Annual Report 2000-2001

National Research Center for Statistics and the Environment

NRCSE

The **NRCSE** was established in 1997 through a cooperative agreement with the United States Environmental Protection Agency which provides the Center's primary funding.

EPA

National Research Center for
Statistics and the Environment

Annual Report 2000-2001
National Research Center
for Statistics and the Environment

1. Summary

The fifth year of operation of NRCSE (Oct. 1, 2000– Sep. 30,2001) has been focusing on proposal writing for continued funding. However, we have also pursued our rich agenda of scientific and outreach activities. During the year we had two workshops/conferences, one on spatial moving averages, organized by Jay Ver Hoef and David Higdon, and a NSF/CBMS Regional Conference on Environmental Statistics, with Richard Smith from the University of North Carolina as the main lecturer.

A serious setback in the communications between the EPA and NRCSE was the departure of Larry Cox, our scientific liaison, from the EPA at the beginning of 2001. We were unable to establish any similar connection inside EPA with another scientist.

The Center responded, in conjunction with Richard Smith and the EPA Northwest PM Center, to the request for public comment on the PM Criteria Document. We found the document seriously flawed in its analysis of the current literature. The special issue of *Environmetrics*, number 6 for the year 2000, produced under the auspices of NRCSE and entirely devoted to statistical aspects of analysis of PM data had, despite efforts both from NRCSE and EPA scientists to make the papers available for the review, been completely overlooked.

On September 10, 2001, NRCSE was informed that the proposal from the University of Chicago had been awarded the next 5-year environmental statistics center award. This ends the EPA funding of NRCSE. A request for a no-cost extension for one year has not yet been acted upon by the EPA. A proposal to NSF for a National Environmetric Institute has proceeded to the second round, and a site visit will take place on October 18–19.

2. Outreach activities

2.1 Seminars

The NRCSE seminars are now nearly always joint with other departments. During the year we had seminars joint with Statistics, Biostatistics, and Atmospheric Sciences. A list of all seminars is in Appendix A.

2.2 Web site

The Center web pages, located at http://www.nrcse.washington.edu, are the main source of information about the Center. During the time period of this report there were 293,606 successful requests for pages from over 15 000 different hosts. On average, about 16 Kb of data were transferred per day. About 30% of the visitors came from .edu domains, 10% from .com, and 80% from .net domains. France, Canada and United Kingdom were the leading non-US domains. About 28% of the hosts could not be identified. The most visited directory was the research directory, where project descriptions and research reports

are kept. The most requested report was TR 60 *Multivariate Receptor Models and Model Uncertainty* by Eun Sug Park, Man-Suk Oh and Peter Guttorp. The 1999-2000 Annual Report and the NRCSE newsletter had about 3000 requests each.

The web site was attacked by foreign hackers, who briefly replaced the welcome page with their own page. Several attempts at denial-of-service attacks failed. There were attempts to infect the server with different computer viruses, but only one (relatively minor) infection was successful.

2.3 Software

During the summer of 2001, several undergraduate students were working on development of user-friendly versions of NRCSE software products associated with nonstationary spatial covariance modeling, multi-objective spatial network design, and visualization. If NRCSE is granted a no-cost extension this work will be completed and published at the NRCSE web site during the 2001-02 academic year.

2.4 Workshops

2.4.1 Spatial moving averages

A workshop on spatial moving averages was organized by Jay Ver Hoef and Dave Higdon and held at the NRCSE from May 20 - 22 2000. Spatial moving average models have surfaced repeatedly in recent years in disparate literatures. They are formed by using a moving average function (or kernel) that operates on an independent spatial process. The goal of the workshop was to bring together authors to share ideas. Talks, followed by discussion, were given by Ron Barry, Nicky Best, Montserrat Fuentes, Dave Higdon, Katja Ickstadt, Doug Nychka, Jean Thiebaux, Jay Ver Hoef, Chris Wikle, Robert Wolpert on topics relating to basic theory, relationships to other spatial methods, estimation methods (classical and Bayesian, large data sets), univariate and multivariate models, and stationary and nonstationary models. The schedule for the workshop can be found in Appendix D1.

2.4.2 NSF/CBMS Regional Conference on Environmental Statistics

The NSF-CBMS regional conference on Environmental Statistics, featuring Richard Smith from North Carolina, took place June 25-29, 2001, at the University of Washington. There were 59 participants. The format had a lecture by Dr. Smith each morning, followed by a guest speaker (Paul Switzer, Stanford University, Jim Zidek, University of British Columbia, Doug Nychka, NCAR Geophysical Statistics Project, and from UW Tilmann Gneiting, Paul Sampson and Peter Guttorp). In the afternoons Dr. Smith gave

his second lecture of the day, followed by a breakout session in which various topics were discussed in a roundtable format.

The conference was extremely well received by the audience. Indeed, some of the participants rated this as the best conference they had ever participated in. Dr. Smith's slides are available on the web at http://www.stat.unc.edu/postcript/rs/envstat/env.html. His lecture notes will be published by the Institute of Mathematical Statistics. The conference program is in Appendix D2.

2.4 Conference presentations

As always, a large number of Center members and graduate students have given presentations and organized sessions at national and international meetings of various scientific organizations. These include the Society for Risk Analysis annual meeting, the Joint Statistical Meetings, and the Environmetrics 2001 conference. At the Joint Statistical Meetings NRCSE organized two sessions: one on network design, and one on receptor modeling. At the Environmetrics 2001 meeting we also organized two sessions: one on work in the IMPACT project, and one on nonstationary spatial covariance modeling. A detailed list of presentations is given in Appendix C.

2.5 Educational activities

As mentioned in previous annual reports, NRCSE supported during the 1998-1999 academic year the development of "The TRUTH about Science," a 5th-8th grade scientific research curriculum available at the NRCSE web site http://www.nrcse.washington.edu/resource/curiculum/truth.asp. During the year this curriculum has been adopted by the Seattle School District, and accepted for publication by NSTA Press. A paper is being prepared for a special issue of *Environmental and Ecological Statistics*,

2.6 Professional service and recognition

As institutional members of the International Environmetric Society (TIES), the Center receives two full memberships, which the executive committee has decided to award to outstanding research assistants. The 2001 awards went to Doris Damian, Biostatistics, and Fadoua Balabdaoui, Statistics.

Samantha Bates was awarded the best student paper award, and the prize for best risk analysis paper at the Environmetrics 2001 conference in Portland, OR in August, 2001.

Peter Guttorp was named a Fellow of the American Statistical Association in Atlanta, GA in August 2001. The award citation read: "For major contributions to the growth of

environmetrics; for research on spatial modeling under nonstationary spatial covariance; for administration of interdisciplinary research groups, especially as Director of the National Research Center for Statistics and the Environment; for service to the profession."

Dennis Lettenmaier was awarded the American Geophysical Union Hydrology Section Award at the Fall Meeting Hydrology Section Reception which was held in December,2000, in San Francisco, CA. This award recognized his outstanding contributions to the science
of hydrology. Dennis has been a key player in the integration of hydrological science with the atmospheric science community on the one hand, and the water resources engineering
community on the other.

2.7 Newsletter

The Center publishes a newsletter about twice a year, with the latest developments, publications, and other items of potential interest to the membership. The newsletters are available at http://www.nrcse.washington.edu/newsletter/newsletter.asp.

