Recent Trends in the Chemistry of Precipitation in the United States. Part 10: Summary of Findings

Shiyng Wu & J.V. Zidek

Technical Report # 156
Recent Trends in the Chemistry of Precipitation in the United States.
Part 10: Summary of Findings

by

Shiying Wu & James V. Zidek

Technical Report 156

May 1990

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

Peter Guttorp
U. of Washington, Statistics

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

Paul Switzer
Stanford University, Statistics

SIMS BOARD OF DIRECTORS

Joel E. Cohen
Rockefeller University
William T. Golden
New York, New York
Ettope 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
ABSTRACT

This is the last of a series of reports based on a study of trends between 1980 and 1986 in the chemistry of wet deposition over the 48 conterminous states of the United States and American Samoa. This report concerns mainly the 1983-86 period. By then there were a substantial number of stations in the NADP/NTN network which produced the data used in our study. A detailed account of our results, given in a series of reports by Wu and Zidek (1989a-g, 1990a-b), concern nine selected ions: hydrogen, sulfate, nitrate, calcium, magnesium, sodium, potassium, chloride, and ammonium. Here we merely summarize our results for the volume weighted monthly mean concentrations levels for these nine ions.

For each ion it is found that there are between one and three contiguous regions of the United States ("clusters") within which wet deposition monitoring sites have similar records. It is found that the distribution of the logarithms of the data is substantially less skewed than that of the untransformed data for each ion and all but our cluster analysis uses the transformed data. A descriptive model and robust data summary for the 1983-86 period, derived by a median polish within each of the clusters, is presented. A confirmatory, site-by-site, nonparametric analysis of trends is also carried out for each of the ions.

Overall, the ions exhibited a variety of concentration histories over the period of study. The levels of hydrogen ion concentration in wet deposition generally increased since 1983 although at certain sites, there were significant down-trends. The same is true for sulfate levels in the more industrialized regions, but not outside these regions where levels remained more or less constant. Levels of nitrate remained approximately
constant over all regions. Those of sodium, potassium, calcium, and magnesium in wet deposition over the period of this study declined until the end of period except for a few individual sites with significant up-trends even in the recent past. There was a marked decline in the levels of chloride and ammonium except in the Eastern states in the latter case. There are a variety of individual site histories and these are presented along with an analysis of seasonal patterns.

KEYWORDS: Acid Rain; median polishing; nonparametric slope estimation; sulfate; nitrate; hydrogen; chloride; ammonium; calcium; magnesium; sodium; potassium
Recent Trends in the Chemistry of Precipitation in the United States.

Part 10: Summary of Findings

Summary

This report is the last of a series on 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 a number of chemical constituents of wet deposition at different locations were the primary purposes of this study which actually consisted of two analyses, one for the data collected between 1980 and 1986, and the other for the data collected between 1983 and 1986. In this way we have tried to differentiate between what happened in the past seven years from what happened in the more recent past. This final report will focus primarily on the latter period because by that time the NADP/NTN network which generated the data for our study includes a large number of monitoring sites.

A secondary purpose of the study was the classification of the stations into homogeneous groups with similar deposition histories. This gives an overview of recent pollution patterns 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 issue will be addressed in a subsequent report.

Our study relies upon the Acid Deposition System (ADS), an integrated, centralized repository of data from monitoring networks in North America. The purposes of ADS are to facilitate access to deposition data collected by different organizations, provide annual statistical summaries of available data, and maintain the data for as-
essment 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 (ibid.), 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. Wet deposition samples must be wet-only samples and any 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 invalid. 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 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, but 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, only the data from stations with 5 or fewer missing observations were used in the study. Thus just over 80 stations provided the 83 - 86 data, on which most of the results cited in this report are based; just over 30 stations provided the 80 - 86 data analyzed in the detail in earlier reports. 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 the NTN which incorporates many existing NADP sites. The two networks are now considered to be a single network with the acronym NADP/NTN. In the sequel, however, we refer to it simply as the NADP network.

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 last of a series of reports based on a study of trends between 1980 and 1986 in the chemistry of wet deposition over the 48 conterminous states of the United States and American Samoa. This report concerns mainly the 1983-86 period. By then there were a substantial number of stations in the NADP/NTN network which produced the data used in our study. A detailed account of our results, given in a series of reports by Wu and Zidek (1989a-g, 1990a-b), concern nine selected ions: hydrogen, sulfate, nitrate, calcium, magnesium, sodium, potassium, chloride, and ammonium. Here we merely summarize our results for the volume weighted monthly mean concentrations levels for these nine ions.

For each ion it is found that there are between one and three contiguous regions of the United States ("clusters") within which wet deposition monitoring sites have similar records. It is found that the distribution of the logarithms of the data is substantially less skewed than that of the untransformed data for each ion and all but our cluster analysis uses the transformed data. A descriptive model and robust data summary for the 1983 - 86 period, derived by a median polish within each of the clusters, is presented. A confirmatory, site-by-site, nonparametric analysis of trends is also carried out for each of the ions.

Overall, the ions exhibited a variety of concentration histories over the period of
study. The levels of hydrogen ion concentration in wet deposition generally increased since 1983 although at certain sites, there were significant down-trends. The same is true for sulfate levels in the more industrialized regions, but not outside these regions where levels remained more or less constant. Levels of nitrate remained approximately constant over all regions. The levels of sodium, potassium, calcium, and magnesium in wet deposition over the period of this study declined until the end of period except for a few individual sites with significant up-trends even in the recent past. There was a marked decline in the levels of chloride and ammonium except in the Eastern states in the latter case.

