SENSITIVITY OF ECOLOGICAL MODELS TO THEIR CLIMATE DRIVERS: STATISTICAL ENSEMBLES FOR FORCING

Montserrat Fuentes,[1,5] Timothy G. F. Kittel,[2,3] and Doug Nychka[4]

[1]*Department of Statistics, North Carolina State University, Box 8203, Raleigh, North Carolina 27695 USA*
[2]*Institute of Arctic and Alpine Research, University of Colorado, Boulder, Colorado 80309 USA*
[3]*Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80526 USA*
[4]*National Center for Atmospheric Research, Boulder, Colorado 80307-3000 USA*

Abstract. Global and regional numerical models for terrestrial ecosystem dynamics require fine spatial resolution and temporally complete historical climate fields as input variables. However, because climate observations are unevenly spaced and have incomplete records, such fields need to be estimated. In addition, uncertainty in these fields associated with their estimation are rarely assessed. Ecological models are usually driven with a geostatistical model's mean estimate (kriging) of these fields without accounting for this uncertainty, much less evaluating such errors in terms of their propagation in ecological simulations. We introduce a Bayesian statistical framework to model climate observations to create spatially uniform and temporally complete fields, taking into account correlation in time and space, spatial heterogeneity, lack of normality, and uncertainty about all these factors. A key benefit of the Bayesian model is that it generates uncertainty measures for the generated fields. To demonstrate this method, we reconstruct historical monthly precipitation fields (a driver for ecological models) on a fine resolution grid for a climatically heterogeneous region in the western United States.

The main goal of this work is to evaluate the sensitivity of ecological models to the uncertainty associated with prediction of their climate drivers. To assess their numerical sensitivity to predicted input variables, we generate a set of ecological model simulations run using an ensemble of different versions of the reconstructed fields. We construct such an ensemble by sampling from the posterior predictive distribution of the climate field. We demonstrate that the estimated prediction error of the climate field can be very high. We evaluate the importance of such errors in ecological model experiments using an ensemble of historical precipitation time series in simulations of grassland biogeochemical dynamics with an ecological numerical model, Century. We show how uncertainty in predicted precipitation fields is propagated into ecological model results and that this propagation had different modes. Depending on output variable, the response of model dynamics to uncertainty in inputs ranged from uncertainty in outputs that matched that of inputs to those that were muted or that were biased, as well as uncertainty that was persistent in time after input errors dropped.

Key words: Bayesian inference; ecological numerical models; kriging; nonstationary process; spatial statistics; variogram.

Introduction

The prediction of a spatial surface from irregularly spaced observations that have a nonstationary spatial dependency structure (in the sense that the spatial correlation function is a function of location) is a common problem in ecology and geosciences. In earlier work, we constructed gridded monthly precipitation and temperature fields for approximately 100 years of the historical record for the coterminous United States (Kittel

et al. 2004). In this paper, we review standard methodology in spatial statistics and present some novel approaches to efficiently predict fine resolution spatial fields. We introduce a Bayesian framework to characterize different sources of uncertainty in the predicted spatial surface. A Bayesian analysis (e.g., Gilks et al. 1996) takes into account prior scientific knowledge and historical data to update (through Bayes theorem) the distribution of the spatial field of interest given all the available sources of information. In this work we present the use of Bayesian statistical ensembles as an alternative approach for characterizing spatial fields, rather than the mean estimate (e.g., kriging predictor) that is more commonly used. Each climate

ensemble member is an equally likely representation of what the actual climate could be accounting for model prediction error. Taken together, ensemble members are then an expression of uncertainty in the estimation of these fields. We evaluate whether this uncertainty has important implications for ecological model simulations that depend on such estimated fields as inputs. We propose that the use of input ensembles is a way to incorporate these errors in ecological model experiments. If ecological model responses to such ensembles are quite variable, our Bayesian framework is ideal for quantifying this variability and assessing the implications of input uncertainty in ecological model results. The spatial-temporal modeling and characterization of uncertainty in climate input fields and the evaluation of the propagation of this error in ecological model output are together the emphasis of this paper.

Alternative Bayesian statistical computer modeling approaches

There is an extensive literature about statistical computer modeling. In a series of papers, O'Hagan and his collaborators (e.g., O'Hagan et al. 1999) use a promising approach to characterize the uncertainty distribution of computer model outputs. In their approach, the output is treated as the variable of interest and it is modeled as a smooth surface with a stationary covariance. This smooth behavior is the property that gives the authors the opportunity to improve on Monte Carlo sampling. In the work by Sacks et al. (1989), the computer output is also modeled as a stationary smooth spatial process. In our approach, we model the input as a nonstationary spatial-temporal process and we compare the model output to an observed process of interest. We are trying to understand the properties and behavior of the model output. Therefore, at this stage we prefer not to make strong assumptions and statements about the smooth potential behavior of the model output. For complex computer models that can be run at different levels of sophistication, the approach introduced by Kennedy and O'Hagan (2000) can be used, so runs from several levels of a code could be used to make inference about the output from the most complex code. The computer model used in this paper is relatively simple, thus, this technique was not used.

Calibration and evaluation of computer models is another important and related area of research. Bayarri et al. (2002), Fuentes and Raftery (2005), and Fuentes et al. (2003), among others, present different Bayesian frameworks to evaluate complex computer models, including a careful study of different sources of uncertainty. Calibration consists of searching and estimating a set of values of the unknown inputs such that the observed data fit as closely as possible, in some sense, to the corresponding outputs of the numerical model. Kennedy and O'Hagan (2001) present a calibration approach for complex models using a Bayesian frame-

work. In our case the approach taken is different, the input at some location is estimated from the known values at other locations. We are avoiding fitting the data closely to the output of the model, since the model output might not be a good representation of the variable of interest. In fact, one of the aims of this study is to determine how realistic the model output might be when the uncertainty of the input is taken into account. The book by Santner et al. (2003) provides an excellent review of different methods and approaches to evaluate and characterize uncertainty in computer models.

STATEMENT OF THE SCIENTIFIC QUESTION

The main objective of this work is to evaluate whether errors in the spatial estimation of climate fields are small or large relative to the sensitivity of ecological models to their climate drivers. This would determine the value of generating statistical ensembles of climate data sets for regional and global ecological simulations such as the Vegetation/Ecosystem Modeling and Analysis Project (VEMAP) and similar efforts (e.g., Cramer 1997, Neilson et al. 1997). VEMAP was a multi-institutional, international ecological model intercomparison to assess the response of biogeography and biogeochemistry to environmental variability in climate and other drivers in both space and time domains across the coterminous United States (VEMAP Members 1995, Schimel et al. 2000, Gordon et al. 2004) that used extensively mean estimated climate fields as model drivers (Kittel et al. 2004).

In a statistical context, it is natural to construct statistical ensembles by sampling from the distribution of possible estimated climate fields given historical data (i.e., the posterior distribution based on model estimated prediction error; see, e.g., Gilks et al. [1996]). These ensembles then represent the uncertainty in the prediction of climate fields. Because the validity of the posterior distribution will depend on how well the spatial process has been modeled, it is very important to account for nonstationarity in space and time. We present in this paper an approach to obtain statistical ensembles for precipitation fields taking into account the complex spatial-temporal dependency structure of these climate drivers.

We evaluate the importance of this approach by comparing the effect on ecological simulations of using such an ensemble representation for model inputs over using the mean estimate. The mean estimate is the mean of the posterior distribution of climate field estimates given observed (i.e., station) climate values and under Gaussian standard assumptions. Determining the mean estimate is actually equivalent to an objective analysis procedure, kriging prediction, which is the process of interpolating data from irregularly spaced locations to a fixed grid. Kriging methods can be used to obtain estimates of point values as well as estimates of block
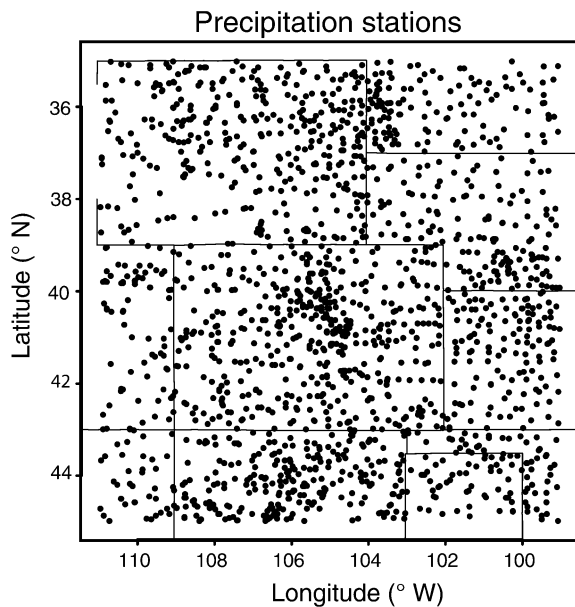
Precipitation stations

FIG. 1. Precipitation stations for the central and southern Rocky Mountains and adjacent Great Plains region, USA, for the period 1895–1996.

averages on a regular grid. To evaluate these two methods, we compare temporal and spatial patterns of the statistical ensembles and the kriging predictor for precipitation fields, especially under conditions of having years with a high number of missing input values. We assess whether these differences significantly affect the simulation of ecological processes. In particular, we consider whether uncertainty (as represented by ensembles) in the mean estimate of input fields is propagated in ecological model outputs with consequences for interpreting the reliability of such simulations.

Description of the climate data

The cumulative number of recording stations for precipitation exceeds 16 000 sites in the coterminous United States. For this study, we selected data for the 102 year period from 1895 to 1996 (National Climatic Data Center 1996). However, during this period not all sites operated at the same time, especially at the beginning when station densities were very low. For regional and global simulations, ecological models are often implemented on a regular grid, e.g., 0.5° longitude/latitude. The driving variables for these models, for instance, monthly precipitation, must then be available on the same grid without gaps in space or time. However, climate data are observed at irregularly spaced locations (see Fig. 1).

Fig. 1 shows the irregularly spaced distribution of precipitation sites for the central and southern Rocky Mountains and adjacent Great Plains region of the United States. Hereafter, we refer to this domain as the Rocky Mountains study area. The methodology pre-

sented in this paper was developed and applied to create fields for the coterminous United States. In this paper, we present results and focus on a smaller, climatically heterogeneous part of this domain, the Rocky Mountains study area. Precipitation fields for this area have been also reconstructed by Johns et al. (2003) using a different statistical framework blending spatial techniques, though without generating ensemble predictions. In 1895, only around 6% of the sites in the Rocky Mountains study area were operating, and by 1960, most sites were reporting. However, even in the 1990s, there are still a large number of missing values at any given time (15%).

