State-Space Analysis of Model Error: A Probabilistic Parameter Estimation Framework With Spatial Analysis Of Variance

PI: Joshua P. Hacker
Department of Meteorology
Naval Postgraduate School
589 Dyer Road, Root Hall Rm 254
Monterey, CA 93943
phone: (831) 656-2722 fax: (831) 656-3061 email: jphacker@nps.edu

Co-PI: Cari G. Kaufman
Department of Statistics
University of California, Berkeley
367 Evans Hall
Berkeley, CA 94720
phone: (510) 643-0915 fax: (510) 642-7892 email: cgk@stat.berkeley.edu

Co-I: James Hansen
Naval Research Laboratory
Bldg 702, Rm 204
7 Grace Hopper Ave.
Monterey, CA 93943
phone: (831) 656-4741 fax: (831) 656-4769 email: jim.hansen@nrlmry.navy.mil

Award Number: N0001410WX20059

LONG-TERM GOALS

An over-arching goal in prediction science is to objectively improve numerical models of nature. Meeting that goal requires objective quantification of deficiencies in our models. The structural differences between a numerical model and a true system are difficult to ascertain in the presence of multiple sources of error. Numerical weather prediction (NWP) is subject to temporally and spatially varying error, resulting from both imperfect atmospheric models and the chaotic growth of initial-condition (IC) error. The aim of our work is to provide new methods that begin to systematically disentangle the model inadequacy signal from the initial condition error signal.

OBJECTIVES

We are engaging a comprehensive effort that uses state-of-the-science estimation methods in data assimilation (DA) and statistical modeling, including: (1) the characterization of existing model-to-model differences via heirarchical spatial modeling methods; (2) the development of a flexible representation for the various spatial and temporal scales of model error; (3) the estimation of parameters to represent those scales using a probabilistic approach to DA, namely the Ensemble Kalman Filter; and (4) the determination of whether incorporation of estimated error structure in

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **2012**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **State-Space Analysis of Model Error: A Probabilistic Parameter Estimation Framework With Spatial Analysis Of Variance**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Department of Meteorology Naval Postgraduate School 589 Dyer Road, Root Hall Rm 254 Monterey, CA 93943**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**SAR**	**9**	

improves short-term forecasts, again using heirarchical methods, this time within a formal testing framework. Research focus is on near-surface winds over both the ocean and land. The method under development are sufficiently general and can apply to a wide range of battlespace environments.

APPROACH

The technical approach includes numerical weather prediction and state estimation efforts at NPS, and statistical modeling efforts at University of California Berkeley (UCB) under sub-contract. At NPS PI Hacker and post-doc Kolczynski are implementing the Navy's Operational Global Atmospheric Prediction (NOGAPS), and two limited-area mesoscale models: the Navy's Coupled Ocean-Atmosphere Mesoscale Prediction System (COAMPS) model, and the open-source Weather Research and Forecast (WRF) model, within a state-of-the-science ensemble Data Assimilation Research Testbed (DART). The NOGAPS-DART provides global ensemble prediction capability that can be consistently applied to the COAMPS and WRF as lateral boundary conditions. Scientific objectives will be met by systematically choosing the WRF or COAMPS as the "truth," which can provide observations for assimilating into the other model. Under this approach, spatio-temporal distributions of uncertainty (error in this context) are available for analysis with special attention to second-order moments. Eventually, we will use the same framework to objectively estimate parameters in statistical models, of NWP model error, developed at UCB. Hypotheses are being formed and formally tested. This work is benefitting from collaboration Co-I James Hansen, Justin McLay, and other NRL staff. Post-doc Walter Kolczynski arrived at NPS in November 2011 to contribute.

UCB PI Cari Kaufman is working to advance the statistical methods needed to provide an objective space-time characterization of the error distributions. Uncertainty is characterized via fitting a hierarchical Bayesian model that captures the important features and variability in the data. The implied distribution from the model will be a valid stochastic spatial process under probability theory. Ideally, fitting the statistical model to different datasets should allow us to capture the significant differences between the different underlying data generating distributions. Moreover, a realistic statistical model can also simulate realistic wind fields quickly which can be beneficial for studying other processes that require surface winds as an input. Graduate student Wayne Lee (unfunded) is contributing substantially to this work. Postdoc Benjamin Shaby began work in September 2011.

