A Regression Strategy for Analyzing Environmental Data Generated by Spatio-Temporal Processes

Steele[1], B.M., Reddy[2], S.K., and Nemani[3], R.R

[1,2]Dept. of Mathematical Sciences, University of Montana, Missoula MT 59812

[3]Mail Stop: 242-4, NASA Ames Research Center, Moffett Field, CA 94035

Correspondence should be directed to B. Steele, Dept. of Mathematical Sciences, University of Montana, Missoula MT 59812, phone: 406-243-5396, fax: 406-243-2674, email: steele@mso.umt.edu

Abstract

Highly complex spatio-temporal environmental data sets are becoming common in ecology because of the increasing use of large-scale simulation models and automated data collection devices. The spatial and temporal dimensions present real and difficult challenges for the interpretation of these data. A particularly difficult problem is that the relationship among variables can vary in dramatically in response to environmental variation; consequently, a single model may not provide adequate fit. The temporal dimension presents both opportunities for improved prediction because explanatory variables sometimes exert delayed effects on response variables, and problems because variables are often serially correlated. This article presents a regression strategy for accommodating these problems. The strategy is illustrated by a case study of simulated net primary production (SNPP) that compares ocean-atmosphere indices to terrestrial climate variables as predictors of SNPP across the conterminous United States, and describes spatial variation in the relative importance of terrestrial climate variables towards predicting SNPP. We found that the relationship between ocean-atmosphere indices and SNPP varies substantially over the United States, and that there is evidence of a substantive link only in the western portions of the United States . Evidence of multi-year delays in the effect of terrestrial climate effects on SNPP were also found.

Keywords: spatio-temporal data, serial correlation, ARIMA models, NPP process models, carbon balance.

1. Introduction

To motivate our regression strategy for analyzing spatio-temporal environmental data, we begin by discussing connections between net primary production and ocean-atmosphere indices. Net primary production (NPP), the process of carbon sequestration via photosynthesis, appears to be an important component of atmospheric carbon variation (Myneni and others 1997; Potter and others 1999, Kimball and others 1997) and global climate change (Melillo and others 1993). Presently, relationships between NPP and climate across large areas are unclear (Nemani and others 2002; Potter and others 1999). There are several reasons for this lack of clarity. Arguably, the complexity of the relationship between climate and NPP is most troublesome because it greatly complicates describing of the relative importance of different climate variables over large areas. Another complication is that NPP is not directly observable over large areas. Two common approaches to the observational problem are to 1) use a remotely sensed surrogate variable such leaf area index in place of NPP, and 2) simulate NPP using a process model. While the surrogate variable approach offers the advantage of using real observations, data are available only for the past 20 years or so. Consequently, the scope of inference drawn from surrogate variables is limited.

In contrast, process models can simulate NPP far into the past, though the accuracy of the estimates depends on model realism, accurate input variables, and accurate land cover type identification. Process models simulate photosynthesis and respiration using nonlinear functions of site-specific variables such as average daily air and soil temperature, leaf area index, available soil moisture, and solar radiation (Hunt and others 1996; Ito 2002; Potter and Klooster 1997; Thornton and others 2002; White and others 2000). While SNPP process models are known and deterministic, they are also complex nonlinear functions involving numerous biophysical variables that vary spatially and seasonally. Consequently, it is difficult to succinctly describe the relative importance of the climate variables as they vary over large geographic regions. Yet this information is important for understanding carbon cycling and global climate processes. Melillo and others (1993) and Nemani and others (2002) have used process models and their estimates to investigate SNPP-climate relationships. These analyses were not statistical in nature, and so did not simultaneously estimate the effects of multiple variables on SNPP. Additionally, they were not able to test the significance of one or more variables after accounting for the effect of others.

Linear regression analysis of the relationship between SNPP and terrestrial climate variables is useful for several reasons. The linear form of the models allows immediate and unequivocal interpretation of the effect of each predictor variable on the response variable. The wealth of statistical methods associated with linear regression analysis provides for tests of significance, measures of model fit, confidence and prediction intervals, and methods for the analysis of residuals and leverage. Statistical models and methods for correlated ecological data have been presented by Epperson (2002) and Jones and Zang (1997), among others, and case studies analyzing spatially and temporally correlated environmental data collected over small areas are discussed in Nychka and others (1998). These articles do not address spatial variation in the relationship between response and predictor variables. In addition, strategies for accommodating seasonal variation are not discussed though seasonality is an important attribute of many biologic response variables.

Our approach to this problem analyzes SNPP using separate regressions for each location among a set of locations systematically sampled across the study area. This approach is quite general as it can be used for problems in which observations are directly observed over large areas and over time, such remotely sensed data, or problems in which the models generating the data are stochastic (for example, Keane and others 2002). In this case study, the regression analysis approach also permits investigation of temporal aspects of the relationship between SNPP and terrestrial climate that are not explicit components of the original NPP process model. Finally, a quantitative measure of the extent to which a linear terrestrial climate model can approximate SNPP is a useful benchmark against which to judge the information content of alternative variables such as ocean-atmosphere indices towards explaining variation in SNPP.

The article is organized as follows. In Section 2, we discuss the case study and our general approach to accommodating seasonal and spatial variation in modeling NPP. Section 3 presents the k-th order autoregressive regression model and its application to analyzing the relationship between NPP and predictor variables. We also discuss methods of assessing model fit and measures of trend. Section 4 presents results of the case study. Section 5 compares recent empirical NPP modeling methods to the regression strategy presented herein, and the article concludes in Section 5.

2. The Case Study

Evidence of connections between terrestrial and ocean climate has been presented by Dai and others (1997), Dima and others (2001), Hurrell (1995), and Hurrell (1996). Ocean climate effects on terrestrial climate appear to be delayed in some situations (Dima and others 2001; Hurrell 1995; Potter and others 1999; Sutton and Allen 1997). Connections between ocean climate and vegetation production have been addressed by Anyamba (1994), Anyamba and Eastman (1996), Cane and others (1994) and Myneni and others (1996). With respect to the conterminous United States, Li and Kafatos (2000) and Mennis (2001) present evidence of correlation between seasonal means of normalized difference vegetation index and equatorial Pacific sea temperatures between 1982 and 1992. Links between ocean-atmosphere indices and terrestrial climate suggest similar links between ocean-atmosphere indices and SNPP (Nemani and others 2002), though the extent and strength of these links are largely unknown. Investigation into the connections between SNPP and ocean-atmosphere indices are motivated by the need to further understand large-scale ecosystem interactions.

The objectives of the case study were to 1) compare the extent and strength of the relationship between SNPP and two sets of predictors, ocean-atmosphere indices and terrestrial climate, and 2) investigate the effect of scale on model fit for the two sets of predictors. Our approach was to empirically model SNPP as a function of three terrestrial climate variables (temperature, precipitation and solar radiation), and of three ocean-atmosphere indices series [the Southern Oscillation Index (SOI), the Pacific Decadal Oscillation

(PDO), and the North American Oscillation (NAO)], and compare model fit. The analysis was conducted at two scales; the finer scale data consisted of a sample of 135 pixels from a ½° lattice cover of the United States, and the coarse scale data consisted of seven regional averages computed from the pixel-scale data. The pixel-scale terrestrial climate variables pose a rigorous standard against which to compare ocean-atmosphere indices because the terrestrial climates variables are spatially explicit inputs to the process model. In contrast, the ocean-atmosphere indices indirectly measure climate at great distances from the pixel locations, and so are, at most, moderately associated with terrestrial climate (Hurrell 1995, 1996).

