Selected Methods for Analyzing Trends
In Environmental Data Series

Shiying Wu & James V. Zidek

Technical Report #127

SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY
Selected Methods for Analyzing Trends
In Environmental Data Series

by

Shiying Wu & James V. Zidek

Technical Report 127

January 1989

Prepared in part under support to SIMS from the United States Environmental Protection Agency.

Department of Statistics, University of British Columbia, 2021 West Mall, Vancouver, British Columbia, CANADA V6T 1W5.
This is one of a series of technical reports produced under a Cooperative Agreement, "Study on Statistics and Environmental Factors in Health (Acid Deposition, Human Exposure, Quality Assurance, and Global Climate Change)" sponsored by the U.S. Environmental Protection Agency. The study is supported by a grant to SIMS (Societal Institute of the Mathematical Sciences) from the EPA. Cooperating members of the study are Rand Corporation, Stanford University (Departments of Statistics and Applied Earth Sciences), University of British Columbia (Department of Statistics), University of Washington (Department of Statistics), and SIMS.

SIMS is a nonprofit corporation founded in 1973 to "foster the application of mathematics to societal problems". To achieve this objective, SIMS has organized appropriate research and education studies and conferences, with mathematics and statistics playing a principal role.

Support for SIMS' programs has been forthcoming from private foundations, industry, and government agencies. Governance of SIMS resides with a Board of Directors from academe, government/industry, and the private sector. SIMS' research results have been recorded in a series of Technical Reports, which have been distributed to those requesting them. Most of the Technical Reports have been published in scientific journals.

SIMS Technical Reports are available from: SIMS Publications
97 Parish Road South
New Canaan, Connecticut
06840

PRINCIPAL INVESTIGATORS

Donald L. Thomsen, Jr.
Societal Institute of the Mathematical Sciences

Naihua Duan
The Rand Corporation

A. John Petkau,
U. of British Columbia, Statistics

Peter Guttorp
U. of Washington, Statistics

Paul Switzer
Stanford University, Statistics

SIMS BOARD OF DIRECTORS

Joel E. Cohen
Rockefeller University
William T. Golden
New York, New York
Ettore F. Infante
University of Minnesota
William H. Kruskal
University of Chicago
Brockway McMillan
Bell Telephone Laboratories
Emanuel R. Piore
Rockefeller University

Howard Raiffa
Harvard University
Mina S. Rees
City University of New York
Fred S. Roberts
Rutgers University
I. Richard Savage
Yale University
Donald L. Thomsen, Jr.
Societal Institute of the Mathematical Sciences
Arthur C. Upton
New York University Medical Center
ACKNOWLEDGEMENT

The idea of using median polish in the analyses of environmental data was suggested by Dr. Peter Guttrop in a proposal prepared for the Environment Protection Agency in 1985 through the Societal Institute for the Mathematical Sciences. We are much indebted to Dr. Tony Olsen and his colleagues at Battelle's Pacific Northwest Laboratories for providing data summaries so expeditiously. We thank the members of the SIMS group of investigators and various individuals from the EPA for their suggestions and encouragement.
Selected Methods for Analyzing Trends in Environmental Data Series

Summary

This report is one of a series giving the results of a study of trends between 1980 and 1986 in the chemistry of wet deposition.

The detection and estimation of possible temporal trends in the levels of various chemical constituents of wet deposition at different locations are the primary purposes of this study which actually consists of two analyses, one for the data collected between 1980 and 1986 (the "historical data"), and the other for the data collected between 1983 and 1986 (the "recent data"). In this way we have tried to differentiate between what happened in the past seven years from what happened in the last four years.

A second purpose of the study is the classification of the stations into homogeneous groups with similar deposition records. This will lead to a better qualitative picture of the overall situation in the continental United States. As well, clustering the stations in this way is a useful first step toward reducing the number of stations in the network and hence the cost of maintaining the network without losing too much information. This design issue will be addressed in a subsequent report.

Our study relies upon the extremely valuable Acid Deposition System (ADS), an integrated, centralized repository for data from monitoring networks in North America. The purposes of ADS are (1) to facilitate access to deposition data collected by different organizations, (2) to provide annual statistical summaries of available data, and (3) to maintain the data for assessment of long-term trends. A complete description of ADS is available in a system design and user's code manual by Olsen and Slavich (1986).
In general, according to Olsen and Slavich (1986), ADS requires networks to document their operation and provide a minimum level of information with each sample. Information about network protocols and data screening procedures along with sample results are incorporated into the ADS database. All sampling periods for which it is known that no precipitation occurred are considered valid sample periods. Wet deposition samples must be a wet-only sample. Wet deposition samples that have insufficient precipitation to complete a chemical analysis for a specified ion are considered invalid for that specific ion species. Any individual ion species concentration accompanied by a comment code designating the measurement as "suspect" or "invalid" is declared as an invalid sample. The actual sampling period for a wet deposition sample must be close to the network's protocol sampling period. Data completeness measures for each summary are computed. Criteria based on data completeness measures and site representativeness are selected for reporting a specific data summary. For more details, see Olsen and Slavich (1986).

The data set used in this study is obtained from ADS Data Summary (Olsen and Slavich, 1986) and it contains only the data from the National Atmospheric Deposition Program network, the largest and most important network among the networks represented in the ADS Data Base. It was established by the Association of State Agricultural Experiment Stations in 1978 to monitor trends in the exposure of various ecosystems to acidic deposition in the United States and, in cooperation with various research agencies to do research on atmospheric deposition. A major program objective is to discover and characterize biologically important geographical and temporal trends in the chemical climate of North America through the continued development and maintenance of a deposition monitoring network. Since its inception the network has grown from 22 operational sites during 1978 to 222 sites in 1986.

Over time, the NADP network came to include more than 200 monitoring stations. Many of them either started to operate after 1980 or have a lot of missing data. For
convenience and with little apparent loss of information, we consider only the stations with 5 or fewer missing observations. The resulting sets of stations are different for various chemicals under consideration. Roughly speaking, there are just over 80 stations providing the recent data and just over 30 stations providing the historical data. The stations are located throughout the United States and there are more stations in the East than in the West.

The Deposition Monitoring Task Group of the Interagency Task Force on Acid Precipitation was charged in the National Acid Precipitation Assessment Plan with developing a National Trends Network (NTN). The 150-station National Trends Network is intended to provide at least 10 years of monitoring at sites across the United States representing broad regional characteristics of the chemistry of wet deposition. Robertson and Wilson (1985) describe the design of NTN. Many existing NADP sites were incorporated in the NTN network, the two networks are now considered to be a single network and given the acronym NADP/NTN. In the following, however, we refer to it simply as NADP.

NADP monitoring is based on weekly (Tuesday to Tuesday) wet only sampling. The NADP program adheres to strict sample collection and analysis requirements which assure uniformity in siting criteria, sampling protocol, analytical chemistry techniques, data handling, and overall network operation. All NADP precipitation chemistry samples are analyzed by the Central Analytical Laboratory at the Illinois State Water Survey. For specific details the reader may consult existing publications on siting criteria (NADP 1984a), site operation and collection protocol (NADP 1982), overall quality assurance plans (NADP 1984b), and analytical procedures (NADP 1980).

