Final Report *A Statistics Program at the National Center for Atmospheric Research*

1 Overview

DMS-9815344

The award DMS-9815344 made by the Statistics Program, Division of Mathematical Sciences has provided significant support for the Geophysical Statistics Project (GSP), a statistical research and training group embedded within the National Center for Atmospheric Research (NCAR — NCAR has a staff of more than a 1000 members including several hundred Ph.D. scientists and receives more than half of its funding from the NSF Division of Atmospheric Sciences). GSP has been a successful model for making contributions to atmospheric and oceanographic sciences through the development and application of new statistical methodology. This has been achieved by supporting a critical mass of statistics visitors at NCAR, training young researchers, and maintaining close collaboration with NCAR geophysical scientists.

The long-term impact on statistics has been substantial; not only contributing a large body of published research and software, but also through an extended network of university researchers who have participated in GSP as visitors and postdoctoral fellows (PF). These connections to the larger statistics community has been reinforced by two workshops held at NCAR and supported by this grant. GSP's impact on geophysical researchers has been through *statistics by example*.

Successful collaboration with NCAR scientists has provided case studies where new statistical methods have tackled problems and questions that otherwise would not have been attempted. Scientific applications identified through GSP have also suggested new statistical models and theory. One measure of success is GSP's involvement in two NCAR strategic inititiatives, Data Assimilation and Climate and Weather Impacts Assessment. These large research efforts represent new directions for the Center and have provided additional support for statistics PFs and graduate students. Finally, on October 1, 2004 the Institute for Mathematics Applied to Geosciences (IMAGe) was formed at NCAR with GSP as a founding section and gives statistical science a permanent home at NCAR. Part of the vision of IMAGe came from the ability of GSP to transfer cutting edge statistical methodology across the breadth of NCAR science.

Besides this overview report more details on the products from this grant can be found at the GSP web page (www.cgd.ucar.edu/stats) including annual science reports, useful datasets, and R-based statistics software.

1.1 Scientific Accomplishments

Because of the emphasis on collaborative research, GSP scientific results have appeared in leading statistical, geophysical, and environmental statistical publications. A noteworthy feature of this research is that the majority of publications are co-authored by both geophysical and statistical scientists. In order to succinctly report this wide range of research; project publications, submitted manuscripts and technical reports have been organized into some broad categories. This outline is complemented by more detailed highlights of some key scientific contributions.

Statistical methodology and theory
Time series analysis [8],[7],[22],[30],[31],[34],[49],[54],[64],[62],[66],[90],[89],[100],[116],[124],[125],[129],[102];
Spatial statistics [26],[27],[37],[113],[70],[76],[78],[83],[106],[51];
Space/time processes [16],[104],[109],[107];
Nonlinear time series and dynamical systems [6],[93];
Nonparametric regression [18],[48],[22],[50],[79],[80],[103],[119],[121],[122],[123],[128];

Bayesian methodology [35],[36];
Analysis of functional data [59],[60].

Geophysical Applications
Analysis of the climate record [52],[72],[85],[87],[120],[126];
Time series analysis [1],[17],[38],[44],[45],[53],[63],[77],[82],[81],[98],[112],[101],[118];
Spatial fields [20],[24],[55],[65];
Validation and statistical analogs of numerical models [2],[4],[5],[3],[19],[69],[111],[28];
Extreme events [43],[41],[47],[57],[67],[68];
Climate change [15],[61],[14],[56],[58],[75],[97],[127];
Forecasting and data assimilation [9],[10],[12],[21],[88],[95],[96],[92].

Overviews and surveys [11],[13],[29],[39],[40],[42],[71],[73],[74],[108],[110],[105].

1.2 Research Highlights

The following sampling of GSP research accomplishments helps emphasize the substantial collaboration among the postdoctoral fellows/visiting scientists and resident NCAR scientists. These highlights are meant to be illustrative; they only constitute a subset of GSP research activities.

Statistics for large spatial datasets. A challenge to spatial statistics is to scale methods from the moderate-size datasets considered in the statistics literature to the large and substantive records assembled to study the atmosphere and oceans. As a testbed problem, GSP has constructed spatially and temporally complete monthly meteorological records for the coterminous U.S. based on historical meteorological data ([37]). The work implemented local covariance models that blended sample correlations with a global covariance model, successfully handled a large number of spatial locations (8000-12000) over many time points (1200 months) and provided cross-validated measures of uncertainty of the infilled estimates.

Motivated by large geophysical datasets, two related projects are to adapt multi-resolution models for nonstationary ([76]) and non-Gaussian fields ([64]). The compact tapering of covariance functions to introduce sparsity in a covariance matrix is a successful method in data assimilation to handle large covariance matrices. This method has been studied from a statistical point of view ([114], [115]) and helps to explain why tapering works. Moreover, using sparse matrix methods and tapered covariances, this project has demostrated that large spatial datasets can be analyzed in a high-level, interactive environment such as R.

Statistical analogs of atmospheric numerical models. The size and complexity of output produced by state-of-the-art geophysical models often require statistical descriptions to understand and summarize the overall behavior and specific aspects of the simulated system. Cloud-resolving models (CRM) are fine-scale convective models for the atmosphere that can create and evolve individual clouds. Extreme value distributions were used to model maximum vertical velocity ([68]) in CRM output as a stochastic representation of processes that are important for large-scale atmospheric general circulation models (GCM). One result was a quantification of the impact of wind exceedances on the total mass flux in the CRM domain. Also based on model output, multi-resolution methods have been designed to extract features representing the organization of convection (e.g. squall lines) ([69]), providing an objective classification of the types of convection in a model simulation. Current GCMs rely on heuristic schemes to overlap cloud layers which may introduce biases in the amount of absorbed and reflected radiation in the atmosphere. As an alternative, a statistically based cloud overlap model was estimated from CRM experiments and was shown to give better simulation of the radiative fluxes ([111]).

