PERGAMON

Atmospheric Environment 33 (1999) 5023–5036

ATMOSPHERIC ENVIRONMENT

www.elsevier.com/locate/atmosenv

Climatology and forecast modeling of ambient carbon monoxide in Phoenix, Arizona

Andrew C. Comrie*, Jeremy E. Diem

Department of Geography and Regional Development, University of Arizona, Tucson, AZ 85721, USA

Received 19 November 1997; received in revised form 18 June 1999; accepted 18 June 1999

Abstract

We perform a climatology of factors influencing ambient carbon monoxide (CO), in which we examine the relationships between meteorology, traffic patterns, and CO at seasonal, weekly, and diurnal time scales in Phoenix, Arizona. From this analysis we identify a range of potentially important variables for statistical CO modeling. Using stepwise multivariate regression, we create a suite of models for hourly and 8-h ambient CO designed for daily operational forecasting purposes. The resulting models include variables and interaction terms related to anticipated nocturnal atmospheric stability as well as antecedent and climatological CO behavior. The models are evaluated using a range of error statistics and skill measures. The most successful approach employs a two-stage modeling strategy in which an initial prediction is made that may, depending on the forecast value, be followed by a second prediction that improves upon the first. The best models provide accurate daily forecasts of CO, with explained variances approaching 0.9 and errors under 1 ppm. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Inversion; Climate; Urban air quality; Air pollution forecast; Statistical modeling

1. Introduction

1.1. Background and aims

Multivariate statistical models for predicting daily air pollution concentrations are widely used by urban and regional air quality agencies. Such models typically link a suite of meteorological conditions and emission patterns to pollutant concentrations. The vast majority of these models have been developed for ozone, primarily a summer pollutant, but the number of such models for other pollutants (especially those predominant in winter) is quite small. Remarkably few examples exist in the literature regarding detailed models of ambient carbon monoxide (CO) concentrations and links to meteorological conditions. This paper outlines a climatological analysis and the subsequent development of a set of operational forecast models for CO in Phoenix, Arizona.

Like many cities in the western United States, the greater Phoenix area (1995 population $\approx 2.6 \times 10^6$) is a fast-growing metropolitan center with a dispersed layout, a low population density, a high dependency on motor vehicles, and a relative lack of public transportation. Thus, Phoenix is characterized by large travel distances and little option to walk or use public transport; as a result, the city has an annual per capita gasoline use that is over 1.5 times higher than that for New York City (Newman and Kenworthy, 1989). When combined with a relatively stable winter atmosphere, this situation provides the potential for high ambient CO concentrations.

In 1984, CO concentrations in Phoenix exceeded the national ambient air quality standard (9 ppm averaged over 8 h) on 99 days (US EPA, 1996). The situation improved dramatically in the late 1980s due to fleet modernization, reformulated fuel, and other pollution control measures. In the last decade exceedances have typically been near zero each year, but Phoenix continues to be designated by the EPA as a CO non-attainment area. One consequence of this is a desire by the state air

* Corresponding author.

1352-2310/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S1352-2310(99)00314-3

quality agency to develop a daily CO forecast model as part of an advisory system for potential high CO concentrations.

This study was initiated to provide such a model. The broad goals are twofold. First, to understand the factors that influence CO pollution in Phoenix, we examine variability in motor vehicle traffic patterns (as a proxy for CO emissions) and we identify important meteorological variables and antecedent atmospheric conditions, as they all relate to ambient CO concentrations. Second, we develop a set of models (and associated error measures) that incorporate this knowledge to forecast winter CO concentrations accurately for central Phoenix.

1.2. CO and atmospheric conditions

CO is a primary pollutant resulting from incomplete combustion. In urban areas, transportation is the major CO source, with the principal removal mechanism being the reaction with the OH radical. CO concentrations are typically greatest over urban areas, reaching their highest levels adjacent to busy roads. Seasonally, CO concentrations are highest in winter during stagnant conditions. In the western United States, high air-pollution potential is associated with several synoptic-scale features: the most important feature is an anticyclone centered over the Great Basin, usually observed between mid-latitude cyclones, associated with light winds and shallow mixing depths during the fall and winter (Holzworth, 1962; Comrie 1996). Colucci and Begeman (1969) compared CO concentrations in Detroit, New York, and Los Angeles and found that Los Angeles' high levels are related to frequent atmospheric temperature inversions and lower wind speeds. Robinson and Boyle (1979) used a synoptic weather classification scheme to classify regional weather patterns in the St. Louis area, identifying weather patterns corresponding with high CO levels. Bardeschi et al. (1991) noted the importance of antecedent concentrations, emissions, and meteorological conditions during the hours prior to the high CO concentrations.

During the 1970 New York City taxi strike, CO emissions in midtown Manhattan were reduced by 34%; however, due to coincident low wind speeds, CO concentrations were reduced by only a few percent (Norbeck et al., 1979). At low wind speeds, a substantial reduction in vehicular emissions in a midtown area can lead to only a slight improvement in air quality, thus wind speed affects CO levels significantly. Nkendirim (1988) also determined that atmospheric ventilation factors are important with respect to the air pollution process. In Phoenix, ventilation factors such as surface wind speed and the strength of the low-level inversion influence the diurnal behavior of CO (Tiao et al., 1989).

1.3. CO modeling

An attempt at predicting hourly and daily maximum CO concentrations in the Los Angeles basin was undertaken by McCollister and Wilson (1975), using antecedent CO concentrations as predictors in a linear stochastic model. This one-dimensional model predicted extreme events poorly, and the authors emphasized the importance of including meteorological variables and pattern recognition techniques (e.g., synoptic circulation patterns) in the models. Tiao et al. (1975) modeled the variation in CO concentrations in downtown Los Angeles by examining traffic densities, wind speeds, and inversion heights, and concluded that the effect of a change in the inversion height on CO is much stronger at low inversion heights. Since CO is primarily emitted within approximately 1 m of the surface, an extremely low inversion base height and a steep temperature gradient within the inversion layer will keep the pollutant close to the surface.

Aron and Aron (1978) forecast daily maximum CO concentrations throughout the Los Angeles area and discovered that the most important variables were the preceding day's CO concentration, pressure differences between nearby stations, 500 mb height, surface temperature, 1000 and/or 950 mb height, day of the week, length of daylight, potential solar radiation, and the height of the inversion base. Maximum mixing height did not prove significant in any of the models, and thus afternoon conditions have a low forecasting potential. Jakeman et al. (1991) used a hybrid deterministic/statistical model to predict seasonal (rather than daily) extremes of 1-h average urban CO concentrations, and in a similar article Miles et al. (1991) noted the limited general ability of deterministic models to predict high-percentile extreme events. Glen et al. (1996) developed an empirical model of monthly CO for long-term trend assessment. It was successfully applied to 15 cities, and it was based on emissions estimates, wind speed, and mixing height.

