A SPATIAL APPROACH TO ESTIMATING SOIL CARBON STOCKS
AT THE FIELD LEVEL

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAI'I IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

AGRONOMY AND SOIL SCIENCE

MAY 2007

by

Laura B. Delisle

Thesis Committee:

Russell Yost, Chairperson
Goro Uehara
Tomoaki Miura
We certify that we have read this thesis and that, in our opinion, it is satisfactory in scope and quality as a thesis for the degree of Master of Science in Agronomy and Soil Science.

THESIS COMMITTEE

[Signatures]

Russell Yost
Chairperson

[Signatures]
# TABLE OF CONTENTS

List of Tables ..................................................................................................................... v
List of Figures ................................................................................................................... vi

Introduction ........................................................................................................................ I

Background ...................................................................................................................... 1
Problem .............................................................................................................................. 2
Non-spatial Approach ........................................................................................................ 3
Spatial Approach: Geostatistics ...................................................................................... 4
Simulation ......................................................................................................................... 17
Bayesian Geostatistics ....................................................................................................... 20
The Data ............................................................................................................................. 26
Objectives .......................................................................................................................... 27
Hypothesis ........................................................................................................................ 27

Materials and Methods ..................................................................................................... 28

Dataset Descriptions ......................................................................................................... 28
Geographic Information ..................................................................................................... 38
Non-spatial Statistics ......................................................................................................... 39
Spatial Analysis ................................................................................................................ 40
Replication ........................................................................................................................ 42

Results ............................................................................................................................... 45

Descriptive Results .......................................................................................................... 45
Spatial Analysis ................................................................................................................ 56
Predicted Values

Simulated Values

Global Estimation

Discussion

Precision

Bayesian Analysis

Remaining Issues

Recommendations

Conclusions

Appendix: Example R Code

References
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Geographic Coordinate Locations of Farms</td>
<td>29</td>
</tr>
<tr>
<td>2. Area, Number of Observations, and Maximum Sampling Distances</td>
<td>37</td>
</tr>
<tr>
<td>3. Bulk Density</td>
<td>38</td>
</tr>
<tr>
<td>4. Trends</td>
<td>43</td>
</tr>
<tr>
<td>5. Transformations</td>
<td>44</td>
</tr>
<tr>
<td>6. Classical Statistics</td>
<td>55</td>
</tr>
<tr>
<td>7. Posterior Summaries for the Range Parameter</td>
<td>83</td>
</tr>
<tr>
<td>8. Posterior Summaries for the Partial Sill Parameter</td>
<td>84</td>
</tr>
<tr>
<td>9. Median Posterior Values for the Relative Nugget and Trend Parameters</td>
<td>85</td>
</tr>
<tr>
<td>10. Estimated Mean Carbon Concentration and 95 % Probability Interval Limits</td>
<td>121</td>
</tr>
<tr>
<td>11. Total Estimates of Carbon and 95 % Probability Interval Limits</td>
<td>124</td>
</tr>
<tr>
<td>12. Carbon per Hectare</td>
<td>125</td>
</tr>
<tr>
<td>13. Comparison of Spatial vs. Non-spatial Estimates</td>
<td>127</td>
</tr>
<tr>
<td>14. Spatial Relative to Non-spatial Precision</td>
<td>129</td>
</tr>
<tr>
<td>15. % Error</td>
<td>132</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Variogram Cloud and Binned Variogram</td>
<td>7</td>
</tr>
<tr>
<td>2.</td>
<td>Location of Villages</td>
<td>30</td>
</tr>
<tr>
<td>3 – 7</td>
<td>Field Perimeters and Sampling Configurations</td>
<td>31-35</td>
</tr>
<tr>
<td>8 – 13</td>
<td>Magnitude of Sampled Values</td>
<td>48-53</td>
</tr>
<tr>
<td>14 – 24</td>
<td>Empirical and Estimated Variograms</td>
<td>60-70</td>
</tr>
<tr>
<td>25 – 30</td>
<td>Histograms of the Posterior Range Parameter</td>
<td>73-78</td>
</tr>
<tr>
<td>31 – 41</td>
<td>Predicted Values</td>
<td>89-99</td>
</tr>
<tr>
<td>42 – 52</td>
<td>Simulated Values</td>
<td>102-112</td>
</tr>
<tr>
<td>53 – 57</td>
<td>Probability Density Plots</td>
<td>115-119</td>
</tr>
</tbody>
</table>
INTRODUCTION

Background

Carbon dioxide is one of several “greenhouse” gases implicated in global warming and climate change (Prentice et al., 2001). In 1997 the Kyoto Protocol established the basis for a global carbon market in which producers of CO₂ exceeding their quotas for greenhouse gas emissions can purchase “carbon credits” from entities that reduce emissions of CO₂ or sequester carbon from the atmosphere in sinks such as forests or the ocean (UNFCCC, 1998).

The largest land-based sink of carbon is the soil. Over the last several centuries much of the world’s grasslands have been plowed, depleting soils of organic matter and releasing carbon dioxide to the atmosphere. Changes in farm management that reverse this process and increase the amounts of stable organic matter in the soil sequester carbon, eventually reaching steady-state carbon levels. It is estimated that the world’s soils could provide a sink for carbon over the next twenty-five to fifty years, offsetting a quarter to a third of the annual increase in atmospheric CO₂ (Lal, 2004). Sequestration of carbon in soil may mitigate greenhouse gas emissions from other sources and allow time for alternative technologies to be developed.

Farmers in developing countries may be able to participate in the carbon trading market if they are able to show that they have sequestered carbon in their soils. There would be a double benefit in this as increasing the amount of organic matter in degraded soils would improve soil physical properties and availability of nutrients to plants, at the same time increasing plant productivity and food security, which is a serious concern in
many third-world countries. Carbon sequestration projects that specifically target
degraded dryland soils in developing countries are among the priorities of the Food and
Agriculture Organization of the United Nations (FAO) as they present one of the greatest
opportunities for improving food security for large numbers of people while simultane­
ously helping to mitigate global warming (FAO, 2004). Aspects of dryland soils that
make them particularly interesting for carbon sequestration projects include long
residence times of carbon in these soils due to the lack of moisture for decomposing
organisms, and generally low carbon saturation status due to degradation. Thus, there is a
potential to increase the carbon stored in these soils and at the same time, the sequestered
carbon can be expected not to degrade quickly.

Problem

Garcia-Oliva and Masera (2004) discuss the challenges that soil carbon sequestra-
tion projects face in attempting to participate in carbon mitigation projects under the
Kyoto Protocol. One of the challenges is how to measure carbon stocks and amounts of
carbon sequestered in soil accurately and precisely enough for the purposes of carbon
trading. Carbon typically represents only about 2% of the weight of the soil. In the Sahel
region of West Africa, where the farms in this study are located, it is particularly low,
often less than 1% of the weight of the soil (Pieri, 1992). Normal processes of carbon
change in soil are generally slow, thus amounts of carbon that have been sequestered over
periods of time that are of interest to carbon traders are quite small, and difficult to
measure precisely. Many samples are required to achieve an adequate precision but
sampling and chemical analysis are expensive.
The focus of this thesis is the spatial variability of carbon measurements and whether an explicitly spatial approach to sampling and data analysis can improve the precision or accuracy of estimates.

**Non-spatial Approach**

A straightforward approach to estimating the total amount of a specified soil attribute within a region is to base the calculation on the average of a representative set of measurements (or "observations") collected within the region. In the case of carbon, the total amount of carbon in the region can be calculated as the average concentration of carbon in the soil (by weight) times the total weight of the soil, which is in turn, the total volume of soil times the bulk density of the soil. The total amount of carbon stocks in tonnes (Mg) is calculated as:

\[
\text{Area (m}^2\text{)} \times \text{Depth (m)} \times \text{Bulk density (Mg m}^{-3}\text{)} \times \% \text{C} / 100 \quad [1]
\]

Various strategies exist for the placement of the locations where the observations will be collected, in other words for deciding what constitutes a representative sample. However, in the traditional, non-spatial approach, the individual observations are not usually referenced to their physical locations; the dataset consists of a set of measurements that are unlabeled as to the locations where they were collected within the region. A measure of the confidence with which the mean of the dataset can be claimed to reflect the true mean is provided by the variance of the dataset, which theoretically reflects the variability of the value of the attribute in the region. In the case of a simple random sample, a 95% confidence interval for the estimated mean would be the estimate plus or
minus approximately two times the standard error of the mean. The standard error is the square root of the variance divided by the square root of the number of observations.

An aspect of soil that makes it somewhat different than other types of phenomena to which statistical methods are applied, is its continuous nature. A region of soil is in fact not a population of individuals, but instead usually a single continuous entity. Because of this, agronomists and soil engineers often prefer to sample systemically (for example with sampling locations placed on a grid). Because the sampling locations are not chosen randomly (except for possibly the first grid node) the variance of the sample must be approximated (Webster and Oliver, 2001, p. 33). Sampling on a grid, or other strategies in which the observations represent equal subareas of the total area, are not the only possible sampling designs. Other strategies exist that provide both reasonable coverage and randomization at the same time. For example the region could be divided into a number of equal-sized subregions and at least two observations collected randomly within each of the subregions (Webster and Oliver, 2001, p. 32).

**Spatial Approach: Geostatistics**

A spatial approach to soil sampling and analysis takes into account the physical locations of the individual observations in a dataset and their configuration with respect to one another. Geostatistics is a specific collection of spatial methods that has become popular over the last twenty or so years in soil science and other fields such as geology, petroleum engineering, and hydrology. Two important applications in soil science are “precision agriculture”, in which spatially variable fertilizer requirements and yields are
calculated, and environmental assessment of soil contamination by pollutants or heavy metals.

Most commonly, geostatistical methods are used to interpolate values at unsampled locations, that is, to make maps or point predictions of the value of a soil attribute at locations where it is unknown. In this study, I have applied a geostatistical approach to making global estimates, that is estimates of the total amount of carbon within sampled regions.

The geostatistical approach explicitly models the continuous nature of soil. In general, we expect that pairs of observations that are collected in close proximity to one another should be more similar to one another than pairs separated by greater distances. In other words, observations that are close in proximity to one another are likely to be correlated, and this correlation should become weaker and weaker as the separation distance between observations increases, until it disappears altogether. In the geostatistical approach, a model of the spatial correlation between pairs of observations at increasing separation distances from one another is first constructed, based on the data. Then, depending on the purpose of the analysis, values at unsampled locations in the region are interpolated conditional to the known data values and the model of spatial correlation.

Spatial correlation is usually referred to in geostatistical jargon as "spatial dependence" to emphasize the concept that correlated observations are not "independent" of one another. Independence is an important assumption in many classical statistical methods. Spatial correlation (or spatial dependence) is usually modeled by means of a graph of the data called a "semivariogram," also called simply a "variogram" (Fig. 1).
The empirical variogram value, or "semivariance," is one half the squared difference in value between each pair of observations:

$$\gamma(u_i,u_j)=\frac{1}{2}[z(u_i)-z(u_j)]^2,$$

[2]

where $\gamma(u_i,u_j)$ is the semivariance for a pair of locations $u_i$ and $u_j$, and $z(u_i)$ and $z(u_j)$ are the values of the attribute at $u_i$ and $u_j$ respectively. The semivariance is plotted against the separation distance between the pairs.

Because a plot of the semivariance values for all paired combinations of the data results in a "cloud" of variogram points (Fig. 1, top graph), these are are usually grouped into separation distance classes called "lags" or "bins" and the resulting variogram is called a "binned" variogram (Fig. 1, bottom graph). The value of the semivariance for a particular lag is the average semivariance of the pairs in the lag, and the separation distance is their average separation distance.

A theoretical model, chosen from one of several "authorized models" (Webster and Oliver, 2001, p. 109-121) that guarantee positive-definiteness to the covariance matrices in the interpolation step, is fit to the empirical variogram. In Fig. 1 I have fit an exponential function (bottom graph, solid line). The basic exponential model is:

$$\gamma(h)=\sigma^2[1-\exp(-\frac{h}{\varphi})],$$

[3]

where $\gamma(h)$ is the theoretical semivariance for separation distance ($h$), and $\sigma^2$ and $\varphi$ are parameters of the function: $\sigma^2$ is the variance of the spatial component, and $\varphi$ is the range parameter. These are further explained below. A third parameter is required to completely describe the model in Fig. 1, the nugget variance, $\tau^2$, which determines the y-intercept.
Fig. 1. Variogram cloud and binned variogram. Top: variogram cloud; bottom: binned variogram for the same data, with 12 bins; line is exponential model fitted by weighted least squares (weights are number of pairs in each bin).
Under the assumption of stationarity, which implies among other things that the mean of the spatial process is expected to be invariant throughout the region, or in other words that the value of carbon is expected to fluctuate above and below an average value which is constant for the region, the semivariogram (standardized to a scale of 0 – 1) has the following relationship to the correlation function, $\rho(h)$:

$$\gamma(h)_{\text{standardized}} = 1 - \rho(h) \quad .$$

[4]

Semivariance increases with distance between pairs, and correlation decreases. We can also state the relationship in terms of the non-standardized semivariance and the non-standardized correlation, which is the covariance function $C(h)$:

$$\gamma(h) = \sigma^2 - C(h) \quad .$$

[5]

Traditionally, spatial correlation is modeled by means of the variogram, but interpolation algorithms are based on covariances; this is because matrices of covariance values are more stable than matrices of semivariance values.

The complete variogram model can be described by the range parameter ($\varphi$), the nugget ($\tau^2$), and the total sill ($\sigma^2 + \tau^2$). The range is the separation distance at which measurements are assumed to be no longer correlated. A practical range for the exponential model, $3\varphi$, is defined as the separation distance at which the function reaches 95% of its asymptotic sill; the semivariance value at which the function ceases increasing and flattens out. The nugget is the y-axis intercept, theoretically it is the variability expected for repeated measurements of the attribute at a single location in space; in practice it also includes spatial variation at separation distances less than the distance between the most closely-spaced pairs of observations. The total sill is the nugget variance plus the
variance of the spatial process (the partial sill, \( \sigma^2 \)). The term "nugget" originated when geostatistical methods were developed for the gold mining industry.

Partitioned parts of the total sill are often referred to as "variances" as they describe the contributions of the nugget and spatial processes (of which there may be more than one, with different ranges of spatial correlation) to the total variance. The term "semivariance" refers to the fact that the variance between pairs of points has been divided in half. This turns out to be convenient because the model now corresponds to a breakdown of the variances on an individual point basis. Under the stationarity assumption, as the size of the region approaches infinity in size the total sill is equal to the total variance of values within the region (the "dispersion variance"). The dispersion variance in a finite region, especially one with dimensions smaller than the range of spatial correlation, will usually be smaller than the theoretical variance (the total sill).

To give an example, in the variogram in Fig. 1 (bottom graph), the model parameters for the fitted theoretical model are: nugget variance (\( \tau^2 \)) = 0.0342, partial sill (\( \sigma^2 \)) = 0.0308, and range parameter (\( \varphi \)) = 100 meters (which corresponds to a practical range of 300 meters). The total sill, \( \tau^2 + \sigma^2 \), is 0.0342 + 0.0308 = 0.065. The nugget variance can be seen to correspond to the value of the y-intercept in the graph; the spatial component (partial sill) increases from that value to the total sill of 0.065, which is reached off the right side of the graph. The value of the function at a distance of 300 meters (practical range) is approximately 95% of the total sill (0.95 * 0.065 = 0.062). The dataset variance is 0.055, smaller than the theoretical variance of 0.65.

Variogram modeling requires a number of decisions on the part of the modeler. For example, when the mid-points of the distance class intervals are chosen differently,
the empirical variogram plot has an altered appearance (not shown). This means that the choice of lag size affects the variogram shape to which the model will be fit. Also, because the variogram values are considered unreliable at much more than half of the longest separation distance (because of lack of replication of pairs separated by the largest distances), the variogram is often not modeled beyond that distance. The maximum distance to which the analysis is limited is a decision of the modeler. The model may be fitted by one of several methods, including fitting by eye, least squares, or weighted least squares. In weighted least squares, the weights are commonly determined by the number of pairs in each lag distance class, as this varies. The model in Fig. 1 (bottom graph, solid line) was fitted by weighted least squares, with the weights being the number of pairs in each lag.

In general, as the size of the dataset increases the empirical variogram becomes more reliable and the effect of modeler decisions on the resulting model becomes less crucial. The ideal dataset size for optimal variogram modeling has not been defined, because it would vary depending on the specifics of the region being sampled. Webster and Oliver (2001, p. 90) suggest that variogram modeling is unreliable at dataset sizes of less than 100. The variogram in Fig. 1 was plotted from a dataset which contains 21 measurements; too small, according to the above.

