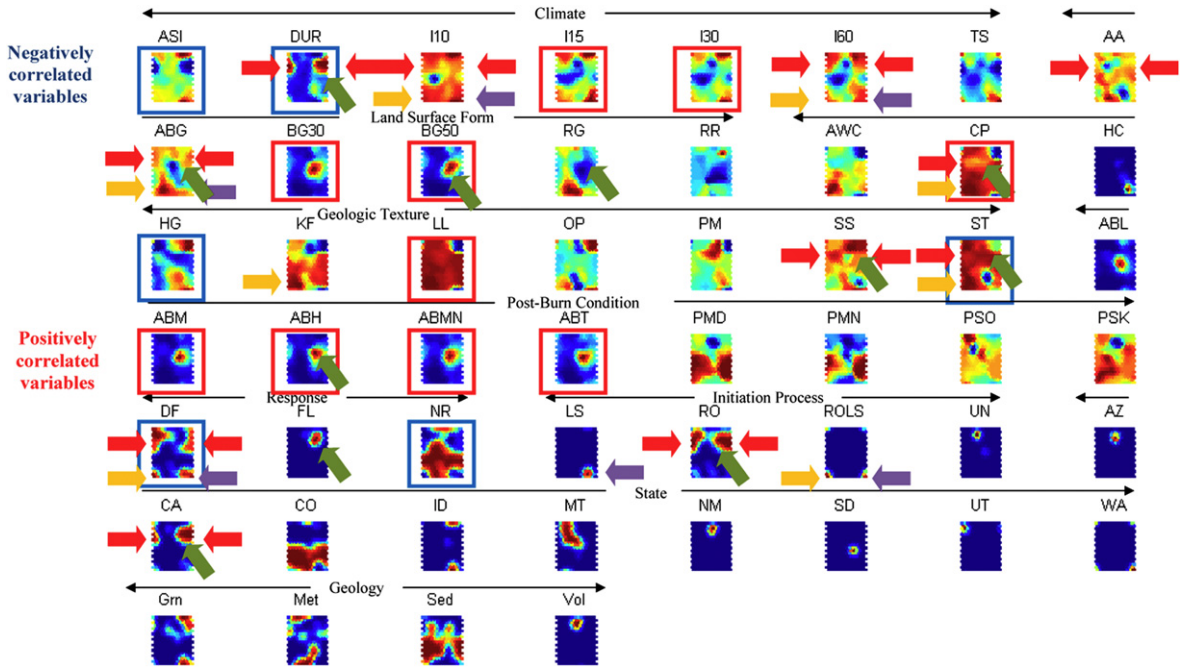


Fig. 3. Component planes plot for the post-fire response data set (blue = low and red = high values). Same pattern and color temperatures reveal positive correlation (for example, BG30 and BG50); and opposite color temperatures a negative correlation (for example, ASI and DUR). In many cases, the patterns among component planes are only partially similar. The reduced correlation is related to spatial heterogeneity. Similarly colored arrows reveal correlation among variables at adjacent locations in different categories [*Climate*: ASI = Average storm intensity in mm per h; DUR = rainstorm duration in h; I10, I15, I30, I60 are the 10-, 15-, 30-, and 60-min per h rainfall intensities; TS = total storm rainfall in mm. *Land surface form*: AA = average basin aspect in degrees, ABG = average gradient in percent, BG30 and BG50 = basin areas in m^2 with slopes greater or equal to 30 and 50%, RR = relief ration, and RG = ruggedness. *Geologic texture*: AWC = Available water capacity in mm per mm, CP = clay content in percentage, HC = hydric capacity, HG = hydrologic group, KF = erodibility factor in tons per unit, LL = liquid limit in percent moisture by weight, OP = organic matter in percent by weight, PM = hydraulic permeability in inches per h, ST = soil thickness, and SS = soil slope. *Post-burn condition*: ABL, ABM, ABH, ABMH, ABT are the burn severities characterized as area burned low, medium, high, medium plus high, and total in m^2; PSK, PSO, PMN, PMD = burned soil diameter skewness, sorting, mean, median statistics. *Response*: DF = debris flow, FL = flood response, NR = no response. *Initiation process*: LS = landslide, RO = runoff, and ROLS = runoff plus landslide, UN = unknown. *States*: AZ = Arizona, CA = California, CO = Colorado, ID = Idaho, MT = Montana, NM = New Mexico, SD = South Dakota, UT = Utah, and WA = Washington. *Geology*: Grn = igneous, Met = metamorphic, Sed = sedimentary rock, and Vol = volcanic.]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Results and discussion

4.1. Relations among variables across the post-fire landscape

4.1.1. Component planes

A plot of component planes characterizing the post-fire landscape of western U.S. is shown in Fig. 3. In this figure, multiple variables exhibit strong correlations among those in the same category. For example, a strong positive correlation (same color pattern; identified by red boxes) exists among climate variables, such as the 15- and 30-min peak rainfall intensities; land surface form variables, such as the area with slopes greater than 30 and 50°; geologic texture variables, such as the percent clay content and liquid limit; and post-fire condition variables, such as the area of burn severity at low, moderate, high, moderate plus high, and total burned. These observations suggest that, at this scale of investigation, variables comprising a group in these categories have the same relative influence (or are being influenced the same) across the post-fire landscape. For that reason, only one of the related variables in each category may be needed to adequately describe the post-fire relations and behavior. In the case of burn severity, the greatest values (red) are concentrated at the same location revealing their association with that same group. However, the distribution of light blue and yellow color differs among the component planes implying that they each have minor relations to variables in other groups. Whereas a reduction in the number of strong and positively correlated variables may save field time and money, understanding the implications with respect to estimation and prediction uncertainty needs to be further evaluated.

According to Rallo et al. (2002), one of the elements necessary for accurate SOM estimation is model diversity; this includes variables with a strong negative correlation (same pattern but opposite colors; identified by blue boxes in Fig. 3). For example, a strong negative correlation exists for climate variables, such as the average storm intensity and rainstorm duration, and geologic texture variables, such as the hydrologic group and soil thickness. A negative correlation also appears among the binary variables in the response, initiation, state, and geology categories. The pattern differs because it is the sum of these variables that characterize a particular category (presence and absence are indicated by red and blue colors, respectively).

Within a particular binary category, the presence of a category (indicated by the spatial extent of red) reflects the extent of its association with variables in other categories and relative number of the associated data. With respect to association, the debris flow variability appears greater than flooding and similar to no response. The debris flow variability also relates to separate groups of initiation processes. For example, the upper left and right corner anomalies (indicated by red arrows) relate to runoff and runoff plus landslides, the lower right anomaly (indicated by a purple arrow) relates to landslides and runoff plus landslides, and the lower left anomaly (indicated by an orange arrow) relates to runoff plus landslides. Similarly, the presence of debris flows occurring in each state and geologic environment appear as anomalies occurring at the same location in the component planes.

Comparing component planes for the post-fire response types reveals that there is a greater variability in debris flow and no response events than for flooding. This suggests that debris flows and no response events occur under a wider range of basin conditions than does flooding. In the case of post-fire initiation processes, the runoff appears to have the greatest variability. By contrast, the conditions under which landslides and runoff plus landslides occur appear to be more restricted. In terms of states, California and Colorado appear to have the greatest variability in post-fire variables, Idaho and Montana a moderate amount; and Arizona, New Mexico, South Dakota, Utah, and Washington the least.

Other weaker but still positive component plane correlations exist among variables in the same categories. These weaker correlations appear as similar patterns but with different intensities. Some examples of these types of climate variables include the 10-, 15-, 30-, 60-min peak rainfall intensities, total storm rainfall, and land surface form variables, such as the average aspect and average basin gradient, and ruggedness and relief ratio. There are two groups of correlated variables in the geologic texture category. In the first group, the variables are available water capacity, percent clay content, and erodibility factor. In the second group, the variables are hydrologic group, percent organic matter, and permeability. Correlated variables in the post-fire condition category include median, mean, sorting, and skewness statistics (of the burned surface soil particle diameters).