The study area is the conterminous United States and the study period is 1901 to 1993. This area is large enough to be an important source of carbon on a global scale; moreover, long-term climate data is available from the United States Historical Climate Network (USHCN) (Knapp and Smith 2001), and NPP process model estimates are available. In particular, we obtained daily climate and NPP process model estimates from the BIOME-BGC simulation model (Kittel and others 1997; Schimel and others 2000; Thornton and others 2002) for 1901 through 1993. The BIOME-BGC model simulates terrestrial ecosystem carbon, nitrogen, and water cycles for each of 3168 pixels in a ½° lattice coincident with the VEMAP terrestrial climate lattice (Kittel and others 1995). A model of daily leaf canopy photosynthesis is used to estimate carbon absorption by vegetation. Total respiration is simulated as the sum of maintenance and growth respiration estimates, and the difference between gross carbon uptake and total respiration is SNPP. To reduce the computational demands of using all 3168 pixels, we systematically sampled the ½° lattice by projecting a 2° lattice over ½° lattice cover and selecting those ½° pixels with centers that coincided with 2° lattice centers (Figure 1). A significant portion of several sampled pixels were predominantly covered by water; these pixels were replaced by nearby pixels that were not appreciably covered with water. Lastly, we refer to the SNPP data for a particular pixel, or a region, as a *series* in recognition of the possibility that the data obtained from a particular pixel or region may be serially correlated.

2.1. Spatial and Seasonal Variation in SNPP

The relationship between NPP and terrestrial climate varies across the conterminous United States. For example, the onset of the growing season varies with latitude, as does the effect of March precipitation on NPP. Consequently, the analysis of SNPP at a particular location depends in part on identifying a time period that encompasses most of the total annual net production of carbon, but also is short enough that the relationship between predictor variables and SNPP is nearly time-invariant throughout the period. We determined these periods, henceforth referred to as growing seasons, in two stages. First, an exploratory analysis of spatial variation in calendar monthly means of absolute SNPP was conducted. The results are described immediately as they provide motivation for the methods that follow. A star plot (Johnson 1998) shows that absolute NPP is greatest in the eastern half of the United States and that maximum absolute NPP occurs in June and July in the north, and roughly a month earlier in southern regions (Figure 2). SNPP decreases rapidly after midyear except in Florida where monthly mean SNPP is relatively constant from May through October. Principal components analysis of between-pixel variation in monthly mean SNPP indicated that five different sources account for most (82.2%) of the spatial variation in SNPP. Star plots showing the eigenvector coefficients, or principal component loadings, for the five most dominant eigenvectors were constructed, and adjacent pixels with similar eigenvector coefficients were aggregated to form preliminary regions.

The second stage of defining growing season refined the preliminary regions through an iterative process of aggregating pixels as regions, assessing the fit of regional regression models of growing season SNPP, and reassigning pixels to regions to improve model fit. The predictor variables used in the regression models were growing season means of precipitation, temperature and solar radiation, and previous winter and previous summer means of precipitation, temperature and solar radiation. Hence, there were $3 \times 3 = 9$ predictor variables and $93 - 1 = 92$ observations for each model. These SNPP models were used only for assessing model fit; more thorough analyses involving variable selection, analysis of residuals, and interpretation of model coefficients were carried out only after the regions were finalized. Seven regions were identified, and are henceforth referred to as the Interior West, Pacific Coast and Southwest, Great Plains, Mid-Atlantic, Northeast, Southcentral, and Southeast (Figure 1). The regional definitions of seasons are shown in Table 1, and calendar monthly means are plotted against calendar month in Figure 3. Those months that were excluded from a regional growing season are collectively referred to as a winter season. Seasonal patterns are generally similar throughout the conterminous U.S., with most regions undergoing rapid increases in monthly mean NPP between April and May, and declines from July onward. The monthly standard deviations of SNPP exhibit a pattern similar to the means (small values in the winter months relative to the summer months). For instance, for the Southwest and Northcentral regions, the January standard deviations were 0.066 and 0.030 (g /m^2 /day), whereas the June values were 1.790 and 2.018 (g /m^2 /day). Because SNPP monthly means are nearly constant in the winter, we infer that winter climate has little immediate effect on SNPP over much of conterminous United States. This observation implies that fitted relationships between annually measured biophysical variables and annual climate summaries are potentially subject to a loss of accuracy because winter climate variation becomes part of annual climate variation, yet winter climate may not be directly connected to the biophysical response variable.

The regional and pixel SNPP series were standardized for regression analysis so that the parameter estimates obtained for different locations would be comparable. Specifically, the SNPP seasonal series for each pixel consisted of total growing season SNPP (g/m^2) for the 93 years of observation. These totals were standardized using the sample mean and standard deviation of the 93 values. Regional series were obtained by first computing unstandardized regional series by averaging unstandardized pixel totals within region for each growing season; then, the unstandardized regional series were standardized using the sample mean and standard deviation of the unstandardized series.

2.2. Ocean-Atmosphere Indices and Terrestrial Climate Variables

Ocean climate is represented by three series of monthly values obtained from National Centers for Environmental Prediction. The first of these series, the Pacific Decadal Oscillation (PDO), is a measure of tropical and Northern Hemisphere extratropical Pacific basin sea surface temperature derived from the monthly Historical Sea Surface Temperatures Dataset (Mantua et al. 1997). The Southern Oscillation Index (SOI) series is a measure of standardized sea level pressure difference between Tahiti and Darwin, Australia, obtained from NOAA/NECP Climate Prediction Center, The third series, the North Atlantic Oscillation (NAO), measures the difference in normalized sea level pressures between The Azores and Iceland (Hurrel 1995). Monthly ocean-atmosphere indices were averaged to produce seasonal series for each region. For example, the mean of the PDO observations for the region r growing season for year t is $G_{rt}^{PDO} = n_r^{-1} \sum_{m \in G_r} x_{tm}^{PDO}$ where x_{tm}^{PDO} denotes PDO for year t and month m, G_r denotes the set of growing season months in region r, and n_r is the number of months in the growing season for region r. We also constructed lagged versions of ocean-atmosphere indices to account for delayed effects on SNPP. Herein, we say that a particular predictor of SNPP (say PDO) for year t is represented by a set of components. These components are the current growing and previous winter season mean and up to three previous growing and winter season means. Notationally, W_{rt}^{PDO} is the seasonal mean of PDO for the winter preceding growing season t for region r, and $G_{r,t-1}^{PDO}$ is the growing season mean for the year preceding t, and so on.

Observations on monthly mean daily temperature, monthly total precipitation and monthly mean daily solar radiation were obtained for each of the 135 pixels sampled from the VEMAP terrestrial climate ½° lattice cover of the conterminous United States. Mean growing season temperature, G_{it}^{temp}, for pixel i and year t was computed by averaging over those days belonging to the regional growing season. Mean growing season solar radiation was computed similarly and is denoted by G_{it}^{solar}. Regional means follow the same notation as established for ocean-atmosphere indices; for example, W_{rt-1}^{precip} denotes the total winter precipitation for region r and the growing season preceding year t. Plots of the regional means of these climate variables against year show few notable patterns; in particular, there was little visual evidence of trend over time in any of the series.

3. Regression Analysis

Our model fitting strategy was organized around the objectives of describing the relative importance of the terrestrial climate variables and ocean-atmosphere indices towards predicting SNPP across the study area, measuring the information content of ocean-atmosphere indices and terrestrial climate variables towards predicting SNPP at two scales (pixel and region), and quantifying evidence of SNPP trend during the study period. Models were fit using variables from one of two sets of predictor variables: ocean-atmosphere indices and terrestrial climate variables.

In general, data collected over time or space should not be treated as independent without careful consideration. Observations are dependent when near observations (near in time or space) are more alike than distant observations, and a consequence of ignoring spatial and temporal correlation is a lost opportunity to exploit similarities among observations and thereby improve the accuracy of the parameter estimates. In addition, if correlation is ignored, then significance tests used for selecting predictor variables will tend to yield incorrect observed significance levels (p-values). When observations are positively correlated, the observed significance levels are usually biased downwards (i.e., too small). In this analysis, both serial and spatial correlation is a concern. It is possible to account for either source of correlation provided that the correlation structure is relatively simple, and provided that the relationship between the response variable and the explanatory variables does not vary temporally or spatially. In this situation, it is reasonable to assume that the relationship between SNPP and the predictors at a particular location is nearly time-invariant. On the other hand, we previously argued that there is considerable doubt that the relationship between response and explanatory variables is spatially invariant over the study area. Consequently, our strategy was to model data originating from a single spatial unit (pixel or region) while accounting for serial correlation. Data from different spatial units was not used to fit a particular model because we could not justifiably assume that data from different spatial units follow a common model. As there are 135 pixels and 7 regions, the adopted strategy of fitting separate models for each spatial unit required many models, and an even larger number of significance tests. Methods of controlling experiment-wise Type I error when testing the significance of explanatory variables were not used because these methods are overly cautious when many tests are carried out (Ott and Longnecker, 2001, p. 438); instead, emphasis was placed on describing the fitted models and summarizing model fit in terms of percent variation in NPP explained by a particular set of predictor variables.

