Contract Report 2003-07

Feasibility of Forecasting Surface Ozone Concentrations in the Chicago Area

by

Donald F. Gatz

Prepared for the
Illinois Department of Natural Resources
Department of Commerce and Community Affairs

April 2003

Illinois State Water Survey
Atmospheric Environment Section
Champaign, Illinois

A Division of the Illinois Department of Natural Resources

Feasibility of Forecasting Surface Ozone Concentrations in the Chicago Area

Donald F. Gatz
Illinois State Water Survey
2204 Griffith Dr., Champaign, Illinois 61820

Support for this work was provided by the State of Illinois,
through the Department of Natural Resources
and the Department of Commerce and Community Affairs

Executive Summary

Episodic controls on sources of ozone precursor gases have been suggested as an alternative to continuous controls as a strategy for reducing ozone concentrations to meet current air quality standards. To show the feasibility of episodic controls to meet ozone air quality standards, it is first necessary to show that it is feasible to forecast surface ozone concentrations with sufficient accuracy and sufficient lead time that episodic controls can be instituted.

This study examined the feasibility of a statistical forecast of surface ozone concentrations in the Chicago area (Lake, Cook, and DuPage Counties), based on current concentrations and current and expected weather conditions. Forecast methods were developed using historical data on surface ozone concentrations and meteorological variables measured from 1990-1995. Overall, the study included:

- An extensive literature review and summary.
- Documentation of forecast methods used to call Ozone Action Days.
- Analysis of Ozone Action Days called in 1995-1997.
- Creation of air quality and meteorological databases.
- Examination of bivariate relationships between ozone and meteorological variables, including back trajectories on days with high ozone concentrations.
- Development of four forecasting approaches involving regression equations and two methods of adjusting or enhancing the results of the regression equations.
- Analyses of forecasts based on the four approaches.

The literature contains a substantial number of papers on ozone forecasting methods. Most papers used some form of regression analysis, but uses of neural network methods and classification and regression trees also were reported. There was considerable interest in forecasting ozone concentrations in the late 1970s and early 1980s, and again recently.

Ozone Action Days in the Chicago area are currently called by meteorologists from Illinois, Indiana, Wisconsin, and Michigan. Decisions to call an Ozone Action Day for the following day are based on their expert judgment of current and expected weather and air quality conditions, after discussions during a morning conference call. Between 1995 and 1997, 38 Ozone Action Days were called in the Chicago area. During this time, nine exceedances of the 125 parts per million (ppm) 1-hour standard were observed, seven of which occurred on Ozone Action Days. This rate of exceedances is similar to that observed in the Chicago area since about 1990. Exceedances have tended to occur more frequently on Thursday, Friday, and especially on Saturday, than on other days of the week. This pattern has been seen throughout the period for which we have data, beginning in 1981. The frequency of exceedances was higher on Saturdays during 1995-1997, but the total number of occurrences was relatively low, and the differences from the historical period were not statistically significant.

Considerable spatial variation of ozone concentrations on days when one or more high values were observed may be evidence of a strong influence of local precursor sources on ozone concentrations in the Chicago area.

Bivariate scatterplots of observed ozone concentrations versus potential predictor variables show strong direct relationships between ozone and high temperature, dew point, and solar radiation. There was also a strong direct relationship with the previous day's ozone concentration. An inverse relationship with wind direction was observed. A plot of back trajectories on high ozone days showed transport winds predominantly from the southwest, but there was an appreciable fraction of winds from the east and other directions as well.

Linear regression equations were developed using 70% of the observations of daily 1-hour or 8-hour maximum ozone concentrations and daily meteorological variables. The equations were then applied to the remaining 30% of the data, and the "predictions" were compared to observed ozone to test the performance of the regression equations. For both 1-hour and 8-hour ozone, regressions were developed for a base case (Approach 0) using the full set of predictor variables. The remaining regressions used only forecasted variables. Approach 1 was identical to Approach 0, but the data excluded solar radiation, the only nonforecasted variable. Approaches 2a and 2b applied different adjustments to the ozone concentrations calculated by Approach 1. Approach 3 developed a regression equation using only data with high observed ozone concentrations. Approach 4 developed a fourth-order (nonlinear) polynomial equation using the daily high temperature as the only predictor variable.

The predictor variables with the strongest influences on ozone concentrations in Approaches 1 and 3 were maximum temperature, the previous day's maximum 8-hour ozone concentration, and relative humidity. Other variables, including visibility, wind speed, weekday/weekend occurrence, and sky cover, had smaller influences.

Overall results of comparing predicted and observed ozone concentrations may be summarized in terms of R^2, the square of the correlation coefficient. Agreement between predicted and observed concentrations is also expressed by the root mean square difference or error (RMSE) between the two. For 1-hour ozone, Approach 2a achieved the highest R^2 (0.711) and the lowest RMSE (12.44 parts per billion or ppb) when applied to the full 30% data set. When forecast high temperatures were substituted for observed high temperatures in Approaches 1, 3, and 4, R^2 decreased and RMSE increased. As the temperature forecast lengthened from 1-3 days, R^2 decreased and RMSE increased further.

For 8-hour ozone, Approach 2a again produced the highest R^2 (0.687) and the lowest RMSE (10.55 ppb) when applied to the full 30% data set. As with 1-hour ozone, substitution of the forecasted high temperature for the observed high temperature reduced R^2 and increased the RMSE. As the temperature forecast lengthened from 1-3 days, R^2 based on Approach 4 decreased and RMSE increased. However, for Approaches 1 and 3, changes in both R^2 and RMSE were minimal as the temperature forecast lengthened from 1-3 days.

Skill in forecasting exceedances was assessed in terms of seven separate quantitative measures. Forecasting skill was computed for two different forecasting strategies, one in which an exceedance was forecast only when the regression equation predicted a value above the respective standard, and one in which the threshold for prediction of an exceedance was somewhat lower than the standard. For 1-hour ozone, the standard is 125 ppb, and the lower threshold selected for the second strategy was 100 ppb. For 8-hour ozone, the standard is 85 ppb, and the lower threshold selected was 65 ppb.

For both 1-hour and 8-hour ozone, the probability of detection was markedly better using the lower prediction threshold, and three more complex measures of forecasting skill showed higher scores with the lower threshold as well. Of course, this strategy increased the false alarm rates. These results suggest that one may choose a prediction threshold somewhat below the standard to achieve either a predetermined probability of detection (POD) level or a predetermined false alarm rate (FAR) level. By fine-tuning the prediction threshold between those that achieve predetermined POD and FAR levels, one may find a value that yields an acceptable combination of these two skill measures. The more complex skill scores also can provide information that may help to suggest an acceptable prediction threshold. Based on these results, it appears quite likely that one or more of the approaches used in this work can be fine-tuned to give forecast techniques for 1-hour and 8-hour ozone that will predict exceedances with equal or greater skill than the current method. At least, quantitative statistical predictions would be useful as one form of input to decisions on whether or not to call an Ozone Action Day in Chicago.

Table of Contents

Feasibility of Forecasting Surface Ozone Concentrations in the Chicago Area

Donald F. Gatz
Illinois State Water Survey
2204 Griffith Dr., Champaign, Illinois 61820

1. Introduction

To show the feasibility of episodic controls to meet ozone air quality standards, it is first necessary to show that it is feasible to forecast surface ozone concentrations with sufficient accuracy and sufficient lead time that episodic controls can be instituted.

This study examined the feasibility of forecasting daily maximum ozone concentrations in the Chicago area (Lake, Cook, and DuPage Counties). The forecast methods were developed using historical data on surface ozone concentrations and meteorological variables measured from 1990-1995. Overall, the study included:

1. An extensive literature review and summary.
2. Documentation of forecast methods used to call Ozone Action Days.
3. Analysis of Ozone Action Days called in 1995-1997.
4. Creation of air quality and meteorological databases.
5. Examination of bivariate relationships between ozone and meteorological variables. including back trajectories on days with high ozone concentrations.
6. Development of five forecasting approaches involving regression equations and two methods of adjusting or enhancing the results of the regression equations.
7. Analyses of forecasts based on the five approaches.

2. Review of Pertinent Literature and Contacts

2.1 Literature Review

Table 1 lists 17 papers on forecasting of daily ozone concentrations, grouped according to the method used for forecasting. Fourteen papers used some form of regression analysis, mostly stepwise regression. Two papers used classification and regression trees (CART), and three papers used neural network methods. Two papers used two or more different methods and compared their results.

Judging from publication years, there was considerable interest in forecasting ozone concentrations in the late 1970s and early 1980s, and again recently. The seven papers published

Table 1. Summary of Papers on Forecasting of Daily Ozone Concentrations.

Forecast method	*Authors*	*Year*	*Forecast location*
Regression analysis	Hubbard and Cobourn	1998	Louisville, KY
	Bloomfield et al.	1996	Chicago
	Ryan	1995	Baltimore
	Comrie	1995	Seattle, Pittsburgh, Chicago, Atlanta, Charlotte, Boston, Tucson, Phoenix
	Eder et al.	1994	Birmingham, AL
	Feister and Balzer	1991	Germany
	Robeson and Steyn	1990	British Columbia, Canada
	Clark and Karl	1982	Northeast U.S.
	Prior et al.	1981	St. Louis, MO
	Karl	1979	St. Louis, MO
	Wolff and Lioy	1978	Northeast U.S. (including Chicago)
	Revlett	1978	Los Angeles
	Aron and Aron	1978	Los Angeles
	Tiao et al.	1976	Los Angeles
CART			
	Burrows et al.	1995	Vancouver, Montreal, and Atlantic region of Canada
	Ryan	1995	Baltimore
Neural network			
	Yi	1996	
	Comrie	1995	Seattle, Pittsburgh, Chicago, Atlanta, Charlotte, Boston, Tucson, Phoenix
	Ruiz-Suarez	1995	Mexico City

between 1978 and 1982 all used regression analysis. Their target areas for forecasts were Los Angeles (three papers), the northeastern United States (two papers), and the St. Louis area (two papers). The remaining ten papers were published in the 1990s, six of them in 1995 and 1996, and one in 1998. The recent papers have targeted forecasts at a wide variety of locations in the United States, including Chicago (twice), and internationally as well, including locations in Canada, Germany, and Mexico.

Brief descriptions of some of the more significant and relevant papers follow, including those that targeted Chicago or Cook County. Results from these papers are tabulated in Table 2, and include such measures of the agreement between observed and forecasted concentrations as R^2, also known as the coefficient of determination (or square of the correlation coefficient), and the root mean square error (RMSE). The table shows results separately for observed and forecasted input variables.

Wolff and Lioy (1978) used linear stepwise multiple regression to develop a best fit equation relating maximum afternoon ozone concentrations to meteorological and air quality conditions along a 24-hour upwind air parcel trajectory. The equation was developed using data from measurement locations in and upwind of northern New Jersey, and it was applied to five other locations in the Northeastern quadrant of the United States, including Cook County, Illinois. The four variables included in the equation were the maximum upwind ozone on the previous day, the maximum temperature on the day forecast, the maximum temperature upwind on the previous day, and the average wind speed in the mixed layer. Estimated upwind emissions of hydrocarbons and nitrogen oxides were also tested, but did not improve the multiple correlation coefficient.

Table 2 shows results for both the calibration site in northern New Jersey and for Cook County. The R^2 value for the New Jersey location, based on observed meteorological data from 50 cases, was 0.92, with a root mean squared (RMS) difference between observed and "forecasted" concentrations of 10 parts per billion (ppb). Where forecasted input data were used, however, the R^2 values dropped to between 0.06 and 0.23, and the RMS difference ballooned up to 29-35 ppb, depending on which one of three forecast data options was used. The mean percent difference between observed and forecasted concentrations over 14 test cases was 6% using observed meteorological variables, but this jumped to 23-29%, depending on forecast option, when forecasted variables were used.

The same equation was applied to forecasting the highest 1-hour ozone concentration at any of 45 sites in Cook County, with 22 cases tested. Using observed (not forecasted) meteorological variables, R^2 was 0.78, and the RMS difference was 17 ppb. The mean percent difference was 13%, with a maximum of 27%.

Karl (1979) also used linear stepwise regression to develop equations for forecasting 1-hour maximum ozone concentrations at St. Louis. Separate equations were developed for 1) the next day, and 2) the second day, for three subgroups of 25 sampling sites. The sites were

Table 2. Results of Ozone Forecasting Methods from the Literature, 1978-1996

		R²		*RMSE, ppb*		*Other measures*	
Authors	*Location*	*Forecast meteorology*	*Observed meteorology*	*Forecast meteorology*	*Observed meteorology*	*Forecast meteorology*	*Observed meteorology*
						Mean difference, percent	
Wolff and Lioy (1978)	Northern New Jersey	Opt :1 :0.15 Opt. 2: 0.23 Opt. 3: 0.06	0.92	Opt :1: 35 Opt. 2: 29 Opt. 3: 32	10	Opt. 1: 29 Opt. 2: 23 Opt. 3: 23	6
	Cook Co, IL		0.78		17		Mean diff = 13%, max = 27%
Karl (1979)	St. Louis					Standard error of est., ppb	
	24 h, inner sites	0.41				27	
	24 h, trans sites	0.51				30	
	24 h, outer sites	0.55				23	
	48 h, inner sites	0.35				28	
	48 h, trans sites	0.42				31	
	48 h, outer sites	0.45				25	
Clark and Karl (1982)	27 sites in 9 states, NE U.S.	Mean = 0.41 Max = 0.50 Min = 0.22		Mean = 32 Min = 18 Max = 50		± 20%, 50% of the time ± 40%, 77% of the time	
Ryan (1995)	Baltimore					Standard error, ppb	
	Regression, all data	0.53	0.62			17	20
	CART, all data	0.19	0.34			23	18
	Regression, high ozone	0.31	0.34			13	13
	CART, high ozone	0.28	0.24			14	14
Eder et al. (1994)	Birmingham, AL						
	Composite model		0.59		13		
	Range of 7 clusters		0.30 - 0.58		10 -15		
Bloomfield et al. (1996)	Chicago		0.80		8		± 5 ppb 50% of the time; ± 16 ppb 95% of the time.
Hubbard and Cobourn (1998)	Louisville, KY		0.70				± 7.6 ppb 50% of the time; ± 15 ppb 80% of the time.

Notes: R^2 = square of the correlation coefficient. RMSE = root mean square error.

grouped according to their distance from the city center. The input data for the forecast equations were model output statistics (MOS) derived from the National Meteorological Center's Limited-Area Fine Mesh (LFM) model, and thus were considered to represent forecasted input variables. Separate equations were also developed to forecast the probability of ozone concentrations exceeding the 1971 National Ambient Air Quality Standard of 80 ppb. Results are shown in Table 2. For the next day forecast, R^2 ranged from 0.41 to 0.55, and the standard error of estimate from 23-30 ppb. For the second day forecast, R^2 ranged from 0.35-0.45, with standard errors of estimate between 25 and 31 ppb. The forecasts were shown to be better than could be obtained by chance, or by using persistence or seasonality (climatology).

Clark and Karl (1982) developed linear multiple regression equations for each of 27 ozone monitoring sites in nine Northeastern states to forecast the next day's maximum 1-hour average ozone concentrations. Six of the 27 sites were rural or remote, and the rest were urban or suburban. Again, most of the input variables required by the prediction equations were derived from the LFM model, so the results in Table 2 are listed in the "forecasted input variables" category. The variables selected nine or more times in the 27 equations were: maximum temperature, the quadrant of back-trajectory approach, atmospheric pressure change, and the boundary-layer u-component of the wind. A "predictand enhancement technique," designed to reduce the typical underprediction of high values and overprediction of low values, was used in this work. Results in Table 2 show that, over the 27 sites, the mean R^2 was 0.41, with a range of 0.22-0.50. The mean RMS difference was 32 ppb, with a range of 18-50 ppb. The R^2 values from the prediction equations were greater than those from persistence forecasts, and the RMS differences were less than those from persistence; even so these results appear to be only marginally useful as prediction tools. This was confirmed by the overall finding that only 50% of the predictions at all sites were within 20% of the observed ozone concentration, and only 77% were within 40%.

A pilot forecast program was undertaken in Baltimore (Ryan, 1995) to support possible episodic emission control efforts. A CART method was used, as well as standard regression analysis and a human expert. All approaches tended to underpredict ozone at forecast lead times of 24 hours. The low bias varied from 5 ppb for the expert forecast to 7 ppb for the CART forecast. Separate equations were developed for 1) 46 days between June 15 and July 31, 1993, and 2) 24 days during that period with ozone concentrations greater than 100 ppb. A helpful feature of this paper is that results were also expressed in terms of seven different measures of forecast skill.

For the full data set, the R^2 for the regression was 0.62 with observed variables, and dropped to 0.53 with the forecasted variables. For the CART method, R^2 for the observed variables was only 0.34, and it dropped to 0.19 for the forecasted variables. The standard error for the regression equation was 21 ppb using observed variables, and somewhat less, 17 ppb, using forecasted variables. For the CART method, the standard error was 18 ppb using the observed variables, and 23 ppb with the forecasted variables.

For the high-ozone data set, the respective R^2 values were all somewhat less than for the full data set, except for that of the CART method with forecasted variables. In this instance, it increased from 0.19 to 0.28, still a relatively low value. The standard errors all decreased relative to those of the full data set, however, and fell in a narrow range between 13 and 14 ppb.

One of the principal sources of error was underprediction of mid-day surface temperature by the standard meteorological models during extremely warm episodes. The author judged that improvements could be made by using more recent data for initialization of the ozone regression forecasts and the use of local forecasts to supplement temperature forecasts during warm periods.

Eder et al. (1994) investigated the dependence of ozone concentrations on meteorological conditions at Birmingham, Alabama. A clustering approach was used to identify seven statistically distinct meteorological regimes. Then stepwise regression was applied to develop ozone prediction equations for each of the regimes. Finally, the equations for the seven regimes were amalgamated into a single composite model that exhibited a significantly larger R^2 and smaller RMS difference when compared to an overall model in which the meteorological data were not clustered.