This is the first report of the series presenting the results of our study of trends. In it we briefly describe the various statistical methods which are used in our investigation,
namely, clustering, median polishing, nonparametric trend testing plus slope estimation, and Kriging. Various data transformation are also used but these are not be discussed. The use of these methods is illustrated on the 1980 - 86 ADS data series for SO₄. The focus of the presentation is the methods under review rather than on the substantive results of the analysis. The latter constitutes the subject of the other reports of this series. A brief discussion of the various methods in the light of their performance is the topic of the last section of this report.

It is expected that this report will provide a useful introduction to these valuable statistical methods for other EPA investigators. In any case, this resume is necessary as background to the forthcoming series of SIMS technical reports which provide the substantive conclusions derived from this study of trends in wet deposition.
ABSTRACT

This report is one of a series giving the results of a study of trends between 1980 and 1986 in the chemistry of wet deposition. The primary purpose of that study is to detect and estimate the possible temporal trends in the levels of various chemical constituents of wet deposition at different locations. Secondarily, this study is leading to a classification of stations into homogeneous groups. This is leading to an overall qualitative impression of trend patterns over the continental United States. And forming these groups is a first step toward recommending the design of more economical future monitoring networks.

This report is specifically devoted to a review of the statistical methods which were found to be useful in the study. They are briefly described with examples for the possible use of environmetricians, or resource engineers in other context. The methods we present are: clustering, nonparametric analysis of slope, median polish, and Kriging.

KEYWORDS: Environmetrics; Acid rain; Clustering; Nonparametric Slope; Median Polish; Kriging.
Selected Methods for Analyzing Trends in Environmental Data Series

1. Introduction

This report is one of a series summarizing the results of a study of trends between 1980 and 1986 in the chemistry of wet deposition. The detection and estimation of possible temporal trends in the levels of various chemical constituents of wet deposition at different locations are the primary purposes of this study. A secondary purpose is the classification of the stations into homogeneous groups having similar deposition histories. This provides a qualitative overview of deposition patterns in the continental United States. It is also a natural first step toward reducing the number of stations in the network, and hence the cost of maintaining the network, without losing too much information.

The results of this study will be published elsewhere. However, in the belief that the statistical methods used in the study will be equally valuable in other environmental investigations, we briefly describe in the next section the various statistical methods which are used in our investigation, namely, clustering, median polishing, nonparametric trend testing with slope estimation, and Kriging. Various data transformations were used but these are not be discussed here. In Section 3 we illustrate the use of these methods on a data series for SO4. The focus of the presentation in Section 3 will be the methods under review rather than the substantive results of the analysis which will be reported elsewhere. A brief discussion of the various methods in the light of their performance in Section 3 is the subject of the last section of this report.

2. Methodology Background
2.1 Clustering. In the single sample problem, let \((x_{i1}, \ldots, x_{in})\) represent the ith observed \(n\)-dimensional sample vector, \(x_i, i = 1, \ldots, m\), which may well be heterogeneous. In our study \(x_i\) represents the series of levels of a given chemical measured over \(n\) time points at station \(i\). The aim of cluster analysis is to group these samples into \(g\) homogeneous classes where \(g\) is unknown, \(g \leq m\). The multi-sample variant involves \(x_{ik}\), \(k = 1, \ldots, n_i\), the observations of the ith sample \(x_i, i=1, \ldots, m\) but the aim is the same (Mardia, Kent, Bibby, 1979).

Here, we consider only hierarchical methods which after creating \(p\) clusters crates \(p+1\) groups at the next level of the hierarchy according to the distances between the samples, dividing one group into two and keeping the others the same. One way to do this is to start with \(\{C_1(1) = x_i, i = 1, \ldots, m\}\). Suppose \(\{C_i(p), i = 1, \ldots, m - p + 1\}\) are the clusters at step \(p\). Define \(D_{ij}(p)\) to be a measure of distance between \(C_i(p)\) and \(C_j(p)\). Let

\[
D_{12}(p) = \min\{D_{ij}(p) : i, j = 1, \ldots, m-p+1, i \neq j\}.
\]

Then set

\[
C_1(p+1) = C_1(p) \cup C_2(p)
\]

and

\[
C_{i+1}(p+1) = C_i(p), \quad i = 1, \ldots, m-p+1.
\]

Continue this procedure until all the inter-cluster distances are greater than \(D_0\) where \(D_0\) is an arbitrary threshold value.

If \(d_{rs}\) is taken to be the distance between \(x_r\) and \(x_s\) (the distance can be defined in various ways), and if

\[
D_{ij}(p) = \min\{d_{rs} : (r, s) \text{ such that } x_r \in C_i(p), x_s \in C_j(p)\}
\]
the method is called a \textbf{single linkage method}. If
\[
D_{ij}(p) = \max \{d_{rs}: (r, s) \text{ such that } x_r \in C_i(p), x_s \in C_j(p)\}
\]
then the method is called a \textbf{complete linkage method}.

The clusters obtained by a single linkage method are "rod" type elongated clusters without nuclei. This leads to a chaining effect. Chaining can be misleading if items at opposite ends of the chain are quite dissimilar (Johnson and Wichern, 1982).

On the other hand, the clusters obtained by the complete linkage method tend to be compact clusters without a chaining effect. Thus the within-group distances of the resulting groups are all less than the threshold value $D_0$.

Since one of the ultimate purposes of the clustering used here is to reduce monitoring costs by choosing representatives from each cluster and since we expect the inter-group distances to be small, we choose complete linkage in this study. For the same reason, the distances used in this study are defined as
\[
d_{ij}^2 = \sum_{k=1}^{n}(x_{ik}-x_{jk})^2/(n-1), \quad \text{for } i, j = 1, \cdots, m,
\]
where $n = n_i$.

In the single sample case, such a $d_{ij}$ is equivalent to Euclidean distance since the method is invariant under any monotonic transformation of $d_{ij}$. In multi-sample problems where the samples are one dimensional, if $x_i$ has mean $\mu_i$ and variance $\sigma_i^2 (i = 1, \cdots, m)$, and if the covariance of $x_i$ and $x_j$ is denoted by $\rho_{ij}$, $d_{ij}^2$ is an estimate of
\[
E[(x_i-x_j)^2] = \sigma_i^2 + \sigma_j^2 - 2 \rho_{ij} \cdot (\mu_i - \mu_j)^2, \quad \text{for } i, j = 1, \cdots, m.
\]
Thus a small $d_{ij}$ requires that both $\rho_{ij}$ be large and that $(\mu_i - \mu_j)^2$ be small. This means that the two samples vary in a similar fashion.

2.2 Nonparametric Monotone Trend Test and Slope Estimator. Let $x_1, \ldots, x_n$ be a sequence of observation ordered by time. We are interested in the null hypothesis,

$H_0$: the observations are randomly ordered, i.e., $x_1, \ldots, x_n$ are i.i.d. samples,

and the alternative hypothesis,

$H_1$: there is a monotone trend over time, i.e., $F_{x_i}(x) \geq (or \leq) F_{x_j}(x)$

for all $i < j$ with at least one strict inequality,

where $F_{x_i}(x)$ is the cumulative distribution function of the random vector $x_i$.

Let

$$\text{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases}$$

Mann (1945) proposes the following test statistic:

$$S = \sum_{i<j} \text{sgn}(x_j - x_i).$$

Under $H_0$, it has mean 0 and variance,

$$\sigma^2 = \frac{n(n-1)(2n+5)}{18},$$

while $S/\sigma$ is asymptotically $N(0,1)$.