WORK COMPLETED

Work in FY2012 continued toward finalizing tasks 1 and 2, which include the primary technical development for the production data sets. Work also addressed several theoretical and practical issues underlying tasks 4 and 5, which include the primary error estimation methodologies. Partial and limited funding increments led us to redirect efforts toward away from the technical developments in tasks 1 and 2. Development did not stop, but slowed considerably. Instead, we chose to focus on leveraging existing data sets to ensure progress on error-estimation methods.

At NPS we completed an initial implementation of the COAMPS-DART infrastructure, and completed a month-long WRF-DART simulation using NOGAPS-DART to provide lateral boundary conditions. Initial results from the NOGAPS-WRF-DART simulations were presented at the joint Canadian Meteorological and Oceanographic (CMOS) and American Meteorological Society (AMS) 21st Numerical Weather Prediction meetings in Montréal (May 2012). Collaboration with NRL staff was critical during this phase. Analysis of the WRF-DART, driven from NOGAPS-DART during the Oct 2009 simulation period, also proceeded. We adopted a methodology called Self-Organizing Maps

(SOMs), and wrote code to compute them. SOMs have emerged during the last decade, provide an objective method for classification, and have not previously been extended to study predictability and model inadequacy. Initial results are presented below. NPS also completed error estimation and parameter estimation experiments and analysis with the single-column version of the WRF model in DART, with results also presented below.

At UCB focus was on addressing challenges associated with applying heirarchical Bayesian techniques to large, multivariate, and non-stationary datasets typical of NWP. Several methods were proposed and tested, leading to eventual development of a canonical correlation analysis (CCA) based methodology to estimate spatially and temporally varying model errors. Progress is documented in the following results sections.

RESULTS

1. Mesoscale model results

Systematic increments from data assimilation are a linear function of model errors integrated over the assimilation interval. The challenge is to intepret the increments in space and time to reveal the scales of model errors. We have recently examined Self-Organizing Maps (SOMs; Kohonen, 1988) as a tool for identifying the coherent systematic error structures. SOMs are unattended machine-learning algorithms that produce low-dimensional "maps" of possible state vectors called "nodes" organized in a way so that nearby nodes are similar. The location of the nodes is specified and may be arranged in any pattern, though rectangular and hexagonal grids are most common. SOMs are often considered a non-linear analogue of principle component analysis (PCA). The method has been used for cluster analysis in the past, but not been used in a data assimilation context or to understand model inadequacy.

SOMs are produced through an iterative process. During each iteration a random state is chosen from the training dataset. The random state is compared to each node to identify the node closest to the chosen state based on a cost function (often the root-mean-squared error). That node is adjusted closer to the random state. Nearby nodes are also adjusted closer to the same random state (to a lesser degree). The magnitude of the adjustment and the size of the neighborhood are both reduced with each iteration, so that initially large changes are made to a large portion of the map become small changes to individual nodes.

We have applied SOMs to the ensemble-mean increment in meteorological fields produced by DART-NOGAPS-WRF, and have preliminary results. An example of 2-m temperature increments is shown in Fig. 1. The right-hand nodes are dominated by night-time, and the left-hand nodes are dominated by daytime. The increments consistently warm the temperatures at higher elevations, and cool the California Central Valley. Errors are in the opposite sense. cool errors persist over the Sierras, but errors over the Rockies are modulated by other factors. At this point it is not clear whether those other factors vary on seasonal or synoptic time scales. Results demonstrate that SOMs objectively classify the increments, and no *a priori* data conditioning is needed. Methods for characterizing the spatial structures of the error, such as those being developed at UCB in this project, are still needed. Additionally, we plan to use data from the full ensemble (the entire distribution) to determine if the SOM can separate forecasts based on the variance of the ensemble, which is a proxy for the certainty of the forecast.

