Environmental Statistics –A Personal Review

Peter Guttorp

NRCSE

Technical Report Series

NRCSE-TRS No. 074

December 3, 2002

The NRCSE was established in 1996 at the University of Washington.

Environmental Statistics

–A Personal View

Peter Guttorp
Department of Statistics
Box 354322
University of Washington
Seattle, WA 98195-4322
USA

ABSTRACT

The field of environmental statistics is one of rapid growth at the moment. Environmental decision-making is prevalent in much of the world, and politicians and other decision makers are requesting new tools for understanding the state of the environment. In this paper, three case studies involving water pollution, air pollution, and climate change assessment are presented, together with brief descriptions of some other areas of environmental statistics. A discussion of future directions of the field concludes the paper.

Key words: Model assessment, heterogeneous correlation, hidden Markov model, general circulation models, downscaling, risk assessment.

1. Introduction

The field of environmental statistics is relatively young. The term "environmetrics" was apparently introduced in an NSF proposal by Philip Cox in 1971 (Hunter, 1994). During the last decade, the field has achieved some recognition, in that there now are three journals wholly or partially devoted to the field (*Environmetrics* ***published*** by Wiley; *Environmental and Ecological Statistics* published by Kluwer, and *Journal of Agricultural, Biological and Environmental Statistics* published by the American Statistical Association). There are three regularly occurring conferences with environmental statistics focus: the annual international Environmetrics meeting, the (roughly) biennial SPRUCE (Statistics for Public Resources, Utilities, and in Care for the Environment), and the triennial International Meetings on Statistical Climatology. There is an international Environmetric Society (TIES), The Royal Statistical Society has an Environmental Statistics Study Group, the American Statistical Association has a section on Statistics and the Environment, and the International Statistical Institute is currently discussing such a section. Volume 12 of the series *Handbook of Statistics* (Patil and Rao, 1994) was devoted to the topic of environmental statistics. Its 28 chapters constitute an interesting overview over the field. Wiley recently published a four-volume *Encyclopedia of Environmetrics* (El-Shaarawi and Piegorsch, 2001), with 530 entries written by 490 researchers. The Center for Statistical Ecology and Environmental Statistics at the Pennsylvania State University has been operating since 1969 with funding from the U.S. National Oceanic and Atmospheric Administration (NOAA). The US Environmental Protection Agency has funded a five-year National Research Center for Statistics and the Environment at the University of Washington, and is currently funding a similar Center for Integrating Statistical and Environmental Science at the University of Chicago.

In spite of all this activity, there are only a handful of courses taught in environmental statistics at US universities. As far as I know, only three schools in the United States (Cornell, Tulane and Penn State) offer Master's degrees in environmental statistics, al-

though several other departments allow an emphasis in environmental statistics for a general graduate degree in Statistics. Perhaps this is due to the relative rarity of Ph.D. level environmental statisticians; perhaps it is due to the complexity of many environmental problems; perhaps it is due to lack of critical mass even in the Statistics departments with a strong emphasis in environmental research. The field of environmental statistics has not yet reached the level of recognition within the statistical community that is needed to obtain the status of a separate subfield. Rather, it seems that the majority of statisticians consider environmental statistics as just one particular application of statistics. I will return to this discussion in the final section of this article.

Following my teaching strategy when teaching environmental statistics at the University of Washington, I will try to illustrate some aspects of the field by presenting three case studies, covering some aspects of environmental statistics that I and some of my coworkers currently are involved in. The first case deals with biological monitoring of water quality. Here, the methods are largely borrowed from classical compositional data models. The second deals with source-receptor modeling, and is using tools from multivariate statistics. Finally, I describe some aspects of stochastic downscaling of general circulation models to model local and mesoscale precipitation. The modeling techniques in this case originated in ecology. Common to these situations is considerable spatial and temporal dependence. These cases are some that I have been directly involved with. They do not span the range of work that could go under the title environmental statistics (and the references are in no way comprehensive), so in the penultimate section of the paper I present an overview of some of the other areas of research in the field, and sketch some of the methodology needed for that work. Four major areas: assessment of deterministic models (Fuentes et al, 2003); estimation of health effects (Dominici et al, 2003); data assimilation (Bertino and Wackernagel, 2003); and hierarchical modeling of complex systems (Wikle, 2003) are covered in other articles in this issue. The final section of this paper concerns the future of environmental statistics.

