J5.3

A SPATIAL TIME SERIES FRAMEWORK FOR MODELING DAILY PRECIPITATION AT REGIONAL SCALES

Phaedon C. Kyriakidis[1] * Norman L. Miller[2] Jinwon Kim[3]
[1]University of California Santa Barbara, Santa Barbara, CA
[2]Lawrence Berkeley National Laboratory, Berkeley, CA
[3]University of California Los Angeles, Los Angeles, CA

1. INTRODUCTION

Estimates of precipitation at regional scales constitute one of the most important input parameters for hydrologic impact assessment studies. At these scales, Limited Area Models (LAMs) provide an emerging means for enhancing the accuracy of precipitation predictions (Giorgi and Mearns, 1991; Kim and Soong, 1996; Miller and Kim, 1996; Kim et al., 1998). Dynamic downscaling using LAMs yield precipitation predictions which are physically and dynamically consistent with other atmospheric variables produced in the downscaling procedure. Dynamical downscaling, however, is computationally expensive and not error-free due to limited spatial resolution and model parameterizations. Stochastic characterization of rainfall fields based on rain gauge data and ancillary information, e.g., terrain elevation, still provides one of the basic tools for constructing rainfall maps at regional scales (Bras and Rodríguez-Iturbe, 1985; Seo et al., 2000; Kyriakidis et al., 2002), even though the physical and dynamic consistency of such maps is not guaranteed.

Time domain approaches for modeling daily precipitation typically involve vectors of time series, e.g., multivariate autoregressive (AR) models. Such models exploit the typically better informed time domain, but are limited to predictions only at rain gauge locations (Wilks, 1998; von Storch and Zwiers, 1999). This limitation hinders the all important task of spatiotemporal mapping. More recently, time series approaches have been generalized to a continuous spatial domain and maps of precipitation levels are constructed at any arbitrary location via interpolation of time series model parameters (Johnson et al., 2000).

In this paper, a framework for stochastic spatiotemporal modeling of daily precipitation in a hindcast mode is presented. Observed precipitation levels in space and time are modeled as a joint realization of a collection of space-indexed time series, one for each spatial location. Time series model parameters are spatially varying, thus capturing space-time interactions. Stochastic simulation, i.e., the procedure of generating alternative precipitation realizations (synthetic fields) over the space-time domain of interest (Deutsch and Journel, 1998), is employed for ensemble prediction. The simulated daily precipitation fields reproduce a data-based histogram and spatiotemporal covariance model, and identify the measured precipitation values at the rain gauges (conditional simulation). Such synthetic precipitation fields can be used in a Monte Carlo framework for risk analysis studies in hydrologic impact assessment investigations (Bras and Rodríguez-Iturbe, 1985; Kyriakidis et al., 2001).

*Corresponding author address: Phaedon C. Kyriakidis, University of California Santa Barbara, Dept. of Geography, Ellison Hall 5710, CA 93117; e-mail: phaedon@geog.ucsb.edu

2. SPATIAL TIME SERIES

In the proposed methodology, daily precipitation is modeled as a collection of spatially correlated time series, $\{Z(\mathbf{u},t), \mathbf{u} \in D, t \in T\}$, one per location $\mathbf{u} \in D$; here $\mathbf{u} = (u_1, u_2)$ denotes the 2D spatial coordinate vector, D denotes the study area, and T the time span of interest. That spatiotemporal process is decomposed into:

$$Z(\mathbf{u},t) = M(\mathbf{u},t) + R(\mathbf{u},t), \ \forall \mathbf{u} \in D, \ \forall t \in T \qquad (1)$$

where $M(\mathbf{u},t)$ is a stochastic space-time component modeling some "average" smooth variability of the spatiotemporal process $Z(\mathbf{u},t)$, and $R(\mathbf{u},t)$ is stationary residual component, independent of $M(\mathbf{u},t)$, modeling higher frequency fluctuations around that trend in both space and time.

The trend component typically characterizes long-term temporal patterns, for example precipitation variability attributed to climatic factors. Other patterns of variability, e.g., those linked to local weather conditions, are typically accounted for by the stochastic residual component. It should be stressed that the dichotomy of equation (1) is a (subjective) modeling decision: there is no "true" temporal trend component, since there are no trend data. The resulting residual component is thus a collective term for all components of variability that are not included in the trend model (Thiébaux, 1997).

The temporal characteristics of precipitation profiles are not stationary in space. For example, spatially varying weather conditions can lead to different patterns of precipitation temporal variability in regions near the ocean than in orographically isolated areas. It is critical to consider spatially non-stationary patterns of temporal variability in the modeling procedure, as well as to account for the influence of ancillary information on the spatial distribution of these parameters.

In this paper, local parametric models for the temporal trend of daily precipitation are first established at the rain gauges. The joint spatial distribution of the temporal trend model parameters is then characterized in a stochastic mode via a vector random function (RF) or random field model (Wackernagel, 1995). Estimates of these parameters are constrained by additional information, such as terrain elevation and its interaction with large-scale specific humidity derived from an assimilated data product from the National Centers for Environmental Prediction and the National Center for Atmospheric Research (NCEP/NCAR reanalysis).

