METEOROLOGICAL FACTORS OF OZONE PREDICTABILITY AT HOUSTON, TEXAS

Roland R. Draxler
NOAA Air Resources Laboratory, 1315 East West Hwy, Silver Spring, MD 20910

Journal of the Air and Waste Management Association, 50:259-271, February 2000.

IMPLICATIONS

The purpose of the paper is to provide some understanding of the uncertainties meteorological data introduce into the calculation of maximum ozone concentrations. The paper is intended to provide guidance to managers on the resources that might be needed to develop an operational ozone forecast system in the light of meteorological uncertainty and the tradeoffs involved regarding model complexity, accuracy, and precision.

ABSTRACT

Several ozone modeling approaches were investigated to determine if uncertainties in the meteorological data would be sufficiently large to limit the application of physically realistic ozone forecast models. Three diagnostic schemes were evaluated for the period of May through September of 1997 for Houston, Texas. Correlations between measured daily maximum and model calculated ozone air concentrations were found to be 0.70 using a linear regression model, 0.65 using a non-advective box model, and 0.49 using a 3-Dimensional transport and dispersion model. Although the regression model had the highest correlation, it showed substantial under-estimates of the highest concentrations. The box model results were the most similar to the regression model and did not show as much underestimation. The more complex 3-D modeling approach yielded the worst results, likely resulting from ozone maxima that were driven by local factors rather than by the transport of pollutants from outside of the Houston domain. The highest ozone concentrations at Houston were associated with light winds and meandering trajectories. A comparison of the gridded meteorological data used by the 3-D model to the observations showed that the wind direction and speed values at Houston differed most on those days in which the ozone underestimations were the greatest. These periods also tended to correspond with poor precipitation and temperature estimates. It is concluded that better results are not just obtained through additional modeling complexity but there needs to be a comparable increase in the accuracy of the meteorological data.

INTRODUCTION

Incidents of elevated summer tropospheric ozone over the Middle Atlantic Region are typically attributed to precursor pollutants accumulating near associated quasi-stationary high pressure systems with strong inversions and no cloud cover[1]. These conditions occur regularly from May

through September, which has become known as the *ozone season*[2]. However, Davis et al.[3] found that in Houston, Texas high ozone concentrations are also associated with migratory anticyclones which can extend the ozone season through October. Air concentrations during these ozone events usually depend upon only a few meteorological variables, such as temperature and wind speed, and therefore many of the current forecasting approaches have found that regression methods[4] are satisfactory in situations where the ozone maxima depend primarily on local meteorological variables for which data are available. Although more physically realistic models can provide information on the timing and spatial distribution of ozone concentrations, excessive computational demands and inherent variability in meteorological processes have limited their broad application in ozone forecasting.

In anticipation of the development of ozone forecasting models based upon more complete descriptions of physical mechanisms, emissions data, and forecast meteorological parameters, an investigation into the prediction ability of such forecast systems has been conducted. There are several factors to be considered. First, the design of the ozone calculation scheme in conjunction with a dispersion model should be simple and fast enough to be run operationally and still produce realistic results. Second, the system should be as easily usable with historical data or meteorological forecasts for both diagnostic and prognostic evaluation. Finally, the prognostic system should perform better than simpler statistical methods. Houston, Texas was selected for this evaluation primarily because it experiences some of the highest ozone values in the United States[5]. Daily forecast ozone maxima, calculated using a non-advective box model and a 3-dimensional transport and dispersion model have been compared with measured maximum ozone values for the summer of 1997. To obtain the maximum computational speed a simple 3-equation steady-state chemical scheme, used to compute ozone concentrations from precursor pollutants, was incorporated into the two models. This same chemical model is used in all of the meteorological modeling variations tested, except for the linear regression approach which is used to determine a predictability baseline.

The intent of this analysis is not to compare various ozone chemical calculation schemes, because that has been addressed in past studies[6], but to evaluate the ozone prediction sensitivity to the meteorological data and the corresponding modeling approach. Considering that a meteorological forecast will not be as accurate as the meteorological analysis fields, it is possible to determine a baseline level of predictability using the analysis fields. The baseline is relevant in the sense that the analysis fields are the best representation of the actual meteorological conditions, but yet they may not be sufficiently accurate to obtain acceptable predictions of ozone air quality.

The following sections will review the specifics of the three modeling approaches and their data requirements, the general results for each model's calculation of daily maximum ozone concentration for the entire summer, and the results of a two week case study analysis of hourly ozone values and their relation to the meteorological conditions.

MODELS

Although only one ozone calculation scheme and one set of meteorological data will be used in the evaluation, an archive derived from the National Centers for Environmental Prediction (NCEP) Eta forecast model analysis fields (EDAS - Eta Data and Analysis System), the available meteorological data will range from using meteorological values at a single point in the regression and the non-advective box models, to using the meteorological data at all grid points in the fully 3-dimensional transport and dispersion model.

Ozone Model

The ozone prediction method, a semi-empirical approach called the Integrated Empirical Rate (IER) model developed by Johnson[7], is incorporated as the ozone calculation scheme in two of the modeling approaches evaluated in this study. The IER model, based on outdoor smog chamber studies for conditions typical of Australian cities, is usually applied in conjunction with monitoring data, to evaluate the age of photochemical smog[8] and to determine whether the air parcel is in the NO_x-limited or light-limited regimes. In the light-limited regime the smog produced is only a function of the accumulated incident light and the hydrocarbon concentration. In the NO_x-limited regime, the concentration of smog produced is assumed to be independent of the amount of incident light and the smog concentration can only increase by increasing the amount of NO available. An updated version of the IER model[9] was derived from the more complete Generic Reaction Set (GRS[10]) and includes the loss of NO_x to stable nitrate. Like IER, GRS is also a semi-empirical approach, but it is written in terms of a chemical mechanism using the reaction rate equations for the seven species. The GRS is not limited by the photo-stationary steady-state assumption of the IER model. The IER model, as used in this study, is described in more detail in the Appendix. In terms of implementation of either the IER or GRS models within the framework of a meteorological transport and dispersion model, the primary difference between the approaches is that the IER integration of the smog produced is linear and can easily be incorporated into a Lagrangian model, while the GRS solution is non-linear and requires adding the precursor contributions from all sources before integration.

The GRS model has been evaluated by Venkatram et al.[11] and Tonnesen and Jeffries[12]. The smog production component of the IER approach was evaluated by Chang et al.[13] with respect to NO_x or VOC control strategies. They suggested some modifications to the parameters, but felt that the IER observational approach, relating NO_x to smog using measured O_3, NO, and NO_y , could be competitive with more complete 3-D modeling approaches. The accuracy of more complex models tends to be limited by uncertainty in the input data and atmospheric processes. In the first complete modeling application of the IER scheme, Wratt et al.[14] applied the IER equations to an air parcel following a trajectory to determine the impact of a proposed power plant near Auckland, New Zealand.