Bloomfield et al. (1996) examined ozone concentrations and meteorology in the Chicago area using data for 1981-1991. Ozone data were hourly averages from 45 sites. The surface meteorological data were from O'Hare Airport, and upper air data were from the nearest radiosonde station at Peoria, Illinois. The median of the daily 1-hour maximum values at each sampling site was selected as the response variable for the development of a nonlinear regression model. The model was intended to be used to estimate that part of the trend of ozone concentrations that cannot be accounted for by trends in meteorology, and to "adjust" observed ozone concentrations for anomalous weather conditions. It was not used for forecasting daily ozone concentrations. Ozone concentrations computed from the model differed from actual values by up to ± 5 ppb about half the time and up to ± 16 ppb about 95% of the time. An R^2 value of 0.80 and an RMS difference of about 8 ppb were reported. However, it appears that these results were obtained with the same data used to develop the model, not an independent data set.

After an investigation very similar to the one described in this report, Hubbard and Cobourn (1998) published results for Louisville, Kentucky. An ozone forecasting capability was sought to support episodic emission controls on Ozone Action Days (OADs). The dependent variable was an area daily maximum 1-hour ozone concentration observed from among three sampling sites. The independent variables were selected from surface meteorological data to develop a ten-parameter, multiple linear regression model. The data used for development of the model were collected between May and September from 1993 through 1996. Ozone concentrations computed from the model differed from actual values by up to ± 7.6 ppb about half the time and up to ± 15 ppb about 80% of the time. An R^2 value of 0.70 was reported. Again, it appears that these results were obtained with the same data used to develop the model, not an independent data set.

The results of these previous studies, in terms of their ability to predict ozone concentrations accurately 24-48 hours in advance, may be summarized by the R^2 values, RMS differences, standard errors, or mean differences they achieved. Table 2 summarizes these parameters for a number of studies in the literature. The studies listed in Table 2 were either at Illinois-relevant locations (e.g., Chicago or St. Louis), or they describe particularly interesting methods that might be applied to forecasting ozone in Illinois locations.

In general, R^2 values obtained using observed meteorological input variables were greater than those in which forecasted variables were used, as would be expected. The highest R^2 values were reported in studies that used observed (i.e., as opposed to forecasted) meteorological independent variables, and in which the models were tested on the same data used to develop them. Wolff and Lioy (1978) reported an R^2 of 0.92 using the same observed data from northern New Jersey used to develop the model. With forecasted input variables, however, the R^2 results were not as good. The same model equation applied to an independent data set from Cook County still achieved an R^2 of 0.78, however, using observed (not forecasted) input variables. Bloomfield et al. (1996) achieved an R^2 of 0.80 when their model was tested on the data used in its development. Direct comparison of R^2 values obtained using observed versus forecast input variables on the same data set is also possible for the results of Ryan (1995). In three of four cases, R^2 with observed variables was greater than R^2 with forecasted variables, although the differences were much less striking than in the Wolff and Lioy data set.

Root mean square differences are another commonly used relative measure of forecast accuracy. The RMS is analogous to the standard deviation of a set of measurements in which the differences are taken with respect to the mean value of the set. When observed meteorological variables are used as input to the equations, RMS differences in Table 2 range from 8-17 ppb. The Wolff and Lioy (1978) data set provides the only comparison of RMS differences from using observed meteorological variables and forecasted variables on the same data set. When the observed variables gave an RMS difference of 10 ppb, the values using forecasted variables were three times larger, and ranged from 29-35, depending on the forecast data option. The RMS differences of roughly the same magnitude were reported by Clark and Karl (1982) for forecasted input variables at 27 sites in the Northeastern United States.

Two of the papers summarized in Table 2 provide a comparison of forecast accuracy with observed variables versus forecasted variables, based on somewhat different measures of accuracy. Wolff and Lioy (1978) reported mean percent differences between observed and forecast ozone concentrations using both observed and forecasted input variables. With the observed variables, a mean difference of 6% was obtained, but that value rose to between 23 and 29% when the forecasted variables were used.

Ryan (1995) provided a similar comparison based on the standard error. In four cases based on both regression analysis and the CART technique, the standard errors were equal in two cases. In one case the observed variables produced a lower standard error, and the forecasted variables gave the lower value in another case.

To summarize previous ozone forecasting efforts, although some of the results in the literature using observed values of the input variables offer the hope of rather accurate forecasts, comparison with results based on forecasted input variables shows that forecast accuracy often drops significantly when the input variables are forecasted.

The literature related to the Lake Michigan Ozone Study (LMOS) is also of interest in connection with ozone forecasting efforts. The field observation phase of the LMOS took place during the summer of 1991. Four episodes of high ozone concentrations, covering 21 days, were analyzed for this study (Hanna and Chang, 1995). All four episodes were associated with large polluted regions with dimensions of 1000-2000 kilometers (km), located on the western side of high pressure systems in the Eastern United States. The air coming into the LMOS region on these occasions was traced to upwind source regions from St. Louis through the Ohio valley, and as far east as the large cities of the Northeastern United States. The local sources in the Chicago area (Gary to Milwaukee) add to the incoming polluted air mass to affect ozone concentrations. Local meteorological conditions associated with Lake Michigan and the layer of relatively cool air immediately above the lake surface also play a significant role in the transport of ozone and its precursors, especially downwind of Chicago in Wisconsin and Michigan, where ozone concentrations are often observed to be higher than in the Chicago area.

2.2 Contacts

Terry Sweitzer of the Illinois Environmental Protection Agency (IEPA), who directs the IEPA OAD program, provided the following information. The program has been in effect for three summers. It is a coalition of IEPA, business and industry, and environmentalists. It targets ozone in the Chicago area, although the forecast for an OAD applies to all of Illinois and Indiana. Actually four states are working together: Illinois, Indiana, Wisconsin, and Michigan. Illinois and Indiana are one region for which the forecast is made, and Wisconsin and Michigan do their own forecasts. The forecasts are based on a conference call between meteorologists from the four states. Bob Swinford is the meteorologist for Illinois. The call is made in the morning before the target day. The meteorologists each look at their maps and use their experience to discuss the situation and decide, based on the discussion, whether to call an OAD for the next day. They do not use a mathematical equation to predict the ozone concentrations. When an OAD is called, it is announced to the mass media and the commercial and industrial "partners" who have agreed to take some measures to reduce emissions. From the media outreach, the hope is that the public will reduce automobile travel, use car pools, not cut lawns until evening, not refuel during the heat of the day, and similar measures.

The OAD forecasting procedures are outlined in a document called "Lake Michigan Ozone Weather Forecasting Protocol" (see Appendix 1). The forecasts also are used to trigger episodic enhanced monitoring of ozone, nitrogen oxides (NOx), volatile organic compounds (VOCs), and meteorological parameters in the Chicago area. The document also indicates that high ozone concentrations are favored by high temperatures, high humidity, light-to-moderate winds, low cloud cover, and little or no precipitation.

Surveys are used to gauge responses from the partner companies and the public. The survey of 500 citizens conducted in August 1997, indicated that 77% have heard about OADs. About 65% took one or more ozone-reducing actions: 43% limited their driving, 21% deferred mowing, and 17% avoided certain household products. Of those who had heard about OADs, 50% heard from TV, 26% from radio, 20% from newspapers, and 3% from Illinois Department of Transportation highway signs.

The National Oceanic and Atmospheric Administration (NOAA) Air Resources Laboratory currently provides hourly forecasts of surface ozone concentrations at its World Wide Web site (http://www.arl.noaa.gov/ready/ozone.html). The Hysplit-4 trajectory model is used to compute hourly ozone concentrations for the eastern half of the United States for each hour from 2:00 p.m. to 8:00 p.m. EDT, plus a time-lapse "movie" of these hourly spatial patterns. These forecasts are made each day for the same day and the next day. Users are cautioned that the forecasts are experimental and unverified, and that the underlying assumptions may change without warning. Further details of the forecast methods are given in Appendix 2.

3. Methods

3.1 Methodology for Gathering Air Quality and Meteorological Data

3.1.1 Ozone Data

The IEPA retrieved ozone data from the USEPA AIRS database and provided 1-hour average ozone concentrations in parts per billion for this research. Data were obtained for 21 Illinois sampling sites. These included four sites in central Illinois that we anticipated using as indicators of air quality in air approaching the Chicago area. However, ozone forecasting activities were focused on 17 sampling sites in Lake, Cook, and DuPage Counties. A complete list of all 17 sites by county is:

Cook County: Alsip, Calumet City, Chicago-CTA, Chicago-Jardine, Chicago-SE Police, Chicago-South Water Filtration Plant (SWFP), Chicago-Taft, Chicago-University, Cicero, DesPlaines, Evanston, and Lemont.

DuPage County: Lisle

Lake County: Deerfield, Libertyville, Waukegan, and Zion.

Ozone measurements were taken in the field once every minute and averaged for the entire hour to produce the 1-hour ozone average. These hourly averages were then used to determine the daily maximum 1-hour ozone concentration, as well as the derived variables such as the daily maximum 8-hour mean ozone concentration. The maximum 8-hour concentration is the highest mean concentration over all possible combinations of eight consecutive hours each day.

3.1.2 Meteorological Data

Hourly meteorological data from the National Weather Service (NWS) reporting station at O'Hare Airport in Chicago were obtained for use in this research. The hourly data were used to compute daily summary data of various kinds (e.g., daily maximum, mean, or total), as detailed in Section 3.3, for use in the regression analyses. Data on the ozone concentrations and some meteorological variables were available for 1990-1997, but full meteorological data sets were available only for 1990-1995.

Daily maximum temperature forecasts prepared by the National Weather Service's National Centers for Environmental Prediction (NCEP) in Silver Spring, Maryland, were obtained for use in comparing ozone concentrations predicted from observed and forecast maximum temperatures. Forecasts for maximum temperature one, two, and three days in advance were obtained. The temperature forecasts used were those produced by the Coded Cities Forecast (CCF) for Chicago (O'Hare Airport). This product comes directly from an individual NWS Forecast Office and is produced twice daily, once around 5 a.m. and once in the afternoon around 3 p.m.. The on-duty forecast meteorologist uses personal knowledge of available forecast models and current weather conditions to produce this product. Variables forecast are maximum and minimum temperatures out to 60 hours in advance and probability of precipitation for 36 hours in advance.

3.2 Databases Created

Excel and Quattro Pro spreadsheet databases were created for use in plotting graphs and statistical calculations. Each row contains data for an individual day. The columns contain information about the date and day of the week of the observations, data for each of the meteorological variables, 1-hour and 8-hour maximum ozone concentrations from all 17 area sites for each day (the dependent variables) and for the day before, and a smoothed long-term daily mean of the area's maximum 1-hour ozone concentration. Thirty percent of the days in this data set were randomly selected and set aside as a separate independent data set (referred to later as the "30%" or "independent" data set) for testing regression equations developed from the remaining 70% of the data (the "70%" or "development" data set). In addition, we created a separate spreadsheet of the hourly ozone concentrations for the 1990-1997 ozone seasons (April 1 through October 31), measured at 21 sampling stations in northern and central Illinois.

Table 3. Predictor Variables Considered in the Stepwise Regression Analyses.

Predictor	*R^2 for bivariate regression*	
	Vs. 1-hr O_3	*Vs. 8-hr O_3*
3-county 1-hour maximum O_3 concentration, lagged 1 day, ppb.	0.401	0.373
3-county 8-hour maximum O_3 concentration, lagged 1 day, ppb.	0.416	0.403
Smoothed historical mean 1-hr maximum ozone concentration, by day of O_3 season.	0.256	0.264
Weekday/weekend. (Weekday = 0, Saturday or Sunday = 1)	0.961E-3**	0.281E-2
O'Hare daily maximum temperature, F*.	0.511	0.471
O'Hare daily mean dew point, F*.	0.251	0.197
O'Hare daily mean relative humidity, percent*.	0.089	0.093
O'Hare daily mean wind speed, mi/hr*.	0.097	0.073
O'Hare daily mean u-component of the wind, mi/hr*.	0.90E-5	0.40E-5
O'Hare daily mean v-component of the wind, mi/hr*.	0.035	0.033
O'Hare daily mean surface pressure, inches of Hg*.	0.73E-3	0.002
O'Hare daily mean total sky cover, tenths*.	0.075	0.087
O'Hare daily total solar radiation, MJoules/m^2	0.288	0.344
O'Hare daily mean horizontal visibility, miles*.	0.83E-2	0.002

Notes: * Variables for which forecasts are routinely made. ** Scientific notation. E-3 means 10 to the power (-3). Thus 0.961E-3 = 0.000961.

3.3 Description of Forecast Method: Regression Analysis

Linear least-squares stepwise multiple regression was the primary technique used to develop forecast equations. This was the method of choice for most of the publications reviewed in Section 2.1. The computational procedures are relatively straightforward, and software was readily available. These were distinct advantages because of the time constraints of this investigation. The computations were carried out using Systat statistical software (Wilkinson, 1997). The regression was carried out in a backward mode, using a probability value of 0.10 to remove a variable. The potential predictors considered for the regression analyses are listed in Table 3, along with their respective values of R^2 for bivariate linear regression with a constant.

The lagged ozone concentrations are the previous day's 1-hour or 8-hour maximum ozone concentrations in the three-county area. The smoothed historical mean ozone concentration is the average 1-hour maximum ozone concentration for each day of the ozone season. Daily values were obtained by averaging overall sampling stations for 1990-1997, and smoothed using a lowess (locally-weighted scatterplot smoother) procedure (Cleveland, 1985). The unsmoothed daily mean values are shown as a function of time in Figure 1.

The weekday/weekend variable takes the value 0 for Monday-Friday and 1 for Saturday and Sunday. An alternative scheme (Thursday, Friday, and Saturday = 1, otherwise 0) was tested, but found less useful.

Wind direction is typically expressed in terms of the u- and v-components of the wind direction vector when used as variables in regression analyses. The u-component is the east-west component, positive when the wind is out of the west. The v-component is the north-south component, positive for winds out of the south. Use of these components avoids the problems inherent in expressing near-north winds as both low values (near 0 degrees) and high values (near 360 degrees).

Five separate approaches were used to develop ozone forecasting methods. Approach 0 is the baseline case, using all available potential predictor variables. To be useful for operational forecasting, however, an ozone forecast equation must be based on variables that are either known or routinely forecasted. Thus, in Approach 1, regressions were carried out using only known or routinely forecasted variables. In the list in Table 3, those variables with asterisks were known (the first four) or forecasted. Among the variables in Table 3, only solar radiation was considered to be not routinely forecast.

Approach 2 consists of two separate "enhancements" of results from Approach 1. The first is that of Hubbard and Cobourn (1998), who used a quadratic fit of the relationship between predicted and observed ozone concentrations to adjust computed values. The forecast ozone concentration $\tilde{Y}$ (using their original notation) is given by:

$$\tilde{Y} = a_0 + a_1\hat{Y} + a_2\hat{Y}^2$$

where $\hat{Y}$ is the unadjusted forecast and the *a* terms are the regression coefficients from the quadratic fit.

The second enhancement method is that of Clark and Karl (1982). It has the effect of increasing predicted high values and decreasing low values. In the Clark and Karl notation, the equation is:

$$P_E = \frac{P - \overline{P}}{R} + \overline{P}$$

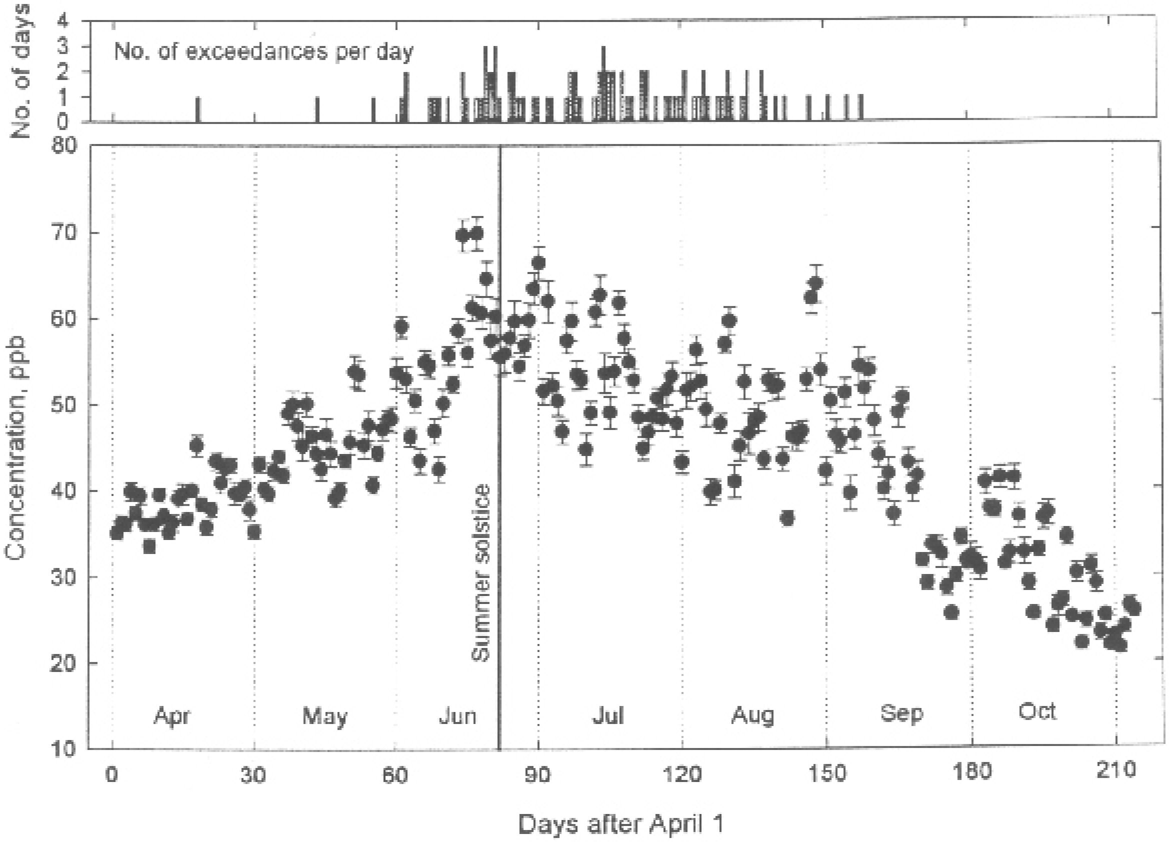

Figure 1. Lower panel: Mean of the daily maximum 1-hour ozone concentration at each of 17 sampling sites in Lake, Cook, and DuPage Counties during 1990-1997, by day of the ozone season. Monthly indicators are approximate. Error bars are standard errors. Upper panel: Number of exceedances for each day of the ozone season, 1981-1997.

where P_E is the enhanced predictand, P is the predictand obtained from the regression equation, $\overline{P}$ is the average value of the predictands determined from the dependent data sample, and R is the multiple correlation coefficient of the predictand with the predictors in the regression equation.

