NISS

Meteorologically-dependent Trends in Urban Ozone

Li-Shan Huang and Richard L. Smith

Technical Report Number 72
November, 1997

National Institute of Statistical Sciences
19 T. W. Alexander Drive
PO Box 14006
Research Triangle Park, NC 27709-4006
www.niss.org

Meteorologically-dependent Trends in Urban Ozone

Li-Shan Huang
huang@stat.fsu.edu
Department of Statistics
Florida State University
Tallahassee, FL 32306-4330, USA

Richard L. Smith
rs@stat.unc.edu
Department of Statistics
University of North Carolina
Chapel Hill, NC 27599-3260, USA

National Institute of Statistical Sciences
Research Triangle Park, NC 27709, USA

This research was supported in part by the U.S. Environmental Protection Agency under Cooperative Agreement #CR819638-01-0 at the National Institute of Statistical Sciences. Li-Shan Huang's work was carried out in part while a graduate student at the University of North Carolina in Chapel Hill, and was subsequently supported by a First-Year Assistant Professor Research Award of Florida State University.

Meteorologically-dependent Trends in Urban Ozone

Summary

Ozone concentrations are affected by precursor emissions and by meteorological conditions. As part of a broad study to assess the effects of standards imposed by the U.S. Environmental Protection Agency (EPA), it is of interest to analyze trends in ozone after adjusting for meteorological influences. Previous papers have studied this problem for ozone data from Chicago, using a variety of regression techniques. This paper presents a different approach, in which the meteorological influence is treated nonlinearly through a regression tree. A particular advantage of this approach is that it allows us to consider different trends within the clusters produced by the regression tree analysis. The variability of trend estimates between clusters is reduced by applying an empirical Bayes adjustment. The results show that there is downward overall trend throughout the process, but this trend is stronger at higher levels of ozone.

Key words: ANOVA; Empirical Bayes; Regression tree.

1 Introduction

Urban ozone arises as a consequence of the emissions of nitrous oxides and hydrocarbons into the atmosphere, but it is also heavily dependent on meteorological conditions. For example, it is known that the daily ozone concentrations are related to temperature. In analyzing ozone data, one aim is to separate genuine trends, that might be explained by changes in the levels of emissions, from the effects of meteorological fluctuations. This problem has been discussed in Chapter 2 of the report of the National Research Council (1991). Three major approaches were mentioned, including the "classification" techniques. Particularly, Stoeckenius (1990) used the CART (classification and regression trees, Breiman *et al.* 1984) method, to study the pattern of meteorological conditions on ozone-conducive days, and successfully highlighted the results of unusual meteorological conditions in the high-ozone years (1983, 1987, and 1988) for data from the Philadelphia and southwestern Connecticut networks. Segal (1994) proposed the use of regression trees to study representative curves for longitudinal data with an illustration of atmospheric ozone data.

Regression tree methods are useful when there are a large number of explanatory variables and a complex relationship is expected between the response variable and covariates. In these cases, tree methods may be more adept at capturing non-additive behavior and sometimes easier to interpret and discuss than linear models. The CART algorithm and the `tree` functions in Splus (see Clark and Pregibon 1990) have made tree methods a popular tool in applications, e.g. LeBlanc and Crowley (1993) in analysis of censored survival data and Elsner *et al.* (1996) for classifying the meteorological conditions of hurricanes. In the present paper, based on the tree regression approach, a different analysis is given of the problem of trends for the Chicago ozone data 1981-91. The same problem was addressed in Bloomfield, Royle, Steinberg, and Yang (1996), using nonlinear ozone modeling while separating out the meteorological effects, Smith and Huang (1993), on characterizing the probability of exceeding a high threshold, and Gao, Sacks, and Welch (1996) based on semi-parametric modeling. For the problem of trends, Bloomfield *et al.* (1996) and Smith and Huang (1993) built the trends in the models with a covariate "YEAR". Gao *et al.* (1996) considered trends for four time periods, May 15 - June 15, June 15 - July 15, July 15 - Aug 15, and Aug 15 - Sep 15, and concluded that there are significant downward ozone trends for the periods of June 15- July 15 and Aug 15 - Sep 15, and no significance for the periods of May 15 - June 15 and July 15- Aug 15.

As mentioned in Smith (1989), besides the meteorological effects, the trends may not be the same at all levels of the process, and hence it may be more reasonable to assume different trends at different levels of the ozone concentrations after adjusting for meteorological effects. Tree modeling is an attempt to analyze the trends within different clusters of the data. The clusters are formed by recursively partitioning the data space into two groups (growing a tree), such that the two groups are the most different from each other, until the difference is no longer significant, and then finding the right size of tree as in assessing the bias and variance trade-off for common regression problems (pruning the tree). Using the meteorological variables as covariates and ozone level as the response variable, the regression tree method could identify the pattern of meteorological conditions at different ozone levels. Based on the selected tree, we then analyze the data by a variety of different ANOVA models which contain a trend parameter for each cluster and the cluster effect (the overall meteorological effects). An empirical Bayes adjustment is used to reduce the variability of the estimated trend coefficients. Our results confirm a downward trend, and moreover the decline is stronger as ozone level increases. The last point has not been made in previous analyses of the Chicago data.

