Journal of Vegetation Science 20: 639–649, 2009

Optimal spatio-temporal hybrid sampling designs for ecological monitoring

Hooten, Mevin B.[1]*; Wikle, Christopher K.[2]; Sheriff, Steven L.[3]; Rushin, John W.[4]

[1]*Department of Mathematics and Statistics, Utah State University, Logan, UT 84322-3900;*
[2]*Department of Statistics, University of Missouri, Columbia, MO 65211; E-mail wiklec@missouri.edu;*
[3]*Missouri Department of Conservation, Resource Science Center, 1110 South College Avenue, Columbia, MO 65201; E-mail Steve.Sheriff@mdc.mo.gov;* [4]*Department of Biology, Missouri Western State College, St. Joseph, MO 64507-2294; E-mail rushin@missouriwestern.edu;*
**Corresponding author; E-mail mevin.hooten@usu.edu*

Abstract

Question: Static sampling designs for collecting spatial data efficiently are being readily utilized by ecologists, however, most ecological systems involve a multivariate spatial process that evolves dynamically over time. Efficient monitoring of such spatio-temporal systems can be achieved by modeling the dynamic system and reducing the uncertainty associated with the effect of design choice at future observation times. However, can we combine traditional techniques with dynamic methods to find optimal dynamic sampling designs for monitoring the succession of a herbaceous community?

Location: Lower Hamburg Bend Conservation Area, Missouri, USA (40°34′42″ lat. 95°45′38″ long.).

Methods: The dynamic nature of the system under study is modeled in such a way that uncertainty in the measurements and temporal process can both be accounted for. Both fixed and roving monitoring locations were used in conjunction with a spatio-temporal statistical model to efficiently determine optimal locations of roving monitors over time based on the reduction of uncertainty in predictions.

Results: During the first 3 years of the study, roving monitors where held at fixed locations to allow for statistical parameter estimation from which to make predictions. Optimal monitoring locations for the remaining 2 years were selected based on the overall reduction in prediction uncertainty.

Conclusions: The dynamic and adaptive vegetation monitoring scheme allowed for the efficient collection of data that will be utilized for many future ecological studies. By optimally placing an additional set of monitoring locations, we were able to utilize information about the system dynamics when informing the data collection process.

Keywords: Dynamic sampling design; environmental monitoring; vegetation restoration; Kalman filter.

Introduction

Background: Hamburg Bend vegetation monitoring project

The Missouri River was historically a braided channel that was extremely dynamic. Channelization efforts removed all side channels along the river in northwest Missouri, USA. This channelization also produced large tracts of accreted lands adjacent to the Missouri River. These lands normally developed into forests, but many were later cleared and converted to agricultural fields.

Recently, the U.S. Army Corps of Engineers (COE) and the Missouri Department of Conservation (MDC) teamed together to acquire and manage an extensive area in the Missouri River flood plain in northwest Missouri. Currently, between St. Joseph, Missouri and the Iowa border, the COE has purchased over 24 km^2 within the Missouri River flood plain (Fig. 1). All of this land has been passed to MDC to be managed for fish and wildlife habitat. Future land acquisitions by the COE could add even more land that can be developed into fish and wildlife habitat. Such efforts have presented managers with both challenges and opportunities for vegetation management in these areas.

Chute restoration has become a priority management option in the area, and, as such, chutes are cut through an area to create shallow, flowing water habitat for fish and an island for wildlife habitat. The backwaters and shallow sloughs provide spawning grounds for fish, and the heavily vegetated islands make habitat for wildlife. However, as succession occurs on the islands, the wildlife habitat becomes less diverse and is only able to serve a limited number of species. The number of management strategies available to managers is quite limited for

