Environ Monit Assess
DOI 10.1007/s10661-006-9322-6

ORIGINAL ARTICLE

Evaluating Kriging as a Tool to Improve Moderate Resolution Maps of Forest Biomass

Elizabeth A. Freeman · Gretchen G. Moisen

Received: 28 September 2005 / Accepted: 27 February 2006

Abstract The USDA Forest Service, Forest Inventory and Analysis program (FIA) recently produced a nationwide map of forest biomass by modeling biomass collected on forest inventory plots as nonparametric functions of moderate resolution satellite data and other environmental variables using Cubist software. Efforts are underway to develop methods to enhance this initial map. We explored the possibility of modeling spatial structure to make such improvements. Spatial structure in the field biomass data as well as in residuals from the map was investigated across 18 ecological zones in the Interior Western U.S. Exploratory tools included directional graphs of summary statistics, three dimensional maps, Moran's I correlograms, and variograms. Where spatial pattern was present, field and residual biomass were kriged, and predictions made for an independent test set were evaluated for improvement over predictions in the initial biomass map. While kriging has some potential benefit when analyzing the field data and exploring spatial structure, kriging residuals resulted in little or no improvement in the initial biomass map developed using Cubist software. Stationarity assumptions, variogram behavior, and appropriate model fitting strategies are discussed.

Keywords Biomass · Correlograms · Forest Inventory and Analysis · Kriging · Moran's I · Variograms

E. A. Freeman (✉) · G. G. Moisen
U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Forestry Sciences Laboratory, 507 25th Street, Ogden, UT 84401, USA
email: eafreeman@fs.fed.us

1 Introduction

Spatial depictions of forest biomass are important for identifying areas with large fire risk and managing forest carbon offsets. The USDA Forest Service, Forest Inventory and Analysis program (FIA) produced a nationwide map of forest biomass (Blackard et al., In review). The map was created by modeling biomass collected on forest inventory plots as nonparametric functions of moderate resolution satellite data using tree-based models implemented in Cubist software (www.rulequest.com). Efforts are underway to develop methods to enhance the initial map. Here, we explored the possibility of modeling spatial structure to make such improvements in the Interior Western U.S. The Cubist models used to create the nationwide forest biomass maps (hereafter referred to as the Cubist biomass models) did not explicitly include a spatial component. However, several of the predictor variables included in the Cubist biomass models, such as elevation and aspect, are spatially correlated. Our objective was to deter-

mine if these variables adequately account for a spatial pattern, or if the addition of an explicit spatial model, such as kriging, would improve the predictions, resulting in a more accurate map.

Geostatistical techniques have become popular tools for modeling spatial structure of ecological data as they offer many advantages. For example, unlike simple smoothing algorithms, kriging uses the sample variogram to define the nature of the spatial interaction and determine the appropriate size of the interpolation window (Burrough, 1986, p.155). Kriging and conditional simulation have been used on FIA data from the northeastern region of the U.S. (Riemann Hershey, Ramirez, & Drake, 1997; Moeur & Riemann Hershey, 1999; Riemann Hershey, 2000). Other applications of geostatistical techniques in forestry include mapping of forest resources (Samra, Gill, & Bhatia, 1989; Gunnarsson, Holm, Holmgren, & Thuresson, 1998), analysis of insect pests (Aleong, Parker, Skinner, & Howard, 1991; Liebhold, Holn, & Gribko, 1993; Speight, Hails, Gilbert, & Foggo, 1998), assesment of water stressors in streams (Yuan, 2004), and modeling of the effects of deer herbivory (Gribko, Holn, & Ford, 1999). Numerous other examples and applications can be found in Cressie (1991).

In this study, we investigated improving the point prediction accuracy of the nationwide forest biomass map by using kriging in conjunction with the existing Cubist biomass models. We explored and modeled spatial structure in the biomass data collected on forest inventory plots, as well as in the residual biomass from the Cubist biomass models, across 18 ecological zones in the Interior Western U.S. We began with exploratory techniques such as directional graphs of summary statistics and three-dimensional maps of the data to check if the assumption of stationarity required by kriging and other spatial analyses was appropriate for our dataset. Next, Moran's I correlograms and variograms were constructed to discover if spatial patterns were present in the data, either in the field biomass or in the residual biomass from the Cubist biomass models. In cases where a spatial pattern was present, kriging models were built to incorporate this spatial information into the model. Finally, we used independent test data to assess whether kriging, either of the field or of the residual biomass, offered substantial improvement to the existing biomass map.

2 Materials and Methods

2.1 Forest inventory and analysis (FIA) biomass study

FIA conducts inventories of status and trends in forested ecosystems nationwide (http://www.fia.fs.fed.us/). A network of sample plots has been established across the country at an intensity of approximately one plot per 6,000 acres. Biomass data collected on a total of 20,325 forested FIA plots in the Interior Western U.S. were used in this study. Correct spatial coordinates, as opposed to the privacy protected coordinates available on public data bases, were available for this work. This study used a mixture of current and historic data to ensure enough training plots across the entire study region, with collection dates ranging between 1990 and 2003. Data is currently collected annually on a rotating panel system, At each sample plot, data on forest stand structure is sampled in a cluster of four circular sub-plots, one center and three satellite sub-plots regularly spaced around the center sub-plot. Each sub-plot covers a radius of 7.3 m and the total area represented by one sample plot is approximately 0.6 ha. Historically, data was collected periodically (every 5 to 20 years) for each state, with plots consisting of a cluster of between 4 and 9 sub-plots. For both current and historic data the cluster average was used to determine each plot's biomass. Ten percent of the plots were withheld for testing model accuracy, as described later. USGS mapping zones based on Homer and Gallant (2001) were used as subpopulations for constructing individual models and analyses. The 67 mapping zones covering the U.S. and Alaska divide the nation into ecologically similar regions. The 18 zones covering the Interior West are shown in Figure 1. These zones are defined for the entire U.S.; therefore, the 18 zones in the Rocky Mountains are not numbered sequentially from 1 to 18.

We analyzed the spatial structure of the forest biomass, both the 'field biomass,' which is the above ground live biomass derived for each forested FIA plot as a function of individual tally trees, and also the 'residual biomass,' which is the difference between the field biomass and the predicted biomass from the Cubist biomass models developed for the nationwide forest biomass mapping project (Blackard et al., In review). In the nationwide mapping project, field biomass collected on FIA plots was modeled as functions of numerous 250-m predictor layers, includ-

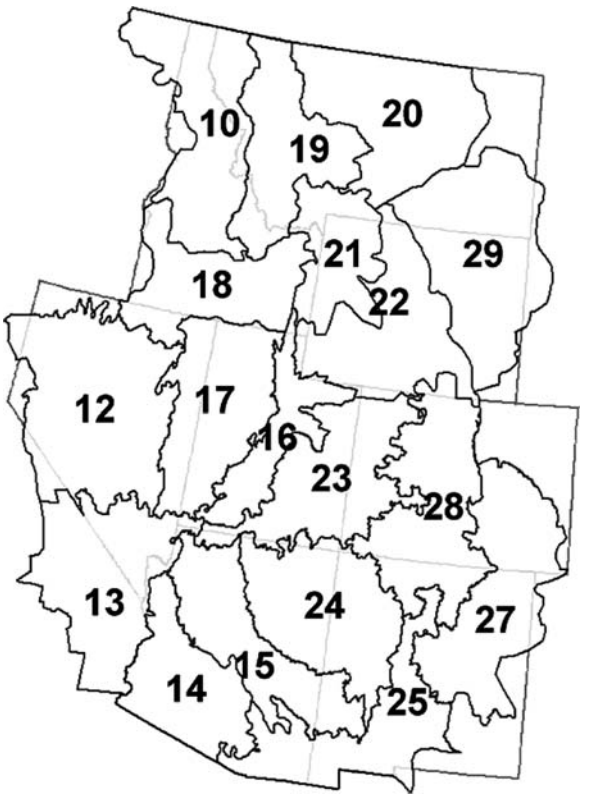

Figure 1 USGS mapping zones of the Rocky Mountain region.

ing 16-day Moderate Resolution Imaging Spectrometer (MODIS) composites, associated vegetation indices, and percent tree cover; vegetative diversity and type synthesized from the National Land Cover Dataset (NLCD); topographic variables derived from Digital Elevation Models (DEMs); monthly and annual climate parameters; and other ancillary variables. Nonparametric functions relating forest biomass to the predictor layers were built using Cubist software. Cubist is a variant on regression tree modeling, or recursive partitioning regression (Breiman, Friedman, Olshen, & Stone, 1984). Trees subdivide the space spanned by the predictor variables into regions for which the values of the response variable are most similar, and then estimate the response variable by a model parameter specific to each of these regions. After fitting a regression tree, Cubist fits a piecewise non-overlapping regression to improve tree results. Using resampling methods, a committee of models is fit (Breiman, 1996), and final model predictions are obtained by averaging committee model results. Details on the biomass modeling effort are available in Blackard, et al. (In review). An independent test set was created by randomly selecting 10% of the available 20,325 forested FIA plots. This same test set was withheld from the spatial modeling analyses and used to test model performance.