3. Research activities

3.1 Internal funding

The internal NRCSE research funds are allocated annually by the Executive Committee after a competitive application process. Criteria for awards include

* Scientific merit.
* Relevance to the Center's agenda.
* Evidence of involvement by EPA scientists.
* Feasibility of project and likelihood of substantial products.
* Results and EPA contacts of previous NRCSE support

3.1.1 Ranked set sampling: costs and applications

Center member: Loveday L. Conquest
EPA collaborator: Barry Nussbaum
W ESTAT collaborator: David Marker
Research assistant: Rebecca Buchanan

The paper, "Incorporating Human Judgment into Ecological Sampling" by Mode, Conquest and Marker had been presented at the Fourth International Chemometrics/Environmetrics Meetings in Las Vegas, Nevada, in September, 2000. This pa-

per has since been revised and accepted for publication in the journal, Environmetrics. QERM graduate student Rebecca Buchanan developed extensions of the balanced design cost models presented in "Ranked Set Sampling for Ecological Research: Accounting for the Total Costs of Sampling" by Mode, Conquest and Marker (Environmetrics 1999). These extensions include considerations for unbalanced designs.

Dr. Conquest was successful in participating in an EPA STAR grant with Dr. Don Stevens of Oregon State University. The UW portion is "Model-assisted Design for Ecological Sampling". The research will be done by QERM graduate student Rebecca Buchanan, UW/NRCSE post-doc J.Y. Courbois, and Dr. Conquest. Designing sampling schemes for sampling river networks must take into account such network processes as correlation running downstream (flow direction) and also upstream (biological processes, such as salmon migration). Using model-assisted designs, we intend to develop sampling strategies that estimate model parameters and, at the same time, address traditional monitoring purposes, tracking biological, chemical, and geological responses through time.

3.1.2 Using evolutionary algorithms to calculate multi-criteria assessment of process models

Center member: David Ford
Outside collaborator: Joel Reynolds, Alaska Fish & Game
Research assistant: Marianne Turley, Rie Komuro

During the past year our work has developed on two fronts: developing practical methods for multi-criteria assessment for two competing models and the testing of the evolutionary computation program developed to calculate the Pareto set.

Ms. Turley has made comparison of two competing models of plant competition using multiple criteria. The models were for one- and two-sided competition where large plants affect small ones but there is no reciprocal influence (one-sided), and where there is reciprocal influence (two-sided). She has shown the importance of how multiple criteria are selected and developed in order to calculate a Pareto Set where different model parameterizations satisfy different groups of criteria. Two types of criteria are important: measures of location for principal output such as mean, median, quartiles; and measures describing important data characteristics such as frequency distributes, and metrics of spatial structure. She was also able to compare models with just a single parameter difference and reject the more complex model when that parameter solved as zero. This work has illustrated that selection of assessment criteria, and deciding upon the range within which a criterion might be considered as satisfied, are as important as model formulation – though both are frequently relegated to an after thought of model development. This has considerable significance for the development and use of environmental models. Ms Turley was awarded a doctorate for her work.

Calculation of the Pareto set, on which multi-criteria assessment depends, requires an efficient evolutionary algorithm that is fast and does find the complete possible set. Ms Komuro has compared our algorithm, Pareto_Evolve, originally developed by Dr. Joel Reynolds, with some other algorithms. Performance for some simple tests was satisfactory (Komuro and Ford 2001; NRCSE TRS 62). However, the standard tests used are for a two criterion problem whereas in multi-criteria model assessment the number is likely to be greater than two. Further, most tests use continuous functions–frequently with well known solutions. Our recent work has concentrated on the segmentation of a known data stream into multiple criteria, each representing different characteristics, and to calculate the effectiveness with which solutions are found as the number of criteria are changed. This work is showing that the choice of criteria must be designed to test particular aspects of model function.

3.1.3 Monitoring network design

Center member: Paul Sampson, Peter Guttorp
EPA collaborator: Dave Holland
Research assistant: San-San Ou
Undergraduate research assistants: Brooke Hoem, Friedrich Kuchling and Lean Richmond

U.S. EPA guidelines for air quality monitoring network design specify four explicit aims

1. to determine highest concentrations expected to occur in the area covered by the network;
2. to determine representative concentrations in areas of high population density;
3. to determine the impact on ambient pollution levels of significant sources or source categories;
4. to determine general background concentration levels.

However, the statistical literature on optimal network design seems far removed from these aims, considering almost exclusively the optimization of a single criterion, such as (some function of) kriging predictive variances. In view of the fact that practical policy decisions require consideration of (at least) these four aims, we initiated a project to develop a methodology for "Pareto optimal" monitoring network design for multiple objective criteria. We argue that an attractive alternative to optimization of a single (possibly composite) design criterion is the identification and consideration of the space of Pareto optimal designs for a set of objective functions. Consideration of this "Pareto frontier" of designs will allow better understanding of the trade-offs necessary to obtain greater relative efficiency with respect to the optimization of a single criterion such as a (possibly weighted) spatial average of kriging variances. We have successfully employed a sequence of three different undergraduates (Brooke Hoem (ACMS, now graduated), Friedrich Kuchling (computer engineering), and Leah Richmond (ACMS)) to adapt for this purpose the "Pareto Evolve" software developed at NRCSE for multi-criteria assess-

ment of ecological process models. Pareto-Evolve uses genetic algorithms (evolutionary computation) to identify candidate parameterizations in the "Pareto Frontier". In this context, each parameterization represents a monitoring network.

Work to date has involved (a) coding of simple geostatistical design criteria such as maximum and average kriging variances, as well as a spatial coverage criterion for use with the Pareto-Evolve software; (b) modification of some details of the evolutionary computation algorithm, and (c) a preliminary demonstration of the successful application of the evolutionary computation algorithm for a toy design problem. This work was the subject of invited presentations by Paul Sampson at the 2001 Joint Statistical Meetings, Aug 5-9, the First Spanish Workshop on Spatio-Temporal Modeling of Environmental Processes in Benicassim, Spain, October 27-30, and a forthcoming EPA Spatial Data Analysis Technical Exchange Workshop in RTP, NC, Dec 3-5. The methodology is discussed in the proceedings publication Sampson et al., 2001 (see section 3.3)

Current research plans include the application of this methodology to practical network (re)design calculations using, first, the example of the CASTNET monitoring network. Modeling of data from this network in preparation for estimation of a spatial covariance model to be used as a basis for spatial estimation criteria was carried out by graduate R.A. San-San Ou.

3.1.4 Statistical adjustment of ozone for meteorological variables

Center members: Mary Lou Thompson, Peter Guttorp, Tilmann Gneiting, Don Percival and Paul Sampson
EPA collaborator: Larry Cox
Outside collaborators: Joel Reynolds, Alaska Fish Game, Hans Wackernagel and Christian Lajaunie, Centre de Géostatistique, France
Research assistants: Fadoua Balabdaoui, Sinjini Mitra

The review paper on meteorological adjustment of ozone (NRCSE TR 26) has appeared in *Atmospheric Environment*.

In conjunction with the French arm of the EU-funded IMPACT project, we carried out extensive analyses of the space-time structure of Paris region (Isle-de-France) ozone concentration monitoring data for one year. After many months delay, long sought meteorological data were provided and some exploratory analyses were made relating ozone to meteorology. There were clear indications of increased ozone on the south-west part of the network in connection with northeasterly winds, indicating a direct link with the transport of the air mass across the metropolitan region.

