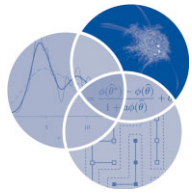
Focus Article



Spatio-temporal processes

Jane L. Harvill*

We present an overview of the literature on the analysis of spatio-temporal processes with a nonseparable covariance structure. We focus on those methods that rely heavily on computing for the estimation or inference. Topics are classified into frequentist approaches, which rely on expectation–maximization algorithms, and hierarchical Bayesian approaches, which rely on Markov chain Monte Carlo. We also present discussions on other computational issues related to the analysis of spatio-temporal data. © 2010 John Wiley & Sons, Inc. *WIREs Comp Stat* 2010 2 375–382 DOI: 10.1002/wics.88

Keywords: spatio-temporal processes; Kalman filter; EM algorithm; Markov chain Monte Carlo; parallel computing

As the name implies, 'spatio-temporal processes' are the processes which evolve in time and space. The modeling of spatio-temporal distributions resulting from these dynamic processes is critical in many scientific and engineering fields such as environmental sciences, climate prediction and meteorology, population biology, epidemiology, criminology, image analysis, and agriculture, to name a few. Many approaches for modeling spatio-temporal processes have been proposed. For instance, one can separate the time and spatial dimensions so that a separate spatial analysis is at each point in time, or vice versa. However separating the time and spatial elements is ill-advised (see, e.g., Refs 1 or 2). In this article, we focus on computational aspects of methods for analyzing spatio-temporal data that do not separate the time and spatial dimensions.

THE JOINT SPACE–TIME FRAMEWORK

Consider a time domain $\mathcal{T} \subset \Re^1$, a finite time-dependent spatial domain $\mathcal{D}(t) \subset \Re^d, t \in \mathcal{T}$, and without loss of generality, let $d = 2$. A spatio-temporal random process $Y(\boldsymbol{s}, t)$ is a random variable that can take a series of outcome values at any location $\boldsymbol{s} \in \mathcal{D}(t)$ and instant in time $t \in \mathcal{T}$, i.e., $\{Y(\boldsymbol{s}, t) : \boldsymbol{s} \in \mathcal{D}(t),\ t \in \mathcal{T}\}$. It is common to model the variability in $\{Y(\boldsymbol{s}, t)\}$ as a random field.[1] It is also commonly assumed that $\mathcal{D}(t) = \mathcal{D}$, although it is often not necessary.

Until the mid- to late-1990s, covariance functions most commonly used for describing spatio-temporal processes were separable, i.e., the covariance function might decompose into the sum or the product of a purely spatial and a purely temporal covariance functions; see, for example, Refs 3,4. This separability does not allow for space–time interactions. Moreover, as noted by Stein,[5] separable covariance functions generally imply that small changes in the locations of observations can lead to large changes in the correlations between certain linear combinations of observations.

Analyzing a spatio-temporal process as spatial data with an 'extra dimension' is another approach that has been proposed. Consider, however, that there are fundamental differences between time and space. Time has a clear ordering—a past, present, and future—while space does not. Because of the intrinsic ordering in time, isotropy, which is well-defined in the spatial context, has no meaning in the space–time context. This additional complexity, along with a lack of computational tools in standard software, leads many to over-simplify and proceed by separating the two elements of space and time and to perform either a (1) separate spatial analysis for each time point or (2) separate temporal analysis of each location. These conditional approaches isolate a particular time point or location and apply standard techniques for the resulting data. An additional stage of the analysis usually includes combining the results of the first stage of the analysis, e.g., nonlinear mixed model applications for clustered data. While these two stage approaches are appealing, they have serious drawbacks.[6]

We discuss methods that employ the joint analysis of spatio-temporal data. As with most areas in statistics, approaches to modeling spatio-temporal data are varied, but can be categorized into two schools of thought: frequentist and Bayesian. In keeping with the scope of this volume, we focus on those results that rely heavily on computational techniques.

*Correspondence to: jane_harvill@baylor.edu

Department of Statistical Science, Baylor University, Waco, TX, USA

DOI: 10.1002/wics.88

 © 2010 John Wiley & Sons, Inc.