Strong seasonality for all the ions is also revealed by our analysis. Nonparametric monotone trend tests are performed on the data and described here for the period from 1983 to 1986. The results in all cases are consistent with those from the median polish approach. A nonparametric slope estimate of the trend is obtained for each monitoring station. Geographical pattern in the data is suggested by hierarchical clustering and by median polishing.
Recent Trends in the Chemistry of Precipitation in the United States.

Part 10: Summary of Findings

1. INTRODUCTION

This paper describes a study of the chemistry of wet deposition in the United States between 1980 and 1986. Possible trends in the levels of various chemical constituents of precipitation were of particular interest. Of secondary interest was broad spatial pattern. So the stations in the National Atmospheric Deposition Program (NADP) Network and National Trends Network (NTN), which generated the data used in our study, were clustered into homogeneous subgroups of stations having similar deposition histories. This was deemed to be a useful first step toward the possible reduction, without significant loss of information, of the size of the network and hence of its operating cost.

Our study relies upon a record of monthly volume weighted mean concentrations and monthly median concentrations (in mg/l except for those of hydrogen ions which are in μg/l) supplied to us by Battelle's Pacifac Northwest Laboratory which maintains an integrated, centralized repository for data from monitoring networks in North America. The database, called the Acid Deposition System (ADS), is described in detail in a system design and user's code manual by Olsen and Slavich (1986).

Although the study examined monthly median and volume weighted mean levels of concentration (see Wu and Zidek, 1989a-i and 1990a-b for details), there was little overall difference in the results for the two analyses and for brevity only those for means will be presented.

Additional brevity is gained by concentrating only on the the second and arguably
more relevant of the two time periods, 1980–86 and 1983–86 studied. Moreover, the number of stations without significant amounts of missing data had increased to a more substantial 81 and, in the case of hydrogen ion concentrations, 86 stations from the approximately 30 stations in the longer period. The trends suggested by the 80-to-86 data for the 1983-86 period were, however, entirely consistent with those suggested by the 1983 to 1986 data, the subject of this paper.

The codes from the ADS manual for the stations in our study are given in the Appendix and their location is shown in Figure 1. Except for the analysis of hydrogen ion concentrations which used the data from all of these stations, only data from the 81 stations obtained by deleting Stations 007a, 015a, 016a, 062a, and 339a from the long list were used.

In the next section, the main results of our analyses of volume weighted mean concentration levels (mg/l, and µg/l for hydrogen) in the 1983-86 period are given for the nine ions in this study, hydrogen (H), sulfate (SO₄), nitrate (NO₃), sodium (Na), potassium (K), calcium (Ca), magnesium (Mg), chloride (Cl), and ammonium (NH₃). The statistical methods involved in the study are discussed in the first report of this series (see Wu and Zidek, 1989h).

We investigated transformations to achieve a symmetrical distribution of the transformed data values for each ion. As indicated Section 3, the logarithm proved best overall among those investigated and was adopted for simplicity and to enable comparisons between the various ions. But the data for some ions like Cl proved to have such heavy tails that the log-transformation was not satisfactory.

Our cluster analysis is discussed in Section 2. For some of the ions like Cl, there
seemed to be no natural partitions in the data whereas for others like $SO_4$, there were three.

Section 4 presents the results of three way median polishes of the data for each ion and cluster. The twelve "month-effects", the 81 (and for hydrogen, 86) "site-effects", and the 4 "year-effects" (1983 - 1986 inclusive) are all given along other information derived from the median polish.

While the median polishes of Section 4 provide a good overall view of temporal trends, they do not give a good view of the complex spatial pattern underlying them. In this direction, a nonparametric trend analysis (including significance testing and slope estimation) was carried out for each site and each ion. The results for the 1983-86 period are described in Section 5.

In the previous reports in this series are presented the results of kriging the nonparametric slopes estimates to determine spatial pattern. But for brevity the results from this kriging analysis will not be given in this report.

Our tentative conclusions are given in Section 6 with discussion.
2. CLUSTER ANALYSES

A cluster analysis was performed for each of the ions to achieve a overview of the spatial pattern of each ion's concentration levels. For any given ion, the separation between any given pair of sites was taken to be the square root of the average of squares of the differences between their consecutive monthly average concentrations. To be close then, the records of two stations would have to have similar overall averages, variability, trends, seasonality, and so on. However, this measure of separation is highly sensitive to extreme values and these must be deleted before clustering to avoid misleading results like a station isolated in a one site cluster simply because of a single aberrant value.

The separation of any two groups or "clusters" of stations was taken to be the maximum separation between any pair of stations, one from each cluster. This ensured that the resulting clusters were compact rather than strung out through a series proximate station pairs.

For any given ion, clustering began by identifying the closest pair of stations. These were combined to form the first cluster. An examination of Figure 2.1(a) for hydrogen, for example, reveals that stations 074a (Olympic National Park, Washington), and 172a (American Samoa, American Samoa) had very similar data records over the four years under study. Either would be a potential target for elimination if cost reductions were necessary.

Next the separation from this first cluster to its "closest" outside neighbouring station was determined. This was compared with the separation of the next closest pair of stations at least one of which was outside the first cluster. Referring again to
the hydrogen example, it is found that Station 061a (H.J.Andrews, Oregon) is very close to the established cluster but its separation from that cluster exceeds slightly the separation between stations 037a (Glacier National Park, Montana) and 271a (Headquarters, Idaho). So the second cluster in the hierarchy was formed out of the latter pair of stations.

Now Station 061a joins the first cluster so that at the next level in the hierarchy, there are two small clusters of three and two stations, respectively. Analysis proceeds in this way until at the top of the cluster diagram we see three and ultimately just one cluster, the latter including all the 86 stations involved in the hydrogen ion analysis.

