ENVIRONMETRICS
*Environmetrics* 2005; **16**: 507–522
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/env.718

# Dynamic design of ecological monitoring networks for non-Gaussian spatio-temporal data

Christopher K. Wikle[1,*,†] and J. Andrew Royle[2,‡]

[1]*Department of Statistics, University of Missouri, 146 Middlebush, Columbia, MO 65211, U.S.A.*
[2]*U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, MD 20708, U.S.A.*

SUMMARY

Many ecological processes exhibit spatial structure that changes over time in a coherent, dynamical fashion. This dynamical component is often ignored in the design of spatial monitoring networks. Furthermore, ecological variables related to processes such as habitat are often non-Gaussian (e.g. Poisson or log-normal). We demonstrate that a simulation-based design approach can be used in settings where the data distribution is from a spatio-temporal exponential family. The key random component in the conditional mean function from this distribution is then a spatio-temporal dynamic process. Given the computational burden of estimating the expected utility of various designs in this setting, we utilize an extended Kalman filter approximation to facilitate implementation. The approach is motivated by, and demonstrated on, the problem of selecting sampling locations to estimate July brood counts in the prairie pothole region of the U.S. Copyright © 2005 John Wiley & Sons, Ltd.

## 1. INTRODUCTION

Information about annual waterfowl production is important in the management of North American waterfowl populations. At smaller scales, such information can be utilized to improve understanding of biological mechanisms that affect recruitment and to guide habitat management activities. One source of information on production is based on brood counts from a fixed-wing aerial survey of linear segments conducted by the U.S. Fish and Wildlife Service (FWS), Canadian Wildlife Service and other organizations in July of each year over much of the U.S. Prairie Pothole Region (PPR) and Canadian Prairie Parklands. This survey generates counts of the number of broods on a large number of 18 mile linear segments. Although these counts are incomplete in the sense that not all broods present are detected, and no attempt to account for detection probability is made, they are believed to represent relative variation in the distribution of brood density.

*Correspondence to: C. K. Wikle, Department of Statistics, University of Missouri, 146 Middlebush, Columbia, MO 65211, U.S.A.
†E-mail: wikle@stat.missouri.edu
‡E-mail: aroyle@usgs.gov

Contract/grant sponsor: National Science Foundation; contract/grant number: DMS-0139903.

*Received 16 February 2004*
*Accepted 28 September 2004*

In an attempt to generate detection-bias adjusted estimates of brood density, the U.S. FWS is currently evaluating alternative, ground-based sampling protocols which will provide estimates of actual density by properly accounting for imperfect detection. Sampling protocols under consideration include double-observer sampling (Nichols *et al.*, 2000), distance sampling (Rosenstock *et al.*, 2002) or simple point counts (Royle, 2004). Such protocols require considerably more effort because the sampling is ground-based (i.e. observers must visit individual wetlands) and in some cases requires multiple samples of individual linear segments. Consequently, there is a need to evaluate methodologies for making informed sampling decisions—i.e. determining which segments should be sampled each year. Lacking information on actual brood density, we consider using the fixed-wing aerial brood counts as a surrogate for abundance in order to evaluate sample design methodology. Specifically, our goal is to select a small number (e.g. 20) of the 139 segments for ground-based sampling in a manner that yields the most information about the spatial distribution of broods over the prairie pothole region (PPR).

This problem requires a choice of sampling locations in the future, called a design, based on observed information in the past. Model based design in this dynamic spatio-temporal setting will require a spatio-temporal dynamic model, with a non-Gaussian data model. We propose a design strategy for such a model and illustrate its use on the brood-sampling problem described above. Section 2 describes the non-Gaussian spatio-temporal dynamic design approach, followed by an implementation of the approach on the brood example in Section 3. Section 4 contains discussion and conclusion.

## 2. NON-GAUSSIAN SPACE–TIME DYNAMIC DESIGN

There has been substantive work in the statistics literature devoted to the issue of obtaining spatial designs for environmental processes (e.g. Federov and Müller, 1989; Haas, 1992; Guttorp *et al.*, 1993; Cox *et al.*, 1996; Oehlert, 1996; Nychka and Saltzman, 1998; Bueso *et al.*, 1998; Sansó and Müller, 1998; Müller and Pázman, 2003). For comprehensive overviews of optimal spatial design see Atkinson and Federov (1988), Federov and Hackl (1997) and Müller (2000). Given that most environmental and ecological processes are dynamic in the sense that the underlying process evolves in space and time, static designs based on spatial correlation only may not be as efficient as designs that consider the joint spatio-temporal dependence. In general, the same methodologies used to construct optimal designs in the presence of spatial correlation can be used in the presence of spatio-temporal correlation if the design is to remain static for all time (for example, Arbia and Lafratta, 1997; Le and Zidek, 1994; Federov and Nachtsheim, 1995). However, as demonstrated in Wikle and Royle (1999), one can gain significant efficiencies in the spatio-temporal context by allowing the design to change in time (if the monitoring protocols allow for such network changes). We refer to this problem as the dynamic (or adaptive) design problem. Although related, this is different from the traditional statistical notion of 'adaptive sampling' (e.g. Thompson and Seber, 1996). For example, Chao and Thompson (2001) discuss optimal adaptive selection in the spatial case, but they are not concerned with the issues of temporal dependence in spatial settings.

As shown by Wikle and Royle (1999), dynamic designs for Gaussian spatio-temporal processes (i.e. Gaussian data model and Gaussian state process) show greater improvement relative to static designs as the temporal correlation increases, where the measure of efficiency is based on the average prediction variance across space. This is intuitive in that if one took observations at location $s$ and time $t$ then it would be largely redundant to take observations at the same location $s$ at time $t + 1$ if there is

 

strong temporal correlation for location $s$. Thus, design efficiency is improved by moving the observation locations at time $t+1$. Furthermore, there can be large increases in efficiency for dynamic designs if the spatial correlation is strong. They show that in the case of non-separable spatio-temporal dependence one still achieves substantial improvement in design efficiency relative to static designs by allowing measurement locations to vary with time. However, the optimal designs in this case are less intuitive due to the non-separable spatio-temporal dependence.

In the non-Gaussian setting, the dynamic designs are also likely to be less intuitive. Unlike the Gaussian setting, these designs are dependent on the observed process rather than just the dependence structure of the process. Furthermore, the dependencies of the design on the anticipated future observations makes the calculation of the design criterion more complicated. For this reason, we favor the simulation based design approach, as outlined in Müller (1999). Before describing this approach in the context of spatio-temporal dynamic design, we discuss the general non-Gaussian spatio-temporal dynamical model under consideration.

### *2.1. Non-Gaussian spatio-temporal dynamic models*

Consider spatio-temporal data $\mathbf{y}_t \equiv (y_t(r_1^t), \ldots, y_t(r_{m_t}^t))'$ for spatial locations $r_i^t \in \{s_1, \ldots, s_n\}$ in $\mathcal{R}^2$ and times $t = 1, \ldots, T$. Note that the time superscript on the spatial locations $r_i$ is necessary since in general not all possible sampling locations have data for a given time $t$. Let $\boldsymbol{\alpha}_t \equiv (\alpha_t(s_1), \ldots, \alpha_t(s_n))'$ be a spatio-temporal process at spatial locations $\{s_1, \ldots, s_n\}$. Furthermore, for each time $t$ define a matrix of covariates $\mathbf{X}_t$ which is of dimension $m_t \times p$ along with associated $p \times 1$ parameter vector $\boldsymbol{\beta}$. Let $\mathbf{Y}_1^t \equiv \{\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots, \mathbf{y}_1\}$ and $\boldsymbol{\alpha}_1^t \equiv \{\boldsymbol{\alpha}_t, \boldsymbol{\alpha}_{t-1}, \ldots, \boldsymbol{\alpha}_1\}$.

In the brood count design problem of interest here, we have areal survey brood count data for July of each year, time $t = 1, \ldots, T$. These are assumed to be available for all locations $s_i, i = 1, \ldots, n$, at all times $t = 1, \ldots, T$. Furthermore, the primary covariates in our case are areal survey pair counts, taken in May of each year. Again, these are available at all locations $s_i$ and times $t$. The design problem is to select a subset of the spatial locations $s_1, \ldots, s_n$ at which to take ground truth brood counts when the $T+1$ areal survey is under way. Thus, we seek a subset of spatial segments at $T+1$, $\{r_i^{T+1} \in \{s_1, \ldots, s_n\}\}$.