The reduced strength in correlation among variables may be attributed, in part, to nonuniqueness. Specifically, in the first geologic texture group, the low component plane values (dark blue color) for percent clay content generally relate to low values of erodibility factor and available water capacity. These relations are not unique because high values (red color) of percent clay content also can result in low susceptibility to rill erosion if runoff entrains sediment and velocities are high, or in the presence of high silt content with predominantly clay sized particles. Also, the available water capacity varies depending on soil properties that affect the retention of water and depth of the root zone, such as organic matter, soil texture, bulk density, and soil structure. In the second geologic texture group, the presence of high percent organic matter typically is associated with low percent clay content and moderate permeability, and therefore increased infiltration capacity leading to decreased runoff potential as indicated by high values in the hydrologic group. Given that model diversity is important for accurate SOM predictions, it is instructive to maintain (or incorporate) correlated variables when their relationship is considered nonunique.

Another potential reason for reduced strength in correlation is attributed to spatial variability. Specifically, the similarity among component plane patterns often times is disrupted at one or more locations by an opposite color. This relation indicates that a variable is associated with other groupings of variables. One example of this heterogeneity is present at the center of the average aspect (orange color) and average basin gradient (blue color) component planes. Collectively, the differences among variables reveal complex nonlinear basin controls on post-fire hydrologic and geomorphic hazards. In the following sections, we explore the component plane relations of climate, susceptibility and initiation process variables to the post-fire responses: debris flows, flooding, and no response.

First, consider the influence of climate and susceptibility variables on the post-fire debris flow response appearing in four groups (upper left, upper right, lower left, lower right) of the component plane (Fig. 3). In evaluating climate variables, the 10- and 60-min peak rainfall intensities appear to dominate in terms of their influence on all four of these groups (indicated by red arrows upper left and right; orange arrow lower left, and purple lower right). By contrast, the high values of storm duration relate only to anomalies in the upper portion of the debris flow component plane (indicated by red arrows). The fact that the storm duration and peak rainfall intensities relate to different parts of the debris flow component plane underscores the influence of differing geographical climate controls.

In contrast to climate, there is considerably more regional variability indicated by component plane differences when comparing the four debris flow anomalies to susceptibility component planes (land surface form, geologic texture, and post-fire condition categories). For example, all of the debris flow groups appear related to the average basin gradient, but this is not true for any of the other land surface forms. Likewise, the geologic texture depicts a preferential association of variables for each group. For example, the upper left debris flow group is related to high values of clay content, soil slope, and soil thickness; intermediate values of erodibility; and low values of available water capacity, hydric capacity, hydrologic group, organic material, and permeability. The upper right debris flow group is related to high values of hydrologic group, organic material, and soil slope; and low values of available water capacity, clay content, hydric capacity, erodibility, permeability. The lower left debris flow group is related to high values of available water capacity, clay content, erodibility, organic matter, and soil thickness; and low values of hydric capacity, hydrologic group, permeability, and soil slope. The lower right debris flow group is related high values of hydric capacity, hydrologic group, organic potential, and soil slope; and low values of available water capacity, erodibility, and permeability.

Some of the positive susceptibility relations to the lower right debris flow anomaly actually account for only half of the debris flow group. This implies that there is at least one other group that is more difficult to discern from the component planes. Likewise, the high values for burn severities appear to be associated only with a portion of the upper right debris flow group, whereas the descriptors for burned soil show marked differences. These four primary debris flow groups appear to be related to differing geologic terrain. Specifically, the geology associated with these debris flow groups are at the upper left: metamorphic and sedimentary; upper right: granite and metamorphic; lower right granite; and lower left sedimentary.

The primary component plane groupings of debris flows also depict relations with various initiation processes. For example, runoff-initiated debris flows are associated with the upper two groups (indicated by red arrows). In contrast to runoff, the other two debris flows groups appear weakly related to landslides (lower right purple arrow) and runoff plus landslides (lower right purple arrow and lower left orange arrow) as indicated by the comparatively small coincident regions of red color (orange and purple arrows). These observations agree with the relative number of actual observations, where debris flows are initiated by 181 runoff events, 23 landslide events, and 17 landslide plus runoff events. This finding supports the notion that most burned basins will initiate debris flows through the runoff process (Iverson, 2000). These debris flow initiation processes can be related back to the basin susceptibility variables previously discussed.

Regional landscape controls on debris flows are evident when inspecting component planes associated with the various western states. For example, the long rainstorm duration and low values of the 10- and 30-min peak rainfall intensities and total storm rainfall appear related to runoff-initiated debris flows in California and Utah; the 10- and 60-min peak rainfall intensities appear related to runoff-initiated debris flows in California, Colorado, and Montana; and low values of rainstorm duration and 10-, 15-, 30-, and 60-min peak rainfall intensities appear related to runoff-initiated debris flows in Arizona and New Mexico basins underlain by volcanic rocks. The initiation of debris flows by landslides or runoff plus

landslides are limited to two states but with different basin controls. For example, in Idaho intermediate to high magnitude storm intensities and high total storm amounts appear to promote landslide-initiated debris flows that are primarily controlled by a land surface form with intermediate values of average basin gradient, and low values of average aspect ratio, area of slopes greater than 30 or 50°, and relief ratio; geologic texture with high values of percent organic matter, permeability, and soil slope and low values of available water capacity, percent clay content, liquid limit, and erodibility factor. In Washington, the 10-min peak rainfall intensity and average storm rainfall intensity appear to promote runoff plus landslide-initiated debris flows controlled primarily by geologic texture with high values of percent clay content, erodibility factor, and soil thickness; and low values of available water capacity and hydrologic group. In both cases, these controls occur in basins underlain by granite.

Second, consider the influence of susceptibility variables on post-fire flood response. Inspection of the component planes indicates a strong relation between rainstorm duration (of the storm intensities, only the 10-min peak storm intensity is related), and runoff in California (see green arrows in Fig. 3). This implies that floods tend to occur in California under conditions of runoff originating from rainstorms that are of long duration, low storm totals, and low peak rainfall intensity. Some direct controls on floods are land surface form variables, such as intermediate values of average basin gradient and low values of ruggedness; geologic texture variables, such as high values of soil slope; and low values of soil thickness, erodibility, and available water capacity. The post-fire condition variables include some with high values for burn severity.

Third, consider the post-fire influence of climate variables resulting in no response. In this scenario, it is important to note that the majority of weight in the no post-fire response component plane appears across the middle portion and minor weight at the upper edge. The middle region is related to low values of rainstorm duration, total storm rainfall, and average storm rainfall intensity and intermediate to high values of the 10-min peak rainfall intensity, whereas the upper edge is related to low to intermediate rainstorm intensities. This suggests that regions subject to these climate conditions are least likely to produce a debris flow or flood response. Other basin controls contributing to this response in parts of Colorado and Montana include land surface form variables with low to moderate values of average basin gradient and ruggedness; geologic texture variables with large values of percent clay content, soil slope, and soil thickness and low to moderate values of erodibility factor and hydrologic group; and post-fire condition variables with high values of all area burn severities and burned surface soil particle diameter statistics.