Descriptive statistics for each of the mapping zones are shown in Table I. The number of plots per zone varied widely, depending on both the total size and the proportion of forested area. For example, Zone 10 (northern Idaho and northwestern Montana) had 4,671 plots, while Zone 13 (southern Nevada and southeastern California) had only 64 plots. Mean field biomass for the entire region was 28.5 tons per acre but for each zone ranged from 1.3 tons per acre for Zone 14 (southwestern Arizona) to 48.1 tons per acre for Zone 10 (northern Idaho and northwestern Montana). The mean predicted biomass from the Cubist biomass models for the entire region was 27.2, and in fact, the Cubist biomass models slightly underpredicted the mean biomass in every zone, suggesting a slight bias in this model.

2.2 Stationarity assumptions

Variograms, Moran's I, and kriging all rely on some form of stationarity assumption. A descriptive use of

Table I Summary statistics for the zones

Zone	Number of plots	Mean field biomass	Standard dev. field biomass	Mean predicted biomass (Cubist)	Mean absolute residual biomass
10	4,671	48.1	38.8	45.7	23.8
12	1,002	11.2	10.8	10.9	6.5
13	64	13.0	16.4	12.7	4.8
14	168	1.3	2.7	1.0	0.9
15	2,262	17.7	21.7	16.5	8.2
16	2,057	25.8	24.6	24.8	14.1
17	806	12.3	11.1	12.2	6.5
18	247	24.6	24.5	23.4	12.7
19	2,060	37.8	27.4	36.0	17.3
20	207	23.3	20.8	23.1	9.6
21	1,366	35.9	30.8	35.5	18.3
22	135	12.3	16.8	11.3	5.7
23	913	13.9	15.9	12.9	7.4
24	1,160	10.0	11.7	9.5	5.1
25	636	11.1	16.5	10.4	4.4
27	277	9.8	10.4	9.3	4.0
28	1,681	30.9	26.7	29.9	12.9
29	613	22.1	20.8	21.3	9.4
All Zones	20,325	28.5	30.3	27.2	13.8

variograms only requires the less stringent *intrinsic assumption*, that the study area has a constant variance, though possibly a nonconstant mean. However significance tests of Moran's I and kriging both require the stronger assumption of *second-order stationarity*, that is, both the mean and covariance are constant throughout the study area. (Legendre & Legendre, 1998, p. 718–719).

If an unrecognized spatial trend is present in the population (e.g. stationarity assumptions were not justified) variograms and Moran's I correlograms may be reflecting this trend rather than fine-scale autocorrelation, leading to mis-identification of the ecological processes (Lichstein, Simon, Shriner, & Franzreb, 2002). Therefore, exploratory data analyses were conducted to identify any trends in the data before testing for small-scale spatial pattern.

To assess trend, directional graphs of the means, medians, and variances of the response variable were constructed for the entire Interior West region and for each zone. Directional graphs were constructed by dividing the map into strips and calculating the statistics for each strip (Figure 2). Both the mean and median were calculated because the median is more resistant to outliers (Isaaks & Srivastava, 1989, p. 17–18). In addition, three-dimensional maps (where the response variable is plotted on the *z*-axis) were constructed, including contour maps, perspective plots, and gray scale images.

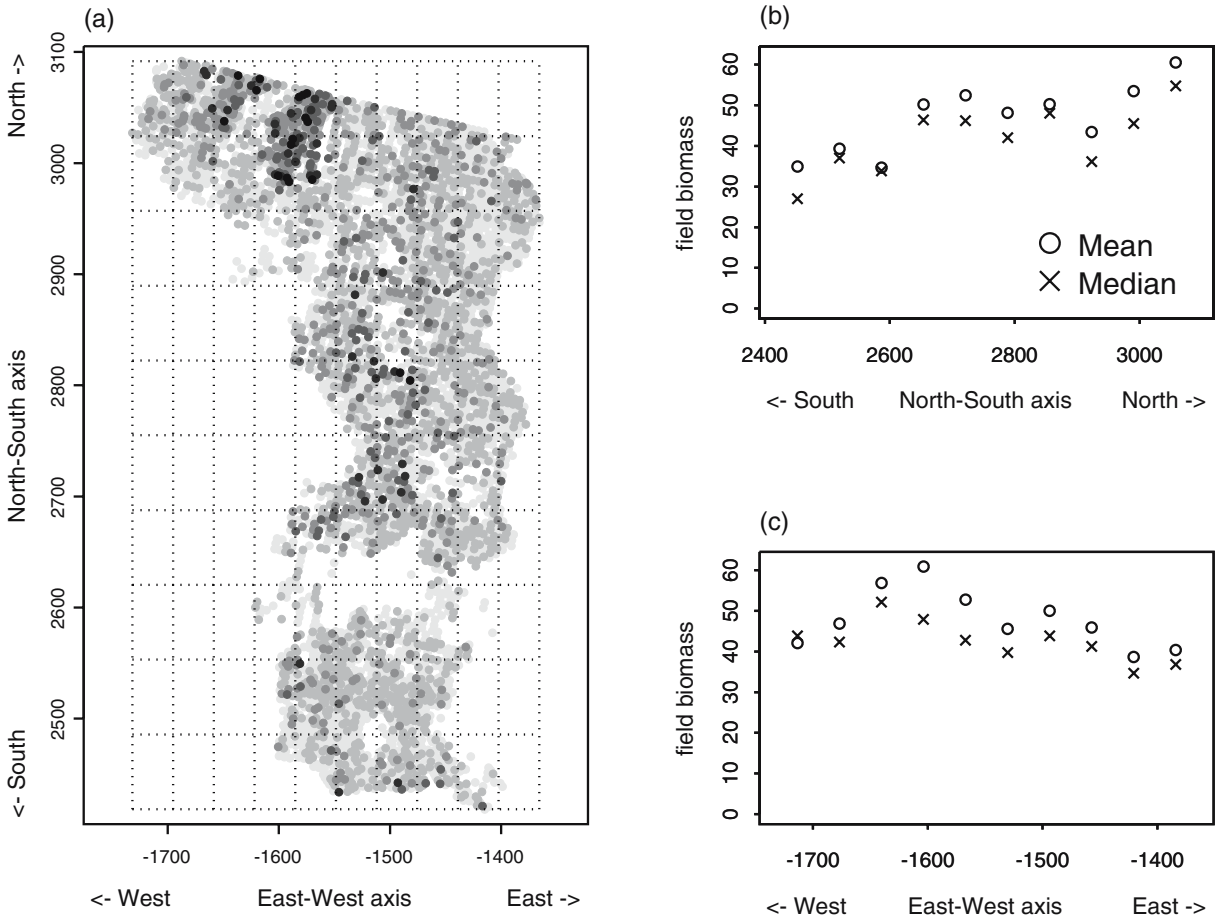

Figure 2 An example of the use of directional means and medians to check the assumption of stationary biomass. Even after splitting the Rocky Mountain region into subpopulations, several zones still show evidence of trend in their field biomass. In Zone 10 (Northern Idaho and Northwestern Montana), there is a North–South gradient, with increasing biomass to the north. Field biomass as a function of location for Zone 10: **a** map of plot locations, with *darker color* indicating higher field biomass, and *dotted lines* showing the *vertical and horizontal strips* used to calculate the directional means and medians. Mean and median field biomass along the **b** North–South axis and **c** East–West axis.

2.3 Moran's I correlograms

Moran's I correlograms (Moran, 1950; Legendre & Legendre, 1998, p. 714–728) are a valuable tool both for detecting autocorrelation and for identifying the range of the correlation when modeling the variogram. Note that significance tests of Moran's I assume second-order stationarity.

For all mapping zones, 20 lags (or distance classes) were used, evenly spaced at every 5 km, the most common grid spacing in this study, as suggested by Isaaks and Srivastava (1989, p. 144,146–149). The lag tolerance was one-half the lag width. Thus, the first lag would include all pairs between 2.5 and 7.5 km, and the last lag includes pairs between 97.5 and 102.5 km.

The raw Moran's I statistic for each lag was then calculated as:

$$I_{raw} = \left(\frac{n}{S_o}\right)\left[\frac{\sum_i \sum_j w_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\sum_i (z_i - \overline{z})^2}\right]$$

where n is the number of points, z_i is the value at point i, w_{ij} is the weight corresponding to the effect of point i on point j, and S_o is the sum of the weights w_{ij}.