A simple test of the IER chemistry code would be to compare the model results to one of the many smog chamber data sets for which the model was initially formulated. The results for

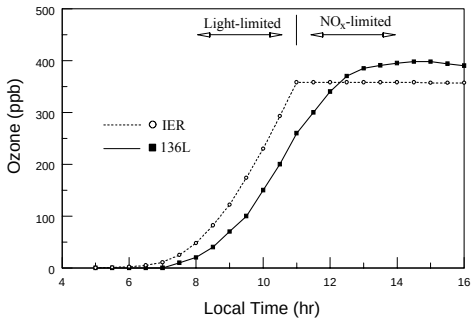

Figure 1. Smog chamber results[7] for experiment 136L and the results from the IER model using observed chamber temperatures.

experiment 136L, with initial conditions of 112 ppb-C VOC, 109 ppb NO, and 110 ppb NO_x, are shown in Fig. 1 The test was conducted at an outdoor chamber in Sydney, Australia, on November 11, 1980 between 0527 and 1615 local time. Naturally, the model results should be very similar to the data, but more importantly, the calculation tests the chemical model linkage with the position, date, solar elevation, and cloud-cover algorithms of the parent meteorological transport and dispersion model. In addition, the results clearly illustrate the initial ozone formation in the light-limited regime, the rate of which is sensitive to the VOC concentration, solar radiation, and the subsequent constant ozone concentrations in the NO_x limited regime. The results also show that differences of about 10% between the IER model and measurements can be expected even under the conditions for which the model was developed. This result is consistent with the other IER simulations and smog chamber results[7].

Although Tonnesen and Jeffries[12] concluded that GRS compares well with more complete descriptions of chemical reactions such as the Carbon Bond Four (CB4[15]) mechanism for VOC/NO_x ratios greater than 8, they questioned the applicability of GRS for ambient urban ozone predictions. Typical initial ambient values for NO_x of 10 ppb and VOC of 50 ppb-C yield maximum ozone values (from Tonnesen and Jeffries, Figs. 1 and 2) of 120 ppb from CB4 and 80 ppb from GRS, indicating a bias toward lower concentrations from the GRS/IER approach than from the more commonly applied CB4. This may not be a serious concern, in that a predictive ozone forecast system does not necessarily need to contain all the chemical kinetic mechanisms or detailed meteorological parameters. However the chemical model should respond properly to the range of meteorological variables and reproduce realistic ozone air quality estimates for those conditions. As long as the predictive system responds properly to changes in the primary forecast parameters, a bias correction can be tolerated and statistically corrected, considering the less demanding computational requirements of a simpler chemical prediction scheme.

Regression Model

Regression models are one of the most common approaches to ozone prediction and can vary considerably in complexity [3,4,16,17]. In this analysis, the regression model results are only used to demonstrate how much of the variance in ozone concentration can be explained by the meteorological variables and will be the standard against which the other two more physically realistic modeling approaches will be evaluated. Although it can be applied in a prognostic mode, in this application the regression model is only intended to be used as an explanatory or diagnostic tool. Subsequent modeling approaches will need to show a performance better than the regression baseline. A simple linear model was applied to calculate the maximum daily ozone value (ppb),

$$O_3 (max) = a\ U + b\ T + c\ Rh + d\ Z_i + e, \quad \textbf{(1)}$$

where U (m/s) is the wind speed, T (^oK) the temperature, Rh (%) the relative humidity, and Z_i (m) the mixed layer depth, computed with the same algorithm used in the 3-D dispersion model. The regression coefficients (a,b,c,d,e) were obtained using the daily maximum measured ozone and meteorological values from the 925 hPa level of the EDAS archive averaged each day between 0700 and 1900 LST for the period of May through September at the ozone measurement location. Because the regression coefficients are obtained from the same ozone data for which all the other models will be tested, the regression model results are not independent and are only used for comparative purposes.

Non-Advective Box Model

The next level of modeling complexity is the "box" approach, in which it is assumed that both horizontal and vertical mixing are instantaneous and emissions are uniform over the ground surface of the box. Although the box approach may appear simple, the corresponding chemistry may be quite complex. An early application of a box model by Schere and Demerjian[18] used 24 species in 38 reactions in an Eulerian framework, with a 20 km box centered over St. Louis, always aligned with the wind direction. Measured ozone provided the upwind boundary condition. An example of a Lagrangian box model is that of Eliassen et al.[19], where the box is carried along a pre-computed trajectory accumulating emissions. Ozone was formed using 40 species in 100 reactions. In box model approaches, the resulting prediction is a spatial average and may not be representative of a value at a specific measurement point. The resolution of sub-grid cell details is a limitation of all box methods whether they use single cells or multiple cells in complex 3-D Eulerian models.

In the application of the IER solution using the box model approach, it is assumed that one box, fixed and centered about the sampling location, accumulates emissions from all pollutant sources within the box. The vertical extent of the box is the mixed layer depth (Z_i). The air concentrations for all precursor species are computed from the sum of their masses in the volume of the box,

$$C(t) = \sum_{0}^{t} q(t)\ V(t)\ [\ \Delta x\ \Delta y\ Z_i (t)\ ]^{-1}\ \Delta t\ , \quad \textbf{(2)}$$

where q is the emission rate, V is the box ventilation factor, and Δt is the integration time step. The product of the box sides, $\Delta x\ \Delta y$, represents the fixed box horizontal area and it should be sufficiently large to incorporate most of the emission sources. Although emissions accumulate with time, they are vented out of the box by the ventilation-advection factor,

$$V(t) = 1 - k\ U(t)\ \Delta t\ \ \Delta x^{-1}\ , \quad \textbf{(3)}$$

computed each time step from the scalar wind speed (U) at the 925 hPa level of the EDAS archive. The last term of Eq. 3 represents the fraction of the pollutants within the box that are

depleted each time step. The value of "k", a free parameter chosen to give the best overall fit to the measurements, was empirically determined to be 0.6. The "k" factor incorporates all the remaining un-modeled meteorological effects such as wind shear and advection of pollutants into the box, pollutant removal by clouds and boundary layer venting, errors in the emission estimates and deposition, and even biases in the chemical model. Without the ventilation factor, pollutants would accumulate within the box for the duration of the simulation resulting in substantial over-estimation.

The optimum grid cell size for the box model was determined through a series of sensitivity tests with various box sizes. Results are shown in Table 1. When going from a 50 km to 100 km box size, an increase in surface area of a factor of 4, the emissions also increase by a factor of 4 (VOCs slightly more). However going from a 100 km box to a 200 km box results in an emission increase of less than a factor of 2 for NO_x and VOC and about a factor of 4 for isoprene. The ozone concentration increases substantially only between the 50 and 100 km simulations and then decreases again for the 200 km simulation. The box model calculation is expected to be relatively insensitive to the grid cell size as long as the additional emissions are in the same proportion as the change in surface area. Concentration increases should be compensated by an increasing box volume. Further increases in cell size lowers the concentration due to the fact that the increased dilution factor is not compensated by increased emissions; the larger cell boundaries are no longer within the dense urban emission area. The only substantial change in air concentration, between the 50 and 100 km box simulations, are due to a lower wind speed ventilation (Eq. 3), i.e., typical wind speeds of a few meters per second do not provide a sufficiently long residence time for precursor pollutants to reach the peak ozone concentrations in the smaller box size simulations.

Table 1. Sensitivity of average maximum ozone (ppb) and pollutant emissions (10^3 kg/day) to horizontal grid cell size for Houston.

Cell size	50 km	100 km	150 km	200 km
O_3	20	56	61	52
NO_x	170	680	1000	1100
VOC	300	1400	2100	2400
Isoprene	120	480	1100	1900

Three-Dimensional Transport-Dispersion Model

The integration of a chemical model with a Lagrangian transport and dispersion model is a more complex adaptation of an advective box model, where the chemical transformations occur in a box following a pre-computed trajectory. In a fully 3-dimensional model, calculations are performed with temporally-varying gridded meteorological data, using either archive or forecast fields. The model used in this study is the HYbrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT)[20,21] model. Pollutant particles or puffs are transported through the domain

and the dispersion is calculated along the trajectory from the vertical diffusivity profile, wind shear, and horizontal deformation of the wind field. Air concentrations are calculated on a grid from the sum of the masses of each puff or particle within the cell.