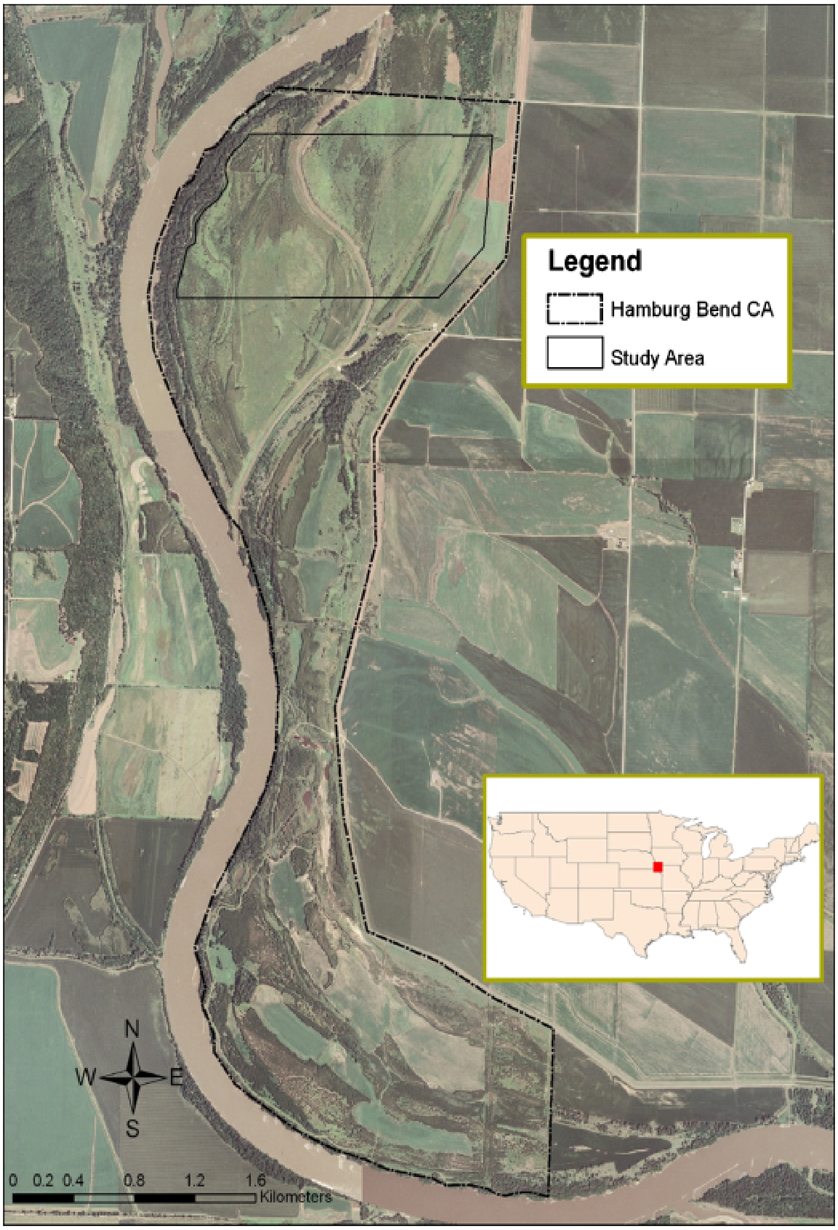

Fig. 1. Aerial photo of the Hamburg Bend chute under construction.

maintaining a diversity of habitats due to financial limitations and inaccessibility. That is, due to the physical attributes of the chute, swift river currents, and the fact that much of the study area is an island, it is not feasible to perform much mechanical manipulation of habitat.

The current management strategy on the newly created islands is simply to observe succession passively. The question then, is what happens to the vegetation under such a strategy and how might the vegetation change when alternative strategies, such as occasional burning of the island, are implemented? Also, a considerable amount of meandering of the chute shoreline is expected. This meandering will also remove vegetation and transport seeds.

For this study, the overall objective was to develop and implement a vegetation sampling protocol that would provide baseline information about the pre-chute development and immediate effects after the chute is restored at the Lower Hamburg Bend Conservation Area, on the Missouri side of the state line near Hamburg, Iowa (see Fig. 1). The specific objective for this statistical study associated with the overall project was to develop a hybrid sampling approach that would combine design-based sampling procedures with dynamic sampling techniques in order to reduce spatial and temporal uncertainties in the time series of data collected on this area. The reason that this hybrid approach was needed is because no previous data were available; nor was an underlying spatial/temporal model available that reflected the situation at Lower Hamburg Bend. An important aspect of the approach is that the data collected will ultimately be utilized in a variety of studies with topics ranging from successional community dynamics to invasibility of exotic species. Therefore, it is of interest to collect the data in such a manner that the overall uncertainty about vegetation processes is reduced so that the resulting data have maximal value for future research efforts.

Hybrid sampling designs

A hybrid sampling design, in this sense, refers to a set of monitoring locations that consist of both a fixed subset and a roving subset. The fixed subset contains monitoring locations that do not change from one observation period to another (i.e. are static in time), while the roving subset of monitors are allowed to change location over time (i.e., dynamic in time). Both static and dynamic designs (e.g. Cochran 1977; Wikle & Royle 1999, 2005) have been employed in the past for monitoring natural processes and each have advantages and disadvantages. For example, fixed designs are generally more convenient and economically feasible, while dynamic designs can better capture time-evolving behavior in the process, but generally require constant attention and more resources. Depending on the relative numbers of fixed versus roving monitors, a hybrid design can conserve resources by using the fixed monitors to efficiently gather baseline data and static behavior in conjunction with as many roving monitors as can be afforded in the project to capture dynamic behavior and reduce overall uncertainty in the measured process.

In the case of the Hamburg Bend vegetation monitoring project, the fixed monitoring portion of the proposed hybrid sampling approach used a systematic sample procedure. The systematic sample consisted of two independent grids of vegetation plots. Vegetation plots were placed in aligned squares for each grid (Cochran 1977, p 227–228), and each plot was 125 m from its neighbors. Each grid was randomly placed on the area using a random two-dimensional starting point rounded to the nearest meter within the 125 m × 125 m square on the southwest corner of the study area (Fig. 2). Any

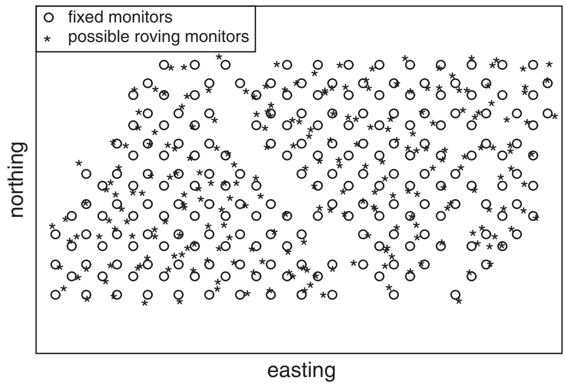

Fig. 2. Spatial plot of fixed design locations (circles) and possible locations for roving monitors (stars).

point outside of the study area was eliminated from further consideration during data collection. The boundaries of the study area were established as the area inside permanent levees on the west and east sides of Lower Hamburg Bend Conservation Area, with the east side boundary being about 150 m from the east levee. The study area is located at 40°34′42″ latitude and 95°45′38″ longitude. The fixed portion of the design consisted of two vegetation sampling grids, one containing 117 sampling plots and the other containing 115 plots. The purpose of using two sampling grids was to improve sampling coverage over the entire area and supply more information about the variability and spatial autocorrelation among plots than would be available through the use of a single grid with all sampling points an equidistance from neighboring points (Cochran 1977).

In addition to the fixed set of monitoring locations, an equal number of potential roving monitor locations were selected randomly throughout the study area. Given that it would be too costly to monitor at all of these additional locations each year, an economically feasible subset could be moved from year to year based on the characteristics (represented by a 'design criterion,' as discussed in the next section) of the process being studied. An added benefit of using this hybrid design (i.e., both fixed and roving monitors) is that more information is made available concerning the inherent spatial autocorrelation in the process being studied because the roving monitor locations occur at varying distances from fixed monitoring locations (Müller 2000). Moreover, roving monitors that are allowed to move throughout the study area are capable of capturing important behavior in the process being studied by taking advantage of its dynamic properties. For example, depending on the design criterion of choice (see the 'General Methodology' section for additional detail), one may be able to reduce measurement uncertainty in future monitoring periods by placing roving monitors closer to where change is most likely to occur. Similarly, depending on the dynamics of the process being studied and choice of design criterion, it may be of benefit to locate future monitoring efforts furthest from each other and possibly to move them from year to year to avoid sampling redundant behavior.