3.1. Autoregressive Models

Autoregressive models provide a suitable framework for regression analysis with serially correlated response variables. Harvey (1989), Ramsey and Schafer (2002, p. 436), and Shumway and Stoffer (2001) discuss statistical aspects of regression with autoregressive models, and Epperson (2000) and Manly (2001, p. 212) discuss ecological and environmental applications. To set up the generic model, let z_t denote an observation on a response variable at time $t = 1, \dots, T$, let $x_{1,t}, \dots, x_{p,t}$ denote the values of p predictor variables at time t, and let $\varepsilon_1, \dots, \varepsilon_T$ denote independent and normally distributed residuals with mean 0 and constant variance. Two sets of parameters are used. The autoregressive parameters $\phi_1, \dots, \phi_k$ are used to model the correlation structure among the response variables, and the regression parameters $\beta_1, \dots, \beta_p$ are used to model the relationship between the response and predictor variables. The AR(k) model specifies that z_t depends on the past k observations z_{t-1}, $\dots, z_{t-k}$ according to

$$z_t = \sum_{r=1}^{k} \phi_r z_{t-r} + \sum_{s=1}^{p} \beta_s x_{s,t} + \varepsilon_t. \tag{1}$$

If the AR(k) model is correctly specified, then the term $\sum \phi_r z_{t-r}$ accounts for serial correlation and the residuals are independent. Note that the first k observations $z_1, \ldots, z_k$ cannot be modeled via equation (1) because a complete set of past observations is unavailable for these observations.

Often, trend over time can be accounted for in equation (1) by introducing a polynomial function of time. In this study, there was relatively little visual evidence of trend in the pixel and regional mean NPP series, so we considered only second-order polynomial functions of time. Thus, the two trend components were the linear component $x_{1,t} = t$, and the quadratic component $x_{2,t} = t^2$. Because our interests differ with regard to trend versus ocean-atmosphere indices and terrestrial climate variables, we use the term *explanatory variables* to refer to any of these terms and *predictors* to refer to ocean-atmosphere indices and terrestrial climate variables.

The full ocean climate model for the ith pixel specifies that z_{it}, the standardized total growing season SNPP for pixel i for the tth year after $1903, t = 1, \ldots, T = 90$, is a function of PDO, SOI and NAO given by

$$\begin{aligned} z_{it} = {} & \phi_1 z_{i,t-1} + \cdots + \phi_k z_{i,t-k} \\ & + \beta_1 G_{it}^{PDO} + \beta_2 W_{it}^{PDO} + \beta_3 G_{i,t-1}^{PDO} + \cdots + \beta_8 W_{i,t-3}^{PDO} \\ & + \beta_9 G_{it}^{SOI} + \beta_{10} W_{it}^{SOI} + \cdots + \beta_{16} W_{i,t-3}^{SOI} \\ & + \beta_{17} G_{it}^{NAO} + \beta_{18} W_{it}^{NAO} + \cdots + \beta_{24} W_{i,t-3}^{NAO} \\ & + \beta_{25} t + \beta_{26} t^2 + \varepsilon_{it}. \end{aligned} \tag{2}$$

An intercept is not included in the model because the standardized NPP series have mean 0. The full AR(k) model as a function of the terrestrial climate variables is

$$\begin{aligned} z_{it} = {} & \phi z_{i,t-1} + \cdots + \phi_k z_{i,t-k} \\ & + \beta_1 G_{it}^{temp} + \beta_2 W_{it}^{temp} + \beta_3 G_{i,t-1}^{temp} + \cdots + \beta_8 W_{i,t-3}^{temp} \\ & + \beta_9 G_{it}^{precip} + \beta_{10} W_{it}^{precip} + \cdots + \beta_{16} W_{i,t-3}^{precip} \\ & + \beta_{17} G_{it}^{solar} + \beta_{18} W_{it}^{solar} + \cdots + \beta_{24} W_{i,t-3}^{solar} \\ & + \beta_{25} t + \beta_{26} t^2 + \varepsilon_{it}. \end{aligned} \tag{3}$$

The AR(k) regional models of NPP differ only in that the response and predictor variables are regional means computed by averaging pixel values.

3.2. Model Fitting

We used a Cochran-Orcutt scheme for fitting the AR(k) regression model (Venables and Ripley 1999, p. 429). Our strategy was to first search for the smallest value of k necessary to remove autocorrelation, and then select explanatory variables using conventional multiple regression methods. The search for k began by computing the maximum likelihood estimates of the parameters for the full AR(1) model (equation [1] with $k = 1$). Then, the pixel (or regional) NPP series and the explanatory variables were pre-whitened using the maximum likelihood estimate $\widehat{\phi}_1$ (Box and Jenkins 1976, p. 379). Pre-whitening eliminates serial correlation by subtracting the portion of z_t and $x_{s,t}$, $s = 1, \ldots, p$, attributable to the past observations from both sides of formula (1) by computing $z_t^* = z_t - \widehat{\phi}_1 z_{t-1}$, and $x_{s,t}^* = x_{s,t} - \widehat{\phi}_1 x_{s,t-1}$. Conventional multiple regression methods, including variable selection techniques, can be used without modification provided that pre-whitening has succeeded and the residuals are normally distributed. Checking the residuals for serial correlation and non-normality is a necessary step in the application of the AR(k) model.

Model residuals were tested for autocorrelation using the Box-Ljung Q-statistic (also called the Portmanteau statistic) (Harvey 1993, p. 79). If the Q-statistic did not present evidence against the null hypothesis of no autocorrelation (that is, the p-value associated with Q was greater than 0.05), then the AR(1) model was adopted. If there was evidence of autocorrelation, then the AR(2) model was tentatively adopted and maximum likelihood estimates $\widehat{\phi}_1$ and $\widehat{\phi}_2$ were computed. The data were pre-whitened by computing $z_t^* = z_t - \widehat{\phi}_1 z_{t-1} - \widehat{\phi}_2 z_{t-2}$ and $x_{s,t}^* = x_{s,t} - \widehat{\phi}_1 x_{s,t-1} - \widehat{\phi}_2 x_{s,t-2}$. Regression parameter estimates were recomputed and the new residuals assessed for autocorrelation If these residuals were autocorrelated, then higher orders of k were to be investigated in the same fashion. It was not necessary to do so, though. The second stage of the model fitting (described in detail below) sequentially removed non-significant variables from the model.

An alternative approach to the Cochran-Orcutt scheme simultaneously computes maximum likelihood estimates of both the regression and autoregressive parameters. Consequently, the maximum likelihood approach is slightly simpler to execute than the Cochran-Orcutt scheme. However, the maximum likelihood estimators and associated likelihood ratio tests may be subject to substantial bias unless the number of observations is much larger than the number of model parameters (e.g., $T > 30p$) . Because the series used in this analysis were relatively short ($T = 93$), we used the Cochran-Orcutt scheme.

3.3. Variable Selection and Assessment of Model Fit

Our variable selection procedure was applied to each of the three predictor variables representing ocean-atmosphere indices, or terrestrial climate, one at a time. Initially, the full model [equation (2) or (3)] was adopted. In the case of the ocean-atmosphere indices, we first assessed the significance of PDO by successively removing the highest-order lag components until the remaining lower-order terms were jointly significant. That is, first $W_{i,t-3}^{PDO}$, then $G_{i,t-2}^{PDO}$, and so on, were removed from the model until the remaining components were found to be significant at the $\alpha = 0.05$ level. The significance test, often called the extra-sums-of-squares F-statistic (Ramsey and Schafer, 2002, p. 281), compares the error sums-of-squares between the model containing a particular set of explanatory variables and the model containing none of the terms while accounting for differences in numbers of explanatory variables between the two models. After analyzing PDO, the significance of SOI was investigated using the full model [equation (2)] except that non-significant PDO components were omitted. Lastly, NAO was treated in the same way, by adopting the final SOI model as the initial model. Generally, the order in which variables are tested can affect the outcome of variable selection; in this analysis, though, the effects of order were negligible.