A complete and detailed description of the model is available by Draxler and Hess [20,21]. However, a brief summary follows. The trajectory of a particle or puff, or the change of its position vector (**P**) with time,

$$\mathbf{P}(t+\Delta t) = \mathbf{P}(t) + 0.5\,[\,\mathbf{W}(P,t) + \mathbf{W}(\,\{\mathbf{P}(t) + \mathbf{W}(\mathbf{P},t)\,\Delta t\},\, t+\Delta t\,)\,]\,\Delta t, \qquad \textbf{(4)}$$

is computed from the average of the three-dimensional velocity vectors (**W**) at their initial and first-guess positions. Particle dispersion is computed by adding an additional velocity term to the advection equation (4) that includes a contribution from a turbulent velocity component,

$$W'(t+\Delta t) = R(\Delta t)\,W'(t) + W''\,(\,1 - R(\Delta t)^2\,)^{0.5}, \qquad \textbf{(5)}$$

which depends upon the turbulent velocity component at the previous time W'(t), a velocity auto-correlation coefficient (R), and a computer-generated random component (W"). The vertical turbulent velocity component is computed from the diffusivity profile which is a function of the profiles of wind and temperature and the mixed-layer depth. The mixed-layer depth is estimated to be the height at which the potential temperature exceeds the surface value by 2 degrees.

In HYSPLIT, puffs are released from each emission cell that has a non-zero value with an initial horizontal radius that equals the area of the emission cell (50 km square) and with the corresponding pollutant mass for each precursor species emitted. If the chemistry is linear, then transformations can occur from one species to the other independently on each puff. The pollutant's air concentration at a location is simply the sum of the contribution from all the nearby puffs. A non-linear process, such as ozone formation, requires a hybrid approach[22]. In this situation, the transformation from one species to another is also a function of the local concentration of one or more pollutants, i.e. information is required about all the puffs in the vicinity of the one for which the chemical transformations are being computed. This was the approach used in an earlier version of HYSPLIT[23] to compute sulfate deposition. To accomplish the non-linear chemical transformation, the current version of HYSPLIT is set up with two concentration grids, one with the averaging time desired for the air concentration output and the other with no averaging for instantaneous (over one time-step) air concentrations of each species. The instantaneous values are used in the chemical transformation equations which are solved in an Eulerian framework on the concentration grid. The modeling system becomes a hybrid of a Lagrangian-Eulerian approach; the transport and dispersion is computed along the pollutant particle or puff trajectory, while the chemical transformations are solved on a grid. This is similar to the approach used by Chock and Winker[24] in coupling their particle model with a chemical system.

In the implementation of the IER model within HYSPLIT, the total smog product (SP), is integrated on the puff (Eq. A2) each time step, according to the local temperature and

incoming solar radiation. Contributions from the precursor species of all the pollutant puffs are summed to the concentration grid to obtain their air concentrations. Ozone is then calculated on the concentration grid as a cell-average from Eqs. A9, A10, and A12. Note that the integration of SP is independent of solution of the pollutants on the concentration grid.

Model Input Data

Precursor emissions of non-methane volatile organic hydrocarbons (VOC), NO_x, and NO are defined hourly based upon the 1985 National Acid Precipitation Assessment Program (NAPAP) inventory[25] for a summer weekday and include all identified area, point, and mobile emissions. More recent emissions data were not available at the time this study was designed. In the box model calculation, all sources within the box emit the appropriate fractional mass each time step. In the HYSPLIT simulations, puffs are released every hour from each emission grid cell. Although the NAPAP emissions are given at a resolution of 20 km, they were aggregated into 50 km cells to speed the computation by reducing the number of emitted puffs, to reduce the model's sensitivity to the emission inventory, and to provide a certain degree of comparability between the resolution of the non-advective box model and the 3-dimensional modeling approach. However due to the nature of the 3-D model, its sensitivity to the emissions inventory will always be a factor as compared with the box model. Only emissions in the region between 25 N to 35 N and 105 W to 90 W were included in the simulations.

Estimates of isoprene emissions were added to the VOC pollutant mix using a simple formulation derived from Jacob et al.[26], where the emission flux,

$$Q_I \ (\text{kg km}^{-2} \text{ hr}^{-1}) = 2 \sin(\alpha) \exp[\ 0.096\ (T-298)\ ], \qquad \textbf{(6)}$$

is adjusted from a typical forest value (2 kg km^{-2} hr^{-1}) according to the local solar elevation angle (α) and air temperature. Isoprene emissions are only computed from those grid cells defined to have a forest land-use category based upon a 1 degree latitude-longitude inventory[27].

As noted earlier, meteorological data for all model simulations were derived from the EDAS [28] archive. The 3-hourly EDAS is an intermittent assimilation system consisting of successive 3-h Eta model forecasts and Optimum Interpolation (OI) analyses for a pre-forecast period on a 38 level, 29 km grid. The 3-h analysis updates allow for the use of high frequency observations, such as wind profilers, radar, and aircraft data. For this study, the 29 km data are archived on a 80 km, Lambert Conformal grid, covering the continental United States.

Hourly ozone measurements for the period 1 May through 30 September, 1997 were obtained from a single monitor (CAMS-35, 29.67 N, 95.13 W) selected by the Texas Natural Resources Conservation Commission (TNRCC - private communication) to be representative of the Houston, Texas environment. Measurements from the monitor were compared with model calculations of the average daily ozone maximum (always defined between 0700 and 1900 LST) calculated each day. The hourly measurements also were compared to the individual hourly model calculations.

GENERAL RESULTS

Daily maximum ozone predictions are compared to measurements for the regression model, the box model, and the HYSPLIT model. It is assumed that the more complex models, which include physics of the ozone formation chemistry, should be able to explain more of the variance in the measured concentrations than a regression model based only upon meteorological parameters. The non-advective box model is a substantial step up in complexity from the regression model. This model is almost as simple to run and it is not based upon empirical coefficients derived from dependent data. The regression model and box model use the same 3-h meteorological data, extracted at one point at the mid-boundary-layer level. The most complex approach considered, HYSPLIT, is comparable to an Eulerian grid model. The regional emission pattern and temporal and spatial variations in meteorology are all included in the calculation. The basic meteorological data, emissions, and chemical model, are the same as used in the non-advective box model, although with considerably more spatial definition.

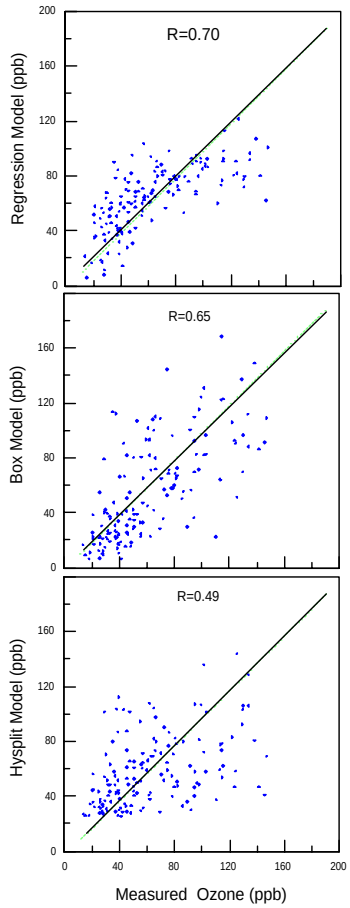

Figure 2 Daily maximum ozone measured at Houston from 1 May 1997 through 30 September 1997 and calculated ozone using the linear regression model (top panel), the non-advective box model (middle panel), and the HYSPLIT-IER model (bottom panel).

