

DISSERTATION

ESTIMATION OF CATCHMENT-SCALE SOIL MOISTURE PATTERNS FROM  
TOPOGRAPHY AND RECONSTRUCTION OF A PRESERVED ASH-FLOW  
PALEOTOPOGRAPHY

Submitted by

Michael Lee Coleman

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2012

Doctoral Committee:

Advisor: Jeffrey D. Niemann

José D. Salas

Timothy R. Green

Stephanie K. Kampf

## ABSTRACT

### ESTIMATION OF CATCHMENT-SCALE SOIL MOISTURE PATTERNS FROM TOPOGRAPHY AND RECONSTRUCTION OF A PRESERVED ASH-FLOW PALEOTOPOGRAPHY

This dissertation consists of three parts, two of which examine methods for estimating spatial soil moisture patterns while the third investigates the reconstruction of a fluvially-eroded paleotopography. Part I of the dissertation evaluates unsupervised machine-learning techniques' effectiveness for estimating soil moisture patterns and compares them with linear regression. Physical processes that impact soil moisture are typically expressed as nonlinear functions, but most previous research on the estimation of soil moisture has relied on linear techniques. In the present work, two machine learning techniques, a spatial artificial neural network (SANN) and a mixture model (MM), that can infer nonlinear relationships are compared with multiple linear regression (MLR) for estimating soil moisture patterns using topographic attributes as predictor variables. The methods are applied to time-domain reflectometry (TDR) soil moisture data collected at three catchments with varying characteristics (Tarrawarra, Satellite Station, and Cache la Poudre) under different wetness conditions. The methods' performances with respect to the number of predictor attributes, the quantity of training data, and the attributes employed are compared using the Nash-Sutcliffe Coefficient of Efficiency (NSCE) as the performance measure. The performances of the methods are dependent on the site studied, the average soil moisture and the quantity of training data provided. Although the methods often perform similarly, the best performing method overall is the SANN, which incorporates additional predictor variables more effectively than the other methods.

Next, Part II of the dissertation presents the development and testing of a new conceptually-based model for estimating soil moisture patterns and describes the investigation of the climatic, vegetation and soil characteristics that affect pattern organization and temporal stability with the model. Soil moisture is a key hydrologic state variable for the Earth's surface affecting both energy and precipitation partitioning. Additionally, the nonlinear dependence of hydrologic processes on soil moisture means that not only is the average moisture condition important for many applications, but the spatial patterns of soil moisture are also important. At the catchment scale, soil moisture patterns have been observed to exhibit different types of dependence on topography. Some catchments have their wettest locations in the valley bottoms, while others have their wettest locations on hillslopes that are oriented away from the sun. Additionally, some catchments have moisture patterns that maintain a similar organization at all times (time stability), while other catchments have soil moisture patterns that change through time (time instability). Although these tendencies are well known, the reasons for their occurrence at a particular catchment are not well understood. In this paper, we investigate the conditions under which the different types of topographic dependence and different degrees of time instability occur through the use of a new conceptual model. The type of topographic dependence and the degree of instability are quantified by two metrics that are also introduced in the paper, and the effects of soil, vegetation, and climatic parameters on these metrics are then evaluated. The evaluations indicate that saturated horizontal hydraulic conductivity, pore disconnectedness, vegetation evapotranspiration efficiency, and an exponent relating the horizontal hydraulic gradient to the topographic slope have the strongest effects on the organization and instability of the soil moisture patterns. In contrast, annual potential evapotranspiration alone does not strongly impact the organization or its stability.

Finally, Part III of the dissertation describes the modification of a previously-developed interpolation scheme for fluvial topography and the reconstruction of a paleotopography that may be potentially important to groundwater movement by the modified method. Many applications in geology require estimation of the depth and thickness of lithologic layers based on limited observations. The boundaries of such layers are typically estimated using Kriging or other estimation methods that produce smooth surfaces. In some cases, however, smooth surfaces may be inappropriate. A boundary that is formed by a preserved hillslope and valley paleotopography, in particular, is expected to exhibit drainage characteristics and inherent roughness that are not consistent with standard estimation methods. This paper discusses the generalization of a technique originally designed to interpolate fluvially-eroded topography. The method incorporates a simple river basin evolution model to generate realistic topography and adjusts an erodability parameter in space to match observed elevations. The method is generalized to allow flow to enter from outside the interpolation region, which is a likely scenario when reconstructing paleotopography. The method is then applied to the lower boundary of the Tshirege Member of the Bandelier Tuff, which underlies Los Alamos National Laboratory and Bandelier National Monument in north-central New Mexico. The method produces surfaces with major valleys that are consistent with previous observations. The method is also applied in a framework that estimates the likelihood that any particular point within the interpolation region drains through a specified boundary. Although the surfaces vary between simulations, most portions of the interpolation domain drain through consistent boundaries.

## ACKNOWLEDGEMENTS

First, I would like to express my sincere appreciation to my advisor Dr. Jeffrey Niemann for his support, mentoring, patience and encouragement that have made this work possible. Special thanks also go to Dr. Timothy Green for providing comments and suggestions on the work in all parts of this dissertation as well as data used in Parts I and II. I would also like to thank the other members of my graduate committee, Dr. José Salas and Dr. Stephanie Kampf, for their suggestions and comments for improving the work in this dissertation.

I also express my thanks to the U.S. Army Research Office and the Colorado State University Center for Geosciences/Atmospheric Research for their support of my research. In addition, I want to thank the different scholarship sponsors at the Department of Civil and Environmental Engineering at Colorado State University for their support.

Finally, I want to thank my family and friends for their support and encouragement and all of the members of the Niemann Research Group and the B10 laboratory for making my time at CSU an enjoyable experience.

## TABLE OF CONTENTS

GENERAL INTRODUCTION.....	1
PART I: AN EVALUATION OF NONLINEAR METHODS FOR ESTIMATING CATCHMENT-SCALE SOIL MOISTURE PATTERNS BASED ON TOPOGRAPHIC ATTRIBUTES .....	4
INTRODUCTION .....	5
METHODS .....	9
MULTIPLE LINEAR REGRESSION (MLR) .....	9
SPATIAL ARTIFICIAL NEURAL NETWORK (SANN).....	9
MIXTURE MODEL (MM).....	12
EXPERIMENTAL DESIGN .....	13
APPLICATION SITES AND DATA.....	14
TARRAWARRA.....	15
SATELLITE STATION .....	16
CACHE LA POUFRE.....	17
TOPOGRAPHIC ATTRIBUTES.....	18
RESULTS .....	19
TYPICAL RESULTS .....	19
NUMBER OF PREDICTOR VARIABLES .....	21
SAMPLE SIZE .....	23
PREDICTOR SETS CHOSEN BY METHODS .....	25
CONCLUSIONS.....	28
ACKNOWLEDGMENTS.....	29

REFERENCES .....	30
PART II: CONTROLS ON TOPOGRAPHIC DEPENDENCE AND TIME-INSTABILITY IN CATCHMENT-SCALE SOIL MOISTURE PATTERNS.....	35
NOTATION.....	36
INTRODUCTION .....	37
MODEL DEVELOPMENT .....	40
APPLICATION SITES.....	46
MODEL CALIBRATION AND PERFORMANCE RESULTS .....	51
DEVELOPMENT OF PATTERN AND INSTABILITY METRICS .....	59
RESULTS FOR PATTERN AND INSTABILITY METRICS .....	62
CONCLUSIONS.....	69
ACKNOWLEDGMENTS.....	72
REFERENCES .....	73
APPENDIX: EMT MODEL DERIVATION.....	77
PART III: RECONSTRUCTION OF HILLSLOPE AND VALLEY PALEOTOPOGRAPHY BY APPLICATION OF A GEOMORPHIC MODEL .....	90
INTRODUCTION .....	91
STUDY SITE.....	93
METHOD DESCRIPTION .....	95
METHOD APPLICATION.....	98

RESULTS.....	103
CONCLUSIONS.....	111
ACKNOWLEDGMENTS.....	112
REFERENCES .....	113
SUMMARY.....	117

## LIST OF TABLES

Table 1 - Parameter sets that produce soil moisture estimates with maximum average NSCE for each catchment .....	52
Table 2 - Statistics of the NSCE values from all available dates at the three catchments when the best-performing parameter sets are used. ....	54
Table 3 - Statistics of mean absolute deviation [ <i>Ahmadzadeh and Petrou, 2001</i> ] between simulated and observed elevations.. ....	106

## LIST OF FIGURES

Figure 1 - (a) Elevations relative to the low point at each study site and spatial patterns of soil moisture for (b) dry, (c) medium and (d) wet dates used. ....16

Figure 2 - (a) Observed soil moisture patterns at each study site, and soil moisture patterns estimated by (b) MLR, (c) SANN and (d) MM. The dots indicate training data locations. ....20

Figure 3 - Median NSCE for each estimation method as the number of topographic attributes used increases.....22

Figure 4 - Box and whisker plots characterizing the performance of each estimation method as the amount of training data used increases. Plus symbols indicate outliers, which are defined as values that are more than 1.5 times the distance between the upper and lower quartiles away from the box limits. ....24

Figure 5 - Proportion of times each topographic attribute is selected by the best-performing models as a function of the number of topographic attributes used.....26

Figure 6 - Topography observed at the Tarrawarra, Satellite Station and Cache la Poudre catchments. Shading shows elevation (m) relative to each catchment’s outlet, and lines show elevation contours.....48

Figure 7 - Representative observed (a, b, c), EMT model (d, e, f), and implicit (equation 2.6) (g, h, i) soil moisture patterns for Tarrawarra. The observed soil moisture patterns are from 14 Feb. 1996, 27 Sep. 1995, and 2 May 1996, and the spatial-average moisture values are 0.26, 0.38, and 0.42 V/V for a, b, and c, respectively. NSCE values for the EMT model are 0.12, 0.54, and 0.53.....49

Figure 8 - Representative observed (a, b, c), EMT model (d, e, f), and implicit (equation 2.6) (g, h, i) soil moisture patterns for Satellite Station. The observed soil moisture patterns are from 18 Feb. 1999, 25 Mar. 1998, and 30 Oct. 1999, and the spatial-average moisture values are 0.25, 0.32, and 0.40 V/V for a, b, and c, respectively. NSCE values for the EMT model are 0.19, 0.24, and 0.14. ....50

Figure 9 - Representative observed (a, b, c), EMT model (d, e, f) and implicit (equation 2.6) (g, h, i) soil moisture patterns for Cache la Poudre. The observed soil moisture patterns are from 24 June 2008, 12 June 2008, and 28 May 2008, and the spatial-average moisture values are 0.04, 0.11, and 0.19 V/V for a, b, and c, respectively. NSCE values for the EMT model are 0.07, 0.11, and 0.00. ....51

Figure 10 - Calibrated LFI and ETI patterns for the Tarrawarra, Satellite Station, and Cache la Poudre catchments. ....57

Figure 11 - Calibrated weights plotted as a function of spatial average soil moisture for (a) Tarrawarra, (b) Satellite Station and (c) Cache la Poudre. Markers indicate observed spatial-average soil moisture values and the associated process weights. ....59

Figure 12 - Soil moisture variance introduced by the LFI and ETI patterns and their sum for (a) Tarrawarra, (b) Satellite Station and (c) Cache la Poudre. Markers indicate observed spatial-average soil moisture values and the associated variances. ....64

Figure 13 - Difference between soil moisture variances introduced by the LFI and ETI patterns (n) as a function of the spatial average soil moisture ( $\bar{\theta}$ ) for the Tarrawarra, Satellite Station and Cache la Poudre catchments. ....65

Figure 14 - Behavior of the pattern metric as the model parameters are changed at the Tarrawarra catchment. Bars indicate the median metric value and boxes span the IQR. ....66

Figure 15 - Behavior of the instability metric as the model parameters are changed at the Tarrawarra catchment. Bars indicate the median metric value and boxes span the IQR. 69

Figure 16 - Simplified upper stratigraphy of Pajarito Plateau (after Broxton and Vaniman [2005]). Western limit in figure is beginning of Jemez Mountains and eastern limit is White Rock Canyon.....94

Figure 17 - Exposed (a) valley and (b) hillslope cross-sections in lower boundary of Tshirege Member in Alamo Canyon, BNM. Scales shown are approximate. ....95

Figure 18 - Surveyed hillslope profiles in lower boundary of Tshirege Member from (a) a road cut on New Mexico Highway 4 near White Rock and (b) an exposed outcrop in Alamo Canyon in BNM. Second degree polynomial regression curves are of form  $z = ax^2 + bx + c$ .  $D$  values are estimated from regressions by formula  $D = -1/2a$ . Elevation values given are relative to an arbitrary datum for each site..... 100

Figure 19 - Relationships between local slope and contributing area from two basins (Frijoles and Water Canyons) on present topography of the Pajarito Plateau. (a) Slope-area relationship and linear regression used to estimate  $m$  and  $\beta$  for interpolation method. (b) Area-slope relationships for same basins and regressions used to estimate  $\gamma$  parameter for assigning contributing area to closed-boundary inlets. In both plots, circles and solid lines refer to Frijoles Canyon while stars and dashed lines refer to Water Canyon..... 102

Figure 20 - Map showing the locations where the elevation of the lower boundary of the Tshirege Member has been observed. Points are labeled according to their LANL designations. .... 104

Figure 21 - Estimates of lower boundary of Tshirege Member based on (a) a completely-regularized spline and (b) physically-based interpolation surface with base parameters. Circles represent elevation measurement locations, and vertical lines show deviation between observed and simulated elevations. .... 105

Figure 22 - Map of likelihood each point ultimately drains through southern boundary. Circles indicate elevation measurement locations. .... 108

Figure 23 - Surfaces generated by physically-based interpolation method when individual parameters are changed. Altered parameter values are: (a)  $\phi = 0.0035$ , (b)  $\phi = 0.0055$ , (c)  $m = 0.32$ , (d)  $m = 0.45$ , (e)  $\gamma = 1.15$ , and (f)  $\gamma = 1.70$ . Circles represent elevation measurement locations, and vertical lines show deviation between observed and simulated elevations. .... 109

Figure 24 - Maps of likelihood that each point drains through southern boundary. In each map, an individual parameter value was changed, and 100 simulations were conducted. Altered parameters are: (a)  $\phi = 0.0035$ , (b)  $\phi = 0.0055$ , (c)  $m = 0.32$ , (d)  $m = 0.45$ , (e)  $\gamma = 1.15$ , and (f)  $\gamma = 1.70$ . Circles indicate elevation measurement locations. .... 110

## **General Introduction**

This dissertation is the result of studies of methods to estimate spatial soil moisture patterns and to interpolate topography formed by fluvial processes. The dissertation is divided into three parts: Part I is an evaluation of nonlinear techniques for estimating spatial patterns of soil moisture; Part II describes the development and evaluation of a new model for estimating spatial soil moisture patterns; and Part III describes the modification of a method that uses a landscape evolution model to interpolate topography and the application of the modified method to reconstruct a preserved paleotopography potentially important to groundwater transport.

The soil water content (soil moisture state) of terrestrial surfaces is an important state variable for the hydrologic cycle at multiple scales. Incoming radiation and heat are partitioned at the earth's surface into either latent heat in the vapor phase from evaporation of surface water (including soil moisture) or into sensible heat due to temperature changes. The resulting partition and its effects on surface and air temperatures have important ramifications for weather patterns such as the persistence of droughts. Furthermore, the soil moisture state affects the partitioning of precipitation falling onto terrestrial surfaces into infiltration and runoff and, therefore, downstream hydrographs. In addition to hydrologic processes, the soil moisture also influences other earth surface processes, including vegetation patterns, weathering, erosion and soil production.

Not only is the overall soil moisture content important to hydrologic processes but spatial patterns of moisture influence them as well. For example, one mechanism for runoff generation occurs when precipitation falls on surfaces already saturated with water. In this case, the presence of a hydraulically-connected pattern of saturated areas can lead to greater runoff than if

the saturated areas were in the form of multiple isolated subareas. As a result, knowledge of the spatial distribution of soil moisture is important for hydrologic applications.

Unfortunately, it is currently difficult, if not impossible, to measure surface (top 5 to 30 cm) soil moisture values with both extents and resolutions (spatial and temporal) appropriate for many applications. Soil moisture at approximately 1 km resolutions can currently be inferred from remote-sensing techniques; however, soil moisture content may vary substantially within the grid-cell resolution of those methods. Alternatively, manual measurement techniques may be employed to obtain point values of moisture, but the resources necessary to measure values representing useful spatial extents for meaningful durations can be prohibitive. Therefore, accurate estimation of soil moisture patterns over useful extents with adequate resolution is highly desirable for hydrologic uses and is the objective of the first two parts of this dissertation. Because many hydrologic processes affecting soil moisture are conceived as nonlinearly-related to the moisture state and because the great majority of previous research has focused on linear estimation methods, Part I of this work focuses on evaluating nonlinear techniques for estimating soil moisture patterns. Two nonlinear methods and multiple linear regression are tested and their performance compared using soil moisture datasets from three topographically and climatically diverse catchments. In Part II, a conceptual model for estimating soil moisture from topography and the average spatial soil moisture content is developed and tested. The model in Part II was developed partially due to the nonlinear techniques' marginal improvement in pattern estimation relative to linear estimation techniques. The model is applied to the same three catchment datasets as used in Part I and its performance is comparable to that of the other methods; however, the model is somewhat simpler to apply than the nonlinear techniques in Part I and is more flexible in its data requirements.

Part III of the dissertation modifies a method previously developed to interpolate fluvially-eroded topography [Niemann, 2000]. The method was modified to allow for incoming flow across area boundaries, which permits the modified method to be used for areas that do not represent entire watersheds. The interpolation method is based on a landscape evolution model, which ties Part III to the first two parts of the dissertation, albeit loosely, because as mentioned previously soil moisture states affect landscape development processes such as erosion and soil development. The modified model is applied to provide feasible reconstructions of a paleotopography that indicates evidence of fluvial erosion and is preserved between two volcanic ash-flow geologic layers. Fluvial drainages within the paleotopographic layer have been thought to be potentially important hydraulic conduits for groundwater and as such their spatial characteristics would be valuable for modeling purposes.

**Part I: An evaluation of nonlinear methods for estimating catchment-scale soil moisture patterns based on topographic attributes**

## Introduction

Soil moisture is an important hydrologic state variable owing to its influence on a variety of surface hydrologic processes and land surface-atmospheric interactions. For example, soil moisture affects both the partitioning of radiation into sensible and latent heat [Entekhabi *et al.*, 1996] and the partitioning of rainfall into infiltration and runoff [Dunne and Black, 1970]. Additionally, soil moisture influences vegetation patterns [Eagleson, 1978], land surface erosion processes [Moore *et al.*, 1988], and soil development [Hillel, 1998]. Spatial patterns of soil moisture and their characteristics, such as connectivity of wet areas, are also important for hydrologic considerations [Dunne and Black, 1970; Dunne *et al.*, 1975; Hewlett and Hibbert, 1967; Western *et al.*, 2001]. Soil moisture is typically observed using remote-sensing techniques or manual techniques, such as time-domain reflectometry (TDR). Unfortunately, neither approach is practical for observing soil moisture patterns within catchments at suitable resolutions [Robinson *et al.*, 2003; Vereecken *et al.*, 2008]. The desire for accurate characterization of moisture patterns coupled with the difficulties in observing the patterns has led to many efforts to estimate moisture patterns [Bardossy and Lehmann, 1998; Nyberg, 1996; Sulebak *et al.*, 2000; Western *et al.*, 1999a; Yates and Warrick, 1987].

Significant research has been devoted to investigating the correlations between soil moisture and topographic attributes and the effectiveness of multiple linear regression (MLR) for estimating soil moisture using topographic attributes as predictive data. Topographic attributes have been used because organized patterns of soil moisture resemble patterns of topography and because surface elevation data are readily available for nearly all parts of the world. Zaslavsky and Sinai [1981] explained 81% of soil water variation two weeks after rainfall by curvature in an agricultural field near Beer-Sheba, Israel. Moore *et al.* [1988] found that 33% of the soil moisture variation on a transect of a 7.5 ha catchment in Australia could be explained by the

wetness index, which is defined as the ratio of the specific contributing area and the local slope, and that 41% of the variation could be explained by using both the wetness index and the topographic aspect. Nyberg [1996] explained between 15% and 42% of soil moisture variation by correlations with elevation, slope, wetness index, and the logarithm of the contributing area. Western et al. [1999a] were able to explain up to 61% of spatial soil moisture variation under relatively wet conditions but only 22% in drier conditions by a combination of potential solar radiation index (PSRI), which is the ratio of the potential insolation of a surface with a particular slope and aspect to a hypothetical horizontal surface at the same location, and either the wetness index or the logarithm of the contributing area. Sulebak et al. [2000] found the combination of slope, aspect, and profile curvature could explain 70% of moisture variation at two locations in Sweden. Green and Erskine [2004] found the highest correlations between soil moisture and topographic attributes for agricultural fields in Colorado on the wettest date considered and the strongest correlation with slope even though that attribute only explains approximately 20% of the variance. Despite some instances where MLR is effective, a conceptual inconsistency exists in linearly regressing soil moisture on topographic attributes because those attributes are generally associated with physical processes that relate nonlinearly to soil moisture [Rodriguez-Iturbe, 2000]. Such nonlinearities might produce nonlinearity in the relationships between soil moisture and topographic attributes. For example, Western et al. [1999a] noted that scattergraphs indicate a possible nonlinear relationship between the wetness index and soil moisture data from organized patterns in their dataset.

Geostatistical techniques [Chiles and Delfiner, 1999; Journel and Huijbregts, 1978; Kitanidis, 1993] have also been applied to the tasks of soil moisture characterization and estimation. Yates and Warrick [1987] used cokriging with bare soil temperature and percent sand content as

ancillary variables to estimate soil moisture in Arizona. They found that if the ancillary variable is well correlated with soil moisture then the cokriging estimates are better than estimates from ordinary kriging. Bardossy and Lehmann [1998] estimated soil moisture patterns from a 630 ha catchment in Germany with several geostatistical techniques. They found that Bayes-Markov Updating (BMU), a simplified form of Bayes-Markov Kriging [Zhu and Journel, 1993], has the lowest errors of all the tested methods for the conditions analyzed. BMU can incorporate ancillary data in a nonlinear manner and its performance with either the wetness index or land use as ancillary data is better than both ordinary kriging and external drift kriging. The method performs slightly better with the wetness index than with land use for that dataset. A common assumption made in geostatistical analyses is that of a stationary random field, but previous research has indicated that soil moisture patterns are not random but exhibit spatial organization [Dunne et al., 1975; Rodriguez-Iturbe et al., 1995; Western et al., 1999a].

In addition to those standard methods, other methods have also been developed and/or employed for estimating soil moisture patterns. Western et al. [1999] applied LOWESS regression [Hirsch et al., 1993] to estimate soil moisture based on topographic attributes and found that it does not substantially improve the amount of variance explained compared with linear regression.

However, the extent of their LOWESS analyses in that work is not clear. Wilson et al. [2005] developed a linear estimation method with coefficients that vary nonlinearly with the spatial average soil moisture and applied the method to data from the Maharungi catchment in New Zealand. They found that the use of terrain attributes alone does not estimate realistic spatial moisture patterns but that the inclusion of a spatially stable residual pattern improves the estimates and concluded that factors other than topography are also important to soil moisture patterns. Finally, Perry and Niemann [2007; 2008] used empirical orthogonal functions (EOFs)

to decompose a time series of spatial moisture patterns into patterns of covariation that are present to some extent on every date and interpolated those stable patterns with linear regressions against topographic attributes.

Our objectives are to evaluate the abilities of selected nonlinear estimation techniques for estimating soil moisture from sparse soil moisture observations and to compare the results of those techniques to the results of MLR. We hypothesize that the nonlinear techniques will avoid some of the previously mentioned shortcomings of other methods and improve estimation of spatial soil moisture patterns. The nonlinear estimation methods used in this research are a spatial artificial neural network (SANN) [Shin and Salas, 2000a] and mixture modeling (MM) with multivariate Gaussian distribution functions [McLachlan and Peel, 2000]. These methods were selected because they are unsupervised machine learning techniques that do not assume *a priori* any specific form of relationship between soil moisture and the predictor data. In addition, the SANN and MM are both kernel density estimation methods, which attempt to estimate the joint probability density function between soil moisture and the predictor variables. Finally, Green et al. (2007) applied the SANN to crop yields, which are closely related to soil moisture, and found good performance. Most previous applications of machine learning techniques in hydrology have focused on time-series data [Lin et al., 2006; Liu et al., 2008; Wang et al., 2009], including those focused specifically on soil moisture [Ahmad et al., 2010; Elshorbagy and El-Baroudy, 2009; Gill et al., 2006]. Here, the SANN and MM are applied to three study areas representing diverse climates and landscape characteristics to test their predictive abilities when different processes may dominate the moisture pattern formation.

## Methods

### Multiple linear regression (MLR)

MLR is used as a baseline estimation method due to its simplicity and common use. The general model for linear regression can be written in matrix form as:

$$\boldsymbol{\theta} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (1.1)$$

where  $\boldsymbol{\theta}$  is an  $n$  by 1 vector of observed responses (soil moisture in this analysis),  $n$  is the number of observations,  $\mathbf{X}$  is an  $n$  by  $d+1$  matrix,  $d$  is the number of predictor variables (topographic attributes),  $\boldsymbol{\beta}$  is a  $d+1$  by 1 vector of unknown coefficients, and  $\boldsymbol{\varepsilon}$  is an  $n$  by 1 vector of residuals, or errors. Use of the ordinary least squares (OLS) criterion for estimating the coefficient vector leads to the following equation for the coefficient estimates:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\theta} \quad (1.2)$$

Ordinary least-squares provides the minimum-variance unbiased coefficient estimates regardless of the distribution properties of the errors [Draper and Smith, 1981]. The main drawback of linear regression for this application is that it cannot account for possible nonlinear relationships between the topographic attributes and soil moisture.

### Spatial artificial neural network (SANN)

The SANN method was developed by Shin and Salas [2000a] and can be viewed as a specific implementation of the Nadaraya-Watson model, or kernel regression [Bishop, 2006; Nadaraya, 1964; Watson, 1964]. The SANN has been used previously for regional drought analysis [Shin and Salas, 2000b] and to estimate crop yields from topographic attributes [Green et al., 2007]. Martinez et al. [2004] investigated the sensitivity of the SANN to its internal parameters using grain yield data and found optimal parameter values, which were subsequently used by Green et al. [2007].

The SANN is similar to kernel density estimation using multivariate Gaussian kernels. We can represent soil moisture,  $\theta$ , at some location in space as a random variable in a  $d$ -dimensional domain by  $\theta(\mathbf{x})$ , where  $\mathbf{x} = [x_1, x_2, \dots, x_d]$  is a vector of topographic attributes associated with the same spatial location. The optimal estimator of the soil moisture value is then the conditional expectation given by [Bishop, 1995]:

$$E[\theta(\mathbf{x}) | \mathbf{x}] = \frac{\int_{-\infty}^{\infty} \theta(\mathbf{x}) p(\mathbf{x}, \theta) d\theta}{\int_{-\infty}^{\infty} p(\mathbf{x}, \theta) d\theta} \quad (1.3)$$

where  $p(\mathbf{x}, \theta)$  is the joint probability density function of  $\mathbf{x}$  and  $\theta$ . The probability density function is estimated using multivariate Gaussian kernel density estimation [Specht, 1991]. If we observe the soil moisture  $\theta(\mathbf{x})$  and a vector of topographic attributes  $\mathbf{x}$  at  $N$  locations given by  $[\mathbf{x}_n | n=1, \dots, N]$ , then the Gaussian kernel density estimator at any point  $\mathbf{x}$  in the domain is given by:

$$p(\mathbf{x}, \theta) = \frac{1}{N} \sum_{n=1}^N G(\mathbf{x} | \boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n) G(\theta | \mu_\theta, \sigma_\theta) \quad (1.4)$$

where

$$G(\mathbf{x} | \boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n) = \frac{1}{(2\pi)^{d/2} |\boldsymbol{\Sigma}_n|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_n)^T \boldsymbol{\Sigma}_n^{-1} (\mathbf{x} - \boldsymbol{\mu}_n)\right] \quad (1.5)$$

with  $\boldsymbol{\mu}_n$  and  $\boldsymbol{\Sigma}_n$  the mean and covariance, respectively, of the Gaussian kernel associated with the  $n^{\text{th}}$  observation. For simplicity, the variance of each kernel function is taken as equal in all dimensions of the predictor variable subspace so that all diagonal entries of  $\boldsymbol{\Sigma}_n$  are equal. Also, the covariances (off-diagonal entries) are assumed to be zero. The diagonal entries of  $\boldsymbol{\Sigma}_n$  are then denoted  $\sigma_n^2$ . Also,

$$G(\theta | \mu_\theta, \sigma_\theta) = \frac{1}{(2\pi)^{1/2} \sigma_\theta} \exp \left[ -\frac{(\theta - \mu_\theta)^2}{2\sigma_\theta^2} \right] \quad (1.6)$$

where  $\mu_\theta$  and  $\sigma_\theta$  are the mean and standard deviation, respectively, of the soil moisture. After substituting Equation (1.4) into Equation (1.3) and simplifying, the result is:

$$\hat{\theta}(\mathbf{x}) = \frac{\sum_{n=1}^N \theta(\mathbf{x}_n) G(\mathbf{x} | \boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)}{\sum_{n=1}^N G(\mathbf{x} | \boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)} \quad (1.7)$$

which may be used as a point estimator for  $\theta$ . Note that in the simplification the numerator terms involving  $G(\theta | \mu_\theta, \sigma_\theta)$  become the  $\theta(\mathbf{x}_n)$  terms inside the summation in Equation (1.7) while in the denominator each term involving  $G(\theta | \mu_\theta, \sigma_\theta)$  integrates to a value of 1.

In order to use Equation (1.7) to estimate soil moisture, the widths of the kernel functions need to be specified. The width of the kernel function centered on observation  $n$  is denoted  $\sigma_n$  and is calculated by:

$$\sigma_n = \frac{RMSD_n}{F} \quad (1.8)$$

where  $RMSD_n$  is the root mean-squared Euclidean distance (measured in the attribute domain) between data point  $n$  and its nearest  $P$  neighbors. The number of neighbors  $P$  and the factor  $F$  are the two parameters of the SANN. Both parameters help determine the spatial scale of the kernel function and their values must be specified prior to the SANN training. The parameter  $P$  relates directly to the kernel widths, and the effects of  $P$  on the individual kernel widths depend on the data configuration and density. The  $F$  parameter is related inversely to the kernel widths and affects all kernels to the same degree. We tested the effects of adjusting each parameter and found that comparable results were achieved through manipulation of either parameter.

Therefore, in the present implementation, the value of  $F$  was fixed at 2.5, which is the value recommended by Martinez et al. [2004] for large datasets, but  $P$  was considered a free parameter and various values were tested (see below). One potential drawback of the SANN is that it requires all the observations to be stored in order to make future estimates, which can make evaluation slow if the quantity of data is large.

### Mixture model (MM)

The MM method, like the SANN, is capable of capturing nonlinear relationships between the topographic attributes and soil moisture. Operationally, it is also similar to the SANN in that estimates are made by conditioning a multivariate density function with the values of the predictor variables. However, the number of kernel functions employed by MM to form the model is less than the number of observations, and the method employs the Expectation-Maximization (EM) algorithm [Dempster et al., 1977; McLachlan and Peel, 2000] to identify optimal locations for the kernels. In the present implementation, Gaussian functions were used as the kernel functions in the MM, and we will refer to them as components, which is common terminology in the MM literature. The multivariate density function developed by MM has the form [Bishop, 2006]:

$$p(\mathbf{x}, \boldsymbol{\theta}) = \sum_{k=1}^K \pi_k G(\{\mathbf{x}, \boldsymbol{\theta}\} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (1.9)$$

where  $K$  is the number of components in the model and is the only parameter of this method,  $\pi_k$  is the mixing coefficient associated with the  $k^{\text{th}}$  component, and  $\boldsymbol{\mu}_k$  and  $\boldsymbol{\Sigma}_k$  are the mean and covariance, respectively, of the  $k^{\text{th}}$  component. The mixing coefficient values lie between 0 and 1, and they sum to 1. The form of each Gaussian component is:

$$G(\{\mathbf{x}, \boldsymbol{\theta}\} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{1}{(2\pi)^{(d+1)/2} |\boldsymbol{\Sigma}_k|^{1/2}} \exp \left[ -\frac{1}{2} (\{\mathbf{x}, \boldsymbol{\theta}\} - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\{\mathbf{x}, \boldsymbol{\theta}\} - \boldsymbol{\mu}_k) \right] \quad (1.10)$$

where  $d$  is the number of predictor variables. The values of  $\boldsymbol{\pi}_k$ ,  $\boldsymbol{\mu}_k$  and  $\boldsymbol{\Sigma}_k$  are determined by the EM algorithm, which maximizes the likelihood of the model. Equation (1.3) is used again to estimate the soil moisture given the values of the ancillary data for any given location. The MATLAB statistics toolbox [Mathworks, 2009] version of MM, *gmdistribution*, was used to implement the MM method. As with the SANN, the density function is then conditioned on the observed value of the topographic attribute to determine the soil moisture estimate.

