European Journal of Soil Science

European Journal of Soil Science, October 2009, **60,** 757–769 doi: 10.1111/j.1365-2389.2009.01158.x

Spatial monitoring of a non-stationary soil property: phosphorus in a Florida water conservation area

B. P. MARCHANT[a], S. NEWMAN[b], R. CORSTANJE[a], K. R. REDDY[c], T. Z. OSBORNE[c] & R. M. LARK[a]

[a]*Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK*, [b]*Marsh Ecology Research Group, South Florida Water Management District, West Palm Beach, FL, USA*, and [c]*Soil and Water Science Dept, University of Florida, Gainsville, FL 32611, USA*

Summary

Ecological restoration plans in the Florida Everglades require detailed information about the status and change of the nutrient content of the soil. The soil total phosphorus (TP) content is of particular importance as the system is naturally P limited and the TP enrichment has led to changes in the wetland vegetation communities. One way to provide the relevant information is by geostatistical prediction from sampled data. However, conventional geostatistical models assume that properties being monitored are realizations of second-order stationary random functions. The assumption of second-order stationarity is not appropriate for soil TP in Water Conservation Area 1 (WCA-1) of the Florida Everglades because the mean and variance of soil TP are larger at sites adjacent to the canals which bound WCA-1 and deliver P to the system than at sites in the interior of the region.

We develop a novel linear mixed-model framework for spatial monitoring of a property for which this assumption is not valid. Specifically we use this non-stationary model to map the status and change of TP within WCA-1 from surveys carried out in 1991 and 2003. We fit the parameters of the model by residual maximum likelihood (REML) and compare the effectiveness of this non-stationary model with the conventional stationary model.

Conventional second-order stationary models fail to represent accurately the large uncertainty in predictions of TP adjacent to the canals. The non-stationary model predicts an invading front of P entering the interior of the region which is not evident in the predictions from the stationary model. Tests on the log-likelihood and the standardized squared prediction error of the fitted models provide further evidence in favour of the non-stationary model.

The sampling intensity required to ensure a certain precision of TP predictions varies across WCA-1 with the variance of TP. Therefore we apply a spatial simulated annealing optimization algorithm to design future monitoring surveys based upon our non-stationary model which ensure that the status and change are efficiently and effectively predicted across the region.

Introduction

Spatial monitoring of non-stationary soil properties

The standard geostatistical approach to mapping the change that occurs between surveys of non-colocated observations assumes that the surveys are realizations of coregionalized and second-order stationary random functions (Papritz & Flühler, 1994). We denote the random functions at each time as $Z_1(\mathbf{x})$ and $Z_2(\mathbf{x})$, where $\mathbf{x}$ denotes location. Random functions $Z_i(\mathbf{x})$ $i = 1$, 2 are jointly second-order stationary if the mean value of each function is constant across the study region, i.e.

$$\mathrm{E}\,[Z_i(\mathbf{x}) - Z_i(\mathbf{x}+\mathbf{h})] = 0, \quad \forall \boldsymbol{x}, \tag{1}$$

Correspondence: B. P. Marchant. E-mail: ben.marchant@bbsrc.ac.uk
Received 7 August 2008; accepted after revision 23 January 2009

and if the auto- and cross-variograms are functions of lag $\mathbf{h}$ only, i.e.

$$\mathrm{E}\,[\{Z_i(\mathbf{x}) - Z_i(\mathbf{x}+\mathbf{h})\}\{Z_j(\mathbf{x}) - Z_j(\mathbf{x}+\mathbf{h})\}] = 2\gamma_{ij}(\mathbf{h}), \ \forall \boldsymbol{x}, \tag{2}$$

where $i = 1, 2$; $j = 1, 2$; $\mathbf{h}$ is a lag vector that represents the separation in space between a pair of observation sites and $\gamma_{ij}(\mathbf{h})$ is the auto-variogram of $Z_i(\mathbf{x})$ if $i = j$ or the cross-variogram of $Z_i(\mathbf{x})$ and $Z_j(\mathbf{x})$ if $i \neq j$.

Papritz & Flühler (1994) predict the change in the property by cokriging the temporal differences $Z_2(\mathbf{x}) - Z_1(\mathbf{x})$ where the coregionalization for $Z_1(\mathbf{x})$ and $Z_2(\mathbf{x})$ is represented by a linear model (LMCR). They fit the LMCR to point estimates of the

auto- and cross-variograms — or pseudo cross-variogram (Clark *et al*., 1989; Myers, 1991) if the surveys do not observe the property at the same sites — of $Z_1(\mathbf{x})$ and $Z_2(\mathbf{x})$ under constraints which ensure that the coregionalization structure is positive definite. Fitting such a model is not trivial since the number of variogram parameters is greater than in the univariate case and the precision of the point estimates vary over different lags and at different times.

Stationarity is a property of the random functions $Z_i(\mathbf{x})$ and not of the observed data. However, exploratory analysis of the observed data can indicate that the assumption of stationarity is not plausible. For example, initial analyses of surveys of total phosphorus (TP) content in Water Conservation Area 1 (WCA-1) suggest that both conditions of second-order stationarity may be invalid. The assumption of a stationary mean (Equation 1) can be relaxed if the soil property is represented by a linear mixed model (Lark *et al*., 2006). This allows the mean of each $Z_i(\mathbf{x})$ to vary across the study region or according to an auxiliary property. The parameters of the linear mixed model may be fitted with minimum bias by residual maximum likelihood (REML). Marchant & Lark (2007) demonstrated that LMCRs could be included in a linear mixed model and the parameters may be fitted by REML. The maximization of the residual likelihood function is a rigorous statistical criterion for fitting a LMCR which fully accounts for the spatial and temporal correlations between observations.

Another advantage of REML is that because it is model-based it allows appropriate parametric variance models to be fitted in which the assumption of stationarity in the variance (Equation 2) can be relaxed. Lark (in press) explored simple methods of modelling the spatial structure of a single property with a variogram that varied along a transect. He assumed that the spatial correlation of the property was fixed across the transect but the variance varied. He divided the transect into two sub-transects along which the property had different constant variances. The parameter defining the boundary between the two sub-transects was fitted by REML along with the other variogram parameters. In this paper we extend this approach to include two-dimensional study regions and multiple spatially correlated properties via the LMCR. Thus the resulting non-stationary linear mixed model is suitable for the general problem of monitoring properties for which neither stationarity assumption is valid.

Monitoring soil properties in the Florida Everglades

Effective ecosystem restoration to mitigate environmental degradation requires a sound and thorough survey of the state of the ecosystem prior to the implementation of the restoration plan, and a continual monitoring of the system as the restoration measures take effect.

The Florida Everglades is an example of a system in which the ecosystem degradation due to hydrological and nutrient impacts has led to the implementation of large-scale and long-term restoration plans. Restoration plans in the Florida Everglades are designed to maintain and restore characteristic landscape features such as the patterns of indigenous wetland vegetation communities (e.g. tree islands, ridge and slough communities). Two primary drivers in establishing the indigenous Everglades ecosystems are hydrology and soil nutrient content. The Everglades is historically a nutrient-limited system, particularly P-limited, and the ecosystem has evolved within these constraints (Davis, 1991). Nutrient influxes from neighbouring agriculture have increased the soil P content. Where the concentration of soil P has exceeded 450–500 mg kg^{-1} the result is a change in composition of plant communities with a shift to a predominance of cattail (*Typha* spp.). These visible changes in the plant community composition have been linked to changes in the soil nutrient status.

Restoration consists of diversion or treatment of the agricultural runoff and restoration of the historical hydrological patterns. As part of the restoration, the baseline spatial patterns of soil nutrient content need to be established. This has been done repeatedly for sections of the Everglades such as WCA-1 (Newman *et al*., 1997; Corstanje *et al*., 2006) and WCA-2a (DeBusk *et al*., 2001; Rivero *et al*., 2007). All of these studies assumed that the soil properties being mapped were second-order stationary. This study is concerned with mapping the change in soil TP within WCA-1 (Figure 1) based upon surveys (Figure 2) carried out in 1991 (Newman *et al*., 1997) and 2003 (Corstanje *et al*., 2006). Phosphorus is known to be delivered to WCA-1 by the canals which bound the area, leading to different patterns of TP variation adjacent to these canals from those in the interior of the region.

Monitoring TP within a non-stationary framework

The different patterns of TP variation adjacent to the bounding canals and within the interior of WCA-1 invalidate the assumption of stationarity. Therefore we model TP within our novel non-stationary framework in a manner that accounts for variation in the means of $Z_1(\mathbf{x})$ and $Z_2(\mathbf{x})$ and the auto- and cross-variograms of $Z_1(\mathbf{x})$ and $Z_2(\mathbf{x})$. Unlike Lark's (in press) case study the observations of TP are made within a two-dimensional region. Therefore it is more complicated to divide this region into sub-regions within which the variance of TP is stationary. We compare two approaches. The first divides the region into two sub-regions based upon the shortest distance to a bounding canal. The second approach divides the region into two sub-regions based on auxiliary information: specifically, an interpolated map of *Typha* spp. coverage. In both cases a parameter defining the boundary between the two sub-regions is fitted by REML.