2. Biological monitoring

Traditionally, most environmental monitoring has been based on chemical analysis of water, air and soil. This allows for the identification of sources and pathways of pollution, and the thought is that some of the chemical changes are "biologically relevant." Biologists have generally not been convinced that the concept of biologically relevant chemistry is a valid one, and have proposed direct biological rather than indirect chemical monitoring methods (Cairns, 1979; Marmorek et al, 1988). The basic idea of biological monitoring is to study the effect of environmental insults on the biota by measuring changes in population composition, biodiversity, and so-called *indicator species*, or species that are particularly sensitive to certain kinds of pollution. There has also been an effort towards deriving indexes of pollution effect or biological integrity of a stream or watershed (e.g., Fore et al., 1996)

The US Environmental Protection Agency (EPA) started in 1989 an ambitious monitoring program called EMAP (Environmental Monitoring and Assessment Program). This was intended to create a "report card" for the state of the US environment. While EMAP has many components (see its home page at http://www.epa.gov/emap), I will here mainly be concerned with an effort to characterize changes in estuarine environments on the East coast of the United States. Specifically, I consider the Delaware Bay in the mid-Atlantic region, and in this bay the population of benthic (or bottom-dwelling) organisms.

The spatial design of the EMAP study (Overton et al., 1990) is a hexagonal grid, with a random starting point and a side of 27 km, resulting in 12,600 grid points over the continental United States, of which 25 fall in the Delaware Bay. The EMAP protocol required revisiting some of the sites on a rotating 3-year basis. The measurements made at each site (three times each summer) included a bottom grab sample of benthic organisms, together with measurements of covariates such as temperature, depth, and salinity.

The basic biological tenet behind this sampling scheme is that environmental insults affect the distribution of organisms, in that, with increased pollution levels, pollution tolerant species would tend to get a larger proportion of the sample than do sensitive species. Looking at the benthic fauna we expect bottom-feeding organisms to exhibit the clearest response to harmful changes in the sediment, while organisms that feed in the water column, or have mixed feeding strategies, would not react as much to changes in sediment quality. Both groups of organisms would be expected to react to changes in water quality. Consultation with biologists led to a classification scheme (see Billheimer et al., 1997, for details) in which we look at three groups of taxa: a pollution tolerant bottom-feeding group, a pollution sensitive bottom-feeding group, and a control-like group with mixed feeding strategy.

A simple model for this type of data would be a multinomial model with constant proportions of the three categories across the entire observation region. Some data analysis (Billheimer et al., 1997) indicates, however, that there is more variability between samples at the same site than can be explained by this simple model. Instead, we assume that the proportions follow a spatial autoregressive model in logit space, and that the observed counts are conditionally multinomial, given the actual proportions. This is a state space model in which the (unobserved) state is the actual proportion at the site, and the observation is the counts in the three categories. In the case where the true state follows a Markov process, this is called a hidden Markov model. I will return to that in Section 4. Letting $p(x) = (p_1, p_2, p_3)$ denote the vector of proportions at site x, $n(x) = (n_1, n_2, n_3)$ the corresponding counts, and $g(p) = \log(p_1/p_3, p_2/p_3)$, we write

$$n \mid p \sim \text{Mult}(1^{\text{T}}n, p)$$

$$g(p(x)) \mid p(y); y \neq x \sim \text{N}(\square(x), \square(x)) \qquad (1)$$

where $\mu(x) - \beta = \sum_{y \sim x} a_{x,y}(\mu(y) - \beta)$, is the conditional mean of $g(p(x))$, given all the other proportions; $y \sim x$ means that y is a neighbor of x,; and $\Sigma(x)$ is a block diagonal matrix depending on the number of observed neighbors of site x, the details of which are given in Billheimer et al. (1997). In other words, we have a conditional autoregressive model (CAR; see Besag 1974). The mean model can be made to depend on covariates by replacing the term β by a (usually linear) function of the covariate at the appropriate site.

This is a generalization of the work by Aitchison (1986) on iid proportions (see Billheimer, 1995). Model fitting is done using Markov chain Monte Carlo methods (Billheimer and Guttorp, 1996). Another approach to proportional data was developed by Grunwald et al. (1993) with an application to biological monitoring given by Guttorp (1993).

The analysis in Billheimer et al. (1997) provides a baseline measure of the 1990 variability of benthic counts in the three categories. The analysis showed a dependence on water salinity, with saltier water having proportionally more pollution intolerant species and fewer water feeders. The salinity gradient does not, however, explain all the spatial dependence. The spatial autoregressive parameter was deemed different from zero. Model checking was done by fitting the model to the data, leaving out one site at the time, and then estimating a prediction region for the observations at the left out site. For all sites except one, the three observations fell inside the prediction region. For the exceptional site, however, all three observations were well outside the prediction region. On-site inspection revealed a nuclear power plant coolant exhaust near the observation site!