2. Data and methods

2.1. Data

Hourly CO concentrations from five winter seasons (October–February of 1990/1991, 1991/1992, 1992/1993, 1993/1994, and 1994/1995) for three central Phoenix air quality monitoring stations were obtained from the EPA's AIRS database (Table 1). These three stations have the highest historical CO levels in the Phoenix metropolitan area and they had 8-h average CO concentrations representing, respectively, the 16th highest value (10.1 ppm), 26th highest value (9.1 ppm), and 33rd highest value (8.8 ppm) reported by all US monitors for 1995 (US EPA, 1996). Overall frequency distributions and

Table 1
Site details for the three CO stations. All monitors are instrumental non-dispersive infrared

Site	Address	Latitude	Longitude	Surroundings
0016	3315 W. Indian School	33°29′40″N	112°07′48″W	Residential, in urban core
0019	3847 W. Earll	33°29′02″N	112°08′32″W	Residential, in urban core
3002	1845 E. Roosevelt	33°27′31″N	112°02′36″W	Residential, in urban core

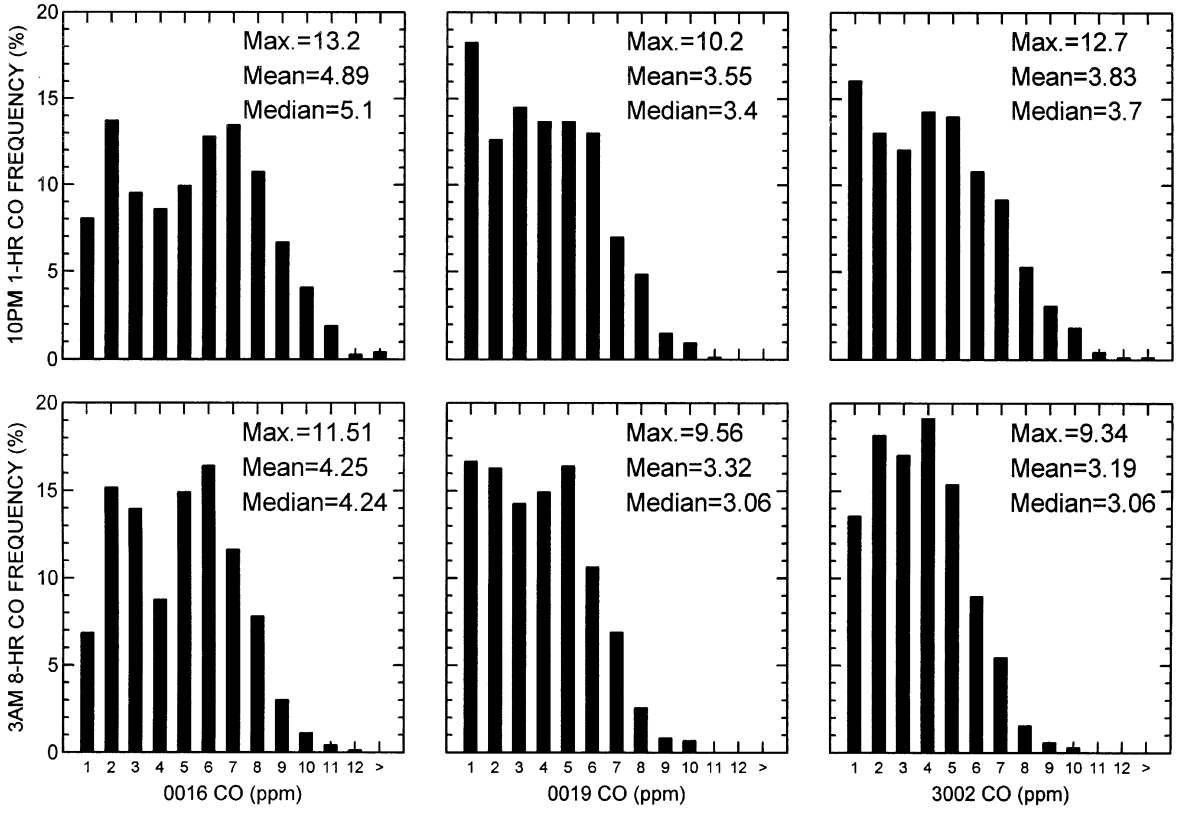

Fig. 1. Histograms summarizing CO concentrations at the three study sites, including maximum, mean, and median values (ppm). The 1-h and 8-h CO values are typically highest at 10 p.m. and 3 a.m. respectively.

descriptive statistics for 1-h and 8-h values at the three sites are presented in Fig. 1. The topography of central Phoenix is essentially flat, and all three sites experience broadly similar atmospheric conditions (see Fig. 2 for location map). While all three sites are in or adjacent to urban residential areas (Table 1), site 0016 is near a busy road, site 0019 is on a quieter residential street, and site 3002 is atop a low building in a neighborhood but in sight of a freeway. The specific siting details thus differ for each monitor, including proximity to vehicle traffic on busy roads, and microscale variability in meteorological factors due to nearby buildings and instrument exposure. The CO levels in Fig. 1 therefore result from a complex combination of local factors related to emissions and weather superimposed upon general regional similarities.

Hourly wind speed and wind direction values for the same stations over the same time period were acquired from AIRS. Hourly ΔT, an index of near-surface lapse rate expressed as the difference in temperature from 9.1–2.4 m (30–8 ft) a.g.l. measured at site 0019, was also extracted; however, these ΔT data were only available for the first three winter seasons. Hourly atmospheric pressure, temperature, dewpoint temperature, relative humidity, wind speed, and wind direction data for Phoenix Sky Harbor Airport (Fig. 2) were obtained from the National Climatic Data Center. A full multi-year set of hourly

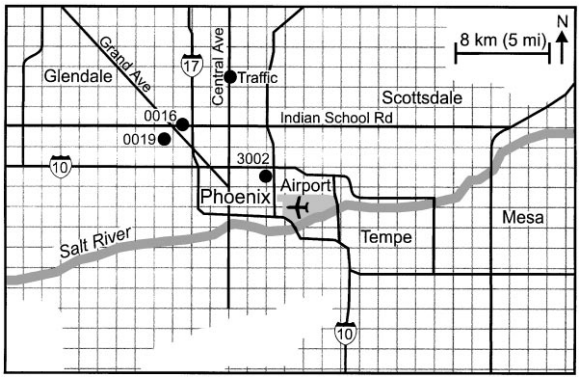

Fig. 2. Location map of central Phoenix, Arizona, showing CO monitor locations (0016, 0019, 3002), traffic count site, Sky Harbor International Airport, and selected highways and major roads.

traffic count data for central Phoenix was unavailable (real-time data were also unavailable), so the most complete and representative site was obtained from the Maricopa County Department of Transportation (1992) (data from Central at Montebello; see Fig. 2). Within the above constraints, missing data were no more than a few percent in each data set. Such cases were omitted rather than interpolated, and calculations (e.g., means) were adjusted accordingly.

2.2. Climatological methods

Initial exploratory analyses were performed to reveal general relationships in the Phoenix data. These included diurnal variations of all variables, and day-of-the-week differences in hourly CO concentrations and motor vehicle traffic counts. Frequencies of 8-h average CO concentrations exceeding 9 ppm were determined for each month, week, hour, and day of the week.

A compositing analysis was performed to discern the surface and near-surface synoptic scale atmospheric conditions associated with high and low 8-h average CO concentrations during winter mornings in central Phoenix. Diurnally, the highest 1-h CO value typically occurs at 10 p.m. (average CO from 9 to 10 p.m.); however, the end-hour for the highest 8-h mean CO (the value ultimately to be forecast) usually occurs around 3 a.m. the following morning because of the averaging procedure. Note that these and all other times are quoted in local standard time (L.S.T.). The following is a description of the compositing procedure (low CO parameters appear in parentheses):

1. The top 5% (bottom 10%) of all 8-h average CO concentrations for each station and month were identified. If a morning (3 a.m. $\pm$ 2 h; i.e., 1–5 a.m.) had a high (low) concentration at all three stations, then that day was deemed high (low) and extracted;
2. Means of 8-h average CO concentrations and hourly ΔT, atmospheric pressure, wind speed, relative humidity, and dewpoint temperature were determined for the 50 h preceding and 30 h following each high (low) CO morning, by month;
3. Mean 850 mb pressure–height synoptic maps were created for the day of the high (low) CO morning, the two days preceding high (low) CO mornings, and the day following the high (low) CO morning, by month and for the entire winter season.