We use the estimated model in the best linear unbiased predictor (BLUP) to predict the concentration of TP at unsampled sites on each survey date and to predict the change in TP concentration at these sites and the corresponding prediction variances. We compare these predictions with those from a second-order stationary model in terms of statistical diagnostics and properties of the predicted maps.

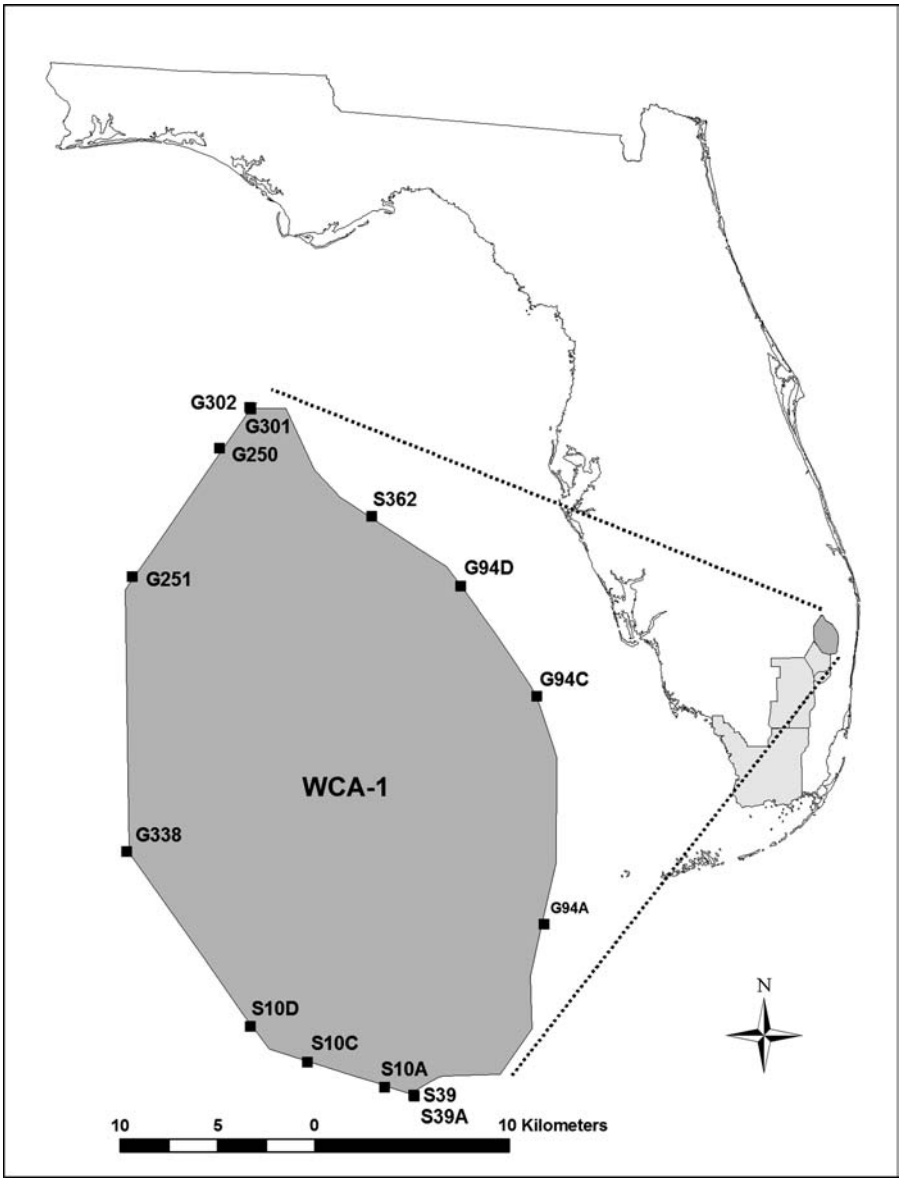

Figure 1 Map of Florida indicating position of WCA-1.

Future TP surveys will be required to assess the effectiveness of restoration strategies. The optimal configuration of sampling locations for these surveys will depend upon the coregionalization structure of TP. We assume that this structure will remain broadly similar to the model fitted here and use this model within a spatial simulated annealing algorithm (van Groenigen *et al*., 1999) to select the configuration of sampling locations which minimize the mean root prediction variance across WCA-1.

In summary, the aim of this paper is to devise a linear mixed model framework to represent the spatial variation of properties such as TP in WCA-1 which are non-stationary in the mean and/or variance. This model may then be used to predict the property at unsampled locations or to optimize the location of observations in future surveys.

Theory

Geostatistical prediction of non-stationary variables

It is generally well understood that the assumption of stationarity in the mean of a soil property can be relaxed by representing the property as a linear mixed model

$$\mathbf{z} = \mathbf{M}\boldsymbol{\beta} + \boldsymbol{\eta}, \tag{3}$$

where $\mathbf{z}$ is a vector of n observations of the property, $\mathbf{M}$ is an $n \times p$ design matrix containing values of p auxiliary variables or the fixed effects, $\boldsymbol{\beta}$ is the length p vector containing the coefficients of the fixed effects and $\boldsymbol{\eta} \sim \mathcal{N}(0, \mathbf{V})$ is a vector of spatially correlated and normally distributed random residuals with covariance matrix $\mathbf{V}$. In this paper we will assume that

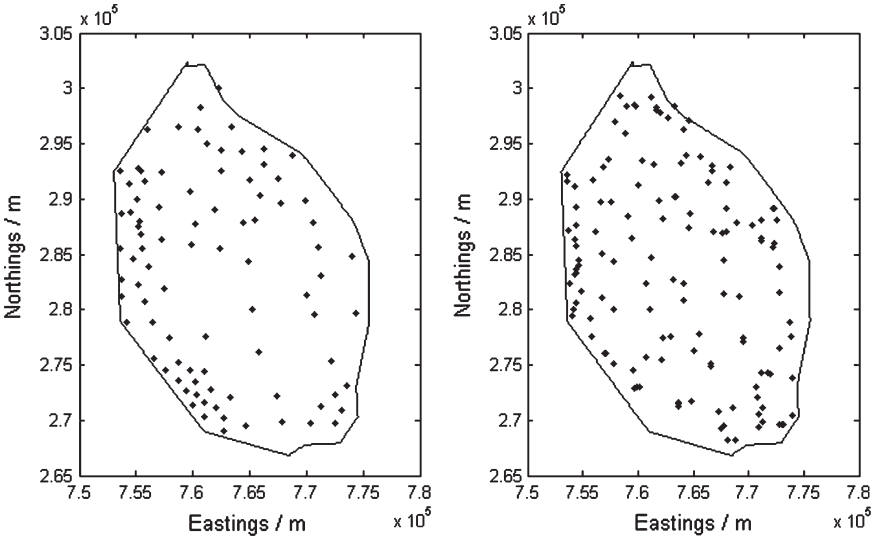

Figure 2 Locations of observations in 1991 (left) and 2003 (right).

the variogram of the residuals is isotropic (i.e. is a function of the length of lag vector **h** but not the direction) and may be modelled by a nested nugget and exponential model,

$$\gamma(h) = c_0 + c_1 \exp\left(\frac{-h}{a}\right), \tag{4}$$

where h is the length of **h**, c_0 is the nugget variance, c_1 the partial sill variance and a a distance parameter. Under this assumption the residuals are second-order stationary, and the covariance function exists and can be obtained from the variogram parameters, $\boldsymbol{\alpha} = (c_0, c_1, a)$.