Preliminary regression results underpredicted very high ozone concentrations. To test the hypothesis that equations developed from samples with high ozone concentrations would be better predictors of the highest ozone concentrations, separate regression equations were developed from the subset of samples from the 70% data set with maximum 1-hour ozone concentrations $\geq$ 80 ppb and maximum 8-hour ozone concentrations $\geq$ 60 ppb. This was Approach 3.

All of the results of linear regression analysis were compared with ozone concentration estimates made from a nonlinear fourth order polynomial, using only maximum temperature as an independent variable (Approach 4):

$$y = y_0 + bx + cx^2 + dx^4$$

where: y is the dependent variable, the forecasted 1-hour maximum ozone concentration,
x is the independent variable, in this case the maximum temperature T_{max},
y_0 is a constant, and
b, c, and d are the regression coefficients.

To partially simulate the effects of using actual forecasted variables to estimate ozone concentrations, the equations developed using only forecasted variables were tested using both observed maximum temperatures and with 1-day, 2-day, and 3-day forecasted maximum temperatures.

All of these analyses were carried out separately for 1-hour and 8-hour maximum ozone concentrations.

4. Results of Investigation

4.1 Summary and Analysis of Ozone Concentrations in Northeast Illinois on Ozone Action Days, 1995-1997

Table 4 summarizes the observed maximum 1-hour ozone concentrations in Cook, Lake, and DuPage Counties on OADs. In 1995, OADs were called on 20 occasions when violations of the ozone standards were considered likely. Ozone concentrations equaled or exceeded the 1-hour standard of 125 ppb on four of these days; no exceedances occurred on non-OADs. In 1996, OADs were called on six occasions. Ozone concentrations in excess of the standards were observed on one of these six days, as well as on two days when OADs were not called. In 1997,

Table 4. Summary of Ozone Concentrations in Lake, Cook, and DuPage Counties on Ozone Action Days, 1995-1997.

Year	*Month*	*Date*	*Day*	*Observed 1-hour maximum O_3 concentration, ppb*	*Violations*
1995	June	6	Tu	95	
		7	W	101	
		15	Th	93	
		16	F	108	
		17	Sa	113	
		18	Su	119	
		19	M	95	
		20	Tu	72	
		21	W	82	
		24	Sa	166	X
	July	12	W	124	
		13	Th	130	X
		14	F	104	
		15	Sa	149	X
		30	Su	116	
		31	M	109	
	August	12	Sa	143	X
		13	Su	78	
		14	M	77	
		29	Tu	56	

1995 summary: 20 OADs; four exceedances occurred, all on OADs.

Table 4. Concluded.

Year	*Month*	*Date*	*Day*	*Observed 1-hour maximum O_3 concentration, ppb*	*Violations*
1996	June	28	F	125	X
		29	Sa	92	
	August	5	M	75	
		6	Tu	80	
		7	W	76	
	September	6	F	78	

1996 summary: 6 OADs; 1 exceedance on an OAD; 2 exceedances on other days.

Year	*Month*	*Date*	*Day*	*Observed 1-hour maximum O_3 concentration, ppb*	*Violations*
1997	June	24	Tu	88	
		27	F	87	
		28	Sa	109	
		29	Su	128	X
		30	M	72	
	July	1	Tu	104	
		12	Sa	110	
		13	Su	79	
		14	M	77	
		17	Th	104	
		26	Sa	157	X
		27	Su	88	

1997 summary: 12 OADs; two exceedances occurred, both on OADs.

12 OADs were called; on two of these days concentrations in excess of the standard were observed. No exceedances occurred on non-OADs in 1997.

Over the three-year period, 38 OADs were called. Exceedances of the standards were observed on seven of these days (18% of the called OADs). Exceedances also occurred on two days that were not called OADs. In other words, from 1995 to1997 nine exceedances occurred in the three-county area, seven (78%) of which were on OADs.

One can interpret these data in two ways. One may say that OADs are substantially overpredicted. The other interpretation is that they are extremely effective in reducing ozone concentrations. The truth may lie somewhere between these extremes. It is useful to examine Figure 2, which shows a histogram of observed maximum 1-hour ozone concentrations in Lake, Cook, and DuPage Counties on OADs called during 1995-1997. As noted earlier, observed 1-hour ozone concentrations exceeded the 125 ppb standard on seven OADs in the 3-county area. (Note that the 1-hour ozone standard is written in terms of 0.12 parts per million (ppm). Although 0.12 ppm is equivalent to 120 ppb, when the usual rules concerning rounding and significant figures are applied, the lowest concentration considered to be an exceedance is 125 ppb.)

It seems likely that, if the maximum 1-hour ozone concentrations would have been in excess of the 125 ppb standard if an OAD had not been called, calling an OAD might result in a reduction in maximum ozone concentration of probably not more than 25%. If that were the case, observed values of about 100 ppb or more (50% of the called OADs) might have occurred on OADs when the standard would have been exceeded if an OAD had not been called; but values less than 100 ppb (the other 50%) would not have been in violation even if an OAD had not been called. If this reasoning is approximately correct, the distribution of observed 1-hour maximum ozone concentrations in Figure 2 suggests that OADs are overpredicted by roughly 50%.

Whether OADs are overpredicted or not, we should try to determine whether they have been effective in reducing the number of exceedances of the 1-hour ozone standard. It is not possible to know what ozone concentrations would have been observed on OADs if the OAD had not been called, but we can get some indications of the effectiveness of the OADs by comparing numbers of exceedances during the three years OADs have been forecasted with historical values. Because we have three years of OAD forecasts, it is convenient to compare the number of exceedances during those three years with all available historical three-year totals.

Figure 3 shows running three-year totals for Lake, Cook, and DuPage Counties plotted at the middle year of the three. The historical data set covered the years 1981-1994. Results for the three years (1995-1997) of OAD forecasts are plotted at 1996. Compared to the bulk of the historical data, the seven exceedances during the three years of OAD forecasts look quite good. However, emissions reductions during the 1980s appear to have reduced the number of exceedances, so that by the early 1990s the number of exceedances were reduced considerably.

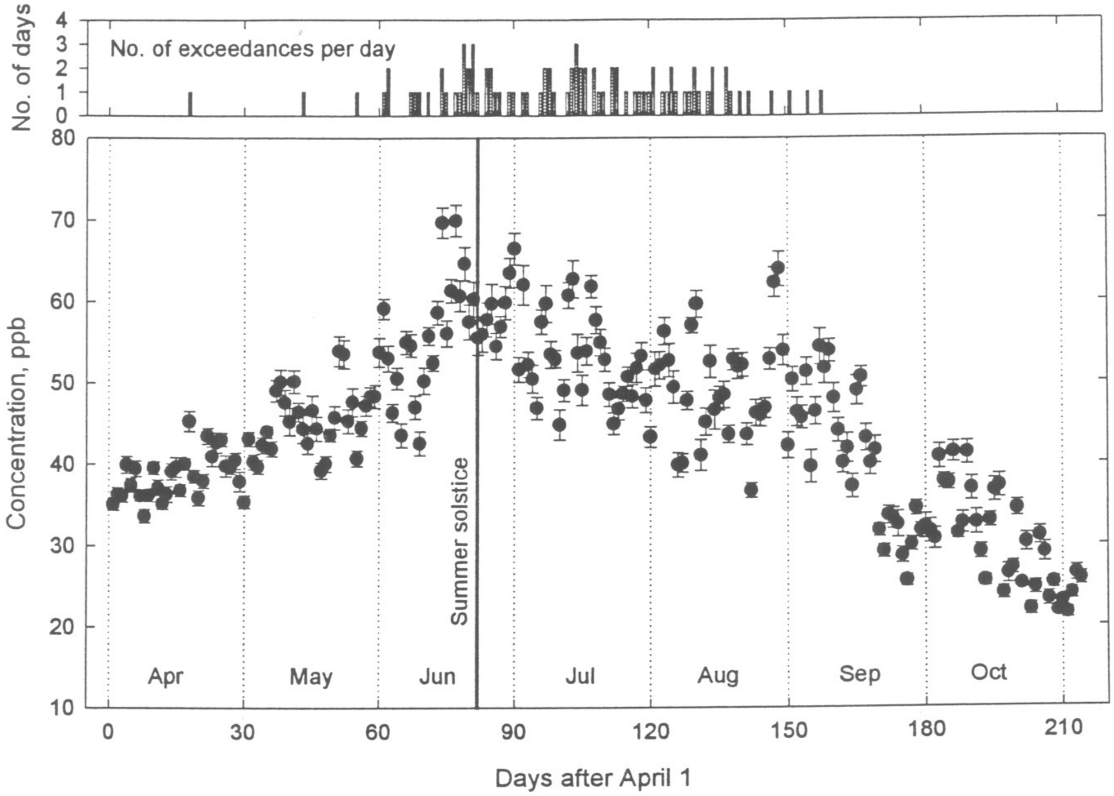

Figure 2. Summary of observed 1-hour maximum ozone concentrations on Ozone Action Days in Lake, Cook, and DuPage Counties, 1995-1997.

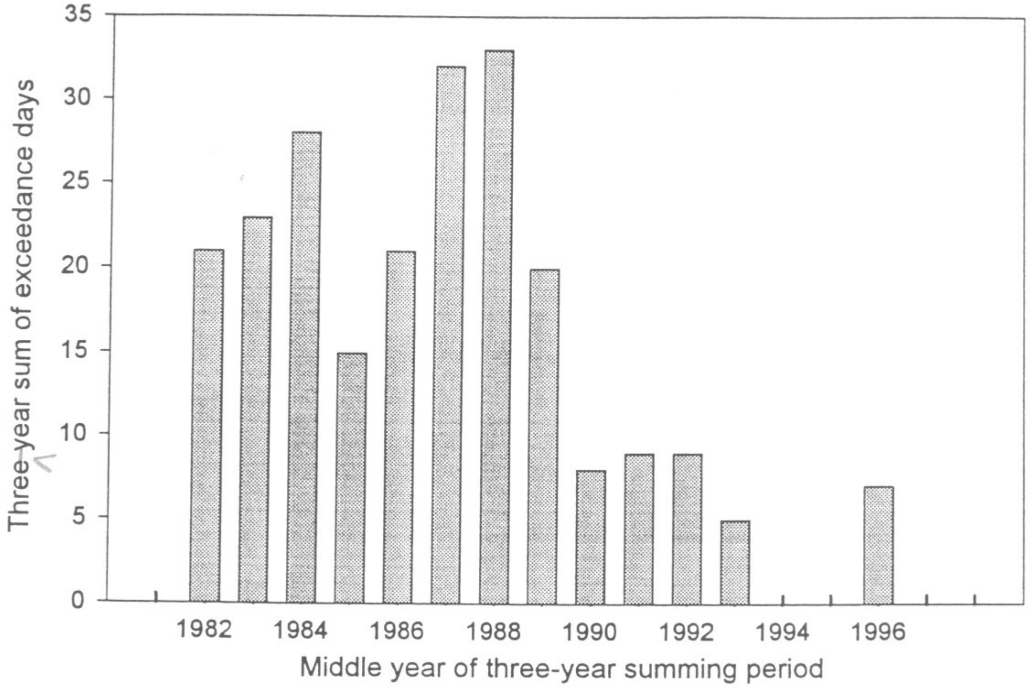

Figure 3. Running three-year sum of 1-hour ozone exceedance days in Lake, Cook, and DuPage Counties. Three-year sums are plotted at the middle year. Value plotted at 1996 represents the total exceedances during the three-year period of Ozone Action Day forecasts.

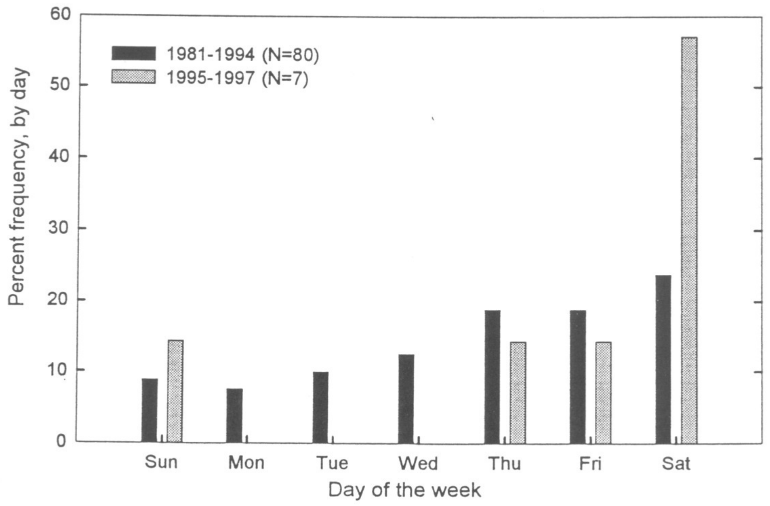

Figure 4. Percent frequency of 1-hour ozone exceedance days by day of the week. N is the number of exceedances in each period.

Compared to the running three-year values since 1990, it is difficult to conclude that the OAD forecasts in 1995, 1996, and 1997 had any significant effect.

Another analysis with a possible bearing on the effectiveness of the OAD forecasts is the distribution of the days of the week when exceedances occurred. If the exceedance days were strictly random, one would expect they should occur uniformly throughout the week at a rate of one-seventh (14.3%) each day of the week.

Figure 4 shows the distribution of exceedances by day of the week for both the historical period and the OAD forecast period. Eighty exceedances were observed during the historical period (1981-1994). The frequency of these exceedances increases from a minimum on Monday to a maximum on Saturday, then returns to near the minimum on Sunday. Note that six of the seven exceedances during the OAD period (1995-1997) occurred on Thursday, Friday, or Saturday, the three days of highest frequencies of exceedances during the historical period; and more than half of these days occurred on Saturday. Both distributions in Figure 4 were tested against the uniform distribution (same frequency of occurrence on each day of the week) using the Kolmogorov-Smirnov one-sample test and found to be significantly (P=0.000) different. The two distributions were tested for differences from each other using two separate nonparametric tests. Neither the sign test (P=0.453) nor the Wilcoxon signed ranks test (P=0.498) found a significant difference in the distributions by day of the week between the historical period and the OAD forecast period.

Differences between weekday and weekend ozone concentrations have been reported since the 1970s (Cleveland et al., 1974; Lebron, 1975; Graedel et al., 1977). Recently, Altshuler et al. (1995) reported higher ozone concentrations and higher frequencies of exceedances of standards on weekends in San Francisco and other areas of northern California. On the other hand, Walker (1993) reported lower ozone concentrations and fewer exceedances on weekends than on weekdays in the Atlanta area. Based on the Empirical Kinetic Modeling Approach (EKMA), Altshuler et al. hypothesized that the presence of the weekend effect in northern California and its absence in the Southeastern United States is explained by differences in relative changes in NO_x and reactive organic gas emissions from weekdays to weekends. In simple terms, NO_x emissions are lower on weekends than on weekdays because of lower vehicle emissions. With less ozone scavenging by NO_x, weekend ozone concentrations are higher. The EKMA approach also explained why the weekend ozone effect was greater in the 1990s than in the 1980s in San Francisco.

The weekend ozone effect also has been observed in the Chicago area (LADCO, 1996), where it also has been attributed to weekend-weekday differences in emissions. The pattern of historical exceedance frequencies in Figure 4 is one of increasing frequencies from Monday through Saturday, with the highest frequencies occurring on Thursday, Friday, and Saturday. This pattern appears to be more complex than simple weekend-weekday differences. The maximum frequency occurs on Saturday, but the Sunday frequency is much lower. If lower NO_x emissions on Saturday lead to higher ozone because of less ozone scavenging by the NO_x, how do we explain the relatively high frequencies on Thursday and Friday? As Altshuler et al. (1995) emphasized, day-to-day differences in ozone concentrations and exceedance frequencies may provide information regarding the relative effectiveness of NO_x versus VOC controls in reducing ambient ozone concentrations and exceedances. Further analysis of this phenomenon in the Chicago area may be fruitful.

The suggestion that VOC controls would be more effective than NO_x controls argues for the importance of local emissions in determining ozone concentrations in the Chicago area. In this regard, it is interesting to examine the data record for indication of the degree of spatial variations in the three-county area on days with high ozone concentrations. Minimal spatial variation on days with some high concentrations would suggest that distant sources were dominant on those days, because pollutant plumes would be broad and well mixed at long distances from the sources. Conversely, high spatial variation on days with some high concentrations would imply a stronger role for local sources on those days.

To characterize spatial variation, we examined data from the 87 days between 1990 and 1997 when ozone concentrations > 100 ppb were observed at one or more of the 17 Chicago area sampling sites. On 53 (61%) of the 87 days, concentrations >100 ppb were observed at only one or two sites. On 70 (81%) of the 87 days, concentrations >100 ppb were observed at four or fewer sites. These results show that, when high ozone concentrations occur, they tend to occur at only a few sites. This would appear to be additional evidence for the strong influence of local sources on ozone concentrations in the Chicago area. Additional investigations along these lines could be fruitful.

4.2 Scatterplots

Scatterplots of daily maximum 1-hour ozone concentrations in the three-county area versus selected potential predictor variables are shown in Figure 5. In general, many of the potential predictor variables show some relationship to the ozone variable, but each relationship also exhibits considerable variability in that relationship.

A good example of this is the relationship between today's highest observed 1-hour ozone and yesterday's highest 8-hour value (Figure 5). Today's observed 1-hour ozone maximum generally increases as the previous day's 8-hour value increases, but there is wide scatter in the relationship. Some of the highest values observed on the previous day were followed by moderate-to-low values the next day. Observed maximum ozone generally increases as the long-term daily mean ozone maximum increases (Figure 5), but the scatter is especially pronounced at the higher values of the long-term mean.