2.3. Modeling methods

The final portion of the study encompassed the modeling of hourly and 8-h average CO concentrations, for the three years with available ΔT data. The modeling approach employed forward stepwise regression analysis ($\alpha < 0.05$) on variables identified in the climatological part of the study. Expanded terms (cross-products) were included to account for possible nonlinearities and the various interactions that might occur among the variables.

The objective of the modeling was to develop a suite of models to predict the 3 a.m. 8-h average CO concentration at each air quality monitoring station (the value of primary interest to the forecaster). The models utilized data available by 10 a.m. of the day in question to forecast evening CO. Eight regression equations were created for each of the eight individual hours included in the 3 a.m. 8-h average (i.e. 8 p.m.–3 a.m.). The eight predicted hourly values are averaged to produce a more robust predicted value for the 3 a.m. 8-h average CO concentration, and yet the forecaster can examine the eight individual hourly predictions for greater information. To increase the predictive power of the individual hourly CO models, a separate model was created to predict 11 p.m. ΔT values (because of the importance of the low-level nocturnal inversion). The resulting ΔT forecasts were included as possible predictors in a subset of hourly CO models. As a final step, two additional models were created for use under extreme high and low 8-h average CO concentrations, to increase the forecasting ability nearer the tails of the CO distribution. Models were developed on an 80% subset of the data, with the remaining (randomly selected) 20% retained for cross-validation and comparison using a suite of model evaluation statistics.

3. Results

3.1. Diurnal patterns

Motor vehicle exhaust is the primary source of CO in most urban areas. In Phoenix, over 98% of CO emissions are derived from mobile sources, the greatest proportion

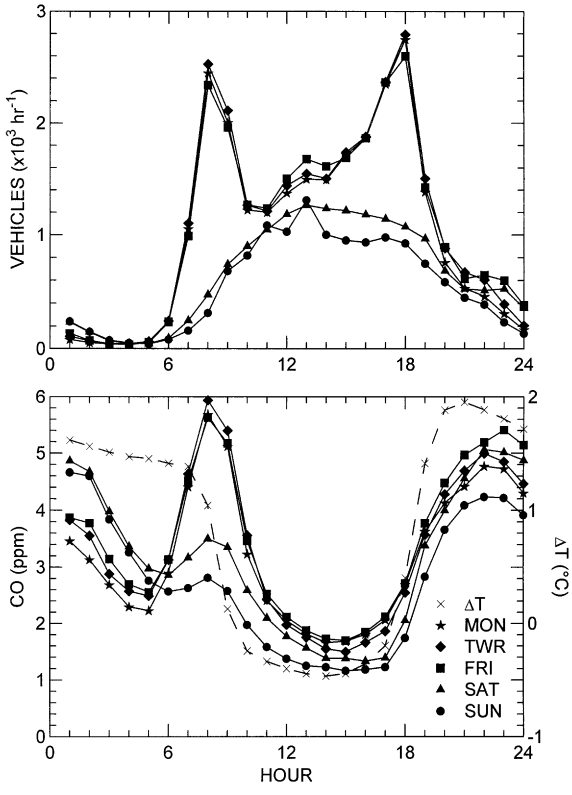

Fig. 3. Mean hourly winter traffic counts at an intersection (Central and Montebello) in central Phoenix (upper panel), and matching mean hourly ΔT (Station 0019) and CO concentrations at Station 0016 (lower panel), all stratified by day of the week. In the legend, TWR refers to the mean of Tuesday, Wednesday, and Thursday values.

of this from on-road vehicles (ADEQ, 1997). In the central part of the city, off-road, stationary, and area CO sources are minimal, and diurnal traffic patterns closely represent CO emissions. During a typical weekday, hourly traffic counts are highest during the rush hour periods around 8 a.m. and again around 6 p.m. (Fig. 3). On weekends, the highest traffic counts occur during early afternoon. A combination of these traffic-related emission patterns and meteorological influences ultimately controls ambient hourly CO concentrations in central Phoenix.

The dominant meteorological factor influencing CO concentrations in Phoenix is the nocturnal low-level temperature inversion, in particular, its timing and strength (vertical temperature gradient). The increased stability of air at the surface effectively traps motor vehicle emissions near the ground and causes increased ambient CO concentrations. Usually, low-level inversions are still present during the morning rush hour, while they are strongest ($\Delta T > 1.7^{\circ}$C) at 10 p.m. during most winter months (Fig. 3). This results in two slightly different ambient CO peaks each day: one that coincides with the morning rush hour, and another nighttime peak that occurs several hours after the evening rush hour. Because the inversion disappears within a few hours after sunrise, the morning CO peak has a relatively brief duration, as compared to the longer nocturnal peak when the inversion strengthens overnight and increases evening concentration of ambient CO (see Fig. 3). Thus, from the regulatory and forecasting standpoints, the 8-h average CO concentration is usually higher overnight (typically the 8 p.m.–3 a.m. average) than the period covering the morning rush hour.

3.2. Weekly patterns

With respect to days of the week, the midweek days (Tuesday, Wednesday, and Thursday) have the highest morning hourly CO values, but Friday has the highest evening hourly CO concentrations (occurring overnight and into the early hours on Saturdays) due to additional late-night leisure traffic occurring after evening commuter traffic (Fig. 3). Because of the averaging effect described above, Friday nights thus have the highest 8-h CO concentrations of the whole week. The weekend CO concentrations show the importance of the timing of traffic levels rather than sheer amount, because the traffic maximum occurs in the middle of the day when mixing is usually greatest. Instead, the morning and evening CO maxima (although at lower concentrations) remain synchronized with ΔT behavior. Thus, on weekends, diurnal ambient CO concentrations rise when the near-surface inversion coincides with moderate but declining traffic emissions in the early evening. Ambient CO levels decrease somewhat through the night, but rise again in the early morning as traffic increases slightly before the inversion dissipates. The somewhat higher leisure traffic on Saturday night results in a matching CO increase, as compared to Sunday night when traffic levels are lowest. Ground-level atmospheric stability, in concert with vehicle traffic, therefore plays a critical role in determining ambient CO concentrations.

3.3. Seasonal patterns

Over the course of a typical winter season, a temporal clustering of CO exceedance days occurs during the first three to four weeks of December (preceding the Christmas and New Year holiday period). There are several causes for this pattern, but it appears to be due mainly to earlier sunset time and the early formation time for the low-level inversion, which at this time of year shifts to approximately 5 p.m. near the start of the evening rush hour (Fig. 4). Thus, the CO emissions from almost the entire evening rush hour period accumulate in the stable air near the ground, leading to increased ambient CO concentrations. The other factors that influence CO at

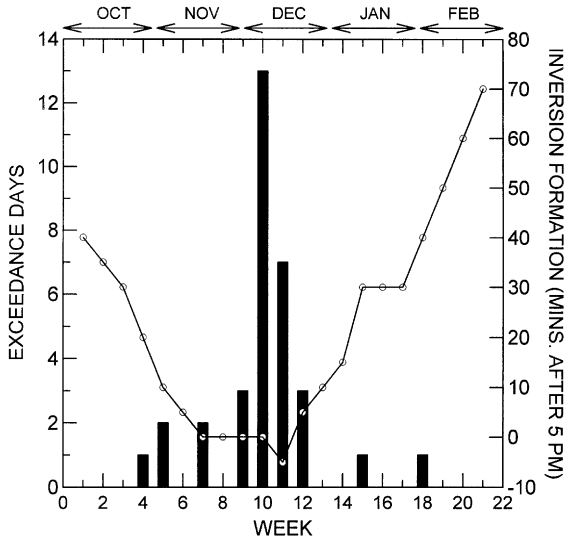

Fig. 4. Total days per week with at least one 8-h CO value at any station exceeding 9 ppm (bars) illustrating widespread occurrence in early December, and mean near-surface inversion formation times (line), for each week of winter (1–7 October is week 1, 8–14 October is week 2, etc.). The CO data are for five winters, and the inversion data are for three winters (see text).

this time of year include approximately 20% higher recorded traffic counts in these weeks (as compared to those weeks immediately before or after this period), and the possibility of greater CO emissions from domestic wood-burning in the evening hours (although there are regulations prohibiting this activity in many parts of Phoenix). Generally, then, the worst-case scenario for widespread, extreme high-CO events occurs on a Friday night in the first few weeks of December.

