Nonlinear relationships between atmospheric aerosol and its gaseous precursors: Analysis of long-term air quality monitoring data by means of neural networks

Igor B. Konovalov

[1] {Institute of Applied Physics, Russian Academy of Sciences, Nizhniy Novgorod, Russia}

Abstract
The nonlinear features of the relationships between concentrations of aerosol and volatile organic compounds (VOC) and oxides of nitrogen (NOx) in urban environments are derived directly from data of long-term routine measurements of NOx, VOC, and total suspended particulate matter (PM). The main idea of the method used for the analysis is creation of special empirical models based on artificial neural networks. These models which are in essence the nonlinear extension of commonly used linear statistical models are believed to provide the best fit for the real (nonlinear) PM-NOx-VOC relationships under different atmospheric conditions. It is believed that such models may be useful in context of various scientific and practical problems concerning atmospheric aerosols. The method is demonstrated by the example of two empirical models created with independent data-sets collected at two air quality monitoring stations at South Coast Air Basin, California. It is shown that in spite of considerable distance between the monitoring stations (more than 50 km) and thus substantially different environmental conditions, the empirical models manifest several common qualitative features. Specifically, it is found that, under definite conditions, the decrease of the level of NOx or VOC may lead to the increase of mass concentration of aerosol. It is argued that these features are caused by the nonlinear dependence of hydroxyl radical on VOC and NOx.

1 Introduction
Atmospheric aerosol is an important environmental factor having impact on human health, affecting visibility, and playing a significant role in atmospheric chemistry and climate. The mass concentration of aerosols about the surface is regulated by national air quality standards, implementation of which requires a reliable scientific assessment of possible strategies for its reduction. Such an assessment, however, represents an extremely difficult problem because the evolution of aerosol is driven by complex physical and chemical processes which, in turn, are influenced by a number of environmental factors.

In particular, significant components of typical aerosol in the boundary layer such as sulfate, nitrate, and organic compounds may reach particulate phase as a result of photochemical processes involving certain gaseous species (aerosol precursors) such as NOx, SO2, NH3, and volatile organic compounds (VOC) (e.g., Senfeld and Pandis, 1998). Specifically, the important potential source of the airborne particulate matter (PM) is formation of sulfuric acid in a gas-phase chain mechanism which can be represented by the combined reaction (Stockwell et al., 1990)

\[ \text{SO}_2 + \text{OH} + (+\text{M}, \text{O}_2, \text{H}_2\text{O}) \rightarrow \text{H}_2\text{SO}_4 + \text{HO}_2. \]  \hfill (1)

Another potential source of atmospheric PM is the reversible reaction of ammonia with nitric acid

\[ \text{NH}_3(g) + \text{HNO}_3(g) \leftrightarrow \text{NH}_4\text{NO}_3(p), \]  \hfill (2)
where (g) and (p) denote gas and particulate phases, respectively. The HNO$_3$, in turn, is formed in the gas-phase reaction

$$\text{OH} + \text{NO}_2 \rightarrow \text{HNO}_3.$$  \hspace{1cm} (3)

Condensable compounds may be also formed in reactions of certain VOCs with OH, O$_3$, or NO$_3$ (Odum et al., 1996; 1997).

It is obvious from the aforesaid that formation of PM is driven to a significant extent by the hydroxyl radical (OH), the behavior of which is known to be determined by a large number of chemical processes (see, e.g., Stockwell et al., 1997). It is clear also that due to the crucial role of OH in processes of ozone formation, the gas-to-particle conversion processes are closely linked with complex photochemistry of ozone (e.g., Meng et al. 1997).

During two last decades the significant efforts have been devoted to modeling evolution of aerosol from the first principles; that is, the attempts have been made to describe explicitly all important processes affecting formation and removal of air-borne particles (see, e.g., Pilinis and Seinfeld, 1988; Wexler et al., 1994; Binkowski and Shankar, 1995; Jacobson, 1997; Meng et al., 1998; Ackermann et al., 1998; Sun and Wexler, 1998; Nenes et al., 1999). Nonetheless, in spite of the great progress in understanding and representation of aerosol evolution, the comparisons of model results with measurements bear evidence that models still contain significant uncertainties. Moreover, it still remains unknown to which extent the chemical-transport (CT) models are accurate in predictions of responses of aerosol to changes of emissions of potential PM precursors.

Taking into account close chemical coupling between aerosol and ozone, it is useful to note that evaluation of adequacy of CT model predictions of effectiveness of various emission reduction scenarios represents a rather difficult problem in researches of surface ozone as well (e.g., Sillman et al., 1999). In ozone studies, that problem arises, principally, due to the essential non-linearity of dependence of ozone production on concentrations of its precursors (National Research Council (NRC), 1991). In view of this problem, several methods for assessing ozone sensitivity to changes of its precursors have been developed recently (see, e.g., Kleinman et al., 2000) which are based predominantly on the use of data of ambient measurements. On the contrary, similar methods for aerosol that have not yet been proposed. Moreover, though some nonlinear features of PM-NO$_x$-VOC relationships have been discussed (e.g., Meng et al., 1997), it still remains unclear to which extent these features are general.

Thus, at present, there exists the definite need in better understanding of relationships between aerosol and its gaseous precursors, as well as the need in developing additional methods for their identification in the real environment. This paper presents results of the first attempt to study nonlinear features of PM-NO$_x$-VOC relationships directly on the basis of data of ambient measurements. The analysis is conducted by means of the novel original method, the essence of which is creation of empirical nonlinear statistical models based on artificial neural networks. Similar method have been successfully applied recently to studying nonlinear relationships between surface ozone and its precursors (Konovalov, 2002). Though neural networks are widely used in atmospheric researches (see, e.g., the overview by Gardner and Dorling, 1998), they are commonly considered as a "black box" which bears little understandable information of scientific or practical value. However, as it was shown in our previous work (Konovalov, 2002), the possibility of obtaining useful knowledge about physical or chemical properties of a natural system exists if the neural network is trained and analyzed properly.

In this study, neural networks are applied for analysis of data of routine long-term measurements (1980-1995 years) of total suspended particulate (TSP) in two air quality monitoring stations at
California. From the point of view of practical usefulness the analysis of fine fraction of aerosol (<2.5 µm) could be of most interest (taking into account US and European standards for PM$_{2.5}$). Unfortunately, at present time the available records of PM$_{2.5}$ measurements accompanied with simultaneous measurements of NO$_x$ and VOC are not long enough for building neural network models aimed at retrieval of PM-NO$_x$-VOC relationships. Nonetheless, we believe that features of these relationships discussed in this paper are fairly robust and applicable to all fractions of PM. It is hoped also that the method suggested in this paper may be successfully applied to fine fraction of aerosol in not very far future.

