Published paper
Zito, P., Chen, H.B. and Bell, M.C. (2008) *Predicting real-time roadside CO and NO2 concentrations using neural networks,* IEEE Transactions on Intelligent Transportation Systems, Volume 9 (3), 514 - 522.

Predicting Real-Time Roadside CO and NO_2 Concentrations Using Neural Networks

Pietro Zito, Haibo Chen, and Margaret C. Bell

Abstract—The main aim of this paper is to develop a model based on neural network (NN) theory to estimate real-time roadside CO and NO_2 concentrations using traffic and meteorological condition data. The location of the study site is at a road intersection in Melton Mowbray, which is a town in Leicestershire, U.K. Several NNs, which can be classified into three types, namely, the multilayer perceptron, the radial basis function, and the modular network, were developed to model the nonlinear relationships that exist in the pollutant concentrations. Their performances are analyzed and compared. The transferability of the developed models is studied using data collected from a road intersection in another city. It was concluded that all NNs provide reliable estimates of pollutant concentrations using limited information and noisy data.

Index Terms—Multilayer perceptron (MLP), pollutant concentration prediction and air quality, radial basis function (RBF).

I. INTRODUCTION

THE POLLUTANT concentrations in urban areas that are produced by traffic emissions depend on vehicle characteristics, traffic and weather conditions, geographic and built environment characteristics of the local site, etc. [14], [16]. Carbon monoxide (CO) and nitrogen oxides [collectively known as NO_x and include nitrogen monoxide (NO) and nitrogen dioxide (NO_2)] are mostly from vehicle exhausts and, hence, the focus of this paper. When emitted into the atmosphere, CO has a stable behavior, which means that it does not change its structure through chemical reactions with other pollutants or substances [6]. In contrast, NO reacts with ozone (O_3) to form NO_2. NO_2 can revert to NO by energy gained from sunlight. It has been shown that NO_2 causes adverse effects in the pulmonary function when inhaled at high concentrations [6]. The estimates of the pollutant concentration near road intersections are the focus of air pollution regulations.

The use of deterministic models to forecast pollution levels in urban areas is made rather difficult due to the complex topography and heat phenomena, which characterize the urban

P. Zito is with the Department of Transportation Engineering, University of Palermo, 90128 Palermo, Italy (e-mail: zito@ditra.unipa.it).

H. Chen is with the Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, U.K. (e-mail: H.Chen@its.leeds.ac.uk).

M. C. Bell is with the Department of Civil Engineering and Geosciences, University of Newcastle, Newcastle upon Tyne NE1 7RU, U.K. (e-mail: margaret.bell@ncl.ac.uk).

environment. The advantage of using the neural network (NN) lies in the fact that it learns to identify the intrinsic relationships between the input and output data by being repeatedly presented with the data during its training phase. The NN can approximate highly nonlinear functions with little information about the nature of these relationships hidden in the data. The use of the NN allows us to better understand the complex intrinsic and nonlinear relationships, which may not be easily modeled by conventional statistical approaches. As a result, the NN often presents an improvement over existing air quality models [6], [14].

Similar applications of NNs to forecast pollutant concentrations and to model dispersion phenomena have widely been used for a number of years. Moseholm *et al.* [14] studied the usefulness of the NN to understand the relationships between traffic parameters and CO concentrations measured near an intersection. Dorzdowicz and Benz [5] developed a dispersion model based on the NN to estimate hourly mean concentrations of CO in the urban area of Rosario City, Argentina. Their analysis considered 11 input variables (including flow, wind speed and direction, solar radiation, humidity, pressure, rain fall, and temperature). Gardner and Dorling [6] developed a multilayer perceptron (MLP) NN model with hourly NO_x and NO_2 and meteorological condition data of Central London, U.K. Their results showed that the NN model outperforms the regression models developed by Shi and Harrison [18] using the same study site. Similar work was done in Santiago, Chile [16], and in Perugia, Italy [20]. Furthermore, Grivas and Chaloulakou [7] used the NNs to predict PM_{10} hourly concentrations in the metropolitan area of Athens, Greece, by comparing their performance with a multivariate regression model, whereas Pelliccioni and Tirabassi [15] showed that the integrated use of dispersion models and NNs can improve the prediction performance of models. Questions remain on whether the NN models developed at a site can still perform well at other locations and can enhance the pollutant prediction models during peak periods with negative impacts on human health. In this analysis, we developed a methodological approach to achieve the transferability of neural models on other sites. In particular, the proposed approach does not depend on the geometry of the road intersection. In fact, by aggregation of the inbound and outbound traffic parameters of each road link and by testing of the trained NN on another intersection with unseen data, we were able to achieve transferability. In addition, we conducted a comparison between different NN paradigms, namely, the MLP, the radial basis function (RBF), and the modular network, to estimate CO and NO_2 concentrations at roadside during peak and off-peak periods, achieving different results. A sensitivity

analysis was carried out to study the formation of the pollutant concentrations to understand the interrelationships between traffic, weather conditions, and pollutant concentrations and to evaluate the strength of these relationships. Two classes of input variables, namely, traffic and meteorological condition data, are taken into account. The traffic data, which are extracted from the instrumented City (iC) database [2], include flow, delay, stops, and congestion for each direction of the chosen road intersection. The meteorological condition data considered are wind speed, wind direction, and ambient temperature. The sensitivity analysis of the input parameters is to determine the most important inputs, in particular, understanding which are the dominant links of the road intersection to the formation process of pollutant concentrations. Their knowledge is crucial for decision makers in planning more effective traffic control strategies to decrease congestion and the resulting pollution, improving environmental sustainability.

II. NEURAL MODELS

A. MLP

The MLP consists of a set of simply interconnected neurons. The output of each neuron (except those in the input layer) is scaled by the connecting weights, modified by a transfer or activation function, which can be either linear or nonlinear, and fed forward to be an input to the neurons in the next layer of the network.