3.4. Compositing analyses

The exploratory analyses above highlight relationships between CO and various individual controlling factors. The compositing analyses identify the suite of atmospheric factors associated with different levels of CO. The highest overnight 8-h average CO concentrations (usually ending around 3 a.m.; see methods) are associated with a winter high-pressure system ridging to the east of the study area (Fig. 5), similar to the Type 3 synoptic pattern of Comrie (1996). This kind of circulation pattern results in a combination of strong low-level inversions, large diurnal temperature ranges, rising dewpoints, low wind speeds, and high pressure (Fig. 6). These are fairly

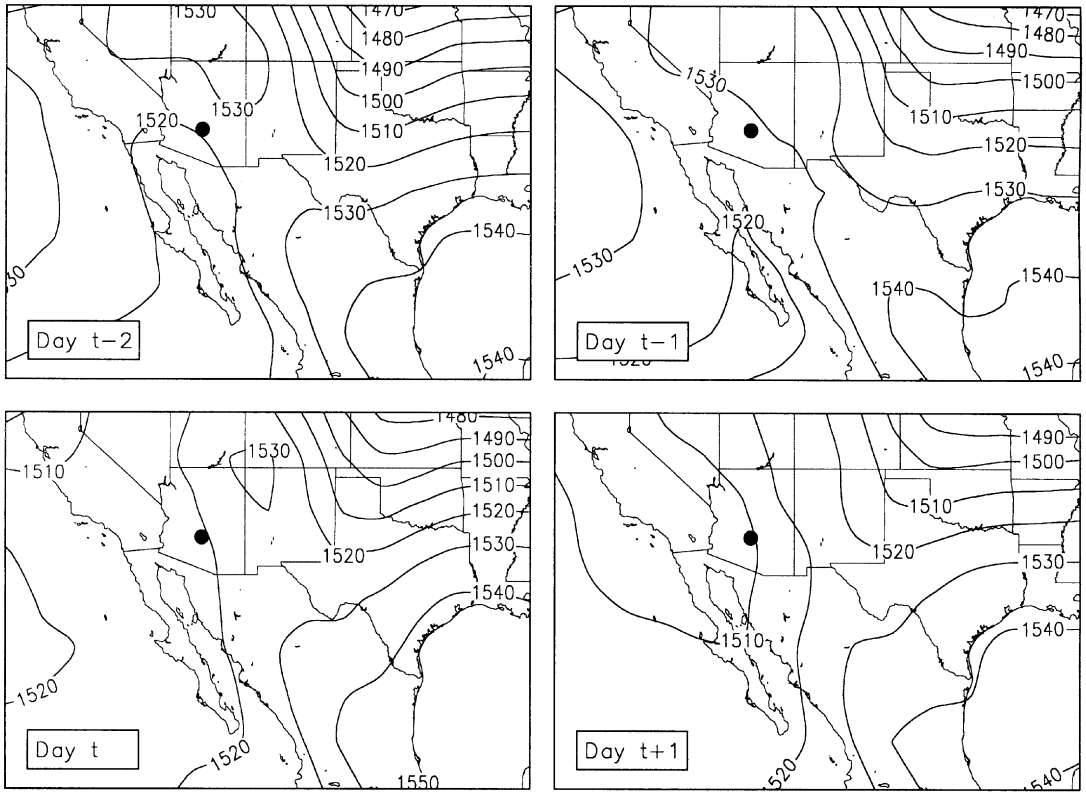

Fig. 5. Four-day sequence of composite 850 mb maps (gpm) for high-CO days (top 5% of 8-h values) in Phoenix (dot) during the study period (high CO occurs on Day t).

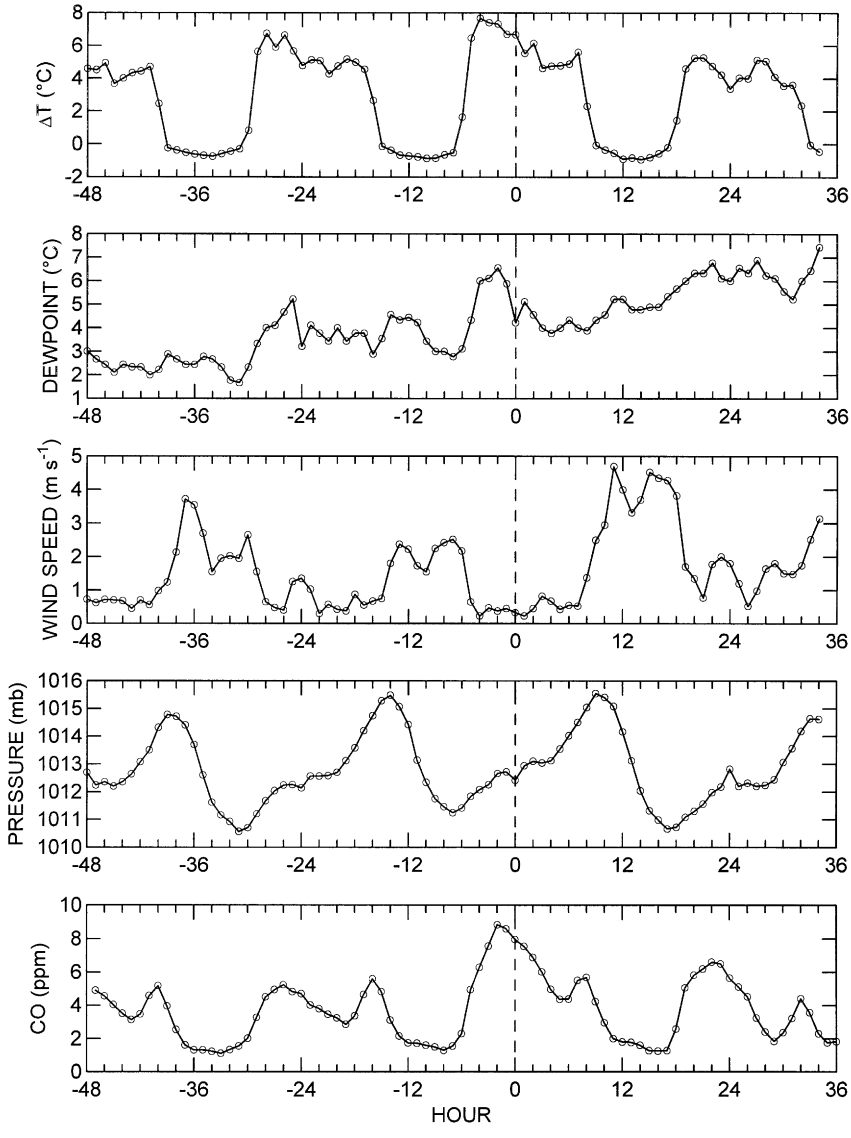

Fig. 6. October example of composite hourly surface atmospheric conditions (upper four panels) associated with highest 5% of 8-h CO concentrations at Station 0016, and the associated hourly CO (lower panel). The ΔT data are from Station 0019, the dewpoint, wind speed, and pressure data are from Sky Harbor Airport. Hour zero is the time of overnight 8-h CO maximum (see text).

classic low-ventilation conditions, and they quantify the average situation leading to extreme high CO, providing key behavior of variables to be included in the forecast models. Not surprisingly, low overnight CO concentrations are linked with weak inversions, small diurnal temperature ranges, strong winds, and a relatively moist atmosphere associated with a low-pressure trough (often containing a cold front) passing through Arizona.