A binary function of distance for the neighbor weights w_{ij} was used. Here, any pairs of points within a given lag were given a weight of 1, while pairs outside this lag were given a weight of 0. As a result, each inter-point distance was used in the calculation

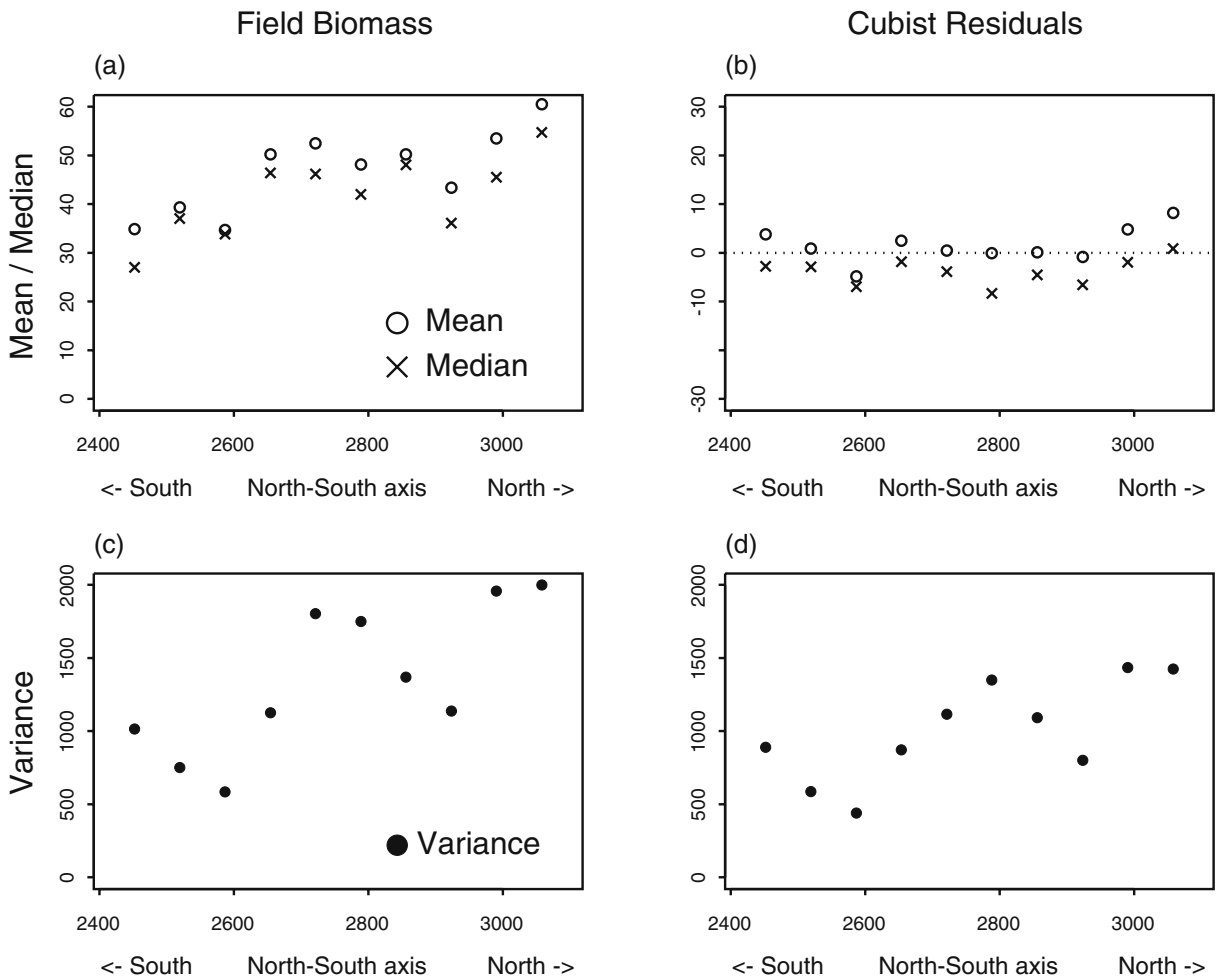

Figure 3 Directional means, medians, and variances for field biomass and residual biomass from the Cubist biomass models for Zone 10 (Northern Idaho and Northwestern Montana). The field biomass for Zone 10 has trend in the mean as well trend in the variance. The Cubist biomass models accounted for the trend in the mean, but the variance still exhibits trend.

of Moran's I for exactly one of the lags, and S_o for each lag equals the number of pairs in that lag.

As in any measure of autocorrelation, the numerator of Moran's I is a measure of covariance, and the denominator is a measure of variance (Cliff & Ord, 1981, p.13–17). Values of I_{raw} near 0 indicate no autocorrelation. However, unlike a correlation coefficient, I_{raw} does not vary strictly between −1 and 1. The Cauchy–Schwartz inequality shows that I_{raw} has a maximum value of:

$$|I_{max}| = \left(\frac{n}{S_o}\right)\left\{\frac{\sum_i\left[\sum_j w_{ij}\left(z_j-\overline{z}\right)\right]^2}{\sum_i\left(z_i-\overline{z}\right)^2}\right\}^{\frac{1}{2}}$$

Standardized Moran's I values were also calculated so that they varied between −1 and 1 (Cliff & Ord, 1981, p. 21–22; Haining, 1990, p. 234–235; Lichstein et al., 2002, Appendix):

$$I_{std} = \frac{I_{raw}}{I_{max}}$$

Monte Carlo tests for significance were performed where the standardized Moran's I statistic from the observed biomass was tested against 999 random permutations. For each permutation, the plot locations were kept constant, but the observed values were randomly shuffled amongst these fixed locations, and Moran's I is calculated for each lag of this permuted dataset.

Probabilities for a one-tailed test as presented in Legendre and Legendre (1998, p. 22–24) were calculated, with a null hypothesis of no autocorrelation *versus* an alternative hypothesis of positive autocorrelation such that:

$probability = \frac{N_{permutation}\left(I_{permutation} > I_{observed}\right)+1}{N_{permutation}+1}$ where $N_{permutation}$ is the number of Monte Carlo permutations (in our case 999), and $N_{permutation}(I_{permutation} > I_{observed})$ is the number of permutations whose Moran's I for a given lag is greater the Moran's I for the observed data.

With α=0.05, a progressive Bonferroni correction (Legendre & Legendre, 1998, p. 671– 673) was used to test each distance lag. The progressive Bonferroni correction is appropriate to determine the range of significant autocorrelation, when autocorrelation, if present at all, is expected to be found in the smallest distance classes. In this test, the *i*th lag is tested at $\alpha=\alpha/i$. A one-sided test was used because the autocorrelation, if present, was expected to be positive.

As an example, with the progressive Bonnferoni correction, the second lag will be tested at a corrected alpha level of 0.025, so if 19 of the 999 permutations for this lag give a Moran's I greater than that of the observed data, resulting in a *P* value of 0.02, then this lag will be declared significant.

2.4 Variograms and kriging

2.4.1 Variograms

Variograms (actually semi-variograms, but commonly referred to as variograms for simplicity) are used to identify and describe spatial structure. Variograms are the first step toward modeling the small-scale spatial structure for use with kriging.

Semi-variance is computed by:

$$\gamma(h) = \frac{1}{2|N(h)|}\sum_{N(h)}\left(z_i - z_j\right)^2$$

where $\gamma(h)$ is the estimated semivariance for lag h, $N(h)$ is the set of all pair-wise distances ($i - j = h$),

Table II Distance (in meters) of largest lag with significant spatial autocorrelation from Moran's I correlograms for the field biomass and for the residual biomass from the Cubist biomass models

Zone	Range of significant spatial autocorrelation	
	Field biomass	Residual biomass
10	100	75*
12	5	5
15	60	30
16	15	10
17	40*	–
19	20	5
21	30	–
23	90*	15
24	40	40*
25	70	15
28	100	25
29	80	30*

* indicates zones with four or more non-significant lags at smaller distances.

– indicates zones with no significant autocorrelation at any distance.