At larger spatial scales, a combination of nonlinear time series models and clustering algorithms have been used to summarize the dynamical properties of the large-scale atmospheric circulation from a long integration of the NCAR Community Climate Model (CCM0) ([93]). The statistical model quantified the system's nonlinear behavior and identified regimes in phase-space whose dwell times were found to be consistent with the timescales of atmospheric motions.

Data assimilation and forecasting. Data assimilation refers to combining information from a geophysical model with observations to give an improved estimate of the state. This process lies at the heart of operational numerical weather prediction and has formal relationships to nonlinear filtering. Current work in the atmospheric sciences community is on a (Monte Carlo based) Ensemble Kalman Filter (EnKF). GSP statistical contributions include extensions to non-Gaussian distributions for the state ([9],[10],[112]). By applying multivariate statistical theory, it was possible to analyze and better characterize the dependence of EnKF performance on ensemble size ([28]).

Aircraft are substantially affected by the presence of clear air turbulence. This localized atmospheric phenomenon is very difficult to predict using conventional weather forecast models. A successful approach trained a flexible discriminant model on pilot reports of turbulence using turbulence indices derived from a numerical weather prediction model as predictors ([95],[96]). These statistical forecasts outperform other operational approaches currently tested at NCAR.

Analysis of long-term climate series. Long-term variation in solar irradiance can be confounded with possible climate change due to human activities; consequently, it is important to investigate the relationship between solar flux and surface temperatures. A multi-resolution analysis was applied to several proxy measures of solar flux and to proxies for temperature to determine the relationship among these climate time series ([77]). One main result indicates evidence for the forcing of temperature by solar flux on a timescale consistent with the Gleissberg cycle (period of ≈ 85 years). Another influence on temperature records are large, explosive volcanic eruptions that inject substantial amounts of sulfur dioxide into the lower stratosphere. A multiprocess state-space model was estimated and was able to identify and quantify the impulse-like volcanic signal in proxy records, thus distinguishing those temperature proxy series that are more sensitive to aerosol loading ([1],[63]).

Spatial extremes. Extremes in geophysical processes can play a signficant role when considering the impact of weather and climate on society. During this period two important and long term threads of reaserach were initiated that apply the statistics for extremes to climate model output and historical observations. Some initial work on climate model projections analyzes the sources of decreases in number of frost days ([57]) and the increase in heat waves ([56]) under the scenario of a warmer climate due to increased greenhouse gas emissions. At an intermediate time scale, extreme behavior in daily ozone air quality was modeled with the intent of understanding the spatial distribution of exceedences of the EPA standard. The standard is based on an extreme order statistic (fourth highest out of approximately 184 days). The spatial field for this statistic was inferred using a Monte Carlo approach based on a space-time model fit to daily data [32] and also by fitting a generalized Pareto distribution to the individual station data and then extrapolating the distribution parameters over space [117]. Despite the very different perspectives that underlie these approaches the results agree for a case study based on monitoring data for the Southeast. Another component of this research area is the analysis of precipitation data to determine the distribution of extreme events. An 'exceedance over threshold' approach has been applied to hourly data from Colorado with the intent of generating a precipitation atlas of return levels, and also applied to regional climate model output in order to estimate extreme quantiles from short model integrations.

1.3 Postdoctoral and Graduate Training

The highest priority of the project has been to recruit and train new recipients of Ph.D.'s in statistics and probability. GSP has been able to attract high quality individuals from nationally-ranked programs in statistics and has placed the majority of GSP fellows in university positions or in other research environments (the second line in the list below indicates the fellow's subsequent position). One milestone that indicates the growing maturity of GSP is the presence of Tebaldi as a Project Scientist and Gilleland as an Associate Scientist on on the permanent NCAR scientific staff.

Former postdoctoral fellows and placement:

Claudia Tebaldi, Statistics, Duke U., 1997; GSP 12/1997–6/2000
Project Scientist, Environmental and Societal Impacts Group, NCAR.

Philippe Naveau, Statistics, Colorado State U., 1998; GSP 7/1998–12/2000
Assistant Professor, Dept. of Applied Mathematics, U. of Colorado - Boulder

Craig Johns, Statistics, U. of California - Davis, 1999; GSP 8/1999–7/2001
Assistant Professor, Dept. of Mathematics, U. of Colorado - Denver.

Thomas Bengtsson, Dept. of Mathematics, U. of Missouri, 1999; GSP 8/2000–8/2003
Jerzy Neyman (visiting) Assistant Professor, U. of California at Berkeley.

Enrica Bellone, Statistics, U. Washington, 2000; GSP 9/2000–9/2002
Research Associate, Dept. of Civil and Environmental Engineering, Imperial College, London.

Hee-Seok Oh, Statistics, Texas A&M, 1999; GSP 9/2000–8/2002
Assistant Professor, U. of Alberta.

Sarah Streett, Statistics, Colorado State U., 2000; GSP 9/2000-8/2003
Mathematical Statistician, National Institute of Standards and Technology, Boulder CO.

Brandon Whitcher, Statistics, U. Washington, 1998; GSP 9/2000–9/2002
Senior Researcher, Translational Medicine & Technology, GlaxoSmithKline, London.

Current postdoctoral fellows and research interests:

Reinhard Furrer, Statistics, Swiss Federal Institute of Technology, 9/2002– *Spatial statistics, filtering.*

Jarrett Barber, Statistics, North Carolina State U., 9/2002– *Misaligned data, nonstationary fields.* (75% appointment Duke U.)

