Agris on-line Papers in Economics and Informatics

Volume VI Number 2, 2014

A Neural Network Model for Forecasting CO_2 Emission

C. Gallo, F. Contò, M. Fiore

Department of Economics, University of Foggia, Italy

Abstract

Air pollution is today a serious problem, caused mainly by human activity. Classical methods are not considered able to efficiently model complex phenomena as meteorology and air pollution because, usually, they make approximations or too rigid schematisations. Our purpose is a more flexible architecture (artificial neural network model) to implement a short-term CO2 emission forecasting tool applied to the cereal sector in Apulia region – in Southern Italy - to determine how the introduction of cultural methods with less environmental impact acts on a possible pollution reduction.

Key words

Air pollution, CO_2 emission, artificial neural network, autoregressive, cereal sector.

Introduction

Air pollution is a critical element in environmental damage, mainly caused by anthropogenic factors. Human activities are influencing both air quality and climate change at regional and global scales. The rapid growth of population and the increase in energy demand are the main causes of the emission of large amounts of harmful pollutants and greenhouse gases, with serious consequences on health and environment. There are a lot of empirical evidences showing that the resolution of the problems of over-exploitation of natural resources goes through their replacement with forms of investment in intensive physical capital, and therefore through actions of innovation and technological diffusion (Contò, Prota, 2004). Substantial progress has been made over the past few decades to prevent and control air quality in many parts of the world through a combination of technology improvements and policy measures (Dentner et al., 2006; Esty et Al., 2005). Classical statistical approaches are not considered suitable to model such complex phenomena efficiently because they often lead to approximations or too rigid frameworks (Stanley, 1997; Hussain, Reynolds, 1995). The adoption of flexible architectures as artificial neural networks results in a more effective and rewarding approach in air quality control (Hornik et Al., 1989; Ojha, Singh, 2002). The optimization of the use of environmental resources also meets the dual aim of reducing the impact of economic activities and streamlines the company‘s costs. In this work we focus on the complex problem of controlling air pollution (particularly due to CO_2 emission) and mitigating its adverse effects on human health and the environment by proposing an artificial neural network model for forecasting CO_2 emission to help environmental policy decisions. The adopted model falls in the category of temporal linear stochastic models, in particular of the “non-linear auto-regressive moving average with exogenous inputs” type. The model was applied to a group of companies located in Southern Italy – Apulia region - participants in an Integrated Project of Food Chain (IPF ‘Oritalia’) which provided for the testing and introduction of forms of conservation agriculture in these cereal farms. Specifically, it is the introduction of the no-tillage farming techniques that allows a strong reduction of agronomic inputs, thus ensuring greater protection of agricultural soils. The model has been applied to the Apulian cereal sector to determine how the introduction of cultural methods with less impact (e.g. tillage) acts on a possible reduction of emissions. In the following sections we briefly expose the rationale behind the mathematical model adopted, followed by the illustration of the dataset of Apulian cereal sector CO_2 emissions and related methods of analysis and investigation. Finally results and conclusions are drawn.

Atmospheric dispersion models

The most important activity for air quality management consists in collecting environmental data and analysing these data by fitting it into a mathematical model that can be subsequently

used to predict air quality under different scenarios and also to quantify the health and environmental risks (Mori, 2000; Holmes, Morawska, 2006). Air quality modelling is a well-researched and fully developed area, and several works are available that deal with this subject (Nema and Tare, 1989; O'Neils et Al., 2008; Krzyzanowski, 2008; Ramanathan, Feng, 2008). Popular statistical approaches have been recently adopted to cope with these problems, leading to the use of specific techniques such as ANN, fuzzy logic, and time series analysis (Comrie, 1997; Haykin, 2001). In the absence of any rigorous modelling, however, time series can be used locally for useful forecasts of concentrations of air pollutants.