2 Data

This study utilizes data-sets of measurements of TSP, O$_3$, NO$_x$, and total volatile organic compounds (this quantity is referred to below as VOC) made in two air quality monitoring stations at South Coast Air Basin, California during 1980-1995 (California Air Quality Data, 1998). The stations were selected a priori based on announced spatial scale of observations (only stations with the urban scale (4-50 km) of ozone and NO$_x$ observations were considered) and availability of synchronous measurements of all components noted above. The urban scale means that both distribution of sources of pollutants and topographical situation in the neighborhoods of the station are such that the appearance of strong spatial gradients of the corresponding pollutants is unlikely. The only stations satisfying to these criteria turned out to be Riverside-Rubidoux (33°58'55"N, 117°22'21"W), Azusa (34°08'10"N, 117°55'26"W) and Reseda (34°11'57"N, 118°31'58"W). However, since a series of TSP measurements at Reseda is significantly shorter than those at the other two stations (at Reseda, the measurements were conducted only in 1980-1986), the data-set from this station was not analyzed. Note that data series used in this analysis ended in 1995, as the measurements of VOC at South Coast Air Basin were cancelled that year. The measurements of TSP (they present daily values of mass concentration) were conducted rather irregularly. The most frequent interval of subsequent measurements were 6 days; however, there is a number of much more extended gaps in the data series. The measurements of gaseous species were much more regular and were conducted usually each hour. Nonetheless, corresponding data series also contain a number of gaps. The hourly data for gaseous species were averaged from 10 to 18 o’clock each day of measurements. Such averaging were performed in order to get unique value of each component per a day and to suppress effects of possible random variations of species concentrations within the course of the day. Note also that since the aim of this study is to analyze manifestations of photochemical processes which are most active during light period of day, it is natural to consider dependence of PM on the day-time rather than night-time or daily mean concentrations of gaseous pollutants. No doubt, it would be advantageous to use TSP data for the same period of a day as NO$_x$ and VOC data, but they are not available. Nonetheless, it is reasonable to hope that neural networks themselves are able to retrieve a part of variations of daily mean TSP caused by variations of day-time concentrations of NO$_x$ and VOC.

The air quality data were matched by daily mean meteorological data from Climate Prediction Center, National Ocean and Atmospheric Administration (http://www.cpc.ncep.noaa.gov/). Again, it is reasonable to assume that the use of the meteorological data averaged over the same period of day as gaseous species data would facilitate to the better quality of the empirical models. Unfortunately, free hourly meteorological data were not found.

Finally, the data-sets analyzed in this study include 1880 and 1032 days of measurements at Azusa and Riverside-Rubidoux stations, respectively. The data have not been pre-selected in any way.
3  Methodology

The basic assumption in creation of the empirical models for aerosol is that mass concentration of TSP (this quantity is referred to below simply as PM) can be approximately represented as a sum of a nonlinear function \( f \) of \( n \) other measured values \( x_i, i=1,\ldots, n \) and some random value \( \delta \), variations of which cannot be accounted for by variations of the arguments of \( f \):

\[
PM = f(x_1,\ldots,x_n) + \delta,
\]

(4)

The proper choice of the arguments \( x_i \) is the most critical point in providing usefulness and reliability of the empirical models. Obviously, the arguments should include, in particular, such values, the dependence on which is to be studied. In this study, such values are gaseous precursors of PM, that is, NO\(_x\) and VOC. It would be of interest to consider other important gaseous precursors such as NH\(_3\) and SO\(_2\), but adequate contemporary measurements of these species are not available.

It is prudent also to take into account other important factors which, presumably, can significantly influence studied relationships. However, a number of such factors should not be too large. Indeed, it is obvious that larger number of arguments of a function, more data are needed to approximate it with the same accuracy. Taking into account the aforesaid, we have included in the PM empirical models only four meteorological values: temperature, specific humidity, and two horizontal components of wind speed. The empirical models with an additional argument, for example, total cloud cover which was included in our ozone empirical models (Konovalov, 2002), were found to demonstrate higher diversity between them in properties discussed in Section 4 than the models without this argument, what is indicative of the deficit of data. Thus we believe that the models with seven arguments are less adequate and thus such models will not be considered here. Nonetheless, it was found that empirical models with seven arguments retain the most important features of PM-NO\(_x\)-VOC relationships retrieved from the models with six arguments.

It is assumed further that \( f \) can be approximated by means of a neural network of the perceptron type with one hidden layer (see, e.g., Bishop, 1995). That the perceptron with just one hidden layer (given large enough number of neurons) is capable of approximating any measurable function to any desired degree of accuracy has been proven by Hornik et al. (1989). Mathematically, the perceptron \( p(X) \) with one hidden layer can be represented as follows:

\[
p(X) = \sum_{j}^{N} \left( w_j g_j(X) + w_0 \right),
\]

(5)

where \( X=(x_1,\ldots, x_n) \) is a vector of input values (that is, a vector of arguments of the function \( f \)), \( w_j, w_0, w_{ij}, w_{0j} \) are weight coefficients, \( N \) is a number of neurons in the network.

A procedure by which appropriate values of weights are determined is one of the key elements of the methodology of this study. Such a procedure is usually referred to as a training algorithm. Here a standard backpropagation algorithm (e.g., Reed and Marks, 1999) is used in combination with the early stopping method (Nelson and Illingworth, 1991; Sarle, 1995). The early stopping is usually used to provide such an important property of a neural network as generalization. For most applications generalization means that outputs of the network are able to approximate target values with given inputs that are not in the training set (Sarle, 2002). In case of this study, generalization means that a neural net learns the regular relationships \( f \) which are valid for any arbitrary chosen subset of the considered data-set and disregards noisy variations \( \delta \). Implementation of early stopping in this study includes the following steps. First, the sequence of days in the data-sets was randomly mixed. Thereafter each of the data-sets were divided into two equal subsets: training and
validation ones. The net is trained with the training subset but the error of prediction which is
defined as

\[ \varepsilon = \left[ \frac{2}{N_d} \sum_{i=1}^{Nd/2} \left( \frac{\text{PM}^{\text{ob}}_i - \text{PM}^{n}_i}{\text{PM}^{n}_i} \right)^2 \right]^{1/2}, \]  

(6)

where PM^{ob} and PM^{n}=p(X) are observed and predicted PM levels, and N_d is number of training
cases (that is, the total number of days presented in the data-set), is controlled with the validation
subset. The iterations are stopped as \( \varepsilon \) reaches minimum.