In the following sections, we present the general model specification that allows for hybrid sampling designs and discuss the formulation for the monitoring project at Hamburg Bend, and the specific design criterion chosen by managers of this Conservation Area, as well as its properties. It should be noted that many other types of design (both fixed and dynamic) and design criteria are available for use, each with its own properties. We refer the interested reader to Le & Zidek (2006) for additional detail and implications on other design methods and criteria not discussed here.

General Methodology

In this section, we outline the methodological approach for optimal hybrid design of monitoring locations for natural processes. If a reduction of the uncertainty in the measurements over time is desired, one needs to have a good understanding of the dynamic nature of the process under study. That is, since many natural processes can be thought of as systems that evolve (i.e. change) over time, a model needs to be constructed that captures the dynamic behavior of the system under study in order to predict what it is likely to do during future monitoring periods. If we can predict its behavior, we can place roving monitors in areas that help to reduce the uncertainty in the measured process. This is the fundamental notion behind the methods that follow. Thus, to implement the hybrid design, the following steps need to be taken:

(1) Collect baseline data whose purpose is to help estimate the dynamic characteristics of the process (i.e. plant community).
(2) Specify and fit a model that adequately represents the observational (i.e. data) and latent dynamic processes (i.e. underlying laws governing community spatio-temporal dynamics).
(3) Estimate the parameters involved in the model.
(4) Predict the likely future state of the latent dynamic process.
(5) Use the information from (4) to select locations for roving monitors that reduce the uncertainty involved in the prediction.
(6) Collect data at new locations and repeat steps (3)–(6) for future monitoring periods.

In what follows, we address the specific methodological aspects of steps (2)–(6), presented above, in a general manner so that they may be applied to other situations. The details on the specific attributes of the data under consideration here (e.g. normalizing transformations), and their utilization in the aforementioned general framework, can be found at the beginning of the Application section.

General spatio-temporal model

Realistic models for naturally occurring spatio-temporal processes must account for both temporal and spatial dependence in the system. The dynamic nature of the evolving system can often be modeled in an autoregressive fashion. Such models are common for characterizing dynamic processes and are discussed in detail in the temporal modeling literature (e.g. Caswell 2001 and Shumway & Stoffer 2006). In particular, for the application here, consider a linear formulation describing the evolution of the system ($\boldsymbol{\alpha}_t$, an $[n \times 1]$ vector, where n is the dimension of the full system of interest):

$$\boldsymbol{\alpha}_t = \mathbf{M}_t(\phi)\boldsymbol{\alpha}_{t-1}, \quad t = 1, \ldots, T\ . \tag{1}$$

One can think of this vector $\boldsymbol{\alpha}_t$ as a time-evolving spatial process in the system being studied; for example, if the natural system being studied is plant abundance, then $\alpha_{1,t}$ represents the abundance at the first location of interest during time t and $\alpha_{2,t}$ at the second location of interest, for $\boldsymbol{\alpha}_t = [\alpha_{1,t}, \alpha_{2,t}, \ldots]'$. Note that the $[n \times n]$ evolution matrix $\mathbf{M}_t(\phi)$ allows for system dynamics that vary in time; that is, $\mathbf{M}_t(\phi)$ modifies the vector representing the process at the previous time ($\boldsymbol{\alpha}_{t-1}$) to yield the process at the current time ($\boldsymbol{\alpha}_t$). Also called a 'propagator' or 'transition' matrix, $\mathbf{M}_t(\phi)$ is capable of exhibiting very complex nonlinear spatio-temporal behavior depending on the specification of its parameter ϕ (a vector that could be any dimension, but often less than n^2 for parsimony). In fact, various parameterizations of $\mathbf{M}_t(\phi)$ can be specified such that they approximate physical process models for theoretical dynamic systems (e.g. integral and differential equations; see Wikle 2003 and Hooten et al. 2007 for details).

Assuming that any chosen model [i.e. parameterization of $\mathbf{M}_t(\phi)$] is incapable of perfectly representing a natural system, one may want to accommodate such uncertainty. Thus, consider the process model in (1) in the presence of model error:

$$\boldsymbol{\alpha}_t = \mathbf{M}_t(\phi)\boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t, \quad t = 1, \ldots, T, \qquad (2)$$

such that, $\boldsymbol{\eta}_t$ in (2) represents the process error and is distributed normally, with mean zero and dependence structure given by the $[n \times n]$ covariance matrix $\Sigma\eta$.

Observations of the process ($\boldsymbol{\alpha}_t$) could be made in a multivariate fashion (i.e. with more than one piece of information monitored per spatial location and time), thus let $\mathbf{Y}_t$ denote the multivariate and spatial observation matrix of dimension $[m_t \times p]$, with m_t being the number of locations monitored at time t, and p being the number of measurements taken at each location in the presence of measurement error (say, $\boldsymbol{\varepsilon}_t$). Thus, the observation portion of the model can be written:

$$\mathbf{Y}_t = \mathbf{K}_t\boldsymbol{\alpha}_t\boldsymbol{\psi}'_t + \boldsymbol{\varepsilon}_t\boldsymbol{\psi}'_t, \quad t = 1, \ldots, T, \qquad (3)$$

where $\boldsymbol{\psi}_t$ is a $[p \times 1]$ vector suitable for linearly transforming the data matrix to a univariate spatial observation. This transformation should be scientifically motivated and could range from very simple and supervised (i.e. $\boldsymbol{\psi}_t$ is a vector of ones) to quite complex and unsupervised (i.e. $\boldsymbol{\psi}_t$ is the result of a spectral decomposition such as that used in principal components analysis). In a situation where the data represent ecological community structure, such as a matrix of ones and zeros denoting species occupancies at each location and time, the transformation where $\boldsymbol{\psi}_t$ is a vector of ones implies that the process being modeled is species richness (i.e. a simple measure of biodiversity). The $[m_t \times n]$ matrix $\mathbf{K}_t$, then, maps elements of the observed data matrix to the process vector; note that the row dimensionality (m_t) of the observed matrix ($\mathbf{Y}_t$) need not be the same dimension (n) as the process vector ($\boldsymbol{\alpha}_t$). Let the measurement error, $\boldsymbol{\varepsilon}_t$, be normally distributed with mean zero and $[m_t \times m_t]$ covariance matrix Σ_ε.