The same strategy was applied to the trend components. First, we tested for significance of both the linear and quadratic components; if these were not jointly significant, then we tested for the significance of the linear component alone. However, for the purpose of describing directional trend (upwards or downwards) in NPP, we also fit models containing the significant predictor variables and the linear trend component (regardless of significance) while excluding the quadratic term (regardless of significance) and using the unstandardized, but pre-whitened SNPP observations. An estimate of percent annual change after accounting for significant predictors was computed by dividing the linear trend coefficient by the SNPP series mean.

A possible alternative to jointly testing lower-order lag components is to test individual lag components separately using a t-test. This strategy is counter to our presumption that if a predictor variable is related to SNPP, then it affects SNPP through a complete set of lag components up to some order. Thus, singly testing each component may lead to implausible models if a higher order term is judged to be significant, but one or more lower-order terms are not.

Two coefficients of determination were used to measure the percent variation in the response variable explained by the fitted model. The coefficient of determination R_1^2 measures the percent variation in SNPP (not pre-whitened) explained by the final model; hence, these models may contain one or both of the trend components. This statistic was computed as $R_1^2 = (\mathrm{SST} - \mathrm{SSE})/\mathrm{SST}$, where SST is the total sums-of-squares about the series mean $\bar{z} = 0$, $\mathrm{SSE} = \sum(z_t - \widehat{z}_t)^2$ is the total error sums-squares, and $\widehat{z}_t$ is the tth fitted value. If one or more trend components were retained in the final model, then R_1^2 would reflect the contributions of both the predictor variables and trend components towards explaining variation in SNPP. Because our interests lie primarily with the predictor variables, a second coefficient of determination R_2^2 was computed after removing the trend components, regardless of significance, and refitting the model. The coefficient R_2^2 measures the percent variation explained by only the ocean-atmosphere indices, or terrestrial climate predictors. Note that if

$R_1^2 > R_2^2$, then one or both of the trend components were retained in the model used to compute R_1^2.

4. Results

The analysis of the regional series is discussed first, followed by the analysis of the individual pixel series. Ocean climate explained 35.9, 51.6, and 31.8% of the variation in SNPP for the Interior West, Pacific Coast and Southwest, and Mid-Atlantic regions, respectively, as measured by R_2^2 (Table 2). For the remaining four series, none of the ocean-atmosphere index series were found to be significant, and so no variation in SNPP was explained by ocean-atmosphere indices. Table 2 also shows that terrestrial climate variables were very good predictors of SNPP given that R_2^2 varied from 70.9% for the Pacific Coast and Southwest to 89.3% for the Southcentral region. Based on the Q-statistics (Table 2), the AR(1) model was judged to be adequate with respect to the assumption of independently distributed residuals. Normal probability plots indicated that the residuals were normal in distribution.

Table 3 shows the final model coefficients for the regional ocean climate models. PDO was a significant predictor of SNPP for the Interior West and Pacific Coast and Southwest regions, NAO was significant for the Mid-Atlantic region, and SOI was significant for all three regions. There is clear evidence of delayed effects of ocean-atmosphere indices on SNPP given that the complete set of lag components was retained in the fitted models for four of the six predictor variables. It should be noted that the magnitude of the higher order SOI lag coefficients are substantially smaller than the current growing season coefficient. This result suggests that the SOI effect is comparatively short in duration. Table 3 shows that lag coefficients vary sign when comparing lag coefficients for a particular ocean-atmosphere index, and this suggests that the relationship between ocean climate and SNPP is highly complex. Finally, similarities between Pacific Coast and Southwest and Interior West ocean-atmosphere index coefficients imply similar ocean climate effects across the western third of the United States.

Linear trend coefficients and estimated annual percent change in SNPP are shown in Table 4. After accounting for significant predictors of SNPP, there is statistical evidence of linearly increasing trend between 1901 and 1993 in all regional series except the Mid-Atlantic and Southeast. Both the ocean-atmosphere index models and the terrestrial climate models yielded the greatest annual percent change estimates for the Great Plains and Southcentral regions. Estimates of 0.119% and 0.150% were obtained for the Great Plains, and 0.136% and 0.171% for the Southcentral regions (Table 4) from the terrestrial climate and ocean climate models, respectively. Differences between the two types of models are attributable to different predictor variables comprising the models.

Terrestrial climate explained a substantial portion of the variation in SNPP at the pixel scale. Specifically, Table 5 shows that the regional averages of the terrestrial climate model coefficients of determination (R_2^2) varied between 64.5 and 85.6%. In contrast, ocean-atmosphere indices were substantially less useful for modeling SNPP. Table 5 also shows that ocean-atmosphere indices were found to be significant (i.e., $R_2^2 > 0$) for only 44 of the 135 series, and that when significant, the regional means of R_2^2 statistics varied between 8.9 and 27.0%. The regional means of R_2^2 are not comparable with the corresponding means for R_1^2 because different sets of pixel values were averaged in computing the two statistics. Spatial variation in R_2^2 over the conterminous United States is substantial; in particular, Figure 4 shows that the majority of pixels for which $R_2^2 > 0$ are located in the southwestern portion of the United States. In contrast, there was very little spatial variation in the fit of the terrestrial climate models, and a corresponding figure for these models is not shown.

The autoregressive modeling strategy appears to have been successful at accounting for serial correlation given that the Q-statistic was significant at the 0.05 level for only 6.7% of the 135 terrestrial climate model residual series and 11.6% of the ocean climate model residual series. The assumed Type I error rate (or significance level) of 5% implies that approximately 5% of all residual series that are truly free of serial correlation will be (incorrectly) found to be significant; hence, the observed fraction of significant Q-statistics is consistent with there being little serial correlation in the residual series.

Figures 5, 6 and 7 summarize the estimated coefficients for the terrestrial climate variables. For each region and lag component, we averaged those coefficients that were found to be significant and plotted the averages against lag. Figure 5 shows that current growing season temperature to be negatively associated with SNPP for all regions, and that the largest effects (in terms of magnitude) of SNPP were observed for the Southeast (-0.237) and Southcentral (-0.235) regions. In contrast, the coefficient means associated with the previous year growing season temperature (lag 2) were positive for all regions. In general, winter season temperatures are relatively less important than growing season temperatures, and inconsistent in sign among regions. Figure 6 shows estimated effects of current growing season and previous winter season precipitation were positive for all regions, though previous winter season precipitation is far more important, based on the magnitude of the coefficients. The effect of previous winter precipitation on SNPP, as measured by the mean of the previous winter coefficients, was greatest in the arid regions of the United States, namely, the Great Plains, Interior West, and Pacific Coast and Southwest regions. The coefficient means associated with current growing season solar radiation are negative for all regions, and the greatest impact on SNPP are associated with the Northeast, Southcentral and Pacific Coast and Southwest regions (Figure 7); in contrast, previous growing season solar radiation was found to be positively associated with SNPP for all regions. As expected, winter season solar radiation has little effect on SNPP, regardless of lag. Because the terrestrial climate and SNPP series were standardized, the coefficients associated with different climate variables can be compared to assess the relative importance of these variables. Comparing Figures 5, 6, and 7 shows that the effect of previous season precipitation is an order of magnitude greater than all other effects, including temperature and solar radiation. For example, the average of the regional means of the previous season precipitation coefficients was 4.2, whereas the current growing season mean of the temperature and solar radiation coefficients were -0.10 and -0.19, respectively. Finally, a corresponding set of three coefficient plots was constructed using coefficients obtained from the regional terrestrial climate models. These plots are similar in appearance to Figures 4, 5, and 6, though the general patterns are not as striking.

In contrast to the analysis of the regional SNPP series, there was substantially less evidence of trend at the pixel scale after accounting for significant predictors. Because differences between R_1^2 and R_2^2 are attributable to the trend components, a comparison of the regional means of these statistics derived from the terrestrial climate models (Table 5) shows little evidence of trend in SNPP at the pixel scale after accounting for terrestrial climate variables. Comparisons of the regional means of R_1^2 and R_2^2 computed from the ocean climate models cannot be drawn from Table 4 because different pixels were used in computing the two regional means. However, trend components were not consistently important when fitting the ocean climate models, given that trend was found to be significant for 47.4% of the 135 pixels. We attribute this result, at least in part, to weak power stemming from unexplained variation in the pixel series. To illustrate, when predictor variables were ignored and percent annual change was computed for all pixels, then the median percent annual change was 0.135%, and the regional medians of percent annual change varied between 0.047% for the Mid-Atlantic to 0.168% for the Southcentral. These values are consistent with the estimates of annual change obtained from the regional series (Table 4).