Regression Model

The regression model is intended to demonstrate how much of the variance in the ozone can be explained by the meteorological variables. Subsequent modeling approaches need to show a performance better than the regression baseline. The regression model is not designed to be a predictive tool because the coefficients are derived from the observations and there is no assurance that the model's performance (with the same coefficients) will be comparable in other years. The regression coefficients for Eq. 1 were computed from all the daily maximum ozone values from May through September of 1997 and the daytime (1200 - 0000 UTC) average meteorological values from the 925 hPa level at the location of the Houston sampler. The values of the coefficients, sorted in the order that they could explain the ozone variance, are a=-7.47, c=-1.08, b=-0.10, d=-0.02, and e=258. Although the overall correlation coefficient of 0.70 explains about half the variance, the scatter diagram shown in the top panel of Fig. 2 illustrates that the regression model substantially underestimates the highest concentrations. Underestimation of peaks is a common feature of regression methods[17].

More complex forecasting approaches, such as neural networks, step-wise regression, and lagged regression (using the previous day's ozone as a predictor) were compared with

measurements at 8 different cities by Comrie[4]. He found that the simple regression approach explained from 0.15 to 0.59 of the variance, with the majority of values above 0.5. It is certainly possible that better regression models may be employed for the Houston data that would capture its more complex meteorological environment. For instance, Bloomfield et al.[16] were able to explain 80% of the variance in maximum ozone concentrations at Chicago using a non-linear model. Other regression methods were also evaluated at Houston, such as using the previous day's maximum ozone value as a predictor, but they did not substantially raise the correlation, nor change the basic structure of underestimation of the highest concentrations. A regression approach using principal component analysis by Davis et al. [3] was able to explain 54% of the overall variance of ozone maxima at Houston, while results for individual meteorological clusters was as high as 73%. The main result for this analysis is that large scale meteorological factors can explain about half the variance in ozone at Houston. This appears to be consistent with other studies, and provides a baseline upon which future predictive schemes can be judged. Clearly, a better and well-conditioned statistical model may still be a superior choice over a chemical model for a customized prediction scheme.

Box Model

The non-advective box model includes the same single-point meteorological assumptions as the regression approach, but the IER chemical model is used to compute the ozone concentrations. The box model is almost as simple to run as the regression model and it is not based upon empirical coefficients derived from the dependent data. The box model uses the same single-station 3-hr extract meteorological data. However, in this case, the model is integrated (at 15 min time steps) to produce hourly ozone values from which a maximum daily predicted value is obtained. The results of the box model computation are shown as a scatter diagram in the middle panel of Fig. 2. Although the computation shows a slightly lower correlation (0.65) than the regression model (0.70), the results do not show any bias at the highest concentrations. Moreover the scatter appears to be much greater above 100 ppb. Because the model is derived independently from the Houston ozone data, it can reasonably be expected to be comparable in performance for other years. One important difference between the regression model and the box model is that the box model does not assume a linear relationship between the meteorological conditions and the resulting ozone maxima.

HYSPLIT-IER

The HYSPLIT-IER model combination is comparable to a 3-dimensional Eulerian model in that the Lagrangian component is used to compute the advection and dispersion while the chemical equations are solved on the concentration grid. The regional emissions and both temporal and spatial variations in meteorology are included in the calculation. The emissions and chemical model are the same as are used in the non-advective box model. In the Houston HYSPLIT-IER simulations, pollutant puffs were released each hour from every non-zero emission grid cell with appropriate masses of VOC, NO_x, and NO, that correspond to the total area, point, and mobile emissions in that cell. Puffs were released at the first time step of each hour's integration at 429 cells over the south central U.S. (primarily Texas and excluding Mexico) modeling domain. In

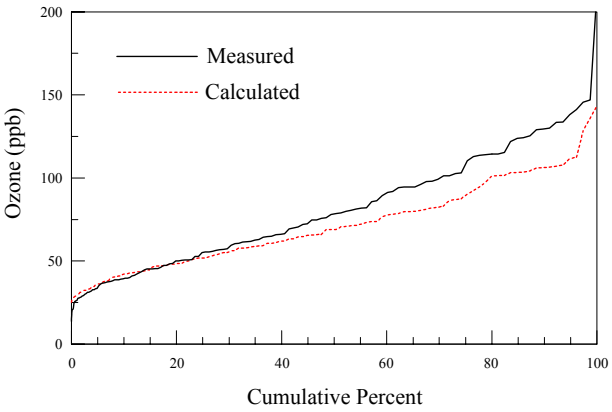

Figure 3 Cumulative frequency distribution of maximum ozone, measured and calculated at Houston using the HYSPLIT-IER, Box, and Regression models.

addition, biogenic hydrocarbon (Eq. 6) emissions were added to each cell that is dominated by a forest land-use category. Pollutant particles were then tracked and dispersed for 48 hours according to the meteorological analysis fields (EDAS) from the Eta model. The model's computational domain, centered at 30 N and 97.5 W, covers about the same area as the emissions grid (a 1200 km square). The concentration grid was defined at about a 50 km horizontal (0.5 deg) resolution and in 5 vertical layers (0-500 m, 500-1000 m, 1000-2000 m, 2000-5000 m, 5000-10000 m) AGL. An ozone background concentration of 20 ppb was assumed for all hours. Mueller[2] reported background levels (assumed to be at the 5th percentile) in Great Smoky National Park (GSNP) on the order of 10 to 30 ppb, but perhaps 20 ppb might be too low for the Houston region.

The HYSPLIT-IER scatter diagram results are shown in the bottom panel of Fig. 2 and they show the least correlation (0.49) of the three methods, only explaining about 25% of the variance. In addition, the highest concentrations are underestimated as well. The marked difference between the box model and an advective model (HYSPLIT) results cannot be explained by the chemical prediction scheme, identical for both methods, nor by unrepresentativeness of the driving meteorological data, both sets from the EDAS. The primary difference is that the HYSPLIT-IER approach includes a more sophisticated advection process. The wind in the box model is used only to ventilate pollutants out of the box. There are no sources defined outside of the box to provide an upwind boundary condition for precursor pollutants or ozone. The HYSPLIT approach transports pollutants over the entire modeling domain, permitting them to enter and exit the Houston concentration grid cell. The fact that the additional modeling complexity provided a less satisfactory result suggests that the high ozone events at Houston are dominated by local emissions occurring when winds are weak -- a common scenario for the ozone season.

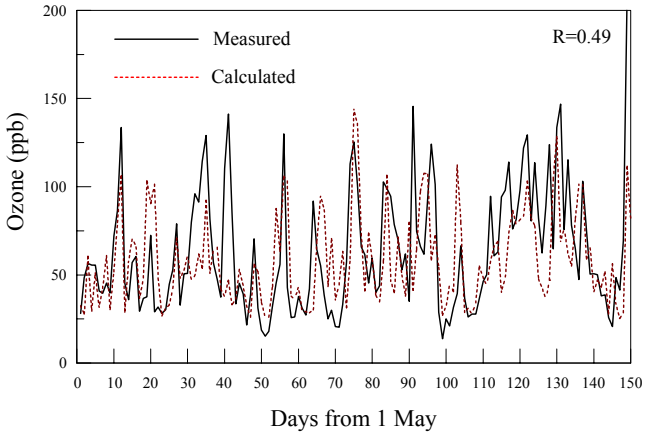

Figure 4 Time series from 1 May 1997 through 30 September 1997 of daily maximum ozone concentrations measured at Houston and calculated using the HYSPLIT-IER models.