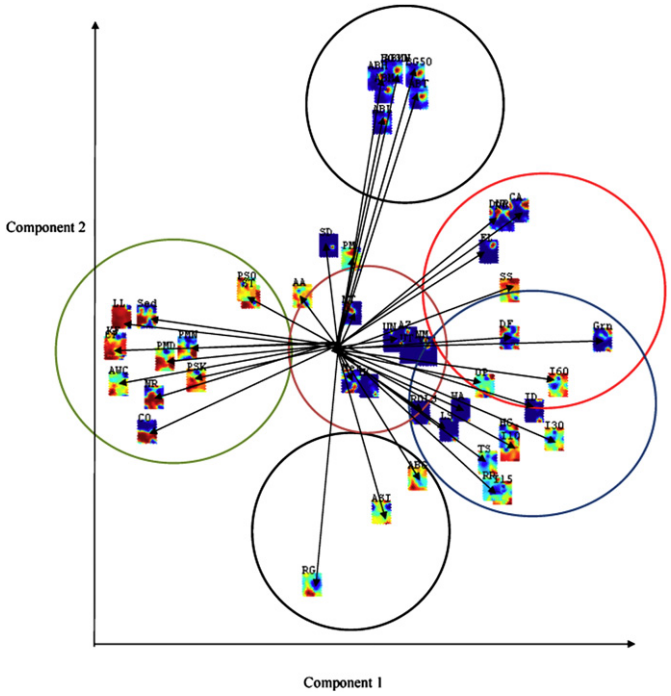


Fig. 4. Principal component analysis loading plot of component planes characterizing post-fire response landscape data set (blue = low and red = high values). Component planes that group together are positively correlated (for example, CO and AWC; or CA and DUR) and those on opposite ends of an axis are negatively correlated (for example, CO and I60; CA and ASI). Those variables at the center contribute less to overall post-fire model variance and therefore are associated with other principal components [*Climate*: ASI = Average storm intensity in mm per h; DUR = rainstorm duration in h; I10, I15, I30, I60 are the 10-, 15-, 30-, and 60-min per h rainfall intensities; TS = total storm rainfall in mm. *Land surface form*: AA = average basin aspect in degrees, ABG = average gradient in percent, BG30 and BG50 = basin areas in m^2 with slopes greater or equal to 30 and 50%, RR = relief ration, and RG = ruggedness. *Geologic texture*: AWC = Available water capacity in mm per mm, CP = clay content in percentage, HC = hydric capacity, HG = hydrologic group, KF = erodibility factor in tons per unit, LL = liquid limit in percent moisture by weight, OP = organic matter in percent by weight, PM = hydraulic permeability in inches per h, ST = soil thickness, and SS = soil slope. *Post-burn condition* (blue = low, red = high): ABL, ABM, ABH, ABMH, ABT are the burn severities characterized as area burned low, medium, high, medium plus high, and total in m^2; PSK, PSO, PMN, PMD = burned soil diameter skewness, sorting, mean, median statistics. *Response*: DF = debris flow, FL = flood response, NR = no response. *Initiation process*: LS = landslide, RO = runoff, and ROLS = runoff plus landslide, UN = unknown. *States*: AZ = Arizona, CA = California, CO = Colorado, ID = Idaho, MT = Montana, NM = New Mexico, SD = South Dakota, UT = Utah, and WA = Washington. *Geology*: Grn = igneous, Met = metamorphic, Sed = sedimentary rock, and Vol = volcanic.]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.1.2. Principal components

Application of principal component analysis (PCA) to the prototype vectors reveals relations among the two distinct axes (Fig. 4). The appearance of climate, land surface form, geologic texture, post-fire condition, geology, initiation process, and response variables at the end of these axes underscores their importance when describing the post-fire landscape in western U.S.. Also, there are certain state landscape groups (closer variables indicate stronger relative relation) that appear along principal component axis one (from left to right); for example, California and Idaho at the right, and Colorado at the left end.

The first group (see red circle in upper quadrant at right-hand side of axis 1) suggests that in California, climate variables characterized by high rainstorm duration and intermediate values of 60-min peak rainfall intensities (and low values of average storm intensity, total storm rainfall, and other peak rainfall intensities) are most likely to initiate runoff-dominated flooding and then runoff-dominated debris flows. Landscape controls on these events include moderate values of average basin gradient and average aspect (and comparatively low values of ruggedness and relief ratio). The common geologic texture associated with both event types is a high value of soil thickness. Other controls on flooding are basins underlain by sedimentary rock with high values of soil slope, clay percentage, and permeability; whereas debris flows in basins underlain by granite have intermediate values of soil slope (low values of available water capacity, erodibility factor, hydric capacity, organic content, and permeability). Post-burn severity includes low values for debris flows and high values for flooding (low soil statistics for both). These findings, while more explanatory, are consistent with statements by Coe et al. (2008a).

The second group (see blue circle in lower quadrant at right-hand side of axis 1) suggests that in Idaho and Washington, climate variables characterizing intermediate to high values of all

peak rainfall intensities (and to a less extent average storm intensity) will initiate landslide- (or runoff plus landslide) initiated debris flows in basins underlain by granite. Primary landscape controls on these events are moderate values of the average basin gradient (low values of ruggedness and relief ratio). The geologic texture controls are associated with high values of hydrologic group and percent organic matter (and to a less extent the soil slope and thickness); but low values of available water capacity, percent clay content, erodibility factor, and liquid limit.

The third group (see green circle in lower quadrant at left-hand side of axis 1) suggests that in Colorado, climate variables characterizing low to high peak rainfall intensities will most likely initiate no response but debris flows also can occur. Landform controls in these basins (underlain predominantly by sedimentary but also some metamorphic rock) include low to high values of average basin gradient, ruggedness and relief ratio. The geologic texture has high values of available water capacity, percent clay content, erodibility factor, and liquid limit. The post-fire condition characterizing these basins includes high values of soil diameter statistics.

Notable relations occurring along the second axis include the climate condition described by the average storm intensity; landscape forms, such as area of slopes greater than 50° and average basin gradient; and post-fire conditions, such as all of the area burn severities (see black circles at either end of axis 2). These findings indicate that their contribution is divided among post-fire hazards occurring in California, Colorado, Idaho and Washington. At the middle of the two principal axes (see black purple circle at the center of axes 1 and 2) are variables associated with other orthogonal axes 3–6 (not physically visible in this projection plane). For example, some variables include the states Arizona, New Mexico, Montana, and Utah; geologic texture, such as hydric capacity; rock types, such as metamorphic and volcanic rock; and initiation process, such as unknown. One interpretation is that those variables comprising combinations of the nonlinear variables along the first six components describe the majority (>99%) of variation in the post-fire landscape system. Because variables associated with other components add no little (or no) information, the overall problem is confined to the subspace defined by these six principal components. Practically speaking, however, the first three principal components at this scale explain about 96% of the post-fire landscape system and those associated with the other axes have a higher degree of uncertainty that is mostly attributed to fewer observations.

4.2. Conceptual multivariate regional-scale models across the post-fire landscape

In this section, we identify conceptual post-fire regional-scale landscape models from the SOM analysis for future research and development of empirical models. Each conceptual model describes the likely importance of variables associated with the post-fire landscape system for that region. The model elements include system boundaries (state and underlying geology), input (climate), parameters (landscape form, geologic texture, post-burn condition), and output (initiation process and response). The statistical grouping of SOM nodes is done by k-means cluster analysis with natural merges identified using the Davies–Bouldin criteria.

Eight classes are identified based on the Davies–Bouldin criteria as being a natural number of distinct combinations for variables reflecting regional post-fire landscapes. A summary of median variable values (missing values are estimated beforehand using the SOM; see Section 2.3) comprising each conceptual model is presented in Table 3. The matrix designators refer to a likelihood group

Table 3
Regional conceptual models and their relation to post-fire hydrologic and geomorphic variables [*Climate* (blue = low, red = high): ASI = Average storm intensity in mm per h; DUR = rainstorm duration in h; I10, I15, I30, I60 are the 10-, 15-, 30-, and 60-min per h rainfall intensities; TS = total storm rainfall in mm. *Land surface form* (blue = low, red = high): AA = average basin aspect in degrees, ABG = average gradient in percent, BG30 and BG50 = basin areas in m^2 with slopes greater or equal to 30 and 50%, RR = relief ration, and RG = ruggedness. *Geologic texture* (blue = low, red = high): AWC = Available water capacity in mm per mm, CP = clay content in percentage, HC = hydric capacity, HG = hydrologic group, KF = erodibility factor in tons per unit, LL = liquid limit in percent moisture by weight, OP = organic matter in percent by weight, PM = hydraulic permeability in inches per h, ST = soil thickness, and SS = soil slope. *Post-burn condition* (blue = low, red = high): ABL, ABM, ABH, ABMH, ABT are the burn severities characterized as area burned low, medium, high, medium plus high, and total in m^2; PSK, PSO, PMN, PMD = burned soil diameter skewness, sorting, mean, median statistics. *Response* (blue = no; red = yes): DF = debris flow, FL = flood response, NR = no response. *Initiation process* (blue = no, red = yes): LS = landslide, RO = runoff, and ROLS = runoff plus landslide, UN = unknown. *States*: AZ = Arizona, CA = California, CO = Colorado, ID = Idaho, MT = Montana, NM = New Mexico, SD = South Dakota, UT = Utah, and WA = Washington. *Geology*: Grn = igneous, Met = metamorphic, Sed = sedimentary rock, and Vol = volcanic.].