The residuals from these local trend models are regarded as a realization of a stationary spatiotemporal process. Realizations of this process are generated via conditional stochastic simulation and added to the estimated trend component to produce alternative conditional realizations of the spatiotemporal distribution of daily precipitation.

2.1. Station-specific temporal trend models

The sample precipitation profile $\{z(\mathbf{u}_\alpha,t_i), i\in T_\alpha\}$ at each rain gauge location with coordinate vector $\mathbf{u}_\alpha$ is regarded as a realization of a random process $\{Z(\mathbf{u}_\alpha,t_i), i\in T_\alpha\}$, where T_α is the time span of measurements at $\mathbf{u}_\alpha$. This random process $\{Z(\mathbf{u}_\alpha,t_i),\ i\in T_\alpha\}$ is decomposed as:

$$Z(\mathbf{u}_\alpha,t_i) = m(\mathbf{u}_\alpha,t_i) + R(\mathbf{u}_\alpha,t_i),\ \ i=1,\ldots,T_\alpha \qquad (2)$$

where $\{m(\mathbf{u}_\alpha,t_i), t_i\in T_\alpha\}$ is a deterministic temporal trend, and $\{R(\mathbf{u}_\alpha,t_i), t_i\in T_\alpha\}$ is a stationary, zero mean, stochastic residual process.

The deterministic trend at each rain gauge location $\mathbf{u}_\alpha \in D$ is modeled as the sum of $(K+1)$ basis functions of time $f_k(t)$:

$$m(\mathbf{u}_\alpha,t_i) = \sum_{k=0}^{K} b_k(\mathbf{u}_\alpha) f_k(t_i), \quad i=1,\ldots,T_\alpha \qquad (3)$$

where $b_k(\mathbf{u}_\alpha)$ is the coefficient (intensity) associated with the k-th function $f_k(t_i)$, with $f_0(t_i)=1$ by convention.

Each basis function $f_k(t)$ is independent of the spatial location $\mathbf{u}$, and should ideally have a physical interpretation pertinent to the entire study region. Periodicities, especially when physically interpretable, should be incorporated in the deterministic trend $\{m(\mathbf{u}_\alpha,t_i), t\in T_\alpha\}$ as a Fourier series. Alternatively, such basis functions could be identified to a set of orthogonal factors derived via Empirical Orthogonal Function (EOF) analysis of the rain gauge precipitation profiles (von Storch and Zwiers, 1999), or to the spatial average of the latter.

The $(K+1)$ temporal trend coefficients $\mathbf{b}_\alpha = [b_k(\mathbf{u}_\alpha), k=0,\ldots,K]'$ are modeled at each rain gauge location $\mathbf{u}_\alpha$, independently from one location to another, using multiple regression; here superscript $'$ denotes a vector (or matrix) transpose. More precisely, the precipitation data at rain gauge $\mathbf{u}_\alpha$ are expressed as:

$$\mathbf{z}_\alpha = \mathbf{F}\mathbf{b}_\alpha + \mathbf{r}_\alpha, \quad \alpha\in(n) \qquad (4)$$

where $\mathbf{z}_\alpha = [z(\mathbf{u}_\alpha,t_i), i=1,\ldots,T_\alpha]'$ is a $(T_\alpha\times 1)$ vector of observations available at location $\mathbf{u}_\alpha$, $\mathbf{F}$ is a $(T_\alpha\times(K+1))$ design matrix whose k-th column is the k-th basis function $\mathbf{f}_k = [f_k(t_i), i=1,\ldots,T_\alpha]'$, and $\mathbf{r}_\alpha = [r(\mathbf{u}_\alpha,t_i), i=1,\ldots,T_\alpha]'$ is a $(T_\alpha\times 1)$ vector of residuals at location $\mathbf{u}_\alpha$; n is the number of rain gauges.

The vector of coefficients $\mathbf{b}_\alpha$ is expressed as a weighted linear combination of the data vector $\mathbf{z}_\alpha$: $\mathbf{b}_\alpha = \mathbf{H}_\alpha\mathbf{z}_\alpha$, where $\mathbf{H}_\alpha$ is a $((K+1)\times T_\alpha)$ matrix of weights assigned to each of the T_α data. If the matrix $\mathbf{F}$ is of full rank, the above system has a unique solution and the resulting matrix of weights $\mathbf{H}_\alpha$ is given by the ordinary least squares (OLS) solution: $\mathbf{H}_\alpha = (\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$ (Searle, 1971).

Once the $(K+1)$ coefficients $\mathbf{b}_\alpha$ specific to each rain gauge location $\mathbf{u}_\alpha$ are determined, the temporal trend model $\{m(\mathbf{u}_\alpha,t_i), t_i\in T_\alpha\}$ at that location is given by expression (3), and the corresponding residual series are obtained as:

$$r(\mathbf{u}_\alpha,t_i) = z(\mathbf{u}_\alpha,t_i) - \sum_{k=0}^{K} b_k(\mathbf{u}_\alpha) f_k(t_i), \quad i=1,\ldots,T_\alpha \qquad (5)$$

In this work, the $(K+1)$ station-specific temporal trend coefficients are defined via the algorithm adopted for their construction (e.g., OLS); these coefficients are treated as precise data.