Kendall (1975) gives the mean and variance of $S$ under $H_0$ when there may be ties in the $x$ values:
\[ E(S) = 0, \]
\[ \text{Var}(S) = \left\{ n(n-1)(2n+5) - \sum_{t} t(t-1)(2t+5) \right\}/18, \]

where \( t \) is the number of x's involved in the associated tie and \( \sum \) denotes the summation over all ties.

Both Mann and Kendall (\textit{ibid.}) derive the exact distribution of \( S \) for \( n \leq 10 \) and show that even for \( n = 10 \) the normal approximation is excellent, provided one uses a continuity correction, i.e., computes the standard normal variate \( Z \) by

\[
Z = \begin{cases} 
(S-1)(\text{Var}(S))^{-1/2}, & S > 0 \\
0, & S = 0 \\
(S+1)(\text{Var}(S))^{-1/2}, & S < 0.
\end{cases}
\]

Then in a two-sided test, a positive value of \( Z \) indicates an up-trend and a negative value of \( Z \) indicates a down-trend. This test is commonly called the Mann-Kendall test. Bradley (1968, p228) notes that when this test is used as a test of randomness against normal regression alternatives, this test has an asymptotic relative efficiency of 0.98 relative to the parametric test based on the regression slope coefficient.

Kendall's (1970) \( \tau \) test for correlation deals with a more general case. Suppose \((x_1, y_1), \ldots, (x_n, y_n)\) are a sequence of bivariate observations, ordered by time. The hypothesis to be tested is:

\[ H_0: \tau = 0 \]

versus
$H_1: \tau \neq 0,$

where $\tau = 2P(X_j > X_i | Y_j > Y_i) - 1$. In Kendall's procedure, a point estimator of $\tau$ is given by

$$\hat{\tau} = \sum_{i<j} \text{sgn}[(x_j - x_i)(y_j - y_i)].$$

For Mann's test, $y_i = i$, $i = 1, \ldots, n$.

Sometimes the time series of interest exhibits seasonal patterns and thus the hypotheses stated above may be too restrictive. However, different procedures are needed for dealing with such cases.

Let $X=(X_1, \ldots, X_p)$, where $X_g=(x_{1g}, \ldots, x_{ng})^t$ is a subsample for season $g$ and $x_{ig}$ is the observation obtained in the $i$th year and the $g$th season, $g = 1, \ldots, p$. A procedure proposed by Dietz and Killen (1981) can be used to look for trends in this case. The null hypothesis is that the $p$-vectors are randomly ordered versus the alternative hypothesis that there is a monotone trend in one or more of the $p$ variables. Let

$$S = (S_1, \ldots, S_p)^t,$$

and

$$\Sigma = (\sigma_{gh}),$$

where

$$S_g = \sum_{i<j} \text{sgn} (x_{ig} - x_{ij}), \quad g=1,\ldots,p,$$

$$\sigma_{gg} = \frac{n(n-1)(2n+5)}{18}, \quad g = 1, \ldots, p.$$
and
\[
\sigma_{gh} = \frac{1}{3} \left\{ \sum_{i<j} \text{sgn} [(x_{ij} - x_{ig})(x_{jh} - x_{ih})] + \sum_{(i,j,k)} \text{sgn} [(x_{ij} - x_{ig})(x_{jh} - x_{kh})] \right\},
\]
g \neq h, \text{ the estimated covariances of } S_g \text{ and } S_h. \text{ Then } S^t \Sigma^{-1} S \text{ is asymptotically } \chi^2_q, \text{ where } \Sigma^{-1} \text{ is any generalized inverse of } \Sigma \text{ and } q \leq p \text{ is the rank of } \Sigma.

Farrell (1980), following Sen (1968a), proposes another test procedure in which the data are "deseasonalised" first. Let \( y_{ij} = x_{ij} - \bar{x}_j \) where \( \bar{x}_j = \frac{1}{n} \sum_i x_{ij} \), \( R_{ij} \) is the rank of \( y_{ij} \) among all \( np \) \( y_{lm} \)'s, \( l = 1, \ldots, n, \ m = 1, \ldots, p \), and \( H_0 \) represents the hypothesis of no trend. Then
\[
T = 12p^2 \left( \frac{n(n+1)}{2} \sum_{i,j} (R_{ij} - R_{j.})^2 \right)^{-1} \left( \sum_i \left[ (i - (n+1)/2) (R_{i.} - (np+1)/2) \right] \right)
\]
is asymptotically \( N(0,1) \), where \( R_{j.} = \frac{1}{n} \sum_i R_{ij} \), \( R_{i.} = \frac{1}{p} \sum_j R_{ij} \).

Hirsh et al. (1982) define a multivariate extension of the Mann-Kendall test called the Seasonal Kendall Test. Let \( H_0 \) and \( H_1 \) be the hypotheses given, respectively, by

\( H_0 \): \( (x_{1j}, \ldots, x_{nj}) \) are independent and identically distributed random samples for \( j = 1, \ldots, p \) and the \( x_{ij} \)'s are independent,

and

\( H_1 \): the \( (x_{1j}, \ldots, x_{nj}) \)'s are not independent and identically distributed random samples.

Let
\[
S_g = \sum_{i<j} \text{sgn}(x_{ij} - x_{ig}) \ , \ g = 1, \ldots, p
\]
and
\[ S' = \sum_{g=1}^{p} S_g . \]

Then \( E(S') = 0 \), and
\[ \text{Var}(S') = \sum_{j=1}^{p} \text{Var}(S_j) . \]

Let
\[ Z' = \begin{cases} (S'-1)(\text{Var}(S'))^{-1/2}, & S' > 0 \\
0, & S' = 0 \\
(S'+1)(\text{Var}(S'))^{-1/2}, & S' < 0 ; \end{cases} \]

then \( Z' \) is asymptotically \( N(0,1) \) under \( H_0 \). The authors show that the normal approximation is quite accurate even for sample sizes as small as \( n=2 \), \( p=12 \).

The Seasonal Kendall Test is similar to a test proposed by Jonckheere (1954) as a multivariate extension of the sign test for the case when the number of observations is greater than 2. In the case that \( n_j \)'s are equal, the seasonal test and Jonckheere's (ibid.) test are equivalent.

Hirsch and Slack (1984) modify their seasonal Kendall test using Dietz and Killen's estimator of the covariances of \( S_g \) and \( S_h \). So the variance of \( S' \) becomes
\[ \text{Var}(S') = \sum_{g=1}^{p} \text{Var}(S_g) + \sum_{g \neq h} \text{cov}(S_g, S_h) . \]

The approximation is good in this test when \( p = 12 \) and \( n > 10 \). Compared with the Seasonal Kendall Test, this test is less powerful but more robust against serial correlation.
As van Belle and Hughes (1984) suggest in their paper, the procedures proposed by Hirsch (1982) and by Farrell (1980) may be derived from the same model, namely,

$$x_{ij} = u + a_i + b_j + e_{ij}, \quad i = 1, \ldots, n, \; j = 1, \ldots, p,$$

where

$$\sum_{i} a_i = \sum_{j} b_j = 0,$$

$a_1$ is the yearly component, $b_j$ is the seasonal component and $e_{ij}$ are i.i.d. with $E(e_{ij}) = 0$. The common hypothesis being tested is

$$H_0: a_1 = \cdots = a_n = 0$$

against

$$H_1: a_1 \leq \cdots \leq a_n \text{ or } a_1 \geq \cdots \geq a_n$$

with at least one strict inequality. These are also the hypotheses underlying the modified procedure of Hirsch and his colleagues in the 1982 article cited above.

van Belle and Hughes (1984) conclude that Farrell’s (1980) procedure is more powerful when there are no missing data since ranking all the data together preserves the relative ranks between seasons while these are lost in the Seasonal Kendall Test, while that of Hirsch et al (1982) is easier to compute when there are missing data.