Category	Variable	Conceptual post-fire regional models							
		1	2	3	4	5	6	7	8
Climate	ASI	Mod	Low	High	Mod	High	Mod	Mod	High
	DUR	Mod	High	Low	Low	Mod	Mod	Mod	High
	I10	Mod	Mod	High	Mod	High	Mod	Low	High
	I15	Mod	Mod	High	Mod	High	Low	Low	High
	I30	Mod	Mod	High	Low	High	Mod	Low	High
	I60	Mod	Mod	High	Low	High	Mod	Low	High
	TS	Mod	Low	High	Mod	High	Mod	Mod	High
Land surface form	AA	Mod	Mod	Low	Mod	Low	High	Low	Mod
	ABG	Mod	High	Mod	Mod	Mod	Low	Mod	High
	BG30	Low	High	Mod	Low	Mod	High	Mod	Mod
	BG50	Low	High	Mod	Low	Mod	High	Mod	Mod
	RG	Mod	Low	Mod	High	High	Low	Mod	Mod
	RR	High	Low	Mod	Mod	Mod	Low	Mod	Mod
Geologic texture	AWC	Mod	Mod	Low	High	Mod	Mod	Low	Low
	CP	Mod	Low	Low	High	High	Mod	Low	Low
	KF	Mod	Mod	Low	High	Mod	Mod	Low	Low
	OP	Low	Mod	Mod	Mod	Mod	Mod	Low	High
	PM	Mod	High	High	Mod	Low	High	High	Low
	ST	Mod	High	Low	Mod	High	Mod	Mod	Low
	HG	High	Low	High	Mod	Low	Mod	High	High
	SS	Mod	Low	High	Low	Mod	Mod	High	High
	LL	High	Low	Low	High	Mod	Mod	Low	Low
	HC	Mod	High	Low	High	Low	Mod	Low	Low
Post-fire	ABL	Mod	High	High	Low	Low	High	Mod	High
	ABM	Mod	High	High	Low	Low	High	Low	High
	ABH	Low	High	Mod	Low	Mod	High	Mod	Mod
	ABMH	Mod	High	Mod	Low	Mod	High	Mod	Mod
	ABT	Mod	High	Mod	Low	Mod	High	Mod	Mod
	PMD	High	Mod	Low	High	Mod	Mod	Low	Low
	PMN	High	Low	Mod	Mod	Mod	Mod	Low	Mod
	PSO	Mod	Mod	Low	Mod	High	Mod	Mod	Low
	PSK	Mod	Low	Low	High	High	Mod	Low	Low
Geology	Grn	Mod	Mod	High	Low	Low	High	Mod	High
	Met	High	Mod	Low	Mod	High	Mod	Mod	Mod
	Sed	Mod	Low	Low	High	Mod	Mod	Low	Mod
	Vol	Low	Low	High	Low	Low	Low	High	Low
State	AZ	Low	Low	Low	Low	Mod	Low	Mod	Low
	CA	High	High	Mod	Low	Mod	Mod	Mod	High
	CO	High	Low	Low	High	Mod	Mod	Low	Low
	ID	Low	Low	High	Low	Low	Low	Mod	Mod
	MT	Mod	Mod	Mod	Mod	High	High	High	Low
	NM	Low	Low	High	Low	Low	Low	Mod	Low
	SD	Low	Low	Low	Low	Low	High	Low	Low
	UT	Low	Low	Low	Low	Mod	Low	High	Low
	WA	Low	Low	High	Low	Low	High	Low	Mod
Initiation process	LS	Low	Low	High	Low	Low	Mod	High	Low
	RO	Low	High	Mod	Low	High	Mod	Mod	High
	ROLS	Low	Low	Mod	Low	Low	Mod	Low	Mod
Response	DF	Low	High	Mod	Low	High	Mod	Mod	High
	FL	Low	High	Low	Low	Low	Mod	High	Mod
	NO	High	Low	Mod	High	Low	Mod	Mod	Low

in which the median values indicate their relative importance for that model, for example, low likelihood (0–33%), moderate likelihood (34–67%), or high likelihood (68–100%). A summary of these cluster-based conceptual regional landscape models (RLM) is presented next.

RLM-1: This conceptual model includes basins in California and Colorado that are mostly underlain by metamorphic rock. In this region, there is a high likelihood for no post-fire hydrologic or geomorphic response. The independent landscape features that best characterize this region are land surface forms with large relief ratios; and moderate basin sizes, gradients, and ruggedness. The geologic texture has high values of hydrologic group (indicating low runoff potential), liquid limit, and organic content. The climate and post-fire variables have moderate values for all variables.

RLM-2: This conceptual model includes basins in California and Montana that are mostly underlain by metamorphic and granitic rock. These basins have a high likelihood for an equal proportion of post-fire runoff-initiated debris flow and flood events. The independent landscape features that best characterize this region are land surface forms with high values of basin area, average basin gradient, and basin gradients exceeding 30 or 50%, but small values of relief ratio and ruggedness. The geologic texture reflects soil with little to no clay, moderate erodibility, and high hydraulic permeability. Despite the comparatively high hydraulic permeability, the thick soils with moderate slopes and impermeable underlying rocks contribute to high values of hydric capacity and low hydrologic group contributing to the high runoff potential. The climate variables include high values of storm duration and moderate values of recurrent rainfall, whereas post-fire variables are associated with high area values for all burn severities.

RLM-3: This conceptual model includes areas of Idaho, New Mexico, and Washington, and less likely areas of California and Montana. These basins are typically underlain by granitic or volcanic rock, and have a moderate likelihood for debris flows or no response. When debris flows occur, they are most likely to be initiated by runoff and less likely by the runoff–landslide combination. The independent landscape features that best characterize this region are land surface forms with moderate values of basin area, basin gradient, and basin gradients exceeding 30 or 50%, relief ratio, and ruggedness. The geologic texture reflects soil with high values of hydraulic permeability and soil slopes. This combination together with low values of clay content, erodibility, water capacity, and soil thickness are believed to contribute to the low runoff potential. This is the only landscape region in which all climate variables reflect high values affecting the region. In terms of post-fire variables, basins reflect predominantly areas of low and moderate burn severities but also some basins of high burn severities.

RLM-4: This conceptual model likely includes Colorado basins and some basins from Montana that are underlain by sedimentary and metamorphic rocks. In these areas, basins will mostly likely experience post-fire runoff without initiating any hydrologic or geomorphic response. The independent landscape features that best characterize this region are land surface forms with small basins with moderate gradients, and relief ratio that are highly ruggedized. The geologic texture reflects soil with high values of clay content, water capacity, soil thicknesses; moderate hydraulic permeability values; and low soil slopes that collectively contribute to high values of hydric capacity and erodibility. This region has climate variables that are moderate in value and low area values of post-fire burn severity.

RLM-5: This conceptual model includes mostly Montana basins but also some basins from Arizona, California, Colorado, and Utah. These basins typically are underlain by metamorphic and sometimes sedimentary rocks that have a high likelihood for post-fire runoff-initiated debris flow events. Independent landscape features that best characterize this region are land surface forms similar to RLM-4 but also include moderate values of basin gradients exceeding 30 and 50%. This region has climate variables that are high in value and post-fire moderate likelihood for high and moderate plus high burn areas.

RLM-6: This conceptual model includes basins mostly from Montana, South Dakota, and Washington, but also some basins from California and Colorado. These basins are typically underlain by granitic and sometimes metamorphic or sedimentary rock that have a moderate likelihood to experience post-fire debris flow, flood, or no events with a moderate likelihood to be initiated by landslides, runoff, or runoff and landslides, or an unknown mechanism. The independent landscape features that best characterize this region are land surface forms with large basins and gradients, and low relief ratio and ruggedness. The geologic texture reflects soils that are moderate in value accept for the high hydraulic permeability. The climate variables are moderate in value, whereas post-fire variables have high values for all burn severities (similar to RLM-2).