### Experimental design

A Monte Carlo cross-validation methodology, which provides an empirical measure of the methods' generalization abilities, was employed [Bishop, 2006; McLachlan and Peel, 2000] to assess the effectiveness of the different methods for estimating soil moisture. The cross-validation methodology divides each set of observations into training and testing sets. The training sets were created by randomly sampling the data available at each site 30 times for each of five sampling rates: 10%, 25%, 50%, 75%, and 95%. For each training set, the observations at unsampled locations comprise the testing set. Because all the methods are empirical, they require some amount of observations to develop a model. Each of the three estimation methods was used to develop a model of the relationship between topographic attributes and soil moisture values based on the training sets. After a model was developed from a specific training set, soil moisture was estimated with the model at all locations, but the Nash-Sutcliffe Coefficient of Efficiency (NSCE) [Nash and Sutcliffe, 1970] was calculated based only on the testing set to measure the model performance. The NSCE is defined as:

$$NSCE = 1 - \frac{\sum_{n=1}^N (\theta_n - \hat{\theta}_n)^2}{\sum_{n=1}^N (\theta_n - \bar{\theta})^2} \quad (1.11)$$

where  $N$  is the number of data points in the testing set,  $\theta_n$  is the  $n^{\text{th}}$  observed soil moisture,  $\hat{\theta}_n$  is the model estimate of the  $n^{\text{th}}$  observation, and  $\bar{\theta}$  is the average of the observations. Note that the NSCE has a maximum value of 1, for which all observed variance is explained, but there is no minimum value. Because 30 realizations of the training data are produced for every sampling rate, the typical performance of a particular estimation method is characterized by the median NSCE calculated from the 30 realizations.

A large number of estimation scenarios were performed to compare the methods. For each of the 3 methods, all 63 possible subsets of the 6 topographic attributes were tested as predictor variable sets. For each of those subsets of predictor variables, 30 sets of training data were supplied to the method for each of the 5 sampling rates given above. This collection of estimation problems was repeated for all 3 wetness conditions at all 3 study sites. The SANN and MM methods additionally have parameters that were adjusted. The  $P$  parameter in the SANN was tested with values of 6, 9, 15, 20, 30, and 45. Larger values of  $P$  lead to larger kernel widths and a smoother estimated density function. For the MM method, the adjustable parameter is  $K$ , the number of Gaussian components, which was varied from 1 to 4. The number of components controls the potential complexity of the resulting density function and the inferred relationship between the topographic attributes and the soil moisture.

### **Application Sites and Data**

Soil moisture data were compiled for three application sites: the Tarrawarra catchment in Southeastern Australia, the Satellite Station site in the Maharungi catchment on the North Island of New Zealand, and the Cache la Poudre site in north-central Colorado in the United States. Soil moisture patterns from three dates were used for each site. The dates were chosen to represent dry, moderate, and wet conditions as defined by the range of spatial mean soil moisture

values  $\Theta$  observed at each study site. All the soil moisture data were measured with TDR probes.

### **Tarrawarra**

The first soil moisture dataset is from the Tarrawarra catchment located near Melbourne, Australia and was originally described by Western and Grayson [1998]. The catchment has a temperate climate with average annual rainfall of 820 mm, average annual potential evapotranspiration of 830 mm, and a rainfall deficit in summer and excess in winter. The catchment is covered by pasture for cattle grazing. Soils generally have a silty loam A horizon and a B horizon with higher clay content and soil depths vary from 40 cm in the upper catchment to over 2 m in the low areas.

Soil moisture data were collected in the top 30 cm of the soil with TDR probes on thirteen dates between 27 September 1995 and 29 November 1996 on a 10 m by 20 m sampling grid [Western *et al.*, 1999b]. We filtered the dataset to remove locations with missing values on any date, so the remaining dataset includes 454 points. 14 February 1996, 27 September 1995, and 3 July 1996 were selected as the dry, medium, and wet dates, respectively. The spatial average soil moisture ( $\Theta$ ) values for those dates are 26.4, 38.0, and 45.3% volume of water/total volume (V/V), respectively. A digital elevation model (DEM) with 5 m resolution is available for this site. The topography consists of undulating hills (Figure 1) with no incised drainage channels and the catchment area is 10.5 ha. Total relief for the Tarrawarra catchment is 29 m. The dry moisture pattern does not visually exhibit any organization while the moderate and wet dates show strong and moderate degrees of organization, respectively, with wetter sites tending to occur in valley bottoms (Figure 1).

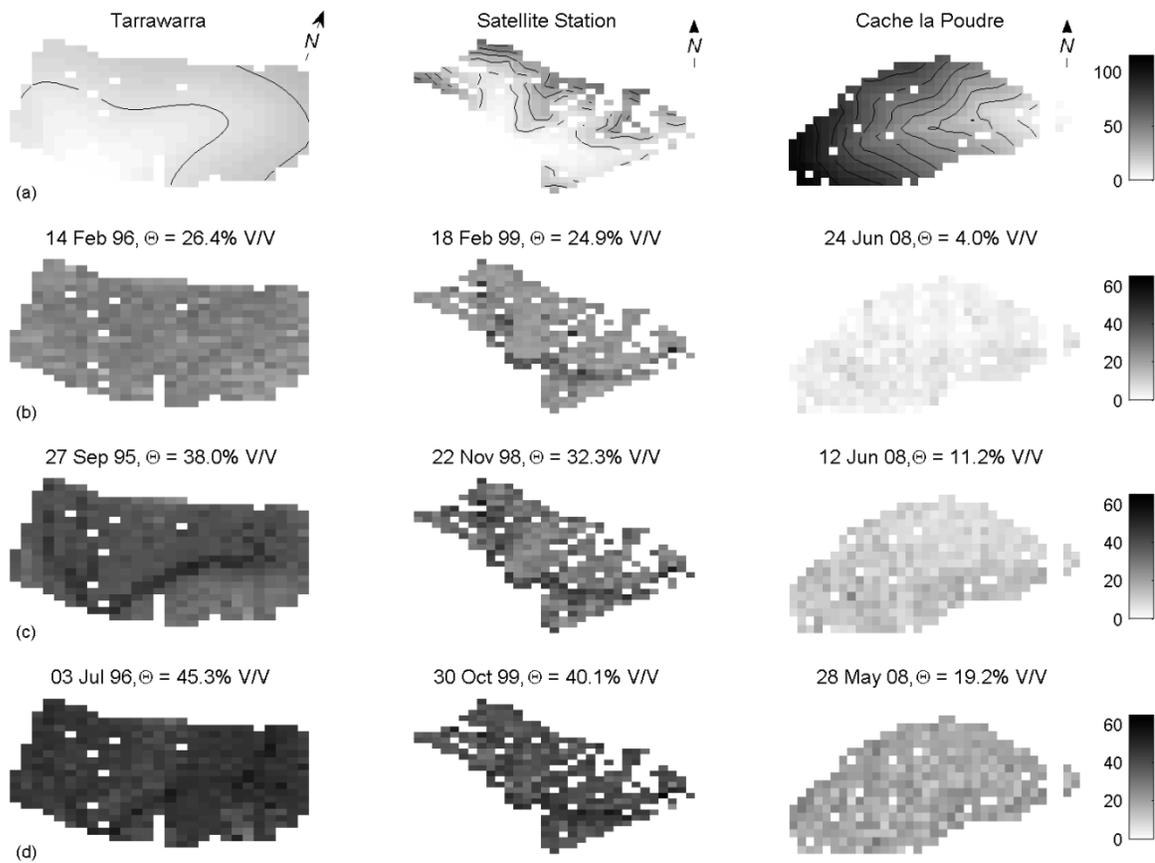


Figure 1 - (a) Elevations relative to the low point at each study site and spatial patterns of soil moisture for (b) dry, (c) medium and (d) wet dates used.

### Satellite Station

The second dataset is from the Satellite Station field site of the Maharungi River Variability Experiment (MARVEX) conducted in New Zealand [Wilson *et al.*, 2003; Woods *et al.*, 2001]. The Satellite Station site has a humid climate with annual rainfall approximately 1,200 mm and annual pan evapotranspiration of approximately 1,300 mm [Wilson *et al.*, 2003]. Hillslope soils are generally silty clay loam, while lowland valley soils are relatively deep alluvial fill with high clay content. The dataset used here is comprised of TDR measurements collected at 0-30 cm depth on a 40 m grid for six dates between April 1998 and November 1999. The dataset was again filtered so that it includes only locations with measurements on all dates (322 locations).

The dates chosen for this study are 18 February 1998, 25 March 1998, and 30 October 1999, with spatial average soil moisture values of 24.9, 32.3, and 40.1% V/V, respectively. The topographic data used was from a 10 m resolution DEM. The Satellite Station site has undulating terrain with a total area of approximately 60 ha, and it contains two subcatchments within it (Figure 1). The relief of the site is approximately 80 m. Visually, the moisture patterns at Satellite Station are more stable through time than those at Tarrawarra with the valley bottoms remaining wetter than the upland areas as the overall wetness condition changes (Figure 1).

### **Cache la Poudre**

The final dataset is from a catchment near Rustic, Colorado that is part of the Cache la Poudre River basin [*Lehman and Niemann, 2008*]. The catchment has a semi-arid climate with an annual precipitation of about 400 mm and annual potential evapotranspiration of about 930 mm. The vegetation on the north-facing slope of the catchment is predominantly coniferous forest with scattered shrubs in open areas and near ridges. The south-facing slope has scattered shrubs with a few coniferous trees. The site has thin sandy soil on the south-facing hillslope and thicker mineral soils overlaid with organic matter on the north-facing hillslope.

The sampling strategy for the Poudre site consisted of collecting manual TDR measurements immediately after a sequence of spring rainfall events and during subsequent periods of drying, which resulted in a total of 12 soil moisture patterns. Due to the shallow soils, particularly on the south-facing hillslope, surface soil moisture was measured in the top 5 cm of the soil once any litter layer was temporarily removed. After filtering to produce a consistent dataset for all sampling dates, a total of 350 locations remain for this dataset. The dates chosen for application of the interpolation methods are 24 June 2008, 12 June 2008, and 28 May 2008. The spatial average soil moisture values on these dates are 4.0, 11.2 and 19.2% V/V, respectively, which represent dry, medium, and wet conditions for this dataset. The DEM for the catchment has 15

m resolution, and a local coordinate system offset approximately 1 degree from north. The total relief for the Poudre site is approximately 124 m (Figure 1). The catchment consists of the headwater area for one incised channel with both steep and flat portions and the catchment area is approximately 8 ha. Visually, the soil moisture patterns exhibit slightly wetter conditions on the north-facing slope than the south-facing slope (Figure 1).

### **Topographic attributes**

The form of the topography for each catchment was characterized using elevation and five additional topographic attributes: slope, cosine of the topographic aspect ( $\cos A$ ), the logarithm of the specific contributing area ( $\log SCA$ ), the sum of the plan and profile curvatures ( $Curv$ ), and the potential solar radiation index (PSRI). These attributes are related to different processes and variables that affect soil moisture. Specifically, surface slope is related to the horizontal hydraulic gradient of subsurface flows and to insolation, a primary driver of evapotranspiration and snowmelt [Western *et al.*, 1999a]. Aspect also affects insolation [Western *et al.*, 1999a]. The specific contributing area (SCA) is a measure of the upslope area that can potentially contribute flow to a unit length of contour on the surface [Western *et al.*, 1999a]. The profile curvature is related to the change in the hydraulic gradient and hence the velocity of flow [Mitasova and Hofierka, 1993], and the plan curvature is related to the degree of surface and flow convergence along an elevation contour [Mitasova and Hofierka, 1993]. The number of attributes was intentionally limited to allow testing of all possible attribute combinations in the analyses below. The wetness index [Beven and Kirkby, 1979] was not included because it combines  $\log SCA$  and slope in a predetermined way. If wetness index is important, the nonlinear methods should be able to identify that importance using the underlying variables.

The SCA and the slope were calculated using the  $D_\infty$  algorithm [Tarboton, 1997] as implemented in TauDEM [Tarboton, 2008]. SCA was initially used as a predictor variable but its large

skewness resulted in numerical complications and led us to use log-transformed values instead. The aspect for each DEM cell was calculated according to the formula in Moore et al. [1991] with units of degrees clockwise from north. The aspect values were then cosine-transformed, which provides a continuous range of values from -1 to 1 and separates north-facing aspects from south-facing aspects by sign [Green et al., 2007].

The sum of the profile and plan curvatures [Mitasova and Hofierka, 1993] was the only curvature measure used. The profile curvature is the curvature of the intersection of the topographic surface with a vertical plane oriented in the downslope direction. The plan curvature is the curvature of the intersection of the surface with a horizontal plane. The sum of the two curvatures was chosen based on a screening analysis designed to select the single curvature measure with the highest overall potential for estimating soil moisture.

The PSRI depends on the day of the year, latitude, local slope, and aspect [Dingman, 2002; Western et al., 1999a]. No allowances were made for other factors affecting actual insolation, such as shading by vegetation or atmospheric attenuation, in the PSRI calculation. The PSRI was calculated individually for each sampling date, so it is the only topographic attribute that changes between different dates at the same catchment.

## **Results**

### **Typical results**

Figure 2 presents observed soil moisture patterns and moisture patterns estimated by each method when 25% of the available data is used for training. The patterns shown are from the set of topographic attributes that produces the highest median NSCE for each method and the training data that produce the results closest to that median NSCE among the different samples (using those topographic attributes). The  $P$  value used in the SANN is 45, which gave the best performance among all tested values. The MM method uses 2 component densities.

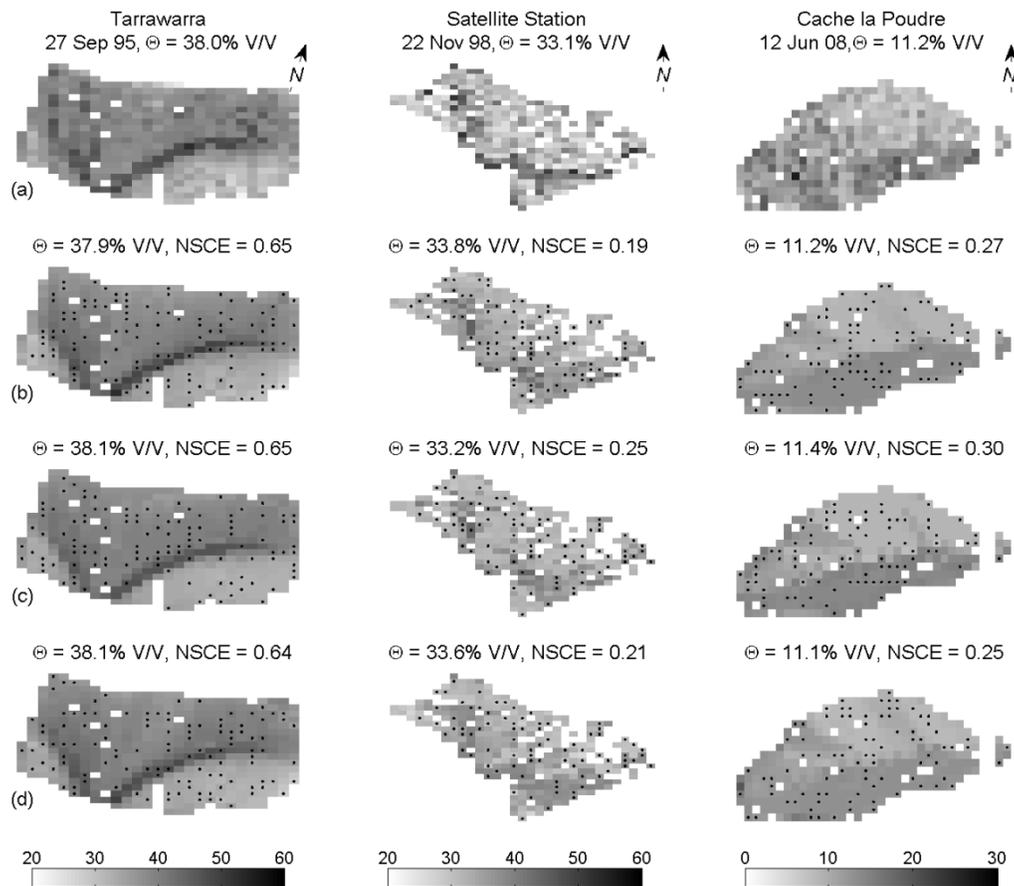


Figure 2 - (a) Observed soil moisture patterns at each study site, and soil moisture patterns estimated by (b) MLR, (c) SANN and (d) MM. The dots indicate training data locations.

In almost all situations, one component density provided slightly superior performance, but the MM method with one component is very similar to linear regression. The results for 2 components are shown here to highlight the method's ability as a potentially nonlinear method. All of the methods reproduce the main features observed in the patterns. The soil moisture pattern for Tarrawarra exhibits a pronounced organization with wet areas occurring in the valley bottoms and on the south-facing hillslope, and all of the methods capture this general tendency. The Satellite Station pattern is more weakly organized than Tarrawarra, but the estimation methods capture some of this organization. The Cache la Poudre site has an aspect-dependent pattern, which is the main feature that is reproduced by the estimation methods. The NSCE

values included in the figure are calculated only from the associated testing dataset. For all three sites, the SANN has the highest NSCE value although all methods perform very similarly for Tarrawarra. The NSCE values are all significantly lower for the other two sites than for Tarrawarra, but the SANN remains the best-performing method. The MM is the second-best method for Satellite Station, while MLR is the second-best method for Cache la Poudre. The SANN method likely performs better than the other methods because it has the most flexibility in the type of relationship that it can infer from the data. Such flexibility would allow the SANN to include subtleties in the relationships to the topographic attributes that are ignored by the other two methods.

### **Number of predictor variables**

Figure 3 plots the median NSCE for the three methods as the number of topographic attributes used varies. All three wetness conditions are shown for the three applications sites. In all cases, 25% of the observations were used as training data, and the lines in the figure indicate the median NSCE values from all 30 samples at that sampling rate (calculated from only the testing locations). For each subset size, the NSCE shown is from the attribute set with the highest median value. For the SANN and MM methods, the parameter values are those that produce the best possible performance in each case. No condition is used to determine whether the additional predictor variables explain a statistically significant portion of the variance in the soil moisture observations for any of the methods. In nearly all scenarios in Figure 3, the performance of the three methods is similar. The broadest range of performance between methods is observed at Satellite Station for the dry and medium dates. In both of those cases, the SANN has the best performance. The SANN also performs the best in most other cases in the figure, and it typically performs the best for other sampling rates (not shown).

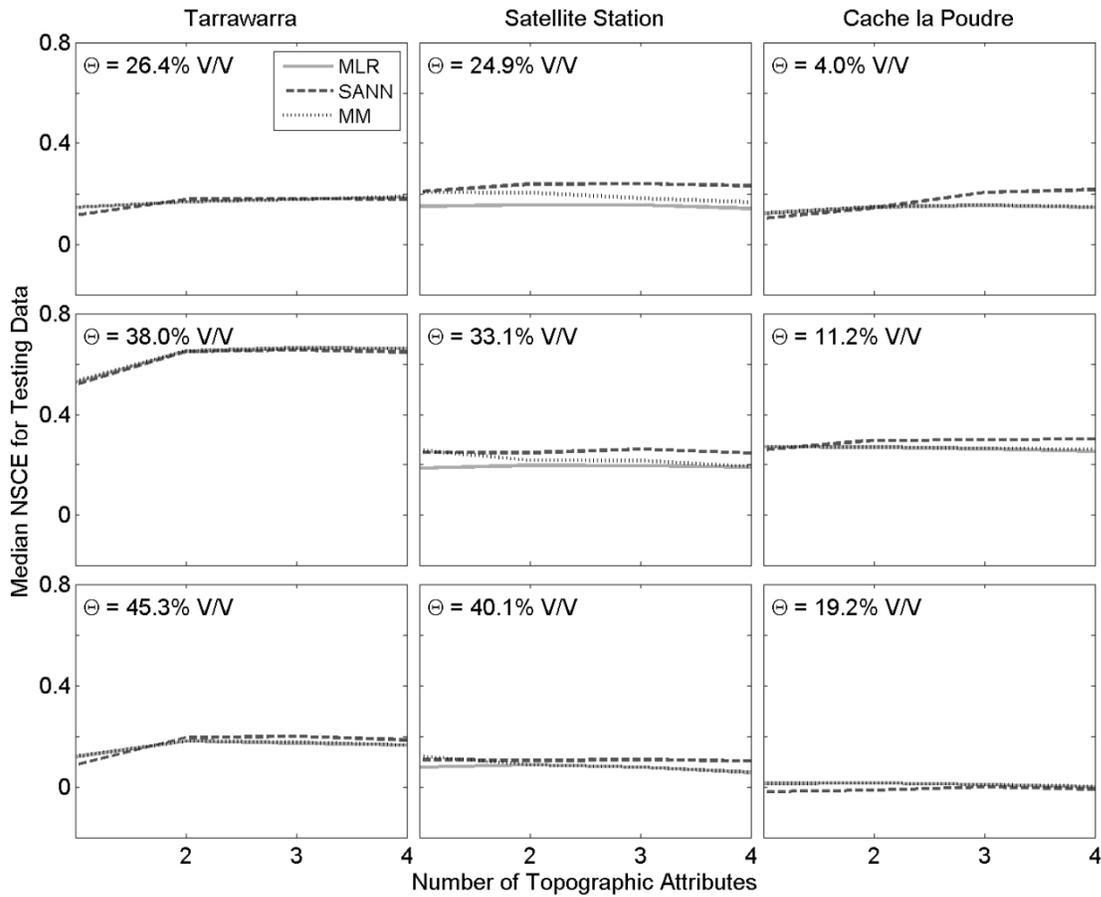


Figure 3 - Median NSCE for each estimation method as the number of topographic attributes used increases.

All of the methods perform best for the moderate wetness condition at the Tarrawarra and Cache la Poudre sites, but at the Satellite Station site the methods perform similarly for the dry and moderate conditions and perform worst for the wettest condition. Reduced performance in dry and wet conditions is consistent with reduced spatial structure in the soil moisture patterns and greater importance of local controls (e.g. porosity) relative to topography under those conditions [Grayson *et al.*, 1997; Western *et al.*, 1999a]. The similarity of the results of the MLR, MM, and SANN indicates that allowing nonlinearity in the relationships between soil moisture and topographic attributes does not substantially improve pattern estimation. This result likely implies that the relationships (where meaningful) are close to linear. For most scenarios in the

figure, the methods show little variation in NSCE as predictor variables are added. The exceptions are the moderate and wet dates at Tarrawarra as well as the dry date at Cache la Poudre. For those two dates at Tarrawarra, the NSCE values increase substantially with the use of a second predictor variable and then remain roughly constant. For the dry date at Cache la Poudre, the SANN method improves consistently with the addition of more predictor variables while the other two methods show little change. Overall, the NSCE for the SANN usually remains roughly constant or improves with additional predictor variables, while the other two methods generally decrease slightly with additional predictor variables. The lack of substantial improvement with additional variables suggests that one or two attributes represent the major effects of the dominant physical processes. Where they occur, negative trends in NSCE with additional variables indicate overtraining, which leads to reduced generalization capability. When higher sampling rates are used (not shown), the NSCE values are less likely to decrease as the number of predictor variables increases. That result is expected because higher sampling rates likely include more information that could justify the inclusion of additional attributes.

### **Sample size**

The effects of the size of the training dataset on the performance of the methods are presented in Figure 4. The figure shows the results for the best set of two predictor variables for each of the methods, where the best set is defined as the set that produces the maximum value for the median NSCE. The rows in Figure 4 show the results for a given method while the columns show the results for a given study site. Only the moderate wetness condition is shown for each site, but the results are consistent for all three conditions. The box-and-whisker plots in the figure characterize the variation in the performance of each method among the 30 training sets generated at a given sampling density. The upper and lower limits of the boxes represent the 25<sup>th</sup> and 75<sup>th</sup> percentiles, respectively, for the NSCE and the horizontal line within the box represents

the median value. Whiskers represent the limits of observed NSCE values that are not considered outliers while outliers are marked by plus signs.

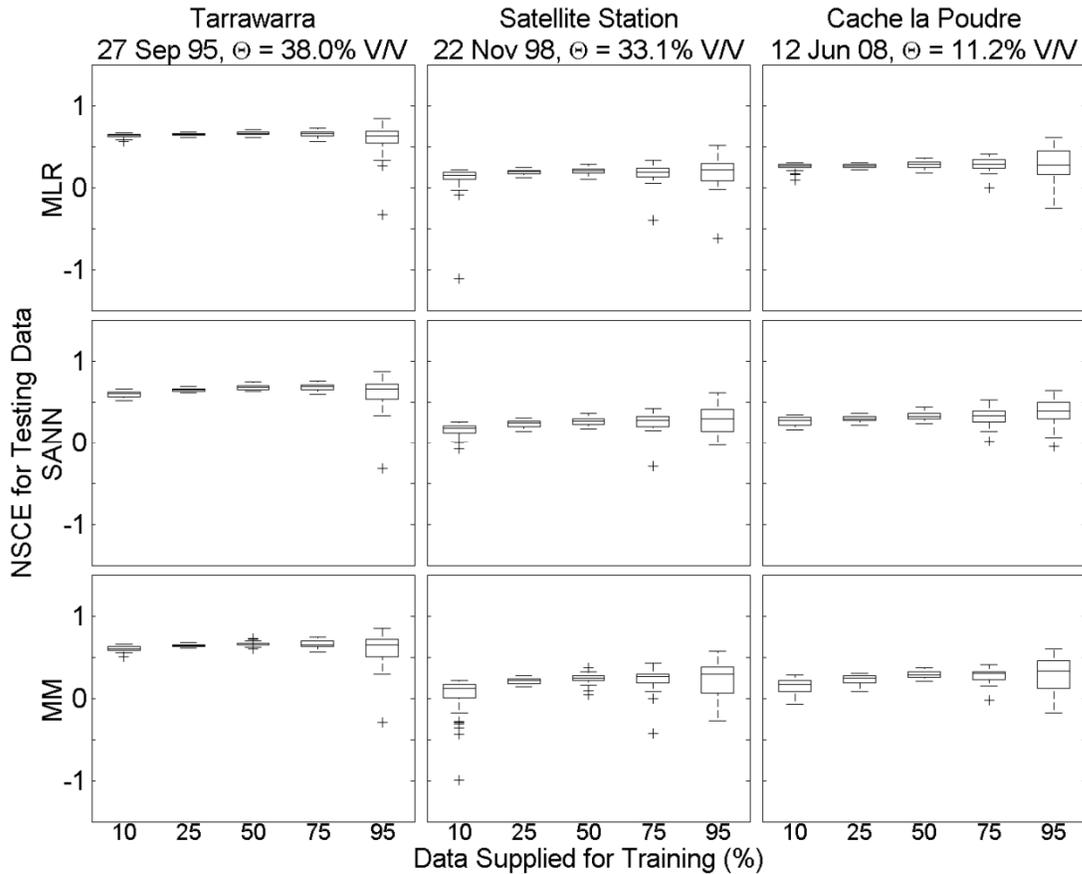


Figure 4 - Box and whisker plots characterizing the performance of each estimation method as the amount of training data used increases. Plus symbols indicate outliers, which are defined as values that are more than 1.5 times the distance between the upper and lower quartiles away from the box limits.

For all the methods, the performance does not increase much as more observations are supplied for training, and in some cases performance can decline with increasing data due to overtraining.

The highest median NSCE values for the SANN and MM that are achieved as one varies the sampling rate are higher than for MLR (for all study sites). As a specific example, at the Cache la Poudre site the highest median NSCE for MLR is 0.29, which occurs at the 75% sampling rate.

For the SANN and MM, the values are 0.39 and 0.33, respectively, and they both occur at the

95% sampling rate. However, the sampling rates for which the median NSCE reaches its highest value with the MLR method for the Tarrawarra, Satellite Station and Cache la Poudre sites are 50%, 95%, and 75%, respectively, while the rates for the SANN are 75%, 95%, and 95%, and for MM they are 50%, 95%, and 95%. Overall, the nonlinear methods' best performances were obtained by using more data than the MLR, reflecting their ability to continue extracting useful information as the amount of data increases. However the highest median NSCE occurs for all methods and sites with at least 50% of the data being used to train the estimator, which represents a significant data collection effort. The numbers of measurements represented by the 10% and 25% sampling rates are more feasible for regular manual collection. At those sampling rates the SANN has the highest median NSCE values at the Satellite Station and Cache la Poudre sites while MLR has the highest median values at Tarrawarra. Overall, the SANN achieves equivalent or higher NSCE values than MLR when the sampling rate is at least 25%, and the MM also outperforms MLR in most cases when the sampling rate is at least 50%.

### **Predictor sets chosen by methods**

The SANN and MM can model nonlinear relationships between topographic attributes and soil moisture, so it is possible that they perform best when using different topographic attributes than the MLR uses. To investigate this possibility, we analyzed the frequency with which each attribute is included in models with relatively high NSCE values for each estimation method. Specifically, the top 10% of the models in terms of NSCE were identified for each method when 25% of the observations were used as training data, and the fraction of those models that contains each topographic attribute was calculated. For example, for models with 2 predictor variables, there are 15 combinations of predictor variables, and 30 samples were tested for each combination, for a total of 450 results. From those results, the 45 cases with the highest NSCE

values were analyzed. In Figure 5, the height of the bar associated with each attribute represents the proportion of those cases that contain that attribute.

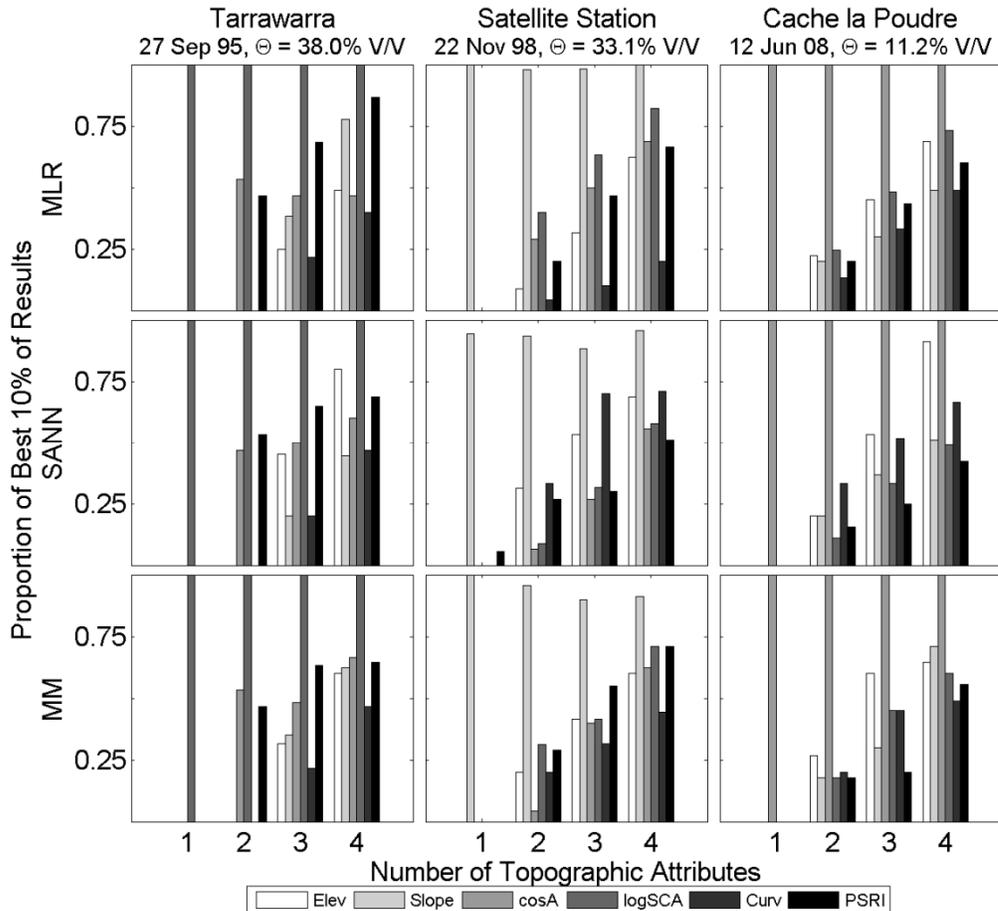


Figure 5 - Proportion of times each topographic attribute is selected by the best-performing models as a function of the number of topographic attributes used.

For each study site, all three methods perform best using the same attribute in the one-variable case although the attribute is different for each site. That single attribute remains the most frequently used as the number of variables increases. The best-performing single attribute is logSCA for Tarrawarra, slope for Satellite Station, and cosA for Cache la Poudre. The logSCA is related to the convergence of flow due to topography and suggests the importance of water redistribution at Tarrawarra. The Satellite Station topography is generally divided between highland areas with high slopes and lowland areas with low slopes, and the soil moisture values

reflect that division as well. The preference for cosA at Cache la Poudre site reflects the distinct vegetation on the two opposing hillslopes at this site. For the two-predictor variable scenario at Tarrawarra, the second selected variable is split approximately equally between cosA and PSRI for all methods, indicating the importance of radiation-driven evapotranspiration variations at this site. There is not a dominant secondary variable chosen by any of the methods in the two-predictor variable scenario at the Satellite Station or Cache la Poudre sites. Beyond the two-variable case, no consistent preferences for additional attributes are observed at any site.