Century model description

We used a terrestrial biogeochemical model, Century (Parton et al. 1994), to evaluate the use of climate field ensembles. Century is a monthly time-step model of carbon and nutrient states and flows originally developed for grassland and cropping systems (see, e.g., Parton et al. [1987]), but also implemented for woody systems including forest and savanna environments. Model comparison exercises such as that by VEMAP members (1995) and Smith et al. (1997) have shown Century to be among a group of highly effective simulation models. Century has been shown to be largely responsive to changes in climate, soil texture, and plant tissue chemistry. To run Century, we used 0.5° gridded soil characteristics from the VEMAP database (Kittel et al. 2004). Vegetation and other site parameters were set based on field data and estimates which had been applied before with success (Kelly et al. 2000). Gridded climate inputs were ensembles of monthly precipitation (developed in this paper) and kriging predicted monthly mean minimum and maximum temperature (Kittel et al. 2004) covering the 100-yr period. Century requires an initialization period to equilibrate soil organic matter pools to vegetation and management or disturbance patterns. Because soil organic matter represents a relatively large reservoir of carbon with a long turnover time, a 3000-year equilibration is standard for grassland systems.

TOOLS IN SPATIAL STATISTICS

Local semivariograms

In this section, we present a useful visualization tool for spatial problems to detect lack of stationarity, we call it the local semivariogram, which is different from the local relative semivariogram introduced by Cressie (1985a) to handle spatial heteroscedascity. First, we introduce some definitions and valuable tools in spatial statistics.

Definition of semivariogram.—Consider a spatial process $\{Z(\mathbf{x}), \mathbf{x} \in \mathbb{R}^2\}$ where $\mathbb{R}^2$ is the set of real numbers in two dimensions. The process Z is said to be second-order stationary, or weakly stationary, if the

mean is constant and the covariance function C satisfies

$$\text{cov}[Z(\mathbf{x}_1), Z(\mathbf{x}_2)] = C(\mathbf{x}_1 - \mathbf{x}_2) \quad \text{for all } \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2.$$

In general, the covariance is a function of location, i.e., $C(\mathbf{x}_1, \mathbf{x}_2)$, but for a second-order stationary process it is only a function of the relative distance between $\mathbf{x}_1$ and $\mathbf{x}_2$. The covariance is a useful tool to study the local spatial-dependency structure of a spatial process. Other distance-based measures of spatial correlation include the correlogram and semivariogram. The semivariogram measures the local spatial variation of a random field by describing how sample data are related to each other as a function of distance and direction and is a key parameter in geostatistics (Journel and Huijbregts 1978). The semivariogram is also needed to obtain the kriging predictor (see Cressie 1993). In general, a semivariogram shows increasing variance with distance, reflecting that two closely neighboring points are more likely to have similar values than two points farther apart. The semivariogram function γ is defined as

$$\gamma(\mathbf{v}) = \frac{1}{2} \text{var}[Z(\mathbf{x} + \mathbf{v}) - Z(\mathbf{x})] \quad (1)$$

where $\mathbf{v}$ is the vector distance (considering direction) separating two points in space. Where the covariance function, C, of the process Z exists, the semivariogram can also be written as

$$\gamma(\mathbf{v}) = C(0) - C(\mathbf{v}). \quad (2)$$

The observational or empirical semivariogram can be efficiently computed for data on a regular grid from the spectral density function (the Fourier transform of the covariance) using the fast Fourier transformation (FFT) and smoothed periodogram. Unfortunately, the FFT cannot be applied, at least easily, on most geophysical datasets, which are often characterized by unequally spaced sampling or incomplete grids. In these cases, we can use the traditional empirical semivariogram estimate $\hat{\gamma}$ suggested by Matheron (1971), which is

$$\hat{\gamma}(\mathbf{v}) = \frac{1}{2N(\mathbf{v})} \sum_{N(\mathbf{v})} [Z(\mathbf{x}_i) - Z(\mathbf{x}_j)]^2 \quad (3)$$

where $N(\mathbf{v})$ are the number of data pairs $\mathbf{x}_i$ and $\mathbf{x}_j$ separated by $\mathbf{v}$. In Fig. 2, we present the empirical semivariogram for the square root of the monthly precipitation values in the Rocky Mountains study area in February 1996.

A main goal of geostatistical analysis is to construct a semivariogram model that best estimates the autocorrelation structure of the underlying process. Journel and Huijbregts (1978) present several parametric semivariogram models. In this paper, we work with an exponential model:
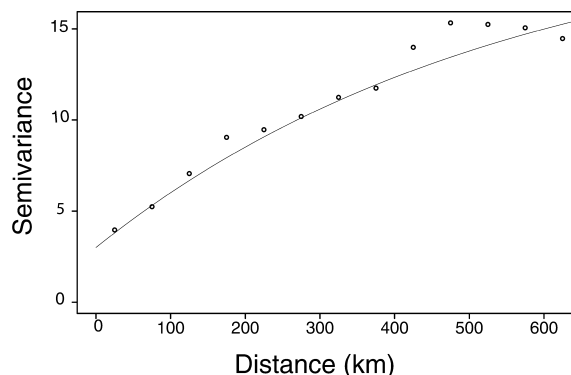

Fig. 2. An exponential semivariogram model (solid line) vs. distance (km), fitted to the empirical semivariogram values (circles) for the squared root of the monthly precipitation (in millimeters) in the Rocky Mountains study area in February 1996.

$$\gamma(h) = \begin{cases} 0 & \text{if } h = 0 \\ c_0 + (\sigma_0^2 - c_0)[1 - \exp(-h/r)] & \text{if } h > 0. \end{cases}$$
$$(4)$$

The parameter c_0 is the nugget ($c_0 \geq 0$), that represents intrinsic microscale variation or measurement error. It can be estimated from the empirical semivariogram as the value of $\gamma(h)$ for $h \to 0$, where h is distance. In cases where replicate samples exist and/or the variation in instrument accuracy is documented, this information is used to estimate the nugget more precisely. The parameter σ_0^2 denotes the sill, and it is the value of $\gamma(h)$ for $h \to \infty$, representing the variance σ_0^2 of the random field, i.e., at distances with no autocorrelation. The range parameter is r, and it is the distance at which the data are no longer autocorrelated at some specified level (this level depends on the model). The range for the exponential model is the distance at which the covariance drops $1/e$ of its value at zero. The difference $\sigma_0^2 - c_0$ is generally referred to as the partial sill. In Fig. 2, we fitted an exponential model to the empirical semivariogram values. The nugget parameter is 1.6 mm, the range is 300 km, and the partial sill is 11 mm. These parameters are estimated using weighted nonlinear least squares WNLS (Cressie 1985b). The WNLS method provides the best semivariogram model, in the least squares sense, that fits the empirical semivariogram values.

Visualization tools for spatial problems: local semivariograms.—The semivariogram is a useful tool to summarize the spatial continuity of certain random field. Fig. 2 is a common representation of a semivariogram function plotted versus distance in a region of interest. However, with this kind of graph, we do not get an appreciation of possible nonstationarity in space within the region.

We define a local semivariogram in a neighborhood L as follows:
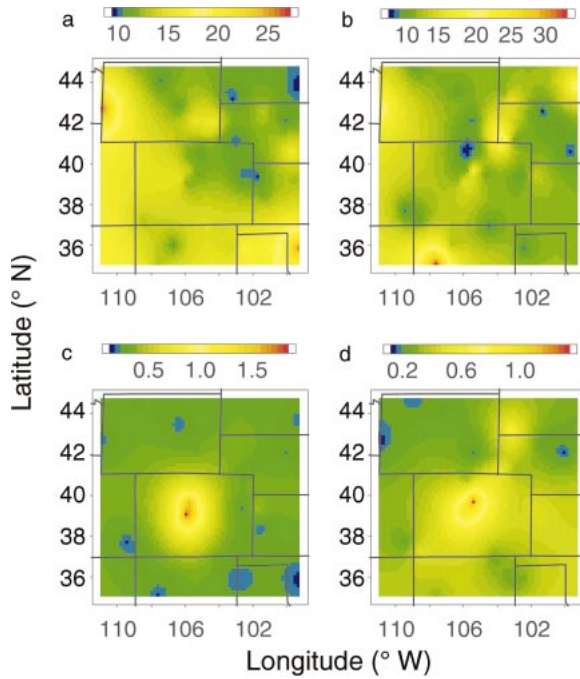
FIG. 3. Semivariograms for precipitation anomalies (units for the anomalies are $\sqrt{1/100}$ inches, 1 inch = 2.54 cm) and for standardized anomalies at a fixed distance of 40 km (with a tolerance of 5 km) in the Rocky Mountains study area. The semivariogram for the anomalies is shown for (a) January 1995 and (b) July 1996. The semivariogram for the standardized anomalies is shown for (c) January 1995 and (d) July 1996. We used a spline smoothing to smooth the image. The horizontal axis is longitude (° W), and the vertical axis is latitude (° N).

$$\hat{\gamma}_L(\mathbf{v}) = \frac{1}{2N_L(\mathbf{v})} \sum_{N_L(\mathbf{v})} [Z(\mathbf{x}_i) - Z(\mathbf{x}_j)]^2 \qquad (5)$$

where $N_L(\mathbf{v})$ are the number of data pairs $\mathbf{x}_i$, $\mathbf{x}_j \in L$, separated by a distance $\mathbf{v}$. We could add some tolerance vector ε to the distance $\mathbf{v}$ in case there are not enough points in L separated by exactly a vector distance $\mathbf{v}$.

A useful display to analyze the possible spatial non-stationarity in a region R is a map of the values of $\hat{\gamma}_L(\mathbf{v})$ for neighborhoods L_i which will cover the domain R, for a fixed value of $\mathbf{v}$.

Fig. 3a and b shows some local semivariograms for the Rocky Mountains study area, where we fixed the distance $\mathbf{v}$ to 40 km and tolerance ε to 5 km. We selected neighborhoods L_i in the Rocky Mountains study area, the neighborhoods L_i are circles with a radius of at least 0.3° (we increased the radius in some situations to have at least 25 observations in the neighborhood) centered in a regular grid of 0.5° × 0.5° in the region of interest.

The local semivariogram introduced here is different from the semivariogram surface (e.g., Gribov et al. 2005), which shows the value of the average squared differences for data points that are in a distance bin

corresponding to the distance to the center of the semivariogram surface graph and in a direction matching their location in the graph.

Computational methods for large spatial problems

In this section, we present an approach to efficiently compute the optimal spatial interpolator (e.g., based on a semivariogram model) for large data sets. The approach presented here uses an iterative algorithm to approximate the inverse of a large matrix.

Spatial prediction refers to predicting Z at a new location $\mathbf{x}_0$ from data $Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n)$ observed at known spatial locations $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Kriging is an optimal spatial predictor, in the sense that the kriging predictor minimizes the mean-squared prediction error,

$$E\left[Z(\mathbf{x}_0) - \sum_{i=1}^{n} \lambda_i Z(\mathbf{x}_i)\right]^2 \qquad (6)$$

for the class of linear predictors

$$\hat{Z}(\mathbf{x}_0) = \sum_{i=1}^{n} \lambda_i Z(\mathbf{x}_i) \qquad (7)$$

where λ_i for $i = 1, \ldots, n$ are the parameters, that satisfy $\Sigma_{i=1}^{n} \lambda_i = 1$. Because the kriging predictor is determined by the second-order covariance (semivariogram) properties of the process Z, the validity of kriging will depend on how well the semivariogram γ for the process Z has been modeled. To guarantee that the kriging predictor is optimal, in the sense that it minimizes the mean squared error, we need to assume that Z is a stationary Gaussian random field.

