PERGAMON

Atmospheric Environment 35 (2001) 815–825

ATMOSPHERIC ENVIRONMENT

www.elsevier.com/locate/atmosenv

Neural networks and periodic components used in air quality forecasting

M. Kolehmainen*, H. Martikainen, J. Ruuskanen

Department of Environmental Sciences, University of Kuopio, P.O. Box 1627, FIN-70211 Kuopio, Finland

Received 8 March 2000; accepted 25 July 2000

Abstract

Forecasting of air quality parameters is one topic of air quality research today due to the health effects caused by airborne pollutants in urban areas. The work presented here aims at comparing two principally different neural network methods that have been considered as potential tools in that area and assessing them in relation to regression with periodic components. Self-organizing maps (SOM) represent a form of competitive learning in which a neural network learns the structure of the data. Multi-layer perceptrons (MLPs) have been shown to be able to learn complex relationships between input and output variables. In addition, the effect of removing periodic components is evaluated with respect to neural networks. The methods were evaluated using hourly time series of NO_2 and basic meteorological variables collected in the city of Stockholm in 1994–1998. The estimated values for forecasting were calculated in three ways: using the periodic components alone, applying neural network methods to the residual values after removing the periodic components, and applying only neural networks to the original data. The results showed that the best forecast estimates can be achieved by directly applying a MLP network to the original data, and thus, that a combination of the periodic regression method and neural algorithms does not give any advantage over a direct application of neural algorithms.

Keywords: Nitrogen dioxide; Self-organizing maps; Multi-layer perceptron; Model comparison; Residual

1. Introduction

In recent years, air quality has emerged as a major factor contributing to the quality of living in urban areas, especially in densely populated and industrialised areas. Air pollution control is needed to prevent the situation from becoming worse in the long run. On the other hand, short-term forecasting of air quality is needed in order to take preventive and evasive action during episodes of airborne pollution. In this way, by influencing people's daily habits or by placing restrictions on traffic and industry it should be possible to avoid excessive medication, reduce the need for hospital treatment and even prevent premature deaths. This is especially essential where certain sensitive groups in the population are concerned, such as children, asthmatics and elderly people (Schwartz, 1996; Tiittanen et al., 1999).

The trend in recent years has been to use more statistical methods instead of traditional deterministic modelling. A number of linear methods have been applied to time-series for air pollutants, especially to ozone forecasting (Simpson and Layton, 1983), including comparisons with neural network methods (Yi and Prybutok, 1996; Comrie, 1997). NO_2 time-series have also been investigated using linear methods (Ziomas et al., 1995; Shi and Harrison, 1997) and comparisons with neural networks (Gardner and Dorling, 1999). In their overview of applications of neural networks in the atmospheric sciences, Gardner and Dorling (1998) concluded that neural networks generally give as good or better results than linear methods.

So far, little attention has been paid to combining linear methods with neural networks in order to enhance the power of the latter. A general rule for applying neural networks says that the phenomenon to be learned by

* Corresponding author. Fax: + 358-171-63191.

E-mail address: mikko.kolehmainen@uku.fi (M. Kolehmainen).

1352-2310/01/$ - see front matter
PII: S1352-2310(00)00385-X

the network should be as simple as possible and all advance information should be utilised by pre-processing (Haykin, 1994). Natural candidates for this kind of advance information are periodic variations due to seasonal and community-originating fluctuations. In a more general perspective, this means investigating the spectral properties of time-series. An example that pursues this line can be found in Schlink et al. (1997), where Kalman filtering is applied to the forecasting of SO_2 concentrations. Even simpler approaches are worth investigating, however, when several methods are to be combined. Therefore, the method selected here was curve fitting using sine and cosine functions as regressors.

Another methodological aspect that deserves further testing is the performance of unsupervised neural methods, especially the self-organizing map (SOM), in relation to more generally used supervised methods such as the multi-layer perceptron (MLP) and the back-propagation (BP) algorithm. The principal benefit of the SOM is the possibility to visualise the weights of the trained network. This makes it easier to understand the relationships between the variables than with black-box-like modelling of MLPs. It should also be noted that less parameters need to be tuned in a SOM application.

The objective of this work was to investigate the forecasting capability of the following methods: regression using periodic components at the year, week and day levels, the MLP and SOM methods applied to the original time series and the MLP and SOM methods applied to the residual of the periodic components. A time series of NO_2 concentrations was used as an example.

2. Materials and methods

2.1. Data

The data were collected in the city of Stockholm, Sweden during the years 1994–1998. The metropolitan area was covered by means of four measuring stations for NO_2, and the values used here are averages of their readings. The following meteorological variables were also used (station at Hogdalen): temperature, wind speed, wind direction and solar radiation, in addition to which the hour of the day and month of the year were recorded. Note that the wind direction, hour and month variables were dichotomised using the sine and cosine functions. This enabled the neural algorithms to work properly despite discontinuities in the original cyclic signals.

In order to obtain a better understanding of the data and of the role of the different variables, data-mining was used to create descriptive statistics and graphical representations. This was done with Visual Data software, which utilises the NDA neural network toolbox (see Section 2.8 for details).