Uli Schneider, Applied Mathematics, CU - Boulder, 7/2003– *Perfect sampling, extremes.*

Tomoko Matsuo, Atmospheric Science, State U. of New York - Stonybrook 7/2003– *Data assimilation.*

Dorin Drignei, Statistics, Iowa State U., 6/2004– *Analysis of computer experiments.*

GSP-supported graduate students and research interests:

Eric Gilleland, Colorado State U., 9/1999–12/2002 Ph D. Statistics 12/2004, Associate Scientist, Research Applications Program, NCAR 1/2003– *Spatial statistics, extremes.*

Curtis Storley, Colorado State U., 6/2003– *Data mining, feature tracking.*

Dan Cooley, CU - Boulder, 5/2003– *Paleoclimate, extremes.*

1.4 Outreach

Workshops. GSP-sponsored workshops covered statistics relevant to applications in the geophysical sciences and provided an introduction to some outstanding problems in the geosciences. Each workshop was attended by more than 50 participants divided among graduate students, post docs, and senior researchers in statistics and the atmospheric sciences. Additionally, a significant miniworkshop on Climate and Health (5/2004) involved biostatisticians from the John Hopkins School of Public Health.

Statistics for Large Datasets [24–26 July 2000] included speakers on scientific visualization and data mining as well as statistics for geophysical and environmental problems. It was jointly sponsored by the National Research Center for Statistics and the Environment, University of Washington.

Spatio-Temporal Modeling [1–6 June 2003] brought together senior and young researchers in applied mathematics, statistics and environmental sciences and was jointly sponsored by the Statistical and Applied Mathematical Sciences Institute.

Undergraduate programs. GSP and the Data Assimilation Initiative at NCAR sponsored a team of four students in connection with the RIPS summer program at the Institute for Pure and Applied Mathematics , U. of California – Los Angeles. The project was to create a version of the ensemble Kalman filter that is robust to observational outliers. The student team came to NCAR after the completion of the RIPS program to present their work to the data assimilation group.

Statistical software. GSP has been active in developing statistical packages suited to the analysis of geophysical or spatial data. Much of this software are in the form of fully-documented, open source packages in the R langauge and posted on the comprehensive R archive network (CRAN) `www.R-project.org`. The base URL for GSP software is `www.cgd.ucar.edu/stats/software.shtml` and the main postings are:

`fields`: An R package focused on the analysis of spatial data including large datasets and nonstationary covariance models and simulation.

`Waveslim`: A package in R for wavelet analysis.

`RadioSonde`: [R,S] functions to ingest common radiosonde data files and create SKEW-T/log p plots and wind profiles.

`DI: Design Interface`: A graphically-based interactive program in S-PLUS© that allows one to edit and evaluate spatial designs.

`extRemes toolkit`: A package in R for extremes.

`Krisp`: An R package for spatial statistics methods (Kriging) using sparse matrix algebra.

Significant data sets. The GSP web site also created a area with substantive geophysical datasets that can be used as test cases and example for statistical methods. One important data set are monthly average daily min/max temperatures and monthly total precipitation for the US 1985-1997.

Visiting scientists to GSP. To foster connections with the university statistics community, GSP also maintains a shorter term visitor program. Senior visitors have been chosen with the postdoctoral fellows in mind, as they often serve as additional mentors. A typical short-term visitor gives one or more seminars to GSP and local statistics groups and meets individually with each GSP post doc. GSP has supported several longer term visitors from local statistics group to visit on regular basis; e.g. R. Jones and S. Sain (U. Colorado) and T. Lee (Colorado State U.). A number of statistics graduate students have also visited GSP in conjunction with their thesis research or as interns.

Partial visitor list:

Senior visitors and faculty: *W. Eddy*, Carnegie Mellon U.; *D. Ruppert*, Cornell U.; *J. Berger, M. Clyde, D. Higdon, M. West*, Duke U.; *A. Grady, D. Holland*, EPA; *Z.-Q. Lu*, Hong Kong U.; *T.-H. Li*, IBM; *N. Cressie*, Iowa State U.; *A. Chedin*, IPSL-LMD Ecole Polytechnique; *A. Braverman*, JPL; *T. Louis, R. Peng, J. Samet, S. Zeger* Johns Hopkins; *I. Jolliffe*, King's College; *A. Kaplan*, Lamont-Doherty; *L. Sparling*, NASA/Goddard; *G. Huerta*, New Mexico State U.; *M. Genton*, North Carolina State U.; *M. Berliner*, Ohio State U.; *D. Cox, D. Scott*, Rice U.; *J. Red-Horse*, Sandia National Labs; *T. Hastie*, Stanford U.; *S. Morgenthaler*, Swiss Fed. Inst. of Technology; *R. Eubanks*, Texas A&M; *G. Nason*, U. of Bristol; *P. Bickel, J. Rice, B. Yu*, U. of California - Berkeley; *R. Shumway*, U. of California - Davis; *K. Ide*, U. of California - Los Angeles; *B. Sanso*, U. of California - Santa Cruz; *J. Corcoran*, U. of Colorado - Boulder; *K. Kafadar, S. Sain*, U. of Colorado, Denver; *R. Lund*, U. of Georgia; *B. Bailey*, U. of Illinois; *F. Turkman, A. Turkman*, U. of Lisbon; *C. Wikle*, U. of Missouri; *W. Dunsmuir*, U. of New South Wales; *J. Stroud*, U. Pennsylvania; *S. Marron, R. Smith*, U. of North Carolina; *T. Gneiting, P. Guttorp*, U. of Washington; *T. Haas*, U. of Wisconsin.