The value of the parameter $\lambda_0 = (\lambda_1, \ldots, \lambda_n)$ for the kriging predictor (Eq. 7) can be obtained (see, e.g., Cressie 1993:121) by solving the following linear equation (which is a function of the semivariogram):

$$\Gamma_0 \lambda_0 = \gamma_0 \qquad (8)$$

where Γ_0 is a symmetric $n \times n$ matrix,

$$(\Gamma_0)_{ij} \equiv \gamma(\mathbf{x}_i - \mathbf{x}_j)$$

$$\gamma_0 = [\gamma(\mathbf{x}_0 - \mathbf{x}_1), \ldots, \gamma(\mathbf{x}_0 - \mathbf{x}_n)].$$

Thus, Γ_0 and γ_0 are simply a function of the second-order moments of Z, as measured by the semivariogram. Because the parameter λ_0 is the solution of the linear system (Eq. 8), we have $\lambda_0 = \Gamma_0^{-1}\gamma_0$. However, for large data sets, obtaining the inverse of λ_0 is generally difficult.

We propose here to minimize the function $\varphi(\mathbf{x})$, defined by

$$\varphi(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \Gamma_0 \mathbf{x} - \mathbf{x}^\top \gamma_0. \qquad (9)$$

The minimum value of φ is $-\gamma_0^\top \Gamma_0^{-1}\gamma_0/2$, obtained by setting $\mathbf{x} = \lambda_0^{-1}$. Thus, minimizing φ and solving Eq.

8 are equivalent problems. The conjugate gradient algorithm is an iterative method used to minimize Eq. 9 (see, e.g., Golub and Van Loan 1989:516). If λ_0 is a $n \times n$ matrix the conjugate gradient algorithm will converge in n iterations to the solution. This method considers the successive minimization of φ along a set of directions $\{\mathbf{p}_1, \mathbf{p}_2, \dots \}$. If $\mathbf{x}_{k-1}$ is the current approximation to the solution in Eq. 9 for a direction $\mathbf{p}_k$, then $\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha\mathbf{p}_k$ will be the new conjugate gradient iterate, where we choose α to minimize Eq. 9. It is easy to show that to minimize $\varphi(\mathbf{x}_{k-1} + \alpha\mathbf{p}_k)$ with respect to α, we merely set

$$\alpha = \alpha_k = \mathbf{p}_k^\top(\gamma_0 - \Gamma_0\mathbf{x}_{k-1})/\mathbf{p}_k^\top \Gamma_0\mathbf{p}_k.$$

The convergence rate of the algorithm will depend on the starting vector. The method of conjugate gradients works well on matrices that are either well conditioned (with condition numbers near 1) or have just a few distinct eigenvalues (Axelsson 1985). If needed, we could precondition a linear system so that the matrix of coefficients assumes one of these nice forms (see Axelsson 1985).

Bayesian spatial prediction

Kriging is the name frequently used for spatial prediction, though as commonly used, that term refers only to the construction of a spatial predictor in terms of known semivariogram parameters. A Bayesian approach, on the other hand, takes into account uncertainty arising from estimating the covariance parameters used in the spatial prediction, and in that sense, it is more general than traditional kriging. The Bayesian approach leads to the same answers as the standard kriging predictor when the model parameters are known, but also extends to the case where these parameters are unknown.

The problem may be stated in the following form: given observations of a vector field $\mathbf{Z} = \{Z(\mathbf{x}_1), Z(\mathbf{x}_2), \dots, Z(\mathbf{x}_n)\}$, predict the value $Z(\mathbf{x}_0)$, for some $\mathbf{x}_0 \notin \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$.

The Bayesian approach, which leads to the same answers as the standard kriging predictor when the model parameters θ are known, but it also extends to the case where these parameters are unknown.

In the simplest case, where θ (covariance parameters) are all known,

$$\{Z(\mathbf{x}_0) \mid \mathbf{Z}, \theta\} \sim \mathcal{N} \text{ (normal distribution).} \quad (10)$$

Kriging is the expected value (mean) of this normal distribution. We shall now improve upon Eq. 10 by removing the conditioning on θ. We integrate over θ to obtain the posterior predictive distribution (given the data $\mathbf{Z}$):

$$p[Z(\mathbf{x}_0) \mid \mathbf{Z}] = \int p[Z(\mathbf{x}_0) \mid \mathbf{Z}, \theta] p(\theta \mid \mathbf{Z}) \, d\theta \quad (11)$$

where $p(Z(\mathbf{x}_0) \mid \mathbf{Z}, \theta)$ is the posterior predictive distri-

bution of Z at location $\mathbf{x}_0$ and $p(\theta \mid \mathbf{Z})$ is the posterior distribution of θ. We often use Monte Carlo methods to approximate this integral:

$$p[Z(\mathbf{x}_0) \mid \mathbf{Z}] \approx \sum_i^k p[Z(\mathbf{x}_0) \mid \mathbf{Z}, \theta^{(i)}]$$

where $\theta^{(i)}$ are simulated values from $p(\theta \mid \mathbf{Z})$.

All the statistical tools and methods presented in this section are used to conduct our exploratory analysis. Local empirical semivariograms are a powerful tool to understand the spatial structure of the process. Kriging and Bayesian techniques are used to model and predict data fields. Optimization methods and Markov Chain Monte Carlo algorithms simplify numerical solution of these spatial models when working with large data sets.

SPATIAL-TEMPORAL STRUCTURE OF THE INPUT FIELDS: EXPLORATORY ANALYSIS

We present here an exploratory analysis of precipitation data. First, we transform monthly precipitation data by taking the square root. This transformation was chosen based on a preliminary analysis that revealed an increasing linear relationship between the mean field and spatial and temporal variability for precipitation fields in the study area. Then, at each location, the transformed precipitation series is centered (converted into an anomaly series) by subtracting from transformed values their long-term mean seasonal cycle, obtained by taking the corresponding station's means for each month across its record (some sites have less than 102 years of data due to missing values). The seasonal mean pattern observed in precipitation contains most of the spatial behavior which is driven by topography and other physiographic processes. Also, by using anomalies we obtain higher spatial continuity (that is, a smoother field than just using raw precipitation values) and we focus our analysis on spatial processes that are linked to temporal (e.g., interannual) variability.

Monthly anomalies at a fixed location are not highly correlated in time, so we can focus just on their spatial structure for prediction. Spatial covariance structure, however, changes from month to month, which indicates the presence of a spatial-temporal interaction (e.g., Fig. 3a, b). To make the semivariograms stationary with time, we standardize the anomalies at each location by simply dividing the anomaly value by the corresponding monthly standard deviation. As with the calculation of long-term seasonal means, the monthly standard deviation is that of year values for each month at the station of interest across the corresponding record. Semivariograms for standardized anomalies are reasonably stationary over time (e.g., Fig. 3c, d); we observed that the same spatial structure persisted for periods of about 15 years. We did not find any evidence of anisotropy.
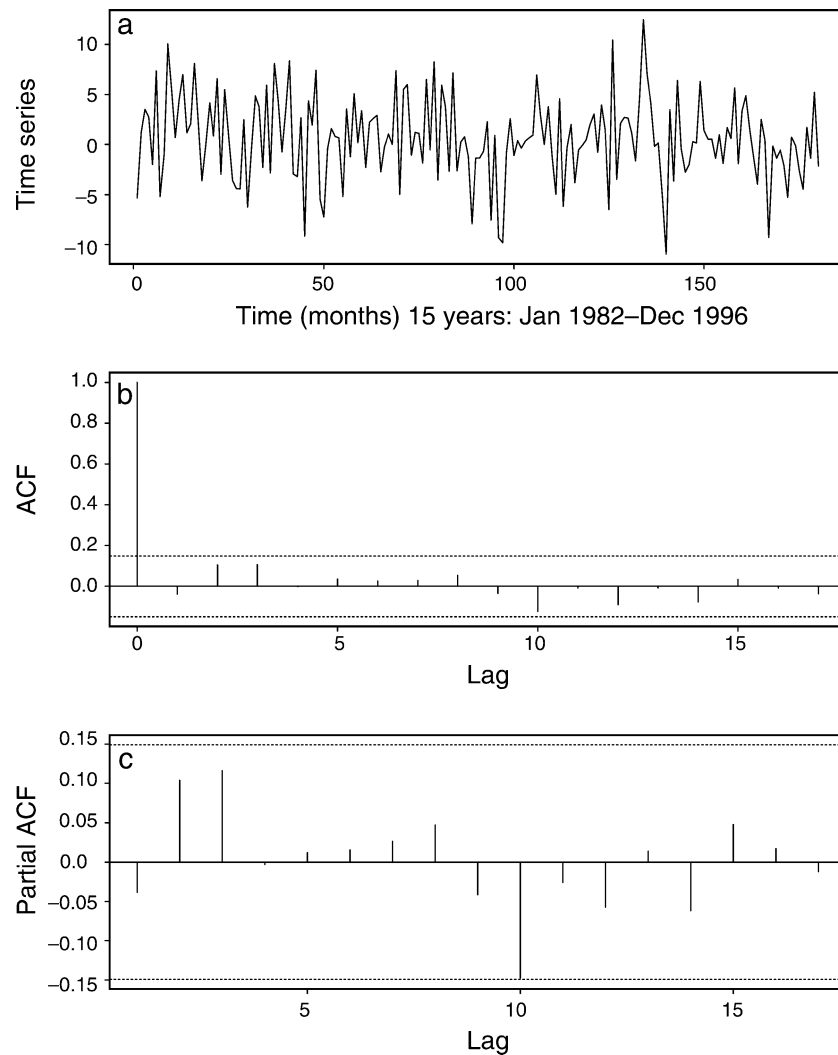
FIG. 4. (a) Time series for the anomalies (units $\sqrt{1/100 \text{ inches}}$; 1 inch = 2.54 cm) for a site in Colorado for 15 years (January 1982–December 1996). (b) Autocorrelation function (ACF) and (c) partial autocorrelation function for the same time series. The lag unit is one year. Dotted horizontal lines delineate the 95% confidence region.

Fig. 4 shows the time series and estimated temporal autocorrelation and partial autocorrelation functions for a site in Colorado for 15 years of data. The partial autocorrelation function at lag k is the correlation coefficient between $Z(t)$ and $Z(t + k)$, where $Z(t)$ is a temporal process at time t, after removing the linear dependence of $Z(t)$ and $Z(t + k)$ on the other intervening variables (the values of the process at other times). Both autocorrelation and partial autocorrelation values are within the 95% confidence regions. Thus, the data do not show any evidence of temporal structure for this time window. We observed similar behavior at other sites and for other 10–15 year time windows.

The standardized anomalies are not a spatial stationary process when time is fixed (i.e., when evaluated for any given month), as Fig. 3c and d suggests. However, it is a temporally stationary process for periods of about 15 years, so once we identify the spatial structure (using semivariograms) for one month, we can take it to be the same in a 15-yr time window. This simplifies considerably the calculations in the reconstruction of precipitation fields across time. Spatial nonstationarity in the standardized anomalies (e.g., Fig. 3c, d) was highly related to changes in elevation (Fig. 5). We can make use of this relationship in the spatial model that we construct in the next section. We also refer to Cramer (1997) for other exploratory analysis of climate drivers.