Only the Dietz-Killen test is valid when the trends in different seasons are heterogeneous but in Hirsch et al (1982) it is argued that it is probably only applicable for sets of data including at least 40 years of monthly values.

All the other tests of seasonal data mentioned above will be misleading if the trends are not homogeneous among seasons especially when there are opposing trends in different
seasons. So van Belle and Hughes (1984) develop a 2-way, ANOVA-like nonparametric
trend test which can test for the homogeneity of trend at different locations and different
seasons. When there is only one location, it resembles a 1-way ANOVA and the statistic
they propose is

$$\chi^2_{\text{homog}} = \chi^2_{\text{total}} - \chi^2_{\text{trend}} = \sum_{g=1}^{p} Z_{g}^2 - \bar{Z}^2,$$

where

$$Z_{g} = S_{g} (\text{Var}(S_{g}))^{-1/2},$$

$$\bar{Z} = \frac{1}{p} \sum_{g=1}^{p} Z_{g},$$

and $S_{g}$ is the Mann-Kendall trend statistic for the $g$th season.

If the trend for each season is in the same direction then $\chi^2_{\text{homog}}$ has a chi-squared
distribution with $(p-1)$ degrees of freedom, $\chi^2_{q-1}$. If $\chi^2_{\text{trend}}$ exceeds a predefined critical
value, then the null hypothesis of homogeneous seasonal trends is rejected in which case
the Seasonal Kendall Test does not apply. However, if that hypothesis is accepted, then
$\chi^2_{\text{homog}}$ is the statistic used to test the hypothesis that the common trend direction is
significantly different from 0.

van Belle and Hughes (1984) point out that the validity of these $\chi^2$ tests require that the
$S_{g}$ be independent. A procedure for testing the contrasts is proposed and the use of the
Newman-Keuls procedure to group the seasons is illustrated in their paper as well.
Sen (1968b) develops a nonparametric procedure to estimate the slope of a possible existing trend with a $100(1-\alpha)\%$ confidence interval and it is robust against gross data errors and outliers. Let

$$Q_{ij} = (x_j - x_i)/(j - i).$$

If there are $N$ such $(i, j)$ pairs, the median is Sen's estimator of the slope.

A simple way to get a $100(1 - \alpha)\%$ confidence interval is by using normal approximation. Suppose $Q_{(1)}, \ldots, Q_{(N)}$ are the order statistics of the $Q_{ij}$'s. Then $[Q_{(m1)}, Q_{(m2+1)}]$ is a $100(1 - \alpha)\%$ confidence interval for the slope estimator, where

$$m1 = (N - Z_{1-\alpha/2}\sqrt{\text{Var}(S)})^{1/2}/2,$$

$$m2 = (N + Z_{1-\alpha/2}\sqrt{\text{Var}(S)})^{1/2}/2,$$

and $Z_{1-\alpha/2}$ is the $(1 - \alpha/2)$ quantile of the standard normal distribution and the $m$'s need not be integers in which case the $Q$'s in the interval are linearly interpolated.

A Seasonal Kendall slope estimator is given by Gilbert (1987). Suppose there are $N_k$ pairs of $x_{ik}, x_{jk}$ such that $i < j$. Then there are $N_k$ $Q_{ijk}$-values where $Q_{ijk} = (x_{jk} - x_{ik})/(j - i)$ for the $k$th season. In all, there are $N = N_1 + \cdots + N_p$ slope estimates, $Q_{ijk}$, for all the seasons combined. The median of these $N Q_{ijk}$'s is the Seasonal Kendall slope estimator and if $Q_{(1)}, \ldots, Q_{(N)}$ are the order statistics of $Q_{ijk}$'s, then $[Q_{(m1)}, Q_{(m2+1)}]$ is a $100(1 - \alpha)\%$ confidence interval for the slope estimator where $m1$ and $m2$ are as defined above with $S$ replaced by $S$.

2.3 Median Polish. Tukey (1977) proposes a procedure called median polish to fit a three-way model:
\[ y_{ijk} = u + a_i + b_j + c_k + a_{ij} + a_{ik} + b_{jk} + e_{ijk} \]  \hspace{1cm} (1)

where

- \( u \) is the common effect,
- \( a_i \) is the \( i \)th row effect,
- \( b_j \) is the \( j \)th column effect,
- \( c_k \) is the \( k \)th layer effect,
- \( a_{ij} \) is the interaction of \( a_i \) and \( b_j \),
- \( a_{ik} \) is the interaction of \( a_i \) and \( c_k \),
- \( b_{jk} \) is the interaction of \( b_j \) and \( c_k \),
- and \( e_{ijk} \) is the random error of \( y_{ijk} \).

Roughly speaking, this procedure uses data medians to estimate the main effects and interactions in the same way as means are used in classical ANOVA. Note that when we use means for fitting, the main effects and the interactions can be found in one "iteration", further iteration leaving the result unchanged. When we use medians, the first iteration may not be adequate.

Tukey (1977) develops the one- and two-way median polish, gives an example showing how three-way median polish proceeds, and points out that the generalization of three-way polish to higher-order cases is straightforward. Here we give a precise description of the three-way median polish procedure.

Suppose \( y_{ijk} \) is the observation for the \( i \)th level of factor a, the \( j \)th level of factor b, and the \( k \)th level of factor c \((i = 1, \ldots, I, j = 1, \ldots, J, k = 1, \ldots, K)\). In general, let \( x_M \) denote
the median of any given set of numbers, $x_1, \ldots, x_k$; these may have additional subscripts like $x_{ij1}, x_{ij2}, \ldots, x_{ijk}$ where now $x_{ijm}$ amounts to computing the median over the last subscript. Finally let

$$r_{ijk} = y_{ijk}, \text{ for all } i, j, k.$$  

We fit model (1) as follows:

(a) let $u^{(0)} = a^{(0)}_1 = b^{(0)}_1 = c^{(0)}_k = ab^{(0)}_{ij} = ac^{(0)}_{ik} = bc^{(0)}_{jk} = 0, r^{(0)}_{ijk} = y_{ijk}$, for all $i, j, k$;