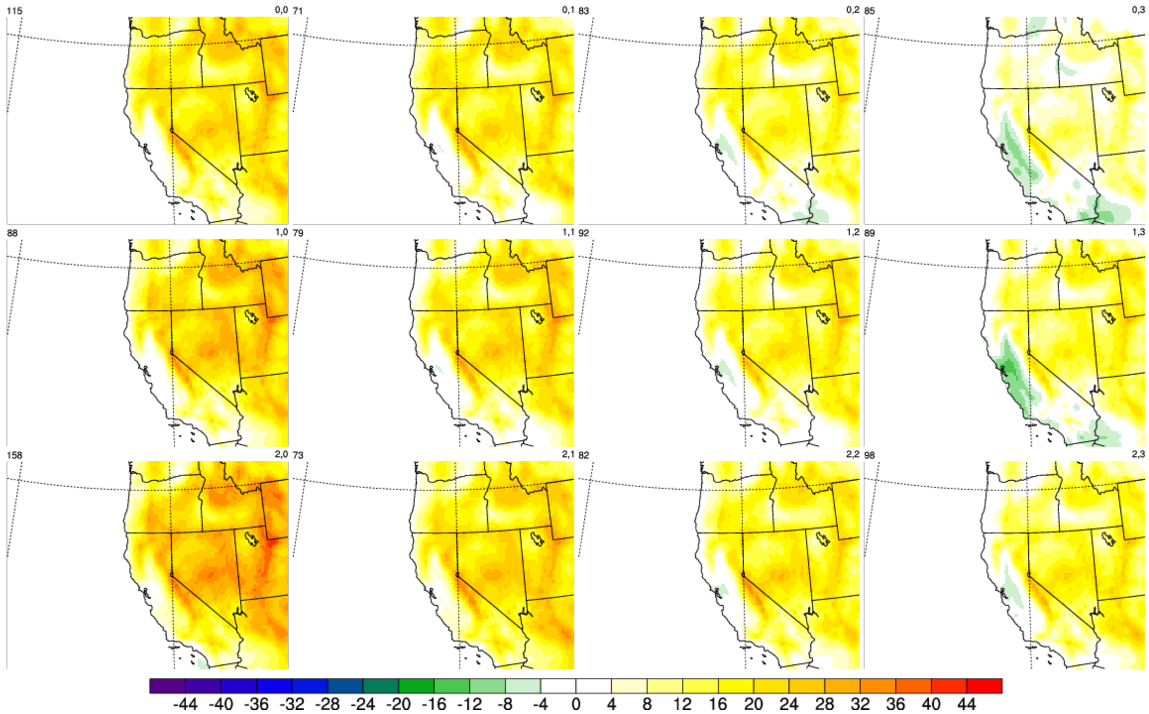

Figure 1: Self-Organizing Map (SOM) for 2-m temperature analysis increments.

In a separate study leveraging other ongoing work, we gained results demonstrating the efficacy of ensemble DA in identifying model inadequacy. Experiments with the single-column implementation of the WRF model provide a basis for deducing land-atmosphere coupling errors in the model. Coupling occurs both through heat and moisture fluxes through the land-atmosphere interface and roughness sub-layer, and turbulent heat, moisture, and momentum fluxes through the atmospheric surface layer. This work primarily addresses the turbulent fluxes, which are parameterized following Monin-Obukhov similarity theory applied to the atmospheric surface layer. By combining ensemble data assimilation and parameter estimation, the model error can be characterized. Ensemble data assimilation of 2-m temperature and water vapor mixing ratio, and 10-m wind components, forces the model to follow observations during a month-long simulation for a column over the well-instrumented ARM Central Facility near Lamont, OK. One-hour errors in predicted observations are systematically small but non-zero, and the systematic errors measure bias as a function of local time of day. Analysis increments for state elements nearby (15-m AGL) can be too small or have the wrong sign, indicating systematically biased covariances and model error. Experiments using the ensemble filter to objectively estimate a parameter controlling the thermal land-atmosphere coupling show that the parameter adapts to offset the model errors, but that the errors cannot be eliminated. Results suggest either structural error or further parametric error that may be difficult to estimate. Experiments omitting atypical observations such as soil and flux measurements lead to qualitatively similar deductions, showing potential for assimilating common in-situ observations as an inexpensive framework for deducing and isolating model errors. These results have been submitted to a peer-reviewed journal (Hacker and Angevine 2012). The methodology could be easily extended to an over-water site with suitable data.