We make conditional independence assumptions for the measurement process such that

$$p(\mathbf{y}_t \mid \boldsymbol{\alpha}_1^t, \mathbf{Y}_1^{t-1}, \boldsymbol{\beta}) = p(\mathbf{y}_t \mid \boldsymbol{\alpha}_t, \boldsymbol{\beta}), \quad \text{for} \quad t = 1, 2, \ldots \tag{1}$$

Analogous to Fahrmeir (1992), we assume that (1) is from the $m_t$-dimensional natural exponential family,

$$\mathbf{y}_t \mid \boldsymbol{\alpha}_t, \boldsymbol{\beta} \propto \exp\{\boldsymbol{\gamma}_t' \mathbf{y}_t - b_t(\boldsymbol{\gamma}_t') - c_t(\mathbf{y}_t)\} \tag{2}$$

where $\boldsymbol{\gamma}_t$ are the natural parameters and are functions of the conditioning variables, $\boldsymbol{\alpha}_t$ and $\boldsymbol{\beta}$. Under the usual regularity assumptions for exponential families we then specify the link function $\mathbf{g}(\ )$ such that

$$\mathbf{g}(\boldsymbol{\mu}_t) = \mathbf{X}_t \boldsymbol{\beta} + \mathbf{Z}_t \boldsymbol{\alpha}_t \tag{3}$$

or

$$\boldsymbol{\mu}_t = \mathbf{h}(\mathbf{X}_t \boldsymbol{\beta} + \mathbf{Z}_t \boldsymbol{\alpha}_t) \tag{4}$$

 

where $\mathbf{h}(\,) = \mathbf{g}^{-1}(\,)$, the inverse link function, $\boldsymbol{\mu}_t$ is the mean of (2), and $\mathbf{Z}_t$ is an $m_t \times n$ model matrix for the random effects, assumed known. A special case, where the data are Poisson and $\mathbf{g}(\,)$ is the log link, is known as a log-Gaussian Cox process (e.g. Møller *et al.*, 1998; Brix and Diggle, 2001; Brix and Møller, 2001; and Christensen and Waagepetersen, 2002).

One also must specify a model for the spatio-temporal random process, $\boldsymbol{\alpha}_t$. For environmental and ecological data, the latent spatio-temporal processes tend to be dynamic, in the sense that a good approximation to the underlying process assumes that the process at time $t$ can be described as a functional modification of the process at time $t-1$ plus noise. The justification for this often comes from the underlying partial differential equations that describe processes with diffusion, advection and/or growth, which are common in the environmental/ecological sciences (e.g. Wikle, 2003). Thus, we assume a first-order Markov structure (i.e. a vector autoregressive model of order one) for the spatio-temporal $\boldsymbol{\alpha}_t$ process:

$$p(\boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_0, \boldsymbol{v}) = p(\boldsymbol{\alpha}_0) \prod_{t=1}^{T} p(\boldsymbol{\alpha}_t \mid \boldsymbol{\alpha}_{t-1}, \boldsymbol{v}) \tag{5}$$

where $\boldsymbol{v}$ are parameters that describe the dynamical evolution. Specifically, we assume

$$\boldsymbol{\alpha}_t \mid \boldsymbol{\alpha}_{t-1}, \mathbf{F}(\boldsymbol{v}), \mathbf{Q}(\boldsymbol{v}) \sim N_n(\mathbf{F}(\boldsymbol{v})\boldsymbol{\alpha}_{t-1}, \mathbf{Q}(\boldsymbol{v})) \tag{6}$$

where the transition or 'propagator' matrix $\mathbf{F}$ and noise covariance matrix $\mathbf{Q}$ depend on the parameters $\boldsymbol{v}$. These parameters may sometimes be treated as 'known' but alternatively may be given a prior distribution $p(\boldsymbol{v})$. Note also that we assume a distribution $p(\boldsymbol{\alpha}_0)$ for the initial condition in (5). Often, this initial condition is also treated as 'known'.

### 2.2. *Simulation based dynamic design*

After defining the non-Gaussian spatio-temporal dynamic model, we can now consider the simulation-based dynamic design framework. For a general review of Bayesian design approaches see Chaloner and Verdinelli (1995). Our description is motivated by the work of Müller (1999).

The design problem of interest is based on the idea that given data $\mathbf{Y}_1^T$ we would like the design $\mathbf{d}_{T+1}$ that optimizes some feature of the distribution $p(\boldsymbol{\alpha}_{T+1}, \boldsymbol{\beta}, \mathbf{y}_{T+1} \mid \mathbf{Y}_1^T, \mathbf{d}_{T+1})$. By 'design' we mean the subset of the $n$ possible spatial measurement locations, $\{r_1^{T+1}, \ldots, r_{m_{T+1}}^{T+1}\}$, at which observations are collected at time $T+1$. Recall, in the brood count design problem, that the design corresponds to the subset of the $n$ segments at time $T+1$ at which we will ground truth the July areal brood count observations. For notational simplification, let $\boldsymbol{\theta}_t \equiv [\boldsymbol{\beta}' \quad \boldsymbol{\alpha}_t']'$. Define $u(\boldsymbol{\theta}_{T+1}, \mathbf{y}_{T+1}, \mathbf{d}_{T+1})$ to be some utility function of interest. We then seek to find the design $\mathbf{d}_{T+1}$ that maximizes the expected utility with respect to $\mathbf{y}_{T+1}$ and $\boldsymbol{\theta}_{T+1}$, conditional upon the data $\mathbf{Y}_1^T$. That is, suppressing the possible dependence on $\boldsymbol{v}$ for now, we seek the design $\mathbf{d}_{T+1}$ from the set of potential designs $D$ that maximizes

$$U(\mathbf{d}_{T+1} \mid \mathbf{Y}_1^T) = \int u(\boldsymbol{\theta}_{T+1}, \mathbf{y}_{T+1}, \mathbf{d}_{T+1}) p_d(\mathbf{y}_{T+1} \mid \boldsymbol{\theta}_{T+1}) p(\boldsymbol{\theta}_{T+1} \mid \mathbf{Y}_1^T)\, \mathrm{d}\boldsymbol{\theta}_{T+1} \mathrm{d}\mathbf{y}_{T+1} \tag{7}$$

where $p(\boldsymbol{\theta}_{T+1} \mid \mathbf{Y}_1^T)$ is just the posterior predictive distribution of $\boldsymbol{\theta}_{T+1}$ and $p_d(\mathbf{y}_{T+1} \mid \boldsymbol{\theta}_{T+1})$ is the data model (e.g. (2)) given design $\mathbf{d}_{T+1}$; we denote the dependence on $\mathbf{d}_{T+1}$ by the $d$ subscript in $p_d(\,)$.

 

Typically, the integral in (7) is intractable, in which case one can either simplify the integrand or approximate the integral. In the case of the latter, one can obtain samples $\{\mathbf{y}^j_{d,T+1}, \boldsymbol{\theta}^j_{T+1}\}_{j=1}^N$ from the joint distribution $p_d(\mathbf{y}_{T+1}, \boldsymbol{\theta}_{T+1} \mid \mathbf{Y}_1^T)$ by sampling from the posterior predictive distribution for $\boldsymbol{\theta}_{T+1}$ and, using that in the data model, assuming design $\mathbf{d}_{T+1}$. That is, $\{\mathbf{y}^j_{d,T+1}, \boldsymbol{\theta}^j_{T+1}\}$ is the $j$th sample from the joint distribution for time $T+1$ given the design $\mathbf{d}_{T+1}$ and the data $\mathbf{Y}_1^T$.

In principle, one can then get a Monte Carlo estimate of the expected utility from

$$\hat{U}(\mathbf{d}_{T+1} \mid \mathbf{Y}_1^T) = \frac{1}{N}\sum_{j=1}^{N} u(\boldsymbol{\theta}^j_{T+1}, \mathbf{y}^j_{d,T+1}) \tag{8}$$

In practice this evaluation can be difficult. For example, in the brood sampling design problem discussed in the Introduction, the objective is to estimate the total number of broods produced in the surveyed region. Thus, if $y_t(s) \mid \mu_t(s) \sim$ ind. Poisson$(\mu_t(s))$, then the total observed counts at time $T+1$, $\sum_s y_{T+1}(s)$, is Poisson$(\sum_s \mu_{T+1}(s))$. Hence, it seems logical to seek a design that minimizes var$(\sum_s \mu_{T+1}(s) \mid \mathbf{Y}_1^T, \mathbf{d}_{T+1})$. Specifically, we are interested in minimizing the variance of the sum, over space, of the predicted mean (ground truth) brood counts for 'next' July (e.g. the mean of $y$ at $T+1$) given areal pair counts in May (e.g. also at $T+1$) and the historical areal survey brood counts (e.g. $y$ at times $t = 1, \ldots, T$). Note that the areal survey May pair count is a covariate $(X_{T+1}(s_i))$ and corresponds to the same yearly time index as the associated July brood count.