The output y_i of a neuron is given by following equation:

$$y_i = f\left(\sum_{k=0}^{N} x_k w_{k,i} + b_i\right) \quad (1)$$

where x_k is the output of a neuron of the previous layer, constituted by N neurons, $w_{k,i}$ is the link weight between the kth previous layer's neuron and the ith neuron of the current layer, and b_i is the bias constant.

The MLP NN, which is used in this paper, learns in a supervised manner and, hence, requires a training data set, which consists of an input vector and an associated target vector. During training, the NN is repeatedly presented with the training data, and the weights in the network are adjusted until the output vector produced by the network does not match the target vector within a certain error. The training process uses this error to adjust the weights of the network according to the gradient descent learning algorithm to minimize the error.

The initial problem using an MLP NN is to decide the number of layers and neurons in each layer. The number of input and output neurons is determined by the nature of the specific problem. Generally, only one hidden layer is required to approximate any smooth measurable function between inputs and outputs [8]. The optimum number of neurons required in the hidden layer depends on the degree of desired accuracy and, generally, has to be found using a trial-and-error approach. Therefore, the goal is to determine the smallest network that is able to adequately capture the relationship to be modeled.

Overtraining may occur if the model learns the noisy details in the training data and results in poor generalization capabilities when new data are presented [3]. For this purpose, our data are divided into three subsets for training, validation, and testing. The validation set is used during training to check the generalization performance. Training can be stopped when the performance on the validation data reaches a maximum or starts to decline. Finally, the performance of the trained NN is assessed using the test data, which are what we report in this paper. This network is widely employed in the literature [19].

B. Modular NN (MNN)

With respect to the previous network, the MNN has, in addition, an integrating unit, called the gating network, that assigns different features of the input space to the different expert networks (subnet). The expert and gating networks receive the input. The gating network has output nodes that are equal to the number of expert networks. Each expert module produces a response corresponding to the input vector, and the output of the MNN is the weighted sum of these responses with the weights equal to the gating network output. The output of the MNN is written as follows:

$$\mathbf{y} = \sum_{i=1}^{N} g_i y_i \quad (2)$$

where $\mathbf{y}$ is the output vector of the MNN, y_i is the output of expert i, and g_i is the ith output of the gating network. Each expert (subnet) is typically a multilayer network. The gating network can also be a multilayer network, with a softmax function as the transfer function in the last layer, i.e.,

$$g_i = \frac{e^{s_i}}{\sum_{j=1}^{N} e^{s_j}} \quad (3)$$

where s_i is the net input to the last layer transfer function. The role of the gate is to find out which weighting should be given to each expert contribution.

The modular network has been chosen since the modular architecture allows the decomposition and assignment of tasks to several modules by its modular architecture (competitive network [10]). The learning algorithm can be the classic error backpropagation.

C. RBF

An interesting alternative to the aforementioned NNs is the RBF NN, which consists of a set of bell-shaped basis functions, each of which has a center and a spread, to approximate the output of the network using the Euclidean distance between the input vector and the center vector [4], [11]. These statements can be mathematically expressed by the following equations:

$$\phi_j(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \boldsymbol{\mu}_j\|^2}{2\sigma_j^2}\right) \quad (4)$$

where $\phi_j(\mathbf{x})$ is the output of the jth basis function. σ is a parameter, the value of which controls the spread of the function and, hence, the smoothness of the approximation. $\mathbf{x}$ is the n-dimensional input vector, and $\boldsymbol{\mu}_j$ is the vector determining the center of the jth basis function ϕ_j.

Each of the basis functions of the RBF NN is connected to the neurons in the output layer by a weight, as shown in the following form:

$$y_k(\mathbf{x}) = \sum_{j=1}^{M} w_{k,j}\phi_j(\mathbf{x}) + w_{k,0} \qquad (5)$$

where $y_k(\mathbf{x})$ is the output of the kth output neurons for the input vector $\mathbf{x}$, M is the number of basis functions in the RBF network, $w_{k,j}$ is the weight between the kth output neuron and the jth basis function neuron, and $w_{k,0}$ is the bias constant for the kth output neuron.

Poggio and Girosi [17] showed that RBF networks possess the property of best approximation. An RBF network is normally trained in two phases. First, the parameters governing the shapes of the basis functions (i.e., σ and $\boldsymbol{\mu}$) are determined through an unsupervised optimization procedure (i.e., no target information required). The basis function centers $\boldsymbol{\mu}_j$ can be considered as prototypes of the input vectors. The second phase is to find the weights $w_{k,j}$ by a supervised training scheme, which resolves a linear matrix equation [3]. The RBF-based self-learning algorithm automatically adds nodes into the hidden layer when needed, until the actual error is below the given error.

III. DATA COLLECTION AND ANALYSIS

The data used in this paper were provided by the iC facility, which is capable of storing the M02 and A02 messages produced by the Split Cycle and Offset Optimization Technique (SCOOT) system [9] that is widely implemented in U.K. cities. The study site was a road junction in Melton Mowbray, which is a town in Leicestershire, U.K. with a population of 25 000. The town has heavy through traffic (with about 7% heavy vehicles), as well as local traffic on its urban roads. The major source of air pollution in Melton Mowbray is from traffic. These characteristics made it an ideal location to study the complex phenomena of vehicular emissions and the resulting air quality. The chosen road junction has seven links: two links on Norman Way (identified by I and E in the SCOOT network diagram), three links on Wilton Road (identified by J, B, and A), and one link on each of Nottingham Road (identified by D) and Asfordby Road (identified by C), as shown in Fig. 1. The roadside pollution monitoring (RPM) unit is located on Norman Way, around 5 m back from the junction with a height from ground of 2 m and a distance from road of 3 m, on a pedestrian island between the left-turning and straight-ahead traffic lanes of westbound traffic. The air inlet of the unit faces the straight-ahead traffic. Prevailingly, the wind direction is to the southwest with an average value of 203°.