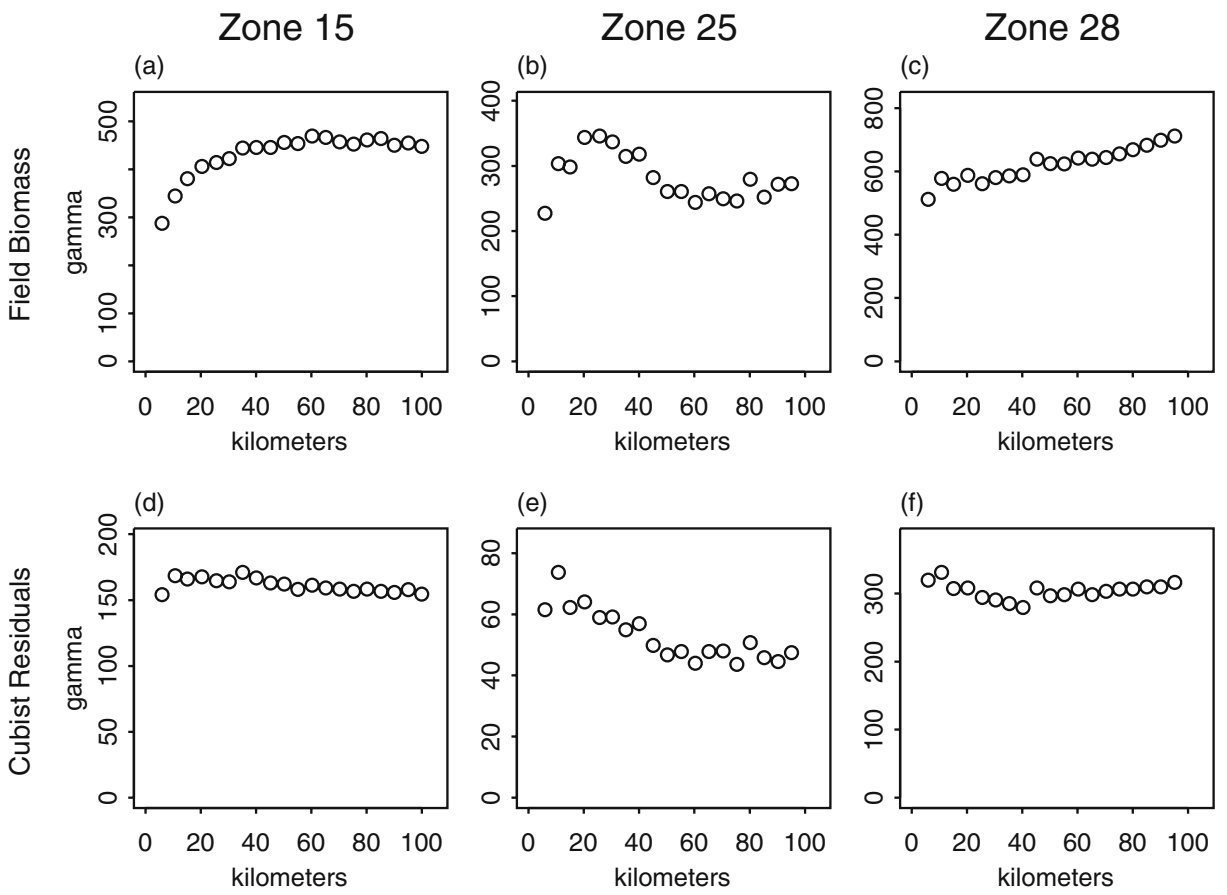

Figure 4 Variograms of field biomass and residual biomass from the Cubist biomass models for the zones where kriging the field biomass was most effective: Zone 15 (Mogollon Rim, Arizona), Zone 25 (Southeastern Arizona and Southwestern New Mexico), and Zone 28 (Central Colorado). Even when the variogram for field biomass was reasonably well behaved, the variogram for residual biomass showed little evidence of spatial auto-correlation.

$|N(h)|$ is the number of pairs in $N(h)$, and z_i and z_j are the data values (of field biomass or residual biomass) at locations i and j. To allow easy comparison, the lags used for the variograms were the same as those used for Moran's I.

For this study, omnidirectional variograms (averaged over all directions) were calculated for both field and residual biomass for all zones with sample sizes greater than $N = 500$. Anisotropic variograms (calculated along given directions) were also examined in an attempt to check if spatial structure changed with direction. However, Isaaks and Srivastava (1989, p. 144) note that if the omnidirectional variogram is poorly behaved without a clearly interpretable structure (as were most of our variograms) the anisotropic variograms are unlikely to exhibit any better behavior. We found this to be true. Even in the few zones where the omnidirectional variogram was relatively well behaved, the anisotropic variograms were indecipherable. Therefore, we relied on omnidirectional variograms for our analyses.

Three parameters are commonly used to describe and model the behavior of variograms: Range, sill, and nugget. The *range* is the distance where the spatial correlation vanishes and the variogram levels off. The height of the variogram after leveling off is known as the *sill*. The height of the variogram as it crosses the y-axis is the *nugget*.

The ratio of the nugget to the sill is known as the *relative nugget effect*. It gives a measure of the percentage of the variability in the data from sources other than spatial autocorrelation. A low relative nugget effect is a sign of strong spatial autocorrelation in the data, and these datasets are more likely to

Table III Mean square error relative to the reference model of mean training biomass, and correlation coefficients between the predicted values and the actual values for the 10% test dataset for the Cubist biomass models and the kriging model

Zone	$\frac{MSE_{cubist}}{MSE_{mean}}$	$\frac{MSE_{krige}}{MSE_{mean}}$	ρ_{cubist}	ρ_{krige}	$\frac{MSE_{cubist+krige}}{MSE_{cubist}}$
10	0.76	0.87	0.52	0.37	0.95
12	0.88	0.94	0.35	0.22	0.99
15	0.39	0.51	0.80	0.72	0.91
16	0.80	0.89	0.45	0.35	0.97
17	0.87	0.99	0.35	0.09	1.00
19	0.74	0.94	0.54	0.30	0.97
21	0.73	0.93	0.54	0.26	1.00
23	0.68	0.94	0.56	0.24	0.99
24	0.71	0.91	0.55	0.42	1.00
25	0.34	0.59	0.82	0.65	1.09
28	0.56	0.75	0.67	0.50	1.00
29	0.59	0.90	0.67	0.59	0.95

The Cubist biomass models did a better job than kriging for all twelve zones. The final column gives the MSE of kriging the residuals of the Cubist biomass models relative to the MSE of the existing Cubist biomass models. The addition of kriging resulted in little or no improvement for most zones, and even in the best case, only resulted in a slight improvement.

benefit from geostatistical techniques. A high relative nugget effect can indicate no autocorrelation in the population, but it can also be caused by high measurement variability, or by spatial pattern at scales smaller than the sampling distance.

The first step in kriging is to fit a model to the variogram. We used a spherical model:

$$\gamma(h) = c_0 + c_1\left\{\frac{3h}{2a} - \frac{1}{2}\left(\frac{h}{a}\right)^3\right\} \quad for \quad 0 < h < a$$

$$\gamma(h) = c_0 + c_1 \quad for \quad h > a$$

where a is the range, h is the lag, c_0 is the nugget, and $c_0 + c_1$ is the sill. The spherical function was fit to the variogram by eye, with the Moran's I correlograms used to suggest a starting point for the range parameter.

2.4.2 Kriging

Kriging was attempted on the 12 zones with sample sizes greater than 500. Ordinary point kriging with a spherical model for the variogram was conducted with a lag length of 5 km and a maximum distance of 100 km. Some zones did contain trends, which would suggest the use of universal kriging. However, Burrough (1986, p. 161) warns that if there is a high relative nugget effect (such as we have in our data) it is difficult to separate true trends from sample variation, and in such cases concludes that universal kriging offers little benefit over ordinary kriging. For this reason, ordinary kriging was used, even when we had reason to believe trend was present.

2.4.3 Model comparisons

First, the point prediction accuracies of kriging and of the Cubist biomass models were examined individually. Both accuracies were evaluated relative to a reference model where the predicted biomass for the test data is simply the mean field biomass from that zone's training data. Second, kriging was applied to the residuals from the Cubist biomass models, and the resulting predictions were evaluated relative to the predictions from the Cubist biomass models alone. Although the focus here is on point prediction accuracies, maps generated using the models in question would be produced at 250 m resolution to match resolution of ancillary data used in the Cubist biomass models.

To asses accuracy, the models were first fit to the training dataset, and then used to predict biomass on the independent 10% test data. To evaluate prediction accuracy, relative mean square error (*MSE*) and the correlation coefficient, calculated between the predicted and the field biomass for each zone's test data, were used. The relative MSE was the ratio of the MSE from the new model to that of the old model. A new model that is substantially better than the old one would have a relative MSE considerably less than 1.

3 Results

3.1 Stationarity assumptions

As expected, neither the mean nor the variance of field biomass was uniform over the entire Rocky Mountain region; therefore, pooling the entire dataset was not justified under the assumption of *second-order stationarity*, nor even under the more relaxed *intrinsic assumption*. The Rocky Mountain region is a large and varied study area, with many different habitats and species compositions. Much of the variation in biomass is likely due to multiple populations, rather than genuine trend. After all,

woodland forests of the desert Southwest have lower biomass than the northern conifer forests of Idaho and Montana. Also, biomass is higher in the central mountains than at the edges, where the forests are blending into the desserts and plains. Therefore, the dataset was split into separate populations, and analyses conducted on each zone individually.

Splitting the data into zones reduced the trend in many of the zones by limiting the analysis to single populations. For example, after splitting, Zone 16 (Utah mountains) and Zone 21 (Yellowstone region) had little trend in their directional means and medians.

In some cases, however, there was still evidence of substantial trend even within single zones. For example, in Zone 10 (northern Idaho and northwestern Montana) there was a north–south gradient, with increasing field biomass to the north (Figure 2). And in Zone 15 (Mogollon Rim, Arizona) there was a southwest to northeast gradient, with increasing field biomass to the northeast.