Thus, defining the mean brood count at location $s_i$ and time $t$ as $\mu_t(s_i)$, we are interested in maximizing the following expected utility:

$$\begin{aligned} U(\mathbf{d}_{T+1} \mid \mathbf{Y}_1^T) &= -\,\mathrm{E}\left[\left(\sum_{i=1}^{n} \mu_{T+1}(s_i) - \mathrm{E}\left[\sum_{i=1}^{n} \mu_{T+1}(s_i) \mid \mathbf{y}_{d,T+1}, \mathbf{Y}_1^T\right]\right)^2 \,\middle|\, \mathbf{y}_{d,T+1}, \mathbf{Y}_1^T\right] \\ &= -\,\mathrm{E}\left[\left(\sum_{i=1}^{n} h(\theta_{T+1}(s_i)) - \mathrm{E}\left[\sum_{i=1}^{n} h(\theta_{T+1}(s_i)) \mid \mathbf{y}_{d,T+1}, \mathbf{Y}_1^T\right]\right)^2 \,\middle|\, \mathbf{y}_{d,T+1}, \mathbf{Y}_1^T\right] \end{aligned}$$

which can be estimated in principle by Monte Carlo (MC), given design dependent samples $\{\boldsymbol{\theta}^j_{T+1}, \mathbf{y}^j_{d,T+1}\}$, $j = 1, \ldots, N$:

$$\hat{U}(\mathbf{d}_{T+1} \mid \mathbf{Y}_1^T) = \frac{-1}{N}\sum_{j=1}^{N}\left(\sum_{i=1}^{n} h(\theta^j_{T+1}(s_i)) - \mathrm{E}\left[\sum_{i=1}^{n} h(\theta_{T+1}(s_i)) \mid \mathbf{y}^j_{d,T+1}, \mathbf{Y}_1^T\right]\right)^2 \tag{9}$$

The problem with calculating (9) in practice is the evaluation of the remaining expectation. For example, if one were to estimate this via MC, then it would require evaluation of

$$\hat{\mathrm{E}}\left[\sum_{i=1}^{n} h(\theta_{T+1}(s_i)) \mid \mathbf{y}^j_{d,T+1}, \mathbf{Y}_1^T\right] = \frac{1}{M}\sum_{k=1}^{M}\left(\sum_{i=1}^{n} h(\theta^k_{T+1}(s_i)) \mid \mathbf{y}^j_{d,T+1}, \mathbf{Y}_1^T\right) \tag{10}$$

where the sum over $k$ requires $M$ MC samples, $\boldsymbol{\theta}^k_{T+1}$ from $p(\boldsymbol{\theta}_{T+1} \mid \mathbf{y}^j_{d,T+1}, \mathbf{Y}_1^T)$. This is computationally prohibitive in general given that the MC estimate must be calculated for each sample $\mathbf{y}^j_{d,T+1}$ and design $\mathbf{d}_{T+1}$.

 

Thus, when considering a large design space, one must resort to computationally efficient optimization approaches. For example, Müller (1999) considers a simulation-based approach to explore the design space in which the model is augmented to include the design as a random variable, in addition to the data and parameters. This approach is shown to be equivalent to solving the optimal design problem given above. As noted in Müller (1999) this approach is not guaranteed to give the 'optimal' design, but explores 'good' designs. Of course, this is the case with any practical optimization algorithm. The approach is also complicated by the problem of deducing the mode of the design distribution. Alternatively, we consider computational simplification by approximation of the expectation in (9). Specifically, we get an approximate estimate of the expected value, $\mathrm{E}[\sum_{i=1}^{n} h(\theta_{T+1}(s_i)) \,|\, \mathbf{y}_{d,T+1}^{j}, \mathbf{Y}_1^T]$ by making use of the dynamical properties of the spatio-temporal process, through an extended Kalman filter. That is, as described in the next subsection, we can utilize a computationally efficient extended Kalman filter in the context of the generalized linear dynamic model to obtain an estimate of this expectation.

*2.2.1. Extended Kalman filter for non-Gaussian dynamic models.* It is well-known in the time series and control literature that one can get approximate estimation for non-linear dynamical systems via the extended Kalman filter (EKF) (e.g. Grewal and Andrews, 1993). Fahrmeir and Kaufmann (1991) and Fahrmeir (1992) show that one can use the EKF for multivariate exponential families (2) and dynamical process (6) as well. In our mixed-model context, recall $\boldsymbol{\theta}_t = [\boldsymbol{\beta}' \quad \boldsymbol{\alpha}_t']'$. Thus, the dynamic model (6) can be rewritten,

$$\boldsymbol{\theta}_t = \tilde{\mathbf{F}}\boldsymbol{\theta}_{t-1} + \tilde{\boldsymbol{\eta}}_t \tag{11}$$

where

$$\tilde{\mathbf{F}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{bmatrix}, \quad \tilde{\boldsymbol{\eta}}_t = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\eta}_t \end{bmatrix} \quad \text{where} \quad \mathrm{var}(\tilde{\boldsymbol{\eta}}_t) \equiv \tilde{\mathbf{Q}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix}$$

Furthermore, let $\tilde{\mathbf{X}}_t = [\mathbf{X}_t \;\; \mathbf{Z}_t]$ and $\boldsymbol{\mu}_t = \mathbf{h}(\tilde{\mathbf{X}}_t\boldsymbol{\theta}_t)$. We then define the notation $\boldsymbol{\theta}_{t|t} \equiv \mathrm{E}(\boldsymbol{\theta}_t|\mathbf{Y}_1^t)$, $\boldsymbol{\Sigma}_{t|t} \equiv \mathrm{var}(\boldsymbol{\theta}_t \,|\, \mathbf{Y}_1^t)$, $\boldsymbol{\theta}_{t\,|\,t-1} \equiv \mathrm{E}(\boldsymbol{\theta}_t \,|\, \mathbf{Y}_1^{t-1})$, and $\boldsymbol{\Sigma}_{t\,|\,t-1} \equiv \mathrm{var}(\boldsymbol{\theta}_t \,|\, \mathbf{Y}_1^{t-1})$. One can obtain the following iterative EKF algorithm (Fahrmeir and Kaufmann, 1991):

- *Initialization step:*

$$\hat{\boldsymbol{\theta}}_{0|0} = \boldsymbol{\theta}_0 \qquad \hat{\boldsymbol{\Sigma}}_{0|0} = \tilde{\mathbf{Q}} \tag{12}$$

  where $\boldsymbol{\theta}_0$ and $\tilde{\mathbf{Q}}$ are specified.
- *Prediction step:*

$$\hat{\boldsymbol{\theta}}_{t|t-1} = \tilde{\mathbf{F}}\hat{\boldsymbol{\theta}}_{t-1|t-1} \tag{13}$$

$$\hat{\boldsymbol{\Sigma}}_{t|t-1} = \tilde{\mathbf{F}}\hat{\boldsymbol{\Sigma}}_{t-1|t-1}\tilde{\mathbf{F}}' + \tilde{\mathbf{Q}} \tag{14}$$

  where $\tilde{\mathbf{F}}$ is specified.