Direct application of this technique for creation of empirical models in the form of a single
perceptron (5) meets two main problems. First, it cannot be neglected a priori that perceptrons with
different numbers of neurons (N_N) may give qualitatively different results. Second, since the
training of neural networks is an iteration process, the results may depend on initial assumption
regarding values of weights. To resolve these problems, we considered first, to what degree changes
of N_N may affect quality of approximations and important features of the relationships analyzed. It
was found, in particular, that the dependence of \( \varepsilon \) (see (6)) on N_N, averaged over large number of
single perceptrons with different random initial weights, is saturating such that systematic
differences between perceptrons with different N_N greater, at least, than 20 are quite negligible. The
similar result were obtained for dependencies of derivatives of PM with respect to NOx and VOC,
evaluated from (5), on N_N. Values and, especially, signs of these derivatives are the most important
characteristics in the context of this study. Based on these results, it was concluded that the
dependence of the approximations on N_N is not of any significance if N_N is greater, at least, than 20.
In order to minimize any hypothetical uncertainties associated with choice of definite value of N_N to
maximum degree, it was decided to include in the models perceptrons with different N_N ranging
from 21 to 30.

To get rid of uncertainty associated with initial values of weights, a large enough set (50) of
perceptrons with random initial weights were trained for each value of N_N and a few (5) "best"
(having minimum \( \varepsilon \)) networks selected from that set were included in the empirical model. Finally,
the desired empirical model is represented as the mean of outputs of the set of 50 networks, which
includes five “best” networks for each N_N from 21 to 30:

\[ PM = \frac{1}{50} \sum_{N_N=21}^{30} \sum_{i=1}^{5} p(X, N_N, i), \]  

(7)

where \( i_b \) are the indexes of the best nets.

That the uncertainties in results of the empirical models, associated with the choice of values N_N
and initial values of weights are indeed quite small is illustrated in Fig. 1. This figure presents
comparisons of evaluations of first derivatives of PM with respect to VOC obtained from pairs of
empirical models created in general accordance to (7) but with N_N ranging from 21 to 25 and from
26 to 30. Each point in the figures represents one day in the training sub-set of the data-sets. It is
seen that each pair of independently created versions of the empirical models gives very close
results. In particular, fractions of days for which estimated signs of \( \partial PM/\partial VOC \) values are different
in the different versions are only 2.1% and 3.1% for the Azusa and Riverside models, respectively.
It is reasonable to expect that the uncertainty of the full empirical models (7), associated with the
discussed factors, should be even smaller than that of their component parts.
The empirical model performances in representation of variability of PM are demonstrated in Fig. 2, which show measured values of PM (from the training sub-sets) versus those evaluated by the models. It is quite natural that the models do not reproduce the measured values exactly since they are intended to catch only a part of total variability, associated with changes of the factors represented by the model arguments. It is seen, however, that this part is quite substantial. This fact bear evidence that the model arguments indeed include values which strongly influence the PM level. The results for the validation subsets are very insignificantly worse.

The main idea of the analysis of the empirical models consists in the following. It is assumed that empirical models provide as precise as possible fits of actual PM-NO\textsubscript{x}-VOC relationships under various conditions determined by values of the arguments. The qualitative features of these relationships under the given conditions may be learned via evaluation of various derivatives of PM with respect to VOC and NO\textsubscript{x}. For example, positive first derivative PM with respect to VOC means that the increase of the level of VOC, most probably, would cause the increase of the level of PM under the given conditions; that is, the PM level demonstrates positive sensitivity to changes of the VOC level. It should be noted, however, that positive first derivatives of PM with respect to VOC or NO\textsubscript{x} may reflect simply the natural correlation between the levels of primary PM and gaseous pollutants as they mainly originate from the same emission sources. The positive second derivative of PM with respect to VOC under the same conditions means that the sensitivity would increase with the increase of VOC. The derivatives that present interest may be evaluated, in principle, under any values of argument from the range of observed conditions. However, the most accurate evaluations are expected for conditions observed at days presented in the training sub-set of the data-sets. The special investigation has shown that the evaluations of individual nets forming the empirical models for the conditions that are not in the training subset, including those that are presented in the validation sub-sets, are much more dispersed. This is, presumably, due to the fact that each individual neural network included in the model is learned for the conditions that are in the training sub-set but has more freedom under other conditions. Thus it was decided to analyze only the evaluations made for the days presented in the training sub-sets.

It is rather obvious that PM-NO\textsubscript{x}-VOC relationships in the real environment under windy conditions may be affected by peculiarities of distributions of sources of pollutants. Thus, after creation of the models, it was found to be useful to make the approximation of no wind conditions. That is, all model evaluations were made with values of wind speed in (7) to be zero. Such approximation is possible as similar values of pollutant concentrations, temperature and humidity can be associated with different magnitudes and directions of wind. No wind conditions allow us to identify the effects associated with local processes of formation of secondary PM.

The obtained evaluations are analyzed subsequently in order to find some associations between values (and signs) of different derivatives and argument values. For example, it is considered, whether days with negative values of first derivatives of PM with respect to VOC are more associated with lower temperatures than days with the positive derivatives, or with higher temperatures. If we had only one empirical model, any association of such type would be of little utility. Indeed, it would not be possible to conclude then, whether the observed association is the indication of some real regularity, or it is just a result of poor quality of the model. However, when the same result is found with other model created using the independent data-set collected in a different environment, we have much more grounds to believe that this result is indeed a consequence of some general regularity. It should be taken into account that untrained or trained with the random data-sets neural networks give random outputs and thus cannot demonstrates any significant correlation either between different output values, or between output values and values of arguments.
4 Results

4.1 PM-NO\textsubscript{x} relationship

The rate of gas-to-particle conversion due to the reactions (1)-(3) is proportional to concentration of hydroxyl radical. Thus it can be expected that the increase of the OH level under other constant conditions would lead to the increase of PM production, and vice versa. It is well known that OH itself depends on NO\textsubscript{x} in the essentially non-linear manner (see, for example, Kley et al., 1999). Specifically, OH increases with the increase of NO\textsubscript{x} under low enough NO\textsubscript{x} level, but may decrease under higher NO\textsubscript{x} level. Thus, there is a possibility that the decrease of NO\textsubscript{x} may actually lead to the increase of the level of PM. Though this idea is rather obvious, any experimental demonstration of this possibility, as far as we know, has not been discussed in peer reviewed literature earlier.

In order to check this idea, we have evaluated (using (7)) first derivatives of PM with respect to NO\textsubscript{x} \((\partial \text{PM}/\partial \text{NO\textsubscript{x}})\) for each day presented in the training sub-sets of the data-sets. It was found that though majority of those days are associated with positive \(\partial \text{PM}/\partial \text{NO\textsubscript{x}}\), there are a few of them (8 days from 940 ones in the training sub-set for the Azusa and 6 days from 516 ones for Riverside) associated with \(\partial \text{PM}/\partial \text{NO\textsubscript{x}}<0\). In order to understand the origin of negative \(\partial \text{PM}/\partial \text{NO\textsubscript{x}}\) in the empirical models, some additional analysis was undertaken.