The paper is organized as follows. Section 2 gives the background of the tree method and its implementation in Splus. Section 3 presents an analysis with regression tree, ANOVA models, and the empirical Bayes adjustment. Concluding remarks are given in Section 4.

2 Outline of Regression Tree Approach

We describe in this section the basic idea of tree models in Splus. Details can be found in Clark and Pregibon (1990).

Let Y denote the response variable and $\{X_1, \ldots, X_J\}$ be the explanatory variables. The tree method is to partition the data into homogeneous subsets until it is infeasible to continue (growing a tree), and then to "prune" the tree so that it can be simplified without sacrificing goodness of fit.

2.1 Growing a tree

Assume that Y given $\{X_1, \ldots, X_J\}$ has a normal distribution with non-constant mean (may depend on $X_1, \ldots, X_J$) and common variance σ^2. The loop to grow a tree is as follows. For each $j =$

$1, \ldots, J$, the regressors $\mathbf{X} = X_1 \otimes \ldots \otimes X_J$ are partitioned into two sets $\mathbf{X}_{left}$ and $\mathbf{X}_{right}$ according to the values of X_j, such that the conditional densities $f(y|\mathbf{X}_{left})$ and $f(y|\mathbf{X}_{right})$ are the most different (measured by deviance D). The split that maximizes (over j)

$$\triangle D = D_{parents} - (D_{leftchild} + D_{rightchild})$$

is the split at a given node, where $D_{parents}$ denotes the deviance of the upper node and $D_{leftchild}$ and $D_{rightchild}$ are the deviances of the two split sets. The partitioning will stop if the number of observations in a cluster is less than or equal to 5, or if $\triangle D$ is less than 1% of the $D_{parents}$. Clearly, as partition continues, we may over-fit the tree model.

2.2 Pruning the tree

As in most statistical problems, balancing the bias and variance implies the problem of how to choose a right size tree. A cost-complexity measure is used to assess goodness of fit,

$$D_\alpha(T') = D(T') + \alpha \text{ size}(T'),$$

where $D(T')$ is the deviance of the subtree T', size(T') is the number of terminal nodes of T', and α is the cost-complexity parameter. For a specified α, we can find a tree object or a nested sequence of subtrees that minimizes the cost-complexity measure.

The data that were used to construct the tree should not be used again to evaluate the goodness of fit. Hence a cross-validation approach may be helpful here. First, the original data set is divided into ten (the default) mutually exclusive sets. For each set, a tree is grown to the remaining sets and a subtree sequence is obtained; the set left out is then used to evaluate the sequence. Deviances from each set are accumulated as a function of α. We will find an optimal tree by looking at the plot of deviance vs. size and then select the tree with the smallest or nearly the smallest deviance. The average of the responses falling in a cluster of the optimal tree is used to estimate the cluster mean.

As noted by Ripley (1995), Clark and Pregibon have confirmed that there are some bugs in the current Splus `tree` functions. Ripley has written some S-functions to replace the existing `prune.tree` and `predict.tree` functions. The new functions correct the errors and enhance functionality, and are available in Stat Lib (http://lib.stat.cmu.edu/S/).

3 Analysis

The ozone data consist of hourly averages at 45 stations in the Chicago area. Detailed description of the stations and their locations can be found in Bloomfield *et al.* (1996). For the present study, it is of interest to focus on one variable, e.g. daily maxima for the whole network. However, network maxima are not easy to define since there are many missing values, and different stations have been in operation for different periods of time. Bloomfield *et al.* (1996) applied median polish kriging to interpolate missing values within a subset of 16 stations and then constructed network maxima based on those 16 stations. We use the same data set here. The ozone levels are given for 214 days (April 1 - October 31) for 11 years (1981-91).

The meteorological variables, measured on an hourly basis, are

TOTCOV	total cloud cover %
OPCOV	opaque cloud cover %
CHT	ceiling height m.
PR	barometric pressure mb.
T	temperature, ^oF
TD	dewpoint temperature ^oF
RH	relative humidity %
Q	specific humidity g./kg.
VIS	visibility km.
WSPD	wind speed m./sec
WDIR	wind direction from north

Among these, TOTCOV and CHT were omitted, as discussed by Bloomfield, *et al.* (1996). The noon values are taken as daily representative values to reduce the large volume of hourly data. Motivated by the results in Bloomfield *et al.* (1996), some additional variables were created. In particular, WDIR is replaced by WIND.U and WIND.V.

YEAR	since the beginning of data set (= 1 for 1981)
WIND.U	$-$ WSPD$\times \sin(2\pi$ WDIR $/360)$
WIND.V	$-$ WSPD$\times \cos(2\pi$ WDIR $/360)$

DAY the day within the year, = 1 for January 1, etc.