 

- *Correction step:*

$$\hat{\boldsymbol{\theta}}_{t|t} = \hat{\boldsymbol{\theta}}_{t|t-1} + \mathbf{G}_t(\mathbf{y}_t - \hat{\boldsymbol{\mu}}_t) \tag{15}$$

$$\hat{\boldsymbol{\Sigma}}_{t|t} = (\mathbf{I} - \mathbf{G}_t\hat{\mathbf{H}}_t')\hat{\boldsymbol{\Sigma}}_{t\,|\,t-1} \tag{16}$$

where

$$\mathbf{G}_t = \hat{\boldsymbol{\Sigma}}_{t|t-1}\hat{\mathbf{H}}_t[\hat{\mathbf{H}}_t'\hat{\boldsymbol{\Sigma}}_{t|t-1}\hat{\mathbf{H}}_t + \hat{\boldsymbol{\Sigma}}_t]^{-1} \tag{17}$$

is the Kalman gain, $\hat{\boldsymbol{\Sigma}}_t \equiv \boldsymbol{\Sigma}_t|_{\hat{\boldsymbol{\theta}}_{t|t-1}}$ is the conditional covariance matrix associated with the exponential family (2) evaluated at $\hat{\boldsymbol{\theta}}_{t|t-1}$, $\hat{\boldsymbol{\mu}}_t \equiv \boldsymbol{\mu}_t|_{\hat{\boldsymbol{\theta}}_{t|t-1}}$ is the exponential family conditional mean vector evaluated at $\hat{\boldsymbol{\theta}}_{t|t-1}$ and the matrix $\hat{\mathbf{H}}_t$ is given by

$$\hat{\mathbf{H}}_t = \frac{\partial \mathbf{h}(\tilde{\mathbf{X}}_t\boldsymbol{\theta}_t)}{\partial \boldsymbol{\theta}_t}\Big|_{\boldsymbol{\theta}_t = \hat{\boldsymbol{\theta}}_{t|t-1}} \tag{18}$$

where recall that $\mathbf{h}$ is the inverse link function, $\boldsymbol{\mu}_t = \mathbf{h}(\tilde{\mathbf{X}}_t\boldsymbol{\theta}_t)$.
Like all extended Kalman filters, this filter is approximate, based on local linearization arguments, and can be derived from linear Bayes arguments (e.g. Fahrmeir and Kaufman, 1991). It is assumed that the parameter matrices $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{Q}}$ and the starting value $\boldsymbol{\theta}_0$ are known. In practice these must usually be estimated. We will obtain the estimates from a fully Bayesian analysis, as described below.

*2.2.2. Extended Kalman filter assisted design.* Kalman filter based design for spatio-temporal problems is not new. In statistical environmental spatio-temporal design, Wikle and Royle (1999) show that the Kalman filter framework allows one to easily obtain dynamic designs of Gaussian processes with Gaussian measurement error. That is, one can easily obtain an estimate of $U(\mathbf{d}_{T+1} \mid \mathbf{Y}_1^T)$. The use of the EKF for non-linear filtering and design has also been considered in the literature (e.g. Titterington, 1980). More recently, Berliner *et al.* (1999) use the EKF for adaptive weather observations. As is typical for EKF applications, they applied the EKF methodology to problems with Gaussian data models and non-linear evolution models, rather than our case of non-Gaussian measurement models and Gaussian evolution models. In either case, the design problem is relatively simple once one has the EKF equations. Of course, the EKF approach gives approximate estimates of the posterior predictive mean and covariance since it is based on a local linearization. Thus, there is a tradeoff in terms of efficiency and approximation between the pure MC approach and the EKF approximation of, in our case, the data model. Furthermore, the traditional EKF approach is somewhat limited because one does not in fact know the parameter matrices. Although one can estimate these via the E-M algorithm (e.g. Shumway and Stoffer, 1982) if the dimensionality of the state and measurement process is not too great, in the EKF context the E-M algorithm would only provide estimates for parameters of the approximate (linearized) system.

Our approach is to combine the MC and EKF approaches. First, we recognize that the models of interest are best represented by a fully hierarchical Bayesian model (to account for parameter uncertainty and accommodate the curse of dimensionality in spatio-temporal problems). Thus, it is

 

appropriate to use MCMC approaches to obtain the posterior predictive distribution, $p(\boldsymbol{\theta}_{T+1} \mid \mathbf{Y}_1^T, \boldsymbol{v})$. Then, given a design $\mathbf{d}_{T+1}$, one generates samples of 'data' $\mathbf{y}_{d,T+1}$ from the posterior predictive data distribution. One can then use MCMC estimates of $\mathbf{F}$, $\mathbf{Q}$ and $\boldsymbol{\theta}_0$ in the EKF to get an estimate of the expected value $\mathrm{E}[\sum_{i=1}^{n} h(\theta_{T+1}(s_i)) \mid \mathbf{y}_{d,T+1}^j, \mathbf{Y}_1^T]$. That is, one gets an estimate of $\mathrm{E}(\boldsymbol{\theta}_{T+1} \mid \mathbf{Y}_1^T, \mathbf{F}, \mathbf{Q})$ from (15) and does the appropriate transformation to obtain an estimate of the desired expectation of $\sum_{i=1}^{n} h(\theta_{T+1}(s_i))$. At this point, an MC estimate of the expected utility (9) can be obtained. The hybrid MC-EKF dynamic design algorithm is summarized as follows:

1. Use MCMC to obtain samples $\boldsymbol{\theta}_{T+1}^j$ from the posterior predictive distribution: $p(\boldsymbol{\theta}_{T+1} \mid \mathbf{Y}_1^T, \boldsymbol{v})$
2. Choose design $\mathbf{d}_{T+1}$ from the set of possible designs $D$.
3. Given the design $\mathbf{d}_{T+1}$, generate 'data' $\mathbf{y}_{d,T+1}^j$ from the posterior predictive data distribution $p_d(\mathbf{y}_{T+1} \mid \boldsymbol{\theta}_{T+1}^j)$.
4. Use posterior mean MCMC estimates of $\mathbf{F}$, $\mathbf{Q}$, $\boldsymbol{\theta}_0$ and samples of 'data' from step 3 in the EKF algorithm (i.e. Equation (15)) to obtain an estimate of $\mathrm{E}[\sum_{i=1}^{n} h(\theta_{T+1}(s_i)) \mid \mathbf{y}_{d,T+1}^j, \mathbf{Y}_1^T]$ for each $j, j = 1, \ldots, N$.
5. Use samples from step 1 and 3, as well as the corresponding estimated means from step 4 to evaluate the expected utility (9), $U(\mathbf{d}_{T+1} \mid \mathbf{Y}_1^T)$.
6. Repeat steps 2–5 until the 'optimal' design (the one with the maximum estimated expected utility) is found.

We have not discussed how to choose the design in step 2 above. For simplicity, we choose a basic 'exchange' algorithm. Such algorithms are widely used in practice and many variations on the basic theme exist (e.g. see Cook and Nachtsheim, 1980; Atkinson and Federov, 1988; Nychka and Saltzman, 1998). Although these algorithms can be 'greedy' and may converge to local optima, for relatively small problems they typically find either the true optimum or 'good' designs that are close to the optimum.

## 3. APPLICATION: BROOD SURVEY DESIGN

The design problem of interest can be stated as follows: given areal survey data for July brood counts for 1994–1997 and areal survey May pair counts for 1994–1998 observed at transects in North Dakota, we seek to find the 20 transects out of 139 possible transects that one should sample (ground truth) broods in July 1998 to minimize the sum over all possible transects of the posterior mean counts for July 1998. As discussed in the Introduction, local brood abundance (i.e. on a segment) in the summer should be proportional to the density of breeding pairs the previous spring. Breeding pair counts using the same fixed-wing protocol described in the Introduction are collected in May of each year (Smith, 1995). Because brood sampling occurs in July, this initial observation of the distribution of the breeding population in May should be informative about patterns in brood abundance in July. Given pair density, production is likely to vary in response to landscape structure, and patterns in weather from May to July that are not explicitly accounted for. Thus, we require the spatio-temporal dynamic random effect, $\boldsymbol{\alpha}_t$, to account for these random effects. Figure 1 shows the May pair counts (left column) and July brood counts (right column) for the years 1994 through 1997. The circles are proportional to the counts, with the smallest circle (dot) corresponding to zero counts and the largest circle corresponding to a count of 87. One can see from these plots that there is a fair amount of dependence across time and space for both pair and brood counts. These plots also suggest that the pair to brood relationship is by no means perfect. We note that the sample units actually have areal support as they are counts over 18 mile long by 0.25 mile wide segments (area = 4.5 square miles). However,

 