2.2. Regionalizing temporal trend coefficients

Recall that temporal trend models $\{m(\mathbf{u}_\alpha,t), t\in T\}$ are established independently at each rain gauge location $\mathbf{u}_\alpha$. The resulting temporal trend model parameters $\{b_k(\mathbf{u}_\alpha), \alpha=1,\ldots,n\}$, $k=0,\ldots,K$, are spatially (cross)correlated since they are derived from the same process z-data, themselves correlated in space and time. Spatiotemporal interactions between the $(K+1)$ temporal trend components are characterized via the spatial (cross)correlation of the local trend model parameters.

In this work, a stochastic spatiotemporal trend model $M(\mathbf{u},t)$ is defined by viewing the set of $(K+1)$ trend b_k-coefficients as a joint realization of a set of $(K+1)$ cross-correlated RFs $\{B_k(\mathbf{u}), \mathbf{u}\in D\}$, $k=0,\ldots,K$, i.e.:

$$M(\mathbf{u},t) = \sum_{k=0}^{K} B_k(\mathbf{u}) f_k(t), \quad \forall\mathbf{u}\in D, \forall t\in T \qquad (6)$$

Estimation of the spatiotemporal trend reduces to the joint spatial prediction of the set of $(K+1)$ temporal trend coefficients $\{b_k^{**}(\mathbf{u}), \mathbf{u}\in D\}$, $k=0,\ldots,K$, at any location $\mathbf{u}\in D$ (the use of superscript **, which denotes an estimate as the superscript *, is justified below). Joint modeling is required to account for any cross-correlation between the b_k-coefficients. For example, a negative correlation between intercept and slope fields, $B_0(\mathbf{u})$ and $B_1(\mathbf{u})$, inherent to any line-fitting procedure, should be accounted for in spatial prediction.

Indeed, a set of $(K+1)$ estimated coefficient values $\{b_k^{**}(\mathbf{u}),\ \mathbf{u}\in D\}$, $k=0,\ldots,K$, would yield an estimate of the spatiotemporal trend field $\{m^*(\mathbf{u},t),\mathbf{u}\in D,t\in T\}$ over the space time domain, as:

$$m^*(\mathbf{u},t)=\sum_{k=0}^{K} b_k^{**}(\mathbf{u})\, f_k(t),\quad \forall \mathbf{u}\in D,\ \forall t\in T \qquad (7)$$

Spatial prediction of these coefficients is enhanced by considering relevant ancillary information, such as terrain elevation or lower-atmosphere variables derived from NCEP/NCAR reanalysis. For example, an initial estimate $b_k^*(\mathbf{u})$ of the unknown k-th coefficient $b_k(\mathbf{u})$ at location $\mathbf{u}$ is given by a a regression of the b_k-values derived at the raingauges on the collocated samples of L auxiliary variables; samples of the latter variables are assumed representative of an area equal to the cell size of the prediction/simulation grid.

More precisely, the *n* values of the *k*-th coefficient obtained at the *n* rain gauge locations are expressed as:

$$\mathbf{b}_k=\mathbf{G}\mathbf{q}_k+\mathbf{r}_k \qquad (8)$$

where $\mathbf{b}_k=[b_k(\mathbf{u}_\alpha),\alpha=1,\ldots,n]'$ is a $(n\times 1)$ column vector of samples of the k-th coefficient, $\mathbf{G}$ is a $(n\times(L+1))$ design matrix whose l-th column contains *n* values of the l-th auxiliary variable $\mathbf{g}_l=[g_l(\mathbf{u}_\alpha),\alpha=1,\ldots,n]'$, $\mathbf{q}_k=[q_k(\mathbf{u}_\alpha),k=0,\ldots,K]'$ is a $((L+1)\times 1)$ vector of coefficients, and $\mathbf{r}_k=[r_k(\mathbf{u}_\alpha),\alpha=1,\ldots,n]'$ is a $(n\times 1)$ column vector of residuals.

Once an estimate $\mathbf{q}_k^*$ of the vector $\mathbf{q}_k$ of regression coefficients is obtained by OLS, the regression prediction $b_k^*(\mathbf{u}_\alpha)$ for the *k*-th temporal trend coefficient $b_k(\mathbf{u}_\alpha)$ at any rain gauge $\mathbf{u}_\alpha$ is given as: $b_k^*(\mathbf{u}_\alpha)=\mathbf{G}\mathbf{q}_k^*$. The associated regression residual is then computed as: $r_k(\mathbf{u}_\alpha)=b_k(\mathbf{u}_\alpha)-b_k^*(\mathbf{u}_\alpha)=b_k(\mathbf{u}_\alpha)-\mathbf{G}\mathbf{q}_k^*$.