The HYSPLIT-IER calculation included a predetermined background concentration of 20 ppb and this implicit under-prediction is consistent with known

biases of the GRS/IER chemistry schemes. The box model, also using IER chemistry, did not show such a large bias. The difference is in part due to how the background was treated in each approach. In the box model, the background was implicitly included by the 0.6 factor of Eq. 3, while in the HYSPLIT approach 20 ppb was explicitly added to the ozone calculation result. Frequency distributions for the daily maximum ozone values, measured and calculated, are shown in Fig. 3. The figure shows that the 95th percentile concentration is under-predicted by about 20 ppb by all the models. For HYSPLIT, this additional under-prediction and previously assumed background of 20 ppb gives an overall bias of 40 ppb, comparable to what was expected from the results of Tonnesen and Jeffries[12]. Although the biases of each modeling approach are comparable at the highest ozone concentrations, only the box model shows about the same bias over the entire predictive range, suggesting that it may be the nearest approximation to reality, especially after properly adjusted for ambient background levels.

Although the scatter diagram and associated bias of the HYSPLIT-IER approach suggests that there may be limitations to the modeling approach, the corresponding time series of maximum measured and calculated ozone, shown in Fig. 4, shows a very reasonable agreement between model calculations and measurements for most of the multi-day ozone episodes. Poor agreement, when model results and observations are paired in both space and time, is not uncommon when results are controlled by random atmospheric processes. Although the model bias is still evident in the time series results, the model captures virtually all of the short-term peaks and, although perhaps of less interest in a forecast environment, very faithfully reproduces the longer-term trends in ozone. This can be seen in Fig 5, which shows the correlation between measured and calculated ozone with increasing averaging time. The correlation quickly rises to near 0.7 for multi-day episodes and then stays at that level until it rises to near 0.9 for monthly ozone estimates.

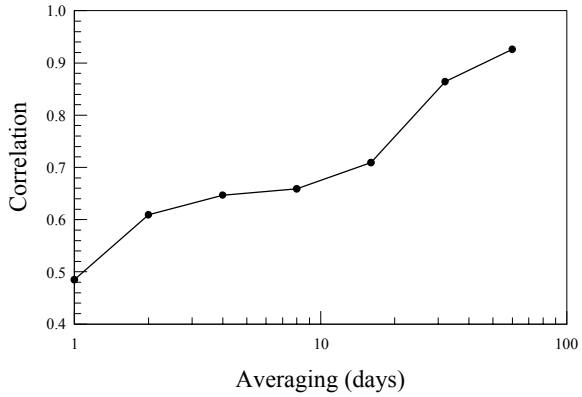

Figure 5 Correlation coefficient between daily measured and HYSPLIT-IER model calculated ozone at Houston as a function of averaging time.

One aspect not directly addressed in this study is the potential degradation of the model's ozone predictive ability when meteorological forecast data are used rather than the EDAS analysis fields. The short-term error for forecasts of 24 h or less is expected to be small. Ryan[29] compared a 24 h regression ozone forecast with the post-facto calculation using observed meteorological data and found that the overall correlation only increased from 0.73 to 0.79.

CASE STUDY ANALYSIS

It is somewhat discouraging that the more complex HYSPLIT-IER simulation does not provide any improvement over the box model. One possible explanation, noted earlier, is that the high ozone events are dominated by local emissions and that the inclusion of the additional

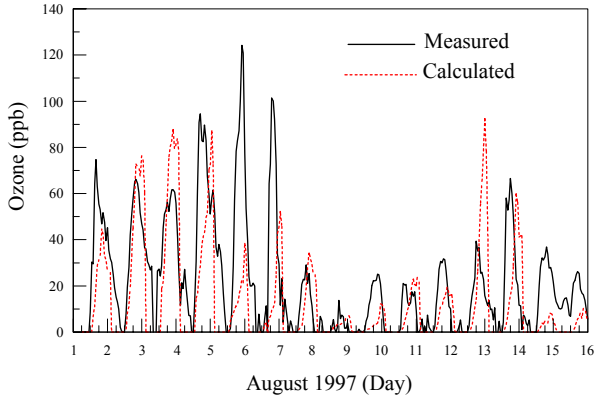

Figure 6 Hourly measured and HYSPLIT-IER calculated ozone at Houston from 1 August 1997 through 15 August 1997.

complexity of advection and dispersion only adds the potential of errors to accumulate in the calculation if the meteorological data do not accurately represent the flow field. In simulations of this kind, perhaps more detailed estimates of emissions, transport direction, and output grids are required for increased accuracy of model predictions at specific locations, perhaps even including consideration of sea-breeze and other mesoscale features. An examination of the hourly predictions may provide some insight to model performance.

The HYSPLIT-IER model simulations were usually run for fortnight intervals because this matched the structure of the meteorological data files. One such period starting on August 1st, 1997 provides an interesting example for study because measured ozone values were high during the first week (60 - 120 ppb) and dropped dramatically (20 - 60 ppb) during the second week. The model and measurement results are shown in Fig. 6 for this period. Note that no background ozone concentration has been added to the hourly calculations - background values were only included for the previous maximum concentration predictions. During the first four days the model performance is quite good with both measurements and predictions of the daily maxima near 60 - 80 ppb. The advection during the last two days of this period is illustrated by 24-h back trajectories, calculated at 6-h intervals, and shown in the top panel of Fig. 7. The trajectories, started at 925 hPa, represent the flow typical of air within the mixed layer and the transport component that was an integral part of the HYSPLIT-IER ozone calculation. Most of the trajectories did not go very far and are consistent with the higher ozone values and good model performance (i.e. good performance corresponds to weak advection).

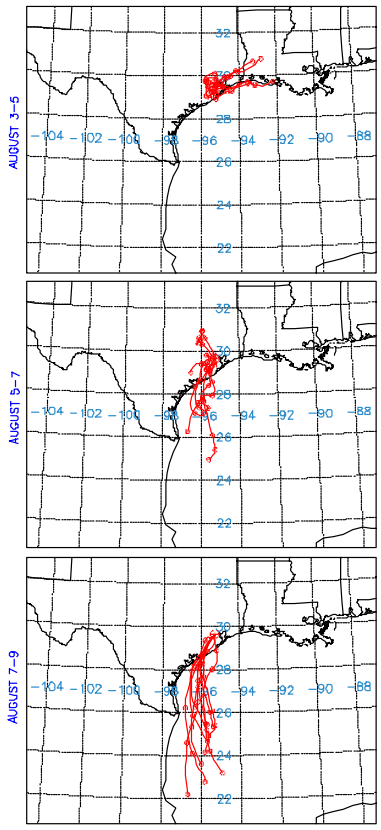

Figure 7 Twenty four hour back trajectories calculated every 6 hours from Houston for the period 1200 UTC to 1200 UTC from 3-5 August (top panel), 5-7 August (middle panel), and 7-9 August (bottom panel).