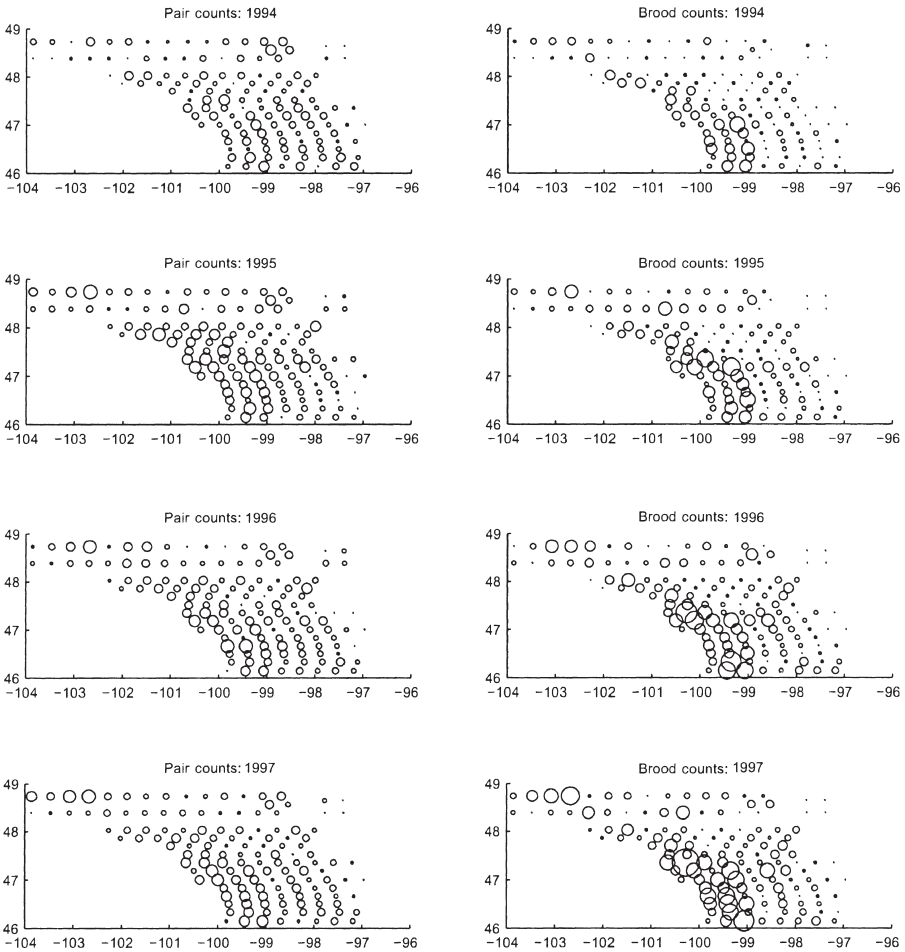


Figure 1. Left column: May pair counts for 1994–1997 for segments in North Dakota. Right column: corresponding July brood counts for 1994–1997. Note that the circle size is proportional to the count with the smallest circle equal to a count of 0 and the largest equal to a count of 87. The *x*-axis corresponds to E. longitude (degrees) and the *y*-axis to N. latitude (degrees)

because of their narrow width, a plot of the polygon boundaries would appear as connected lines and so we have indicated the center points of each segment.

These considerations suggest a log-Gaussian Cox process model (e.g. Møller *et al.*, 1998; Brix and Diggle, 2001) for brood abundance at spatial location $r_i^t$ and time $t$, $y_t(r_i^t)$, of the form

$$y_t(r_i^t) \mid \boldsymbol{\alpha}_t, \boldsymbol{\beta} \sim \text{ind. Poi}(\exp(\mathbf{x}_t'(r_i^t)\boldsymbol{\beta} + \mathbf{z}_t'\boldsymbol{\alpha}_t)) \tag{19}$$

where $\mathbf{x}_t(r_i) = (1 \;\; t \;\; MPC_t(r_i^t))'$ and $MPC_t(r_i^t)$ is the May pair count for time $t$ and location $r_i^t$, and $\mathbf{z}_t$ is an indicator vector (vector of zeros and ones), which extracts the $\alpha_t(s_i)$ element from the $n \times 1$ vector $\boldsymbol{\alpha}_t$ that corresponds to the observation location $r_i^t$. We include the temporal trend component because

 

there was a near linear increase over time of overall brood counts in the middle to late 1990s in response to large-scale climate factors. In vector form,

$$\mathbf{y}_t \mid \boldsymbol{\alpha}_t, \boldsymbol{\beta} \sim \text{ind. Poi}(\exp(\mathbf{X}_t\boldsymbol{\beta} + \mathbf{Z}_t\boldsymbol{\alpha}_t)) \tag{20}$$

where $\boldsymbol{y}_t$ is $m_t \times 1$, $\mathbf{X}_t$ is $m_t \times 3$, $\mathbf{Z}_t$ is $m_t \times n$ and $\boldsymbol{\alpha}_t$ is $n \times 1$. Note that both $\mathbf{X}_t$ and $\mathbf{Z}_t$ are in one-to-one relationship with the design $\mathbf{d}_t$. That is, the $m_t$ locations for which there are observations at time $t$ correspond to $\mathbf{d}_t$. In our case, for times $t = 1, \ldots, T$, $m_t = n$ since all observations are available. Of course, for time $T + 1$, $m_t < n$ as that is the purpose of the design problem. In general, $m_t$ need not equal $n$ for $t = 1, \ldots, T$ to implement this methodology.

The spatio-temporal dynamic process is given by

$$\boldsymbol{\alpha}_t \mid \boldsymbol{\alpha}_{t-1}, \mathbf{F}, \mathbf{Q} \sim \text{N}(\mathbf{F}\boldsymbol{\alpha}_{t-1}, \mathbf{Q}) \tag{21}$$

with $\mathbf{F} = \text{diag}(\mathbf{f})$ and $\mathbf{Q} = \sigma^2_\eta \mathbf{R}(\zeta)$, where the $n \times n$ spatial correlation matrix $\mathbf{R}(\zeta)$ is assumed to follow an exponential correlation model with spatial dependence parameter $\zeta$ (i.e. the correlation at distance dst is given by $r(\text{dst}) = \exp(-\text{dst}/\zeta)$). The relatively simple form of $\mathbf{F}$ in this example is necessary because of the limited number of years for which brood data were available (5 years). Yet, by allowing the diagonal elements of $\mathbf{F}$ to vary with spatial location, non-separable spatio-temporal dynamics can be accommodated. Note also that, by allowing the diagonal of $\mathbf{F}$ to vary with space, the marginal covariance of $\boldsymbol{\alpha}_t$ can be non-stationary. Finally, we assume $\boldsymbol{\alpha}_0 \sim \text{N}(\mathbf{0}, 10\mathbf{I})$.

To account for the variability of the parameters $\boldsymbol{\beta}$ we assign the prior, $\boldsymbol{\beta} \sim \text{N}(\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_\beta)$ and let the elements of $\mathbf{f}$ each be assigned the informative, but relatively vague, prior, $f_i \sim \text{N}(f_0, \sigma^2_0)$. We also specify an inverse gamma prior for $\sigma^2_\eta$, $\sigma^2_\eta \sim \text{IG}(q_\eta, r_\eta)$. Our hyperparameters, corresponding to relatively vague priors, are given as follows: $\boldsymbol{\beta}_0 = [0 \;\; 0 \;\; 0]'$, $\boldsymbol{\Sigma}_\beta = 100\mathbf{I}$, $f_0 = 0$, $\sigma^2_0 = 0.5$, $q_\eta = 2.25$ and $r_\eta = 1.6$ (in our parameterization, these IG parameters correspond to a prior mean of 0.5 and prior variance of 1). Our MCMC results were not overly sensitive to the choice of these priors. Finally, we used the non-informative reference prior for $\zeta$ suggested by Berger *et al.* (2001).

Recall that the EKF algorithm requires estimates of $\boldsymbol{\mu}_t$, $\boldsymbol{\Sigma}_t$ and $\mathbf{H}_t$. In our Poisson data model case, we have $\boldsymbol{\mu}_t = \exp(\tilde{\mathbf{X}}_t\boldsymbol{\theta}_t)$, and thus $\boldsymbol{\Sigma}_t = \text{diag}(\boldsymbol{\mu}_t)$ and $\mathbf{H}_t = \tilde{\mathbf{X}}_t\text{diag}(\boldsymbol{\mu}_t)$.