Interest in the detailed space-time correlation structure led us to investigate the possible application of a new family of spatio-temporal correlation models suggested by Tilmann Gneiting. These models are appropriate when a Lagrangean reference frame is considered for modeling the asymmetric space-time correlations explained (in part) by meteorological systems moving through a region. For hourly ozone concentrations monitored at different sites in the Paris region, we examined temporally lagged cross-correlations as a function of wind speed and direction. In fact, these cross-correlations were (surprisingly) not noticeably temporally asymmetric.

A presentation of the research was made by Peter Guttorp at a special IMPACT session at the Environmetrics 2001 meeting. He also presented the material at the NSF-CBMS Regional Conference on Environmental Statistics at the University of Washington, and at the Fifth Brazilian School of Probability in Ubatuba, Brazil.

3.1.5 Stochastic modeling of precipitation

Center members: Jim Hughes, Peter Guttorp, Paul Sampson
Outside collaborator: Bryson Bates, CSIRO, Australia
Research assistant: Tamre Cardoso

Two main areas of research on this topic involved the development (by Hughes' student Ted Lystig) vastly improved algorithms for fitting hidden Markov models, including algorithms for estimating standard errors (Lystig and Hughes, 2001). In addition, Tamre Cardoso has developed a hierarchical model of precipitation rate, using data from different sources (such as distrometer, rain gauge, and radar reflectivity). The model is currently being fitted to data from northern California.. Hughes gave presentations of the hidden Markov model for precipitation at the Eastern North American Region of the Biometric Society and at the Eighth International Meeting on Statistical Climatology in Germany, while Guttorp presented the research as part of his sequence of talks at the Fifth Brazilian School of Probability.

3.1.6 Statistical aspects of setting and implementing environmental standards

Center members: Peter Guttorp, Paul Sampson, and Mary Lou Thompson
EPA collaborator: Larry Cox
Outside collaborators: Ronit Nirel, Israel, and Bruno Sanso, Venezuela

Work on hypothesis testing approaches to air quality environmental standards continued in the revision of a manuscript now accepted for publication in Environmental and Ecological Statistics (Thompson, Cox and Sampson, 2002, see section 3.3). Current research

plans include the development of explicitly spatial standards (in contrast to current air quality standards that do not address issues of spatial variation). In this context we initiated a project to incorporate the scientific information encoded in deterministic photochemical modeling predictions as prior information in a Bayesian spatial estimation methodology. This project, begun while Ronit Nirel and Bruno Sanso (Universidad Simon Bolivar, Venezuela),were visiting NRCSE in Summer of 1999. It is, however, still in its early stages.

3.1.7 Receptor modeling for air quality data in space and/or time

Center members: Peter Guttorp, Dean Billheimer
Outside collaborators: Ron Henry, USC; Cliff Spiegelman, Texas A&M
Center postdoc: Eun Sug Park

An important problem in environmental science is to identify where pollution comes from given air pollution data. Multivariate receptor modeling aims to achieve this goal by decomposing ambient concentrations of pollutants to components associated with source emissions. This is a difficult problem in its most general form and typically restrictive assumptions are required. One assumption is that the observations are temporally independent, which is inappropriate for most of hourly measurements. We have developed a multivariate receptor model for temporally correlated data, which can incorporate extra sources of variability due to dependence in estimation of model parameters and uncertainty. The work has resulted in a paper, *Multivariate receptor modeling for temporally correlated data by using MCMC*, which will appear in *Journal of the American Statistical Association*, December 2001, Vol. 96, No. 456. An invited session at the Joint Statistical Meetings in Atlanta in August 2001, organized by Peter Guttorp of NRCSE, heard a presentation of this work.

Assumptions on the number of pollution sources and identifiability conditions are the main source of model uncertainty in multivariate receptor models, which is often overlooked. A Bayesian approach based on the marginal likelihood for assessing model uncertainty in multivariate receptor models has been developed. The work has resulted in a paper, *Multivariate Receptor Models and Model Uncertainty*, accepted for publication in *Chemometrics and Intelligent Laboratory Systems.* A different approach uses earlier NRCSE work on spatio-temporal models for compositional data

We are currently focusing on extending receptor models to spatially correlated data obtained from multiple monitoring sites. Two cases, measurements on a single species from multiple monitoring sites, and measurements on multiple species from multiple monitoring sites, are being investigated. The first type of data can be used to locate the major pollution sources by estimating their spatial profiles, while the second type of data is

ideal for characterizing spatial structure of source contributions and errors. The first approach has been applied to an analysis of PM_{10} data for Seoul, Korea, and yielded physically meaningful results, i.e., the resulting estimates for the source spatial profiles seemed to be consistent with our prior expectation about the PM_{10} sources in Seoul. The paper from this research, *Multivariate receptor modeling for air quality data in space and/or time*, was invited to be presented at International Statistical Institute meeting held in August, 2001, Seoul, Korea. The paper will be part of a special issue of *Environmental and Ecological Statistics.* The second approach is the topic of NRCSE Technical Report 71, and uses nonparametric regression on wind direction to infer the source of PM air pollution from data at two locations. This work was also presented at the invited session on Statistical analysis of multivariate air quality data at the Joint Statistical Meetings in Atlanta.

3.1.8 Trend estimation using wavelets

Center members: Don Percival, Peter Guttorp
Research assistant: Peter Craigmile

.

In the last year Peter Craigmile has been continuing to investigate wavelet and spectral based methods for time series analysis.

He finished his Ph.D. dissertation in December 2000. The title was ``Wavelet--based estimation of trend contaminated long memory processes", with Donald Percival and Peter Guttorp. His thesis focuses on a topic in time series analysis, namely estimating a trend component (large scale variations) in the presence of long memory (LM) errors (slowly decaying autocorrelations). Peter has also investigated wavelet-based approximate maximum likelihood estimators for fractionally differenced processes, and established the validity of an exact method for simulating these - and related - processes. The work is a mix of theoretical, methodological and applied statistics (e.g. analyzing Northern Hemisphere temperatures since the mid 1800s).

Since December 2000, he has continued to investigate some of the issues raised in his dissertation, namely parameter estimation of LM processes and trend estimation. He also worked with Abdullah Almasri (a visitor to NRCSE from Sweden) and Donald Percival on tests for periodicity in the presence of LM errors.

In addition, using AR process approximations to the correlation structure of wavelet decompositions of LM processes, he has been investigating some approximate simulation schemes for these processes.

3.1.9 Bayesian methods for assessment of environmental fate and transport models

Center members: Alison C. Cullen, Adrian Raftery
Research assistant: Samantha Bates

Bates' doctoral thesis, titled "Bayesian Inference for Deterministic Simulation Models for Environmental Assessment," was successfully defended. The major component of this work was the development and application of Bayesian methodology for making inference from sequential multicompartment deterministic models, particularly those in environmental assessment, while accounting for uncertainty in the model inputs. In August of 2001, A talk on this aspect of my research was given at the annual meeting of The International Environmetrics Society. It received awards for best student paper (joint) and best risk analysis paper. The paper (Bates, Raftery and Cullen, 2001) has been submitted to *Environmetric*s.

A paper on tools to assess deterministic models in the Bayesian framework is in preparation and follows on from the thesis work. A paper (Bates and Raftery, 2001) has been submitted to the *Journal of Computational and Graphical Statistic*s, presenting a Markov chain Monte Carlo method for sampling distributions, which are ridgelike in high dimensions. Posterior distributions of inputs and outputs to deterministic models may display this behavior.