T2 $(\mathrm{T}-60)^2/10$

3.1 Tree model

Our first attempt is to apply the tree method to find a best tree as described in Section 2 with all the covariates except YEAR, since the aim now is to find the most influential meteorological variables and identify the seasonal effect (DAY). The tree is shown in Figure 1, containing 15 clusters. The cluster index is assigned 1 through 15 from the left to the right. Some observations can be made from Figure 1. Cluster 12 has a very high average ozone level (196.40 ppb) containing only 5 observations, which are June 23, 1983 and July 5-8, 1988 with daily maximum ozone 188, 170, 186, 223, and 215 respectively. The corresponding meteorological conditions for cluster 12 are $T > 88.5$ (90, 96, 96 94, and 95), $WSPD < 4.5$ (3, 4, 2, 3, and 3), $DAY < 196.5$ (174, 187, 188, 189, and 190), and $PR > 1019.54$ (1019.88, 1022.50, 1022,50, 1022,50, and 1020.75). Clearly this is a very unusual set of conditions, since it occurs on only 5 out of 2354 days. Cluster 11 gives the second highest average ozone (143.60 ppb) of 18 observations with the same meteorological conditions as cluster 12 except $PR < 1019.54$. A more careful examination of cluster 11 reveals that there are 16 observations with high ozone levels ranging from 120 to 180 ppb, and 2 observations with only 81 and 94 ppb daily maxima on July 8, 1990, and July 1, 1991. Over the whole data set, there are 143 days with ozone levels ≥ 120 ppb, spread over the different clusters as follows:

Cluster Index	1	2	3	4	5	6	7	
# of observations	315	191	184	230	292	285	95	
# (>= 120 ppb)	0	2	0	2	8	2	3	
% (>= 120 ppb)	0%	1.05%	0%	.87%	2.74 %	.70%	3.16%	
Cluster	8	9	10	11	12	13	14	15
# of observations	148	309	77	18	5	33	62	110
# (>= 120 ppb)	27	11	28	16	5	18	15	7
% (>= 120 ppb)	18.24%	3.56%	36.36%	88.89%	100%	54.55%	24.19%	6.36%

This makes it clear that different meteorological conditions have different probabilities for exceeding the 120 ppb standard. Another noticeable point of the tree is that among all years, 1988 has the lowest number of days falling in clusters 1, 4, 7, and 9 (low levels) and the highest number in clusters 12, 13, and 14 (high levels). Since each cluster corresponds to different meteorological conditions, we may conclude that 1988 has a unusually large number of days that are high-ozone conducive.

The optimal tree has a R^2 (coefficient of multiple determination) 0.6137, which is not much different from .64 and .67 in Smith and Huang (1993) and Bloomfield *et al.* (1996) respectively. Note that they have included YEAR as a covariate while the tree model is fitted excluding the YEAR variable. We further look at the trend within each cluster. Figure 2 gives the ozone levels plotted in time order for each cluster. The solid line indicates the estimated mean from the tree output, and the dash line is the 120 ppb threshold. In several clusters, a downward trend is clearly visible, but it is also apparent that the size of the trend is different over clusters. We now consider more systematic approaches to modeling the trend.

3.2 ANOVA

After partitioning the data into different clusters, we further consider models including the YEAR covariate to assess significance of trends. Let y_{ijk} denote the k'th observation in cluster j and in YEAR i. All the models we consider, with the specific point of looking at differences in trends among different clusters, have the structure

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk}, \qquad i = 1, \ldots 11, j = 1, \ldots 15, k = 1, \ldots n_{ij}, \tag{1}$$

where $\varepsilon_{ijk} \sim N(0, \sigma^2)$ are independent error variables and n_{ij} is the number of observations fall in cluster j and YEAR i. Within (1), consider a sequence of models:

(i) Unconstrained μ_{ij}: this can be fitted as a top-level model against which to test all the others.

(ii) No-interaction model:

$$\mu_{ij} = \alpha_j + \gamma_i, \qquad \text{subject to } \sum_i \gamma_i = 0. \tag{2}$$

This model assumes different means for each cluster and a common trend of YEAR i for all clusters without restricting to be a fixed functional form such as a linear model in the following.

(iii) Linear trends for all clusters:

$$\mu_{ij} = \alpha_j + (i-6)\beta, \tag{3}$$

where $(i-6)$ is used to center about its mean value. Now, extending (3) by allowing different trends corresponding to different levels of ozone gives model (4).

(iv) Separate linear trends within each cluster:

$$\mu_{ij} = \alpha_j + (i-6)\beta_j. \tag{4}$$

For artificial comparison, we include two more models:

(v) The tree model:

$$\mu_{ij} = \alpha_j. \tag{5}$$

(vi) Null model:

$$\mu_{ij} = c_0. \tag{6}$$

An ANOVA table of models (1) through (6) is given below.