First, we considered dependencies of PM on NO\textsubscript{x} for all days with \(\partial \text{PM}/\partial \text{NO\textsubscript{x}}<0\). Such dependencies are shown in Fig. 3. They were calculated from (7) assuming that all model arguments except NO\textsubscript{x} have fixed values corresponding to the days considered. The general features of all these dependencies are the presence of growing parts under lower levels of NO\textsubscript{x} and falling parts under higher levels of NO\textsubscript{x}. The same features are characteristic for the dependence of OH on NO\textsubscript{x} (Kley et al., 1999). Some of the dependencies for the Azusa station possess growing parts also under highest values of NO\textsubscript{x}. Most likely, this growth is an artifact of the empirical model and caused by insufficient number of data with high values of NO\textsubscript{x}. Indeed, the days with NO\textsubscript{x}>120 ppb constitute less than 3% of total number of days in the data-set for the Azusa station.

Second, it was found that at both stations, days with \(\partial \text{PM}/\partial \text{NO\textsubscript{x}}<0\) are associated, on the average, with significantly higher concentrations of NO\textsubscript{x} (107 ppb at the Azusa station and 110 ppb at the Riverside station), higher concentrations of VOC (2.43 ppmC and 2.61 ppmC) and lower temperatures (18.8\textdegree\textsubscript{C} and 14.6\textdegree\textsubscript{C}), as compared with corresponding values averaged over the whole data sets (58.5 ppb and 31.6 ppb; 2.38 ppmC and 2.21 ppmC; 22.2\textdegree\textsubscript{C} and 20.4\textdegree\textsubscript{C} at the Azusa and Riverside stations, respectively). Regarding influence of VOC level on values of \(\partial \text{PM}/\partial \text{NO\textsubscript{x}}\) it was found additionally that at both stations, all days except one at the Azusa station are associated with positive values of mixed second derivative \(\partial^2 \text{PM}/\partial \text{VOC}\partial \text{NO\textsubscript{x}}\). It means that the increase of the VOC level actually facilitates to the less negative values \(\partial \text{PM}/\partial \text{NO\textsubscript{x}}\). Thus the noted above association of days when \(\partial \text{PM}/\partial \text{NO\textsubscript{x}}<0\) with higher concentrations of VOC is, most probably, just a result of positive correlation between the ambient VOC and NO\textsubscript{x} levels. Any common, for both models, association between the signs of \(\partial \text{PM}/\partial \text{NO\textsubscript{x}}\) and second order derivatives of PM with respect to temperature and NO\textsubscript{x} were not found. This fact, most probably, indicates significant uncertainty of evaluations of these derivatives. It is known from the theory (see, for example, Kleinman, 1994; Sillman, 1999) that higher NO\textsubscript{x} level, lower insolation (which can be expected to be associated with lower temperature) and lower VOC level facilitate to the NO\textsubscript{x}-rich regime, when radical level falls with the increase of NO\textsubscript{x}. Thus, the noted results of the empirical models are, at least, do not contradict to the hypothesis that negative sensitivity of PM to NO\textsubscript{x} has the photochemical origin.
Finally, let us discuss the reasons of smallness of a fraction of days with $\partial \text{PM}/\partial \text{NO}_x<0$. We see two possible reasons. On the one hand, as NO$_x$ itself can be incorporated into particulate material according to the mechanism (2), (3), it is obvious that the ridge in dependence of PM on NO$_x$ should be shifted toward higher values of NO$_x$ in comparison with the ridge in dependence of OH on PM under the same conditions, and thus a fraction of days with $\partial \text{PM}/\partial \text{NO}_x<0$ should be smaller than a fraction of days with $\partial \text{OH}/\partial \text{NO}_x<0$. On the other hand, in the real environment, the level of NO$_x$ may correlate with intensity of the primary emission of PM and/or with the levels of other precursors of PM, such as NH$_3$ or SO$_2$, which are not included in the empirical models. Obviously, such correlation would lead to a persistent positive shift in $\partial \text{PM}/\partial \text{NO}_x$ values evaluated by the empirical models. Nonetheless, in our opinion, even with regard of quantitative uncertainties of the empirical model evaluations, the possibility of negative sensitivity of PM level to change of NO$_x$ is established here rather firmly.

4.2 PM-VOC relationship

4.2.1 Empirical results

The first evidence for the essential non-linearity of PM-VOC relationship was found by Meng et al. (1997) in results of their PM-CT model simulation of changes of PM$_{2.5}$ level in response to changes in emission rates of NO$_x$ and VOC at Riverside, California, under conditions of 28 August 1987. They found, in particular, that the decrease of VOC emission rate may lead either to the increase or decrease of PM$_{2.5}$ level, depending on the base case for NO$_x$ and VOC emission rates. Specifically, greater (by absolute value) negative sensitivity of PM level to changes of VOC level was found, in most cases considered, under higher levels of VOC. Unfortunately, fruitful comparison of the theoretical results by Meng et al. (1997) with our empirical results is seriously hindered due to the following reasons. (1) The model by Meng et al. and our empirical models use different measures of PM (PM$_{2.5}$ and TSP, respectively). (2) The arguments of the empirical models are surface concentrations of NO$_x$ and VOC. Thus the empirical models cannot account, for example, for the possible changes of NO$_x$ surface concentration after changes of VOC emission rate due to chemical interaction between VOC and NO$_x$. (3) The results of Meng et al. concern maximum 1-hour average concentration of PM, whereas our empirical models deal with daily averaged level of PM. Nonetheless, the consideration of dependence of PM on VOC given by the empirical model for the Riverside station under conditions of August 28, 1987 (not presented here) reveals that this dependence is also essentially non-monotonous and that high enough values of VOC correspond to negative sensitivity of PM to VOC, whereas the lower ones facilitate to the positive sensitivity.