An alternative method of estimating the parameters (nugget, partial sill, and range) for the variogram is provided by maximum likelihood. In this approach, the variogram parameters may be estimated from all of the data rather than by fitting a model to a graphical display that depends on choices of lag and maximum separation distance, as described above. Given that \( y \) represents the value of the attribute, and \( \theta = (\tau^2, \sigma^2, \varphi) \), the
log likelihood of the joint probability distribution \( f(y, \theta) \) is maximized for values of \( \theta \) when \( y \) is held constant at the observed data values. Maximum likelihood is an optimal estimator, at least for large data sets (Diggle and Ribeiro, 2002), therefore some caution is advised for small datasets. Also, only the optimal estimates of the parameters are derived, no account is taken of the space of uncertainty, that is no account is taken of the probability that the parameters may take other values.

In the interpolation step, values at unsampled locations are predicted as a function of the observed data values, weighted depending on the configuration of the data values with each other and with the point to be interpolated. The various interpolation methods, of which there are several, are called kriging, after D.G. Krige, a mining engineer who pioneered the approach in the gold-mining industry.

According to the theoretical model that underlies the geostatistical approach, the value of an attribute at any particular location in the region is considered to be a random variable and the distribution of values that this variable might take is assumed normal. Its value \( Z(u) \) at a location \( u \) can be predicted as the sum of an underlying stationary mean, \( \mu \), plus a random error, \( \epsilon(u) \):

\[
Z(u) = \mu(u) + \epsilon(u) \quad [6]
\]

The errors are spatially correlated. When extended to include all points in the region and the covariances between them, the model in equation [6] becomes the "random function model" (to earth scientists and engineers) or a "spatial stochastic process" (to statisticians).

The "mean" component may in fact not be stationary. Instead it might be better described by a trend (usually on the x and y coordinates). As a practical matter, trends are
often subtracted from the data and the analysis is performed on the residuals (the trend is
added back to the final predicted values). Whether or not to remove a trend is a decision
of the modeler as it is not possible in reality to prove whether the mean is stationary or
not (Goovaerts, 1997, p. 71). A first order trend on the coordinates has been removed
from the data before plotting the graphs and fitting a model to the data in Fig. 1.

For the two-dimensional case, the spatial component can be visualized as
smoothly varying local peaks and valleys in the value of carbon, fluctuating above and
below a plane representing the stationary mean (in which case the plane is level) or trend
on the coordinates (in which case the plane is tipped).

The basic kriging algorithm is called simple kriging and the simple kriging
estimator is:

\[
\hat{Z}_{sk}(u_0) = \sum_{i=1}^{N} \lambda_i z(u_i) + [1 - \sum_{i=1}^{N} \lambda_i] \mu.
\]  

\(Z_{sk}(u_0)\) is the value of the attribute, \(Z\), at a specific prediction location \(u_0\), predicted by
simple kriging. The \(\lambda_i\) are the simple kriging weights for the data values, \(z\), at the
locations \(u_i\), and \(\mu\) is the known, stationary mean. The predicted value is thus a weighted
average: it is the sum of the weighted data values plus the weighted mean. The weight of
the mean is one minus the sum of the weights of the data values. In other words, the
weights sum to one, assuring unbiasedness, and the weight of the mean increases where
the weights of the data values decrease and decreases where the weights of the data
values increase. Because the weights for the data values increase with nearness to the
prediction location, the mean has little influence on predicting values at locations which
are near to data values and increases in influence at locations far from data values. The weights are determined by solving the system of equations for $\lambda_j$:

$$\sum_{j=1}^{N} \lambda_j C(u_i-u_j) = C(u_i-u_0) \quad \text{for } i = 1 \ldots N . \quad [8]$$

The left-hand side of equation [8] is the sum of the weighted covariances between the data locations $u_i$ and $u_j$. The right-hand side is the covariance between the location at $u_i$ and the prediction location $u_0$. (the expression $u_i - u_j$ is a distance and therefore the covariances are $C(h)$, the covariance for the distance $h$ between $u_i$ and $u_j$). The solved for values of the $\lambda_j$ are unique at each prediction location.

The covariances, $C(h)$, are derived from the theoretical variogram model that was fit to the empirical variogram. As can be seen in the system of equations [8], when the system is solved for the kriging weights, these depend partially on the covariances between the observed data values. The weights increase with nearness of data values to the point to be predicted but also decrease with increasing correlation between observed data values. In this way the influence of clustered data values (which can be seen to supply redundant information) is decreased.

A prediction error variance can be calculated for each of the prediction locations:

$$\sigma^2_{sk}(u_0) = C(0) - \sum_{i=1}^{N} \lambda_i C(u_i-u_0) , \quad [9]$$

where $\sigma^2_{sk}(u_0)$ is the simple kriging error variance for prediction location $u_0$. $C(0)$ is the covariance for $h = 0$, and is equal to the theoretical variance, the total sill. The kriging standard error is simply the square root of the error variance.
Simple kriging is the theoretical basis for the geostatistical approach but other forms of kriging are more common, for example ordinary kriging is a kriging variant which does not assume a known stationary mean but instead implicitly estimates it from the data. The ordinary kriging estimator is:

\[ \hat{Z}_{ok}(u_o) = \sum_{i=1}^{N} \lambda_i z(u_i) \]  \hspace{1cm} [10]

Unlike the simple kriging estimator, equation [10] no longer contains the stationary mean, \( \mu \). A separate constraint must be added to force the weights to sum to one and assure unbiasedness. In order to solve the system of equations with the additional constraint that the sum of the weights be equal to one, a Lagrange multiplier, \( \psi \), is introduced; thus for ordinary kriging, the system to be solved becomes:

\[ \sum_{j=1}^{N} \lambda_j C(u_i-u_j) + \psi = C(u_i-u_0) \]  \hspace{1cm} [11]
\[ \sum_{j=1}^{N} \lambda_j = 1 \hspace{1cm} \text{for} \ i=1 \ldots N \]

The system solves for the Lagrange multiplier along with the weights, \( \lambda_j \), therefore the value of \( \psi \) as well as the values of the \( \lambda_j \) are unique at each prediction location. The ordinary kriging error variance is identical to the simple kriging error variance given in equation [9], except that an additional term, \( -\psi \), is added to the right-hand side:

\[ \sigma_{ok}^2(u_0) = C(0) - \sum_{i=1}^{N} \lambda_i C(u_i-u_0) - \psi \]  \hspace{1cm} [12]

Kriging provides not merely a single predicted value at each location, but rather a distribution, assumed to be normal, of possible values and their associated probabilities. The predicted value is the mean, or expected value, of this distribution and the spread of
the distribution is described by the kriging variance. Since predicted values are means, kriging is a *smoothing* interpolator. Extreme values are smoothed toward the mean.

The prediction error variance increases with the distance between the prediction point and the observed data values, as we might intuitively guess; we have greater certainty about predictions for locations that are near to actual sampled values. In equation [9] we see that the simple kriging error variance is the theoretical variance, \( C(0) \), the total sill, minus the sum of the (weighted) covariances between each of the data values and the prediction point. The ordinary kriging error variance, equation [12], is found in the same way, except that the unique value of the Lagrange multiplier is also subtracted. The covariances decrease with distance. When all data values are beyond the range of spatial correlation from the point being estimated, \( C(h) = 0 \) and the prediction error variance is simply the theoretical variance.

Calculating a prediction error for a global estimate is problematic. We can not simply sum (or average, in the case of a global average) the error variances for the individual point predictions, as we might intuitively suppose, because values at neighboring prediction locations are correlated. The covariances between all of the point estimates would have to be taken into consideration, resulting in a very large calculation.

To save on computational effort, geostatisticians have traditionally preferred to approximate the covariances in various ways. Webster and Oliver (2001, p.178) suggest estimating the global mean as an average of many smaller interpolated "block" values and approximating the global variance as the average of the block error variances.

A prediction error can indeed be calculated for "block" values by means of a form of kriging called "block kriging", which predicts values for areas or volumes rather than
for points. The average of point predictions made on a grid within a block is essentially
equal to the block kriging prediction for the same block; therefore the main advantage of
block kriging is that the error variance for blocks tends to decrease as block size
increases.

Theoretically, the entire region could be considered as a single block. We could
use block kriging to predict a single average value for the region, and a single error
variance. Calculating the prediction error of such a global prediction, however, requires
knowing all of the covariances between the observations and a grid of points chosen to
discretize the region, as well as the covariances among all of the discretized grid points
themselves. These covariances depend on the model of spatial correlation developed from
the empirical variogram, which has several weaknesses in respect to modeling the spatial
covariance of the entire region. Variogram values for pairs of observations at large
distances from one another are considered unreliable, both because of the probable lack
of stationarity across the region, and the few replications of pairs separated by those
distances (Goovaerts, 1997, p. 157).

Another problem with considering the entire region as a single block is that if the
region to be estimated is not rectangular, then the block must be specified as irregular in
size, or else the error variances will include large portions of the unsampled region that
lies outside the boundaries of the irregularly-shaped sampled region, the result of which
will be an increase in uncertainty. Some geostatistical software packages will not
perform block kriging on irregular size blocks.

In summary, the main problems with the conventional geostatistical approach for
making global estimates of carbon include the lack of a way to evaluate uncertainty in the
variogram model parameters, and lack of a method of calculating a global error variance, except by approximation. Conventional geostatistical methods are also undesirable due to the uneven smoothing of kriging predictions and the lack of dependence of kriging variances on nearby observed data values. Predictions made in near proximity to data values tend to be less smoothed, and those at large distances from data values tend to be more smoothed. Kriging error variances depend only on the distance of the prediction point from observed data values and not on the actual value of those observations. For any two particular prediction locations separated by the same distance from data values, the error variance will be the same, regardless of the actual values of their neighboring data.

Regardless of the problem of assessing the global error, the final total carbon estimate for the region can be calculated as in equation [1], except that in the spatial case, the value substituted for carbon concentration in the equation is the average of the interpolated values in the prediction grid (within the region's boundaries), rather than the simple average of the observed data values as it is for the non-spatial estimate.

The above is a very incomplete description of geostatistical methods. In addition to the references already mentioned above, see Isaaks and Srivastava (1989).

Simulation

Simulation methods extend traditional geostatistical interpolation methods by simulating outcomes that include the error term in equation [6]: \( Z(u) = \mu(u) + \epsilon(u) \). Unlike in kriging, the outcome at any given prediction location is not the kriging predicted value but is instead a value drawn from the distribution described by the kriging
mean (the predicted value) and variance. To recreate the full spread of the distribution, it is necessary to draw repeatedly from its probability distribution many times. A single drawing at each location for all nodes of the prediction grid is called a “realization”. The joint distribution of possible values at all locations is simulated by creating multiple realizations. The entire set of simulated values for all locations for all realizations is called an “ensemble”.

Simulation is appropriate where the global, rather than local, characteristics are of interest (Deutsch and Journel, 1998, p.120). Kriging, in contrast, provides “best” (best in the sense that local error variance is minimized) local estimates. Taken altogether, kriging estimates may not produce the ideal map; large values tend to be underestimated and small values overestimated. Also, kriged values tend to be more “smoothed” at locations distant from observed data values and less “smoothed” at locations with nearby data, as mentioned above.

The primary benefit of simulation for our purposes, however, is that it provides a way to evaluate the global uncertainty. Each realization provides an estimate of the global mean (found simply by averaging all of the simulated values in the grid) or any other joint statistic, and the distribution of these values over multiple realizations (all of which are equally probable), provides a measure of the expected variability. For example, the interval between the 0.025 and 0.975 quantiles of the distribution of predicted global means describes a 95 % probability interval for the estimate.

It is necessary to decide among several simulation methods (see Goovaerts, 1997, chapter 8; and Deutsch and Journel, 1998, chapter V); for continuous data, the most appropriate method is sequential Gaussian simulation (Deutsch and Journel, 1998, p. 18).
119), as long as the data are reasonably normally distributed, or transformed to approximate normality.

Multi-Gaussian simulation would appear to be an appropriate approach for estimating global carbon and putting a confidence interval on those estimates. In practice, there are several limitations to the simulation approach.

Sequential Gaussian simulation simulates the multivariate distribution of values for locations in the region. The sequential part of the algorithm is a work-around to simultaneous multivariate prediction at all nodes, instead the prediction follows a random path, utilizing previously predicted nodes as conditioning data, along with the original data, for subsequent prediction at later nodes in the path. Theory requires that distributions at prediction locations are described by the simple kriging means (predicted values) and variances. The simple kriging algorithm requires that the underlying stationary mean is specified, however, in our case at least, it is unknown. It is possible to substitute ordinary kriging for simple kriging, however in this case the error variances will be incorrect. Several work-arounds to this problem are proposed, for example Goovaerts (1997) suggests that the error variance predicted by ordinary kriging can be corrected by multiplying by a factor, determined by trial and error.

A covariance model is still required for input into the simulation algorithm. Especially for sparse datasets, modeling the variogram remains problematic. In particular, the covariance model is unreliable for the largest separation distances. In practice, conditioning data is limited to restricted neighborhoods, even though the multivariate distribution of values separated by all distances, short and long, is theoretically required. The user must make a decision on the size of the neighborhoods.
Bayesian Geostatistics

In the Bayesian approach, the variogram and trend parameters may be modeled as random variables. In contrast, in the conventional geostatistical approach, the variogram parameters are modeled from the data but are then "plugged into" the kriging algorithm as if they were known constants. The final estimation variances for the conventional approach do not include any uncertainty in modeling the variogram parameters. In the Bayesian approach, uncertainty in the variogram parameters is included in the final estimation variances.

Also, in the Bayesian approach the variogram is modeled from all of the data rather than by fitting to a graphical display of the empirical variogram values that depends on choices of lag size and a maximum separation distance, as is the typical practice for the conventional geostatistical approach. One non-Bayesian method that does allow parameter estimation based on all of the data is the maximum likelihood method mentioned above. The Bayesian approach to estimation of the model parameters is related to the maximum likelihood method as it involves the likelihood function. The difference between the two methods is that in the Bayesian approach, the entire likelihood surface for possible values of the parameters is derived. In maximum likelihood only the optimal values are derived and thus no account can be taken of the uncertainty in estimating them.

Bayesian outcomes are equivalent to conventional kriging outcomes when the variogram parameters are modeled as fixed, known constants instead of random variables. When the trend as well as the variogram parameters are modeled as fixed constants, Bayesian outcomes are equivalent to those obtained by simple kriging. When
the variogram parameters are modeled as fixed but the trend parameters are modeled as random, Bayesian outcomes are equivalent to either ordinary kriging or “kriging with a trend” outcomes, depending on the number of trend parameters modeled (a single unknown constant in the case of ordinary kriging, or intercept plus x and y coordinate coefficients in the case of “kriging with a trend”). However these outcomes are for ordinary kriging and “kriging with a trend” outcomes in which the neighborhood of influence of the data has not been restricted.

The underlying theoretical model is similar for the conventional and Bayesian approaches except that in the Bayesian approach, the Gaussian assumption is made explicit, an assumption that Diggle and Ribeiro (2002, p. 133) maintain is implicit in standard kriging practices within the conventional approach. In the Bayesian approach, the underlying theoretical model is:

\[ Y(u) = X(u) \beta + S(u) + \epsilon(u) . \]  

[13]

The notation in equation [13] primarily follows Ribeiro and Diggle (1999). \(Y(u)\) is the value of the attribute at location \(u\), corresponding to \(Z(u)\) in equation [6]. \(X(u)\beta\) is the “mean part of the model”, that is, it is the expectation of \(Y(u)\); \(X(u)\) is a vector of trend covariates, for example, the x and y coordinates; \(\beta\) is a vector of trend parameters: \((\beta_0, \beta_1, \beta_2)\) which are the intercept, x coefficient, and y coefficient of the trend respectively; \(S(u)\) is the value at a location \(u\) of the stationary spatial process; and the \(\epsilon(u)\) are the uncorrelated errors.

Equation [13] can be compared with the conventional geostatistical model in equation [6], \(Z(u) = \mu(u) + \epsilon(u)\). In Ribeiro and Diggle’s formulation, the \(\epsilon(u)\) term has been expanded to \(S(u) + \epsilon(u)\). \(S(u)\) represents the spatially correlated component of
the variation around $\mu$, whereas $\epsilon(u)$ represents the uncorrelated “noise”. Also, $X(u)\beta$ replaces $\mu(u)$. Where $\beta_1$ and $\beta_2$ are zero, the “trend” reduces to intercept only, that is, it reduces to stationary mean. Therefore the equation in [13] describes both the case where there is a trend and the case where there is no trend.