Perhaps the best relationship to the observed ozone maximum is seen with maximum temperature (Figure 5). All of the violations (concentrations >125 ppb) occurred with maximum temperatures >80°F, but again the scatter is especially pronounced at the higher temperatures.

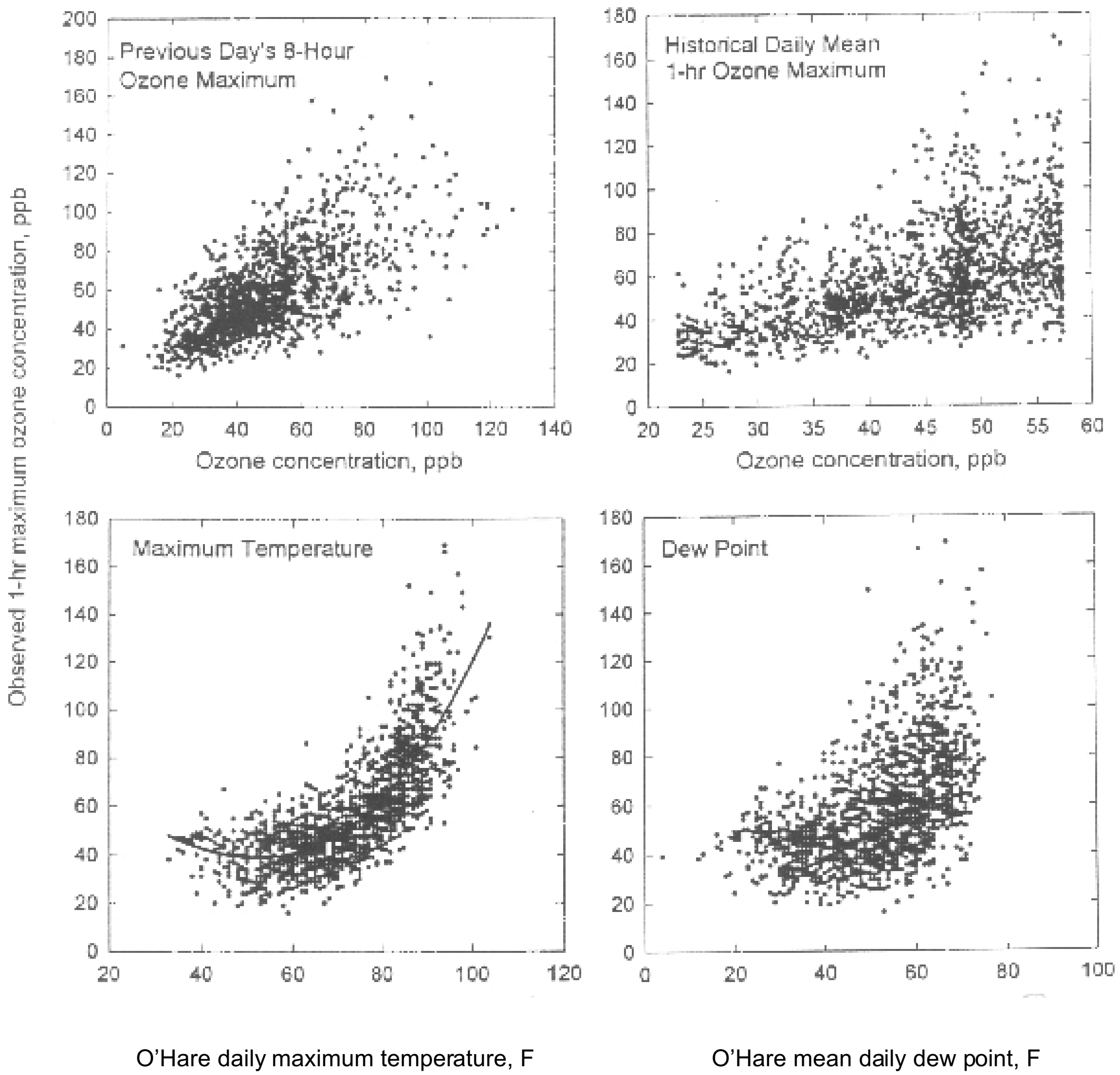

Figure 5. Scatterplots of observed 1-hour maximum ozone concentration vs. potential predictor variables in the three-county area. The curve in the maximum temperature plot is a 4th-order polnomial fit to the data (see text for discussion).

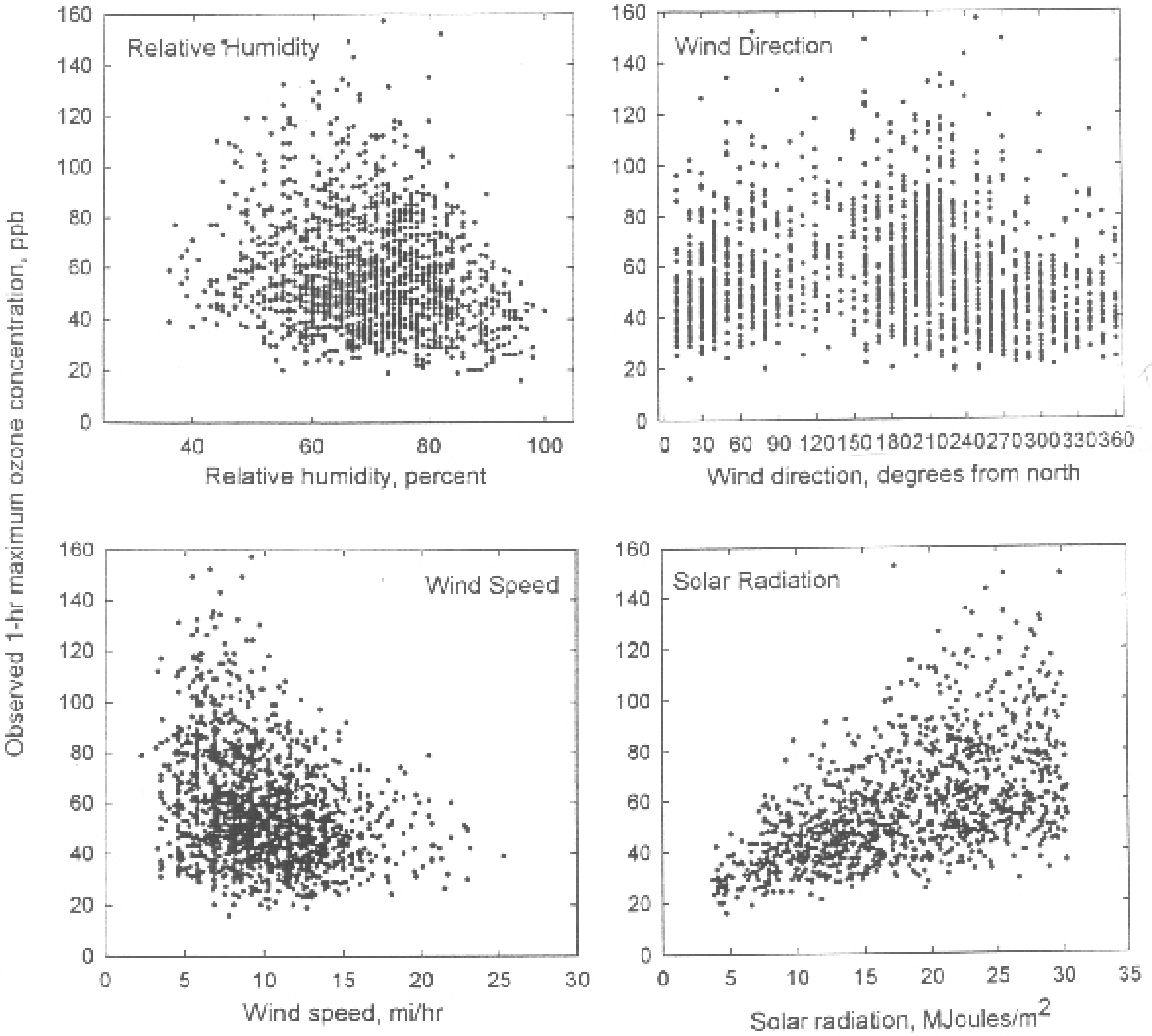

Figure 5. Continued

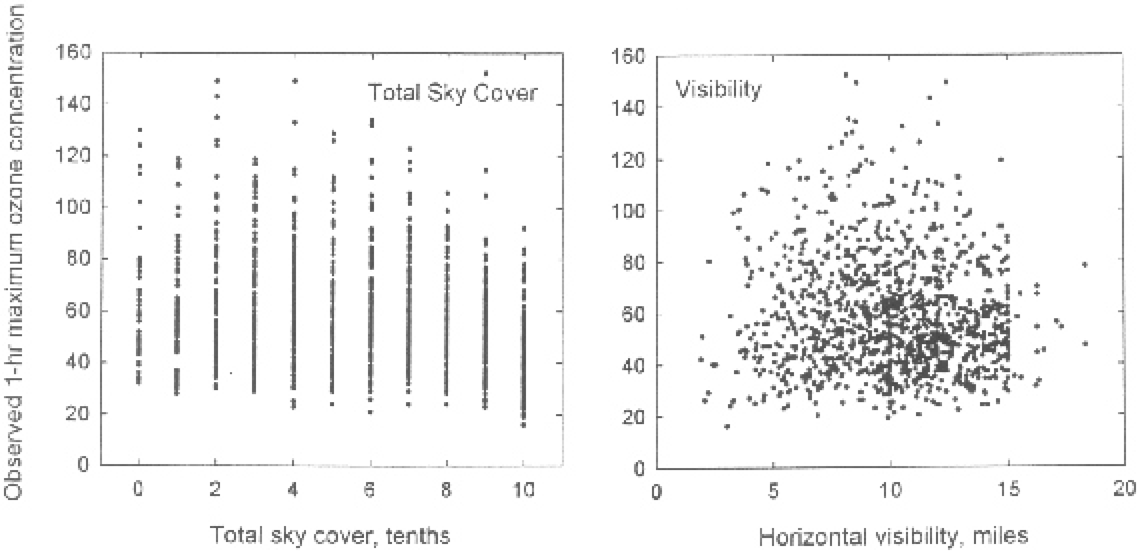

Figure 5. Concluded.

The curve in this panel of Figure 5 is a fourth-order polynomial expressing the observed ozone concentration in terms of daily maximum temperature that was fit to the data. The multiple linear regression results to be presented later will be compared against this nonlinear relationship involving only temperature as the independent variable.

Ozone concentrations also exhibit considerable scatter as a function of the dew point (Figure 5). All the violations appear to occur with dew points of 50°F or more, but the scatter of the ozone concentrations in that dew point range is very large as well.

Overall, the trend of maximum ozone concentrations versus relative humidity (RH) is negative, that is, ozone decreases as RH increases (Figure 5). The highest ozone concentrations occur at humidity values between about 60 and 80%, and ozone maxima decrease substantially as RH increases to 100%.

The variation of maximum ozone concentrations with wind direction is relatively weak. Although we expect high ozone on hot days with winds from the south or southwest when the Chicago area is on the back (west) side of a high pressure center, the data in Figure 5 show that violations (concentrations >125 ppb) occur with a wide range of wind directions from east through south to west. The relationship between maximum ozone and wind speed is inverse and also quite strong, with most violations occurring at wind speeds less than 10 miles per hour (mi/hr).

Maximum ozone concentrations generally increase with solar radiation values (Figure 5), and violations occur with solar radiation greater than about 20 millijoules per square meter (MJ/m^2). Daily fluxes of solar radiation are not routinely forecast, however. Most ozone violations occur with total sky cover of six-tenths or less, and the general trend is for lower ozone concentrations as sky cover increases, but there is considerable scatter in the relationship.

Most ozone violations appear to occur with visibility between about 7 and 13 miles. Ozone concentrations are lower at lower and higher visibilities. Scatter is considerable.

4.3 Back Trajectories on High-Ozone Days

As an aid to understanding airflow regimes prior to the occurrence of high ozone concentrations, it is useful to examine back trajectories. Because ozone forecasts of 1-2 days were likely to be the most useful for episodic control, two-day back trajectories were calculated. Figure 6 shows 2-day back-trajectories from Chicago on all 59 days between 1991 and 1995 when the maximum 1-hour ozone concentration in the three-county area was ≥100 ppb. On 14 days (24%) the 2-day back-trajectory began in Illinois. On 12 days (20%) the back-trajectory began in Missouri; but on most of these occasions, the air also spent much time over Illinois. On nine days (15%) the back-trajectory began in Indiana, and on some of these occasions the air also spent many hours over Illinois. For seven of the occasions (12%) the back-trajectory began in

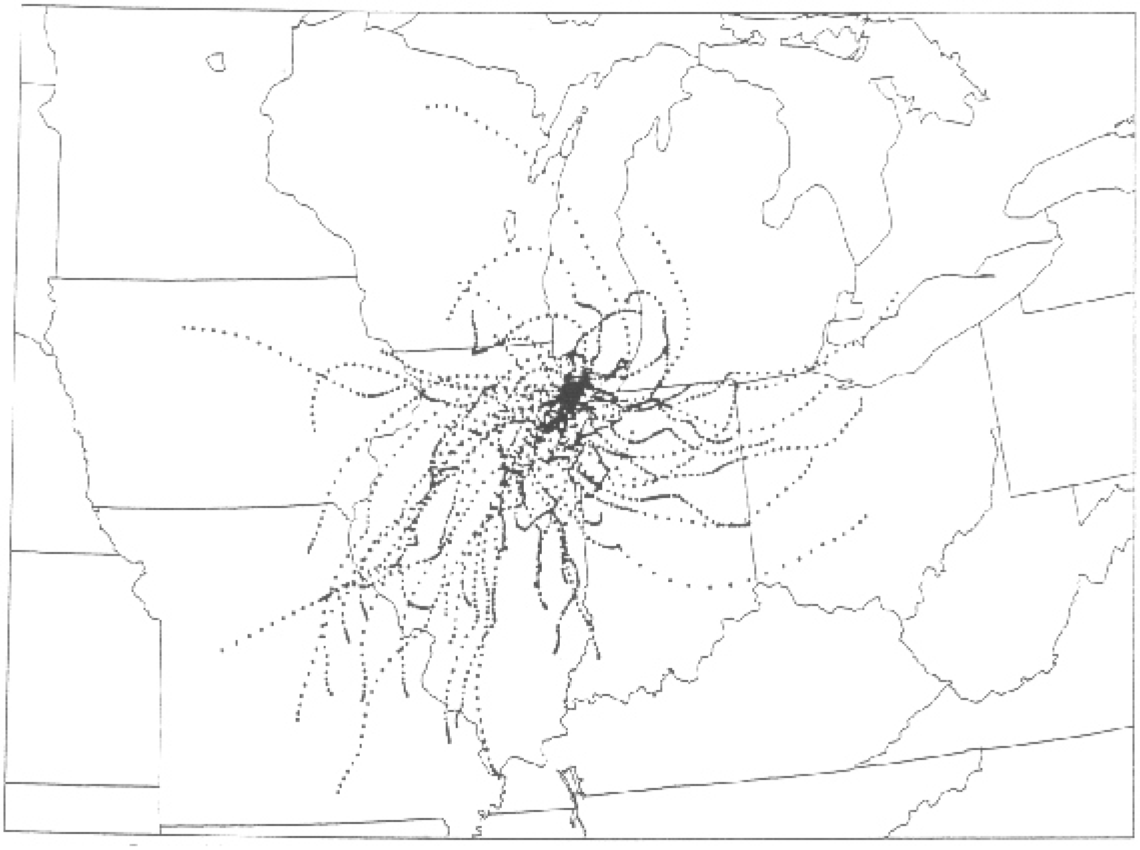

Figure 6. The 48-hour back trajectories for all 59 days between 19919 and 1995 with maximum 1-hour ozone concentrations .100 ppb in Lake, Cook, or DuPage Counties.

Wisconsin. On six occasions the back-trajectory began in Ohio, and in most of these cases the air crossed Indiana as well. In four cases the back-trajectory began over Michigan, and in another four began over Lake Michigan. In some of these cases, the air also crossed Indiana before arriving at Chicago. Iowa was the starting point of the back-trajectory in two cases, and the Province of Ontario, east of Detroit was the origin in one case.

Somewhat surprisingly, the origins of two-day back-trajectories were more uniformly distributed than might have been suspected from the general model of high ozone with Chicago on the back side of a high pressure center located somewhere in the Southeastern United States. The air came from Missouri or Illinois in about 45% of the cases. Ohio and Indiana were the origin in 25% of the cases, and Wisconsin or Michigan were the origin in about 20%. This distribution is consistent with the lack of a strong relationship with surface wind direction seen in the scatterplot in Figure 5. Typically, high ozone concentrations do not occur with trajectories from the north or northwest; however, which also agrees with the surface wind direction analysis in Figure 5. Another feature of the back-trajectories is that, in many of them, the air did not travel very far over the two days, so the winds were generally light. This also agrees with the scatterplot of ozone versus wind speed in Figure 5, which showed that the highest ozone concentrations occurred with surface winds of less than 10 mi/hr.

4.4 Ozone Precursor Emissions

Ambient ozone concentrations depend on precursor emission fluxes as well as weather conditions. If the data used to develop ozone forecasting methods were collected during a period of rapidly changing precursor emissions, the forecasting method could be compromised. Thus, as general background information for our effort to develop statistical forecasting methods, we gathered information on trends of annual emissions. With the time and resources available, it was possible to gather nationwide emissions data. Regional emissions data would have been preferable, but the national data are considered to be adequate for the intended purpose. Annual emissions data are, of course, not intended for use in short-term forecasting of pollutant concentrations, and they were not gathered for that purpose. For use in short-term forecasting, actual emissions over periods of 24 hours or less would be required. Such data are not currently available.

Emission inventories for VOCs and for NO_x, precursors for ozone, were obtained from the USEPA. Table 5 presents the emission values for VOCs by source type for four time periods. Note that natural sources are not included. The total VOC emissions have decreased substantially over the 1970-1995 period. From 1990-1995 there was only a slight decrease (3.1%) in the total; the largest decreases were for the waste disposal and recycling and the miscellaneous categories; and the largest increase for the on-road vehicle category. Natural sources are not likely to have changed substantially in recent years.