If $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are estimated by the method of moments then the $\boldsymbol{\alpha}$ estimate is biased due to the uncertainty in estimates of $\boldsymbol{\beta}$. This bias may be reduced if the parameters are estimated by REML (Patterson & Thompson, 1971) and this approach is becoming more widespread among environmental scientists (e.g. Lark *et al.*, 2006). The method reduces bias in the estimated variogram parameters by projecting **z** into a residual space where all the fixed effects have zero expectation. This projection, denoted **S**, is a function of the design matrix **M** with $\mathbf{SM} = 0$ and $\text{Rank}(\mathbf{S}) = n - \text{Rank}(\mathbf{M})$. The negative log-likelihood of the projected data **Sz** is

$$l_R(\boldsymbol{\alpha}, \mathbf{M}, \mathbf{Sz}) = C(\mathbf{M}) - \frac{1}{2}\ln|\mathbf{V}| - \frac{1}{2}\ln|\mathbf{M}^T\mathbf{V}^{-1}\mathbf{M}| - \frac{1}{2}(\mathbf{z} - \mathbf{M}\hat{\boldsymbol{\beta}})^T\mathbf{V}^{-1}(\mathbf{z} - \mathbf{M}\hat{\boldsymbol{\beta}}), \tag{5}$$

where $C(\mathbf{M}) = \frac{1}{2}\{n - \text{Rank}(\mathbf{M})\}\ln(2\pi) - \ln|\mathbf{M}\mathbf{M}^T|$ and element (r, s) of **V** is given by

$$\mathbf{V}(r, s) = c_0 + c_1 - \gamma(h_{rs}), \tag{6}$$

for observations r and s separated by h_{rs} when $\gamma(h)$ is a bounded (i.e. second-order stationary) variogram function. A numerical algorithm is used to find the $\boldsymbol{\alpha}$ vector which minimizes Equation (5). For each candidate $\boldsymbol{\alpha}$ the covariance matrix **V** is used to estimate $\boldsymbol{\beta}$ by generalized least squares

$$\hat{\boldsymbol{\beta}} = (\mathbf{M}^T\mathbf{V}^{-1}\mathbf{M})^{-1}\mathbf{M}^T\mathbf{V}^{-1}\mathbf{z}. \tag{7}$$

Having fitted a linear mixed model it can then be substituted into the BLUP to form the empirical best linear unbiased predictor (EBLUP) and predict **z** at unsampled locations. The BLUP including fixed effects is often referred to as universal kriging (Webster & Oliver, 2007). If $\mathbf{x}_p$ is a vector of N unsampled target points then the EBLUP estimate for **z** is

$$\hat{\mathbf{z}}(\mathbf{x}_p) = (\mathbf{M}_p - \mathbf{V}_{po}\mathbf{V}^{-1}\mathbf{M})\hat{\boldsymbol{\beta}} + \mathbf{V}_{po}\mathbf{V}^{-1}\mathbf{z}, \tag{8}$$

where $\mathbf{M}_p$ is the design matrix for the prediction sites and $\mathbf{V}_{po}$ is the covariance matrix of the observed **z** with the values at the target sites.

The covariance matrix of the prediction errors is given by

$$\mathbf{C} = (\mathbf{M}_p - \mathbf{V}_{po}\mathbf{V}^{-1}\mathbf{M})\mathbf{G}^{-1}(\mathbf{M}_p - \mathbf{V}_{po}\mathbf{V}^{-1}\mathbf{M})^T + \mathbf{V}_{pp} - \mathbf{V}_{po}\mathbf{V}^{-1}\mathbf{V}_{po}^T, \tag{9}$$

where $\mathbf{G} = \mathbf{M}^T\mathbf{V}^{-1}\mathbf{M}$ and $\mathbf{V}_{pp}$ is the covariance matrix for **z** at the target sites. Elements of $\mathbf{V}_{pp}$ and $\mathbf{V}_{po}$ are calculated in the same manner as **V** (Equation 6).

Recently, Lark (in press) proposed that this approach could be extended to properties where the residuals are non-stationary in the variance by scaling the elements of a second-order stationary covariance matrix **Q** such that $\mathbf{V} = \mathbf{RQR}$ for diagonal matrix **R**. Such a covariance matrix **V** is positive definite if **Q** is positive definite and the elements of **R** are real valued. This form of covariance matrix allows the variance of **z** to vary across the study region. Lark (in press) tested this approach on a survey of the slope of a soil surface along a transect and

demonstrated that the random variation in the slope may be effectively modelled if a boundary on the transect is fitted and the elements of **R** take one of two values depending on which side of the boundary they lie. However **R** can arise from any parametric model such as one which uses an auxiliary variable. The parameters of such a model, including the boundary position, may be fitted by REML. Similarly the EBLUP may be applied with $\mathbf{V}_{po}$ replaced by $\mathbf{R}_p\mathbf{Q}_{po}\mathbf{R}_o$ and $\mathbf{V}_{pp}$ replaced by $\mathbf{R}_p\mathbf{Q}_{pp}\mathbf{R}_p$ in Equations (8) and (9).

Marchant & Lark (2007) showed that the REML-EBLUP can be extended to N_v coregionalized properties by including observations of the N_v properties in the **z** vector and predictions of the N_v properties in the $\hat{\mathbf{z}}$ vector. The corresponding covariance matrices must account for the cross-correlation between the different properties. In this study we assume that each auto- and cross-variogram can be represented by a nested nugget and exponential model and we denote the parameters of auto-variogram i by $c_0^{i,i}$, $c_1^{i,i}$ and $a^{i,i}$ and the parameters of the cross-variogram between variables i and j by $c_0^{i,j}$, $c_1^{i,j}$ and $a^{i,j}$. Element (r, s) of the second-order stationary covariance matrix **V** is given by

$$\mathbf{V}_{rs} = c_0^{i,j} + c_1^{i,j} - \gamma^{i,j}(h_{rs}), \tag{10}$$

where the rth entry of **z** is an observation of $z_i(\mathbf{x})$, the sth entry of **z** is an observation of $z_j(\mathbf{x})$, and $h_{r,s}$ is the lag separating the rth and sth observations.

Sufficient conditions to ensure that the covariance structure is positive definite are that the auto- and cross-variograms are all represented by the same authorized model such as the nested exponential and nugget model (Equation 4), the spatial parameters for each auto- and cross-variogram are equal and

$$c_0^{1,2} \leq (c_0^{1,1} c_0^{2,2}), \tag{11}$$

$$c_1^{1,2} \leq (c_1^{1,1} c_1^{2,2}). \tag{12}$$

In this study the coregionalized properties are the same property measured at different times. We can predict the change in this property at site $\mathbf{x}_0$ as the difference between the EBLUP predictions at each time, i.e.

$$\hat{z}_c(\mathbf{x}_0) = \hat{z}_2(\mathbf{x}_0) - \hat{z}_1(\mathbf{x}_0), \tag{13}$$

where $\hat{z}_c$ denotes the prediction of the change and $\hat{z}_1$ and $\hat{z}_2$ denote the predictions of the property at each time which we calculate from Equation (8).

The prediction variance for the change is given by

$$\sigma_c^2(\mathbf{x}_0) = \sigma_1^2(\mathbf{x}_0) + \sigma_2^2(\mathbf{x}_0) - 2\mathrm{Cov}\left[\hat{z}_1(\mathbf{x}_0), \hat{z}_2(\mathbf{x}_0)\right], \tag{14}$$

where $\sigma_c^2(\mathbf{x}_0)$ is the prediction variance of the change, $\sigma_i^2(\mathbf{x}_0)$ is the prediction variance of $\hat{z}(\mathbf{x}_0)$ at time i and $\mathrm{Cov}[\hat{z}_i(\mathbf{x}_0), \hat{z}_j(\mathbf{x}_0)]$ is the prediction covariance between the predictions of z at times i and j. The $\sigma_i^2(\mathbf{x}_0)$ and $\mathrm{Cov}[\hat{z}_i(\mathbf{x}_0), \hat{z}_j(\mathbf{x}_0)]$ values can be extracted from the prediction covariance matrix **C** (Equation 9).

In this paper we further extend this approach by using the REML-EBLUP to predict the change in properties with non-stationary residuals by assuming that the covariance matrix may be written $\mathbf{V} = \mathbf{RQR}$ for second-order stationary covariance matrix **Q** and diagonal matrix **R**.

Selection of covariance models and fixed effects

Once a general linear mixed model has been fitted to the data we can decide whether each covariance parameter or fixed effect within the model is necessary by comparing the full fitted model with nested sub-models by means of likelihood ratio tests. Details of these tests are given in the Appendix.

Cross-validation of models

We cross-validate our models via the standardized square prediction error

$$\boldsymbol{\theta}(\mathbf{x}_i) = \frac{\left\{z(\mathbf{x}_i) - \hat{z}^{-}(\mathbf{x}_i)\right\}^2}{\sigma^2(\mathbf{x})}, \tag{15}$$

where $\hat{z}^{-}(\mathbf{x}_i)$ is the EBLUP prediction of z at $\mathbf{x}_i$ when $z(\mathbf{x}_i)$ is removed from the observation vector **z**. The models under investigation are fitted to the full observation vector. If the prediction variance correctly describes the uncertainty in the EBLUP then $\bar{\boldsymbol{\theta}}$, the expectation of $\boldsymbol{\theta}$ over all observation locations, is 1.0 and assuming normally distributed prediction errors, the expected value of $\tilde{\boldsymbol{\theta}}$, the median of $\boldsymbol{\theta}$, is 0.455. Lark (2000) suggests that $\tilde{\boldsymbol{\theta}}$ should be used to validate fitted models because it is more robust to outliers that may be present. We can deduce confidence limits on $\tilde{\boldsymbol{\theta}}$ by simulating multiple realizations of the fitted linear mixed model by Cholesky factorization (Deutsch & Journel, 1998) and then calculating $\tilde{\boldsymbol{\theta}}$ for each realization.