Once a background model is determined, it becomes feasible to look for changes in the proportions of the different groups as a function of time. Data from 1991-94 were predicted, using observed salinity and the fitted model from 1990 (Silkey, 1997). The results were not very surprising: in essence, the data did not indicate any departures from the baseline model over the three years of further data. Comparing data to a baseline set on an estu-

ary with significant human pollution already present is a much less sensitive comparison than if the baseline can be set on relatively undisturbed waters. Unfortunately, there is not much of the latter available in North America.

3. Source-receptor modeling

3.1 General description

The chemical composition of air pollution at a given site can yield important clues as to the origin of the pollution. The problem of identifying these sources, and of attributing the proportion of air pollution at a site to a particular course has been studied in the chemometric literature under the name of multivariate receptor modeling (Henry, 1991, contains a review).

Assuming that q sources contribute concentrations $\boldsymbol{y}_t = (y_{1t},...,y_{pt})$ of p chemical species to the receptor at time t, a basic chemical relation is the mass balance equation

$$\boldsymbol{y}_t = \square_t \boldsymbol{P} \qquad (2)$$

where $\square_t = (\square_{1t},...,\square_{qt})$ is the vector of contributions from each of the q sources, and $\boldsymbol{P}$ is a $q \square p$ matrix with each column being the source composition profile (chemical fingerprint) of a particular source, Notice that $\square_t$ is a q-dimensional proportion.

A simple statistical model based on the mass balance equation (2) is obtained by adding measurement error,

$$\boldsymbol{y}_t = \square_t \boldsymbol{P} + \square_t . \qquad (1)$$

If the errors are iid, we can think of this as a factor analysis model in which $\boldsymbol{P}$ corresponds to the loading matrix and $\square_t$ to the factor scores. This model requires attention to be focused on chemical species which are not being highly reactive in the atmosphere, so that the fingerprint emitted at the source is advected by the wind and received relatively unattenuated at the receptor.

The traditional chemometric approach to the receptor modeling problem has been factor analysis. There are three major difficulties: (a) identifiability–unless the chemical fingerprints are sufficiently different we cannot separate pollution from different sources; (b) number of sources (generally unknown); and (c) temporal dependence. Park et al. (2001a) develop Bayesian methodology for estimating the number of sources and judging different identifiability conditions, while Park et al. (2001b) develop time series methodology for source-receptor models.

3.2 Spatial source-receptor modeling

Conventional receptor models depending on chemical mass balance equations are mainly used for multivariate air pollution data (a series of measurements on multiple chemical species, e.g. VOC species) collected at a single receptor. When the effect of meteorology is fairly constant (or negligible) over the region covered by receptors, a multivariate receptor model treating receptor sites as variables can be applied to locate spatially distinct source regions (Park et al., 2002a; Park et al., 2002b). Here, source composition profiles are interpreted as spatial profiles. Consider the profile of one chemical species (source type). The profile represents the relative amounts of that species (from the source) conveyed to receptor sites. The closer the receptor site is to the source, the higher the relative amount in the

spatial profile is. This approach was used for Seoul PM_{10} data (Park et al., 2002b), where a distinct source, only present in the winter, was detected and attributed to wood smoke.

The relationship of the concentration of air pollutants to wind direction can also be determined by nonparametric regression (Henry et al., 2002). The results are smooth curves with error bars that allow for the accurate determination of the wind direction where the concentration peaks, and thus, the location of nearby sources. The method was applied to cyclohexane data from 1997 at two sites near a heavy industrial region in Houston, Texas, USA. According to published emissions inventories, 70 percent of the cyclohexane emissions are from one source. Nonparametric regression correctly identified the direction of this source from each site. The location of the source determined by triangulation of these directions was less than 500 m from that given in the inventory.

Within the next few years, multivariate air pollution data obtained from multiple receptors will become available. Current receptor models do not fit into that framework. This calls for developing/extending multivariate receptor models to account for spatial variability in the data. When multiple species are measured at multiple receptors, incorporating spatial variability in modeling is a challenging problem. For simplicity, I assume no temporal dependence.

To account for spatial dependence in the data, we need to specify the spatial dependence structure in Φ and Γ, since the spatial correlations among the $\boldsymbol{y}$ are induced by those in Φ and Γ. This can be done either by a conditional autoregression as in the previous section, or by a geostatistical approach.

An alternative to the mass balance equation (2) is to consider the vector $\mathbf{y}_t$ as a composition, ignoring the measured masses of each chemical species. The methodology in section 2 can then be applied directly (Billheimer, 2001).