After examining the residual biomass from the Cubist biomass models, we found that in most zones the spatially correlated predictor variables of the models were successful in accounting for the large-scale trends in the mean and median values. In Zone 10 (northern Idaho and northwest Montana) the

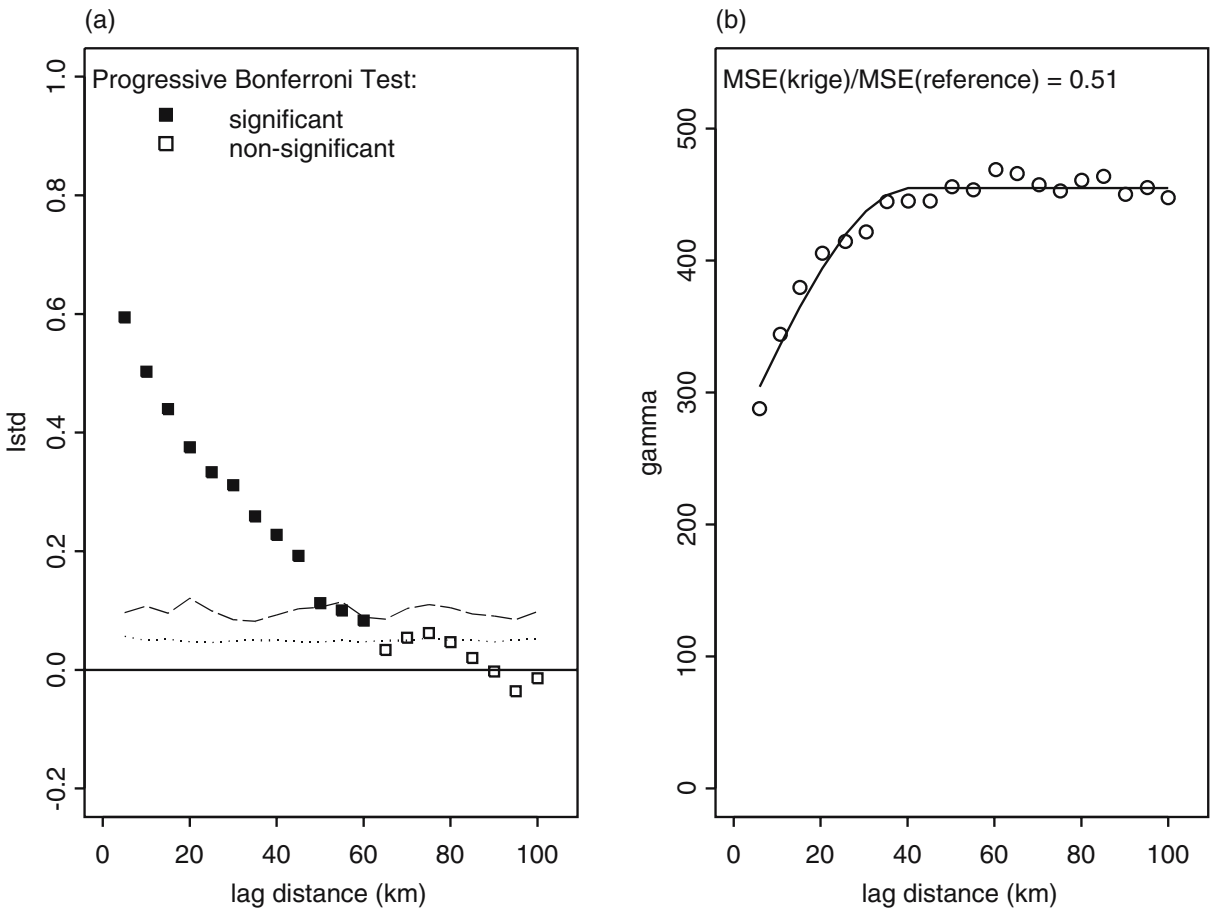

Figure 5 Zone 15 (Mogollon Rim, Arizona) field biomass **a** standardized Moran's I correlogram and **b** variogram modeled with a spherical covariance function. Confidence envelopes for Moran's I are from the 950th and the 999th largest simulations at each lag (not Bonferroni). Kriging was most effective in zones with strongly significant Moran's I and well behaved variograms. For Zone 15 MSE of kriging relative to the reference model was 0.51. Though even in this case, kriging still was not as effective as the Cubist biomass models. For comparison, the MSE of the Cubist model relative to the reference model was 0.39.

north–south gradient from the field biomass was nearly absent in the residual biomass (Figure 3). Even in Zone 15 (Mogollon Rim, Arizona), the trend in the mean values was much improved.

However, stationarity assumes not just stationary means, but also stationary variances. The Cubist biomass models were not completely successful in correcting for the trends in the variance. For example, the residual biomass for Zone 10 (northern Idaho and northwest Montana) still had a north–south trend in its variance (Figure 3).

3.2 Moran's I correlograms

Zones with small sample sizes showed little or no pattern in their Moran's I correlograms; however, the field biomass had significant autocorrelation for the 12 zones with sample sizes greater than 500. The residual biomass had significant autocorrelation in 10 of these 12 zones, though this autocorrelation was reduced in both strength and range when compared to that of the field biomass (Table II).

Many zones that did not show trend in the preliminary analysis, such as Zone 16 (Utah mountains) and Zone 21 (Yellowstone region), did have significant small-scale spatial autocorrelation. In Zone 21 the Cubist biomass models were able to successfully account for this autocorrelation, leaving no significant autocorrelation in the residual biomass. For Zone 16, the Cubist biomass models substantially reduced the autocorrelation.

In zones with strong trends in field biomass (Zone 10, Zone 15), the Cubist biomass models did not completely account for the spatial autocorrelation.

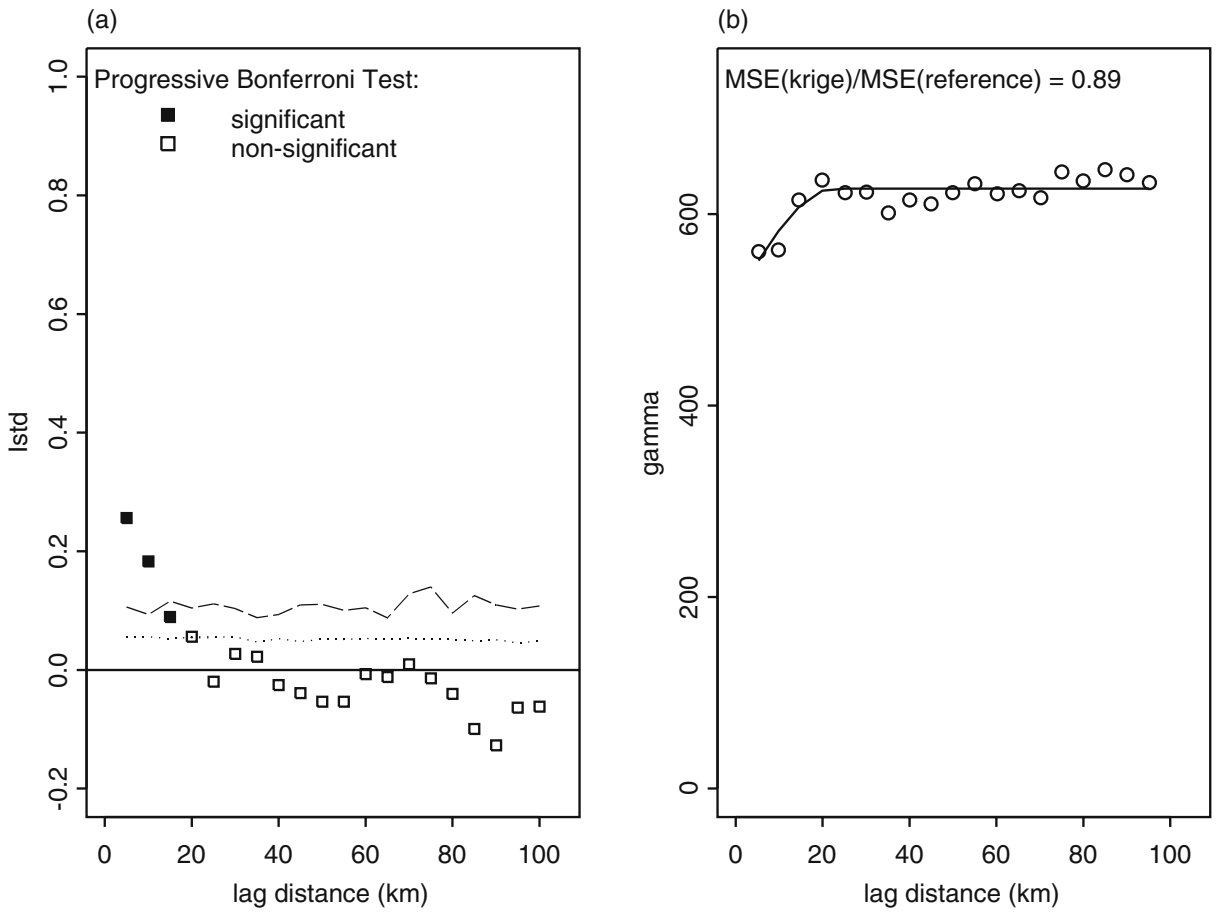

Figure 6 Zone 16 (Utah Mountains) field biomass **a** standardized Moran's I correlogram and **b** variogram modeled with a spherical covariance function. Confidence envelopes for Moran's I are from the 950th and the 999th largest simulations at each lag (not Bonferroni). Zone 16 had little trend, and a much shorter range of small-scale spatial autocorrelation, but it still had a reasonably well-behaved variogram, and kriging was able to effect at least a slight improvement over the reference model.

However, it did succeed in reducing both the range and strength of spatial autocorrelation.