5. Comparison to Empirical NPP Modeling Methods

Among the promising recent approaches to empirical modelling of NPP are modified production efficiency models (PEMs) driven by remotely sensed data (Goetz and others 1999). The PEM approach exploits remotely sensed data for modeling, thereby reducing the requirements for initial state variable specification and but also predictive utility. Zamolodchikov and Karelin (2001) developed a regression-based model of carbon flux for the Russian tundra that used climate and GIS-based biophysical variables to predict NPP. The advantage of their approach is that direct measurements of carbon flux were used to develop the models, and that regression methods were used. However, extension is problematic to other environments because of the difficulty of measuring NPP. Jiang and others (1999) developed an NPP modeling approach that calibrated simple process models via regression at each point in a lattice. Alexandrov and others (2002) propose an "inversion" method of constructing empirical models by calibrating NPP process models (such as the Miami model [Lieth 1975]) using biome-averaged NPP estimates. Their method produces simple global models, though model accuracy is a concern because relatively few, and fundamentally different data are used in model fitting .

The regression strategy proposed in this article yields empirical models of SNPP that may be used for predicting NPP. Our method departs from other empirical modeling approaches by using climate data collected over many years to fit the predictive models. By using SNPP data generated over many years, a substantial amount of process variation is realized and (partially) accounted for by the fitted models. A weakness of our approach is that process model errors appearing in the SNPP data will be carried over to the regression models. An advantage common to Jiang and others (1999) is that our method generates spatially explicit models across a lattice and allows for a high degree of spatial resolution. Finally, our approach uses well-understood statistical methods though out, and exploits seasonality and delayed climate effects to improve model accuracy.

6. Discussion and Conclusion

This article has presented a strategy for analyzing spatio-temporal biophysical data. A distinctive and novel feature of this strategy is that separate autoregressive regression models are fit at each location among a set of locations regularly distributed across the study area. By doing so, spatially induced differences in the relationship between the response variable and the predictor variables can be investigated. This strategy significantly advances our abilities to analyze the behavior of complex ecological models.

The case study analysis showed, unsurprisingly, that the three terrestrial climate variables provide good linear models of SNPP. The ocean-atmosphere index series PDO and SOI yield some predictive information at the regional scales for the Interior West and Pacific Coast and Southwest, while NAO provides weak information for the Mid-Atlantic region. At the pixel scale, ocean-atmosphere indices provide roughly similar predictive information as at the regional scale, and over a similar geographic area. For remainder of the conterminous United States, the methods and data used in this study did not find consistent and useful associations between SNPP and ocean-atmosphere indices.

At the regional scale, there is statistical evidence of trend in SNPP for all regions except the Mid-Atlantic and Southeast. Moreover, estimated percent annual change (Table 4) was found to be fairly consistent among regions. Upward trend in SNPP is consistent with trend in atmospheric CO_2 during the 20th century; however, changes in land use during the study period complicate comparisons of SNPP and atmospheric CO_2. At the pixel scale, percent annual change in the SNPP series tended to be relatively large (ignoring predictor variables). The linear trend coefficients, however, tended not to be significant when terrestrial climate predictors were accounted for. This result implies that some of this change is at least partially accounted for by terrestrial climate.

Regression modeling of SNPP as a function of ocean and terrestrial climate is greatly improved by lagging the predictor variables. Two explanations for the usefulness of lagged variables are: 1) climate variables have long-term effects on NPP, and 2) some climate patterns are multi-year in duration (Latif and Barnett 1994; Hurrell 1995). Winter precipitation and growing season solar radiation coefficients consistently reversed sign when comparing current and previous years effects on NPP (Figures 6 and 7). We suggest that this result is attributable to using linear functions of these variables to approximate nonlinear relationships between soil moisture availability and SNPP. An example of a linear approximation that produces coefficients of different sign is given by a response variable y that responds monotonically to x according to the model $e^y = x$. Fitting a second-order polynomial function of x to y will yield a positive coefficient for x and a negative coefficient for x^2.

5. Acknowledgments

We thank NASA Earth Science Enterprise for their support of this research.

6. References

Anyamba, A. 1994. Retrieval of ENSO signal from vegetation index data. Earth Observer: 6, 24-6.

Anyamba A., Eastman J.R. 1996. Interannual variation of NDVI over Africa and its relation to El Nino/Southern oscillation. Int. J. Rem. Sens., 17: 2533-2548.

Alexandrov, G. A., Oikawa, T., Yamagata, Y. The scheme for globalization of a process-based model explaining gradations in terrestrial NPP and its application. Ecological Modelling, 148: 293-306.

Box, G.E. and Jenkins, G.M. 1976. Time Series Analysis Forecasting and Control. Prentice-Hall, New York. 575 pp.

Cane, M.A., Eshel, G., Buckland, R.W. 1994. Forecasting Zimbabwean maize yield using eastern equatorial Pacific sea temperature. Nature, 370: 204-205.

Dai, A., Fung. I.Y., Del Genio, A.D. 1997. Surface observed global land precipitation variations during 1900-88. Journal of Climate, 10: 2943-2962.

Dima, M., Rimbu, N. Stefan, S., Dima, I. 2001. Quasi-decadal variability in the Atlantic basin involving tropics-midaltitudes and ocean-atmospheric interactions. J. Climate, 14: 823-832.

Epperson, B.K. 2000. Spatial and space-time correlations in ecological models. Ecological Modeling, 132: 63-76.

Goetz, S.J., Prince, S.D., Goward, S.N., Thawley, M.M., Small, J. 1999. Satellite remote sensing of primary production: an improved production efficiency modeling approach. Ecological Modeling, 122: 239-255.

Harvey, A.C. 1993. Time Series Models, 2nd Ed. MIT Press, Cambridge, 308 pp.

Harvey, A.C. 1989. Forecasting, Structural Time Series Models, and the Kalman Filter. Cambridge University Press, Cambridge, 554 pp.

Hunt Jr., E.R., Piper, S.C., Nemani, R., Keeling, C.D., Otto, R.D., Running, S.W. 1996. Global net carbon exchange and intra-annual atmospheric CO_2 concentrations predicted by an ecosystem process model and three-dimensional atmospheric transport model. Global Biochemical Cycles, 10:431-456.

Hurrell, J.W. 1995. Decadal trends in North Atlantic oscillation: regional temperatures and precipitation. Science, 269: 676-679.

Hurrell, J.W. 1996. Influence of variations in extropical wintertime teleconnections on Northern Hemisphere temperature. Geophysical Res. Lett., 23: 665-668.

Ito, A., Oikawa, T. 2002. A simulation model of the carbon cycle in land ecosystems (Sim-CYCLE): a description based on dry-matter production theory and plot-scale validation. Ecological Modelling, 151, 143-176.

Jiang, H. Appsa, M.A., Zhanga, Y. Penga, C. Woodard, P.M. 1999. Modelling the spatial pattern of net primary productivity in Chinese forests. Ecological Modelling, 122: 275-288.

Jones, R.H., Zang, Y. 1997. Models for continuous stationary space-time processes. In: Modelling Longitudinal and Spatially Correlated Data, Gregoire, T.G., Brillinger, D.R., Diggle, P.J., Russek-Cohen, E.,Warren, W.G., Wolfinger, R.D. (Ed.), Modelling Longitudinal and Spatially Correlated Data, Springer, New York, 289-298.

Johnson, D.E. 1998. Applied Multivariate Methods for Data Analysis. Duxbury Press, New York. 567 pp.

Keane, R.E., Parsons, R.A., Hessburg, P.F. 2002. Estimating historical range and variation of landscape patch dynamics: limitations of the simulation approach. Ecological Modelling, 151, 29-49.