4. Downscaling of general circulation models

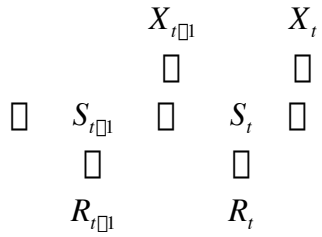
The problem of the possible global warming due to increased anthropogenic production and emission of greenhouse gases has been the subject of intensive research over the last decade or so (Intergovernmental Panel on Climate Change, 1995, 2001). Typically, the effect of changes in greenhouse gas concentrations is assessed by running the changed scenario through a (deterministic) model for the general atmospheric circulation and interaction with land and water surfaces. Modern general circulation models (GCMs) also contain sophisticated analyses of the oceanic circulation. Because of the complexity of the processes involved in GCMs, the resolution of such models is generally quite limited, such as 3 x 3 degree grid squares. In order to assess the effect of climate change on processes such as precipitation, which take place on a scale much smaller than the resolution of a GCM, it is necessary to somehow model subgrid variability for at least some of the processes involved in the GCM. The general term for this type of modeling is downscaling. For precipitation, there are two main approaches: the extremely compute-intensive mesoscale meteorological models (the Penn State/NCAR fifth-generation mesoscale model MM5, Grell et al., 1995, has been successfully applied in a variety of situations), or stochastic models for precipitation (e.g. Hughes and Guttorp, 1994).

The application of Markov models to precipitation has a long history, starting with Quetelet (1852). For some modern approaches, see Coe and Stern (1982) and Woolhiser (1992). Generally speaking the Markov chain approach suffers from two drawbacks: it is

unable to reproduce observed dry and wet spells, and it does very poorly in attempting to fit spatial patterns (e.g., Guttorp, 1995, section 2.12).

Instead of assuming that the patterns of precipitation follow a Markov model, the hidden Markov approach used in the benthic population model above can be used. Here the hidden states are thought to summarize the meteorological situation. There are two different main approaches to this: the weather states can be estimated from data, or they can be determined *a priori* from meteorological considerations. While the latter would be preferable from a scientific point of view, experiences thus far have not been overly successful. We (Hughes, 1993, Hughes and Guttorp, 1994a,b, Hughes et al., 1998, Bellone et al., 2000) have found a maximum likelihood approach to a hidden Markov model a reasonable way of determining suitable weather states. The Viterbi algorithm (e.g., Guttorp, 1995, p.112) can be used to estimate the most likely sequence of states, given the observations. An important aspect of this is that summaries of the atmospheric variables for each state have been found to agree with the understanding of meteorologists, although the weather states determined in this fashion seem to provide a less detailed distinction between atmospheric patterns than would the meteorological *a priori* determination.

In more detail, assume that there are a finite number S of weather states, which progress according to a temporally non-homogeneous Markov chain S_t. The transition probabilities for the Markov chain are assumed to depend (in a logistic fashion) upon atmospheric summary measures X_t. Given a weather state, the precipitation occurrences R_t at the network of stations is taken to follow an autologistic spatial model. A simple way of including the amounts into this model is to assume that amounts are conditionally independent between stations within weather states. Schematically, the dependence structure can be depicted thus:



The likelihood is evaluated using a Markov chain Monte Carlo approach (Hughes et al., 1998), combined with the standard forward-backward algorithm (Baum et al., 1970) for estimation in hidden Markov models. A recent algorithm developed by Lystig (2001) improves on the Baum algorithm, and allows for easy evaluation of derivatives of the log likelihood for use in estimating standard error of parameter estimators. In order to determine what atmospheric measures to use, and how many weather states are needed, we use a Bayes factor approach (Kass and Raftery, 1995).

Current work by Bellone (presentation at TIES 2002 conference in Genova, Italy) deals with precipitation in southeastern United States. There are 175 stations in the region with high quality data available for the entire period April-June 1965-1987.

Fitting the model outlined above to these data is computationally infeasible with current computing equipment. Rather, a subset of 30 stations was selected to fit a preliminary model. This model was then used to divide the area into three homogeneous regions: the coastal area, the mountainous area, and the inland area. Separate models, chosen to have five weather states each, were then fitted in the three regions. As is often the case with this type of precipitation models, there is a state with high probability of precipitation everywhere, and one with low probability everywhere. The remaining three states show substantial geographic variability of precipitation probabilities: one with rain in the south, one in the north, and one near the mountainous interior but not the coast.

The resulting model displayed some spatial dependence between regions, which could not be explained simply by large-scale atmospheric variables. A promising approach

uses a three-dimensional state space. Looking at the resulting dependence structure, there is some indication that two sites may be located in the wrong region.