The Bayesian analysis proceeds in three stages. First, prior probability distributions for the variogram and trend parameters are posited. This amounts to a statement of what is known prior to collection of the data about the probable values the parameters might take and may simply be a statement of prior ignorance. The priors are then “updated” by conditioning to the data, yielding “posterior” parameter distributions. Finally, the values of the attribute at locations where it is unknown are predicted; or rather their predictive distributions are created.

The prior distributions for the parameters may be posited as flat or uniform, indicating prior ignorance, or they may incorporate previously known information. Updating the priors to the data requires the likelihood function, $L(\theta)$ and the definition:

$$L(\theta|y) \equiv pr(y|\theta),$$

where $\theta$ is the set of parameters $(\beta, \sigma^2, \tau^2, \phi)$, which were each defined previously, and $y$ is the observed data. Vertical lines indicate a conditional relationship and $pr(\theta)$ indicates a probability distribution.

Also necessary is the relation:

$$pr(\theta|y) \propto pr(\theta) pr(Y|\theta).$$

In words, the posterior distribution for $\theta$, conditional (i.e., updated) to the data $y$, is proportional to the prior distribution for $\theta$ times the distribution for $Y$ given $\theta$. A prior distribution for $\theta$ can be posited as described above, and the distribution for $Y$ conditional
to $\theta$ can be derived by means of the relationship in equation [14]. The likelihood of $\theta \mid y$ is substituted for $pr(y \mid \theta)$ in equation [15]. Expanded, equation [15] becomes:

$$
pr(\beta, \sigma^2, \tau^2, \phi \mid y) \propto pr(\beta, \sigma^2, \tau^2, \phi \mid \tau^2 I + V_y(\sigma^2, \phi))^{-1/2} \exp\left\{ -\frac{1}{2} (y - X \beta)' \left( \tau^2 I + V_y(\sigma^2, \phi) \right)^{-1} (y - X \beta) \right\}
$$

[16]

The first part of the right-hand side of equation [16] is the prior for $\theta$, expanded to $(\sigma^2, \tau^2, \beta, \phi)$, and the second part of the right-hand side is the likelihood of $\theta$ given the data, $y$. $I$ is the identity matrix, and the $V_y$ are the covariance matrices defined by the partial sill $(\sigma^2)$ and range $(\phi)$ parameters.

In words, the joint posterior distribution of the model parameters is found by multiplying their (posited) joint prior distribution times their likelihood given the data.

When all of the parameters are modeled as random the posterior and predictive distributions can not be derived analytically. Monte Carlo methods are used to sample these distributions, or in other words, to simulate them. The “sampling” proceeds in a number of steps:

1. A sensible range of discrete values (support points) for $\phi$ and $\nu^2$ (where $\nu^2$ is defined as $\tau^2/\sigma^2$, the relative nugget) are chosen, for example the range parameter might plausibly fall between 0 and twice the maximum sampling distance between data locations, in steps of 20 meters: 0, 20, 40, 60 meters, etc. The larger the set of support points chosen, the larger the computation requirement will be.

2. The discrete marginal and joint prior distributions of $\phi$ and $\nu^2$ are updated by conditioning to the observed data values, as in equation [16], above.

3. A value is sampled from the posterior distribution: $pr(\phi, \nu^2 \mid y)$
4. The value sampled in (3) is attached to the posterior distribution: \( pr(\sigma^2, \beta \mid \varphi, v^2, y) \).

5. Steps 3 and 4 are repeated many times, resulting in the joint posterior distribution:

\( pr(\sigma^2, \beta, \varphi, v^2) \).

Similarly, the predictive distributions, \( Y(u_0) \) (or more simply, \( Y_0 \)) are simulated by sampling. The predictive distribution is given by:

\[
pr(y_0|y) = \int pr(y_0|y, \theta) pr(\theta|y) \, d\theta,
\]

in which the first probability distribution after the integral is basically the same as conventional kriging (fixed \( \theta \)) and the second is the joint posterior distribution: \( pr(\sigma^2, \beta, \varphi, v^2) \) from step 5, above.

The predictive distributions for \( Y(u_0) \) are essentially an average of conventional kriging outcomes for particular values of \( \sigma^2, \beta, \varphi, \) and \( v^2 \) weighted according to their joint posterior distribution.

The algorithm produces many simulated surfaces (in the two-dimensional case) of \( Y(u_0) \), all consistent with the data and equally probable, as in sequential Gaussian simulation. Taken together, these present a visual assessment of the spatial uncertainty, as the individual simulations may vary substantially. Unlike in sequential Gaussian simulation, the uncertainty in the model parameters is included in the simulations. For different realizations of the same data, trends may appear to be strong or weak, patches of continuous values of the attribute may be large (long range) or small (short range), and overall, the surface may be smooth (small nugget) or noisy (large nugget).

The prediction distribution for any specified location (i.e., for any specific prediction node) is given by the set of values from the different realizations. The mean of
this set is the expected value, but it is also possible to compute any desired quantile or other summary of the predictive distribution. In addition, the spread of the distribution of realized values provides a measure of the uncertainty for the predicted value. Unlike in conventional kriging, these uncertainties are conditional to the observed data values, and thus reflect the influence of nearby data.

A predicted value for any combination of prediction locations can be calculated from the point predictions, including block values and associated uncertainties, and predictions for any regular or irregular subregion. Finally, the global average and uncertainty can be derived (as described above for non-Bayesian simulation); in effect these are the "block" estimate and uncertainty for the region considered as a single block. Block kriging and point kriging results may thus be derived from a single output of the algorithm. No approximations of the covariances among the point estimates are required.

The Bayesian approach was applied to the estimation of carbon for the datasets included in this study within a specific statistical software environment called "R" (R Development Core Team, 2006) and a particular "package" within R, called geoR (Ribeiro and Diggle, 2001). The Bayesian approach to geostatistics as implemented in geoR is described in Ribeiro and Diggle (1999) and Diggle and Ribeiro (2002). The discussion above is based on the particular application of Bayesian methods to geostatistics described in these papers.
The Data

The data are existing datasets that were collected as part of a project\(^1\) which had as one of its goals to increase soil fertility and food productivity in the degraded soils of small farms in several countries of West Africa. Another of the goals of the project was to increase soil carbon stocks and to develop a protocol to measure, predict, and monitor soil organic carbon over time and space. For the farms which are included in this analysis, a change in soil management was implemented: a system of permanent and impermanent soil ridges (sometimes referred to as “ridge-tillage”) was installed along the contours of the fields in order to reduce rainfall runoff and erosion and increase soil moisture retention. The impermanent ridges are installed between the permanent ridges and are reconstructed yearly (M.A. Doumbia, personal communication, 2007).

I have included here an analysis of soil carbon data for five farms in Mali, each of which is located in a different village in that country. The data were collected from two soil depths and on several sampling dates, separated by two or four year intervals. Some of the datasets are quite small, including as few as 13 carbon measurements. The largest includes 75 measurements. The values are % soil organic carbon (by weight).

\(^1\) The Carbon from Communities Project, a joint project of the Sustainable Agriculture and Natural Resource Management and the Soils Management (SM) sub-programs of the Collaborative Research Support Program (CRSP). Funding for the Carbon from Communities Project is from NASA (National Aeronautic and Space Administration; the USDA; Virginia Polytechnic Institute and State University; the University of Georgia; and SM-CRSP, University of Hawai‘i.)
The data were collected with a geostatistical analysis in mind. Geographic coordinates were recorded for the soil sampling locations as well as for the perimeters of the regions within the farms for which estimates of total soil organic carbon were desired.

To calculate the total amount of carbon for a specified region as in equation [1], we need to know the soil bulk density. Bulk density values are available for four of the farms, but no information is available on the variability of these measurements.

Objectives

The objectives of the study are to:

1) Apply the Bayesian geostatistical approach to estimate the global carbon concentration for each the datasets, and calculate total carbon stocks where bulk density values are available.

2) Provide a measure of uncertainty for each of the above estimates.

3) Compare spatial estimates and uncertainty intervals with non-spatial estimates and intervals for the same datasets.

Hypothesis

Where spatial correlation and trends exist, including these in a model to estimate soil carbon should increase the precision of estimates. Precisions will also vary depending on factors such as non-correlated variability, the strength of the spatial correlation, strength of trends, number of observations, and sampling configurations.
MATERIALS AND METHODS

Dataset Descriptions

Location of Villages

The datasets were collected in five farms located in five villages in southern Mali, in West Africa. The locations of the villages in Mali can be seen in Fig. 2. Each farm is referred to in this paper by the village in which it is located: Oumarbougou, Fansirakouro, Sougoumba, Sikasso, and Siguidolo. Geographic coordinate locations for each of the farms are given in Table 1. The farthest distance between any two of the farms included in the study is 332 kilometers: this is the distance from either Fansirakouro to Sougoumba or Fansirakouro to Sikasso. The farms in Oumarbougou and Sougoumba are nearest in proximity, at a distance of about 2 kilometers from one another.

Climate, Elevation, Vegetation, and Soil

Available information includes the following: Rainfall decreases from west to east and from north to south, with Fansirakouro being the wettest among the villages; the annual rainfall in Oumarbougou and Sougoumba is approximately 1000 mm and in Siguidolo is 800 mm. The elevation is approximately 400 meters in Sougoumba. All of the soils are Paleustalfs; the soils in Fansirakouro and Siguidolo are Plinthic Paleustalfs. Common crops include cotton (*Gossypium hirsutum*, L.), sorghum (*Sorghum bicolor* (L.) Moench), millet (*Pennisetum glaucum*, L.), and maize (*Zea mays*, L.); these are usually rotated among the several fields in a farm. In the farm at Sikasso, which has the most
extensive system of ridges, the first permanent ridges were installed in 1995 and more ridges have been added, through 1999.

Table 1. Geographic Coordinate Locations of Farms.

<table>
<thead>
<tr>
<th>Farm</th>
<th>Longitude</th>
<th>Latitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oumarbougou</td>
<td>5° 7' 57.33&quot; W</td>
<td>12° 11' 45.87&quot; N</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td>8° 5' 15.83&quot; W</td>
<td>12° 57' 22.82&quot; N</td>
</tr>
<tr>
<td>Sougoumba</td>
<td>5° 8' 53.71&quot; W</td>
<td>12° 11' 10.15&quot; N</td>
</tr>
<tr>
<td>Sikasso</td>
<td>5° 35' 7.08&quot; W</td>
<td>11° 14' 3.17&quot; N</td>
</tr>
<tr>
<td>Siguidolo</td>
<td>6° 47' 37.93&quot; W</td>
<td>12° 54' 54.92&quot; N</td>
</tr>
</tbody>
</table>

Field Perimeters

Maps in Fig. 3-7 show the boundaries of the regions for which carbon estimates will be made, and sampling configurations. Location of the permanent ridges are also shown in the maps. For Oumarbougou, the estimation region sampled in 2002 was expanded in 2004 to include an equally-sized control region that was not under the ridge-tillage system (Fig. 3). For this farm two estimation regions were defined: one for the original region sampled in 2002, and one for the expanded region sampled in 2004. A separate estimate for the 2004 dataset was made within the 2002 subregion. For the other farms, one estimation region was defined for all sampling dates.
Fig. 2. Location of villages in Mali.
Fig. 3. Field perimeters and sampling configurations, Oumarbougou: a) 2002 sampling locations; b) 2004 sampling locations. Dotted lines are locations of permanent ridges. The hashed line in the left figure is an arbitrary boundary defining the region that will be estimated for 2002 data and for 2004 data for the purpose of comparison.

Geographic Area, Number of Observations, and Distances between Sampling Locations

The areas of the regions range from 2 hectares (Oumarbougou) to greater than 33 hectares (Sikasso). The number of observations (n) included in any one dataset ranges from 13 to 75. This information is given in Table 2, along with the maximum distances between any two sampling locations in each sampling configuration.

Sampling Configurations

These differ between sampling dates, although all were collected on a preferential sampling scheme, that is, neither randomly nor systematically. The 2000 sampling
Fig. 4. Field perimeter and sampling configurations, Fansirakouro: a) 2002 sampling locations; b) 2004 sampling locations. Dotted lines are locations of permanent ridges.

configurations have few sampling locations and large unsampled subregions (the purpose that year was primarily exploratory). The Sougouumba 2000 datasets (Fig. 5a) were collected on a “field central” strategy, avoiding sampling near either the outer perimeters of the farm or the boundaries between individual fields within the farm (the internal field boundaries are not shown in the map). The Siguidolo 2000 datasets (Fig. 7a) have a small ring of sampling locations in the southwest corner that seem to have been chosen to describe some particular feature.

The overall 2002 strategy was to increase the number of sampling locations and to include sufficient closely-spaced samples to define the variogram well at small separation distances. As a result, sampling configurations for that year tend to have
subregions that are either clustered or sparse (see Fig. 4a and 5b). The 2004 sampling configurations are better dispersed but have fewer closely-spaced pairs. In 2004 an emphasis was also placed on including sampling locations at varying distances from trees (the location of trees are not shown in the maps).

No boundaries were defined before sampling other than farm or field perimeters. I have defined several arbitrary boundaries to avoid extreme extrapolations to unsampled subregions. Some very sparsely sampled subregions still remain. In the case of Fansirakouro 2002, sampling locations do not extend into either the southwest “toe” or into the northwest or northeast subregions of that farm (Fig. 4a). Several of the 2000
Fig. 6. Field perimeter and sampling configurations, Sikasso: a) 2000 sampling locations; b) 2004 sampling locations. Dotted lines are locations of permanent ridges.

datasets have subregions that are poorly represented by sampling (see Fig. 6a, Sikasso 2000; and 7a, Siguidolo 2000).

Sampling Dates and Methods

Soil samples were collected in December 2000 (Oumarbougou, Sikasso, and Siguidolo), March 2002 (Oumarbougou, Fansirakouro, and Sougoumba), and in March 2004 (all five farms). Only the farm in Sougoumba was sampled at all three sampling dates. Samples were taken from two soil layers: the 0- to 20-cm layer cm and 20- to 40-
cm layer, at the same sampling locations. The amount of soil collected from each soil layer at each sampling location was approximately 500 g.

Carbon Analysis

Soil samples were analyzed for total % carbon content (per mass of soil) by combustion, with a LECO Analyzer. As the soils generally have pH values below 6, they were assumed to contain negligible amounts of carbonate C. The values are thus assumed to represent organic carbon (referred to in this paper simply as carbon, or abbreviated as C). The analysis was done by the staff of the Agricultural Diagnostic Service Center (ADSC), College of Tropical Agriculture and Human Resources,
University of Hawai‘i at Mānoa, Honolulu, Hawai‘i. Values were reported as “% C”. Soil attribute values are often reported in units such as mg kg⁻¹, which clarifies the meaning by specifically referring to the weight of the soil. I have used the % C values here but specified “by weight of soil” (as opposed to volume).

Measurement Error

Measurement error among our samples was assessed with 20 soil samples chosen at random from 493 samples that were collected at 7 farms in March 2004 (some of which are not included in this study). Two replications of each of the 20 samples were reanalyzed for % C at the ADSC in April 2005. The average standard deviation for the replicated samples was 0.01 and the average variance was 0.0004 (after removal of one outlier), agreeing with the example standard deviation, 0.01, provided by the LECO company for carbon values of similar magnitude (around 0.5 %). The measurement error variance is of interest when modeling the components of spatial variance, because it provides an estimate of the theoretical minimum nugget variance.

Bulk Density

Values of bulk density (apparent bulk density, dry soil) for each farm are given in Table 3. The values were obtained in February 2006 from one or two locations in each farm. No value is available for the farm in Sougoumba.
Table 2. Area, Number of Observations, and Maximum Sampling Distances.

<table>
<thead>
<tr>
<th>Farm and year</th>
<th>Area</th>
<th>Number of observations</th>
<th>Maximum distance†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m²</td>
<td>0-20 cm</td>
<td>20-40 cm</td>
</tr>
<tr>
<td>Oumarbougou</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>19,648</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>2004</td>
<td>42,216</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>47,666</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>2004</td>
<td>&quot;</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Sougoumba</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>188,526</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>2002</td>
<td>&quot;</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>2004</td>
<td>&quot;</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>Sikasso</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>331,957</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>2004</td>
<td>&quot;</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>Siguidolo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>156,963</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>2004</td>
<td>&quot;</td>
<td>35</td>
<td>35</td>
</tr>
</tbody>
</table>

† Maximum distance between sampling locations.
Table 3. Bulk Density.