GENERATING STATISTICAL ENSEMBLES OF THE INPUT
FIELDS FOR ECOLOGICAL MODELS

We generate a statistical ensemble of different versions of the precipitation fields to represent uncertainty in the spatial modeling process. In the next section, we
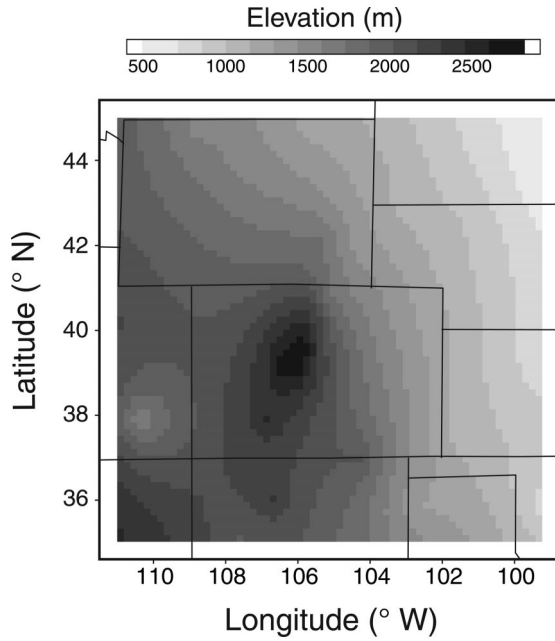
Fig. 5. Elevation (in meters) for the Rocky Mountains study area.

introduce a statistical model for precipitation (as the climate driver of interest) and then we explain how to generate the statistical ensembles.

Space-time statistical model

We developed a statistical model for precipitation fields supported by the exploratory data analysis. If $Z(\mathbf{x}, t)$ is the space-time process of interest, the anomaly of the square root of precipitation at location $\mathbf{x} = (x_1, x_2)$ and time t, we propose the following model:

$$Z(\mathbf{x}, t) = \mu(\mathbf{x}, t) + \sigma(\mathbf{x}) \times \varepsilon(\mathbf{x}, t) \qquad (12)$$

where μ is the spatially varying temporal trend (spatial-temporal trend) that explains the location-specific long-period (>15 years) nonstationarity with time, σ captures the time-independent spatial heteroscedascity (which is associated to elevation E), and ε is the spatially correlated error term.

The function σ is modeled as follows:

$$\log[\sigma(\mathbf{x})] = \beta_x E(\mathbf{x}) + \varepsilon_\sigma(\mathbf{x})$$

where the error term ε_σ is white noise and β_x is a spatially varying coefficient. To avoid overparametrization we model β_x as a linear function of location (invariant over time):

$$\beta_\mathbf{x} = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2$$

where α, α_1, and α_2 are unknown parameters. We use a normal distribution as the prior distribution for these three parameters.

The spatially correlated error term ε (in Eq. 12) is stationary in space but shows different variability over time (i.e., the sill changes over time):

$$\text{cov}[\varepsilon(\mathbf{x}, t), \varepsilon(\mathbf{y}, t)] = C(\mathbf{x} - \mathbf{y})\sigma_t^2 \qquad (13)$$

where σ_t is the sill and is time dependent. The covariance model used for ε is an exponential model. The prior distribution for the range parameter is a uniform distribution defined on [50 km, 500 km]. For the partial sill, we use an inverse prior, i.e.,

$$p(\sigma_t^2) \propto 1/\sigma_t^2.$$

For the nugget we use a uniform prior with support [0 mL, 10 mL]. The prior distributions are based on previous experience obtained by analyzing similar precipitation data sets.

The spatial-temporal trend is modeled using a space-time dynamic model:

$$\mu(\mathbf{x}, t) = \gamma_{\mathbf{x},t} + \gamma_t$$

where γ_t is an overall (spatially averaged) temporal trend, and $\gamma_{\mathbf{x},t}$ is the point-specific deviation from that trend, modeled as

$$\gamma_{\mathbf{x},t} = \frac{1}{|M(t)|} \sum_{t_i \in M(t)} Z(\mathbf{x}, t_i)$$

where $M(t)$, if for instance $t =$ January 1995, are all the January months over the 100-year period, and $|M(t)|$ is the cardinal of that set. The function γ_t is modeled as a linear function of time, because there was no evidence that more complex temporal structure was needed.

We use θ to denote the parameter vector with all the mean and covariance parameters in the statistical model in Eq. 12.

Bayesian spatial prediction: ensemble generation.— We sample values from the posterior predictive distribution:

$$p[Z(\mathbf{x}_0, t) \mid \mathbf{Z}]$$

at locations $\mathbf{x}_0$ on a fine-resolution grid ($0.5°$), where $\mathbf{Z} = \{Z(\mathbf{x}_1, t_1), \ldots, Z(\mathbf{x}_n, t_i), 1 \le T\}$, T is the number of observations over time, and n is the number of location sites. Each sampled value (at locations $\mathbf{x}_0$ of interest) constitutes a statistical ensemble of Z and is obtained using a multiple-stage Gibbs sampling approach (e.g., Gilks et al. 1996). We cycle through three stages. In stage 1, we estimate the posterior distribution of the covariance parameters for the error term ε. In stage 2, we estimate the parameters that explain the spatial heteroscedascity, σ. And in stage 3, we obtain the posterior distribution of the parameters that explain the mean of the process μ. Thus, we obtain $\{\theta^{(i)}\}_{i=1}^N$, which are N simulated values from the posterior distribution of the vector parameter θ. The posterior predictive distribution of process Z at the location $\mathbf{x}_0$ and time t_0 given all the available data $\mathbf{Z}$ is

$$p[Z(\mathbf{x}_0, t_0) | \mathbf{Z}] \propto \int p[Z(\mathbf{x}_0, t_0) | \mathbf{Z}, \theta] P(\theta | \mathbf{Z}) \, d\theta.$$

This posterior predictive distribution is approximated by the following (see Gelfand and Smith 1990):

$$p[Z(\mathbf{x}_0, t_0) | \mathbf{Z}] = \frac{1}{N} \sum_{i=1}^{N} p[Z(\mathbf{x}_0, t_0) | \mathbf{Z}, \theta^{(i)}].$$

The sample values from this posterior predictive distribution are obtained by conditioning on $\theta^{(i)}$, which for $i = 1, \ldots, k$ are k plausible values of the parameters in the statistical model. Thus, we approximate $p(Z(\mathbf{x}_0, t) | \mathbf{Z})$ using $p(Z(\mathbf{x}_0, t) | \mathbf{Z}, \theta^{(i)})$ as in *Bayesian spatial prediction.* The sample values from the posterior predictive distribution in Eq. 13 constitute a statistical ensemble of Z, in our case, the climate input for our ecological model runs. A corresponding ensemble of ecological model output is obtained by running Century for each simulated ensemble member of Z.

An ensemble of precipitation fields

In a geostatistical framework, reconstructed precipitation fields would be the kriging spatial predictor for the anomalies transformed back to the original scale at some locations of interest. In a Bayesian framework, we create a statistical ensemble of different versions of the anomaly fields, by simulating values from the posterior distribution for the anomalies. The transformation of the anomaly ensembles gives us a suite of reconstructed precipitation fields in the original scale. Figs. 6 and 7 show six statistical ensemble members for February 1996 and 1895 precipitation, respectively, for the Rocky Mountains study area.

Analysis of spatial and temporal variation of ensemble members

Given that the number of climate stations generally increased over time, most observations are missing at the beginning of our data set. Fig. 8 shows the observations (as anomalies) for February 1895 and February 1996 in the Rocky Mountains study area. Fig. 8 also shows the corresponding kriging point predictor for these months. The values on this graph are point predictions in a fine resolution grid ($0.5°$). This kriging predictor is expected to be more spatially homogeneous when we have fewer observations. In 1895, there were too few observations for the kriging predictor to capture the level of detailed spatial structure of the anomalies as found in 1996. The space-time model used here (Eq. 12) has a range of autocorrelation for the exponential semivariogram function that is invariant over time (see Eq. 13). Therefore, the spatial behavior of the predictive surfaces in Fig. 8 is not due to a change in the range parameters, because the semivariograms of the standardized precipitation anomalies used to generate the graphs in Fig. 8 had the same range of autocorrelation over time

(150 km). That is the reason why we do the spatial analysis and prediction of the stationary standardized anomalies and then we transform back to the original scale, rather than working directly with nonstationary precipitation values. The prediction error for the kriging predictor in 1895 at any fixed location should be much larger than in 1996, due to the lack of information in 1895 to obtain an accurate prediction. This is reflected in greater variability among ensemble members for any given month in 1895 (Fig. 9a) than in 1996 (Fig. 9b). The statistical ensembles are simulations from the posterior distribution. Thus, we should expect to observe less consistency in the spatial pattern among ensemble members in 1895 than in 1996. In Fig. 7 we can clearly appreciate the larger variation from statistical ensemble to statistical ensemble, due to the larger prediction error in 1895. On the other hand, the statistical ensembles in Fig. 6 are very similar, this suggests less prediction error in 1996.

Ensemble estimation versus standard kriging prediction

In this section, we compare the features of individual statistical ensemble members with the kriging predictor. Fig. 9 shows clearly the larger prediction error in early years for a grassland site in northeastern Colorado ($40.8°$ N, $104.8°$ W).