We comment briefly below on the results of the analysis for each ion and refer the reader to the the reports by Wu and Zidek (1989a-g, 1990a-b) for more detail.

HYDROGEN. Figure 2.1a shows the hierarchical cluster structure for the 86 sites depicted in Figure 1 which generated the data used in our analysis of this ion. The sites partition naturally into three clusters, labelled by cluster number and shown in Figure 2.2a. Cluster 3, which covers the subregion of highest industrial concentration, is surrounded by Cluster 1. Cluster 2 covers the rest of the United States. This shows in particular that the acidity of wet deposition in the industrial areas has a different character from that of the other areas.

SULFATE. Before clustering the weighted mean data, two outliers were deleted these being single observations for the Stations 035a (Lamberton, Minnesota) and 040a (Aurora, New York). They correspond to January, 1985 and September, 1986, respectively. The first, 365.56, was much larger than 6.39, the second largest observation at the same site. The cluster diagram in Figure 2.2(b) shows that the sites
partition naturally into three clusters except for the two stations in Texas, 070a and 071a. For simplicity these two sites were assigned to cluster 3, the closest one according to Figure 2.1b. The locations of the sites labelled by cluster are shown in Figure 2.2b. The overall pattern revealed by this analysis is similar to the one in Figure 2.2a: Cluster 1 is located in the subregion of highest industrial concentration, Cluster 2 surrounds it, and Cluster 3 is spread out over the rest of the United States. We believe this is because emissions of $SO_2$ are higher in the industrial areas than elsewhere.

A cluster analysis was also done without deleting outliers. The clusters are the same except for the stations identified in the last paragraph, 035a and 040a which are then isolated from all the clusters.

**NITRATE.** Before clustering an outlier was deleted from the data of station 040a; the same station and month gave rise to the outlier identified above for the sulfate series. Figure 2.1c shows the hierarchical cluster structure for these data. The sites first partition naturally into two clusters: a large one on the left and a smaller one on the right. The large cluster can be further subdivided into Clusters 1 and 2 where for simplicity Stations 035a (Lamberton, Minnesota), 268a (ZWSW, Arkansas) and 283a (Lake Bubay, Wisconsin), have been assigned to the nearer of the two clusters, Cluster 1, as shown in Figure 2.2c. Cluster 3 is over the subregion of highest industrial concentration, while Clusters 1 and 2 are spread out over the rest of the United States. The overall pattern is similar therefore to that of hydrogen and sulfate above, with similar implications.

Our analysis was also done without deleting outliers with the expected result that
all the clusters are the same except for Station 040a which is singled out.

**SODIUM.** Figure 2.1d shows the hierarchical cluster structure of these data. There are two major clusters plus a few small clusters consisting of one or two sites. For convenience, all the small clusters were grouped into one cluster called Cluster 3. The locations of the sites labelled by cluster are shown in Figure 2.2d Cluster 2 consists of sites scattered haphazardly over the United States. All the sites in Cluster 1 are either on an island or along the coast unlike the sites in Cluster 2. This indicates, not surprisingly, that sodium concentration levels are related to geographical location, and more specifically, proximity to the sea.

**POTASSIUM.** Figure 2.1e indicates that the 81 sites which generated these data cannot be split into a small number of homogeneous clusters. There is one very large homogeneous group in the Figure with Station 024a (NIARC, Illinois) on its left, but then the remaining 11 stations fall into no less than 9 separate clusters. In Section 4, we analyse these sites as a single group.

**CALCIUM.** Figure 2.1f shows the cluster structure for these data. There are two major clusters plus a few small clusters consisting of just one or two sites which for simplicity, are combined in a single cluster labelled 3. The locations of the sites labelled by cluster are shown in Figure 2.2f. Of the two major clusters, one is in fact rather small and has just 11 sites in it. And eight of these lie along a North-South line centered at Station 011a (Manitou, Colorado); the remaining three cluster around Argonne, Illinois. The biggest cluster, 2, is widely scattered over most of the United States like the heterogeneous but small Cluster 3.

**MAGNESIUM AND CHLORIDE.** Figures 2.1g and h show the cluster struc-
ture for these of these ions. In each case there appears to be no natural partition of
the sites into a small number of relatively homogeneous groups. Thus our analysis in
Sections 3 and 4 treats these sites as one large cluster.

AMMONIUM. Two major clusters plus a few small clusters consisting one
or two sites are identified in Figure 2.1i. For simplicity, all the small clusters are
combined in Cluster 1. The locations of the sites labelled by cluster are shown in
Figure 2.2i; Cluster 2 covers the subregion of highest industrial concentration.
3. DATA TRANSFORMATIONS.

In our study it was not the raw data but rather monthly aggregates, medians in the one case and volume weighted means in the other which were to be assessed. We were therefore surprised that the distribution of the series of monthly values was nevertheless, highly right skewed and heavy tailed for each ion. The data histograms tended to have one large bar over the interval adjoining the origin. It is difficult to see patterns in such data so to spread it out, and make it more symmetric it was decided to transform the data for each ion.