2.2. Handling missing data

Missing data items were filled in using the weighted Nearest-neighbour method described as the "Normal" method in Dixon (1979). This handles a measurement line of n variables as a co-ordinate in an n-dimensional space and takes the missing values from the nearest neighbour (measurement line) in that space where they are available, at the same time weighting the Euclidean distance metrics in proportion to the number of missing values in each line. Thus, measurement lines having more missing values are weighted as being further away than they actually are in order to compensate for their lesser reliability.

The original data set of hourly average values included 43,824 lines and nine variables. The total of 3698 missing values represented 0.94% of the whole data set. Because of this low percentage, no estimation of the effect of the replacement technique on the results was necessary. In any case, the point of the work was to make comparisons between two neural network methods and the removal of periodic components.

2.3. Extracting periodic components and trend

Since the factors mainly contributing to the NO_2 concentration are connected with the source activity (e.g. traffic or heat production) and periodic variations in nature (e.g. photochemical reactions in the atmosphere), it was natural to expect periodic components to be found in the NO_2 time series. These were expected at the year and week levels and in the form of daily variations, and it was also anticipated that both a constant factor and a trend would be detectable. In order to extract these components, a model was constructed by fitting regressors. The regressors selected for this purpose included zero and first degree curves (a constant and a line) and sine and cosine terms of different multifolds (see Eq. (1))

$$y_t = \sum_{k=1}^{m} (a_k \sin k\omega_0 t + b_k \cos k\omega_0 t) + \sum_{l=0}^{n} c_k t^l, \qquad (1)$$

where m is the maximum multifold of the sine and cosine terms, n the maximum multifold of curve fitting, ω_0 the basic frequency (e.g. $\omega_{\text{year}} = 2\pi/(365 \times 24)$, since the resolution of measurement is 1 h), a_k, b_k, c_l are coefficients to be resolved by fitting and t the time.

2.4. Multi-layer perceptron

The multi-layer perceptron (MLP) is the most commonly used type of feed-forward neural network. Its structure consists of processing elements and connections (Hecht-Nielsen, 1991). The processing elements, called neurons, are arranged in layers, the input layer, hidden

layer(s) and output layer. An input layer serves as a buffer that distributes input signals to the next layer, which is a hidden layer. Each unit in the hidden layer sums its input, processes it with a transfer function and distributes the result to the output layer. It is also possible for there to be several hidden layers connected in the same fashion. The units in the output layer compute their output in a similar manner.

The most common supervised learning algorithm is the back-propagation (BP) algorithm (Haykin, 1994). This is a gradient descent algorithm that is normally used to train a MLP network. The learning algorithm used here was Levenberg–Marquardt back-propagation of Matlab Neural Network toolbox. The transfer functions selected for the layers were sigmoid for the hidden layer and linear for the output layer. The number of neurons in the hidden layer was the optimum found by experimentation. An introduction and overview of MLP applications in the atmospheric sciences can be found in Gardner and Dorling (1998).

2.5. Self-organizing maps

SOM is one of the best known types of unsupervised neural learning algorithms. The aim is to find prototype vectors that can represent the input data set and at the same time achieve a continuous mapping from the input space to a lattice (Kohonen, 1997). This lattice can be used for visualisation or for estimating unknown parameters using the known parameters as input.

The SOM learning algorithm is initiated by assigning random values to the weight vectors of the network . The training patterns (measured data lines) are fed to the network one-by-one, and this is repeated a pre-determined number of times (epochs) for each of them. At each training step the best-matching unit (BMU) is found by comparing the input and weight vectors of the neurons in terms of Euclidean distance metrics. The weights of the BMU and its four neighbours (in this implementation) are then adjusted towards the input vector according to an update rule in which the learning rate factor decreases monotonically towards the end of learning. Finally, the training patterns are assigned to the neurons according to a stabilised SOM weight configuration. This enables averages and other statistics to be calculated for each neuron.

A variation on the SOM was used here (Koikkalainen, 1994). The software implementation of the tree-structured SOM (TS-SOM) consists of a number of SOMs organised hierarchically in several layers in a pyramid-like fashion. This arrangement results in the number of neurons at a larger level being four times the number at the previous level. The levels, and thus the numbers of neurons selected for estimation, were determined experimentally.

2.6. Application of the methods

The periodic components of the NO_2 concentrations were first extracted from the data for the years 1994–1997. The number of multifolds used in regressor fitting was 11 at the year level, 6 at the week level and 10 at the day level (Eq. (1)). These were selected to represent a sufficient number of multifolds but not to overlap with the next time level. The coefficients were then used to calculate an estimate for the year 1998.

The NO_2 residual for the years 1994–1997 was obtained by subtracting the fitted periodic components from the original time series, and the SOM and MLP networks were then trained using the residual and meteorological and timing variables. The number of neurons for the SOM was 4096 and the number in the hidden layer of the MLP was 24. The SOM and MLP models were then applied to the estimation of the residual NO_2 concentration for the year 1998, after which the residual estimate was added to the estimate given by the periodic components and the sum was used as a new estimate.

Finally, the SOM and MLP networks were trained with data for the years 1994–1997. The number of neurons for the SOM was 16,384 and that in the hidden layer of the MLP was 16. The models were used to estimate the NO_2 concentration directly for the year 1998.