Students: *T. Shao*, U. of Alberta, CA; *C. Kaufman*, Carnegie Mellon U.; *R. Buchberger, M. Eschenberg*, Colorado State U.; *C. Paciorek*, Carnegie Mellon U.; *W. Chang*, California Insit. Tech.; *E. Gilleland, C. Storlie*, Colorado State U.; *R. Paolo*, Duke U.; *P. Abbit*, Iowa State U.; *M. Vrac*, IPSL-LMD Ecole Polytechnique. *A. Nail, K. Madsen*, North Carolina State U.; *S. Smith*, Spellman College; *B. Fournier*, Swiss Fed. Inst. of Technology; *Y. Munoz*, Texas A&M; *L. Welty*, U. Chicago; *H. Naderi*, U. of California - Berkeley; *A. Pignotti*, U. of California - Santa Cruz; *U. Schneider, D. Cooley*, U. of Colorado; *B. Shi*, Georgia Inst. Tech.; *V. Bulaevskaya*, U. of Minnesota; *P. Caragea*, U. of North Carolina; *B. Das*, U. of Washington.

References

GSP articles published or in review

[1] C. Ammann and P. Naveau. Statistical analysis of tropical explosive volcanism occurrences over the last 6 centuries. *Geophysical Research Letters*, 30(5), 2003. 1210, doi: 10.1029/2002GL016388.

[2] B. A. Bailey, L. M. Berliner, W. Collins, D. W. Nychka, and J. T. Kiehl. Neural networks: Cloud parameterizations. In L. M. Berliner, D. Nychka, and T. Hoar, editors, *Studies in the Atmospheric Sciences*, volume 144 of *Lecture Notes in Statistics*, pages 97–116, New York, 2000. Springer-Verlag.

[3] B. A. Bailey, L. M. Berliner, W. Collins, D. W. Nychka, and J. T. Kiehl. Spatio-temporal neural network modeling of cloud cover. *Technometrics*, 2003. In review.

[4] B. A. Bailey and S. C. Doney. Quantifying the effects of noise on biogeochemical models. In *Computing Science and Statistics*, volume 32 of *Proceedings of the 32nd Symposium of the Interface*, pages 447–453, 2000.

[5] B. A. Bailey, S. C. Doney, and I. D. Lima. Quantifying the effects of dynamical noise on the predictability of a simple ecosystem model. *Environmetrics*,15:337-355, 2004. In press.

[6] B. A. Bailey, D. W. Nychka, and S. Ellner. A central limit theorem for local Lyapunov exponents. *Stochastic Processes and their Applications*, 2003. In review.

[7] T. Bengtsson and J. Cavanaugh. A simulation-based model selection criteria for the state-space framework. *Journal of Computational and Graphical Statistics*, 2002. In review.

[8] T. Bengtsson and J. Cavanaugh. State-space discrimination using Kullback information measures. *Technometrics*, 2002. In review.

[9] T. Bengtsson and D. W. Nychka. Adaptive methods in numerical weather. In *Proceedings of the 1st Spanish Workshop on Spatio-Temporal Modeling of Environmental Processes.*, Benicassim, Castellon (Spain), 2001.

[10] T. Bengtsson, C. Snyder, and D. W. Nychka. Toward a nonlinear ensemble filter for high-dimensional systems. *Journal of Geophysical Research*, 108(D24), 2003. 8775, doi:10.1029/2002JD002900.

[11] L. M. Berliner. Hierarchical Bayesian modeling in the environmental sciences. *Allgemeines Statistisches Archiv, Journal of the German Statistical Society*, 84:141–153, 2000.

[12] L. M. Berliner. Monte Carlo based ensemble forecasting. *Statistics and Computing*, 11:269–275, 2001.

[13] L. M. Berliner. Physical-statistical modeling in geophysics. *Journal of Geophysical Research*, 108(D24), 2003. 8776, doi: 10.1029/2002JD002865.

[14] L. M. Berliner. Uncertainty in climate change. *Statistical Science*, 2003. To appear.

[15] L. M. Berliner, R. A. Levine, and D. J. Shea. Bayesian climate change assessment. *Journal of Climate*, 13:3805–3820, 2000.

[16] L. M. Berliner, R. F. Milliff, and C. K. Wikle. Bayesian hierarchical modeling of air-sea interaction. *Journal of Geophysical Research*, 108(C4), 2003. 3104, doi:10.1029/2002JC001413.

[17] 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, 2000.

[18] D. J. Cummins, T. G. Filloon, and D. Nychka. Confidence intervals for nonparametric curve estimates: Toward more uniform point-wise coverage. *Journal of the American Statistical Association*, 96:233–246, 2001.

[19] J. M. Davis, D. W. Nychka, and B. A. Bailey. A comparison of the regional oxidant model (ROM) with observed data. *Atmospheric Environment*, 34:2413–2423, 2000.

[20] S. C. Doney, D. M. Glover, S. McCue, and M. Fuentes. Mesoscale variability of Sea-viewing Wide Field-of-view Sensor (SeaWiFS) satellite ocean color: Global patterns and spatial scales. *Journal of Geophysical Research, Oceans*, 108(C2), 2003. 3024, doi:10.1029/2001JC000843.

[21] R. M. Errico, L. Fillion, D. W. Nychka, and Z.-Q. Lu. Some statistical considerations associated with the data assimilation of precipitation observations. *Quarterly Journal of the Royal Meteorological Society*, 126(562):339–359, 2000. Part A.

[22] Y. Fan and B. Whitcher. A wavelet solution to the spurious regression of fractionally differenced processes. *Applied Stochastic Models in Business and Industry*, 19(3):171–183, 2003.