The choice of input variables is generally based on *a priori* knowledge of the physics and chemistry processes, which may explain the pollutant concentration levels. According to the literature [13], roadside pollutant concentration levels depend on many factors, such as vehicle characteristics (e.g., fuel type and engine power), vehicle categories (e.g., cars, lorries, buses, and motorcycles), vehicle age, maintenance level, traffic conditions (e.g., speed, flow, delay, stops, and acceleration), meteo-
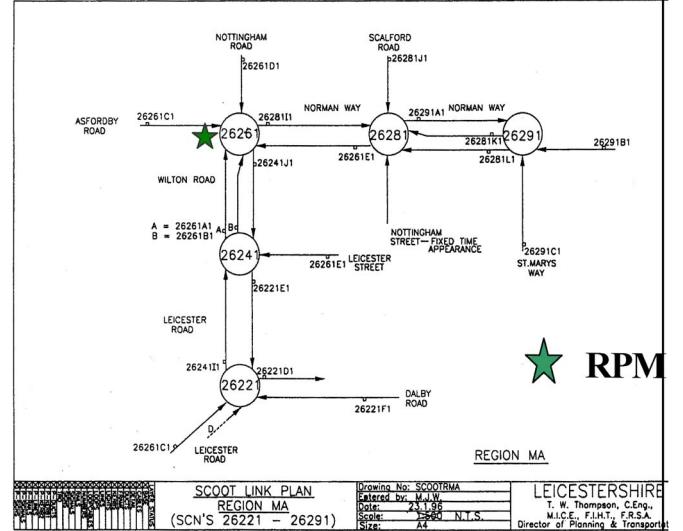

Fig. 1.   Study site showing neurons, links, and location of RPM.

rological conditions (e.g., temperature, wind speed, and wind direction), and site characteristics (e.g., topology of network and canyon streets). All these factors can cause a significant change in the levels of roadside pollutant concentrations.

Traffic and meteorological condition data were available as input variables in this paper. The traffic data (i.e., flow, delay, stops, and congestion) were collected by the SCOOT system every 5 min.

Flow is the number of vehicles that cross the stop line on a road link (number of vehicles/hour). Stop is the number of vehicles that stopped at least once along a road link (number of vehicles/hour). Delay is an estimate of the total delay, in one-tenth vehicle-hours per hour, experienced by all vehicles arriving at the stop line. Congestion is the total number of 4-s intervals in a time interval during which the loop on a road link was continuously occupied (time interval/hour). The prevailing meteorological condition data were recorded every 15 min by a facility installed 5 mi away from the study site. The CO and NO_2 concentrations were measured every minute by the RPM unit and averaged into 5-min intervals to synchronize with the traffic data before presented to the NNs.

All the data used are regularized into 5-min intervals. Data between January 1, 2001, and December 31, 2001, inclusive, are collected and preprocessed to remove invalid values for the training and testing of NNs. The initial data set of input variables is made up of 31 parameters (four traffic parameters from each of the seven links plus three meteorology parameters) and consists of 115 000 cleaned records (65 000 for 2001 and 50 000 for 2002). Initially, the data were analyzed by using traditional statistical methods to investigate whether any obvious correlations were present. The linear regression analysis carried out by SPSS highlighted low coefficients of correlation that are equal to 0.656 and 0.538 for CO and NO_2, respectively. The determination of the number of appropriate input variables is particularly important for complex problems such as air-quality prediction because a large number of input variables can increase the size of the NN model. Larger networks have a number of disadvantages, including decreased processing speed

and an increase in the amount of data required to estimate the model parameters (i.e., weights). An appropriate input set was determined using a sensitivity analysis.

IV. METHODOLOGY

The number of input and output neurons in the NNs developed was driven by the nature of the problem under study, and the number of neurons in the hidden layer governs the required degree of accuracy and is therefore a parameter in formulating an NN model. The 2001 data set was divided into three different data subsets: the training set with 50% of all of the records and the validation and test sets with 25% of the records, respectively, picking the sets as equally spaced points throughout the original data. Data were linearly scaled to a range that is appropriate for the transfer function used in the NNs. The normalization of data ensures that values of different input variables are in the same range of the transfer functions used and also avoids overflows due to very large or very small weights.

Six NNs were developed in MATLAB [1]: one MLP, one RBF, and one MNN for each of the modeled pollutants (i.e., CO and NO$_2$). A pruning approach was used to determine the optimal neuron number of hidden layers for the MLP and MNN, whereas for the RBF, the training algorithm automatically adds nodes into the hidden layer when needed until the actual error is below the given error. All six NNs consist of an input layer of 31 input variables and an output layer of one neuron corresponding to either CO or NO$_2$. The MLPs each have ten neurons in the hidden layer; the MNNs each have two hidden layers with 12 and six neurons, respectively; whereas the RBFs each have a hidden layer of 183 and 268 neurons for CO and NO$_2$, respectively. Our results showed that the use of more neurons or more hidden layers did not improve the performance of the NNs. The logistic and linear transfer functions were used in hidden and output layers, respectively, for all NNs. The learning algorithm was the back-error propagation for the MLP and MNN. The performance index used during the training process was the mean sum of squared errors between the estimated t_i and the actual values a_i, and N is the number of observations, according the following equation:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (t_i - a_i)^2. \tag{6}$$