All estimates were then compared with the known values for NO_2 in 1998. The outputs of the SOM models were based on the mean value of the NO_2 concentration of the best-matching neuron for the meteorological and temporal variables at the predicted time. The MLP models were constructed by splitting the data from the years 1994–1997 into training and test sets (sizes 90 and 10%, respectively).

2.7. Performance indicators

We used selected statistical indicators to provide a numerical description of the goodness of the estimates (Willmott, 1982; Willmott et al., 1985). One of the most common indicators used with neural networks is the root mean square error (RMSE). This is calculated according to Eq. (2).

$$\mathrm{RMSE} = \left(\frac{1}{N} \sum_{i=1}^{N} [P_i - O_i]^2 \right)^{1/2}, \tag{2}$$

where N is the number of data points, O_i the observed data point and P_i the predicted data point.

The coefficient of determination (R^2) tells us how much of the observed variability is accounted for by the estimate model. Even though it has its known defects in certain situations (Comrie, 1997), this measure was used in order to maintain compatibility with other studies.

R^2 is calculated according to Eq. (3).

$$R^2 = \frac{\sum_{I=1}^{N}[P_i - \bar{O}]^2}{\sum_{i=1}^{N}[O_i - \bar{O}]^2}, \tag{3}$$

where $\bar{O}$ is the average of observed data.

To describe how much the model underestimates or overestimates the situation, the bias was calculated according to Eq. (4).

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N}[P_i - O_i]. \tag{4}$$

A relative measure of error called the index of agreement (d) was calculated according to Eq. (5). This is a dimensionless measure that is limited to the range of 0–1.

$$d = 1 - \frac{\sum_{i=1}^{N}(P_i - O_i)^2}{\sum_{i=1}^{N}(|P_i'| - |O_i'|)^2}, \tag{5}$$

where $P_i' = P_i - \bar{O}$ and $O_i' = O_i - \bar{O}$.

The RMSE can be further divided into systematic and unsystematic components. This is achieved by first fitting a line by least-squares regression and then decomposing

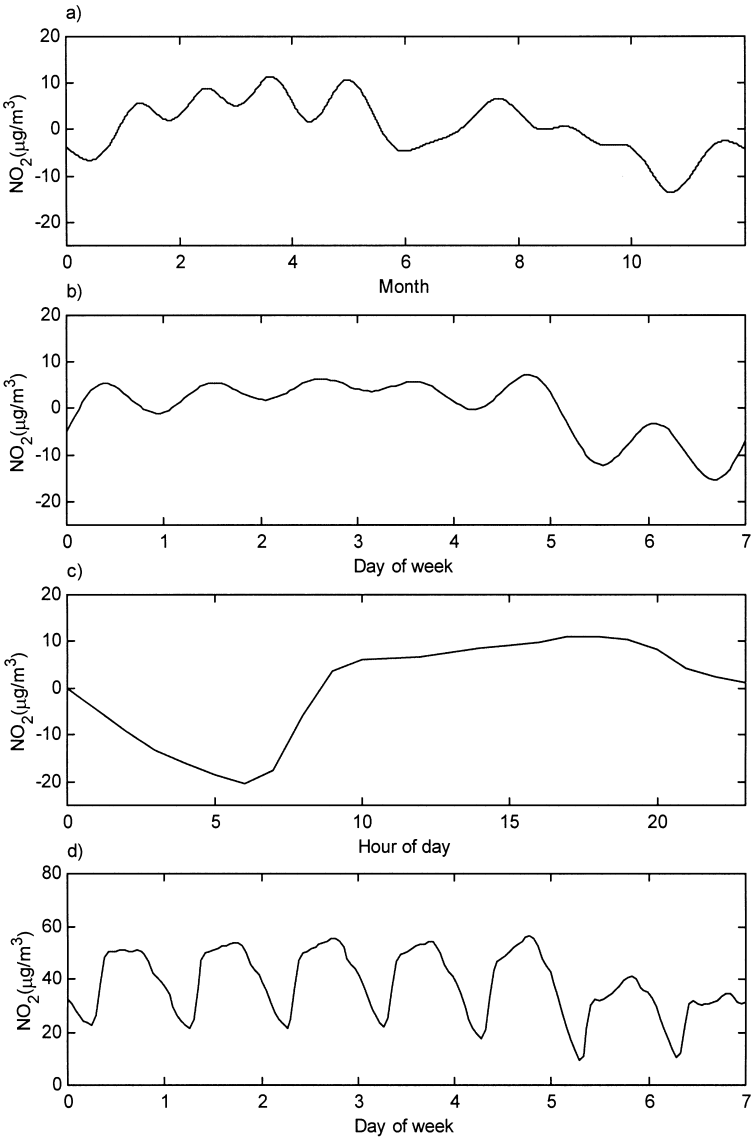

Fig. 1. Periodic components of NO_2 concentration of the data from the years 1994–1997: year-level component (a), week-level component from Monday to Sunday (b), day-level component (c) and a typical week composed of periodic components and an average (d).