Model	RSS	df
1	571957	2206
2	643223	2329
3	658476	2338
4	649422	2324
5	669152	2339
6	1732422	2353

The following gives the results on testing goodness of fit of the models at a 5% level.

H_0 vs. H_1	F-value	Pr(F)	
Model (6) vs. Model (5)	265.4736	≈ 0	rejection
Model (5) vs. Model (3)	37.90574	≈ 0	rejection
Model (3) vs. Model (2)	6.136548	≈ 0	rejection
Model (3) vs. Model (4)	2.314263	.00368	rejection
Model (4) vs. Model (1)	2.531995	≈ 0	rejection
Model (2) vs. Model (1)	2.234676	≈ 0	rejection

The first point to note is that models (5) and (6) both fit much less well than the other four. However, it is not so clear which of models (1) to (4) is best. Model (3) is nested within both (2) and (4), and is rejected in favor of either of them by the F-test. Model (2) is slightly superior to model (4) as judged by the R^2 values of .6287 and .6251 respectively, but either is rejected by the F test in favor of model (1). On the other hand the latter model, with its unrelated μ_{ij} coefficients, is not easy to interpret in terms of an overall trend. Figure 3 gives estimated values for γ_i under model (2), for $i = 1, ..., 11$. They show a steadily decreasing trend except the years 1984, 1987, and 1988. It is tempting to fit a linear trend to these values, even though the F test shows them significantly different from a linear trend, as a means of measuring the overall effect of decreasing ozone over the 11-year time span.

In view of these considerations we have decided to focus on model (4) for our subsequent analysis. In this model a separate linear trend with slope β_j is fitted to each cluster. The model is readily interpretable because differences among the β_j's correspond to differences in trend according to different meteorological conditions. Figure 4(a) plots $\hat{\beta}_j$ against $\hat{\alpha}_j$, together with the 95% confidence interval for each $\hat{\beta}_j$, computed from model (4). Nearly all the $\hat{\beta}_j$ estimates are negative, confirming that there is an overall downward trend after adjusting for meteorology. This much is only confirming the conclusions reached by other authors, in particular, Bloomfield *et al.* (1996). However it also appears from Figure 4(a) that $\hat{\beta}_j$ is decreasing as a function of $\hat{\alpha}_j$ — in other words, as the mean level of ozone increases, the downward trend becomes stronger. This appears to be a new interpretation of these data. The picture is complicated by cluster 12 which has the highest $\hat{\alpha}_j$ and also the highest (positive) $\hat{\beta}_j$, but on the other hand this is based on just 5 days' data and the standard error of $\hat{\beta}_{12}$ is very large. This suggests we should try to adjust the plotted values to reduce the standard errors. The most obvious way to do this is through an empirical Bayes analysis, to which we turn next.

3.3 Empirical Bayes adjustment

Let $\hat{\theta}_j = (\hat{\alpha}_j, \hat{\beta}_j)^T$ and correspondingly denote $\theta_j = (\alpha_j, \beta_j)^T$ for $j = 1, ..., 15$. Assume θ_j's are independent observations from a bivariate normal distribution with mean vector μ and covariance matrix S_0, and $(\hat{\theta}_j - \theta_j)$ is independent of θ_j and is normal with mean vector 0 and covariance matrix S_j. In other words, $(\theta_j^T, \hat{\theta}_j^T)^T$ is modeled as a joint normal random variable with mean $(\mu^T, \mu^T)^T$ and covariance matrix partitioned as

$$\begin{pmatrix} S_0 & S_0 \\ S_0 & S_0 + S_j \end{pmatrix}.$$

Here S_j is assumed to be known and equal to the sample covariance matrix $\hat{S}_j$ obtained from Model (4). What we are interested is the conditional mean of θ_j given $\hat{\theta}_j$:

$$\mu + S_0(S_0 + \hat{S}_j)^{-1}(\hat{\theta}_j - \mu).$$

For estimating S_0, since $E\{\sum_{j=1}^{15}(\hat{\theta}_j - \bar{\hat{\theta}})(\hat{\theta}_j - \bar{\hat{\theta}})^T\} = 14\left(S_0 + \frac{1}{15}\sum_{j=1}^{15}\hat{S}_j\right)$ with $\bar{\hat{\theta}} = \sum_{j=1}^{15}\hat{\theta}_j/15$, an unbiased estimator $\hat{S}_0$ of S_0 is

$$\frac{1}{14}\sum_{j=1}^{15}(\hat{\theta}_j - \bar{\hat{\theta}})(\hat{\theta}_j - \bar{\hat{\theta}})^T - \frac{1}{15}\sum_{j=1}^{15}\hat{S}_j.$$