Empirical model evaluations of the first derivative of PM with respect to VOC ($\partial \text{PM}/\partial \text{VOC}$) shows that some of the days presented in the data-sets are associated with negative sensitivity of PM to changes of VOC and others with positive one. These evaluations are plotted in Fig. 4 and 5 against temperature and concentration of VOC observed at corresponding days, respectively. It is seen, in particular, that noticeable fractions of the days (namely, 6.7% at the Azusa station and 31.4% at the Riverside station) have negative values of $\partial \text{PM}/\partial \text{VOC}$. Furthermore, the points with $\partial \text{PM}/\partial \text{VOC}<0$ in Fig. 4 are not distributed uniformly but tend “to avoid” definite ranges of moderately warm temperatures, thus forming a “butterfly-like” structures. Solid curves represent polynomial (of the 4th order) fits of the data and illustrate non-linear character of dependence (in the statistical sense) of $\partial \text{PM}/\partial \text{VOC}$ on temperature. These features allow us to hypothesize that different processes drive behavior of PM under low and high temperatures. This hypothesis finds further confirmation in Fig. 5, where points are separated into high-temperature and low-temperature “modes”: the threshold temperatures are chosen to be equal to 18°C at the Azusa station and 22°C at the Riverside station and marked in Fig. 4 by dash-dot vertical lines. Specifically, it is seen that, under low enough VOC level, the high-temperature mode is associated predominantly with positive $\partial \text{PM}/\partial \text{VOC}$, whereas
the majority of the points with negative $\partial PM/\partial VOC$ belong to the low-temperature mode. In contrast, under high enough VOC level, the most part of the points with $\partial PM/\partial VOC<0$ belong to the high-temperature mode. The noted features give an idea to perform separate analysis of PM-VOC relationships for the different temperature modes.

Table 1 presents average observed values of PM, VOC, NO$_x$, T, and RH, calculated separately for high temperature and low temperature modes, separated additionally with regard to the sign of predicted value of $\partial PM/\partial VOC$ in the respective days. This table presents also day fractions (of total number of days in the given temperature mode) satisfying to certain conditions concerning values of temperature, the sign of $\partial PM/\partial VOC$, and the sign of corresponding second order derivatives of PM. It is worth to note the following general for both models features. First, in the low (high) temperature mode, the average values of PM, VOC, NO$_x$, and T is less (greater) in days with negative $\partial PM/\partial VOC$ than in the days with positive $\partial PM/\partial VOC$. Second, the minority (majority) of days have negative values of $\partial^2 PM/\partial VOC^2$ and $\partial^2 PM/\partial T \partial VOC$ in the low (high) temperature mode. Values of relative humidity do not demonstrate general for both models changes with changes of the sign of $\partial PM/\partial VOC$. There may be only two reasons for that. First, at least one of the models did not “learn” the relations between $\partial PM/\partial VOC$ and RH satisfactorily. Second, these relations are indeed qualitatively different in different environments. At the given state of knowledge of these relationships, it is not possible to determine which of two reasons is true. The similar situation takes place with the behavior of $\partial^2 PM/\partial NO_x \partial VOC$. The only common feature is that the ratio of number of days having both negative $\partial^2 PM/\partial NO_x \partial VOC$ and $\partial PM/\partial VOC$ to number of days with negative $\partial PM/\partial VOC$ (this ratio can be obtained by dividing the number in the 8th row of the table to that in the 1st row) is significantly less than 0.5 at the high temperature mode. An attempt to identify physical and chemical reasons for noted features is described in the following section.

4.2.2 Discussion of PM-VOC relationship

First of all, it is necessary to note that the following discussion does not pursue the notion of exact identification of the processes which cause the noted above features of PM-VOC relationships. Such identification is out of the scope of this paper and would require in-detail consideration of complex processes that drive evolution of PM at the given site. Here, we would like to formulate and substantiate, wherever possible, the most plausible hypotheses for the explanation of the empirical relationships.

We believe that the main factor driving formation of secondary PM in the high-temperature mode is the behavior of OH radical. The behavior of OH radical is directly related to the behavior of ozone and underlines qualitative nonlinear features of O$_3$-NO$_x$-VOC relationships which were extensively studied during two last decades (see, for example, NRC, 1991). Correspondingly, the key processes responsible for the essential general nonlinear features of the dependence of OH on VOC and NO$_x$ have also been identified earlier. In particular, those features were interpreted by Sillman (1990, 1995) by means of the mechanistic model based on the equation of the radical budget. It is demonstrated below that the main empirical features of PM-VOC relationships discussed above are closely associated with features of OH-VOC relationships given by the Sillman’s model. Note that the use of the mechanistic model in this study appears to be quite natural, as our desire here is to appeal to the most general qualitative properties of the tropospheric photochemistry. For the sake of clarity of the following discussion, it is worth to give here brief description of that model, with inessential modifications.
It includes the schematic representation of the sequence of the essential processes determining the sources and sinks of the radicals OH, HO$_2$ and RO$_2$. (The latter denotes various organic radicals.) Specifically, RO$_2$ is formed in the reaction of a hydrocarbon (RH) with OH,

\[
\text{(R1)} \quad \text{RH} + \text{OH} \rightarrow \text{RO}_2 + \text{H}_2\text{O}
\]

The instantaneous equilibrium of RO$_2$ radicals is provided by the reaction

\[
\text{(R2)} \quad \text{RO}_2 + \text{NO} \rightarrow \text{HO}_2 + \text{NO}_2 + \text{products},
\]

which also provides the source of HO$_2$ radical, whose concentration is balanced by reaction

\[
\text{(R3)} \quad \text{HO}_2 + \text{NO} \rightarrow \text{OH} + \text{NO}_2
\]

A magnitude of the sum of concentrations of OH, HO$_2$ and RO$_2$ radicals (this sum is called as odd hydrogen) is determined by the equation of the radical budget,

\[
S_H = (\text{R4})[\text{NO}_2][\text{OH}] + \left[ \frac{(\text{R1})[\text{RH}]}{(\text{R3})[\text{NO}]} \right]^2 + \left[ \frac{(\text{R6})[\text{R1}]^2[\text{RH}]^2}{(\text{R3})(\text{R2})[\text{NO}]^2} \right][\text{OH}]^2,
\]

where it is assumed that the source of odd hydrogen, $S_H$, is balanced by its sinks provided by the reaction of OH with NO$_2$ (see (3)) with the rate coefficient (R4), and the radical-radical reactions,

\[
\text{(R5)} \quad \text{HO}_2 + \text{HO}_2 \rightarrow \text{H}_2\text{O}_2 + \text{O}_2
\]

\[
\text{(R6)} \quad \text{RO}_2 + \text{HO}_2 \rightarrow \text{ROOH} + \text{O}_2.
\]

In Eq. (8), square brackets denote concentrations of species. $S_H$ may be further parameterized as

\[
S_H = S_1 + S_2[\text{RH}],
\]

where it is assumed that the main processes responsible for the radical production are photolysis of ozone and aldehydes.