Table 1
Regressor coefficients of the curve fitting using training data from the years 1994–1997[a]

l or k	Curves c_l	Year		Week		Day	
		a_k	b_k	a_k	b_k	a_k	b_k
0	0.04	—	—	—	—	—	—
1	0.00	0.22	−0.17	−0.28	−0.14	−0.59	−0.26
2	—	0.09	−0.16	−0.13	0.07	−0.12	0.17
3	—	0.05	0.08	0.01	0.07	0.15	0.03
4	—	−0.04	−0.03	−0.03	−0.03	−0.02	−0.05
5	—	0.02	0.07	−0.15	−0.01	−0.04	0.06
6	—	−0.10	−0.03	−0.01	0.15	0.03	0.01
7	—	−0.05	0.03	—	—	0.01	0.02
8	—	−0.02	−0.04	—	—	−0.01	0.00
9	—	−0.10	0.01	—	—	0.01	0.00
10	—	0.04	0.02	—	—	0.01	0.00
11	—	0.03	0.00	—	—	—	—

[a] l = multifold for curve fitting, k = multifold for sine and cosine fitting, a_k = coefficient for sine terms, b_k = coefficient for cosine terms and c_l = coefficient for curves.

the RMSE using Eqs. (6) and (7).

$$\mathrm{RMSE_S} = \left(\frac{1}{N}\sum_{i=1}^{N}(\hat{P}_i - O_i)^2\right)^{1/2}, \tag{6}$$

$$\mathrm{RMSE_U} = \left(\frac{1}{N}\sum_{i=1}^{N}(\hat{P}_i - P_i)^2\right)^{1/2}, \tag{7}$$

where $\hat{P}_i$ is a least-squares estimate.

Finally, the proportion of systematic error (PSE) was calculated according to equation (8):

$$\mathrm{PSE} = \frac{\mathrm{RMSE_S^2}}{\mathrm{RMSE}^2}. \tag{8}$$

2.8. Tools

The software used for calculation of the periodic components and for neural processing using the MLP was Matlab version 5.4. The neural processing using SOM and the pre-processing of the data were carried out with the neural data analysis (NDA) software package (http://erin.math.jyu.fi). The NDA software was run in batch mode in the Linux operating system. Preliminary data mining was performed with Visual Data software (http://www.visipoint.fi), which utilises the NDA toolbox in a Windows 95/98/NT environment.

3. Results

3.1. Periodic components of NO_2

The constant factor, trend and periodic components were extracted from the NO_2 signal, and the resulting components at the year, week and day levels are illustrated in Fig. 1. The corresponding coefficients are given in Table 1. The year level component can be seen to follow a known periodicity in which NO_2 concentrations are at their highest in winter and at their lowest in summer (Fig. 1(a)). The sub-monthly variation that could be detected is most probably due to overfitting of the data. No trend over the years could be detected. The week-level component seems to represent a community rhythm (probably traffic) in which lower concentrations can be found at weekends (Fig. 1(b)). The daily variation also follows the activities of the city (Fig. 1(c)). The pattern for a typical week in the year presented in Fig. 1(d) was constructed by summing the mean and all of the periodic components.

3.2. Some examples of predicted versus observed signals

In order to visualise the performance of the different methods and their combinations, estimated versus actual signal plots were made for representative weeks during the validation period (the year 1998). Plots representing the effects of using the periodic components for estimation purposes (a), applying the SOM algorithm to the residual after the periodic component (b) or to the raw signal (c) and using the MLP network to produce an estimate from the residual (d) the raw signal (e) are presented in Fig. 2. These plots apply to the 17th week of the year 1998, a time when all the estimates can be seen to be fairly accurate. An opposite example (the 41st week) is contained in Fig. 3, where the especially high concentrations (peaks) were difficult to find by any of these methods, although the MLP network could be seen to follow the observed signal reasonably well even in that week.