After the NNs were trained and their performances were evaluated, the input variables were examined to see which of them is the most important input to the estimation of the roadside concentrations. The determination of appropriate input variables was carried out using the following sensitivity analysis: Let x_p be an input variable for $p = 1, \ldots, 31$; each of the input variables was increased by a certain percentage (e.g., 5%) in turn, and the change in the output caused by the increase in the input was calculated. The sensitivity of each input is given by

$$S(x_p) = 100 \times \frac{\Delta V_{\text{output}}}{\Delta V_{\text{input}}}. \tag{7}$$
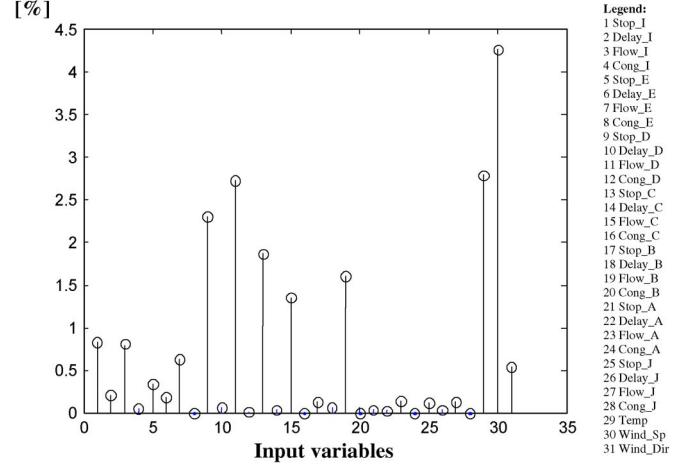

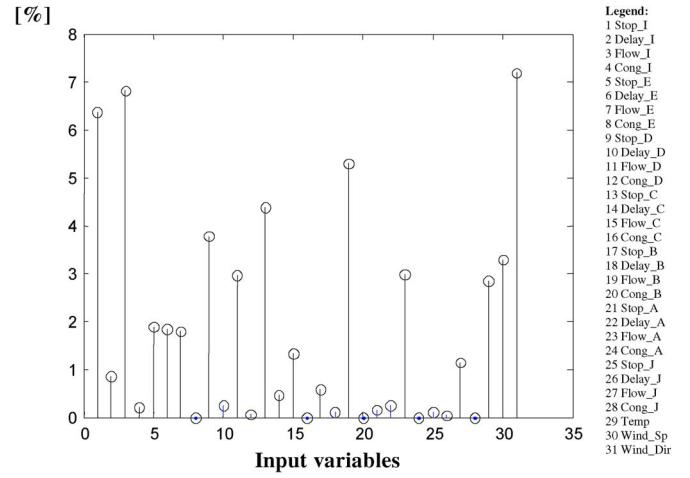Fig. 2. CO: Sensitivity analysis of the 31 input variables using the 2001 data set.

Fig. 3. NO$_2$: Sensitivity analysis of the 31 input variables using the 2001 data set.

The sensitivity analysis was used to study the response of the NN to small and equal increments of input variables. The conclusion from the sensitivity analysis was that if an increase in an input variable causes a significant change (either positively or negatively) to the output variable, this input variable is regarded as an important input and should be retained in the model.

V. RESULTS

A. Modeling of Roadside Concentrations

Following the initial training of the NNs, a sensitivity test was carried out on the individual input variables in turn to assess their contributions to the output (i.e., CO or NO$_2$) and, hence, decide whether they should be retained in the models. As can be seen in Figs. 2 and 3, many of the 31 input variables give low ($< 0.5\%$) contributions to the model outputs and were removed from the model. In particular, the congestion parameters for all the links hardly make contributions (i.e., a value of zero), whereas the meteorological condition parameters highlight a remarkable importance, as do vehicle stops and flow.

It also shows that links D and C are the most important inputs to the determination of pollutant concentrations. By removing

TABLE I
NN MODEL PERFORMANCE FOR 2001 (WITH 18 INPUTS)

	CO concentrations [ppm]				NO$_2$ concentrations [ppb]			
Model	MLP	RBF	MNN	LR	MLP	RBF	MNN	LR
RMSE	0.512	0.492	0. 556	0.807	9.600	9.100	10.053	12.671
R	0.774	0.776	0.721	0.656	0.651	0.685	0.603	0.538

Remarks: RMSE – root of the mean squared error, R – correlation coefficient and LR – linear regression.

TABLE II
NN MODEL PERFORMANCE FOR APRIL 2002 (WITH 18 INPUTS)

	CO concentrations [ppm]				NO$_2$ concentrations [ppb]			
Model	MLP	RBF	MNN	LR	MLP	RBF	MNN	LR
RMSE	0.373	0.350	0.390	0.406	11.778	11.741	11.953	12.097
R	0.651	0.671	0.630	0.566	0.442	0.445	0.420	0.405

Remarks: RMSE – root of the mean squared error, R – correlation coefficient and LR – linear regression.

the least important input variables from the models, 18 input variables were retained for further studies (in particular, STOP of links I, E, D, C, B, and J; DELAY of links I and E; FLOW of links I, E, D, C, B, A, and J; TEMP; WIND_SD; and WIND_DIR).

Table I shows the performance of the NNs compared by using the root-mean-square error (RMSE) and the correlation coefficient as statistical indexes of the goodness of fit between real and estimated pollutant concentrations. These statistical indexes have been calculated by taking into account the entire data set 2001 (hence, training, validation, and test sets).

It was concluded that the CO concentrations are more related to the traffic characteristics and the meteorological conditions than the NO$_2$ concentrations. These indexes highlight good performances for all neural nets; moreover, they give a measure of error on the roadside pollutant concentration estimates with respect to real values. The generalization capabilities of the NNs were tested using an independent data set collected in 2002. Table II shows the performance of the NNs calculated over the period (April 2002).

The lower performances are due to the fact that the 2002 data set was found to be corrupted by the SCOOT system faults. In fact, the traffic parameters of link E for all of 2002 were completely missing. This causes an error with estimated values and results in a low generalization capability of the NNs used. However, the achieved performances highlight that both MLP and RBF NNs still performed well, even with the training data having such a large number of missing values. Moreover, MNNs perform worse than MLPs and RBFs. This shows that the NNs can be used as a reliable tool to predict the roadside pollutant concentrations in real time, where missing data are inevitable.

In all cases, the RBF produced better prediction than the MLP and the MNN. The performances of MNNs are shown in the tables but not plotted in the figures because they are significantly lower than those of RBFs and MLPs. Figs. 4 and 5 show the comparison between actual and predicted CO concentrations by the RBF and MLP networks, respectively, considering a single month under observation (April 2001).