<table>
<thead>
<tr>
<th>Farm and soil layer</th>
<th>Bulk density Mg m⁻³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oumarbougou</td>
<td></td>
</tr>
<tr>
<td>0-20 cm</td>
<td>1.69</td>
</tr>
<tr>
<td>20-40 cm</td>
<td>1.63</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td></td>
</tr>
<tr>
<td>0-20 cm</td>
<td>1.53</td>
</tr>
<tr>
<td>20-40 cm</td>
<td>1.66</td>
</tr>
<tr>
<td>Sikasso</td>
<td></td>
</tr>
<tr>
<td>0-20 cm</td>
<td>1.64</td>
</tr>
<tr>
<td>20-40 cm</td>
<td>1.58</td>
</tr>
<tr>
<td>Siguidolo</td>
<td></td>
</tr>
<tr>
<td>0-20 cm</td>
<td>1.66</td>
</tr>
<tr>
<td>20-40 cm</td>
<td>1.67</td>
</tr>
</tbody>
</table>

Geographic Information

Geographic Coordinate Information

Geographic coordinate information for the boundaries of fields and the location of sampling points was recorded with the Trimble GeoExplorer II (in 2000) or III (in 2002 and 2004), a DGPS (Differential Geographic Positioning System) receiver. Some of the GPS data files were differentially corrected to a known base station; others were not. The GeoExplorer III is accurate to 1-5 meters with base station differential correction; thus the files that were not corrected to a base station can be assumed to have a poorer accuracy than this. Geostatistical theory assumes that spatial locations are known exactly; that is, it does not allow for uncertainty in the coordinate positions of sampling locations, however,
we expect positional uncertainty to have much less impact on global estimates than it would have on point estimates.

Processing of Geographic Information

The raw GPS data was processed with Trimble software: Pathfinder Office (differential correction was made for some of the files) and exported to ESRI ArcGIS shapefile format. ESRI ArcGIS is a geographic information system (GIS) database and processing software. Processing in ESRI ArcMap included projection change from latitude and longitude to UTM (Universal Transverse Mercator). The villages of Fansirakouro and Siguidolo are located in UTM zone 29, while Sougoumba, Oumarbougou, and Sikasso are located in UTM zone 30. The change of projection was to facilitate calculation of distances in meters. Carbon data was associated with sampling locations within the ArcMap database and then exported from ArcMap to text files containing x and y coordinates (UTM eastings and northings) and carbon values. The coordinates for the vertices of boundary polygons were similarly exported. The geographic areas of the regions within the polygons were calculated in ArcMap. Maps showing the magnitude of sampled values were also created in ArcMap.

Non-spatial Statistics

Non-spatial statistics were calculated in R (R Development Core Team, 2006) version 2.4.1, including descriptive statistics and Student’s t-test.
Spatial Analysis

The spatial analysis was also completed in R, primarily in a “package” of R called “geoR” (Ribeiro and Diggle, 2001). Bayesian geostatistical analysis was completed in geoR by means of a function called “krige.bayes”. The description for krige.bayes given in R is as follows: “'krige.bayes' performs Bayesian analysis of geostatistical data allowing specifications of different levels of uncertainty in the model parameters. It returns results on the posterior distributions for the model parameters and on the predictive distributions for prediction locations.”

Decisions required in implementing krige.bayes include: whether or not to include a trend, and if so, which trend; whether to transform the data and which transformation, the choice of correlation function, cell size of the prediction grid, and whether to model anisotropy or include more than one spatial structure in the variogram. The algorithm also requires the prior distributions for the parameters to be specified. A vector of discrete values (“support points”) for the prior distributions must also be provided for the range and relative nugget parameters.

Trends that were included in the model are reported in Table 4. These were decided on the basis of data exploration. Transformations are reported in Table 5; skew coefficients for the datasets are included in the table. Transforming the data is an issue of concern; on the one hand transforming the data to near-normality is desirable, given the normality assumption of the model. On the other hand it is not clear that the estimates are optimal given transformation (Diggle and Ribeiro, 2002, p.133).

The theoretical correlation model was the exponential function, described in the introduction. This is a commonly used function in geostatistical analysis. The primary
difference between the exponential function and any other authorized model that might have been appropriately chosen is the shape of the function near the origin. Because the most closely-spaced pairs of data are usually the most scarce, I judged the data insufficient to distinguish among different shapes of the model near the origin. It is possible to model as random a parameter describing the shape of the function near the origin by using the Matern family of functions, of which the exponential function is a member (Diggle and Ribeiro, 2002, p.132). This would be preferable to “guessing” the shape of the model. However this feature is not currently implemented in geoR.

Cell sizes of the prediction grids were: Oumarbougou 2002, 5 meters; Oumarbougou 2004, 8 meters; Fansirakouro, 8 meters; Sougouomba, 15 meters; Sikasso, 20 meters; Siguidolo, 15 meters (“cell size” is the length of one side of a square cell). Experimentation showed that cell size made little difference to outcomes (at least for global values). Computation time increased dramatically as the cell sizes decreased however.

Anisotropy and additional spatial structures were not included in the model as both of these increase the requirement for data. Including a directional component (anisotropy) requires sufficient data for each direction modeled. Modeling more than one spatial structure (in other words modeling spatial processes with different ranges of spatial correlation) also generally requires more data.

For the purposes of this study the objective was to choose priors that were non-informative. Posited prior distributions for the trend intercept ($\beta_0$) and coefficients ($\beta_1, \beta_2$) were “flat”; for the range parameter ($\varphi$) the prior was “uniform”; and for the partial sill parameter ($\sigma^2$) the prior was “reciprocal”: $\text{pr}(\sigma^2) = 1/\sigma^2$. These choices reflect the
posited shapes of the distributions prior to conditioning to the data. Flat and “uniform”
can be seen to be non-informative, but the choice of reciprocal prior for $\sigma^2$ is also an
expression of prior ignorance according to Diggle and Ribeiro (2002, p.142). The prior
distribution chosen for the relative nugget parameter was also “reciprocal”: $pr (u^2) = 1/
u^2$. Because this parameter has been re-parameterized as $u^2 = \tau^2/\sigma^2$, deciding on its prior
distribution is non-intuitive. In retrospect, the choice of $pr (u^2) = 1/u^2$ did not meet the
objective of being non-informative. The choice of uniform prior for $u^2$ was not clearly
appropriate either, but might have been more conservative.

The support points chosen for the discrete distribution of $\phi$ was a sequence of 51
values between 0 and two times the maximum distance between the data locations. The
support points for the prior for $u^2$ was the sequence between 0 and 10 in increments of
0.2. Increases in the number of support points dramatically increase computation time.

The empirical data variograms are plotted with the midpoints of the bins being 13
equal intervals between 0 and the maximum distance between sampling locations. In
several cases the first bin is empty (no pairs of sampling locations in that bin).

1000 simulations were created.

Replication

In order to assess the variation in outcomes each time the model is run, five
replications were made for each dataset with an identical model. The average standard
deviation of the resulting estimated global means for the five replications was 0.0007, an
amount smaller than the precision with which the final estimates are reported.
Table 4. Trends.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>0-20 cm</th>
<th>20-40 cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oumarbougou</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>2004</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>No trend</td>
<td>No trend</td>
</tr>
<tr>
<td>2004</td>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td>Sougoumba</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>No trend</td>
<td>y</td>
</tr>
<tr>
<td>2002</td>
<td>y</td>
<td>y</td>
</tr>
<tr>
<td>2004</td>
<td>x</td>
<td>No trend</td>
</tr>
<tr>
<td>Sikasso</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>x</td>
<td>No trend</td>
</tr>
<tr>
<td>2004</td>
<td>x and y</td>
<td>x and y</td>
</tr>
<tr>
<td>Siguidolo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>No trend</td>
<td>No trend</td>
</tr>
<tr>
<td>2004</td>
<td>No trend</td>
<td>No trend</td>
</tr>
</tbody>
</table>

† Trends on the x and/or y coordinates that were included in the models for the spatial approach.
<table>
<thead>
<tr>
<th>Dataset</th>
<th>Transformation†</th>
<th>Skew coefficient‡</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-20 cm</td>
<td>20-40 cm</td>
</tr>
<tr>
<td>Oumarbougou</td>
<td>log</td>
<td>log</td>
</tr>
<tr>
<td>2002</td>
<td>none</td>
<td>log</td>
</tr>
<tr>
<td>2004</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>2002</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>2004</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Sougoumba</td>
<td>log</td>
<td>log</td>
</tr>
<tr>
<td>2000</td>
<td>log</td>
<td>log</td>
</tr>
<tr>
<td>2002</td>
<td>log</td>
<td>log</td>
</tr>
<tr>
<td>2004</td>
<td>log</td>
<td>log</td>
</tr>
<tr>
<td>Sikasso</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>2000</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>2004</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Siguidolo</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>2000</td>
<td>log</td>
<td>log</td>
</tr>
</tbody>
</table>

† Choices of dataset transformation that were included in the model for the spatial algorithm.
‡ Skew coefficients for original datasets before transformation.
RESULTS

Overall, results include maps showing the magnitude of sampled values, descriptive statistic summaries, spatial analysis including variogram plots and posterior parameter summaries, predicted value maps, example simulations, and final global estimates.

Descriptive Results

Magnitude of Sampled Values

The symbols in the maps in Fig. 8-13 show the magnitude of carbon values (% C) at sampling locations, in 5 equally-spaced class intervals. All of the maps for each farm are displayed in the same scale to facilitate comparisons.

Values appear to be slightly larger in the 0- to 20-cm soil layer than in the 20- to 40-cm layer, but this is not true for Siguidolo (Fig. 13). This agrees with field observations that the soil appears to be more uniform for the entire 0- to 40-cm soil layer in Siguidolo than in the other villages.

For Oumarbougou (Fig. 8), concentrations of carbon appear to be larger toward the east side of the region, corresponding to the side of the region that is lowest in elevation and agreeing with observed evidence of erosion in the field. We expect erosion to displace surface soil that is higher in organic matter content from the top of slope to the bottom of it.

In Fansirakouro 2004 (Fig. 9, c and d), there are prominent large values in the toe-shaped subregion in the south west, where permanent ridges have been installed.
However this same subregion is not well represented by sampling locations in the earlier 2002 datasets (Fig. 9, a and b), making it difficult to say whether carbon is increasing there.

Carbon values are lower in the Sougoumba 2000 datasets (Fig. 10, a and b), than for the later datasets for that farm (Fig. 10, c and d and Fig. 11, a and b). The largest values for Sougoumba occur in 2004 (Fig. 11, a and b). This suggests that carbon is increasing in this farm. However the earliest dataset was collected on a field-central strategy, where trees and bushes are fewer, while the latest included sampling locations along the field perimeters as well as locations that were nearer to trees than in prior sampling. Carbon values are likely to be higher where there is a greater density of trees and bushes. The change in strategy makes a time comparison difficult. Apparent increases for Sougoumba do not appear to be associated with the location of the permanent ridges.

The sampling configuration for Sikasso 2000 is very scant (Fig. 12, a and b); a few large values occur in the northwest corner. The 2004 datasets for Sikasso (Fig. 12, c and d) show a clear trend from large to small value going from west to east, corresponding to an elevation gradient in the farm from high to low elevation. This is the opposite tendency to that observed for Oumarbougou, where the highest carbon values are found at the bottom of the gradient. In Sikasso, the permanent ridges have been in place longest at the top of the gradient and the most recent ridges have been installed lower down the gradient. It is tempting to speculate that carbon values are higher at the top of the field because the permanent ridges have been there longest. However, we have no information on the pattern of carbon occurrence in the field before the ridges were installed. Carbon values were possibly highest at the top of the field even before the ridges were installed.
The small ring of sampling locations in the southwest corner of the Siguidolo 2000 datasets have moderate values (Fig. 13, a and b) which appear possibly inconsistent with the 2004 data (Fig. 13, c and d), adding evidence that this ring of sampling locations may have been intended to represent some particular feature that is otherwise atypical for that subregion of the field.

The 2004 Siguidolo datasets have a cluster of large values in a fairly small subregion in the north part of the farm, in an area which is low in elevation. Again, this agrees with the expectation that carbon is likely to accumulate at lower topographical positions in the landscape.
Fig. 8. Magnitude of sampled values, Oumarbougou. The locations of permanent ridges are shown as dotted lines.
Fig. 9. Magnitude of sampled values, Fansirakouro. The locations of permanent ridges are shown as dotted lines.
Fig. 10. Magnitude of sampled values, Sougoumba 2000 and 2002. The locations of permanent ridges are shown as dotted lines.
Fig. 11. Magnitude of sampled values, Sougoumba 2004. The locations of permanent ridges are shown as dotted lines.
% Organic Carbon
- 0.14 - 0.30
- 0.31 - 0.46
- 0.47 - 0.62
- 0.63 - 0.78
- 0.79 - 0.95

Fig. 12. Magnitude of sampled values, Sikasso. The locations of permanent ridges are shown as dotted lines.
Fig. 13. Magnitude of sampled values, Siguidolo. The locations of permanent ridges are shown as dotted lines.
Mean carbon concentrations (Table 6, columns 6 and 7), were a little larger in the surface than in the subsurface soil layer, as might be expected, except in Siguidolo. Also, mean concentrations at later sampling dates were slightly larger; except in Siguidolo.

The average carbon concentration when all of the datasets are considered together was 0.40% for the surface layer and 0.35% for the subsurface soil layer, a difference of 0.05 in value between the soil layers. The largest mean value occurred in Fansirakouro, 0.71%, and the lowest occurred in Siguidolo, 0.23%. The highest value of carbon at a single sampling location, 1.5%, was found in Fansirakouro in the surface soil layer in 2004. The lowest value, 0.11%, was found in Siguidolo in 2000 in both soil layers.

Standard deviations of the datasets tended to be higher where the carbon values were higher; that is, in the datasets for the surface layer and also in farms where carbon values were higher. The average standard deviation was 0.15 for the surface layer and 0.11 for the subsurface, a difference in the value of the standard deviation between the two soil layers of 0.04. The largest standard deviation, 0.32, occurred in the Fansirakouro 2004, 0- to 20-cm layer dataset, which also had the largest mean and maximum values. The lowest standard deviation among the datasets was 0.06 for Oumarbougou in the subsurface layer in 2004.

The mean values of carbon in Table 5 represent the non-spatial estimates of global carbon concentration for the estimation regions, which will be compared with spatial estimates for the same regions.
Table 6. Classical Statistics.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>n</th>
<th>Min. 0-20 cm</th>
<th>Max. 0-20 cm</th>
<th>Mean 0-20 cm</th>
<th>Std. Dev. 0-20 cm</th>
<th>Min. 20-40 cm</th>
<th>Max. 20-40 cm</th>
<th>Mean 20-40 cm</th>
<th>Std. Dev. 20-40 cm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0-20 cm</td>
<td>20-40 cm</td>
<td>0-20 cm</td>
<td>20-40 cm</td>
<td>0-20 cm</td>
<td>20-40 cm</td>
<td>0-20 cm</td>
<td>20-40 cm</td>
</tr>
<tr>
<td>Oumarbougou</td>
<td>2002</td>
<td>24, 25†</td>
<td>0.14</td>
<td>0.15</td>
<td>0.68</td>
<td>0.46</td>
<td>0.33</td>
<td>0.26</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>22</td>
<td>0.18</td>
<td>0.16</td>
<td>0.61</td>
<td>0.42</td>
<td>0.34</td>
<td>0.25</td>
<td>0.13</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td>2002</td>
<td>20, 18†</td>
<td>0.27</td>
<td>0.24</td>
<td>0.95</td>
<td>0.88</td>
<td>0.66</td>
<td>0.52</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>21</td>
<td>0.34</td>
<td>0.29</td>
<td>1.5</td>
<td>1.1</td>
<td>0.71</td>
<td>0.63</td>
<td>0.32</td>
</tr>
<tr>
<td>Sougoumba</td>
<td>2000</td>
<td>16</td>
<td>0.21</td>
<td>0.17</td>
<td>0.50</td>
<td>0.44</td>
<td>0.29</td>
<td>0.25</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>2002</td>
<td>38</td>
<td>0.17</td>
<td>0.16</td>
<td>0.51</td>
<td>0.48</td>
<td>0.32</td>
<td>0.28</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>49</td>
<td>0.17</td>
<td>0.14</td>
<td>0.76</td>
<td>0.55</td>
<td>0.32</td>
<td>0.31</td>
<td>0.12</td>
</tr>
<tr>
<td>Sikasso</td>
<td>2000</td>
<td>13</td>
<td>0.14</td>
<td>0.17</td>
<td>0.86</td>
<td>0.63</td>
<td>0.42</td>
<td>0.35</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>75</td>
<td>0.26</td>
<td>0.22</td>
<td>0.95</td>
<td>0.74</td>
<td>0.51</td>
<td>0.43</td>
<td>0.15</td>
</tr>
<tr>
<td>Siguidolo</td>
<td>2000</td>
<td>14,13†</td>
<td>0.11</td>
<td>0.11</td>
<td>0.39</td>
<td>0.40</td>
<td>0.27</td>
<td>0.27</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>35</td>
<td>0.12</td>
<td>0.15</td>
<td>0.61</td>
<td>0.46</td>
<td>0.23</td>
<td>0.25</td>
<td>0.11</td>
</tr>
</tbody>
</table>

† Number of observations for the 0- to 20-cm layer and the 20- to 40-cm layer respectively.
Spatial Analysis

The spatial analysis includes empirical variograms and variogram models estimated by the Bayesian algorithm, and histograms and summaries of the posterior parameter distributions.