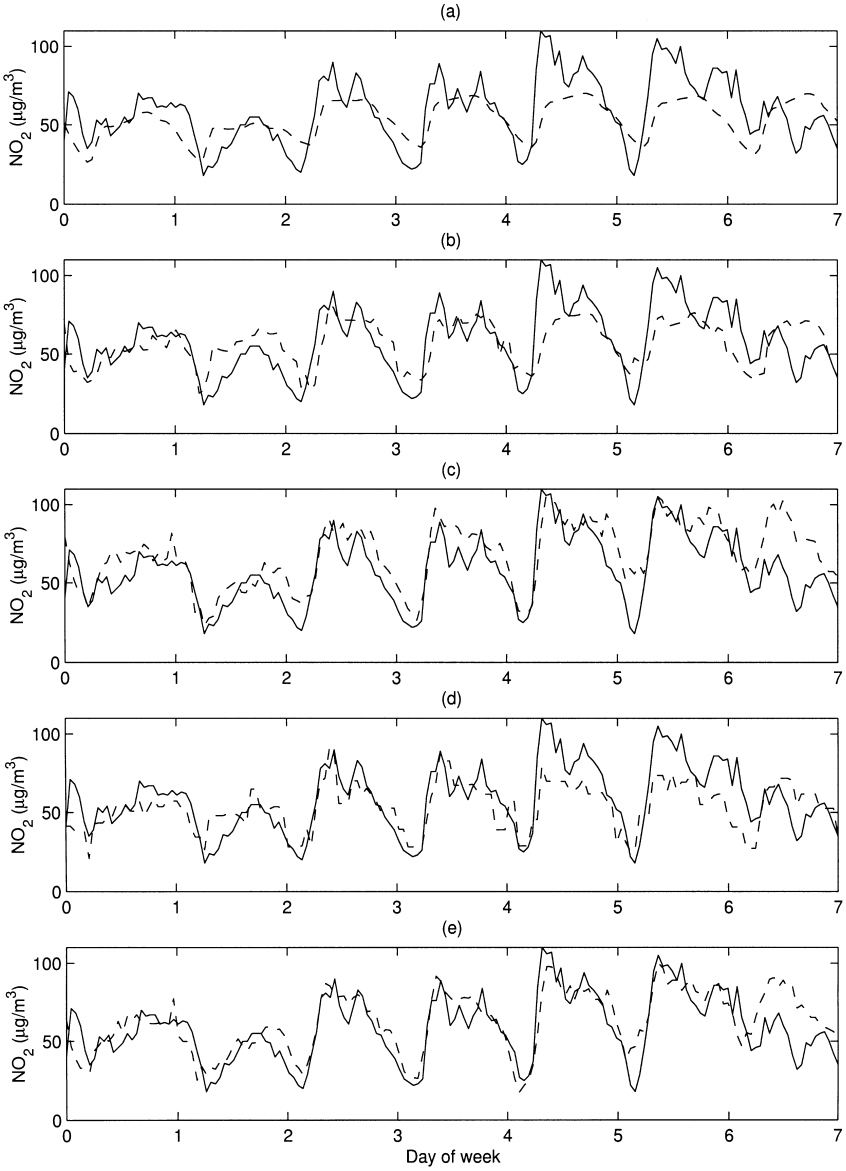

Fig. 2. Predicted (dashed line) versus observed (solid line) signals of NO_2 for the 17th week of 1998. (a) periodic fitting, (b) SOM applied to the residual, (c) SOM, (d) MLP applied to the residual and (e) MLP.

3.3. Histograms of errors

In order to achieve a more precise understanding of the goodness of the estimates, histograms of error residuals (estimated minus actual values) for the year 1998 were plotted for each of the five cases (Fig. 4). A good model constitutes of a shape where the number of hits near zero is maximised and where the total range is as narrow as possible. The shapes of the error histograms for the periodic components and for the SOM applied to the residual of periodic fitting are clearly the least satisfactory, while the best profile was achieved by applying the MLP algorithm to the raw signal, where the number of hits is the highest, especially with low error cases.

3.4. Indicators of statistical performance

Statistical indicators of error for the validation data (the year 1998), namely the root mean square error (RMSE), its systematic ($RMSE_S$) and unsystematic ($RMSE_U$) components, the coefficient of determination (R^2), index of agreement (d), proportion of systematic