Table 5. Volatile Organic Compound Emissions by Principal Source Category for the United States.

Principal Source	*Volatile Organic Compound Emissions* *(thousands of short tons)* 1970	1980	1990	1995
Solvent Utilization	7,174	6,584	5,975	6,394
On-Road Vehicles	12,972	8,979	6,854	6,104
Waste Disposal & Recycling	1,984	758	2,262	2,411
Non-Road Sources	1,542	1,869	2,120	2,252
Storage & Transport	1,954	1,975	1,759	1,803
Chemical & Allied Product	1,341	1,595	1,526	1,617
Petroleum & Related Industries	1,194	1,440	643	628
Fuel Combustion, Other	541	848	749	539
Miscellaneous	1,101	1,134	1,069	446
Other Industrial Processes	270	237	401	422
Fuel Combustion, Industrial	150	157	135	135
Metals Processing	394	273	72	77
Fuel Combustion, Electric	30	45	36	35
Totals	30,647	25,894	23,601	22,863

Table 6 presents the emission values for NO_x for four time periods. Again, natural sources are not accounted for in these figures, but they are not likely to have changed substantially. The total NO_x emissions increased from 1970-1980, then decreased from 1980-1995. From 1990-1995 there were small increases in the on-road and non-road categories, a small decrease in the fuel combustion (industrial) category, and a large decrease in the fuel combustion (electric utilities) category. The overall decrease in NO_x emissions from 1990-1995 was about 5.5%. These small changes in precursor emissions during the 1990-1996 data period are not likely to have substantially impacted the development of forecast equations.

4.5 Regression Model Results

Stepwise regression analyses were carried out, as described earlier, using the 70% data for the three-county area. The dependent variables to be estimated were the maximum daily 1-hour and 8-hour ozone concentrations observed at any of the 17 sampling sites in Lake, Cook, or DuPage Counties. Separate regressions were carried out using the five approaches and the baseline case described earlier.

Hubbard and Cobourn (1998) indicate that, to obtain optimal statistical properties using ordinary least squares regression, the regression model should satisfy three assumptions: 1) the model is correctly specified with respect to retaining the important variables and utilizing appropriate functional transformations when a straight-line fit is inadequate; 2) model error terms have zero means, are uncorrelated, and have constant variances; and 3) model error terms follow

Table 6. Nitrogen Oxide Emissions by Principal Source Category for the United States.

Principal Source	*Nitrogen Oxide Emissions* *(thousands of short tons)* 1970	1980	1990	1995
Solvent Utilization	NA	NA	2	3
On-Road Vehicles	7,390	8,621	7,488	7,605
Waste Disposal & Recycling	440	111	82	85
Non-Road Sources	1,628	2,423	2,843	2,996
Storage & Transport	NA	NA	2	3
Chemical & Allied Product	271	216	276	283
Petroleum & Related Industries	240	72	100	91
Fuel Combustion, Other	836	741	712	707
Miscellaneous	330	248	373	228
Other Industrial Processes	187	205	306	323
Fuel Combustion Industrial	4,325	3,555	3,256	3,137
Metals Processing	77	65	81	84
Fuel Combustion, Electric	4,900	7,024	7,516	6,233
Totals	20,624	23,281	23,037	21,778

a normal probability distribution (Gunst and Mason, 1980). A square root transformation was applied to the dependent variable to improved the normality and homoscedasticity (constant variance) of the residuals. This transformation increased the R^2 of the regression, and satisfied the necessary conditions for the ordinary least squares fit.

4.5.1 One-hour Ozone Concentrations

Table 7 shows results of the multiple linear regression using all potential predictor variables (Approach 0). The table gives the regression coefficient and its standard error for each of the seven variables selected. The table also shows the t-statistic and the probability associated with each variable. Large t-statistics and a small (<0.10) probability indicate that the variable contributes significantly to the regression equation. The variables are listed in order of the absolute value of their t-statistic.

Because of correlations among the variables, the signs of the regression coefficients do not always agree with the bivariate relationships between dependent and independent variables indicated by the scatter plots (Figure 5). To minimize this effect, some (e.g., Bloomfield et al., 1996) have rescaled variables around their mean values. In this analysis, we eliminated as potential variables some parameters (e.g., total opaque sky cover) that were highly correlated with others (total sky cover). The signs of the coefficients of most variables in Table 7

Table 7. Regression Results for 1-hour Maximum Ozone Concentrations in the Three-county Area Using All Available Predictor Variables (Approach 0). F-ratio = 276, R^2 = 0.686.

Variable	*Coefficient*	*Standard error*	*t-statistic*	*Probability of larger |t|*
Constant	4.789	0.382	12.536	0.000
Maximum temperature	0.044	0.003	15.919	0.000
8-hr ozone, lag 1 day	0.022	0.002	11.772	0.000
Solar radiation	0.041	0.006	7.032	0.000
Horizontal visibility	-0.067	0.011	-5.895	0.000
Relative humidity	-0.018	0.003	-5.732	0.000
Wind speed	-0.044	0.008	-5.550	0.000
Weekday/weekend	0.153	0.059	2.615	0.009

correspond to their relationships to ozone concentrations plotted in Figure 5. The positive signs on the coefficients for weekday/weekend, temperature, solar radiation, and lagged ozone concentrations, and the negative signs on those of relative humidity and wind speed, indicate the expected direct and inverse relationships, respectively. The scatter plot for visibility (Figure 5), which has a negative coefficient, shows little evidence for either a direct or inverse relationship with ozone concentrations.

In terms of variables with the highest t-values, results in Table 8 are similar to those of Table 7, bearing in mind that solar radiation was excluded from the data set of Approach 1. Among the variables with the smallest t-values, Table 8 includes the daily mean, total sky cover, and the u-component, in addition to wind speed and weekday/weekend. The signs (positive or negative) of the regression coefficients are mostly what one would expect from the relationships between maximum ozone concentrations and the variables, shown in Figure 5. Indeed, the magnitudes of the coefficients are also rather similar to those in Table 7, for those variables present in both Tables 7 and 8. This is not really surprising, because only the solar radiation variable was dropped from the input data set as not routinely forecasted.

Because preliminary results showed a tendency to underpredict the highest ozone concentrations, it seemed potentially useful to carry out a multiple regression analysis on the subset of samples with the highest 1-hour maximum ozone concentrations. Table 9 shows results

Table 8. Regression Results for 1-hour Ozone in the Three-county Area Using Only Forecasted Variables (Approach 1). F-ratio = 199, R^2 = 0.668.

Variable	*Coefficient*	*Standard error*	*t-statistic*	*Probability of larger |t|*
Constant	5.699	0.373	15.267	0.000
Maximum temperature	0.046	0.003	14.371	0.000
8-hr ozone, lag 1 day	0.020	0.002	10.371	0.000
Relative humidity	-0.027	0.003	-8.723	0.000
Horizontal visibility	-0.071	0.012	-6.079	0.000
Wind speed	-0.040	0.009	-4.674	0.000
Daily mean O_3	0.015	0.004	3.584	0.000
Total sky cover	-0.035	0.011	-3.029	0.003
Weekday/weekend	0.143	0.060	2.367	0.061
U-component	-0.007	0.004	-1.876	0.045

for maximum ozone concentrations of 80 ppb or more, using forecast variables only. Again, the signs of the coefficients are much as expected from the scatterplots of Figure 5.

Results of testing the stepwise regression equations on the independent 30% of the data are shown in Table 10. The first column lists the data used to develop the regression equations for the various approaches described earlier. All five approaches are listed, including the baseline method (Approach 0), which used all samples in the 70% data set, and all potential predictor variables. Approaches 1-4 used only variables for which routine forecasts are available. Of all the available potential predictor variables, only solar radiation was dropped from consideration because it is not routinely forecast.

The second and third columns of Table 10 list the type of independent predictor variables used to test the ozone concentrations computed from the regression equations against observed ozone concentrations, and the number of samples, N. The variable types are either observed variables, or forecasted variables, or a combination of the two. The only variable for which actual forecasted values were used was maximum temperature T_{max}. When the "Type" column indicates that the forecasted T_{max} was used, observed values were used for any other variables in the equation. The fourth-order polynomial is the only method that used exclusively forecasted

Table 9. Regression Results for Days with Maximum 1-hour Ozone Concentrations ≥ 80 Ppb in the Three-county Area, Using Only Forecasted Variables (Approach 3). F-ratio = 14, R^2 = 0.409.

Variable	*Coefficient*	*Standard error*	*t-statistic*	*Probability of larger |t|*
Constant	6.456	0.977	6.607	0.000
Wind speed	-0.089	0.023	-3.928	0.000
Maximum temperature	0.036	0.011	3.268	0.001
8-hr ozone, lag 1 day	0.009	0.003	3.154	0.002
Horizontal visibility	-0.061	0.021	-2.903	0.004
Total sky cover	-0.066	0.024	-2.792	0.006
Weekday/weekend	0.283	0.115	2.454	0.016
Daily mean O_3	0.022	0.012	1.852	0.066

variables, i.e., the single variable T_{max}. The full 30% independent data set consisted of 395 samples. However, forecasted T_{max} values were available for only 276 of these samples.

The next two columns show the R^2 between the observed and predicted ozone concentrations, and the root mean square error, which is equivalent to the standard deviation of the differences between the observed and predicted ozone concentration. The last three columns report results in terms of the number of exceedances of the daily maximum 1-hour ozone standard predicted correctly, as well as the numbers for which the prediction was too high (overprediction) or too low (underprediction).

These results may be compared with previous results from the literature (Table 2) in terms of the same parameters. The range of R^2 values (0.618-0.711) obtained from the various methods applied to the full (N=395) independent data set, using observed values of the variables, falls toward the upper end of the range of literature values (0.24-0.92) for observed meteorology. The enhancement method of Hubbard and Cobourn (1998) achieved the highest R^2 value, at 0.711. The Clark and Karl (1982) method produced an R^2 value of 0.693; this is equal to that of the regression based on forecast variables only, which produces the predicted ozone concentrations that both enhancement techniques are applied to. The fourth-order polynomial gave an R^2 value only slightly lower, at 0.670. The regression equation developed from the samples with the highest ozone concentrations produced the lowest R^2 value, at 0.618.

Table 10. Tests of Regression Equations for 1-hour Maximum Ozone Concentrations on the Independent 30% Data Set.

	Test data		*Results*		*Exceedance predictions*		
Portion of 70% data used to develop the regression equation (Approach)	*Type*	*N*	*R^2*	*RMSE (ppb)*	*Correct*	*Too high*	*Too low*
All samples, all variables (Approach 0)	Obsvd	395	0.708	12.55			
All samples, forecast variables (Approach 1)	Obsvd	395	0.697	12.83	0	0	8
	Obsvd T_{max}	276	0.696	12.35			
	1-day T_{max} forecast	276	0.656	13.14			
	2-day T_{max} forecast	276	0.658	13.12			
	3-day T_{max} forecast	276	0.640	13.44			
All samples, forecast variables, enhanced using method of Hubbard & Cobourn (1998) (Approach 2a)*	Obsvd	395	0.711	12.44	1	1	7
All samples, forecast variables, enhanced using method of Clark & Karl (1982) (Approach 2b)	Obsvd	395	0.693	12.98	1	1	7
Samples with maximum 1-hr ozone ≥80 ppb (Approach 3)	Obsvd	395	0.618	23.45	0	0	8
	Obsvd T_{max}	276	0.600	23.74			
	1-day T_{max} forecast	276	0.556	24.14			
	2-day T_{max} forecast	276	0.548	24.18			
	3-day T_{max} forecast	276	0.526	24.22			
All samples, 4th-order polynomial in T_{max} (Approach 4)	Obsvd T_{max}	395	0.670	13.32	1	0	7
	Obsvd T_{max}	276	0.685	12.54			

Table 10. Concluded.

	Test data		*Results*		*Exceedance predictions*		
Portion of 70% data used to develop the regression equation (Approach)	*Type*	*N*	R^2	*RMSE (ppb)*	*Correct*	*Too high*	*Too low*
	1-day fcst	276	0.588	14.29			
	2-day fcst	276	0.567	14.66			
	3-day fcst	276	0.508	15.58			

Note: * Since actual ozone levels must be known to employ approach 2a, its use for prospective forecasting depends on the relationship between projected and observed ozone being consistent over time.

When observed variables were used on the partial data set (N=276), the respective R^2 values were slightly higher for the fourth-order polynomial method (0.685 versus 0.670), but slightly lower (0.600 versus 0.618 and 0.696 versus 0.697) for the other two methods in which such a comparison can be made. Generally, R^2 decreased when observed values of T_{max} were replaced with forecasted values, and it decreased further as the forecast target day increased from one to three days in advance.

The RMSE values in Table 10 fall about in the middle of the range of those in Table 2 for observed meteorology (8-17 ppb), except for those associated with the regression equation based on samples with $O_3 \geq 80$ ppb, which are much higher (23.4 - 24.2). In general, RMSE values in-creased as R^2 values decreased, as expected. Ranking the various methods in terms of low RMSE values produces the same order as listed above based on high R^2. Note that the RMSE values associated with the regression equation developed from the subset of highest ozone concentrations were much higher than the others; the reason for this will be clear when we discuss Figure 7.

Performance of the various ozone forecasting techniques with respect to exceedance predictions is also shown in Table 10. The best performance was one exceedance correctly predicted and seven underpredicted, by the fourth-order polynomial in T_{max}. The two enhancements of the regression for all samples with only forecasted variables also had one correct prediction and seven underpredictions, but they both also had one overprediction. The remaining two methods both had no correct exceedance predictions, and eight underpredictions.

Performance is also indicated in Figure 7, which shows scatter plots of predicted versus observed 1-hour maximum ozone concentrations for four prediction approaches. Of the two enhancement methods in Table 10, only the Hubbard and Cobourn (1998) method was chosen to be included in Figure 7, because of its slightly higher R^2 and slightly lower RMSE, as compared

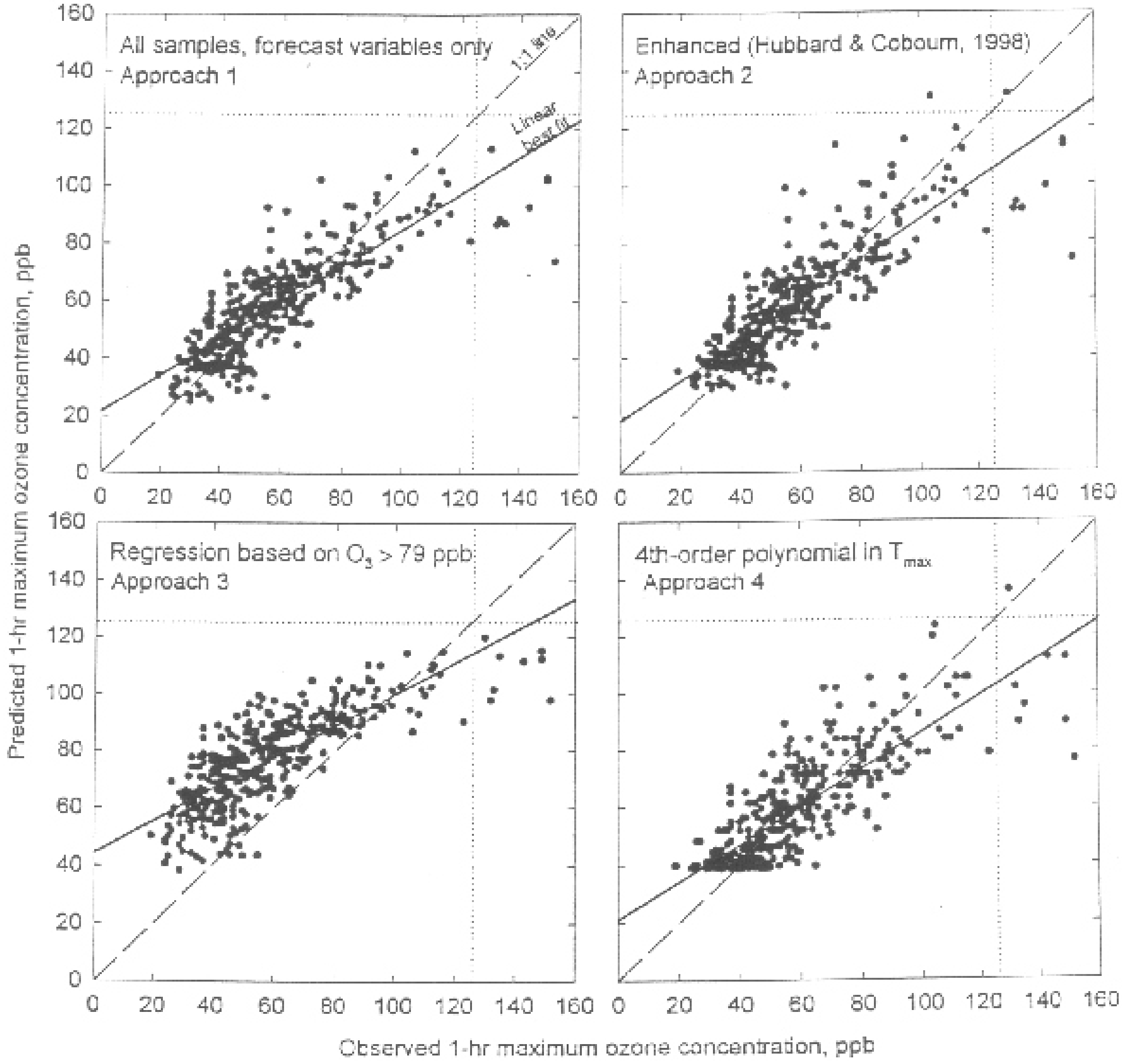

Figure 7. Predicted vs. observed 1-hour maximum ozone concentrations for Lake, Cook, and DuPage Counties, based on four different prediction methods. Dotted lines at the 1-hour standard of 125 ppb on both the observed and predicted axes help identify correct predictions as well as over- and under predictions. Predictions were based on multiple linear regression equations developed from 70% of the available ozone season database for 1990-1995.

to the Clark and Karl (1982) method. As noted above, all of the methods produced R^2 values between 0.60 and 0.71, with three of the four falling at the upper end of that range.