3.1.10 ORCA: A visualization toolkit for high-dimensional data

Center member: Thomas Lumley
Postdoc: Pip Courbois
Other investigators: Dianne Cook, Nicholas Lewin-Koh and Zach Cox (Iowa State), Peter Sutherland (Neomorphic, Inc.), and Tony Rossini (UW)
Undergraduate assistants: Renata van Dienst,, Zach Frazier

The ORCA development team is implementing a variety of visualization tools. The web page for the project is http://software.biostat.washington.edu/orca.

3.1.11 Bayesian estimation of nonstationary spatial covariance structure

Center members: Paul D. Sampson, Peter Guttorp
Research assistant: Doris Damian

Doris Damian is in the final stages of completing her Ph.D. thesis on a Bayesian approach to modeling and estimation of the spatial correlation structure of spatio-temporal environmental monitoring data using the spatial deformation model of Sampson and Guttorp. The first publication on this work appeared in Environmetrics (Damian et al., 2001, re-

ported in the previous annual report). Since that publication, Doris has revised the approach to parametrizing the thin-plate splines used for the spatial deformation in order to specify prior probability models on both the affine and non-affine components of the spatial deformation. (Previous "bending-energy" priors addressed only the extent of non-affine variation in the spatial deformation while being flat over the space of affine transformations.) In addition, she extended the modeling so as to accommodate the (temporal) variance of the space-time process varying spatially according to a nonstationary pattern according with the same spatial deformation assumed to underlie the spatial correlations.

This project also employed an undergraduate major from the U.W. program in Applied, Computational and Mathematical Sciences, Gabriel Johnson, to port the code for model estimation using McMC algorithms from her Unix version to a version running under a Windows PC operating environment. In addition to the porting of the computational algorithms, with substitution and testing of mathematical support libraries as necessary, Gabriel implemented a user interface that will greatly benefit our intended release of the software in early 2002.

Invited presentations on this work were given by Paul Sampson at the geoENV 2000 Meeting in Avignon, France, Nov 22-24, at the NSF/CBMS Environmental Statistics lecture series sponsored by NRCSE here in Seattle, June 25-29, and at the First Spanish Workshop on Spatio-Temporal Modeling of Environmental Processes in Benicassim, Spain, October 27-30. Peter Guttorp gave an invited presentation on this subject at the 2001 annual meeting The International Environmetric Society (TIES) in Portland, Oregon, Aug 13-17. Publications based in part on this work include Sampson et al. (2001), Sampson, Damian and Guttorp (2001) and Sampson (2001) (see section 3.3).

3.1.12 Fast and exact simulation of fractional Brownian motion

Center member: Tilmann Gneiting
Research assistant: Peter Craigmile
Outside collaborator: Martin Schlather

Long-memory dependence plays crucial roles in the assessment of environmental concerns such as global warming. In this context, fast and exact simulations of long-memory processes are desirable. The best technique presently available is the Davies-Harte algorithm. Craigmile (2000) validates this algorithm for broad classes of long-memory processes. Schlather (2001) made software publicly available; Gneiting has been a consultant on this project, which is still being developed. Gneiting and Schlather (2001) develop new classes of long-memory processes, simulate from these processes, and suggest new statistical tools for their analysis. Theoretical background material motivated by this project is discussed in Gneiting, Sasvári and Schlather (2001).

3.1.13 Decision-Making Under Uncertainty: Prioritizing Freshwater Habitat Restoration for Salmon Recover in the Columbia River Basin

Center member: Ray Hilborn
Research assistant: Jody Brauner

This research focused on data collection and model development to better understand the linkages between riparian management/restoration and habitat carrying capacity for salmonids. Regional data were collected on riparian management regulations, stream surveys for LWD, channel morphology, and pool characteristics, as well as age-specific salmonid habitat preferences (coho and steelhead). An existing wood recruitment model (Riparian Aquatic Interaction Simulator) was linked to a forest growth and yield model (Organon) to generate a matrix of wood loading in streams (pieces/m) as a function of channel width, riparian management practices, stand age and density. The resulting wood recruitment profiles were subsequently used as input to a model of pool formation and habitat carrying capacity for salmonids. Parameters in the pool formation model were estimated using standard linear regression techniques and then compared to parameter estimates based on a posterior probability distribution. The purpose of this comparison was to illustrate the effects of incorporating estimation uncertainty on the distribution of consequences under different riparian management scenarios. Ongoing work is focused on the incorporation of three additional types of uncertainty in the wood recruitment and pool formation models - process, observation and model uncertainty.

3.1.14 Is there a contradiction between apparent long-term increases in the frequency of extreme precipitation over the coterminous U.S. and the absence of flood trends?

Center member: Dennis Lettenmaier
Center researcher: Caren Marzban, National Severe Storms Laboratory

Among the potential consequences of climate change to society, implications for the availability of water in inhabited areas are among the most often quoted. A particular concern voiced recently in many scientific for has been the possibility that acceleration of the global hydrological cycle that is expected to accompany ongoing increases in greenhouse gases might lead to increases in hydrologic extremes, including floods. This possibility is given prominence, for instance, in the recent Third Assessment Report of the Intergovernmental Panel on Climate Change. On the other hand, published studies in the hydrologic literature that have attempted to determine whether changes in flood frequency have occurred over the U.S. show varying results, from essentially no evidence of changes in one study to a conclusion of demonstrable links between increases in precipitation intensity and flood frequency in another.

To address this question in more detail, we have assembled a set of approximately 500 river basins defined by U.S. Geological Survey stream gauges with at least 50 years of observations and minimal effects of water management. For each of these river basins, we have located similarly lengthy precipitation observation records within or close to the basins. We are attempting to answer questions like 1) for appropriately defined time series of moderate to large floods (defined as having recurrence intervals of roughly 1/3 yr^{-1}; somewhat less than used in previous studies) is there evidence of trends at more stations across the continental U.S. that would be expected by chance, and b) for those river basins for which there are statistically significant trends in floods, are there identifiable increases in daily precipitation with similar return periods. In addition to use of the "peaks over threshold" approach to identify the streamflow and precipitation time series, we have utilized various methods of attempting to assure that the precipitation and streamflow events are causally connected. This is accomplished by examining trends in the subclass of conditioned large precipitation and streamflow events. That is, we examine the conditional probability of a flood, given that a large precipitation event precedes it. Although the analysis is ongoing, preliminary results suggest an absence of evidence that trends in precipitation extremes are accompanied by trends in the accompanying “causally related” streamflow. Examination of possible reasons for this apparently anomalous result is currently ongoing. Among the possible explanations are lack of power (due both to small sample sizes and high natural variability) to detect modest trends in the available data, and the predominance of trends in precipitation at times of year (e.g., summer) when relatively dry antecedent conditions dictate that relatively few extreme precipitation events occur.

3.1.15 Modeling multiple pollutants at multiple sites, with application to acute respiratory studies

Center member: Jon Wakefield
Other collaborator: Gavin Shaddick, Imperial College, UK

In cooperation with South East Institute of Public Health in the U.K., a multivariate Gaussian model was developed to model multiple pollutants measured at a number of sites over time. This model was applied to four pollutants measured at eight sites in London. We found very little spatial variability in the pollutants; the temporal variability dominated. This lead to the paper "Modelling Daily Multivariate Pollutant Data at Multiple Sites" by Shaddick, G. and Wakefield, J. (NRCSE TRS 70) that has been accepted for publication in Applied Statistics.