Residual r_k-values from the above regression procedure are most likely auto- and cross-correlated in space. Consequently, their spatial estimation calls for modeling the (cross)covariance matrix of the vector RF $\{R_k(\mathbf{u}),\mathbf{u}\in D\}$, $k=0,\ldots,K$ modeling the joint spatial correlation of these regression residuals. The geostatistical prediction algorithm of cokriging is adopted for this joint prediction task (Wackernagel, 1995). The simple cokriging (SCK) estimate $r_0^*(\mathbf{u})$ for the unknown intercept regression residual $r_0(\mathbf{u})=b_0(\mathbf{u})-b_0^*(\mathbf{u})$, for example, at any location $\mathbf{u}\in D$ is expressed as:

$$r_0^*(\mathbf{u})=\sum_{k=0}^{K}\mathbf{w}_k'\mathbf{r}_k \qquad (9)$$

where $\mathbf{r}_k=[r_k(\mathbf{u}_\alpha),\alpha=1,\ldots,n]'$ denotes the $(n\times 1)$ vector of regression residual values for the *k*-coefficient, and $\mathbf{w}_k=[w_{\alpha k}(\mathbf{u}),\alpha=1,\ldots,n]'$ the $n\times 1$ vector of cokriging weights assigned to these data for prediction of the regression residual $r_0(\mathbf{u})$ at location $\mathbf{u}$, and obtained per solution of the SCK system of equations:

$$\begin{bmatrix}\mathbf{C}_{00}\cdots\mathbf{C}_{0K}\\ \vdots\ \ddots\ \vdots\\ \mathbf{C}_{K0}\cdots\mathbf{C}_{KK}\end{bmatrix}\begin{bmatrix}w_{00}\\ \vdots\\ w_{0K}\end{bmatrix}=\begin{bmatrix}\mathbf{c}_{00}\\ \vdots\\ \mathbf{c}_{0K}\end{bmatrix} \qquad (10)$$

where $\mathbf{C}_{kk'}$ denotes the $n\times n$ matrix of auto or cross-covariance values between any pair of regression residuals $r_k(\mathbf{u}_\alpha)$ and $r_{k'}(\mathbf{u}_\beta)$, and $\mathbf{c}_{0k}$ denotes the $(n\times 1)$ vector of auto or cross-covariance values between any regression residual $r_k(\mathbf{u}_\alpha)$ and the unknown residual $r_0(\mathbf{u})$. Similar equations can be written for the spatial prediction of residuals related to other b_k-coefficients, i.e., for $k\neq 0$.

An estimate $b_k^{**}(\mathbf{u})$ of the unknown *k*-th coefficient $b_k(\mathbf{u})$ at any location $\mathbf{u}\in D$ is finally obtained as:

$$b_k^{**}(\mathbf{u})=b_k^*(\mathbf{u})+r_k^*(\mathbf{u}) \qquad (11)$$

and is then used in equation (7) to yield an estimated spatiotemporal trend component $m^*(\mathbf{u},t)$ at any location $\mathbf{u}\in D$ and for any day $t\in T$.

2.3. Simulation of space-time precipitation

The spatiotemporal residual *r*-values resulting from equation (5) are modeled as a realization of a stationary space-time process $\{R(\mathbf{u},t),\mathbf{u}\in D,t\in T\}$. Stochastic characterization of this process calls for modeling the spatiotemporal covariance of these *r*-residuals. In this work, this covariance is modeled as a sum of a purely temporal and a purely spatial component, plus a common spatiotemporal component. The latter captures stochastic space-time interactions via the definition of a generalized distance metric: $d=\sqrt{u_1^2+u_2^2+t^2}$, see Kyriakidis and Journel (1999) for details.

Simulation of the residuals in space and time proceeds by generating alternative realizations of the residual field $R(\mathbf{u},t)$ conditional on the residual data and their spatiotemporal covariance model. To this respect, sequential Gaussian simulation is used (Deutsch and Journel, 1998) for generating a S-member ensemble of residual realizations $\{r^{(s)}(\mathbf{u},t),\mathbf{u}\in D,t\in T\}$, $s=1,\ldots,S$.

A set of S simulated precipitation realizations $\{z^{(s)}(\mathbf{u},t),\mathbf{u}\in D,t\in T\}$, $s=1,\ldots,S$, is finally built by adding the single estimated trend $\{m^*(\mathbf{u},t),\mathbf{u}\in D,t\in T\}$ and the S simulated residual $\{r^{(s)}(\mathbf{u},t),\mathbf{u}\in D,t\in T\}$ fields. For a more elaborate procedure, which also accounts for the uncertainty in the estimated trend component $m^*(\mathbf{u},t)$, the reader is referred to Kyriakidis and Journel (2001). Note that any missing values in the rain gauge precipitation profiles are in-filled by simulation. The set of S alternative, equally probable, realizations $\{z^{(s)}(\mathbf{u},t),\mathbf{u}\in D,t\in T\}$ provide a model of uncertainty for the unknown precipitation levels in both space and time, which can be used for hydrologic impact assessment studies (Kyriakidis et al., 2001).

3. CASE STUDY

The study domain is a $300 \times 360 km^2$ area of the northern California coastal region, which is characterized by complex terrain and extreme seasonal variation in precipitation. Annual precipitation varies from 200mm/year in the Central Valley (east of the Coastal Range) to over 1300mm/year in the Santa Cruz Mountains (north of the Monterey bay). Precipitation in the region generally originates from stratiform clouds due to orographic lifting of the westerly flow over the western slope of the Coastal Range. Occasionally, strong convection embedded within the stratiform clouds generates intense local precipitation.