The sloping solid lines in each panel are best-fit regression lines for predicted versus observed concentrations. Each panel of Figure 7 shows a clear relationship between observed and predicted maximum ozone concentrations, with considerable scatter about the regression line. However, the main result desired, an accurate forecast of ozone standard violations (i.e., concentrations ≥ 125 ppb), was not achieved; the high-ozone cases were clearly underpredicted by all four approaches. Recall also that the plotted results are based on observed values of the meteorological input variables. Use of forecasted values of all variables can be expected to yield forecasts with even greater uncertainty. The results above (Table 10), when just the forecasted value of the maximum temperature was used instead of the observed value, provide just an indication of the minimum degradation in the prediction that could be expected if forecasted values of all the meteorological variables were used, as they would be in an operational ozone forecasting situation.

Clearly, use of a subset of samples with high ozone concentrations to develop the regression equation was not effective. Careful comparison of the upper left and lower left panels in Figure 7 shows that the regression based on the subset of samples with high ozone concentrations does predict higher concentrations, but not high enough. Further, the cases of low observed ozone concentrations were substantially overpredicted. This is a typical result when a prediction equation developed from a narrow range of independent variables is applied using variable values outside that range. The large overpredictions of low ozone values were also the cause of the high RMSE values for this approach. Note that the nonlinear approach (4) using only T_{max} had rather similar results to the other approaches. We discuss relative forecast skill later.

4.5.2 Eight-hour Ozone Concentrations

Table 11 shows results of the multiple linear regression analysis to predict 8-hour ozone concentrations, using all potential predictor variables (Approach 0). The variables chosen here are very similar to those of the corresponding 1-hour regression (Table 7). The seven variables with the highest t-values are the same in both cases, but the 8-hour results also include total sky cover and both wind components among the lower t-values. Signs of the coefficients are much as expected; the magnitudes of the coefficients are similar to those for the same variables in the equation for the 1-hour ozone concentrations (Table 7).

Table 12 gives results of the regression analysis using only forecasted variables. The variables selected are the same as those for the corresponding equation for 1-hour ozone, although the order is slightly different here. Again, the magnitudes of the corresponding coefficients in the 1-hour and 8-hour equations are similar.

Table 11. Regression Results for 8-hour Maximum Ozone Concentrations in the Three-county Area Using All Available Predictor Variables (Approach 0). (A Square Root Transformation Was Applied to the Dependent Variable.) F-ratio = 181, $R^2 = 0.670$.

Variable	*Coefficient*	*Standard error*	*t-statistic*	*Probability of larger |t|*
Constant	4.969	0.363	13.692	0.000
Maximum temperature	0.033	0.003	11.260	0.000
8-hr ozone, lag 1 day	0.018	0.002	9.977	0.000
Solar radiation	0.059	0.006	9.182	0.000
Relative humidity	-0.024	0.003	-7.663	0.000
Horizontal visibility	-0.072	0.011	-6.781	0.000
Weekday/weekend	0.206	0.055	3.739	0.000
Wind speed	-0.020	0.008	-2.613	0.009
Total sky cover	0.021	0.012	1.842	0.066
V-component	0.007	0.004	1.831	0.067
U-component	-0.006	0.003	-1.741	0.082

Table 13 gives results of the regression analysis based on days with 8-hour mean concentrations of 60 ppb or more. Many of the chosen variables are the same as in the equation based on 1-hour ozone concentrations ≥ 80 ppb, but RH appears only here, where it achieved the second-highest t-statistic.

Results of testing the equations for 8-hour ozone concentrations on the independent 30% of the data are shown in Table 14, which has the same format as the 1-hour results in Table 10. The R^2 values in Table 14 are slightly lower than those in Table 10 for the respective test data of Approaches 1, 2, and 4; but they are mostly slightly higher for those of Approach 3. The RMSE values in Table 14 are somewhat lower than those in Table 10 for all five approaches, which indicates greater overall accuracy of the predictions for 8-hour ozone. Compared to Table 10, the RMSE values of Approach 3 in Table 14 are lower, although still highest of all five approaches. As earlier, R^2 decreased and RMSE increased when forecast T_{max} was used instead of observed T_{max}. The same changes occurred for Approach 4 as forecast target time increased from one to

Table 12. Regression Results for 8-hour Ozone in the Three-county Area Using Only Forecasted Variables. (A Square Root Transformation Was Applied to the Dependent Variable.) F-ratio = 185, R^2 = 0.652.

Variable	*Coefficient*	*Standard error*	*t-statistic*	*Probability of larger |t|*
Constant	6.032	0.351	17.173	0.000
Relative humidity	-0.035	0.003	-11.916	0.000
Maximum temperature	0.035	0.003	11.728	0.000
8-hr ozone, lag 1 day	0.018	0.002	9.544	0.000
Horizontal visibility	-0.073	0.011	-6.657	0.000
Daily mean O_3	0.023	0.004	5.853	0.000
Weekday/weekend	0.208	0.057	3.670	0.000
Total sky cover	-0.035	0.011	-3.263	0.001
Wind speed	-0.021	0.008	-2.645	0.008
U-component	-0.006	0.003	-1.736	0.083

three days. For Approaches 1 and 3 there were only slight changes in either direction as forecast target time increased.

The performance of the regression equations with respect to predicting exceedances of the 85 ppb 8-hour standard were markedly better than those for the 1-hour standard. Table 14 shows that, of 19 observed exceedances, Approach 1 had eight of them correct, with 11 underpredictions. Approach 1 also had one overprediction. Approaches 2 and 3 did even better, with nine or ten correct predictions of exceedances, and nine or ten underpredictions. Approach 3 also had five overpredictions, while Approach 2a (Hubbard and Cobourn, 1998) had seven and Approach 2b (Clark and Karl, 1982) had ten. Approach 4 produced six correct predictions, 13 underpredictions, and four overpredictions. Although the number of underpredictions still exceeded the number of correct predictions, their ratio is close to one for a number of these approaches, which is a considerably better ratio than was obtained in the case of the predictions for 1-hour ozone (Table 10). Forecast skill for the various approaches will be discussed later.

Figure 8 shows results for four approaches in terms of scatter plots of predicted versus observed 8-hour ozone concentrations. The improvement in terms of predicting exceedances is readily apparent in Figure 8, with more cases falling in the upper right "correct prediction"

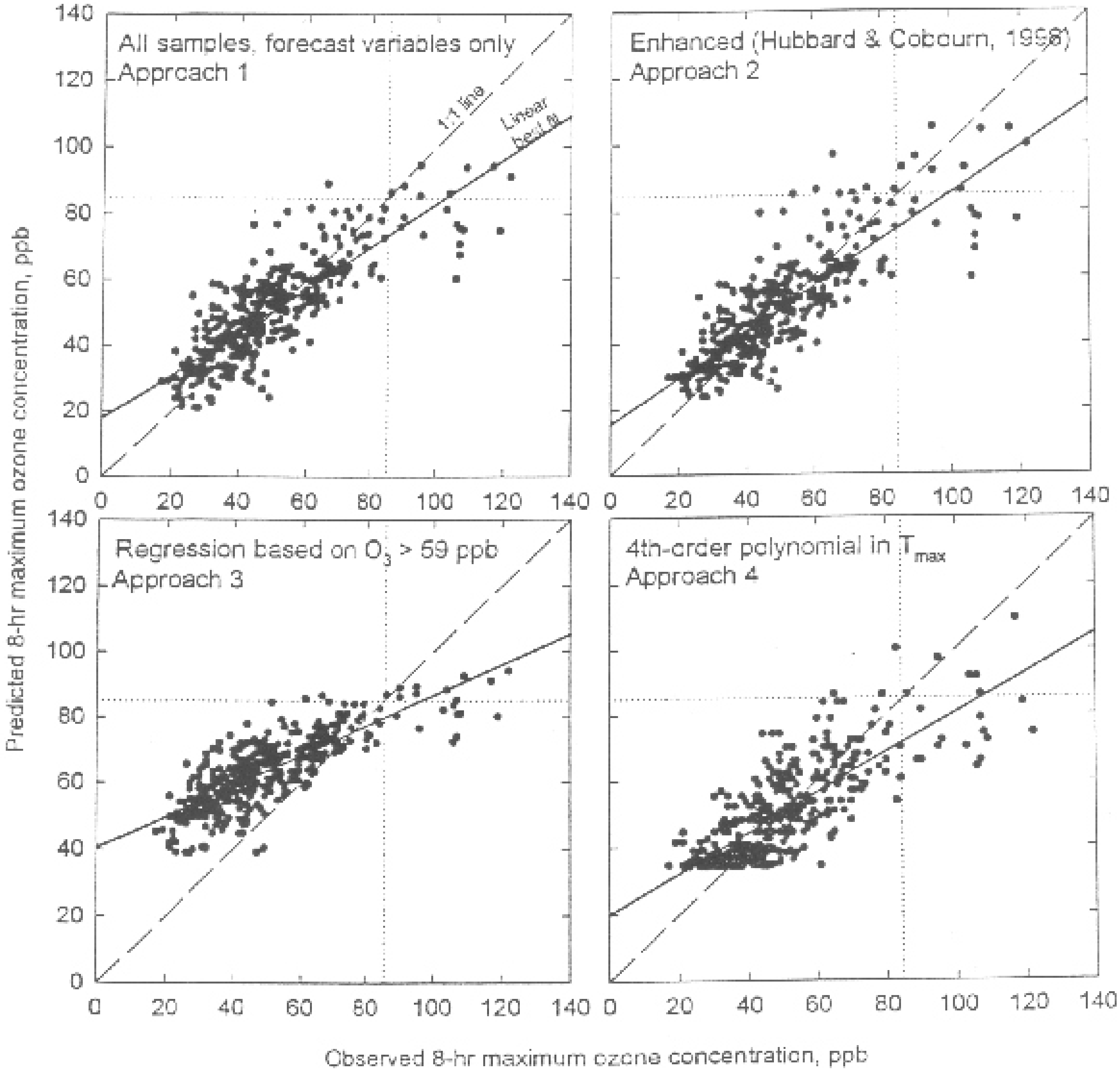

Figure 8. Predicted vs. observed 8-hour maximum ozone concentrations for Lake, Cook, and DuPage Counties, based on four different prediction methods. Dotted lines at the 8-hour standard of 85 ppb on both the observed and predicted axes help identify correct predictions as well as over- and under predictions. Predictions were based on multiple linear regression equations developed from 70% of the available ozone season database for 1990-1995.

Table 13. Regression Results for Days with Maximum 8-hour Ozone Concentrations $\geq$ 60 Ppb in the Three-county Area, Using Only Forecasted Variables (Approach 3). (A Square Root Transformation Was Applied to the Dependent Variable.) F-ratio = 20, R^2 = 0.417.

Variable	*Coefficient*	*Standard error*	*t-statistic*	*Probability of larger* $\lvert t \rvert$
Constant	7.960	0.636	12.513	0.000
Relative humidity	-0.021	0.004	-4.957	0.000
Wind speed	-0.053	0.013	-4.232	0.000
8-hr ozone, lag 1 day	0.009	0.002	4.128	0.000
Horizontal visibility	-0.053	0.016	-3.369	0.001
Maximum temperature	0.018	0.006	3.036	0.003
Weekday/weekend	0.232	0.078	2.975	0.003
Daily mean O_3	0.020	0.007	2.938	0.004
Total sky cover	-0.041	0.015	-2.701	0.007

quadrant formed by the intersecting dotted lines. Results for all approaches were better for 8-hour than 1-hour ozone. For 8-hour ozone, Approaches 2 and 3 seem to be particularly useful. The nonlinear approach appears promising in that it was able to achieve results almost as good as the best using only a single predictor variable. However, all of the results in Figure 8 were based on observed values of the predictor variables. The use of forecast values of the variables in an operational situation would not be expected to be as successful.

4.5.3 Measures of Forecasting Skill

The thrust of this work has been to forecast exceedances of the National Ambient Air Quality Standards (NAAQS) for ozone, so it is useful to express the results in terms of the success in categorical forecasting. Ryan (1995) listed seven separate measures of forecasting skill, all based on a contingency table like that of Table 15. Note that the A, B, C, and D values in Table 15 correspond to the number of events (solid circles) occurring in the quadrants defined by the intersecting dotted lines in Figures 7 and 8. The number of cases in the upper right quadrant, for which exceedances were both predicted and observed, corresponds to A. The number of cases in the lower right quadrant corresponds to B, in which an exceedance occurred

Table 14. Tests of Regression Equations for 8-hour Maximum Ozone Concentrations on Independent Data Set.

	Test data		*Results*		*Exceedance predictions*		
Data used to develop the regression equation	*Type*	*N*	R^2	*RMSE (ppb)*	*Correct*	*Too high*	*Too low*
All samples, all variables (Approach 0)	Obsvd	395	0.691	10.48			
All samples, forecast varbls (Appr. 1)	Obsvd	395	0.675	10.77	8	1	11
	Obsvd T_{max}	276	0.678	10.61			
	1-day T_{max} forecast	276	0.646	11.12			
	2-day T_{max} forecast	276	0.648	11.06			
	3-day T_{max} forecast	276	0.642	11.17			
All samples, forecast variables, enhanced using method of Hubbard & Cobourn (1998) (Approach 2a)*	Obsvd	395	0.687	10.55	9	7	10
All samples, forecast variables, enhanced using method of Clark & Karl (1982) (Approach 2b)	Obsvd	395	0.675	11.18	10	10	9
Samples with maximum 8-hr ozone ≥ 60 ppb (Approach 3)	Obsvd	395	0.631	18.36	10	5	9
	Obsvd T_{max}	276	0.620	18.67			
	1-day T_{max} forecast	276	0.596	18.84			
	2-day T_{max} forecast	276	0.597	18.82			
	3-day T_{max} forecast	276	0.595	18.81			
All samples, 4th-order polynomial in T_{max} (Approach 4)	Obsvd T_{max}	395	0.620	11.62	6	4	13
	Obsvd T_{max}	276	0.624	11.43			

Table 14. Concluded.

Data used to develop the regression equation	Test data		Results		Exceedance predictions		
	Type	*N*	R^2	*RMSE (ppb)*	*Correct*	*Too high*	*Too low*
	1-day fcst	276	0.523	12.85			
	2-day fcst	276	0.495	13.24			
	3-day fcst	276	0.455	13.77			

Notes: * Since actual ozone levels must be known to employ approach 2a, its use for prospective forecasting depends on the relationship between projected and observed ozone being consistent over time. N is the number of cases, R^2 is the square of the correlation coefficient., and RMSE is the root mean square error.

but was not forecast. The number of cases in the upper left quadrant corresponds to C, in which an exceedance was predicted, but did not occur. The number of cases in the lower left quadrant corresponds to D, in which exceedances were not predicted and none occurred.

Appendix 3 (from Ryan, 1995) provides definitions of the following seven measures of forecasting performance in terms of the entries A, B, C, and D in the contingency table. Briefly, the probability of detection (POD) expresses the ability of a given method to forecast an exceedance accurately. It is simply the fraction of all exceedances that were correctly forecast, and it ranges from 0 to 1. The false alarm rate (FAR) measures overprediction. It is the fraction of all exceedance forecasts in which an exceedance did not occur. It ranges from 0 to 1. The miss rate (MISS) measures underprediction. It is equal to 1-POD, or the fraction of all exceedances that were not correctly forecast; its range is 0 to 1. The rate of correct predictions of nonexceedances is given by correct null forecasts (CNULL). It is the fraction of nonexceedances that were correctly forecast, and ranges from 0 to 1. Three additional measures of forecasting skill are the critical success index (CSI), the true skill score (TSS), and the Heidke skill score (S). The CSI combines forecast occurrences and observed occurrences without regard to successful null forecasts; it is the ratio of correct exceedance forecasts to the sum of the observed exceedances plus the number of exceedances forecast but not observed.

The TSS includes the success of null forecasts in the form of a ratio of observed skill to perfect forecast skill. This measure is not dependent on the relative frequency of occurrence and nonoccurrence or the number of trials. If all forecasts are correct, TSS = 1; if all are incorrect, TSS = -1.

For rare events, the S is appropriate. This is a measure of the skill of a set of forecasts compared to the skill of a random forecast.

Table 15. Contingency Table Used for Verification of Ozone Forecasts.

Exceedance	*Forecast of exceedance*	*Forecast of non-exceedance*
Observed	A	B
Not observed	C	D

Notes: A is the number of cases where an exceedance was both forecast and observed. B is the number of cases in which an exceedance occurred when it was not forecast. C is the number of cases where an exceedance was forecast but none occurred. D is the number of cases in which no exceedance was forecast and none occurred.

Values of these measures for predictions of 1-hour maximum ozone concentrations are given in Table 16. This table includes results for two types of verifications, defined by the thresholds used for the prediction and the observation. Typically, these thresholds are the same; when the equation gives a maximum 1-hour ozone concentration of 125 ppb or greater, we forecast an exceedance. If the observed maximum 1-hour ozone concentration is at least 125 ppb, the forecast verifies. However, because the regression equations tend to underpredict ozone, it is also useful to examine how often an exceedance occurs when we forecast an exceedance and the regression equation predicts a value somewhat less than 125 ppb. Note that selection of an artificially low prediction threshold is equivalent to dropping the horizontal dotted lines in Figures 7 and 8 to the new threshold value. This changes the number of events in the various quadrants defined by the dotted lines, which in turn affects the various skill scores. Clearly, one could drop the prediction threshold sufficiently so that all the observed exceedances were above the dotted line (POD = 1.00), but that would simultaneously increase the FAR. In this case 100 ppb was arbitrarily selected as an artificial low threshold for prediction of an exceedance.

Forecasting skill scores in Table 16 for predicted and observed thresholds of 125 ppb for 1-hour maximum ozone concentrations quantify the poor results seen in Figure 7. The best POD was 0.13 for Approaches 2a, 2b, and 4, with correspondingly high MISS rates. Two of the FAR values were not computable. Approaches 2a and 2b had FARs of 0.50, but Approach 4 had no false alarms. The three detailed skill scores ranged from 0.00-0.22, all quite poor. Overall, Approach 4 was probably the best of this group; they had a POD and MISS equal to those of Approaches 2a and 2b, but they had slightly higher values for the detailed skill scores CSI, TSS, and S.