RLM-7: This conceptual model includes mostly basins from Montana and Utah but also some from Arizona, California, Idaho, and New Mexico that are typically underlain by granitic with metamorphic, volcanic with sedimentary, volcanic, and sometimes granitic or metamorphic rocks. These basins that have a high likelihood to experience post-fire flood events that are mostly landslide-initiated but may also be initiated by runoff. The independent landscape features that best characterize this region are land surface forms that are similar to RLM-3. The differences primarily are associated with the climate variables that include total storm, average storm intensity, and duration. In terms of post-fire variables, this region includes basins that reflect an equal likelihood for moderate values that span all severities.

RLM-8: This conceptual model most likely includes basins of California and less likely Idaho and Washington that are underlain by granitic with metamorphic or sedimentary rocks. These basins have a high likelihood for post-fire debris flow and moderate likelihood for flood events. The initiation mechanisms are most likely runoff but also can be a runoff and landslide combination. The independent landscape features that best characterize this region are land surface forms with large areas and average gradients, moderate values of basins with gradients exceeding 30 or 50%, aspect, relief ratio, and ruggedness. The geologic texture reflects soil with high values of soil slope, hydrologic group, and organic content. All of the other variables have low values. This landscape region has climate variables that reflect high values. In terms of post-fire variables, basins are predominantly areas of low and moderate burn severities but also basins of high burn severities.

4.3. Bias and uncertainty in simultaneous predictions across the post-fire landscape

4.3.1. Cross-validation

The number of Monte Carlo trials used in formulating the cross-validation statistics is restricted to 30 because each took about 6 h. Despite these limitations, this process indicates that global mean absolute errors for response and initiation process events are on the order of 10^{-1}. This result signifies that SOM predictions are unbiased when collectively considering all of the initiation processes (landslide, runoff, and runoff and landslide combination) or responses (debris flow, flooding, and no events). In addition to the global model bias, individual response biases also are evaluated.

Variability and uncertainty in the SOM precludes the prediction of pure binary (zeros and ones) relations. Rather the model computes simultaneous prediction values for each initiation process and response that ranges between 0 and 1 (Fig. 5). For an unbiased model, the actual number of events having prediction values equal to 0.5 is the same as the 50th percentile (median) value. The maximum difference between known and median number of debris flow predictions is about 5% indicating a reasonably unbiased prediction for this response. The respective actual and median number of recorded debris flow events is 216 and 210, but depending on the cut-off prediction response the number of predicted events ranges from about 30 to 530. The number of events associated with values that are less than or greater than 0.5 reveals false negatives (under predicting) and false positives (over predicting). In contrast to debris flow events, the respective difference between known and median predicted flood events is about 21% indicating a tendency for more biased model predictions.

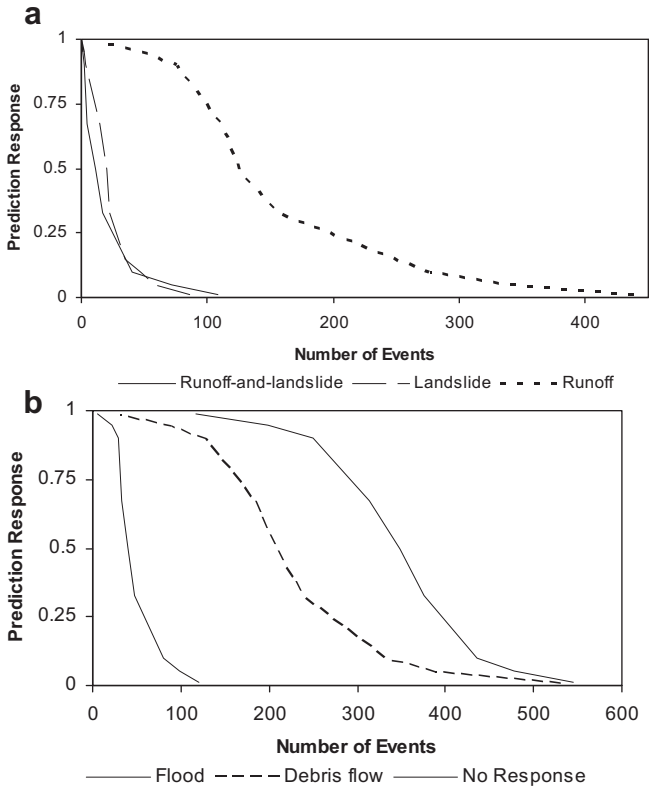


Fig. 5. Predicted response when using the post-fire SOM for western U.S.. (a) Initiation processes (runoff, landslide, and landslide and runoff); (b) erosional response (debris flow, flood, no response) and initiation process. The actual number of flood, debris flow, and no response events were 21, 48, and 346 respectively and are similar to the median values.

For this case, the respective actual and median number of recorded flood events is 46 and 38, but the probable range in predicted number is from 4 to 121 events. The least biased response occurs when predicting no response events. This finding is anticipated given that it has the largest number of measurement data for training. For this case, the maximum difference between known and median number of no response event predictions is less than 1% indicating unbiased predictions. The actual number of recorded no response events is 346, but the probable range in predicted number is from 117 to about 545 events.

In addition to response predictions, bias also is evaluated with respect to the three initiation process events: runoff, landslide, and runoff plus landslide combination. The least model bias is for landslides with a difference between actual and median number of about 21%. For this case, the actual number of recorded events is 23 as compared to the median value of 19. The probable range in predicted number is from 0 to 86 events. Model bias for the runoff process predictions is similar to landslides with a difference between actual and median number of about 23%. For this case, the actual number of recorded events is 154 as compared to the median value of 125. The probable range in predicted number is from 13 to 439 events. The runoff and landslide process combination has the greatest prediction bias with a difference between actual and median number of about 55%. For this case, the actual number of recorded events is 17 as compared to the median value of 11. The probable range in predicted number is from 0 to 109 events. Based on both the response and initiation process results, a prediction response of 0.5 is considered reasonable to use as the cut-off value in the split-sample validation and forecasting.

4.3.2. Split-sample validation

Ten percent (60 out of 600) of the basins and their data are randomly selected and withheld during training of the post-fire SOM. In this way, an evaluation of the SOM model's predictive ability is extended to basins with known initiation processes and responses but that is outside the training set. Evaluation of the results are based on a binary assessment where prediction values above 0.5 indicate the likelihood for an event to occur and those below not for an event to occur. A summary of the post-fire model prediction performance results at outlets of the 60 basins in western U.S. is presented in Table 4.

The results for geomorphic and hydrologic hazards are variable, but they all exhibit accuracy of 78% or better. For example, the debris flow has the lowest number of correct event predictions (78%). Predicting runoff and no response events has corresponding accuracy levels of 83% and 87% correct, whereas the flooding,

Table 4
Post-fire artificial neural network (ANN) model performance results. Prediction results are for geomorphic and hydrologic responses and initiation processes observed at 60 basins in the southwestern United States that were not used in model development.

Summary	Model performance					
	Response			Initiation process		
	Debris flow	Flood	None	Landslide	Runoff	Landslide and runoff
Number of basins	60	60	60	60	60	60
Number events observed	28	4	28	2	26	0
Number events predicted	20	5	34	2	18	0
Number of non-events observed	32	56	32	58	34	60
Number of non-events predicted	40	55	26	58	42	60
Number of false positives	3	3	7	0	2	0
Number of false negatives	10	2	1	0	8	0
Correct, %	78	92	87	100	83	100

Table 5
The forecast effects of climate change on post-fire initiation processes and associated responses were evaluated using the ANN model. The respective influence of El Nino (Wet condition) and La Nina (Dry condition) were characterized by a factor of two increase (100%) and decrease in the associated climate factors [Event: 1 = yes, blank = none; *Response*: DF = debris flow, FL = flood response, NR = no response. *Initiation process*: LS = landslide, RO = runoff, and ROLS = runoff plus landslide, UN = unknown. *States*: AZ = Arizona, CA = California, CO = Colorado, ID = Idaho, MT = Montana, NM = New Mexico, SD = South Dakota, UT = Utah, and WA = Washington.].