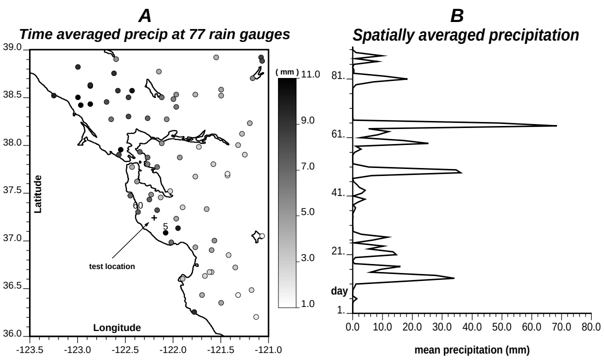

Figure 1: *Time average (from 11/01/1981 to 01/31/1982) of observed daily precipitation at 77 rain gauges (***A***), and space average of precipitation profiles for the same 92 days (***B***).*

The rainfall data set used in this study consists of 77 rain gauge precipitation measurements of daily rainfall during the 92 days from November 1 1981 to January 31 1982, see Figure 1. The original daily precipitation values constitute a subset of the Cooperative Observer (COOP) and first-order precipitation stations, obtained from the National Oceanic and Atmospheric Administration (NOAA); for details see Pandey et al. (1999). The proportion of rain gauge data above the threshold of 0.01mm (indicating a wet day) over all 92 days is 0.39. Wet-day precipitation amounts range from 0.25mm to 291.38mm, with a mean of 14.98mm and a median of 6.35mm indicating a positively skewed precipitation distribution. The standard deviation and coefficient of variation of the wet-day precipitation amounts is 23.88mm and 1.59, respectively, indicating a significant spatiotemporal variability. The objective of this study is to generate ensemble predictions of precipitation on a 300×360 grid of cell size $1km^2$ for the period 11/01/1981 to 01/31/1982, using all relevant information available for this region.

The first step is to establish a set of local temporal trend models of precipitation at each rain gauge, see Section 2.1. To this respect, two basis functions are used as temporal precipitation predictors at each rain gauge: $\mathbf{f}_0 = [f_0(t_i) = 1, i = 1,\ldots,92]'$, and $\mathbf{f}_1 = [\frac{1}{n}\sum_{\alpha=1}^{n} z(\mathbf{u}_\alpha, t_i), i = 1,\ldots,92]'$, see equation (4). In other words, the spatial average $\mathbf{f}_1$ of the precipitation profiles from the 77 rain gauges, Figure 1B, is used as the temporal precipitation predictor at each rain gauge. Two temporal trend coefficients are thus available at each rain gauge $\mathbf{u}_\alpha$: an intercept coefficient $b_0(\mathbf{u}_\alpha)$ and a slope coefficient $b_1(\mathbf{u}_\alpha)$, see Figure 2. Rain gauges with near zero intercept and near unit slope values (see the eastern part of the study domain and the south Bay Area) indicate precipitation profiles very similar to the spatially averaged profile $\mathbf{f}_1$.

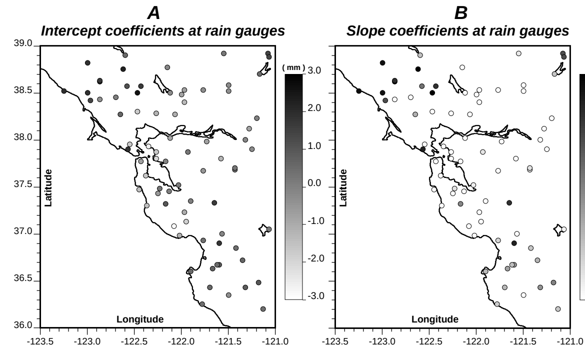

Figure 2: *Coefficients, intercept (***A***) and slope (***B***), of local temporal trend models established at the 77 rain gauges.*

A measure of the predictive ability of the spatially averaged precipitation profile $\mathbf{f}_1$ is the regression coefficient of determination (R^2) computed at each rain gauge. The spatial variability of these R^2-values is shown in Figure 3A. The average R^2 value is 0.58, with a minimum of 0.09 and a maximum of 0.87 (Figure 3B), indicating that the proportion of temporal precipitation variance accounted for by the spatially-averaged precipitation profile $\mathbf{f}_1$ changes significantly from one rain gauge to another. Precipitation profiles at rain gauges with high R^2-values (located in the northern part of the study area and in the Santa Cruz mountains) can be adequately characterized by a linear regression on the spatially-averaged profile $\mathbf{f}_1$.