3.5. Model development and variables

The CO models used variables suggested by the above analyses, with the constraint that the variables had to be routinely available to the forecaster (from local agency measurements, such as antecedent CO and ΔT, or from the National Weather Service at Sky Harbor Airport). Note that the models were developed for operational (i.e. forecasting) purposes, and not for inferential purposes, so

that inclusion of non-independent variables such as antecedent CO concentrations was permissible. The coefficient of determination (R^2) was used as a preliminary indicator of utility, but a range of model evaluation statistics was computed for the final set of cross-validated model results. All models were developed using only three seasons of data (1990/1991, 1991/1992, and 1992/1993) because the important ΔT variable was unavailable for the other two winters.

Potential model variables comprised several subgroups, the first being general variables. The CO Index reflects baseline "climatological" behavior of CO concentrations. It is based on a 5-winter composite of 10 p.m. CO data from Station 0016, averaged by week-of-the-season and then multiplied by a day-of-the-week factor (e.g., 10 p.m. on a Friday during the tenth week of the winter season). Time of inversion formation was averaged by week for the five winters as the second general variable, and the presence of the high-CO synoptic pressure pattern was the third general variable.

There were ten potential atmospheric variables, of which seven were observed data: 9 a.m. ΔT, 9 a.m. atmospheric pressure, 9 a.m. wind speed (at each of the three sites), previous 10 p.m. ΔT, and 5 a.m. ΔT. The other three atmospheric variables selected for potential inclusion were predicted values. Overnight temperature range (e.g., today's forecast high-temperature minus tomorrow morning's forecast low) and forecast 11 p.m. wind speed are both available from the National Weather Service. A predicted 11 p.m. ΔT value was used because of the importance of low-level stability for CO. This value was obtained from a separate regression model (cross-validated $R^2 = 0.67$) based on the observed 8 a.m. and forecast 11 p.m. meteorological data and antecedent ΔT (5 a.m., previous 10 p.m.).

Fourteen potential variables relating to observed CO were also used (comprising antecedent CO, dummy variables and interaction terms): 9 a.m. CO if on a weekday (all three stations), 5 a.m. CO (all three stations), previous 10 p.m. CO (all three stations), previous 10 p.m. CO if on a Sunday (all three stations), mid-week (Tuesday, Wednesday, or Thursday) dummy variable, and Friday dummy variable. Thus, several of these terms were effectively "used" only for particular days, pre-selecting for high or low CO concentrations. In addition, the full set of interaction terms (cross-products) was included with the above potential variables to maximize model flexibility. The modeling analysis indicated that squared terms were confounding, and they were omitted.

3.6. Selected model variables

The CO forecast models were divided into four categories, Models A, B, C, and D. Model A is the base 1-h model that does not use the predicted 11 p.m. ΔT value. Model B is an alternate form of Model A that includes the ΔT prediction as a potential input variable in the regressions. There were 24 A and B models, respectively (eight 1-h models for each of the three stations). Model C was used only if the predicted 3 a.m. 8-h average CO concentration from the A or B models was greater than a threshold, the upper 25% of values for a station. It uses cases from the upper tail of the distribution to produce a single "improved" 8-h average CO prediction (individual hours are not predicted). Model D is similar to Model C, except that it is developed on cases falling within the lowest 25% of CO concentrations at each station. Models ACD and BCD are thus combined two-stage models that substitute predictions from Models C and D for the Model A or B predictions, if the A or B results should fall in the upper or lower quartiles of the distribution. To recap, each site has two sets of eight 1-h predictions, each set averaged to an 8-h value; if that value falls in the upper or lower quartile, then another prediction is substituted using a model based on only the data for that quartile.

The forward stepwise regression procedure reduced the large list of potential variables down to between 4 and 12 of the most statistically significant covariates for the A and B models, and between 1 and 9 covariates for the C and D models. Final variables are summarized in Table 2, which includes the letter codes used to detail the individual models in Table 3 (A and B models) and Table 4 (C and D models). Many of the final model predictors are interaction terms that include pairs of individual variables appearing in different combinations and frequencies. Standardized coefficients (β) are provided as a guide to the importance of the terms in each model. Table 2 includes a summary of these values across all like models and sites via a weighted frequency. This measure is calculated as the sum of the absolute values, $\Sigma|\beta|$, expressed as a percent of the total. Both of the variables in interaction terms are counted, and variables occurring alone (not in combination) are double-counted to approximate an equivalent weighting. This weighted frequency is a rough and simple summary of the complex individual models, but it provides a sense of the overall utility of each variable. Thus, while all the final variables were sufficiently important for inclusion in the models, some were used repeatedly in many of the individual models. For the 24 A models, these most useful variables with the highest weighted frequencies include forecast 11 p.m. wind speed, forecast overnight temperature range, the CO index, inversion formation time, and previous 10 p.m. CO. The 24 B models made most use of the 11 p.m. ΔT prediction and forecast 11 p.m. wind speed, along with the terms listed for the A models. The most useful variables in the three C models included 5 a.m. CO, predicted 11 p.m. ΔT, the CO index, and previous 10 p.m. ΔT. In the D models, forecast 11 p.m. wind speed and 9 a.m. ΔT were the most useful variables, followed by several of the others mentioned above. To generalize, the

Table 2
Variable letters and names used in Tables 3 and 4, with weighted percent frequencies of use in CO models (singly or within a cross-product) across all sites for each model type. See text for details

Letter	Weighted frequency (%)				Variable name
	A	B	C	D	
A	0.87	1.26	2.28	18.45	9 a.m. ΔT
B	5.70	1.36	2.90	0.00	9 a.m. Pressure
C	14.33	8.16	6.75	6.60	Forecast overnight temperature range
D	4.85	5.41	3.13	2.79	Synoptic Type 3
E	2.32	1.79	3.19	5.07	9 a.m. CO (if weekday)
F	2.56	4.67	2.61	5.53	9 a.m. Wind speed
G	3.23	4.05	9.85	8.80	Previous 10 p.m. ΔT
H	6.43	2.60	4.01	3.24	5 a.m. ΔT
I	12.32	8.54	12.61	0.75	CO index
J	6.04	7.69	23.40	0.00	5 a.m. CO
K	18.62	16.32	0.00	28.45	Forecast 11 p.m. wind speed
L	9.45	7.15	2.90	5.94	Inversion formation time
M	9.03	8.08	4.83	6.79	Previous 10 p.m. CO
N	0.18	1.09	0.00	0.00	Previous 10 p.m. CO (if Sunday)
O	0.00	17.01	15.70	7.59	Modeled 11 p.m. ΔT
P	4.07	4.81	5.83	0.00	9 a.m. CO (if Sunday)

Table 3
Variables ranked by standardized coefficients for Type-A and B CO models, by site. Letter codes are referenced in Table 2. Dual letters indicate an interaction term (e.g., AB is variable A multiplied by variable B)