### *3.1. Implementation*

The MCMC implementation was straightforward. We used a Gibbs sampler with Metropolis updates for $\zeta$, $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}_t$, $t = 1, \ldots, T$, and conjugate Gibbs updates for all remaining parameters (e.g. see Royle *et al.*, 2002; Wikle, 2003). The MCMC was run for 100 000 iterations, the first 20 000 of which were discarded for burn-in. The pair count and time covariates were centered to facilitate convergence. The MCMC samples show no evidence of non-convergence based on trace plots and simple univariate convergence diagnostics. For purposes of the Monte Carlo estimation of the expected utility, 4000 samples were used, based on every 20th sample from the 80 000 post burn-in samples, to minimize Markov chain induced correlation in the MCMC samples. The initial design in the exchange algorithm was selected randomly from the 139 possible sampling locations. The exchange algorithm considered the 10 closest sampling locations for exchange. Computationally, selection of the 'optimal' design, including the cost of running the MCMC, was less than 15 min on a high-end 2003-vintage desktop workstation.

 

Table 1. Posterior summary of model parameters

| Parameter | Mean | Standard deviation | 2.5% tile | 50% tile | 97.5% tile |
|---|---|---|---|---|---|
| $\beta_1$ | 1.05 | 0.078 | 0.91 | 1.05 | 1.22 |
| $\beta_2$ | 0.34 | 0.050 | 0.25 | 0.34 | 0.44 |
| $\beta_3$ | 0.073 | 0.024 | 0.023 | 0.074 | 0.12 |
| $\sigma_\eta^2$ | 0.37 | 0.056 | 0.27 | 0.36 | 0.49 |

First, we consider the spatial dependence parameter $\zeta$. The posterior mean of this parameter was found to be 3.76 (miles) and the 2.5 and 97.5 percentiles were 1.74 and 5.96 miles, respectively. Given that the distance range of the observation locations is about 496 miles, it is clear that this posterior suggests there is very little residual spatial dependence in the $\boldsymbol{\alpha}$ process. (Note that the shortest distance between any two segment centers is about 11 miles, which corresponds to a minimal correlation of 0.05 and 0.16 at the mean and 97.5 percentile for $\zeta$, respectively.) Consequently, given that there is little evidence for residual spatial structure, subsequent results and discussion are based on the simpler model in which $\mathbf{Q} = \sigma_\eta^2 \mathbf{I}$. This result, suggesting little or no residual spatial structure, seems reasonable in the present context because the model describing variation in brood counts is conditioned on observed May pair counts. Apparently, much of the landscape structure contributing to relative variation in brood density is inherited through this conditioning on pair counts. Note that May pair counts will generally be obtained for the entire sample frame each year and thus there is no need (in the present context) to consider incomplete pair counts.

Table 1 shows a summary of the posterior samples for the scalar parameters. These summaries show that the linear time trend (corresponding to $\beta_2$) and the centered May pair counts ($\beta_3$) are significant in the sense that the 95 per cent credible intervals do not cover zero. In addition, Figure 2 shows the posterior means for the spatially varying autoregressive parameters ($f(s_i)$). These show substantial variability, but coherent spatial structure, implying that the spatio-temporal process ($\boldsymbol{\alpha}_t$) is non-separable, as expected. Note that this spatial structure in the $f$ parameters is in contrast to the lack

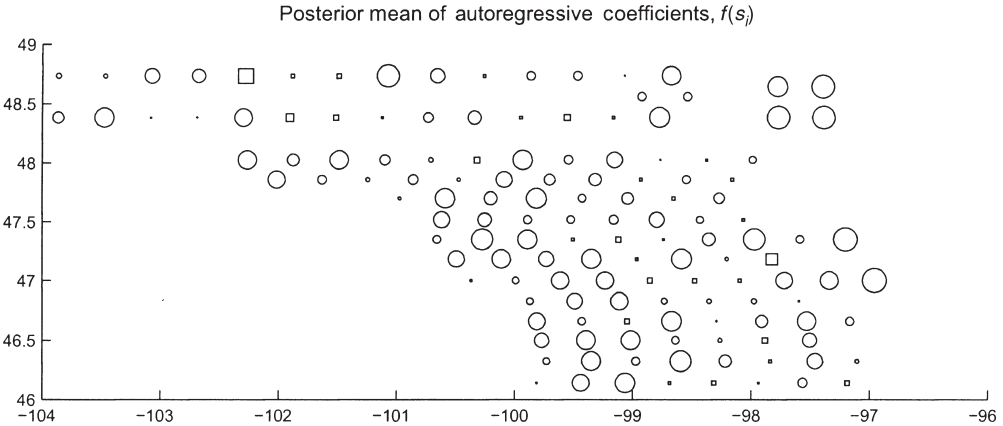


Figure 2. Posterior mean of spatially referenced autoregressive coefficients, $f(s_i)$, in the space-time dynamic model (21). Note that the circles and squares correspond to positive and negative values of the posterior mean of $f(s_i)$, respectively. The largest circle corresponds to a value of $f(s_i)$ equal to 1.2 and the largest square corresponds to a value of $f(s_i)$ equal to $-0.78$. The $x$- and $y$-axes are defined as in Figure 1

 

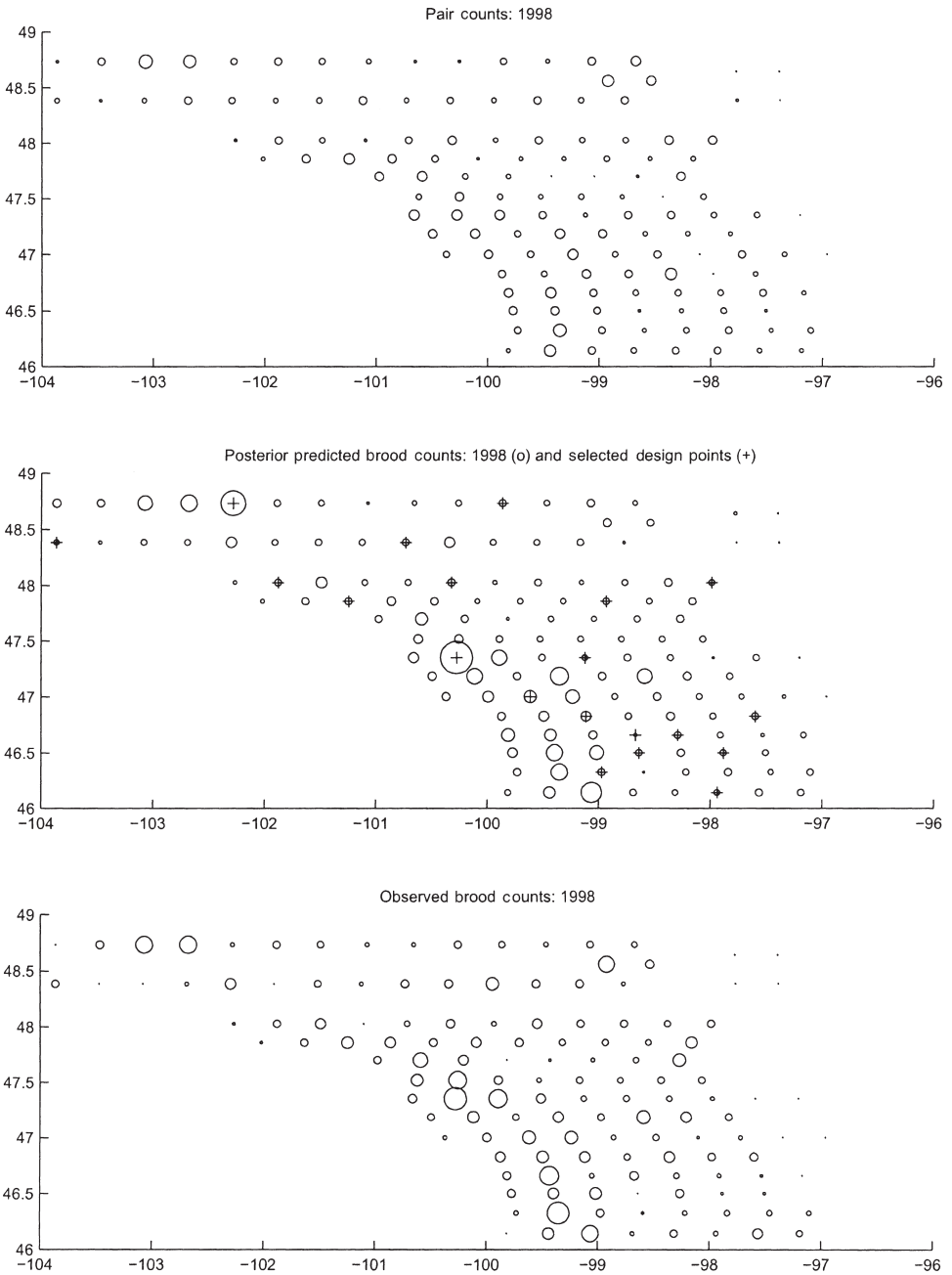


Figure 3. Top panel: May pair counts for 1998. Center panel: circles represent posterior mean of predicted brood counts for July 1998 based on brood counts from 1994–1997 and pair counts from 1994–1998. The pluses show the 20 locations that were chosen as the best design based on the hybrid MC-EKF algorithm. Bottom panel: observed July brood counts for 1998. Circles and axes are defined as in Figure 1

 

of structure in **Q** discussed above. We also note that the larger values of $f(s_i)$ correspond to locations that show relatively persistent high or low brood counts, but that are not completely captured by the May pair counts. Note, however, that the very large values of $f$ (greater than 1) occur in the upper right portion of the map, corresponding to sites that showed very few brood counts over the period of study.