Assuming that the dependence of ratios of NO$_2$ and NO to NO$_x$ on VOC and NO$_x$ can be neglected, concentration of OH can be obtained by resolving the quadratic equation. With a little manipulation Eq. (8) can be represented in the most general, non-dimensional form as

\[
1 + \hat{\alpha} x = \alpha \nu + \frac{x^2 \nu^2}{y^2},
\]

where

\[
\nu = z[\text{OH}], \quad z = \left[ \frac{[\text{RH}]}{[\text{NO}]} \right]_0 \left( \frac{(\text{R5})}{(\text{R3})(\text{R2})} \right)^{1/2},
\]

\[
x = \frac{[\text{RH}]}{[\text{RH}]}_0, \quad y = \frac{[\text{NO}_x]}{[\text{NO}_x]}_0,
\]

\[
\alpha = \frac{(\text{R4})[\text{NO}_2]}{zS_1}, \quad \beta = S_2[\text{RH}]_0/S_1,
\]
and subscript “0” denotes some typical values of the species concentrations.

In typical polluted environments, the sources of radicals due to photolysis of ozone and aldehydes are of the same order of magnitude, so typical value of $\beta$ is about of unity. Thus the dependence of $\text{OH}$ on $\text{VOC}$ is almost exclusively determined by the parameter $\alpha$.

Specifically, when $\alpha \gg 1$, $\text{OH}$ increases with the increase of $\text{VOC}$ ($\partial v/\partial x = \beta/\alpha y > 0$). However, as $\alpha$ decreases, $\partial \text{OH}/\partial \text{VOC}$ can become negative. For example, when $\alpha << 1$, $v = y(1 + \beta x)^{1/2}/x$ and $\partial v/\partial x < 0$. It is evident from (11) that less value of $\alpha$ corresponds to situations with greater overall reactivity of hydrocarbons (i.e., greater $R_1$) and greater insolation and ozone concentration (i.e., greater $S_1$). It is clear that in the real atmosphere, those are associated with greater, on average, temperature. These features may be responsible for the appearance of the negative sensitivity of $\text{PM}$ to changes of $\text{VOC}$ at the high-temperature mode, the association of days with $\partial \text{PM}/\partial \text{VOC}$ with higher temperatures and the dominance of negative values of $\partial^2 \text{PM}/\partial T \partial \text{VOC}$ at those days.

It is obvious also from (10) that a similar “switch” from $\partial v/\partial x < 0$ to $\partial v/\partial x > 0$ under fixed parameter values can be caused by either an increase of $x$, or a decrease of $y$. This feature may be responsible for the empirical facts that the high-temperature mode is mostly associated with negative values of $\partial^2 \text{PM}/\partial \text{VOC}^2$, and days with $\partial \text{PM}/\partial \text{VOC} < 0$ at the high-temperature mode are associated with higher values of $\text{VOC}$ and, mostly, with positive values of $\partial^2 \text{PM}/\partial \text{NO}_x \partial \text{VOC}$. Note that $x$ and $y$ are involved in (10) asymmetrically. Specifically, in the situation with $\alpha \gg 1$, the increase of $y$ leads to the decrease of $\partial v/\partial x$, that is $\partial^2 v/\partial x \partial y < 0$, in contrast to the situation with $\alpha$ about unity, when $\partial^2 v/\partial x \partial y > 0$. At the same time, it can be seen that $\partial^2 v/\partial x^2$ is negative in both these situations and becomes positive only when $\alpha << 1$. This observation, with the additional assumption that in the real situations $\alpha$ is greater or insignificantly less than unity, may account for the empirical fact that domination of $\partial^2 \text{PM}/\partial \text{VOC}^2 < 0$ takes place independently on the sign of $\partial \text{PM}/\partial \text{VOC}$ at the high-temperature mode, whereas domination of $\partial^2 \text{PM}/\partial \text{NO}_x \partial \text{VOC} > 0$ is found only in days with $\partial \text{PM}/\partial \text{VOC} < 0$.

We believe that conditions associated with the low-temperature mode correspond, generally, to $\alpha \gg 1$. Indeed, the CT model evaluations (see, for example, Milford et al., 1994) show that significant areas at South Coast Air Basin (SCAB) are under $\text{NO}_x$-saturated conditions (that is, in terms of our model, with $\alpha > 0$) even in the hot and sunny days. Thus, it is reasonable to expect that significant decrease of insolation in colder seasons would drive the photochemistry of SCAB into conditions with $\alpha \gg 0$. Under these conditions concentration of $\text{OH}$ is relatively small and, as a result, the behavior of $\text{PM}$ in the low-temperature mode is affected by non-chemical processes to much greater extent than that in the high-temperature mode. In particular, it can be speculated that dominating positive values of $\partial^2 \text{PM}/\partial \text{VOC}^2$ in the low-temperature mode are associated, in part, with a physically evident possibility that the increase of total concentration of hydrocarbons is associated with the increase of number of individual semi-volatile hydrocarbons reaching the saturation. Consequently, the increase of measured total concentration of volatile hydrocarbons (that is, $\text{VOC}$) may be associated with the increase of the ratio of total concentration of hydrocarbons in the condensed phase to concentration of $\text{VOC}$, and as a result, with positive values of $\partial^2 \text{PM}/\partial \text{VOC}^2$. However, the same effect can be caused by the chemical processes of formation of $\text{PM}$ via oxidation of volatile hydrocarbons [Odum et al., 1996; 1997]. The rate of formation of $\text{PM}$ due to these processes is proportional to the product of concentrations of $\text{VOC}$ and $\text{OH}$, that is, in terms of Eq. (8), it is proportional to $x \nu$. It is easy to see that in situation with $\alpha \gg 1$, $\partial^2 x \nu/\partial x^2 > 0$. This process may account also for substantial fraction of days with $\partial^2 \text{PM}/\partial \text{VOC}^2 > 0$ at the high-temperature mode.
The most speculative point of this discussion concerns the reasons for negative sensitivity of PM to VOC under very low, relatively, temperatures. We failed to find any chemical reason for this, consistent with majority of the empirical facts. It can be assumed that the increase of concentration of VOC, associated with the increase of concentration of condensed semi-volatile hydrocarbons, hinders condensation of inorganic semi-volatile species such as \( \text{NH}_4\text{NH}_3 \) by providing a hydrophobic film on the surface of aerosol particles and thus leads, effectively, to the increase of saturation vapor pressure, or to the decrease of value of the Henry’s law constants. At present, we do not aware of experimental or theoretical data that would allow us to estimate actual possibility and significance of this effect. Under higher temperatures, which are associated, presumably, with less \( \alpha \) and, consequently, larger values of \( \partial \text{OH}/\partial\text{VOC} \) (\( \partial\nu/\partial x = \beta/\alpha y \) when \( \alpha >> 1 \)), this effect is suppressed, probably, by the more rapid increase of concentration of semi-volatile inorganic species with the increase of VOC. This effect may be suppressed also because of the increase of fraction of organic aerosol with the increase of concentration of VOC under the same temperature. Anyway, it is evident that results of the empirical models concerning the low-temperature conditions pose very interesting questions for further researches.