The top panels in Fig. 9 present the values of 10 statistical ensembles for each month in 1895 and 1996. There is more variability among statistical ensembles in 1895 than in 1996. However, the time series of the mean estimator (the mean of ensemble members; Fig. 9, bottom panel) for this site shows more temporal variability at the end of the dataset than at the beginning. In the VEMAP historical data set, this behavior was also common for kriging predicted climate fields early in the record in areas where station densities were very low (Kittel et al. 2004). This is expected because there are few observations at the beginning of the data set (Fig. 8, top left). The kriging predictor becomes more spatially heterogeneous at the end when there are more observations. This kriging predictor, the mean of the posterior, is approximately the same as the mean of the statistical ensembles, that are simulated values from the posterior. Fig. 10 presents monthly precipitation values transformed back from the statistical ensemble of square-root precipitation anomalies for this site. We have 11 versions of this input from the 10 statistical ensemble members and from the mean of the statistical ensemble. The bottom right time series shows the square root of the estimated mean squared error (MSE) for precipitation, calculated from the ensemble members. We can clearly appreciate in Fig. 10 that the largest errors, as represented by the MSE values, occur at the beginning of the time series when there are the fewest observations. Greater errors arise
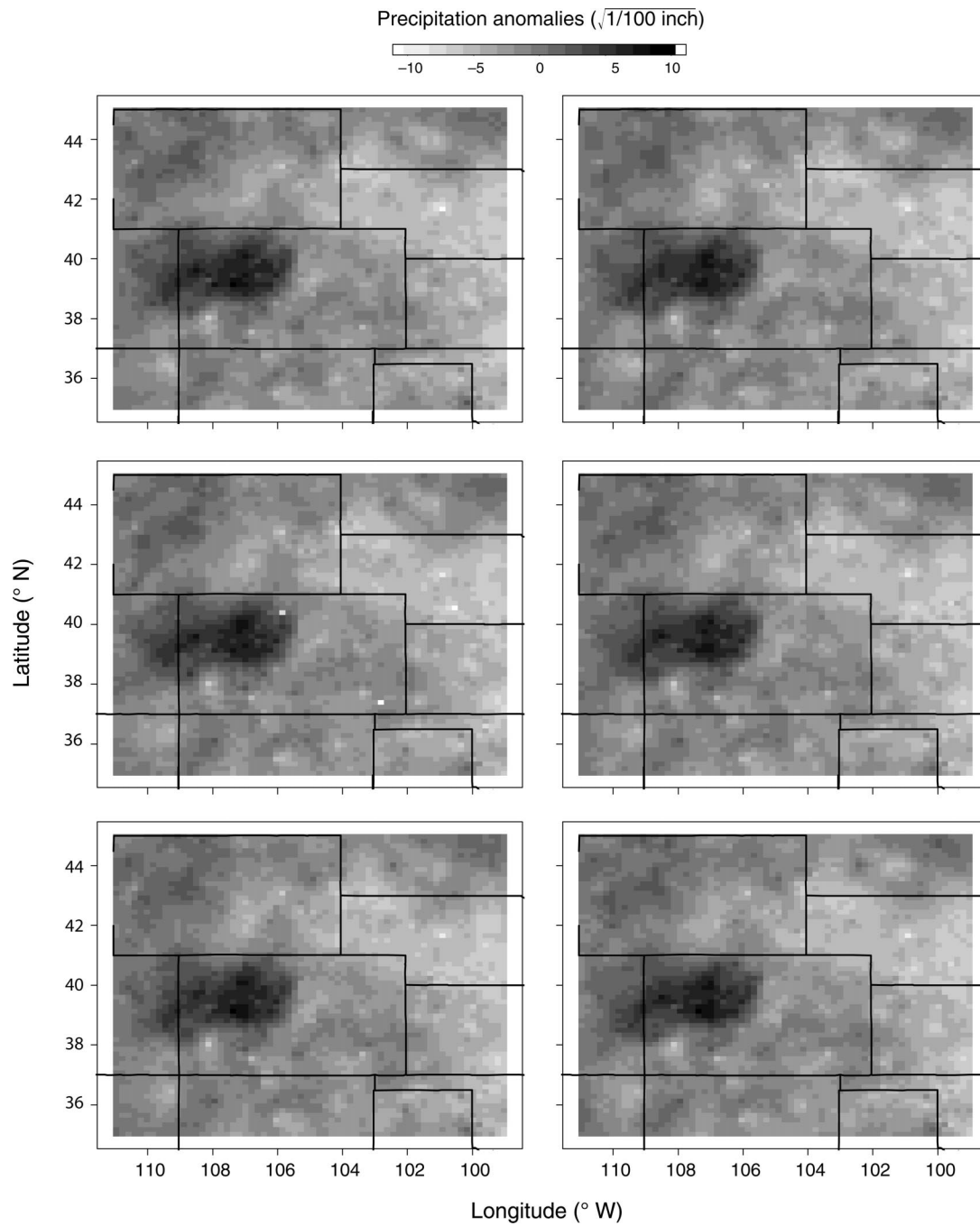
Precipitation anomalies ($\sqrt{1/100\ \text{inch}}$)
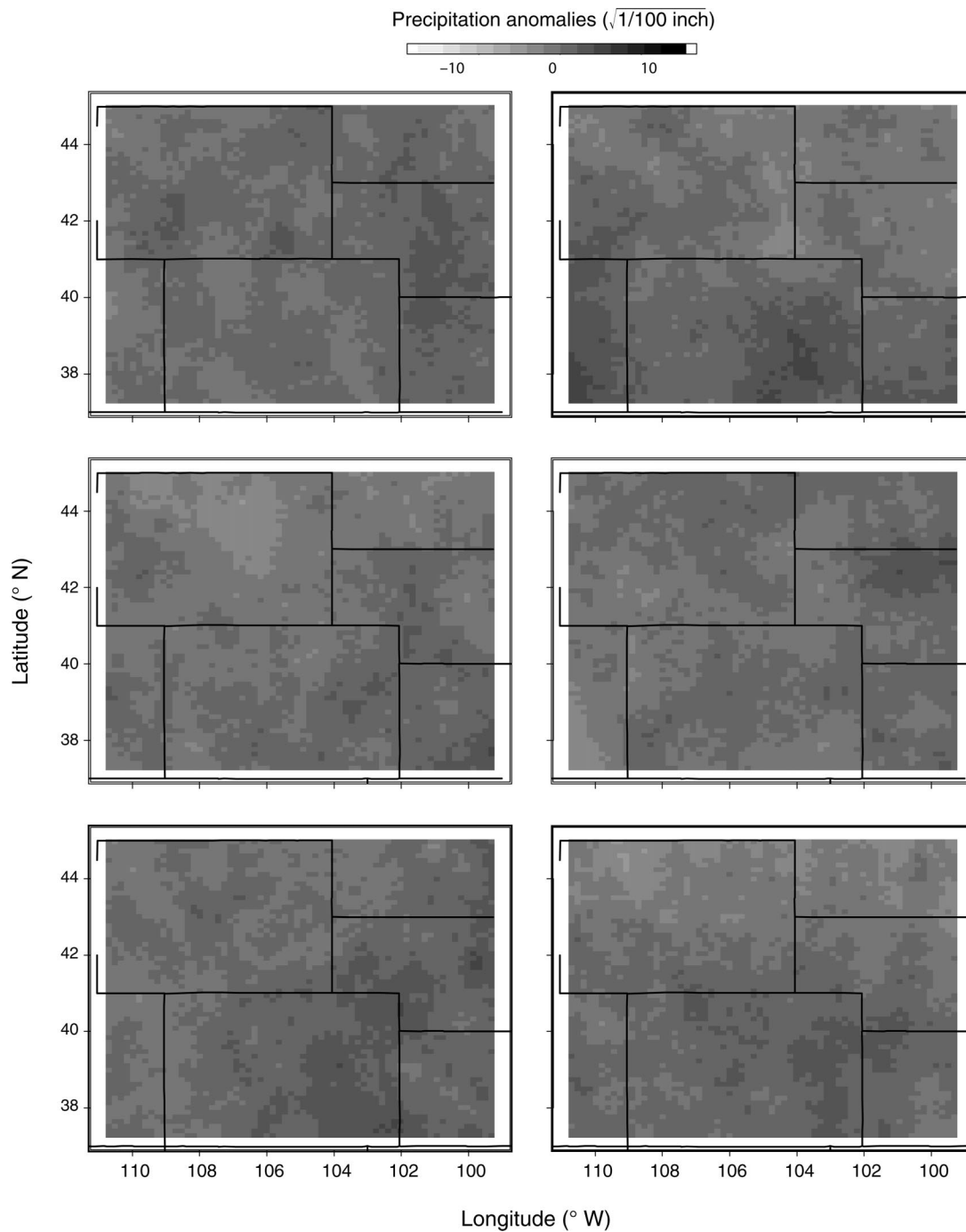

FIG. 6. Ensemble members for precipitation anomalies (units $\sqrt{1/100\ \text{inches}}$; 1 inch = 2.54 cm) for February 1996 in the Rocky Mountains study area.

when station density is low because the prediction model becomes based on a sparser, more dispersed network of observations, so that the signal from different climate variability regimes are mixed together to predict points in between. This mixing results in a dampened anomaly signal in the mean estimate, as seen in Fig. 9c. On the other hand, the individual ensemble members are more successful than the mean predictor at capturing interannual variability during the early record, as seen in the plots in Fig. 10 for the ensemble members. The trade-off is that there is instead a large uncertainty associated with each ensemble representation (Fig. 10, MSE plot). This reflects a major benefit of using ensembles.

FIG. 7. Ensemble members for precipitation anomalies (units $\sqrt{1/100 \text{ inches}}$; 1 inch = 2.54 cm) for February 1895 in the Rocky Mountains study area.

SENSITIVITY OF ECOLOGICAL MODELS TO CLIMATE FIELD PREDICTION ERROR

The spatial prediction error for reconstructed precipitation fields can be very large in some situations (Fig. 10). Thus, simulated responses by ecological models, such as Century, of carbon and nutrient dy-namics for different types of ecosystems could be in-accurate and lead to misleading conclusions about the historical behavior of an ecosystem. Therefore, it is crucial to assess the sensitivity of the ecological models to uncertainty in their inputs and then to determine what is the tolerance region for the prediction error of the
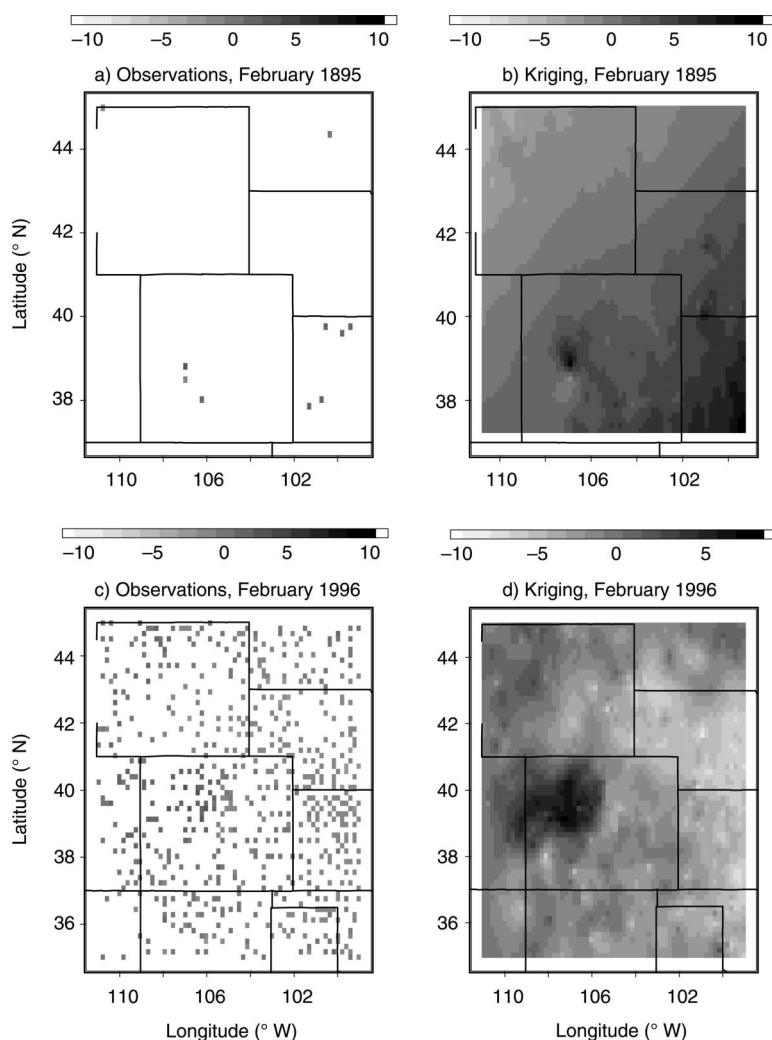
FIG. 8. (a) Observed values and (b) kriging predictor for precipitation anomalies (units $\sqrt{1/100 \text{ inches}}$; 1 inch = 2.54 cm) for February 1895 and (c,d) analogous plots for February 1996.

reconstructed fields. These regions will depend on the scientific questions being addressed by a given model experiment.