Variograms

The variograms in Fig. 14 to 24 are empirical data variograms (represented as points in the graphs) overlain with summary variograms (represented as lines) that are the outcome of the Bayesian algorithm (krige.bayes function in geoR). Decisions of trend and transformation that were included in the krige.bayes models are given in the captions for the figures.

Although the krige.bayes models for some of the datasets included trends, trends have not been subtracted from the data before plotting the empirical data variograms. In contrast, in the conventional geostatistical approach we usually subtract a trend from the data before creating the empirical variogram for the residuals and fitting a model to it. In the Bayesian approach, the model is not fit to a graphical display of the residual variogram, but instead estimated by the algorithm. Also the trend parameters ($\beta_0, \beta_1, \beta_2$), are random and may take many values. If we were to retroactively subtract the trend estimated by the algorithm from the data before plotting, we would have to decide which values of the trend parameters to choose.

My purpose in plotting the Bayesian outcome variograms against the empirical data variograms is to check whether the outcomes seem plausible. Also, by plotting
several relevant quantiles of the posterior variograms, uncertainty in the estimation of the variogram parameters can be visually assessed.

The posterior variogram summaries in Fig. 14 to 24 are only three out of many that could be plotted. Their values are “summaries” of the posterior distributions of the variogram parameters; specifically they represent the 0.05, 0.5, and 0.95 quantiles of the posterior distributions of the partial sill, range, and relative nugget parameters. Other summaries are possible: one could plot for example the mean, mode, minimum, maximum, or other quantile summaries. I chose the 0.05 (lower dotted lines in the graphs) and 0.95 (upper dotted lines in the graphs) quantiles to represent a 90% probability envelope for an estimate of the true variogram parameters, and specifically chose the 90% level rather than some other level only because it seemed the best one to describe the uncertainty visually. Many of the upper lines of a 95 % probability envelope, which I might have chosen instead, were off the tops of the graphs.

For several of the datasets, even the upper lines of the 90 % probability envelopes are off the top of the graph, or barely appear. The “poorest” in this regard are for Fansirakouro 2002 (Fig. 16); and Siguidolo 2000, the 0- to 20-cm layer dataset (Fig. 23, top graph). Many of the empirical variograms have points that appear quite scattered, for example those just mentioned as well as the Sougoumba 2000 datasets (Fig. 18). In general we can say that for many of the datasets, there appears to be quite a large uncertainty in the modeled variograms.

In the Siguidolo 2004 variograms, semivariances appear to be decreasing rather than increasing with separation distance. This behavior may have resulted from most of the closely-spaced pairs (of which there are few) having been sampled in the small
subregion which has large values of carbon, in the north part of the farm (see Fig. 13, c and d). In subregions where values are large, the local variances are often large as well, resulting in large semivariances in the variogram at small values of separation distance, if those lags contain pairs from the high-valued subregion only. In this case it appears that in addition the small subregion is relatively isolated, with nearby values of carbon being much lower in value and contributing to high variances between pairs at moderately short separation distances.

When the data contain a trend, the points of the empirical variogram often appear to increase with increasing separation distance and do not level out at a sill. For example, in the variogram for the Oumarbougou 2002, 0- to 20-cm layer dataset (Fig. 14, top graph), the points of the empirical variogram increase from left to right. If we were modeling this variogram by the conventional method, we might ignore points for separation distances beyond about half of the longest distance, that is everything beyond about 100 meters. Even considering the points to 100 meters, there appears to be a trend. The median variogram (solid line), which is the output of the Bayesian algorithm, does not follow this trend but instead levels out at a lower semivariance than it might. This means that, at least for the median values of the variogram parameters, the trend is fairly strong and the variogram, which is modeling the spatial correlation, follows the residuals of that trend.

In practice we never know for sure whether a trend exists or not. The behavior of the data might be explained better by spatial correlation than by a trend. In the variogram in Fig. 14 which we are examining, the upper line of the envelope includes all points to
the 100 meter distance. This shows that there is uncertainty whether spatial correlation might explain the behavior of the data better than a trend might.

Of course in the Bayesian approach the information from all pairs of data values are considered by the likelihood function, not just those up to half the longest separation distance (and not just their average semivariances based on choices of “bins” as for the empirical variograms).

We can visually assess the strength of trends in the variograms in Fig. 14 to 24 by the extent to which the median variogram and upper probability envelopes do not contain the apparent trends. Among the datasets for which trends were modeled, those in Oumarbougou (especially Oumarbougou 2002, see Fig. 14, both graphs), Sikasso (for example Sikasso 2000, the 0- to 20-cm layer; Fig. 21, top graph), and Fansirakouro 2004 (Fig. 17, both graphs) appear stronger than those in Sougoumba (Fig. 18 to 20), which appear very weak to non-existent.
Fig. 14. Empirical sample variograms (points) and estimated variograms (lines), Oumarbougou, 2002; top: 0- to 20-cm soil layer, bottom: 20- to 40-cm soil layer. Models for both included a trend in x and both datasets were log transformed.
Fig. 15. Empirical sample variograms (points) and estimated variograms (lines), Oumarbougou, 2004; top: 0- to 20-cm soil layer, bottom: 20- to 40-cm soil layer. Top: data was not transformed; bottom: data was log transformed. Models for both included a trend in x.
Fig. 16. Empirical sample variograms (points) and estimated variograms (lines), Fansirakouro, 2002; top: 0- to 20-cm soil layer, bottom: 20- to 40-cm soil layer. No trends or transformations were included in the models for either dataset.
Fig. 17. Empirical sample variograms (points) and estimated variograms (lines), Fansirakouro, 2004; top: 0- to 20-cm soil layer, bottom: 20- to 40-cm soil layer. Top: model included a trend in x; bottom: model included a trend in y. Neither dataset was transformed.
Fig. 18. Empirical sample variograms (points) and estimated variograms (lines), Sougoumba, 2000; top: 0- to 20-cm soil layer, bottom: 20- to 40-cm soil layer. Top: model did not include a trend; bottom: model included a trend in y. Both datasets were log transformed.
Fig. 19. Empirical sample variograms (points) and estimated variograms (lines), Sougoumba, 2002; top: 0- to 20-cm soil layer, bottom: 20- to 40-cm soil layer. Models for both include a trend in $y$; both datasets were log transformed.
Fig. 20. Empirical sample variograms (points) and estimated variograms (lines), Sougoumba, 2004; top: 0- to 20-cm soil layer, bottom: 20- to 40-cm soil layer. Top: model included a trend in x; bottom: no trend was included in the model. Both datasets were log transformed.
Fig. 21. Empirical sample variograms (points) and estimated variograms (lines), Sikasso, 2000; top: 0- to 20-cm soil layer, bottom: 20- to 40-cm soil layer. Top: model included a trend in x; bottom: no trend was included in the model. Neither dataset was transformed.
Fig. 22. Empirical sample variograms (points) and estimated variograms (lines), Sikasso, 2004; top: 0- to 20-cm soil layer, bottom: 20- to 40-cm soil layer. Models for both include a trend in x and y; neither dataset was transformed.
Fig. 24. Empirical sample variograms (points) and estimated variograms (lines), Siguidolo, 2004; top: 0- to 20-cm soil layer, bottom: 20- to 40-cm soil layer. No trend was included in either of the models; however, the datasets were both log transformed.
Histograms of the Posterior Range Parameter

The histograms in Fig. 25 -30 represent the posterior distributions for the range parameter that were estimated from the data by the Bayesian algorithm. They are a visual representation specifically of the variability in the value of the range parameter, whereas the probability envelopes overlaying the variograms show visually the joint uncertainty of the range, relative nugget and partial sill parameters. All of the histograms for any one farm are displayed in the same scale.

The spread and the shapes of the histograms show the uncertainty in estimating the range parameter. Some of the histograms appear almost uniform in shape, indicating great uncertainty (all values of the range are equally likely). However, some of them although wide, have a distinct peak (for example Sougoumba 2002, the 20- to 40-cm layer, Fig. 27, lower right-hand graph), and for several, the shape of the histogram is strongly positively skewed with the tallest bar of the histogram to the left (toward smaller values of range). These last are the best defined and include Sikasso 2004, both the 0- to 20-cm and the 20- to 40-cm datasets (Fig. 29, right-hand-side graphs), and Siguidolo 2004 (Fig. 30, right-hand-side graphs), most strongly in the 0- to 20-cm soil layer (upper graph).

In general, we could say that for many of the datasets, the range parameter appears poorly defined, however for those just mentioned it appears reasonably well-defined.

Particular values for the range are difficult to determine from the histograms. Some useful summaries of this parameter are given in Table 7. I have not included histograms for the partial sill, relative nugget, or trend parameters as the histogram shapes for each of these are very similar across datasets (positively skewed for the partial sill and
relative nugget parameters and normal for the trend parameters). I give some relevant summaries for these parameters in Table 8 (partial sill) and Table 9 (relative nugget and trend parameters).
Fig. 25. Histograms of posterior range parameter, Oumarbougou. Top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer; left: 2002; right: 2004.
Fig. 26. Histograms of posterior range parameter, Fansirakouro. Top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer; left: 2002; right: 2004.
Fig. 27. Histograms of posterior range parameter, Sougoumba. Top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer; left: 2000; right: 2002.
Fig. 28. Histograms of posterior range parameter, Sougoumba, 2004. Top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer.
Fig. 29. Histograms of posterior range parameter, Sikasso. Top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer; left: 2000; right: 2004.
Fig. 30. Histograms of posterior range parameter, Siguidolo. Top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer; left: 2000; right: 2004.
Posterior Parameter Summaries

Table 7 shows summaries of the posterior range parameter, including the 0.05 and 0.95 quantiles, the median, and the mode. These directly correspond to the histogram distributions in Fig. 25 – 30 and to the range parameter values for the variogram models in Fig. 14 – 24 (except that the mode is not shown in the variograms). The maximum distance between sampling locations for each of the datasets is given in Table 7 for comparison with range parameter values.

Considering that the practical range is three times the value of the range parameter, the median values of the range parameter in the table are all quite large relative to their respective maximum sampling distances. It appears that more moderate range distances (for example in the neighborhood of 100 to 150 meters, practical range), which we expected to find based on our experience, are not well defined by the data. The spread between the 0.05 and 0.95 quantiles show the large uncertainty in estimating values of the range parameter. This is consistent with our observation that many of the range parameter histograms are widely spread.

On the basis of our experience, the values of the mode look more reasonable. Possibly the mode of the posterior distribution is a better estimator for this parameter [under specific circumstances, the mode corresponds to parameter estimates made by the method of restricted maximum likelihood (REML) (Diggle and Ribeiro, 2002, p.136)]. However, the histograms show that in some cases the distributions are almost uniform; the mode is obviously unreliable in these cases.

In Table 8, posterior parameter summaries are given for the partial sill parameter. Dataset variances are given in the table for comparison; these are the variances of the...
transformed data for those datasets that were log-transformed. The summary values in the table correspond to the partial sills of the variogram models in Fig. 14 – 24.

Again, the spread between the values of the 0.05 and 0.95 quantiles shows the large uncertainty in estimating this parameter. The mode of the posterior for the partial sill parameter is shown for comparison with the median; in many cases it is larger than the median. The values of the mode predict that the spatial component contributes a larger part of the variance than the median values suggest. Modes of the relative nugget (not shown) also tend to be smaller than their median counterparts. This is logical; where the relative nugget decreases, the partial sill generally increases. As with the modes for the posterior range parameter (which were generally more plausible than the median range values), the modes of the partial sill and relative nugget give greater indication of spatial correlation than the medians.

In Table 9, median values of the posterior distributions of the relative nugget and trend parameters are given. The trend coefficients are included in the table but the trend intercepts are not shown. Also not shown are the 0.05 and 0.95 quantile values for these parameters; however, the spread in these is wide, as it was in the case of the range and partial sill parameters.

The values of the median relative nugget correspond to the values of the median variogram models plotted in Fig. 14 – 24, although this is more difficult to see than for the other parameters. The relative nugget is the ratio of the nugget variance to partial sill. A relative nugget of 1.0, for example, means that the nugget variance and partial sill each represent half of the total sill variance.
In general the relative nuggets are larger than expected, or at least the median relative nugget values are larger than expected. Values of the mode of the relative nugget parameter (not shown) are around 0.2 and very uniform between datasets, only because of the limited values that were chosen for the support points of the discrete prior for this parameter.

The values of median relative nugget (in Table 9) can be compared between datasets. Several of the datasets have values of around 0.3 or 0.4, for example several of the Sikasso datasets. The largest median relative nugget value, 2.6, is for the Siguidolo 2000, 0- to 20-cm layer dataset. Smaller values of relative nugget (reflecting smaller nugget variance) present stronger evidence for spatial correlation.

The value of the nugget variance can be calculated by multiplying the relative nugget times the partial sill. For example for the Fansirakouro 2004, 20- to 40-cm layer dataset, the median relative nugget value is 0.4 (Table 9, column 4). The median partial sill value for this dataset is 0.054 (Table 8, column 7), therefore the nugget variance is: 0.4 x 0.054 = 0.022. The estimated total sill then is: 0.022 + 0.054 = 0.076. This can be compared with the dataset variance of 0.064 (Table 8, column 6), which is less than the total sill as we generally expect, at least in the case of stationarity. In the case of a trend, the dataset variance also includes variance contributed by the trend.

We can also compare the nugget variance, 0.022, which we calculated above, with the measurement error variance, 0.0004, which we calculated by replicating the chemical analysis for identical soil samples. The measurement error variance is very small in comparison with the estimated nugget variance, at least for this dataset (but this is generally expected to be the case). The nugget variance of the datasets in our study is
most likely primarily comprised of short-range spatial variation or also quite possibly short- to medium-range variation not well enough described by the data. (However, measurement error might also consist of more than variation in the chemical analysis, it could also consist of variability in sampling technique, for example.)

The median posterior trend coefficients give indication of the steepness and direction of trends. A positive sign for the x-coefficient means that carbon increases from west to east with the x-coordinate value; a positive sign for the y-coefficient means carbon increases south to north with the y-coordinate. Negative signs indicate the reverse of these. The magnitude of the trend coefficients indicate the steepness of the trend.

The largest trend coefficient is for the Oumarbougou, 2000, 0- to 20-cm layer dataset, which has a value of 0.006 for the x-coefficient (Table 9, column 2). For this dataset, we saw a clear trend in the variogram. The sign of the coefficient is positive, therefore carbon increases in the west to east direction. This agrees with what we saw in the map of the magnitude of sampled values (Fig. 8a).