20:00h		21:00h		22:00h		23:00h		00:00h		01:00h		02:00h		03:00h	
0016A															
HI	0.237	K	−0.302	K	−0.366	K	−0.406	K	−0.410	K	−0.423	K	−0.383	CJ	0.249
B	0.214	GH	0.300	CI	0.170	GH	0.203	EI	0.297	DJ	0.249	CI	0.224	K	−0.240
CG	0.213	CE	0.170	HI	0.169	CI	0.193	GH	0.277	CI	0.216	CJ	0.194	CI	0.198
I	0.165	BI	0.120	FI	−0.142	DJ	0.170	EH	−0.261	GH	0.159	CD	0.168	JK	−0.166
HK	−0.157	DG	0.116	CE	0.134	EI	0.097	DJ	0.208	FI	−0.114	CF	−0.123	B	0.161
FK	−0.139	F	−0.112	FJ	0.125	FI	−0.057	CI	0.175	AD	−0.098			DH	0.140
EI	0.113	B	0.108	DG	0.125			AD	−0.086					CF	−0.076
DG	0.086	DK	0.104	DK	0.052			FI	−0.085						
0019A															
IM	0.290	IM	0.261	KL	−0.274	KL	−0.367	KL	−0.401	KL	−0.416	KL	−0.392	CH	0.342
KL	−0.248	KL	−0.232	IM	0.260	IM	0.271	IM	0.266	IM	0.226	IM	0.203	KL	−0.223
CI	0.217	CI	0.204	CI	0.232	CI	0.203	CI	0.212	CH	0.161	CE	0.164	HK	−0.212
EI	0.201	B	0.186	B	0.180	CD	0.142	CD	0.113	CD	0.128	BC	0.139	CD	0.175
BL	−0.197	CE	0.082	DI	0.096							CD	0.124	CJ	0.167
I	−0.166	DM	0.067											B	0.135
B	0.155													FN	−0.120
DM	0.116														
FJ	−0.094														
3002A															
CM	0.552	CM	0.573	BM	0.313	CM	0.594	PI	0.192	CJ	0.286	CJ	0.373	CJ	0.460
CH	0.337	KM	−0.361	KL	−0.305	KL	−0.427	KL	−0.415	KL	−0.464	KL	−0.336	PI	0.396
KM	−0.278	HI	0.221	CI	0.223	IM	−0.360	PK	−0.183	PM	0.230	PI	0.271	PK	−0.366
L	−0.262	L	−0.167	KM	−0.207	AG	−0.263	CJ	0.168	AJ	−0.184	PK	−0.133	J	−0.205
HM	−0.236	B	0.149	CH	0.198	PL	0.252	CH	0.164	FG	0.081	J	−0.098	B	0.195

(*continued overleaf*)

Table 3 (*continued*)

20:00h		21:00h		22:00h		23:00h		00:00h		01:00h		02:00h		03:00h	
3002A															
PG	−0.149	GM	−0.094	DF	0.147	L	−0.196	CM	0.162			PH	−0.072	PH	−0.071
DF	0.096	DJ	0.093			PK	−0.184	DK	0.108						
						DP	0.174	AJ	−0.037						
						AH	0.166								
						HJ	0.131								
						FG	0.053								
0016B															
O	0.395	O	0.380	K	−0.360	K	−0.420	K	−0.380	IK	−0.380	K	−0.391	O	0.442
BI	0.277	EI	0.332	EI	0.210	CI	0.231	DJ	0.237	DJ	0.231	CI	0.274	IJ	0.266
FI	−0.238	CK	−0.215	GH	0.198	DJ	0.206	O	0.225	O	0.217	DJ	0.258	DH	0.195
EF	0.180	EO	−0.179	CI	0.193	GI	0.173	GO	0.157	BI	0.209	CF	−0.184	JK	−0.176
CK	−0.165	DG	0.177	FI	−0.133	FI	−0.161	I	0.138	GO	0.150	GO	0.159	FH	−0.143
DG	0.131	IO	0.130	DG	0.120	EF	0.100	EJ	0.132	AO	−0.118	AO	−0.107	K	−0.119
		F	−0.114	DK	0.106	AO	−0.086	AD	−0.118	FI	−0.101			AJ	−0.104
				AO	−0.100			EO	−0.113					KO	−0.090
				FJ	0.097			FI	−0.083						
0019B															
O	0.332	O	0.260	O	0.456	O	0.346	O	0.436	O	0.419	O	0.374	O	0.427
FM	0.230	IK	−0.226	JM	0.299	KL	−0.298	KL	−0.240	KL	−0.274	KL	−0.333	CK	−0.297
L	−0.206	CM	0.220	GK	−0.251	DM	0.176	DI	0.176	KO	−0.137	KM	0.128	IJ	0.154
IK	−0.200	L	−0.188	DM	0.162	JM	0.175	KM	0.154	DI	0.133	CD	0.126	CD	0.132
DM	0.120	FM	0.186	FN	−0.098	FM	−0.098	KO	−0.127	MO	0.129	GN	−0.125	GN	−0.101
MO	0.114	GM	−0.183					FN	−0.108	GN	−0.107				
FG	−0.087	FO	0.123												
		DM	0.102												
3002B															
CM	0.415	CM	0.474	KL	−0.453	CM	0.361	KL	−0.352	KL	−0.464	CJ	0.378	CJ	0.444
CH	0.349	KM	−0.374	CI	0.280	KL	−0.297	CM	0.215	CJ	0.286	KL	−0.301	PI	0.408
KM	−0.333	HI	0.184	HM	0.225	PL	0.222	PI	0.189	PM	0.230	PI	0.273	PK	−0.408
L	−0.203	L	−0.142	DF	0.137	KM	−0.206	PK	−0.169	AJ	−0.184	PM	0.209	B	0.230
PG	0.191	PJ	0.141	PJ	0.100	DP	0.164	CJ	0.165	FG	0.081	PK	−0.203	J	−0.212
GH	−0.145	CO	0.127	FM	0.055	HJ	0.160	CH	0.154			J	−0.159	JN	−0.120
DF	0.133	DF	0.110	IJ	0.050	PK	−0.146	DK	0.105			PG	−0.137	KO	0.066
JK	0.115	JK	0.038	AJ	0.012	L	−0.135	KM	−0.097			DO	−0.126	PH	−0.060
PK	−0.098					AG	−0.117	AJ	−0.032			FJ	−0.053		
DO	−0.085					FG	0.026								
FM	0.056														
AF	0.021														

most frequent and useful predictors include anticipated nocturnal stability (wind speed, temperature range, ΔT, inversion formation time) and antecedent or climatological CO behavior (previous 10 p.m. CO, 5 a.m. CO, CO index).

3.7. CO model evaluation

Model evaluation statistics and prediction skills were calculated using the data reserved for cross-validation. The evaluation statistics compare observed versus predicted average ambient CO for the 8-h period ending at 3 a.m. (Table 5). Recall that final forecast CO values in these model evaluation statistics are complexly determined from numerous sub-model predictions. For example, in Model ACD a final forecast CO value is the mean of eight individual Model A hourly predictions if it falls between the 25th and 75th percentiles, but it is the result of an individual 8-h prediction if it is higher or lower than these bounds.