Dynamic and hybrid design methodology

The optimal design of experiments is a longstanding area of research in statistics (e.g. see reviews in Atkinson & Federov 1988; Federov & Hackl 1997). In the environmental and ecological sciences, there has also been a strong research emphasis on network design for monitoring spatial processes (e.g. Le & Zidek 1994; Nychka & Saltzman 1998). In addition, work has been done to consider spatially adaptive designs (e.g. Chao & Thompson 2001). The notion of adaptive design is similar to the design used here for the Hamburg Bend monitoring project, except there is the additional complication that the spatial process is changing over time. Thus, we seek to find the optimal design, dynamically, so that information from the previous times at which there were observations plays a role in deciding the optimal designs for future monitoring. Our approach is based on the space–time dynamic design methodology first developed by Wikle & Royle (1999, 2005).

The specification in the previous section allows the observation and process to have differing dimensionality as well as time-varying $\mathbf{K}_t$, thus it permits the sampling design to change over time. That is, suppose $\boldsymbol{\alpha}_t$ represents the process at a large but finite set of n spatial locations. Additionally, suppose that the set of n spatial locations is partitioned into two sets, a fixed set of m_f locations where the process will be measured at all T times, and an additional set of m_p locations at which the process may be monitored at a given time (thus, $n = m_f + m_p$).

The row dimension of the observation matrix $\mathbf{Y}_t$ can then be partitioned as: $m_t = m_f + m_{r,t}$ (where the subscript r merely indicates roving locations). The process will then be observed at the fixed set of locations each time, while a possible additional set of $m_{r,t}$ roving locations will be observed at a given time t. Therefore, the matrix $\mathbf{K}_t$ controls the sampling design because it links the observations and process.

Numerous methods have been developed for fitting such models; in particular, hierarchical Bayesian methods have been used to simultaneously estimate model parameters, variance components, the underlying process, and the predicted process (e.g. Wikle et al. 2001; Hooten & Wikle 2007). The uncertainty pertaining to such predictions can be evaluated with a criterion of choice and, ideally, minimized by finding an appropriate sampling design for future times (i.e. $T+j$, for $j = 1, \ldots, J$).

Unfortunately, such methods are prohibitively computationally intensive when the space of possible sampling designs is large (Chaloner & Verdinelli 1995; Müller 1999). In cases where $\boldsymbol{\Sigma}_\varepsilon$, $\boldsymbol{\Sigma}_\eta$, and $\mathbf{M}_t(\phi)$ are known or can be estimated (in a similar fashion as in traditional spatial prediction) a Kalman filtering approach can be used (Wikle & Royle 1999) and is much more efficient. Kalman methods have enjoyed much success in the fitting of sophisticated spatio-temporal models, and we refer the interested reader to Shumway & Stoffer (2006, p. 330–339) for detailed information on its derivation and utility for general purposes, as a full treatment of the method is beyond the scope of this paper. The important aspect of using Kalman filtering for the situation described here is that it allows us to estimate the latent process (i.e. $\boldsymbol{\alpha}_t$) as well as its uncertainty (which is critical for choosing a sampling design that reduces the uncertainty and thus improves predictive precision).

One key to taking the Kalman filtering approach here is that, because optimal sampling designs are the main focus, we need only to evaluate a design criterion that is based on the prediction error variance. The error variance for the jth prediction (i.e. $\text{var}(\boldsymbol{\alpha}_{T+j}|\mathbf{y}_T,\ldots,\mathbf{y}_1)$) can be found, in this case, through the following recursive Kalman equations:

$$\begin{aligned}\mathbf{A}_t &\equiv \text{var}(\boldsymbol{\alpha}_t|\mathbf{y}_t,\ldots,\mathbf{y}_1),\\ &= \mathbf{B}_t - \mathbf{B}_t\mathbf{K}_t'(\mathbf{K}_t\mathbf{B}_t\mathbf{K}_t' + \Sigma_\varepsilon)^{-1}\mathbf{K}_t\mathbf{B}_t,\end{aligned} \tag{4}$$

where,

$$\begin{aligned}\mathbf{B}_t &\equiv \text{var}(\boldsymbol{\alpha}_t|\mathbf{y}_{t-1},\ldots,\mathbf{y}_1),\\ &= \mathbf{M}_t\mathbf{A}_{t-1}\mathbf{M}_t' + \boldsymbol{\Sigma}_\eta,\end{aligned} \tag{5}$$

where the initial value for $\mathbf{A}_0$ is generally specified to be the unconditional covariance matrix of the process (i.e. $\boldsymbol{\alpha}_t$) at the prediction locations. Additionally, in the absence of significant information concerning the process, various simplifying assumptions must be made to aid in parameter estimation. For example, the measurement errors are often assumed to be independent (i.e. $\boldsymbol{\Sigma}_\varepsilon = \sigma_\varepsilon^2\mathbf{I}$), the process error can be parameterized to account for spatial dependence (i.e. $[\boldsymbol{\Sigma}_\eta]_{i,j} = \sigma_\eta^2 \exp(d_{i,j}/\theta)$, where $d_{i,j}$ represents the Euclidean distance between locations i and j), and the propagator matrix is parameterized to be diagonal with a single autoregressive parameter ϕ modeling the temporal dependence.

For situations where the focus is on one-step ahead sampling designs, equations (4) and (5) are iterated for $t = 1,\ldots,T$, after which $\mathbf{B}_{T+1}$ is calculated, a new sampling design is chosen, and $\mathbf{K}_{T+1}^{(i=1)}$ is created. The prediction error variance ($\mathbf{A}_{T+1}^{(i=1)}$) is then calculated using (4) and then utilized to create the design criterion $q_d^{(i=1)}$. In this case, let $q_d^{(i)} = \text{avg}(\text{diag}(\mathbf{A}_{T+1}^{(i)}))$.