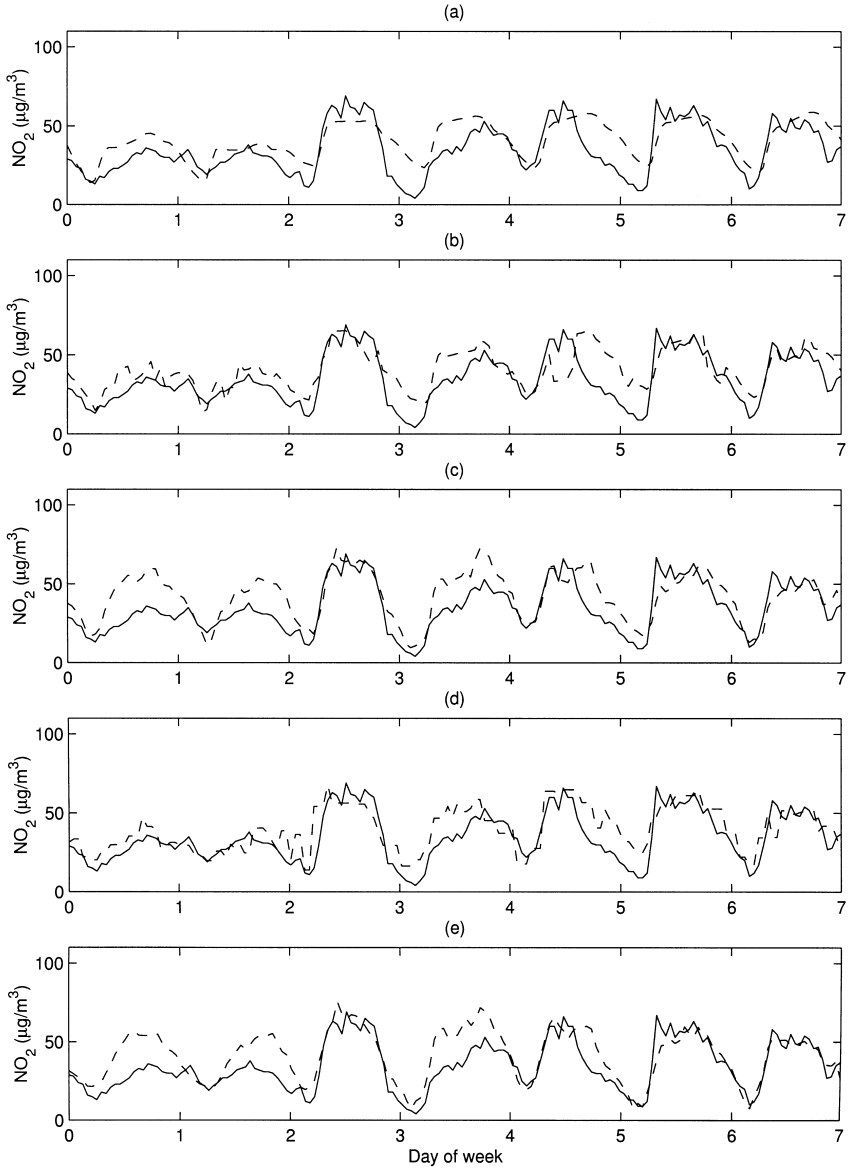

Fig. 3. Predicted (dashed line) versus observed (solid line) signals of NO_2 for the 41st week of 1998. (a) periodic fitting, (b) SOM applied to the residual, (c) SOM, (d) MLP applied to the residual and (e) MLP.

error (PSE) and bias, were calculated for each of the five methods (Table 2). The partitioning of the RMSE into systematic and unsystematic components is illustrated in the plots of observed vs. predicted values in Fig. 5. The systematic component is thus the difference between the perfect fit and the least-squares fit. The unsystematic component of the error is then the remaining error not explained by the systematic component. The plots are further enhanced with a LS fitting line and a line showing the perfect fit. The coefficients for this fitting are shown in Table 3.

It can be seen that the application of a MLP neural network to the raw data clearly gave the best results according to almost all the indicators, with the exception of Bias, where the SOM is the best. A low systematic RMSE value ($RMSE_S$) in particular is considered an indicator of a good model (Willmott, 1982). As expected, the application of neural algorithms has an advantage over linear methods, as represented by the regressor fitting of periodic components. However, the combina tion of linear and neural methods did not yield any advantage over the direct application of neural

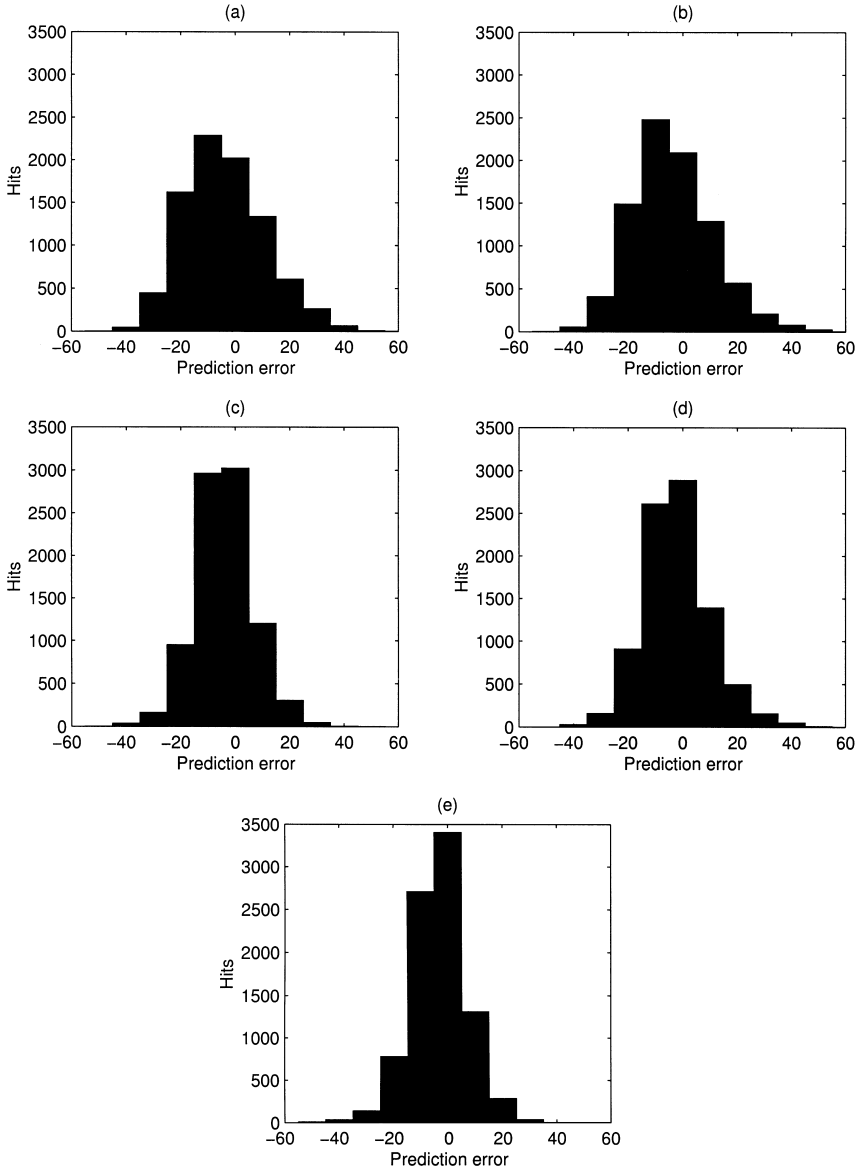

Fig. 4. Histograms of error in NO_2 predictions (signal–estimate) for the year 1998. (a) periodic fitting, (b) SOM applied to the residual, (c) SOM, (d) MLP applied to the residual and (e) MLP.