The top panel of Figure 3 shows the pair counts observed for 1998. The middle panel shows the posterior predicted mean brood counts from the MCMC and the associated best design obtained from the MC-EKF hybrid algorithm. The bottom panel shows the observed brood counts for 1998 for comparison. These 1998 brood count data were not included in the analysis. From the MCMC prediction, it seems that the space–time dynamical model is doing a reasonable job of predicting the general locations for which the brood counts were high. It is also clear that the selected 20 design locations are not necessarily intuitive. However, as shown by Wikle and Royle (1999) through simulation, in the presence of non-separable spatio-temporal dependence, designs are not typically intuitive, even in the case of Gaussian data. Note that we also considered 1000 designs corresponding to 20 locations selected randomly (with equal probability) from the 139 potential sampling sites. The estimated design criterion ranged from 43 379 to 268 730 out of these sample designs (with a median of 219 450). For comparison, the 'best' design obtained from the exchange algorithm and shown in Figure 3 had an estimated design criterion of 38 052. Figure 4 shows the MC samples of EKF estimated total (summed over space) mean counts under the chosen 'best' design. The mean and median from these samples are 1155 and 1094, respectively.

Finally, for comparison, Figure 5 shows the 'best' designs given that 5, 10 and 15 sites could be sampled. The estimated design criterion in these cases was 66 527, 41 572 and 39 808, respectively. Thus, even with only five monitoring locations chosen according to this procedure, one obtains better results relative to the chosen utility function than the vast majority of the random designs with 20 locations. Furthermore, these analyses suggest that there may be little to gain in terms of our design criterion if one considers 20 monitoring locations rather than 15 (i.e. the design with 15 locations is 96 per cent as efficient as the design with 20 locations.)

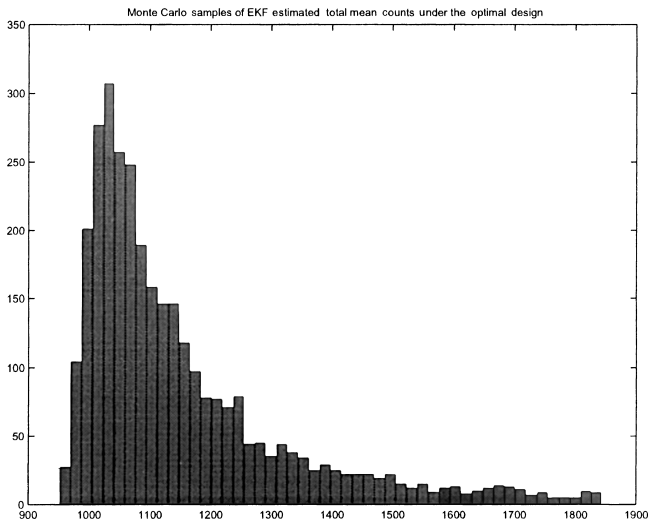


Figure 4. Histogram of Monte Carlo samples of the EKF-estimated total (over space) mean pair counts for July 1998, given by the selected 'best' design shown in Figure 3

 

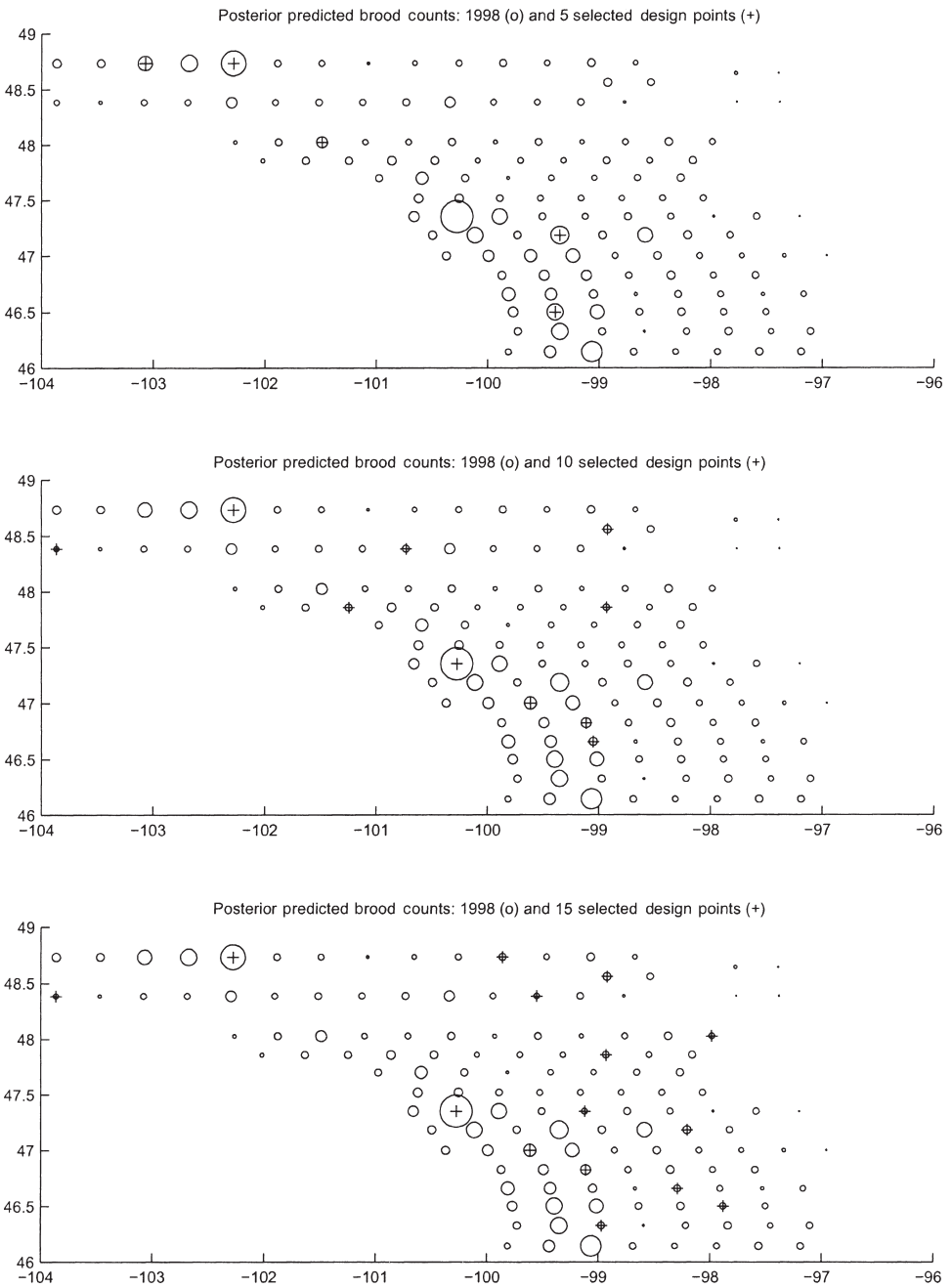


Figure 5. Top panel: MCMC-predicted brood counts for July 1998 and the optimal design (+) chosen for 5 sampling locations. Middle panel: same as top panel except design is based on 10 sampling locations. Bottom panel: same as top panel except design is based on 15 sampling locations. Circles and axes are defined as in Figure 1

 

## 4. DISCUSSION

The dynamic design problem for spatial sampling networks and non-Gaussian spatio-temporal data sets is relatively easy to specify in terms of a simulation-based approach. However, implementation of such approaches can be very difficult and time-consuming from a computational perspective. Thus, it is important to have relatively efficient practical procedures that can facilitate computation. The hybrid MC-EKF simulation-based design approach described here is a relatively efficient procedure for obtaining near optimal dynamic designs for non-Gaussian spatio-temporal data. That is, we provide an approximate solution to the problem that is practical to implement.