Nonseparable covariance functions

The overarching theme of developing nonseparable covariance functions is that the dynamic evolution in time and space adds complexity such that the dependence structure across both space and time is nontrivial. In a seminal paper, Cressie and Huang[7] developed nonseparable space–time stationary covariance functions that model space–time interactions. However, their approach depends upon Fourier transform pairs in $\Re^d$, and is thus restricted to a small class of functions for which a closed-form solution to the d-variate Fourier integral is known. Gneiting[8] avoids the limitation of Ref 7, providing a more general class of valid space–time covariance models. In a series of papers, Iaco et al.,[9–11] continued the development of new classes of stationary space–time covariance functions on $\Re^d \times \Re$, and Iaco et al.[11] provided algorithms for fitting such models. Fuentes et al.[12] use the spectral representation of a process to develop a general, flexible parametric class of nonseparable space–time covariance models. Stein[5] further extends this body of work on nonseparable covariance functions by developing nonseparable covariance functions that are smooth everywhere except possibly at the space–time origin, and which allow for different degrees of smoothness away from the space–time origin. Stein[5] shows how one can generate space–time covariance functions that are spatially isotropic but not fully symmetric in space–time by taking derivatives of spatially isotropic fully symmetric models. A good review of the literature on covariance functions for spatio-temporal processes can be found in Refs 13 and 14.

Following Xu and Wilke,[15] the covariance function $C(\boldsymbol{s}_1, \boldsymbol{s}_2; t_1, t_2)$ and semivariogram $\gamma(\boldsymbol{s}_1, \boldsymbol{s}_2; t_1, t_2)$, assuming they exist, are defined by

$$\begin{aligned} C(\boldsymbol{s}_1, \boldsymbol{s}_2; t_1, t_2) &= \mathrm{Cov}\{Y(\boldsymbol{s}_1, t_1), Y(\boldsymbol{s}_2, t_2)\} \\ &= \mathrm{E}[\{Y(\boldsymbol{s}_1, t_1) - \mathrm{E}Y(\boldsymbol{s}_1, t_1)\}\{Y(\boldsymbol{s}_2, t_2) \\ &\quad - \mathrm{E}Y(\boldsymbol{s}_2, t_2)\}] \\ \gamma(\boldsymbol{s}_1, \boldsymbol{s}_2; t_1, t_2) &= \frac{1}{2}\mathrm{Var}\{Y(\boldsymbol{s}_1, t_1) \\ &\quad - Y(\boldsymbol{s}_2, t_2)\}, \boldsymbol{s}_1, \boldsymbol{s}_2 \in \mathcal{D}, \quad t_1, t_2 \in \mathcal{T} \end{aligned}$$

Properties and characterizations of the covariance function and semivariogram can be found in, e.g., Ref 15. Specifically, for the covariance function to be well-defined, it is necessary to assume that $\mathrm{Var}[Y(\boldsymbol{s}, t)] < \infty$ for all $(\boldsymbol{s}, t) \in \mathcal{D} \times \mathcal{T}$, which implies that the first two moments exist. Then the process is second-order stationary in space–time if $\mathrm{E}Y(\boldsymbol{s}_1, t_1) = \mathrm{E}Y(\boldsymbol{s}_2, t_2)$ and $C(\boldsymbol{s}_1, \boldsymbol{s}_2; t_1, t_2)$ depends only on $\boldsymbol{s}_1 - \boldsymbol{s}_2$ and $t_1 - t_2$. In this case, the corresponding covariance function can be denoted by $C(\boldsymbol{s}, t) = \mathrm{Cov}\{Y(\boldsymbol{s}_0, t_0), Y(\boldsymbol{s}_0 + \boldsymbol{s}, t_0 + t)\}$. The process is intrinsically stationary in space–time if $\mathrm{E}\{Y(\boldsymbol{s}_1, t_1) - Y(\boldsymbol{s}_2, t_2)\} = 0$ and $\gamma(\boldsymbol{s}_1, \boldsymbol{s}_2; t_1, t_2)$ depends only upon $\boldsymbol{s}_1 - \boldsymbol{s}_2$ and $t_1 - t_2$. In the case of intrinsic stationarity, the semivariogram can be written as, simply, $\gamma(\boldsymbol{s}, t)$. Moreover, if the process $Y(\boldsymbol{s}, t)$ is second-order stationary, it is easily shown that the semivariogram can be written as

$$\gamma(\boldsymbol{s}, t) = C(\boldsymbol{0}, 0) - C(\boldsymbol{s}, t) \quad \text{for all } (\boldsymbol{s}, t) \in \mathcal{D} \times \mathcal{T}$$

Spatio-temporal models

Huang and Cressie[16] were among the first to develop a full implementation of a space–time dynamic model with a Kalman filter for a separable space–time structure. Cane et al.[17] implemented a reduced dimension space–time dynamic model from a truncated set of multivariate empirical orthogonal basis functions derived from a long model run without assimilation. Mardia et al.[18] provide the details and full implementation of the general reduced dimension model they call the 'kriged Kalman filter.' These reduced dimension space–time Kalman filter approaches provide techniques for modeling spatially nonstationary and space–time nonseparable process, and can be applied to large data sets. However, they did not provide for a small-scale spatial structure that does not evolve temporally. Wikle and Cressie[19] considered not only the measurement error and space–time dynamic components, but also a non-dynamic term that captures small-scale spatial variability.

A state-space approach

Following Xu and Wilke,[2] let $\boldsymbol{Z} = (Z(\bar{\boldsymbol{s}}_1, t), \ldots, Z(\bar{\boldsymbol{s}}_{m_t}, t))'$ be an $m_t \times 1$ vector containing data values at m_t spatial locations $\boldsymbol{s}_i$ at time t. The vector $\boldsymbol{Y} = (Y(\boldsymbol{s}_1, t), \ldots, Y(\boldsymbol{s}_n, t))'$ is an $n \times 1$ vector for an unobservable spatio-temporal state process at some fixed network of locations $\boldsymbol{s}_1 \ldots, \boldsymbol{s}_n$. The state process $\boldsymbol{Y}$ is of primary interest, and the two sets of spatial locations $\bar{\boldsymbol{s}}_i, \boldsymbol{s}_i \in \mathcal{D}$ need not be the same. Then $\boldsymbol{Z}$ and $\boldsymbol{Y}$ can be expressed by the system of equations

$$\boldsymbol{Z}_t = \boldsymbol{K}_t \boldsymbol{Y}_t + \boldsymbol{\varepsilon}_t \tag{1}$$

$$\boldsymbol{Y}_t = \boldsymbol{H}\boldsymbol{Y}_{t-1} + \boldsymbol{\eta}_t \tag{2}$$

for $t = 1, \ldots, T$. In this representation, Eq. (1) is referred to as the 'measurement equation,' and Eq. (2) the 'state equation.' The matrix $\boldsymbol{K}_t$ is a known $m_t \times n$

 © 2010 John Wiley & Sons, Inc.

19390068, 2010, 3, Downloaded from https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.88 by University of Massachusetts Dartmouth, Wiley Online Library on [22/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

matrix that maps the data $\boldsymbol{Z}_t$ to the process $\boldsymbol{Y}_t$. The measurement noise $\boldsymbol{\varepsilon}_t$ is zero-mean, uncorrelated in time and Gaussian with an $m_t \times m_t$ covariance matrix $\boldsymbol{R}_t$. The dynamics are described in the state equation (2) via a first-order Markov process with transition matrix $\boldsymbol{H}$. The shocks $\boldsymbol{\eta}_t$ are spatially colored, temporally white, and Gaussian, with mean zero and a common $n \times n$ covariance matrix $\boldsymbol{Q}$. The process starts with a Gaussian spatial process $\boldsymbol{Y}_0$ with mean $\boldsymbol{\mu}_0$ and $n \times n$ covariance matrix $\boldsymbol{\Sigma}_0$. The parameters of the system in Eqs. (1) and (2) are represented as $\boldsymbol{\Theta} = \{\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \boldsymbol{H}, \boldsymbol{Q}, \boldsymbol{R}_t\}$. Often, researchers use Bayesian hierarchical models (see the section following) for dealing with the high-dimensionality, i.e., problematic in most space–time applications. Xu and Wilke[2] show it is often possible to fit such models and estimate $\boldsymbol{\Theta}$ using a Kalman filter and expectation–maximization (EM) algorithm.

If the values of the parameters $\boldsymbol{\Theta}$ are known, then the Kalman filter and Kalman smoother[20–22] can be used to obtain the conditional mean and covariance of the state variable $\boldsymbol{Y}_t$. Xu and Wilke[2] give a nice review. On the other hand, if they are not known, they can be estimated using method of moments. Those estimates are then 'plugged into' Eq. (1) to implement the Kalman filter.[19] However, the Kalman recursion has, as a by-product, the computed value of the likelihood. From this, the maximum likelihood estimates of $\boldsymbol{\Theta}$ can be computed numerically[23] or via an EM algorithm (details in Ref 24). The EM iteration consists of two steps: an 'E-step' and an 'M-step.' In the context of spatio-temporal data, and the state space model in Eqs. (1) and (2), the complete data vector is given by $(\boldsymbol{Y}_0, \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_T, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_T)$ with likelihood $L_{Y,Z}$, say. Given the current value of the parameters $\boldsymbol{\Theta}^{(j-1)}$, the E-step computes the expected value of $L_{Y,Z}$. Denote this expectation $g(\boldsymbol{\Theta} \mid \boldsymbol{\Theta}^{(j-1)})$. In the M-step, an update $\boldsymbol{\Theta}^{(j)}$ is chosen to ensure that the likelihood increases monotonically, i.e., $\boldsymbol{\Theta}^{(j)}$ is chosen to satisfy $g(\boldsymbol{\Theta}^{(j)} \mid \boldsymbol{\Theta}^{(j-1)}) < g(\boldsymbol{\Theta}^{(j-1)} \mid \boldsymbol{\Theta}^{(j-1)})$. If the likelihood is bounded, then the iterates $\boldsymbol{\Theta}^{(j)}$ eventually converge to the maximum likelihood estimator. If $\boldsymbol{\Theta}^{(j)}$ is also a minimum of $g(\boldsymbol{\Theta} \mid \boldsymbol{\Theta}^{(j-1)})$, the algorithm is the 'standard EM algorithm.' Otherwise, it is known as the 'general EM algorithm.'

In the case of the state-space model for spatio-temporal applications, the data vector $\boldsymbol{Z}_t$ is usually of high dimension, and correspondingly, the parameter $\boldsymbol{\Theta}$ is also of high dimension. Xu and Wilke[2] provide a method for dimension reduction that reparameterizes $\boldsymbol{\Theta}$ by exploiting the structure of the process. They provide several approaches for specifying realistic models for $\boldsymbol{R}_t, \boldsymbol{Q}$, and $\boldsymbol{H}$. Their first approach uses an expectation-conditional maximization algorithm, in which they replace the M-step with two simpler conditional maximization steps. The second approach is a general EM algorithm based on one Newton–Raphson step proposed in Ref 25. In that same paper, Xu and Wilke[2] provide algorithms for the parameterizations of the $\boldsymbol{R}_t$ matrix, the $\boldsymbol{Q}$ matrix, and the $\boldsymbol{H}$ matrix.

HIERARCHICAL BAYESIAN SPACE–TIME MODELS

In the Bayesian framework, modeling spatio-temporal processes requires the implementation of hierarchical Bayesian methodology, also referred to as Bayesian maximum entropy. Many examples of hierarchical Bayesian space–time modeling are found in the literature.[26–36] In general, these models take into account the uncertainty in the observations, in the specification of the spatio-temporal process, and in the knowledge of the parameter values which describe the dependence in space and time. Because of the high level of complexity in a joint specification of all of these degrees of uncertainty, hierarchical models rest on the ability to factor the joint distribution of the data, process, and parameters into a product of conditional distributions. This Bayesian hierarchical structure allows the process to be modeled in terms of means at various stages, as opposed to an unwieldy joint covariance matrix. The models also provide the analyst with opportunities to explore trade-offs between rich time-lagged site-specific time series models and models based on more direct specification of spatial structure, but at short time lags.[28]

The paper of Handcock and Wallis[26] is one of the seminal papers which applies Bayesian methodology to spatio-temporal data. The authors take a traditional approach, choosing priors for which numerical integration can be performed to determine posteriors. However, Handcock and Wallis[26] give no mention of the Markov chain Monte Carlo (MCMC) approach for simulating approximate posteriors. Wilke et al.[28] were the first to implement a fully Bayesian (hierarchical) space–time dynamic model. However, their approach does not allow for the case when there is a significant small-scale spatial structure that does not have a discernible temporal evolution. Such behavior is commonly seen in physical and biological processes. Recognizing this shortcoming, Wikle and Cressie[19] consider explicitly the measurement error and space–time dynamic components, as well as a non-dynamic term that captures small-scale spatial variability. They derive a general empirical Bayesian predictor based on their model, which has as special cases, the various space–time models of the approaches in the references

 © 2010 John Wiley & Sons, Inc.

19390068, 2010, 3, Downloaded from https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.88 by University of Massachusetts Dartmouth, Wiley Online Library on [22/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

they cite within their paper, many of which are included here. We explain their approach in the following paragraphs.

Following Wikle and Cressie,[19] assume that data $Z(s_1, t_1), \ldots, Z(s_N, t_N)$ are obtained from an observable and spatially continuous process $\{Z(s,t)\}$, where $s \in \mathcal{D}$ and $t \in \mathcal{T}$, where $\mathcal{T}$ is a discrete index of times. The observable process has a component of measurement error expressed through the measurement equation:

$$Z(s,t) = Y(s,t) + \varepsilon(s,t)$$

where $Y(s,t)$ is an unobservable process, i.e., 'smoother' than $Z(s,t)$. The goal is to predict the process $Y(s,t)$ regardless of when and where $Z(s_1, t_1), \ldots, Z(s_N, t_N)$ are observed. Further assume that the process $Y(s,t)$ can be written as:

$$Y(s,t) = Y_K(s,t) + \nu(s,t)$$

where $\nu(s,t)$ is a component of variance representing small-scale spatial variation that does not have a temporally dynamic structure and $Z_K(s,t)$ evolves according to the state equation:

$$Y_K(s,t) = \int_{\mathcal{D}} w_s(u) Y_K(u, t-1)\, du + \eta(s,t)$$

Here $\eta(s,t)$ is a 'spatially descriptive' component, $w_s(u)$ is a function representing the interaction between the state process Y_K at location u and time $t-1$ and Y_K at location s and time t, called the 'temporally dynamic' component. Through a series of linear representations and reasonable model assumptions, Wikle and Cressie[19] develop a predictor that minimizes mean-square prediction error for the class of all linear predictors, even when the Gaussian assumption is not appropriate. The predictor can be expressed recursively in terms of a Kalman filter.

The approach for model parameter estimation in Ref 19 corresponds to viewing the Kalman filter as an empirical Bayesian technique. Fully Bayesian hierarchical approaches, as in Ref 28, can also be implemented. But as pointed out in Ref 19, there is a trade-off between computational efficiency, with the empirical Bayesian approach, and statistical precision, with the fully Bayesian approach. The approach proposed in Ref 19 is motivated by the need to model large spatio-temporal data sets, and so they focus more on the computationally efficient algorithm.

These two approaches[19] and Wikle[28] are the fundamental underpinnings of the majority of Bayesian methodology applied to spatio-temporal data. The underlying assumption is that the process $\{Y(s,t)\}$ is a Gaussian random field. Work has been done that expands the approaches to non-Gaussian processes. See, e.g., Refs 32,36,37 or 38.

Finally, Fahrmeir et al.[34] propose extensions of penalized spline generalized additive models for analyzing space–time regression data and study them from a Bayesian perspective. Using their approach they illustrate inferential methods using either full Bayes or empirical Bayesian posterior analysis. For the fully Bayesian technique, their MCMC techniques are only a slight extension of previous work. For inference based on empirical Bayesian methodology, they develop a computationally efficient solution on the basis of a generalized linear mixed model representation. This representation can be viewed as posterior mode estimation and is closely related to penalized likelihood estimation in the frequentist setting.

Markov chain Monte Carlo algorithms

The high dimension and complexity of the models discussed above require the use of some form of an MCMC algorithm to obtain estimates of posterior and predictive quantities. However, while MCMC is a powerful and useful technique, rich models in very high dimensions lead to technical problems in their implementation. As a result, while discussing the implementation of MCMC to spatio-temporal problems, different authors propose different variations on the MCMC theme. For example, Waller et al.[27] note that the Gibbs sampler[39] is not well-suited to their approach, and so enhances it by using a Metropolis algorithm[40,41] to obtain the necessary samples. They begin with a univariate version of the algorithm, associating each parameter with a normal candidate density centered at the current parameter value with a variance chosen to provide a Metropolis acceptance ratio between 25 and 50%. Motivated by the fact that the likelihood by itself cannot inform about spatial heterogeneity and clustering parameters, but only their sum,[27] use an elementary transformation method[42] to update parameter values. Because their model features many parameters identified only by a vague prior, they report slow convergence.

A similar approach is found in Ref 28. They use a Gibbs sampler approach to MCMC. The choice of Gaussian distributions with conjugate priors made the derivation and implementation of the full conditional distributions straightforward, with one exception—the full conditional distributions for each of the Markov random fields spatial dependence parameters. To account for these, Wilke et al.[28] employ a Metropolis-Hastings step in the Gibbs sampler

 © 2010 John Wiley & Sons, Inc.

19390068, 2010, 3, Downloaded from https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.88 by University of Massachusetts Dartmouth, Wiley Online Library on [22/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

by sampling from the 'pseudo full conditional likelihood,' which makes use of the pseudo-likelihood approach to classical Markov random field estimation[43] for these parameters. They too report slow convergence. The study of Wikle and Hooten[37] contains a good overview of computational methods, and includes a sketch of an MCMC program.

EM VERSUS HIERARCHICAL BAYESIAN APPROACHES

The question of which approach—EM/generalized expectation-maximization (GEM) or Bayesian—is 'better' naturally arises. Generally speaking, the Bayesian approach is most useful when there is some understanding about the process dynamics, in particular, regarding $\boldsymbol{H}$. When it is reasonable to assume the parameters are spatially varying (as in Ref 32), the hierarchical Bayesian approach is more appropriate than GEM. Xu and Wilke[2] conclude that when the model complexity increases, or when one has significant prior knowledge about the dynamics, a fully Bayesian (MCMC) approach is superior to a GEM approach. However, if the model is relatively simple, and there is little prior knowledge, then the EM/GEM approach is more reasonable. They go on to say that the EM/GEM approach is of limited use if the parameter space is highly dimensional because algorithmic convergence is problematic.

SIMULATING SPATIO-TEMPORAL DATA

Literature on methods for simulating data from a spatio-temporal process is sparse. One common means begins with estimation of the covariance function. Once the covariance function is estimated, if the process is assumed Gaussian, then simulating a realization from a spatio-temporal model can be accomplished by simulating from a d-dimensional Gaussian process. If the process is stationary, the Fourier representation gives an efficient tool for this (see Refs 12, 44–46, and the references therein).

LARGE OR MULTIPLE DATASET CHALLENGES

In addition to the slow convergence encountered because of high-dimensionality, it is also common that the data set is large, contributing to long computing times. Another challenge that arises is that of combining more than one, very different, datasets to estimate a surface, i.e., changing through time; see, e.g., Ref 30. Hoar et al.[33] provide a solution to both challenges by migrating the Bayesian hierarchical model from a workstation-class implementation to a (massive) parallel architecture. They outline a number of ways in which running jobs on a (single-processor) workstation is fundamentally different from using a shared, high-performance computing environment. They illustrate those ideas by tracing through the production of a value-added surface wind dataset based on combining two very different wind data sets—one with irregularly spaced gaps in space and time, the other on a regularly spaced spatio-temporal grid, but that lacked sufficient detail. They successfully migrate software from a workstation-class implementation to a server-class implementation and outline the planning in terms of hardware consideration, data stream management, and software customization.

CONCLUSION

Computational methods play a huge role in the analysis of spatio-temporal data. Methods for estimation and inference typically involve some sort of iteration. Moreover, cross-validation is often used to assess if a model fits a data set well (see, e.g., Ref 38). Data sets are often large, requiring sophisticated, efficient algorithms. Often computer hardware—memory, storage, etc.—also play a role in the many facets that must be considered. The goal in writing this article is to provide an overview of the most current, common, and innovative literature in computational applications to spatio-temporal processes. Other topics of interest that provide more insight or additional, supplementary methods may be found in references 47–59, that are elaborated in Further Reading.

REFERENCES

1. Stein A, Groenigen JV, Jeger M, Hoosbeek M. Space-time statistics for environmental and agricultural related phenomena. *Environ Ecol Stat* 1998, 5:155–172 (Special Issue on Space-time Methods).
2. Xu K, Wilke CK. Estimation of parameterized spatio-temporal dynamic models. *J Stat Plan Inference* 2007, 137:567–588.
3. Carroll RJ, Chen R, George EI, Li T, Newton HJ, Schmiediche H, Wang N. Ozone exposure and

 © 2010 John Wiley & Sons, Inc.

19390068, 2010, 3, Downloaded from https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.88 by University of Massachusetts Dartmouth, Wiley Online Library on [22/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

population density in Harris County, Texas. *J Am Stat Assoc* 1997, 92:392–404.

4. Mitchell MW, Gumpertz ML. Spatio-temporal prediction inside a free-air co_2 enrichment system. *J Agric Biol Environ Stat* 2003, 8:310–327.
5. Stein ML. Space-time covariance functions. *J Am Stat Assoc* 2005, 100:310–321.
6. Schabenberger O, Gotway CA. *Statistical Methods for Spatial Data Analysis Texts in Statistical Science*, Chapter 9. Boca Raton, FL: Chapman & Hall/CRC; 2005.
7. Cressie NA, Huang H. Classes of nonseparable, spatio-temporal stationary covariance functions. *J Am Stat Assoc* 1999, 94:1330–1340.
8. Gneiting T. Nonseparable, stationary covariance functions for space-time data. *J Am Stat Assoc* 2002, 97:590–600.
9. Iaco SD, Myers D, Posa D. Space-time analysis using a general product-sum model. *Stat Probab Lett* 2001, 52:21–28.
10. Iaco SD, Myers D, Posa D. Nonseparable space-time covariance models: some parametric families. *Math Geol* 2002, 34:23–42.
11. Iaco SD, Myers D, Posa D. The linear coregionalization model and the product-sum space-time variogram. *Math Geol* 2003, 35:25–38.
12. Fuentes M, Chen L, Davis JM. A class of nonseparable and nonstationary spatial temporal covariance functions. *Environmetrics* 2008, 19:487–507.
13. Kyriakidis PC, Journel AG. Geostatistical space-time models: a review. *Math Geol* 1999, 31:651–684.
14. Porcu E, Mateu J, Saura F. New classes of covariance and spectral density functions for spatio-temporal modeling. *Stoch Environ Res Risk Assess* 2008, 22(Suppl 1):S65–S79.
15. Ma C. Families of spatio-temporal stationary covariance models. *J Stat Plan Inference* 2003, 116:489–501.
16. Huang HC, Cressie NA. Spatio-temporal prediction of snow water equivalent using the Kalman filter. *Comput Stat Data Anal* 1996, 22:159–175.
17. Cane MA, Kaplan A, Miller RN, Tang B, Hackert EC, Busalacchi AJ. Mapping tropical pacific sea level: data assimilation via a reduced state space Kalman filter. *J Geophys Res* 1996, 101:22599–22617.
18. Mardia KV, Goodall C, Redfern EJ, Alonso FJ. The kriged Kalman filter. *Test* 1998, 7:217–285.
19. Wikle CK, Cressie NA. A dimension-reduced approach to space-time Kalman filtering. *Biometrika* 1999, 86:815–829.
20. Kalman RE. A new approach to linear filtering and prediction problems. *Trans ASME* 1960, 82:35–45.
21. Shumway RH, Stoffer DS. An approach to time series smoothing and forecasting using the em algorithm. *J Time Ser Anal* 1982, 3:253–264.
22. West M, Harrison J. *Bayesian Forecasting and Dynamic Models*. 2nd ed. Springer Series in Statistics. New York: Springer; 1997.
23. Gupta NK, Mehra RK. Computational aspects of maximum likelihood estimation and redicution in sensitivity function calculations. *IEEE Trans Automat Contr* 1974, 19:774–783.
24. Shumway RH, Stoffer DS. *Time Series Analysis and Its Applications with R Examples*. 2nd ed. Springer Series in Statistics. New York: Springer; 2006.
25. McLachlan GJ, Krishnan T. *The EM Algorithm and Extensions*. 2nd ed. Wiley Series in Probability and Statistics. Hoboken, NJ: John Wiley & Sons; 2008.
26. Handcock MS, Wallis JR. An approach to statistical spatial-temporal modeling of meterological fields. *J Am Stat Assoc* 1994, 89:368–378.
27. Waller LA, Carlin BP, Xia H, Gelfand AE. Hierarchical spatio-temporal mapping of disease rate. *J Am Stat Assoc* 1997, 92:607–617.
28. Wikle CK, Berliner LM, Cressie N. Hierarchical Bayesian space-time models. *Environ Ecol Stat* 1998, 5:117–154 (Special issue on space-time methods).
29. Sansó B, Guenni L. Venezuelan rainfall data analysed by using a Bayesian space-time model. *J R Stat Soc Ser., C: Appl Stat* 1999, 48:345–362.
30. Wikle CK, Milliff RF, Nychka D, Berliner LM. Spatiotemporal hierarchical Bayesian modeling: tropical ocean surface winds. *J Am Stat Assoc* 2001, 96:382–397.
31. Kibria BG, Sun L, Zidek JV, Le ND. Bayesian spatial prediction of random space-time fields with applications to mapping pm-2.5 exposure. *J Am Stat Assoc* 2002, 97:112–124.
32. Wikle CK. Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology* 2003, 84:1382–1394 (Special feature article).
33. Hoar TJ, Milliff RF, Nychka D, Wikle CK, Berliner LM. Winds from a Bayesian hierarchical model: computation for atmospheric-ocean research. *J Comput Graph Stat* 2003, 12:781–807 (Special issue on statistical analysis of massive data streams).
34. Fahrmeir L, Kneib T, Lang S. Penalized structured additive regression for space-time data: a Bayesian perspective. *Stat Sin* 2004, 14:731–761 (Special issue on emerging issues in longitudinal and functional data analysis).
35. Catelan D, Biggeri A, Dreassi E, Lagazio C. Space-cohort Bayesian models in ecological studies. *Stat Modelling* 2006, 6:159–173.
36. Hooten MB, Wikle CK. A hierarchical Bayesian non-linear spatio-temporal model for the spread of invasive species with application to the eurasian collared-dove. *Environ Ecol Stat* 2008, 15:59–70 (Special issue on the new development of statistical analysis in wildlife, fisheries, and ecological research).

 © 2010 John Wiley & Sons, Inc.

19390068, 2010, 3, Downloaded from https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.88 by University of Massachusetts Dartmouth, Wiley Online Library on [22/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

37. Wikle CK, Hooten MB, eds. Hierarchical Bayesian spatio-temporal models for population spread. In: Clark JS, Gelfand AE, eds. *Hierarchical Modelling for the Environmental Sciences, Statistical Methods and Applications*. New York: Oxford University Press; 2006, 145–169.

38. Zheng Y, Zhu J. Markov chain Monte Carlo for a spatial-temporal autologistic regression model. *J Comput Graph Stat* 2008, 17:123–137.

39. Gelfand AE, Smith AF. Sampling-based approaches to calculating marginal densities. *J Am Stat Assoc* 1990, 85:398–409.

40. Chib S, Greenberg E. Understanding the Metropolis-Hastings algorithm. *Am Stat* 1995, 49:327–335.

41. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equations of state calculations by fast computing machines. *J Chem Phys* 1953, 21:1087–1091.

42. Gelfand AE, Sahu SK, Carlin BP. Efficient parameterization for normal linear mixed models. *Biometrika* 1995, 82:479–488.

43. Cressie NA. *Statistics for Spatial Data*, Revised Edition, Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons; 1993.