Results using a prediction threshold of 100 ppb and a verification threshold of 125 ppb for 1-hour ozone are also shown in Table 16. Note the improved POD and corresponding drop in the MISS. Of course, the FAR increased as well. The detailed skill scores also improved across the board. Overall, approaches 3 and 4 appear to have given the best results with the low prediction threshold, with PODs of 0.75 and 0.50, respectively, and detailed skill scores among the highest.

Table 16. Forecasting Skill for 1-hour Maximum Ozone.

	Threshold		*Measure of forecast skill*						
Approach	*Pred*	*Obs*	*POD*	*FAR*	*MISS*	*CNULL*	*CSI*	*TSS*	*S*
1	125	125	0.00	N/A*	1.00	1.00	0.00	0.00	0.00
2a**	125	125	0.13	0.50	0.88	1.00	0.11	0.12	0.19
2b	125	125	0.13	0.50	0.88	1.00	0.11	0.12	0.19
3	125	125	0.00	N/A	1.00	1.00	0.00	0.00	0.00
4	125	125	0.13	0.00	0.88	1.00	0.13	0.13	0.22
1	100	125	0.38	0.63	0.63	0.99	0.23	0.36	0.36
2a	100	125	0.38	0.80	0.63	0.97	0.15	0.34	0.24
2b	100	125	0.38	0.81	0.63	0.97	0.14	0.34	0.23
3	100	125	0.75	0.79	0.25	0.94	0.19	0.69	0.30
4	100	125	0.50	0.71	0.50	0.97	0.22	0.47	0.35
OAD, 1995-1997	125	125	0.78	0.82	0.22	0.93	0.18	0.71	0.27

Notes: * Value cannot be calculated. Both numerator and denominator are zero. ** Since actual ozone levels must be known to employ approach 2a, its use for prospective forecasting depends on the relationship between projected and observed ozone being consistent over time. See text for definitions of forecast measures.

There is a penalty, of course, with FARs rising to between 70 and 80%, but that is actually slightly better than the record of calling OADs in the 1995-1997 period (Table 4), which had a FAR of 82% (see Table 16).

Calculated forecasting skill measures for 8-hour ozone concentrations are given in Table 17. The upper half of the table shows results for both thresholds at the 8-hour standard of 85 ppb; the lower half of the table is for an arbitrary prediction threshold of 65 ppb and an observation threshold of 85 ppb. With both thresholds at 85 ppb, the POD ranged from 0.32-0.53 over the various approaches; this is a marked improvement over the corresponding results for 1-hour ozone. The FARs were relatively modest, at 0.14-0.47 All of the more complex skill scores were much improved over the 1-hour results. Overall, Approach 3 had the best results, with a tie for the highest POD (0.53); it also had less than half the FAR of Approaches 2a and 2b, which also had high PODs. Approach 3 also had the highest of the more complex skill scores.

Table 17. Forecasting Skill for 8-hour Maximum Ozone.

	Threshold		*Measure of forecast skill*						
Approach	*Pred*	*Obs*	*POD*	*FAR*	*MISS*	*CNULL*	*CSI*	*TSS*	*S*
1	85	85	0.42	0.11	0.58	1.00	0.40	0.42	0.56
2a*	85	85	0.47	0.40	0.53	0.98	0.36	0.46	0.51
2b	85	85	0.53	0.47	0.47	0.98	0.36	0.50	0.50
3	85	85	0.53	0.17	0.47	0.99	0.48	0.52	0.63
4	85	85	0.32	0.40	0.68	0.99	0.26	0.31	0.39
1	65	85	0.95	0.67	0.05	0.90	0.33	0.85	0.45
2a	65	85	0.95	0.67	0.05	0.90	0.33	0.85	0.45
2b	65	85	0.95	0.76	0.05	0.85	0.23	0.79	0.33
3	65	85	1.00	0.89	0.00	0.57	0.11	0.57	0.11
4	65	85	0.95	0.73	0.05	0.87	0.26	0.82	0.37

Notes: * Since actual ozone levels must be known to employ approach 2a, its use for prospective forecasting depends on the relationship between projected and observed ozone being consistent over time. See text for definitions of forecast measures.

With the prediction threshold at 0.65 ppb (Table 17), the PODs were no worse than 0.90, with all FARs except that for Approach 3 less than those of the 1995-1997 OAD program. The fraction of CNULL decreased somewhat, compared to results for both thresholds at 0.85 ppb, with that of Approach 3 only 0.57. Some of the more complex skill scores were better, and some were worse with the prediction threshold at 65 ppb rather than 85 ppb, depending on the approach and the specific skill score. The TSS values all were higher with the 65 ppb threshold; The CSI and S values all were lower than the corresponding values for the 85 ppb prediction threshold. Overall, Approaches 1 and 2a may have the best results of this group. Their POD was almost perfect, their FAR was lowest (0.67), and their complex skill scores were highest.

The choice of 65 ppb as the prediction threshold was arbitrary, and the rather variable results suggest that it may well have been too low. What is clear, however, is that one may choose a prediction threshold somewhat below the NAAQS to achieve either a predetermined POD level or a predetermined FAR level. By fine-tuning the prediction threshold between those that achieve predetermined POD and FAR levels, one may find a value that yields an acceptable

Table 18. Forecasting Skill for 8-hour Maximum Ozone Using Observed and Forecasted Predictor Variables, Based on Approach 4 (Non-linear 4th-order Polynomial in T_{max}).

	Threshold		*Measure of forecast skill*						
T_{max}	*Pred*	*Obs*	*POD*	*FAR*	*MISS*	*CNULL*	*CSI*	*TSS*	*S*
Observed	85	85	0.36	0.00	0.64	1.00	0.36	0.36	0.51
1-day fcst	85	85	0.29	0.00	0.71	1.00	0.29	0.29	0.43
2-day fcst	85	85	0.29	0.00	0.71	1.00	0.29	0.29	0.43
3-day fcst	85	85	0.14	0.00	0.86	1.00	0.14	0.14	0.24
Observed	65	85	1.00	0.68	0.00	0.89	0.32	0.89	0.44
1-day fcst	65	85	0.79	0.68	0.21	0.91	0.30	0.70	0.42
2-day fcst	65	85	0.71	0.76	0.29	0.88	0.22	0.60	0.31
3-day fcst	65	85	0.71	0.71	0.29	0.90	0.26	0.62	0.36

Notes: See text for definitions of forecast measures.

combination of these two skill measures. The more complex skill scores also can provide information that may help to suggest an acceptable prediction threshold.

The question of changes in forecast skill as observed predictor variables are replaced by forecasted variables and, as forecast target dates move from 1-3 days, is addressed in Table 18.

Table 18 shows that the exceedances POD decreased and MISS increased accordingly as observed variables were replaced with forecasted variables and as the forecast target increased from 1-3 days; the same was true both for 85 ppb and 65 ppb prediction thresholds. The FAR and CNULL measures were constant for the 85 ppb prediction threshold and showed no strong trend for the 65 ppb threshold. For the 85 ppb threshold, all of the more complex skill scores decreased as the observed temperature was replaced by the forecast temperature and the forecast increased in length. The same was true at the 65 ppb threshold out to a 2-day forecast, but the skills increased slightly again at the 3-day forecast.

5. Summary and Conclusions

The literature contains a substantial number of papers on ozone forecasting methods. Most papers used some form of regression analysis, but uses of neural network methods and classification and regression trees also were reported. There was considerable interest in forecasting ozone concentrations in the late 1970s and early 1980s, and again recently.

Ozone action days in the Chicago area are currently called by meteorologists from Illinois, Indiana, Wisconsin, and Michigan. Decisions to call an OAD for the following day are based on their expert judgment of current and expected weather and air quality conditions, after discussions during a morning conference call. Between 1995 and 1997, 38 OADs were called in the Chicago area. During this time, nine exceedances of the 125 ppm standard were observed, seven of which occurred on OADs. This rate of exceedances is similar to that observed in the Chicago area since about 1990. Exceedances have tended to occur more frequently on Thursday, Friday, and especially on Saturday, than on other days of the week. This pattern has been seen throughout the period for which we have data, beginning in 1981. The frequency of exceedances was higher on Saturdays during 1995-1997, but the total number of occurrences was relatively low, and the differences from the historical period were not significant.

Considerable spatial variation of ozone concentrations on days with one or more high values may be evidence for a strong influence of local precursor sources on ozone concentrations in the Chicago area.

Bivariate scatter plots of observed ozone concentrations versus potential predictor variables show strong direct relationships between ozone and high temperature, dew point, and solar radiation. There was also a strong direct relationship with the previous day's ozone concentration. An inverse relationship with wind direction was observed. A plot of back trajectories on high ozone days showed transport winds predominantly from the southwest, but there was an appreciable fraction of winds from the east as well.

Linear regression equations were developed using 70% of the observations of daily 1-hour or 8-hour maximum ozone concentrations and daily meteorological variables. The equations were then applied to the remaining 30% of the data, and the "predictions" were compared to observed ozone to test the performance of the regression equations. For both 1-hour and 8-hour ozone, regressions were developed for a base case (Approach 0) using the full set of predictor variables. The remaining regressions used only forecasted variables. Approach 1 was a repeat of Approach 0, but excluding solar radiation, the only nonforecasted variable. Approaches 2a and 2b applied different adjustments to the ozone concentrations calculated by Approach 1. Approach 3 developed a regression equation using only data with high observed ozone concentrations. Approach 4 developed a fourth-order (nonlinear) polynomial equation using the daily high temperature as the only predictor variable.

The most significant predictor variables in Approaches 1 and 3 were maximum temperature, the previous day's maximum 8-hour ozone concentration, and RH. Other variables, including visibility, wind speed, weekday/weekend occurrence, and sky cover, were less significant.

Overall results of comparing predicted and observed ozone concentrations may be summarized in terms of R^2, equivalent to the fraction of variance explained by the regression. Agreement between predicted and observed concentrations is also expressed by the RMS difference (or error) between the two. For 1-hour ozone, the Hubbard and Cobourn (1998) enhancement (Approach 2a) achieved the highest R^2 (0.711) and the lowest RMSE (12.44 ppb) when applied to the full 30% data set. When forecast high temperatures were substituted for observed high temperatures in Approaches 1, 3, and 4, R^2 decreased and RMSE increased. The R^2 decreased and the RMSE increased further as the temperature forecast lengthened from 1-3 days.

For 8-hour ozone, the Hubbard and Cobourn (1998) enhancement (Approach 2a) again produced the highest R^2 (0.687) and the lowest RMSE (10.55 ppb) when applied to the full 30% data set. As with 1-hour ozone, substitution of the forecasted high temperature for the observed high temperature reduced R^2 and increased the RMSE. As the temperature forecast lengthened from 1-3 days, R^2 based on Approach 4 decreased and RMSE increased. However, for Approaches 1 and 3, changes in both R^2 and RMSE were minimal as the temperature forecast lengthened from 1-3 days.

Skill in forecasting exceedances was assessed in terms of seven separate quantitative measures. Forecasting skill was computed for two different forecasting strategies, one in which an exceedance was forecast only when the regression equation predicted a value above the relevant standard, and one in which the threshold for prediction of an exceedance was somewhat lower than the standard. For 1-hour ozone, the standard is 125 ppb, and the lower threshold selected for the second strategy was 100 ppb. For 8-hour ozone, the standard is 85 ppb, and the lower threshold was selected as 65 ppb.

For both 1-hour and 8-hour ozone, the probability of detection was markedly better using the lower prediction threshold, and three more complex measures of forecasting skill showed higher scores with the lower threshold as well. Of course, this strategy increased the FARs. These results suggest that one may choose a prediction threshold somewhat below the standard to achieve either a predetermined POD level or a predetermined FAR level. By fine-tuning the prediction threshold between those that achieve predetermined POD and FAR levels, one may find a value that yields an acceptable combination of these two skill measures. The more complex skill scores can also provide information that may help to suggest an acceptable prediction threshold. Based on these results, it appears quite likely that one or more of the approaches used in this work can be fine-tuned to give forecast techniques for 1-hour and 8-hour ozone that will predict exceedances with equal or greater skill than the current method. At least, quantitative predictions should be used as one form of input to decisions to call or not call an OAD in Chicago.

6. Suggestions for Future Research

Other methods for ozone forecasting that have been reported in the literature should be explored. A nonlinear approach similar to that of Bloomfield et al. (1996) holds a great deal of promise for improved forecasting success, especially in view of the relative success of Approach 4, a nonlinear approach using only maximum temperature as a forecast variable. It is not reasonable when using forecast variables to expect an R^2 of 0.80 or a RMS difference of 8 ppb (Table 2), which Bloomfield et al. achieved at Chicago using observed meteorological variables; However, nonlinear methods appear to be quite promising and deserve additional investigation.

The neural network approaches reported in the literature have been about as successful as linear regression in many cases. This approach also should be tried for forecasting both 1-hour and 8-hour ozone concentrations at Chicago.

Based on tests using forecasted high temperatures, it appears that the accuracy of ozone forecasts will degrade if forecast variables are used instead of observed variables. The forecasts mostly appear to degrade further as the forecast target date lengthens beyond one day. The accuracy of the various approaches using a full set of forecast variables should be tested.

Further exploration of spatial variability and the distribution of exceedance occurrences by day of the week may be useful in elucidating the relative importance of local and distance precursor emissions on ozone concentrations in the Chicago area.

Probably the ultimate approach to forecasting ozone in the Chicago area is to develop or adapt an existing photochemical transport model for the midwest region, or to create a hybrid approach involving both photochemical modeling and statistical methods.

7. Acknowledgments

Support for this work was provided by the State of Illinois, through the Department of Natural Resources and the Department of Commerce and Community Affairs. The views expressed in this report are those of the author and do not necessarily reflect the views of the Illinois State Water Survey or the sponsors. Gary Stensland and Allen Williams provided valuable suggestions that helped to guide this work. John Marshall and Luther Smith provided valuable statistical advice. Byron Gleason obtained the ozone and meteorological data and developed spreadsheets of the data. Roger Claybrooke obtained the data on forecasted variables and prepared the spreadsheet used for their analyses. Carl Lonnquist assisted with data acquisition from archive tapes. Gary Stensland and Clyde Sweet reviewed the manuscript. Eva Kingston and Agnes Dillon edited the manuscript.

8. References

Altshuler, S.I., T.D. Arcado, and D.R. Lawson. 1995. Weekday vs. weekend ambient ozone concentrations: discussion and hypotheses with focus on northern California. *JAPCA* **45**(12): 967-972.

Aron, R.H., and I.-M. Aron. 1978. Statistical forecasting models: II. Oxidant concentrations in the Los Angeles basin. *JAPCA* **28**(7):684-688.

Bloomfield, P., J.A. Rogle, L.J. Steinberg, and Q. Yang. 1996. Accounting for meteorological effects in measuring urban ozone levels and trends. *Atmos. Environ.* **30**(17):3067-3077.

Burrows, W.R., M. Benjamin, S. Beauchamp, E.R. Lord, D. McCollor, and B.Thomson. 1995. CART decision tree statistical analysis and prediction of summer maximum surface ozone for the Vancouver, Montreal, and Atlantic Regions of Canada. *J. Appl. Meteorol.* **34**:1848-1862.

Clark, T.L., and T.R. Karl. 1982. Application of prognostic meteorological variables to forecasts of daily maximum one-hour ozone concentrations in the northeastern United States. *J. Appl. Meteorol.* **21**:1662-1671.

Cleveland, W.S. 1985. *The Elements of Graphing Data.* Wadsworth Advanced Books and Software, Monterey, CA.

Cleveland, W.S., T.E. Graedel, B. Kleiner, and J.L. Warner. 1974. Sunday and workday variations in photochemical air pollutants in New Jersey and New York. *Science* **186**:1037-1038.

Comrie, A.C. 1995. Comparing neural networks and regression models for ozone forecasting. *JAWMA* **47**:653-663.

Eder, B.K., J.M. Davis, and P. Bloomfield. 1994. An automated classification scheme designed to better elucidate the dependence of ozone on meteorology. *J. Appl. Meteorol.* **33**:1182-1199.

Feister, U., and K. Balzer. 1991. Surface ozone and meteorological predictors on a subregional scale. *Atmos. Environ.* **25A**(9):1781-1790.

Graedel, T.E., L.A. Farrow, and T.A.Weber. 1977. Photochemistry of the "Sunday effect," *Environ. Sci. Technol.* **11**(7):690-694.

Gunst, R.F., and R.L. Mason. 1980. *Regression Analysis and Its Application.* Marcel Dekker, New York, p. 170.

Hanna, S.R., and J.C. Chang. 1995. Relations between meteorology and ozone in the Lake Michigan region. *J. Appl. Meteorol.* **34**:670-678.

Hubbard, M.C., and W. G. Cobourn. 1998. Development of a regression model to forecast ground-level ozone concentration in Louisville, KY. *Atmos. Environ.* **32**(14/15):2637-2647.

Karl, T.R. 1979. Potential application of model output statistics (MOS) to forecasts of surface ozone concentrations. *J. Appl. Meteorol.* **18**:254-265.

LADCO. 1996. Lake Michigan Ozone Study: 1995 data analysis report, version 1.1. Report prepared by Lake Michigan Air Directors Consortium, Des Plaines, IL, April.

Lebron, F. 1975. Comparison of weekend-weekday ozone and hydrocarbon concentrations in the Baltimore-Washington metropolitan area. *Atmos. Environ.* **9**:861-863.

Prior, E.J., J.R. Schiess, and D.S. McDougal. 1981. Approach to forecasting daily maximum ozone levels in St. Louis. *Environ. Sci. Technol.* **15**:430-436.

Revlett, G.H. 1978. Ozone forecasting using empirical modeling. *JAPCA* **28**(4):338-343.

Robeson, S.M., and D.G. Steyn. 1990. Evaluation and comparison of statistical forecast models for daily maximum ozone concentrations. *Atmos. Environ.* **24B**(2):303-312.