An estimate of μ may be formed by a weighted average of $\hat{\theta}_j$'s with weights proportional to the reciprocal of its variance (diagonal components of $\hat{S}_0 + \hat{S}_j$), or simply $\bar{\hat{\theta}}$. It turns out that the two estimates are nearly the same: $\bar{\hat{\theta}} = (92.9533, -1.0111)^T$ and the weighted average is $(92.6359, -.9831)^T$. Therefore $\bar{\hat{\theta}}$ is adopted, and an "shrinkage" estimator for the conditional mean is

$$\hat{\hat{\theta}} = \bar{\hat{\theta}} + \hat{S}_0(\hat{S}_0 + \hat{S}_j)^{-1}(\hat{\theta}_j - \bar{\hat{\theta}}).$$

Similarly one can obtain the estimated conditional covariance matrix $\hat{S}_0 - \hat{S}_0(\hat{S}_0 + \hat{S}_j)^{-1}\hat{S}_0$. The results are summarized in Figure 4(b) using a 95% confidence interval. The adjusted $\hat{\hat{\beta}}_{12}$ is negative (-1.0555) but it still stands out when looking at the confidence interval. An explanation is that even empirical Bayes analyses are affected by outliers. In this case, it seems obvious that the positive $\hat{\beta}_{12}$ is meaningless, considering that one observation of the cluster is from 1983 and the other four from 1988. Given that cluster 12 appears to be an outlier, the empirical Bayes analysis was repeated with cluster 12 omitted altogether. The resulting analysis is also plotted on Figure 4(b). This makes little difference to the estimates and standard errors for most of the β_j's, but in the case of β_{11} and β_{13}, both the estimate and standard error are changed somewhat, increasing the apparent evidence for an overall downward trend in β_j against α_j.

The value of $10 \times \frac{\hat{\beta}_j}{\hat{\alpha}_j}$ is an estimate for the percentage of decrease per decade for j-th cluster subject to the significance of $\hat{\beta}_j$. To give an overall percentage of decrease, the weighted average of $10 \times \frac{\hat{\beta}_j}{\hat{\alpha}_j}$ is -9.55%/decade with standard error 1.73%/decade, where weights are proportional to the number of observations in the clusters. Similarly the weighted average based on empirical Bayes estimates $10 \times \frac{\hat{\hat{\beta}}_j}{\hat{\hat{\alpha}}_j}$ is -9.66% with standard error 1.65% and -9.33% with standard error 1.56% if excluding cluster 12. The analysis of Bloomfield *et al.* (1996) gives -9.5% per decade with standard error 1.8%/decade. Results by Gao *et al.* (1996) are -18.9%/decade for June 15 - July 15 with standard error 3.4%/decade and -9.2% for Aug 15- Sep 15 with standard error 3.5%. All of the results support the significance of a downward ozone trend.

4 Summary and Conclusion

As a method of describing the influence of meteorology urban ozone, regression trees form a viable alternative to continuous nonlinear regression models. The goodness of fit, as measured by R^2, is comparable with those found in previous studies of the same data, but the results are simpler to describe. The main advantage that has been highlighted of the regression tree approach is that it allows us to estimate different trends in different clusters. Although a linear trend does not represent the best-fitting trend as determined by ANOVA comparisons, it is the simplest kind of trend to interpret and so has been adopted for subsequent comparisons. The result shows that apart from one cluster which appears to be a clear outlier, the downward trend is greater at higher ozone levels. This result is sharpened by using an empirical Bayes approach to adjust the trend estimates.

The results show the effectiveness of the EPA standards, which are aimed particularly at curbing exceedances of ozone at high levels. After adjusting for meteorology, there is a clear downward trend at all levels, but particularly at the higher levels of ozone. However, the city of Chicago was judged to be out of compliance with the EPA standards, because several monitors showed more than three exceedances of the 120 ppb standard during July 1988 (the period around the outlying cluster 12). The analysis of the present paper reinforces the conclusion that this was due solely to extreme meteorological conditions and does not reflect the clear improvements in overall ozone conditions which occurred in Chicago during the 1980s.

References

Bloomfield, P., Royle, J.A., Steinberg, L.J. and Yang, Q. (1996). 'Accounting for meteorological effects in measuring urban ozone levels and trends'. *Atmospheric Environment* **30**, 3067-3077.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). *Classification and regression trees.* London: Chapman and Hall.

Clark, L.A. and Pregibon, D. (1990). Tree-based models. In *Statistical Models in S*, eds. J.M. Chambers and T.J. Hastie. Pacific Grove, Calif. : Wadsworth & Brooks/Cole, 377-419.

Elsner, J.B., Lehmiller, G.S. and Kimberlain, T.B. (1996). 'Objective classification of Atlantic hurricanes'. *Journal of Climate* 11, 2880-2889.

Gao, F., Sacks, J. and Welch, W.J. (1996). 'Predicting urban ozone levels and trends with semiparametric modeling'. *Journal of Agricultural, Biological, and Environmental Statistics* **1**, 404-425.

LeBlanc, M. and Crowley, J. (1993). 'Survival trees by goodness of split'. *Journal of the American Statistical Association* **88**, 457-467.