The need for practicality is essential for the ecological brood estimation problem considered here. Ground-based sampling of broods is very time-consuming and expensive. Thus, given that only a small number of samples can be selected, it is critical that thought be given to sample selection. As shown in the previous section, 'random samples' (i.e. designs chosen from random selection of sampling locations) do much more poorly than the near optimal design approach presented here for these data.

More generally, with regard to the design criteria, there is much that we do not know or understand about choosing a utility (i.e. design criterion) for estimating and predicting ecological count data. In the present example, the goal of the motivating management activity is to obtain estimates of the total number of broods occurring in a geographic area, and thus minimizing the variance of the estimated total given an incomplete sample seems logical. In general, however, there is a need to study more closely other potential utility functions. For example, if desired, the utility function chosen for this problem could account for monetary considerations (e.g. Sansó and Müller, 1998).

Clearly, more complicated dynamical models (e.g. Wikle, 2003) than that utilized in the brood sampling example presented here could be considered in this framework. For example, it would be simple to allow the $\boldsymbol{\beta}$ parameters to be time varying in this context, if necessary. In addition, alternative design selection algorithms could be considered, as opposed to the exchange algorithm. Finally, the recent inhomogeneous Markov chain approach to simulation design, discussed in Müller *et al.* (2002) could be considered here. These issues will be the subject of future work.

ACKNOWLEDGEMENT

Wikle's research was supported under National Science Foundation grant DMS-0139903. The authors would like to thank Mark Berliner and two anonymous reviewers for helpful comments on an early version of this work.

REFERENCES

Arbia G, Lafratta G. 1997. Evaluating and updating the sample design in repeated environmental surveys: monitoring air quality in Padua. *Journal of Agricultural, Biological and Environmental Statistics* **2**(4): 451–466.

Atkinson AC, Federov VV. 1988. Optimum design of experiments. In *Encyclopedia of Statistics* (Supplemental Volume), Kotz S, Johnson NI (eds). Wiley: New York; 107–114.

Berger JO, De Oliveira V, Sansó B. 2001. Objective Bayesian analysis of spatially correlated data. *Journal of the American Statistical Association* **96**: 1361–1374.

Berliner LM, Lu Z-Q, Snyder C. 1999. Statistical design for adaptive weather observations. *Journal of the Atmospheric Sciences* **56**: 2536–2552.

Brix A, Diggle PJ. 2001. Spatiotemporal prediction for log-Gaussian Cox processes. *Journal of the Royal Statistical Society B* **63**: 823–841.

Brix A, Møller J. 2001. Space-time multi type log Gaussian Cox processes with a view to modelling weed data. *Scandinavian Journal of Statistics* **28**: 471–488.

 

Bueso JM, Angulo JM, Alonso FJ. 1998. A state-space model approach to optimum spatial sampling design based on entropy. *Environmental and Ecological Statistics* **5**: 29–44.
Chaloner K, Verdinelli I. 1995. Bayesian experimental design: a review. *Statistical Science* **10**: 273–304.
Chao C-T, Thompson SK. 2001. Optimal adaptive selection of sampling sites. *Environmetrics* **12**: 517–538.
Christensen OF, Waagepetersen R. 2002. Bayesian prediction of spatial count data using generalized linear mixed models. *Biometrics* **58**: 280–286.
Cook RD, Nachtsheim CJ. 1980. A comparison of algorithms for constructing exact D-optimal designs. *Technometrics* **22**: 315–324.
Cox DD, Cox LH, Ensor KB. 1996. Spatial sampling for the environment. *Environmental and Ecological Statistics* **4**: 219–233.
Fahrmeir L. 1992. Posterior mode estimation by extended Kalman filtering for a multivariate dynamic generalized linear models. *Journal of The American Statistical Association* **87**: 501–509.
Fahrmeir L, Kaufmann H. 1991. On Kalman filtering, posterior mode estimation and Fisher scoring in dynamic exponential family regression. *Metrika* **38**: 37–60.
Federov VV, Hackl P. 1997. *Model-Oriented Design of Experiments*. Lecture Notes in Statistics. Springer-Verlag: New York.
Federov VV, Müller W. 1989. Comparison of two approaches in the optimal design of an observation network. *Statistics* **20**(3): 339–351.
Federov VV, Nachtsheim C. 1995. Optimal designs for time-dependent responses. In *Proceedings of MODA4*, Kitsos CP, Muller WG (eds). Physica Verlag: Heidelberg.
Grewal MS, Andrews AP. 1993. *Kalman Filtering, Theory and Practice*. Prentice Hall: Englewood Cliffs, New Jersey.
Guttorp P, Le ND, Sampson PD, Zidek JV. 1993. Using entropy in the redesign of an environmental monitoring network. In *Multivariate Environmental Statistics*, Patil GP, Rao CR (eds). Amsterdam: North Holland; 175–202.
Haas TC. 1992. Redesigning continental-scale monitoring networks. *Atmospheric Environment* **26A**: 3323–3333.
Le ND, Zidek JV. 1994. Network designs for monitoring multivariate random spatial fields. In *Recent Advances in Statistics and Probability*, Vilaplana JP, Puri ML (eds). VSP: Leiden, The Netherlands; 191–206.
Møller J, Syversveen AR, Waagepetersen RP. 1998. Log Gaussian Cox processes. *Scandinavian Journal of Statistics* **25**: 451–482.
Müller P. 1999. Simulation based optimal design. In *Bayesian Statistics 6*, Berger JO, Bernardo JM, Dawid AP, Smith AFM (eds). Oxford University Press: Oxford; 459–474.
Müller WG. 2000. *Collecting Spatial Data: Optimum Design of Experiments for Random Fields*, 2nd edn. Physica Verlag: Heidelberg; 196 pp.
Müller WG, Pázman A. 2003. Measures for designs in experiments subject to correlated errors. *Biometrika* **90**: 423–434.
Müller P, Sansó B, De Iorio M. 2002. Optimal Bayesian design by inhomogeneous Markov chain simulation. *Tec. Report 2002–06, CESMa*.
Nichols JD, Hines JE, Sauer JR, Fallon FW, Fallon JE, Heglund PJ. 2000. A double-observer approach for estimating detection probability and abundance from point counts. *Auk* **117**(2): 393–408.
Nychka D, Saltzman N. 1998. Design of air-quality monitoring networks. In *Case Studies in Environmental Statistics*, Nychka D, Piegorsch W, Cox L, (eds). Springer: New York; 51–76.
Oehlert GW. 1996. Shrinking a wet deposition network. *Atmospheric Environment* **30**: 1347–1357.
Rosenstock SS, Anderson DR, Giesen, KM, Leukering T, Carter MF, 2002. Landbird counting techniques: current practices and an alternative. *Auk* **119**(1): 46–53.
Royle JA. 2004. N-mixture models for estimating population size from spatially replicated counts. *Biometrics* **60**: 108–115.
Royle JA, Link WA, Sauer JR. 2002. Statistical mapping of count survey data. In *Predicting Species Occurrences: Issues of Scale and Accuracy*, Scott JM, Heglund PJ, Morrison M, Raphael M, Haufler J, Wall B (eds). Island Press: Covello, CA; 625–638.
Sansó B, Müller P. 1998. Redesigning a network of rainfall stations. In *Case Studies in Bayesian Statistics*, vol. IV, Springer-Verlag: NY; 383–394.
Shumway RH, Stoffer D. 1982. An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis* **3**: 253–264.
Smith GW. 1995. A critical review of the aerial and ground surveys of breeding waterfowl in North America. *U.S. Dept. of the Interior Biological Science Report 5*.
Thompson SK, Seber AF. 1996. *Adaptive Sampling*. John Wiley & Sons Inc.: New York.
Titterington DM. 1980. Aspects of optimal design in dynamic systems. *Technometrics* **22**(3): 287–299.
Wikle CK, 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology* **84**: 1382–1394.
Wikle CK, Royle JA. 1999. Space-time models and dynamic design of environmental monitoring networks. *Journal of Agricultural, Biological, and Environmental Statistics* **4**: 489–507.

 