Optimization of sampling schemes

Van Groenigen *et al.* (1999) suggested that if the covariance structure of a property is known then the configuration of the sampling sites for prediction of this property can be optimized by spatial simulated annealing (SSA). Spatial simulated annealing is a stochastic algorithm which finds the elements of **X** which minimise some objective function $\phi(\mathbf{X})$. In this context **X** contains the sampling locations and $\phi(\mathbf{X})$ is the mean of the root prediction variance across the study region when observations are made at locations **X**. Van Groenigen *et al.* (1999) found that for second-order stationary variables such an algorithm led to sampling locations being evenly spread across the study region. Brus & Heuvelink (2007) explored how optimal sampling schemes changed when fixed effects were included.

Methods

The survey

The original Everglades was a contiguous system in which water flowed north to south by continuous sheet flow (Parker, 1974). The ecosystem remnant is reduced in size and fragmented into managed impoundments (water conservation areas) and the Everglades National Park, with water moving north to south through a series of canals. Water Conservation Area 1 is the northernmost conservation area. Figure 1 shows a map of WCA-1. It is bounded by a 92-km levée and an internal perimeter canal. This area forms part of the Loxahatchee National Wildlife Refuge (LNWR) and contains much of the ecosystem features desired under the Greater Everglades restoration plans such as wet prairies, sloughs and sawgrass communities and tree islands. As of 1991, the cattail areas had increased to cover more than 10% of the area from $< 1\%$ surveyed in the 1960s. For a more detailed description on the LNWR we refer the reader to Corstanje *et al*. (2006) or Newman *et al*. (1997).

The spatial distribution of nutrients within WCA-1 is greatly influenced by the pattern of water flow through and around the area. The agricultural runoff water from the Everglades Agricultural Area (EAA) was previously pumped into the water conservation area. The EAA is located to the north and west of WCA-1 and these were the main entry points of the nutrient-laden runoff water. This water was pumped into the boundary canals, with no hydrological boundaries between them and the interior of the marsh. The overall direction of the water movement is to the south down a slight gradient of 2–3 cm km^{-1}. In addition the internal perimeter canal drains the northern end and tends to direct inflow around the area rather than through it, as the geology of the interior area is slightly dome-shaped. The introduction of the runoff water is primarily into the boundary areas of WCA-1 and less into the interior. In low-water periods, runoff water flows south through the canals and does not penetrate the marsh interior. However, the boundary region is influenced by hydraulic and nutrient loads from surrounding canals. In high-water periods there is increased pumping, and rainfall and canal water enters the interior marsh. Most of the water exits the area through a series of water control structures on the southern boundary. As a result of this dome shape, the areas of P enrichment do not conform as clearly to enrichment gradients as has been reported for the adjacent area, WCA-2a (DeBusk *et al*., 2001), but are slightly more patchy. In 2001, EAA runoff was either treated or diverted.

Three surveys of TP content in WCA-1 have been made. In 1987 Richardson *et al*. (1990) collected 100 soil samples throughout WCA-1. A second survey of 90 soil samples was made by Newman *et al*. (1997) in 1991. The samples were located where the 1987 survey suggested there was a large spatial gradient in TP. A third survey was made at 120 sites in March 2003 (Corstanje *et al*., 2006). These sites were chosen by stratified random sampling, with the strata based upon the previous surveys and historical, ecological and hydrological data. The 2003 survey also recorded the proportion of coverage by different vegetation types at each site.

The factors described above tend to lead to larger TP content in regions adjacent to the canals. In particular, Newman *et al*. (1997) and Corstanje *et al*. (2006) observed the largest TP content adjacent to the canal on the western boundary. Newman *et al*. (1997) observed that *Typha* spp. incursions tended to occur in these regions where the TP content was largest.

A non-stationary monitoring framework

We propose to model TP in WCA-1 by treating each survey as a realization of a non-stationary coregionalized random function. We express our monitoring framework as a linear mixed model

$$\mathbf{z} = \mathbf{M}\boldsymbol{\beta} + \boldsymbol{\eta},\ \boldsymbol{\eta} \sim \mathcal{N}(0, \mathbf{V} = \mathbf{RQR}), \qquad (16)$$

The vector $\mathbf{z}$ contains observations from the 1991 and 2003 surveys. The data collected in each year are strongly positively skewed (skew = 3.0 for 1991 and skew = 1.96 for 2003). Conventional geostatistical models assume that the realizations of the random function have a normal distribution and hence are not skewed. The introduction of a non-stationary mean and/or variance into the geostatistical model may account for the positive skew but motivated by a desire to use as simple a model as possible we log-transform the data prior to our analysis. After transformation of the data the skew is reduced to 1.65 for 1991 and 0.92 for 2003. We assume that for each survey the transformed TP content is non-stationary but WCA-1 can be split into two regions within which the variance of the transformed TP content is stationary. Region A is in the interior marsh where the water input is predominantly rainwater. Region B is adjacent to canals, from which it receives nutrient-enriched water. We assume that the spatial correlation across WCA-1 may be described by a nested nugget and exponential function but that the variance in Region A is scaled by a factor r_i for survey i. Thus $\mathbf{Q}$ is a second-order stationary covariance matrix and $\mathbf{R}$ is a diagonal matrix with entry r_1 if the corresponding observation is from the first survey and lies in Region A, r_2 if the corresponding observation is from the second survey and lies in Region A and 1 if the observation lies in Region B.

If the spatial extent of regions A and B are known the parameters which determine the covariance matrix $\mathbf{V}$ may be fitted by REML. However the problem of deciding the extent of each area is not trivial. The observations considered by Lark (in press) were made along a transect and he assumed that all observations made on one side of a fitted boundary were in region A and the others were in region B. This boundary parameter was fitted by REML. Our problem is more

complicated because we are investigating a two-dimensional region rather than a transect. We test two different approaches to determining each region. The first is similar to Lark's (in press) approach in that we assume that all observations within distance d^* of the bounding canal are in region B whereas the other observations are within region A. We denote the shortest distance from a site to a bounding canal by d and assume that the variance of TP for $d < d^*$ is scaled by a factor r_i.

The second approach uses auxiliary information about *Typha* spp. incursions. *Typha* spp. coverage (% coverage) was recorded at the observation sites in 2003. We fit a second-order stationary nested nugget and exponential model to these observations by REML and predict the *Typha* spp. coverage at unsampled locations by the EBLUP. The resulting map of *Typha* spp. coverage, which we denote t, is shown in Figure 3. We assume that Regions A and B are defined by a threshold t^* on the *Typha* spp. coverage and that locations with $t < t^*$ lie in Region A. We refer to these two models as the non-stationary distance and non-stationary *Typha* spp. models.

Each of these non-stationary covariance models has ten parameters. Seven of the parameters describe the LMCR, two parameters describe the scaling of the variance in the different regions for each survey and one parameter describes the extent of Region B. These parameters may be fitted by REML but minimization of the negative log residual likelihood function (Equation 5) is computationally demanding due to the highly nonlinear relationships between the parameters and the negative log residual likelihood function and due to the constraints on the parameters which ensure that **Q** is positive definite. Here we apply a stochastic minimization algorithm called simulated annealing (Pardo-Igúzquiza, 1997). However the complexity of the problem means that this does not always converge to the global minimum and therefore we apply the algorithm ten times and use the best solution.

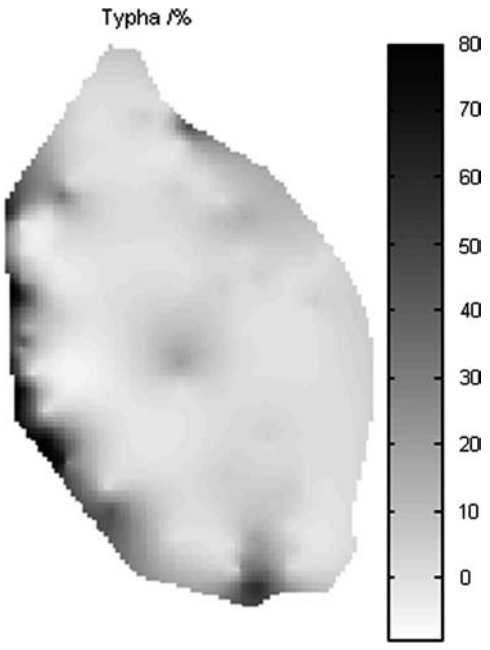

Figure 3 Interpolated map of *Typha* spp. coverage within WCA-1 in 2003.

Table 1 Log likelihood tests for models with subregions determined by distance to boundary

Comparison	d. of f.[a]	D	D_{crit} ($P = 0.01$)[b]
M1 v M2	2	13.56	9.21
M4 v M5	2	0.10	9.21
M2 v M5	sim[c]	29.83	10.43

[a]d. of f.: number of degrees of freedom.
[b]D_{crit} ($P = 0.01$): the critical value of D at $P = 0.01$ level.
[c]sim: critical value of D has been determined by simulation.