For the Sougoumba 2002, 0- to 20-cm and 20- to 40-cm datasets, the sign of the y-coefficient spans zero for the 0.05 and 0.95 quantiles (not shown). This is consistent with what we saw in the variograms; the trend appears to be very weak or non-existent for these two datasets.
Table 7. Posterior Summaries for the Range Parameter.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>0-20 cm</th>
<th>20-40 cm</th>
<th>0-20 cm</th>
<th>20-40 cm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>distance</td>
<td>Median</td>
<td>Quantile.05</td>
<td>Quantile.95</td>
</tr>
<tr>
<td>Oumarbougou</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>185</td>
<td>192</td>
<td>15</td>
<td>355</td>
</tr>
<tr>
<td>2004</td>
<td>247</td>
<td>277</td>
<td>30</td>
<td>474</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>288</td>
<td>300</td>
<td>23</td>
<td>553</td>
</tr>
<tr>
<td>2004</td>
<td>337</td>
<td>364</td>
<td>54</td>
<td>648</td>
</tr>
<tr>
<td>Sougoumba</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>642</td>
<td>616</td>
<td>51</td>
<td>1232</td>
</tr>
<tr>
<td>2002</td>
<td>675</td>
<td>648</td>
<td>27</td>
<td>1324</td>
</tr>
<tr>
<td>2004</td>
<td>751</td>
<td>691</td>
<td>90</td>
<td>1411</td>
</tr>
<tr>
<td>Sikasso</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>802</td>
<td>834</td>
<td>96</td>
<td>1572</td>
</tr>
<tr>
<td>2004</td>
<td>829</td>
<td>332</td>
<td>33</td>
<td>1426</td>
</tr>
<tr>
<td>Siguidolo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>557</td>
<td>646</td>
<td>67</td>
<td>1091</td>
</tr>
<tr>
<td>2004</td>
<td>568</td>
<td>250</td>
<td>23</td>
<td>1023</td>
</tr>
</tbody>
</table>

† Summaries of the distributions predicted by the spatial algorithm for values of the range parameter of the correlation model.
‡ Maximum distance between sampling locations.
Table 8. Posterior Summaries for the Partial Sill Parameter.†

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Sample Variance</th>
<th>Median</th>
<th>Q.05</th>
<th>Q.95</th>
<th>Mode</th>
<th>Variance</th>
<th>Median</th>
<th>Q.05</th>
<th>Q.95</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oumarbougou</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>0.200</td>
<td>0.047</td>
<td>0.009</td>
<td>0.278</td>
<td>0.074</td>
<td>0.110</td>
<td>0.030</td>
<td>0.006</td>
<td>0.207</td>
<td>0.050</td>
</tr>
<tr>
<td>2004</td>
<td>0.018</td>
<td>0.007</td>
<td>0.001</td>
<td>0.046</td>
<td>0.012</td>
<td>0.060</td>
<td>0.017</td>
<td>0.004</td>
<td>0.110</td>
<td>0.038</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>0.040</td>
<td>0.021</td>
<td>0.004</td>
<td>0.130</td>
<td>0.032</td>
<td>0.026</td>
<td>0.013</td>
<td>0.003</td>
<td>0.082</td>
<td>0.024</td>
</tr>
<tr>
<td>2004</td>
<td>0.103</td>
<td>0.057</td>
<td>0.008</td>
<td>0.231</td>
<td>0.203</td>
<td>0.064</td>
<td>0.054</td>
<td>0.009</td>
<td>0.152</td>
<td>0.064</td>
</tr>
<tr>
<td>Sougoumba</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.040</td>
<td>0.026</td>
<td>0.005</td>
<td>0.138</td>
<td>0.037</td>
<td>0.070</td>
<td>0.043</td>
<td>0.007</td>
<td>0.217</td>
<td>0.070</td>
</tr>
<tr>
<td>2002</td>
<td>0.050</td>
<td>0.023</td>
<td>0.006</td>
<td>0.143</td>
<td>0.044</td>
<td>0.070</td>
<td>0.078</td>
<td>0.011</td>
<td>0.241</td>
<td>0.063</td>
</tr>
<tr>
<td>2004</td>
<td>0.110</td>
<td>0.164</td>
<td>0.030</td>
<td>0.392</td>
<td>0.299</td>
<td>0.100</td>
<td>0.043</td>
<td>0.010</td>
<td>0.293</td>
<td>0.352</td>
</tr>
<tr>
<td>Sikasso</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.044</td>
<td>0.014</td>
<td>0.002</td>
<td>0.074</td>
<td>0.018</td>
<td>0.023</td>
<td>0.023</td>
<td>0.005</td>
<td>0.065</td>
<td>0.019</td>
</tr>
<tr>
<td>2004</td>
<td>0.022</td>
<td>0.014</td>
<td>0.004</td>
<td>0.055</td>
<td>0.013</td>
<td>0.011</td>
<td>0.006</td>
<td>0.001</td>
<td>0.024</td>
<td>0.006</td>
</tr>
<tr>
<td>Siguidolo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.008</td>
<td>0.003</td>
<td>0.001</td>
<td>0.027</td>
<td>0.009</td>
<td>0.005</td>
<td>0.003</td>
<td>0.001</td>
<td>0.021</td>
<td>0.005</td>
</tr>
<tr>
<td>2004</td>
<td>0.140</td>
<td>0.114</td>
<td>0.022</td>
<td>0.437</td>
<td>0.109</td>
<td>0.070</td>
<td>0.044</td>
<td>0.008</td>
<td>0.219</td>
<td>0.059</td>
</tr>
</tbody>
</table>

† Summaries of the distributions predicted by the spatial algorithm for values of the partial sill parameter of the correlation model.
‡ For log transformed datasets this is the variance of the transformed data.
Table 9. Median Posterior Values for the Relative Nugget and Trend Parameters.†

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Median posterior values</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>relative nugget</td>
<td>x-coefficient</td>
<td>y-coefficient</td>
</tr>
<tr>
<td></td>
<td>(v²) ‡</td>
<td>(β₁)</td>
<td>(β₁ or β₂)</td>
</tr>
<tr>
<td></td>
<td>0-20 cm</td>
<td>20-40 cm</td>
<td>0-20 cm</td>
</tr>
<tr>
<td>Oumarbougou</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>1.6</td>
<td>0.006</td>
<td>1.8</td>
</tr>
<tr>
<td>2004</td>
<td>2.0</td>
<td>0.001</td>
<td>2.6</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>1.6</td>
<td></td>
<td>1.8</td>
</tr>
<tr>
<td>2004</td>
<td>0.8</td>
<td>-0.002</td>
<td>0.4</td>
</tr>
<tr>
<td>Sougoumba</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>1.2</td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>2002</td>
<td>2.0</td>
<td>0.001</td>
<td>0.4</td>
</tr>
<tr>
<td>2004</td>
<td>0.4</td>
<td>0.001</td>
<td>2.0</td>
</tr>
<tr>
<td>Sikasso</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>1.0</td>
<td>-0.001</td>
<td>0.3</td>
</tr>
<tr>
<td>2004</td>
<td>0.4</td>
<td>-0.001</td>
<td>-0.00003</td>
</tr>
<tr>
<td>Siguidolo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>2.6</td>
<td></td>
<td>1.4</td>
</tr>
<tr>
<td>2004</td>
<td>0.4</td>
<td></td>
<td>1.2</td>
</tr>
</tbody>
</table>

† Median values of the distributions predicted by the spatial algorithm for the relative nugget parameter of the correlation model, and for the trend parameters (the x- and y-coefficients).
‡Relative nugget (v²) = nugget variance (σ²) / partial sill (σ²).
Predicted Values

The predicted value maps in Fig. 31 - 41 show the smoothed interpolated surfaces that predict the values of the underlying trend plus spatial component, \( X(u)\beta + S(u) \). Predicted values are the expected values (in other words the means) of the Bayesian prediction distributions. The values are \( \% C \) (by weight of soil).

All of the maps for any one farm are displayed in the same scale. The values are classified in 12 equally-spaced intervals between the minimum and maximum values found in the four (or six in the case of Sougoumba) prediction surfaces for each farm, and displayed on a color scale with red representing the smallest values of \( \% C \) and white representing the largest values. Contour lines for a similar number of levels of carbon values overlay the color maps. Some of the maps have a "flat" appearance because they have a smaller range in value than the other datasets for the same farm and therefore have fewer than 12 intervals or contours. Typically the lower soil layer has a smaller range in value than the surface layer, but this is not true for all of the farms. The lower soil layer for the Sougoumba 2000 and 2002 datasets, for example, does not have a smaller range in predicted value than the surface layer. However, the 2004 datasets for Sougoumba do show a narrower range of value in the lower soil layer than in the surface layer.

The Fansirakouro 2002 and Siguidolo 2000 maps are flat in appearance for both soil layers, indicating a narrower range in predicted values for the earlier sampling date than for the later sampling date (2004) for the two farms. The standard deviations of the original datasets also increase from the earlier to the later sampling dates (Table 6), most strongly for Fansirakouro. The explanation for the increase in variability for the Fansirakouro datasets may be the increase in the extent of the sampling area in
Fansirakouro between 2002 and 2004 (Fig. 4, compare a and b). In Siguidolo the explanation appears to be that the very large carbon values in the small subregion in the north part of the farm in 2004 but not in 2000 increase the variability for the 2004 datasets.

Many of the features discussed in relation to the maps showing the magnitude of sampled values and variograms can be seen in the predicted surfaces, for example trends can be seen in the Oumarbougou, Sikasso, and Fansirakouro 2004 maps, especially in the 0- to 20-cm soil layer.

Maximum and minimum predicted values for each farm are reported in the captions for the figures. These are generally smoothed in comparison to the original maximums and minimums in Table 6.

My purpose in including the predicted value maps here is to make general observations, rather than to make predictions of the value of carbon for particular points in the regions. The uncertainty in estimating the value of carbon at individual locations should be larger than the uncertainty in estimating an average or total value for the region. The average standard error for point predictions, which is simply the average of the standard errors for all prediction nodes in the grid, gives an idea of the size of this uncertainty. Among the datasets, the largest average standard error for the point estimates was 0.144, for the Fansirakouro 2004, 0- to 20-cm dataset; and the smallest average standard error was 0.028 for the Oumarbougou 2002, 20- to 40-cm dataset. The largest standard error for a single prediction node was 0.227, for a node in the Fansirakouro 2004, 20- to 40-cm dataset. A margin of error (at the 95 % level) for this node would be approximately two times the standard error, or 0.45; this can be compared with the
average value of carbon for this farm in the subsurface soil layer which is around 0.6 %
(giving an interval of 0.6 plus or minus 0.45).

The average of the point predictions for all of the nodes in each predicted value
grid provides an estimate of the average carbon concentration for each dataset. The final
estimates, however, will be based on the average of 1000 simulated means. The estimates
arrived at by these two alternative methods should be very close in value. In comparison
to the method of averaging the values in the predicted value grids, the simulations
provide a distribution of probable estimates; which will provide a measure of the
uncertainty of the estimates.
Fig. 31. Predicted values (% C), Oumarbougou, 2002; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2002 (this page) and 2004 (following page) are in the same scale; minimum value: 0.15; maximum value 0.53.
Fig. 32. Predicted values (% C), Oumarbougou, 2004; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2002 (previous page) and 2004 (this page) are in the same scale; minimum value: 0.15; maximum value 0.53.
Fig. 33. Predicted values (% C), Fansirakouro, 2002; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2002 (this page) and 2004 (following page) are in the same scale; minimum value: 0.37; maximum value 1.2.
Fig. 34. Predicted values (% C), Fansirakouro, 2004; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2002 (previous page) and 2004 (this page) are in the same scale; minimum value: 0.37; maximum value 1.2.
Fig. 35. Predicted values (% C), Sougoumba, 2000; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000 (this page), 2002 and 2004 (following pages) are in the same scale; minimum value: 0.16; maximum value 0.47.
Fig. 36. Predicted values (% C), Sougoumba, 2002; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000 (previous page), 2002 (this page) and 2004 (following page) are in the same scale; minimum value: 0.16; maximum value 0.47.
Fig. 37. Predicted values (% C), Sougoumba, 2004; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000, 2002 (previous pages), and 2004 (this page) are in the same scale; minimum value: 0.16; maximum value 0.47.
Fig. 38. Predicted values (% C), Sikasso, 2000; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000 (this page) and 2004 (following page) are in the same scale; minimum value: 0.13; maximum value 0.81.
Fig. 39. Predicted values (% C), Sikasso, 2004; top: 0-20 cm; bottom: 20-40 cm. Map values for 2000 (previous page) and 2004 (this page) are in the same scale; minimum value: 0.13; maximum value 0.81.
Fig. 40. Predicted values (% C), Siguidolo, 2000; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000 (this page) and 2004 (following page) are in the same scale; minimum value: 0.16; maximum value 0.33.
Fig. 41. Predicted values (% C), Siguidolo, 2004; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000 (previous page) and 2004 (this page) are in the same scale; minimum value: 0.16; maximum value 0.33.
Simulated Values

Whereas predicted values are the expected values of the prediction distributions, simulated values are values “sampled” from those distributions according to their probabilities. They are samples of $Y_0$, probable outcomes of the underlying process $X(u)\beta + S(u) + \varepsilon(u)$. Simulated values include the error term. Many simulations are required to describe the prediction distributions fully. The values in the maps are % C, by weight of soil.

A single simulation for each dataset is shown in Fig.42-52. As for the predicted value maps, the values are classified in 12 equally-spaced intervals between the minimum and maximum values found in the four (or six) prediction surfaces for each farm, and the color scale is from red (small values of carbon) to white (large values of carbon). Contour lines overlay the color maps.

A great deal of variability between simulations for the same dataset is possible, although each simulation is equally probable. We need to be careful about drawing conclusions from a single simulation. However, many of the features discussed previously are visible. For example, trends can be seen in the maps for Oumarbougou and Sikasso and for Fansirakouro in 2004. Trends are not evident in the Sougoumba datasets.

Some features of spatial correlation are also evident; these are best expressed in the simulations for Sikasso 2004 (Fig. 50), especially in the 0- to 20-cm layer. They can be seen as “patches” perhaps 100 – 200 meters in diameter (this would be easier to see if the trends were removed). As we saw in the spatial analysis, the Sikasso 2004 datasets had better defined range and relative nugget parameters than many of the other datasets.
The modes of the range parameter were 33 meters for this farm (approximately 100 meters practical range) for both soil layers (Table 7).

Each simulation provides an estimate of the global mean carbon concentration. This is calculated as the arithmetic mean of the simulated point values in the grid: \( \frac{\sum Y_0}{N(u_0)} \), where \( N(u_0) \) is the number of simulated grid nodes within the region. The average of the 1000 simulated means for each dataset provides the final estimate of the true mean carbon concentration, and the variation in the distribution of the 1000 simulated means provides a measure of the uncertainty of that estimate.
Fig. 42. Simulated values (% C), Oumarbougou, 2002; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2002 (this page) and 2004 (following page) are in the same scale; minimum value: 0.04; maximum value 0.77.
Fig. 43. Simulated values (% C), Oumarbougou, 2004; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2002 (previous page) and 2004 (this page) are in the same scale; minimum value: 0.04; maximum value 0.77.
Fig. 44. Simulated values (% C), Fansirakouro, 2002; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2002 (this page) and 2004 (following page) are in the same scale; minimum value: 0.12; maximum value 1.2.
Fig. 45. Simulated values (% C), Fansirakouro, 2004; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2002 (previous page) and 2004 (this page) are in the same scale; minimum value: 0.12; maximum value 1.2.
Fig. 46. Simulated values (% C), Sougoumba, 2000: top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000 (this page), 2002, and 2004 (following pages) are in the same scale; minimum value: 0.12; maximum value 0.65.
Fig. 47. Simulated values (% C), Sougoumba, 2002; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000 (previous page), 2002 (this page) and 2004 (following page) are in the same scale; minimum value: 0.12; maximum value 0.65.
Fig. 48. Simulated values (% C), Sougoumba, 2004; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000, 2002 (previous pages), and 2004 (this page) are in the same scale; minimum value: 0.12; maximum value 0.65.
Fig. 49. Simulated values (% C), Sikasso, 2000; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000 (this page) and 2004 (following page) are in the same scale; minimum value: 0.06; maximum value 0.86.
Fig. 50. Simulated values (% C), Sikasso, 2004; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000 (previous page) and 2004 (this page) are in the same scale; minimum value: 0.06; maximum value 0.86.
Fig. 51. Simulated values (% C), Siguidolo, 2000; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000 (this page) and 2004 (following page) are in the same scale; minimum value: 0.07; maximum value 0.60.
Fig. 52. Simulated values (% C), Siguidolo, 2004; top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer. Map values for 2000 (previous page) and 2004 (this page) are in the same scale; minimum value: 0.07; maximum value 0.60.
Global Estimation

Probability Density Plots

Probability density curves for the distributions of the 1000 simulated global means for each of the datasets are shown in Fig. 53 - 57 (values are % C). Each simulated global mean is the arithmetic mean of the simulated point values in the grid: \((\sum Y_o) / N(uo)\), where \(N(uo)\) is the number of simulated grid nodes within the region. Because they are distributions of means, the shapes of the distributions are much closer to normal than the original dataset distributions.

All of the plots for any one farm are displayed in the same scale. Solid lines in the graphs represent an earlier sampling date (either 2000 or 2002, depending on the farm) and dashed lines represent a later sampling date (2002 or 2004, depending on the farm). In the graphs for Sougoumba, dotted lines represent the third and most recent sampling date (2004).

A taller, narrower distribution shape indicates greater precision, and a shorter, wider distribution shape indicates less precision. We expect to be able to make a more precise estimate, in other words provide an estimate with a smaller uncertainty interval, for those datasets with narrower distributions of simulated means.

In general, the distributions tend to be narrower, in the 20- to 40-cm soil layer than in the 0- to 20-cm layer. The distributions also tend to be narrower for large datasets than for small datasets as might be expected; for example compare the curves for Sikasso 2000 (\(n = 13\)) with Sikasso 2004 (\(n = 75\)) in Fig. 56.