5 Summary and conclusions

This paper presents results of the first attempt of analysis of nonlinear relationships between airborne particulate matter and its gaseous precursors based on observational data. The analysis of the long term monitoring data at South Coast Air Basin, California, is performed by means of original method which assumes creation of the empirical models based on the artificial neural networks. Such models does not involve any a priori assumptions regarding character of analyzed relationships. The reliability of results of this study is confirmed by the fact that two empirical models created using independent data-sets which are obtained in two different monitoring stations separated by a distance more than 50 km manifest a number of common qualitative features of the relationships between PM, VOC and \( \text{NO}_x \). The models are assumed to provide the best fit for the real PM-VOC-\( \text{NO}_x \) relationships under different observed values of temperature, relative humidity, and horizontal wind speed. In order to suppress effects caused by transport of PM from different areas and to emphasize the features of PM-VOC-\( \text{NO}_x \) relationships determined by the local formation of PM, the approximation of zero wind speed is enforced in the models. With this approximation, the models are used to evaluate different derivatives of PM with respect to VOC and \( \text{NO}_x \) under conditions corresponding to a number of days presented in the data-sets. These evaluations are used further to identify general, in the statistical sense, regularities of PM-VOC-\( \text{NO}_x \) relationships.

It is found, in particular, that there are conditions when PM demonstrates a negative sensitivity to changes of \( \text{NO}_x \), or VOC (that is, when first derivatives of PM with respect to \( \text{NO}_x \) or VOC are negative). It is found also that negative sensitivity of PM to \( \text{NO}_x \) is facilitated by higher ambient \( \text{NO}_x \) concentrations and lower temperatures and VOC concentrations. These features are consistent with the hypothesis that the non-linearity of PM-\( \text{NO}_x \) relationships is caused by the corresponding non-linearity of the dependence of concentration of hydroxyl radical (OH) on \( \text{NO}_x \).

The qualitative features of relationships between PM and VOC are found to be dependent on temperature. In view of this dependence it is turned out to be fruitful to separate analyzed days into two “modes” with respect to value of temperature (T) observed at those days: the high-temperature mode and the low-temperature one, with the threshold temperature to be about 20°C. Consequently, the PM-VOC relationships are analyzed for each temperature mode separately. It is found that in the low-temperature mode, the appearance of negative sensitivity of PM with respect to VOC is facilitated by lower temperatures and VOC concentrations, whereas positive and greater values of the sensitivity correspond to higher temperatures and VOC concentrations. In the high-temperature
mode, the overall picture is inverse: the appearance of negative sensitivity of PM with respect to VOC is facilitated by higher temperatures and VOC concentrations, whereas positive and greater values of the sensitivity correspond to lower temperatures and VOC concentrations.

It is argued that the most significant features of PM-VOC-NO\textsubscript{x} relationships in the high-temperature mode are closely associated with the similar features of OH-VOC-NO\textsubscript{x} relationships, whereas the behavior of aerosol in the low temperature mode is strongly perturbed by peculiarities of the processes of condensation of primary and secondary semi-volatile compounds. In our opinion, the empirical results of this study provide the rich “food” for the further theoretical analysis.

The main condition of future successful applications of the methodology discussed in this paper is availability of long enough concurrent measurements of PM and its main gaseous precursors. Up to now, there were only a few stations satisfying to this criteria. Moreover, unfortunately, the necessary measurements of VOC at air pollution monitoring stations at California were almost totally cancelled in 1995 and thus it is not possible to apply this methodology for assessment PM-NO\textsubscript{x}-VOC relationships at the present-day situation. Nonetheless, in the near future, the same methodology may be applied to the analysis of data of PM\textsubscript{2.5} and PM precursors measurements at the Photochemical Assessment Monitoring Stations (PAMS) at US (US Environmental Protection Agency, 1994), which have begun to operate in recent years. Anyway, the results of this study emphasize the importance and fruitfulness of concurrent and long-term measurements of coupled parameters of the atmospheric environment.

The results of this study indicate that the elaboration of the effective strategy for reduction of the PM level can present a rather difficult problem in view of essential non-linearities of relationships between PM and its gaseous precursors. These non-linearities should be studied further.

Acknowledgements
This work was supported in part by the Russian Foundation for Basic Researches, grants 99-02-16162 and 02-02-17080.

References
Ackermann, I. J., H. Hass, M. Memmesheimer, A. Ebel, F. S. Binkowski, and U. Shankar, Modal aerosol dynamics model for Europe: Development and first applications, Atmos. Environ. 32, 2981-2999, 1998.

Binkowski F. S., and U. Shankar, The regional particulate matter model, 1. Model description and preliminary results, J. Geophys. Res., 100, 26,191-26,209, 1995.

Bishop, C.M., Neural Networks for Pattern Recognition, Clarendon Press, Oxford, 1995.

California Ambient Air Quality Data 1980-1997 (CA AQD), CD No. PTSD-98-010-CD, Air Resources Board, California Environmental Protection Agency, December 1998.

Finlayson-Pitts, B.J. and Pitts, J.N., Jr., Atmospheric Chemistry: Fundamentals and Experimental Techniques, John Wiley and Sons, New York, NY, 1986.

Jacobson, M. J., Development and application of a new air pollution modeling systemII. Aerosol module structure and design. Atmos. Environ., 31, 131-144, 1997.
Hornik, K., M. Stinchcombe, and H. White, Multilayer feedforward networks are universal approximators, *Neural Networks*, 2, 359-366, 1989.

Gardner, M. W. and S. R. Dorling, Artificial neural networks (the multilayer perceptron)-A review of applications in the atmospheric sciences, *Atmos. Environ.*, 32, 2627-2636, 1998.

Kleinman, L., Seasonal dependence of boundary layer peroxide concentration: the low and high NO<sub>x</sub> regimes, *J. Geophys. Res.*, 96, 20,721-20,733, 1991.

Kleinman, L. I., Low and high NO<sub>x</sub> tropospheric photochemistry, *J. Geophys. Res.*, 99, 16,831-16,838, 1994.

Kleinman, L. I., Ozone processes insights from field experiments – part II: Observation-based analysis for ozone production, *Atmos. Environ.*, 34, 2023-2033, 2000.

Kley, D., M. Kleinmann, H. Sanderman, and S. Krupa, Photochemical oxidants: state of science. *Environ. Pollution*, 100, 19-42, 1999.

Konovalov, I.B., Application of neural networks for studying nonlinear relationships between ozone and its precursors, *J. Geophys. Res.*, 107, D11, 10.1029/2001JD000863, 2002.

Meng Z., D. Dabdub, J. H. Seinfeld, Chemical coupling between atmospheric ozone and particulate matter, *Science*, 277, 116-119, 1997.