We may also expect the mean value of TP to be different within the two regions. It is not possible to fit a model for which the fixed effects vary with t^* or d^* by REML unless t^* or d^* is known. Otherwise the projection of the data **S** would vary with parameters that are being fitted. Thus we only fit d^* and t^* for models which are stationary in the mean. We then assume the same values for models which are non-stationary in the mean. It seems unrealistic for the mean TP to jump abruptly at the boundary between Regions A and B so when the mean is non-stationary we assume that it is stationary in Region A, that the mean at the boundary between Regions A and B is equal to the mean within Region A and that the mean increases linearly within Region B with decreasing distance to a canal. Specifically each row of the non-stationary design matrix **M** for the distance model is equal to

$$\begin{array}{llll} [1 & 0 & d_A & 0] \quad \text{if } i = 1 \text{ and } d < d^*, \\ [1 & 0 & 0 & 0] \quad \text{if } i = 1 \text{ and } d \geq d^*, \\ [0 & 1 & 0 & d_A] \quad \text{if } i = 2 \text{ and } d < d^*, \\ [0 & 1 & 0 & 0] \quad \text{if } i = 2 \text{ and } d \geq d^*, \end{array} \tag{17}$$

where $i = 1$ if the observation was made in 1991 and $i = 2$ if the observation was made in 2003 and d_A is the shortest distance from the observation to a site in Region A. The design

Table 2 Log likelihood tests for models with subregions determined by *Typha* spp. proportion

Comparison	d. of f.[a]	D	D_{crit} ($P = 0.01$)[b]
M1 v M2	2	2.29	9.21
M4 v M5	2	11.19	9.21
M2 v M5	sim[c]	37.54	9.80

[a]d. of f.: number of degrees of freedom.
[b]D_{crit} ($P = 0.01$): the critical value of D at $P = 0.01$ level.
[c]sim: critical value of D has been determined by simulation.

Table 3 Fitted linear mixed model parameters for log(TP) using the distance approach

Model	M1	M2	M3	M4	M5	M6
$c_0^{1,1}/(\log(\text{mg kg}^{-1}))^2$	0.22	0.28	0.20	0.08	0.09	0.08
$c_1^{1,1}/(\log(\text{mg kg}^{-1}))^2$	0.14	0.25	0.12	0.02	0.09	0.09
$c_0^{2,2}/(\log(\text{mg kg}^{-1}))^2$	0.09	0.15	0.09	0.06	0.05	0.05
$c_1^{2,2}/(\log(\text{mg kg}^{-1}))^2$	0.08	0.26	0.07	0.04	0.15	0.15
$c_0^{1,2}/(\log(\text{mg kg}^{-1}))^2$	0.01	0.03	0.02	0.05	0.00	0.01
$c_1^{1,2}/(\log(\text{mg kg}^{-1}))^2$	0.09	0.24	0.08	0.024	0.12	0.11
a /m	5 535	5 535	5 535	5 535	5 535	5 535
r_1	0.43	0.37	0.45	1.00	1.00	1.00
r_2	0.70	0.50	0.71	1.00	1.00	1.00
l_R	23.39	26.97	13.41	42.79	56.81	56.71
$\bar{\boldsymbol{\theta}}$	0.99	0.99		1.01	0.98	
$\tilde{\boldsymbol{\theta}}$	0.39	0.34		0.33	0.29	
$\tilde{\boldsymbol{\theta}}^{95-a}$	0.32	0.33		0.33	0.33	
$\tilde{\boldsymbol{\theta}}^{95+b}$	0.62	0.63		0.62	0.63	

[a] $\tilde{\boldsymbol{\theta}}^{95-}$: lower boundary of 95% confidence interval of $\tilde{\boldsymbol{\theta}}$.
[b] $\tilde{\boldsymbol{\theta}}^{95+}$: upper boundary of 95% confidence interval of $\tilde{\boldsymbol{\theta}}$.

matrix for the *Typha* spp. model is identical with the threshold d^* replaced by t^*. Each of these design matrices can be partitioned into $\mathbf{M} = [\mathbf{M}_0\ \mathbf{M}_1]$ where $\mathbf{M}_0$ is the stationary mean model, i.e. the first two columns of $\mathbf{M}$. Projection $\mathbf{S}$ corresponds to fixed effects matrix $\mathbf{M}$ and projection $\mathbf{S}_0$ corresponds to fixed effects matrix $\mathbf{M}_0$.

Thus for each approach we have designed a linear mixed model with non-stationary mean and variance. We refer to this model as M1 and we estimate its parameters by minimizing the likelihood function $l_R(\boldsymbol{\alpha}^F, \mathbf{M}, \mathbf{Sz})$. Here $\boldsymbol{\alpha}^F$ denotes that the threshold parameter (d^* or t^*) in $\boldsymbol{\alpha}$ is fixed since it cannot be estimated for a non-stationary mean. The fixed value of the threshold parameter is that fitted to a model that is stationary in the mean and non-stationary in the variance (M2 below). In the interests of parsimony we wish to confirm that non-stationary in both the mean and variance is required. Therefore we compare M1 with M2, a model where the mean is stationary but the variance is non-stationary. This model is fitted by minimizing $l_R(\boldsymbol{\alpha}, \mathbf{M}_0, \boldsymbol{S}_0\mathbf{z})$. This comparison is achieved using the likelihood ratio statistic defined in Equation (21) which requires us to fit model M3 with likelihood function $l_R(\boldsymbol{\alpha}^F, \mathbf{M}_0, \mathbf{Sz})$. Furthermore, we compare model M2 with a model that is stationary in both the mean and the variance which we denote by M5 and fit by minimizing $l_R(\boldsymbol{\alpha}_0, \mathbf{M}_0, \boldsymbol{S}_0\mathbf{z})$.

Table 4 Fitted linear mixed model parameters for log(TP) using the *Typha* spp. approach

Model	M1	M2	M3	M4	M5	M6
$c_0^{1,1}/(\log(\text{mgkg}^{-1}))^2$	0.13	0.15	0.13	0.09	0.09	0.08
$c_1^{1,1}/(\log(\text{mgkg}^{-1}))^2$	0.69	0.83	0.61	0.04	0.09	0.04
$c_0^{2,2}/(\log(\text{mgkg}^{-1}))^2$	0.11	0.12	0.10	0.07	0.05	0.06
$c_1^{2,2}/(\log(\text{mgkg}^{-1}))^2$	0.16	0.29	0.15	0.09	0.16	0.09
$c_0^{1,2}/(\log(\text{mgkg}^{-1}))^2$	0.01	−0.02	0.00	0.07	0.00	0.00
$c_1^{1,2}/(\log(\text{mgkg}^{-1}))^2$	0.25	0.39	0.23	0.06	0.12	0.06
a /m	3 852	5 535	3 701	5 535	5 534	5 534
r_1	0.27	0.27	0.28	1.00	1.00	1.00
r_2	0.68	0.60	0.69	1.00	1.00	1.00
l_R	35.60	19.27	16.98	62.62	56.81	45.62
$\bar{\boldsymbol{\theta}}$	1.01	1.01		1.01	0.98	
$\tilde{\boldsymbol{\theta}}$	0.36	0.40		0.27	0.29	
$\tilde{\boldsymbol{\theta}}^{95-a}$	0.31	0.32		0.30	0.30	
$\tilde{\boldsymbol{\theta}}^{95+b}$	0.60	0.65		0.60	0.62	

[a] $\tilde{\boldsymbol{\theta}}^{95-}$: lower boundary of 95% confidence interval of $\tilde{\boldsymbol{\theta}}$.
[b] $\tilde{\boldsymbol{\theta}}^{95+}$: upper boundary of 95% confidence interval of $\tilde{\boldsymbol{\theta}}$.

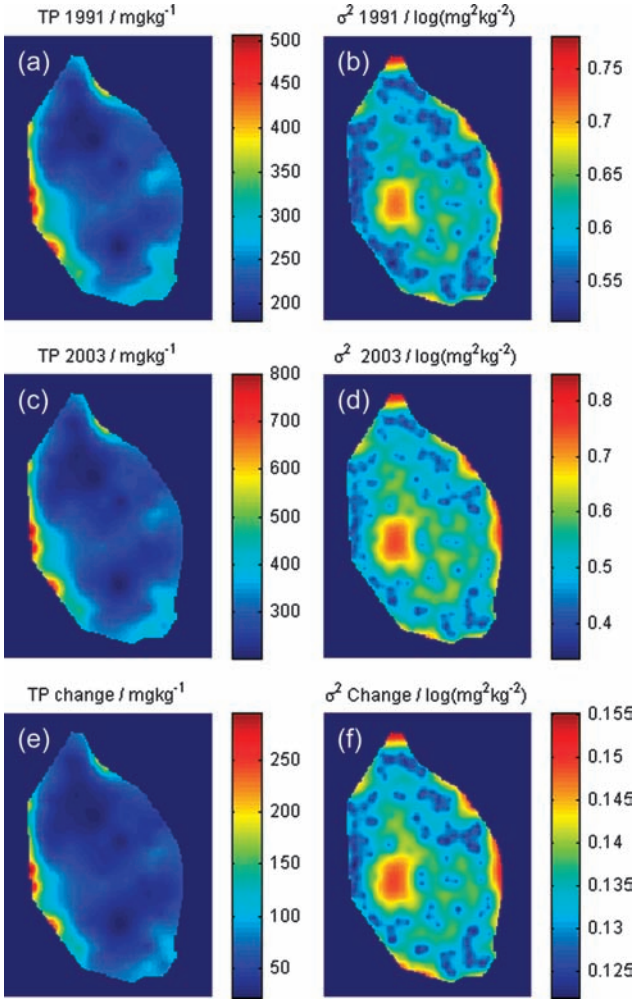

Figure 4 Predictions based upon a second-order stationary model of variation (M5) of TP /mgkg^{-1} and prediction variance (σ^2) / (log(mgkg^{-1}))^2 of log(TP) across WCA-1. (a) and (b): predictions and prediction variances from 1991. (c) and (d): predictions and prediction variances from 2003. (e) and (f): predictions and prediction variances for change between the two surveys.