The 6th and 7th of August showed the highest measured ozone during the period and the largest model under-prediction. The trajectories (Fig. 7 - middle panel) show a flow that is slightly stronger than before and with a more consistent directional component from regions of low emissions (the Gulf of Mexico), potentially explaining some of the model's under-prediction. Davis et al.[3] also found that Houston's lowest ozone concentrations were associated with flow from the Gulf

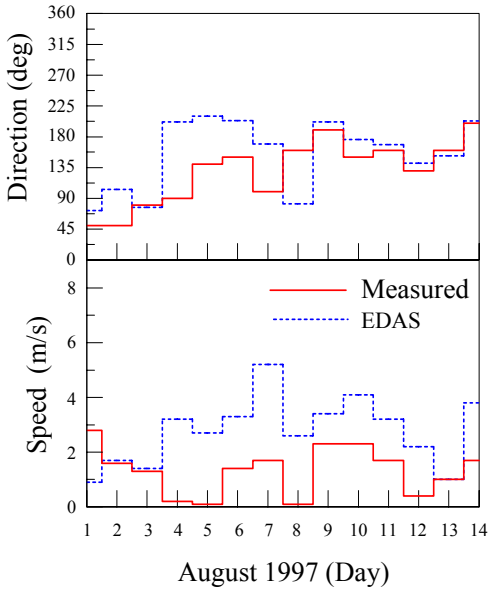

Figure 8 Daily vector average 10 m level wind direction and speed measured at Houston and the corresponding value from the EDAS meteorological data archive.

of Mexico. This suggests that the flow fields from the EDAS archive may not be representative of the true conditions in Houston during this time and that the ozone is still controlled by emissions in the Houston-Galveston corridor.

The third period, starting on the 8th of August shows a large drop in both measured and predicted ozone, and the trajectories (Fig. 7 - bottom panel) are long, indicating strong winds, from the south with little curvature from a region of little precursor emissions. Other studies, summarized by Stohl[30], have shown that computed trajectories are much more accurate in strong wind situations than during light winds.

Clearly there is a pattern in the direction of the upwind flow and associated ozone measurements, but the correspondence of the trajectory directions with model performance is not at all clear, considering that these same trajectories are the basis of the ozone prediction. It is possible that the wind direction error (or representativeness) of the EDAS archive is different in each flow regime, a suggestion consistent with previous trajectory studies[31]. One explanation of a potential source of trajectory error is shown in Fig. 8, where the 10 m level daily vector average EDAS wind is compared with the measured vector average at Houston. The largest differences between measured and model winds, in both direction and speed, occur between August 4th and 8th, on the order of 90° in direction and 2 to 4 m/s in speed. More significantly, the wind speed is substantially over-predicted, supporting the earlier contention that the 3-D model computed too much advection of cleaner air into the Houston cell. This certainly does not explain all the error, as the period of the most consistent trajectories (7th - 9th) also showed errors in the wind prediction. However, when trajectories originate from low emissions regions, directional errors have little consequence.

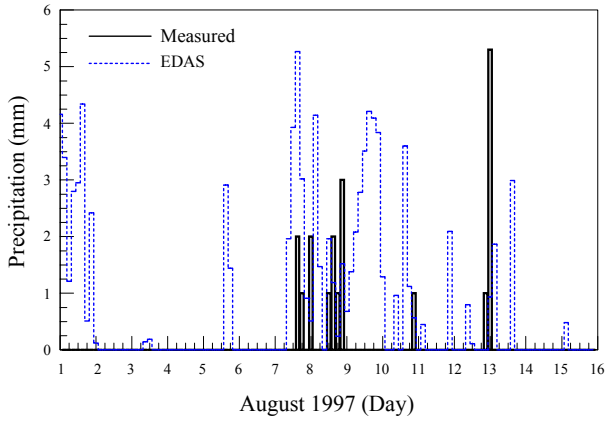

Figure 9 Precipitation accumulated into 3 hour bins from the EDAS archive and measured at Houston. Trace measurements were assumed to equal 1 mm.

Although the model calculation did not include any wet or dry deposition, the occurrence of rain would certainly be expected to lower measured ozone values, while model predictions during periods of model rain (which may or may not match observations) will still have some effect in reducing ozone values due to the increased cloud amounts and reduction in

photolysis. Rainfall amounts measured at Houston are compared with the EDAS values on Fig. 9. Rainfall prediction is probably one of the most difficult challenges for any meteorological model. The results show a general over-prediction of rain. Not unexpectedly, the largest rainfall amounts, both measured and predicted, are associated with the lowest ozone concentrations, especially during the 8th and 9th of August. The model's sensitivity to rain is surprising in that precipitation was not explicitly treated and sensitivity tests with the box model showed only a slight dependence on cloud amount. It is more likely that the model's rain dependence is only an indicator of the quality of the entire meteorological prediction, including wind direction, temperature, cloud amount -- all factors that will affect the subsequent ozone calculation. Perhaps it is only coincidence, but the largest ozone under-prediction (on the 6th) is associated with model-predicted rain but no measured rain, while the largest ozone over-prediction on the 13th is associated with the largest measured rainfall but only a small rainfall prediction for that day.

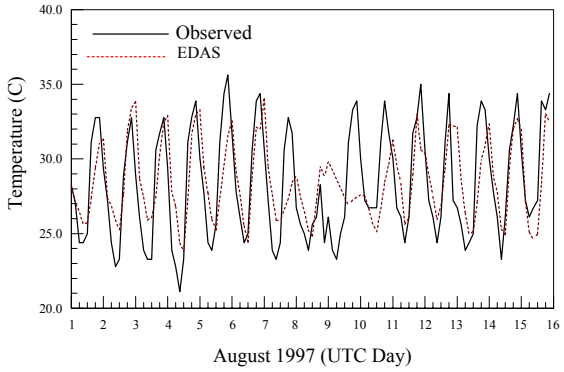

Figure 10 Hourly air temperature measured at 2 m at Houston and from the EDAS archive from 1 - 15 August 1997.

The relationship of the rain prediction to the other meteorological parameters can be seen in Fig. 10 as well, where the 2 m EDAS air temperature is compared with temperatures measured at Houston. In general the EDAS estimates are quite good - especially the maximum values - and the EDAS archive does not appear to have the same under-predictive bias found by Ryan[29] for NOAA's Nested Grid Model. However, minimum EDAS temperatures are consistently higher than the observations. The period between August 8th and 11th shows the poorest model performance from the EDAS and also corresponds with the period of over predicted rainfall.

SUMMARY AND CONCLUSIONS

Several ozone modeling approaches were investigated to determine if the uncertainties in the meteorological data would be sufficiently large to discourage the application of more physically realistic ozone forecast models. Three computational schemes were evaluated: a linear regression model to provide a baseline to show how much of the variability in maximum ozone can be attributed to meteorological factors, a non-advective box model that couples the local meteorological conditions with a simple ozone formation scheme (the IER[7] model), and a fully 3-dimensional transport and dispersion model also coupled with the IER scheme. Correlation between measured daily maximum ozone and model predicted ozone was found to be 0.70 for the regression model, 0.65 for the box model, and 0.49 for the 3-D model. Absolute error statistics (in terms of ppb) are summarized in Table 2 for all daily maximum ozone concentrations and for those days that the measured ozone maximum exceeded 100 ppb. Although in general the errors are large, the root mean square error and the mean absolute error are smaller for the regression approach than for the other two methods. However, it is important

to note that all of the models have a free parameter which is chosen to give the best overall fit to the measurements: the regression model by definition, the box model through the ventilation factor, and the 3-D model through the selection of a background concentration. For the highest concentrations, the box model error results were more similar to the regression model than the 3-D model and showed less underestimation of the mean ozone concentrations.

Table 2. Model Evaluation Statistics (ppb) for 5 months of daily maximum ozone for each day (All) and for those days where the observed ozone maximum exceeded 100 ppb (>100).

Model	Regression		Box		3-D	
	All	>100	All	>100	All	>100
Observed mean	63.6	126	63.3	126	63.7	126
Predicted mean	63.6	88.4	55.8	95.8	58.5	81.1
Mean absolute error	18.7	37.2	23.3	40.1	23.8	48.7

The results are consistent with those found by other researchers using more complex models. A median correlation of 0.5 for maximum daily ozone was reported[32] over a 120 station European network between 1989-1996 using a Lagrangian box model with a more complete chemical prediction scheme (45 species, 100 reactions)[33]. Although the more complex 3-D HYSPLIT-IER model could only account for 25% of the variance in the daily ozone maxima, similar to the European network results, the 5 month time series of measured and predicted concentrations showed that the model predictions were very good for multi-day episodes (correlations approaching 0.7).