Then, to reduce uncertainty in the measured process, we seek to minimize the sampling design criterion that is based on prediction error variance. Thus, we need to search the space of possible designs for spatial arrangements of roving monitors that provide small values of $q_d^{(i)}$. This can be accomplished iteratively by incrementing $i = 1,\ldots,n_d$ and recalculating $\mathbf{A}_{T+1}^{(i)}$ and $q_d^{(i)}$ such that all possible (n_d) sampling designs will be evaluated and their criteria compared to find the minimum. In practice, the space of sampling designs is much too large to loop through sequentially, and in such cases an approximately optimal design is sought. A cumbersome approach is to evaluate a finite set of designs randomly. Fortunately, algorithms have been developed to search for near optimal designs more efficiently. Here we employ an exchange algorithm that iteratively swaps roving monitor locations with neighboring locations and retains the design if the swapping reduces the criterion (e.g. Nychka & Saltzman 1998). The exchange algorithm can be iterated until the change in criterion drops below a given threshold or until no better designs can be found.

Once the optimal design is chosen for time $T+1$, new data can be collected ($\mathbf{y}_{T+1}$), parameters can be re-estimated, the Kalman equations (4) and (5) can be iteratively computed, and the design space for time $T+2$ can be searched for optimality. This process can be repeated indefinitely, thus providing optimal sampling designs for each time at which data are collected. In practice, it may be necessary to retain the same sampling design for the first few times so that model parameters can be effectively estimated.

Application: Hamburg Bend Vegetation Monitoring Project

The methods presented (generally) in the previous section serve as a framework that can be applied to a specific situation. In this section, we present details concerning the specific application of the methodological framework to the Hamburg Bend data. As such, we discuss the necessary transformations of the data as well as project-specific model specifications. Additionally, we present a method of comparison between a fixed sampling design approach versus the hybrid approach under a large number of sampling schemes. The purpose of this comparison is to illustrate the reduction of uncertainty possible when utilizing designs with dynamic components for monitoring systems that dynamically evolve over time.

Hybrid sampling designs for the Hamburg Bend vegetation monitoring project

The methods presented in the previous section have been applied to the Hamburg Bend vegetation

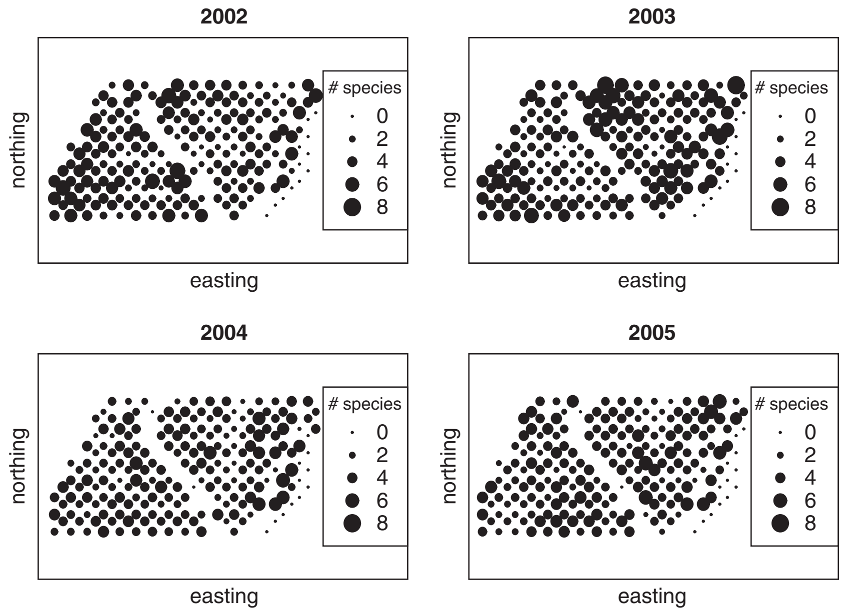

Fig. 3. Species richness data ($\mathbf{y}_t$) at fixed locations for the years 2002–2005.

monitoring study, which has collected data annually from 2002 to 2006, although the project is ongoing. Specifically, the data used here consist of the observed occupancy of a set of species ($\mathbf{Y}_t$) at a fixed number of locations (Fig. 2) in the Hamburg Bend Conservation Area (spanning approximately 12 km^2) for the first 3 years. Due to time constraints on this particular project, an initial monitoring period of only 3 years was possible; however, it should be noted that longer periods for preliminary data collection will yield improved parameter estimates and thus will allow for improved monitoring designs to be found. Part of the rationale for employing an optimal dynamic design in the Hamburg Bend project was related to the limited resources, and thus, managers sought to obtain as much information as possible out of the data.

In this case, $\mathbf{Y}_t$ is a binary data matrix as discussed in 'General spatio-temporal model,' and a simple transformation vector of ones (i.e. $\boldsymbol{\psi}_t = \mathbf{1}$) was used to convert the occupancy data to species richness data (Fig. 3). Thus, let $\mathbf{y}_t = \mathbf{Y}_t \psi_t$ (where, $\mathbf{y}_t$ is $[m_t \times 1]$), then as a normalizing and autoregressive modeling transformation, we let $\mathbf{z}_t = \sqrt{\mathbf{y}_t} - (1/T) \sum_{t=1}^{T} \sqrt{\mathbf{y}_t}$. Note that this additional nonlinear data transformation is used here mainly for normalization and still fits into the framework presented in the previous section. The result is that we are modeling the dynamics on a normalized and centered scale (a fairly common technique for autoregressive modeling). Now we can write the observation portion of the model:

$$\mathbf{z}_t = \mathbf{K}_t \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \ldots, T, \tag{6}$$

A few (i.e. seven) sampling locations were omitted from this analysis due to inaccessibility in all monitoring years, thus after the preliminary 3-year period of monitoring at the remaining 225 fixed locations, a subset of 17 locations were allowed to 'rove' and selected optimally from an additional set of 225 possible roving locations in each of the following years (Fig. 2). Thus, for years beyond the 3-year preliminary period, the fixed set consists of 208 locations, and the 17 roving monitors were selected from an additional set of 225 locations. Note that there is much flexibility in the choice of the numbers of fixed monitors versus roving monitors. More roving monitors allows for a better characterization of dynamic behavior but is often more expensive and time demanding. In the case of the Hamburg Bend project, funding only allowed for an additional 17 locations to be monitored in each observation period.