In this work, we focus on the Century model, the impact of spatial prediction error in precipitation fields, and simulated biogeochemical responses for the grassland site in northeastern Colorado. However, the same analysis could be done with other ecological models, other spatial inputs (e.g., temperature), and other locations and ecosystem types. We used mean estimators for other climate inputs required by Century, monthly mean maximum and minimum temperature (see *Century model description*). In addition to spatial prediction error considered here, we note that there are other sources of error in climate datasets that are associated with data collection and transcription (including instrumentation biases and station changes); these must also be considered in a full evaluation of uncertainty in ecological model inputs.

Research sites and output variables

The main goal is to study the sensitivity of the models in different ecosystems. Thus, we selected some sites representative of different climates and vegetation. In this paper, we only discuss model results for the grassland site in northeastern Colorado. This is the location of the USDA Central Plains Experimental Range and the Shortgrass Steppe Long-Term Ecological Research (LTER) site. Output variables from Century that we analyze here are net primary productivity, evapotranspiration, total stored carbon, vegetation carbon, and net nitrogen mineralization. These variables reflect the dynamics of different aspects of biogeochemical cycling in ecosystems and are each controlled by different complex interactions among plant and soil processes, often with different responses (and response times) to changes in moisture and temperature.
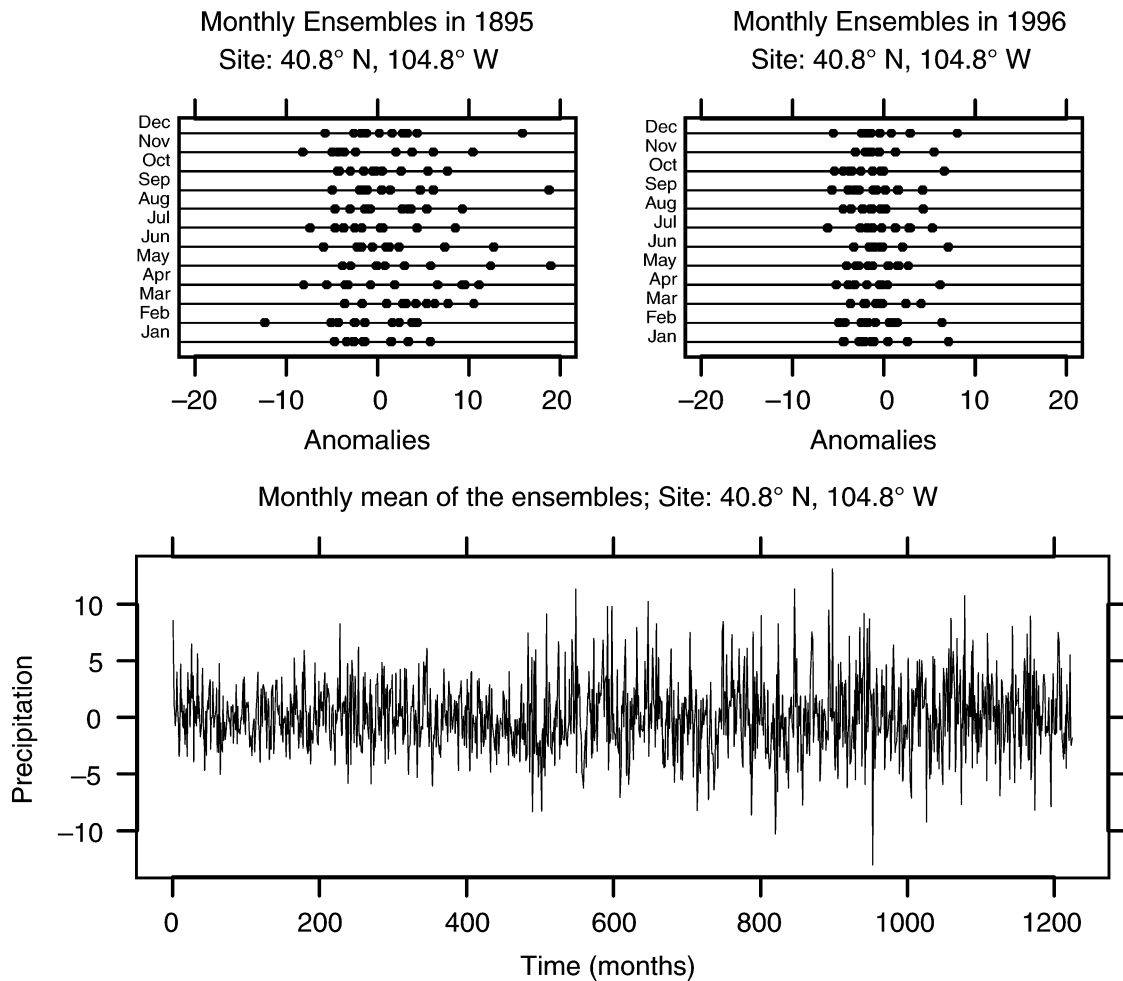
FIG. 9. (a) Monthly statistical ensembles (units $\sqrt{1/100}$ inches; 1 inch = 2.54 cm) at the grassland site in northeastern Colorado for 1895 and (b) monthly statistical ensembles at the same site for 1996. (c) Time series for the mean of the simulated statistical ensembles (units $\sqrt{1/100}$ inches; at the site in Colorado from January 1895 to December 1996.

Statistical analysis of the variability in the models output

For the Colorado grassland site, we ran the Century model on the 10 precipitation ensemble members and on the mean of the ensemble for 1188 months from 1895 to 1993 (the period was limited because the temperature data only ran through 1993) (Fig. 10). We study the sensitivity of the model outputs with respect to precipitation fields, by estimating the coefficient of variation (CV) of the outputs from the different statistical ensemble members. The CV is a statistical measure of the deviation of a variable from its mean (standard deviation divided by the mean) and is presented here in units of decibels (dB). The dB is a logarithmic unit used to describe a ratio, in this case the ratio of the standard deviation to the mean:

$$dB = 10 \log(CV)$$

where the log is base 10. The dB can describe very large ratios using numbers of modest size. In our application, due to the large variability in the CV values for the different variables, the logarithmic scale made the comparison easier. The result of this logarithmic basis for the scale is that increasing the CV by a factor of 10 raises its level by 10 dB; increasing it by a factor of 100 raises its level by 20 dB; by 1000, 30 dB; and so on. We present here how the CV (rather than the standard deviation) changes over time for different model outputs, with the objective of making more clear the comparison between different outputs in terms of understanding the temporal variability, because the outputs have very different means and different units.

Fig. 11 shows 11 versions of simulated net primary production at the Colorado site from the 10 different precipitation ensemble members and from the mean of the statistical ensemble. The bottom right panel in Fig. 11 shows the square root of the estimated MSE (SQRT MSE) for this variable. Fig. 12 shows the coefficient
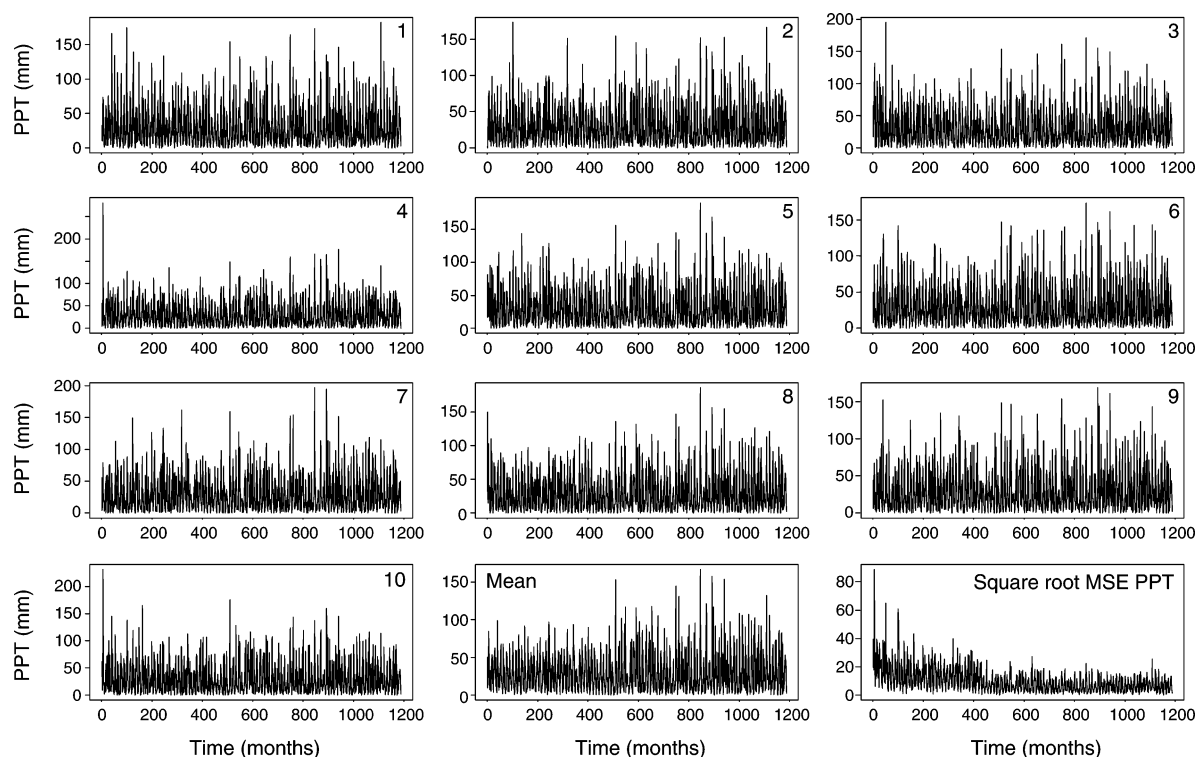
FIG. 10. Precipitation (PPT) input for the Century model for the site in Colorado (11 versions of this input from 10 statistical ensemble members [panels 1–10] of anomaly fields and from the mean of the ensemble). The bottom right panel shows the square root of the estimated mean squared error for PPT. Each graph is a time series with 1188 months of data (from January 1895 to December 1993). Note the different *y*-axis scales for PPT.

of variation across time for the five Century outputs we consider here and the precipitation field ensemble. The variability, over time, in the output variables is always higher at the beginning of the record, largely reflecting the influence of the instability (due to the lack of observations) of the ensembles during this period. We analyze in the next section the sensitivity of the output with respect to the variability in the predicted input.

Ecological model response to uncertainty in predicted precipitation.—As noted earlier, we see in Fig. 10 that the SQRT MSE for the precipitation ensemble is high early in the record relative to precipitation values. The coefficient of variation (CV) of precipitation varies around −2 dB at this time, or roughly 60%, and generally declines to the vicinity of −5 dB (∼30%) by the end of the record (Fig. 12). We evaluate the ecological model response to the high level of uncertainty in this input (and time dependence in this uncertainty) in terms of the following questions:

1) Do ecological outputs have comparable levels of uncertainty to that of the precipitation inputs, or are they amplified or muted?