This comparison is made using the likelihood ratio statistic defined in Equation (18). However since $\boldsymbol{\alpha}_0$ is not nested in $\boldsymbol{\alpha}$ we calculate the critical value of this statistic via a Monte Carlo approach. Finally we compare a model with non-stationary mean and stationary variance — denoted M4 and fitted by minimizing $l_R(\boldsymbol{\alpha}_0, \mathbf{M}, \mathbf{Sz})$ — with M5. This comparison is made using the likelihood ratio statistic defined in Equation (21) which requires that M6 with likelihood function $l_R(\boldsymbol{\alpha}_0, \mathbf{M}_0, \mathbf{Sz})$ is fitted. Additional comparisons between models are not possible due to the pattern of nesting between models. However these tests are sufficient to determine whether the data are best described by models with non-stationary mean and/or non-stationary variance. For each comparison we test the null hypothesis — that the data are best described by the simpler model — at the $P = 0.01$ level.

Optimization of future surveys

We assume that the primary aim of the next survey of TP in WCA-1 will be to predict the change in TP since 2003. We optimize a 100 observation survey for this purpose by SSA. The objective function is the mean prediction variance of the change in TP (Equation 14). To apply the SSA algorithm we require a model of the spatial and temporal variation of TP. We select the model which our likelihood ratio tests showed to be most appropriate for modelling the 1991 and 2003 surveys. We have fitted the model parameters for 2003. The parameters for a future survey are unknown, therefore we assume that they are equal to the 2003 parameters. We select the two cross-variogram parameters such that the cross-correlation between this future survey and 2003 is the same as that between the 2003 and 1991 surveys.

The location of observations in the optimized scheme defined above may be heavily influenced by the location of points in the 2003 survey. Therefore we also optimized a static survey. This survey is optimized in the same manner except WCA-1 is sampled at the same locations for each survey.

Results

The critical values of d and t which determined the extent of Regions A and B for the distance and *Typha* spp. approaches were $d^* = 1103$ m and $t^* = 23.7\%$. These values were determined from fitting M2. The likelihood ratio tests show that for both the distance (Table 1) and *Typha* spp. (Table 2) approaches, the non-stationary variance and stationary mean model (M2) describes the data better than the stationary variance model (M5). Furthermore the tests suggest that the best model for the distance approach is the non-stationary variance, non-stationary mean model (M1) whereas the non-stationary variance, stationary mean (M2) is best for the *Typha* spp. approach. Therefore we use these models to predict TP status and change.

The details of the fitted models are shown in Tables 3 and 4. The M2 model for the *Typha* spp. approach achieves a lower negative log residual likelihood than for the distance approach. The M1 models for each approach can not be compared because they have different fixed effects and hence undergo different projections. For each model $\bar{\boldsymbol{\theta}}$ is close to 1.0. However the $\tilde{\boldsymbol{\theta}}$ value for the stationary model (M5) lies outside the 95% confidence limits which were calculated by a Monte-Carlo method. The $\tilde{\boldsymbol{\theta}}$ values for all four models that have non-stationary variance (M1 and M2 for both distance and *Typha* spp. approaches) lie inside the 95% confidence limits. The model which is stationary in the variance and non-stationary in the mean (M4) lies outside the 95% confidence interval for the distance approach and marginally inside the 95% confidence interval for the *Typha* spp. approach.

In the maps of TP predictions shown in Figures 4–6 the log(TP) predictions have been back transformed to the median

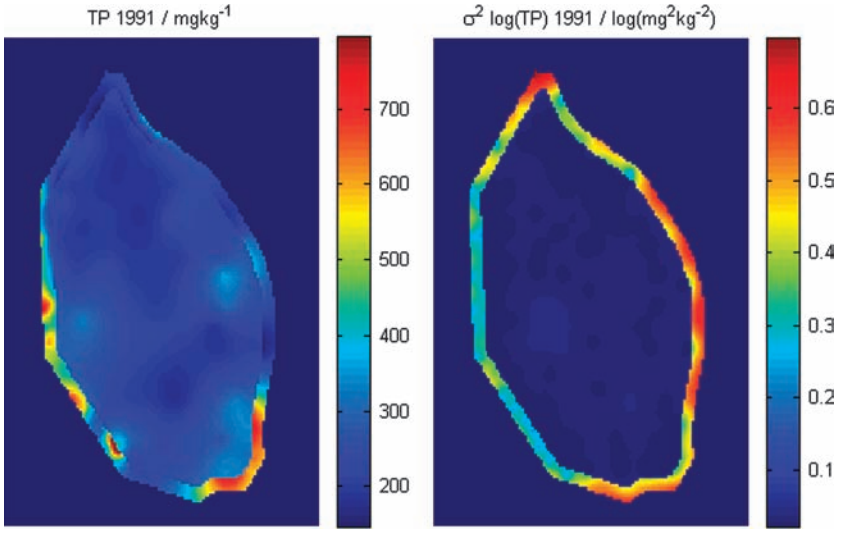

Figure 5 Predictions based upon a distance model which is non-stationary in the mean and variance (M1) of TP /mgkg^{-1} and prediction variance (σ^2) /(log(mgkg^{-1}))^2 of log(TP) across WCA-1 in 1991.

predictions in mg kg^{-1}. The prediction variances have not been back transformed since no unbiased back transform exists. The prediction maps from the stationary model (Figure 4) show large TP values adjacent to the western boundary of WCA-1 and, to a lesser extent, adjacent to the north east and south east boundaries. The predictions smoothly decrease towards the interior of the region. The predictions for the non-stationary variance models (Figures 5–6) show a sharper decrease in TP values away from the boundaries. The prediction variances for the stationary model are relatively small adjacent to the western boundary since this area is more intensely sampled than other parts of WCA-1. This pattern of prediction variances is unrealistic because TP is more variable within this part of WCA-1. The non-stationary models have relatively large prediction variances on this boundary. For the 1991 predictions with the distance approach (Figure 5) the area of larger prediction variances is a strip around the boundary of WCA-1 whereas the area of larger prediction variances is more irregular for the *Typha* spp. approach (Figure 6). We suggest that this is a more appropriate pattern of uncertainty since the variation of TP is not uniform at all points on the edge of WCA-1 but is largest down-stream from the pumps which transfer water from the bounding canals.

The map of change predicted from the stationary model (Figure 4e) is positive across WCA-1 and the spatial pattern is similar to the pattern of TP variation in 2003. This similarity is due to the magnitude of the spatially correlated variance in 1991 being less than that in 2003. This means that the predicted map for 1991 is relatively flat and the variations in it are dwarfed by those from 2003 when we calculate the change. The map of change in TP predicted from the non-stationary *Typha* spp. model (Figure 6e) shows a different pattern. Generally TP does increase across the region but the increases are relatively small adjacent to the western canal. This area had the largest increase on the stationary model. The largest increase in the non-stationary model occurs at the edge of the *Typha* spp. incursion suggesting that there is an invading front of P entering the interior region. There is an area to the south of WCA-1 where TP decreases but we suspect this is an artefact due to the area not being sampled in 1991. We note that the largest prediction variances occur in this area. The pattern of change adjacent to the western canal is patchy. This may again be an artefact due to the large variance within this area meaning that a larger sampling intensity is required.

We optimized the sample schemes for the M2 model of the *Typha* spp. approach. The left plot in Figure 7 shows the optimized scheme for a future survey designed to predict the change since the 2003 survey. The sampling is more intense in the regions where *Typha* spp. is dominant and there is some distortion caused by a preference to sample close to the 2003 sampling locations. The right hand plot shows the optimized static survey. Our predictions of change based on the M2 *Typha* spp. model suggested that there was an invasive front of P and the position of the front was correlated with the edge of the *Typha* spp. incursion. Therefore in this optimization we decreased the *Typha* spp. threshold to 10% to ensure that the region in danger of shifting to *Typha* spp. was intensely sampled. In this plot observation sites are dispersed throughout WCA-1 but the sampling is more intense within the *Typha* spp. region. Thus by employing this sample scheme we expect the artefacts within the map of change in TP to be reduced.