Kimball, J.S., Thornton, P.E, White, M.E., Running, S.W. 1997. Simulating forest productivity and surface-atmosphere carbon exchange in the BOREAS study region. Tree Physiology, 17:589-599.

Kittel, T.G.F., Rosenbloom, N.A., Painter, T.H., Schimel, D.S. 1995. The VEMAP integrated database for modeling United States ecosystem/vegetation sensitivity to climate change. Journal of Biogeography, 22:857-862.

Knapp, A.K. and Smith, M.D. 2001. Variation among biomes on temporal dynamics of aboveground primary production. Science, 291: 481-484.

Latif, M., Barnett, T.P. 1994. Causes of decadal climate variability over the North Pacific and North America. Science, 266:634-637.

Li, Z., and Kafatos. M. 2000. Interannual variability of vegetation in the United States and its relation to El Nino/Southern oscillation. Rem. Sens. Environ., 71: 239-247.Lieth, 1975.

Lieth, H. 1975. Modelling the primary productivity of the world. In: H. Lieth and R.H. Whittaker, Editors, Primary Productivity of the Biosphere, Springer-Verlag, New York, pp. 237–263.

Manly, B.F.J,. 2001. Statistics for Environmental Science and Management. Chapman Hall, Boca Raton, FL, 362 pp.

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M., Francis, R.C. 1997. A Pacific interdecadal oscillation with impacts on salmon production. Bull. Am. Meteorol. Soc., 78: 1069-1079.

Melillo, J.M., McGuire, A.D., Kicklighter, D.W., Moore III, B., Vorosmarty, C.J., Schloss, A.L. 1993. Global climate change and terrestrial net primary production. Nature, 363: 234-240.

Mennis, J. 2001. Exploring relationships between ENSO and vegetation vigour in the southeast USA using AVHRR data. Int. J. Rem. Sens., 22:3077-3092.

Myneni, R.B., Los, S.O., Tucker, C.J. 1996. Satellite-based identification of linked vegetation index and sea surface temperature anomaly areas from 1982-1900 for Africa, Australia and South America. Geophys. Res. Lett., 23: 727-732.

Myneni, R.B., Keeling, C.D., Tucker, C.J., Asrar, G., Nemani, R.R. 1997. Increased plant growth in the northern high latitudes from 1981 to 1991. Nature, 386: 698-702.

Myneni R.B., Tucker, G.A., Keeling, C.D. 1998. Interannual variations in satellite-sensed vegetation index data from 1981 to 1991. J. Geophysical Res., 103: 6145-6160.

Nemani, R., White, M., Thornton, P., Nishida, K., Reddy, S., Jenkins, J., Running, S. 2002. Recent trends on hydrologic balance have enhanced the terrestrial carbon sink in the United States. Geophysical Res. Lett., 29: 106-109.

Nychka, D., Piegorsch, W.W. and Cox, L.H. 1998. Case Studies in Environmental Statistics. Springer, New York, 196 pp.

Ott, L.R. and Longnecker, M. 2001. An Introduction to Statistical Methods and Data Analysis, 5th Ed. Duxbury Press, New York, 1152 pp.

Potter, C.S., Klooster, S.A.. Brooks, V. 1999. Interannual variability in terrestrial net primary production: exploration of trends and controls on regional to global scales. Ecosystems, 2: 36-48.

Potter, C.S. and Klooster, S.A. 1997. Global model estimates of carbon and nitrogen storage in litter and soil pools: response to change in vegetation quality and biomass allocation. Tellus. 48 1-17.

Ramsey, F.L. and Schafer, D.W. 2002. The Statistical Sleuth, 2nd Ed. Duxbury Press, Belmont CA, 742 pp..

Schimel, D., Melillo, J., Tian, H., McGuire, D.A., Kicklighter, D., Kittel, T., Rosenbloom, N., Running, S., Thornton, P., Ojima, D., Parton, W., Kelly, R., Sykes, M., Neilson, R., Rizzo, B., 2000. Contribution of increasing CO_2 and climate to carbon storage by ecosystems in the United States. Science, 287: 2004-2006.

Shumway, R.H. and Stoffer, D.S. 2001. Time Series Analysis and Its Applications. Springer, New York, 549 pp.

Sutton, R.T., Allen, M.R. 1997. Decadal predictability of North Atlantic sea surface temperature and climate. Nature, 388: 563-567.

Thornton, P.E., Law, B., Gholz, H., Clark, K., Fagle, E., Ellsworth, D. 2002. Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. Agricultural and Forest Meteorology. Forthcoming.

Venables, W.N. and Ripley, B.D. 1999. Modern Applied Statistics with S-PLUS. Springer, New York, 501 pp.

White, M.A., Thornton, P.E., Running, S.W., Nemani, R.R. 2000. Parameterization and sensitivity analysis of the BIOME_BGC terrestrial ecosystem model: Net primary production controls. Earth Interactions, 4:1-84.

Zamolodchikov, D.G., Karelin, D.V. 2001. An empirical model of carbon fluxes in Russian tundra. Global Change Biology, 7: 1365-2486.

Tables

Table 1. Regional growing seasons.

Region	Growing Season
Interior West	March - November
Pacific Coast and Southwest	March - October
Great Plains	April - October
Mid-Atlantic	February - October
Northeast	March - October
Southcentral	April - October
Southeast	April - December

Table 2. Summary statistics for the regional NPP models. Tabled values are percent variation in regional series explained by the fitted models (R_1^2 and R_2^2)[1], Box-Ljung Q-statistic, and an approximate p-value (p)[2] for Q.

Region	Ocean Climate Models				Terrestrial Climate Models			
	R_1^2	R_2^2	Q	p	R_1^2	R_2^2	Q	p
Interior West	35.9	33.4	6.7	.74	79.6	79.6	11.7	.30
Pacific Coast and Southwest	51.6	51.6	6.7	.75	70.9	70.9	7.6	.67
Great Plains	19.1	0	8.4	.59	87.6	81.3	2.6	.99
Mid-Atlantic	31.8	31.8	9.2	.51	81.2	81.2	6.6	.75
Southcentral	9.0	0	9.1	.52	91.1	89.3	16.6	.08
Northeast	9.8	0	7.6	.67	75.6	74.1	7.4	.68
Southeast	0	0	13.1	.17	82.4	82.4	14.8	.14

[1] R_1^2 is percent variation explained by all significant predictive variables whereas R_2^2 is the percent variation explained by all significant predictive variables excluding trend components.
[2] The approximate p-value is $p = P(\chi_{10}^2 \geq Q)$, where χ_{10}^2 denotes a chi-square random variable with $k = 10$ degrees of freedom.

Table 3. Estimates of the ocean-atmosphere index coefficients for the Interior West, Pacific Coast and Southwest, and Mid-Atlantic regions. Estimates were computed for each region after eliminating non-significant components. The ocean-atmosphere index identified in the first row and the lag components are identified in the first column.

	PDO		SOI			NAO
Season	Int. West	Pac. SW	Int. West	Pac. SW	Mid-Atlantic	Mid-Atlantic
G_j	0.125	0.232	−0.201	−0.162	0.191	0.295
W_j	·	−0.306	−0.081	−0.032	·	−0.218
G_{j-1}	·	0.384	−0.049	−0.075	·	−0.192
W_{j-1}	·	−0.037	−0.004	0.017	·	0.070
G_{j-2}	·	−0.010	−0.170	−0.017	·	0.224
W_{j-2}	·	−0.070	0.021	−0.001	·	0.003
G_{j-3}	·	−0.025	−0.157	0.066	·	−0.349
W_{j-3}	·	−0.222	−0.027	−0.015	·	0.276

Table 4. Coefficient estimates of linear trend derived from the regional NPP models. These coefficients estimate change in standardized NPP per year after accounting for significant predictors of NPP (ocean-atmosphere indices or terrestrial climate variables). The estimated annual percent change in the unstandardized NPP series is also shown. Those estimates that were not found to be significant at the $\alpha = 0.05$ level are identified by an asterisk.