The more complex 3-D modeling approach showed poorer results than the non-advective box model because the Houston ozone maxima were dominated by local factors rather than transport of pollutants from the surrounding region and the transport was not always modeled correctly. One conclusion is that better results are not just obtained by the incorporation of additional modeling complexity but that there needs to be a comparable increase in accuracy. A comparison of the EDAS wind direction and speed values at Houston showed the largest difference from the observations on those days in which the ozone underestimates were the greatest. These periods also tended to correspond with poor precipitation and temperature estimates. Perhaps some of these predictions may be improved with better spatial resolution in the meteorological data. Gaza[34] showed that it is important for ozone forecast models to simulate the formation and movement of smaller scale features such as the sea-breeze front and troughs, although he suggested that many of the currently available mesoscale models may not be able to accurately handle these features. The issue of accuracy and resolution is not just a meteorological one, but includes improvements in the spatial definition of emissions and higher resolution sampling grids. All the factors that influence the transport of pollutants from source to receptor will not be well resolved for larger grid cell sizes and may degrade the calculation of more complex modeling approaches.

There are still considerable uncertainties to be addressed in future studies, including the importance of accurately modeling small scale flow features, the importance of better emissions inventories (considering the population of Houston more than doubled between the 1980 and 1990 census), and how a more complete chemical scheme, although at greater computational cost, might improve the ozone forecast. The results of this study suggest that in regions where there are extensive existing measurements, it would be difficult to design a more physically realistic system that would perform better than a site specific regression technique although the box model approach seemed to have a superior performance over a broader range of concentration levels. Perhaps a hybrid regression-box model methodology can be developed that would take advantage of the strengths of each method. More physically realistic models, if run at sufficiently fine resolution with input data at comparable resolution, have the potential to address these modeling complexities. However, it would still be necessary to demonstrate that the underlying variability of the atmospheric would not limit the precision of such a system.

ACKNOWLEDGEMENTS

The author wishes to thank Dale Hess from the BoM Australia, Martin Cope and Peter Hurley from CSIRO Aspendale, for providing valuable feedback and discussions regarding the implementation of the IER/GRS modeling approaches; Stu McKeen from NOAA's Aeronomy Laboratory for providing the 1985 NAPAP emissions data already processed for modeling applications; and Dan White from TNRCC for providing ozone concentration data.

REFERENCES

1. Ryan, W.F.; Doddridge, B.G.; Dickerson, R.R.; Morales, R.M., Hallock, K.A.; Roberts, P.T.; Blumenthal, D.L.; Anderson, J.A.; Civerolo, K.L. "Pollutant transport during a regional O_3 episode in the Mid-Atlantic States," *J. Air & Waste Manage. Assoc.* **1998**,*48*, 786-797.

2. Mueller, S.F. "Characterization of ambient ozone levels in the Great Smoky Mountains National Park," *J. Appl. Meteor.* **1994**, *33*, 465-472.

3. Davis, J.M.; Eder, B.K.; Nychka, D.; Yang, Q. "Modeling the effects of meteorology on ozone in Houston using cluster analysis and generalized additive models," *Atmos. Environ.* **1998**, *32*, 2505-2520.

4. Comrie, A.C. "Comparing neural networks and regression models for ozone forecasting," *J. Air & Waste Manage. Assoc.* **1997**, *47*, 653-663.

5. U.S. Environmental Protection Agency, *National Air Pollutant Emission Trends, 1990-1994*, EPA-454/R-95-011, **1995**, Office of Air Quality Planning and Standards, Research Triangle Park, NC.

6. Hess, G.D.; Carnovale, F.; Cope, M.E.; Johnson, G.M. "The evaluation of some

photochemical smog reaction mechanisms — I. Temperature and initial composition effects," *Atmos. Environ.* **1992**, *26A*, 625-641.

7. Johnson, G.M. "A simple model for predicting the ozone concentration of ambient air," In *Proc. 8th Int. Clean Air Conf., 2*, 715-731, Eds. H.F. Hartmann, J.N. O'Heare, J. Chiodo, and R. Gillis, Clean Air Soc. of Australia and New Zealand, Melbourne, Australia, 7-11 May **1984**.

8. Johnson, G.M.; Quigley, S.M. "A universal monitor for photochemical smog," *Proceedings of the 82nd Annual Meeting of the Air and Waste Management Assoc*., June 25-30, **1989**, Anaheim, CA, Paper 89-29.8

9. Azzi, M.; Johnson, G.M. "Notes on the derivation, The Integrated Empirical Rate Model, Version 2.2," 20 March **1992**, CSIRO Division of Coal and Energy Technology, PO Box 136, North Ryde, NSW 2113, Australia.

10. Azzi, M.; Johnson, G.M. "Airtrak: new developments," *Clean Air* **1993**, *27*, 191-193.

11. Venkatram, A.; Karamchandani, P.; Pai, P.; Goldstein, R. "The development and application of a simplified ozone modeling system (SOMS)," *Atmos. Environ.* **1994**, *28*, 3665-3678.

12. Tonnesen, S.; Jeffries, H.E. "Inhibition of odd oxygen production in the carbon bond four and generic reaction set mechanisms," *Atmos. Environ.* **1994**, *28*, 1339-1349.

13. Chang, T.Y.; Chock, D.P.; Nance, B.I.; Winkler, S.L. "A photochemical extent parameter to aid ozone air quality management," *Atmos. Environ.* **1997**, *31*, 2787-2794.

14. Wratt, D.S.; Hadfield, M.G.; Jones, M.; Johnson, G.M. "Predicting the impact of a proposed gas fired power station on photochemical pollution levels around Auckland," *Proceedings, 10th Int. Conference of the Clean Air Soc. of Australia and New Zealand*, Auckland, N.Z., **1990**, pp 159-164.

15. Gery, M.W.; Whitten, G.Z.; Killus, J.P.; Dodge, M.C. "A photochemical kinetics mechanism for urban and regional scale computer modeling," *J. Geophys. Res.* **1989**, *94(D10)*, 12,925-12,956.

16. Bloomfield, P.; Royle, J.A.; Steinberg, L.J.; Yang, Q. "Accounting for meteorological effects in measuring urban ozone levels and trends," *Atmos. Environ.* **1996**, *30*, 3067-3077.

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

18. Schere, K.L.; Demerjian, K.L. "A photochemical box model for urban air quality simulation," *Proceedings, 4th Joint Conference on Sensing of Environmental Pollutants* **1978**, pp 427-433, American Chemical Society.

19. Eliassen, A.; Hov, O.; Isaksen, I.S.A.; Saltbones, J.; Stordal, F. "A Lagrangian long-range transport model with atmospheric boundary layer chemistry," *J. Appl. Meteor.* **1982**, *21*, 1645-1661.

20. Draxler, R.R.; Hess, G.D. "Description of the HYSPLIT_4 Modeling System," *NOAA Tech. Memo. ERL ARL-224,* Dec. **1997**, 24 p.

21. Draxler, R.R.; Hess, G.D. "An Overview of the HYSPLIT_4 Modelling System for Trajectories, Dispersion, and Deposition," *Aust. Meteor. Mag.* **1998**, *47,* 295-308.