Three commonly used members of the Box–Cox family, the square root, the fourth root and the logarithmic transformation, were compared on an ion–by–ion basis for each of the two time periods in our study for the volume weighted means data only. No clear winner emerged, although all three dramatically reduced the skewness of our data distributions. The tails of those for hydrogen, sulfate, are relatively light and all three of the transformations helped to bring the data close to a Gaussian form as exemplified in Figure 3.1 for sulfate; arguably, the square or the fourth root transformations was the best. For sodium and potassium the logarithm worked best among the three transformations and made the resulting distribution of transformed values quite symmetric. For calcium and ammonium something between the fourth root and logarithmic transform seemed preferable while for magnesium and chloride, the ions with distributions having the heaviest tails, the logarithm was the best of the three. However in the case of the latter two ions, the resulting distribution remained somewhat right skewed although much less so than that of the original data. This is illustrated in Figure 3.2 where the results for chloride concentrations are shown.
Choosing a single transformation for all ions seemed desirable for simplicity and to enable comparisons, for example between the nonparametric slope estimates in Section 5. So the logarithmic transformation was ultimately adopted for our analysis. Overall, it performed quite well and is a commonly used transformation. In fact the usual pH scale of acidity is obtained through a logarithmic transformation of $H^+$. 

The distribution of the log transformed volume weighted means data within homogeneous clusters have lighter tails than that of the unclustered values.

We will now turn in Section 4 to the analysis of the logarithmically transformed data.
4. MEDIAN POLISH

There were two factors of primary interest in our study, time trends and spatial variations in concentration levels of the various ions at the sites represented in our data. A secondary factor of interest was seasonal variation since the latter, if not factored out, could mask the factors of primary interest.

We explored the log-transformed data by a three way median polish, a method analagous to a three way analysis of variance (ANOVA) except that wherever the latter computes averages, the median polish computes medians. But unlike the average, the median is not a linear functional of the data. Thus to extract the main effects (column, row and layer medians), along with the two-way interaction effects and finally the residuals, the three-way interactions, it is necessary to use an iterative calculation which recycles through the data several times. The details are given in Wu and Zidek (1989h), including the requisite formulae for this interative computation.

The median polish was chosen because it is a well established tool for data exploration and is highly robust, a very desirable feature of the method, since even the transformed data have fairly heavy tails and lots of outliers. The validity of the ANOVA would be in great doubt because the average unlike the median, is susceptible to gross distortions from a few outlying observations. Of course, the median polishing method does not, unlike the ANOVA, have an associated confirmatory (significance testing) theory. But then it is very unlikely that the model assumptions underlying the latter would be met anyway.

In median polishing, we are seeking indications of trend and not attempting to prove that trends exist in any formal sense. Our estimated site effects are likewise just
indications of spatial differences. In the next section by contrast, we do use nonparametric testing theory in a site by site study of trends and these results complement the more global picture of time trends described in this section.

Our results are expressed as graphical displays. Year effects are plotted cluster by cluster in Figures 4.1a-i, those for month effects in Figures 4.2a-i, and the site effects in Figures 4.3a-i.

HYDROGEN. All three clusters show up-trends; during the period 1983 to 1986, the concentration of $H^+$ in wet deposition increased for most of the United States. Monthly concentrations are high from June to September, low from November to April, with a peak in August. There is a similarity of the month effects from cluster-to-cluster except for an anomalous dip in April for Cluster 2. For cluster 3, the stations located in the center of the industrial region tend to have large effects. The largest of these is 056a, Caldwell, Ohio. In constrast, the station with the lowest site effect in the cluster, 2, with the lowest common mean effect, is 278a, Give Out Morgan, Montana.

Overall, the median polish model fitted the data for Cluster 3 best. Although the boxplots of residuals for individual stations could not be included here, those for Stations 035a (Lamberton, Minnesota) and 038a (Mead, Nebraska) are relatively broad indicating that the data from these two stations do not fit the model well. And there were extreme outliers for this ion and all those presented below. These may warrant further study.

SULFATE. The three plots of the yearly effects of the clusters for weighted means reveal similar patterns. Clusters 1 and 2 show up-trends while cluster 3 seems
to indicate a decline. The concentration of $SO_4$ in wet deposition increased for the most industrialized subregion of the United States during the period of 1983 to 1986. Concentrations are high from June to September and low from November to February with a peak in August. The seasonal pattern seems stable over space as it is similar from cluster to cluster. As in the case of $H^+$, the stations in the center of the industrial region (Cluster 1) tend to have larger effects, 021a, Argonne, Illinois and 056a, Caldwell, Ohio sharing the top spot. At the other extreme in Cluster 3, we find Stations 061a, 074a, and 271a, H.J. Andrews, Oregon, Olympic National Park, Washington, and Headquarters, Idaho.

The variation in station effects in Figure 4.4b.3 are relatively large, perhaps because we assigned Stations 070a and 071a to this cluster. In general, the variation in year effects for 1983 to 1986 are small in magnitude and there is no uniformity in direction from site to site.

**NITRATE.** There appears to be a downtrend in the levels of concentration for the sites in Cluster 1 and no obvious trends in the other two clusters. These levels are high from January to August and low from September to December with two jumps around September and December plus a gap around March. The stations located in the middle of the United States (Cluster 1), have large positive effects while the stations located along the West Coast have large negative effects. The effects for stations in Cluster 3, close to the center of the most heavily industrialized region of the United States, tend to be positive, for those in this cluster far from the center of this region, negative. In general, the closer a station is located to the center of the industrial region, the more likely it is to have a large positive effect. At the
positive extreme are Stations 040a (Aurora, New York) and 046a (Bennett Bridge, New York), in Cluster 3, which have about the same level, while at the other extreme is 172a (American Samoa, American Samoa) in Cluster 1.

The variation of the station effects in Cluster 1 is large as a consequence of assigning Stations 035a, 268a and 283a to this cluster. Among the other effects, the variations in year effects and the year-by-station interactions are relatively small. For Cluster 2, the variation in the yearly effects and the interactions of yearly effects with other effects are all small. This suggests that for the stations in this cluster there was little year to year variation in their data patterns.