2) Does uncertainty in ecological dynamic have a nonlinear, biased, or threshold response to uncertainty in inputs, such that there are disproportional responses at certain levels of uncertainty in precipitation?

3) Are there long-term effects of uncertainty in precipitation inputs early in the record, that are reflected in the latter part of the record?

The coefficients of variation for net primary production (NPP) and net nitrogen mineralization (NNM) are muted relative to that for precipitation, but follow a similar, proportional pattern of decline from −10 to −13 dB (10% to 5%, a decrease of roughly half as found for precipitation) (Fig. 12). The CV of evapotranspiration (EVAP) is at the same level as precipitation CV and its long-term decline is in proportion to that for precipitation (from on order of −2 to −5 dB; Fig. 12). This is expected given the strong dependence of EVAP on precipitation in this dryland ecosystem. At a fine temporal scale, however, evapotranspiration CV appears to have a nonlinear, skewed response with frequently repeating low values (to −20 to −30 dB or 1.0–0.1%). How this nonlinear response is tied to precipitation uncertainty requires a more detailed model analysis than we undertake here, but we note that EVAP response to precipitation in a given month can be strongly constrained by temperatures, which do not vary among ensemble runs.
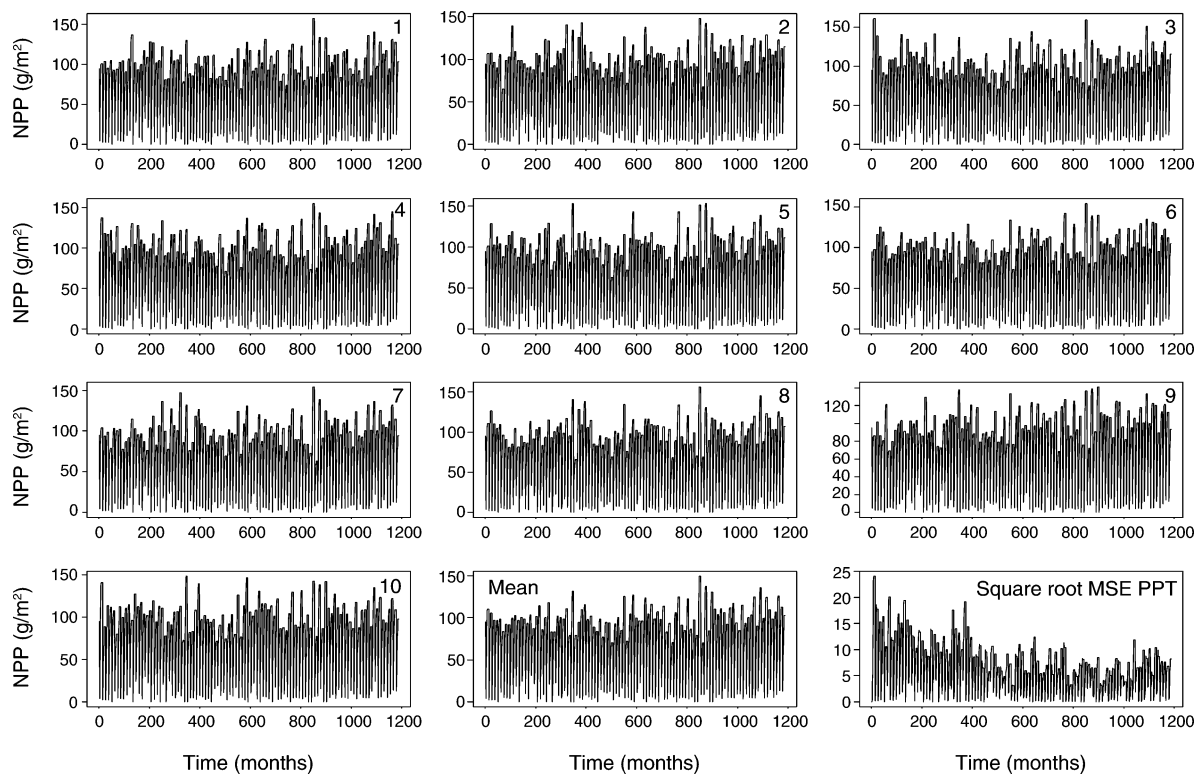
FIG. 11. Net primary production (NPP; units are g C·m^{-2}·mo^{-1}) output from the Century model for the site in Colorado (11 versions of this output from 10 statistical ensemble members [panels 1–10] of precipitation fields and from the mean of the ensemble). The bottom right panel shows the square root of the estimated mean squared error for NPP. Each graph is a time series with 1188 months of data (from January 1895 to December 1993). Note the different y-axis scales for NPP.

The overall pattern of vegetation carbon (VEGC) CV is similar to that of NPP and NNM (Fig. 12). This follows because NPP directly (and NNM indirectly) contributes to the accumulation of vegetation carbon. However, high frequency variability in vegetation carbon CV is much reduced. This likely reflects that VEGC (or plant biomass) is a slower-moving variable than production and mineralization, changing at seasonal and longer time scales, so that its response to precipitation uncertainty is going to be in these time frames.

The CV of total carbon (TOTC, which includes vegetation and soil carbon) shows a similar reduced high frequency response. However, the long-term pattern is quite different from that of precipitation and other response variables. Total carbon CV exhibits a decrease in the first half of the record which roughly follows that of precipitation, from on the order of −3 dB to −4.5 dB, but then stays in the range of −4 to −5 dB, rather than continuing to decrease as does precipitation CV. This is because in the first third of the period (up to month 400), when precipitation uncertainty is high (Fig. 10, MSE plot), TOTC traces of the 10 ensemble members are quite different (Fig. 13). As precipitation differences among ensemble members decrease the traces start to parallel each other (especially by the midpoint, month 600), but are offset because they have

different starting points coming out of the period of high uncertainty. After the midpoint, the CV stays in a set range because of the offset (Fig. 13). This indicates that critical differences in precipitation among ensemble runs early in the period sets the level of carbon accumulation for the rest of the period.

Analysis of variability in the model's output.—Using the various precipitation ensembles, Century effectively simulated ecosystem characteristics of the Central Plains Experimental Range in Colorado. Measured net nitrogen (N) mineralization is approximately 2.0 to 2.5 g N·m^{-2}·yr^{-1}, and our simulations range between 1.75 and 3.0 g N·m^{-2}·yr^{-1}. Simulated aboveground net primary production (NPP) is between 60 and 120 g C·m^{-2}·mo^{-1} and the field observations fall within the same range (Lauenroth and Sala 1992). Century estimates soil organic matter between 2200 and 2300 g C/m^2. Soil organic matter measurements generally encompass a wide range that includes those simulated by Century. Our simulations of total soil organic matter indicate a slow but steady increase that may indicate that the system was not completely equilibrated by the beginning of the simulation, rather than a response to climate drivers. Total system carbon, influenced primarily by the increase in soil organic matter (particularly the slow response soil organic matter pool), in-
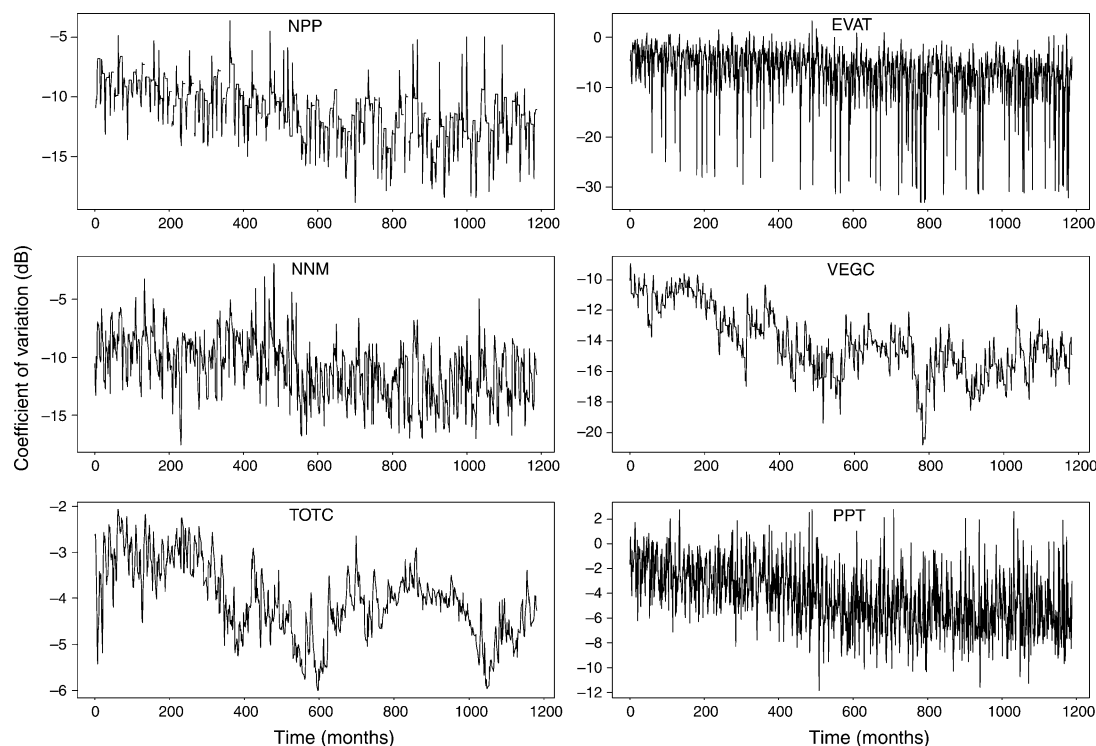
FIG. 12. The coefficient of variation, CV (in decibels) for the Century outputs: net primary production (NPP), evapotranspiration (EVAT), net nitrogen mineralization (NNM), vegetation carbon (VEGC), and total carbon (TOTC). The CV is a measure of relative dispersion and is given by the standard deviation divided by the mean. We also present the CV for precipitation ensemble members (PPT), inputs to the Century simulations. Each graph is a time series with 1188 months of data (from 1895 to 1993) for the site in Colorado.

creased over the period of simulation. Measurements of actual evapotranspiration (Lapitan and Parton 1996) range between 1 to 10 cm H_2O/mo, with annual peaks between 8 and 10 cm/mo. Results of the 11 simulations indicate sometimes higher peaks, but generally fall within the same range. Therefore, the suite of simulations all result in reasonable output when compared to observations. We surmise that Century gives a reasonable output (within expected ranges) despite variability in these simulations derived from uncertainty in predicted inputs (as driven by the precipitation ensemble).

SCIENTIFIC CONCLUSIONS AND FINAL REMARKS

From a statistical perspective, we would like to emphasize the effectiveness of the computational and visualization tools presented in this paper, such as the local semivariogram, to study nonstationarity patterns and then determine the validity of the posterior distribution used to simulate the ensembles. The flexibility of the Bayesian framework proposed here allows us to model and efficiently estimate different sources of uncertainty about the data and the parameters in the statistical model.