As can be seen, all NNs were able to predict well the CO levels that were less than 4 ppm but failed to predict the high CO concentrations. This may be because the intention of NNs, like many other statistical analysis techniques, is to model
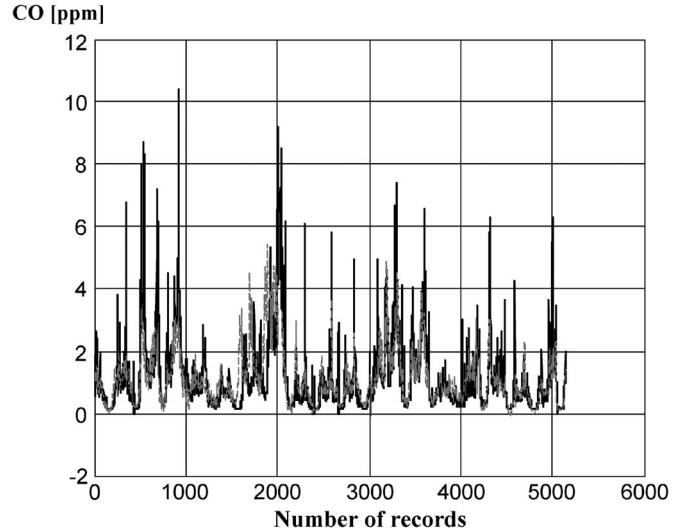

Fig. 4. RBF performance for April 2001. (Black line) Actual values. (Gray dashed line) Predicted values.
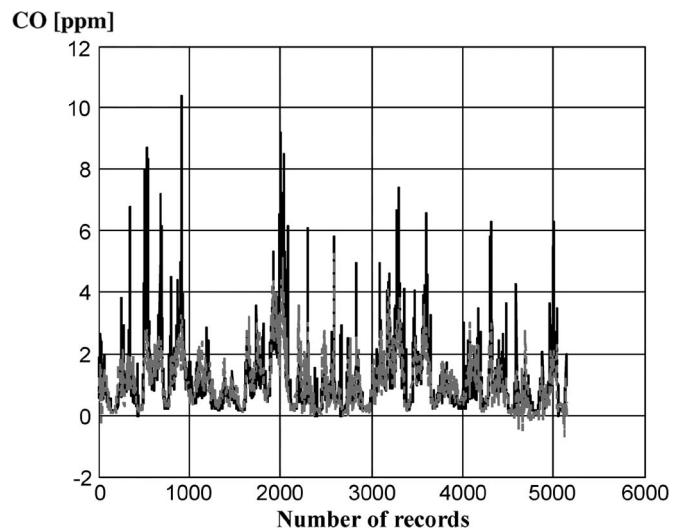

Fig. 5. MLP performance for April 2001. (Black line) Actual values. (Gray dashed line) Predicted values.

the average behavior of a process by assuming that data are normally distributed. Such an assumption is unlikely to be true with CO data.

TABLE III
NN MODEL PERFORMANCE FOR PEAK PERIODS

Model	CO concentrations [ppm]				NO₂ concentrations [ppb]			
	MLP	RBF	MNN	LR	MLP	RBF	MNN	LR
RMSE	0.494	0.278	0.545	0.903	12.777	12.641	12.779	13.543
R	0.783	0.979	0.761	0.641	0.361	0.397	0.360	0.331

Remarks: RMSE – root of the mean squared error, R – correlation coefficient and LR – linear regression.

TABLE IV
NN MODEL PERFORMANCE FOR OFF-PEAK PERIODS

Model	CO concentrations [ppm]				NO₂ concentrations [ppb]			
	MLP	RBF	MNN	LR	MLP	RBF	MNN	LR
RMSE	0.442	0.402	0.492	0.771	7.523	6.033	7.553	12.553
R	0.803	0.839	0.782	0.684	0.701	0.749	0.695	0.541

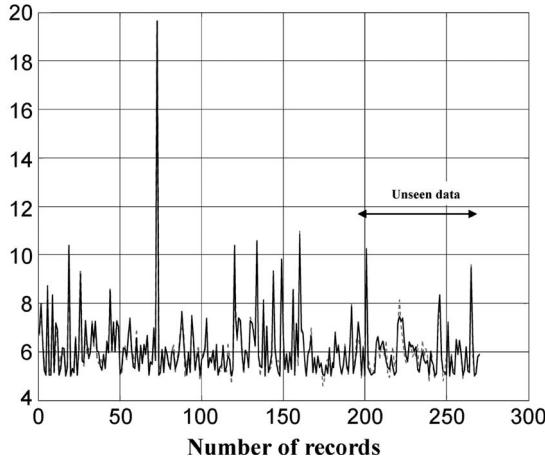Remarks: RMSE – root of the mean squared error, R – correlation coefficient and LR – linear regression.

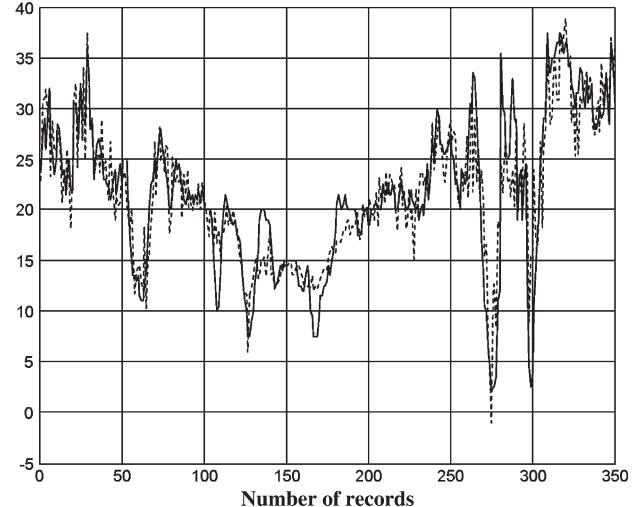Fig. 6. RBF performance for peak periods. (Black line) Actual values. (Gray dashed line) Predicted values.