[23] B. Fournier, R. Furrer, T. Gsponer, T., and E.-M. Restle. *Proceedings of the 13th European Young Statisticans Meeting*, September 21-26, Ovronnaz, Switzerland, Stmpfli AG, ISBN 3-908152-17-8, 2003.

[24] M. Fuentes, S. C. Doney, D. M. Glover, and S. J. McCue. Spatial structure of the SeaWiFS ocean color data in the North Atlantic Ocean. In L. M. Berliner, D. Nychka, and T. Hoar, editors, *Studies in the Atmospheric Sciences*, volume 144 of *Lecture Notes in Statistics*, pages 153–171, New York, 2000. Springer-Verlag.

[25] M. Fuentes, T. G. F. Kittel, and D. Nychka. Sensitivity of ecological models to spatial-temporal estimation of their climate drivers: Statistical ensembles for forcing. *Ecological Applications*, 2004 To appear.

[26] R. Furrer. Covariance estimation under spatial dependence. *Journal of Multivariate Analysis*, 2004. To appear.

[27] R. Furrer, M. G. Genton, D. Nychka. Covariance Tapering for Interpolation of Large Spatial Datasets. *Journal of Computational and Graphical Statistics*, 2004. in revision.

[28] R. Furrer, T. Bengtsson. Evaluation of Forecast and Analysis Covariances in Kalman Filter Variants. *Manuscript*, 2004. in preparation.

[29] R. Gençay, F. Selçuk, and B. Whitcher. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics.* Academic Press, San Diego, September 2001. ISBN: 0-12-279670-5.

[30] R. Gençay, F. Selçuk, and B. Whitcher. Asymmetry of information flow between fx volatilities across time scales. *Econometrica*, 2003. In review.

[31] R. Gençay, F. Selçuk, and B. Whitcher. Multiscale systematic risk. *Journal of International Money & Finance*, 2003. In review.

[32] E. Gilleland and D. Nychka. Statistical Models for Monitoring and Regulating Ground-level Ozone. *Environmetrics*, 2004. To appear.

[33] E. Gilleland and D. Nychka. *Statistical Models for the Ozone Air Quality Standard.* Department of Statistics, Colorado State University Ph D. Dissertation, Fort Collins, CO

[34] G. K. Grunwald and R. H. Jones. Markov models for time series with mixed distribution. *Environmetrics*, 11:327–339, 2000.

[35] T. J. Hoar. Bayes meets big iron. In *Proceedings of the 1999 Joint Statistical meetings*, Journal of the American Statistical Association, 1999.

[36] T. J. Hoar, R. F. Milliff, D. W. Nychka, C. K. Wikle, and L. M. Berliner. Wind realizations from a Bayesian hierarchical model: Computation for atmosphere-ocean research applications. *Journal of Computational and Graphical Statistics*, 12:781-807, 2003.

[37] C. J. Johns, D. W. Nychka, T. G. F. Kittel, and C. Daly. Infilling sparse records of spatial fields. *Journal of the American Statistical Association*, 98:796–806, 2003.

[38] R. H. Jones and A. L. Jones. Testing skill in earthquake predictions. *Seismological Research Letters*, 74(6);753, 2004.

[39] A. Kaplan, M. A. Cane, and Y. Kushnir. Reduced space approach to the optimal analysis interpolation of historical marine observations: Accomplishments, difficulties, and prospects. In *WMO Guide to the Applications of Marine Climatology*, volume WMO/TD-1081, Geneva, Switzerland, 199–216, 2003. World Meteorological Organization.

[40] R. W. Katz. Sir Gilbert Walker and a connection between El Niño and statistics. *Statistical Science*, 17:97–112, 2002.

[41] R. W. Katz. Stochastic modeling of hurricane damage. *Journal of Applied Meteorology*, 41:754–762, 2002.

[42] R. W. Katz. Techniques for estimating uncertainty in climate change scenarios and impact studies. *Climate Research*, 20:167–185, 2002.

[43] R. W. Katz, M. B. Parlange, and P. Naveau. Statistics of extremes in hydrology. *Advances in Water Resources*, 25:1287–1304, 2002.

[44] R. W. Katz, M. B. Parlange, and C. Tebaldi. Stochastic modeling of the effects of large-scale circulation on daily weather in the southeastern U.S. *Climatic Change*, 60:189–216, 2003.

[45] T. Kittel, P. Thornton, J. A. Royle, and T. N. Chase. Climates of the Rocky Mountains: Historical and future patterns. In J. Baron, editor, *Rocky Mountain Futures: An Ecological Perspective.*, pages 59–82, Covelo, CA, 2002. Island Press.

[46] T. Kupper, J. D. Berset, R. Etter-Holzer, R. Furrer, and J. Tarradellas. Concentrations and specific loads of polycyclic musks in sewage sludge originating from a monitoring network in Switzerland. *Chemosphere*, 54(8):1111–1120, 2003.

[47] J. Lee and R. B. Lund. Trends in extreme United States temperatures. *Journal of the American Statistical Association*, 2003. In review.

[48] T. C. M. Lee and H.-S. Oh. Automatic polynomial wavelet regression. *Statistics and Computing*, 14(4): 337–341, 2004.

[49] T.-H. Li and H.-S. Oh. Wavelet spectrum and its characterization property for random processes. *IEEE Trans. on Information Theory*, 48:2922–2937, 2002.

[50] Z.-Q. Lu. Multivariate local polynomial fitting for Martingale nonlinear regression models. *Annals of the Institute of Statistical Mathematics*, 51(4):691–706, 1999.