22. Peters, L.K.; Berkowitz, C.M.; Carmichael, G.R.; Easter, R.C.; Fairweather, G.; Ghan, S.J.; Hales, J.M.; Leung, L.R.; Pennell, J.M.; Potra, F.A.; Saylor, R.D.; Tsang, T.T. "The current state and future direction of Eulerian models in simulating the tropospheric chemistry and transport of trace species: a review," *Atmos. Environ.* **1995**, *29*, 189-222.

23. Rolph, G.D.; Draxler, R.R.; de Pena, R.G. "Modeling sulfur concentrations and depositions in the United States during ANATEX," *Atmos. Environ.* **1992**, *26A*:73-93.

24. Chock, D.P.; Winkler, S.L. "A particle grid air quality modeling approach, 2. Coupling with chemistry," *J. Geophys. Res.* **1994**, *99*, 1033-1041.

25. U.S. Environmental Protection Agency, "The 1985 NAPAP emissions inventory (Version 2): Development of the Annual Data and Modular Tapes," **1989**, EPA-600/7-89-012a.

26. Jacob, D.J.; Logan, J.A.; Yevich, R.M.; Gardner, G.M.; Spivakovsky, C.M.; Wofsy, S.C.; Munger, J.W.; Sillman, S.; Prather, M.J.; Rodgers, M.O.; Westberg, H.; Zimmerman, P.R. "Simulation of summertime ozone over North America," *J. Geophys. Res.* **1993**, *98*, 14,797-14,816.

27. Wilson, M.F.; Henderson-Sellers, A. "A global archive of land cover and soils data for use in general circulation climate models," *J. Clim.* **1985**, *5*, 119-143.

28. Rodgers, E.; Black, T.L.; Deaven, D.G.; DiMego, G.J.; Zhao, Q.; Baldwin, M.; Junker, N.W.; Lin, Y. "Changes to the operational 'early' Eta analysis/forecast system at the National Centers for Environmental Prediction," *Wea. and Forecast.* **1996**, *11*, 391-413.

29. Ryan, W.F. "Forecasting severe ozone episodes in the Baltimore metropolitan area," *Atmos. Environ.* **1995**, *29*, 2387-2398.

30. Stohl, A. “Computation, accuracy and applications of trajectories — a review and bibliography,” *Atmos. Environ.* **1998**, *32*, 947-966.

31. Kahl, J.D.W. “On the prediction of trajectory model error,” *Atmos. Environ.* **1996**, *30*, 2945-2957.

32. Co-operative programme for monitoring and evaluation of the long range transmission of air pollutants in Europe (EMEP), “Calculations of tropospheric ozone and comparison with observations,” MSC-W Status Report **1998**, Norwegian Meteorological Institute, P.O. Box 43, Blindern, N-0313, Oslo, Norway.

33. Simpson, D. “Photochemical model calculations over Europe for two extended summer periods: 1985 and 1989. Model results and comparisons with observations,” *Atmos. Environ.* **1993**, *27A*, 921-943.

34. Gaza, R.S. “Mesoscale meteorology and high ozone in the Northeast United States,” *J. Appl. Meteor.* **1998**, *37*, 961-977.

35. Azzi, M.; Johnson, G.M.; Cope, M. “An introduction to the generic reaction set photochemical smog mechanism,” In *Proceedings of the International Conference of Clean Air Soc. of Aust. and N.Z.* **1992**, *2*, 451-462.

ABOUT THE AUTHOR

Roland Draxler is a research meteorologist with the Air Resources Laboratory of the National Oceanic and Atmospheric Administration at 1315 East West Highway, Silver Spring, MD 20910. His recent focus has been on modeling the long-range transport of pollutants.

APPENDIX

The Integrated Empirical Rate (IER) model is an algebraic solution of the Generic Reaction Set (GRS), and is based upon the definition of a quantity called the smog produced (SP),

$$\{SP\}_t = \{O_3\}_t - \{O_3\}_o + \{NO\}_o - \{NO\}_t \qquad \textbf{(A1)}$$

which is defined as the ozone produced plus the oxidized nitric oxide. The formation of SP is linear with the precursor concentrations and sunlight. In the light-limited regime

$$\{SP\}_a = \int_0^t Z\, k_1\, dt \qquad \textbf{(A2)}$$

where the integration represents the photolysis of nitrogen dioxide to nitric oxide. In the NO_x limited regime

$$\{SP\}_b = \beta\, \{NO_x\}_0 \qquad \textbf{(A3)}$$

and where the {SP} at any time will be the minimum from the two regimes,

$$\{SP\}_t = MIN\, [\, \{SP\}_a\, ,\, \{SP\}_b\,], \qquad \textbf{(A4)}$$

which may be limited by the amount of NO_x available. The NO_x stochiometric coefficient , β, for the maximum SP formation equals 4.09. The "Z" term is defined by

$$Z = (\, A_1\, \{ROC\} + A_2\, \{ISOP\}\,)\, e^{-1000\,\gamma\,(1/T\, -\, 1/316)}\, , \qquad \textbf{(A5)}$$

where γ is an empirical constant (4.71) derived from the smog chamber studies, T (^oK) is temperature, and A_1 is the average activity coefficient $(0.0067)^{35}$ for non-methane hydrocarbons and A_2 is the activity coefficient (0.0117 - Cope, private communication) for isoprene/voc mixtures. The rate coefficient for NO_2 photolysis [18] is given for three different zenith angles by

$$\begin{aligned} k_1 &= [4.23 \times 10^{-4} + (1.09 \times 10^{-4} / \cos\theta)\,]\, S && 0^o \le \theta < 47^o \\ k_1 &= [5.82 \times 10^{-4}]\, S && 47^o \le \theta < 64^o \\ k_1 &= [-0.997 \times 10^{-4} + 1.20 \times 10^{-3}\,(1-\cos\theta)\,]\, S && 64^o \le \theta \le 90^o \end{aligned} \qquad \textbf{(A6)}$$

where θ is the solar zenith angle, S (W m^{-2}) is the total solar radiation at the ground. The SP concentration is also used to compute the loss of NO_x which is assumed to occur due to conversion of smog to Stable Non-Gaseous Nitrate (SNGN) particles and is defined as a linear proportion to the SP concentration

$$\{SNGN\}_t = 0.125\, \{SP\}_t \qquad \textbf{(A7)}$$

and where the losses are limited by the NO_x concentration

$$\{NO_x\}_t = \{NO_x\}_0 - MIN[\, \{NO_x\}_0\, ,\, \{SNGN\}\,], \qquad \textbf{(A8)}$$

and NO_x is defined by

$$\{NO_x\}_t = \{NO\}_t + \{NO_2\}_t \, . \qquad \textbf{(A9)}$$

The IER systems of equations is solved using the photo-stationary state equation assumption to balance the formation and destruction (nitric oxide-ozone titration reaction) of ozone,

$$\{O_3\}_t = k_1 \, \{NO_2\}_t \, / \, k_4 \, \{NO\}_t \, , \qquad \textbf{(A10)}$$

and where the k_4 titration reaction constant[35] is

$$k_4 = 9.24 \text{x} 10^5 \, T^{-1} \, e^{-1450/T} \, . \qquad \textbf{(A11)}$$

The remaining species are determined from the algebraic solution of Eqs. A1, A9, and A10:

$$\{NO\} = 0.5[(k_1/k_4+\{C\})^2 + 4k_1/k_4\{NO_x\}]^{0.5} - 0.5(k_1/k_4+\{C\}), \qquad \textbf{(A12)}$$

where

$$\{C\} = \{SP\}_t + \{O_3\}_o - \{NO\}_o \, . \qquad \textbf{(A13)}$$