Ruiz-Suarez, J.C. 1995. Short-term ozone forecasting–neural network. *Adv. Eng. Software* **23** (3):143.

Ryan, W.F. 1995. Forecasting severe ozone episodes for the Baltimore metropolitan area. *Atmos. Environ.* **29**(17):2387-2398.

Tiao, G.C., M.S. Phadke, and G.E.P. Box. 1976. Some empirical models for the Los Angeles photochemical smog data. *JAPCA* **26**(5):485-490.

Walker, J.S. 1993. Tropospheric ozone concentration trends by day of the week. Paper FM2-III6, *Proceedings* of the International Conference and Course: Regional Photochemical Measurement and Modeling Studies, Air & Waste Management Association, Pittsburgh.

Wilkinson, L. 1997. *Systat 7.0.* SPSS, Inc., Chicago, IL.

Wolff, G.T., and P.J. Lioy. 1978. An empirical model for forecasting maximum daily ozone levels in the Northeastern U.S. *Atmos. Environ.* **11**:967-983.

Yi, J. 1996. *Environ. Pollut.* **92**(3):349

Appendix 1

LAKE MICHIGAN
OZONE WEATHER FORECASTING PROTOCOL

Purpose Statement

Forecasts of ozone-conducive weather conditions are needed during the summertime in the Lake Michigan area to support the Ozone Action Day (OAD) programs and special Photochemical Assessment Monitoring Station (PAMS) sampling. This document reviews the weather forecasting procedures and the criteria for calling an OAD and a special PAMS sampling day. Because 1998 will be a transition year for implementing the new 8-hour National Ambient Air Quality Standard (NAAQS) for ozone, both 1-hour and 8-hour ozone concentrations need to considered in calling an OAD and a special PAMS sampling day.

Ozone Action Day (OAD) Program

The primary objectives of the State OAD programs are to maintain and improve air quality (for public health reasons), and to promote public education. The OAD programs will encourage voluntary actions by business, industry, and the public on days when ozone-conducive weather conditions are forecast. By avoiding or curtailing certain activities, ozone precursor emissions resulting from these activities can be reduced. It should be noted, however, that the OAD programs are not a mandatory attainment strategy. Notification of ozone-conducive conditions may also help to reduce exposure to high ozone levels.

While each State is responsible for its own OAD program, the four States have worked together to coordinate their programs. In particular, the program name and activities are similar for each State. Such coordination is necessary given the regional nature of the ozone problem in the Lake Michigan area.

Photochemical Assessment Monitoring Station (PAMS) Program

Pursuant to the Clean Air Act Amendments of 1990, USEPA promulgated regulations requiring enhanced ambient ozone and ozone precursor monitoring. A regional PAMS program was developed, and approved by USEPA, for the Lake Michigan area. The following data are collected from June 1 to August 31[1]:

- ozone, Nox, and meteorological data at all sites on a continuous basis:
- upper air meteorological data (wind speed and wind direction) at the Waukegan SODAR site and a new Illinois upper air site on a continuous basis:
- VOC data according to the following schedule
  - on a continuous basis at the Chicago-Jardine, Northbrook, Camp Logan, Braidwood, Holland, and Gary-IITRI sites;
  - either four 3-hour samples every 3 days (Milwaukee-UWMN site) or three 3-hour samples every 6 days (Harrington Beach);
  - one 24-hour sample every 6 days year-round at Chicago-Jardine, Gary-IITRI, and Milwaukee-UWMN sites) and from June 1 to August 31 (all other PAMS sites).
- Carbonyl data according to the following schedule:
  - either four 3-hour samples every 3 days (Chicago-Jardine, Gary-IITRI, and Milwaukee-UWMN sites) or once every 6 days (all other PAMS)
- VOC and carbonyl data during episodic conditions at the two Wisconsin sites for the same four (or three) 3-hour intervals on forecasted days of high ozone (Note: the window for these episodic measurements is May 15 - September 15.) It is anticipated that there be as many as 15 episodic monitoring days per ozone session.

Given the establishment of auto-GCs at most PAMS sites in t he region (and the plans for establishing yet another auto-GC at one of the Wisconsin PAMS sites), the only “special PAMS sampling” still supported by the forecasting are the episodic ground-based VOC/carbonyl measurements at the Harrington Beach site (and, until the auto-GC is operational, the Milwaukee site), and aircraft measurements by the State of Wisconsin and R.B. Jacko. These measurements are referred to in t his document as “special PAMS sampling”.

---

[1] Additional (noon-PAMS) regional measurements include aircraft sampling (along the upwind boundary of the Lake Michigan area and over Lake Michigan), tall building monitoring (Sears Tower), and over-water ozone monitoring (Badger Ferry).

Forecasting Criteria

The meteorologists will focus on ozone-conducive weather conditions, not specific ambient concentration levels. (Note, however, that the criteria for calling an OAD or special PAMS sampling also requires consideration of expected ozone levels.) Table 1 identifies the conditions associated with historical ozone exceedances in the Lake Michigan area. In general, the weather conditions of most concern are high temperatures, high humidities, light to moderate winds, low cloud cover, and little (or no) precipitation.

The following factors should be considered in assessing ozone-conducive conditions:

- synoptic weather conditions, including upper air pattern (especially, with the Lake Michigan on th e back-side of a High, and a ridge aloft);
- local weather conditions, including surface winds and temperatures (especially, light to moderate winds with a southerly component, and maximu afternoon temperatures greater than 90°F);
- previous day's ozone levels; and
- upwind (boundary) ozone levels.

Forecasting Procedures

The ozone weather forecasts will be made by the meteorologists in the following State agencies:

Bureau of Air Management, Wisconsin Department of Natural Resources (WDNR), Madison, Wisconsin;

Bureau of Air, Illinoiso Environmental Protection Agency (IEPA), Springfield, Illinois.

Air Quality Division, Michigan Department of Environmental Quality (MDEQ), Lansing, Michigan; and

Office of Air Management, Indiana Department of Natural Resources (IDEM), Indianapolis, Indiana

Each of these offices has access to real-time (or near real-time) meteorological data, weather forecast products, and satellite imagery through meteorological information servers available on the Internet.

Each State meteorologist will prepare a weather forecast for every day from approximately mid-May through mid-September. The meteorologists will prepare a consensus weather forecast for the Lake Michigan area on those days when a forecast conference call is scheduled. Regular conference calls will be held at 10:30 am CDT[2] every Monday and other days, as necessary. The first call of the season will be held in late May. A preliminary call may be held in early May to review and test the forecasting procedures. To support the forecasting effort, LADCO will prepare a map of peak 1-hour ozone concentrations from the day before and fax it to the forecasters prior to the call.

Table 2 contains a draft timetable of the daily events regarding the preparation and dissemination of weather forecast information, and the PAMS and OAD decision-making.

Decision-Making Criteria

An OAD should be called for the next day if ozone-conducive conditions are forecast and if the next day is expected to be part of a broad-based, multi-day event with peak 1-hour ozone concentrations equal to or greater than 100 ppb[3]. The greater the magnitude, spatial extent, and temporal duration of high ozone, the more certain the decision should be to call an OAD.

A special PAMS sampling day should be called for the next day if ozone-conducive conditions are forecast and the next day is expected to be part of a broad-based, multi-day event with peak 1-hour ozone concentrations equal to or greater than 125 ppb. The greater the magnitude, spatial extent, and temporal duration of high ozone, the more certain the decision should be to call a special PAMS sampling day. (Earlier notification will be given whenever possible to facilitate the scheduling of personnel and to ensure the availability of sampling equipment.) A special PAMS sampling day should also be called on "ramp-up" days of expected multi-day episodes. If high ozone concentrations fail to materialize as expected, then the sampling should be terminated and the equipment readied for the next event.

It is possible to have an OAD without having a special PAMS sampling day. An OAD may be associated with broad-based, multi-day events exceeding the 8-hour NAAQS, while special PAMS sampling may be associated with broad-based, multi-day events exceeding the 1-hour NAAQS.

---

[2] The forecaasters may change the time of the call to a more convenient time, if desired. For example, weekend calls have generally occurred earlied in the morning.

[3] A 1-hour concentration of 100 ppb is assumed to provide a reasonable proxy for an 8-hour concentration of 85 ppb (i.e., the 8-hour NAAQS).

Given the regional nature of the ozone problem in the Lake Michigan area, it is desirable to call an OAD in all four States if meteorological conditions are conducive to “high” ozone in the region. Certain conditions, however, may occur where it may be appropriate to call an OAD in fewer than all four States (e.g., high ozone is expected in only one or two States).

The final decision to call an OAD and a special PAMS sampling day will be made at least one day at a time. An OAD or special PAMS sampling day may be called for several consecutive days during times of stagnation or other persistent conditions. On Saturdays, Sundays, and holidays media contacts will continue to be notified one day at a time. Business contacts, on the other hand, will be notified on a Friday that the next day (Saturday) is an ozone action day, if appropriate, and will be advised as to the likelihood that Sunday and/or Monday will be an OAD. (If there is the likelihood that Monday will be an OAD, then businesses will also be advised that confirmation will be provided first thing Monday morning.

Decision-Making Authority

The decision to call a special PAMS sampling day shall be made by LADCO, with input by the States. Among the factors to consider in making this decision is the availability of equipment, platforms, and personnel.

The decision to call an OAD shall be made by a group of State representatives, including the State meteorologists and other State personnel participating on the conference call. The OAD decision should be made separate from the PAMS decision.

Table 1

Overview

Typical Meteorological Conditions Associated
With High Ambient Ozone Levels
Within 50 Miles of Lake Michigan
In Wisconsin, Illinois, Indiana and Michigan

The precise nature of meteorological processes that contribute to the transformation and transport of substantial amounts of ozone and ozone precursors in areas near Lake Michigan are extremely complicated and not completely understood at this time. Lake Michi Ozone Study (LMOS) results have improved the general understanding of relevant ozone formation/transport mechanisms in the region.

Additionally, the Lake Michigan state air pollution control agencies have gained a qualitative understanding of prevailing weather conditions that are identified with high ozone in the state's proximity to the Lake. An overview of these general circumstances i listed below. Ozone levels can be expected to increase with persistence in these weather conditions (i.e., ozone episodes).

Parameter	SE/E Wisconsin, NE Illinois, NW Indiana	W Michigan
Surface synoptic pressure system:	Strong high centered over Indiana or Ohio	Strong high centered over Indiana or Ohio
Pressure tendency (all states):	Steady or slightly decreasing	
Surface Temp °C (°F)		
- Land maximum:	> 28 (82)	> 27 (80)
- Land - Lake surface temp difference	> 6 (10)	> 6 (10)
Nearcoast Winds		
- Speed m/s (mph):	≤ 4 (9)	< 12 (27)
- Dir (early/mid morning):	190-250°	180-250°
- Dir. (mid-afternoon):	110-160°	180-250°
- Time span in dir. shift:	< 2 hr	Speed is more important
Solar Radiation (langleys/day):	> 550	> 550
low,mid (< 15000 ft) cloud cover (tenths):	< 3	< 6
Precip. (previous 24 hrs, inches):	< 0.01	< 0.01
Probability of precip - next 24 hrs, (%):	< 10	< 10
Visibility km (mi):	< 13 (8)	< 13 (8)
Obstruction to vision:	Haze (mod.-to-heavy)	Haze (mod.-to-heavy)
Persistence factor (all states):	A minimum of two (2) days of the same conditions	

May 11, 1998

Table 2. Draft Timetable for Lake Michigan Ozone Weather Forecasting and PAMS/OAD Decision-Making

Time (CDT)	Task Description
until 1030 [a]	Each State meteorologist will prepare weather forecast for both next day (detailed) and following several days (general)
	Each State will provide LADCO with a summary of peak 1-hour ozone concentrations from the day before by 9:00 am CDT. LADCO will then prepare a map showing the magnitude and location of these concentrations and fax it to the meteorologists prior to the call.
1030 [b]	Conference call will be held, as necessary. The call will be structured to consist of the following three parts:
	Part 1 (Forecasting): A "lead" meteorologist from either IEPA or WDNR will provide a summary of their forecast. Following input from the other meteorologists, a consensus regional forecast will be prepared.
	Part 2 (PAMS Decision): Based on the forecast, LADCO, with input from the States, will decide whether to call for special PAMS sampling the following day(s) in accordance with the protocol criteria.
	Part 3 (OAD Decision): Based on the forecast, State representatives (including the forecasters) will decide whether to call an OAD the following day(s) in accordance with the protocol criteria.
	If no OAD or special PAMS sampling day is called, then the meteorologists must decide when to have the next conference call.
	If an OAD is called for the following day, then each States' meteorologist will notify their OAD contact. A conference call will be held the following day.
	If a special PAMS sampling day is called for the following day, then each States' meteorologist will notify their PAMS monitoring contact and any other appropriate field people (e.g., aircraft pilots). A conference call will be held the following day.

---

[a] Task performed during every scheduled office day (and possibly weekends) of the PAMS season.

[b] Task performed during every day for which a conference call is scheduled.

Appendix 2

Available Details of NOAA Ozone Forecasts on the World Wide Web

The following was copied from http://www.arl.noaa.gov/ready/ozone.html#assump on 17 August 1998.

Experimental Forecast Assumptions

Ozone concentrations (ppb) for the eastern U.S. are summarized above for today and tomorrow. Concentrations are given as hourly average maps between the hours of 1800-2400 UTC (1400-2000 EDT) each day and as an hourly time series at a grid point nearest to a user selected location. Air concentrations are computed based upon a 50 km resolution grid and a 500 m vertical layer. An ozone background concentration of 10 ppb is assumed for all hours. Emissions of VOC and NOx are based upon the 1985 NAPAP inventory reduced to a resolution of 100 km. Transport and dispersion is computed using meteorological fields from NOAA NCEP's Eta model, once each day, based upon the 0000 and 1200 UTC forecasts (to 48 h). The calculation is started 24 hours prior to the forecast initialization using Eta archive fields and therefore each day's forecast represents a 48 h pollutant accumulation. Depending on system load, forecasts should be completed between 1230 to 1330, and 0130 to 0230 EDT.

Pollutant particles are released each hour from each emission grid cell with the appropriate mass of VOC, NOx, and NO, corresponding to the appropriate total area, point, and mobile emissions in that cell. In addition, biogenic hydrocarbon (isoprene) emissions are added to each cell that is dominated by a forest land-use category. Pollutant particles are then tracked and dispersed according to the meteorological fields from the Eta model. At each time step the ozone formation equations, summarized below, are integrated for the total smog produced (SP), according to the local temperature and incoming solar radiation. Ozone is then calculated from the photostationary state equation and summed to the concentration grid.

Ozone Model Details

The simplified version of the Generic Reaction Set (GRS) by Graham Johnson (CSIRO-Sydney), the Integrated Empirical Rate (IER) model, has been incorporated into the Hysplit code for the purpose of calculating Ozone air concentrations.

The GRS equations can be summarized as follows:

1) ROC + hv -> RP + ROC
2) RP + NO -> NO_2
3) NO_2 + hv -> NO + O_3

4) $NO + O_3 \rightarrow NO_2$
5) $RP + RP \rightarrow RP$
6) $RP + NO_2 \rightarrow SGN$
7) $RO + NO_2 \rightarrow SNGN$

The IER model is an algebraic solution of the GRS, that depends upon the definition of the smog produced (SP), the definition of NO_x, and the photo-stationary state equation, all defined below:

$[SP]t = [O_3]t - [O3]o + [NO]o - [NO]t$
$[O_3]t = k1\ [NO_2]t / k4\ [NO]t$
$[NO_x]t = [NO]t + [NO_2]t$

The interface between Hysplit and the IER is accomplished by releasing a series of puffs or particles at each emission location. Each particle contains mass associated with 6 different pollutant species, where only SP is the result of a temporal integration and the others are defined by their emission:

$[SP]t\ \ [ROC]\ [isoprene]\ \ [NO_x]\ \ [O_3]\ \ [NO]$

Initial values of NO are converted to NO_2 according to the concentration of ROC and the temporal integration of SP. The remaining species:

$[NO_2]t\ [NO]t\ \ [O_3]t$

have a analytic solution according to the IER equations and are computed each time step and summed on the model's concentration grid for output.

Future Activities

The model results are currently being compared with measured data, and many of the assumptions are being tested through a variety of sensitivity trials. A paper describing the results is in preparation.

References

Johnson, G.M. 1984: A simple model for predicting the ozone concentration of ambient air. In Proc. 8th Int. Clean Air Conf., Melbourne, Australia, May 1984.

Appendix 3

Definitions of Measures of Forecast Skill (after Ryan, 1995)

Verifications of the ozone forecasts were based on a standard continency table (Table 15). Forecast skill was expressed in seven different ways, as follows. The probability of detection (POD) is the fraction of ozone events that were correctly forecast, and it is given by:

POD = A/(A + B).

The miss rate (MISS) is the fraction of ozone events that occurred but were not forecast, and it is given by:

MISS = 1 - POD = B/(A + B).

The false alarm rate (FAR) is the fraction of ozone event forecasts that were wrong, and it is given by:

FAR = C/(C + A).

Skill in forecasting nonevents is expressed by the correct null forecast (CNULL), which is the fraction of forecasted nonevents that were correct, and it is given by:

CNULL = D/ (D + C).

Several more detailed skill scores are also useful. The critical success index (CSI) combines forecast occurrences and observed occurrences without regard to successful null forecasts, and it is given by:

CSI = A / (A + B + C).

The true skill score (TSS) includes the success of null forecasts in the form of a ratio of observed skill to perfect forecast skill. This measure is not dependent on the relative frequency of occurrence and nonoccurrence or the number of trials, and it is expressed as:

TSS = (AD - BC) / (A + B)(C + D) = POD + CNULL -1.

If all forecasts are correct, TSS = 1; if all incorrect, TSS = -1.

For rare events, the Heidke skill score (S) is appropriate. This is a measure of the skill of a set of forecasts compared to the skill of a random forecast, and it is given by:

$S = 2(AD - BC) / (B^2 + C^2 + 2AD = (B + C)(A + D))$.

Illinois State
WATER
SURVEY

ILLINOIS
DEPARTMENT OF
NATURAL
RESOURCES

ILLINOIS