	Ocean Climate Models		Terrestrial Climate Models	
Region	Coefficient	Annual Change (%)	Coefficient	Annual Change (%)
Interior West	0.065	0.102	0.032	0.051
Pacific Coast and Southwest	0.071	0.114	0.023	0.037
Great Plains	0.141	0.150	0.115	0.119
Mid-Atlantic	0.026*	0.032	0.048*	0.059
Southcentral	0.107	0.171	0.085	0.136
Northeast	0.119	0.077	0.067	0.043
Southeast	0.050*	0.073	0.047*	0.075

Table 5. Number of pixels per region (n_i) and regional means of the coefficients of determination (R_1^2 and R_2^2) for the ocean and terrestrial climate models. Only pixels for which at least one of the predictor variables were significant were used in computing the regional means. The number of pixels used to compute the regional averages are shown in parentheses for the ocean climate models. For the terrestrial climate models, the numbers of pixels used to compute the regional means were the same as number of pixels per region, and are not shown.

		Ocean Climate Models		Terrestrial Civatte Models	
Region	n_i	R_1^2	R_2^2	R_1^2	R_2^2
Interior West	23	17.0 (14)	26.8(7)	69.5	64.5
Pacific Coast and Southwest	28	24.3 (14)	25.6(13)	64.0	65.0
Great Plains	33	11.9 (25)	19.2 (9)	70.3	70.9
Mid-Atlantic	7	8.9 (4)	8.9 (4)	74.6	74.9
Southcentral	22	12.1 (15)	27.0 (4)	80.4	80.7
Northeast	12	14.7 (11)	21.0 (4)	62.2	60.3
Southeast	10	13.6 (3)	13.6 (3)	85.6	85.6

Figure 1. Regions and pixel locations. Pixels are numbered according to region, and are not drawn to scale. Interior West=1, Pacific Coast and Southwest=2, Great Plains=3, Mid-Atlantic=4, Southcentral=5, Northeast=6, Southeast=7.

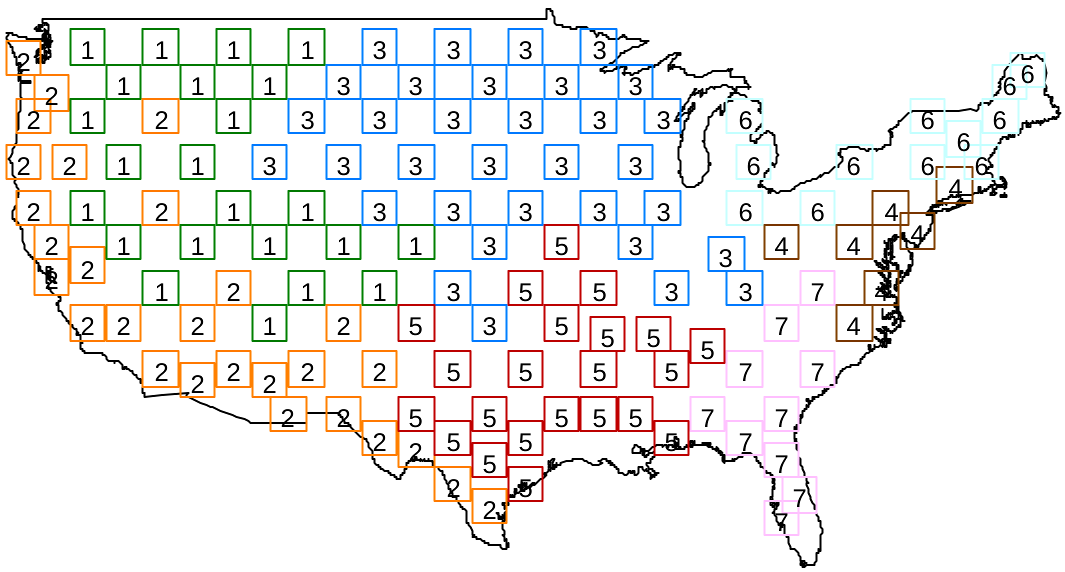


Figure 2. A star plot of showing monthly means of absolute daily NPP. These values, showing total amount of transfer of carbon (positive and negative), are plotted in lieu of NPP because star plots cannot accommodate both positive and negative values. January means are positioned at 3 o'clock, and the calendar months advance in counterclockwise direction. The length of the star radii for the is proportional to monthly mean carbon exchange.

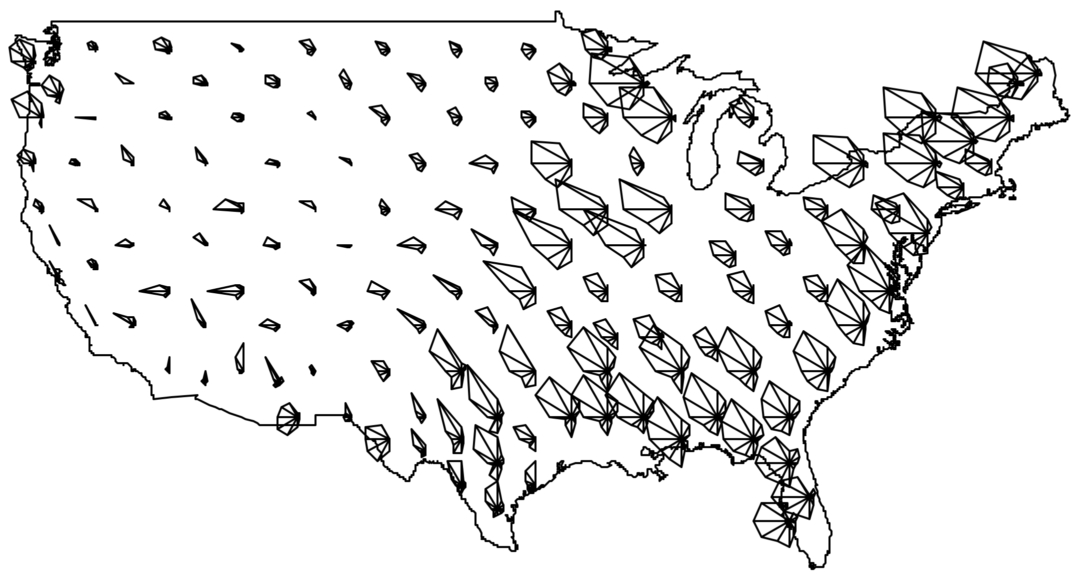


Figure 3. Calendar monthly means of NPP plotted by region.

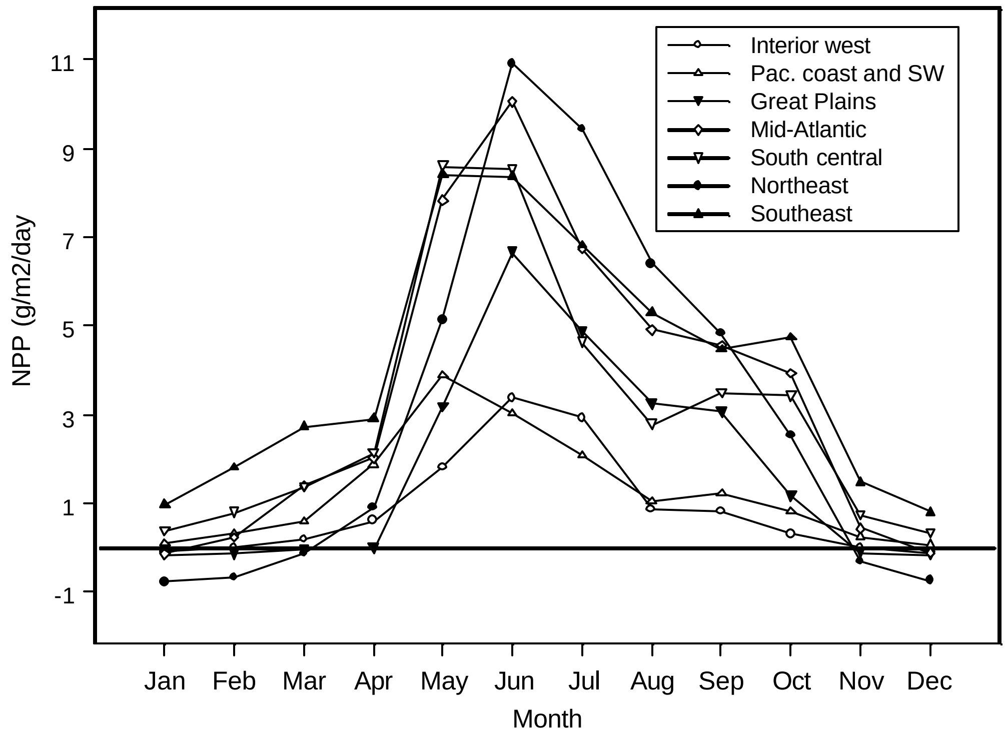
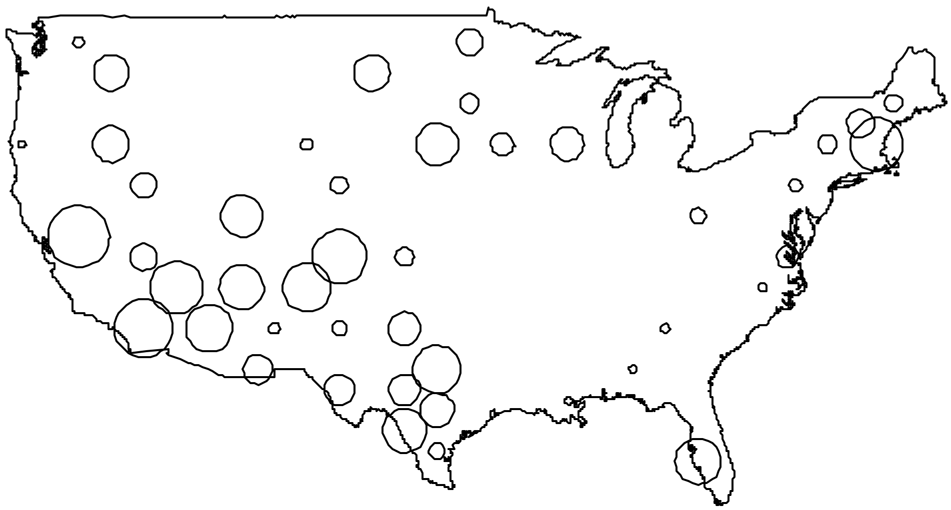