Basin	State	Forecast											
		Response						Initiation process					
		Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry
		DF	DF	FL	FL	NR	NR	LS	LS	RO	RO	ROLS	ROLS
1	MT	1	1							1			
2	CA	1	1							1	1		
3	MT	1					1						
4	CA			1			1			1			
5	CO					1	1						
6	CA	1					1			1			
7	MT	1					1			1			
8	MT					1	1						
9	CO					1	1						
10	CA			1			1			1			
11	CO					1	1						
12	CO	1					1						
13	UT	1	1							1	1		
14	CA			1			1			1			
15	CO					1	1						
16	MT		1			1					1		
17	CA	1					1			1			
18	CO					1	1						
19	CA	1	1							1	1		
20	ID	1	1					1	1				
21	CO					1	1						
22	NM		1			1					1		
23	MT		1			1					1		
24	CO					1	1						
25	CA		1	1						1	1		
26	MT	1					1			1			
27	CA	1					1						
28	CO					1	1						
29	CO					1	1						
30	CO		1			1					1		
31	CO					1	1						
32	NM					1	1						
33	MT					1	1						
34	CO					1	1						
35	ID					1	1						
36	MT	1	1							1	1		
37	ID					1	1						
38	MT	1					1			1			
39	CO					1	1						
40	MT	1					1			1			
41	CA	1								1	1		
42	CO					1	1						
43	CO					1	1						
44	CO					1	1						
45	CO		1			1					1		
46	CA	1					1			1			
47	NM	1			1					1	1		
48	MT					1	1						
49	CO		1			1					1		
50	CA	1	1							1	1		
51	CO					1	1						
52	CA	1					1			1			
53	CA	1					1			1			
54	CA	1					1						
55	CA	1	1						1	1			
56	MT	1					1			1			
57	CO					1	1						
58	MT	1					1			1			
59	ID					1	1						
60	ID	1			1			1			1		
Number of events		26	15	4	2	30	41	2	2	24	14	0	0

landslide, and runoff plus landslide are predicted with a 92%, 100%, and 100% accuracy respectively. With the exception of landslide and runoff plus landslide, the predicted events result in both false positives and false negatives. Both debris flows and runoff are associated with the greatest number of false negatives (10 and 8); that is, they tend to under report the number of likely events following a storm. The fact that most of these false negatives are associated with a no response results in a higher level of false positives for no response; that is, they tend to overreport the number of no response events. These results are particularly encouraging given that these random basin records are sparsely populated (see Table 2).

4.4. Forecasting effects of climate variability on responses across the post-fire landscape

In this section, we use the SOM to model climate variability effects on annual post-fire responses across burned basins in the western U.S.. Some metrics of success are the ability to perform simultaneous forecasts for multiple dependent variables and that forecasts for each basin set of initiation processes (and responses) sum to one (see Section 4.3). Achieving these metrics is important because traditional modeling approaches predict only one dependent variable, and they require the use of special constraints in their development. In the post-fire SOM for western U.S., there are multiple dependent variables and no physical equations to constrain the forecasts (only relations among variables described by the underlying multivariate probability density).

The effects of climate variability are forecast across the 60 independent basins (not used in the training process; see Section 4.3) reflecting likely conditions imposed by El Niño and La Niña episodes. These episodes represent end members of a climate continuum based on conditions observed following 100 years of sea surface temperature monitoring in the tropical Pacific. In general, a La Niña episode brings dry (drought) conditions to the south and wet conditions to the north, whereas an El Niño episode brings dry conditions to the north and wet conditions to the south. In performing forecasts, we increase the climate variables by a factor of two replicating wet conditions and then decrease variables by a factor of one-half replicating dry conditions. Other variables associated with land surface form, geologic texture, and post-fire condition categories are held constant. The separate forecast results for wet and dry climate scenarios are summarized in Table 5.

Forecasts among the 60 basins reveal that it is possible to perform simultaneous and sensible predictions of multiple binary dependent variables. The fact that individual basin forecasts of initiation process (and response) variables each sum to 1 suggests that relations in the underlying multivariate density provide a conditional constraint at this scale of study. Further evidence supporting this assertion is the decreasing landscape trend in the number of debris flows, flooding, and runoff events under dry conditions. For example, the respective number of total debris flows, flooding, and runoff events decreases from 26 to 15, from 4 to 2, and from 24 to 14 respectively.

In contrast to the general trend, there is considerable variability with significant differences among the individual post-fire basin forecasts under wet and dry conditions. For example, twenty-three basins in Colorado (5, 9, 11, 15, 18, 21, 24, 28, 29, 31, 34, 39, 42, 43, 44, 51, 57), Idaho (35, 37, 59), NM (32), and Montana (8, 33, 48) yield no response under the wet or dry conditions. Fifteen basins in California (6, 17, 27, 46, 52, 53, 54), Colorado (12), and Montana (3, 7, 26, 38, 40, 56, 58) result in runoff-initiated debris flows under wet conditions but have no response under dry conditions. Nine basins in California (2, 19, 50, 55), Idaho (20), Montana (1, 16, 36), and Utah (13) result in debris flows under wet and dry conditions but their initiation processes differ. Specifically, the debris flow initiation process in one California (55) basin is runoff under wet conditions and a landslide under dry conditions; and in Idaho (2) the debris flows occurring under wet and dry conditions are initiated by landslides. Five basins in Colorado (30, 45, 49), New Mexico (22), and Montana (23) result in runoff-initiated debris flows under wet conditions but have no response under dry conditions. Three basins in California (4, 10, 14) result in runoff-initiated flooding under wet conditions but have no response under dry conditions. Another basin in California (25) has runoff-initiated flooding under wet conditions that switches to a runoff-initiated debris flows under drier conditions. In other California (41) and New Mexico (47) basins, runoff-initiated debris flows occur under wet conditions but then change to runoff without flooding and runoff-initiated flooding. In one Idaho basin (60), wet conditions result in a landslide-initiated debris flow, but under dry conditions runoff-initiated flooding.

As in other traditional modeling studies, assessing the performance of a forecast will require a post-audit analysis. The purpose, however, is not to focus on forecast accuracy; rather it is to establish a method for modeling local real-time and future regional trends. The ability to forecast a sequential set of model scenarios can be profitable for planning and management of basin resources, particularly in view of the increasing number of wildfires (Running, 2006), long-term basin recovery process (Stickney and Campbell, 2000), and future climate change projections (Westerling and Swetnam, 2003) for the western U.S..

5. Conclusions

In this study, we found that a type of unsupervised artificial neural network, called the SOM, can be developed for modeling at the western U.S. scale using published data from 540 burned basins in nine states. Some of the noteworthy conclusions are below:

(1) It is possible to identify relations among variables in post-fire landscape system elements (climate, susceptibility and initiation processes, and responses) at the western U.S. scale. Many variables in certain categories are highly correlated indicating an unnecessary redundancy in describing system contributions. The general lack of direct relations between climate and response variables is attributed to the intervening nonlinear and coupled basin controls represented by susceptibility and initiation process variables. These susceptibility variables serve as a basin transfer function that modifies the effect of climate on the hydrologic and geomorphic responses.
(2) Regional landscape controls can be identified using component planes and multivariate statistics. The application of principal component analysis to component planes is useful for identifying state groupings of variables for California, Colorado, Idaho, and Montana. k-means clustering of SOM neurons together with the Davies–Bouldin criteria is valuable for identifying realistic conceptual post-fire regional models. We identified eight conceptual multistate (regional) post-fire landscape models (representing combinations of variables characterizing climate, land surface forms, geologic texture, post-fire conditions, underlying geology, states, initiation processes, and responses) that can be useful for focusing field research, identification of risk areas following fires, and future empirical model development.
(3) The SOM performance can be evaluated using the new stochastic cross- and traditional split-sample validation approaches. The stochastic cross-validation approach reveals that simultaneous SOM-based predictions of post-fire hydrologic and geomorphic hazards are globally unbiased with the

occurrence of both false positives and false negatives. The degree of prediction bias for individual initiation process or response events is dependent on the relative number of observations. Therefore, the addition of observations to the database can provide improvements to the SOM model. Because up to 20% of predictions may be misclassified emphasizing the importance to extend analysis to basins with known responses and initiation processes outside the training set.