**SODIUM.** Clusters 1 and 2, containing most of the sites under study, show slight nearly identical down-trends. Overall, for most of the sites in the United States the level of sodium concentration in wet deposition declined or remained flat between 1983 to 1986. The effects for Cluster 3 also show a down-trend. But since the data in this cluster are not homogeneous, a naive interpretation could be misleading. Monthly concentration levels are high from December to April, low from May to August, and change smoothly from month-to-month. The seasonal pattern for Cluster 1 is different from that of Cluster 2 in that former seems to have two very distinct cycles, the peaks being in April and November, and the troughs in August and January, respectively. The seasonal pattern of Cluster 3 differs markedly from those of Clusters 1 and 2. Two sites, American Samoa (0172a), on the island of American Samoa, and NACL, Massachusetts(251a) on the East Coast, have very large positive site effects. These effects are even more striking if we compare the scales of Figures 4.3d.1 – 3. Within Cluster 2, the sites close to the Coast(349a, 281a and 277a) also have the
largest positive effects. In addition, the common or "median-effect" of Cluster 1 is -0.81, significantly higher than those of Clusters 2 and 3, which are -2.68 and -2.48 respectively. These results reflect the fact that some sodium comes from the sea.

For all three clusters, the variation in month effect and site effect was substantially larger than the yearly variation, indicating that the effect of trend is small compared with that of seasonal and and site. The distribution of the residuals from Cluster 2 is slightly left skewed while that of Cluster 1 is right skewed, showing that besides the average level, there is a difference in the distributions of the clusters. The extremely large differences among the site effects is due to their inhomogeneity.

POTASSIUM. There is a persistent downtrend in the levels of potassium concentration in wet precipitation during the period of this study. Monthly effects are highest in April, May and June, November and December the lowest, with a smooth transition from June to November. American Samoa (172a) has the largest site-effect for potassium and the runner up is NACL, Massachusetts (251a). The sites with the smallest and second smallest site effects are Bridgton, Maine (164a) and Greenville Station, Maine (030a). Again, no geographical pattern is revealed except that the one with the highest level is in the middle of the Pacific ocean and most of the those with the lowest of levels are located in the Northeast.

The combined distribution of the residuals is markedly right skewed with heavier than normal tails. The estimated main effects are symmetrically distributed about zero. The year-to-year variation is smaller than that of the seasonal and spatial effects. So, conclusions about trend must be regarded as tentative. The substantial interaction effects too suggests caution in interpreting the main effects.
**CALCIUM.** All three clusters show down-trends. Overall, for most of the sites in the United States the concentration level of calcium in wet deposition declined between 1983 to 1986. For all three clusters, the monthly concentration levels of Ca$^{++}$ tend to be high from March to August and low from October to January with a possible peak around April. For Cluster 1, the four stations with negative site effects, 280a, 011a, 279a and 160a, are located, suggestively, close to each other in Colorado and New Mexico.

The combined variation in month and site effects is larger than that of the year effects indicating that trends, if any are relatively small. The distributions of interactions and residuals for Clusters 1 and 2 are symmetric, but not those for Cluster 3.

**MAGNESIUM.** There seems to have been an increase in the levels of magnesium concentration in wet precipitation from 1983 to 1984 followed by a decrease from 1984 to 1986. The monthly concentration levels of magnesium are high from March to July and low from October to December, the highest level being in April, the lowest level in November with a smooth transition from June to November. American Samoa (172a), NACL, Massachusetts (251a) and Give Out Morgan, Montana (278a) have the highest magnesim site effects followed by Alsea, Oregon (059a) and Victoria, Texas (071a) The sites with the smallest and second smallest site effects are Headquarters, Idaho (271a) and Hubbard Brook, New Hampshire (039a). No apparent geographical pattern is revealed in this figure except that among the 5 sites with the highest effects only one (278a) is located inland and all of the ones with the lowest level are located inland.
The combined distribution of residuals from the median polish is quite symmetric with heavier than normal tails. Each of the estimated effects also has a fairly symmetric distribution if we ignore a few outliers. The site-to-site variation is larger than yearly and seasonal variations and the seasonal patterns may vary from site to site. The variation of the month-site interactions is also large. So, as has been emphasized already conclusions about trend must be regarded as tentative. The substantial interaction effects also suggest caution in interpreting the main effects.

CHLORIDE. There appears to be a more or less persistent downtrend in the levels of chloride concentration in wet precipitation. The corresponding estimated monthly effects derived from the median polish are high in winter and low in summer with a minimum in July or August. The site having the maximum (mean) site effect for magnesium in the last paragraph American Samoa (172a), has also the highest site-effect here. The runner up is NACL, Massachusetts (251a) with Alsea, Oregon (059a) in third place. At the other extreme, the sites with the smallest and second smallest site effects are Headquarters, Idaho (271a) and Glacier National Park, Montana (037a).

The combined distribution of residuals from the median polish is symmetric but has heavier than normal tails. The site effects, unlike most of the other displayed effects (except possibly Yr), are not symmetrically distributed about zero. There is a substantial month-to-month variation which tends to dominate any year-to-year variation so conclusions about trend must be regarded as tentative. The substantial interaction effects also suggests caution in interpreting the main effects.

AMMONIUM. Clusters 2 and 3, comprising most of the sites under study,
show down-trends. Overall, for most of the sites in the United States the concentration level of wet deposition of ammonium declined or remained flat between 1983 to 1986. Monthly concentration levels were high from April to August and low from October to February, with a peak around April and May. For Cluster 2, the stations located away from the center of the industrial region tend to have large negative effects for ammonium. The sites with large effects like 273a, Konza Prairie, Kansas, may warrant further study. For Cluster 3, Alsea, Oregon (059a), H.J.Andrews, Oregon (061a), Olympic National Park, Washington (074a), and American Samoa (172a) have substantial negative effects and these stations are either close to coast or on an island.

The combined distribution of the estimated principal effects and of the residuals from the median polish on weighted means indicate that the variation in month and site effects are larger than that of year effects indicating that trends, if any, are relatively small. The displays for clusters 1 and 2 are similar. The distributions of interactions and residuals for Cluster 3 are quite symmetric unlike those for Clusters 1 and 2 which are not.
5. NONPARAMETRIC TREND ANALYSES

Although "trend" is a much used concept in environmetrics, it is not a well defined concept. It can mean that a data series follows, approximately, a parametric function like a line. In that case the slope would index the magnitude of the trend. At the other extreme, for highly irregular data series like those considered in our study where even the seasonal patterns vary markedly from site to site, a more primitive, nonparametric notion of trend seems called for. A brief survey of nonparametric approaches to defining and analyzing methods is given in Wu and Zidek (1989h). A much more extensive discussion is provided by Gilbert (1987).

The well known Mann-Kendall test (c.f. Wu and Zidek, *ibid.*) interprets, say increasing, trend to mean that a series rises more than it falls. The test uses the sum of the signs, negative and positive, of changes from beginning to end over periods of a given length. A large value of this sum relative to its expected stochastic variation in the absence of trend would be an indication of the presence of a positive trend.

But there is no analogue of the notion of "slope" in the Mann-Kendall interpretation of trend. So we have also computed at each site for each ion, Sen's nonparametric slope statistic. This is a very robust statistic which is the median of the ratios of change from beginning to end over intervals of every (discrete) length in the series to their length. Confidence intervals for this statistic are easily found and these (at the 80% level of confidence) are presented in Tables 5a to i along with the slopes themselves, the values of the Mann Kendall (normalized) test statistic and their associated p-values.

Observe that in each case, a large number of sites have significant trends at the
p=0.2 level of significance, substantially more, as we argue in Section 6, than the 20% of significant sites which would be expected by chance alone.

The results of nonparametric slope estimation are plotted in Figures 5a-i. If the lower 80% confidence bound of the slope estimate is greater than 0 at any site, a "+" (significant up-trend) is plotted; if the corresponding upper confidence bound is less than 0 a "-" (significant down-trend) is plotted. Otherwise, a "0" (no-trend) is plotted.

Notes on the results for individual ions now follow.

**HYDROGEN.** Observe that in Figure 5a, 10 sites, located in the mid Eastern United States, show down-trends. Slightly more than half of the sites show no-trend and the rest, an up-trend. Those sites are dispersed throughout the United States. This confirms the impression given by the median polish that overall there has been up-trend in levels of $H^+$ during the 1983-1986.

**SULFATE.** In Figure 5b, about 1/4 of the sites, mainly in the Eastern states exhibit up-trends, 1/4, mainly in the Western and middle states, down-trends while the rest, distributed throughout the United States, no trends. The patterns in this figure is consistent with the results of median polish since the sites with ",-"'s belong mainly to cluster 3 and those with "+"'s, to clusters 1 and 2. This indicates that although the concentration levels for the stations located in the West are declining, most of those in the East are experiencing either no change or up-trends between 1983 to 1986.

**NITRATE.** In Figure 5c, 12 sites, mainly in the Eastern United States show up-trends while 15 sites mainly in the middle and Western United States show down-
trends. The rest of the sites distributed throughout the United States exhibit no-trends. This suggests that although the trends in concentration for the stations located in the Western United States declined, most of the stations located in the Eastern United States experienced either no change or an up-trend, in broad agreement with the findings of Section 4.

SODIUM. In Figure 5d, about 1/2 of the sites show down-trends and are scattered over the U. S. except the West Coast. Three sites show up-trends and the rest of the sites, which are scattered over most of the United States other than the Middle West, show no trends. If we compare Figure 5d with Figures 2.1d it seems that site-trends do not determine the outcome of clustering as much as their concentration-level.

POTASSIUM. The trends depicted in Figure 5f are predominantly down; there is only 1 "+" that being at Horton's Station, Virginia (073a). The downtrend "-"s are well dispersed over the 48 states.

CALCIUM. In Figure 5f about 2/3 of the sites, quite uniformly distributed across the country, show down-trends and the remaining sites, no trends. A comparison of this Figure with 2.1f, suggests that trends do not contribute much to the observed cluster patterns.

MAGNESIUM. In Figures 5g all the significant trends depicted are down and well dispersed over the 48 states.

CHLORIDE. Here the trends depicted in Figure 5h are predominantly down; there are just 4 "+"s at Bull Run, Oregon (281a), Walker Branch, Tennessee (171a), Wye, Maryland (350a) and Parsons, West Virginia (075a). The downtrend "-"s are
well dispersed excluding perhaps just a small region in the middle Eastern states.

AMMONIUM. In Figure 5i about 1/3 of the sites show down-trends and are located mainly in the Western and central parts of the USA. Except for 4 sites in the Eastern States which show up-trends, the remaining sites scattered throughout the United States show no trends. Comparing Figures 5i and 2.1i suggests that site trends do not contribute as much to the observed cluster patterns as level of concentration, probably the most important determinant of the cluster patterns.
6. DISCUSSION AND TENTATIVE CONCLUSIONS

An investigation, much of which is described in this report has examined in parallel, monthly medians and volume weighted means of the concentrations in wet deposition over the United States of nine selected ions during the 1980–86 period. Our results add to the rapidly growing body of information about recent global deposition patterns. A comprehensive review is not feasible here but we would single out a number of very recent papers to give an indication of the nature of some of the complementary studies which have been completed.

Sanden, Grimvall and Lohm (1987) apply variations of the nonparametric methods of Section 5 to analyse precipitation chemistry over the 1969–82 period in relation to the acidification of surface waters of Sweden. Among their many finding are significant declines in sulfate ion concentrations and in pH in one of the three study areas in each case.

Zemba, Golomb, and Fay (1988), like the authors of the present paper, rely on the ADS database for their analysis of spatial pattern over the 1982–85 period. They look at the data from 77 stations concentrated in the Eastern part of North America. However, they base their descriptive analysis on averages of the four values of various aggregate statistics like seasonal average, obtained over the four years of observation. Thus they do not attempt to determine temporal trends if any.

Ro, Tang, Chan, Kirk, Reid and Lusis (1988) examine sulfur and nitrogen deposition in Ontario for the 1980–85 period. Relying on informal comparisons among various data averages, the authors compare levels of wet and dry deposition levels as well as levels of sulfur and nitrogen over that period.
A comprehensive study of the deposition of acidifying components over the Netherlands during the 1980–86 period, is the subject of the paper of Ensman, de Leeuw and van Asløt (1989). Wet and dry deposition levels are investigated along with sources and trends.

Spatial patterns over North America of pH and mean $H^+$ concentrations during 1981–82 are the subject of the paper of Wisniewski and Kinsman (1988). They rely on the data provided by a large number (130) of stations from various networks and present their findings as contour plots.

Finally we would cite the paper of Parungo, Nagamot, Hoyt and Bravo (1990) who give the results derived from data gathered during a research cruise of the Gulf of Mexico in the summer of 1986. Their study unlike those cited above looks at all the ions we consider in this report (except for hydrogen) in aerosol as well as rain samples but, of course, over a very limited time scale.

In our study monthly rather than say, weekly values were selected as the basis of our investigation to avoid significant numbers of missing data values and hence ad hoc, potentially controversial adaptations of conventional statistic techniques to allow for these missing values. The cost has been the loss of some information in the so-called "raw" data. The "optimum" level of aggregation of the raw data has not been explored, however; the choice of monthly values was a somewhat arbitrary but convenient and apparently satisfactory choice.

It would be expected that aggregating data would have the side benefit of yielding sample distributions with greater symmetry and lighter tails than that of the raw data. We have not systematically assessed the degree to which this side benefit was
realized. But as indicated in Section 3, the sample distributions of the monthly values are highly asymmetric even after their logarithms have been taken. As well, there are a substantial number of outlying monthly values and the tails of the corresponding sample distributions are quite heavy.

Means and medians were analyzed in parallel in our study primarily because it did not seem to the authors that either of these very different indicators of the monthly level, in wet deposition, of the concentration of any given ion could be chosen over the other; ultimately the choice would seem to depend on substantive rather than statistical issues. To have made an arbitrary choice of one of them might well have invited controversy and weakened the results of the investigation. Finally, it was thought that the two competing analyses would serve as useful checks on one another and on any conclusions which might derive from them.

Although there is broad agreement between the results of the two analyses, notable and interesting differences obtain in special cases and point to the need for further investigation.

Another duality in our study, not in this report, is in the separate analyses of the 80 - 86 data and the 83 - 86 data. The latter involves far more participating sites than the former, but the former covers a longer time span and hence gives a broader historical perspective. Overall the results presented here for the shorter, recent period strongly confirm those for the longer one. In particular, as noted by Wu (1988) there has been a rise in the levels of concentrations of $H^+$, $SO_4^{2-}$, and $NO_3^-$ since 1983 after an apparent decline between 1980 and 1983 over the subregion of highest industrial concentration. This conclusion, perhaps the most striking result
of this investigation, must be viewed with caution. While median polishing is a
great, robust exploratory approach, the indication of a recent overall upward trend
is just that – it is not a conclusion based on testing a statistical hypothesis of no
upward trend during that period. Furthermore, as indicated in Section 2, year-to-
year variations in concentration levels are smaller than naturally occurring month to
month variations. While small year to year trends may have very negative long term
impacts, they are extremely difficult to detect and indications of recent overall trends
like those observed in Section 4 are more suggestive of the need for close monitoring in
the future rather of the need to sound “alarm bells.” Further caution in interpreting
our findings is suggested by the substantial interactions among the main factors in
our analysis seen in Section 4. Of course, overall patterns can be misleading in special
cases.

In summary, there has been a rise in concentrations of HYDROGEN since 1983
after an apparent decline between 1980 and 1983. The same is true for both SUL-
FATE and NITRATE over the subregion of highest industrial concentration while over
the rest of the United States concentrations of these two ions seem to have levelled
off. Levels of SODIUM, POTASSIUM, CALCIUM, MAGNESIUM, and CHLORIDE
seemed to be decreasing while those of AMMONIUM have levelled off.

For the 1983 to 1986 data a substantial number of significant trends were observed
at individual sites. Among 81 (or in the case of \( H^+ \), 86) sites it would be expected that
by chance alone there would be something like \( 0.2 \times 80 = 16 \) significant slopes. In fact,
as noted in Section 5, on the order of 30 to 40 sites exhibited significant trends for the
various ions. If the 81 tests were independent (they are not), the standard deviation in
the number of sites found to have significant slopes would be about 4. Since the test statistics seem likely to be positively correlated over fairly broad regions, 4 might well be a conservatively large value for the true SD. Thus 40 would seem to be significantly larger than the 16 rejections of no trend predicted by chance.