5. Other topics in environmental statistics

The US Environmental Protection Agency is committed to assessing environmental problems using risk analysis. Traditionally, this has been done by putting down a deterministic model of the relationship between level of pollutant and effect. The typical risk function is a differential equation, with parameters that are determined from a variety of sources, such as laboratory experiments, measurements on exposed individuals, or scientific consensus. When the model has to do with human health effects, the basis for the risk function is more often than not experiments on animals, which are then rescaled to provide a risk functions for humans using a fairly arbitrary scaling factor.

Recently much emphasis has been put on uncertainty analysis of these risk assessments. Primarily it has been noted that the values of the parameters in the model is subject to uncertainty, which then propagates through the whole assessment and results in uncertainty about the final risk. The method of probabilistic risk analysis (Cullen and Frey, 1999) assigns what a statistician would call a prior distribution to each of the parameters. Typically, the parameters are treated as independent *a priori,* with simple marginal distributions such as uniform or normal. The analysis is done by simulating values from the prior distributions and summarized by producing simulated confidence intervals for quantiles from the resulting risk distribution. Current work aims at assessing the uncertainty more accurately by looking at the entire model uncertainty (e.g., Poole and Raftery, 2002, Clyde 2000, Fuentes and Raftery 2001). This includes, in addition to the uncertainty of the parameters mentioned above, uncertainty of the data used to fit and/or assess the model, and uncertainty of the model selection process itself. Bates et al. (2002) apply these models to a risk assessment for dredging a heavily polluted bay.

The detailed understanding of the health effects of a pollutant (Dominici et al., 2003) is one of the tools needed for setting scientifically valid standards for environmental compliance. Currently, most standards are set by requiring all sites in an area to remain below a limit for a given time period. As an example, the US standard for ozone requires that all sites in a region have an expected number of annual maximum daily 1-hour exceedances of 120 ppb of not more than one. Such a standard is not enforceable, since the expected number of exceedances is not directly measurable, and measurements cannot be taken at all sites. Rather, the rule was implemented by requiring that each site in an approved monitoring network, there would be no more than 3 exceedances in 3 years. In fact, this implementation was motivated in the regulation by applying the law of large numbers to n=3. The concept of statistically realizable ideal standards was introduced by Barnett and O'Hagan (1998). Their idea is to combine an ideal standard with a statistically based rule of implementation. A statistical approach to the problem of setting scientifically valid environmental standards is likely to borrow tools from industrial process control (Thompson et al., 2002) as well as from the cell level analysis of dose-response relationships (Leroux et al., 1996).

An area of considerable importance in all of modern statistics is the management, display and analysis of massive data sets. Land use data from satellite-based sensors, automated air quality sensors, continuous water flow meters are among a variety of new measurement devices producing vast amounts of data. The fields of data mining (Hastie et al., 2001) and geographic information systems (**ref**) provide some valuable tools. We are lacking tools for displaying uncertainty measures for spatially expressed data (see Lindgren and Rychlik, 1995, for one approach to confidence intervals for contour lines). Recent advances in three-dimensional visualization allow a viewer to consider spatially expressed multivariate data (essentially having a rotating scatter plot at each observational site; see Sutherland et al., 2000).

The field of environmental statistics was first implemented in the areas of sampling techniques and monitoring design. Modern sampling techniques that are finding increasing use in environmental problems include composite and ranked set sampling. Composite sampling, defined as the pooling of field samples prior to measurement or laboratory analysis, is a simple and straightforward method of enhancing sampling programs in situations where estimates of variability are less important. It is particularly useful when trying to assess whether or not a pollutant has spread into a given area. Ranked set sampling is a two-phase sampling procedure involving initial ranking of each of m samples of size m (often via a relatively cheap or fast method of measurement), followed by observing (often using a more accurate and more expensive method of measurement) the first order statistic from the first sample, the second order statistic from the second sample, and so on, until the m^{th} order statistic from the m^{th} sample yields a secondary sample of size m from the initial m^2 data points.

Environmental monitoring design deals mainly with two quite different sorts of design problems: monitoring for trend, where spatial and temporal dependence is of importance, and monitoring for "hot spots", or regions of local high intensity. The basic theory of optimal design for spatial random fields is outlined in Ripley (1981, Chapter 3). Among the popular designs are systematic random sampling designs, in which a point is chosen uniformly over the study area, and a regular design (consisting of squares, triangles, or hexagons) is put down starting at the chosen point. When the sample mean is used to estimate the spatial mean over a region, the regular sampling plans are most efficient (Matérn, 1960, Chapter 5). As an example, the EMAP monitoring program mentioned in section 2 above, implemented a hexagonal design with a random starting point. The hexagonal design requires fewer sampling sites than a square or triangular one to cover the same area, but does not take into account spatial covariance heterogeneity or temporal nonstationarity. The covariance mapping technique (Fuentes et al., 2003) can be used to deal with spatial heterogeneity, by implementing a spatial design in the transformed space. There is a need for de-

signs aimed at multivariate measurements. While optimality criteria can be useful in choosing between alternate designs, it is more important to develop designs that are easily explained to field workers and planners.