The density curves for the lower soil layer, except those for Siguidolo, tend to be shifted toward the left (toward smaller values of carbon) in comparison with the curves.
for the upper soil layer. In other words, carbon values are smaller in the 20- to 40-cm soil layer than in the 0- to 20-cm layer (except in Siguidolo). The curves also appear to be shifted farther to the right (toward larger values) for later sampling dates than for earlier sampling dates, again except for Siguidolo, where instead the curve for the later sampling date is shifted toward the left (toward smaller values). In other words carbon values appear to be higher in later sampling dates than in earlier sampling dates (except in Siguidolo). These observations are all consistent with previous observations.
Fig. 53. Probability density plots for 1000 simulated means (% C), Oumarbougou. Top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer; solid lines: 2002; dashed lines: 2004. Grey dashed line: 2004 subregion corresponding to 2002 region.
Fig. 54. Probability density plots for 1000 simulated means (% C), Fansirakouro. Top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer; solid lines: 2002; dashed lines: 2004.
Fig. 56. Probability density plots for 1000 simulated means (% C), Sikasso. Top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer; solid lines: 2000; dashed lines: 2004.
Fig. 57. Probability density plots for 1000 simulated means (% C), Siguidolo. Top: 0- to 20-cm soil layer; bottom: 20- to 40-cm soil layer; solid lines: 2000; dashed lines: 2004.
Global Mean Estimates

The means of the distributions of simulated global means, discussed above, provide the final estimates of the global mean concentration of carbon for each of the farms. A measure of uncertainty for these final estimates is provided by the 0.025 and 0.975 quantiles of the simulated means, which are the upper and lower limits of a 95% probability interval, respectively. As we have actual distributions of 1000 values, probability intervals are preferred over confidence intervals, which infer the uncertainty from the standard deviation.

The 95% level is common for reporting scientific results; this does not mean that it is necessarily the most appropriate level in all circumstances. However, the Intergovernmental Panel on Climate Change suggests that it is good practice to estimate uncertainties at the 95% confidence level for the purposes of “estimating, measuring, monitoring and reporting on carbon stock changes” for Land Use, Land-Use Change and Forestry (LULUCF) activities under the Kyoto Protocol (IPCC, 2003, chapter 4, p. 37). The LULUCF is not the only program under which carbon trading may take place. However, it seems reasonable to follow this guideline here.

The final estimates of mean carbon concentration (% C) and the associated upper and lower limits of the 95% probability intervals for these estimates are shown in Table 10.

As observed in the probability density curves, the carbon estimates tend to be larger in the later sampling dates as compared with the earlier sampling dates, for all of the farms except Siguidolo. The differences between sampling dates are largest in the subsurface soil layer for Fansirakouro, Sougoumba, and Sikasso.
Table 10. Estimated Mean Carbon Concentration and 95 % Probability Interval Limits.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>0-20 cm Estimate</th>
<th>0-20 cm Lower limit</th>
<th>0-20 cm Upper limit</th>
<th>20-40 cm Estimate</th>
<th>20-40 cm Lower limit</th>
<th>20-40 cm Upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oumarbougou</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>0.29</td>
<td>0.26</td>
<td>0.33</td>
<td>0.23</td>
<td>0.21</td>
<td>0.25</td>
</tr>
<tr>
<td>2004 (subreg.)</td>
<td>0.32</td>
<td>0.26</td>
<td>0.38</td>
<td>0.25</td>
<td>0.22</td>
<td>0.28</td>
</tr>
<tr>
<td>2004</td>
<td>0.35</td>
<td>0.31</td>
<td>0.40</td>
<td>0.25</td>
<td>0.23</td>
<td>0.28</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>0.66</td>
<td>0.55</td>
<td>0.76</td>
<td>0.52</td>
<td>0.43</td>
<td>0.60</td>
</tr>
<tr>
<td>2004</td>
<td>0.70</td>
<td>0.60</td>
<td>0.81</td>
<td>0.63</td>
<td>0.54</td>
<td>0.71</td>
</tr>
<tr>
<td>Sougoumba</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.28</td>
<td>0.25</td>
<td>0.31</td>
<td>0.24</td>
<td>0.22</td>
<td>0.27</td>
</tr>
<tr>
<td>2002</td>
<td>0.30</td>
<td>0.28</td>
<td>0.33</td>
<td>0.26</td>
<td>0.23</td>
<td>0.28</td>
</tr>
<tr>
<td>2004</td>
<td>0.31</td>
<td>0.28</td>
<td>0.34</td>
<td>0.29</td>
<td>0.27</td>
<td>0.32</td>
</tr>
<tr>
<td>Sikasso</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.46</td>
<td>0.37</td>
<td>0.54</td>
<td>0.34</td>
<td>0.27</td>
<td>0.41</td>
</tr>
<tr>
<td>2004</td>
<td>0.50</td>
<td>0.48</td>
<td>0.53</td>
<td>0.43</td>
<td>0.41</td>
<td>0.45</td>
</tr>
<tr>
<td>Siguidolo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.26</td>
<td>0.19</td>
<td>0.32</td>
<td>0.27</td>
<td>0.22</td>
<td>0.32</td>
</tr>
<tr>
<td>2004</td>
<td>0.22</td>
<td>0.19</td>
<td>0.24</td>
<td>0.24</td>
<td>0.22</td>
<td>0.26</td>
</tr>
</tbody>
</table>

† Estimates of mean carbon concentration (by weight of soil) made by the spatial method. Lower and upper limits of the 95% probability interval are the 0.025 and 0.975 quantiles, respectively, of the distributions of simulated spatial means.
Total Amounts of Carbon

Estimates of the total amount of carbon in the specified soil layer for each of the datasets are presented in Table 11 for four of the farms. The estimates are based on the measurements of bulk density given in Table 3, and do not include assessment of the variability of bulk density values in the sampled regions. Total amounts of carbon in tonnes (Mg) were calculated as in Equation [1]: Area (m$^2$) $\times$ 0.2 (m) Depth $\times$ Bulk density (Mg m$^{-3}$) $\times$ $\%$ C / 100, where $\%$ C are the estimated carbon concentrations.

The total amount of carbon in each farm obviously depends on the size of the farm. The smallest estimate is 15 tonnes for the Oumarbougou 2002, 20- to 40-cm layer. The largest estimate is 548 tonnes of carbon for the Sikasso 2004, 0- to 20-cm layer. Increases in tonnages over the sampling intervals for the two soil layers combined are: 4 tonnes for Oumarbougou (for the subregion that was sampled in both 2002 and 2004), 24 tonnes for Fansirakouro, and 154 tonnes for Sikasso. The change for Siguidolo is a decrease of 42 tonnes.

It is not one of my objectives to make time comparisons between the sampling dates. Changes in carbon may be explained by changes in the sampling strategies. However the values in Table 11 may be used to illustrate how the probability interval limits may be used to make time comparisons. Significant change entails that the upper limit for the earlier sampling date does not overlap the lower limit for the later sampling date. The amount of change that can be considered significant is the difference between these two. Among the values in Table 11, assuming a 95 $\%$ probability level and that the data are not biased by preferential sampling, significant change occurs only in Sikasso in the 20- to 40-cm soil layer. The amount of change that is significant at the 95 $\%$ level is 7
tonnes, the difference between the lower limit of the 2004 estimate, 435 tonnes, and the upper limit of the 2000 estimate, 428 tonnes. A much larger amount would be significant if the earlier dataset was as precise as the later one (see Fig. 56, bottom graph).

The estimates converted to tonnes carbon per hectare (Mg ha\(^{-1}\)) are presented in Table 12, allowing comparisons among the farms. At the smallest, the 0- to 20-cm soil layer in Siguidolo in 2004 contained an estimated 7.2 tonnes carbon per hectare. At the largest, the 0- to 20-cm soil layer in Fansirakouro contained an estimated 21.5 tonnes carbon per hectare.
Table 11. Total Estimates of Carbon and 95% Probability Interval Limits.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Total estimated C †</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-20 cm</td>
<td>20-40 cm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Estimate</td>
<td>Lower limit</td>
<td>Upper limit</td>
<td>Estimate</td>
<td>Lower limit</td>
<td>Upper limit</td>
<td>Estimate</td>
<td>Lower limit</td>
</tr>
<tr>
<td>Oumarbougou</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>19</td>
<td>17</td>
<td>22</td>
<td>15</td>
<td>13</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004 (subregion)</td>
<td>22</td>
<td>18</td>
<td>25</td>
<td>16</td>
<td>14</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>50</td>
<td>44</td>
<td>57</td>
<td>35</td>
<td>32</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fansirakouro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>96</td>
<td>80</td>
<td>112</td>
<td>82</td>
<td>68</td>
<td>95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>102</td>
<td>87</td>
<td>118</td>
<td>100</td>
<td>86</td>
<td>112</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sikasso</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>497</td>
<td>406</td>
<td>587</td>
<td>352</td>
<td>283</td>
<td>428</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>548</td>
<td>521</td>
<td>574</td>
<td>455</td>
<td>435</td>
<td>474</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Siguidolo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>136</td>
<td>102</td>
<td>167</td>
<td>143</td>
<td>117</td>
<td>169</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>112</td>
<td>100</td>
<td>126</td>
<td>125</td>
<td>115</td>
<td>137</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† Estimated total amounts of C in Mg (tonnes) in the specified soil layer within the region perimeters. Calculated from the estimate and lower and upper limits of the 95% probability interval for estimated mean C concentration by substitution into the following equation: Area (m²) × 0.2 (m) depth × bulk density (Mg m⁻³) × % C / 100. Not calculated for Sougoumba as a soil bulk density value is not available for that farm.
Table 12. Carbon per Hectare.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Area</th>
<th>0-20 cm</th>
<th>20-40 cm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ha</td>
<td>Mg ha⁻¹</td>
<td></td>
</tr>
<tr>
<td>Oumarbougou</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>2.0</td>
<td>9.8</td>
<td>7.5</td>
</tr>
<tr>
<td>2004 (subregion)</td>
<td>2.0</td>
<td>11.0</td>
<td>8.2</td>
</tr>
<tr>
<td>2004</td>
<td>4.2</td>
<td>11.9</td>
<td>8.3</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>4.8</td>
<td>20.1</td>
<td>17.1</td>
</tr>
<tr>
<td>2004</td>
<td>4.8</td>
<td>21.5</td>
<td>21.0</td>
</tr>
<tr>
<td>Sikasso</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>33.2</td>
<td>15.0</td>
<td>10.6</td>
</tr>
<tr>
<td>2004</td>
<td>33.2</td>
<td>16.5</td>
<td>13.7</td>
</tr>
<tr>
<td>Siguidolo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>15.7</td>
<td>8.7</td>
<td>9.1</td>
</tr>
<tr>
<td>2004</td>
<td>15.7</td>
<td>7.2</td>
<td>8.0</td>
</tr>
</tbody>
</table>

† Estimated C per hectare (Mg ha⁻¹) in the specified soil layer. Calculated as: estimated total amount of carbon in the soil layer (see Table 11) divided by the size of the region in hectares. Not calculated for Sougoumba.
Spatial vs. Non-spatial Estimates

In order to test the question of whether the spatial approach improves the accuracy of estimates over the non-spatial method, we need to know the true values. Lacking that, we can state that kriging is known to decrease the influence of clustered data and provide an unbiased result. The non-spatial method is capable of delivering an unbiased estimate as well, given a random sample or other unbiased design. However, here the sampling was not random or designed to be unbiased. Under those circumstances, the non-spatial method is not guaranteed to be unbiased. Therefore, for the present study, we expect that the spatial estimates should be more accurate than the non-spatial estimates.

The spatial and non-spatial estimates are compared in Table 13. For most of the datasets, the estimates made by the two methods are close in value. The largest differences between spatial and non-spatial estimates occur in the Oumarbougou 2002 datasets; calculated from the values in the table, for that farm and year the non-spatial estimate is 14% larger than the spatial estimate in the surface layer and 13% larger in the subsurface layer. For the Sikasso 2000, 0- to 20-cm dataset, the non-spatial estimate is about 9% smaller than the spatial estimate. Both the Sikasso and the Oumarbougou datasets are characterized by having clear trends.
Table 13. Comparison of Spatial vs. Non-spatial Estimates.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Spatial</th>
<th>Non-spatial</th>
<th>Spatial</th>
<th>Non-spatial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-20 cm</td>
<td>20-40 cm</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Oumarbougou</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>0.29</td>
<td>0.33</td>
<td>0.23</td>
<td>0.26</td>
</tr>
<tr>
<td>2004</td>
<td>0.35</td>
<td>0.34</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>0.66</td>
<td>0.66</td>
<td>0.52</td>
<td>0.52</td>
</tr>
<tr>
<td>2004</td>
<td>0.70</td>
<td>0.71</td>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td>Sougoumba</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.28</td>
<td>0.29</td>
<td>0.24</td>
<td>0.25</td>
</tr>
<tr>
<td>2002</td>
<td>0.30</td>
<td>0.31</td>
<td>0.26</td>
<td>0.28</td>
</tr>
<tr>
<td>2004</td>
<td>0.31</td>
<td>0.32</td>
<td>0.29</td>
<td>0.31</td>
</tr>
<tr>
<td>Sikasso</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.46</td>
<td>0.42</td>
<td>0.34</td>
<td>0.35</td>
</tr>
<tr>
<td>2004</td>
<td>0.50</td>
<td>0.51</td>
<td>0.43</td>
<td>0.43</td>
</tr>
<tr>
<td>Siguidolo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.26</td>
<td>0.27</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>2004</td>
<td>0.22</td>
<td>0.23</td>
<td>0.24</td>
<td>0.25</td>
</tr>
</tbody>
</table>

† Spatial and non-spatial estimated mean carbon concentration (by weight of soil). The spatial estimate is the outcome of a Bayesian geostatistical algorithm and the non-spatial estimate is simply the mean of the values in the sample dataset.
Spatial vs. Non-spatial Precision

The relative precision of spatial vs. non-spatial estimates is presented in Table 14.
The measure of uncertainty for the spatial method is the width of the 95% probability interval; this is the lower limit of the probability interval subtracted from the upper limit. The non-spatial uncertainty measure is the width of the interval between the lower and upper limits of the 95% confidence interval as calculated by the t-test (interval widths are not shown). The relative precision of spatial estimates compared with non-spatial estimates is the width of the spatial interval divided by the width of the non-spatial interval.

Calculated with the values given in the table, the gain in precision of the spatial method over the non-spatial method is 16%. For all of the datasets except those for Fansirakouro 2002, Sougoumba 2002, and Siguidolo 2000, the spatial estimates have greater precision (smaller values of relative precision) than the non-spatial estimates. The distinctive aspect that the three datasets just mentioned have in common is that their sampling configurations are poorly dispersed.

The largest improvements in precision for the spatial method over the non-spatial method are for the Oumarbougou 2002 datasets; the improvement for the 0- to 20-cm soil layer is 45% and for the 20- to 40-cm layer the improvement is 36%. The Sikasso 2000, 0- to 20-cm dataset improved 35% in precision over the non-spatial method. From our previous discussion, the distinctive aspect of these three datasets is the presence of trends.

The relative precisions of the remaining datasets are all very similar, around 0.70 to 0.80, or an improvement of about 20 to 30%.
Table 14. Spatial Relative to Non-spatial Precision.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>0-20 cm</th>
<th>20-40 cm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>Oumarbougou</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>24, 25‡</td>
<td>0.55</td>
</tr>
<tr>
<td>2004</td>
<td>22</td>
<td>0.79</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>20,18‡</td>
<td>1.15</td>
</tr>
<tr>
<td>2004</td>
<td>21</td>
<td>0.74</td>
</tr>
<tr>
<td>Sougoumba</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>16</td>
<td>0.84</td>
</tr>
<tr>
<td>2002</td>
<td>38</td>
<td>1.13</td>
</tr>
<tr>
<td>2004</td>
<td>49</td>
<td>0.76</td>
</tr>
<tr>
<td>Sikasso</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>13</td>
<td>0.65</td>
</tr>
<tr>
<td>2004</td>
<td>75</td>
<td>0.71</td>
</tr>
<tr>
<td>Siguidolo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>14,13‡</td>
<td>1.19</td>
</tr>
<tr>
<td>2004</td>
<td>35</td>
<td>0.69</td>
</tr>
</tbody>
</table>

† Spatial interval width divided by non-spatial interval width. The spatial interval width is the lower limit of a 95% probability interval for the spatial estimate subtracted from the upper limit. The non-spatial interval width is the lower limit of a 95% confidence interval for the non-spatial estimate, as calculated by the t-test, subtracted from the upper limit.
‡ Number of observations in the 0-20-cm and 20-40-cm soil layers respectively.
DISCUSSION

Precision

Likely reasons for the overall improvement in precision include the presence of trends, spatial correlation, and possibly correction for skew in the original datasets. The long tails of skewed datasets inflate the variance; measures of uncertainty that depend on the variance assume a normally-shaped distribution. Poor dispersion of sampling locations may result in a precision for the non-spatial method that is underestimated; in these cases the non-spatial method may appear to do better than spatial method, although the spatial estimate of the uncertainty may be more accurate. The non-spatial method takes no account of the spatial configuration of sampling locations while the spatial method takes into account the large uncertainty of poorly sampled subregions.