Conclusions

This case study illustrates how recent developments in geostatistics based upon linear mixed models may be extended to monitor non-stationary soil properties over two-dimensional study

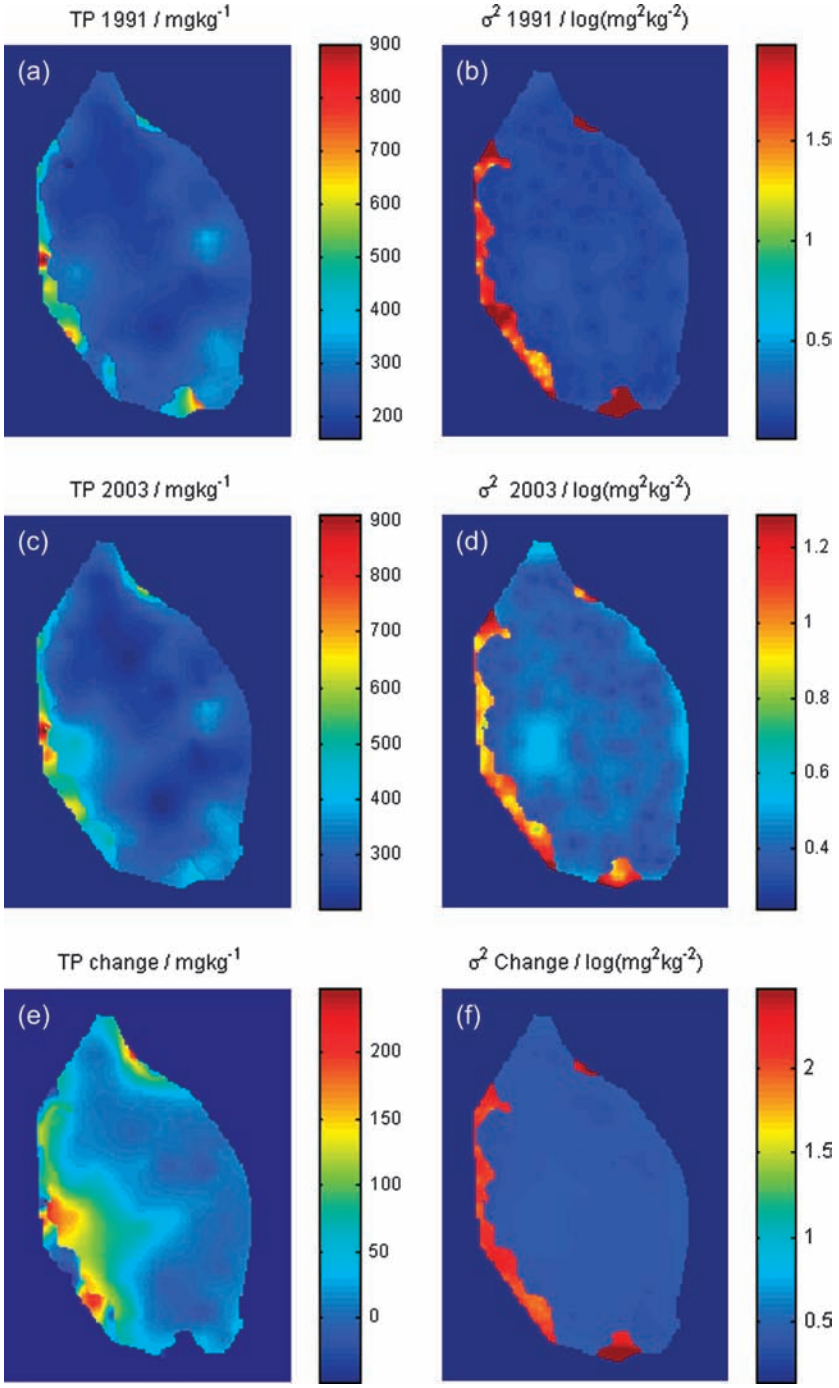

Figure 6 Predictions based upon a *Typha* spp. model which is non-stationary in the mean but stationary in the variance (M2) of variation of TP /mg kg^{-1} and prediction variance (σ^2) / $(\log(\text{mg } kg^{-1}))^2$ of log(TP) across WCA-1. (a) and (b): predictions and prediction variances from 1991. (c) and (d): predictions and prediction variances from 2003. (e) and (f): predictions and prediction variances for change between the two surveys.

regions. Specifically, the non-stationary model proposed by Lark (in press) has been extended to model spatial variation of multivariate rather than univariate soil properties over two-dimensions rather than one-dimension. Furthermore the SSA algorithm for the optimization of sample schemes suggested by van Groenigen *et al*. (1999) may also be extended to this non-stationary and multivariate framework. The non-stationary models of spatial variation are expressed as linear mixed models and the parameters are fitted by REML. Had the models been fitted by the conventional method of moments it would not have been possible to fully account for the complex spatial covariance between observations. The method of moments fits models to point estimates of the variogram by some form of weighted least squares. It is not clear how suitable weights could be selected for the complex behaviour described here.

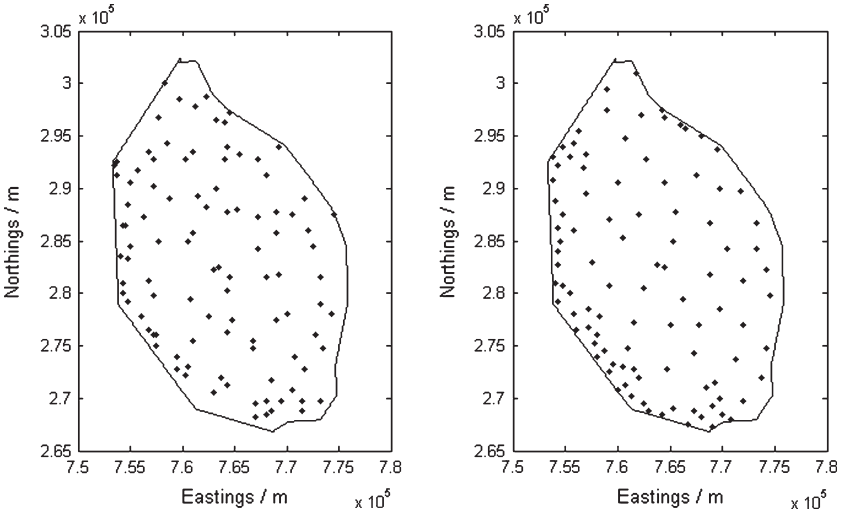

Figure 7 Optimized sampling scheme for predicting change in TP since 2003 (left) and optimized static scheme for predicting change in TP (right).

The second-order stationary model of spatial variation is shown to be unsuitable to monitor TP in WCA-1 since it cannot account for larger variances in TP content close to canals bounding WCA-1 than are observed in the interior of WCA-1. The areas of large variance cause the prediction variance to increase throughout WCA-1. Thus the prediction variances in the interior of WCA-1 are too large and the prediction variances adjacent to the boundary are too small. This leads to the median of $\boldsymbol{\theta}$ being less than the lower bound of the 95% confidence limit.

A non-stationary model based upon distance from a bounding canal leads to large prediction variances in a distinct strip adjacent to the boundary of the region. This is an over-simplification because the variance of TP varies around the boundary of WCA-1. The behaviour is more realistically accounted for if we divide the region into two sub-regions based upon the current *Typha* spp. coverage at each site. The non-stationary variance model based upon *Typha* spp. coverage achieves a smaller negative log-likelihood than the corresponding model based upon distance from the boundary. Modified likelihood ratio tests demonstrate that this non-stationary model of spatial variation represents the TP data better than stationary models. Furthermore, when non-stationary models are applied the median of $\boldsymbol{\theta}$ lies inside the 95% confidence limit. The predicted map of change in TP concentrations exhibits an invading front of P moving into the interior of the region. This is in contrast to the stationary model where the spatial pattern of the predicted change is similar to the pattern of TP variation seen in 2003.

The *Typha* spp. coverage provides auxiliary information of the spatial extent of TP enrichment. At present the information regarding *Typha* spp. coverage is interpolated from observations made in 2003. This is not entirely satisfactory since the *Typha* spp. coverage varies with time and the varying prediction variance of the interpolated map is not accounted for. However the situation can be improved in future phases of the monitoring survey if the *Typha* spp. coverage is re-surveyed using an optimized sampling scheme. Also it may be possible to use alternative exhaustive sources of auxiliary information of TP enrichment such as satellite images or to consider the sites at which P enters WCA-1 and the direction of water flow from these sites. In theory more detail may be included in our non-stationary model of TP. For example the extent of the two sub-regions could vary between the different surveys. However there may be practical difficulties in fitting these more complicated models to a dataset of less than 400 observations. Such problems will be reduced if the sample scheme is optimized for the specific model being fitted in the manner described in this paper.

