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

A R T I C L E I N F O

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

A B S T R A C T

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.

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: