A Kernel-Based Spectral Approach for Spatiotemporal Dynamic Models *

Christopher K. Wikle
University of Missouri. Department of Statistics. Columbia, Missouri 65211 USA

Abstract

Spatiotemporal processes can often be written as hierarchical state-space processes. In situations with complicated dynamics such as wave propagation, it is difficult to parameterize state transition functions for high-dimensional state processes. Although in some cases, prior understanding of the physical process can be used to formulate prior models for the state transition, this is not always possible. Alternatively, for processes where one considers discrete time and continuous space, integro-difference equations suggest that complicated dynamics can be modeled by allowing the associated redistribution kernel to vary with space and/or time. We show that this is indeed the case and that by considering a spectral implementation of such models, one can formulate a spatiotemporal model with relatively few parameters that can accommodate complicated dynamics.

Keywords: kernel, integro-difference equations, propagation.

1 Introduction

There has been much interest in recent years in modeling spatiotemporal processes in the environmental and physical sciences. Methods have considered geostatistical approaches (see the review [1]), multivariate time series approaches (e.g., [2]), space-time dynamic models (e.g., [3], [4],[5], [6], [7]), and hierarchical approaches (e.g., [8], [9], [10]).

In particular, we are interested in processes that have coherent dynamical interactions such as exhibited by geophysical or ecological processes with wave behavior. One approach to accounting for realistic dynamic structure in such complicated spatiotemporal settings is to consider underlying explicit theoretical relationships. For example, in an atmospheric/oceanic context one has well-specified deterministic relationships (partial differential equations, PDEs) that describe, to an extent,

*This research was made possible by a grant from the U.S. Environmental Protection Agency's Science to Achieve Results (STAR) program, Assistance Agreement No. R827257-01-0.

the process of interest. Such information can be used in traditional state-space settings as demonstrated in the "data assimilation" literature in atmospheric/ocean science (e.g., [11]). Alternatively, one can consider the physics as "approximate" and actually use the PDEs as a framework for developing prior distributions in a hierarchical model (e.g., [12],[10]). This approach works very well when one has an understanding of the underlying physical relationships from which to develop the PDEs and thus, the priors. However, it is often the case that the scientific knowledge for a specific problem is not well-developed and one does not have easily-described physically-based priors. In that case, one may still need efficient methodologies that can model dynamical behavior with relatively few parameters.

In this paper, an approach for modeling complicated dynamical spatiotemporal processes is described. This approach is based on a stochastic integro-difference equation, where the redistribution kernel is allowed to vary with space and/or time. This methodology makes use of recent developments in nonstationary spatial modeling using convolution kernels (e.g., [13]), as well as theoretical results concerning deterministic integro-difference equations and their ability to model dynamical processes. We show that by considering such models from a spectral perspective (e.g., [7], [14]), one naturally obtains dimension reduction in many circumstances.

2 Background

As described by [15], integro-difference equations have recently seen increasing utility in population ecology. They are discrete-time, continuous-space models that take the form

$$y_{t+1}(s) = \int_{-\infty}^{\infty} k_s(r) g(y_t(r)) dr \tag{1}$$

where $y_t(s)$ is a spatiotemporal process at spatial location s and time t, $g(\cdot)$ is some (possibly nonlinear) function of the y-process, and $k_s(r)$ is a "redistribution kernel" that describes how the process at the previous time at location r is related to the process at location s at the next time. That is, the function $g(y_t(r))$ describes the growth of the process at some location r between time t and $t+1$, and the integro-difference equation describes how the resulting growth is redistributed in space. For example, in ecological settings, one might consider a nonlinear growth function given by the Beverton-Holt stock-recruitment curve (e.g., [15]). Often, $g(\cdot)$ is assumed to be linear and the redistribution kernel is assumed not to change with spatial location, but only depends on the difference in location. In that case, we rewrite (1) as the following convolution integral

$$y_{t+1}(s) = g'(0) \int_{-\infty}^{\infty} k(s-r) y_t(r) dr, \tag{2}$$

where $g'(\ )$ is the first-derivative of $g(\ )$. Integro-Difference equations of this type are relatively simple, yet can model relatively complicated dynamics. For example, Fisher's "reaction-diffusion" partial differential equation [16], which might describe the advance of an invading species (among other things) as a wave with known speed, can also be described by an integro-difference equation with a Gaussian kernel, where the speed of propagation depends on the kernel variance. Different choices of kernels lead to different dynamical behavior such as different speeds of the dispersive wave (e.g., [15], [17]).

These results suggest that more complicated or "extra-diffusive" dynamical behavior can be modeled by altering the kernel. Specifically, we are interested in propagation of features across space through time. For example, consider the one-dimensional Gaussian spatial kernel

$$k_s(r, \theta_1, \theta_2) = \frac{1}{\theta_2\sqrt{2\pi}} \exp\{-.5(r - \theta_1 - s)^2/\theta_2\} \tag{3}$$

where the kernel is centered at $\theta_1 - s$ and thus is shifted by θ_1 spatial units relative to location s, and θ_2 is the scale parameter. We refer to θ_1 as the translation parameter and θ_2 as the dilation parameter, analogous to the usual translation and dilation in the description of wavelet basis functions. Presumably, in the integro-difference kernel context, these parameters influence the dynamical evolution of the y process. We investigate this influence with some simple simulations, as discussed below.

Figure 1 shows a successive integration of (2) with a 2-D Gaussian kernel analogous to (3), where the kernel is centered at location s (i.e., translation parameters are zero for all spatial locations). Note that over time, this suggests that the process propagates in a diffusive fashion, but the center of the disturbance does not propagate; we refer to this as diffusive propagation.

Figure 2 shows a similar simulation but with the kernel translated to the left and down relative to s, identically for all s. In this case the disturbance shows diffusive propagation but the center of the disturbance also propagates to the right and up, with speed proportional to the translation distance; we refer to this as extra-diffusive propagation. Notice that in this case the spread (dilation) of the kernel is also larger than in the previous figure and thus the process diffuses at a greater rate. Now, consider a spatially-varying kernel, in which the translation parameters are allowed to vary with space. Such models have recently been considered by [13] for spatial problems in which a convolution of white noise is used to generate nonstationary spatial covariance models. In the integro-difference equation setting, such spatially-varying (heterogeneous) kernels can capture more complicated dynamics than homogeneous kernels. For example, Figure 3 shows arrows that indicate the propagation direction at each location implied by slowly spatially-varying translation parameters. That is,

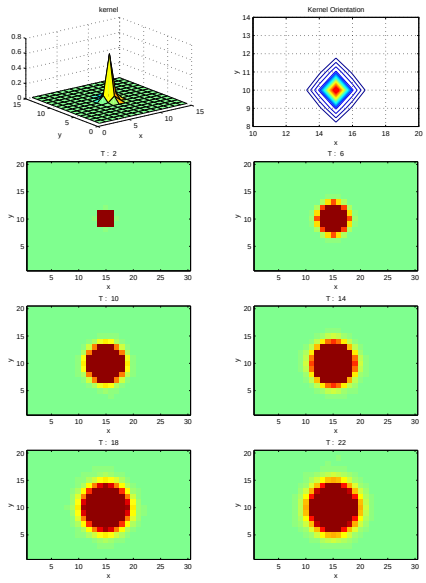

Figure 1: Simulation using kernel in 2-d space with no translation.

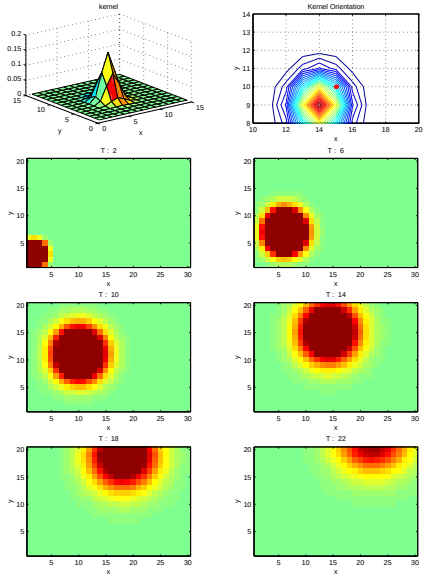

Figure 2: Simulation using kernel in 2-d space with kernel translated to the left and down, implying propagation up and to the right.

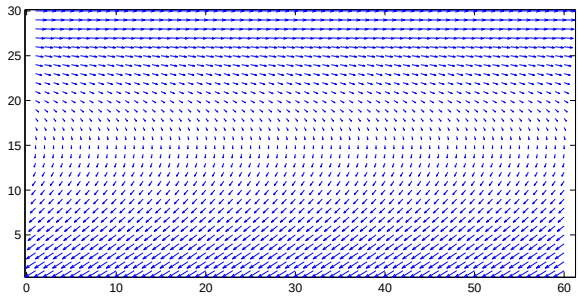

Figure 3: Vectors indicating spatial variation in translation parameter in 2-d spatial kernel. Note that the arrow points in the direction in which the propagation is suggested by the translation (i.e., actual translation vector is in the opposite direction as shown).

the kernel is translated in the opposite direction (relative to s) shown by the arrows. Figure 4 shows the resulting propagation of a disturbance; note how the disturbance propagates in a circular fashion.

Clearly, different kernels can give rise to very complicated dynamics from a deterministic point of view. As is the case for PDE models, we seldom expect real-world processes to behave as purely deterministic processes. Thus, to model realistic processes, we must allow for random kernels and additive noise-forcing. The next section describes such a model.

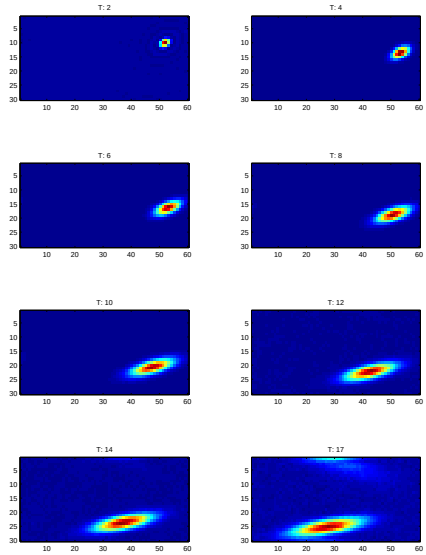

Figure 4: Propagation suggested by the spatially-varying translation parameters shown in Figure 3.

3 Methodology

Consider the stochastic integro-difference equation

$$y_{t+1}(s) = \gamma \int k_s(r; \boldsymbol{\theta}_s) y_t(r) dr + \tilde{\eta}_{t+1}(s) \tag{4}$$

where $\tilde{\eta}$ is a spatially-colored noise process and the parameter γ is analogous to the derivative term in (2) but is used in this context to control (and allow for) explosive growth. A similar model was considered by [18] and [7] as a "spatially descriptive, temporally dynamic" model, and by [19], [20] as "blur-generated" space-time models. Although these authors recognized that this modeling framework is powerful in that it can easily accommodate non-separable space-time covariance structures, the flexibility of such models to accommodate complicated extra-diffusive dynamics was not considered in the statistical context.

3.1 Spectral Formulation

As is often the case (e.g., [14]), the model (4) is easier to implement in the heterogeneous kernel case if we work in the spectral domain. Thus, as in [7] we expand

the kernel and the process in terms of some spectral basis functions $\phi_i(s)$:

$$k_s(r;\boldsymbol{\theta}_s) = \sum_i b_i(s;\boldsymbol{\theta}_s)\phi_i(r) \tag{5}$$

$$y_t(s) = \sum_j \alpha_j(t)\phi_j(s) \tag{6}$$

where basis functions are complete and orthonormal

$$\int_D \phi_k(s)\phi_l(s)ds = \begin{cases} 1 & \text{for } k = l, \\ 0 & \text{otherwise.} \end{cases}$$

If the process of interest has non-trivial dynamics, then the kernel has significant spread and thus can be represented as the linear combination of a relatively small set of spectral basis functions. In that case, the sum in (5) is truncated at I and upon substitution of (5) and (6) into (4), we get

$$y_{t+1}(s) = \gamma \mathbf{b}'(s;\boldsymbol{\theta}_s)\boldsymbol{\alpha}_t^{(1)} + \tilde{\eta}_{t+1}(s) \tag{7}$$

where $\mathbf{b}(s;\theta_s) \equiv [b_1(s;\theta_s)\dots b_I(s;\theta_s)]'$ and $\boldsymbol{\alpha}_t^{(1)} \equiv [\alpha_1(t)\dots\alpha_I(t)]'$. So, for spatial locations $\{s_1,\dots,s_n\}$ (which need not be data locations if there is a "measurement" model that conditions the data on y),

$$\mathbf{y}_{t+1} = \gamma \mathbf{B}'_{\boldsymbol{\theta}}\boldsymbol{\alpha}_t^{(1)} + \tilde{\boldsymbol{\eta}}_{t+1}$$

where $\mathbf{y}_{t+1} \equiv [y_{t+1}(s_1)\dots y_{t+1}(s_n)]'$ and $\mathbf{B}_{\boldsymbol{\theta}} = [\mathbf{b}(s_1;\theta_{s_1})\dots\mathbf{b}(s_n;\theta_{s_n})]$. Thus,

$$\boldsymbol{\alpha}_{t+1}^{(1)} = \boldsymbol{\Phi}'_{(1)}\mathbf{B}'_{\boldsymbol{\theta}}\boldsymbol{\alpha}_t^{(1)} + \boldsymbol{\eta}_{t+1}^{(1)} \tag{8}$$

$$\boldsymbol{\alpha}_{t+1}^{(2)} = \boldsymbol{\Phi}'_{(2)}\mathbf{B}'_{\boldsymbol{\theta}}\boldsymbol{\alpha}_t^{(1)} + \boldsymbol{\eta}_{t+1}^{(2)}, \tag{9}$$

where $\boldsymbol{\alpha}_t^{(2)} \equiv [\alpha_{I+1}(t)\ \dots\ \alpha_n(t)]'$, $\boldsymbol{\Phi}_{(1)} \equiv [\boldsymbol{\phi}_1\ \dots\ \boldsymbol{\phi}_I]$, $\boldsymbol{\Phi}_{(2)} \equiv [\boldsymbol{\phi}_{I+1}\ \dots\ \boldsymbol{\phi}_n]$, and if we assume that $\tilde{\boldsymbol{\eta}}_t \sim N(\mathbf{0},\mathbf{C}_{\tilde{\eta}})$, then $\boldsymbol{\eta}_t^{(1)} \sim N(\mathbf{0},\mathbf{C}_\eta^{(1)})$, $\boldsymbol{\eta}_t^{(2)} \sim N(\mathbf{0},\mathbf{C}_\eta^{(2)})$, where $\mathbf{C}_\eta^{(j)} \equiv \boldsymbol{\Phi}'_{(j)}\mathbf{C}_{\tilde{\eta}}\boldsymbol{\Phi}_{(j)}$ for $j = 1,2$.

Note that the evolution of $\boldsymbol{\alpha}^{(2)}$ in (9) depends on the past value of the $\boldsymbol{\alpha}^{(1)}$ process rather than the $\boldsymbol{\alpha}^{(2)}$ process. Depending on the dimension reduction (i.e., the dilation of the kernel) and the underlying process, one might assume that $\boldsymbol{\alpha}_t^{(2)}$ is a non-dynamic spatiotemporal component without much loss of predictive power (e.g., [7]).

3.2 Modeling $\mathbf{B}_{\boldsymbol{\theta}}$

The methodology of Section 3.1 does not depend explicitly on the functional form of the kernel, with the exception of the assumption that the kernel can be modeled reasonably well as a finite sum of orthogonal basis functions. Indeed, since the final model formulation is in spectral space, a nonparametric estimation procedure need only find estimates for the parameters in $\mathbf{B}$ (where we drop the θ subscript in this case). One may consider this problem using moment-based estimators in a Kalman filter framework as in [7]. Likelihood estimates obtained from E-M algorithm approaches can be used as well provided the number of parameters is reasonably small (e.g., [21]).

In cases where $\mathbf{B}$ is high-dimensional, we can model $\mathbf{b}_i \equiv \{b_i(s_1), \ldots, b_i(s_n)\}'$ in terms of another spectral decomposition. That is, $\mathbf{b}_i = \boldsymbol{\Psi}\boldsymbol{\beta}_i$, where $\boldsymbol{\Psi}$ is an $n \times J$ matrix of basis functions (which may be the same as we used before, $\psi_i(s) = \phi_i(s)$, but need not be), and $\boldsymbol{\beta}_i \equiv [\beta_i(1), \ldots, \beta_i(J)]'$, where $J << n$. In this case, we can write (8) as

$$\boldsymbol{\alpha}_{t+1}^{(1)} = \boldsymbol{\Phi}_{(1)}'\boldsymbol{\Psi}\boldsymbol{\beta}\boldsymbol{\alpha}_t + \boldsymbol{\eta}_{t+1}^{(1)} \tag{10}$$

where $\boldsymbol{\beta} \equiv [\boldsymbol{\beta}_1 \ldots \boldsymbol{\beta}_I]$ is an $J \times I$ matrix of parameters. Thus, in this approach the dimensionality of the parameters has been reduced from $n \times I$ to $J \times I$. If the dimensionality has been sufficiently reduced, it may be possible to use likelihood-based estimation approaches efficiently.