Fig. 7. RBF performance for off-peak periods. (Black line) Actual values. (Gray dashed line) Predicted values.

To predict these extreme values to improve modeling performance, CO and NO₂ data sets were split into peak and off-peak periods (threshold values were 5 ppm for CO and 38 ppb for NO₂, taking into account short-term human heath effects [21]). For both pollutant concentrations, the AM and PM peaks during the average day occurred at about 08:20 and 16:40, respectively, reflecting their strong correlation with traffic peaks. Therefore, we retrained the neural models considering these four subdata sets, whose results are summarized in Tables III and IV.

It should be noted that we have an improvement in performance, reaching a correlation of 0.979 for CO and peak periods (Fig. 6), whereas for NO₂, we have an increase during off-peak periods (a correlation up to 0.75; Fig. 7) and a decrease of the correlation for peak periods. This different behavior is probably due to NO₂, which is a secondary pollutant. In fact, off-peak values are more correlated to traffic parameters than peak values, since the latter depend on other factors, such as ozone and direct solar radiation, which are responsible for the high levels of NO₂.

It is crucial to model pollutant extreme values to identify where and when they occur so that sufficiently accurate traffic demand management strategy models can be developed to mitigate their impact.

B. Prediction of Roadside Pollutant Concentrations

While recognizing the importance of knowing the relationship between traffic characteristics and roadside pollutant concentrations is useful, it is more powerful to use the relationship to forecast in real time the roadside pollutant concentrations in the future (e.g., in 5, 10, 15, and 20 min). The prediction abilities of the NNs were tested through a set of simulations using the current traffic and meteorological condition data (18 inputs) to predict the roadside pollutant concentrations in 5, 10, 15, 20 min, etc., in the future, using the same structure of NNs. Fig. 8 shows the forecasting performance of MLP NNs for the CO concentration. As can be seen, the indexes R and RMSE show the good prediction capability of the neural net. The forecast is less accurate when the forecast time lag increases, as expected. The decision makers can plan more effective traffic control strategies to decrease the acute traffic congestion phenomena and the resulting pollution by the use of a calibrated prediction tool.

C. Transferability Study of the Models

The main purpose of this paper is to study a generic approach that allows the NNs developed at the selected study site to
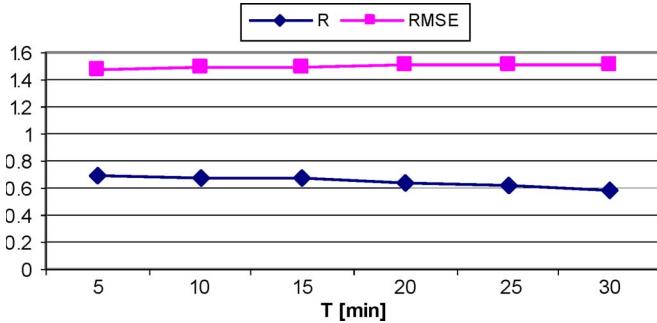
Fig. 8. CO: Forecasting performance. R is the coefficient of determination, and RMSE is the root-mean-square error.
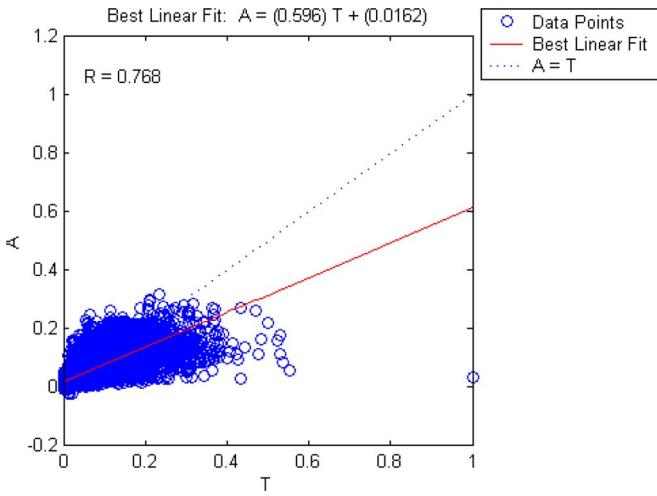

Fig. 9. CO: Goodness of fit between estimated values A and actual values T for 2001.

be transferred to other sites. The topologies of intersections in urban areas vary from one to another. For example, the most common four road intersections normally have eight traffic directions (four inbound and four outbound), whereas the intersection chosen for this paper has only seven traffic directions. To study the transferability of the models developed, it is hence necessary to aggregate all the flow, stops, and delay on inbound and outbound links, respectively, to produce six input variables. Together with the meteorological condition data (six traffic plus three meteorological parameters), nine input variables were used to retrain an NN with the data. Fig. 9 shows the goodness of fit of the network after retraining. The correlation coefficient and RMSE between actual and estimated values are 0.7682 and 0.5169, respectively, for CO, which are very similar to the performance of the NN with 18 input variables, as used earlier. This implies that link-based traffic data could be aggregated to generate network data to reduce the dimension of the input vector of the NN. Thus, the neural model with aggregated input requires less computational time than that with disaggregated input to calibrate the weight matrix. Furthermore, it is independent of the geometric characteristics of the road intersection with nine aggregated input parameters.