[51] Z.-Q. Lu, L. M. Berliner, and C. Snyder. Experimental design for spatial and adaptive observations. In L. M. Berliner, D. Nychka, and T. Hoar, editors, *Studies in the Atmospheric Sciences*, volume 144 of *Lecture Notes in Statistics*, pages 65–78, New York, 2000. Springer-Verlag.

[52] R. Lund, L. Seymour, and K. Kafadar. Temperature trends in the United States. *Environmetrics*, 12(7):673–690, 2001.

[53] R. B. Lund and J. Reeves. Detection of undocumented changepoints: A revision of the two-phase regression model. *Journal of Climate*, 15:2547–2554, 2002.

[54] R. B. Lund, Q. Shao, and I. V. Basawa. Parsimonious periodic time series modeling. *Australian & New Zealand Journal of Statistics*. In review.

[55] T. Matsuo, A. D. Richmond, and D. W. Nychka. Modes of high-latitude electric field variability derived from DE-2 measurements: Empirical Orthogonal Function (EOF) analysis. *Geophysical Research Letters*, 29(7), 2002. 1107, doi:10.1029/2001GL014077.

[56] G. A. Meehl and C. Tebaldi. More intense, more frequent and longer lasting heat waves in the 21st Century. *Science*, 305(5686):994–997, 2004.

[57] G. A. Meehl, C. Tebaldi, and D. W. Nychka. Changes in frost days in simulations of 21st century climate. *Climate Dynamics*, 23(5):495–511, 2004. doi: 10.1007/s00382-004-0442-9, 2004.

[58] G. A. Meehl, W. M. Washington, C. Ammann, J. M. Arblaster, T. M. L. Wigley, and C. Tebaldi. Combinations of natural and anthropogenic forcings in twentieth-century climate. *Journal of Climate*, 17(19):3721–3727, 2004.

[59] W. Meiring. Seasonal variation in stratospheric ozone levels, a functional data analysis study. In L. M. Berliner, D. Nychka, and T. Hoar, editors, *Studies in the Atmospheric Sciences*, volume 144 of *Lecture Notes in Statistics*, pages 79–96, New York, 2000. Springer-Verlag.

[60] W. Meiring. Ozonesonde mid-latitude stratospheric ozone variability, with emphasis on the quasi-biennial oscillation: A functional data analysis approach. *Journal of the American Statistical Association*, 2004. In review.

[61] J. F. B. Mitchell, D. J. Karoly, et al. Detection of climate change and attribution of causes. In J. Houghton et al., editors, *Climate Change 2001: The Scientific Basis.*, pages 695–738, 2001. L. M. Berliner (Contributing Author). Cambridge Press University.

[62] P. Naveau. Almost sure relative stability of the maximum of a stationary sequence. *Advances in Applied Probability*, 35(3):721–736, 2003.

[63] P. Naveau, C. Ammann, H.-S. Oh, and W. Guo. An automatic statistical methodology to extract pulse-like forcing factors in climatic time series: Application to volcanic events. In A. Robock, editor, *Volcanism and the Earth's Atmosphere*, Geophysical Monographs, 139:177–186, 2003. doi:10.1029/139GM11.

[64] P. Naveau, P. Brockwell, and D. W. Nychka. Dependence structure of random wavelet coefficients in terms of cumulants. In A. Cohen, C. Rabut, and L. Schumaker, editors, *Curve and Surface Fitting*, pages 343–350, Nashville, TN, 1999. Vanderbilt University Press. ISBN 0-8265-1357-3.

[65] P. Naveau, R. Furrer, and P. Keckhut. The spatio-temporal influence of the vortex on Arctic total column ozone variability. In J. Mateu, D. Holland, and W. Gonzalez-Manteiga, editors, *The ISI International Conference on Environmental Statistics and Health*, pages 131–139. Universidade de Santiago de Compostela, 2003.

[66] P. Naveau, M. Genton, and X. Shen. The skewed Kalman filter. *Journal of Multivariate Statistics*, 2004. To appear.

[67] P. Naveau, R. W. Katz, and M. Moncrieff. Extremes and climate: an introduction and a case study. *Notes de l'Institut Pierre Simon Laplace*, 11:4–22, 2001. ISSN 1626-8334.

[68] P. Naveau and M. Moncrieff. A statistical formulation of convective mass fluxes. *Quarterly Journal of the Royal Meteorological Society*, 129(592):2217–2233, 2003. Part A.

[69] P. Naveau, M. Moncrieff, J.-I. Yano, and X. Wu. Exploratory statistical analysis of tropical oceanic convection using discrete wavelet transforms. In L. M. Berliner, D. Nychka, and T. Hoar, editors, *Studies in the Atmospheric Sciences*, volume 144 of *Lecture Notes in Statistics*, pages 117–132, New York, 2000. Springer-Verlag.

[70] P. Naveau and H.-S. Oh. Polynomial wavelet regression for images with irregular boundaries. *IEEE Transactions on Image Processing*, 13(6):773–781, 2004.

[71] P. Naveau and D. Paillard. Briefing note on statistical downscaling methodologies for WP3. In *The EC BIOCLIM Project (2000-2003) 5th Euratom Framework Programme*, 2002.

[72] P. Naveau, M. Vrac, M. Genton, A. Chedin, and E. Diday. Two statistical methods for improving the analysis of large climatic data sets: General skewed Kalman filters and distributions of distributions. In *Proceedings of The Fourth European Conference on Geostatistics for Environmental Applications*, Barcelona, 2002. geoENV2002.

[73] D. Nychka. Challenges in understanding the atmosphere. *Journal of the American Statistical Association*, 95:972–975, 2000.