2. Modeling Uncertainty in Surface Wind Fields

One of our goals has been to construct a probabilistic model to characterize the dependence structures in surface wind fields. This will allow us to compare wind fields from different models within a hierarchical Bayesian framework. In the previous report, we described our incorporation of the geostrophic relationship into a hierarchical model to relate surface winds to pressure gradients. This simplified the multivariate aspect of wind and efficiently captured the dependencies between different wind components. We also allowed the geostrophic coefficients to vary spatially, which was an improvement upon the original model proposed by Royle et al. (1999). However, in implementing this model, we faced computational issues when generating samples from the posterior distribution, due to the large size of the data.

We proposed to use a Gaussian Markov random field (GMRF) approximation to the original spatial correlation function (Lindgren and Rue, 2011). Boundary regions were problematic while dealing with the precision matrix instead of the covariance matrix, because the conditional relationship for the boundary points is different than for interior points. We examined a strategy of constructing the precision matrix by using the original covariance for the boundary points and embedding the interior points using the conditional relationship. This method is still computationally intensive, and the embedding showed a consist bias in estimating the parameters in our toy examples. For example, Figure 1 shows a comparison between the original Matérn model and the GMRF approximation while estimating the consistently estimable parameter $\sigma^2/\rho^{2\nu}$ in the Matérn model. We published this simulation study in Lee and Kaufman (2011).

We also examined the effectiveness of our hierarchical model on a global climate model (the Japanese Model MIROC3.2 at medium resolution under the pre-industrial experiment scenario). This dataset has 8200 spatial locations over the globe over 40 years. The spatial locations are on a regular latitude/longitude grid and thus are suitable for the GMRF approximation method. To address the highly correlated posterior samples, we originally proposed to use slice sampling. However, slice sampling in multiple dimensions is more complicated to code and harder to tune than the adaptive metropolis Hastings algorithm. It may also suffer from high autocorrelation (Agarwal and Gelfand 2005). Therefore, we implemented the adaptive metropolis Hastings algorithm (Shaby and Wells 2010), which requires little tuning, is less prone to coding errors, and returns less autocorrelated parameters than the naive Gibbs sampler. Finally, we proposed a flexible nonstationary prior based on topography, but few papers have shown the benefits of introducing such complexity into the model. In particular, Reich et al. (2011) showed little gain in prediction accuracy from introducing a flexible covariance model that depends on covariates. Note this is different from introducing covariates into the mean, such as we have done with the geostrophic component.

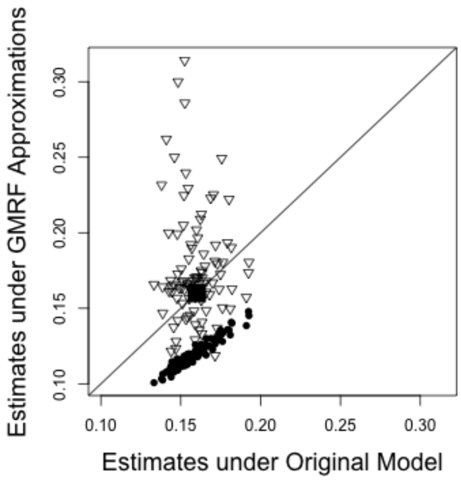

Figure 2: Estimates for $\sigma^2/\rho^{2\nu}$ with effective range of 10 on a 20x20 unit grid and with true values $\nu = 1$ and $\sigma^2 = 2$. The solid square represents the true parameter value, the solid circles represent the results from embedding, and the hollow triangles represent the result from simply ignoring the boundary issue on a small grid with a relatively large effective range.

3. Separating spatial scales with model-based EOFs

Mismatch between model output and observations, or between competing models, is often the result of multiple processes or features. These error fields typically are analyzed using empirical orthogonal functions (EOFs), which partition the fields into orthogonal components that maximally account for variance. It may be the case that components of the error fields vary at characteristic time scales. By taking advantage of these different characteristic time scales, we can study the constituent sources of the error. Unfortunately, there is no way to direct EOF computations to consider different periods of variation. In contrast, we use hierarchical Bayesian models to explicitly separate error fields that occur at different time scales.