All of these factors can possibly be seen in the relative precisions of the datasets in Table 14. The datasets with the smallest relative precisions have strong trends, large skew coefficients, and/or small relative nuggets. Those with relative precisions greater than 1.0 have poorly dispersed sampling locations. The number of observations appears to have no connection to the ability of the spatial method to improve the precision. Even very small datasets (for example the Sougoumba 2000 datasets) improve in precision for the spatial method over the non-spatial as long as the sampling locations are reasonably well-dispersed.

This is not to say that the small datasets have precisions as good as the large datasets, only that their precisions can be improved over the non-spatial method just as well. To compare the precisions of the final estimates made by the spatial method among the datasets, we need a measure of the uncertainty that is standardized. In Table 15 the
confidence or probability interval widths have been divided by two and standardized by the estimates to yield a measure which can be called “error”. The uncertainty in the estimates can now be expressed as plus or minus the error (as a percentage of the estimate). The larger datasets have smaller errors.

Apparently, the size of the dataset is more important than the method of analysis. For example, the most precise dataset among all of the farms is the Sikasso 2004, 20- to 40-cm dataset. The % error for this dataset is 4.3% and the number of observations is 75. If an error of less than 5% is desired, it appears that it is necessary to collect a sample size as large as 75 observations. However the non-spatial error for this dataset is still only 5.5%. With a sample size of 75 observations the precision is fairly good regardless of the method of analysis.
Table 15. % Error.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>n</th>
<th>Spatial</th>
<th>Non-spatial</th>
<th>Spatial</th>
<th>Non-spatial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-20 cm</td>
<td></td>
<td></td>
<td>20-40 cm</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
<td>---------</td>
<td>-------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Oumarbougou</td>
<td>2002</td>
<td>24,25</td>
<td>12.2</td>
<td>19.4</td>
<td>10.1</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>22</td>
<td>13.4</td>
<td>17.4</td>
<td>9.3</td>
</tr>
<tr>
<td>Fansirakouro</td>
<td>2002</td>
<td>20,18</td>
<td>16.4</td>
<td>14.0</td>
<td>17.0</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>21</td>
<td>15.5</td>
<td>20.6</td>
<td>13.0</td>
</tr>
<tr>
<td>Sougoumba</td>
<td>2000</td>
<td>16</td>
<td>10.7</td>
<td>12.7</td>
<td>11.6</td>
</tr>
<tr>
<td></td>
<td>2002</td>
<td>38</td>
<td>9.0</td>
<td>7.7</td>
<td>9.6</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>49</td>
<td>8.7</td>
<td>11.0</td>
<td>8.4</td>
</tr>
<tr>
<td>Sikasso</td>
<td>2000</td>
<td>13</td>
<td>18.2</td>
<td>30.1</td>
<td>20.6</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>75</td>
<td>4.8</td>
<td>6.8</td>
<td>4.3</td>
</tr>
<tr>
<td>Siguidolo</td>
<td>2000</td>
<td>14,13</td>
<td>24.2</td>
<td>19.6</td>
<td>18.0</td>
</tr>
<tr>
<td></td>
<td>2004</td>
<td>35</td>
<td>11.6</td>
<td>15.5</td>
<td>8.8</td>
</tr>
</tbody>
</table>

† The % error for estimates of mean C concentration was calculated as: \( \frac{(95\% \text{ confidence or probability interval width} / 2)}{\text{estimated mean}} \times 100 \).
Bayesian Analysis

The Bayesian geostatistical approach offers clear improvements over conventional geostatistics. With the Bayesian approach, there is no need to model the variogram from the graphical display, instead the model is created from all of the data. The uncertainty in the variogram parameters is included in the final estimation error. The Bayesian algorithm is an improvement over non-Bayesian simulation in a number of aspects. A known mean does not have to be specified, as it is in the case of simulation with simple kriging as the basis of the algorithm, nor corrections made to the kriging errors as in the case of simulation with ordinary kriging as the basis. In the Bayesian approach, the covariance model is estimated from all of the data, including those separated by large distances, and conditioning data do not need to be limited to specified neighborhoods. The Bayesian algorithm is able to provide "block" values for entire irregular regions and also to provide the block estimation error, without approximating covariances.

Remaining Issues

The most delicate issue in the Bayesian approach is specifying the prior distributions and discrete support points. However, the problem is not so much that these require "ad hoc" judgments as that they require knowledgeable judgments. Primarily what is required is a good understanding of the implications of the various choices of prior distributions.

The appropriateness of the Gaussian assumption and the effect of transformation on the reliability of the estimates are questions that have not been addressed here. Also,
validation of the results is not possible as this would require the "true" values to be known. In point interpolation, cross-validation is possible, but this method is not available for validating the global estimate. Some questions remain about the appropriateness of the choices of trend included in the model, although presumably if these are not ideal, the estimation error will be larger than it might be otherwise. And finally, the question remains of how large the dataset needs to be in order for the results to be reliable. Because the variogram model parameters are not estimated from the graph of the data, it is possible that the dataset size may be smaller than it would need to be otherwise. If there is no spatial correlation, or spatial correlation can not be determined from the data, then the model should return the same results as the non-spatial method, or perhaps with a larger uncertainty interval. However, this is a question that should be tested.

**Recommendations**

Validation datasets should be created to test some of the questions mentioned above and additional questions such as whether inclusion of anisotropy or more than one structure in the variogram would improve accuracy or precision. What is needed is an "exhaustive" dataset, that is one that is sampled very densely, which will provide a set of "true" values against which to test. Smaller subsamples can be drawn from the exhaustive set. Tested in this way, appropriate model choices should predict the "true" values without bias and also have the narrowest error bounds that still contain the "true values".

Sampling should be objective, consistent in design from year to year, and provide even coverage. Kriging is an "unbiased" estimator in that it reduces the influence of
clustered samples. However it can not resolve bias contributed by preferential choices of sampling locations. Consistency in sampling design and objectivity are necessary if claims of increases in carbon stocks are to be successfully defended. Uneven coverage is undesirable as it increases the estimation error and thus reduces the precision of estimates.

Increasing the number of closely-spaced pairs or the sampling intensity might define the variogram better at a shorter range of spatial correlation. However it is not clear how much improvement in precision might be gained. The extra sampling may not be worth the extra effort. The question of ideal sampling intensity and design could be explored by means of validation sampling as described above.

Information on the variability of bulk density measurements will certainly be required for carbon estimates made for the purpose of carbon trading, as this variability will contribute uncertainty to estimates. Spatial correlation between carbon and bulk density should be explored.

**Conclusions**

Improvement in precision was found for the spatial approach over the non-spatial approach, however it was a very modest improvement; for datasets for which the sampling configuration was reasonably well-dispersed, the improvement in precision was in the neighborhood of 20 – 30%. Where trends were strong, inclusion of these in the models appeared to improve the precision still further; however the relative contributions of trend, spatial correlation, and possible correction for skew to improving the precision have not been quantified here. The size of the dataset remains probably a more important
factor in the precision of estimates. We have not solved here the essential problem of the sampling variability of carbon, however, we have made an improvement.

Judging from the large uncertainty in variogram parameter estimates, spatial correlation was not strongly defined in most of the datasets, although at least some evidence of spatial correlation was evident in almost all cases. The largest datasets (75 observations) showed the best evidence of spatial correlation.

If questions of the efficiency and cost of making a spatial analysis as compared with a non-spatial analysis are set aside, the specifically Bayesian spatial approach may possibly be useful even where the dataset sizes are small and trends and spatial correlation are not well described by the data. The model includes uncertainty in the estimation of the parameters of the trend and spatial correlation. However the validity of the approach should be tested against exhaustive validation datasets. Also a cost comparison between the methods should certainly be made. It is doubtful that the improvement in precision here was worth the extra effort in collecting geographic coordinates and in making a lengthy spatial analysis.

The GIS aspects of the approach are certainly useful, however. Determining accurate geographic areas for the regions for which estimates are needed is not trivial for example. Also, the approach affords insights into the spatial behavior of carbon, both on the level of the specific farms that have been sampled, and in terms of generalizations about carbon behavior that might be made.

Questions remain about the Gaussian assumption, and the effects of data transformation and decisions of prior distributions of model parameters on outcomes; these should be further explored.
The approach can certainly be extended, resulting in much greater gains in precision. For example the correlation between sampling years should be exploited; if carbon changes in soil are slow, then the correlation between sampling dates should be strong. A simple time model would consist of extending the model to a third dimension. Correlation between values of carbon in the two soil layers could be taken advantage of in the same way (a fourth dimension?). The model can also be extended to include co-variables. Multidimensional analysis beyond the 2nd dimension and the possibility of including co-variables are currently not implemented in geoR (although it is possible to include co-variables in addition to, or in place of, the x and y coordinates as “trend” variables.) Another geostatistical package in R, gstat, does include multivariate and multidimensional analysis, however it does not presently include the Bayesian algorithm.
APPENDIX

Example R Code

R is a language and environment for statistical computing and graphics which is available as free software via the Internet site: http://www.r-project.org/. For the work presented here, I have run R on a Windows platform, but it can also be run on UNIX and MacOS platforms. A concise introduction to R (most of which is also available in the “help” function in R) is Venables et al. (2004). A more complete resource may be found in Venables and Ripley (2002), which, although specifically written for the statistics software “S”, is nevertheless mostly applicable to R.

The following code for R includes most of the basic functions mentioned in this paper, but much more can be done in R, including creation of code to batch process datasets, or create multiple-step functions or complex graphical output. The geostatistical functions mentioned here are found in the package “geoR” which must be acquired from the R site along with R itself. A command to load geoR is given at the beginning of each session involving geospatial analysis.

Load geoR package:

library(geoR)

Read documentation on krig.e.bayes function:

help(krig.e.bayes)
Import datafile into R and name it "example.data" (example is for Sikasso 2000 data):

```r
example.data<-read.table("example.datafile.txt",
                   header = T, sep = "")
```

View new R file:

```r
example.data
```

The output for which is:

<table>
<thead>
<tr>
<th></th>
<th>x</th>
<th>y</th>
<th>c020</th>
<th>c2040</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>217759.0</td>
<td>1242956</td>
<td>0.32</td>
<td>0.36</td>
</tr>
<tr>
<td>2</td>
<td>217666.1</td>
<td>1242971</td>
<td>0.27</td>
<td>0.32</td>
</tr>
<tr>
<td>3</td>
<td>217554.3</td>
<td>1242878</td>
<td>0.44</td>
<td>0.24</td>
</tr>
<tr>
<td>4</td>
<td>217752.0</td>
<td>1243303</td>
<td>0.31</td>
<td>0.30</td>
</tr>
<tr>
<td>5</td>
<td>217656.3</td>
<td>1243381</td>
<td>0.45</td>
<td>0.17</td>
</tr>
<tr>
<td>6</td>
<td>217573.7</td>
<td>1243255</td>
<td>0.75</td>
<td>0.25</td>
</tr>
<tr>
<td>7</td>
<td>217581.1</td>
<td>1243135</td>
<td>0.30</td>
<td>0.26</td>
</tr>
<tr>
<td>8</td>
<td>217798.2</td>
<td>1242850</td>
<td>0.23</td>
<td>0.48</td>
</tr>
<tr>
<td>9</td>
<td>217723.6</td>
<td>1242865</td>
<td>0.30</td>
<td>0.24</td>
</tr>
<tr>
<td>10</td>
<td>217650.9</td>
<td>1242865</td>
<td>0.14</td>
<td>0.19</td>
</tr>
<tr>
<td>11</td>
<td>217516.2</td>
<td>1243545</td>
<td>0.54</td>
<td>0.57</td>
</tr>
<tr>
<td>12</td>
<td>217460.6</td>
<td>1243544</td>
<td>0.58</td>
<td>0.63</td>
</tr>
<tr>
<td>13</td>
<td>217375.0</td>
<td>1243532</td>
<td>0.86</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Create "geodata" file:

```r
example.geodata<-as.geodata(example.data, data.col = 3:4)
```

Read in polygon (boundary) datafile:

```r
example.poly<-read.table("example.polygon.txt",
                   header = T, sep = "")
```
Create prediction grid (for grid cell size 20):

```r
e.example.grid<-expand.grid(seq(min(example.grid$x), max(example.grid$x), 20), seq(min(example.grid$y), max(example.grid$y), 20))
```

Limit prediction grid to nodes within boundary polygon:

```r
e.example.grid.poly<-locations.inside(e.example.grid, e.example.poly)
```

Set model control options ( lambda = 1 is no transformation, lambda = 0 is log transformation):

**Trend in x, no transformation:**

```r
mc<-model.control(trend.d=example.geodata$coords[,1],
                  trend.l=example.grid.poly[,1], lambda=1)
```

**Trend in y, no transformation:**

```r
mc<-model.control(trend.d=example.geodata$coords[,2],
                  trend.l=example.grid.poly[,2], lambda=1)
```

**Trend in x and y, no transformation:**

```r
mc<-model.control(trend.d="lst", trend.l="lst", lambda=1)
```

Define prior control for relative nugget:

```r
pc<-prior.control(tausq.rel.prior="recip",
                  tausq.rel.discrete= seq(0, 10, l=51))
```
Run krige.bayes function (for the data in column 3 in the example above). To save simulations leave out "output = output.control (mean.var = T)"

```r
example.kbayes<-krige.bayes(example.geodata, 
    data = example.geodata$data[,1], model = mc, prior = pc, 
    locations = example.grid, 
    output = output.control(mean.var = T), 
    borders = example.poly)
```

Query krige.bayes output file for names of components:

```r
names(example.kbayes)
```

The output for which is:

```
[1] "posterior" "predictive" "prior" "model" ".Random.seed"
[6] "max.dist" "call"
```

"Predictive" subcomponents:

```r
names(example.kbayes$predictive)
```

The output for which is:

```
[1] "mean" "variance" "distribution"
[4] "mean.simulations" "variance.simulations" "sim.means"
```

Expected values at nodes:

```r
example.kbayes$predictive$mean
```
Global means of simulations (1000 values) and the mean of these (global estimate):

```r
example.kbayes$predictive$sim.means
mean(example.kbayes$predictive$sim.means)
```

**Relevant quantiles of distribution of simulated means:**

```r
quantile(example.kbayes$predictive$sim.means, 0.025)
quantile(example.kbayes$predictive$sim.means, 0.975)
```

**Posterior parameter summaries:**

```r
example.kbayes$posterior$sigmasq$summary
example.kbayes$posterior$phi$summary
example.kbayes$posterior$tausq.rel$summary
example.kbayes$posterior$beta$summary
```

**Plot empirical data variogram (this is not an output of krigebayes):**

```r
plot(variog(example.geodata, data = example.geodata$data[,1],
    trend = "cte", lambda = 1))
```

**Add to above plot variogram lines from the simulations (output of krigebayes):**

```r
quant.summary <- function(x){quantile(x, prob = c(0.05, 0.5, 0.95))}
lines(example.kbayes, summ = quant.summary, ty = "1",
    lty = c(2,1,2), col = 1)
title(main = "Median variogram and 90% probability envelope")
```
Expected values map with contours:

```r
image(example.kbayes)
min <- min(example.kbayes$predictive$mean)
max <- max(example.kbayes$predictive$mean)
contour(example.kbayes,zlim = c(min,max),add = T,nlevels = 13)
```

Density plot of distribution of simulated means:

```r
plot(density(example.kbayes$predictive$sim.mean),
     main = "Density plot",
     xlab = "Simulated mean (%C)"
)
```

Map of simulated values (for simulation #1 of 1000) and contours:

```r
image(example.kbayes.sims, val = "simulation", number.col = 1)
contour(example.kbayes.sims,val = "simulation", number.col = 1,
        add = T,nlevels = 13)
```

Histogram of posterior range parameter:

```r
hist(example.kbayes$posterior$sample$phi,xlim = c(0,1800),
     ylim = c(0,175),breaks = 10,
     main = "Histogram of posterior range parameter",
     xlab = "Range parameter")
```

Classical statistics:

```r
length(example.data$c020)
var(example.data$c020)
```
sd(example.data$c020)

mean(example.data$c020)

t.test(example.data$c020)$conf.int[1]

t.test(example.data$c020)$conf.int[2]
REFERENCES