Acknowledgement

BPM and RML were funded by the Biotechnology and Biological Sciences Research Council through its core grant to Rothamsted Research.

References

Brus, D.J. & Heuvelink, G.B.M. 2007. Optimization of sample patterns for universal kriging of environmental variables. *Geoderma*, **138,** 86–95.

Clark, I., Basinger, K.L. & Curran, P.J. 1989. MUCK - a novel approach to cokriging. In: *Proceedings of the Conference on Geostatistical, Sensitivity and Uncertainty Methods for Groundwater Flow and Radionuclide Transport Modelling* (ed. B.E. Buxton), 473–493, Battelle Press, Columbus, OH.

Corstanje, R., Grunwald, S. Reddy, K.R. Osborne, T.Z. & Newman, S. 2006. Assessment of the spatial distribution of soil properties in a northern everglades marsh. *Journal of Environmental Quality*, **35,** 938–949.

Davis, S.M. 1991. Growth, decomposition and nutrient retention in of *Cladium jamaicense* Crantz and *Typha domingensis* Pers., in the Florida Everglades. *Aquatic Botany*, **40,** 203–224.

DeBusk, W.F., Newman, S. & Reddy, K.R. 2001. Spatial-temporal patterns of soil phosphorus enrichment in Everglades Water Conservation Area 2A. *Journal of Environmental Quality*, **30,** 1438–1446.

Deutsch, C.V. & Journel, A.G. 1998. *GSLIB: Geostatistical Software Library and Users Guide*, 2nd ed., pp. 146–147. Oxford University Press, New York.

Lark, R.M. 2000. A comparison of some robust estimators of the variogram for use in soil survey. *European Journal of Soil Science*, **51,** 137–157.

Lark, R.M. in press. Kriging a soil variable with a simple non-stationary variance model. *Journal of Agricultural, Biological and Environmental Statistics*.

Lark, R.M., Cullis, B.R. & Welham, S.J. 2006. On spatial prediction of soil properties in the presence of a spatial trend: the empirical best linear unbiased predictor (E-BLUP) with REML. *European Journal of Soil Science*, **57,** 787–799.

Marchant, B.P. & Lark, R.M. 2007. Estimation of linear models of coregionalization by residual maximum likelihood. *European Journal of Soil Science*, **58,** 1506–1513.

Myers, D.E. 1991. Pseudo cross-variograms, positive definiteness and cokriging. *Mathematical Geology*, **62,** 29–43.

Newman, S., Reddy, K.R., DeBusk, W.F., Wang, Y., Shih, G. & Fisher, M.M. 1997. Spatial distribution of soil nutrients in a northern everglades marsh: water conservation area 1. *Soil Science Society of America Journal*, **61,** 1275–1283.

Papritz, A. & Flühler, H. 1994. Temporal change of spatially autocorrelated soil properties - optimal estimation by cokriging. *Geoderma*, **62,** 29–43.

Pardo-Igúzquiza, E. 1997. MLREML: A computer program for the inference of spatial covariance parameters by maximum likelihood and restricted maximum likelihood. *Computers & Geosciences*, **23,** 153–162.

Parker, G.G. 1974. Hydrology in predrainage system of the Everglades in South Florida. In: *Environments of South Florida, Past and Present, II*. (ed. P.J. Gleason), pp 18–27. Miami Geological Society, Coral Gables, FL.

Patterson, H.D. & Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika*, **58,** 545–554.

Richardson, J.R., Bryant, W.L., Kitchens, W.M., Mattson J.E. & Pope, K.R. 1990. An evaluation of refuge habitats and relationships to water quality, quantity and hydroperiod. A synthesis report. A.R.M. Loxahatchee National Wildlife Refuge, Boynton Beach, FL.

Rivero, R.G., Grunwald, S., Osborne, T.Z., Reddy, K.R. & Newman, S. 2007. Characterization of the spatial distribution of soil properties in Water Conservation Area -2A, Everglades, Florida. *Soil Science*, **172,** 149–166.

van Groenigen, J.W., Pieters, G. & Stein, A. 1999. Constrained optimization of soil sampling for minimisation of the kriging variance. *Geoderma*, **87,** 239–259.

Webster, R. & Oliver, M.A. 2007. *Geostatistics for Environmental Scientists*. 2nd ed. Wiley, Chichester.

Welham, S.J. & Thompson, R. 1997. Likelihood ratio tests for fixed model terms using residual maximum likelihood. *Journal of the Royal Statistical Society Series B*, **59,** 701–714.

Appendix

Selecting the covariance model

A likelihood ratio test may be applied to determine whether a particular covariance model better describes the data than a nested covariance model. The full vector of covariance parameters can be partitioned into two sub-vectors, i.e. $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_0 \ \boldsymbol{\alpha}_1]$. Thus the nested covariance model has parameter vector $\boldsymbol{\alpha}_0$ whereas the full model has additional covariance parameters $\boldsymbol{\alpha}_1$. The standard likelihood ratio test explores whether the full covariance model describes the data better than the sub-model. If $\boldsymbol{\alpha}_0$ is nested in $\boldsymbol{\alpha}$ then

$$D = -2\{l_{\mathrm{R}}(\hat{\boldsymbol{\alpha}}, \mathbf{M}, \mathbf{Sz}) - l_{\mathrm{R}}(\hat{\boldsymbol{\alpha}}_0, \mathbf{M}, \mathbf{Sz})\}, \qquad (18)$$

will be asymptotically distributed as χ^2 with q degrees of freedom where q is the difference between the number of parameters in $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}_0$. This gives a standard likelihood ratio test for the null hypothesis that all elements in $\boldsymbol{\alpha}_1$ are zero.

Lark (in press) compared his non-stationary model with a stationary model. However the parameter describing the boundary between the two regions in the non-stationary model could take any value in the stationary model and the two models were not truly nested. Therefore he modified the standard likelihood ratio test. One approach was to obtain an empirical distribution of D under the null hypothesis by Monte Carlo simulation and compare this distribution with the value of D from real data. Thus he simulated multiple realizations of the stationary model at the sites of the observations by Cholesky factorization (Deutsch & Journel, 1998), fitted both the stationary and non-stationary models to the simulated data and calculated D for each realization.

Selecting the fixed effects

The conventional likelihood ratio test is not valid for comparing nested design matrices for models estimated by REML because the data undergo different projections for each design matrix and thus the residual likelihoods are not comparable. Welham & Thompson (1997) suggest a solution to this problem based upon a single projection of the data. They describe two such tests. Each assumes that the model contains two sets of fixed effects which partition the design matrix $\mathbf{M} = [\mathbf{M}_0 \ \mathbf{M}_1]$ and parameters $\boldsymbol{\beta} = [\boldsymbol{\beta}_0 \ \boldsymbol{\beta}_1]$. The full fixed effect model is $\mathrm{E}[\mathbf{z}] = \mathbf{M}\boldsymbol{\beta}$ and the submodel is $\mathrm{E}[\mathbf{z}] = \mathbf{M}_0\boldsymbol{\beta}_0$. If we were fitting these models by REML the full model would use projection $\mathbf{S}$ whereas the sub-model would use projection $\mathbf{S}_0$. The test used here fits both $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_0$ using projection $\mathbf{S}$.

Using projection $\mathbf{S}$ the residual log likelihood of the submodel is

$$l_{\mathrm{R}}(\boldsymbol{\alpha}_0, \mathbf{M}_0, \mathbf{Sz}) = c(\mathbf{M}) - \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}\log|\mathbf{M}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{M}| - \frac{1}{2}(\mathbf{z} - \mathbf{M}_0\hat{\boldsymbol{\beta}}_0)^{\mathrm{T}}\mathbf{V}^{-1}(\mathbf{z} - \mathbf{M}_0\hat{\boldsymbol{\beta}}_0), \qquad (19)$$

with

$$\hat{\boldsymbol{\beta}}_0 = (\mathbf{M}_0^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{M}_0)^{-1}\mathbf{M}_0^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{z}, \qquad (20)$$

and Welham & Thompson's (1997) residual likelihood test statistic is

$$D = -2\{l_{\mathrm{R}}(\hat{\boldsymbol{\alpha}}_0, \mathbf{M}_0, \mathbf{Sz}) - l_{\mathrm{R}}(\hat{\boldsymbol{\alpha}}, \mathbf{M}, \mathbf{Sz})\}. \qquad (21)$$

This gives a standard likelihood ratio test for the null hypothesis that all elements in $\boldsymbol{\beta}_1$ are zero. Under the null hypothesis and for large data sets, D is approximately distributed as a chi-square distribution with q degrees of freedom where q is the number of fixed effect parameters in $\boldsymbol{\beta}_1$. We decide whether the $\mathbf{X}_1$ fixed effects should be included in the model by comparing D with a chi-squared distribution to test the null hypothesis that all entries of $\boldsymbol{\beta}_1$ are zero.