Much of the work in environmental statistics is aimed at providing decision tools for planners and decision makers at different levels. The problem of communication with non-statisticians therefore is very important, perhaps more so here than in other fields of application. Research specifically aimed at improving such communication is important, and not much has been done from the statistical point of view. Methods of display (including uncertainty) of monitoring data; methods for effects analysis and risk assessment that make sense to non-technical end users; methods for involving people with a stake in the decisions being made; and methods aimed at collaborative decision analysis are some of the areas in which current work is ongoing. In particular, much work is being done in several countries at the moment on the problem of regional summary measures of environmental quality, so-called environmental report cards. Other interesting areas involving decision tools include natural resource management, and environmental economics.

6. The future of environmental statistics

In each of the cases discussed above there has been important aspects of the work that originated outside statistics. In fact, most environmental statistics projects (and, indeed, most applied statistical projects) are driven by scientific questions that can only be solved in cooperation with other scientists. In the past, statisticians have often played the role of an outside expert, brought in to deal with the inferential aspects of a statistical problem. Our profession has frequently complained about this role, and urged scientists to bring in statisticians at an earlier stage: while designing their study. However, even then the statistician is viewed as an outside specialist, not really as a part of the research team. This must change. Not only in environmental statistics, but in most areas to which we apply our tools, we must start to learn to maintain a dialogue. This means that not only do we need to explain to other

scientists what we as statisticians consider important, but we need to learn their scientific language, so that we become useful members of research teams, rather than outside experts. Environmental problems are by their very nature multi-disciplinary, and statistics is just one of several of the disciplines involved. This has, obviously, important implications for how we teach our graduate students.

Many of the important environmental problems involve directly multi-dimensional, spatially heterogeneous, and temporally non-stationary random processes. My personal belief is that the development of statistical research tools (Cressie, 1993, is a good place to find some of the background material) in this area may prove to be the most useful development in the field of environmental statistics. The multivariate aspect, in particular, is very important, in that there are few symmetries in space and time that can be used in setting up models for realistic situations. As an example, if we are studying the joint distribution of SO_2 and SO_4 during situations of similar meteorology, we will find different space-time correlations for positive and negative time lags, since most of the SO_4 is produced from SO_2 emissions.

With an increased emphasis on the use of monitoring data to make environmental decisions, it is likely that environmental statistics will be a fruitful area of employment for statisticians. In addition to employment in governmental agencies and consulting firms, I would expect that large corporations will need some of the tools discussed in the previous section for their own environmental decision-making, and thus may need environmental statisticians to tailor these tools to their particular situation.

There seems to be an increasing demand for environmental statisticians, and, thus, for more graduate programs with environmental statistics pathways. In addition, there is a need for more statistics departments with groups of researchers focusing in environmental problems. As I mentioned above, there is a need for crossdisciplinary work, where statisticians work with other environmental scientists on an equal footing. I do not believe that we need departments of environmental statistics, or of environmetrics. The focus should be on

the multi-disciplinary nature of the environmental field, not on creating more subfields of existing fields.

The area of environmental statistics, albeit not yet a full-fledged subfield on its own, is full of interesting and complicated problems. There is much to be done, and we can use many more statisticians to do it.

Acknowledgments:

The research in Section 2 was done by a group including Dean Billheimer and Mariabeth Silkey. The work in Section 3 was performed jointly with Eun Sug Park. The work in Section 4 was done jointly with James Hughes and Enrica Bellone. For each of these projects the assistance of numerous colleagues in working groups at the National Research Center for Statistics and the Environment is gratefully acknowledged. The work described in Section 4 had partial support from the U S. National Science Foundation from grant DMS-9524770. Although the research described in this article also had partial support from the U. S. Environmental Protection Agency under cooperative agreements CR 821799 and CR 825173-01-0 with the University of Washington, it has not been subjected to the Agency's required peer and policy review, and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

References:

Aitchison, J. (1986): *The Statistical Analysis of Compositional Data*. London: Chapman & Hall.

Barnett, V. and O'Hagan, A. (1997): *Setting Environmental Standards*. London: Chapman & Hall.

Bates, S.C., Cullen, A.C. and Raftery, A.E. (2002). Bayesian Uncertainty Assessment in Multicompartment Deterministic Simulation Models for Environmental Risk Assessment. To appear in Environmetrics.