3.2.1 Parametric Formulation

A parsimonious representation of $\mathbf{B}$ can be obtained if one considers a specific parametric kernel function. This approach has the additional advantage that we can control the dynamics by altering the translation and dilation of the kernel in terms of a relatively few number of parameters. Although easily implemented in a two-dimensional spatial setting, for illustration we focus on the 1-d spatial case here.

Consider the Gaussian kernel (3). Let $\phi_i(s)$ be Fourier basis functions. Of course, the Fourier transform of the Gaussian kernel is just its characteristic function

$$\begin{aligned} b_j(s;\theta_1(s),\theta_2(s)) &= \exp\{i\omega_j(\theta_1(s)+s) - .5\omega_j^2\theta_2(s)\} && (11)\\ &= \cos\{\omega_j(\theta_1(s)+s)\}\exp\{-.5\omega_j^2\theta_2(s)\} && (12)\\ &\quad + i\sin\{\omega_j(\theta_1(s)+s)\}\exp\{-.5\omega_j^2\theta_2(s)\} \end{aligned}$$

where ω_j is the spatial frequency. Thus, the real and imaginary coefficients of the characteristic function correspond to cosine and sine Fourier basis functions, respectively. These coefficients are completely determined if we know the kernel

parameters $\theta_1(s)$ and $\theta_2(s)$ at each spatial location s. That is we need to specify the spatial fields $\boldsymbol{\theta}_1 = [\theta_1(s_1) \dots \theta_1(s_n)]'$ and $\boldsymbol{\theta}_2 = [\theta_2(s_1) \dots \theta_2(s_n)]'$. For example, we might assume $\boldsymbol{\theta}_1 \sim N(\mu_1, \mathbf{C}_{\theta_1})$ and $\boldsymbol{\theta}_2 \sim LN(\mu_2, \mathbf{C}_{\theta_2})$. Since we expect the dynamics to be relatively slowly varying over space, the spatial structure in these fields might be relatively simple.

3.3 Temporally-Varying Kernels

To model processes with complicated dynamics, time-varying kernels might be considered. Thus, in addition to letting θ_1 and θ_2 vary spatially, we allow them to vary temporally as well; that is, $\boldsymbol{\theta}_1(t)$ and $\boldsymbol{\theta}_2(t)$. In that case, the model (8) becomes $\boldsymbol{\alpha}_{t+1}^{(1)} = \boldsymbol{\Phi}'_{(1)}\mathbf{B}'_{\theta_t}\boldsymbol{\alpha}_t^{(1)} + \boldsymbol{\eta}_{t+1}$. The θ processes might be modeled using one of the recently developed classes of space-time covariance functions ([22], [23]), if the dimensionality is not an issue. Alternatively, one might consider a spatiotemporal dynamical model for these parameters such as:

$$\boldsymbol{\theta}_i(t+1) = \mathbf{H}\boldsymbol{\theta}_i(t) + \boldsymbol{\zeta}_i(t+1), \tag{13}$$

where the subscript i refers to the i-th parameter of the redistribution kernel. One must be careful in this setting not to simply replace one complicated spatiotemporal problem by another of equal (or greater) complexity! Typically, one would expect that the spatiotemporal dynamics of the parameters are substantially less complicated than the dynamics of the original process. Dimensionality might be reduced by another spectral decomposition of these parameter processes. That is, one might project the parameter vector on some orthogonal basis set $\boldsymbol{\Upsilon}$ $(n \times p)$, such that $\boldsymbol{\xi}_i(t) = \boldsymbol{\Upsilon}'\boldsymbol{\theta}_i(t)$, then model $\boldsymbol{\xi}_i(t)$. Ideally, one could select p such that most of the variability in $\boldsymbol{\theta}_i(t)$ is captured by a few spectral coefficients (i.e., $p << n$). This is a realistic possibility as we typically expect the spatial variation in $\boldsymbol{\theta}_i(t)$ to be relatively large scale (i.e., we don't expect the dynamical behavior to change dramatically over small spatial scales).

4 Example: Modeling Cloud Intensity over Time

The propagation of clouds is a complicated nonlinear function of various atmospheric state processes. In fact, cloud dynamics are still the subject of intense research in the atmospheric science community. Cloud parameterizations are a fundamental component of atmospheric General Circulation Models (GCM's) used to study climate and weather. Our interest here is whether the dynamics of such cloud processes can be modeled adequately by the kernel-based integro-difference equation model

outlined above. To examine this, we obtained cloud intensity information from a regional climate model as discussed in [24].