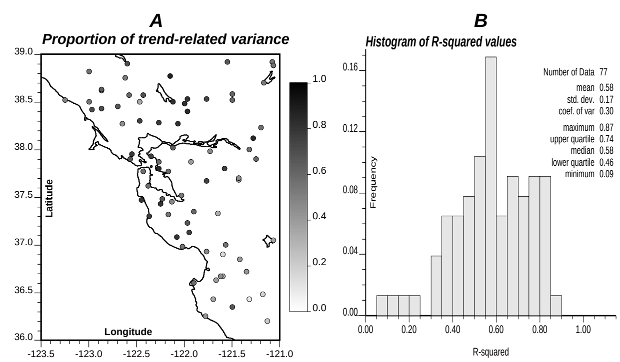

Figure 3: *Proportion of variance, as quantified by the regression coefficient of determination (R^2), of precipitation temporal variability accounted for by local temporal trend models at the 77 rain gauges (***A***), and histogram of R^2-values (***B***).*

Joint spatial prediction of intercept b_0 and slope b_1 coefficients at any location $\mathbf{u}$ in the study domain D is enhanced by accounting for their relation with terrain elevation and its interaction with specific humidity derived from NCEP/NCAR reanalysis data, see Section 2.2. A smoothed version of a United States Geological Survey (USGS) digital elevation model was used in this study. The smoothing window of $13 \times 13 km^2$ was determined by maximizing the correlation between time averaged precipitation (Figure 1A) and smoothed elevation, see Kyriakidis et al. (2002) for details. Time averaged specific humidity integrated over $850-1000 hPa$ was derived by interpolation from the 9 NCEP/NCAR reanalysis nodes closest to the study domain, and represents the large-scale availability of moisture in the lower atmosphere over the time span of interest.

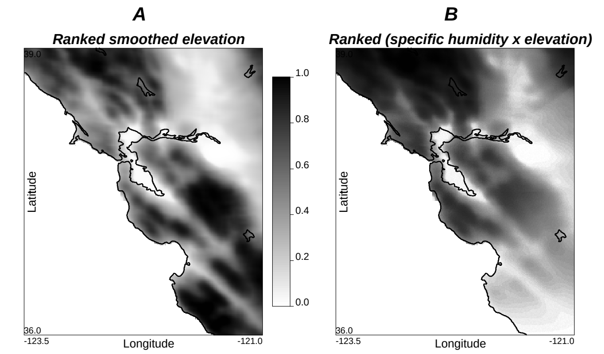

Figure 4: *Maps of rank-transformed window averaged elevation (**A**), and rank-transformed interaction of smoothed elevation with large-scale specific humidity derived from NCEP/NCAR reanalysis nodes (**B**).*

The rank transform of the window averaged elevation (Figure 4A) was used as an auxiliary variable in the spatial prediction of intercept b_0-coefficients. Similarly, the rank transform of the product (interaction) of specific humidity with the smoothed terrain elevation (Figure 4B) was used as an auxiliary variable in the spatial prediction of slope b_1-coefficients. The R^2-values for the regression of intercept b_0-coefficients (Figure 2A) on collocated rank-transformed smoothed elevation values (Figure 4A), and of slope b_1-coefficients (Figure 2B) on rank-transformed humidity-elevation interaction values (Figure 4B) were 0.1 and 0.34 respectively, see equation (8). Both regression models were statistically significant at the 95% level.

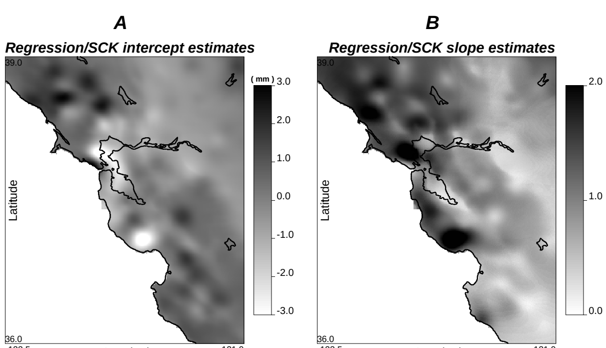

Figure 5: *Maps of estimated temporal trend coefficients, intercept (**A**) and slope (**B**), derived respectively by regression on elevation and its interaction with NCEP/NCAR specific humidity, followed by simple cokriging (SCK) of the resulting residuals.*

Simple cokriging was used for the joint spatial prediction of the resulting regression residuals r_0 and r_1, see Section 2.2 and equations (9) through (10). All auto- and cross-covariance functions of these residuals (not shown) were jointly modeled using the linear model of coregionalization (Wackernagel, 1995). The maps of estimated temporal trend coefficients, intercept b_0-values and slope b_1-values are shown in Figures 5A and B, respectively. Note that (co)kriging is an exact interpolator, which implies that regression residual r_k-values, hence temporal trend coefficient b_k-values, are reproduced at their respective rain gauge locations.

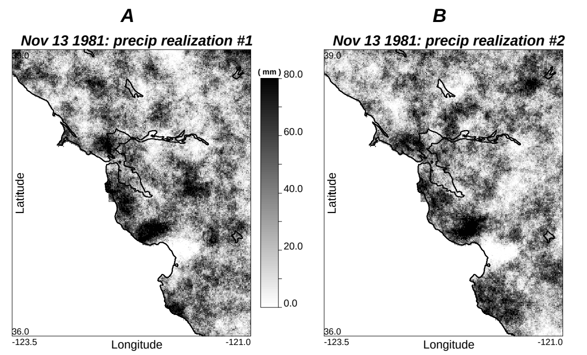

Figure 6: *Two (out of 30) synthetic precipitation fields for November 13 1981 (**A-B**) generated by conditional stochastic simulation.*

A set of 30 alternative realizations of daily precipitation over the 300×360 grid of cell size $1 km^2$ for the 92 days from November 1 1981 to January 31 1982 were generated using conditional stochastic simulation, see Section 2.3. Two of these realizations for November 13 1981 are shown in Figure 6. Conditioning entails that areas around high (low) rain gauge precipitation values (see Figure 8B) appear also as areas of high (low) precipitation in all simulated realizations.