[74] D. W. Nychka. Challenges in understanding the atmosphere. In A. Raftery, M. Tanner, and M. Wells, editors, *Statistics in the 21st Century*, pages 199–206, New York, NY, 2001. Chapman and Hall/CRC.

[75] D. W. Nychka and C. Tebaldi. Comment on "calculation of average, uncertainty range and reliability of regional climate changes from AOGCM simulations via the 'reliability ensemble averaging' (REA) method". *Journal of Climate*, 16(5):883–884, 2002.

[76] D. Nychka, C. K. Wikle, and J. A. Royle. Multiresolution models for nonstationary spatial covariance functions. *Statistical Modelling: An International Journal*, 2:315–331, 2002.

[77] H.-S. Oh, C. Ammann, P. Naveau, D. Nychka, and B. Otto-Bliesner. Multi-resolution time series analysis applied to solar irradiance and climate reconstructions. *Journal of Atmospheric and Solar-Terrestrial Physics*, 65:191–201, 2003.

[78] H.-S. Oh and T.-H. Li. Estimation of global temperature fields from scattered observations by a spherical-wavelet-based spatially adaptive method. *Journal of the Royal Statistical Society: Series B*, 66:221–238, 2004.

[79] H.-S. Oh, P. Naveau, and G. Lee. Polynomial boundary treatment for wavelet regression. *Biometrika*, 88:291–298, 2001.

[80] H.-S. Oh, D. W. Nychka, T. M. Brown, and P. Charbonneau. Period analysis of variable stars by robust smoothing. *Journal of the Royal Statistical Society: Series C*, 2004. (Applied Statistics). In press.

[81] F. Parisi and R. Lund. Seasonality and return periods of landfalling Atlantic hurricanes. *Australian & New Zealand Journal of Statistics*, 42:271–282, 2000.

[82] M. B. Parlange and R. W. Katz. An extended version of the Richardson model for simulating daily weather variables. *Journal of Applied Meteorology*, 39:610–622, 2000.

[83] J. A. Royle. Multivariate spatial models. In L. M. Berliner, D. Nychka, and T. Hoar, editors, *Studies in the Atmospheric Sciences*, volume 144 of *Lecture Notes In Statistics*, pages 23–44, New York, 2000. Springer-Verlag.

[84] S. Sain, S. Jagtap, L. M. Mearns, and D. Nychka. A multivariate spatial model for soil water profiles *Journal of Agricultural Biological and Environmental Statistics*, 2004. In review.

[85] B. D. Santer, T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. W. Nychka, D. E. Parker, and K. E. Taylor. Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *Journal of Geophysical Research, Atmosphere*, 105:7337–7356, 2000.

[86] U. Schneider and J. N. Corcoran (2004). Perfect sampling for Bayesian variable selection in a linear regression model. *Journal of Statistical Planning and Inference*, 126:153–171.

[87] E. E. Small, L. C. Sloan, and D. W. Nychka. Changes in surface air temperature caused by desiccation of the Aral sea. *Journal of Climate*, 14:284–299, 2001.

[88] G. Sneddon. A statistical perspective on data assimilation in numerical models. In L. M. Berliner, D. Nychka, and T. Hoar, editors, *Studies in the Atmospheric Sciences*, volume 144 of *Lecture Notes in Statistics*, pages 7–21, New York, 2000. Springer-Verlag.

[89] S. B. Streett, R. A. Davis, and W. T. M. Dunsmuir. Maximum likelihood estimation for an observation driven model for Poisson counts. *Methodology and Computing in Applied Probability*, 2003. In review.

[90] S. B. Streett, R. A. Davis, and W. T. M. Dunsmuir. Observation driven models for Poisson counts. *Biometrika*, 2003. To appear.

[91] C. Tebaldi and D. Nychka. Invited discussion to "Calibrated probabilistic mesoscale weather field forecasting: the geostatistical output perturbation (GOP) method". *Journal of the American Statistical Association*, 99(467):583–586, 2004.

[92] C. Tebaldi. Looking far back vs. looking around enough: operational weather forecasting by "spatial" composition of recent observations. In *Proceedings of the 16th Conference on Probability and Statistics in the Atmospheric Sciences*, AMS Annual Meeting, Orlando, FL, January 2002.

[93] C. Tebaldi, G. Brantstator, and D. W. Nychka. Looking for nonlinearities in the large scale dynamics of the atmosphere. In *Proceedings of the First SIAM International Conference on Data Mining*, Chicago, IL, April 2001.

[94] C. Tebaldi, L. O. Mearns, D. W. Nychka, and R. L. Smith. Regional Probabilities of Precipitation Change: a Bayesian Analysis of Multimodel Simulations. *Geophysical Research Letters*, In press.

[95] C. Tebaldi, D. W. Nychka, B. Brown, and B. Sharman. Predicting clear-air turbulence. In L. M. Berliner, D. Nychka, and T. Hoar, editors, *Studies in the Atmospheric Sciences*, volume 144 of *Lecture Notes in Statistics*, pages 133–151, New York, 2000. Springer-Verlag.

[96] C. Tebaldi, D. W. Nychka, B. G. Brown, and B. Sharman. Flexible discriminant techniques for forecasting clear-air turbulence. *Environmetrics*, 13(8):859–878, 2002.

[97] C. Tebaldi, R. L. Smith, D. W. Nychka, and L. O. Mearns. Quantifying uncertainty in projections of regional climate change: a Bayesian approach to the analysis of multimodel ensembles. *Journal of Climate*, 2003. Accepted.

[98] C. Tebaldi, M. West, and A. F. Karr. Statistical analyses of freeway traffic flows. *Journal of Forecasting*, 21(2):39–68, 2002.