We consider the data $Z_t(s_i)$, where $i = 1, \ldots, n$ and $t = 1, \ldots, T$. Specifically, we selected $n = 60$ and $T = 80$. The spatial locations are evenly spaced at a resolution of 52 km and represent a 1-D spatial (longitudinal) domain over the central U.S. The temporal sampling is every 6 hours and is representative of large scale forcing in late March 1979. The data values are cloud water content in grams of water per kilogram of air (g/kg). Thus, the data have positive integer support at each spatial location. For the purposes of this study, we remove 40% of the data by randomly selecting from the 60 spatial locations, assuming the data are missing from all times at the selected pixels. The goal is to model the cloud intensity (cloud water content) and thus predict the cloud water content over the whole domain, including the locations for which data are missing.

4.1 Non-Gaussian Hierarchical Spatiotemporal Dynamic Model

Similar to the non-Gaussian spatial modeling approach of [25], we assume that conditional on the Poisson intensity process at all spatial and temporal locations of interest, the data are distributed as independent Poisson random variables:

$$Z_t(s_i)|\boldsymbol{\lambda}_t \sim Poi(\mathbf{k}'_{i,t}\boldsymbol{\lambda}_t) \tag{14}$$

where $\boldsymbol{\lambda}_t \equiv [\lambda_t(s_1), \ldots, \lambda_t(s_n)]'$ is the Poisson intensity process at all spatial locations for time t, and $\mathbf{k}_{i,t}$ is an incidence vector indicating whether a prediction location has an associated observation. We then let $\mathbf{u}_t \equiv \log(\boldsymbol{\lambda}_t)$ and assume

$$\mathbf{u}_t|\mu, \nu, \mathbf{y}_t, \sigma_\epsilon^2 \sim N(\mu\mathbf{1} + \nu\boldsymbol{\Phi}\boldsymbol{\alpha}_t, \sigma_\epsilon^2\mathbf{I}), \tag{15}$$

where $\mathbf{y}_t = \boldsymbol{\Phi}\boldsymbol{\alpha}_t$, μ is the overall mean effect, ν is a scaling parameter, and σ_ϵ^2 represents extra-Poisson variability. We then make use of the model (8) and (9),

$$\begin{aligned}
\boldsymbol{\alpha}_t^{(1)}|\boldsymbol{\alpha}_{t-1}^{(1)}, \mathbf{B}_\theta, \mathbf{C}_\eta^{(1)} &\sim N(\boldsymbol{\Phi}'_{(1)}\mathbf{B}'_\theta\boldsymbol{\alpha}_{t-1}^{(1)}, \mathbf{C}_\eta^{(1)}) \\
\boldsymbol{\alpha}_t^{(2)}|\boldsymbol{\alpha}_{t-1}^{(1)}, \mathbf{B}_\theta, \mathbf{C}_\eta^{(2)} &\sim N(\boldsymbol{\Phi}'_{(2)}\mathbf{B}'_\theta\boldsymbol{\alpha}_{t-1}^{(1)}, \mathbf{C}_\eta^{(2)}).
\end{aligned}$$

Note that we have assumed $\gamma = 1$ in this example. One of our primary interests with these data is to determine the overall tendency of the spatially-varying translation parameter. Thus, we assume $\boldsymbol{\theta}_1$ is relatively smooth and let $\boldsymbol{\theta}_1 = \boldsymbol{\Upsilon}\mathbf{f}$, with $\mathbf{f} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_f)$, where $\boldsymbol{\Upsilon}$ are the first p ($p = 5$) eigenvectors of an exponential correlation matrix with relatively strong spatial dependence ($c_\theta(h) = \exp(-h/30)$ for distances $h = 0, \ldots, 60$) and $\boldsymbol{\Sigma}_f$ is the associated diagonal matrix. For this