The main building block for our models is the observation of Tipping and Bishop (1999) that EOF construction is equivalent to maximizing a particular probability model with respect to the data. EOF computations decompose errors into p spatial fields, which we denote as $\mathbf{m}_1,\ldots,\mathbf{m}_p$, each scaled by p time series, which we denote as $\mathbf{z}_1,\ldots,\mathbf{z}_p$. The standard probability model corresponding to EOFs assumes that each spatial field $\mathbf{m}_i$ is independent across locations and that each time series $\mathbf{z}_i$ is independent through time. Our Bayesian method works by assigning a prior distribution to each $\mathbf{m}_i$ that encourages nearby locations to behave similarly, and a prior distribution on each $\mathbf{z}_i$ to encourage temporal structures that exhibit characteristic frequencies. These prior distributions are modeled as Gaussian processes with prescribed covariance structures. For example, Figure 2 shows covariance functions that induce time series with periods of five (left) and twenty (right) years. Figure 3 shows the results of applying this model to simulated data. The left panel shows the true constituent error fields (a), the estimated error fields from the Bayesian model (b), and the estimated error fields from standard EOF analysis. The right panel shows the true error time series (black line) with time time scales of five and twenty years, the estimated time series from the Bayesian model (red), and the estimated time series from standard EOF analysis (green). This simulation shows that when the error arises from spatial processes with different characteristic time scales, the Bayesian model successfully partitions the error, while traditional EOFs falter.

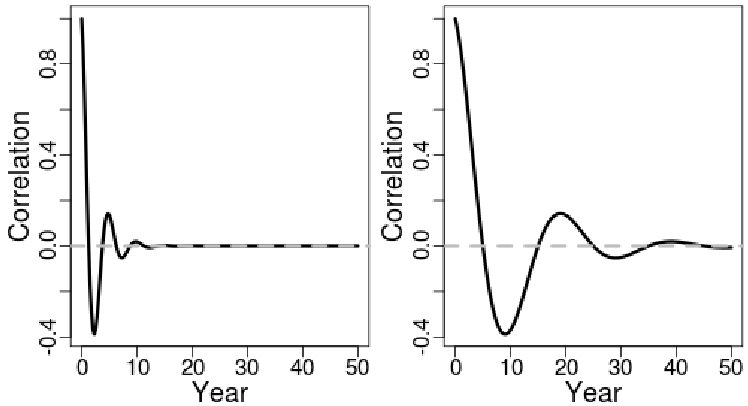

Figure 3: Covariance functions that induce periodicity of five (left) and twenty (right) years in EOF time series.

In our future work, we plan to extend the model to canonical correlation analysis (CCA), which is similar in spirit to EOF analysis but tries to maximize correlations across datasets, rather than variability within a single dataset. The advantage of a probabilistic CCA to this project is that it will allow us to characterize components of explained variability within a data assimilation framework. Specifically, by using datasets consisting of 1) ensemble analyses and 2) differences between ensemble analyses at the next time step and the corresponding forecasts, we can describe the spatial components of the analysis that are maximally correlated with subsequent structural model errors. This will allow a better understanding of the sources of these errors.