3.1.16 Evaluating the Benefits of an Ecological Study

Center member: Jon Wakefield

Jon Wakefield has been working on a framework for ecological studies and in particular to aid in determining the benefits of a specific study. Ecological bias is discussed with respect to confounding, both within and between areas, and within-area variability in risk. It is argued that more energy should be placed on such issues, rather than refining models for spatial dependence. The paper "A critique of ecological studies" (NRCSE TRS 72) is awaiting submission

3.1.17 Applications of Bartolucci's theorem

Center member: Julian Besag

NRCSE funding has enabled Julian Besag to establish a fruitful collaboration with Francesco Bartolucci at the University of Perugia, Italy, and this contributed to the latter's very recent tenure promotion. One development has been perfect block Gibbs sampling for synergistic autologistic models and this is being incorporated into our Biometrika submission, which has already been favorably reviewed. Further applications are in progress.

3.1.18 Particulate matter field study in Slovakia

Center member: Alison Cullen
EPA collaborator: John Vandenberg, Sarita Noyt
Other collaborators: Michael Brauer, UBC, Canada; Eleanora Fabianova, Eva Mikhalikova, Kvetoslava Koppova, Marek Drimal, Silvia Vojtekova, Frantiska Hruba, SUHE, Slovakia.

A series of projects in a Slovak-US collaboration have focused on the link between PM air pollution and child hospital admissions in Central Europe. These have involved application of GIS technologies to combine individual demographic, health and residential location data with PM levels and PM source information in an attempt to tease out the strength of the relationship between pollution exposure and health effects in children.

In April a group of collaborators met in Boulder, CO, to develop further plans for the project. The main emphasis will be to develop and/or implement tools for air pollution modeling in the very complex terrain in the Banska Bystrica region.

3.1.19 Composite sampling

Center member: Gerald van Belle
WESTAT collaborator: David Marker

3.1.20 Temporal information in biomarker based exposure inference

Center researchers: Rafael Ponce and Elaine Faustman
EPA collaborator: Anne Jarabek
UW collaborator: W. C. Griffith
Research assistant: Scott Bartell

Biomarker measurements from single time points are often used to make inferences about longer periods of toxicant intake. However, toxicant exposures rarely, if ever, occur under steady-state conditions, and biomarkers are typically most sensitive to recent toxicant exposures. Moreover, toxicant exposures are often episodic and vary in magnitude over time. The initial work involved investigation of the potential magnitude of the error introduced by fallacious steady-state assumptions when using biomarkers to estimate average rates of exposure. Examples were developed using measurements of blood and hair mercury content in human adults. Current efforts are focused on statistical estimation of key exposure parameters using non-steady state models.

3.2 External funding

NRCSE has considerable funds available for researchers at other institutions. These are used in two different categories: funds for visiting researchers, who spend substantial time at the Center, and funds for subcontracts, where the bulk of the work is done at the researcher's home institution.

3.2.1 Visiting researchers

Name	Arrival	Departure	Affiliation	NRCSE collaborator
Bjorkestol, Kirsten	000821	010630	Agder University College, Norway	Visiting student
Marzban, Caren	000901	011201	Natl. Severe Storms Lab	Peter Guttorp
Murtaugh, Paul	000901	010715	Oregon State	Sabbatical leave
Fuentes, Monteserrat	001012	001021	North Carolina State	Guttorp, Sampson
Shaddick, Gavin	001101	001201	Imperial College School of Medicine	Wakefield
Almasri, Abdullah	010101	010615	Lund University	Percival
Linder, Ernst	010215	010615	University of New Hampshire	Guttorp, Sampson
Gelfand, Alan	010304	010324	University of Connecticut	Guttorp, Sampson, Raftery
Golinelli,Daniela	010705	010801	USC	Guttorp

3.2.2 Particulate matter air pollution

Richard Smith from University of North Carolina, jointly with Peter Guttorp and Lianne Sheppard from NRCSE, provided an extensive comment on the draft PM Criteria Document which was made available in March. The comment, together with a public comment

produced by researchers at the EPA NW PM Center and an opinion piece by Guttorp and Smith, can be found in NRCSE Technical Report Series no.66.

3.2.3 USC

A subcontract with the University of Southern California has enabled Ron Henry and a graduate student to pursue research (with Cliff Spiegelman, Texas A&M) on source location estimation from multi-receptor networks of PM monitors (see section 3.1.7), resulting in NRCSE Technical Report Series no.71.

3.2.4 Slovak Health Institute

3.2.5 IMPACT

In cooperation with the European Council NRCSE has been participating in the IMPACT project. Research has focused on covariance estimation in the presence of covariates, and an invited session at the Environmetrics 2001 meeting in Portland, OR (Appendix C), was devoted to IMPACT work.

3.3 Submitted and published research papers

Bartell, S., Griffith, W.C. and. Faustman, E.M (2001) Temporal fallacy in biomarker based average exposure inference. Submitted to *Journal of Exposure Analysis and Environmental Epidemiology.*

Bartolucci, F. and Besag, J.E. (2002). A recursive algorithm for Markov random fields. Tentatively accepted by *Biometrika.*

Bates, S.C., Cullen, A.C. & Raftery, A.E. (2001) Bayesian Uncertainty Assessment in Multicompartment Deterministic Simulation Models for Environmental Risk Assessment. Submitted to *Environmetrics*.

Bates, S.C. & Raftery, A.E. (2001) An Efficient Markov Chain Monte Carlo Proposal Distribution for Ridgelike Target Distributions Using Nearest Neighbors. Submitted to *Journal of Computational and Graphical Statistics*.

Besag, J.E. (2001) Invited discussion of ``Conditionally specified distributions'', by Arnold,
Castillo and Sarabia. *Statistical Science*, in press.

Besag, J.E. (2002) Likelihood analysis of binary data in space and time. Volume edited by P.J. Green, N. Hjort and S. Richardson. In press.

Craigmile, P. F. (2001) Simulating a class of stationary Gaussian processes using the Davies-Harte algorithm, with application to long memory processes. Submitted to *Journal of Time Series Analysis*.

Gertler, N. and. Cullen, A (2000) Effects of a Transient Cancer Scare on Property Values: Implications for Risk Valuation and the Value of Life. *Human and Ecological Risk Assessment* **6** (5): 731–745

Doberstein, C. P,, Karr, J. R,., Conquest, L. L. (2000) The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. *Freshwater Biology* **44**:1–17.

Gneiting, T.,. Sasvári, Z. and Schlather, M. (2000) Analogies and correspondences between variograms and covariance functions. *Advances in Applied Probability* **33**: 617–-630.

Gneiting, T. and Schlather, M (2001) Stochastic models which separate fractal dimension and Hurst effect", submitted to *SIAM Review.*

Hruba F., Fabianova E., Koppova K., Vandenberg J. (2001) Childhood respiratory symptoms, hospital admissions and long-term exposure to particulate matter. *Journal of Exposure Analysis and Environmental Epidemiology*; **11**:33–40.

Jankowski, P. and Nyerges, T. (2001) *Geographic Information Systems for Group Decision Making* London: Taylor & Francis.

Levy, D., Lumley, T., Sheppard, L., Kaufman, J., Checkoway, H. (2001) Referent selection in case-crossover analyses of health effects of air pollution. *Epidemiology* **12**.