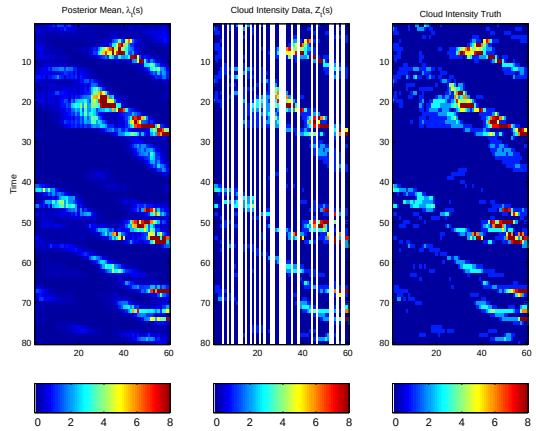

Figure 5: Left Panel: Posterior mean of $\boldsymbol{\lambda}_t$; Center Panel: Cloud intensity "Data"; Right Panel: Cloud intensity "truth".

study, we assume that $\boldsymbol{\theta}_2$ is fixed such that $\theta_2(s_i) = 4$ for all $i = 1, \ldots, n$. We let $\mathbf{C}_\eta^{(j)} = \boldsymbol{\Phi}'_{(j)}\mathbf{C}_\eta\boldsymbol{\Phi}_{(j)}$ where $\mathbf{C}_\eta$ is a Gaussian covariance matrix with fixed parameters ($c_\eta(h) = \exp(-h^2/200)$ for $h = 0, \ldots, 60$). We also specify uniform priors for $\mu \sim Unif[-10, 10]$ and $\nu \sim Unif[.5, 5]$ and let $\sigma_\epsilon^2 \sim IG(q_\epsilon, r_\epsilon)$, where $q_\epsilon = 3$ and $r_\epsilon = 5$. Finally, we must specify a prior distribution for the initial state for the $\boldsymbol{\alpha}^{(1)}$ process; we let $\boldsymbol{\alpha}_0^{(1)} \sim N(\mathbf{0}, .05\mathbf{I})$.

Implementation is via a Gibbs sampler, with straightforward conjugate updating with the exception of $\mathbf{u}_t$ and $\mathbf{f}$, which are updated via Metropolis-Hastings steps. The Gibbs sampler was run for 10000 iterations after burn-in. Standard errors were computed by batching, to account for the correlation in the Markov chain.

4.2 Results

Figure 5 shows the posterior mean of the intensities $\boldsymbol{\lambda}_t$ as well as the "data" $\mathbf{Z}_t$ and the truth. These plots only consider 1-D space and are interpreted by considering the x-axis as 1-D space (e.g., "longitude") and the y-axis as time, such that time increases from top to bottom. Thus, if one sees a diagonal stripe slanted to the right, it suggests propagation to the right (or East since the x-axis represents longitude); similarly, a left-slanted stripe suggests propagation to the left (West).

The model is able to fill in the missing information in a dynamically reasonable fashion. There is some suggestion in the visual inspection of the spatiotemporal cloud process that the speed of propagation is slower on the Eastern end of the domain than

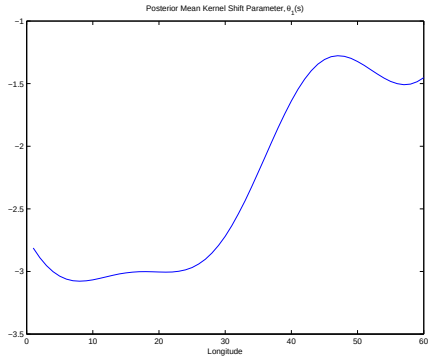

Figure 6: Posterior mean of $\boldsymbol{\theta}_1$.

in the Western. Figure 6 shows the posterior mean of the $\boldsymbol{\theta}_1$ translation parameter spatial process. This posterior mean confirms that the dynamical propagation is to the right over the entire domain, but is slower from about $x = 40$ to $x = 60$ than in the Western portion of the domain.

5 Discussion

We have presented a spatiotemporal hierarchical model that is based on a stochastic integro-difference equation with heterogeneous redistribution kernels. We showed that by letting kernels exhibit spatially-varying translation and dilation, complicated dynamics can be modeled. The methodology was demonstrated on cloud-intensity data from a regional climate model. Although the modeling approach works well for the cloud water content data, one can imagine that over longer time-spans the kernels should change with time as well. For example, the dynamics in the mid-latitudes are different in the spring than in the summer due to the annual migration of the jet stream and associated semipermanent high and low pressure systems. Thus, the kernel translation and dilation will change accordingly. As discussed, the modeling approach outlined here can easily accommodate time-varying kernels. In fact, one might allow the kernels to be influenced by other atmospheric variables, suggesting a possible approach to dynamic parameterization of clouds in climate models.