This analysis was repeated for the dry and wet dates at each site. For the dry date (not shown), the most significant difference from the moderate wetness date is that the models now perform best when using PSRI at Tarrawarra, which indicates greater importance of evapotranspiration compared with flow convergence. Additionally, elevation is a clear secondary attribute selected at Satellite Station by the SANN and MM but not the MLR. For the wet date (not shown), the only noteworthy difference from the moderate wetness scenario is the lack of a clearly preferred attribute at the Cache la Poudre site for the SANN. However, the MLR and MM methods preferentially selected logSCA as the first predictor. Overall, the methods tend to select the same attributes for soil moisture estimation at each catchment. Therefore, performance differences between the methods are due to differences in the forms of the modeled relationships between the attributes and soil moisture rather than due to the selection of different attributes. We also implemented each method in a stagewise manner using the same topographic attributes and sampling scheme. In this implementation, the best single predictor variable was chosen first, and its estimate of soil moisture was retained. Then, this predictor variable was removed from subsequent consideration, and the remaining predictor variables were evaluated in their ability to explain the residuals. This process was repeated until all predictor variables were used. We

found that the stagewise method does not offer any improvement in performance over the original implementation.

## Conclusions

1. The SANN method consistently outperforms the MM method and MLR and is the best overall method tested. For the majority of locations, wetness conditions, and levels of training data, it provides higher NSCE values than the other two methods. In most scenarios, the increase in performance is not large, but the improvement is consistent, and the method never performs much worse than the other methods. Another positive aspect of the SANN method is its superior performance when using multiple predictor variables. All the methods tend to perform the best at a given site when using the same topographic attribute as a single predictor variable, but a different attribute is best for each of the sites. However, because SANN performs better than the other methods when using multiple predictor variables, the *a priori* selection of one or two attributes for soil moisture estimation would not be required for that method. Thus, one could use a single, larger set of topographic attributes at a variety of sites. MLR may also be able to be used in a similar fashion if a suitable test for statistical significance is evaluated before a predictor variable is added to the model.
2. The estimation accuracy of all of the methods depends on both the catchment and wetness condition for which estimates are made. All methods perform best at Tarrawarra under the moderate wetness condition and perform worst at Cache la Poudre for the wet condition. Thus, not only do individual site characteristics determine how well the methods estimate soil moisture from topographic attributes, but those characteristics also affect the wetness condition under which the methods perform best.

3. In order for the SANN and MM to consistently perform better than MLR, relatively large training datasets are required. More training data leads to higher NSCE values for most scenarios tested. For all scenarios except the wet condition at Cache la Poudre, SANN performs equivalently or better than MLR for sampling rates greater than or equal to 25%, and for most scenarios, MM outperforms MLR at the 50% sample rate and above. The superior performance of the SANN and MM methods suggest that the relationships between soil moisture and topographic attributes might be weakly nonlinear, but more data are required for the estimation methods to discern any nonlinearity and for the improved performance of these methods to be realized.

Several directions are open for future research. These include testing the abilities of the methods to estimate soil moisture under circumstances (e.g., wetness conditions, locations, or DEM resolutions) different than those for which the methods were trained. In addition, other statistical learning techniques such as support vector machines and genetic algorithms could be evaluated for estimating soil moisture patterns.

## **Acknowledgments**

The authors acknowledge the Army Research Office and the Center for Geosciences and Atmospheric Research for their financial support. The authors thank all parties involved in MARVEX, including researchers from NIWA and the University of Melbourne, as well as Tim Green from the Agricultural Research Service for helpful suggestions. The authors also thank Brandon Lehman, Josh Melliger, Jonathan Freed, and Rob Erskine for collecting the Cache la Poudre data, and the anonymous reviewers for their suggestions for improving this paper.

## References

- Ahmad, S., A. Kalra, and H. Stephen (2010), Estimating soil moisture using remote sensing data: A machine learning approach, *Adv. Water Resour.*, 33(1), 69-80, doi: 10.1016/j.advwatres.2009.10.008.
- Bardossy, A., and W. Lehmann (1998), Spatial distribution of soil moisture in a small catchment. Part 1: geostatistical analysis, *J. Hydrol.*, 206, 1-15.
- Beven, K. J., and M. J. Kirkby (1979), A physically-based, variable contributing area model of basin hydrology, *Hydrol. Sci. Bull.*, 24(1), 43-69.
- Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, 1st ed., 482 pp., Oxford University Press, New York.
- Bishop, C. M. (2006), *Pattern Recognition and Machine Learning*, 1st ed., 738 pp., Springer, New York.
- Chiles, J.-P., and P. Delfiner (1999), *Geostatistics: Modeling Spatial Uncertainty*, 695 pp., John Wiley and Sons, New York, NY.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), Maximum likelihood from incomplete data via EM algorithm, *J. R. Stat. Soc. Ser. B-Methodol.*, 39(1), 1-38.
- Dingman, S. L. (2002), *Physical Hydrology*, 2nd ed., 646 pp., Prentice Hall, Upper Saddle River, New Jersey.
- Draper, N. R., and H. Smith (1981), *Applied Regression Analysis*, 2nd ed., 709 pp., John Wiley & Sons, New York.
- Dunne, T., and R. D. Black (1970), Partial area contributions to storm runoff in a small New England watershed, *Water Resour. Res.*, 6(5), 1296-&.
- Dunne, T., T. R. Moore, and C. H. Taylor (1975), Recognition and prediction of runoff-producing zones in humid regions, *Hydrol. Sci. Bull.*, 20, 305-327.
- Eagleson, P. S. (1978), Climate, Soil and Vegetation .3. Simplified Model of Soil-Moisture Movement in Liquid-Phase, *Water Resour. Res.*, 14(5), 722-730.
- Elshorbagy, A., and I. El-Baroudy (2009), Investigating the capabilities of evolutionary data-driven techniques using the challenging estimation of soil moisture content, *Journal of Hydroinformatics*, 11(3-4), 237-251, doi: 10.2166/hydro.2009.032.
- Entekhabi, D., I. Rodriguez-Iturbe, and F. Castelli (1996), Mutual interaction of soil moisture state and atmospheric processes, *J. Hydrol.*, 184, 3-17.

- Gill, M. K., T. Asefa, M. W. Kemblowski, and M. McKee (2006), Soil moisture prediction using support vector machines, *Journal of the American Water Resources Association*, 42(4), 1033-1046.
- Grayson, R. B., A. W. Western, F. H. S. Chiew, and G. Bloschl (1997), Preferred states in spatial soil moisture patterns: Local and nonlocal controls, *Water Resour. Res.*, 33(12), 2897-2908.
- Green, T. R., and R. H. Erskine (2004), Measurement, scaling, and topographic analyses of spatial crop yield and soil water content, *Hydrol. Processes*, 18, 1447-1465.
- Green, T. R., J. D. Salas, A. Martinez, and R. H. Erskine (2007), Relating crop yield to topographic attributes using Spatial Analysis Neural Networks and regression, *Geoderma*, 139, 23-37.
- Hewlett, J. D., and A. R. Hibbert (1967), Factors affecting the response of small watersheds to precipitation in humid areas, in *Forest Hydrology: Proceedings of a National Science Foundation Advanced Science Seminar Held at the Pennsylvania State University*, edited by W. E. Sopper and H. W. Lull, p. 813, Pergamon Press, New York.
- Hillel, D. (1998), *Environmental Soil Physics*, 771 pp., Academic Press, New York.
- Hirsch, R. M., D. R. Helsel, T. A. Cohn, and E. J. Gilroy (1993), Statistical analysis of hydrologic data, in *Handbook of Hydrology*, edited by D. R. Maidment, pp. 17.11-17.55, McGraw-Hill, New York.
- Journel, A. G., and C. J. Huijbregts (1978), *Mining Geostatistics*, 600 pp., Academic Press, San Francisco, CA.
- Kitanidis, P. K. (1993), Geostatistics, in *Handbook of Hydrology*, edited by D. R. Maidment, p. 1424, McGraw-Hill, Inc., New York.
- Lehman, B. M., and J. D. Niemann (2008), Spatial Patterns of in a Semi-Arid Montane Catchment with Aspect-Dependent Vegetation, paper presented at First International Conference on Hydrogeology, State College, Pennsylvania.
- Lin, J. Y., C. T. Cheng, and K. W. Chau (2006), Using support vector machines for long-term discharge prediction, *Hydrol. Sci. J.-J. Sci. Hydrol.*, 51(4), 599-612.
- Liu, H. B., D. Xie, and W. Wu (2008), Soil water content forecasting by ANN and SVM hybrid architecture, *Environmental Monitoring And Assessment*, 143(1-3), 187-193, doi: 10.1007/s10661-007-9967-9.
- Martinez, A., J. D. Salas, and T. R. Green (2004), Sensitivity of spatial analysis neural network training and interpolation to structural parameters, *Math. Geol.*, 36(6), 721-742.
- Mathworks (2009), Matlab, edited, The Mathworks, Natick, MA.

- McLachlan, G., and D. Peel (2000), *Finite Mixture Models*, 419 pp., John Wiley & Sons, Inc., New York.
- Mitasova, H., and J. Hofierka (1993), Interpolation by regularized spline with tension: II. Application to terrain modeling and surface geometry analysis, *Math. Geol.*, 25(6), 657-669.
- Moore, I. D., G. J. Burch, and D. H. Mackenzie (1988), Topographic effects on the distribution of surface soil water and the location of ephemeral gullies, *Trans. Am. Soc. Ag. Engr.*
- Moore, I. D., R. B. Grayson, and A. R. Ladson (1991), Digital terrain modeling - a review of hydrological, geomorphological, and biological applications, *Hydrol. Processes*, 5(1), 3-30.
- Nadaraya, E. A. (1964), On estimating regression, *Theory Prob. and Apps.*, 9(1), 141-142.
- Nash, J. E., and J. V. Sutcliffe (1970), River forecasting through conceptual models: Part I, a discussion of principles, *J. Hydrol.*, 10, 282-290.
- Nyberg, L. (1996), Spatial variability of soil water content in the covered catchment at Gardsjon, Sweden, *Hydrol. Processes*, 10, 89-103.
- Perry, M. A., and J. D. Niemann (2007), Analysis and estimation of soil moisture at the catchment scale using EOFs, *J. Hydrol.*, 334(3-4), 388-404, doi: 10.1016/j.jhydrol.2006.10.014.
- Perry, M. A., and J. D. Niemann (2008), Generation of soil moisture patterns at the catchment scale by EOF interpolation, *Hydrol. Earth Syst. Sci.*, 12(1), 39-53.
- Robinson, D. A., S. B. Jones, J. M. Wraith, D. Or, and S. P. Friedman (2003), A Review of Advances in Dielectric and Electrical Conductivity Measurement in Soils Using Time Domain Reflectometry, *Vadose Zone Journal*, 2(4), 444-475.
- Rodriguez-Iturbe, I., G. K. Vogel, R. Rigon, D. Entekhabi, F. Castelli, and A. Rinaldo (1995), On the spatial organization of soil moisture fields, *Geophys. Res. Lett.*, 22(20), 2757-2760.
- Rodriguez-Iturbe, I. (2000), Ecohydrology: A hydrologic perspective of climate-soil-vegetation dynamics, *Water Resour. Res.*, 36(1), 3-9.
- Shin, H.-S., and J. D. Salas (2000a), Spatial analysis of hydrologic and environmental data based on artificial neural networks, in *Artificial neural networks in hydrology*, edited by R. S. Govindaraju and A. R. Rao, pp. 259-286, Kluwer Academic Publishers, Boston.
- Shin, H.-S., and J. D. Salas (2000b), Regional drought analysis based on neural networks, *J. Hydrol. Engr.*, 5(2), 11.
- Specht, D. F. (1991), A general regression neural network, *IEEE Trans. Neural Net.*, 2(6), 568-576.
- Sulebak, J. R., L. M. Tallaksen, and B. Erichsen (2000), Estimation of areal soil moisture by use of terrain data, *Geog. Ann. Ser. A Phys. Geog.*, 82A(1), 89-105.

- Tarboton, D. G. (1997), A new method for the determination of flow directions and upslope areas in grid digital elevation models, *Water Resour. Res.*, 33(2), 309-319.
- Tarboton, D. G. (2008), TauDEM, edited, Utah State University, Logan, UT.
- Vereecken, H., J. A. Huisman, H. Bogaert, J. Vanderborght, J. A. Vrugt, and J. W. Hopmans (2008), On the value of soil moisture measurements in vadose zone hydrology: A review, *Water Resour. Res.*, 44, doi: W00d06  
10.1029/2008wr006829.
- Wang, W. C., K. W. Chau, C. T. Cheng, and L. Qiu (2009), A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series, *J. Hydrol.*, 374(3-4), 294-306, doi: 10.1016/j.jhydrol.2009.06.019.
- Watson, G. S. (1964), Smooth regression analysis, *Sankhya Indian J. Stat. A*, 26, 359-372.
- Western, A. W., and R. B. Grayson (1998), The Tarrawarra data set: Soil moisture patterns, soil characteristics, and hydrological flux measurements, *Water Resour. Res.*, 34(10), 2765-2768.
- Western, A. W., and G. Bloschl (1999), On the spatial scaling of soil moisture, *J. Hydrol.*, 217(3-4), 203-224.
- Western, A. W., R. B. Grayson, G. Bloschl, G. R. Willgoose, and T. A. McMahon (1999a), Observed spatial organization of soil moisture and its relation to terrain indices, *Water Resour. Res.*, 35(3), 797-810.
- Western, A. W., R. B. Grayson, and T. R. Green (1999b), The Tarrawarra project: high resolution spatial measurement, modelling and analysis of soil moisture and hydrological response, *Hydrol. Processes*, 13, 633-652.
- Western, A. W., G. Bloschl, and R. B. Grayson (2001), Toward capturing hydrologically significant connectivity in spatial patterns, *Water Resour. Res.*, 37(1), 83-97.
- Wilson, D. J., A. W. Western, R. B. Grayson, A. A. Berg, M. S. Lear, M. Rodell, J. S. Famiglietti, R. A. Woods, and T. A. McMahon (2003), Spatial distribution of soil moisture over 6 and 30 cm depth, Mahurangi river catchment, New Zealand, *J. Hydrol.*, 276(1-4), 254-274.
- Wilson, D. J., A. W. Western, and R. B. Grayson (2005), A terrain and data-based method for generating the spatial distribution of soil moisture, *Adv. Water Resour.*, 28(1), 43-54.
- Woods, R. A., R. B. Grayson, A. W. Western, M. Duncan, D. J. Wilson, R. I. Young, R. Ibbitt, R. Henderson, and T. A. McMahon (2001), Experimental design and initial results from the Mahurangi River Variability Experiment: MARVEX, in *Land Surface Hydrology, Meteorology, and Climate: Observations and Modeling, Water Science and Application*, edited by V. Lakshmi, et al., pp. 201-213, American Geophysical Union, Washington, D.C.
- Yates, S. R., and A. W. Warrick (1987), Estimating soil water content using cokriging, *Soil Sci. Soc. Am. J.*, 51(1), 23-30.

Zaslavsky, D., and G. Sinai (1981), Surface Hydrology .1. Explanation of Phenomena, *J. Hydr. Div. ASCE*, 107(1), 1-16.

Zhu, H., and A. G. Journel (1993), Formatting and integrating soft data: Stochastic imaging via the Markov-Bayes algorithm, in *Geostatistics Troia '92*, edited by A. Soares, pp. 1-12, Kluwer, Dordrecht.

## **Part II: Controls on Topographic Dependence and Time-Instability in Catchment-Scale Soil Moisture Patterns**

## Notation

$A$  – upslope contributing area

$c$  – contour length

$E$  – evapotranspiration rate

$E_p$  – potential evapotranspiration rate

$F$  – infiltration rate

$G$  – drainage rate to saturated zone

$I_p$  – potential solar radiation index

$K_{s,h}$  – horizontal saturated hydraulic conductivity

$K_{s,v}$  – vertical saturated hydraulic conductivity

$L$  – lateral unsaturated flow rate

$S$  – topographic slope

$w_G$  – weight applied to drainage to saturated zone pattern

$w_L$  – weight applied to lateral flow moisture pattern

$w_R$  – weight applied to radiative evapotranspiration pattern

$w_A$  – weight applied to aerodynamic evapotranspiration pattern

$\alpha$  – proportion of aerodynamic evapotranspiration relative to radiative evapotranspiration

$\beta_a$  – exponent describing aerodynamic evapotranspiration moisture limitation

$\beta_r$  – exponent describing radiative evapotranspiration moisture limitation

$\eta_h$  – horizontal pore disconnectedness index

$\eta_v$  – vertical pore disconnectedness index

$\delta_0$  – soil depth for location with no curvature, mm

$\varepsilon$  – exponent describing relationship between topographic slope and horizontal hydraulic gradient

$\theta$  – soil moisture content at a location

$\bar{\theta}$  – spatial average soil moisture

$\theta_G$  – soil moisture pattern when drainage to saturated zone is dominant process

$\theta_L$  – soil moisture pattern when lateral flow is dominant process

$\theta_R$  – soil moisture pattern when radiative evapotranspiration is dominant process

$\theta_A$  – soil moisture pattern when aerodynamic evapotranspiration is dominant process

$\kappa$  – topographic curvature

$\kappa_{\min}$  – minimum curvature for which soil is present

$\Lambda$  – spatial average of lateral flow index

$\Pi$  – spatial average of evapotranspiration index

$\phi$  – soil porosity

## Introduction

The land surface wetness state, as represented by volumetric soil water content (soil moisture), plays a large role in determining the effects of hydrologic processes. For example, soil moisture affects both the partitioning of precipitation into infiltration and runoff [Dunne and Black, 1970] as well as the partitioning of radiation into evapotranspiration (ET) and latent heat flux [Entekhabi *et al.*, 1996]. Soil moisture also influences vegetation patterns [Eagleson, 1978], erosion processes [Moore *et al.*, 1988], and soil development [Hillel, 1998]. In addition, the nonlinearity of many hydrologic processes with respect to soil moisture means that spatial patterns of soil moisture influence the spatially-averaged or spatially-integrated effects of the processes [Dunne and Black, 1970; Dunne *et al.*, 1975; Hewlett and Hibbert, 1967; Western *et al.*, 2001].

Catchment-scale soil moisture values have been found to be correlated with a variety of topographic, soil, and vegetation characteristics [Lin, 2006; Niemann and Edgell, 1993; Western *et al.*, 1999a]. Because topographic information in the form of digital elevation models (DEMs) is readily available, relationships between topographic attributes and soil moisture are particularly desirable for estimating soil moisture. Soil moisture exhibits at least two pattern types related to topography. The first type of pattern is organized similarly to the catchment drainage pattern with wetter locations occurring in the valley bottoms and will be called a valley-dependent pattern in this paper. Valley dependence can be represented quantitatively by topographic attributes such as the contributing area or the wetness index [Beven and Kirkby, 1979; Moore *et al.*, 1988]. Those attributes have exhibited reasonably strong correlations with soil moisture at some times and locations [Burt and Butcher, 1985; Nyberg, 1996; Perry and Niemann, 2007; Western *et al.*, 1999a] but not at others [Crave and Gascuel-Oudou, 1997; Ladson and Moore, 1992]. The second pattern type has an organization that is related to the

hillslope orientations with wetter locations occurring on the slopes that are more shaded from the sun and will be called a hillslope-dependent pattern. Hillslope dependence can be represented quantitatively by aspect-related attributes, including the cosine of the aspect and potential solar radiation index (PSRI) [Dingman, 2002]. These attributes have also been found to be correlated with soil moisture values in some cases [Famiglietti et al., 1998; Gomez-Plaza et al., 2001]. In some instances, topographic attributes related to both the valley-dependent and hillslope-dependent patterns have been found to improve soil moisture estimates [Gomez-Plaza et al., 2001; Western et al., 1999a; Wilson et al., 2005]. It remains unclear why different catchments exhibit different types of topographically-dependent soil moisture patterns.

The dependence of soil moisture on topography or other characteristics at a catchment can also change through time [Western et al., 1999a; Wilson et al., 2005]. Such behavior can lead to changes in the spatial structure of the soil moisture pattern, which is known as time instability. Vachaud et al. [1985] introduced the concept of time stability and its analysis as a means to reduce the number of observations required to characterize the behavior of spatially-averaged soil moisture. Several studies in various catchments have since used temporal stability analyses to confirm the presence of time-stable locations [Brocca et al., 2009; Cosh et al., 2006; Grayson and Western, 1998; Martinez-Fernandez and Ceballos, 2005]. The physical origins of time stability in soil moisture patterns are not as widely studied. Gomez-Plaza et al. [2000] studied soil moisture along three transects in a semi-arid climate and found greater time instability on the transect consisting of two hillslopes with different aspects. They attributed the increased instability to greater heterogeneity of that transect's characteristics relative to the other transects' characteristics. The heterogeneity included not only the different aspects but also different vegetation densities. On a larger scale, Mohanty and Skaggs [2001] analyzed pixels from the

Southern Great Plains 1997 (SGP97) experiment and found that the footprint with sandy loam soil was more time-stable than the footprints with silty loam soil. Additionally, they found that flat topography and variable land cover were associated with greater time instability. Time stability has been observed at several study sites where the spatial variations in soil moisture are correlated with soil properties [*da Silva et al.*, 2001; *Grant et al.*, 2004; *Lin*, 2006; *Vachaud et al.*, 1985], but it has also been observed at sites where the patterns depend on topography [*Kachanoski and DeJong*, 1988]. At the Tarrawarra catchment, where topographic attributes exhibit a strong influence on soil moisture variations, time instability has been documented [*Western et al.*, 1999a]. *Grayson et al.* [1997] suggested that the tendency of the soil moisture patterns at Tarrawarra to change spatial organizations through time is related to the relative importance of vertical and horizontal fluxes, which depend on the relative amounts of precipitation and potential ET (PET) in different seasons. Overall, however, the catchment characteristics that promote time instability remain poorly understood.

The objectives of this research are to determine: (1) the local soil, vegetation, and climatic characteristics that most influence the type of topographic-dependence that occurs in the soil moisture patterns at a catchment and (2) the factors affecting the temporal stability of the soil moisture patterns at a catchment. We begin by deriving a conceptual model, the Equilibrium Moisture from Topography (EMT) model, for estimating soil moisture patterns from physically-based considerations of the processes affecting soil moisture. The processes are represented in the model as functions of topographic attributes. All other characteristics are assumed to be homogeneous within a given catchment. The model uses the spatial-average soil moisture as input, which allows for different types of topographic dependence to emerge as the overall wetness changes. Three catchments are used to test the model: Tarrawarra in Southeastern

Australia, Satellite Station in the Maharungi catchment on the North Island of New Zealand, and Cache la Poudre in north-central Colorado. These three catchments have very different climatic and terrain conditions as well as soil moisture patterns with a range of topographic dependences and time–instability characteristics. The type of topographic dependence and the degree of time instability produced by the model for these catchments are quantified using two metrics. The responses of those metrics to different values of model parameters are used to assess the influence of catchment characteristics on the moisture pattern type and the time-instability.

## Model Development

The model is developed by considering the land area that drains through an arbitrary length of an elevation contour in the catchment (or equivalently the downslope edge of a DEM grid cell). In particular, the water balance is evaluated for the vadose zone within this area. Four processes are included to transport water to or from the vadose zone: infiltration  $F$  [ $L T^{-1}$ ], deep drainage or recharge to groundwater  $G$  [ $L T^{-1}$ ], lateral unsaturated flow  $L$  [ $L T^{-1}$ ], and ET  $E$  [ $L T^{-1}$ ]. For simplicity, soil moisture is assumed to be constant with depth within the vadose zone. Although the average soil moisture in the catchment will be allowed to vary in time, the soil moisture pattern is derived by considering equilibrium conditions. This approach assumes that the pattern is not affected by whether the catchment is getting wetter or drier and thus precludes consideration of hysteresis. Under these assumptions, one can write the water balance as:

$$\int_A F dA = \int_A G dA + L + \int_A E dA \quad (2.1)$$

where  $A$  is the land area [ $L^2$ ] that can contribute flow to the length of the elevation contour.

The deep drainage process is represented as unsaturated vertical flow of water into deeper levels of the ground. This process is assumed to take place as percolation whereby gravity is the only driving force with no diffusivity contribution. The *Campbell* [1974] equation is used to

determine the unsaturated hydraulic conductivity from the soil moisture. Under those conditions,  $G$  is equivalent to the unsaturated hydraulic conductivity or:

$$G = K_{s,v} \left( \frac{\theta}{\phi} \right)^{\gamma_v} \quad (2.2)$$

where  $K_{s,v}$  is the saturated vertical hydraulic conductivity [ $L T^{-1}$ ],  $\theta$  is the soil moisture in units of volume of water per total volume ( $V/V$ ),  $\phi$  is the soil porosity, and  $\gamma_v$  is the pore disconnectedness index [Dingman, 2002].

The lateral unsaturated flow is determined by calculating a horizontal specific discharge from Darcy's Law. In that expression, the unsaturated hydraulic conductivity is again described using the Campbell [1974] equation, and the hydraulic gradient is assumed to be related to the topographic slope [Beven and Kirkby, 1979; Temimi et al., 2010]. Then, the volumetric lateral flow  $L$  is determined by multiplying the specific discharge by the area through which it occurs. That area is the arbitrary length of the contour multiplied by the depth of the layer through which the flow is occurring. That depth is represented as a function of topographic curvature based on an expression for soil depth in Heimsath et al. [1999]. In the end, the resulting equation for  $L$  is:

$$L = \delta_0 \left( \frac{\kappa_{min} - \kappa}{\kappa_{min}} \right) c K_{s,h} \left( \frac{\theta}{\phi} \right)^{\gamma_h} S^\varepsilon \quad (2.3)$$

where  $\delta_0$  [L] is the layer thickness for a location with a curvature of zero,  $\kappa_{min}$  [ $L^{-1}$ ] is the minimum curvature for which the hydrologically active layer is present,  $\kappa$  [ $L^{-1}$ ] is the local surface curvature,  $K_{s,h}$  [ $L T^{-1}$ ] is the saturated horizontal hydraulic conductivity,  $c$  [L] is the contour length under consideration,  $S$  [ $L L^{-1}$ ] is the local topographic slope, and  $\varepsilon$  is a parameter relating the hydraulic gradient to the topographic slope. Note that curvature is defined here as positive for convergent locations and negative for divergent locations.

ET is modeled using a Priestley-Taylor approach [Priestley and Taylor, 1972], which assumes that the aerodynamic term is a specified fraction of the radiation term in the Penman equation [Penman, 1948]. In the present model, topographically-induced local variations in insolation are considered in the radiation term by including the potential solar radiation index (PSRI), which is the ratio of the insolation of the topographic surface relative to that of a horizontal surface at the same location and date. Moisture limitations are included by assuming that the actual ET is the PET multiplied by a power-function of the soil moisture. Combining all these elements, the resulting expression is:

$$E = \frac{E_p I_p}{1 + \alpha} \left( \frac{\theta}{\phi} \right)^{\beta_r} + \frac{E_p \alpha}{1 + \alpha} \left( \frac{\theta}{\phi} \right)^{\beta_a} \quad (2.4)$$

where  $E_p$  [ $L T^{-1}$ ] is the regional PET and  $I_p$  is the PSRI. The first term on the right-hand side of equation (2.4) represents radiative ET and the second term represents aerodynamic ET (saturation deficit ET).  $\alpha$  is the ratio of the aerodynamic term to the radiation term in the Penman equation (i.e., it is equivalent to the Priestley-Taylor coefficient minus one) with values usually in the range [0, 0.74] [Eichinger et al., 1996; Shuttleworth, 1993].  $\beta_r$  and  $\beta_a$  are vegetation-related parameters. In the literature, a single  $\beta$  is typically applied to both ET terms, with  $\beta > 1$  representing bare soil and  $\beta < 1$  representing vegetated soils [Lowry, 1959]. Here, distinct exponents are allowed for each term because moisture limitations might affect the energy and aerodynamic components in different ways. The composition of equation (2.4) also implicitly neglects ground heat flux.

Substituting the expressions for  $G$ ,  $L$ , and  $E$  from equations (2.2), (2.3), and (2.4) into equation (2.1) and assuming for the purposes of this analysis that the values for  $F$ ,  $\phi$ ,  $K_{s,v}$ ,  $K_{s,h}$ ,  $\gamma_v$ ,  $\gamma_h$ ,  $\delta_0$ ,

$\kappa_{min}$ ,  $\varepsilon$ ,  $\alpha$ ,  $\beta_r$ ,  $\beta_a$ , and  $E_p$  are all spatially constant (and therefore represent catchment-wide effective values) produces:

$$F = \frac{K_{s,v}}{\phi^{\gamma_v} A} \int_A \theta^{\gamma_v} dA + \frac{\delta_0 (\kappa_{min} - \kappa) K_{s,h}}{\kappa_{min} \phi^{\gamma_h} A} c S^\varepsilon \theta^{\gamma_h} + \frac{E_p}{(1 + \alpha) \phi^{\beta_r} A} \int_A I_p \theta^{\beta_r} dA + \frac{E_p \alpha}{(1 + \alpha) \phi^{\beta_a} A} \int_A \theta^{\beta_a} dA \quad (2.5)$$

In this expression, the integrals divided by the upslope area represent the averages of the integrands over the upslope area. Assuming for simplicity that those average values can be approximated by the respective local values, equation (2.5) can be written as:

$$F = \frac{K_{s,v}}{\phi^{\gamma_v}} \theta^{\gamma_v} + \frac{\delta_0 (\kappa_{min} - \kappa) K_{s,h}}{\kappa_{min} \phi^{\gamma_h} A} c S^\varepsilon \theta^{\gamma_h} + \frac{E_p}{(1 + \alpha) \phi^{\beta_r}} I_p \theta^{\beta_r} + \frac{E_p \alpha}{(1 + \alpha) \phi^{\beta_a}} \theta^{\beta_a} \quad (2.6)$$

If the value of  $F$  is known, equation (2.6) can be solved for  $\theta$  at each location within the catchment using the topographic attributes and specified soil, vegetation, and climate parameters. However, the equation must be solved iteratively, which makes it inconvenient for estimating the soil moisture pattern if a large number of simulations are required (as is the case later in this work).

To obtain an approximate, explicit solution for  $\theta$ , the following procedure is used. Each term on the right-hand side of equation (2.6) is assumed in turn to be much larger than the other terms. If only a single term is retained in the equation, then an analytical solution can be found. A final estimate of  $\theta$  is then determined as a weighted average of the four analytical solutions as follows:

$$\theta = \frac{w_G \theta_G + w_L \theta_L + w_R \theta_R + w_A \theta_A}{w_G + w_L + w_R + w_A} \quad (2.7)$$

where  $\theta_G$ ,  $\theta_L$ ,  $\theta_R$ , and  $\theta_A$  are the estimates when deep drainage, lateral flow, radiative ET, and aerodynamic ET are assumed to dominate, respectively, and  $w_G$ ,  $w_L$ ,  $w_R$ ,  $w_A$  determine the importance of each estimate in the final calculation of soil moisture. The impact of this approach on the simulated soil moisture patterns will be evaluated later in the paper.