National Research Council (1991). *Rethinking the ozone problem in urban and regional air pollution.* Washington, D.C. : National Academy Service.

Ripley, B. (1995). Treefix: A shar archive of functions to replace prune.tree and predict.tree. `http://lib.stat.cmu.edu/S/`.

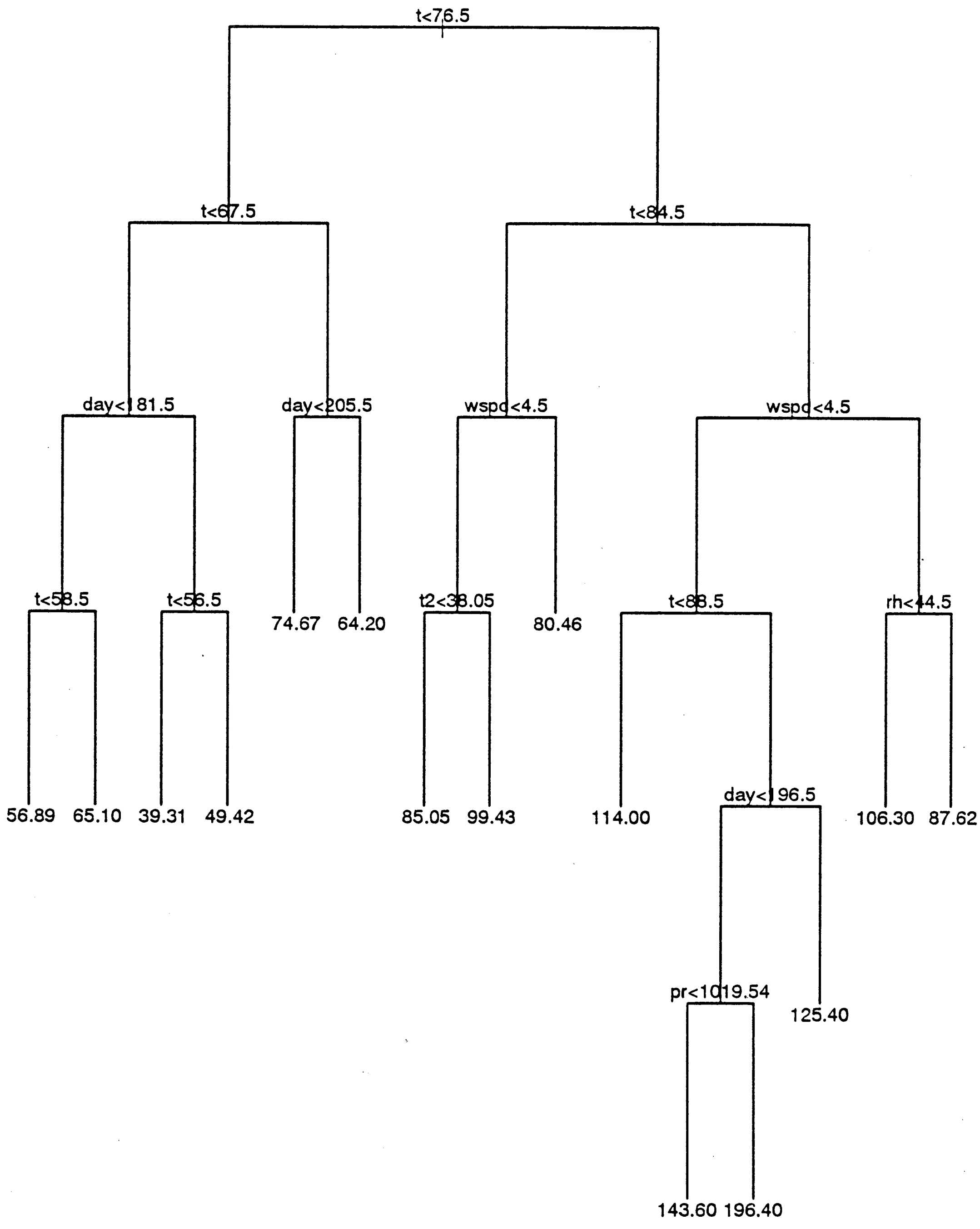
Segal, M.R. (1994). 'Representative curves for longitudinal data via regression trees'. *Journal of Computational and Graphical Statistics* **3**, 214-233.

Smith, R.L. (1989). 'Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone' (with discussion). *Statistical Science* 4, 367-393.

Smith, R.L. and Huang, L.-S. (1993). 'Modeling high threshold exceedances of urban ozone'. Technical Report #6, National Institute of Statistical Sciences.

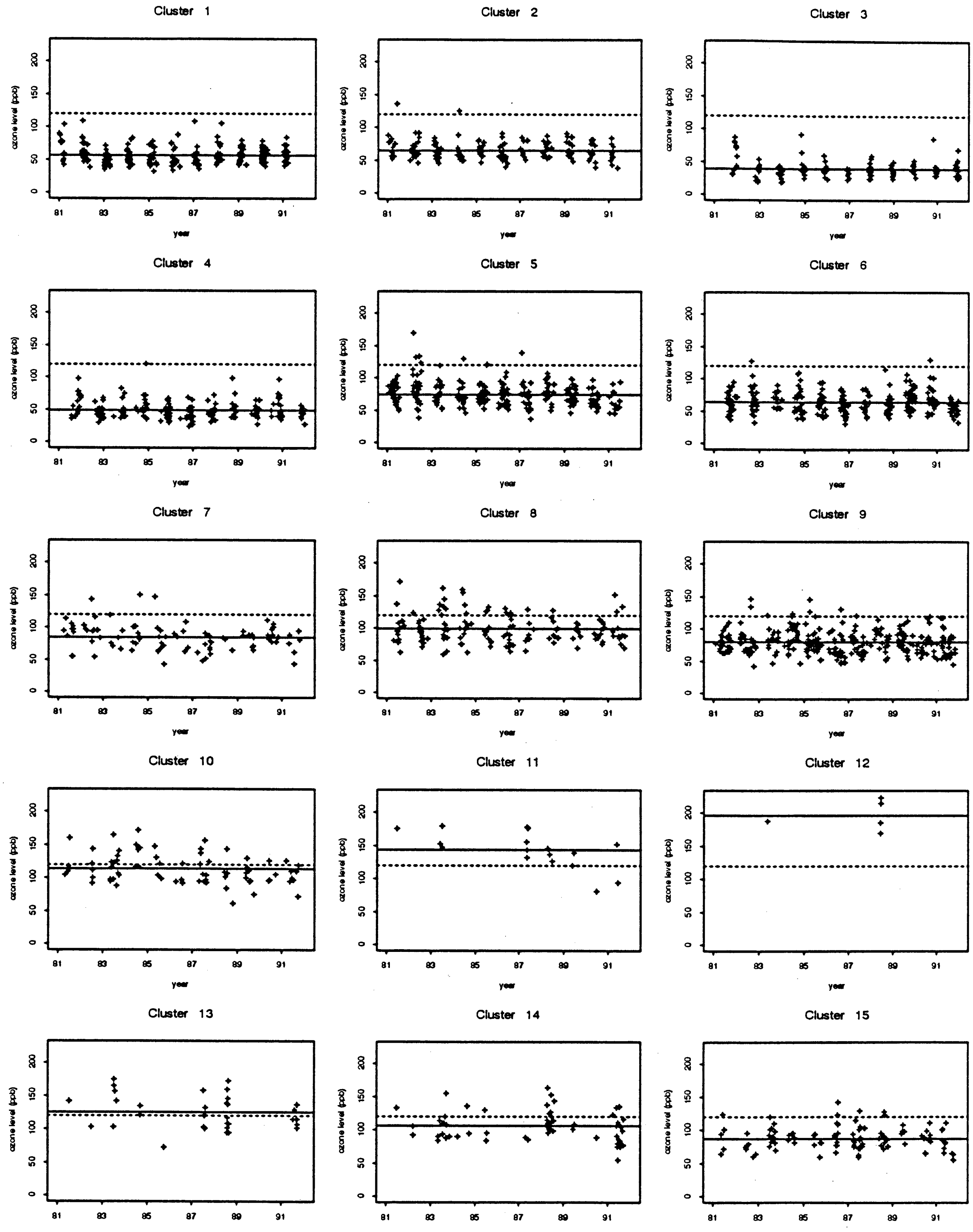
Stoeckenius, T. (1990). 'Adjustment of ozone trends for meteorological variation'. Presented at the Air and Waste Management Specialty Conference, Tropospheric Ozone and the Environment, Los Angeles, CA.

Figure 1



t<76.5
t<67.5
t<84.5
day<181.5
day<205.5
wspd<4.5
wspd<4.5
t<58.5
t<56.5
74.67
64.20
t2<38.05
80.46
t<88.5
rh<44.5
56.89
65.10
39.31
49.42
85.05
99.43
114.00
day<196.5
106.30
87.62
pr<1019.54
125.40
143.60
196.40

Figure 2



Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Cluster 11
Cluster 12
Cluster 13
Cluster 14
Cluster 15
ozone level (ppb)
0
50
100
150
200
81
83
85
87
89
91
year

Figure 3

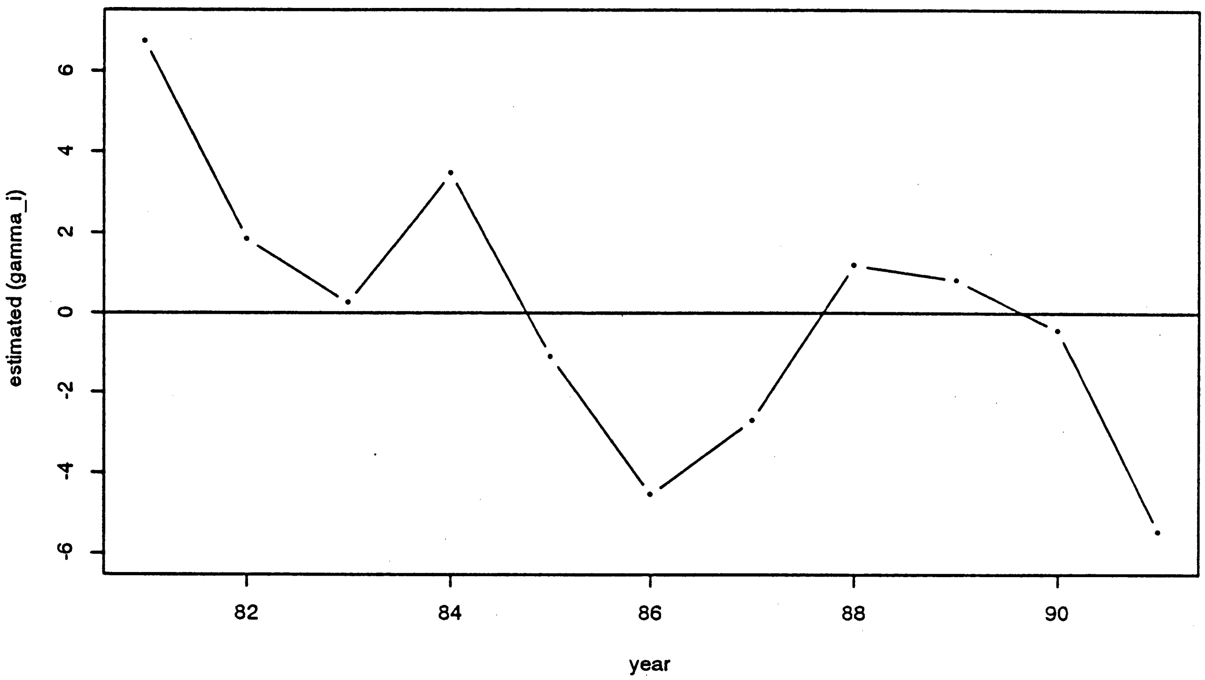


Figure 4(a)

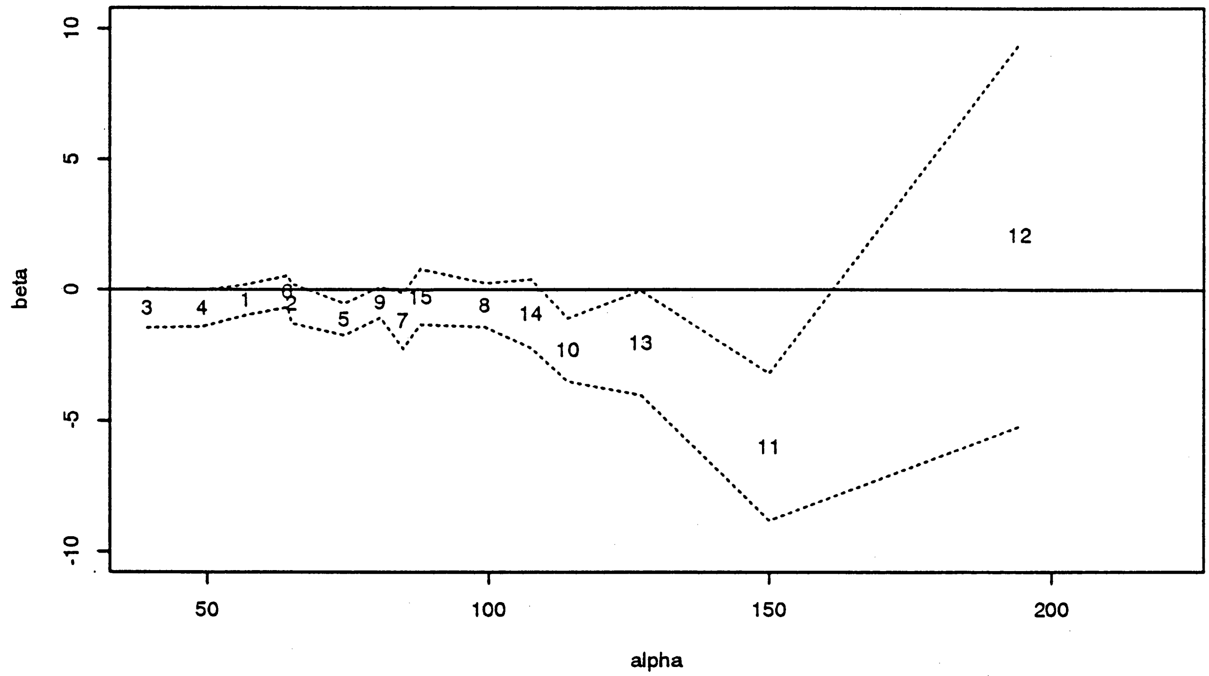


Figure 4(b)