Figure 4. Percent variation (R_2^2) in NPP explained by ocean climate models for each pixel. The circle centers are located at pixel coordinates and circle radius is proportional to R_2^2. Of the 135 pixels, 44 values of R_2^2 were non-zero, and the maximum value was 27.0%.



.

Figure 5. Regional averages of the temperature coefficients plotted against lag. Averages were computed using only those coefficients that were found to be significant at the 0.05 level in the pixel-scale NPP models. Lag 0 coefficients are associated with current growing season, lag 1 coefficients with the previous winter season, and so on.

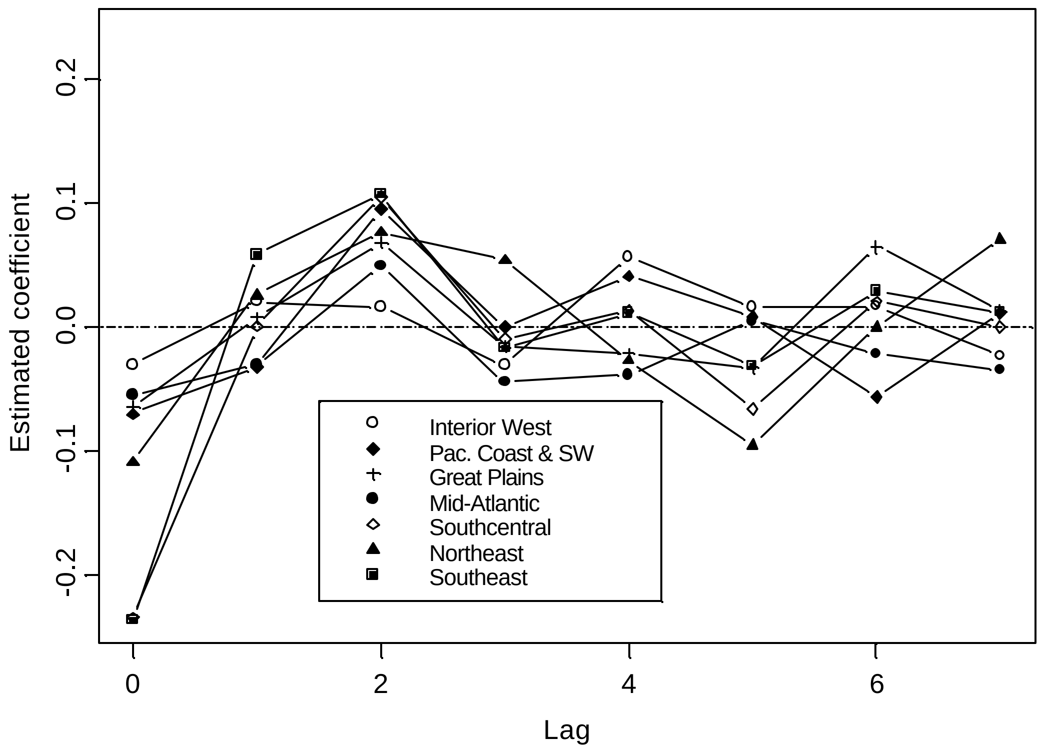


Figure 6. Regional averages of the precipitation coefficients plotted against lag. Averages were computed using only those coefficients that were found to be significant at the 0.05 level in the pixel-scale NPP models. Lag 0 coefficients are associated with current growing season, lag 1 coefficients with the previous winter season, and so on.

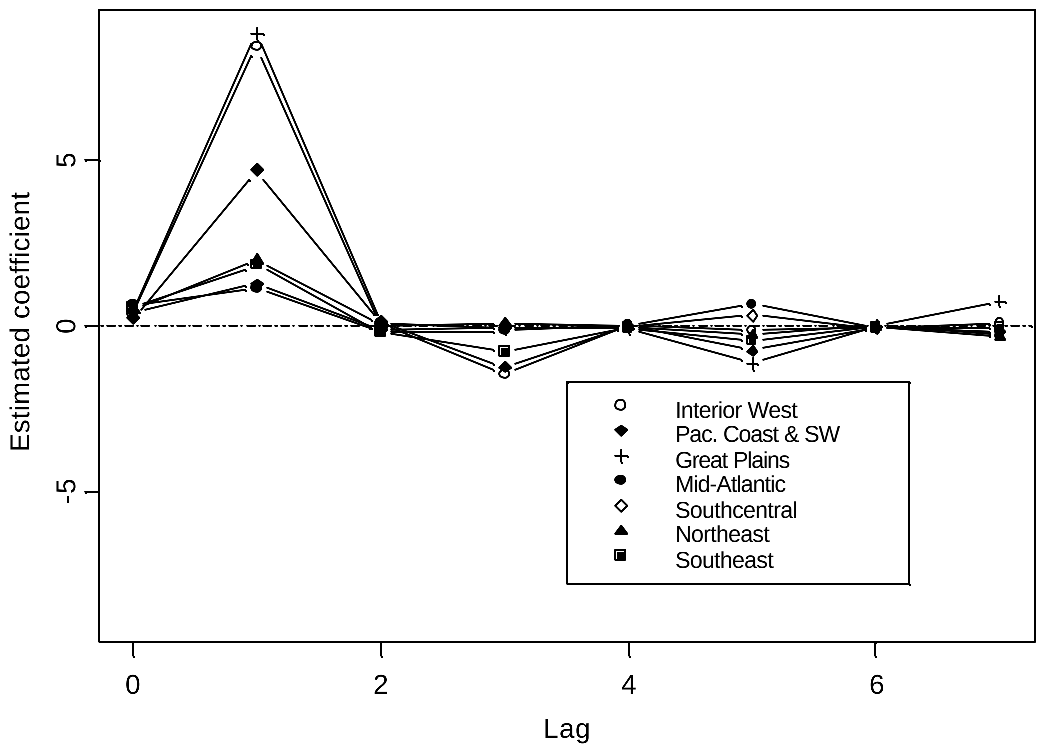


Figure 7. Regional averages of the solar radiation coefficients plotted against lag. Averages were computed using only those coefficients that were found to be significant at the 0.05 level in the pixel-scale NPP models. Lag 0 coefficients are associated with current growing season, lag 1 coefficients with the previous winter season, and so on.