(b) iterate successively the following recursive transformations:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u + c_M \to u$</td>
<td>$u + b_M \to u$</td>
<td>$u + a_M \to u$</td>
</tr>
<tr>
<td>$a_i + ac_{im} \to a_i$</td>
<td>$a_i + ab_{im} \to a_i$</td>
<td>$a_i - a_M \to a_i$</td>
</tr>
<tr>
<td>$b_j + bc_{jm} \to b_j$</td>
<td>$b_j - b_M \to b_j$</td>
<td>$b_j + ab_{mj} \to b_j$</td>
</tr>
<tr>
<td>$ab_{ij} + r_{ijm} \to ab_{ij}$</td>
<td>$ab_{ij} - ab_{im} \to ab_{ij}$</td>
<td>$ab_{ij} - ab_{mj} \to ab_{ij}$</td>
</tr>
<tr>
<td>$c_k - c_M \to c_k$</td>
<td>$c_k + bc_{mk} \to c_k$</td>
<td>$c_k + ac_{mk} \to c_k$</td>
</tr>
<tr>
<td>$ac_{ik} - ac_{im} \to ac_{ik}$</td>
<td>$ac_{ik} + r_{imk} \to ac_{ik}$</td>
<td>$ac_{ik} - ac_{mk} \to ac_{ik}$</td>
</tr>
<tr>
<td>$bc_{jk} - bc_{jm} \to bc_{jk}$</td>
<td>$bc_{jk} - bc_{mk} \to bc_{jk}$</td>
<td>$bc_{jk} + r_{mjk} \to bc_{jk}$</td>
</tr>
<tr>
<td>$r_{ijk} - r_{ijm} \to r_{ijk}$</td>
<td>$r_{ijk} - r_{imk} \to r_{ijk}$</td>
<td>$r_{ijk} - r_{mjk} \to r_{ijk}$</td>
</tr>
</tbody>
</table>

[Observe that in each successive transformation, the quantities involved are those obtained from that immediately preceding it. Thus, for example, the $b_M$ in $u + b_M \to u$ means the
median over \( j \) of the \( b_j \)'s obtained after transformation 1, namely, \( b_j = b_j^{(0)} + b_j^{(0)} \), where the superscript 0 indicates the initial values (0) of the quantities on the right hand side of this assignment statement.\]

(c) repeat (b) until that recursive transformation is reached when

\[
 r_{ijM} = a_{iM} = b_{jM} = c_M = 0
\]

or

\[
 r_{iMk} = a_{iM} = b_{Mk} = b_M = 0
\]

or

\[
 r_{Mjk} = a_{Mj} = a_{Mk} = a_M = 0.
\]

Tukey (1977) suggests that two cycles of (b) are enough -- whether or not we are at the stopping point indicated in (c). He also suggests using the boxplot to display the relative magnitude of the main effects, interactions and the residuals and hence the significant main effects and interactions.

2.4 Kriging and Universal Kriging. Kriging is the name given to a procedure for interpolating a spatial stochastic process; its name comes from D. R. Krige, a mining engineer in South Africa who first proposed the method in a geostatistical context. It is a commonly used method, no doubt in part because of its relative simplicity, it being just an estimate of the best linear unbiased inferential rules obtained by generalized least squares.
**Definition:** A stochastic process $Z(x)$ is said to be **strictly stationary** if the joint probability distribution function of $n$ arbitrary points is invariant under translation (Delhomme, 1976). i.e.,

$$(Z(x_1), \ldots, Z(x_n)) \sim (Z(x_1+h), \ldots, Z(x_n+h))$$

where $h$ is arbitrary and $\sim$ means "is distributed like".

**Definition:** A stochastic process $Z(x)$ is said to be **weakly stationary** if its first two moments are invariant under translation, i.e.,

$$E(Z(x)) = \text{constant for all } x$$

and

$$\text{Cov}(Z(x), Z(y)) = C(x - y),$$

for some positive definite functions and all $x$ and $y$.

**Definition:** The covariance of $Z(x)$ is said to be **isotropic** if

$$\text{Cov}(Z(x), Z(y)) = C(|x - y|),$$

where $|x-y|$ denotes the distance between $x$ and $y$.

**Definition:** A stochastic process $Z(x)$ is said to be **intrinsic** if the first two moments of $Z(x+h) - Z(x)$ depend only on $h$.

**Definition:** $r(h) = \frac{1}{2} \text{Var}(Z(x+h) - Z(x))$ is called the **semivariogram**. $2r(h)$ is called the **variogram**.

Observe that intuitively $r(h)$ is the negative of the covariance function. If
E(Z(x)) = constant,

then

\[ r(h) = \frac{1}{2} E[(Z(x+h) - Z(x))^2] \]

and given sufficient data it may be estimated nonparametrically by (Delfiner, 1975)

\[ r(h) = \frac{1}{N_h} \sum_{i=1}^{N_h} [Z(x_i+h) - Z(x_i)]^2 \]

where \( N_h \) = number of pairs of observations separated by a distance of approximately \( h \) and \( Z(x_i+h), Z(x_i) \) are observations.

There are several commonly used variogram models when \( Z \) is isotropic. One of those is called the generalized covariance model which is used when \( E(Z(x)) \neq \text{constant} \).

**Definition:** \( m(x) = E(Z(x)) \) is called the **drift** of \( Z(x) \).

Let

\[ e(x) = Z(x) - m(x); \]

then

\[ Z(x) = m(x) + e(x), \quad (2) \]

where \( E(e(x)) = 0 \). In (2), \( e(x) \) represents the random fluctuation and \( m(x) \) represents the slowly varying, continuous features of \( Z(x) \). Therefore, \( m(x) \) may be approximated within a restricted neighborhood by

\[ m(x) = \sum_{i=1}^{L} a_p f_p(x) \]
where the $a_p$ are constant coefficients and $f_p$ are arbitrary functions of the point $x$. In particular, if the \{f_i(x)\} are spatial polynomials, Kriging is called "Universal Kriging".

It is easily seen that to estimate $m(x)$ from the data we need to know the covariance of $e(x)$, and to estimate $e(x)$ we need to know $m(x)$. But usually in practice, neither is known and both must be estimated from the data.

Matheron (1973) and Delfiner (1975) have developed a technique to estimate $m(x)$ and the covariance of $e(x)$ simultaneously; the key concept is the generalized increment.

**Definition**: A generalized increment of order $k$ is a linear combination of the sample values $\sum_i \beta_i Z(x_i)$ for which $\sum_i \beta_i f_i(x_i) = 0$, where $f_i(x_i)$ is a polynomial of order less than or equal to $k$.

In the plane $(x_1 = (x_{11}, x_{21}))$, this condition yields:

$k = 0$, \quad $\sum_i \beta_i = 0$,

$k = 1$, \quad $\sum_i \beta_i = \sum_i \beta_i x_{1i} = \sum_i \beta_i x_{2i} = 0$,

$k = 2$, \quad $\sum_i \beta_i = \sum_i \beta_i x_{1i} = \sum_i \beta_i x_{2i} = \sum_i \beta_i x_{1i}^2 = \sum_i \beta_i x_{2i}^2 = \sum_i \beta_i x_{1i} x_{2i} = 0$.

An increment of order $k$ will filter out a polynomial trend of degree $k$. If $\hat{Z} = \sum_i \lambda_i Z(x_i)$ is a Kriging estimate of $Z(x)$, then $Z(x) - \hat{Z}$, the Kriging error, is a generalized
increment since it can be written as \(-Z(x) + \sum_i \lambda_i Z(x_i)\) and we see that \(\beta_0 = -1\) and \(\beta_i = \lambda_i\) for all \(i\).

**Definition:** If there exists a positive definite function, \(K(h)\), such that for any \(k\)th order generalized increment \(\sum_i \beta_i Z(x_i)\),

\[
\text{Var}(\sum_i \beta_i Z(x_i)) = \sum_{i, j} \beta_i \beta_j K(x_i - x_j)
\]

(3)

then \(K(h)\) is called a **generalized covariance function**, and \(Z(x)\) is known as an intrinsic process of order \(k\).