(4) The SOM is able to forecast the simultaneous effects of changing climate on multiple post-fire hydrologic and geomorphic responses. The fact that 60 independent basin forecasts of initiation process and response variables sum to one suggests that relations in the underlying multivariate density provide a reasonable constraint at this scale of study. Further support for this assertion is the post-fire landscape trend in which these basins indicate a decrease in the number of debris flows, flooding, and runoff events under dry conditions. Whereas the decreasing trend in the forecasted number of events is intuitively anticipated under a drier climate scenario, its ability to simultaneously forecast the individual basin effects is considered more important from a hazards perspective. Forecasting the annual climate effects can be profitable for planning and management of basin resources, particularly because of the long-term basin recovery process and future climate change projections. Likewise, implementation of a SOM-based approach affords the possibility for real-time forecasting of regional post-fire hazards.

Acknowledgments

Financial support for this research was provided by the Assessment of Wildfire-Related Hazards on Human and Ecological Communities: A Demonstration Project in the Front Range of Colorado, U.S. Geological Survey, Denver Colorado.

References

Adaptive Informatics Research Center, 2010. SOM Toolbox. Helsinki University of Technology, Laboratory of Computer and Information Science, Adaptive Informatics Research Center. http://www.cis.hut.fi/projects/somtoolbox/.

Anderson, A.E., Weller, M., Alila, Y., Hudson, R.O., 2009. Dye staining and excavation of lateral preferential flow network. Hydrology and Earth System Sciences 13, 935–944.

Beeson, P.C., Martens, S., Breshears, D.D., 2001. Simulating overland flow following wildfire; mapping vulnerability to landscape disturbance. Hydrological Processes 15, 2917–2930.

Benavides-Solorio, J., MacDonald, L.H., 2001. Post-fire runoff and erosion from simulated rainfall on small plots, Colorado Front Range. Hydrological Processes 15, 2931–2952.

Breiman, L., 1996. Bagging predictors. Machine Learning 24, 123–140.

Cannon, S.H., Gartner, J.E., 2005. Wildfire related debris flow from a hazards perspective (Chapter 15). In: Jacob, M., Hungr, O. (Eds.), Debris-flow Hazards and Related Phenomena. Springer-Praxis Books in Geophysical Sciences, pp. 321–344.

Cannon, S.H., Powers, P.S., Savage, W.Z., 1998. Fire-related hyperconcentrated and debris flows on Storm King Mountain, Glenwood Springs, Colorado, USA. Environmental Geology 35 (2–3), 210–218.

Cannon, S.H., Gartner, J.E., Rupert, M.G., Michael, J.A., 2003a. Emergency Assessment of Debris-Flow Hazards from Basins Burned by the Piru, Simi, and Verdale Fires of 2003, Southern California U.S. Geological Survey Open-File Report 03–481.

Cannon, S.H., Gartner, J.E., Holland-Sears, A., Thurston, B.M., Gleason, J.A., 2003b. Debris flow response of basins burned by the 2002 Coal Seam and Missionary Ridge fires, Colorado. In: Boyer, D.D., Santi, P.M., Rogers, W.P. (Eds.), Engineering Geology in Colorado-Contributions, Trends, and Case Histories. Association of Engineering Geologists Special Publication, vol. 14, p. 31. Colorado Geological Survey Special Publication 55.

Cannon, S.H., Gartner, J.E., Rupert, M.G., Michael, J.A., 2004. Emergency Assessment of Debris-flow Hazards from Basins Burned by the Cedar and Paradise Fires of 2003. U.S. Geological Survey, Southern California. Open-File Report 2004–1011.

Cannon, S.H., 2001. Debris-flow generation from recently burned watersheds. Environmental and Engineering Geoscience 7, 321–341.

Chang, T.-C., Chao, R.-J., 2006. Application of back-propagation networks in debris flow prediction. Engineering Geology 85 (3–4), 270–280.

Christophersen, N., Hooper, R.P., 1992. Multivariate analysis of stream water chemical data the use of principal components analysis of the end-member mixing problem. Water Resources Research 28, 99–107.

Coe, J.A., Cannon, S.H., Santi, P.M. (Eds.), 2008a. Introduction to the Special Issue on Debris Flows Initiated by Runoff, Erosion, and Sediment Entrainment in Western North America. Geomorphology, vol. 96, pp. 247–249.

Coe, J.A., Kinner, D.A., Godt, J.W., 2008b. Initiation conditions for debris flows generated by runoff at Chalk Cliffs, central Colorado. Geomorphology 96, 270–297.

CSIRO Exploration & Mining, 2008. SIROSOM Software. http://www.csiro.au/science/SOM-data-analysis.html.

Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-1, 224–227.

DeBano, L.F., 2000. The role of fire and soil heating on water repellency in wildland environments: a review. Journal of Hydrology 231–232, 195–206.

Dickson, B.L., Giblin, A.M., 2007. An evaluation of methods for imputation of missing trace element data in groundwaters. Geochemistry: Exploration, Environment, and Analysis 7 (2), 173–178.

Ehsani, A.H., Quiel, F., 2008. Geomorphometric feature analysis using morphometric parameterization and artificial neural networks. Geomorphology 99, 1–12.

Elliot, J., Smith, M.E., Friedel, M.J., Litke, D., 2005. Post-fire Hydrologic Hazards Study for the 2002 Hayman, Coal Seam, and Missionary Ridge Wildfires, Colorado U.S. Geological Survey, Science Investigations Report-5300, 125.

Fessant, F., Midenet, S., 2002. Self-organizing map for data imputation and correction in surveys. Neural Computing and Applications 10, 300–310.

Friedel, M.J., 2010. Post-fire debris flow prediction using a two-step hybrid approach. In: 3rd U.S. Geological Survey Modeling Conference, Denver Colorado, June, 2010.

Gartner, J.E., Cannon, S.H., Bigio, E.R., Davis, N.K., Parrett, C., Pierce, K.L., Rupert, M.G., Thurston, B.L., Trebish, M.J., Garcia, S.P., Rea, A.H., 2005. Compilation of Data Relating to the Erosive Response of 606 Recently Burned Basins in the Western U.S. U.S. Geological Survey Open-File Report 2005–1218.

Gartner, J.E., Cannon, S.H., Sanit, P.M., Dewolfe, V.G., 2008. Empirical models to predict the volumes of debris flows generated by recently burned basins in the western U.S. Geomorphology 96, 339–354.

Hanson, R.L., 1991. Evapotranspiration and droughts. In: Paulson, R.W., Chase, E.B., Roberts, R.S., Moody, D.W. (Eds.), Compilers, National Water Summary 1988-89–Hydrologic Events and Floods and Droughts. U.S. Geological Survey Water-supply Paper, vol. 2375, pp. 99–104.

Hong, Y.S., Rosen, M.R., 2001. Intelligent characterisation and diagnosis of the groundwater quality in an urban fractured-rock aquifer using an artificial neural network. Urban Water 3 (3), 193–204.

Hutchinson, J.N., 1988. General report: morphological and geotechnical parameters of landslides in relation to geology and hydrogeology. In: Bonnard, C. (Ed.), Proceedings of the Fifth International Symposium on Landslides, vol. 1. A.A. Balkema, Rottedam, Netherlands, pp. 3–35.

Iverson, R.M., 2000. Landslide triggering by rain infiltration. Water Resources Research 36 (7), 1897–1910.

Iwashita, F., Friedel, M.J., de Souza, C.R., Filho, Fraser, S.J., 2011. Hillslope chemical weathering across Paraná, Brazil: a data mining-GIS hybrid approach. Geomorphology 132 (3–4), 167–175.