Baum, L. E., Petrie, T., Soules, G., Weiss, N. (1970): A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Statist.* **41:** 164–171.

Bellone, E., Hughes, J. P. and Guttorp, P. (2000): A hidden Markov model for relating synoptic scale patterns to precipitation amounts. *Climate Research* **15**: 1–12.

Bertino, L. and Wackernagel, H. (2003): Sequential data assimilation techniques in oceanography. *International Statistical Review*, this issue.

Besag, J. (1977): Efficiency of pseudolikelihood estimation for simple Gaussian fields. *Biometrika* **64:** 616-618.

Billheimer, D. (1995): *Statistical Analysis Of Biological Monitoring Data: State-Space Models For Species Composition.* Ph.D. dissertation, Department of Statistics, University of California.

Billheimer, D. (2001) Compositional receptor modeling. *Environmetrics*:**12**: 451–67.

Billheimer, D., and Guttorp, P. (1996): Spatial models for discrete compositional data. University of Washington Department of Statistics technical report **301.** Available at http://www.stat.washington.edu/tech.reports/tr301.ps.gz

Billheimer, D. , Cardoso, T., Freeman, E., Guttorp, P., Ko, H., and Silkey, M. (1997): Natural variability of benthic species composition in the Delaware Bay. *Env. Ecol. Statist.* **4:** 95–115.

Billheimer, D., Guttorp, P. and Fagan, W. E. (2001): Statistical interpretation of species composition. *J. Amer. Statist. Assoc.* ***96****:* 1205–14.

Burrough, P. A. (2001), GIS and geostatistics: Essential partners for spatial analysis. *Environmental and Ecological Statistics* **8** : 361–377

Cairns, J. Jr. (1979): Biological monitoring: concepts and scope. In Cairns, Patil and Waters (eds.): *Environmental Monitoring Assessment, Prediction and Management*. Fairland, MD: International Cooperative Publishing House, pp.3–20.

Clyde, M. (2000) Model uncertainty and health effect studies for particulate matter. *Environmetrics* **11**: 745–63.

Coe, R. and Stern, R. D. (1982): Fitting models to daily rainfall data. *J. Appl. Met.* **21:** 1024–ß1031.

Cressie, N. (1993): *Statistics for Spatial Data.* 2nd edition. New York: Wiley

Cullen, A. C. and Frey, H. C. (1999): *Probabilistic Techniques in Exposure Assessment*. New York: Plenum.

Dominici, F., Sheppard, L. and Clyde, M. (2003): Health effects of air pollution: A statistical review. *International Statistical Review*, this issue.

El-Shaarawi, A. H. and Piegorsch. W. W. (eds.) (2002): *Encyclopedia of Environmetrics* (4.vols.) Chichester: Wiley.

Fore, L., Karr. J. R., and Wisseman, R. W. (1996): Assessing invertebrate responses to human activities: evaluating alternative approaches. *J. N. Amer. Benthol. Soc.* **15:** 212-231.

Fuentes, M. and Raftery, A. E. (2001) Model Validation and Spatial Interpolation by Combining Observations with Outputs from Numerical Models via Bayesian Melding. Available at http://www.stat.ncsu.edu/~fuentes/combine.ps

Fuentes, M., Challoner, P. and Guttorp, P. (2003): Statistical assessment of numerical models. *International Statistical Review*, this issue.

Grell, G. A., Dudhia, J., and Stauffer, D. R. (1995): *A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model* . NCAR Technical Note NCAR/TN-398+STR. Boulder: National Center for Atmospheric Research.

Grunwald, G. K., Raftery, A. E., and Guttorp, P. (1993): Time series models for continuous proportions. *J. Roy. Statist. Soc. Ser. B* **55:** 103-116/.

Guttorp, P. (1993): Statistical analysis of biological monitoring data. In G. P. Patil and C. R. Rao (eds.): *Multivariate Environmental Statistics*: 165-174. Amsterdam: North Holland.

Guttorp, P. (1995): *Stochastic Modeling of Scientific Data*. London: Chapman & Hall.

Hastie, T., Tibshirani, R. and Friedman, J. H. (2001), *The elements of statistical learning: data mining, inference, and prediction.* Berlin: Springer.

Henry, R. C. (1991): Multivariate Receptor Models, in P Hopke (ed.): *Receptor Modeling for Air Quality Management*, pp. 117–47. Amsterdam: Elsevier.

Henry, R. C., Chang, Y-S. and Spiegelman, C. H. (2002): Locating Nearby Sources of Air Pollution by Nonparametric Regression of Atmospheric Concentrations on Wind Direction. Available at http://www.nrcse.washington.edu/pdf/trs71_regress.pdf

Hughes, J. P. (1993): *A Class of Stochastic Models for Relating Synoptic Atmospheric Patterns to Local Hydrologic Phenomena*. Ph.D. dissertation, Department of Statistics, University of Washington.