References

[1] P.C. Kyriakidis and A.G. Journel. Geostatistical space-time models: A review. *Mathematical Geology*, 31: 651–684, 1999.

[2] R. J. Bennett. *Spatial Time Series*. London, Pion Limited, 1999.

[3] C. Goodall and K.V. Mardia. Challenges in multivariate spatio-temporal modeling. *Proceedings of the XVIIth International Biometric Conference*, Hamilton, Ontario Canada, 8-12 August 1994.

[4] P. Guttorp, W. Meiring, and P.D. Sampson. A space-time analysis of ground-level ozone data. *Environmetrics*, 5: 241–254, 1994.

[5] K.V. Mardia, C.R. Goodall, E.J. Redfern, and F.J. Alonso, The Kriged Kalman filter (with discussion). *Test*, 7:217–285, 1998.

[6] W. Meiring, P. Guttorp, and P.D. Sampson. Space-time estimation of grid-cell hourly ozone levels for assessment of a deterministic model. *Environmental and Ecological Statistics*, 5:197–222, 1998.

[7] C.K. Wikle and N. Cressie. A dimension reduced approach to space-time Kalman filtering. *Biometrika* ,86:815–829, 1999.

[8] C.K. Wikle, L.M. Berliner, and N. Cressie. Hierarchical Bayesian space-time models. *Environmental and Ecological Statistics*, 5:117–154.

[9] L.M. Berliner, C.K. Wikle, and N. Cressie. Long-lead prediction of Pacific SSTs via Bayesian dynamic modeling. *Journal of Climate*, 13:3953–3968.

[10] C.K. Wikle, R.F. Milliff, D. Nychka, and L.M. Berliner. Spatio-temporal hierarchical Bayesian Modeling: Tropical ocean surface wind data. *Journal of the American Statistical Association*, 96:382–397.

[11] M. Ghil, S.E. Cohn, J. Tavantzis, K. Bube, and E. Isaacson. Applications of estimation theory to numerical weather prediction. *Dynamic meteorology: Data assimilation methods*, L. Bengtsson, M. Ghil, and E. Kallen, eds. New York: Springer-Verlag, 139–224, 1981.

[12] J.A. Royle, L.M. Berliner, C.K. Wikle, and R. Milliff. A hierarchical spatial model for constructing wind fields from scatterometer data in the Labrador sea. *Case Studies in Bayesian Statistics IV* , Springer-Verlag, 367–382, 1998.

[13] D. Higdon. A process-convolution approach to modelling temperatures in the North Atlantic Ocean. *Environmental and Ecological Statistics*, 5:173–190, 1998.

[14] M. Fuentes. Spectral methods for nonstationary spatial processes. *Biometrika*, to appear, 2001.

[15] M. Kot, M.A. Lewis, and P. van den Driessche. Dispersal data and the spread of invading organisms. *Ecology*, 77:2027–2042, 1996.

[16] R.A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7:353–369, 1937.

[17] M. Neubert, M. Kot, and M. Lewis. Dispersal and pattern formation in a discrete-time predator-prey model. *Theoretical Population Biology*, 48:7–43, 1995.

[18] C.K. Wikle. Spatio-temporal statistical models with applications to atmospheric processes. Ph.D. dissertation, Iowa State University, Ames, IA, 1996.

[19] P.E. Brown, K.F. Karesen, G.O. Roberts, and S. Tonellato. Blur-generated non-separable space-time models. *J.R. Statistical Society, Series B*, 62:847–860, 2000.

[20] P.E. Brown, P.J. Diggle, M.E. Lord, and P.C. Young. Space-time calibration of radar rainfall data. *Applied Statistics*, 50:221–241, 2001.

[21] R.H. Shumway and D.S. Stoffer. An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis*, 3:253-264, 1982.

[22] N. Cressie, and H.-C. Huang. Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association*, 94:1330–1340, 1999.

[23] T. Gneiting. Nonseparable, stationary covariance functions for space-time data. Under review for *JASA*, 2001.

[24] Z. Pan, J.H. Christensen, R.W. Arritt, W.J. Gutowski, Jr., E.S. Takle, and F. Otieno. Evaluation of uncertainties in regional climate change simulations. Submitted to *Journal of Geophysical Research*, 2001.

[25] P.J. Diggle, J.A. Tawn, and R.A. Moyeed. Model-based geostatistics (with discussion). *Applied Statistics*, 47:299–350, 1998.