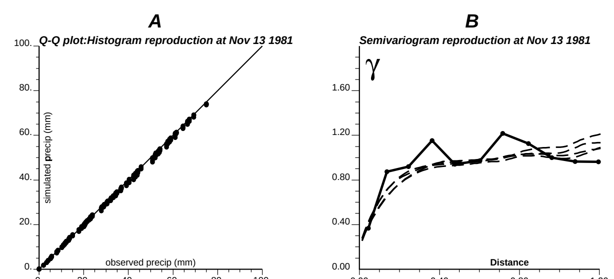

Figure 7: *Reproduction of observed precipitation histogram (**A**) and semivariogram (**B**) from five precipitation realizations for November 13 1981 (solid line: semivariogram of observed precipitation; dashed lines: semivariograms of simulated precipitation realizations).*

The reproduction of the rain gauge precipitation histogram for November 13 1981 by the histograms of five precipitation realizations is shown via the quantile-quantile plot of Figure 7A; a plot aligned along the first bisector implies two nearly identical distributions. The corresponding semivariogram reproduction is shown in Figure 7B; the sample precipitation semivariogram is well approximated by the semivariograms of the five precipitation realizations. Simulated daily precipitation realizations thus provide a realistic synthetic representation of the true (unknown) precipitation field, insofar they reproduce the histogram and semivariogram of observed rain gauge data.

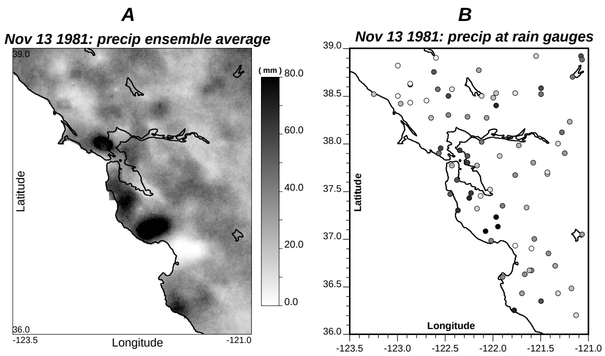

Figure 8: *Precipitation ensemble average for November 13 1981 (***A***) computed from 30 simulated realizations, and observed precipitation at 77 rain gauges for the same day (***B***).*

The ensemble average of simulated precipitation is also given in Figure 8A, along with the contemporaneous rain gauge data (Figure 8A) for comparison. It should be noted that this latter field does not reproduce the statistical properties (histogram, semivariogram) of the rain gauge data. It does reproduce rain gauge precipitation data at their locations, but provides a smooth picture of the spatial distribution of daily precipitation. Ensemble average fields should be used with caution in hydrologic impact assessment studies since they do not accurately depict the spatiotemporal variability of daily precipitation, an input of paramount importance in hydrologic modeling.

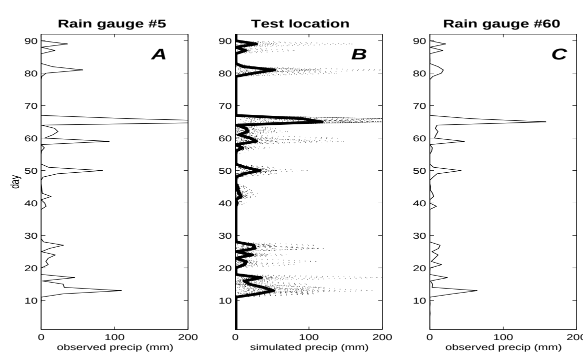

Figure 9: *Reproduction of observed precipitation variability at the test location shown in Figure 1.* **B:** *Thirty-member ensemble of simulated daily precipitation profiles at test location (dotted lines) and their ensemble average (thick solid line),* **A, C:** *precipitation profiles at nearby rain gauges #5 and #60.*

Last, we compare the simulated precipitation profiles at the test location shown in Figure 1, with precipitation profiles at two nearby rain gauges #5 and #60, all located in the same mountainous region. The set of thirty simulated profiles, and their ensemble average, at the test location is shown in Figure 9B. The precipitation profiles at the two nearby rain gauges are shown in Figures 9A and C, respectively. One can appreciate the similarity of the simulated precipitation profiles to the two rain gauge profiles. Note the common rainfall intermittence pattern exhibited by all profiles, and the similarity of the ensemble precipitation average profile (solid line of Figure 9B) to those of the nearby rain gauges. The average correlation coefficient between the simulated precipitation profiles and the precipitation profile of rain gauge #5 is 0.73 with a standard deviation of 0.16. Similarly, that average correlation coefficient with rain gauge #60 is 0.72 with a standard deviation of 0.16. The ensemble average precipitation profile has correlation coefficient 0.91 with the precipitation profile at rain gauge #5, and 0.89 with that at rain gauge #60.