The trained NN was tested with unseen data from another road intersection and RPM on Narborough Road in Leicester, U.K., which a city in Leicestershire county that is so far away from Melton Mowbray that the traffic at these locations hardly
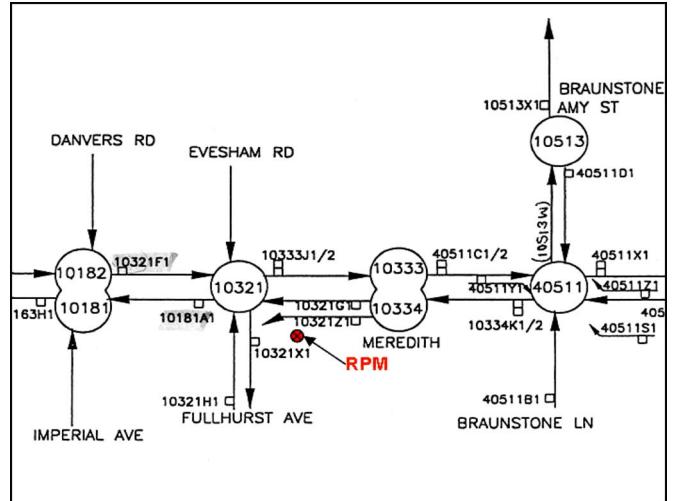

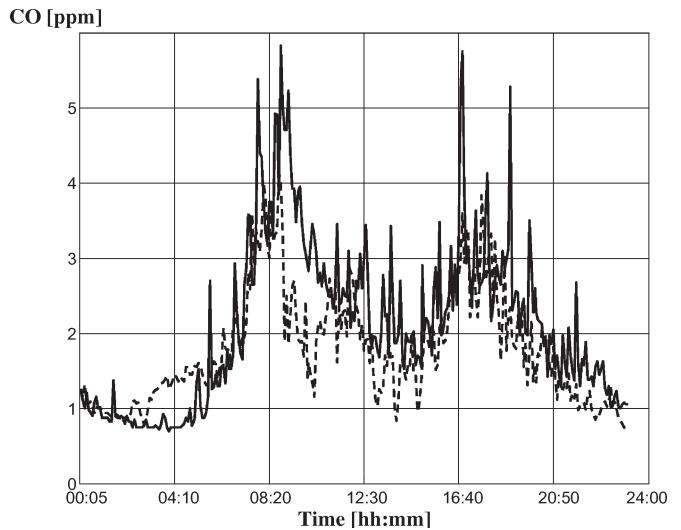
Fig. 10. Narborough Road.

Fig. 11. RBF performance on CO for Narborough Road. (Black line) Actual values. (Gray dashed line) Predicted values.

correlates. This road intersection, which is identified as 10321, has height road links (see Fig. 10). The test carried out gave a correlation coefficient and the RMSE between actual and predicted values of 0.631 and 0.6199, respectively. Fig. 11 shows the comparison between actual and predicted CO concentrations on the test site, highlighting that the approach used still produces a fair performance on transferability.

The difference between actual and predicted values on the test site (Fig. 11) is probably due to the different geometric characteristics of the test site (e.g., distance from RPM). Obviously, the neural model is not able to capture this change since it was trained on another site. However, the NN still performs well. Table V shows the performance of the three models on the test site. For NO_2 concentrations, the test on transferability with calibrated models did not provide significant outcomes since peak values depend on other weather factors (such as ozone and solar radiation), which are directly responsible for the high levels of NO_2. Clearly, this procedure has some limitations since the performance of the model depends on the RPM characteristics, which should be similar to

TABLE V
NN MODEL PERFORMANCE ON THE TEST SITE

Model	CO concentrations [ppm]				NO₂ concentrations [ppb]			
	MLP	RBF	MNN	LR	MLP	RBF	MNN	LR
RMSE	0.689	0.620	0.696	0.716	13.208	12.821	13.644	14.077
R	0.619	0.631	0.605	0.578	0.4101	0.423	0.401	0.391

Remarks: RMSE – root of the mean squared error, R – correlation coefficient and LR – linear regression.

the RPM characteristics of the training site. Nevertheless, by standardizing the geometric characteristics of the monitoring site (e.g., height and distances), this issue can be overcome.

VI. CONCLUSION

A comparison between three different kinds of NNs has been carried out, namely, MLPs, MNNs, and RBFs. The involved input variables are traffic parameters and meteorological condition data. It has been found that the RBF NN in all studied cases performs better than the MLP and the MNN. However, all kinds of NNs produced similar and good estimates of roadside pollutant concentrations. The results highlight how the RBFs can be an effective tool to estimate the pollutant concentrations with comparable performances to the MLPs widely used in the literature. In this paper, the MNNs highlighted a significantly lower performance than the RBFs and MLPs. It was shown that the RBF NN is able to explain about 97% of CO variations for peak periods and 75% of NO₂ variations for off-peak periods. It is conjectured that the remaining variations could be due to the absence of other important parameters (e.g., ozone and solar radiation) in our models. The RBF was shown to be more accurate in predicting pollutant extreme values than the MLP. Furthermore, the calibrated tool allows us to identify where and when extreme values occur so that sufficiently accurate traffic demand management strategy models can be developed to mitigate their impact. A sensitivity analysis has been carried out to determine which of the input variables were most important to the estimation of the pollutant concentrations. This shows that the weather plays the most important role in the formation process of the pollutant concentrations. Moreover, the analysis highlighted how the road links related to Nottingham Road (D) and Asfordby Road (C) were the most important inputs in the determination of pollutant concentrations. Among the traffic inputs, the congestion parameter has the least influence on the process since this parameter produces little effect on the pollutant concentration.

This paper examines the transferability in terms of both time and space. The study of the use of NNs for the short-term prediction of roadside pollutant concentration reveals that there is a fair correlation between current traffic and meteorological condition data and future pollutant concentration. The forecasting of future pollutant concentration (i.e., 15 min) in terms of current traffic and meteorological conditions is crucial since the decision maker can adopt real-time traffic control measures on links that are greatly responsible for the high levels of pollution near the intersection to relieve their effects.

Finally, the transferability of the trained NN has been tested with data from another road intersection in Leicester. The issue concerning various geometries of the intersection (character-ized by different numbers of links) can be overcome by aggregating all the flow, stops, and delays on inbound and outbound links. The results show that the NN still performs well. The results for CO concentrations are encouraging, although the input data set has a high background noise, which does not allow a very high performance of calibrated neural models. For NO₂ concentrations, the test on transferability with calibrated models did not provide significant outcomes since peak values highly depend on other weather factors (such as ozone and solar radiation), which are directly responsible for the high levels of NO₂.