Our hierarchical clustering analysis suggests that for all the ions except $K^+$, $Mg^{++}$, and $Cl^-$ for which there seems to be no natural subdivisions of the monitoring sites, there are two or three contiguous regions of the United States wherein the history of monthly contamination levels are somewhat similar. The choice of the number of clusters is somewhat arbitrary as a review of the cluster diagrams in Section 2 shows. However, there does not appear to be another distinct break in these cluster diagrams at levels of refinement higher than the numbers chosen in Section 2.

It had been hoped that the data distributions within clusters would be more-or-less Gaussian in character. Unfortunately this not even approximately the case and, in fact, the authors ultimately chose a logarithmic transformation of the data to achieve a reasonable shape for the data distribution. Gaussianity itself is not critical to out analysis. The nonparametric and exploratory methods we use do not depend on such a strong assumption. But transforming the data does make distributional patterns easier to see in displays and may well improve the quality of the asymptotic normal approximations used in our analysis.

Although the models developed by median polishing leave outliers, we believe it provides a useful data summary. For some stations the median polish fit is excellent. But as seen in Section 4, there are a substantial number of outliers. Learning their cause might well yield interesting insight into the processes involved in the production
and delivery of airborne contaminations.

There is a marked seasonal pattern in the levels of the ions in our study. For ions like $SO_4$ where there are more than one cluster, there is a reassuring similarity in these patterns among the clusters.

ACKNOWLEDGEMENT.

The idea of using median polish in the analyses of environmental data was suggested by Dr. Peter Guttorm 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 are indebted to Patricia L. Zidek for her invaluable assistance in researching the large background literature in this subject. And we thank the members of the SIMS group of investigators and various individuals from the EPA for their suggestions and encouragement.
BIBLIOGRAPHY


NADP. (1982). National Acid Deposition Program Instruction Manual, Site Operation. natural Resource Ecology Laboratory, Colorado State University, Ft., Collins, CO.


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


No.153, Department of Statistics, University of British Columbia.


Figure 2.1a. Cluster Diagram for the Sites Using H+ Concentrations.
Figure 2.1b. Cluster Diagram for the Sites Using SO4 Concentrations.
Figure 2.1c. Cluster Diagram for the Sites Using NO3 Concentrations.
Figure 2.1e. Cluster Diagram for the Sites Using K Concentrations.
Figure 2.1f. Cluster Diagram for the Sites Using Ca Concentrations.
Figure 2.1g. Cluster Diagram for the Sites Using Mg Concentrations.
Figure 2.1h. Cluster Diagram for the Sites Using Cl Concentrations.
Figure 2.11. Cluster Diagram for the Sites Using NH4 Concentrations.
Figure 2.2a. Labels of Sites by Clusters Based on H+ Concentrations.
Figure 2.2b. Labels of Sites by Clusters Based on SO₄ Concentrations.
Figure 2.2d. Labels of Sites by Clusters Based on Na Concentrations.
Figure 3.1. Histograms of Transformed SO₄.
Figure 3.2. Histograms of Transformed CI.
Figure 4.1a. Yearly Effects of log(H+) for the 3 Clusters.
Figure 4.1b. Yearly Effects of log(SO4) for the 3 Clusters.
Figure 4.1c. Yearly Effects of log(NO3) for the 3 Clusters.
Figure 4.1d. Yearly Effects of log(Na) for the 3 Clusters.
Figure 4.1e. Yearly Effects of log(K).
Figure 4.1f. Yearly Effects of log(Ca) for the 3 Clusters.
Figure 4.1g. Yearly Effects of log(Mg).
Figure 4.1h. Yearly Effects of log(CI).
Figure 4.1i. Yearly Effects of log(NH4) for the 3 Clusters.
Figure 4.2a. Monthly Effects of log(H+) for the 3 Clusters.
Figure 4.2b. Monthly Effects of log(SO4) for the 3 Clusters.
Figure 4.2c. Monthly Effects of log(NO₃) for the 3 Clusters.
Figure 4.2d. Monthly Effects of log(Na) for the 3 Clusters.
Figure 4.2e. Monthly Effects of log(K).
Figure 4.2f. Monthly Effects of log(Ca) for the 3 Clusters.
Figure 4.2g. Monthly Effects of log(Mg).
Figure 4.2h. Monthly Effects of log(CI).
Figure 4.2i. Monthly Effects of log(NH4) for the 3 Clusters.
Figure 4.3a. Site Effects for Hydrogen Ions.
Figure 4.3b. Site Effects for Sulfate.
Figure 4.3c. Site Effects for Nitrate.
Figure 4.3d. Site Effects for Sodium.
Figure 4.3e. Site Effects for Potassium.
Figure 4.3f. Site Effects for Calcium.
Figure 4.3h. Site Effects for Chloride.
Figure 4.3i. Site Effects for Ammonium.
Figure 5a. Significant Trends for log(H) from 1983 to 1986 at the 86 Stations in Our Study.
Figure 5b. Significant Trends for log(SO4) from 1983 to 1986 at the 81 Stations in Our Study.
Figure 5c. Significant Trends for log(NO3) from 1983 to 1986 at the 81 Stations in Our Study.
Figure 5d. Significant Trends for log(Na) from 1983 to 1986 at the 81 Stations in Our Study.
Figure 5e. Significant Trends for log(K) from 1983 to 1986 at the 81 Stations in Our Study.
Figure 5f. Significant Trends for log(Ca) from 1983 to 1986 at the 81 Stations in Our Study.
Figure 5g. Significant Trends for log(Mg) from 1983 to 1986 at the 81 Stations in Our Study.
Figure 5h. Significant Trends for log(Cl) from 1983 to 1986 at the 81 Stations in Our Study.
Figure 5i. Significant Trends for log(NH4) from 1983 to 1986 at the 81 Stations in Our Study.