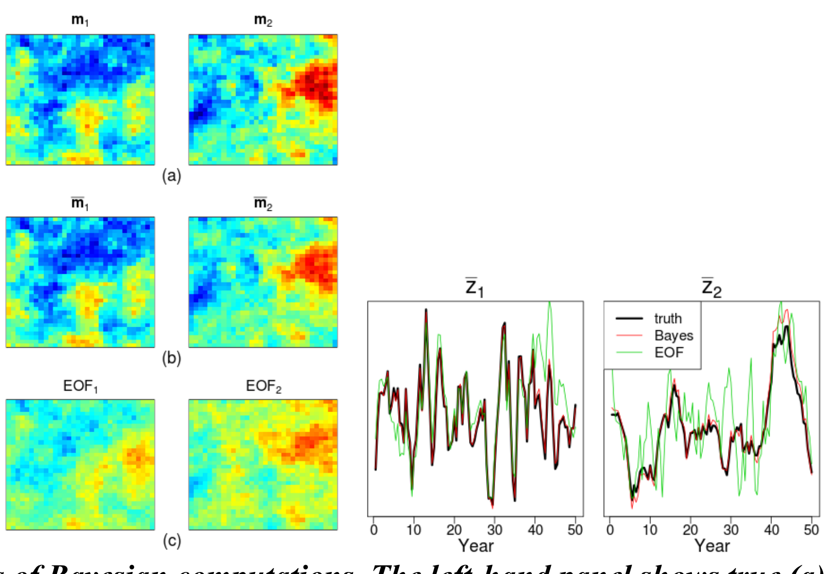

Figure 4: Results of Bayesian computations. The left-hand panel shows true (a) and estimated (b and c) spatial fields. The Bayesian estimates (b) correspond closely to the true fields, while the EOF fields (c) do not separate the components as clearly. The right-hand panel shows the true and estimated error component time series. The Bayesian model (red) tracks the true time series (black) well, while the EOF time series (green) is not able to separate signals from different temporal scales.

IMPACT/APPLICATIONS

The bulk of DoN day-to-day operations rely on accurate predictions of winds, seas, ceiling, and visibility. The focus of the proposed work is to identify inadequacies associated with the modeled atmospheric boundary layer. Any discoveries that enable the improvement of boundary layer modeling will ultimately have a positive impact on Navy warfighters.

The proposed methods have the potential to enable essential improvement in modeling capability. Instead of tuning models based on intuition, we are forming a foundation for objective identification of model errors. Those errors could immediately be accounted for in probablistic forecast systems, and also be subject to physical interpretation by subject experts.

RELATED PROJECTS

The MATERHORN project (http://www.nd.edu/~dynamics/materhorn/index.html), funded by ONR, seeks to improve atmospheric predictability over complex terrain. It is similarly focused on predictions in the atmospheric boundary layer. Rather than a focus on model inadequacy, MATERHORN focuses on field programs aimed at improving models via direct comparison to observations, and quantifying optimal observing strategies for improving predictions. PI Hacker is using some of the technical developments here to aid that effort, and vice versa.

REFERENCES

Agarwal D.K. and A. E. Gelfand, 2005: Slice sampling for simulation based fitting of spatial data models. *Statistics and Computing*, 61-69.

Kohonen, T., 1988: *Self-organization and associative memory*. Springer-Verlag, New York, 312 pp.

Lee, W. and C. Kaufman, 2011: Comment on "An Explicit Link between Gaussian Fields and Gaussian Markov random Fields: The SPDE Approach" by Lindgren, F., Rue, H., and Lindstrom, J. *J. Roy. Stat. Soc., Series B: Methodology*, 479-480.

Lindgren, F. and H. Rue, 2011: An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J. Roy. Stat. Soc., Series B: Methodology*, 423-498.

Reich, B.J., J. Eidsvik, M. Guindani, A.J. Nail, and A. M. Schmidt, 2011: A class of covariate-dependent spatiotemporal covariance functions for the analysis of daily ozone concentration. *Ann. of Appl. Stat.*, **5**, 2425-2447.

Royle, J.A., L.M. Berliner, C.K. Wikle, and R. Milliff, 1999: A hierarchical spatial model for constructing wind fields from scatterometer data in the Labrador Sea. Springer-Verlag.

Shaby, B. and M.T. Wells, 2010: Exploring an Adaptive Metropolis Algorithm. Working Paper, Department of Statistical Science, Duke University, 11-14.

Tipping, M.E. and C.M. Bishop, 1999: Probabilistic principal component analysis. *J. Roy. Stat. Soc., Series B: Methodology*, **61**, 611-622.

PUBLICATIONS

Hacker, J. and W. Angevine, 2012: Ensemble data assimilation to characterize land-atmosphere coupling errors in numerical weather prediction models. Submitted to *Mon. Wea. Rev.*