Consider first the case where lateral drainage is the dominant outflow process. If only this term is retained, equation (2.6) can be solved for  $\theta_L$ :

$$\theta_L = \phi \left( \frac{F}{\delta_0 K_{s,h}} \right)^{1/\gamma_h} \left( \frac{A}{cS^\varepsilon} \right)^{1/\gamma_h} \left( \frac{\kappa_{min}}{\kappa_{min} - \kappa} \right)^{1/\gamma_h} \quad (2.8)$$

Integrating equation (2.8) over the catchment, one obtains:

$$\bar{\theta} = \frac{1}{A_c} \int_{A_c} \theta_L dA_c = \phi \left( \frac{F}{\delta_0 K_{s,h}} \right)^{1/\gamma_h} \frac{1}{A_c} \int_{A_c} \left( \frac{A}{cS^\varepsilon} \right)^{1/\gamma_h} \left( \frac{\kappa_{min}}{\kappa_{min} - \kappa} \right)^{1/\gamma_h} dA_c \quad (2.9)$$

where  $\bar{\theta}$  is the average soil moisture in the catchment and  $A_c$  is the catchment area. The term that remains in the integral,  $(A/cS^\varepsilon)^{1/\gamma_h} [\kappa_{min}/(\kappa_{min} - \kappa)]^{1/\gamma_h}$ , is a compound topographic index, which we call the lateral flow index (LFI). Denoting the average LFI within the catchment by  $\Lambda$  and substituting it into equation (2.9), one obtains:

$$\bar{\theta} = \phi \left( \frac{F}{\delta_0 K_{s,h}} \right)^{1/\gamma_h} \Lambda \quad (2.10)$$

Combining equations (2.8) and (2.10), one obtains:

$$\theta_L = \frac{\bar{\theta}}{\Lambda} \left( \frac{A}{cS^\varepsilon} \right)^{1/\gamma_h} \left( \frac{\kappa_{min}}{\kappa_{min} - \kappa} \right)^{1/\gamma_h} \quad (2.11)$$

If lateral flow dominates, this expression suggests that the soil moisture depends on the average soil moisture in the catchment and the LFI. Similar calculations with the other terms in equation (2.6) lead to the three other estimates for the local soil moisture. In particular,

$$\theta_G = \bar{\theta} \quad (2.12)$$

$$\theta_R = \frac{\bar{\theta}}{\Pi} \left( \frac{1}{I_p} \right)^{1/\beta_r} \quad (2.13)$$

where  $(1/I_p)^{1/\beta_r}$  is another compound topographic attribute, which we call the evapotranspiration index (ETI),  $\Pi$  is the spatial average of the ETI within the catchment and:

$$\theta_A = \bar{\theta} \quad (2.14)$$

Each of these estimates is expected to become more reliable as the associated process becomes more important to the overall water balance. Thus, the weight for each process in the weighted average is determined from the term associated with that process in equation (2.6). For example, identifying the term associated with lateral flow and substituting  $\theta_L$  for  $\theta$  in this term (and simplifying) leads to:

$$w_L = \delta_0 K_{s,h} \left( \frac{\bar{\theta}}{\Lambda \phi} \right)^{\gamma_h} \quad (2.15)$$

The weights for the other processes are determined similarly and are given by the following expressions:

$$w_G = \frac{K_{s,v}}{\phi^{\gamma_v}} \bar{\theta}^{\gamma_v} \quad (2.16)$$

$$w_R = \frac{E_p I_p}{(1 + \alpha) \phi^{\beta_r}} \bar{\theta}^{\beta_r} \quad (2.17)$$

and:

$$w_A = \frac{E_p \alpha}{(1 + \alpha) \phi^{\beta_a}} \bar{\theta}^{\beta_a} \quad (2.18)$$

Finally, substituting equations (2.11) (2.12), (2.13), (2.14), (2.15), (2.16), (2.17), and (2.18), into equation (2.7), and simplifying produces:

$$\theta = \frac{K_{s,v} \left( \frac{\bar{\theta}}{\phi} \right)^{\gamma_v} \bar{\theta} + \frac{\delta_0 K_{s,h}}{\Lambda^{\gamma_h}} \left( \frac{\bar{\theta}}{\phi} \right)^{\gamma_h} \left[ \frac{\bar{\theta}}{\Lambda} \left( \frac{A}{cS^\varepsilon} \right)^{1/\gamma_h} \left( \frac{\kappa_{min}}{\kappa_{min} - \kappa} \right)^{1/\gamma_h} \right] + \frac{E_p}{(1 + \alpha) \Pi^{\beta_r}} \left( \frac{\bar{\theta}}{\phi} \right)^{\beta_r} \left[ \frac{\bar{\theta}}{\Pi} \left( \frac{1}{I_p} \right)^{1/\beta_r} \right] + \frac{E_p \alpha}{(1 + \alpha)} \left( \frac{\bar{\theta}}{\phi} \right)^{\beta_a} \bar{\theta}}{K_{s,v} \left( \frac{\bar{\theta}}{\phi} \right)^{\gamma_v} + \frac{\delta_0 K_{s,h}}{\Lambda^{\gamma_h}} \left( \frac{\bar{\theta}}{\phi} \right)^{\gamma_h} + \frac{E_p}{(1 + \alpha) \Pi^{\beta_r}} \left( \frac{\bar{\theta}}{\phi} \right)^{\beta_r} + \frac{E_p \alpha}{(1 + \alpha)} \left( \frac{\bar{\theta}}{\phi} \right)^{\beta_a}} \quad (2.19)$$

which provides an explicit estimate of the local soil moisture as a function of the average soil moisture in the catchment  $\bar{\theta}$ , two compound topographic indices (the LFI and the ETI), and various climate, soil, and vegetation parameters. It is possible for this expression to produce  $\theta$  values that exceed the porosity  $\phi$  at some locations, particularly if  $\bar{\theta}$  is near  $\phi$ . In such cases,  $\theta$  is reduced to  $\phi$  and the remaining water is interpreted to be flowing or standing on the ground surface. An alternative interpretation could allow for a higher local porosity than the spatially constant effective value represented by the  $\phi$  parameter.

### **Application Sites**

The model is evaluated by applying it to three catchments with extensive soil moisture datasets.

The Tarrawarra catchment is located near Melbourne, Australia and was originally described by

*Western and Grayson* [1998]. The catchment has a temperate climate with average annual

rainfall of 820 mm, average annual PET of 830 mm, a rainfall deficit in summer and a rainfall

excess in winter. The catchment is covered by pasture for cattle grazing. Soils generally have a

silty loam A horizon and a B horizon with higher clay content and soil depths vary from 40 cm

in the upper catchment to over 2 m in the low areas. The topography consists of undulating hills

with no incised channels (Figure 6), the catchment area is 10.5 ha and the total relief is 29 m. A

DEM with a 5 m linear dimension for its grid cells is available for this site and was used to

calculate the topographic attributes that are required by the model. The flow directions and

contributing areas for all catchments were calculated using the  $D_{\infty}$  algorithm [Tarboton, 1997].

Soil moisture data were collected in the top 30 cm of the soil with time-domain reflectometry

(TDR) probes on thirteen dates between 27 September 1995 and 29 November 1996 on a 10 m

by 20 m sampling grid [Western et al., 1999b]. The dataset was filtered to remove locations with

missing values on any date so that the remaining dataset includes 454 points. The  $\bar{\theta}$  values for

these dates range between 0.21 and 0.49. The observed soil moisture patterns at Tarrawarra exhibit more hillslope dependence during drier periods, stronger valley dependence during wetter periods, and a combination of the two patterns during intermediate periods (Figure 7 a,b,c). The Satellite Station field site of the Maharungi River Variability Experiment (MARVEX) is located on the North Island of New Zealand [Wilson *et al.*, 2003; Woods *et al.*, 2001]. The Satellite Station site has a humid climate with annual rainfall of approximately 1,600 mm and annual pan ET of approximately 1,300 mm [Wilson *et al.*, 2003]. Hillslope soils are generally silty clay loam, while lowland valley soils are relatively deep alluvial fill with high clay content. Satellite Station has a total area of approximately 60 ha with undulating terrain and relief of approximately 80 m (Figure 6). The topographic dataset for this location is available as a 10 m resolution DEM. The soil moisture dataset used consists of six sample dates between April 1998 and November 1999 and is comprised of TDR measurements collected at 0-30 cm depth on a 40 m grid.

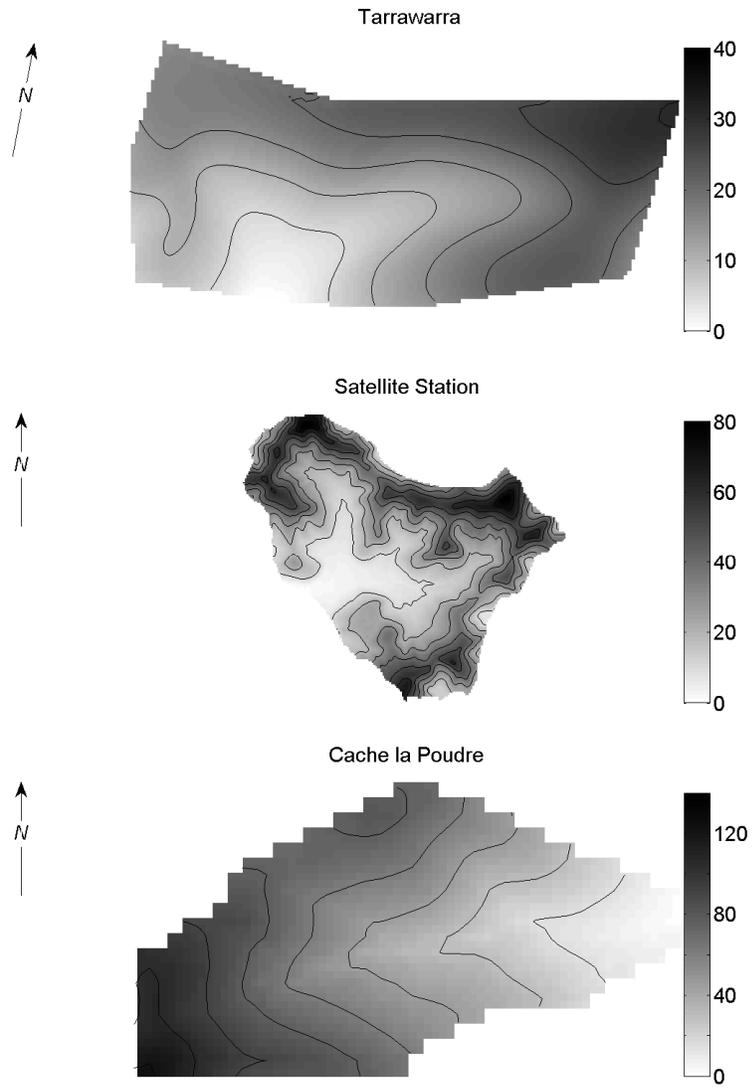


Figure 6 - Topography observed at the Tarrawarra, Satellite Station and Cache la Poudre catchments. Shading shows elevation (m) relative to each catchment's outlet, and lines show elevation contours.

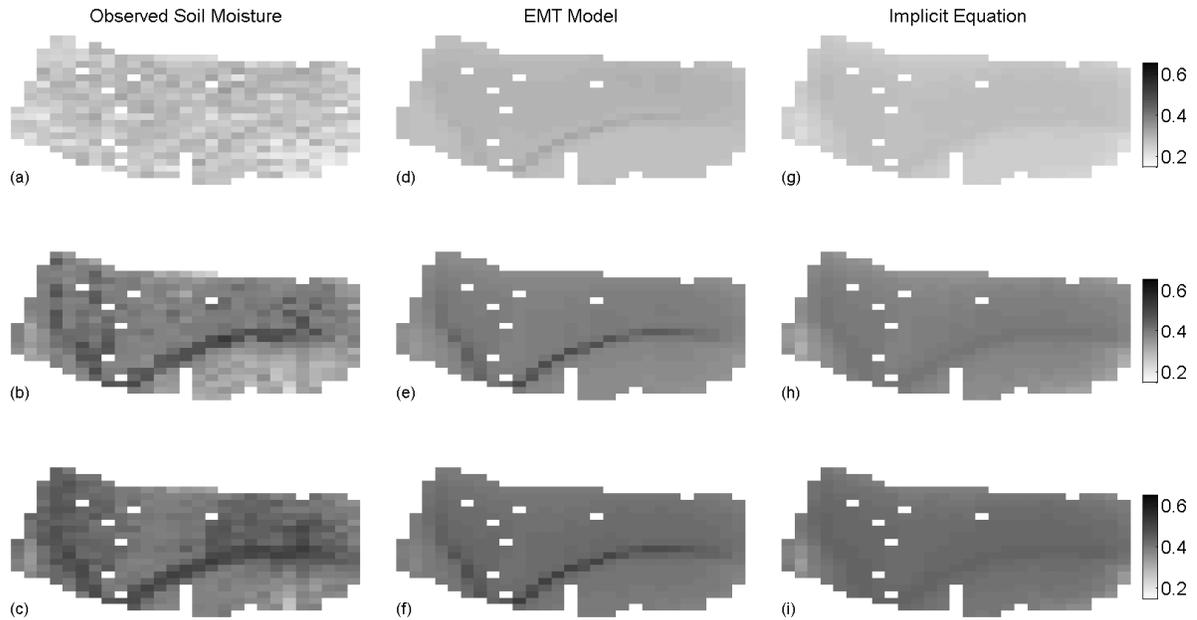


Figure 7 - Representative observed (a, b, c), EMT model (d, e, f), and implicit (equation 2.6) (g, h, i) soil moisture patterns for Tarrawarra. The observed soil moisture patterns are from 14 Feb. 1996, 27 Sep. 1995, and 2 May 1996, and the spatial-average moisture values are 0.26, 0.38, and 0.42 V/V for a, b, and c, respectively. NSCE values for the EMT model are 0.12, 0.54, and 0.53.

The dataset was again filtered so that it includes only locations with measurements on all dates (322 locations). The  $\bar{\theta}$  values for this dataset range from 0.25 to 0.46. Visually, the soil moisture patterns at Satellite Station are more stable through time than those at Tarrawarra with the valley bottoms remaining wetter than the upland areas on all dates (Figure 8a,b,c).

The Cache la Poudre catchment is located near Rustic, Colorado and is part of the larger Cache la Poudre River basin [Lehman and Niemann, 2008]. The catchment has a semi-arid climate with an annual precipitation of about 400 mm and annual PET of about 930 mm. The vegetation is aspect-dependent with shrubs on the south-facing hillslopes and a coniferous forest on the north-facing slope. The soils are also aspect-dependent with thin sandy soil on the south-facing hillslope and thicker mineral soils covered by litter and organic matter on the north-facing hillslope.

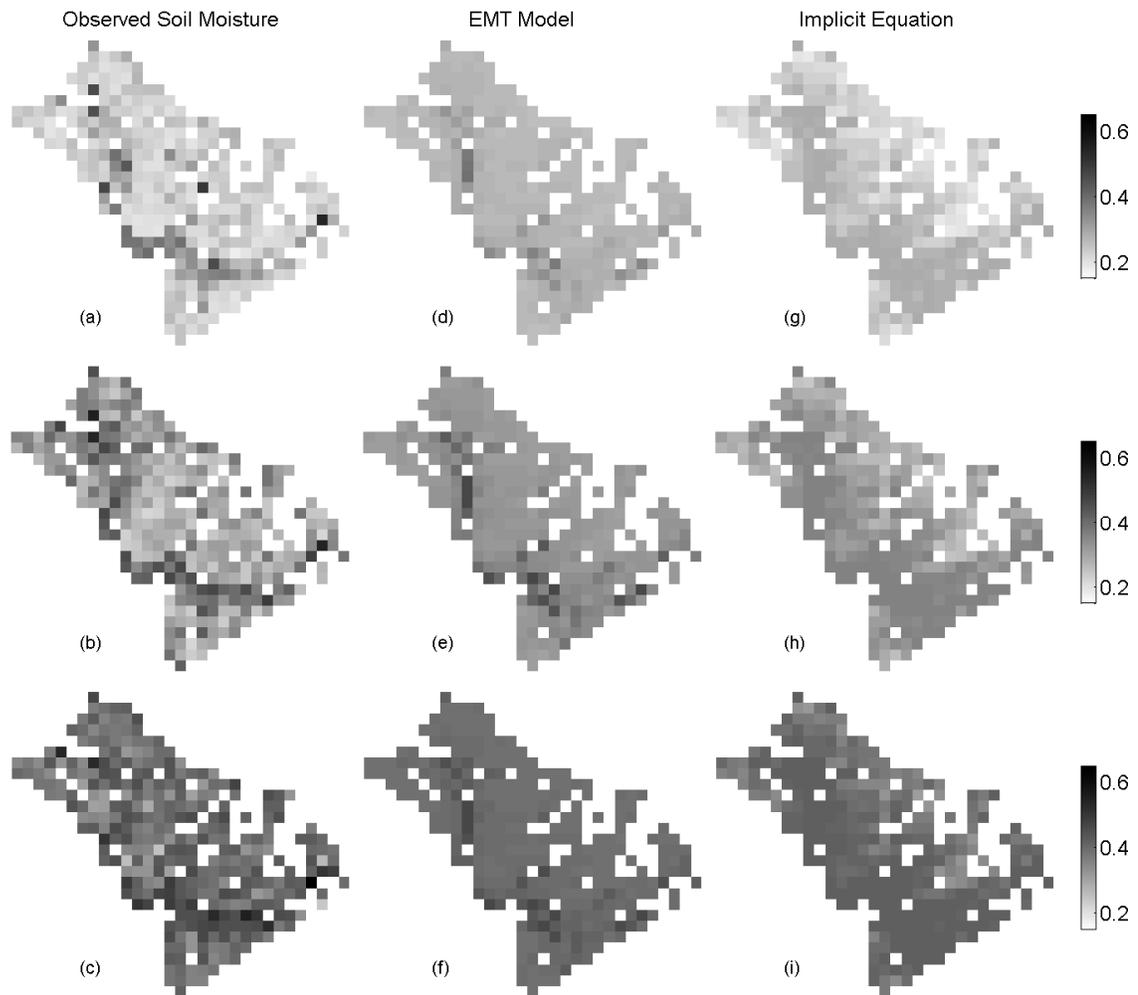


Figure 8 - Representative observed (a, b, c), EMT model (d, e, f), and implicit (equation 2.6) (g, h, i) soil moisture patterns for Satellite Station. The observed soil moisture patterns are from 18 Feb. 1999, 25 Mar. 1998, and 30 Oct. 1999, and the spatial-average moisture values are 0.25, 0.32, and 0.40 V/V for a, b, and c, respectively. NSCE values for the EMT model are 0.19, 0.24, and 0.14.

The catchment is approximately 8 ha, consists of the headwater area for one incised ephemeral channel and has total relief of approximately 124 m (Figure 6). The DEM for the catchment was developed from a total station and GPS survey and has 15 m resolution. Soil moisture data at the Cache la Poudre site were collected on nine dates between 22 April 2008 and 24 June 2008 using TDR instruments. The data were collected immediately after a sequence of spring rainfall events

and during the subsequent drying periods. Due to the shallow soils, particularly on the south-facing hillslope, soil moisture measurements were collected in the top 5 cm of the soil. A total of 350 locations remain for this dataset after removing points that have dates with missing data. Due to the semi-arid climate of this location, the  $\bar{\theta}$  values for this dataset are lower than the other two datasets and range between 0.04 and 0.19. Visually, the soil moisture patterns exhibit more consistent hillslope dependence than the patterns at the other two catchments (Figure 9 a,b,c).

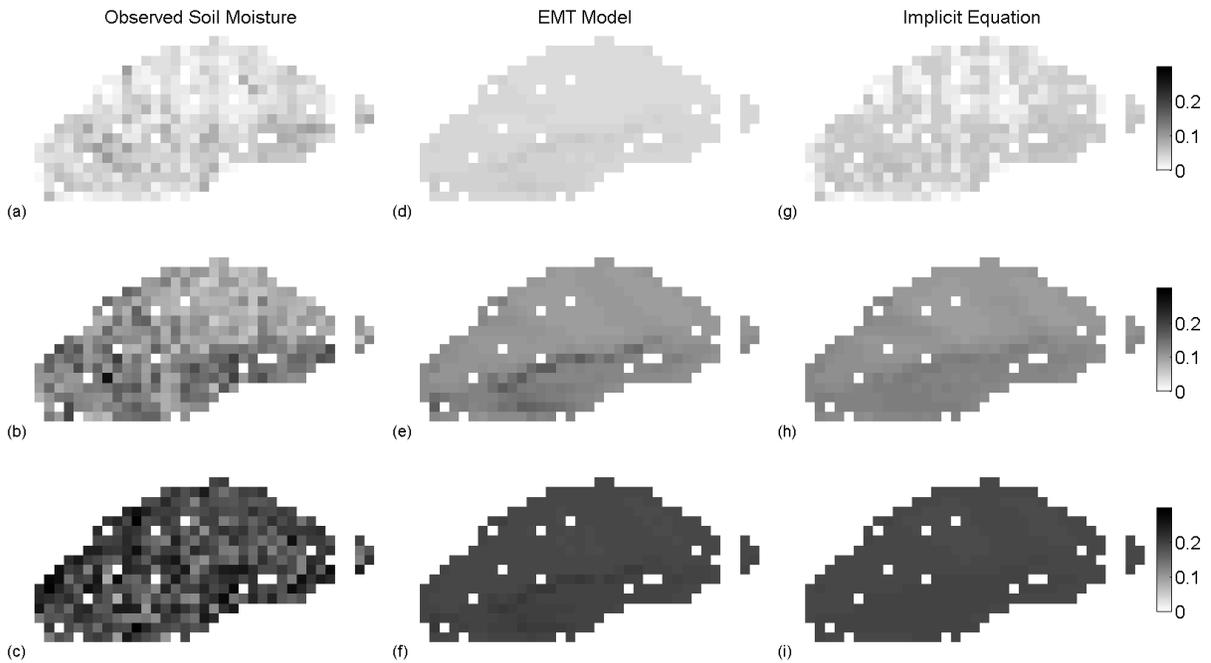


Figure 9 - Representative observed (a, b, c), EMT model (d, e, f) and implicit (equation 2.6) (g, h, i) soil moisture patterns for Cache la Poudre. The observed soil moisture patterns are from 24 June 2008, 12 June 2008, and 28 May 2008, and the spatial-average moisture values are 0.04, 0.11, and 0.19 V/V for a, b, and c, respectively. NSCE values for the EMT model are 0.07, 0.11, and 0.00.

## Model Calibration and Performance Results

A separate model calibration was performed for each catchment. In each case, the calibrated parameter set was found by maximizing the average of the Nash-Sutcliffe Coefficient of

Efficiency (NSCE) for all the dates in the available soil moisture dataset. Multiple initial sets of parameters were used in the optimization algorithm to improve the likelihood of finding the global optimum. These initial parameter sets were chosen by pseudo-random number generation within a reasonable range for each model parameter. The parameter ranges and final, calibrated parameter values are shown in Table 1. The PET values used in the model applications are annual average values. The value for Tarrawarra was obtained from the literature [*Western and Grayson, 1998*], while the values for Satellite Station and Cache la Poudre were calculated from meteorological observations and the Priestley-Taylor equation [*Priestley and Taylor, 1972*].

Table 1 - Parameter sets that produce soil moisture estimates with maximum average NSCE for each catchment

<u>Parameter</u>	<u>Allowed Range</u>	<u>Tarrawarra</u>	<u>Satellite Station</u>	<u>Cache la Poudre</u>
$K_{s,v}$	5 – 50,000 mm day <sup>-1</sup>	203.8 mm day <sup>-1</sup>	396.4 mm day <sup>-1</sup>	608.6 mm day <sup>-1</sup>
$K_{s,h}$	5 – 50,000 mm day <sup>-1</sup>	5.000 mm day <sup>-1</sup>	3647 mm day <sup>-1</sup>	25.42 mm day <sup>-1</sup>
$\phi$	0.25 – 0.70 m <sup>3</sup> m <sup>-3</sup>	0.6037 m <sup>3</sup> m <sup>-3</sup>	0.4834 m <sup>3</sup> m <sup>-3</sup>	0.3361 m <sup>3</sup> m <sup>-3</sup>
$\gamma_h$	4 – 25	7.704	8.530	25.00
$\gamma_v$	4 – 25	14.24	15.03	12.39
$\beta_r$	0.2 – 5.0	3.692	5.000	5.000
$\beta_a$	0.2 – 5.0	0.4809	3.025	4.090
$\alpha$	0 – 0.70	0.0042	0.7000	0.0971
$\delta_0$	300 mm	300 mm	300 mm	300 mm
$\kappa_{min}$	-10 <sup>6</sup> – catchment value m <sup>-1</sup> *	-292.6 m <sup>-1</sup>	-288.5 m <sup>-1</sup>	-320.1 m <sup>-1</sup>
$\varepsilon$	0 – 3	0.1126	3.000	2.980

\*The upper limit for  $\kappa_{min}$  is set slightly lower than the minimum curvature observed at the application catchment to avoid creating locations with no hydrologically active layer.

Also, PSRI was calculated only for the winter solstice at each catchment and used for all dates.

Time-varying values of PET and PSRI could be used in the model, but this level of detail is expected to be unwarranted given the equilibrium assumption made in the model development.

The columns labeled “EMT Model” in Figure 7, Figure 8 and Figure 9 show the simulated patterns for three representative dates at each catchment using the parameter sets with the maximum average NSCE. In Figure 7, which shows Tarrawarra, the EMT model patterns show increasingly prominent wet areas in the valley bottoms as the spatial-average soil moisture increases from (d) to (f). In contrast, the difference in the soil moisture values between the opposing hillslopes decreases. Both of these behaviors are consistent with the observations shown in (a) to (c). Some differences are also observed between the simulated and observed patterns. The observed dry pattern in (a) is much less organized (i.e. there is more speckle) than the simulated pattern. In addition, the model seems to slightly overestimate the importance of the valley bottom to the soil moisture values on this date. For the wetter two dates, the simulated soil moisture patterns exhibit wet locations within a narrower band in the valley bottom than the observed patterns. Overall, however, the type of topographic dependence and its transition between dates is approximately reproduced by the model. In Figure 8, which shows Satellite Station, the observed and simulated patterns all exhibit stable wet areas in the valley bottoms irrespective of the spatial-average soil moisture. Again, the observed patterns exhibit more local variability than the simulated patterns. In Figure 9, which shows the Cache la Poudre catchment, both the observed and the simulated soil moisture patterns tend to have wetter conditions on the north-facing hillslope than the south-facing hillslope, but much more speckle is present in the observed patterns than the simulated patterns. Overall, it appears that the EMT model can produce soil moisture patterns for catchments dominated by either hillslope or valley dependence as well as patterns for catchments that have varying combinations of these dependencies through time.

The ranges of NSCE among the dates in each dataset are shown in Table 2. For a given date, a value of one for the NSCE would mean that the model captures all of the spatial variation present on that date, and a value of zero would mean that the model does not explain any of the spatial variation.

Table 2 - Statistics of the NSCE values from all available dates at the three catchments when the best-performing parameter sets are used.

	<u>Tarrawarra</u>	<u>Satellite Station</u>	<u>Cache la Poudre</u>
Maximum	0.58	0.28	0.22
Average	0.31	0.18	0.08
Minimum	0.06	0.09	-0.02

Comparing between the three catchments, the EMT model explains on average between 8% and 31% of the spatial variation in the soil moisture. The best performance occurs at Tarrawarra where the model can explain up to 58% of the variation on one particular date. These values are similar to the performance of other methods that estimate soil moisture based on topography at these sites [Coleman and Niemann, 2010; Perry and Niemann, 2007; Western et al., 1999a; Wilson et al., 2005]. Approximately 40 percent of the total spatial variation in soil moisture at these scales is noise [Busch et al., 2011] so the model is capturing a substantial portion of the soil moisture dependence on topography. However, these values are relatively low, which implies that other factors such as soil and vegetation variations also impact the spatial variability. An alternative way to evaluate the model performance is to calculate a single NSCE for the entire space-time soil moisture dataset. When calculated in this way, the measure evaluates the model's ability to reproduce both the spatial and temporal variability in the complete dataset. The resulting NSCE values are 0.92, 0.70, and 0.76 for Tarrawarra, Satellite Station, and Cache la Poudre, respectively. The NSCE is higher when calculated in this fashion because the

variations in soil moisture between dates are quite large, and the model successfully reproduces such variations in part because the spatial-average soil moisture is provided as input.

The soil moisture values calculated by the EMT model (equation (2.19)) are approximations of the values that would be calculated by the more complex equation (2.5). To progress from equation (2.5) to (2.6), it was assumed that upslope average soil moisture values could be approximated with the local soil moisture values. To progress from equation (2.6) to (2.19), it was assumed that the implicit equation could be approximated by a linear combination of solutions in which each process is considered dominant. To partially evaluate this solution procedure, the patterns obtained from equation (2.19) are compared to those obtained by iteratively solving equation (2.6). The model parameters that provided the maximum average NSCE value for the EMT model at each catchment were used in both equations. Because equation (2.6) requires specification of  $F$  instead of the spatial-average soil moisture,  $F$  was selected so that equation (2.6) produces a pattern with the same spatial-average as the EMT model. The columns titled “Implicit Equation” in Figure 7, Figure 8 and Figure 9 show the results from equation (2.6) when used in this manner. Note that the EMT model leads to a more exaggerated difference in the soil moisture between the valley bottoms and hillslopes than would occur using the same parameters in the implicit equation. The EMT model also produces wet locations in a narrower band in the valley bottoms than the implicit equation (these wet areas were found to be too narrow when comparing the EMT results to the observations). Because the parameters were calibrated using the EMT model, their values are likely somewhat different than if they had been calibrated using equation (2.6). However, sensitivity tests indicated that the dependence of the EMT results on the model parameters was similar to the dependence of equation (2.6). Model parameters were varied to low, moderate and high values within their

allowed ranges and the resulting effects on the EMT model and the implicit equation results were compared. While the degree of the changes in the two estimates differed in some cases, the response of the patterns from the two equations was generally similar. Thus, the superposition of the individual solutions appears to be a reasonable approximation for the analysis presented later in this paper. It should be noted that the effect of using local values in place of average upslope values in the model remains untested at this writing due to the complexity of implementing equation (2.5).

The EMT model's ability to produce moisture patterns with a diversity of spatial organizations is a product of its use of the LFI and ETI patterns. The calibrated LFI and ETI patterns for each of the three catchments are shown in Figure 10. The LFI and ETI patterns depend on the model parameters but not on  $\bar{\theta}$  so the patterns are fixed once a set of parameters is selected. The LFI patterns for each catchment reflect the valley configuration as shown in Figure 6. Darker regions correspond to convergent locations, locations of higher contributing area, and/or locations with lower slope, which tend to occur in the valley bottoms. If lateral flow is an important process, these locations are expected to have elevated soil moisture. The ETI patterns distinguish the hillslopes that are more exposed to the sun from those that are more shaded due to its dependence on PSRI. The Tarrawarra and Satellite Station patterns have lower ETI values on north-facing slopes because they are in the southern hemisphere, while the behavior is reversed at Cache la Poudre because the catchment is in the northern hemisphere. Darker areas have higher ETI values (i.e., lower potential insolation) and thus are expected to be wetter if radiative ET is important.

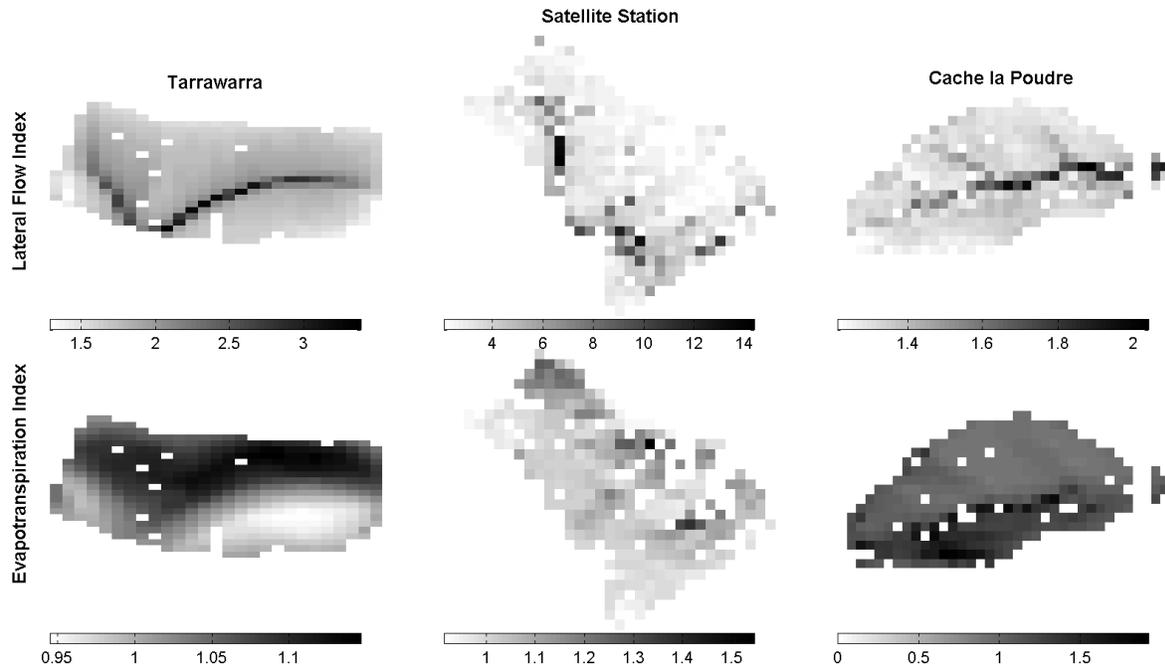


Figure 10 - Calibrated LFI and ETI patterns for the Tarrawarra, Satellite Station, and Cache la Poudre catchments.

The relative importance of the LFI and ETI patterns depends on the spatial-average soil moisture in the catchment. The relative weight for each term in the EMT model is plotted as a function of  $\bar{\theta}$  in Figure 11. For example, the relative weight for lateral flow is  $w_L / (w_G + w_L + w_R + w_A)$ .

The weights for the Tarrawarra catchment are depicted in Figure 11(a). At very low values of  $\bar{\theta}$ , the aerodynamic ET weight is largest, which implies that its estimate of soil moisture ( $\bar{\theta}$ ) is most important. As  $\bar{\theta}$  increases to moderately dry values the radiative ET term becomes largest, the ETI pattern becomes important and the hillslope-dependent soil moisture pattern occurs. As  $\bar{\theta}$  becomes moderately large, the lateral flow weight and influence of the LFI reach their maximums. This situation produces the valley-dependent soil moisture pattern. For very large values of  $\bar{\theta}$ , the deep drainage weight becomes largest and its estimate ( $\bar{\theta}$  again) becomes the

most important. Thus, the soil moisture pattern becomes more spatially constant under very wet conditions.