Matheron (1973) shows that any isotropic \((h = |h|)\) generalized covariance function defines a class of functions which are all equivalent up to the addition of an arbitrary even-powered polynomial of degree less or equal to \(2k\). For example, for \(k=1\), \(K(h) = -h\) and \(K(h) = -h + h^2\) are equivalent. For this reason, only the odd powers are essential when representing a generalized covariance. For orders up to 2, the isotropic polynomial generalized covariance kernels are (Jernigan, 1986):

\[
K(h) = \begin{cases} 
C\delta + a_0 h, & k = 0 \\
C\delta + a_0 h + a_1 h^3, & k = 1 \\
C\delta + a_0 h + a_1 h^3 + a_2 h^5, & k = 2 
\end{cases}
\]

(4)

where \(a_0, a_1, a_2 \leq 0, a_1 \geq -10\sqrt{a_0 a_2 / 3} \text{ in } R^2, a_1 \geq -\sqrt{10a_0 a_2} \text{ in } R^3\), and

\[
\delta = \begin{cases} 
1, & h = 0 \\
0, & h \neq 0 
\end{cases}
\]
The constraints ensure that (3) is nonnegative for any choice of \( \{\beta_1\} \). Polynomial generalized covariances with coefficients satisfying these constraints are said to be admissible.

Our objective is to estimate \( Z(x) \) by \( \hat{Z} = \sum \lambda_i Z(x_i) \) so that

\[
\begin{aligned}
E(\hat{Z}(x) - Z(x)) &= 0 \quad \text{and} \\
\text{Var}(\hat{Z}(x) - Z(x)) &= E(\hat{Z}(x) - Z(x))^2
\end{aligned}
\]

is minimized. \hfill (5)

Consider the case of a planar region \( x = (x_1, x_2) \). Using the previous assumption of a polynomial drift of order less than or equal to \( k \),

\[
m(x) = \sum_p a_p f_p(x),
\]

where the \( f_p(x) \) are given by \( \{1\} \) for \( k = 0 \), \( \{1, x_1, x_2\} \) for \( k = 1 \), and \( \{1, x_1, x_2, x_1^2, x_2^2, x_1 x_2\} \) for \( k = 2 \). It has been shown that the \( \lambda \)'s satisfying (5) can be found by solving the following system of equations, known as the Universal Kriging System:

\[
\begin{aligned}
\sum_j \lambda_i K(x_i - x_j) + \sum_p \mu_p f_p(x_i) &= K(x_i - x) \quad \text{for all } i \\
\sum_j \lambda_j f_p(x_j) &= f_p(x) \quad \text{for all } p
\end{aligned}
\]

(6)

If we know (a) the order of the drift \( k \), and (b) the coefficients of the generalized covariance function, we can solve (6) to obtain \( Z(x) \); (6) has a unique solution in the \( \lambda_i \), provided that the drift functions \( f_p \) are algebraically linearly independent, i.e.,

\[
\sum_p a_p f_p(x) = 0 \quad \text{for all } x \text{ if and only if } a_p = 0 \text{ for all } p.
\]
Here is a procedure to determine the order of drift \( k \) (Devary and Rice, 1982): for different values of \( k \), assuming \( Z(x) \) is an intrinsic function of order \( k \), delete each observation \( Z(x_i) \) in turn and calculate \( Z^{(k)}(x_i) \), the estimate of \( Z(x_i) \), from the remaining observations using the generalized covariance function

\[
K(h) = -h.
\]

This particular generalized covariance function is valid for any value of \( k \). Typically \( k \) is chosen to be 0, 1, or 2.

The best choice of \( k \) depends on the residuals \( \{ Z(x_i) - Z^{(k)}(x_i), \text{all } i\}, k = 0, 1, 2 \).

Three criteria are used to compare the residuals:

1. rank \( \{|Z(x_i) - Z^{(k)}(x_i)|, \text{all } i, k = 0, 1, 2\} \). [The value of \( k \) with the smallest average rank is the preferred order of drift value.];

2. the value of \( k \) with the smallest MSE is preferred, when

\[
\text{MSE}_k = \sum_i (Z(x_i) - Z^{(k)}(x_i)), \quad k = 0, 1, 2;
\]

3. the value of \( k \) with \( |Z_k| \leq 1.96 \) is preferred, when

\[
Z_k = \frac{e_k}{S(e_k)},
\]

\[
\bar{e} = \sum_i (Z(x_i) - Z^{(k)}(x_i)), \text{ and}
\]

\[
S(e) = \left[ \sum_i (Z(x_i) - Z^{(k)}(x_i) - \bar{e})^2/(n - 1) \right]^{1/2},
\]
since under the hypothesis that the true value of \( k \) is less than or equal to \( k \), \( Z_k \) is asymptotically \( \mathcal{N}(0, 1) \).

So, typically, a \( k \) with minimum MSE and/or average rank is selected.

We turn now to the problem of selecting the generalized covariance function. For a chosen \( k \), the model for a polynomial generalized covariance function of order \( k \) is

\[
K(h) = C\delta + \sum_{p=0}^{k} a_p |h|^{2p+1}, \quad \text{where} \quad \delta = \begin{cases} 
1, & h = 0 \\
0, & h \neq 0
\end{cases}
\]

Delete the \( r \)th observation from the data, then estimate \( Z(x_r) \) with the initial estimates of the coefficients \( \{C = 1, a_0 = -1, a_1 = 1, a_2 = -1\} \). For example, if \( k = 1 \), the initial estimate of \( K(h) \) is

\[
K(h) = 1 - h + h^3.
\]

Suppose

\[
Z_r(\beta) = \sum_{j} \beta_{ij}^{(r)} Z(x_j)
\]

is the error of the Kriging estimate of \( Z(x_r) \); by equation (3),

\[
\text{Var}(Z_r(\beta)) = E(Z_r(\beta))^2 = \sum_{i,j} \beta_{ij}^{(r)} \beta_{ij}^{(r)} K(x_i - x_j)
\]

or

\[
E(Z_r(\beta))^2 = C \sum_{j} (\beta_{ij}^{(r)})^2 + \sum_{p=1}^{k} a_p \sum_{i,j} \beta_{ij}^{(r)} \beta_{ij}^{(r)} |x_i - x_j|^{2p+1}.
\]

21
Let
\[ T^r_0 = \sum_j (b_j^{(r)})^2, \]
\[ T^r_p = \sum_{i,j} b_i^{(r)} b_j^{(r)} |x_i - x_j|^{2p+1}, \quad p = 1, \ldots, k. \]

Equation (7) can be written as
\[ E(Z_r(\beta))^2 = C T^r_0 + \sum_{p=1}^k a_p T^r_p \] (8)

which is linear in the coefficients. To determine the coefficients, Devary and Rice (1982) suggest regressing \( E(Z_r(\beta))^2 \) upon \( T^r_0, \ldots, T^r_k \) with the corresponding constraints on the coefficients shown in (4). The set of coefficients obtained by this regression is used to reestimate the observations, yielding a new set of errors and \( T^r \) values. The regression, (8), is performed again and again until the parameters converge. The proposed iteration is to be performed on all of the generalized covariance function of order \( k \).