[99] M. Vrac, E. Diday, P. Naveau, and A. Chedin. Mixture of distributions of distributions. In *Proceedings de la Société Francophone de Classification*, 2001.

[100] B. Whitcher. Wavelet-based estimation for seasonal long-memory processes. *Technometrics*, 46(2):225–238, 2004. In review.

[101] B. Whitcher, S. D. Byers, P. Guttorp, and D. Percival. Testing for homogeneity of variance in time series: Long memory, wavelets and the Nile river. *Water Resources Research*, 38(5), 2002. doi:10.1029/2001WR000509.

[102] B. Whitcher and P. Craigmile. Multivariate spectral analysis using Hilbert wavelet pairs. *International Journal of Wavelets, Multiresolution and Information Processing*, 2(4):567–587, 2004

[103] B. Whitcher, J. B. Weiss, D. W. Nychka, and T. J. Hoar. Stochastic multi-resolution models for turbulence. In M. Akritis and D. Politis, editors, *Recent Advances and Trends in Nonparametric Statistics*, volume forthcoming, 2003.

[104] C. K. Wikle. Hierarchical space-time dynamic models. In L. M. Berliner, D. Nychka, and T. Hoar, editors, *Studies in the Atmospheric Sciences*, volume 144 of *Lecture Notes in Statistics*, pages 45–64, New York, 2000. Springer-Verlag.

[105] C. K. Wikle. Hierarchical models in environmental science. *International Statistical Review*, 71:181–199, 2003.

[106] C. K. Wikle and L. M. Berliner. Combining information across spatial scales. *Technometrics*, 2004. In review.

[107] C. K. Wikle, L. M. Berliner, and R. F. Milliff. Hierarchical bayesian approach to boundary value problems with stochastic boundary conditions. *Monthly Weather Review*, 131:1051–1062, 2003.

[108] C. K. Wikle and N. A. Cressie. Space-time statistical modeling of environmental data. In H. T. Mowrer and R. G. Congalton, editors, *Quantifying Spatial Uncertainty in Natural Resources: Theory and Applications for GIS and Remote Sensing*, pages 213–236, Chelsea, MI, 2000. Ann Arbor Press.

[109] C. K. Wikle, D. Nychka, R. F. Milliff, and L. M. Berliner. Spatio-temporal hierarchical Bayesian blending of tropical ocean surface wind data. *Journal of the American Statistical Society*, 96:382–397, 2001.

[110] C. K. Wikle and J. A. Royle. Spatial statistical modeling in biology. In *Encyclopedia of Life Support Systems*. EOLSS Publishers Co. Ltd., 2002.

GSP manuscripts

[111] E. Bellone, W. Collins, and X. Wu. Statistical models for cloud overlap and amount derived from a cloud-resolving model. 2003. Manuscript.

[112] T. Bengtsson, R. F. Milliff, R. H. Jones, and D. W. Nychka. A state-space model for ocean drifter motions dominated by inertial oscillations. 2003. Manuscript.

[113] B. Das. *Global Covariance Modeling: A Deformation Approach to Anisotropy.* Ph.D. dissertation, University of Washington, Seattle, WA, 2000. Department of Statistics.

[114] R. Furrer, T. Bengtsson, and J. Cavanaugh. Evaluation of forecast and analysis covariance in Kalman filter variants. 2004. Manuscript.

[115] R. Furrer, M. G. Genton, and D. Nychka. Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, 2004. Under Revision

[116] R. Gençay, F. Selçuk, and B. Whitcher. Scale invariance and time-frequency analysis with wavelet hidden markov trees. 2002. Manuscript.

[117] E. Gilleland, D. Nychka, and U. Schneider. Spatial models for the distribution of extremes. In: *Applications of Computational Statistics in the Environmental Sciences: Hierarchical Bayes and MCMC Methods*, ed. J.S. Clark and A. Gelfand. Oxford University Press. To appear.

[118] S. Jain and B. Whitcher. Wavelet analysis of climate-related flood variability. 2003. Manuscript.

[119] T. C. M. Lee and H.-S. Oh. Robust penalized splines. 2003. Manuscript.

[120] P. Naveau, C. Ammann, and H.-S. Oh. A statistical methodology to extract volcanic signals in climatic time series. 2002. Manuscript.

[121] H.-S. Oh and T. C. M. Lee. Hybrid wavelet shrinkage: a method for boundary correction. 2003. Manuscript.

[122] H.-S. Oh, T. C. M. Lee, and D. W. Nychka. Wavelet thresholding for data with outliers. 2003. Manuscript.

[123] H.-S. Oh and D. W. Nychka. A robust selection method of the smoothing parameter for smoothing splines. 2003. Manuscript.

[124] S. B. Streett, R. A. Davis, T. Rydberg, and N. Shephard. The Cbin model for counts: testing for common features in the speed of trading, quote changes, limit and market order arrivals. 2003. Manuscript.

[125] S. B. Streett, D. W. Nychka, and R. A. Davis. Observation driven models for precipitation occurrence. 2003. Manuscript.

[126] C. Tebaldi. Influence of large scale circulation measures on precipitation at local stations in the south-east U.S.. NRCSE Technical Reports Series 55, National Research Center for Statistics and the Environment, 2000.

[127] C. Tebaldi, D. W. Nychka, and L. O. Mearns. An exercise in simple pattern scaling: how regional aggregation impacts uncertainty in the scaled signals. 2003. Manuscript.

[128] B. Whitcher, J. B. Weiss, D. W. Nychka, and T. J. Hoar. A multiresolution census algorithm for calculating vortex statistics. 2003. Manuscript.

[129] B. Whitcher and C. K. Wikle. Spectral analysis of atmospheric time series using Hilbert wavelet pairs. 2002. Manuscript.