At Satellite Station (Figure 11(b)) aerodynamic ET and deep drainage again dominate at the lower and upper ends, respectively, of the  $\bar{\theta}$  range. However, their relative weights are larger than at Tarrawarra, which indicates that these terms are more important at the Satellite Station catchment. The weight for radiative ET peaks at a slightly higher value of  $\bar{\theta}$  than at Tarrawarra, but the lateral flow weight peaks at a slightly lower  $\bar{\theta}$  value than at Tarrawarra. The radiative ET weight is much smaller at Satellite Station than at Tarrawarra, which implies that the hillslope-dependent pattern is much less important at Satellite Station than at Tarrawarra. The relative importance of lateral flow suggested here is consistent with previous analyses of Satellite Station. Although the flow may occur deeper than the root zone, lateral movement of moisture is an important determinant of soil moisture patterns at that catchment [Wilson *et al.*, 2005].

The weights for Cache la Poudre in Figure 11(c) have some apparent similarities to the other two sites, but it is important to note that the plotted range of  $\bar{\theta}$  is much smaller than for the other two catchments. The aerodynamic ET and deep drainage processes again dominate at the extremes of the range, but deep drainage is already the most important process at much lower values of  $\bar{\theta}$ . The importance of deep drainage at this catchment is consistent with field experience. Infiltrated water quickly drains to deeper levels in the soil due to the coarse texture at this site. Figure 11(c) also shows that the radiative ET term is the most important over a substantial portion of the plotted range of  $\bar{\theta}$  and the lateral flow weight is essentially zero for the entire range of  $\bar{\theta}$  that is considered. This result is also expected given the apparent lack of valley dependence in the soil moisture patterns at this catchment.

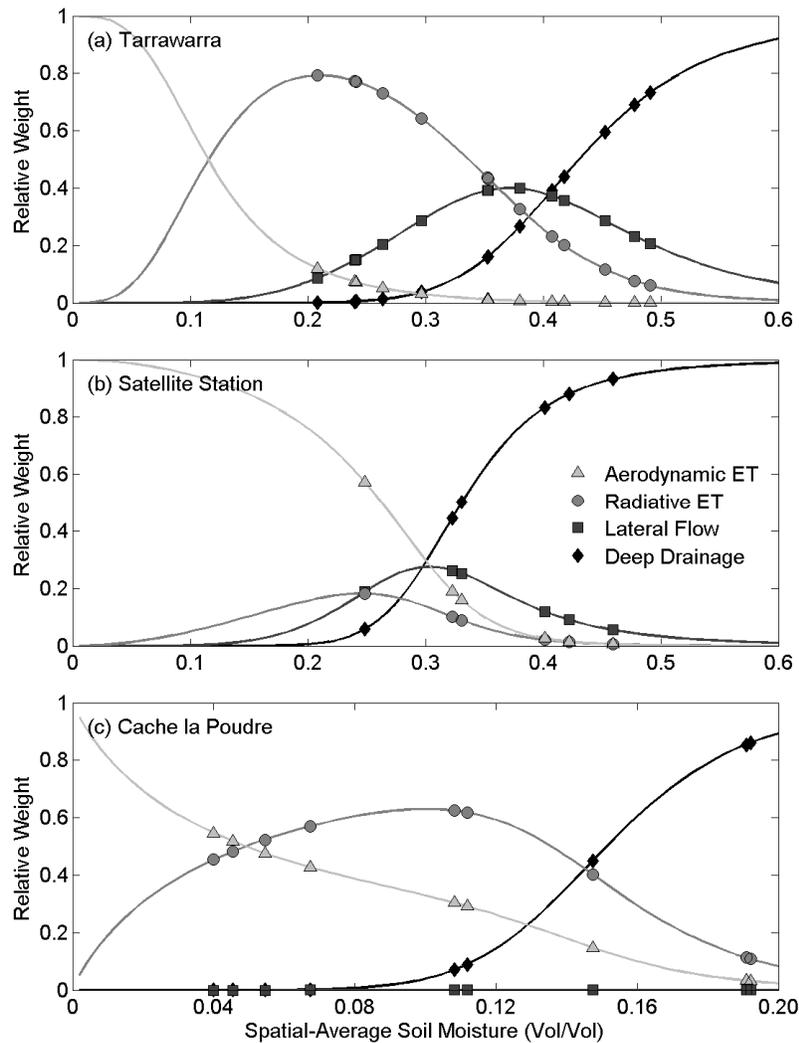


Figure 11 - Calibrated weights plotted as a function of spatial average soil moisture for (a) Tarrawarra, (b) Satellite Station and (c) Cache la Poudre. Markers indicate observed spatial-average soil moisture values and the associated process weights. Note the lower maximum value for spatial average soil moisture in (c).

## Development of Pattern and Instability Metrics

The dominant type of topographic dependence can be investigated by calculating the contribution of each underlying pattern to the total spatial variance of soil moisture. Spatial variation in the EMT soil moisture patterns arises solely from the LFI and ETI patterns and depends linearly on those patterns. Thus, the variance in the modeled patterns can be written as:

$$\sigma_{\theta}^2 = \sigma_L^2 + \sigma_R^2 + 2\rho_{LR}\sigma_L\sigma_R \quad (2.20)$$

where  $\sigma_L^2$  and  $\sigma_R^2$  represent the spatial variance produced by the LFI and ETI patterns, respectively, and  $\rho_{LR}$  is the correlation coefficient between the patterns of variation that are introduced by the LFI and ETI.

Both  $\sigma_L^2$  and  $\sigma_R^2$  can be evaluated by referring back to the EMT model (equation (2.19)) and recognizing that the LFI and ETI patterns are the only components of the model that vary spatially. Specifically:

$$\sigma_L^2 = \left[ \frac{\delta_0 K_{s,h} \left( \frac{\bar{\theta}}{\phi\Lambda} \right)^{\gamma_h} \left( \frac{\bar{\theta}\sigma_{LFI}}{\Lambda} \right)}{K_{s,v} \left( \frac{\bar{\theta}}{\phi} \right)^{\gamma_v} + \delta_0 K_{s,h} \left( \frac{\bar{\theta}}{\phi\Lambda} \right)^{\gamma_h} + \frac{E_p}{1+\alpha} \left( \frac{\bar{\theta}}{\Pi\phi} \right)^{\beta_r} + \frac{E_p\alpha}{1+\alpha} \left( \frac{\bar{\theta}}{\phi} \right)^{\beta_a}} \right]^2 \quad (2.21)$$

$$\sigma_R^2 = \left[ \frac{\frac{E_p}{1+\alpha} \left( \frac{\bar{\theta}}{\phi\Pi} \right)^{\beta_r} \left( \frac{\bar{\theta}\sigma_{ETI}}{\Pi} \right)}{K_{s,v} \left( \frac{\bar{\theta}}{\phi} \right)^{\gamma_v} + \delta_0 K_{s,h} \left( \frac{\bar{\theta}}{\phi\Lambda} \right)^{\gamma_h} + \frac{E_p}{1+\alpha} \left( \frac{\bar{\theta}}{\Pi\phi} \right)^{\beta_r} + \frac{E_p\alpha}{1+\alpha} \left( \frac{\bar{\theta}}{\phi} \right)^{\beta_a}} \right]^2 \quad (2.22)$$

where  $\sigma_{LFI}$  and  $\sigma_{ETI}$  are the standard deviations of the LFI and ETI patterns within the catchment, respectively. For any given  $\bar{\theta}$ , the relative importance of the two patterns can be evaluated by  $n \equiv \sigma_L^2 - \sigma_R^2$ . If  $n$  is larger than zero, more variation is introduced by the LFI than the ETI. If it is smaller than zero, more variation is introduced by the ETI. Note that the variable  $n$  depends on  $\bar{\theta}$ , so its value can change through time in a catchment.

To assess the type of topographic dependence that most commonly occurs at a catchment,  $\bar{\theta}$  is considered a random variable in time and assumed to conform to a uniform distribution in the range  $[0, \phi]$ , where  $\phi$  is the parameter value. A simple uniform distribution is used because very

few values of  $\bar{\theta}$  are available at each catchment to identify an appropriate distribution. In addition, a uniform distribution is simple to modify if the porosity is changed (a scenario that will be encountered later in the paper). The expected value of  $n$ ,  $E[n]$ , could be used to assess the dominant pattern of variation in a catchment, but it is difficult to interpret because its magnitude scales with the average spatial-variability of soil moisture. Thus, we define a final measure for the dominant type of topographic dependence (i.e., the pattern metric  $N_p$ ) as:

$$N_p \equiv \frac{E[n]}{E[\sigma_L^2 + \sigma_R^2]} \quad (2.23)$$

The maximum value of  $N_p$  is 1, which indicates that all spatial variation is due to lateral flow and valley dependence is the dominant organization type. The minimum value of  $N_p$  is -1, which implies that all variation is due to radiative ET and hillslope dependence is dominant.

A measure of time instability can be developed using a similar approach. A soil moisture dataset can be considered to exhibit time instability if the type of topographic dependence changes with  $\bar{\theta}$ . In particular, the valley pattern should be dominant for some range of  $\bar{\theta}$ , and the hillslope pattern should be dominant for a different range of  $\bar{\theta}$ . The strength of the valley dependence over the full range of  $\bar{\theta}$  can be assessed by  $E[n_L] / E[\sigma_L^2 + \sigma_R^2]$ , where:

$$n_L \equiv \begin{cases} n & \text{if } n > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.24)$$

Only positive values of  $n$  are retained here because they indicate a valley-dependent soil moisture pattern. A similar measure  $E[n_R] / E[\sigma_L^2 + \sigma_R^2]$  can be developed to assess the overall strength of the hillslope dependence, where:

$$n_R \equiv \begin{cases} |n| & \text{if } n < 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.25)$$

The absolute value is included so that  $n_R$ , like  $n_L$ , takes only non-negative values. A catchment with time instability must have large values for both  $E[n_L]/E[\sigma_L^2 + \sigma_R^2]$  and  $E[n_R]/E[\sigma_L^2 + \sigma_R^2]$ . Thus, the instability metric  $N_I$  is defined as:

$$N_I \equiv \frac{4E[n_L]E[n_R]}{E[\sigma_L^2 + \sigma_R^2]^2} \quad (2.26)$$

This metric multiplies  $E[n_L]$  and  $E[n_R]$  to emphasize balanced combinations of  $E[n_L]$  and  $E[n_R]$  (when multiplying two numbers that are in the range 0 to 1 and sum to 1, the maximum is reached if both numbers are 0.5). The 4 is included so that the instability metric has a feasible range of [0, 1]. If  $N_I$  takes a value of zero, then the soil moisture patterns are time stable because a single type of topographic dependence always controls the variation. As  $N_I$  becomes large, time instability increases because the contributions of the two patterns become increasingly balanced but occur at distinct ranges of  $\bar{\theta}$ .

## Results for Pattern and Instability Metrics

Before examining the catchment characteristics that affect the values of the pattern and instability metrics, we first analyze these metrics and their origins at the three catchments.

Figure 12 plots the variance induced by the LFI ( $\sigma_L^2$ ), the variance induced by the ETI ( $\sigma_R^2$ ), and the sum of the two variances ( $\sigma_L^2 + \sigma_R^2$ ) as a function of  $\bar{\theta}$  for all three catchments. Correlation coefficients calculated between the calibrated LFI and ETI patterns are -0.06 for Tarrawarra, -0.24 for Satellite Station, and -0.07 for Cache la Poudre. Thus,  $\sigma_L^2 + \sigma_R^2$  is very close to the total simulated soil moisture variance in all cases except Satellite Station (see equation(2.20)). At each of the three catchments,  $\sigma_L^2 + \sigma_R^2$  is lower at the extremes of the  $\bar{\theta}$  range and higher in the middle. The contributions of the LFI and ETI patterns to  $\sigma_L^2 + \sigma_R^2$  vary between the catchments.

At Tarrawarra, both patterns contribute to the sum for a wide range of  $\bar{\theta}$  values. At low  $\bar{\theta}$  values, the ETI-related pattern contributes the majority of the sum. At moderate  $\bar{\theta}$  values, both patterns contribute substantial variance, and at high  $\bar{\theta}$  values, the LFI-related pattern contributes the majority of the sum. Recall that Figure 11 shows the weight associated with radiative ET is larger than the weight associated with lateral flow at Tarrawarra. This result suggests that, overall, the radiative ET process is more important than the lateral flow process. However, in Figure 12, the variance introduced by the lateral flow pattern is much larger than the variance introduced by the radiative ET. Thus, although the ET process is stronger at this site, the spatial variance that it introduces is smaller than the variance introduced by lateral flow. In contrast,  $\sigma_L^2 + \sigma_R^2$  for Satellite Station and Cache la Poudre are each dominated by a single pattern over their entire ranges of  $\bar{\theta}$  (Figure 12). At Satellite Station, the LFI pattern contributes the vast majority of the sum, while at Cache la Poudre, only the ETI pattern is important. These results are consistent with the weights shown previously for these two catchments.

Figure 13 plots  $n$  as a function of  $\bar{\theta}$  for the three catchments. In all three cases,  $n$  is near zero when  $\bar{\theta}$  is small relative to the observed range for each catchment. It then moves away from zero in the middle of the observed range and moves closer to zero again as  $\bar{\theta}$  continues to increase to the upper portion of the range. At Tarrawarra,  $n$  is negative at lower  $\bar{\theta}$  values, indicating more variation is contributed by radiative ET than by lateral flow. At higher  $\bar{\theta}$  values,  $n$  becomes positive indicating that lateral flow now introduces more variance. For Satellite Station,  $n$  is always positive, and for Cache la Poudre,  $n$  is always negative.

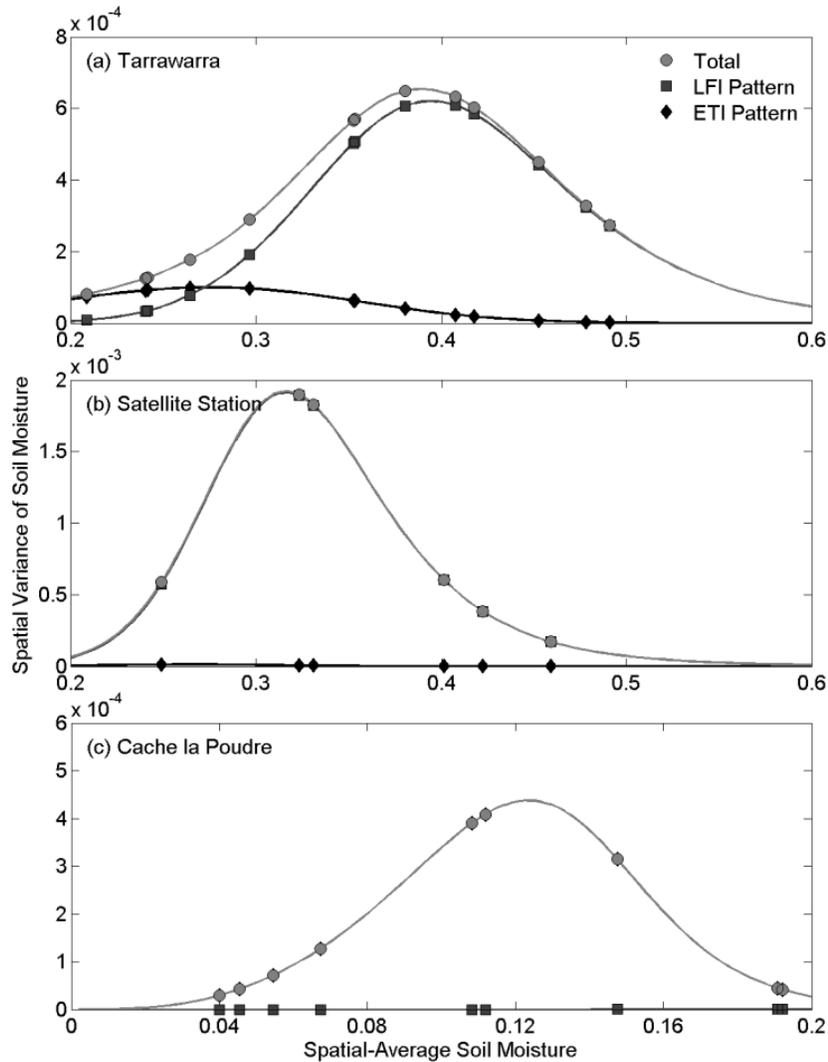


Figure 12 - Soil moisture variance introduced by the LFI and ETI patterns and their sum for (a) Tarrawarra, (b) Satellite Station and (c) Cache la Poudre. Markers indicate observed spatial-average soil moisture values and the associated variances.

Overall,  $N_p = 0.70$  at Tarrawarra,  $N_p = 0.99$  at Satellite Station, and  $N_p = -1.00$  at Cache la Poudre, and these values agree with visual interpretations of the soil moisture patterns at those catchments. The  $N_l$  values for Tarrawarra, Satellite Station, and Cache la Poudre are 0.15, 0.00, and 0.00, respectively. These values are again consistent with the visual interpretations. Also note that while Tarrawarra exhibits visible time instability, its  $N_l$  value is still far from the theoretical maximum of 1.

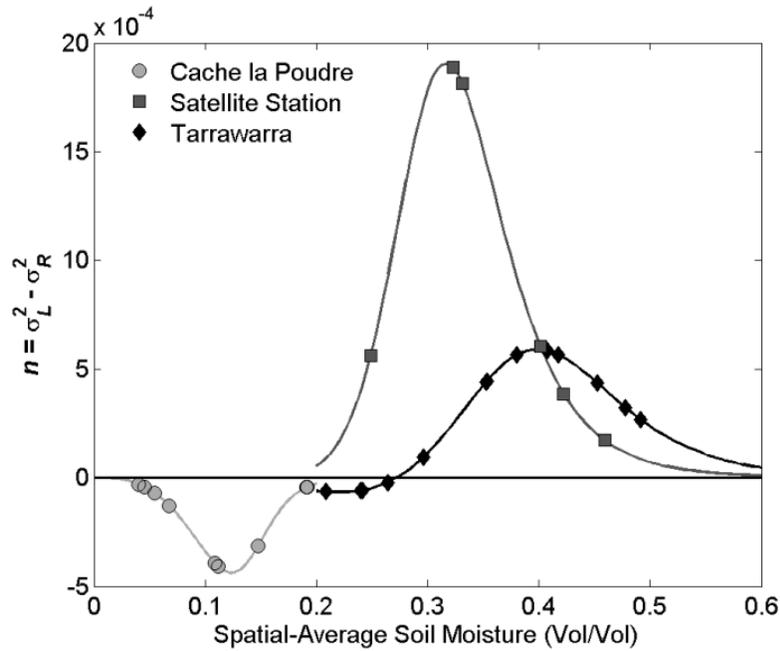


Figure 13 - Difference between soil moisture variances introduced by the LFI and ETI patterns ( $n$ ) as a function of the spatial average soil moisture ( $\bar{\theta}$ ) for the Tarrawarra, Satellite Station, and Cache la Poudre catchments.

To determine the soil, vegetation, and climatic conditions under which a given type of topographic dependence occurs, the pattern metric  $N_P$  is calculated as the model parameters are varied. The Tarrawarra catchment and its topography are chosen as the test case for this analysis. The parameters  $K_{s,v}$ ,  $\phi$ ,  $\gamma$ ,  $\beta_a$ ,  $d_0$ , and  $\kappa_{min}$  are held at their best-performing values while  $K_{s,h}$ ,  $\gamma_h$ ,  $\beta_r$ ,  $\alpha$ ,  $\varepsilon$ , and  $E_p$  are varied over their respective feasible ranges using 10 values approximately uniformly or log-uniformly spaced. The feasible ranges for the parameters are intended to represent broad ranges of soil, vegetation, and climatic conditions, not necessarily conditions that could occur at Tarrawarra. All model parameters were varied in a series of experiments like this one, but we show only the results for parameters whose values are more commonly known at a catchment or have a greater effect on the metrics. For every generated parameter set,  $N_P$  is calculated. Then, a box plot is created for  $N_P$  as each parameter is varied

over its range (Figure 14) to show both the influence of the parameter on  $N_P$  and the extent to which that influence is affected by the other parameters. In Figure 14, the bars indicate the median values of  $N_P$ , and the boxes show the interquartile ranges (IQRs). Whiskers are not included for clarity (they span the allowable range of  $N_P$  for nearly all cases). In Figure 14, the median value of  $N_P$  increases as  $K_{s,h}$  increases through its allowable range. This behavior indicates that soils that are highly conductive in the horizontal direction are more likely to produce valley-dependent soil moisture patterns. The IQRs are relatively small, particularly toward the lower limit of the range, suggesting that  $K_{s,h}$  has a relatively strong influence on the pattern metric. The effect of the pore disconnectedness index  $\gamma_h$  is approximately the opposite of the effect of  $K_{s,h}$ .

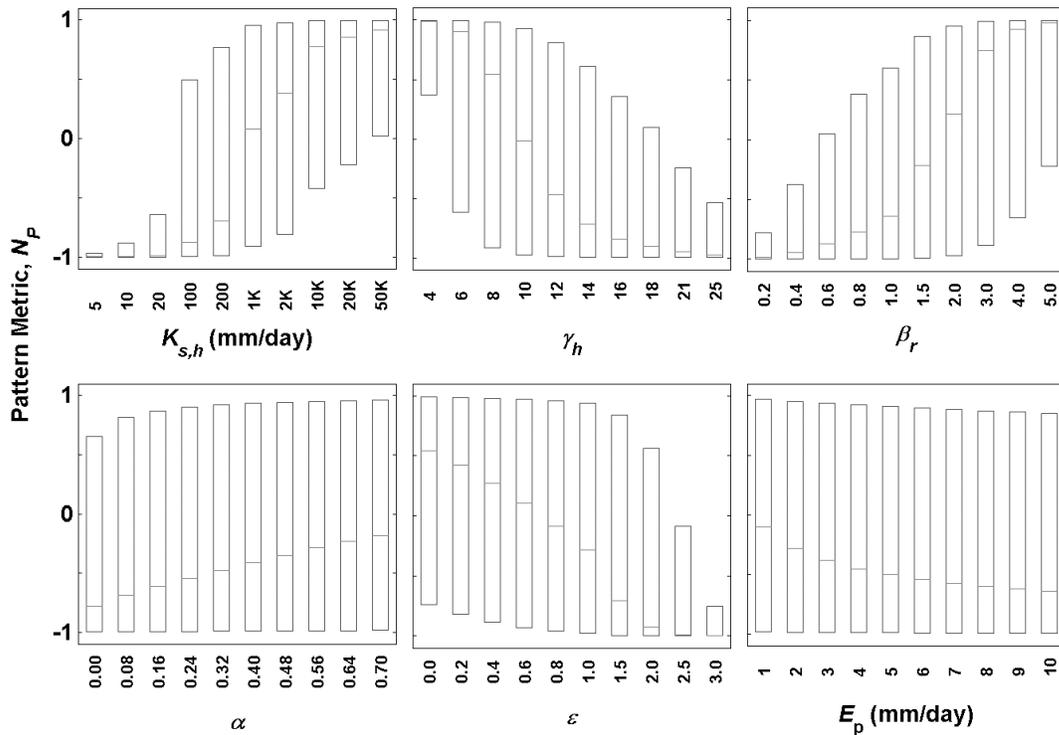


Figure 14 - Behavior of the pattern metric as the model parameters are changed at the Tarrawarra catchment. Bars indicate the median metric value and boxes span the IQR.

As  $\gamma_h$  increases the median  $N_P$  values decrease indicating a shift toward hillslope-dependent soil moisture patterns. As the pores become more disconnected, lateral flow is inhibited and the relative strength of the hillslope dependence is promoted. The effect of  $\beta_r$  on the  $N_P$  values is similar to that of  $K_{s,h}$ . As  $\beta_r$  increases, the vegetation cover decreases, which reduces the influence of radiative ET on the soil moisture and promotes a valley-dependent soil moisture pattern. The ratio of aerodynamic and radiative ET terms  $\alpha$  has only a small effect on the values of  $N_P$ , with the median value increasing as  $\alpha$  increases. The large IQRs indicate that this parameter has a weak connection to the pattern metric. The exponent relating the hydraulic gradient to the surface slope  $\varepsilon$  has a substantial effect on the median  $N_P$  values, with median values decreasing as  $\varepsilon$  increases. As  $\varepsilon$  exceeds a value of one, the moisture pattern becomes more hillslope dependent. Because the topographic slope has values less than one, an increase in  $\varepsilon$  means that the hydraulic gradient and thus lateral flow are reduced, making radiative ET more competitive. Nonetheless, the wide IQRs for all values but the upper end of the range indicate that this parameter is not as strongly connected to the  $N_P$  values as  $K_{s,h}$ , or  $\gamma_h$ . Finally, the PET has an effect on the metric that is similar in magnitude but in the opposite direction to that of  $\alpha$ . The broad IQRs in this plot are somewhat unexpected because they imply that PET does not have a strong influence on the type of topographic dependence in comparison to several other parameters.  $K_{s,h}$  is important in part because it can range over orders of magnitude, while  $\beta_r$ ,  $\gamma_h$ , and, to a lesser extent  $\varepsilon$ , are important in part because they are exponents.  $E_p$  has neither of these properties and thus has a weaker influence. Considering the parameters that are not varied in the figure,  $K_{s,v}$  has a much smaller effect than  $K_{s,h}$  on the pattern metric, but the metric does decrease when  $K_{s,v}$  is at the upper end of its range. The magnitude of  $\gamma_v$ 's effect on the pattern metric is

similar to that of  $K_{s,v}$ , but the metric is highest for the lowest values of  $\gamma_v$  and reaches a minimum for mid-range values.  $\phi$ ,  $\beta_a$ , and  $\kappa_{min}$  have almost no effect on the pattern metric.

The impact of soil, vegetation, and climatic conditions on time instability can be evaluated using the same collection of simulations if  $N_I$  is calculated and plotted instead of  $N_P$ . Figure 15 shows the box plots for the instability metric as six parameters are varied over their allowable ranges.

The plots indicate that  $N_I$  most commonly takes on values less than 0.4. In order for  $N_I$  to achieve a value of 1, a single type of topographic dependence would need to produce all the soil moisture variation for half of the values of  $\bar{\theta}$ , and the other type of topographic dependence would have to produce all of the variation for the other half. This abrupt switching is unrealistic and cannot be easily produced by the EMT model, so one would not expect many  $N_I$  values in the upper half of the range. The largest median values for the  $N_I$  metric occur with  $K_{s,h}$  values greater than 1,000 mm/day,  $\gamma_h$  values in the upper half of its range,  $\beta_r$  values in the lower half of its range, and low values of  $\varepsilon$ . This scenario implies that time instability is more likely for a catchment if the soils are conducive to horizontal flow, have moderately high pore disconnectedness, relatively thick vegetation cover, and sufficient hydraulic gradient to promote horizontal flow. Relatively high values of  $K_{s,h}$  increase instability because they promote a balance in the contributions of lateral flow and radiative ET to the spatial soil moisture variance when integrated over the range of  $\bar{\theta}$ . Such a balance is required if each process is to be dominant for certain  $\bar{\theta}$  ranges. The combination of large  $\gamma_h$  values and small  $\beta_r$  values promote balance between the competing processes, but they also allow lateral flow to become dominant under relatively wet conditions while radiative ET becomes dominant under relatively dry conditions.

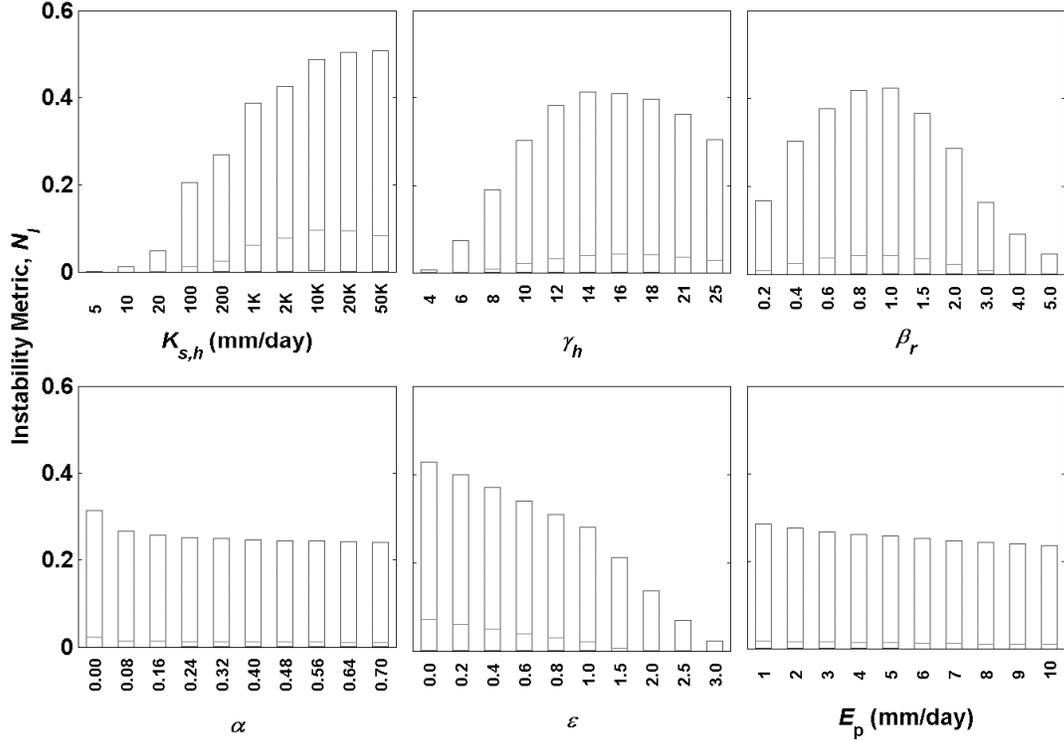


Figure 15 - Behavior of the instability metric as the model parameters are changed at the Tarrawarra catchment. Bars indicate the median metric value and boxes span the IQR.

Small  $\varepsilon$  values accentuate the spatial variance produced by the lateral flow pattern when it occurs. Regarding parameters not shown in Figure 15, low  $K_{s,v}$  values can promote instability but generally the effect is less than that for  $K_{s,h}$ . Variations in  $\gamma$ ,  $\beta_a$ ,  $\phi$ ,  $\alpha$ ,  $\kappa_{min}$ , and  $E_p$  do not strongly affect the values of  $N_j$ .

## Conclusions

1. Realistic soil moisture patterns can be produced by assuming that the spatial structure of the pattern is at equilibrium and inferring the spatial variability of the processes from the topographic attributes. These patterns can reproduce not only the two distinct types of topographic dependence observed at Satellite Station and Cache la Poudre but also the

time instability observed at Tarrawarra. However, these patterns are typically capable of reproducing less than half of the spatial variance in the observed soil moisture patterns.

2. The PET alone is not an important determinant of the type of topographic dependence that occurs in the simulated soil moisture patterns. This result is perhaps unexpected because drier conditions tend to be associated with hillslope-dependent soil moisture patterns in the three analyzed datasets. Thus, one might expect that catchments with higher PET would have more hillslope-dependent soil moisture patterns. If all other catchment characteristics are held constant in the model, higher PET values do promote the occurrence of hillslope-dependent soil moisture patterns. However, the plausible range of PET and its influence on the pattern type are not enough to overcome the influence of other site characteristics.
3. Certain soil properties in the model have a strong influence on the type of topographic dependence that occurs at a catchment. Higher saturated horizontal hydraulic conductivities and greater pore connectedness facilitate lateral flow and thus promote valley-dependent soil moisture patterns. Because saturated hydraulic conductivity can vary over orders of magnitude and because pore connectedness influences a model exponent, these factors have a strong impact on the resulting type of soil moisture pattern. Similarly, the model exponent  $\varepsilon$  that determines the horizontal hydraulic gradient from the topographic slope has a relatively strong impact on the type of topographic dependence.
4. Greater vegetation transpiration efficiency or thicker vegetation cover, represented in the EMT model by lower values of  $\beta_r$ , have a relatively strong influence on the resulting soil moisture patterns. Greater efficiency or thicker vegetation allows the plants to continue

extracting soil water from hillslopes with lower moisture, thus enhancing the role of radiative ET and the importance of the hillslope-dependent pattern. The influence of the vegetation cover on the pattern type is potentially important because it suggests that a modification of the vegetation cover could alter the structure of the soil moisture pattern, which could have significant effects on the catchment's runoff production [Western *et al.*, 2001].

5. Time instability in the model is promoted by parameter values that allow a balance in the overall importance of lateral flow and radiative ET. Such a balance is required for one process to control the soil moisture pattern under certain conditions, while the other process controls the pattern for other conditions. Among other factors, this balance is more commonly achieved with relatively large horizontal saturated hydraulic conductivity values.
6. Time instability in the model also requires parameter values that produce a switching in the dominant process between wet and dry conditions. The combination of relatively high values of pore disconnectedness and high values of vegetation cover (or transpiration efficiency) produces this behavior.