Once several sets of admissible parameters have been determined it remains to choose the best one. Since \( Z_r(\beta) \) is a Kriging error and \( E(Z_r(\beta))^2 \) is an estimate of the Kriging variance, if the choice of \( K(h) \) is correct, \( E(Z_r(\beta))^2 \) should be close to \( \text{Var}(Z_r(\beta)) \) defined by (6). Therefore, the quantity
\[ \rho = \frac{\sum_r (Z_r(\beta))^2}{\sum_r \sigma_r^2} \]
should be close to 1, where $\sigma_r^2 = \sum_{i,j} \beta_i^{(r)} \beta_j^{(r)} K(x_i - x_j)$. In practice, therefore a jackknife estimator of $p$ is recommended by (Delfiner, 1975):

$$p = 2R - (n_1 r_1 + n_2 r_2) / (n_1 + n_2),$$

where

$$R = \frac{\sum_r Z_r(\beta)^2}{\sum_r \sigma_r^2},$$

$$r_1 = \frac{\sum_{r \in I_1} Z_r(\beta)^2}{\sum_{r \in I_1} \sigma_r^2},$$

$$r_2 = \frac{\sum_{r \in I_2} Z_r(\beta)^2}{\sum_{r \in I_2} \sigma_r^2},$$

$n_1$ = number of $r$'s such that $r \in I_1$, and $n_2$ = number of $r$'s such that $r \in I_2$.

It is sufficient to let $I_1 = \{1, \ldots, n/2\}$, $I_2 = \{n/2 + 1, \ldots, n\}$ where $n$ is the total number of observations. The generalized covariance is chosen as the one with the minimum MSE and jackknife statistic $p \in (0.75, 1.25)$. For $n \leq 50$, the minimal MSE criterion should be used.

3. Illustrative applications

The methods described in Section 2 are applied in this section to the Acid Deposition System (ADS), described in the manual of Olsen and Slavich (1986). More specifically
they are applied to the subset of these data derived from the NADP network, the largest and most important of the networks represented in ADS.

Although the NSDP/NTN has come to include more than 200 monitoring stations, we consider for simplicity only those with fewer than 6 missing monthly averages. This yields 30 stations covering the period 1980 - 86 and distributed over all of the continental United States.

The NADP monitoring protocol is based on a weekly (Tuesday to Tuesday) sampling protocol with wet-only sample collection but the summary used in this analysis uses monthly volume weighted averages. The NADP/NTN program has developed and adheres to strict requirements regarding sample collection and analysis. The requirements assure uniformity in siting criteria, sampling protocol, analytical chemistry techniques, data handling, and overall network operation. All NADP/NTN precipitation chemistry samples are analyzed by the Central Analytical Laboratory at the Illinois State Water Survey.

3.1 Transformation and Clustering. A preliminary exploration of the distribution of the data can provide important information which enables us to analyze the data efficiently and accurately. And transforming the data so that they are approximately normally distributed also makes the results of the analysis potentially simpler to interpret. Three commonly used transformations are, \( y = x^{1/2}, \ y = x^{1/4} \) and \( y = \log(x) \) and these were examined separately in this application. Often, and in this example too, the log transformation is best, in the sense that the resulting empirical distribution is approximately symmetric. However the empirical distribution does in this case have relatively long tails compared with the normal. This is similar to previous conclusions about transforming environmental data (c.f. Gilbert, 1987, p152). Since data transformation is relatively simple and routine these days, further details on this aspect of the analysis will be omitted for brevity.
The hierarchical clustering analysis with complete linkage described in Section 2.1 can be applied to either the untransformed or transformed data to cluster the sites under consideration. The difference is that if we cluster the log transformed data we can expect more clusters among the sites with smaller measurements than those with larger measurements. This is because log is a concave function. Since, in general, we are more interested in the sites with larger measurements, and we want to study them more carefully, we decided to cluster the untransformed data. Furthermore, since the precipitation volume weighted monthly means gives the quantity of chemicals in wet deposition rather directly, we apply this method to the volume weighted mean for all the components under study.

Figure 1 Here

The results of clustering depicted in Figure 1 show that at a fairly high level and ignoring Station 040a, the sites break into three clusters. We decided to stop there because there is only a small gap between 5 and 7 clusters and in turn between 7 and 10 so that clustering beyond 3 groups becomes quite arbitrarily. As well the first three clusters in the hierarchy seem fairly natural in that they represent (see Figure 2) fairly contiguous subregions of the country, namely, the subregion with the highest concentration of industry, a roughly concentric area around that region, and the rest of the United States. This result agrees with common sense; industries which emit more SO₂, and thus have more concentrated emission patterns should be more or less contiguous. Station 040a (Aurora, New York) is an outlier with quite a unique history.

Figure 2 Here
After clustering the data, the histogram of the log transformed data was drawn for each cluster. It was hoped that the transformed data within clusters would have histograms which are nearly normal; unfortunately this is not the case. The relatively long tails tend to persist although they do not obtain in every case. This suggests some uncertainty about whether or not the data can be treated as normal. Fortunately, our analysis does not depend critically upon the assumption of normality so the resolution of this issue is not vital.

The remainder of our illustrative analyses are based on the log transformed data which will be called "data" in the sequel for simplicity.

3.2 Trend and Seasonality. Since here and below, the normality of the data is in doubt, the use of parametric methods such as the analysis of variance to estimate trend and seasonality may not be appropriate. But these concerns are of no consequence with respect to the use of median polish. In addition, the median is resistant to outliers, which are quite prevalent. Therefore median polish is used to extract the trend and the seasonal structure of the data.

For sulfate concentration, model (2.3.1) can be rewritten as

\[ C = u + M_i + Y_j + S_k + MY_{ij} + MS_{ik} + YS_{jk} + e_{ijk}, \]  

(3.2.1)

where \( u \) is the common effect, \( M_i \) is the \( i \)th monthly effect, \( i = 1, \ldots, 12 \) standing for Jan., Feb., \ldots, Dec. respectively, \( Y_j \) is the \( j \)th yearly effect, where \( j = 1, \ldots, 7 \) stands for 1980,1981, \ldots, 1986 respectively, and, \( S_k \) is the \( k \)th site effect, \( k = 1, \ldots, 31 \); \( e_{ijk} \) is the residual.

Figures 3, 4 and 5 Here
The estimated main effects, namely, monthly, yearly and site effects are plotted in Figures 3, 4 and 5 respectively. The residuals are displayed using boxplots in Figure 6 where each box corresponds to a site. And a summary of the three-way polish is given in Figure 7 using boxplots where each box represents a source of main effects, interactions or residuals. This summary enables us to see how big the main effects and interactions are compared with the residuals. Further comment on these results is left to Section 4.

Figures 6 and 7 Here

3.3 Nonparametric Test, Slope Estimate and Kriging. Some characteristics of precipitation chemistry data, such as nonnormality, missing data and the limited number of observations over time, create some difficulties in using traditional parametric statistical methods to test for trend. But these characteristics cause no difficulty in using the nonparametric trend tests described in Section 2.2. So these tests are illustrated here and used in the study.