From a biogeosciences view, our key finding is that ecological model simulations driven by an ensemble

of estimated precipitation values show that uncertainty in this input can have subtle and not so subtle impacts on model response. The resultant uncertainty (i.e., the propagated error) in fast-response biogeochemical variables (net primary production, net nitrogen mineralization, and evapotranspiration) tended to track that of precipitation, in some cases with a reduced overall level of uncertainty and others with a skewed response. Uncertainty in slower-response variables (vegetation and total carbon) tended to have muted high-frequency responses relative to that of precipitation. In addition, total carbon response was sensitive to high precipitation uncertainty early in the record that was carried forward to periods of low uncertainty in predicted precipitation (Fig. 13).

This study shows how to estimate uncertainty in the prediction of spatial climate inputs and how these estimates can be used to elucidate uncertainty in resulting ecological simulations. Our results indicate that uncertainty in precipitation fields can be high (such as when station densities are low) with important consequences for interpreting the reliability of ecological model results. In contrast, the mean spatial estimate (i.e., the kriging predictor) of climate fields masks such input errors, precluding an assessment of error propagation, and also generates fields that poorly represent
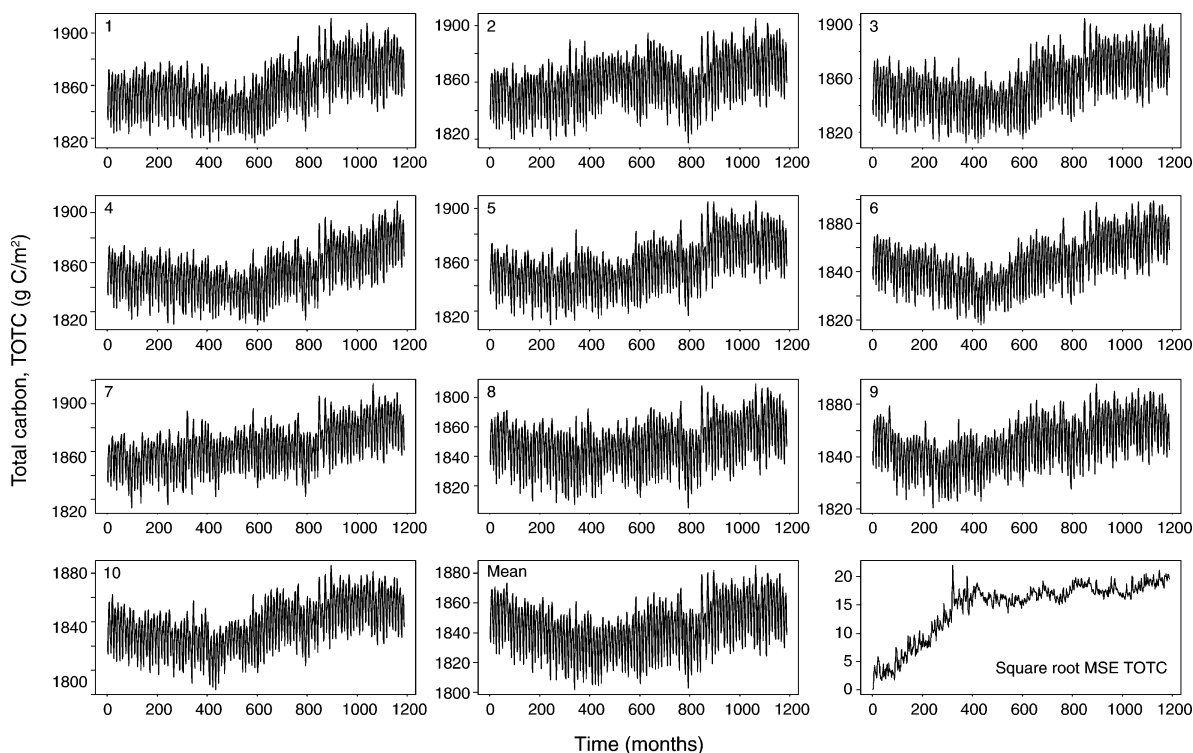
FIG. 13. Ensemble plots of total carbon (TOTC; units are g C/m^2). Each graph is a time series with 1188 months of data (from 1895 to 1993) for the site in Colorado.

spatial heterogeneity and temporal variability under conditions of low station density. Further work needs to be done to determine how many statistical ensembles of the input fields are needed to obtain a good representation of the model input posterior distribution. In situations in which we can assume that the model output is a smooth stationary spatial surface, the methodology proposed by Oakley and O'Hagan (2002) (see *Alternative Bayesian statistical computer modeling approaches*) can be used to reduce the total number of model runs needed. The appropriateness of this technique in our spatial-temporal setting is being investigated.

LITERATURE CITED

Axelsson, O. 1985. A survey of preconditioned iterative methods for linear systems of equations. BIT **25**:166–187.

Bayarri, M. J., J. O. Berger, D. Higdon, M. C. Kennedy, A. Kottas, R. Paulo, and J. Sacks. 2002. A framework for validation of computer models. Technical report number 128. National Institute of Statistical Science, Research Triangle Park, North Carolina, USA.

Cramer, W. 1997. Modeling the possible impact of climate change on broad-scale vegetation structure: examples from northern Europe. Pages 381–401 *in* W. C. Oechel, T. Callaghan, T. Gilmanov, J. I. Holten, B. Maxwell, U. Molau, and B. Sveinbjornsson, editors. Global change and arctic terrestrial ecosystems. Springer-Verlag, New York, New York, USA.

Cressie, N. A. C. 1985*a*. When are relative semivariograms useful in geostatistics? Journal of the International Association of Mathematical Geology **17**:693–702.

Cressie, N. A. C. 1985*b*. Fitting variogram models by weighted least squares. Journal of the International Association of Mathematical Geology **17**:563–586.

Cressie, N. A. C. 1993. Statistics for spatial data. Revised edition. John Wiley and Sons, New York, New York, USA.

Fuentes, M., P. Guttorp, and P. Challenor. 2003. Statistical assessment of numerical models. International Statistical Review **71**:201–221.

Fuentes, M., and A. E. Raftery. 2005. Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. Biometrics **61**:36–45.

Gelfand, A. E., and A. F. M. Smith. 1990. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association **85**:398–409.

Gilks, W. R., S. Richardson, and D. J. Spiegalhalter. 1996. Markov chain Monte Carlo in practice. Chapman and Hall, New York, New York, USA.

Golub, G. H., and C. F. Van Loan. 1989. Matrix computations. Second edition. The Johns Hopkins University Press, Baltimore, Maryland, USA.

Gordon, W. G., J. S. Famiglietti, N. A. Fowler, T. G. F. Kittel, and K. A. Hibbard. 2004. Validation of simulated runoff from VEMAP Phase II using observed streamflow: results from six terrestrial ecosystem models. Ecological Applications **14**:527–545.

Gribov, A., K. Krivoruchko, and J. M. Ver Hoef. 2005. Modified weighted least squares semivariogram and covariance model fitting algorithm. *In* M. J. Yarus and R. L. Chambers, editors. Stochastic modeling and geostatistics. AAPG computer applications in geology. Volume 2. *In press*.

Johns, C. J., D. Nychka, T. G. F. Kittel, and C. Daly. 2003. Infilling sparse records of spatial fields. Journal of the American Statistical Association **98**:796–806.

Journel, A., and C. Huijbregts. 1978. Mining geostatistics. Academic Press, New York, New York, USA.

Kelly, R. H., W. J. Parton, M. D. Hartman, L. K. Stretch, D. S. Schimel, and D. S. Ojima. 2000. Intra- and interannual variability of ecosystem processes in shortgrass steppe: new model, verification, simulations. Journal of Geophysical Research: Atmospheres **105**(D15):20093–20100.

Kennedy, M. C., and A. O'Hagan. 2000. Predicting the output from a complex computer code when fast approximations are available. Biometrika **87**:1–13.

Kennedy, M. C., and A. O'Hagan. 2001. Bayesian calibration of computer models (with discussion). Journal of the Royal Statistical Society, Series B **63**:425–464.

Kittel, T. G. F., et al. 2004. The VEMAP Phase 2 bioclimatic database. I: a gridded historical (20th century) climate dataset for modeling ecosystem dynamics across the conterminous United States. Climate Research **27**:151–170.

Lapitan, R. L., and W. J. Parton. 1996. Seasonal variabilities in the distribution of the microclimatic factors and evapotranspiration in a shortgrass steppe. Agricultural and Forest Meteorology **79**:113–130.

Lauenroth, W. K., and O. E. Sala. 1992. Long-term forage production of North American shortgrass steppe. Ecological Applications **2**:397–403.

Matheron, G. 1971. The theory of regionalized variables and its applications. Les Cahiers du Centre de Morphologie Mathématique, Fasc. 5, Centre de Géostatistique, Fontainebleau, France.

National Climatic Data Center. 1996. Summary of the month—cooperative. NCDC numeric data collection: TD 3220. U.S. Department of Commerce, Washington, D.C., USA.

Neilson, R. P., I. C. Prentice, B. Smith, T. Kittel, and D. Viner. 1997. Simulated changes in vegetation distribution under global warming. Pages 439–456 *in* R. T. Watson, M. C. Zinyowera, and R. H. Moss, editors. The regional impacts of climate change. An assessment of vulnerability. Annex C, A special report of the IPCC working group II. Cambridge University Press, New York, New York, USA.

Oakley, J., and A. O'Hagan. 2002. Bayesian inference for the uncertinaty distribution of computer model outputs. Biometrika **89**:769–784.

O'Hagan, A., M. C. Kennedy, and J. E. Oakley. 1999. Uncertainty analysis and other inference tools for complex computer codes. Pages 503–524 *in* J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, editors. Bayesian statistics 6. Oxford University Press, Oxford, UK.

Parton, W. J., D. S. Schimel, C. V. Cole, and D. S. Ojima. 1987. Analysis of factors controlling soil organic matter levels in Great Plains grasslands. Soil Science Society of America Journal **51**:1173–1179.

Parton, W. J., D. S. Schimel, D. S. Ojima, and C. V. Cole. 1994. A general model for soil organic matter dynamics: sensitivity to litter chemistry, texture and management. Pages 147–167 *in* R. B. Bryant and R. W. Arnold, editors. Quantitative modeling of soil forming processes. Special Publication. Soil Science Society of America, Madison, Wisconsin, USA.

Sacks, J., J. Welch, T. J. Mitchell, and H. P. Wynn. 1989. Design and analysis of computer experiments. Statistical Science **4**:409–435.

Santer, T. J., B. J. Williams, and W. Notz. 2003. The design and analysis of computer experiments. Springer, New York, New York, USA.

Schimel, D., et al. 2000. Contribution of increasing CO_2 and climate to carbon storage by ecosystems of the United States. Science **287**:2004–2006.

Smith, P., et al. 1997. A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments. Geoderma **81**:153–225.

VEMAP Members. 1995. Vegetation/ecosystem modeling and analysis project (VEMAP): comparing biogeography and biogeochemistry models in a continental-scale study of terrestrial ecosystem responses to climate change and CO_2 doubling. Global Biogeochemical Cycles **9**:407–437.