Several notable avenues are available for future research. First, the current analysis assumes that soil, vegetation, and climatic properties are all spatially and temporally constant within the catchment. In real catchments, it is expected that time instability occurs when time-unstable topographic dependence overcomes stable variations produced by other factors. This competition could be explored in the future by including spatial variations in the soil and vegetation characteristics represented in the model. Second, the EMT model is potentially well-suited for downscaling applications because its input is the spatial-average

soil moisture. Further testing of the model should be conducted, including examining its performance when a grid of spatial-average soil moisture values is used and when limited data are available for calibration. Third, the EMT soil moisture estimate can be viewed as a generalization of the wetness index used in TopModel [*Beven and Kirkby, 1979*]. One could potentially use this approach to examine whether this more flexible measure of catchment wetness would allow superior rainfall-runoff simulations, particularly under dry antecedent conditions when the wetness index might not apply.

### **Acknowledgments**

The authors acknowledge the Army Research Office, the Engineer Research and Development Center, and the Center for Geosciences and Atmospheric Research for their financial support.

The authors also thank all parties involved in MARVEX, including researchers from NIWA and the University of Melbourne for supplying their data, as well as Tim Green from the Agricultural Research Service for helpful suggestions. Finally, the authors thank Brandon Lehman, Josh Melliger, Jonathan Freed, and Rob Erskine for collecting the Cache la Poudre data.

## References

- Beven, K. J., and M. J. Kirkby (1979), A physically-based, variable contributing area model of basin hydrology, *Hydrol. Sci. Bull.*, 24(1), 43-69.
- Brocca, L., F. Melone, T. Moramarco, and R. Morbidelli (2009), Soil moisture temporal stability over experimental areas in Central Italy, *Geoderma*, 148(3-4), 364-374, doi: 10.1016/j.geoderma.2008.11.004.
- Burt, T. P., and D. P. Butcher (1985), Topographic controls of soil moisture distributions, *J. Soil Sci.*, 36, 469-486.
- Campbell, G. S. (1974), Simple method for determining unsaturated conductivity from moisture retention data, *Soil Science*, 117(6), 311-314.
- Coleman, M. L., and J. D. Niemann (2010), An evaluation of nonlinear methods for estimating catchment-scale soil moisture patterns based on topographic attributes, *Journal of Hydroinformatics*(in press).
- Cosh, M. H., T. J. Jackson, P. Starks, and G. Heathman (2006), Temporal stability of surface soil moisture in the Little Washita River watershed and its applications in satellite soil moisture product validation, *J. Hydrol.*, 323(1-4), 168-177, doi: 10.1016/j.jhydrol.2005.08.020.
- Crave, A., and C. Gascuel-Oudou (1997), Influence of topography on time and space distribution of soil surface water content, *Hydrol. Processes*, 11, 203-210.
- da Silva, A. P., A. Nadler, and B. D. Kay (2001), Factors contributing to temporal stability in spatial patterns of water content in the tillage zone, *Soil & Tillage Research*, 58(3-4), 207-218, doi: 10.1016/s0167-1987(00)00169-0.
- Dingman, S. L. (2002), *Physical Hydrology*, 2nd ed., 646 pp., Prentice Hall, Upper Saddle River, New Jersey.
- Dunne, T., and R. D. Black (1970), Partial area contributions to storm runoff in a small New England watershed, *Water Resour. Res.*, 6(5), 1296-&.
- Dunne, T., T. R. Moore, and C. H. Taylor (1975), Recognition and prediction of runoff-producing zones in humid regions, *Hydrol. Sci. Bull.*, 20, 305-327.
- Eagleson, P. S. (1978), Climate, Soil and Vegetation .3. Simplified Model of Soil-Moisture Movement in Liquid-Phase, *Water Resour. Res.*, 14(5), 722-730.
- Eichinger, W. E., M. B. Parlange, and H. Stricker (1996), On the concept of equilibrium evaporation and the value of the Priestley-Taylor coefficient, *Water Resour. Res.*, 32(1), 161-164, doi: 10.1029/95wr02920.
- Entekhabi, D., I. Rodriguez-Iturbe, and F. Castelli (1996), Mutual interaction of soil moisture state and atmospheric processes, *J. Hydrol.*, 184, 3-17.

Famiglietti, J. S., J. W. Rudnicki, and M. Rodell (1998), Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas, *J. Hydrol.*, 210, 259-281.

Gomez-Plaza, A., J. Alvarez-Rogel, J. Albaladejo, and V. M. Castillo (2000), Spatial patterns and temporal stability of soil moisture across a range of scales in a semi-arid environment, *Hydrol. Processes*, 14(7), 1261-1277, doi: 10.1002/(sici)1099-1085(200005)14:7<1261::aid-hyp40>3.0.co;2-d.

Gomez-Plaza, A., M. Martinez-Mena, J. Albaladejo, and V. M. Castillo (2001), Factors regulating spatial distribution of soil water content in small semiarid catchments, *J. Hydrol.*, 253(1-4), 211-226, doi: 10.1016/s0022-1694(01)00483-8.

Grant, L., M. Seyfried, and J. McNamara (2004), Spatial variation and temporal stability of soil water in a snow-dominated, mountain catchment, *Hydrol. Processes*, 18(18), 3493-3511, doi: 10.1002/hyp.5798.

Grayson, R. B., A. W. Western, F. H. S. Chiew, and G. Bloschl (1997), Preferred states in spatial soil moisture patterns: Local and nonlocal controls, *Water Resour. Res.*, 33(12), 2897-2908.

Grayson, R. B., and A. W. Western (1998), Towards areal estimation of soil water content from point measurements: time and space stability of mean response, *J. Hydrol.*, 207(1-2), 68-82, doi: 10.1016/s0022-1694(98)00096-1.

Heimsath, A. M., W. E. Dietrich, K. Nishiizumi, and R. C. Finkel (1999), Cosmogenic nuclides, topography, and the spatial variation of soil depth, *Geomorphology*, 27(1-2), 151-172.

Hewlett, J. D., and A. R. Hibbert (1967), Factors affecting the response of small watersheds to precipitation in humid areas, in *Forest Hydrology: Proceedings of a National Science Foundation Advanced Science Seminar Held at the Pennsylvania State University*, edited by W. E. Sopper and H. W. Lull, p. 813, Pergamon Press, New York.

Hillel, D. (1998), *Environmental Soil Physics*, 771 pp., Academic Press, New York.

Kachanoski, R. G., and E. DeJong (1988), Scale dependence and the temporal persistence of spatial patterns of soil-water storage, *Water Resour. Res.*, 24(1), 85-91, doi: 10.1029/WR024i001p00085.

Ladson, A. R., and I. D. Moore (1992), Soil-Water prediction on the Konza Prairie by microwave remote-sensing and topographic attributes, *J. Hydrol.*, 138(3-4), 385-407.

Lehman, B. M., and J. D. Niemann (2008), Spatial Patterns of in a Semi-Arid Montane Catchment with Aspect-Dependent Vegetation, paper presented at First International Conference on Hydrogeology, State College, Pennsylvania.

Lin, H. (2006), Temporal stability of soil moisture spatial pattern and subsurface preferential flow pathways in the shale hills catchment, *Vadose Zone Journal*, 5(1), 317-340, doi: 10.2136/vzj2005.0058.

- Lowry, W. D. (1959), The falling rate phase of evaporative soil-moisture loss - a critical evaluation, *Bulletin of the American Meteorological Society*, 40(12), 605-608.
- Martinez-Fernandez, J., and A. Ceballos (2005), Mean soil moisture estimation using temporal stability analysis, *J. Hydrol.*, 312(1-4), 28-38, doi: 10.1016/j.jhydrol.2005.02.007.
- Mohanty, B. P., and T. H. Skaggs (2001), Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation, *Adv. Water Resour.*, 24(9-10), 1051-1067, doi: 10.1016/s0309-1708(01)00034-3.
- Moore, I. D., G. J. Burch, and D. H. Mackenzie (1988), Topographic effects on the distribution of surface soil water and the location of ephemeral gullies, *Trans. Am. Soc. Ag. Engr.*
- Niemann, K. O., and M. C. R. Edgell (1993), Preliminary analysis of spatial and temporal distribution of soil moisture on a deforested slope, *Physical Geography*, 14(5), 449-464.
- Nyberg, L. (1996), Spatial variability of soil water content in the covered catchment at Gardsjon, Sweden, *Hydrol. Processes*, 10, 89-103.
- Penman, H. L. (1948), Natural evaporation from open water, bare soil and grass, *Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences*, 193(1032), 120-&, doi: 10.1098/rspa.1948.0037.
- Perry, M. A., and J. D. Niemann (2007), Analysis and estimation of soil moisture at the catchment scale using EOFs, *J. Hydrol.*, 334(3-4), 388-404, doi: 10.1016/j.jhydrol.2006.10.014.
- Priestley, C. H. B., and R. J. Taylor (1972), On the assessment of surface heat flux and evaporation using large-scale parameters, *Monthly Weather Review*, 100(2), 81-92.
- Shuttleworth, J. K. (1993), Evaporation, in *Handbook of Hydrology*, edited by D. R. Maidment, pp. 4.1-4.53, McGraw-Hill, Inc., New York, NY.
- Tarboton, D. G. (1997), A new method for the determination of flow directions and upslope areas in grid digital elevation models, *Water Resour. Res.*, 33(2), 309-319.
- Temimi, M., R. Leconte, N. Chaouch, P. Sukumal, R. Khanbilvardi, and F. Brissette (2010), A combination of remote sensing data and topographic attributes for the spatial and temporal monitoring of soil wetness, *J. Hydrol.*, 388(1-2), 28-40, doi: 10.1016/j.jhydrol.2010.04.021.
- Vachaud, G., A. Passerat de Silans, P. Balabanis, and M. Vauclin (1985), Temporal stability of spatially measured soil water probability density function, *Soil Sci. Soc. Am. J.*, 49(4), 822-828.
- Western, A. W., and R. B. Grayson (1998), The Tarrawarra data set: Soil moisture patterns, soil characteristics, and hydrological flux measurements, *Water Resour. Res.*, 34(10), 2765-2768.
- Western, A. W., R. B. Grayson, G. Bloschl, G. R. Willgoose, and T. A. McMahon (1999a), Observed spatial organization of soil moisture and its relation to terrain indices, *Water Resour. Res.*, 35(3), 797-810.

Western, A. W., R. B. Grayson, and T. R. Green (1999b), The Tarrawarra project: high resolution spatial measurement, modelling and analysis of soil moisture and hydrological response, *Hydrol. Processes*, 13, 633-652.

Western, A. W., G. Bloschl, and R. B. Grayson (2001), Toward capturing hydrologically significant connectivity in spatial patterns, *Water Resour. Res.*, 37(1), 83-97.

Wilson, D. J., A. W. Western, R. B. Grayson, A. A. Berg, M. S. Lear, M. Rodell, J. S. Famiglietti, R. A. Woods, and T. A. McMahon (2003), Spatial distribution of soil moisture over 6 and 30 cm depth, Mahurangi river catchment, New Zealand, *J. Hydrol.*, 276(1-4), 254-274.

Wilson, D. J., A. W. Western, and R. B. Grayson (2005), A terrain and data-based method for generating the spatial distribution of soil moisture, *Adv. Water Resour.*, 28(1), 43-54.

Woods, R. A., R. B. Grayson, A. W. Western, M. Duncan, D. J. Wilson, R. I. Young, R. Ibbitt, R. Henderson, and T. A. McMahon (2001), Experimental design and initial results from the Mahurangi River Variability Experiment: MARVEX, in *Land Surface Hydrology, Meteorology, and Climate: Observations and Modeling, Water Science and Application*, edited by V. Lakshmi, et al., pp. 201-213, American Geophysical Union, Washington, D.C.

## **APPENDIX: EMT Model Derivation**

The objective is to develop a model for the soil moisture at any location within a catchment if the spatial-average soil moisture is known. To obtain this estimate, the soil moisture is modeled using a steady-state approximation for the watershed. Specifically, all processes are assumed to occur steadily and the soil moisture is assumed to be at equilibrium so that the total inflow to the soil balances the total outflow from the soil. The only processes that are included to transport water to or from the soil are infiltration  $F$ , deep drainage or recharge to groundwater  $G$ , lateral unsaturated flow  $L$ , and evapotranspiration  $E$ . Thus, one can write the water balance for a given segment of an elevation contour as:

$$\int_A F dA = \int_A G dA + L + \int_A E dA \quad (\text{A.1})$$

where  $A$  is the land area that can contribute flow to the length of an elevation contour and  $F$ ,  $G$  and  $E$  are the infiltration, recharge, and evapotranspiration rates within the area that can contribute flow through the given length of contour.  $L$  is the lateral outflow occurring through the length of the contour.

The deep drainage is modeled as percolation (i.e. the hydraulic gradient is controlled only by gravity meaning this flux can be written as:

$$G = K_v \quad (\text{A.2})$$

where  $K_v$  is the unsaturated vertical hydraulic conductivity. Using the Campbell equation [Campbell, 1974] to describe unsaturated hydraulic conductivity,  $K_v$  can be written:

$$K_v = K_{s,v} \left( \frac{\theta}{\phi} \right)^{\gamma_v} \quad (\text{A.3})$$

where  $K_{s,v}$  is the vertical saturated hydraulic conductivity,  $\theta$  is the volumetric soil moisture,  $\phi$  is the porosity, and  $\gamma$  is the pore disconnectedness index. This implies

$$G = K_{s,v} \left( \frac{\theta}{\phi} \right)^{\gamma_v} \quad (\text{A.4})$$

Lateral flow is described using Darcy's Law and assuming uniform flow with depth in the soil.

One can write the lateral flow as:

$$L = -\delta c K_h \frac{dh}{dx} \quad (\text{A.5})$$

where  $\delta$  is the depth within which lateral flow occurs (soil depth),  $c$  is the contour width being considered,  $K_h$  is the unsaturated horizontal hydraulic conductivity, and  $dh/dx$  is the hydraulic gradient in the horizontal direction. Data in *Heimsath et al.* [1999] indicates that soil depth depends linearly on the topographic curvature:

$$\delta = \delta_0 (\kappa_{\min} - \kappa) / \kappa_{\min} \quad (\text{A.6})$$

where  $\delta_0$  is the soil depth for a location with a curvature of zero and  $\kappa_{\min}$  is the minimum curvature for which soil is present. Note that curvature is defined here as positive for convergent locations and negative for divergent locations. Also, the horizontal hydraulic gradient is taken as a function of the topographic slope, in this case a power function. Using the expression for soil depth, the gradient relationship to topographic slope and the Campbell equation again for unsaturated hydraulic conductivity, the lateral flow can be written:

$$L = \delta_0 \left( \frac{\kappa_{\min} - \kappa}{\kappa_{\min}} \right) c K_{s,h} \left( \frac{\theta}{\phi} \right)^{\gamma_h} S^\epsilon \quad (\text{A.7})$$

where  $K_{s,h}$  is the horizontal saturated hydraulic conductivity,  $\gamma_h$  is the pore size distribution index in the horizontal direction,  $\varepsilon$  is the exponent relating horizontal hydraulic gradient to topographic slope and the horizontal hydraulic gradient is in the opposite direction as the topographic slope.

Evapotranspiration is determined as a function of a local potential evapotranspiration (PET), which is denoted  $E_{p,l}$ . The local PET is calculated from the Penman combination equation:

$$E_{p,l} = \left[ \frac{\Delta}{\Delta + \Gamma} E_r I_p + \frac{\Gamma}{\Delta + \Gamma} E_a \right] \quad (\text{A.8})$$

where  $\Delta$  is rate of change of saturation vapor pressure with air temperature,  $\Gamma$  is the psychrometric constant,  $E_r$  is the energy balance estimate of ET, and  $E_a$  is the aerodynamic estimate.  $I_p$  is the potential solar radiation index, which is the ratio between the insolation of the topographic surface relative to that of a horizontal surface at the same location [Dingman, 2002]. Similarly to the Priestley-Taylor equation, the aerodynamic term is assumed to be a specified fraction,  $\alpha$ , of the energy balance estimate and has values in the range [0, 0.74] [Eichinger et al., 1996; Shuttleworth, 1993]. This implies:

$$E_{p,l} = \left[ \frac{\Delta}{\Delta + \Gamma} E_r I_p + \alpha \frac{\Gamma}{\Delta + \Gamma} E_r \right] \quad (\text{A.9})$$

Note that  $\alpha$  as used here has a value that is 1 less than the coefficient as typically used in the Priestley-Taylor equation. PET can also be calculated similarly for the entire watershed (denoted  $E_p$ ) although over large areas the impact of the topography is neglected implying the following form for regional PET:

$$E_p = \left[ \frac{\Delta}{\Delta + \Gamma} E_r + \alpha \frac{\Gamma}{\Delta + \Gamma} E_r \right] \quad (\text{A.10})$$

Combining equations (A.9) and (A.10), one can write the local PET in terms of the regional PET:

$$E_{p,l} = \frac{E_p I_p}{1 + \alpha} + \frac{E_p \alpha}{1 + \alpha} \quad (\text{A.11})$$

The actual evapotranspiration  $E$  is then calculated by accounting for moisture limitations:

$$E = \frac{E_p I_p}{1 + \alpha} \left( \frac{\theta}{\phi} \right)^{\beta_r} + \frac{E_p \alpha}{1 + \alpha} \left( \frac{\theta}{\phi} \right)^{\beta_a} \quad (\text{A.12})$$

where  $\beta_r$  and  $\beta_a$  are vegetation parameters. In the literature, a single  $\beta$  is applied in a model for evapotranspiration, and  $\beta > 1$  represents bare soil, and  $\beta < 1$  represents vegetated soils. Here, we allow distinct exponents because moisture limitations may affect the energy and aerodynamic estimates in different ways.

We now substitute the expressions for  $G$ ,  $L$ , and  $E$  from equations (A.4), (A.7), and (A.12) back into equation (A.1), which produces:

$$\int_A F dA = \frac{K_{s,v}}{\phi^{\gamma_v}} \int_A \theta^{\gamma_v} dA - \frac{\delta_0 (\kappa_{\min} - \kappa) K_{s,h}}{\kappa_{\min} \phi^{\gamma_h}} c S^\varepsilon \theta^{\gamma_h} + \frac{E_p}{(1 + \alpha) \phi^{\beta_r}} \int_A I_p \theta^{\beta_r} dA + \frac{E_p \alpha}{(1 + \alpha) \phi^{\beta_a}} \int_A \theta^{\beta_a} dA \quad (\text{A.13})$$

In this expression, we have assumed that  $\phi$ ,  $K_{s,v}$ ,  $K_{s,h}$ ,  $\gamma_v$ ,  $\gamma_h$ ,  $\delta_0$ ,  $\kappa_{\min}$ ,  $\varepsilon$ ,  $\alpha$ ,  $\beta_r$ ,  $\beta_a$ , and  $E_p$  are all spatially constant. Similar to TopModel, the infiltration rate is also assumed to be spatially constant, and the equation becomes:

$$F = \frac{K_{s,v}}{\phi^{\gamma_v} A} \int_A \theta^{\gamma_v} dA - \frac{\delta_0 (\kappa_{\min} - \kappa) K_{s,h}}{\kappa_{\min} \phi^{\gamma_h} A} c S^\varepsilon \theta^{\gamma_h} + \frac{E_p}{(1 + \alpha) \phi^{\beta_r} A} \int_A I_p \theta^{\beta_r} dA + \frac{E_p \alpha}{(1 + \alpha) \phi^{\beta_a} A} \int_A \theta^{\beta_a} dA \quad (\text{A.14})$$

We now consider four simplified cases. In each of these cases, one of the four terms on the right (deep drainage, lateral outflow, radiation-based ET or humidity-based ET) is assumed to be much more important than the others. In the case that deep drainage dominates, equation (A.14) can be approximated:

$$F = \frac{K_{s,v}}{\phi^{\gamma_v}} \int_A \theta^{\gamma_v} dA \quad (\text{A.15})$$

In this case, water movement between locations is negligible. Thus, the infiltration that occurs at every point in the catchment must balance the deep drainage at every point. Therefore, the water balance equation can be written locally as:

$$F = \frac{K_{s,v}}{\phi^{\gamma_v}} \theta^{\gamma_v} \quad (\text{A.16})$$

or:

$$\theta = \phi \left( \frac{F}{K_{s,v}} \right)^{1/\gamma_v} \quad (\text{A.17})$$

Our initial objective was to calculate the soil moisture as a function of an available spatial-average soil moisture  $\bar{\theta}$ . We can calculate the spatial average soil moisture  $\bar{\theta}$  by integrating the equation (A.17) over the region for which  $\bar{\theta}$  is available. If  $A_c$  is defined as the area of this region, the spatial average is calculated:

$$\bar{\theta} = \frac{1}{A_c} \int_{A_c} \theta dA_c \quad (\text{A.18})$$

Substituting equation (A.17) into equation (A.18) and noting that all of the terms that determine  $\theta$  are spatially constant, one obtains:

$$\bar{\theta} = \phi \left( \frac{F}{K_{s,v}} \right)^{1/\gamma_v} \quad (\text{A.19})$$

Comparing this expression with equation (A.17), one observes that:

$$\theta = \bar{\theta} \quad (\text{A.20})$$

This expression implies that if deep drainage is the dominant process in determining the spatial pattern of the effective soil moisture, then the effective soil moisture is spatially uniform.

We now consider the second case where lateral drainage is the dominant outflow process. In this case, equation (A.14) simplifies to:

$$F = \frac{\delta_0 (\kappa_{\min} - \kappa) K_{s,h}}{\kappa_{\min} \phi^{\gamma_h} A} c S^\varepsilon \theta^{\gamma_h} \quad (\text{A.21})$$

Solving this expression for  $\theta$ , one obtains:

$$\theta = \phi \left( \frac{F}{\delta_0 K_{s,h}} \right)^{1/\gamma_h} \left( \frac{A}{c S^\varepsilon} \right)^{1/\gamma_h} \left( \frac{\kappa_{\min}}{\kappa_{\min} - \kappa} \right)^{1/\gamma_h} \quad (\text{A.22})$$

Integrating equation (A.22) over the region for which a spatial average is available, one obtains:

$$\bar{\theta} = \frac{1}{A_c} \int_{A_c} \theta dA_c = \phi \left( \frac{F}{\delta_0 K_{s,h}} \right)^{1/\gamma_h} \frac{1}{A_c} \int_{A_c} \left( \frac{A}{c S^\varepsilon} \right)^{1/\gamma_h} \left( \frac{\kappa_{\min}}{\kappa_{\min} - \kappa} \right)^{1/\gamma_h} dA_c \quad (\text{A.23})$$

The term that remains in the integral  $(A / c S^\varepsilon)^{1/\gamma_h} [\kappa_{\min} / (\kappa - \kappa_{\min})]^{1/\gamma_h}$  is a compound topographic index, which we call the lateral flow index (LFI). The average LFI within the region is denoted  $\Lambda$ , and it is simply:

$$\Lambda = \frac{1}{A_c} \int_{A_c} \left( \frac{A}{cS} \right)^{1/\gamma_h} \left( \frac{\kappa_{\min}}{\kappa_{\min} - \kappa} \right)^{1/\gamma_h} dA_c \quad (\text{A.24})$$

Substituting this into equation (A.23), one obtains:

$$\bar{\theta} = \phi \left( \frac{F}{\delta_0 K_{s,h}} \right)^{1/\gamma_h} \Lambda \quad (\text{A.25})$$

Combining equations (A.22) and (A.25), one obtains:

$$\theta = \frac{\bar{\theta}}{\Lambda} \left( \frac{A}{cS^\varepsilon} \right)^{1/\gamma_h} \left( \frac{\kappa_{\min}}{\kappa_{\min} - \kappa} \right)^{1/\gamma_h} \quad (\text{A.26})$$

This expression suggests that the soil moisture depends on the spatial average soil moisture and the LFI if lateral flow dominates.

Next, we consider the case where the radiation-based ET term in equation (A.14). One can simplify the equation as:

$$F = \frac{E_p}{(1 + \alpha)\phi^{\beta_r} A} \int_A I_p \theta^{\beta_r} dA \quad (\text{A.27})$$

Again noting that water movement between locations is negligible, one can rewrite equation (A.27) as a local water balance:

$$F = \frac{E_p}{(1 + \alpha)\phi^{\beta_r}} I_p \theta^{\beta_r} \quad (\text{A.28})$$

And solving for  $\theta$  one obtains:

$$\theta \approx \phi \left( \frac{F(1+\alpha)}{E_p} \right)^{1/\beta_r} \left( \frac{1}{I_p} \right)^{1/\beta_r} \quad (\text{A.29})$$

Integrating over the region to calculate the average soil moisture, one gets:

$$\bar{\theta} = \frac{1}{A_c} \int \theta dA_c = \phi \left( \frac{F(1+\alpha)}{E_p} \right)^{1/\beta_r} \frac{1}{A_c} \int \left( \frac{1}{I_p} \right)^{1/\beta_r} dA_c = \phi \left( \frac{F(1+\alpha)}{E_p} \right)^{1/\beta_r} \Pi \quad (\text{A.30})$$

where  $\Pi$  is the average of an evapotranspiration index (ETI), which is defined as  $(1/I_p)^{1/\beta_r}$ .

Combining equations (A.29) and (A.30), one gets:

$$\theta = \frac{\bar{\theta}}{\Pi} \left( \frac{1}{I_p} \right)^{1/\beta_r} \quad (\text{A.31})$$

Finally, we consider the case where the humidity-related ET term dominates. In this case, equation (A.14) can be written:

$$F = \frac{E_p \alpha}{(1+\alpha)\phi^{\beta_a} A} \int \theta^{\beta_a} dA \quad (\text{A.32})$$

Again, noting that water movement between locations is negligible, one can rewrite equation (A.32) as a local water balance:

$$F = \frac{E_p \alpha}{(1+\alpha)\phi^{\beta_a}} \theta^{\beta_a} \quad (\text{A.33})$$

And solving for  $\theta$  one obtains:

$$\theta \approx \phi \left( \frac{F(1+\alpha)}{E_p \alpha} \right)^{1/\beta_a} \quad (\text{A.34})$$

Integrating over the region to calculate the average soil moisture, one gets:

$$\bar{\theta} = \frac{1}{A_c} \int_{A_c} \theta dA_c = \phi \left( \frac{F(1+\alpha)}{E_p \alpha} \right)^{1/\beta_a} \quad (\text{A.35})$$

Combining equations (A.34) and (A.35) one gets:

$$\theta = \bar{\theta} \quad (\text{A.36})$$

Four estimates for the soil moisture have been produced depending on whether deep drainage, lateral flow, radiation-related ET, or humidity-related ET is the dominant process:

$$\theta_G = \bar{\theta} \quad (\text{A.37})$$

and:

$$\theta_L = \frac{\bar{\theta}}{\Lambda} \left( \frac{A}{cS^\varepsilon} \right)^{1/\gamma_h} \left( \frac{\kappa_{\min}}{\kappa_{\min} - \kappa} \right)^{1/\gamma_h} \quad (\text{A.38})$$

and:

$$\theta_R = \frac{\bar{\theta}}{\Pi} \left( \frac{1}{I_p} \right)^{1/\beta_r} \quad (\text{A.39})$$

and:

$$\theta_A = \bar{\theta} \quad (\text{A.40})$$

A final estimate for  $\theta$  is obtained using a weighted arithmetic average of these four approximations:

$$\theta = \frac{w_G \theta_G + w_L \theta_L + w_R \theta_R + w_A \theta_A}{w_G + w_L + w_R + w_A} \quad (\text{A.41})$$

where  $w_G$ ,  $w_L$ ,  $w_R$ ,  $w_A$  are the weights. One expects the weights to depend on the magnitude of the terms that represent the processes in equation (A.14). In particular, the weight for the deep drainage estimate  $w_G$  should be large if the deep drainage term in equation (A.14) is large. Thus, it is proposed that:

$$w_G = \frac{K_{s,v}}{\phi^{\gamma_v} A} \int_A \theta^{\gamma_v} dA \quad (\text{A.42})$$

but this equation is not easily evaluated because it includes the average soil moisture for the area upslope from the location of interest. If we approximate this average with the local effective soil moisture, then this equation becomes:

$$w_G = \frac{K_{s,v}}{\phi^{\gamma}} \theta^{\gamma_v} \quad (\text{A.43})$$

Next, the estimate of  $\theta_G$  obtained by assuming deep drainage was dominant is then substituted into equation (A.43) to give:

$$w_G = \frac{K_{s,v}}{\phi^{\gamma}} \bar{\theta}^{\gamma_v} \quad (\text{A.44})$$

Similarly, the weight for the lateral flow estimate can be estimated from the associated term in equation (A.14) to arrive at:

$$w_L = \frac{\delta_0 (\kappa_{\min} - \kappa) K_{s,h}}{\kappa_{\min} \phi^{\gamma_h} A} c S^{\epsilon} \theta^{\gamma_h} \quad (\text{A.45})$$

and then equation (A.38) is substituted into equation (A.45) to obtain:

$$w_L = \frac{\delta_0 (\kappa_{\min} - \kappa) K_{s,h}}{\kappa_{\min} \phi^{\gamma_h} A} c S^\varepsilon \left( \frac{\bar{\theta}}{\Lambda} \right)^{\gamma_h} \left( \frac{A}{c S^\varepsilon} \right) \left( \frac{\kappa_{\min}}{\kappa_{\min} - \kappa} \right) \quad (\text{A.46})$$

The weight for the radiation-based ET is:

$$w_R = \frac{E_p}{(1 + \alpha) \phi^{\beta_r} A} \int_A I_p \theta^{\beta_r} dA \quad (\text{A.47})$$

Notice that this expression includes the spatial average of  $I_p \theta^{\beta_r}$  for the region upslope of each point. Again, approximating this average with the local value, one obtains:

$$w_R = \frac{E_p I_p}{(1 + \alpha) \phi^{\beta_r}} \theta^{\beta_r} \quad (\text{A.48})$$

and substituting equation (A.39) into equation (A.48) leads to:

$$w_R = \frac{E_p I_p}{(1 + \alpha) \phi^{\beta_r}} \left( \frac{\bar{\theta}}{\Pi} \right)^{\beta_r} \left( \frac{1}{I_p} \right) \quad (\text{A.49})$$

Finally, the weight for the humidity-related ET is:

$$w_A = \frac{E_p \alpha}{(1 + \alpha) \phi^{\beta_a} A} \int_A \theta^{\beta_a} dA \quad (\text{A.50})$$

Writing this locally and replacing the average with the local value, one obtains:

$$w_A = \frac{E_p \alpha}{(1 + \alpha) \phi^{\beta_a}} \theta^{\beta_a} \quad (\text{A.51})$$

And equation (A.40) is substituted to give:

$$w_A = \frac{E_p \alpha}{(1 + \alpha) \phi^{\beta_a}} \bar{\theta}^{\beta_a} \quad (\text{A.52})$$

Substituting equations (A.37), (A.38), (A.39), (A.40), (A.44), (A.46), (A.49) and (A.52) into equation (A.41) one obtains a combined estimate:

$$\theta = \frac{\frac{K_{s,v}}{\phi^{\gamma_v}} \bar{\theta}^{\gamma_v} \bar{\theta} + \frac{\delta_0(\kappa_{\min} - \kappa) K_{s,h} c S^\varepsilon}{\kappa_{\min} \phi^{\gamma_h} A} \left(\frac{\bar{\theta}}{\Lambda}\right)^{\gamma_h} \left(\frac{A}{c S^\varepsilon}\right) \left(\frac{\kappa_{\min}}{\kappa_{\min} - \kappa}\right) \bar{\theta} \left(\frac{A}{c S^\varepsilon}\right)^{1/\gamma_h} \left(\frac{\kappa_{\min}}{\kappa_{\min} - \kappa}\right)^{1/\gamma_h} + \frac{E_p I_p}{(1+\alpha)\phi^{\beta_r}} \left(\frac{\bar{\theta}}{\Pi}\right)^{\beta_r} \left(\frac{1}{I_p}\right) \bar{\theta} \left(\frac{1}{I_p}\right)^{1/\beta_r} + \frac{E_p \alpha}{(1+\alpha)\phi^{\beta_a}} \bar{\theta}^{\beta_a} \bar{\theta}}{\frac{K_{s,v}}{\phi^{\gamma_v}} \bar{\theta}^{\gamma_v} + \frac{\delta_0(\kappa_{\min} - \kappa) K_{s,h} c S^\varepsilon}{\kappa_{\min} \phi^{\gamma_h} A} \left(\frac{\bar{\theta}}{\Lambda}\right)^{\gamma_h} \left(\frac{A}{c S^\varepsilon}\right) \left(\frac{\kappa_{\min}}{\kappa_{\min} - \kappa}\right) + \frac{E_p I_p}{(1+\alpha)\phi^{\beta_r}} \left(\frac{\bar{\theta}}{\Pi}\right)^{\beta_r} \left(\frac{1}{I_p}\right) + \frac{E_p \alpha}{(1+\alpha)\phi^{\beta_a}} \bar{\theta}^{\beta_a}} \quad (\text{A.53})$$

and simplifying one gets:

$$\theta = \frac{K_{s,v} \left(\frac{\bar{\theta}}{\phi}\right)^{\gamma_v} \bar{\theta} + \frac{\delta_0 K_{s,h}}{\Lambda^{\gamma_h}} \left(\frac{\bar{\theta}}{\phi}\right)^{\gamma_h} \left[\frac{\bar{\theta}}{\Lambda} \left(\frac{A}{c S^\varepsilon}\right)^{1/\gamma_h} \left(\frac{\kappa_{\min}}{\kappa_{\min} - \kappa}\right)^{1/\gamma_h}\right] + \frac{E_p}{(1+\alpha)\Pi^{\beta_r}} \left(\frac{\bar{\theta}}{\phi}\right)^{\beta_r} \left[\frac{\bar{\theta}}{\Pi} \left(\frac{1}{I_p}\right)^{1/\beta_r}\right] + \frac{E_p \alpha}{(1+\alpha)} \left(\frac{\bar{\theta}}{\phi}\right)^{\beta_a} \bar{\theta}}{K_{s,v} \left(\frac{\bar{\theta}}{\phi}\right)^{\gamma_v} + \frac{\delta_0 K_{s,h}}{\Lambda^{\gamma_h}} \left(\frac{\bar{\theta}}{\phi}\right)^{\gamma_h} + \frac{E_p}{(1+\alpha)\Pi^{\beta_r}} \left(\frac{\bar{\theta}}{\phi}\right)^{\beta_r} + \frac{E_p \alpha}{(1+\alpha)} \left(\frac{\bar{\theta}}{\phi}\right)^{\beta_a}} \quad (\text{A.54})$$

which provides an estimate of the local soil moisture as a function of the spatial average soil moisture  $\bar{\theta}$ , two topographic indices (the lateral flow index and the evapotranspiration index), and various climate, soil, and vegetation parameters.