Differences are probably due to the geometric characteristics of the test site. The results presented in this paper are a specific case but are encouraging for further application of the proposed methodology on an enhanced database.

REFERENCES

[1] *Neural Network Toolbox for Use With MATLAB, Reference Guide*, The MathWorks Inc., Natick, MA, 1995.

[2] M. C. Bell, R. G. Evans, S. A. Reynolds, R. E Boddy, and I. Hill, "An introductory guide to the instrumented City facility," *Traffic Eng. Control*, vol. 37, no. 12, pp. 698–703, Dec. 1996.

[3] C. M. Bishop, *Neural Network for Pattern Recognition*. Oxford, U.K.: Oxford Univ. Press, 1995.

[4] H. B. Celikoglu, "A dynamic network loading model for traffic dynamics modeling," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 4, pp. 575–583, Dec. 2007.

[5] B. Dorzdowicz, S. J. Benz *et al.*, *A Neural Network Based Model for the Analysis of Carbon Monoxide Concentration in the Urban Area of Rosario*. Southampton, U.K. Comput. Mech., 1997, pp. 677–685.

[6] M. W. Gadner and S. R. Dorling, "Neural network modelling and prediction of hourly NO$_x$ and NO₂ concentrations in urban air in London," *Atmos. Environ.*, vol. 33, no. 5, pp. 709–719, Feb. 1999.

[7] G. Grivas and A. Chaloulakou, "Artificial neural network models for prediction of PM₁₀ hourly concentrations, in the greater area of Athens, Greece," *Atmos. Environ.*, vol. 40, no. 7, pp. 1216–1229, Mar. 2006.

[8] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989.

[9] P. B. Hunt, D. I. Robertson, R. D. Bretherton, and R. I. Winton, *SCOOT—A Traffic Responsive Method of Co-ordinating Signals*. Crowthorne, U.K.: Transp. Road Res. Lab., 1982.

[10] R. Jacobs, M. Jordan, S. Nowlan, and G. Hinton, "Adaptive mixtures of local experts," *Neural Comput.*, vol. 3, no. 1, pp. 79–87, 1991.

[11] S. Kumarawadu and T. T. Lee, "Neuroadaptive combined lateral and longitudinal control of highway vehicles using RBF networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 500–512, Dec. 2006.

[12] W. Z. Lu, W. J. Wang, X. K. Wang, Z. B. Xu, and A. Y. T. Leung, "Using improved neural network model to analyze RSP, NO$_x$ and NO₂ levels in urban air in Mong Kok, Hong Kong," *Environ. Monit. Assess.*, vol. 87, no. 3, pp. 235–254, Sep. 2003.

[13] G. Marsden, M. Bell, and S. Reynolds, "Towards a real time microscopic emissions model," *Transp. Res. D*, vol. 6, no. 1, pp. 37–60, Jan. 2001.

[14] L. Moseholm, J. Silva, and T. Larson, "Forecasting carbon monoxide concentrations near a sheltered intersection using video traffic surveillance and neural networks," *Transp. Res. D*, vol. 1, no. 1, pp. 15–28, Sep. 1996.

[15] A. Pelliccioni and T. Tirabassi, "Air dispersion model and neural network: A new perspective for integrated models in the simulation of complex situations," *Environ. Model. Softw.*, vol. 21, no. 4, pp. 539–546, Apr. 2006.

[16] P. Perez and A. Trier, "Prediction of NO and NO_2 concentrations near a street with heavy traffic in Santiago, Chile," *Atmos. Environ.*, vol. 35, no. 10, pp. 1783–1789, Apr. 2001.

[17] T. Poggio and F. Girosi, "Networks for approximation and learning," *Proc. IEEE*, vol. 78, no. 9, pp. 1481–1497, Sep. 1990.

[18] J. P. Shi and R. M. Harrison, "Regression modelling of hourly NO_x and NO_2 concentrations in urban air in London," *Atmos. Environ.*, vol. 31, no. 24, pp. 4081–4094, Dec. 1997.

[19] D. Srinivasan, J. Xin, and L. C. Ruey, "Evaluation of adaptive neural network models for freeway incident detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 1, pp. 1–11, Mar. 2004.

[20] P. Viotti, G. Liuti, and P. Di Genova, "Atmospheric urban pollution: Applications of an artificial neural network (ANN) to the city of Perugia," *Ecol. Model.*, vol. 148, no. 1, pp. 27–46, Feb. 2002.

[21] *Air Quality Guidelines for Europe*, World Health Org., Copenhagen, Denmark, 2000, p. 288.

Haibo Chen received the B.Eng. degree in mechanical engineering from Central South University, Changsha, China, in 1984 and the Ph.D. degree in mechatronics from the University of Dundee, Dundee, U.K., in 1997.

He is a Principal Research Fellow with the Institute for Transport Studies, University of Leeds, Leeds, U.K., and has 12 years of experience in the broad field of traffic and environment modeling and analysis of ITS impacts and the development of real-time algorithms for traffic and pollution management applications.

Pietro Zito was born in Palermo, Italy, in 1973. He received the B.S. degree in electrical engineering and the Ph.D. degree in transport engineering from the University of Palermo, in 1999 and 2004, respectively.

During 2005–2006, he was a teacher of the under-degree course in intelligent transportation systems with the University of Palermo, where he is currently a Senior Research Fellow. His research has been mainly focused on transport policy and modeling.

Margaret C. Bell received the B.Sc. (Hons.) degree in physics and the Ph.D. degree in pure science (MCIT, MIHT) from the University of Durham, Durham, U.K., in 1970 and 1974, respectively.

She is a Science City Professor in transport and environment with the School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, U.K.

Dr. Bell was a recipient of the Commander of the Order of the British Empire (CBE) Award for services to sustainable transport on the Queen's 80th Birthday Honours List in 2006.