3.3 Variograms and kriging

Many of the variograms for the field biomass and most of the variograms for the residual biomass were poorly behaved (Figure 4). In general, they did not rise smoothly to a sill, and they jumped around erratically, or even decreased with distance. The anisotropic variograms exhibited still worse behavior. Even when the omnidirectional variogram was reasonably well behaved, the anisotropic variograms were erratic, rising to different sills, and not following any decipherable pattern. Fitting a model to the variogram is problematical with erratic variograms such as these. Because model fitting is the first step in kriging, it is not surprising that when prediction accuracies for the 10% test data were compared, the Cubist biomass models outperformed kriging for all 12 zones. This was true both for the MSE relative to the reference model and for the correlation coefficients. The relative MSE from the Cubist biomass models ranged from 0.34 to 0.88, while the relative MSE from kriging ranged from 0.51 to 0.99. The correlation coefficient for the Cubist biomass models ranged from 0.35 to 0.80 for the 12 zones, while kriging ranged from 0.09 to 0.65 (Table III, Figures 5, 6, 7, 8).

When kriging was added to the Cubist biomass models, there was little, if any, improvement in most zones. At best, the addition of kriging resulted in only slightly improved predictions. The MSE of the Cubist biomass models plus kriging (Cubist + Krige) relative

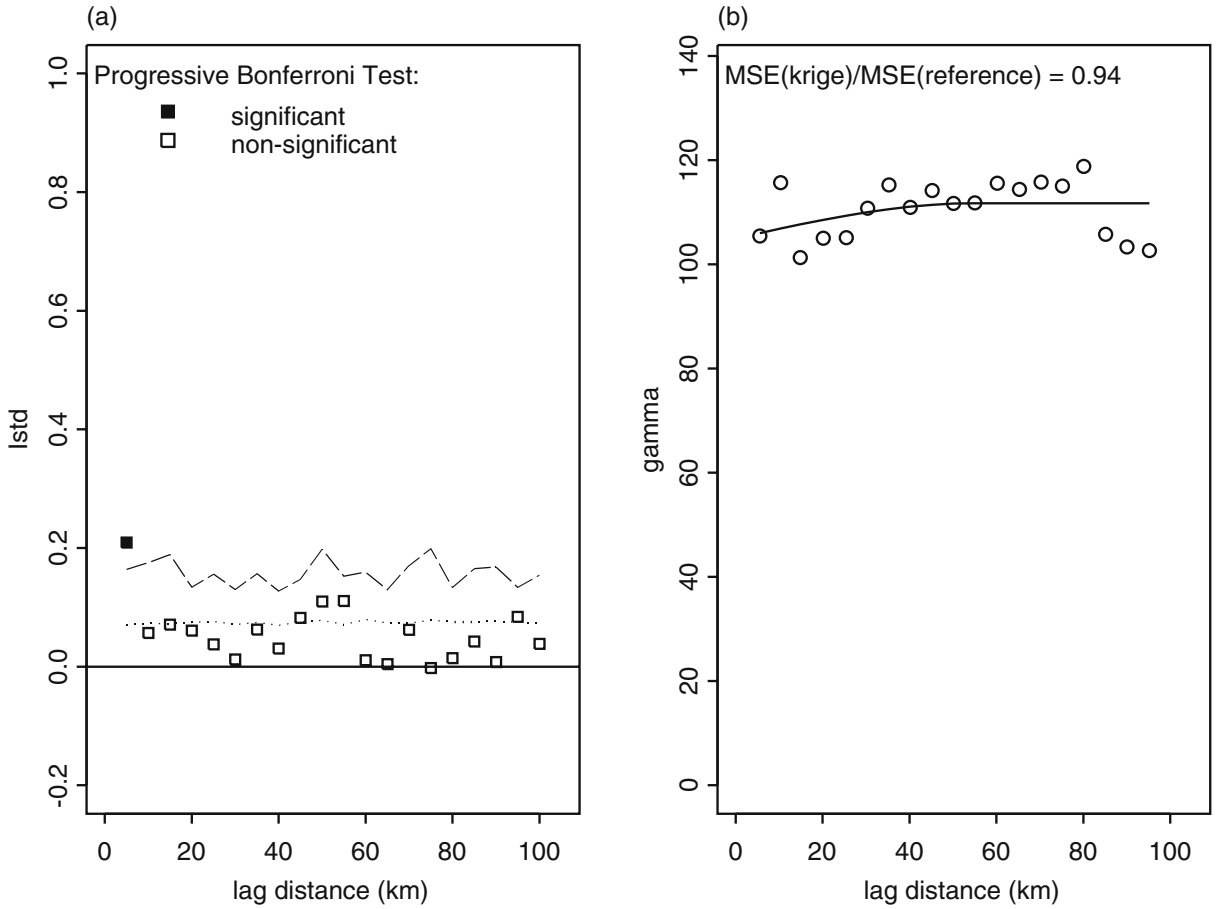

Figure 7 Zone 12 (Nevada) field biomass **a** standardized Moran's I correlogram and **b** variogram modeled with a spherical covariance function. Confidence envelopes for Moran's I are from the 950th and the 999th largest simulations at each lag (not Bonferroni). In zones with little spatial autocorrelation kriging was much less successful at predicting biomass, and resulted in very little improvement over the reference model.

to the original Cubist biomass models varied between 0.91 and 1.09, with only three zones having a relative MSE less than 0.97 (Table III, Figure 9).

4 Discussion

Kriging was most effective in zones with strongly significant Moran's I and well-behaved variograms, such as Zone 15 (Mogollon Rim, Arizona – Figure 5). By comparison, zones with little spatial autocorrelation, such as Zone 12 (Nevada – Figure 7), received much less benefit from kriging.

The behavior of the variogram was at least as important as the presence of strong autocorrelation in determining which zones would respond best to kriging. In zones such as Zone 16 (Utah Mountains – Figure 6), which had smaller strength and range of autocorrelation but had reasonably well-behaved variograms, kriging outperformed the reference model of mean training biomass (though not the Cubist biomass models). In contrast, in zones such as Zone 29 (southeastern Montana and eastern Wyoming – Figure 8), with strong autocorrelation but poorly behaved variograms, kriging offered little improvement in prediction accuracy over the reference model.

The variogram for Zone 25 (Southeastern Arizona and Southwestern New Mexico – Figure 4b) is an interesting example of a poorly behaved variogram. It exhibits what is sometimes termed a 'hole effect' where the variance first increases then decreases as