This dataset is much larger spatially than temporally, thus a parsimonious parameterization, such as $\mathbf{M}(\phi) = \phi$ (yielding a first-order autoregressive process model) is likely necessary. If the process ($\boldsymbol{\alpha}_t$) is known, the following method of moments estimator could be utilized to estimate the autoregressive parameter ϕ:

$$\hat{\phi} = \frac{tr\left(\sum_{t=2}^{T} \boldsymbol{\alpha}_t \boldsymbol{\alpha}'_{t-1}\right)}{tr\left(\sum_{t=2}^{T} \boldsymbol{\alpha}_t \boldsymbol{\alpha}'_t\right)}. \tag{7}$$

In this case, the process is not known, and thus, a surrogate must be used to estimate ϕ. We used the aforementioned centered square-root transformation of the richness data ($\mathbf{z}_t$) and assumed that the estimated autoregressive parameter ($\hat{\phi} = -0.4367$) would be similar for an autoregression in $\mathbf{z}_t$ as in $\boldsymbol{\alpha}_t$. When data are limited temporally, such estimates of autoregressive parameters are less certain; however, as monitoring continues in longer studies, the estimates will improve. Here, with only 3 years of preliminary data, this estimate for ϕ is all we know about the dynamics of the system and has to suffice until more data become available from which to find new estimates.

A geostatistical model was fitted to $\mathbf{z}_{t=3}$ in order to estimate the variance components of the spatio-temporal model (i.e. $\boldsymbol{\Sigma}_\varepsilon$, $\boldsymbol{\Sigma}_\eta$). Specifically, covariance was estimated using weighted least squares and an exponential spatial dependence model with nugget effect (see Waller & Gotway (2004), p. 280–288 for details on fitting such models). We found that exponential spatial dependence provided an adequate fit (via AIC comparison with other spatial models), and the resulting estimated nugget effect (0.0693) was then used as a surrogate for σ_ε^2, and represents measurement error and any remaining small-scale variability not captured by the spatially correlated error model for the process. The estimated sill ($\sigma_\eta^2 = 0.015$) and range ($\theta = 149.8$) parameters were then used to inform the covariance matrix for the process, (i.e. $\boldsymbol{\Sigma}_\eta = \sigma_\eta^2 \exp(d_{i,j}/\theta)$ as defined in the Methods section). Estimation of the inherent spatial and temporal structure of the process under study allows us to more realistically model the dynamic nature of the temporal process (Equations 2 and 3) and thus ultimately choose sampling designs that best reflect the probable future spatio-temporal behavior.

To select the optimal sampling designs, given our estimated model parameters (discussed above) and design criteria of choice for this project (e.g. $q_d^{(i)} = \text{avg}(\text{diag}(\mathbf{A}_{T+1}^{(i)}))$), the Kalman equations were then recursively computed and yielded a set of suboptimal roving locations for 2005. The aforementioned exchange algorithm was employed to efficiently search the space of sampling designs for an approximately optimal configuration. The resulting sampling design selected for 2005 included all 208 fixed locations as well as the additional optimal roving locations. The process was repeated again in 2005 to determine the optimal design to use for vegetation monitoring in 2006. Figure 4 spatially illustrates these optimal sampling designs for the years 2005 and 2006.

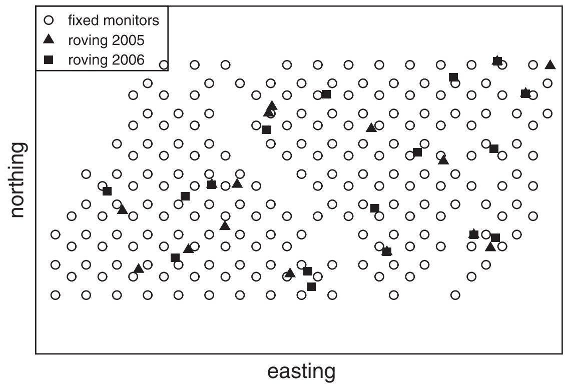

Fig. 4. Spatial plot showing optimal hybrid designs for the years 2005 and 2006 (including fixed locations).

Comparison of static and hybrid sampling designs for Hamburg Bend

One method for comparing the effect of allowing for roving monitors with the effect of fixing the locations of all monitors is through simulation. That is, given a similar spatial domain to the Hamburg Bend study area (i.e. a simulation domain on the same scale with representatively spaced monitoring locations) coupled with a spatio-temporal process that evolves dynamically in a similar fashion (i.e. using parameter estimates from the Hamburg Bend data), we can assess the difference in uncertainty when monitoring the process with a fixed design versus an optimal hybrid design. We provide, in App. 1, a program written in the R Statistical Programming Language (R Development Core Team 2007) that does precisely this. Specifically, we utilize the parameter estimates presented in the previous section and iteratively select numerous sampling designs, then evaluate the prediction uncertainty resulting from each design. This approach for comparing multiple static designs with their optimal hybrid counterparts is best described by the following steps:

(1) Select the locations of all monitors for the purpose of collecting baseline data for a number of observation periods (we used four baseline observation periods in the App. 1 code). This full set of monitoring locations is composed of both the fixed monitors and the roving monitors. The actual selection of locations could be made by hand, utilizing traditional sampling methods (e.g. Cochran 1977), or even optimal static design methods (e.g. Le & Zidek 2006); additionally, the subset of roving monitor locations could follow the same design, or be selected randomly (as in the App. 1 code), from a large but finite number of additional monitoring locations. In either case, the complete design (i.e.