Across the three sites, CO forecasts using the simpler A and B models are only moderately accurate in terms of the coefficient of determination (R^2) or variance explained (Table 5). However, the R^2 values improve from

Table 4
Type-C and Type-D CO models by site. Details as for Table 3

0016C		0016D		0019C		0019D		3002C		3002D	
I	0.360	AG	− 0.522	JM	0.411	K	− 0.583	JO	− 0.517	KL	− 0.571
GJ	0.342	AE	0.487	BL	− 0.247	AH	0.312	CJ	0.500		
EO	0.272	K	− 0.406	CH	0.075	CD	0.268	PG	0.497		
AO	− 0.194	M	0.327			AF	− 0.207	IO	0.354		
		FG	− 0.325					DH	− 0.267		
		CO	0.295					FJ	0.222		
		AO	− 0.246								
		KO	− 0.189								
		CI	− 0.072								

Table 5
Model evaluation statistics for the final CO models for the three stations (units in ppm, where applicable). See text for abbreviations

Model	R^2	$\bar{P}$	$\bar{O}$	MBE	s_P	s_O	RMSE	MAE	D_1	PSE	U_{80}	O_{80}	U_{90}	O_{90}
0016														
A	0.52	3.99	3.99	0.00	1.55	2.22	1.54	1.25	0.60	0.51	0.82	0.02	1.00	0.01
B	0.59	3.97	4.09	− 0.12	1.50	2.27	1.48	1.18	0.62	0.58	0.80	0.02	1.00	0.01
ACD	0.78	4.14	4.07	0.07	2.19	2.31	1.02	0.76	0.78	0.12	0.21	0.06	0.22	0.03
BCD	0.88	4.33	4.32	0.01	2.25	2.33	0.83	0.64	0.84	0.07	0.05	0.06	0.27	0.04
0019														
A	0.51	3.24	3.36	− 0.12	1.37	2.01	1.40	1.12	0.75	0.54	0.75	0.04	0.87	0.00
B	0.66	2.91	3.13	− 0.22	1.30	2.01	1.24	1.01	0.65	0.63	0.64	0.00	1.00	0.00
ACD	0.79	3.50	3.36	0.14	1.86	2.01	0.94	0.69	0.79	0.17	0.18	0.07	0.60	0.01
BCD	0.88	3.56	3.64	− 0.08	1.99	2.15	0.75	0.58	0.84	0.15	0.10	0.11	0.46	0.02
3002														
A	0.53	3.12	3.11	0.01	1.22	1.84	1.57	0.94	0.62	0.56	0.60	0.02	1.00	0.01
B	0.56	3.04	3.01	0.03	1.21	1.77	1.17	0.88	0.63	0.54	0.54	0.02	1.00	0.01
ACD	0.81	2.90	2.91	− 0.01	1.67	1.86	0.81	0.57	0.81	0.19	0.15	0.04	0.25	0.03
BCD	0.86	2.82	2.77	0.05	1.69	1.79	0.68	0.50	0.83	0.11	0.15	0.04	0.25	0.03

about 0.5 and 0.6 for Models A and B to about 0.8 and 0.9 for the two-stage ACD and BCD models. The beneficial effects of including the C and D sub-models are clear in Fig. 7, compressing scatter for lower and upper CO values. For these combined ACD and BCD models, intercept and slope of the best-fit line differ only slightly across models (Fig. 7), with small intercepts and slopes close to 1.0 indicating slight underprediction of high values and slight overprediction of low values (regression tends to bias toward the mean). The mean predicted and observed values ($\bar{P}$ and $\bar{O}$) and their difference, the mean bias error (MBE), show close correspondence between models and data, with average under- and over-predictions between − 0.22 and 0.14 ppm. Standard deviations of predicted values (s_P) and are only slightly less than those for the observed data (s_O) for the ACD and BCD models. The root mean-squared error (RMSE) across the three stations ranges from about 0.68 to 1.02 ppm for these models and between 1.17 and 1.57 ppm for the A and B models. The matching ranges for mean absolute error (MAE, the mean of the unsigned errors) are 0.50 to 0.76 ppm for Models ACD and BCD, and 0.88 to 1.25 ppm for the A and B models. Thus, for the better models, predicted CO concentrations are typically within 1 ppm of the observed concentrations. We also report the index of agreement (D_1), a dimensionless measure of the degree to which a model's predictions are error-free (Willmott, 1981; Comrie, 1997). It ranges from 0.0 (complete disagreement between predicted and observed values) to 1.0 (perfect agreement between predicted and observed values). For the two combined models D_1 ranges between 0.78 and 0.84 versus 0.60 to 0.75 for the simpler models. The proportion of systematic error (PSE) is calculated as MSE_s/MSE, which measures the proportion of overall model mean-squared error that is systematic (Willmott, 1981; Comrie, 1997). The

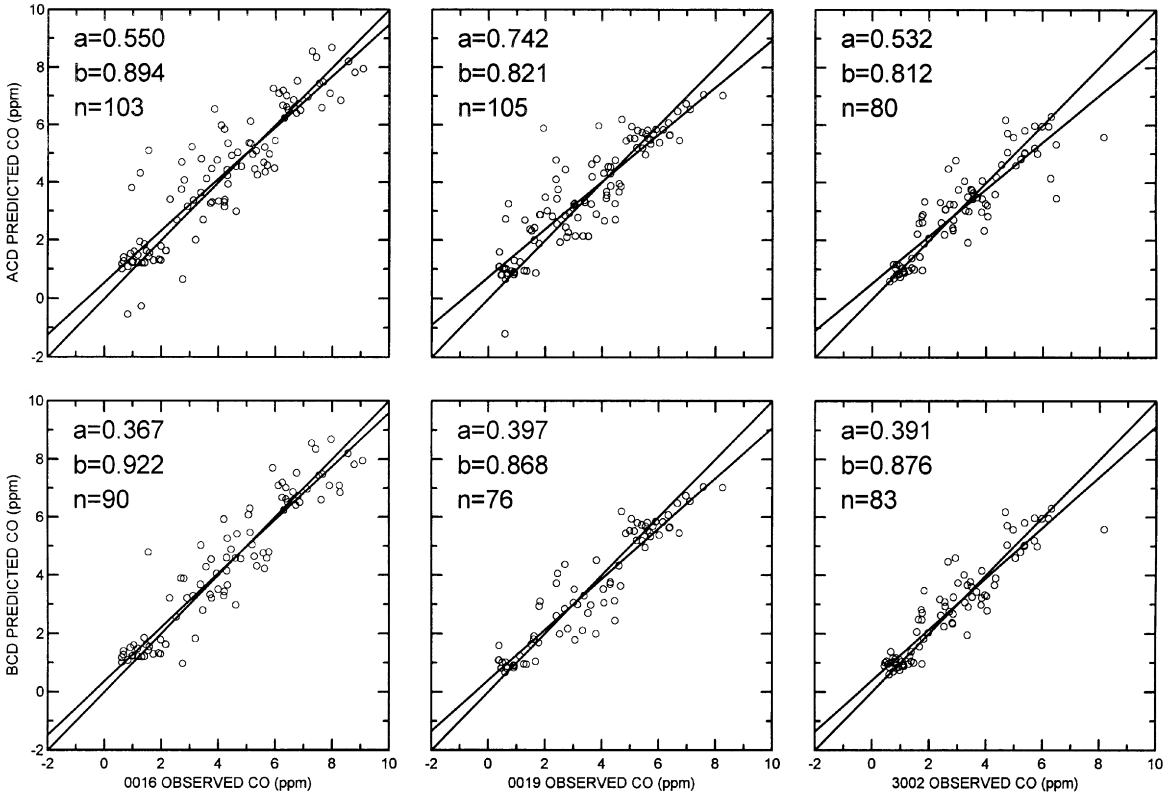

Fig. 7. Scatterplots of observed versus predicted 3 a.m. 8-h CO from models ACD and BCD using reserved cross-validation data for the three sites. The two lines on each panel are the perfect prediction line where predicted values equal observed values, and the best-fit regression line with intercept (a) and slope (b) indicated in the upper left corner.

proportion of unsystematic error is simply 1-(MSE_s/MSE). Systematic error is model-derived while unsystematic error represents the natural variability of the data (stochasticity or "noise") that cannot be reduced by a model. The PSE is between 7 and 19% for the ACD and BCD models, but it is much higher in the A and B models at 51–63%.