Levy, D. , Sheppard, L., Checkoway, H., Kaufman, J., Lumley, T. , Koenig, J. and Siscovick, D. (2001) A case-crossover analysis of particulate matter air pollution and out-of-hospital primary cardiac arrest. *Epidemiology*, **12**:193-199.

Lystig, T. C, Hughes, J. P. (2001) Baum's algorithm extended: Likelihood calculations for hidden Markov models. *Journal of Computational and Graphical Statisics*, in press.

Mode, N. A., Conquest, L. L., Marker, D. A (2000) Incorporating prior knowledge in environmental sampling: ranked set sampling and other double sampling procedures. Submitted to *Environmetrics*.

Percival, D. B., Overland, J E. and Mofjeld, H. O. (2001), Interpretation of North Pacific Variability as a Short and Long Memory Process, *Journal of Climate*, in press.

Sampson, P.D., Damian, D., and Guttorp, P. (2001). Advances in Modeling and Inference for Environmental Processes with Nonstationary Spatial Covariance. In: *GeoENV 2000: Geostatistics for Environmental Applications*, P. Monestiez, D. Allard, R. Froidevaux, eds., Dordrecht: Kluwer, pp. 17-32.

Sampson, P.D., Damian, D., Guttorp, P., and Holland, D.M. (2001). Deformation-based nonstationary spatial covariance modelling and network design. In: *Spatio-Temporal Modelling of Environmental Processes*, Colecció «Treballs D'Informàtica I Tecnologia», Núm. 10., J. Mateu and F. Montes, eds., Castellon, Spain: Universitat Jaume I, pp. 125-132.

Schlather, M. (2001) Random Fields: Simulation and Analysis of Random Fields. Package on random field simulation for R. Posted at http://cran.r-project.or.

Shaddick, G. and Wakefield, J (2002) Modelling Daily Multivariate Pollutant Data at Multiple Sites. *Applied Statistic*s, in press..

Thompson, M.L., Cox, L.H., and Sampson, P.D. (2002) Statistical Hypothesis Testing Formulations for U.S. Environmental Regulatory Standards for Ozone. *Environmental and Ecological Statistics*, in press.

van Belle, G., Griffith, W.C. and Edland, S.D. (2001) Contributions to composite sampling. *Environmental and Ecological Statistics*, **8**:171-180.

Widmann, M. M., Bretherton, C. S. and Salathé Jr., E. P. (2001) Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. Submitted to *J. Climate.*

3.4 PhD students graduated

Peter Craigmile, Statistics.
Co-supervisors: Don Percival, Applied Physics Laboratory, and Peter Guttorp, Statistics.
Dissertation title: *Wavelet-based estimation for trend contaminated long memory processe*s.
Current position: Ohio State, Columbus, OH.

Samantha Bates, Statistics.
Supervisor: Adrian Raftery, Statistics.
Dissertation title: *Bayesian Inference for Deterministic Simulation Models for Environmental Assessment.*

Current position: Virginia Tech, Blacksburg, VA.

4. Administration

4.1 Executive and advisory committees

4.1.1 Executive committee

One of the main issues that the executive committee dealt with during the year was the strategy for continued funding of NRCSE. All executive committee meetings are announced with their agendas on the executive committee web page http://www.nrcse.washington.edu/people/execcom.asp.

4.1.2 Advisory committee

The advisory committee did not meet during the report period.

4.2 Members

Two new members were proposed by the membership and elected by the executive committee during the year. They are Mark Handcock, Statistics and Sociology, and a core faculty member of the recent Center for Statistics in the Social Sciences, and Julian Besag, Statistics.

4.3 Postdocs

Eun Sug Park completed her second year at the Center, and left for a position at the Texas Transportation Institute in College Station, TX. The VIGRE (an NSF program aiming at vertical integration of teaching and research) postdoc Jean-Yves Courbois, continued his active participation with NRCSE, particularly working with Ashley Steel, NMFS/NOAA, Loveday Conquest, and Peter Guttorp.

5. A reflection on the past

In considering the five years of EPA-sponsored research at NRCSE, there are many things to be proud of. The Center has produced a large quantity of research published in peer-reviewed journals. Several monographs have been produced with NRCSE support. Our sequence of workshops has highlighted important research areas in environmetrics, and brought together leading researchers from a variety of fields and geographic regions. Our outreach program has improved the level of scientific education in Seattle middle schools and elsewhere by producing the curriculum product "The Truth About Science." We have had a variety of visitors, both short-and long-term, who have interacted fruitfully with Center members and graduate students. Several Masters and Ph.D. theses of high quality have been submitted by Center research assistants. We have built contacts

with EPA researchers and laboratories, and our work has increased substantially both the activity in and the visibility of the field of environmetrics. We wish the University of Chicago group the best of luck in continuing this endeavor.

Appendix A. Seminars

The NRCSE seminar series during the academic year 2000-01 was maintained as sponsored lectures in other campus units. The following is a list of seminar presentations during the academic year.

Autumn 2000

Friday, October 20, 2000.
Montserrat Fuentes, Statistics Department at NCSU and US EPA
"Spatial Modeling and Prediction of Nonstationary Environmental Processes"

Monday, October 30, 2000. Joint with Statistics and Atmospheric Sciences.
Caren Marzban, NRCSE and the National Severe Storms Lab, NOAA and Department of Physics, University of Oklahoma.
"On the Correlation Between U.S. Tornadoes and Pacific Sea Surface Temperatures"

Thursday, November 30, 2000. Joint with Biostatistics and Statistics.
Jon Wakefield, NRCSE
"Another Solution to the Ecological Inference Problem"

Winter 2001

Monday, January 8, 2001. Joint with Biostatistics and Statistics.
Paul Murtaugh, Oregon State University
"Before-After-Control-Impact Analysis in Ecology"

Spring 2001

Monday, April 30 2001. Joint with Statistics.
Ernst Linder, University of New Hampshire
"Estimating local trends in large environmental spatial temporal databases"

Summer 2001

Tuesday, July 3, 2001. Joint with Statistics.

Gopalan Nair, Curtin University of Technology and University of California, Santa Barbara

"Directed Markov Point Processes"

Appendix B. Technical reports 2000-01

TRS number 71
Locating Nearby Sources of Air Pollution by Nonparametric Regression of Atmospheric Concentrations on Wind Direction
Ronald C. Henry, Yu-Shuo Chang and Clifford H. Spiegelman

TRS number 70
Modelling Daily Multivariate Pollutant Data at Multiple Sites
Gavin Shaddick and Jon Wakefield

TRS number 69
Stochastic models which separate fractal dimension and Hurst effect
Tilmann Gneiting and Martin Schlather

TRS number 68
Journal Quality, Effect Size and Publication Bias in Meta-analysis
Paul Murtaugh

TRS number 67
On Rejection Rates of Paired Intervention Analysis
Paul Murtaugh

TRS number 66
Comments on the Criteria Document for Particulate Matter Air Pollution
Richard Smith, Peter Guttorp, Lianne Sheppard, Thomas Lumley and Naomi Ishikawa

TRS number 65
Interpretation of North Pacific Variability as a Short and Long Memory Process
Donald B. Percival, James E. Overland and Harold O. Mofjeld

TRS number 64
A Markov Chain Model of Tornadic Activity
Caren Marzban and Peter Guttorp

TRS number 63
Nonseparable, Stationary Covariance Functions for Space-Time Data
Tilmann Gneiting