### **Part III: Reconstruction of Hillslope and Valley Paleotopography by Application of a Geomorphic Model**

## Introduction

Determining the location and thickness of lithologic layers is an important task for many applications including mineral and groundwater recovery as well as environmental monitoring and remediation. Practical limitations typically allow observation of layer boundaries only at sparse locations, so boundary elevations are often estimated between observations by mathematical techniques [Jones *et al.*, 1986]. Standard estimation methods such as splines (Briggs, 1974; Hofierka and Cebecauer, 2007; Mitasova and Mitas, 1993; Mitasova *et al.*, 1995; Talmi and Gilat, 1977), inverse-distance weighting [Bartier and Keller, 1996; Chang *et al.*, 2005; Watson and Philip, 1985; Zoraster, 1996], nearest neighbor [Kelley and Kelley, 2004] and geostatistical methods [Chiles and Delfiner, 1999; Journel and Huijbregts, 1978; Stein, 1999], produce smooth interpolation surfaces. However, if the boundary of interest is a paleotopography, then it may exhibit roughness due to drainage patterns preserved in the surface. Fractal-based interpolation methods are also available [Bindlish and Barros, 1996; Bouboulis *et al.*, 2006; Polidori and Chorowicz, 1993; Yokoya *et al.*, 1989], but they do not reproduce the drainage networks observed in most topography.

An alternative approach is to simulate the layer boundary by representing the processes that produced it. Previously, large-scale paleotopographies have been reconstructed using geochemical methods [Craig, 1993; Crowley and Burke, 1998] and used as boundary conditions in atmospheric models to simulate paleo-climate conditions. Also, on a smaller scale historical river channels, which are known to influence the movement of groundwater and its associated constituents, have been delineated or corroborated with a variety of geophysical techniques [Holschuh, 2002; Miller *et al.*, 1999; Sandberg *et al.*, 2001; Williams *et al.*, 2002]. In this paper, we develop a method to simulate a paleotopography at a scale intermediate between the two mentioned above. The surface is the lower boundary of the Tshirege Member of the

Bandelier Tuff, which underlies Los Alamos National Laboratory (LANL) and Bandelier National Monument (BNM) in north-central New Mexico. This boundary may play a hydrologic role by perching water in certain locations below the Pajarito Plateau [*Broxton and Reneau, 1996*], but it is not currently considered to play a major role in regional groundwater movement. The boundary was formed predominantly by fluvial and hillslope processes acting on an ash-flow surface, which was then preserved by a second ash-flow eruption. Similar ash-flows have been observed to produce new, relatively smooth surfaces if the flow has sufficient volume and energy to cover the preexisting landscape features [*Chapin and Elston, 1979; Francis, 1993; Ollier, 1988; Smith, 1960*]. Furthermore, ash-flows can leave the buried landscape relatively intact [*Fisher and Schmincke, 1984; Reneau, 2000; Smith and Bailey, 1966*]. We aim to simulate the boundary of interest based on elevations available at sparse wells and contacts exposed in modern day canyons. The surface will be interpolated by generalizing a topographic interpolation method that incorporates fluvial and hillslope processes to infer drainage patterns from elevation data [*Grimaldi et al., 2004; Grimaldi et al., 2007; Niemann et al., 2003*]. The original method was applied to current topography and requires inclusion of all headwater land areas inside the interpolation region. In this paper, the method's boundary conditions are generalized to overcome the headwater limitation. Field data are used to calibrate the method for this site, and the generalized method is applied to estimate the likelihoods that each location on the paleotopography drains south roughly from LANL toward BNM. Section 2 briefly describes the study area and the paleotopography of interest, and Section 3 gives an overview of the reconstruction method. Section 4 describes the methods used to estimate the parameters and the boundary conditions of the method. The results and conclusions are discussed in Sections 5 and 6, respectively.

## Study Site

The Pajarito Plateau is located in north-central New Mexico and occupies the western side of the Española Basin portion of the Rio Grande rift [*Broxton and Vaniman, 2005; Newman and Robinson, 2005*]. The plateau is bounded on the western side by the Jemez Mountains and on the eastern side by the Rio Grande valley. The modern landscape is one of mesas and canyons lying generally in an east-west direction. Elevations on the mesas vary from approximately 1,900 m near the Rio Grande to approximately 2,400 m near the Jemez Mountains. The relatively large elevation range over the plateau leads to an approximate annual precipitation range of 35 cm to 50 cm with higher precipitation on the western portions of the plateau [*Newman and Robinson, 2005*]. The plateau is the site of the towns of Los Alamos and White Rock as well as LANL and BNM.

The upper stratigraphy of the Pajarito Plateau in the Los Alamos area is shown in Figure 16, which is simplified from Broxton and Vaniman [2005]. The Quaternary-dated Bandelier Tuff is composed of two ash-flow units, the Otowi Member and the Tshirege Member, resulting from separate volcanic events that occurred approximately 1.61 Ma and 1.22 Ma [*Broxton and Reneau, 1996*], respectively. The Otowi Member varies in thickness from 0 m in the eastern portion of the plateau to 125 m in some locations on the western portion of the plateau. The Tshirege Member varies from approximately 60 m thick in the north-central portion of LANL to >180 m thick in the southern portion [*Broxton and Vaniman, 2005*]. The two units are separated by the Cerro Toledo interval, which consists of pumice and ash fall material as well as eroded sediments from the Otowi eruption of the Valles Caldera and other volcanic activity in the area [*Broxton and Vaniman, 2005*].

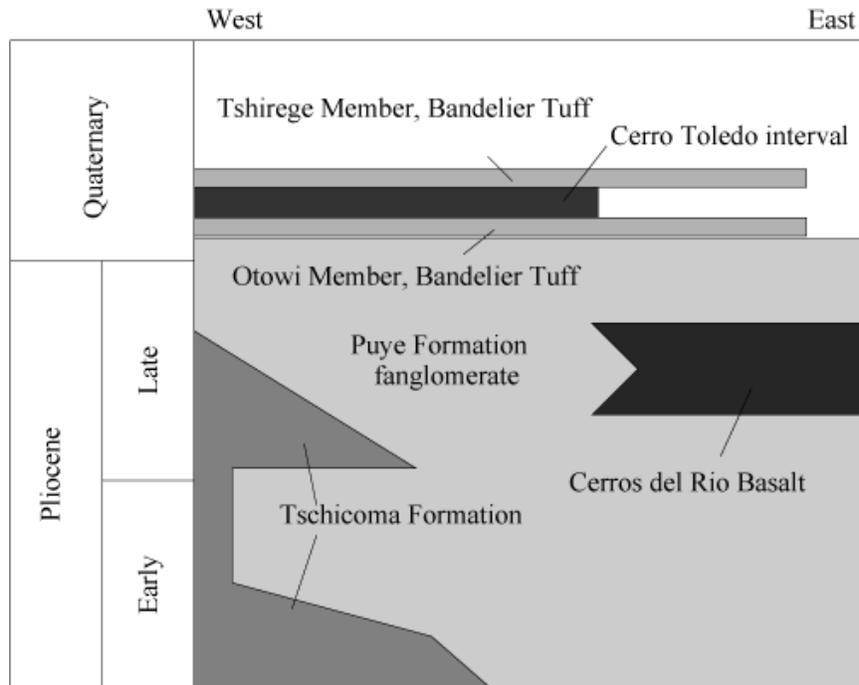


Figure 16 - Simplified upper stratigraphy of Pajarito Plateau (after Broxton and Vaniman [2005]). Western limit in figure is beginning of Jemez Mountains and eastern limit is White Rock Canyon.

We are interested in the lower boundary of the Tshirege Member of the Bandelier Tuff. The emplacement of the Otowi Member by ash flow processes effectively filled in preexisting valleys and created a new surface dipping to the east and southeast [Broxton and Reneau, 1996]. This layer was in place and exposed to fluvial and hillslope erosive processes for approximately 400 ka prior to the emplacement of the Tshirege Member [Izett and Obradovich, 1994; Spell et al., 1996]. Because the Otowi Member is non-welded, the landscape formed during the period between eruptions probably was not similar to the current mesa-canyon landscape but instead consisted of shallower stream valleys and lower overall relief [Broxton and Reneau, 1996]. Broxton and Reneau [1996] used elevation contours for the Tshirege Member base and lithologic characteristics of the underlying alluvium in the Cerro Toledo interval as evidence to conclude that at least four major drainages existed on the landscape prior to the emplacement of the Tshirege Member. The presence of these drainages and the alluvial deposits indicates the

importance of fluvial processes in shaping the topography during this interval. Exposed contacts, such as the one shown in Figure 17(a), indicate that major valleys had narrow floors, which suggests the fluvial processes during the period immediately preceding Tshirege emplacement were commonly erosive rather than depositional. In addition, the crudely parabolic hillslope shapes of exposed contacts like the one in Figure 17(b) suggest that diffusive hillslope processes played a role in the surface formation. The presence of these features in the stratigraphy also suggests that the subsequent emplacement of the Tshirege Member preserved many of the valleys and hillslopes developed during the 400 ka of exposure.

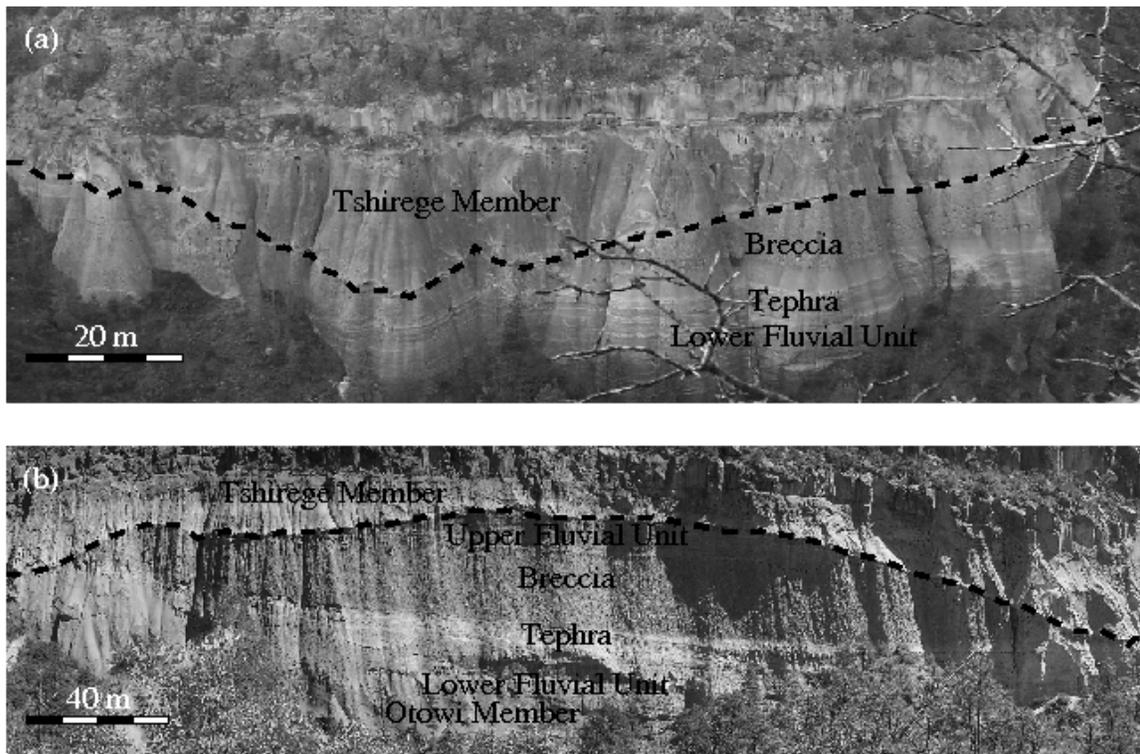


Figure 17 - Exposed (a) valley and (b) hillslope cross-sections in lower boundary of Tshirege Member in Alamo Canyon, BNM. Scales shown are approximate.

## Method Description

This section summarizes the main elements of the method used to simulate the lower boundary of the Tshirege Member including the required parameters. Niemann et al. [2003] and Grimaldi

et al. [2004] provide more detailed descriptions. The first step is to interpolate the surface between the points with known elevations using a completely regularized spline [Mitasova and Mitas, 1993]. The spline produces a smooth surface that passes through the observations, but the surface does not exhibit realistic valley patterns. The spline surface is a first estimate of the paleotopography and provides the initial condition for the geomorphic model used in the second step. The spline requires specification of a tension parameter  $\phi$ , which controls how loosely the surface is draped on the observations. A small random offset in elevation is added to each location on the spline surface (aside from the locations with observed elevations) to help reduce the number of artificially straight drainage patterns produced by the model. Because the landscape evolution model is nonlinear, simulations with different offsets can be significantly different. Any simulation thus represents one potentially feasible surface.

The second step is to modify the spline surface to simulate the effects of the relevant geomorphic processes. This step applies a numerical model to simulate the effects of the processes on the spline surface and uses the resulting differences between simulated and observed elevations to update the estimate of an erodability parameter  $\beta$ , which can vary spatially. The two parts of the step must be iterated to bring the simulated surface close to the observations because changes in  $\beta$  can have unanticipated results as the simulated drainage patterns change. Due to the limited knowledge about the paleotopography, the numerical model includes simple detachment-limited fluvial erosion [Howard, 1994; Moglen and Bras, 1995] and slope-dependent hillslope processes [Kirkby, 1971]. Mathematically, the starting point is:

$$\frac{\partial z}{\partial t} = U - \beta A^m S^n + D \nabla^2 z \quad (3.1)$$

where  $z$  [L] is the elevation of a point in the region,  $t$  [T] is time,  $U$  [L/T] is the base level lowering rate or relative uplift rate,  $\beta$  [L<sup>-2m+1</sup>/T] is the erodability of the point,  $A$  [L<sup>2</sup>] is the

point's contributing area (the land area that can contribute flow to the point),  $S$  [L/L] is the local channel slope,  $m$  and  $n$  are dimensionless fluvial parameters,  $D$  [L<sup>2</sup>/T] is a diffusivity parameter indicating the strength of hillslope processes, and  $\nabla^2 z$  [1/L] denotes the Laplace curvature of the surface. Because no specific transient condition is known, it is assumed that the topography was approximately at equilibrium after the 400 ka of exposure (deviations from this assumption are interpreted by the algorithm as variations in  $\beta$ ). From the equilibrium assumption,  $\partial z / \partial t = 0$ , and  $U = 1$  and  $n = 1$  can be assumed without loss of generality. Implementation of the model requires one to specify the parameters  $\beta$ ,  $m$ , and  $D$  as well as the boundary conditions. The parameter  $\beta$  controls the relief,  $m$  controls the concavity of the long profiles of the channels, and  $D$  controls the hillslope size. The original method allowed flow to exit at all boundaries but did not allow it to enter the region, which implicitly requires the interpolation region to contain all headwater locations of basins. That requirement is expected to be problematic for most paleotopography applications because data typically do not allow reliable delineation of basin boundaries. The generalized method allows individual boundaries to be either open to outgoing flow or closed to outgoing flow but with automatically-identified inlet points. Inlets points occur at any low point along a closed boundary on the initial spline-based surface. The drainage area associated with an inlet is determined from the local slope of the initial surface and an assumed relationship between the slope and the incoming drainage area. The determination of all parameter values including those used to determine the inlet areas is explained in the next section.

After the numerical model has been used to modify the initial surface, the parameter  $\beta$  is updated using a Bayesian procedure to bring the next simulated surface closer to the observed elevations. Although the Otowi member is relatively homogeneous at the outcrop scale,  $\beta$  is allowed to vary

spatially over the region. These variations might be real variations in the material erodability, but they could also represent deviations from the model assumptions described earlier. To reduce the number of  $\beta$  values that must be estimated based on the observed elevations, the spatial field of  $\beta$  is determined from a completely-regularized spline of  $\beta$  values at the points with observed elevations. This spline uses the same tension as the spline used with the elevation data. The estimation of  $\beta$  requires specification of an initial spatially-constant estimate of  $\beta$ ,  $\bar{\beta}$ , the ratio of the error variance to the variance in spline basis function weights,  $\nu$ , and a measure of correlation between the errors in sequential simulations,  $\rho$ . In practical terms,  $\nu$  controls how much the erodability is changed after each simulation, and  $\rho$  controls the number of iterations before the estimation algorithm terminates [Niemann *et al.*, 2003].

It is worth mentioning that the original method [Niemann *et al.*, 2003] incorporated a third step that distorted the final surface produced by the iterative procedure to ensure the simulated elevations exactly match the observed elevations while still preserving the draining characteristic of the surface. The modified method presented here does not incorporate that third step because the elevations obtained from wells and outcrops only provide estimates of the average elevation within each pixel. Thus, the final interpolation surface is allowed to deviate from these values.

## **Method Application**

In order to apply the method to reconstruct the paleotopography, the interpolation parameters  $\phi$ ,  $\nu$ , and  $\rho$  and the geomorphic parameters including  $D$ ,  $m$ , and  $\bar{\beta}$  must be estimated. For this application, the interpolation parameters are determined heuristically as described in Niemann *et al.* [2003]. The tension  $\phi$  is primarily based on the spacing of observations, and  $\nu$  and  $\rho$  are chosen so that the interpolation method is numerically stable. The base values used in this study

are  $\varphi = 0.0045$ ,  $v = 900$ , and  $\rho = 0.005$ . The random offset applied to the spline surface is  $\pm 2.5$  m ( $\approx 1\%$  of overall relief in the observed elevations).

The parameter  $D$  was estimated using hillslope profiles from the paleotopography that are exposed at two different locations. The first site is located along New Mexico Highway 4 between White Rock and the entrance to BNM. A total station and a laser range finder were used to determine the shape of this hillslope profile (Figure 17a). The second site is located in Alamo Canyon in BNM. This site reveals either one or two hillslopes depending on whether one interprets a small decrease in the elevation of the exposed boundary as a valley (see Figure 17b). At both sites,  $D$  can be estimated from Equation (3.1). At equilibrium and using  $U = 1$ , Equation (1) implies:

$$D = \frac{\beta A^m S - 1}{\nabla^2 z}. \quad (3.2)$$

On hillslopes, the fluvial term  $\beta A^m S$  is small and is thus neglected. We do not know the lateral hillslope curvature because only cross-sections are observed. Thus, the Laplace curvature is approximated with the profile curvature  $\partial^2 z / \partial x^2$ , where  $x$  is horizontal distance. With these simplifications, Equation (3.2) implies  $D = -1 / [\partial^2 z / \partial x^2]$ , so  $D$  can be estimated by fitting a second-degree polynomial of the form  $z = ax^2 + bx + c$  to the profile data and calculating  $D = -1 / [2a]$ . Figure 18 shows the fitted parabolas and resulting estimates of  $D$ , which range from 96 to 507. Additional estimates of  $D$  were obtained by noting that the right end of the plot in Figure 17b is clearly a valley bottom because the boundary abruptly increases in elevation beyond the data plotted in the figure. Thus, the horizontal scale of the hillslope from valley bottom to ridge line can be calculated.  $D$  can then be estimated by developing an empirical relationship between the  $D$  value used in the geomorphic model and the resulting length scale of

the modeled hillslopes. From this approach, one estimates  $D$  to be either 490 or 457 depending on whether one interprets Figure 18b to include one or two hillslopes, respectively. Because the estimates from the two methods are consistent and because those data sets represent a hillslope observed from crest to valley whereas the other two data sets do not, a value of  $D = 450$  was selected for use in the method.

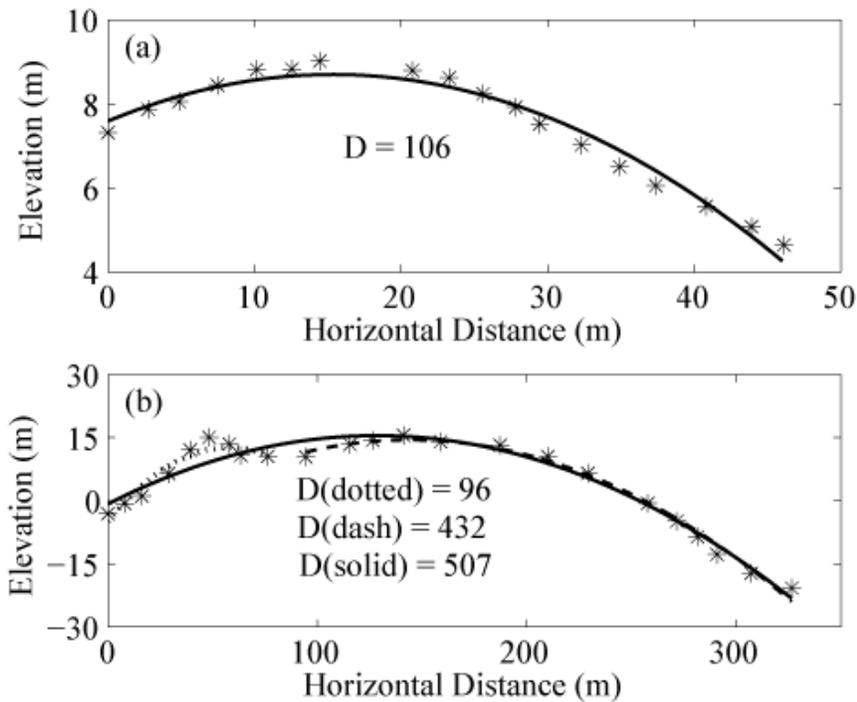


Figure 18 - Surveyed hillslope profiles in lower boundary of Tshirege Member from (a) a road cut on New Mexico Highway 4 near White Rock and (b) an exposed outcrop in Alamo Canyon in BNM. Second degree polynomial regression curves are of form  $z = ax^2 + bx + c$ .  $D$  values are estimated from regressions by formula  $D = -1/2a$ . Elevation values given are relative to an arbitrary datum for each site.

The geomorphic parameters  $m$  and  $\bar{\beta}$  can be estimated from the slope-area relationship [Flint, 1974; Gupta and Waymire, 1989; Hack, 1957] in the fluvial portion of the topography if it is known. For locations with large contributing areas, the fluvial term dominates in Equation (3.1)

and the term with  $D$  can be neglected. With that simplification and assuming equilibrium, one can show that:

$$\log S = \log \frac{1}{\beta} - m \log A \quad (3.3)$$

which suggests that the logarithm of the channel slope is linearly related to the logarithm of the contributing area (see Niemann et al. [2003] for more detailed derivation). The slope of the line is  $m$ , and the intercept is related to  $\beta$  or  $\bar{\beta}$  (recall that  $\bar{\beta}$  is only the starting point for the estimation of a spatially-variable  $\beta$ ). Unfortunately, the sparse data for the lower boundary of the Tshirege Member are not sufficient to determine an associated slope-area relationship. However, Niemann et al. [2003] found that the interpolation method is relatively insensitive to the values  $m$  and  $\bar{\beta}$  because errors in these values can be partially overcome as  $\beta$  is iteratively adjusted.

Furthermore, values of  $m$  typically vary only between 0.3 and 0.6 for different regions [*Moglen and Bras*, 1995; *Snyder et al.*, 2000; *Tarboton et al.*, 1991]. Thus, for simplicity, we estimate these two parameters from the slope-area relationships of two large basins, Frijoles and Water Canyons, on the current Pajarito topography using 30 m USGS DEM data. It is recognized that the paleotopography was likely very different than the current topography, which is dominated by steep-walled canyons [*Broxton and Reneau*, 1996], so the implications of this approach will be analyzed through a sensitivity analysis later in the paper. Due to the uneven distribution of contributing area values within each basin, the contributing areas were grouped into equally-spaced bins in log-scale and the average slope for each bin was calculated (Figure 19a). A linear regression was performed using all binned values with contributing areas greater than or equal to the one where the maximum average slope occurred. The slopes of the regression lines were -

0.26 and -0.25, so 0.26 was used for  $m$ . The intercepts of the regression lines were converted to individual estimates for  $\bar{\beta}$ , which implied an average value of 1.16 for  $\bar{\beta}$ .

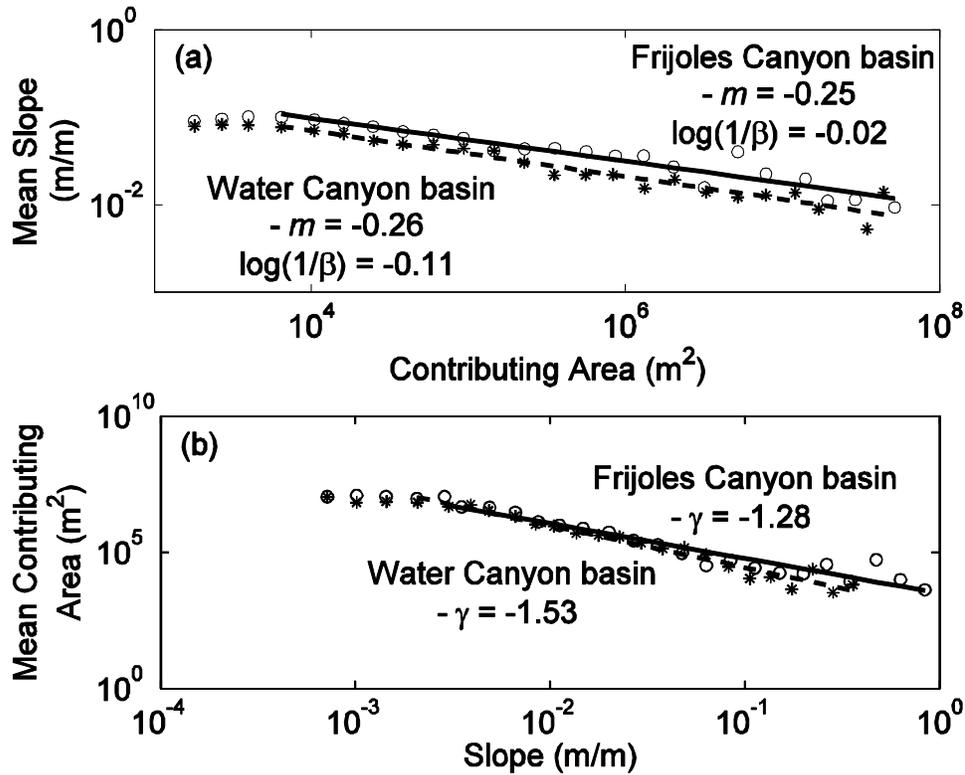


Figure 19 - Relationships between local slope and contributing area from two basins (Frijoles and Water Canyons) on present topography of the Pajarito Plateau. (a) Slope-area relationship and linear regression used to estimate  $m$  and  $\beta$  for interpolation method. (b) Area-slope relationships for same basins and regressions used to estimate  $\gamma$  parameter for assigning contributing area to closed-boundary inlets. In both plots, circles and solid lines refer to Frijoles Canyon while stars and dashed lines refer to Water Canyon.

Finally, in order to apply the interpolation method, one must specify the boundary conditions for the numerical model. A trend in the regional elevation is first specified as downward to the south and east based on the location of the caldera to the north and west and the location of the pre-Tshirege Rio Grande to the south and east [Broxton and Reneau, 1996; Broxton and Vaniman, 2005]. The down-gradient boundaries are open, allowing water to exit, while the up-gradient boundaries are closed to exiting flow but allow flow to enter at the inlets (low points on

the closed boundaries). The local slope of each inlet is used in an area-slope relationship to estimate an associated incoming contributing area. This relationship differs from the slope-area relationship only because slope is treated as the independent variable and contributing area is treated as the dependent variable, which produces a different regression line than the one used earlier. The slope of the new regression line is called  $\gamma$  and is the required parameter to assign the contributing areas to the inlets. Once again, the data for Frijoles and Water Canyons were used, and the regression results can be seen in Figure 19(b). For Frijoles and Water Canyons,  $\gamma$  is 1.28 and 1.53, respectively. The average value of 1.41 was used as the base value for  $\gamma$  in the simulations. The contributing areas inferred from this approach are clearly speculative, but the interpolation results were found to be more sensitive to the inlet locations than the estimated contributing areas.

## **Results**

A 60 m resolution grid was created for the rectangular extent of the available data for the lower boundary of the Tshirege Member, which resulted in a grid 101 rows by 195 columns. In order to avoid undesirable effects by having a boundary immediately adjacent to the domain of interest, the interpolation region was expanded by 4-pixels at all borders. Figure 20 shows the locations of the elevation data used in this study, which were derived from well logs and field surveys of exposed contacts. A total of 49 observations were employed, and each grid cell containing an observation was assigned the elevation of that observation.

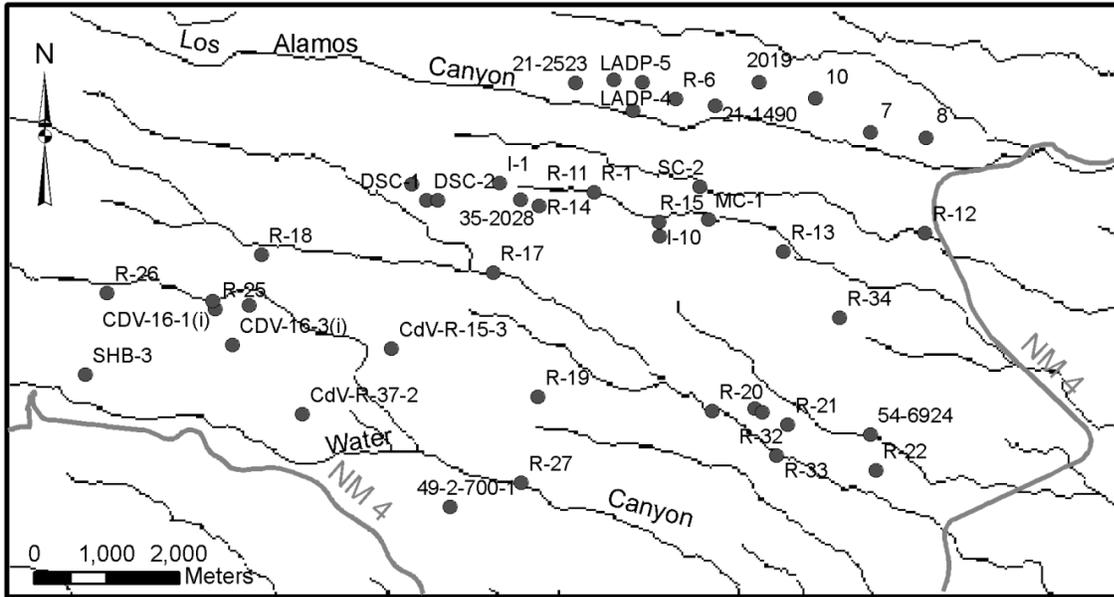


Figure 20 - Map showing the locations where the elevation of the lower boundary of the Tshirege Member has been observed. Points are labeled according to their LANL designations.