Table 2
Estimators for validity of the model for the validation data (1998)

Estimator	Periodic	Periodic + SOM	SOM	Periodic + MLP	MLP
R^2	0.61	0.77	0.72	1.00	0.96
Bias	3.93	3.93	2.17	4.05	3.28
d	0.73	0.77	0.85	0.89	0.90
RMSE	15.38	15.21	12.45	11.41	10.72
$RMSE_s$	10.86	9.34	7.16	5.43	4.72
$RMSE_u$	10.88	12.00	10.19	10.03	9.45
PSE	0.50	0.38	0.33	0.23	0.19

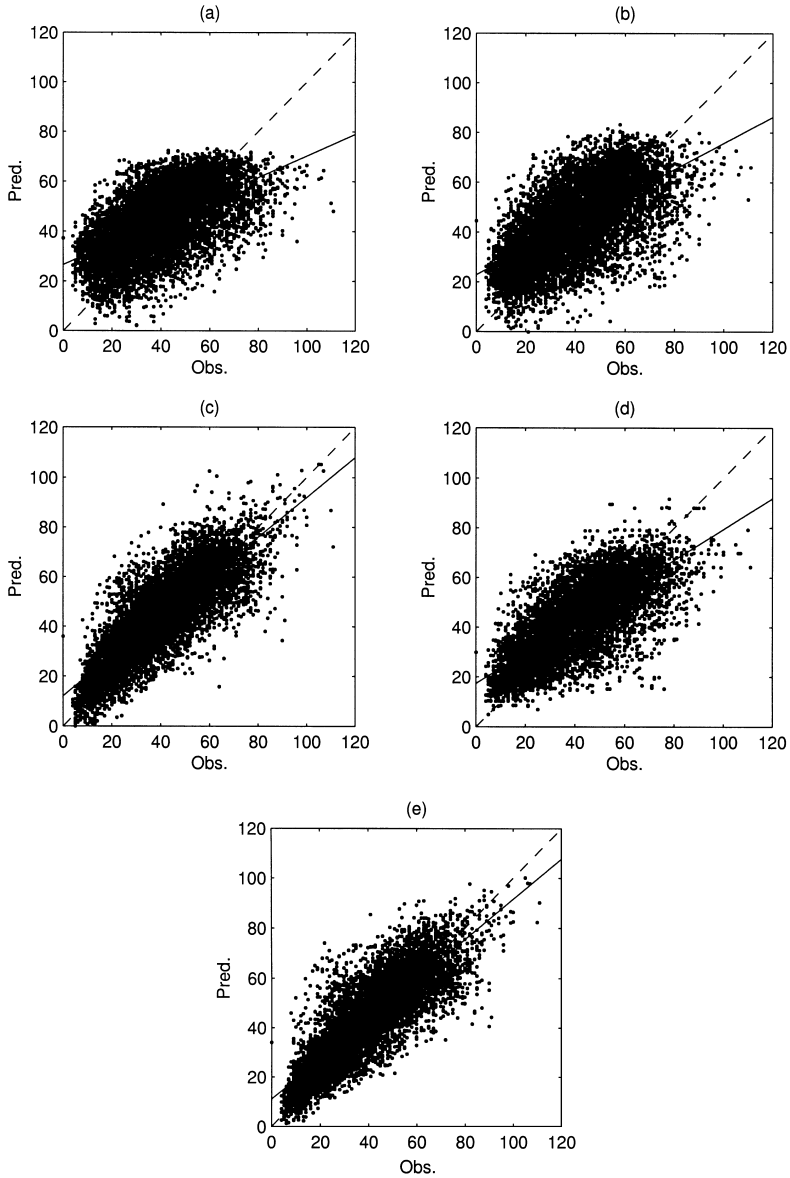

Fig. 5. Plots of predicted versus observed signals of NO_2 for the year 1998 using different models. The plots are further enhanced with a least-squares fitting line (solid) and a line showing the perfect fit (dotted). (a) periodic fitting, (b) SOM applied to the residual, (c) SOM, (d) MLP applied to the residual and (e) MLP.

Table 3
Regressor coefficients of LS fitting for systematic and unsystematic RMSE error for the validation data (1998)

Coefficient	Periodic	Periodic + SOM	SOM	Periodic + MLP	MLP
Bias	26.57	22.87	12.14	17.36	11.08
Slope	0.44	0.53	0.80	0.62	0.80

algorithms. These results are supported by the observed vs. predicted plots in Fig. 5.

3.5. Discussion

The results showed that good forecast estimates of air quality can be achieved by applying neural network methods to the forecasting of time series of NO_2 concentrations. However, as the signal studied shows considerable periodicity, a reasonable estimate can also be recovered simply by using the periodic components as an estimator.