Meng Z., D. Dabdub, and J. H. Seinfeld, Size-resolved and chemically resolved model of atmospheric dynamics, *J. Geophys. Res.*, 103, 3419-3435, 1998.

Milford, J.B., D. Gao, S. Sillman, P.Blossey and A.G. Russel, Total reactive nitrogen (NO<sub>y</sub>) as an indicator of the sensitivity of ozone to reductions in hydrocarbon and NO<sub>x</sub> emissions, *J. Geophys. Res.*, 99, 3533-3542, 1994.

Nelson, M.C. and Illingworth, W.T., *A Practical Guide to Neural Nets*, Reading, MA: Addison-Wesley, 1991.

Nenes, A., S. N. Pandis, C. Pilinis, Continued development and testing of a new thermodynamic aerosol module for urban and regional air quality models, *Atmos. Environ.*, 33, 1553-1560, 1999.

Pilinis, C., and J. H. Seinfeld, Development and evaluation of an Eulerian photochemical gas-aerosol model, *Atmos. Environ.*, 22, 1985-2001, 1988.

Odum, J. R., T. Hoffmann, F. Bowman, D. Collins, R.C. Flagan, and J. H. Seinfeld, Gas/particle partitioning and secondary organic aerosol yields, *Environ. Sci. Technol.*, 30, 2580-2585, 1996.

Odum, J. R., T. P.W. Jungkamp, R. J. Griffin, R.C. Flagan, and J. H. Seinfeld, The atmospheric aerosol-forming potential of whole gasoline vapor, *Science*, 276, 96-99, 1997.

Reed, R.D., and Marks, R.J, Neural Smithing: *Supervised Learning in Feedforward Artificial Neural Networks*, Cambridge, MA: The MIT Press, 1999.

Sarle, W.S., Stopped Training and Other Remedies for Overfitting, Proceedings of the 27th Symposium on the Interface of Computing Science and Statistics, 352-360, 1995.
Sarle W.S., Usenet newsgroup comp.ai.neural-nets, ftp://ftp.sas.com/pub/neural/FAQ3.html, 2002.

Seinfeld, J. H. and Pandis, S.N, Atmospheris Chemistry and Physics: From air pollution to Climate Change, John Wiley & Sons, New York, NY, 1998.

Sillman, S., J. A. Logan, and S. C. Wofsy, The sensitivity of ozone to nitrogen oxides and hydrocarbons in regional ozone episodes, *J. Geophys. Res.*, 95, 1837-1851, 1990.

Sillman, S., The use of NOy, H2O2, and HNO3 as indicators for ozone-NOx-hydrocarbon sensitivity in urban locations, *J. Geophys. Res.*, 100, 14175-14188, 1995.

Sillman S., The relation between ozone, NOx and hydrocarbons in urban and polluted rural environments, *Atmos. Environ.*, 33, 1821-1845, 1999.

Stockwell, W.R., P. Middleton, J. S. Chang and X. Tang, The second generation regional acid deposition model chemical mechanism for regional air quality modeling, *J. Geophys. Res.*, 95, 16,343-16,367, 1990.

Stockwell, W. R., F. Kirchner, M. Kuhn, and S. Seefeld, A new mechanism for the regional atmospheric chemistry modeling, *J. Geophys. Res.*, 102, 25,847-25,879, 1997.

Sun Q., and A. Wexler, Modeling urban and regional aerosols near acid neutralityApplication to the 24-25 June SCAQS episode, *Atmos. Environ.*, 32, 3533-3545, 1998.

US Environmental Protection Agency, Photochemical Assessment Monitoring Station Implementation Manual, EPA-454/B-93/051. Research Triangle Park, NC 27711, March 1994, 1994.

Wexler, A. S., F. W. Lurmann, and J. H. Seinfeld, Modelling urban and regional aerosolsI. Model development, *Atmos. Environ.*, 28, 531-546, 1994.
Table 1. Average characteristics of days with $\frac{\partial PM}{\partial VOC} < 0$ and $\frac{\partial PM}{\partial VOC} \geq 0$ for different temperature modes

|                     | Azusa T<18°C | Azusa T≥18°C | Riverside-Rubidoux T<22°C | Riverside-Rubidoux T≥22°C |
|---------------------|--------------|--------------|---------------------------|---------------------------|
| Fraction of days (%)| 17.2         | 82.8         | 4.0                       | 96.0                      |
| PM (µg/m³)          | 68.7         | 79.8         | 138.5                     | 133.0                     |
| VOC (ppmC)          | 1.96         | 2.28         | 3.73                      | 2.38                      |
| NOₓ (ppb)           | 42.9         | 61.6         | 69.2                      | 58.0                      |
| T (°C)              | 9.8          | 13.6         | 28.1                      | 25.3                      |
| RH (%)              | 52.7         | 68.8         | 45.8                      | 39.7                      |
| Fraction of days with $\frac{\partial^2 PM}{\partial VOC^2} < 0$ | 0.42 | 41.4 | 3.3 | 62.1 |
| Fraction of days with $\frac{\partial^2 PM}{\partial NOₓ \partial VOC} < 0$ | 5.8 | 6.7 | 1.3 | 41.8 |
| Fraction of days with $\frac{\partial^2 PM}{\partial T \partial VOC} < 0$ | 0.8 | 29.2 | 4.0 | 67.3 |
Figure 1. Evaluations of first derivatives of PM with respect to VOC made with the empirical models formed by neural networks with numbers of neurons ($N_N$) ranging from 21 to 25 (axes of abscissas), versus the similar evaluations made with the empirical models with $N_N$ ranging from 26 to 30 (axes of ordinates). Each point corresponds to one day at the training sub-set of the data-set. Solid lines represent best linear fits of the relationships.

Azusa

$R^2 = 0.978$

Riverside-Rubidoux

$R^2 = 0.995$
Figure 2. The empirical model evaluations of the PM concentrations (axes of ordinates) versus actually measured PM concentrations (axes of abscissas). Each point corresponds to one day at the training sub-set of the data-set. Solid lines represent best linear fits of the relationships.
Figure 3. The dependencies of PM concentrations on NO$_x$ obtained from the empirical models for the days, for which negative sensitivity of PM to NO$_x$ is detected. NO$_x$ concentrations observed at these days are shown by crosses.
Figure 4. The empirical model evaluations of first derivatives of concentration of PM with respect to concentration of VOC, plotted against values of ambient temperature. The solid and dashed-dot lines represent 4th order polynomial fits of data and the boundary between two temperature modes, respectively. The dotted line marks the null level of the derivative values.
Figure 5. The empirical model evaluations of first derivatives of concentrations of PM with respect to concentration of VOC, plotted against values of ambient concentration of VOC. Dots and crosses correspond to low-temperature and high-temperature modes, respectively.