TRS number 62
Application of POMAC to the Multiobjective 0/1 Knapsack Problem

Rie Komuro and E. David Ford

TRS number 61
Advances in Modeling and Inference for Environmental Processes with Nonstationary Spatial Covariance
Paul Sampson, Doris Damian and Peter Guttorp

TRS number 60
Multivariate Receptor Models and Model Uncertainty
Eun Sug Park, Man-Suk Oh and Peter Guttorp

TRS number 59
Statistical Hypothesis Testing Formulations for U.S. Environmental Regulatory Standards for Ozone
Mary Lou Thompson, Lawrence H. Cox, Paul D. Sampson and David C. Caccia

TRS number 58
Bayesian Uncertainty Assessment in Deterministic Models for Environmental Risk Assessment
Samantha Bates, Adrian E. Raftery and Alison Cullen

TRS number 57
Simulating a Class of Stationary Gaussian Processes Using the Davies-Harte Algorithm, with Application to Long Memory Processes
Peter F. Craigmile

TRS number 56
Analogies and Correspondences Between Variograms and Covariance Functions
Tilmann Gneiting, Zoltán Sasvári and Martin Schlather
(2000)

Appendix C. Conference presentations

December 2000. T. Gneiting: Positive definite functions: basic facts, applications, and challenges. Universität Tübingen, Germany,

January 2001. T. Gneiting: Criteria of Pólya type for positive definite functions, with applications in analysis, numerical analysis, and statistics. Universität Erlangen, Germany)

March, 2001 J. P.Hughes: Weather simulation methods. Plenary talk at 8th International Meeting on Statistical Climatology, Lüneburg, Germany,

March 2001. J. P. Hughes: Hierarchical models for studying climate variability and climate change in SW Australia. ENAR Annual Meeting, Charlotte, SC,

May 2001. T. Gneiting: Nonseparable covariance models for space-time data. Technical University of Vienna, Austria.

June 2001. T. Gneiting: Nonseparable covariance models for space-time data. GSF Research Center for Environment and Health, Munich, Germany.

June 2001. P. D. Sampson and T. Gneiting: Issues in geostatistical space-time modelling. NSF-CBMS Regional Conference on Environmental Statistics, University of Washington, Seattle.

June, 2001. P. Guttorp: Meteorological adjustment of air pollution data. NSF-CBMS Regional Conference on Environmental Statistics, University of Washington, Seattle.

July 2001. S. M. Bartell, W. C. Griffith, R. A. Ponce, and E. M. Faustman. Temporal fallacy in biomarker based exposure inference. Poster, Environmental Protection Agency STAR Fellowship Conference, Silver Spring, Maryland,

July 2001. T. Gneiting: Correlation models in spatial statistics and positive definite functions. SIAM Annual Meeting, Minisymposium on Spatial Statistics, San Diego, CA.

July 2001. T. Gneiting: Nonseparable, stationary covariance functions and space-time geometry. French Mathematical Research Institute, Luminy, France.

July-August 2001. P. Guttorp: Six lectures on Inference for Stochastic Processes in Environmental Science. Fifth Brazilian School in Probability, Ubatuba, Brazil.

August 2001. Besag, J.E. and Higdon, D.M.: Bayesian analysis of agricultural field experiments. Joint Statistical Meetings, Atlanta, GA.

August 2001. M. Handcock: A Two-part Model for Semicontinuous Spatial Data. Joint Statistical Meetings, Atlanta, GA.

August 2001. T. Lumley: Window Subsampling for Spatially Correlated Censored Data. Joint Statistical Meetings, Atlanta, GA.

August 2001. E. S. Park. Multivariate receptor modeling for temporally correlated data by using MCMC, Joint Statistical Meetings, Atlanta, GA.

August 2001. P. D. Sampson,: Air Quality Monitoring Network Design Using Pareto Optimality Methods for Multiple Objective Criteria. Joint Statistical Meetings, Atlanta, GA.
Eun Sug Park: Multivariate Receptor Modeling for Temporally Correlated Data by Using MCMC. Joint Statistical Meetings, Atlanta, GA.

August 2001. S. Bates: Bayesian Inference for Deterministic Simulation Models for Environmental Assessment. Environmetrics 2001, Portland OR. This received awards for best student paper (joint) and best risk analysis paper.

August 2001. L. Conquest: Ranked Set Sampling and Other Double Sampling Procedures: Incorporating Judgement into Ecological Sampling; Environmetrics 2001, Portland OR.

August 2001. P. Guttorp: Meteorological Adjustment of Air Pollution Data. Environmetrics 2001, Portland OR.

August 2001. P. Guttorp: Bayesian Estimation of Non-stationary Spatial Processes Using the Sampson-Guttorp Deformation Approach. Environmetrics 2001, Portland OR.

August 2001. A. Steel: Applications of Ratios in Monitoring Salmonid Populations: The Problem with Random Denominators. Environmetrics 2001, Portland OR.

August, 2001; P. Craigmile: Wavelet-Based Maximum Likelihood Estimation for Trend Contaminated Long Memory Processes, Recent Developments in Time Series section, European Meeting Of Statisticians 2001, Funchal , Madeira,

August 2001. S. M. Bartell, W. C. Griffith, R. A. Ponce, and E. M. Faustman. Temporal fallacy in biomarker based exposure inference. Poster, Third Annual UC Davis Conference for Environmental Health Scientists, Napa, California

August, 2001. E. S. Park: Multivariate receptor modeling for air quality data in space and/or time, International Statistical Institute meeting, Seoul, Korea

September, 2001. Drimal M., Hruba F., Koppova K.: Analyses of relationship between air pollution and health with use of GIS. Seminar “Air Pollution and Health”, Belusske Slatiny.

Appendix D. Workshop agendas

D.1 Spatial moving averages

Sunday, May 20, 2001

9:15-9:30	Registration
9:30 – 10:00	Welcome/Overview
10:00 – 11:45	Jean Thiebaux and Discussion
11:45 – 1:00	Lunch
1:00 – 2:45	Ron Barry and Discussion
2:45 – 3:00	Break
3:00 – 4:45	Jay Ver Hoef and Discussion

Monday, May 21, 2001

8:00 – 9:45	Dave Higdon and Discussion
9:45 – 10:00	Break
10:00 – 11:45	Montserrat Fuentes and Discussion
11:45 – 1:00	Lunch
1:00 – 2:45	Doug Nychka and Discussion
2:45 – 3:00	Break
3:00 – 4:45	Chris Wikle and Discussion

Tuesday, May 22, 2001

8:00 – 9:45	Robert Wolpert and Discussion
9:45 – 10:00	Break
10:00 – 11:45	Katja Ickstadt & Nicky Best and Discussion