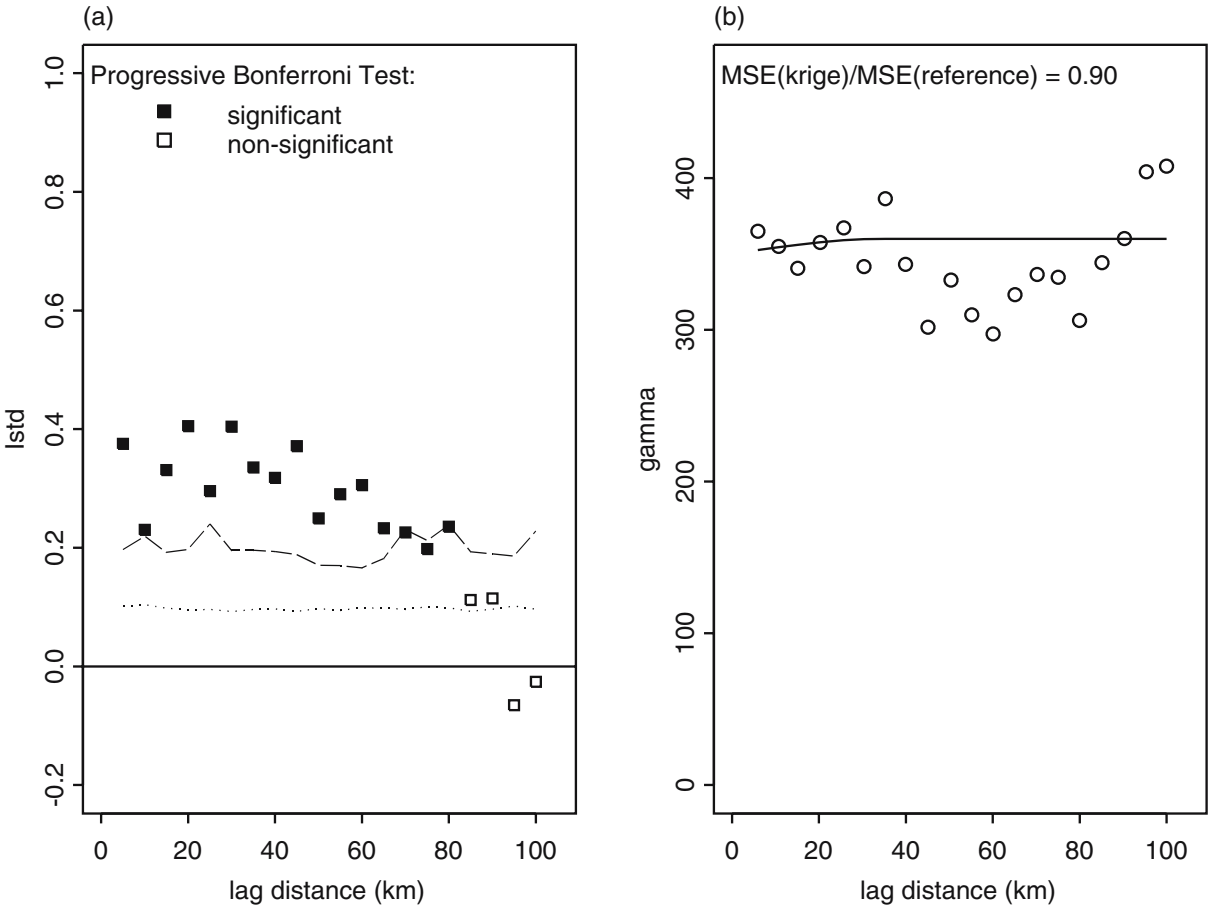

Figure 8 Zone 29 (Southeastern Montana and Eastern Wyoming) field biomass **a** standardized Moran's I correlogram and **b** variogram modeled with a spherical covariance function. Confidence envelopes for Moran's I are from the 950th and the 999th largest simulations at each lag (not Bonferroni). Some zones with strong spatial autocorrelation still had poorly behaved variograms. In these cases kriging was much less successful than the Cubist biomass models. For example in Zone 29, the MSE of kriging relative to the reference model was 0.90, while the MSE of the Cubist model relative to the reference model was 0.59.

the distance increases. This often occurs in mountainous regions, when the mountains are at a relatively large scale relative to the grid size of the plot locations. Thus, for example, a valley plot might have medium distant neighbors on a mountaintop, while further away is another very similar valley (and vice verse). Zone 25 is located in the 'islands in the sky' region of Arizona and New Mexico, where the terrain consists of 'islands' of forested mountain ranges separated by long stretches of desert. Both the 'islands' of forest and the stretches of desert are at a scale much larger than the 5 km spacing of our plot grid. On the other hand, when the plot locations are widely spaced relative to the horizontal scale of the mountains, as in many of our other zones, then the effect is merely to add noise to the variograms. In these cases, there may be two mountains and three valleys within 5 km, and thus the plots surrounding a valley may be mountaintops or may be other valleys, irrespective of their distance.

The above examples are from kriging applied to the field biomass. As noted earlier, variograms for the residual biomass from the Cubist biomass models were even more poorly behaved than those for the field data. As a result, kriging did not work well when analyzing this residual biomass. When considering all zones, kriging was the most successful on the residual biomass of Zone 15 (Mogollon Rim, Arizona – Figure 9), and even then, the MSE of the Cubist biomass models with the addition of kriging relative to the MSE of the existing Cubist biomass models was only 0.91.

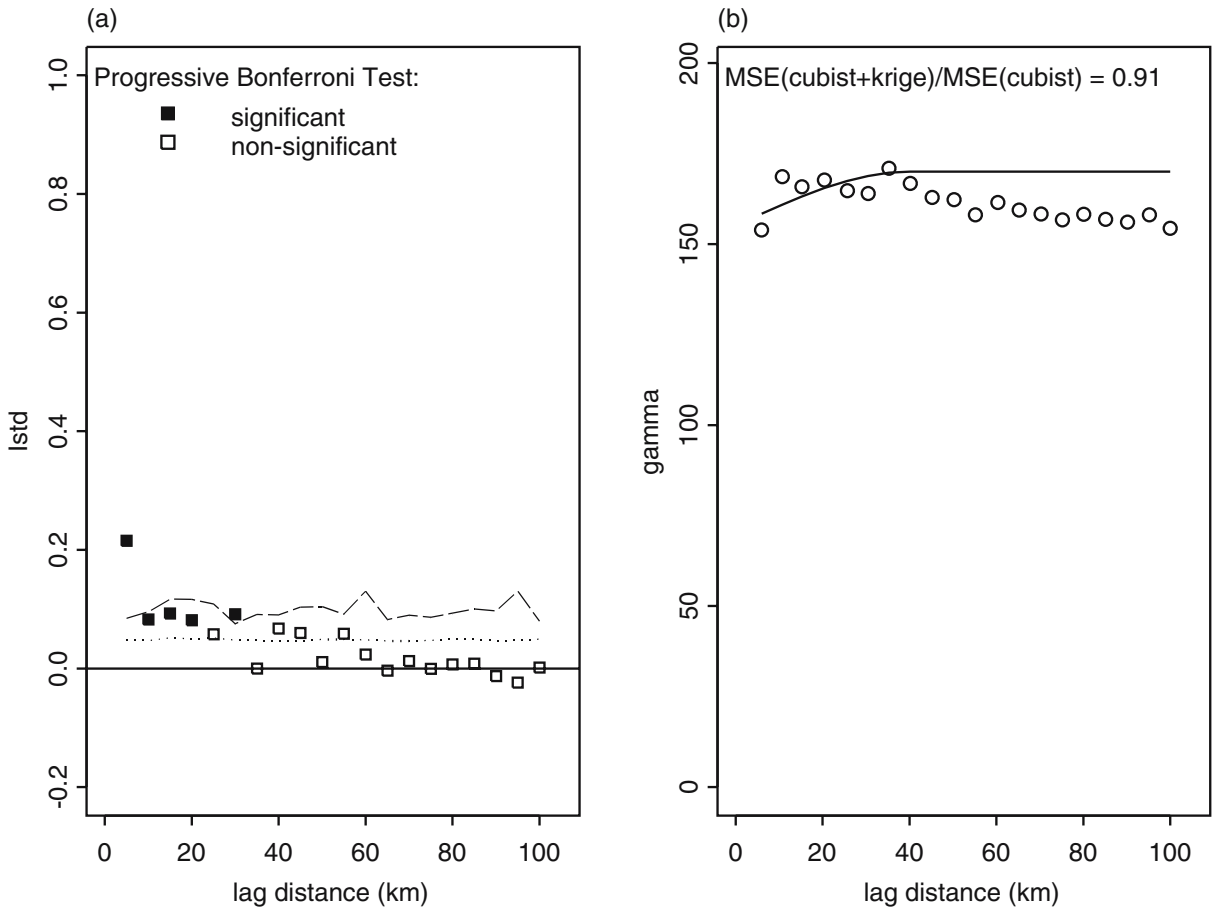

Figure 9 Zone 15 residual biomass: **a** standardized Moran's I correlogram and **b** Variogram modeled with a spherical covariance function. Confidence envelopes for Moran's I are from the 950th and the 999th largest simulations at each lag (not Bonferroni). Kriging the residual biomass was the most successful on Zone 15 (Mogollon Rim, Arizona), and even so the MSE of Cubist + Kriging relative to the existing Cubist biomass models, was only 0.91, a very slight improvement.

As we discovered, successful kriging depends on a well-behaved variogram, and there are many possible causes of the poor variogram behavior seen in this study. For example, inconsistent sampling strategies, scale miss-match between field and remotely sensed data, uneven sample size for each lag, non-normal data, outliers in the data, unjustified assumption of stationarity, and a high relative nugget effect can all cause variograms to behave erratically. All of these problems have the potential to affect results from this study.

Some of these issues are beyond the scope of this paper. However, we attempted corrections for those issues that were under our control. For example, dividing the study area into zones reduced the trend caused by multiple populations. The variogram behavior for our semigridded data was improved by setting the lag length to the most common grid spacing (Isaaks & Srivastava, 1989, p. 144,146–149). Corrections for other potential causes were attempted, including outlier removal, Cressie and Hawkins' (1980) robust estimator of the variogram, and a normal score transform of the response, as utilized on FIA data collected in the northeastern U.S. (Riemann Hershey et al., 1997; Riemann Hershey, 2000). However, with this dataset these further attempts at correction gave little benefit and, in fact, introduced other problems.

For example, most of the zones had extreme data points. Gunst and Hartfield's (1997) preliminary test estimator identified between 0.2 and 23.0% of the plots in each zone as potential outliers, usually for having too high field biomass values. One tendency of the Cubist biomass models that we hope to improve, is that they generally predict biomass values conservatively, so that relatively high biomass values are usually underpredicted. Even if removing the outliers had improved the variograms, this removal would be counter to reducing this underprediction of high biomass values. In fact, neither the Cubist modeling technique nor kriging were well suited to address these high biomass values. An interesting avenue for further study would be to investigate other techniques such as conditional simulation to identify regions of higher uncertainty, or quantile regression to explore the possibility of unmeasured limiting factors in these regions.

Lack of stationarity can make evaluating variograms difficult, and preliminary analyses showed that the zones with strong trends in the mean field biomass also had strong trends in the variance. While the Cubist biomass models did account for the trends in the mean, it did not correct trends in the variance. Therefore, for these zones, neither the field biomass nor the residual biomass met the assumptions of any of the methods of spatial analyses addressed in this paper: Moran's I correlograms, variograms, or kriging.