As the method of fitting the periodic components used in this work was straightforward, it can be developed further in order to get more accurate results. The most obvious way would be to regularise by using less coefficients or using other known techniques like Tikhonov–regularisation. This holds especially for the year level model where sub-monthly variation can be detected.

In certain weeks of the year, none of the methods studied was able to forecast peak values, i.e. episodes when concentrations are at their highest. This is due to the inability of empirical models in general to capture extreme values which is caused by under-representation of these cases in the training data. Another way of expanding the models would be to add more meteorological variables. Variables describing the stability of the atmosphere would be especially good candidates for better explanations of the variations in the NO_2 concentrations and thus also means of achieving better forecasting results. Additionally, time-series data from the previous days could also be used.

Another route also worth exploring would be to consider the chaotic nature of atmospheric phenomena and to alter the model to account better for these non-linearities. An example of this kind of modelling is described by Chen et al. (1998).

4. Conclusions

Five methods of air quality forecasting purposes were evaluated here using hourly time series of NO_2 concentrations and basic meteorological variables from the city of Stockholm collected from years 1994–1998. The models were constructed using curve fitting of periodic signals and neural network methods in order to calculate estimates for the concentration or for the residual after removing the periodic components.

The results showed that fairly good estimates can be achieved by all of the methods, but neural network methods had some advantages. None of the methods was able to account for certain episodes, and therefore the models still need to be improved. It is suggested that meteorological variables describing the stability of the atmosphere should be included in the model as well as time-series data from previous days. Additionally, chaotic phenomena could also be considered.

Acknowledgements

This research was funded by the Technology Development Centre, Finland, Greenwin Oy (Kuopio), the Maj and Tor Nessling Foundation and the University of Kuopio. The staff of Slb Analys in Stockholm, especially Mr. Tage Jonson and Mr. Carl-G. Westerlund, are gratefully acknowledged for supplying the data and for their advice on this work. We would also like to thank the research group on Engineering and Computational Intelligence at the University of Jyväskylä, Finland, for allowing us to use their neural data analysis (NDA) package and Visipoint Oy for use of their Visual Data software for interactive analysis.

References

Chen, J.-L., Islam, S., Biswas, P., 1998. Nonlinear dynamics of hourly ozone concentrations: nonparametric short term prediction. Atmospheric Environment 32, 1839–1848.

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

Dixon, J.K., 1979. Pattern recognition with partly missing data. IEEE Transactions on Systems, Man, and Cybernetics, SMC-9 10, 617–621.

Gardner, M.W., Dorling, S.R., 1998. Artificial neural networks (the multi-layer perceptron) – a review of applications in the atmospheric sciences. Atmospheric Environment 32, 2627–2636.

Gardner, M.W., Dorling, S.R., 1999. Neural network modelling and prediction of hourly NO_x and NO_2 concentrations in urban air in London. Atmospheric Environment 33, 709–719.

Haykin, S., 1994. Neural Networks: A Comprehensive Foundation. Prentice-Hall, Englewood Cliffs, NJ.

Hecht-Nielsen, R., 1991. Neurocomputing. Reprinted with corrections. Addison-Wesley, Reading, MA.

Kohonen, T., 1997. Self-Organizing Maps, 2nd Edition. Springer, Berlin.

Koikkalainen, P., 1994. Progress with the tree-structured self-organizing map. ECAI '94. In: Cohn, A. (Ed.), Proceedings of the 11th European Conference on Artificial Intelligence, August 1994. Wiley, New York, pp. 211–215.

Schlink, U., Herbarth, O., Tetzlaff, G., 1997. A component time-series model for SO_2 data: forecasting, interpretation and modification. Atmospheric Environment 31, 1285–1295.

Schwartz, J., 1996. Air pollution and hospital admissions for respiratory disease. Epidemiology 7, 20–28.

Shi, J.P., Harrison, R.M., 1997. Regression modelling of hourly NO_x and NO_2 concentrations in urban air in London. Atmospheric Environment 31, 4081–4094.

Simpson, R.W., Layton, A.P., 1983. Forecasting peak ozone levels. Atmospheric Environment 17, 1649–1654.

Tiittanen, P., Timonen, K.L., Ruuskanen, J., Mirme, A., Pekkanen, J., 1999. Fine particulate air pollution, resuspended road dust and respiratory health among symptomatic children. European Respiratory Journal 13, 266–273.

Willmott, C., 1982. Some comments on the evaluation of the model performance. Bulletin of the American Meteorological Society 63 (11), 1309–1313.

Willmott, C., Ackleson, S., Davis, R., Feddema, J., Klink, K., Legates, D., O'Donnell, J., Rowe, C., 1985. Statistics for the evaluation and comparison of models. Journal of Geophysical Research 90 (C5), 8995–9005.

Yi, J., Prybutok, V.R., 1996. A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area. Environmental Pollution 92, 349–357.

Ziomas, I., Melas, D., Zerefos, C.S., Bais, A.F., Paliatsos, A.G., 1995. Forecasting peak pollutant levels from meteorological variables. Atmospheric Environment 29, 3703–3711.