the location of all monitors, fixed and roving) will not change over time, and thus will be used for comparison.
(2) After baseline data are collected, parameters in the model are estimated (using methods discussed in 'General spatio-temporal model' and 'Hybrid sampling designs for the Hamburg Bend vegetation monitoring project'). The parameter estimates obtained from the Hamburg Bend data are utilized in the App. 1 code.
(3) Calculate the design criterion (i.e. $q_d^{(i)} = \text{avg}(\text{diag}(\mathbf{A}_{T+1}^{(i)}))$ for current static design i) based on the estimated model parameters (i.e. $\phi, \boldsymbol{\Sigma}_\varepsilon$, and $\boldsymbol{\Sigma}_\eta$) and prediction error variance (i.e. $\mathbf{A}_{T+1}^{(i)}$).
(4) Using the static design as a starting point, utilize the exchange algorithm to assess neighboring sample locations for their ability to reduce the design criterion.
(5) Swap the static locations for the neighboring monitoring locations that reduce the design criterion the most; then, calculate the reduction in criterion value for which to compare static versus hybrid designs. When the number of neighbors searched is large enough, the new design will be approximately optimal and the difference in criteria will represent the amount of uncertainty reduced in the predicted process by using the optimal dynamic design.
(6) For comparison of multiple static designs versus their optimal dynamic counterparts, repeat steps (1)–(5), choosing a new static design each time.

Recall that the design criterion itself is a measure of the prediction uncertainty, thus large positive differences between static and hybrid sampling designs indicate a significant increase in predictive precision by utilizing hybrid designs (via the simulation-based Monte Carlo test). Figure 5 is the result of running the code in App. 1 for comparing a total of 2000 static sampling designs with their optimal hybrid counterparts. Additional details regarding the specific syntax in the code itself can be found in App. 1.

Discussion

The rationale behind this monitoring project is to best utilize available resources for the observation of the natural process. Monitoring the process at fixed locations is physically and economically more efficient, but roving monitoring locations better allow for the observation of a dynamically evolving spatio-temporal process and ultimately re-

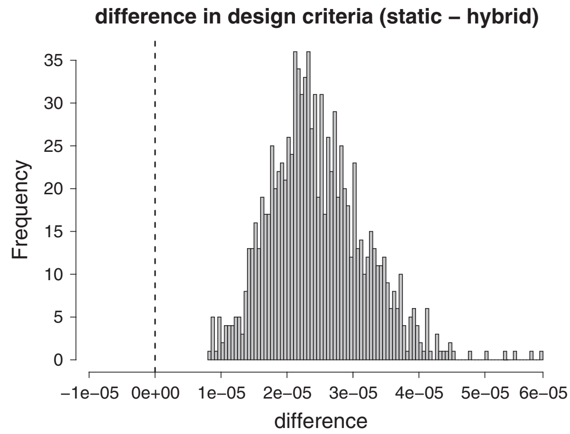

Fig. 5. Histogram showing the difference in sampling design criteria (i.e. units are in terms of prediction error variance for the latent process $\boldsymbol{\alpha}_t$) for 2000 fixed designs versus the optimal hybrid design for each of the 2000 initial design scenarios (x-axis notation: 1e-05 = 0.00001). Note that smaller design criteria imply a reduction of uncertainty in the measured process; thus, a positive difference between fixed and hybrid design, in terms of criteria (shown in the histogram), indicates a reduction of uncertainty (and thus an increase in precision) by allowing some monitors to move. The dashed line at a difference of zero indicates the point below which there would be no improvement in precision by allowing monitors to move.

sult in higher quality data (Wikle & Royle 1999, 2005). In the case of the Hamburg Bend Monitoring Project, enough funding and time was available to monitor 17 additional locations (beyond the 208 fixed monitoring locations) in each observation period. By allowing these additional 17 monitoring locations to vary in space over time, we were able to reduce the prediction uncertainty over the scenario where a static design is used for the entire observational period. An illustration of this fact is provided through the simulation setting, where numerous static sampling designs are compared with their optimal hybrid counterparts (Fig. 5). The design criterion chosen by the managers for this study is a measure of the overall prediction uncertainty (i.e. the average of the prediction error variance given in Equation 4), thus a reduction in the value of the design criterion indicates a reduction in prediction uncertainty. These simulations serve as a Monte Carlo test for the null hypothesis, where static designs and optimal hybrid designs are not different in terms of the design criterion for this study. Since the histogram in Fig. 5 does not overlap zero (i.e. indicating a p-value for the test that is very small) we have reason to believe that, in the case of Hamburg Bend, optimal hybrid designs significantly improve monitoring of the natural process under study. The amount of improvement can be quantified by dividing the difference in criteria by the average prediction error variance of the fixed designs from the simulations (i.e. q_d). In this case, q_d for the simulated fixed designs averaged 0.0058, and thus we were seeing up to approximately a 1% reduction in uncertainty in the simulation without needing to collect any additional data. This could be improved by allowing for more roving monitors, however, budget constraints often control this aspect of a monitoring protocol (as was the case in the Hamburg Bend study). Since a purely fixed design was not employed during the Hamburg Bend data collection, we cannot know the exact reduction of uncertainty, only that it was indeed significantly improved by using a hybrid design, as evidenced by the Monte Carlo simulation test.

The data presented in Fig. 3 for the Hamburg Bend study area show evidence of spatial structure when looking at the individual years of species richness, thus a geostatistical model for describing the covariance of the process error (i.e. $\mathbf{\Sigma}_\eta$) is reasonable. We found that although the spatial structure is similar from year to year, the pattern varies, indicating a dynamic component in the process, and, hence, the use of a spatio-temporal model (i.e. Equations 2 and 3) for describing the system is warranted. When fitting such a model to the Hamburg Bend data, we find that there is negative autocorrelation, suggesting the system is oscillating over time. This oscillatory behavior provides valuable information for choosing locations of roving monitors in future observation periods.

As for the specific designs used in the Hamburg Bend study area, we note that the optimal designs for the roving monitors for 2006 are different than those for 2005 (Fig. 4). This provides one justification for the development and implementation of the methodology, and illustrates further that fixed designs are not optimal for spatio-temporal processes. Based on the simulation study discussed above, we also know that the optimal hybrid designs used for data collection in the Hamburg Bend study area have reduced the uncertainty in the observation of the ecological process, and have thus resulted in higher quality data for use in other studies than would have been available through using a static monitoring protocol.