11:45 – 12:30 Final Discussion and Wrap-up

D.2 NSF-CBMS Regional Conference on Environmental Statistics

Monday, June 25, 2001

8:30 am—9:00	**Continental breakfast**	
9:00 – 10:00	**Richard Smith,** Lecture #1	**Introduction:** Motivated by the question "Is global warming really happening?", I introduce the three principal methodological themes of the series - spatial statistics, time series analysis and extreme values - in the context of climatological time series and some simple questions about the nature of trends.
10:00-10:15	**Discussion**	
10:15-10:45	**Break**	
10:45 – 11:45	**Paul Switzer**	**Air Pollution Epidemiology Using Daily Time Series:** Recent studies have tried to relate daily variations in air Pollution monitoring data to daily variations in mortality, using data from a number of U.S. cities. The goal is to estimate the effect on longevity of putative changes in pollutant levels. Because the relative pollution effects are very small, the modeling of the data plays a critical role in the analysis. The principal tool is a Poisson regression with a mean function that varies daily with pollutant concentrations and important confounding weather variables. Challenging inferential problems arise because of variable selection, linearity and additivity assumptions, measurement error, and seasonality. Pollution effects estimated for different cities show variations that are geographically modeled to account for demographic differences. This lecture will discuss strengths and weaknesses of published reports as well as directions for further research.
11:45 – 12:00	**Discussion**	
12:00 – 1:30	**Lunch**	
1:30 p.m. – 2:30	**Richard Smith,** Lecture #2	**Geostatistical methods I:** Classical methods of spatial statistics (a.k.a. geostatistics) using stationary, isotropic models for spatial processes. Definitions: stationary and intrinsically stationary processes, the variogram, simple parametric models. Estimation of the variogram, and methods of fitting parametric models: Cressie's WLS method, maximum likelihood, REML, Bayesian methods.
2:30 – 2:45	**Discussion**	
2:45 – 3:15	**Break**	
3:15 – 4:30	**Roundtable discus-sions**	

Tuesday, June 26, 2001

8:30 am—9:00	**Continental breakfast**	
9:00 – 10:00	**Richard Smith,** Lecture #3	**Geostatistical methods II:** Spatial prediction and interpolation (kriging). Derivation of the basic equations: allowing for parameter uncertainty: extensions, e.g. reconstructing a surface from observations with measurement error. Applications to atmospheric pollution and meteorology.
10:00-10:15	**Discussion**	
10:15-10:45	**Break**	
10:45 – 11:45	**Jim Zidek Professor, Head of Statistics, University of British Columbia**	**Mapping Urban Pollution Fields From Ambient Monitoring Data:** Some of the problems encountered by my co-investigators (in particular, Nhu D Le and Li Sun) and I in mapping pollution fields, notably in Vancouver and Philadelphia. Interest in mapping such fields stems from the desire to avoid the deleterious effects of measurement error through the prediction of pollution levels down to fairly fine scales of resolution, especially in estimating human exposure and its health impacts. Among the problems are: (1) the inclusion of meteorological effects; (2) systematically and monotone missing data patterns; (3) the inseparability of spatial and temporal correlation in fields corresponding to short time aggregates (hours for example). I will describe approaches to the solution of these problems and illustrate them with applications to both of the cities referred to above. Particulate air pollution will be a focus of attention.
11:45 – 12:00	**Discussion**	
12:00 – 1:30	**Lunch**	
1:30 p.m. – 2:30	**Richard Smith,** Lecture #4	**Nonstationary spatial processes:** Various approaches to spatial modeling that do not assume the standard stationarity and isotropy conditions. Haas's moving windows approach. EOF analysis. Deformation models. Kernel models.
2:30 – 2:45	**Discussion**	
2:45 – 3:15	**Break**	
3:15 – 4:30	**Roundtable discussions**	

Wednesday, June 27, 2001

8:30 am—9:00	**Continental breakfast**	
9:00 – 10:00	**Richard Smith,** Lecture #5	**Models defined by conditional probabilities:** Markov random fields and the Hammersley-Clifford theorem; estimation by likelihood and pseudo-likelihood methods. Modern developments in which MRF models are used as priors within a broader hierarchical structure. The primary emphasis in this section will be on the contrast between models of this structure and the geostatistical approaches more commonly used in environmental statistics
10:00-10:15	**Discussion**	
10:15-10:45	**Break**	
10:45 – 11:45	**Paul Sampson**	**Spatial Covariance Modeling**
11:45 – 12:00	**Discussion**	
12:00 – 1:30	**Lunch**	
1:30 p.m. – 2:30	**Richard Smith,** Lecture #6	**Design of monitoring networks I:** The problem of locating monitors within a network to optimize prediction- or estimation-based criteria. Bayesian approaches to spatial data analysis and their application to network design through entropy criteria. Methods based on optimal design theory. Other approaches. Designs for data assimilation.
2:30 – 2:45	**Discussion**	
2:45 – 3:15	**Break**	
3:15 – 4:30	**Roundtable discussions**	

Thursday, June 28, 2001

8:30 am—9:00	**Continental breakfast**	
9:00 – 10:00	**Richard Smith,** Lecture #7	**Design of monitoring networks II**
10:00-10:15	**Discussion**	
10:15-10:45	**Break**	
10:45 – 11:45	**Doug Nychka**	**Wavelet representations for nonstationary spatial fields.:** Spatial analysis for large nonstationary processes poses challenges in both modeling and computation. A promising way to represent nonstationary covariance structure is by expanding the field in terms of a wavelet basis and then building a simple, sparse model for correlations and variances among the wavelet coefficients. In this talk a nonorthogonal wavelet basis (the W-transform) is presented that not only appears to fit a variety of standard covariance models but is well suited to the computation of Kriging estimates and conditional distributions. From a more conventional perspective, this wavelet-based model provides an reasonable blending between an EOF representation (principle components of the sample covariance matrix) and a stationary, parametric family. This approach is illustrated using output from a run of the Regional Oxidant Model, an EPA pollution simulation.
11:45 – 12:00	**Discussion**	
12:00 – 1:30	**Lunch**	
1:30 p.m. – 2:30	**Richard Smith,** Lecture #8	**Trends in Time Series:** An overview of various strategies for estimating and testing trends in time series, built around the theme of testing the significance of observed trends in global temperature series. ARMA and fractional ARIMA models; spectral approaches; long-range dependence. Extensions to multiple time series.
2:30 – 2:45	**Discussion**	
2:45 – 3:15	**Break**	
3:15 – 4:30	**Roundtable discussions**	

Friday, June 29, 2001

8:30 am—9:00	**Continental breakfast**	
9:00 – 10:00	**Richard Smith,** Lecture #9	**Extreme Values I:** The last two lectures have a somewhat different emphasis, where we look specifically at rare events, their estimation and prediction. However, the discussion will also take in such issues as whether extreme meteorological events are becoming more frequent, and the spatial integration of information about extreme events, thus providing a link with the rest of the course. Specific topics are: the three principal approaches to extreme value analysis based on annual maxima, threshold exceedances and point processes; estimation by moment-based, maximum likelihood and Bayesian methods; diagnostics. Extensions: extreme value regression and trend detection, spatial models for extremes.
10:00-10:15	**Discussion**	
10:15-10:45	**Break**	
10:45 – 11:45	**Peter Guttorp,** **Director, NRCSE**	**Meteorological adjustment of air pollution data:** A variety of statistical methods for meteorological adjustment of ozone have been proposed in the literature over the last decade or so. These can be broadly classified into regression methods, extreme value methods, and space-time methods. Among the crucial issues are questions of variable selection and trend estimation. The end use of the adjustment (e.g., monitoring trend, assessing health effects, etc.) largely determines these issues. I will illustrate the methods with ozone data from the Paris region in France, and particulate matter data from Phoenix, AZ.
11:45 – 12:00	**Discussion**	
12:00 – 1:30	**Lunch**	
1:30 p.m. – 2:30	**Richard Smith,** Lecture #10	**Extreme Values II**
2:30 – 2:45	**Discussion**	
2:45 – 3:00	**Closing remarks**	