Jackson, M., Roering, J.J., 2009. Post-fire geomorphic response in steep, forested landscapes: Oregon Coast Range, USA. Quaternary Science Reviews 28 (11–12), 1131–1146.

Junninen, H., Niska, H., Tuppurainen, K., Ruuskanen, J., Kolehmainen, M., 2004. Methods for imputation for missing values in air quality data sets. Atmospheric Environment 38, 2895–2907.

Kalteh, A.M., Berndtsson, R., 2007. Interpolating monthly precipitation by self-organizing map (SOM) and multilayer perceptron (MLP). Hydrology Science Journal 52 (2), 305–317.

Kalteh, A.M., Hjorth, P., 2009. Imputation of missing values in precipitation-runoff process database. Hydrology Research 40 (4), 420–432.

Kalteh, A.M., Hjorth, P., Berndtsson, R., 2008. Review of the self-organizing map (SOM) approach in water resources: analysis, modeling and application. Environmental Modelling & Software 23, 835–845.

Klock, G.O., Helvey, J.D., 1976. Debris flows following wildfire in North Central Washington. In: Proceedings of the 3rd Federal Inter-agency Sedimentation Conference, March 22–25, Denver, Colorado. Water Resources Council, Denver, CO, pp. 91–98.

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence 2 (12), 1137–1143.

Kohonen, T., 2001. Self-organizing Maps, third ed. Springer, Berlin.

Larsen, I.J., Pederson, J.L., Schmidt, J.C., 2006. Geologic versus wildfire controls on hillslope processes and debris flow initiation in the Green River canyons of Dinosaur National Monument. Geomorphology 81, 114–127.

Lin, G.-F., Wu, M.-C., 2007. A SOM-based approach to estimating design hyetographs of unguaged sites. Journal of Hydrology 339, 216–226.

Lischeid, G., 2003. Taming awfully large data set s using self-organizing maps for analyzing spatial and temporal trends of water quality data. Geophysical Research Abstracts 5, 01879.

Lu, G., Chiu, L.S., Wong, D.W., 2007. Vulnerability assessment of rainfall-induced debris flows in Taiwan. Natural Hazards 43, 223–244.

Maier, H.R., Jain, A., Dandy, G.C., Sudheer, K.P., 2010. Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions. Environmental Modelling & Software 25 (8), 891–909.
Malek, M.A., Harun, S., Shamsuddin, S.M., Mohamad, I., 2008. Imputation of time series data via Kohonen self organizing maps in the presence of missing data. Engineering and Technology 41, 501–506.
Martin, D.A., Moody, J.A., 2001. Post-fire, rainfall intensity-peak discharge relations for three mountainous watersheds in the western USA. Hydrological Processes 15, 2991–2993.
Meyer, G.A., Wells, S.G., 1997. Fire-related sedimentation events on alluvial fans, Yellowstone National Park, U.S.A. Journal of Sedimentary Research 67, 776–791.
Meyer, G.A., Pierce, J.L., Wood, S.L., Jully, A.J.T., 2001. Fire, storms and erosional events in the Idaho Batholith. Hydrological Processes 15, 3025–3038.
Morris, S.E., Moses, T.A., 1987. Forest fire and the natural soil erosion regime in the Colorado Front Range. Annals of the Association of American Geographers 77, 245–254.
Odion, D.C., Hanson, C.T., 2006. Fire severity in conifer forests of the Sierra Nevada, California. Ecosystems 9, 1177–1189.
Omernik, J.M., 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77, 118–125.
Pak, J.H., Kou, Z., Kwon, H.J., Lee, J.-J., 2009. Predicting debris yield from burned watersheds: comparison of statistical and artificial neural network models. Journal American Water Resources Association 45 (1), 210–223.
Pierson, T.C., Costa, J.E., 1987. A rheologic classification of subaerial sediment-water flows: Geological Society of America. Reviews in Engineering Geology 7, 1–12.
Pradhan, B., Lee, S., 2010. Landslide susceptibility assessment and factor effect analysis: back propagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. Environmental Modelling & Software 25 (6), 747–759.
Rallo, R., Ferre-Gine, J., Arenas, A., Giralt, F., 2002. Neural virtual sensor for the inferential prediction of product quality form process variables. Computers and Chemical Engineering 26 (12), 1735–1754.
Ritter, H., Schulten, K., 1986. On the stationary state of Kohonen's self-organizing sensory mapping. Biological Cybernetics 54 (1), 99–106.
Robichaud, P.R., 2000. Fire effects on infiltration rates after prescribed fire in Northern Rocky Mountain forests, USA. Journal of Hydrology 231–232, 220–229.
Rubinstein, R.Y., Kroese, D.P., 2007. Simulation and the Monte Carlo Method, second ed. John Wiley & Sons, New York, ISBN 9780470177938.
Running, S.W., 2006. Is global warming causing more, larger wildfires? Science 313, 927–928.
Sanchez-Martos, F., Aguilera, P.A., Garrido-Frenich, A., Torres, J.A., Pulido-Bosch, A., 2002. Assessment of groundwater quality by means of self-organizing maps: application in a semi-arid area. Environmental Management 30 (5), 716–726.
Seibert, J., McDonnell, J.J., Woodsmith, R.D., 2010. Effects of wildfire on catchment runoff response: a modeling approach to detect changes in snow-dominated forested catchments. Hydrology Research 41 (5), 378–390.
Shanmuganathan, S., Sallis, P., Buckeridge, J., 2006. Self-organising map methods in integrated modeling of environmental and economic systems. Environmental Modelling & Software 21, 1247–1256.
Stickney, P.F., Campbell, R.B., 2000. Database for Early Postfire Succession in Northern Rocky Mountain Forests. General Technical Report RMRS-GTR-61CD. USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO, 21 pp.
Suk, H., Lee, K.K., 1999. Characterization of a ground water hydrochemical system through multivariate analysis: clustering into ground water zones. Ground Water 37, 358–366.
Swanson, F.J., 1981. Fire and geomorphic processes. In: Mooney, H.A., Bonniksen, T.H., Christensen, N.L., Lotan, J.E., Reiners, W.A. (Eds.), Fire Regimes and Ecosystem Properties (USDA Forest Service General Technical Report WO-26, 401-420). US Department of Agriculture, Washington, DC.
U.S. Department of Agriculture, 1994. State Soil Geographic Data Base: Data Use Information, Miscellaneous Publication Number 1492, pp. 35.
U.S. Geological Survey, 2009. , National Elevation Dataset. http://eros.usgs.gov/#Find_Data/Products_and_Data_Available/NED (last viewed 2-15-2010).
Uchida, T., Kosugi, K., Mizuyama, T., 2001. Effects of pipeflow on hydrological process and its relation to landslide: a review of pipeflow studies in forested headwater catchments. Hydrological Processes 15 (11), 2151.
Vesanto, J., Alhoniemi, E., 2000. Clustering of the self organized map. IEEE Transactions on Neural Networks 11 (3), 586–600.
Vesanto, J., 1999. SOM-based data visualization methods. Intelligent Data Analysis 3, 111–126.
Wang, S., 2003. Application of self-organising maps for data mining with incomplete data sets. Neural Computing and Applications 12, 42–48.
Westerling, A.L., Swetnam, T.W., 2003. Interannual to decadal drought and wildfire in the western United States. Eos, Transactions, American Geophysical Union 84 (545), 554–555.
Wieczorek, G.F., Lips, E.W., Ellen, S.D., 1989. Debris flows and hyperconcentrated floods along the Wasatch front, Utah, 1983 and 1984. Bulletin of the Association of Engineering Geologists 26 (2), 191–208.
Winter, T.C., 2001. The concept of hydrologic landscapes. Journal American Water Resources Association 37, 335–349.
Wolock, D.M., Winter, T.C., McMahon, G., 2004. Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses. Environmental Management 34, S71–S88.
Wondzell, S.M., King, J.G., 2003. Post-fire erosional processes in the Pacific Northwest and Rocky Mountain Regions. Forest Ecology and Management, 1775–1787.