These models were developed for operational forecasting, in particular to warn of infrequent high CO concentrations exceeding a threshold at the extreme upper tail of the distribution. This type of model skill is evaluated analogously to Type I and II errors: true or false positive (correctly or incorrectly predicting above the threshold) and true or false negative (correctly or incorrectly predicting below the threshold). Almost none of the cross-validation data were greater than 9 ppm, so we evaluated skill at the 80th and 90th percentiles for each station in order to have data in each of these four categories (80th percentiles were 6.24, 5.08, and 4.84 ppm; 90th percentiles were 7.31, 5.97, and 5.84 ppm, at sites 0016, 0019, and 3002, respectively). These under- and over-prediction results are shown in Table 5, where the abbreviation U_{80} represents the proportion of predicted values under the 80th percentile when the observed value was above it, and O_{80} represents the proportion of predicted values over the 80th percentile when the observed value was below it, and similarly for U_{90} and O_{90} at the 90th percentile.

The forecast skill measures show similar findings to the other statistics. The ACD and BCD models estimate between 5 and 21% of values over the 80th percentile as being under that threshold, whereas these underestimates are between 54 and 82% for the A and B models. At the 90th percentile, the ACD and BCD models estimate 22–27% of values over this threshold as being under it at sites 0016 and 3002, but they underestimate about half of these values at site 0019, while the A and B models underestimate 87–100% of 90th percentile values across all three sites. The reverse phenomenon of predicting a value actually below the threshold as above it (a false alarm for the forecaster) is uncommon across all models at both thresholds, ranging between 1 and 11%.

Together, all these comparisons in Table 5 clearly reveal the compound ACD and BCD models as superior to the conventional regressions in the A and B models. Furthermore, in almost every comparison the BCD

models that include the predicted ΔT variable are marginally better than the ACD models. Examining differences across the three sites, model performance is generally slightly better at site 3002, moderate at site 0019, and slightly worse at site 0016. This ranking is consistent with the means and standard deviations of CO concentrations across the sites, which are low, medium and high, respectively.

4. Concluding discussion

The first part of this study highlights atmospheric conditions and traffic patterns leading to variations in CO concentrations, with an emphasis on situations leading to high CO concentrations. These analyses provided a specific understanding of ambient CO behavior in Phoenix, and guided the selection of variables available for inclusion in CO forecast models for the second part of the study. We developed a suite of multivariate regression models based on variables easily available to the forecaster 12 h prior to the expected high-CO conditions, which enabled accurate forecasts of overnight CO concentrations to be made by 10 a.m. of the day in question. The resulting models included variables and interaction terms related to anticipated nocturnal stability as well as antecedent and climatological CO behavior. Of the various model types considered, a two-stage sequence combining models based on the entire data distribution with models focusing on the tails of the distribution produced the most accurate predictions of 3 a.m. 8-h average CO concentrations. The best of these combined models had R^2 values approaching 0.9, with mean errors under 1.0 ppm and very low proportions of systematic error. Their forecast errors based on skill thresholds were between 5 and 15% at the 80th percentile and 25 and 46% at the 90th percentile. Overall these represent good results in comparison to other air quality forecasting studies.

For operational implementation, the set of models was coded to run on a personal computer, providing easy data entry, the ability for multiple sensitivity runs, and for saving the input and forecast data. The software calculates results for all the final models presented above, so that the human forecaster can make an ensemble forecast utilizing the full range of predictions plus whatever ancillary information might be available. Models A and B each provide a set of eight 1-h values (8 p.m.–3 a.m., each based on a separate regression) plus the 8-h mean, so that CO predictions for individual hours rise and fall differently in each model depending on conditions. Yet, the 8-h means provided more robust CO values (in the same units and averaging time as the NAAQS), but with additional and more accurate 8-h predictions from Models C and D. By knowing the strengths and weaknesses of each model, and by making multiple model runs with slightly different input conditions, the forecaster can make an informed judgement as to the likelihood of a high CO concentration. Finally, although ambient CO does not show any strong secular trends in the last few years, we recommend updating the models regularly as more recent data become available.

Acknowledgements

This work was funded by the Arizona Department of Environmental Quality, Office of Air Quality. Sandra Brazel (Arizona State University, Office of Climatology) assisted in obtaining the Sky Harbor data. Many thanks to Brian Eder (EPA/NOAA), Bill Ryan (University of Maryland), and the anonymous reviewers for their useful comments on earlier drafts of the manuscript.

References

Arizona Department of Environmental Quality (ADEQ), 1997. Metropolitan Phoenix Voluntary Early Ozone Plan. ADEQ, Technical Publication, Phoenix.

Aron, R.H., Aron, I., 1978. Statistical forecasting models: I. Carbon monoxide concentrations in the Los Angeles Basin. Journal of the Air Pollution Control Association 28, 681–684.

Bardeschi, A., Colucci, A., Gianelle, V., Gnagnetti, M., Tamponi, M., Tebaldi, G., 1991. Analysis of the impact on air quality of motor vehicle traffic in the Milan urban area. Atmospheric Environment 25B, 415–428.

Colucci, J.M., Begeman, C.R., 1969. Carbon monoxide in Detroit, New York, and Los Angeles air. Environmental Science and Technology 3, 41–47.

Comrie, A.C., 1996. An all-season synoptic climatology of air pollution in the US–Mexico border region. Professional Geographer 48 (3), 237–251.

Comrie, A.C., 1997. Comparing neural networks and regression models for ozone forecasting. Journal of the Air and Waste Management Association 47, 653–663.

Glen, W.G., Zelenka, M.P., Graham, R.C., 1996. Relating meteorological variables and trends in motor vehicle emissions to monthly urban carbon monoxide concentrations. Atmospheric Environment 30, 4225–4232.

Holzworth, G.C., 1962. A study of air pollution potential for the western United States. Journal of Applied Meteorology 1, 366–382.

Jakeman, A.J., Bai, J., Miles, G.H., 1991. Prediction of seasonal extremes of one-hour average urban CO concentrations. Atmospheric Environment 25B, 219–229.

McCollister, G.M., Wilson, K.R., 1975. Linear stochastic models for forecasting daily maxima and hourly concentrations of air pollutants. Atmospheric Environment 9, 417–423.

Miles, G.H., Jakeman, A.J., Bai, J., 1991. A method for predicting the frequency distribution of air pollution from vehicle traffic, basic meteorology, and historical concentrations to assist urban planning. Environment International 17, 575–580.

Newman, P., Kenworthy, J., 1989. Cities and Automobile Dependence: An International Sourcebook. Gower Technical, Sydney.

Nkemdirim, L.C., 1988. An assessment of the relationship between functional groups of weather elements and atmospheric pollution in Calgary, Canada. Atmospheric Environment 22, 2287–2296.

Norbeck, J.M., Chang, T.Y., Weinstock, B., 1979. Effect of New York City taxi strike on CO concentrations in midtown Manhattan. Journal of the Air Pollution Control Association 29, 845–847.

Robinson, E., Boyle, R.J., 1979. Synoptic weather typing and its application to air quality in the St. Louis, Missouri area. In: Proceedings of the 72nd Annual Meeting of the Air Pollution Control Association, Cincinnati, OH.

Tiao, G.C., Box, G.E.P., Hamming, W.J., 1975. A statistical analysis of the Los Angeles ambient carbon monoxide data 1955–1972. Journal of the Air Pollution Control Association 25, 1129–1136.

Tiao, G.C., Liu, L.-M., Hudak, G.B., 1989. A statistical assessment of the effect of the car inspection/maintenance program on ambient CO air quality in Phoenix, Arizona. Environmental Science Technology 23, 806–814.

United States Environmental Protection Agency (US EPA), 1996. AQS — USA Max Value Report — CO. ⟨http://www.epa.gov/airs/aqsd-co.html⟩.

Willmott, C.J., 1981. On the validation of models. Physical Geography 2, 184–194.