Hughes, J. P. and Guttorp, P. (1994a): A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena. *Water Resourc. Res.* **27:** 493-501.

Hughes, J. P. and Guttorp, P. (1994b): Incorporating spatial dependence and atmospheric data in a model of precipitation. *J. Appl. Meteor.* **33:** 1503-1515.

Hughes, J. P. , Guttorp, P and Charles, S. P. (1998): A nonhomogeneous hidden Markov model for precipitation. To appear, *J. Roy. Statist. Soc. Ser. C.*

Hunter, J. S. (1994): Enivronmetrics: an emerging science. In Patil, G. P. and Rao, C. R. (eds.): *Handbook of Statistics XII: Environmental Statistics*. Amsterdam: North-Holland.

Intergovernmental Panel on Climate Change (1995):*Climate Change 1995, The Science of Climate Change* . Eds. J. T. Houghton, L. G. Meira Filho, B. A. Callander, N. Harris, A Kattenberg and K. Manelli. Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (2001): *Climate Change 2001: Synthesis Report*. Edited by R. T. Watson and the Core Writing Team. Cambridge: Cambridge University Press.

Kass, R. E., and Raftery, A. E. (1995): Bayes factors. *J. Amer. Statist. Assoc.* **90:** 773-795.

Leroux, B. G. , Leisenring, W. M. , Moolgavkar, S. H. and Faustman, E. M. (1996), `A biologically-based dose-response model for developmental toxicology. *Risk Analysis* **16**: 449–458.

Lindgren, G. and Rychlik, I (1995): How reliable are contour curves? Confidence sets for level contours. *Bernoulli* **1:** 301-319.

Lystig. T. C. (2001): *Evaluation of Hidden Markov Models*. Ph.D. dissertation, Department of Biostatistics, University of Washington.

Marmarek, D. R., Berrard, D. P. and Ford, J. (1988): *Biological monitoring for acidification effects: U.S.–Canadian workshop*. Corvallis: Environmental Research Laboratory.

Matérn, B. (1960): *Spatial Variation*. Meddelanden från Statens Skogsforskningsinstitut, **49**, vol. 5. Republished in *Lecture Notes in Statistics*, vol. **36**. New York: Springer.

Overton, W. S., White, D. and Stevens, D. K. (1990): *Design report for EMAP: Environmental Monitoring and Assessment Program*. EPA/600/3-91/053. Washington: Environmental Protection Agency.

Park, E. S., Spiegelman, C. H. and Henry, R. C. (2002a): Bilinear estimation of pollution source profiles and amounts by using multivariate receptor models. To appear, *Environmetrics*.

Park, E.S., Oh, M. S., and Guttorp, P. (2002b): Multivariate receptor modeling and model uncertainty. *Chemometrics and Intelligent Laboratory Systems* **60:** 49–67.

Patil, G. P., and Rao, C. R. (1994): *Handbook of Statistics XII: Environmental Statistics*. Amsterdam: North-Holland.

Poole, D. and Raftery, A. E. (2000): Inference for deterministic simulation models: The Bayesian melding approach. *Journal of the American Statistical Association* **95**: 1244–55

Quetelet, A. (1852): Sur quelques propriétés curieuses que présentent les résultats d'une série d'observations, faites dans la vue de déterminer une constant, lorsque les chances de rencontrer des écarts en plus et en moins sont égales et independantes les unes des autres. *Bull. Acad. Royale Belg*. **19**, Parte 2: 303-317.

Ripley, B. D. (1981): *Spatial Statistics*. New York: Wiley.

Silkey, M. (1997): *Evaluation of a model of the benthic macro invertebrate distribution iof Delaware Bay, Delaware*. Thesis submitted in partial fulfillment of the M.Sc. degree requirements. Graduate program in Quantitative Ecology and Resource Management, University of Washington.

Sutherland, P. , Rossini, A. , Lumley, T. , Lewin-Koh, N. , Dickerson, J. , Cox, Z. , and Cook, D. (2000): Orca: A visualization toolkit for high-dimensional data. *Journal of Computational and Graphical Statistics,* **9** : 509–29

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

Wikle, C. (2003): Hierarchical models in environmental science. *International Statistical Review*, this issue.

Woolhiser, D. A. (1992): Modeling daily precipitation–progress and problems. Chapter 5 in Walden, A. T. and Guttorp, P. (eds.): *Statistics in the environmental and earth sciences*. London: Edward Arnold.