Some of the poor variogram behavior we observed may be due to this trend. However, many zones with little or no trend, such as Zone 16 (Utah Mountains – Figure 6), still had poor variogram behavior. Also, Zone 15 (Mogollon Rim, Arizona – Figure 5), which had a strong trend in mean field biomass, had a well-behaved variogram. Therefore, while this failure to meet the stationarity assumption may be contributing to kriging's relatively poor prediction accuracy, it is not the primary cause of the exhibited variogram behavior.

Most likely, the chief cause of the poor variogram behavior was a high relative nugget effect, probably caused by spatial variation on a scale smaller than the 5 km grid spacing of this study. Many of the zones are in mountainous terrain, where traveling 5 km can result in completely different environmental conditions. Correcting this, unfortunately, would require additional fieldwork with a more intense sampling scheme to capture finer scale variations.

5 Conclusion

Neither the field biomass nor the residual biomass in the Interior Western States proved to be good candidates for kriging. In the case of the field biomass, in many of the zones the assumption of stationary means and variances was unjustified. Also, much of the spatial variation in this mountainous region takes place on a scale smaller than the grid spacing, resulting in high relative nugget effects. Kriging was most successful for zones with strongly significant Moran's I correlograms and well-behaved variograms. Even here, the Cubist biomass models did a much better job of predicting field biomass than kriging alone, as one would expect given that they incorporated multiple predictor variables.

When we attempted to improve the Cubist biomass models by kriging the residual biomass, we found that the Cubist biomass models had already accounted for much of the spatial pattern in the data. Both the large-

scale trends and the small-scale spatial autocorrelation were accounted for, presumably through the use of spatially correlated predictor variables. Therefore, little or no spatial autocorrelation remained in the residuals.

This is not to say that the Cubist biomass models did a perfect job. There is still spatial pattern present in the residual biomass in some zones, particularly in the zones that showed a strong trend. Even in these zones, however, the spatial pattern is much reduced in both strength and range.

Spatial autocorrelation in the residuals from a model is often a sign that an important environmental variable has been overlooked, and therefore spatial analysis of residuals can be used as a check of model fit similar to more traditional residual plots (Cablk, White, & Kiester, 2002; Upton & Fingleton, 1985, p. 336). In this study, even though we found that kriging did not markedly improve our model predictions, the process of spatial analysis provided us with valuable information about our current map, showing that it may be worthwhile to re-examine the Cubist biomass models for any additional variables that could be incorporated to improve their ability to work with spatially autocorrelated data.

In fact, this study, by pinpointing the zones and spatial scales where the Cubist biomass models had the most difficulty, may suggest starting points for discovering missing spatial variables. For example, in many of the zones with strong apparent trends, there was also environmental variation, for example, the elevation change of the Mogollon Rim in Zone 15, where field biomass increases as plots move from the lower elevation mesquite and juniper to the ponderosa pines above the rim. Finding a basis for further subdividing such zones into separate populations might prove more successful than de-trending.

In conclusion, the Cubist modeling technique was a more effective approach than kriging for this dataset, and the addition of kriging to the Cubist biomass models resulted in little gain in the quality of predictions.

References

Aleong, J., Parker, B. L., Skinner, M., & Howard, D. (1991). Analysis of thrips distribution: application of spatial statistics and kriging. *Toward understanding Thysanoptera: proceedings, International Conference on Thrips*, February 21–23, 1989, Burlington, Vermont, USA. Radnor, Pennsylvania: U.S. Dep. of Agriculture, Forest Service, Northeastern Forest Experimental Station, 1991, Gen. Tech. Rep. NE-147.

Blackard, J., Finco, M., Helmer, E., Holden, G., Hoppus, H., Jacobs, D., Lister, A., Moisen, G., Nelson, M., Riemann, R., Ruefenacht, B., Salajanu, D., Weyermann, D. Winterberger, K., Brandeis, R., Czaplewski, R., McRoberts, R., Patterson, P., & Tymcio, R. (In review). Mapping U.S. forest biomass using nationwide forest inventory data and MODIS-based information. *Remote Sensing of Environment*.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *26*, 123–140.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Monterey, California: Wadsworth and Brooks/Cole.

Burrough, P. A. (1986) *Principles of geographical Information Systems for land resource assessment, Monographs on soil and resources survey, Number 12*. Oxford, UK: Oxford Science.

Cablk, M., White, D., & Kiester, A. R. (2002). Assessment of spatial autocorrelation in empirical models in ecology. In J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall & F. B. Samson (Eds.), *Predicting species occurrences: Issues of accuracy and scale* (pp. 429–440). Washington, District of Columbia: Island Press.

Cliff, A. D., & Ord, J. K. (1981). *Spatial processes: Models and applications*. London: Pion.

Cressie, N. A. C. (1991). *Statistics for spatial data*. New York: Wiley.

Cressie, N., & Hawkins, D. M. (1980). Robust estimation of the variogram, I. *Mathematical Geology*, *12*, 115–125.

Gribko, L. S., Hohn, M. E., & Ford, W. M. (1999). White-tailed deer impact on forest regeneration: modeling landscape-level deer activity patterns. In: *Proceedings, 12th Central Hardwood Forest Conference*, February 28–March 1–2, 1999, Lexington, Kentucky, USA, U.S. Dep. of Agriculture, Forest Service, Southern Research Station, 1999, Gen. Tech. Rep. SRS-24.

Gunnarsson, F., Holm, S., Holmgren, P., & Thuresson, T. (1998). On the potential of kriging for forest management planning. *Scandinavian Journal of Forest Research*, *13*(2), 235–237.

Gunst, R. F., & Hartfield, M. I. (1997). Robust semivariogram estimation in the presence of influential spatial data values. In T. G. Gregoire, D. R. Brillinger, P. J. Diggle, E. Russek-Cohen, W. G. Warren & R. D. Wolfinger (Eds.), *Modeling longitudinal and spatially correlated data: Methods, applications, and future directions, lecture notes in statistics (vol 122)* (pp. 265–274). Berlin Heidelberg New York: Springer.

Haining, R. (1990). *Spatial data analyses in the social and environmental sciences*. Cambridge: Cambridge University Press.

Homer, C. G., & Gallant, A. (2001). Partitioning the conterminous United States into mapping zones for landsat TM land cover mapping. USGS Draft White Paper.

Isaaks, E. H., & Srivastava, R. M. (1989). *An introduction to applied geostatistics*. New York, USA: Oxford University Press.

Legendre, P., & Legendre, L. (1998). *Numerical ecology*. Second english edition. Amsterdam, The Netherlands: Elsevier Science.

Lichstein, J. W., Simons, T. R., Shriner, S. A., & Franzreb, K. E. (2002). Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs*, *72*, 445–463.

Liebhold, A., Hohn, M. E., & Gribko, L. S. (1993). Forcasting the spatial dynamics of Gypsy moth defoliation using 3-dimensional kriging. In A. M. Liebhold & H. R. Barrett (Eds.), *Proceedings: Spatial analysis and forest pest management*, April 27–30, 1992, Mountain Lakes, Virginia, USA. Radnor, Pennsylvania: U.S. Dep. of Agriculture, Forest Service, Northeastern Forest Experimental Station, 1993, Gen. Tech. Rep. NE-175.

Moeur, M., & Riemann Hershey, R. (1999). Preserving spatial and attribute correlation in the interpolation of forest inventory data. In K. Lowell & A. Jaton (Eds.), *Spatial accuracy assessment: Land information uncertainty in natural resources* (pp. 419–429). Chelsea, Michigan: Ann Arbor.

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*, 17–23.

Riemann Hershey, R. (2000). Modeling the spatial distribution of ten tree species in Pennsylvania. In H. T. Mowrer & R. G. Congalton (Eds.), *Quantifying spatial uncertainty in natural resources: theory and applications for GIS and remote sensing* (pp. 119–136). Chelsea, Michigan: Sleeping Bear/Ann Arbor.

Riemann Hershey, R., Ramirez M. A., & Drake, D. A. (1997). Using geostatistical techniques to map the distribution of tree species for ground inventory data. In T. G. Gregoire, D. R. Brillinger, P. J. Diggle, E. Russek-Cohen, W. G. Warren & R. D. Wolfinger (Eds.), *Modeling longitudinal and spatially correlated data: methods, applications, and future directions, lecture notes in statistics (vol 122)* (pp. 187–198). Berlin Heidelberg New York: Springer.

Samra, J. S., Gill, H. S., & Bhatia, V. K. (1989). Spatial stochastic modeling of growth and forest resource evaluation. *Forest Science*, *35*(3), 663–676.

Speight, M. R., Hails, R. S., Gilbert, M., & Foggo, A. (1998). Horse chestnut scale (*Pulvinaria regalis*) (*Homoptera*: *Coccidae*) and urban host tree environment. *Ecology*, *79* (5), 1503–1513.

Upton, G. J. G., & Fingleton B. (1985). *Spatial data analysis by example, volume 1: Point pattern and quantitative data*. Chichester, UK: Wiley.

Yuan, L. L. (2004). Using spatial interpolation to estimate stressor levels in unsampled streams. *Environmental Monitoring Assessment*, *94*, 23–28.