Wikle & Royle (1999) used a similar design criterion and found that optimal designs tend to favor locations that have higher variance and low correlation with other sites. Additionally, they show that, for cases with moderate time and spatial dependence, the optimal locations for roving sites should change because information at the same site is somewhat redundant over time. It should be noted that the choice of design criterion will most certainly result in a different optimal design (e.g. Berliner et al. 1998); however, based on the above arguments and the mathematical properties of common design criteria (being measures of uncertainty regarding the system under study), any choice of criterion still reduces uncertainty associated with observation of the process and is thus better than none. This is especially true with temporally evolving systems such as that presented here.

Conclusion

Although becoming more popular in atmospheric and environmental studies, the consideration of dynamic optimal sampling designs in ecological endeavors has not been widely used. It could be said that nearly every ecological monitoring study involves, to some extent, the observation of a process that is evolving in space over time. More sophisticated collection of data will lead to more sophisticated ecological models being employed in the presence of less uncertainty.

Overall, the use of a model-based approach to construct sampling designs can be thought of as both more and less efficient. In terms of an *a priori* time commitment, such sophisticated approaches may require significantly more planning, preparation, and modeling than traditional approaches. On the other hand, when monitoring resources are limited, managers and scientists may want to extract the most information out of the data that they can afford to collect; thus, if efficiency is expressed in terms of the ratio of knowledge gained over the amount of data obtained, then this is where optimal hybrid designs can be more efficient.

Finally, to our knowledge, this is the first time a hybrid fixed/roving spatio-temporal dynamic optimal design approach has been used to collect data in a real-world study. The results presented here are very promising and suggest that this approach can and should be considered in future long-term monitoring projects.

Acknowledgements. The authors would like to thank the Missouri Department of Conservation, and specifically, Craig Scoggins for his assistance and involvement in the project. Kyle Hedges and Nathan Mechlin provided valuable insights and expressed their thoughts as managers of the Hamburg Bend Conservation Area. George Hartman and many students from Missouri Western State University helped with the data collection and management for this project. Additionally, the authors thank the editors and anonymous reviewers for providing numerous helpful comments and suggestions on this manuscript.

References

Anon. (R Development Core Team.) 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at http://www.R-project.org (accessed 31 December 2007).

Atkinson, A.C. & Federov, V.V. 1988. Optimum design of experiments. In: Kotz, S. & Johnson, N.I. (eds.) *Encyclopedia of statistics (supplemental volume)*. pp. 107–114. John Wiley and Sons Inc., New York, NY, USA.

Berliner, L.M., Lu, Z.-Q. & Snyder, C. 1998. Statistical design for adaptive weather observations. *Journal of the Atmospheric Sciences* 56: 2536–2552.

Caswell, H. 2001. *Matrix population models*. 2nd ed. Sinauer Associates Inc. Publishers, MA, USA.

Chaloner, K. & Verdinelli, I. 1995. Bayesian experimental design: a review. *Statistical Science* 10: 273–304.

Chao, C.-T. & Thompson, S.K. 2001. Optimal adaptive selection of sampling sites. *Environmetrics* 12: 517–538.

Cochran, W.G. 1977. *Sampling techniques*. 3rd ed. John Wiley and Sons Inc., New York, NY, USA.

Hooten, M.B. & Wikle, C.K. 2007. Shifts in the spatio-temporal growth dynamics of shortleaf pine. *Environmental and Ecological Statistics* 14: 207–227.

Hooten, M.B., Wikle, C.K., Dorazio, R.M. & Royle, J.A. 2007. Hierarchical spatio-temporal matrix models for characterizing invasions. *Biometrics* 63: 558–567.

Federov, V.V. & Hackl, P. 1997. *Model-oriented design of experiments. Lecture notes in statistics*. Springer-Verlag, New York, NY, USA.

Le, N.D. & Zidek, J.V. 1994. Network designs for monitoring multivariate random spatial fields. In: Vilaplana, J.P. & Puri, M.L. (eds.) *Recent advances in statistics and probability*. pp. 191–206. VSP International Science Publishing, Leiden, NL.

Le, N.D. & Zidek, J.V. 2006. *Statistical Analysis of Environmental Space-time Processes*. Springer, New York, NY, USA.

Müller, P. 1999. Simulation based optimal design. In: Berger, J.O., Bernardo, J.M., Dawid, A.P. & Smith, A.F.M. (eds.) *Bayesian statistics 6*. pp. 459–474. Oxford University Press, Oxford, UK.

Müller, W.G. 2000. *Collecting spatial data: optimum design of experiments for random fields*. 2nd ed. Physica Verlag, Heidelberg, DE.

Nychka, D. & Saltzman, N. 1998. Design of air-quality monitoring networks. In: Nychka, D. & Piegorsch, W. (eds.) *Case studies in environmental statistics*. pp. 51–76. Spring-Verlag, NY, US.

Shumway, R.H. & Stoffer, D.S. 2006. *Time series and its applications*. 2nd ed. Springer-Verlag, New York, NY, USA.

Waller, L.A. & Gotway, C.A. 2004. *Applied spatial statistics for public health data*. John Wiley & Sons Inc., NJ, USA.

Wikle, C.K. 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology* 84: 1382–1394.

Wikle, C.K. & Royle, J.A. 1999. Space-time models and dynamic design of environmental monitoring networks. *Journal of Agricultural, Biological, and Environmental Statistics* 4: 489–507.

Wikle, C.K. & Royle, J.A. 2005. Dynamic design of ecological monitoring networks for non-Gaussian spatio-temporal data. *Environmetrics* 16: 507–522.

Wikle, C.K., Milliff, R.F., Nychka, D. & Berliner, L.M. 2001. Spatiotemporal hierarchical Bayesian modeling: tropical Ocean surface winds. *Journal of the American Statistical Association* 96: 382–397.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix S1. R Code for comparison of static and hybrid sampling designs.

Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Received 10 September 2007;
Accepted 12 September 2008.
Co-ordinating Editor: I. Kühn.