As mentioned in Section 2.2, for data with seasonal patterns, the test procedure which desesanlizes the data first and then ranks them altogether can preserve the relative ranks between seasons; but these are lost if the data are ranked within each season. Consequently, Farrell's (1980) test is quite powerful. But since there are missing data values which would make the computations for this test very complicated, we decided instead to desesanlize the data first and then use the Mann-Kendall Test to test for a monotone trend over time in the desesanlized data. For each site, the data are desesanlized using the one-way median polish by subtracting the median of the values of each month from the data for that month. Then the Mann-Kendall Test is applied to the desesanlized data to obtain the test statistics and the corresponding p-values. Sen's
nonparametric slope estimates are obtained with 80% confidence intervals. Since the alternative hypothesis is that the trend is monotone, for some of these tests our failure to reject the null hypothesis may result from the V-shaped pattern in the data. This is suggested in particular by the plots of the yearly effects obtained from the three-way median polish.

Figure 8 Here

Nonparametric slope estimates are plotted in Figure 8 on the map of the United States. If the lower 80% confidence bound of the slope estimate at a point is greater than 0 then a "+" is plotted at that point. If the upper 80% confidence bound of the slope estimate at a point is less than 0 then a "-" is plotted at that point. Otherwise a "0" is plotted on the point.

In another figure which, for brevity, is not included here, an "x" was plotted on the map of the United States if the p-value of the Mann-Kendall Test is less than 0.2 at a point. The resulting plot is quite similar to that in Figure 8 if we convert "+" and "-" in the latter to an "x". This is not surprising since we expect that the 80% confidence interval of the slope estimate and the p-value of the Mann-Kendall Test at the $\alpha = 0.2$ level to give us similar information.

Based on the slope estimates of the trend, universal Kriging estimates of the trend with estimated standard errors have been obtained at each integer degree grid point of longitude and latitude across the 48 continental states in the United States. The estimates and the estimated standard errors were calculated for the nearest 8 neighbors of the point being estimated. Note that usually the estimates and the standard errors estimated by Kriging are based on the raw data rather than on derived statistics treated as data. So our application of
Kriging is somewhat unconventional and we have to interpret the results with care. The results were plotted to show overall trend patterns. However, by now such plots of the results of Kriging are becoming quite common so again, for brevity, are omitted.

4. Discussion

This paper consists of a review of selected methods which the authors have found useful in a major study of trends in the chemistry of wet deposition in the United States during the seven year period, 1980-86. We hope other investigators will find them useful. They are selected because they are effective, straightforward and relatively free of restrictive distributional assumptions.

The latter is a particularly important feature since in general, environmental data are highly irregular even when, as in the case of the ADS system described in Section 3, great care is taken in the sampling and analytical procedures to ensure high quality. It would appear that inevitably, because of a lack of resources and knowledge of underlying processes, insufficiently many variables are measured in monitoring programs to enable really effective modelling. Variability is therefore hard to track. Heavy distributional tails, outliers, and other such statistical complications like those seen in Section 3 are bound to arise.

The hierarchical clustering method described in Section 2.1 helps to break down extreme variability somewhat. Nonetheless the tails in even the log transformed data like that considered in this paper remain somewhat heavy although not uniformly so over all ions and clusters in the study as a whole.

There is some arbitrariness in the choice of the number of clusters to retain. There did appear to be a distinct break better 3 and 5 in the hierarchy shown in Figure 1. This and the
desire to achieve a simple qualitative picture of the overall spatial pattern of the ion's concentration led to our choice.

Clustering also pointed to the outlying station 040a (Aurora, New York) at a latitude and longitude of 42nd 76 degrees, respectively. Its elevation is 249 (m.). It has been in operation since April 17, 1979 and its SO$_4$ deposition history appears from the cluster analysis to have been differed dramatically from the other sites considered in this report. But this turns out not to be the case as we argue below.

Here and elsewhere in this study, clustering has tended to identify geographically meaningful subregions of the United States. These are depicted in Figure 2. Cluster 2 represents the area of highest industrial concentration.

The median polishing method, which we have described in a directly useable form, is an extremely simple but robust method of summarizing complex data sets like those treated in our study. Its application here reveals, in Figure 3, remarkably consistent seasonal (monthly) patterns from cluster-to-cluster. This offers hope that seasonality might well be modelled out of the sulfate data series.

Median polishing also suggests (see Figure 4) a fairly marked down-trend in SO$_4$ levels at least in the more pristine subregion given by Cluster 3. However, there is a surprising and worrying recent up-trend shown in Figure 4 for both Clusters 1 and 2. This will be discussed in detail in a subsequent report concerned with substantive findings.

In Figure 5 are displayed the station effects for Cluster 2 which is selected for this illustrative application because it seemed to the authors generally to be the most interesting of the three clusters and, in particular, because it includes a dramatic outlier.

The plot shows Stations 075a (Parsons, West Virginia), 020a, and 032a are well below the norm for Cluster 2 while 040a, 058a, 041a (Chautauqua, New York), 021a (Argonne,
Illinois) and 056a (Caldwell, Ohio) are above. Bondville, Illinois (020a) appears to have the best overall record while Caldwell, Ohio (056a) seems to be the worst of the stations in Cluster 2. Of course, these findings are extremely tentative and while suggestive, are in no way to be regarded as firm conclusions. This issue will be further explored in subsequent reports on the results of this study. Here we intend merely to convey the value of the methods under review by showing the kinds of implications to which they can point.

Incidentally, it may be worth emphasizing that the outlier, 040a, does not affect the other values portrayed in Figure 5 because of the robustness of the median polishing procedure. The reason for 040a's distinguished position becomes apparent from Figure 6. There was one extraordinarily large monthly volume weighted mean which led 040a's distinguished position in the hierarchical clustering diagram of Figure 1. It turns out that on eliminating this single aberrant value, 040a rejoins the rest of the stations and easily enters one of the three existing clusters in a completely undistinguished way.

Some care is needed in interpreting the boxplots of Figure 7 because of the varying numbers of effects being plotted in each box. The spread should increase with the number of values. That point notwithstanding, it is clear that there is much less variation between stations and between years than there is between months. Thus trends, if any, are certainly dominated locally by natural variation. But obviously trends attain much more significance in the longer term.

The down-trends in pristine areas which are indicated in Figure 3, turn out to be significant for many of those and other stations according to the results of the Mann-Kendall analysis represented in Figure 8. There are, in fact, no + sign's, i.e. significant up-trends over that period for any of the 31 stations considered here.
REFERENCES


NADP. (1984b). NADP Quality assurance Plan, Deposition Monitoring. Natural Resource Ecology Laboratory, Colorado State University, Ft., Collins, CO.


Clustering of SO4 monthly volume weighted mean based on sqrt(MSE)

1980 - 1986

Figure 1
Clusters of log(SO4) monthly volume weighted mean based on sqrt(MSE)

1980 - 1986 (k=3)

Figure 2
Monthly Effect of log(SO4) for 3 Clusters
(monthly volume weighted mean, 1980 - 1986)
Yearly Effect of log(SO4) for 3 Clusters (monthly volume weighted mean, 1980 - 1986)
Station Effect of log(SO4) for Cluster 2
(monthly volume weighted mean, 1980 - 1986)

Figure 5
Boxplot for the resid. of log(SO4) (80-86, monthly volume weighted mean, clust 2)
Summary of the Effects and Residuals from Median Polish of log(SO4) (80-86, monthly volume weighted mean, clust 2)
Trend of $\log(\text{SO}_4)$ from 1980 to 1986 at the 31 Stations
(calculated by monthly volume weighted mean)

0  no trend
-  down trend
+  up trend

Figure 8