44. Fuentes M. Periodogram and other spectral methods for nonstationary spatial processes. *Biometrika* 2002, 89:197–210.

45. Fuentes M. A formal test for nonstationary of spatial stochastic processes. *J Multivar Anal* 2005a, 96:30–55.

46. Fuentes M. Testing for separability of spatial temporal covariance functions. *J Stat Plan Inference* 2005b, 136:447–466.

47. Jones R, Zhang Y. Models for continuous stationary space-time processes. In: Gregoire T, Brillinger D, Diggle P, Russek-Cohen E, Warren W, Wolfinger R, eds. *Modelling Longitudinal and Spatially Correlated Data*, vol. 122, Lecture notes in Statistics. New York: Springer-Verlag; 1997, 289–298.

48. Christakos G. *Random Field Models in Earth Sciences*. San Diego, CA: Academic Press; 1992.

49. Christakos G. *Modern Spatiotemporal Geostatistics*, vol. 6, International Association for Mathematical Geosciences: Studies in Mathematical Geosciences. New York: Oxford University Press; 2000.

50. Brown PE, Karesen KF, Roberts GO, Tonellato S. Blur-generated non-separable space-time models. *J Roy Stat Soc., Ser B: Stat Methodol* 2002, 62:847–860.

51. Hartfield MI, Gunst RF. Identification of model components for a class of continuous spatiotemporal models. *J Agric Biol Environ Stat* 2003, 8:105–121.

52. Kolovos A, Christakos G, Hristopulos D, Serre M. Methods for generating non-separable spatiotemporal covariance models with potential environmental applications. *Adv Water Resour* 2004, 27:815–830.

53. Ma C. Recent developments on the construction of spatio-temporal covariance models. *Stoch Environ Res Risk Assess* 2008, 22(Suppl 1):S39–S47.

54. Cressie NA, Wikle CK. Space-time Kalman filter. In: El-Shaarawi AH, Piegorsch WW, eds. *Encyclopedia of Environmentrics*, vol. 4. Chichester: John Wiley & Sons; 2002, 2045–2049.

55. Verhein F, Chawla S. Mining spatio-temporal patterns in object mobility databases. *Data Min Knowl Discov* 2008, 16:5–38.

56. Stoffer DS, Wall KD. Bootstrapping state-space models: Gaussian maximum likelihood estimation and the Kalman filter. *J Am Stat Assoc* 1991, 86:1024–1033.

57. Wall KD, Stoffer DS. A state-space approach to bootstrapping conditional forecasts in arma models. *J Time Ser Anal* 2002, 23:733–752.

58. Gelfand AE, Zhu L, Carlin BP. On the change of support problem for spatio-temporal data. *Biostatistics* 2001, 2:31–45.

59. Xu K, Wikle CK, Fox NI. A kernel-based spatio-temporal dynamic model for nowcasting weather radar reflectivities. *J Am Stat Assoc* 2005, 100:1133–1144.

FURTHER READING

Literature on modeling spatio-temporal processes is growing rapidly. In addition to the sources cited in this manuscript, an interested reader might consider the following additional resources. In the development of nonseparable spatio-temporal covariance functions, the reader is referred to Refs 15,47–53, and the references therein. For an overview of space–time Kalman filters, the reader is referred to Refs 19,54. An interesting paper that was difficult to fit into the text of this manuscript is by Ref 55, in which the authors present the STAR-Miner algorithm for finding patterns in how objects move between regions in space over time. Background information for Kalman filtering and cross-validation techniques for space–time processes that are separable (in time and space) are found in Ref 16. Stoffer and Wall[56] describe a simple bootstrap sampling algorithm for parameter estimates in general state-space models, i.e., appropriate for the spatio-temporal setting. Additionally, Wall and Stoffer[57] describe how bootstrap resampling can give estimates of conditional forecast accuracy. Zheng and Zhu[38] present an MCMC approach for a spatio-temporal autologistic

 © 2010 John Wiley & Sons, Inc.

19390068, 2010, 3, Downloaded from https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.88 by University of Massachusetts Dartmouth, Wiley Online Library on [22/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

regression model, proposing a fully Bayesian approach for both model parameter inference and prediction at future time points. They combine a Metropolis-Hastings algorithm with a Gibbs sampler for obtaining the posterior distribution of the model parameters as well as the posterior predictive distributions. Gelfand et al.[58] address the change of support problem for spatio-temporal data, including fully Bayesian kriging. They illustrate the 'judicious specification of the spatio-temporal association' enables manageable computation. Spectral methods are also used for studying spatio-temporal processes in Refs 12,44–46,59 and the references therein.

© 2010 John Wiley & Sons, Inc.

19390068, 2010, 3, Downloaded from https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.88 by University of Massachusetts Dartmouth, Wiley Online Library on [22/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License