The completely-regularized spline surface and an example of the final surface produced by the physically-based method are shown in Figure 21(a) and Figure 21(b), respectively. The circles in each figure indicate the locations of the observations. The smoothness of the spline surface and the more realistic texture, including the drainage patterns, of the physically-based surface are easily seen in the figure. In some applications such as groundwater modeling, these textural differences could be important. For example, paleodrainages might determine some groundwater flow pathways. The physically-based surface exhibits four major basins with outlets on the southern boundary (labeled S1 to S4 in the figure). The westernmost basin is the largest, and it drains area from inlets on the western boundary. The other three basins become progressively smaller as one looks to the east. A major basin also has an outlet on the eastern boundary (labeled E1), and it drains area from inlets on the northern boundary. Broxton and Reneau [1996] identified the possible locations of four major paleodrainages for the pre-Tshirege landscape. Basins S1 and S2 in Figure 21(b) roughly align with the upper portions of two of

those paleodrainages, and basin E1 roughly corresponds to another. The other basin they identified is located outside our interpolation domain.

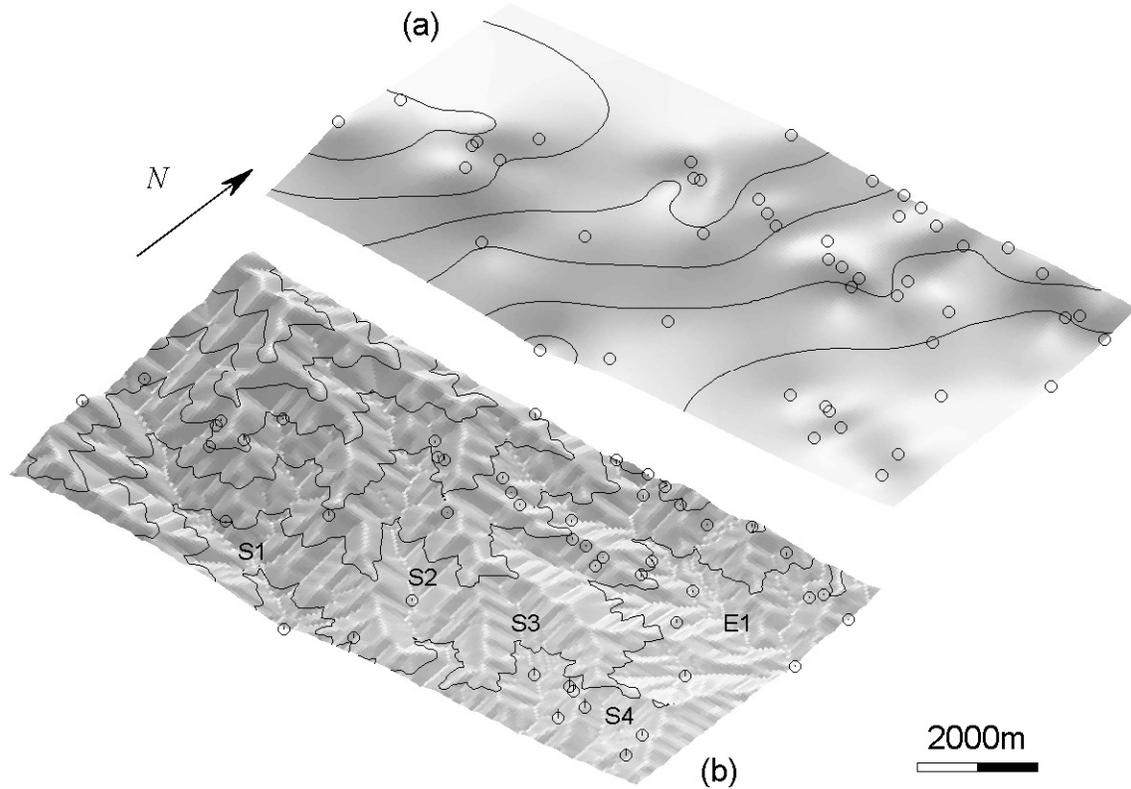


Figure 21 - Estimates of lower boundary of Tshirege Member based on (a) a completely-regularized spline and (b) physically-based interpolation surface with base parameters. Circles represent elevation measurement locations, and vertical lines show deviation between observed and simulated elevations.

One way to evaluate the performance of the method is to quantify its ability to reproduce the observed elevations. One hundred simulations were performed with the base parameters. For each simulation the mean absolute deviation [Ahmadzadeh and Petrou, 2001] between the simulated surface and the observed elevations was calculated. The first column of data in Table 3 characterizes the dataset of *MAD* values from the 100 simulations. The mean value is 15.36 m, which is about 6.5% of the total relief in the observed elevations. It should be noted that a smooth interpolation surface, such as a spline, is able to exactly match all the observations but

does not produce the realistic texture shown in Figure 21(b). The standard deviation given in the table indicates that the variability in the *MAD* between simulations is relatively small. Thus, most simulations have similar performance in terms of their ability to reproduce the observations.

Table 3. Statistics of mean absolute deviation [*Ahmadzadeh and Petrou, 2001*] between simulated and observed elevations. Statistics are calculated using 100 *MAD* values (from 100 simulations) for each parameter set. All parameters remain at their base values except for those listed in each column.

	<u>Base</u>	<u><math>\Phi = 0.0035</math></u>	<u><math>\Phi = 0.0055</math></u>	<u><math>m = 0.32</math></u>	<u><math>m = 0.45</math></u>	<u><math>\gamma = 1.15</math></u>	<u><math>\gamma = 1.70</math></u>
Mean (m)	15.36	13.38	15.99	19.01	36.82	14.09	17.11
Std Dev (m)	2.47	1.82	1.91	2.38	3.93	1.41	2.54

BNM is adjacent to the southern edge of LANL, so one might be interested in determining locations within the interpolation domain that ultimately drain across the southern boundary. For this analysis, drainage directions were determined by assuming flow travels in the direction of steepest descent on the paleotopography. Thus, this analysis does not directly consider groundwater flow directions. The boundary to which each point drains was identified for all one hundred simulations. The likelihood that a given point drains to the southern boundary was then calculated as the weighted proportion of the simulations in which the point drains to the southern boundary. The weight used for each simulation was determined by how closely that simulation matches the observed elevations [*Beven and Binley, 1992*]. In particular, the weight of the  $j^{th}$  simulation  $w_j$  is calculated:

$$w_j = \frac{1/MAD_j}{\sum_{N_{sim}} 1/MAD_j}, \quad (3.4)$$

where  $MAD_j$  is the mean absolute deviation between the simulated and observed elevations of the  $j^{th}$  simulation and the summation is over all 100 simulations ( $N_{sim} = 100$ ). The likelihood that a point  $i$  drains through the southern boundary  $L(i)$  is then calculated:

$$L(i) = \sum_{j=1}^{N_{sim}} W_j, \quad (3.5)$$

where  $W_j = w_j$  if  $i$  drains through the southern boundary in simulation  $j$  and  $W_j = 0$  otherwise.

Figure 22 shows contours of the likelihood that points drain across the southern boundary. For example, a point lying south of the 0.8 contour line has a likelihood greater than 80% of draining through the southern boundary based on the 100 simulations. The contour lines probabilistically identify the drainage divide that separates locations draining to the south and locations draining to the east. The divide is relatively stable between simulations where the contours are closely spaced. It is interesting to note that most of the interpolation domain almost always or almost never drains to the southern boundary. Very few points drain to different boundaries in different simulations, although the northwest region exhibits more variability than the southeast region. These results are likely controlled by the southeast regional tilt in the observed elevations as well as the elevations and locations of the southern and eastern boundaries. Discovery of major drainages crossing the band of contour lines or measurement of low elevations in that area would have a significant impact on the results. Better constraints on the boundary elevations could also affect the results. In order to investigate the robustness of the method, sensitivity analyses were performed individually for the parameters  $\phi$ ,  $m$ , and  $\gamma$ . The parameters  $\phi$  and  $\gamma$  were changed by approximately  $\pm 20\%$  of their base values, and 100 additional simulations were performed with each change. Varying  $\phi$  changes how rapidly the  $\beta$  field can change spatially, and varying  $\gamma$  changes the incoming contributing area associated with an inlet of a given local slope.

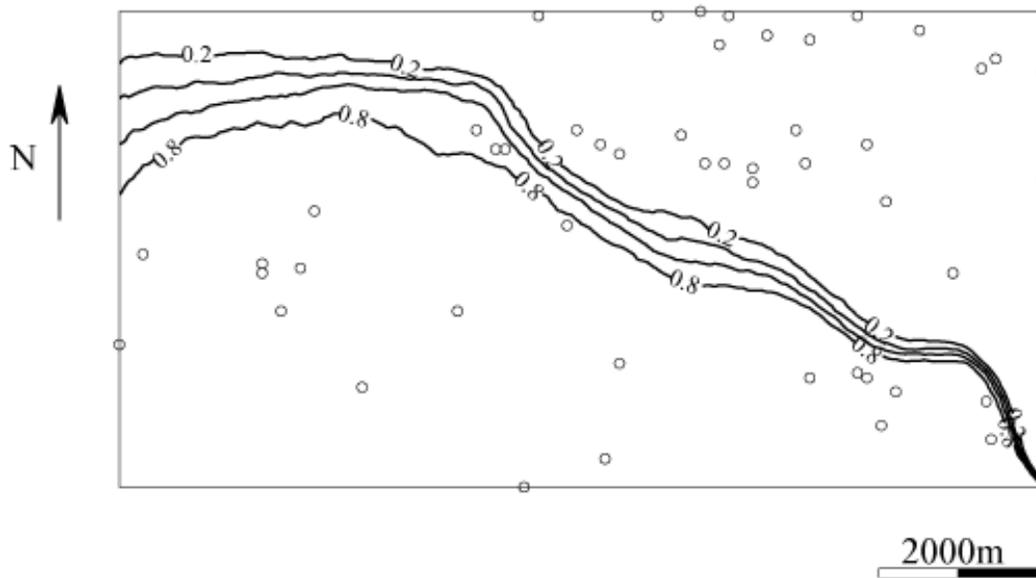


Figure 22 - Map of likelihood each point ultimately drains through southern boundary. Circles indicate elevation measurement locations.

Because the base value of  $m = 0.26$  is small relative to the typical range for this parameter [Moglen and Bras, 1995; Snyder et al., 2000; Tarboton et al., 1991], values of 0.32, which represents a 20% increase from the base value, and 0.45, which is in the middle of the typical range, were used for the sensitivity analysis. A representative surface for each parameter set is shown in Figure 23. Analogous major drainages are present in each surface with the similar orientations. The relative size of some drainage basins, however, differs between some simulations. For example, the area draining through the relatively flat area in the southeast corner of the simulation with  $\phi = 0.0035$  is much greater than in any other simulation. The most significant visual differences occur when  $m$  is varied. As  $m$  increases, the surface appears rougher due to the greater concavity in the long profiles of the channels.

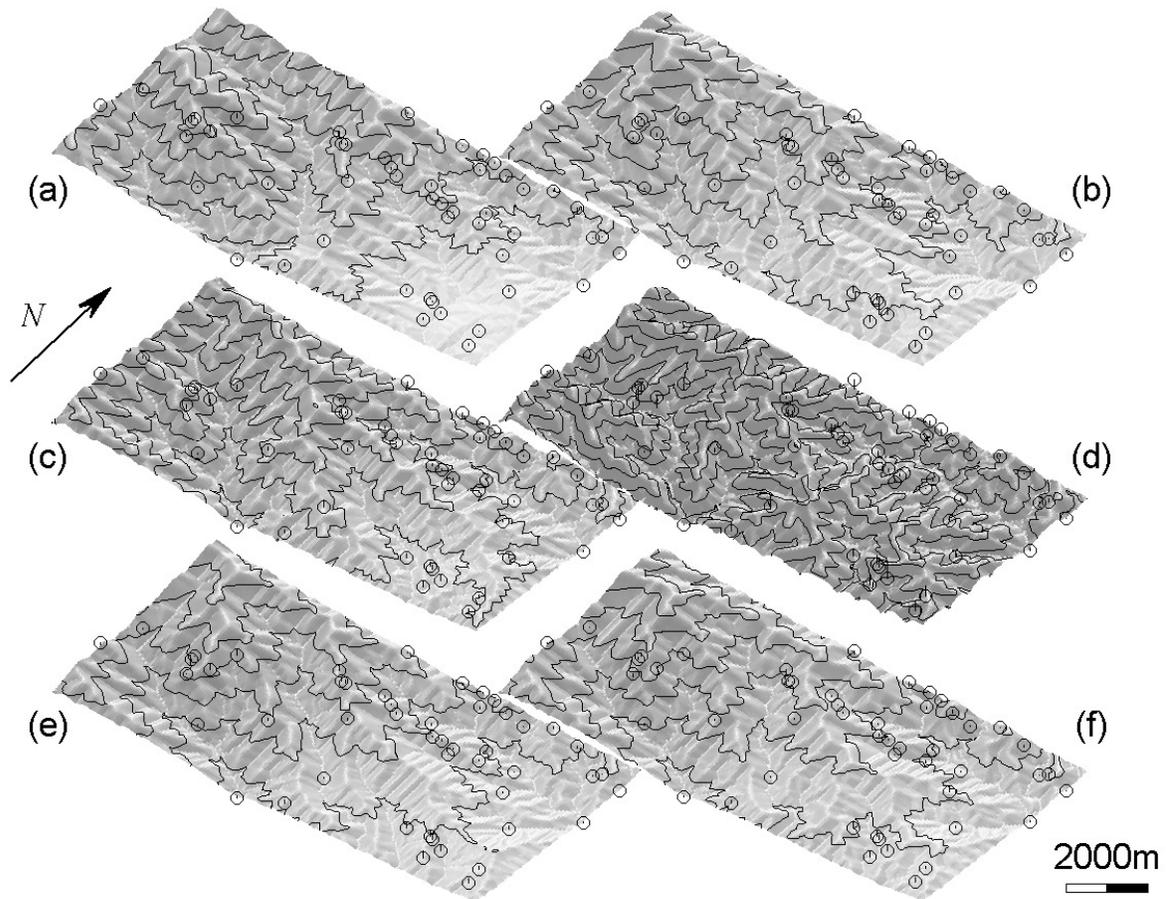


Figure 23 - Surfaces generated by physically-based interpolation method when individual parameters are changed. Altered parameter values are: (a)  $\varphi = 0.0035$ , (b)  $\varphi = 0.0055$ , (c)  $m = 0.32$ , (d)  $m = 0.45$ , (e)  $\gamma = 1.15$ , and (f)  $\gamma = 1.70$ . Circles represent elevation measurement locations, and vertical lines show deviation between observed and simulated elevations.

Table 3 shows the summary of the MAD values for each parameter set. The error summary indicates each parameter set's ability to produce surfaces that honor the observed elevations. In that context, all the parameter sets except the  $m = 0.45$  set perform approximately the same. The error statistics are substantially larger for the  $m = 0.45$  results, but that value also represents a greater change from the base value than the other values used in the sensitivity analyses. In terms of their ability to reproduce the observations, two parameter sets actually perform better than the base set. The  $\varphi = 0.0035$  and the  $\gamma = 1.15$  simulations have mean MAD values that are 5.7% and 6.0% of the observed relief. Decreasing the value of  $\varphi$  increases the radius of

influence of each observation in the estimation of the  $\beta$  surface. This smoothing allows more predictability in the model's response to changes in  $\beta$ , which improves its ability to reproduce the observations. Decreasing the value of the  $\gamma$  parameter reduces the incoming drainage area, which gives the model more flexibility in reproducing the observations.

Likelihood maps generated using the sensitivity analysis parameter sets are shown in Figure 24 - Maps of likelihood that each point drains through southern boundary.. The overall shapes of the contours are quite similar in most cases. The greatest digression from the typical pattern in the southeast region is seen when  $\varphi = 0.0035$  (Figure 24a).

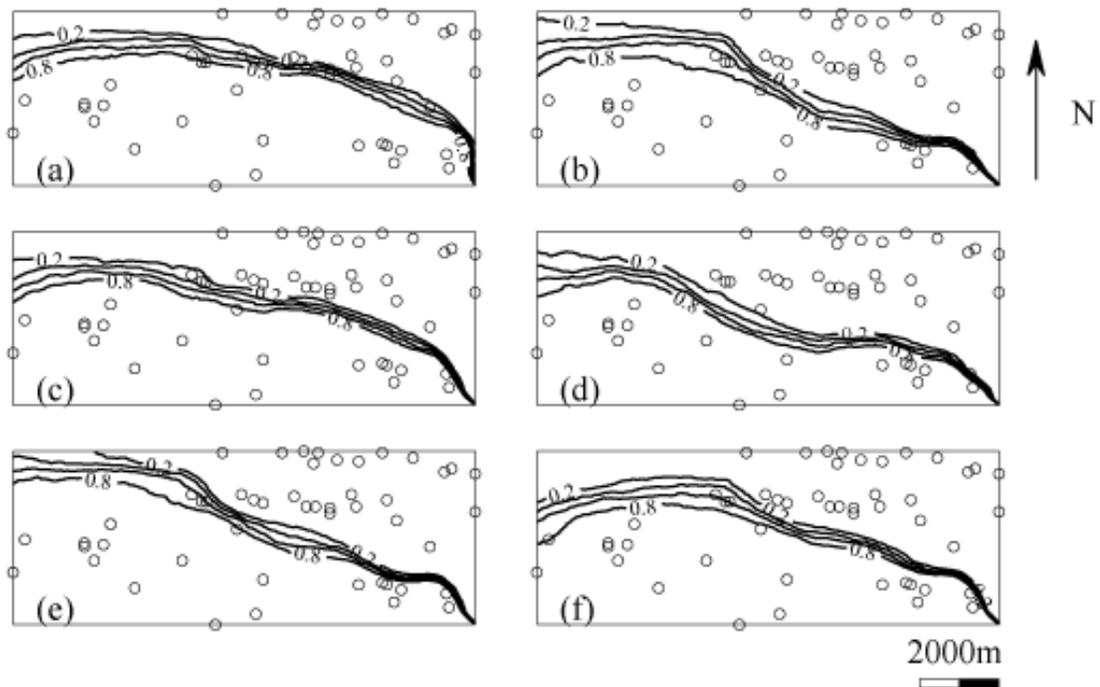


Figure 24 - Maps of likelihood that each point drains through southern boundary. In each map, an individual parameter value was changed, and 100 simulations were conducted. Altered parameters are: (a)  $\varphi = 0.0035$ , (b)  $\varphi = 0.0055$ , (c)  $m = 0.32$ , (d)  $m = 0.45$ , (e)  $\gamma = 1.15$ , and (f)  $\gamma = 1.70$ . Circles indicate elevation measurement locations.

In this case, the contours are pushed further to the north. The value of  $\gamma$  has the most impact on the northwest region, suggesting that the inlet areas have the most impact on the results where elevation data are scarce. The general uniformity of results indicates that large regions within the interpolation area are insensitive to the parameters used in the method, which confirms our interpretation about the importance of the boundary conditions and the observed elevations in determining the drainage patterns.

## **Conclusions**

1. We have demonstrated the feasibility of using a process-based approach to reconstruct a buried hillslope and valley topography. The approach was made possible by generalizing the boundary conditions of a physically-based topographic interpolation method to allow flow to enter the interpolation domain. Depending on the specific parameter values used, the errors between the simulated surface and the observations used in the method are about 5-8% of the total relief in the observed elevations.
2. The application of the method to the lower boundary of the Tshirege Member consistently identifies the occurrence of five major valleys, three of which are similar to those identified by Broxton and Reneau [1996]. Ultimately the number of major valleys produced by the method is closely linked with the drainage density, which was determined based on field surveys of exposed hillslope profiles, and the basin structures, which were inferred by the method's attempt to match the observed elevations.
3. The method also provides guidance for future data collection at the study site. The locations that ultimately drain to the southern edge of the interpolation domain are relatively stable between different simulations with the same parameters and simulations when individual parameters are varied by 20%. Data collection along the south and east boundaries would

provide better constraints on the elevations of these boundaries, which could significantly impact the simulations. Furthermore, data collection along the narrow band where the direction of flow varies between different simulations might identify a channel crossing this region, which could again impact the results.

This research also presents several opportunities for future research. The reconstructed surfaces could be used as part of a comprehensive effort to model groundwater flow and contaminant transport in the region. Such an analysis would need to include the modern day arroyos that produce discontinuities in the lower boundary of the Tshirege Member as well as a detailed representation of the main aquifers. The reconstruction method could also be generalized to accept more diverse information. Grimaldi et al. [2004] presented a different extension of the topographic interpolation method that allows the imposition of known drainage paths. That capability may allow one to include flow directions inferred from material provenance or clast imbrications at exposed contacts.

### **Acknowledgments**

The authors would like to gratefully acknowledge financial support from the Army Research Office Terrestrial Sciences Program. David Broxton from Los Alamos National Laboratory provided well data used in the interpolation model and Bandelier National Monument granted approval for this research under permit #BAND-2007-SCI-0001. The authors would also like to thank two reviewers for their helpful comments.

## References

- Ahmadzadeh, M. R., and M. Petrou (2001), Error statistics for slope and aspect when derived from interpolated data, *IEEE Transactions on Geoscience and Remote Sensing*, 39(9), 1823-1833.
- Bartier, P. M., and C. P. Keller (1996), Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW), *Computers & Geosciences*, 22(7), 795-799.
- Beven, K., and A. Binley (1992), The future of distributed models - model calibration and uncertainty prediction, *Hydrol. Processes*, 6(3), 279-298.
- Bindlish, R., and A. P. Barros (1996), Aggregation of digital terrain data using a modified fractal interpolation scheme, *Computers & Geosciences*, 22(8), 907-917.
- Bouboulis, P., L. Dalla, and V. Drakopoulos (2006), Construction of recurrent bivariate fractal interpolation surfaces and computation of their box-counting dimension, *Journal of Approximation Theory*, 141(2), 99-117.
- Briggs, I. C. (1974), Machine contouring using minimum curvature, *Geophysics*, 39(1), 39-48.
- Broxton, D. E., and S. L. Reneau (1996), Buried early Pleistocene landscapes beneath the Pajarito Plateau, Northern New Mexico, in *New Mexico Geological Society Guidebook, 47th Field Conference*, edited by F. Goff, pp. 325-334, New Mexico Geological Society, Socorro, NM.
- Broxton, D. E., and D. T. Vaniman (2005), Geologic framework of a groundwater system on the margin of a rift basin, Pajarito Plateau, North-Central New Mexico, *Vadose Zone Journal*, 4, 522-550.
- Chang, C. L., S. L. Lo, and S. L. Yu (2005), Applying fuzzy theory and genetic algorithm to interpolate precipitation, *J. Hydrol.*, 314(1-4), 92-104.
- Chapin, C. E., and W. E. Elston (1979), Introduction, in *Geological Society of America Special Paper 180: Ash-flow Tuffs*, edited by C. E. Chapin and W. E. Elston, pp. 1-3, Geological Society of America, Boulder, Colorado.
- Chiles, J.-P., and P. Delfiner (1999), *Geostatistics: Modeling Spatial Uncertainty*, 695 pp., John Wiley and Sons, New York, NY.
- Craig, R. G. (1993), Kinematics of paleo landforms, in *Computers in Geology - 25 Years of Progress*, edited by J. C. Davis and U. C. Herzfeld, pp. 275-286, Oxford University Press, New York, NY.
- Crowley, T. J., and K. C. Burke (Eds.) (1998), *Tectonic Boundary Conditions for Climate Reconstructions*, 285 pp., Oxford University Press, New York, NY.

Fisher, R. V., and H.-U. Schmincke (1984), *Pyroclastic Rocks*, 472 pp., Springer-Verlag, New York, NY.

Flint, J. J. (1974), Stream gradient as a function of order, magnitude, and discharge, *Water Resour. Res.*, 10(5), 969-973.

Francis, P. (1993), *Volcanoes: A Planetary Perspective*, 443 pp., Oxford University Press, New York, NY.

Grimaldi, S., V. Teles, and R. L. Bras (2004), Sensitivity of a physically based method for terrain interpolation to initial conditions and its conditioning on stream location, *Earth Surface Processes and Landforms*, 29(5), 587-597.

Grimaldi, S., F. Nardi, F. Di Benedetto, E. Istanbuluoglu, and R. L. Bras (2007), A physically-based method for removing pits in digital elevation models, *Adv. Water Resour.*, 30(10), 2151-2158.

Gupta, V. K., and E. Waymire (1989), Statistical self-similarity in river networks parameterized by elevation, *Water Resour. Res.*, 25(3), 463-476.

Hack, J. T. (1957), Studies of longitudinal stream profiles in Virginia and Maryland, in *United States Geological Survey Professional Paper 294-B*, edited, pp. 45-97, U. S. Government Printing Office, Washington, D.C.

Hofierka, J., and T. Cebeauer (2007), Spatial interpolation of elevation data with variable density: A new methodology to derive quality DEMs, *IEEE Geoscience and Remote Sensing Letters*, 4(1), 117-121.

Holzschuh, J. (2002), Low-cost geophysical investigations of a paleochannel aquifer in the eastern goldfields, Western Australia, *Geophysics*, 67(3), 690-700.

Howard, A. D. (1994), A detachment-limited model of drainage basin evolution, *Water Resour. Res.*, 30(7), 2261-2285.

Izett, G. A., and J. D. Obradovich (1994), Ar-40/Ar-39 age constraints for the Jaramillo normal subchron and the Matuyama-Brunhes geomagnetic boundary, *Journal of Geophysical Research-Solid Earth*, 99(B2), 2925-2934.

Jones, T. A., D. E. Hamilton, and C. R. Johnson (1986), *Contouring Geologic Surfaces with the Computer*, 1st ed., 314 pp., Van Nostrand, New York, NY.

Journel, A. G., and C. J. Huijbregts (1978), *Mining Geostatistics*, 600 pp., Academic Press, San Francisco, CA.

Kelley, S., and R. E. Kelley (2004), Paleotopography on the Otowi Member of the Bandelier Tuff: Implications for landscape evolution of the Jemez Mountains, New Mexico, paper presented at Geological Society of America Annual Meeting, Denver, CO.

- Kirkby, M. J. (1971), Hillslope processes - response models based on the continuity equation, in *Slopes: Form and Process*, edited by D. Brunsten, pp. 15-30, Institute of British Geographers, Oxford, UK.
- Miller, R. D., J. Xia, C. B. Park, J. Ivanov, and E. Williams (1999), Using MASW to map bedrock in Olathe, Kansas, paper presented at Society of Exploration Geophysicists, Society of Exploration Geophysicists, Houston, TX, Oct 31 - Nov 5, 1999.
- Mitasova, H., and L. Mitas (1993), Interpolation by regularized spline with tension: I. Theory and implementation, *Math. Geol.*, 25(6), 641-655.
- Mitasova, H., L. Mitas, W. M. Brown, D. P. Gerdes, I. Kosinovsky, and T. Baker (1995), Modeling spatially and temporally distributed phenomena - new methods and tools for GRASS GIS, *International Journal Of Geographical Information Science*, 9(4), 433-446.
- Moglen, G. E., and R. L. Bras (1995), The effect of spatial heterogeneities on geomorphic expression in a model of basin evolution, *Water Resour. Res.*, 31(10), 2613-2623.
- Newman, B. D., and B. A. Robinson (2005), The hydrogeology of Los Alamos National Laboratory: Site history and overview of vadose zone and groundwater issues, *Vadose Zone Journal*, 4, 614-619.
- Niemann, J. D., R. L. Bras, and D. Veneziano (2003), A physically based interpolation method for fluvially eroded topography, *Water Resour. Res.*, 39(1), 1017, doi: doi:10.1029/2001WR001050.
- Ollier, C. (1988), *Volcanoes*, 228 pp., Basil Blackwell, Ltd., New York, NY.
- Polidori, L., and J. Chorowicz (1993), Comparison of bilinear and Brownian interpolation for digital elevation models, *ISPRS Journal of Photogrammetry & Remote Sensing*, 48(2), 18-23.
- Reneau, S. L. (2000), Stream incision and terrace development in Frijoles Canyon, Bandelier National Monument, New Mexico, and the influence of lithology and climate, *Geomorphology*, 32(1-2), 171-193.
- Sandberg, S. K., W. Corso, J. R. Levine, G. Newhart, and G. Powell (2001), Mapping a paleochannel system controlling contaminant migration at a wood-treating facility using electromagnetics, paper presented at Symposium on the Application of Geophysics to Engineering and Environmental Problems, Environmental and Engineering Geophysical Society, Denver, CO, March 4-7, 2001.
- Smith, R. L. (1960), Ash Flows, *Bulletin of the Geological Society of America*, 71, 795-842.
- Smith, R. L., and R. A. Bailey (1966), The Bandelier Tuff: a study of ash-flow eruption cycles from zoned magma chambers, *Bulletin volcanologique*, 29, 83-104.

Snyder, N. P., K. X. Whipple, G. E. Tucker, and D. J. Merritts (2000), Landscape response to tectonic forcing: Digital elevation model analysis of stream profiles in the Mendocino triple junction region, northern California, *Geological Society of America Bulletin*, 112(8), 1250-1263.

Spell, T. L., I. McDougall, and A. P. Dougeris (1996), Cerro Toledo Rhyolite, Jemez volcanic field, New Mexico:  $^{40}\text{Ar}/^{39}\text{Ar}$  geochronology of eruptions between two caldera-forming events, *Geological Society of America Bulletin*, 108(12), 1549-1566.

Stein, M. L. (1999), *Interpolation of spatial data: some theory for kriging*, 247 pp., Springer-Verlag, New York, NY.

Talmi, A., and G. Gilat (1977), Method for smooth approximation of data, *Journal of Computational Physics*, 23, 93-123.

Tarboton, D. G., R. L. Bras, and I. Rodriguez-Iturbe (1991), On the extraction of channel networks from digital elevation data, *Hydrol. Processes*, 5(1), 81-100.

Watson, D. F., and G. M. Philip (1985), A refinement of inverse distance weighted interpolation, *Geo-Processing*, 2(4), 315-327.

Williams, B. A., B. Bjornstad, R. Schalla, and B. Webber (2002), Revised hydrogeology for the suprabasalt aquifer system, 200-West Area and vicinity, Hanford Site, Washington, edited by U. DOE, p. 91, Pacific Northwest National Laboratory, Richland, Washington.

Yokoya, N., K. Yamamoto, and N. Funakubo (1989), Fractal-based analysis and interpolation of 3D natural surface shapes and their application to terrain modeling, *Computer Vision, Graphics, and Image Processing*, 46, 284-302.

Zoraster, S. (1996), Imposing geologic interpretations on computer-generated contours using distance transformations, *Math. Geol.*, 28(8), 969-985.

## SUMMARY

The first part of this dissertation evaluated the efficacy of nonlinear estimation techniques for estimating spatial soil moisture patterns relative to multiple linear regression. The nonlinear estimation techniques evaluated were the SANN and MM methods and all of the methods used topographic attributes as auxiliary data. Both of the nonlinear methods are examples of unsupervised machine-learning techniques that do not place *a priori* constraints on the form of the relationship between the auxiliary data and the soil moisture. The SANN method performs best overall and although its performance is consistently better than the other two methods the increase is generally not large. However, the SANN method has a further advantage in that it performs better when using more auxiliary variables than the other two methods, which means it requires less prior knowledge about the attributes to use as predictor variables. Additionally, catchment characteristics and wetness conditions affect the accuracy of all the methods. While the SANN and MM are capable of performing better than multiple linear regression, the improvement generally comes with the cost of needing more observed data; therefore, individual requirements will determine whether the improved performance is warranted. All of the methods used in Part I have some inherent limitations, such as being driven purely by data and not related to identifiable hydrologic processes, which may hamper such techniques' interpretation and transferability.

The model used to estimate spatial soil moisture patterns in Part II overcomes some of the limitations of the nonlinear methods in Part I by having a derivation based in physical hydrologic concepts. This derivation leads to model parameters and outputs that are more identifiable with catchment characteristics and physical processes than those of the previously evaluated nonlinear methods. The model produces realistic soil moisture patterns whose organizational structure can change with the spatial average soil moisture for all three catchments tested. As a result, the

model was used to evaluate the vegetation, climate and soil characteristics that affect moisture pattern organization as well as the temporal stability of the pattern types. Soil characteristics, such as the horizontal saturated hydraulic conductivity and pore connectedness, are strongly related to the pattern organization that occurs in a catchment. In addition, the parameter relating the horizontal hydraulic gradient and the topographic slope is also important to pattern organization. Finally, greater vegetation transpiration efficiency or thicker vegetation cover also has a relatively strong influence on the moisture patterns. Time instability in the model is promoted by parameter values that allow a balance in the overall importance of lateral flow and radiative ET. That balance is commonly achieved with relatively large horizontal saturated hydraulic conductivity values. Furthermore, time instability in the modeled patterns also requires parameter values that allow the dominant process to switch between wet and dry conditions. The combination of relatively high values of pore disconnectedness and high values of vegetation cover (or transpiration efficiency) promotes this behavior.

Part III of the dissertation demonstrated the feasibility of using a process-based approach to reconstruct a buried hillslope and valley topography. The feasibility was shown by application of the approach to a preserved fluvial topography located near Los Alamos, New Mexico in the United States. The method consistently identifies five major drainages in the paleotopography, three of which are similar to drainages identified by previous research. The method could be used to provide guidance for future data collection efforts by indicating the areas in which paleotopographic drainage directions are most uncertain. In the future, the reconstruction method could also be generalized to accept more diverse information, such as known drainage paths.