This latter comparison of temporal profiles of simulated and observed precipitation corroborates the fact that daily precipitation realizations generated via the proposed methodology constitute a realistic synthetic representation of the true (unknown) precipitation field.

4. DISCUSSION

A framework for stochastic spatiotemporal modeling of daily precipitation in a hindcast mode has been presented in this paper. Observed daily precipitation levels are viewed as a joint realization of a collection of spatially correlated time series, thus capitalizing on the typically better informed time domain. The spatiotemporal daily precipitation field is decomposed into a stochastic trend and a stochastic residual component. Parametric temporal trend models are established at all rain gauges, independently from one location to another, and their parameters are (co)regionalized in space to yield an estimate of the space-time trend component at any location for any day. The joint spatial prediction of such temporal trend coefficients accounts for their relation with ancillary information, i.e., a smoothed version of terrain elevation and its interaction with large-scale specific humidity obtained from NCEP/NCAR reanalysis nodes. Simulated realizations of daily precipitation in space and time are obtained by generating alternative realizations of the spatiotemporal residual component and adding them to the estimated trend component.

The case study illustrated the generation of multiple synthetic realizations of daily precipitation on a 300×360 grid of cell size $1km^2$ over a region in northern California for 92 days during the period 11/01/1981 to 01/31/1982. Simulated precipitation realizations were shown to reproduce the histogram and semivariogram model of the rain gauge data. In addition, simulated precipitation profiles compared well with observed profiles at nearby rain gauges.

The set of alternative precipitation realizations constitutes a model of uncertainty regarding the unknown daily precipitation levels in both space and time. Such an uncertainty model can be used in a risk analysis context to study the effect of uncertain precipitation forcing on hydrologic impact assessment investigations.

References

Bras, R. L., and I. Rodríguez-Iturbe, 1985: *Random Functions and Hydrology*. Addison-Wesley, Reading, MA, 559 pp.

Deutsch, C. V., and A. G. Journel, 1998: *GSLIB: Geostatistical Software Library and User's Guide*. 2nd ed. Oxford University Press, New York, 368 pp.

Giorgi, F., and L. O. Mearns, 1991: Approaches to the simulation of regional climate change: A review. *Reviews of Geophysics*, **29**, 191–216.

Johnson, G. L., C. Daly, G. H. Taylor, and C. L. Hanson, 2000: Spatial variability and interpolation of stochastic weather simulation model parameters. *Journal of Applied Meteorology*, **39**, 778–796.

Kim, J., N. L. Miller, A. K. Guetter, and K. P. Georgakakos, 1998: River flow response to precipitation and snow budget in California during the 1994/95 winter. *Journal of Climate*, **11**, 2376–2386.

Kim, J., and S.-T. Soong, 1996: Simulation of a precipitation event in the western United States. *Regional Impacts of Global Climate Change*, S. J. Ghan, W. T. Pennel, K. L. Peterson, E. Rykiel, M. J. Scott and L. W. Vail, Eds., Battelle Press, 73–84.

Kyriakidis, P. C., and A. G. Journel, 1999: Geostatistical space-time models: A review. *Mathematical Geology*, **31**, 651–684.

Kyriakidis, P. C., and A. G. Journel, 2001: Stochastic modeling of atmospheric pollution: A spatial time series approach. Part I: Theory. *Atmospheric Environment*, **35**, 2331–2337.

Kyriakidis, P. C., N. L. Miller, and J. Kim, 2001: Uncertainty propagation of regional climate model precipitation forecasts to hydrologic impact assessment. *Journal of Hydrometeorology*, **2**, 140–160.

Kyriakidis, P. C., J. Kim, and N. L. Miller, 2002: Geostatistical mapping of precipitation from rain gauge data using atmospheric and terrain characteristics. *Journal of Applied Meteorology*, (in press).

Miller, N. L., and J. Kim, 1996: Numerical prediction of precipitation and river flow over the Russian River watershed during the January 1995 California storms. *Bulletin of the Americal Meteorological Society*, **77**(1), 101–105.

Pandey, G. R., D. R. Cayan, and K. P. Georgakakos, 1999: Precipitation structure in the Sierra Nevada of California during winter. *Journal of Geophysical Research*, **104**(D10), 12019–12030.

Searle, S. R., 1971: *Linear Models*, John Wiley & Sons, New York, 532 pp.

Seo, D.-J., S. Perica, E. Welles, and J. C. Shaake, 2000: Simulation of precipitation fields from probabilistic quantitative precipitation forecast. *Journal of Hydrology*, **239**, 203–229.

Thiébaux, H. J., 1997: The power of duality in spatial-temporal estimation. *Journal of Climate*, **10**, 567–573.

von Storch, H. and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge, 484 pp.

Wackernagel, H., 1995: *Multivariate Geostatistics*. Springer-Verlag, Berlin, 256 pp.

Wilks, D. S., 1998: Multisite generalization of a daily stochastic precipitation generation model. *Journal of Hydrology*, **210**, 178–191.