Analytical models

In its simplest representation, an emission source reaching the atmosphere may extend horizontally and vertically through a process called *dispersion*, and the shape drawn is roughly conical and called a *plume*. How it spreads will depend on the meteorological and process conditions prevailing in the atmospheric boundary layer. The complete equation for Gaussian dispersion modelling of continuous air pollution plumes (Beychock, 2005) is shown below:

$$S = \frac{Q}{2\pi u} \cdot \frac{f}{\sigma_y} \cdot \frac{g}{\sigma_z} \quad (1)$$

where:

f = crosswind dispersion parameter $= e^{\frac{-y^2}{2\sigma_y^2}}$;

g = vertical dispersion parameter = $g_1+g_2+g_3$;

g_1 = vertical dispersion with no reflections $= e^{\frac{-(z-H)^2}{2\sigma_z^2}}$;

g_2 = vertical dispersion for reflection from the ground $= e^{\frac{-(z+H)^2}{2\sigma_z^2}}$;

g_3 = vertical dispersion for reflection from an inversion aloft =

$$\sum_{m-1}^{\infty}\left[e^{\frac{-(z-H-2mL)^2}{2\sigma_z^2}} + e^{\frac{-(z+H+2mL)^2}{2\sigma_z^2}} + e^{\frac{-(z+H-2mL)^2}{2\sigma_z^2}} + e^{\frac{-(z-H+2mL)^2}{2\sigma_z^2}}\right]$$

S = concentration of emissions, in g/m^3, at any (x downwind, y crosswind, z above ground) meters located receptor;

Q = source pollutant emission rate, in g/sec;

u = horizontal wind velocity along the plume centreline, in m/sec;

H = height of emission plume centreline above ground level, in m;

σ_z = vertical standard deviation of the emission distribution, in m;

σ_y = horizontal standard deviation of the emission distribution, in m;

L = height from ground level to bottom of the inversion aloft, in m.

When wind is very low, S will tend to increase to infinity. To deal with such situations, the concept of puff modelling has evolved:

$$S = \frac{p}{15.75\sigma^2_{xy}\sigma_z} \cdot e^{-\frac{(x_c-x_r)^2}{2\sigma_{xy}^2}-\frac{(y_c-y_r)^2}{2\sigma_{xy}^2}} \cdot e^{-\frac{(z_c-z_r)^2}{2\sigma_z^2}} + \cdot e^{-\frac{(2z_{inv}-z_c)^2}{2\sigma_z^2}} \quad (2)$$

where x_c, y_c, and z_c are the coordinates of the center point of a puff, x_r, y_r, and z_r are the coordinates of the receptor (the concentration contribution of a single puff at a point), P is puff inventory, σ_{xy} is the puff dispersion parameter in horizontal direction, σ_z is the puff dispersion parameter in vertical direction, and z_{inv} is the height of the inversion lid. These two approaches are useful in understanding the nature of dispersion, where the plume model has been commonly used to analyse steady-state continuous gas diffusion and the puff model for calm wind conditions. By the way, these simplifications cannot be used with variable meteorological conditions and numerical solutions are the only practical possibility (Mori, 2000). Nema and Tare (1989) outline the need of computational simulation procedures to deal with situations of non homogeneous and unsteady conditions which also influence the atmospheric dispersion. Holmes and Morawska (2006) presented a detailed review of dispersion modelling, based on which we may sketch a broad classification of air pollution models, such as Box, Gaussian plume, Gaussian puff, Lagrangian, Eulerian, and Computational fluid dynamics (CFD) which all model dispersion on a local or regional-scale.

Statistical models

Deterministic air quality models have recently been used in support of emission regulation and policy decisions. Their applicability for air quality forecasting (but depending on the quality of inputs) has become useful due to the recent

advances in computational capacity. Comrie (1997) compared deterministic with statistical models in their performance, showing that the latter outperform deterministic models. Statistical models have been developed with the intention of predicting concentrations based strictly on observations. The main types of statistical models are regression-type models and neural network-type models (Gallo et al., 2013). The first are models that use correlations to infer information for future concentrations. In particular, auto-regressive models (Comrie, 1997) are widely used stochastic models for forecasting, providing excellent results when used for long-term prediction, while not satisfactory when used for short-term forecasts and in the presence of non-stationary situations. This can be superseded by the introduction of an exogenous variable in the model, which lowers the erratic component. The model is represented as follows:

$$X_t - c_1 X_{t-1} - c_2 X_{t-2} - \ldots - c_p X_{t-p} = a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} + \beta_1 u_{T_1} + \ldots + \beta_m u_{Tm} \quad (3)$$

where $\beta(.)$ is a polynomial of order m, u_{T_m} represents the $m-$th exogenous input and T_m is the time delay between the output and the m−th input. On the other hand, neural network models "learn" from previous experience, correlating tomorrow's forecast with the outcome of similar patterns and trends.

Artificial Neural Networks (ANN)

As previously indicated, the final objective of any modelling is to determine the concentrations of a given pollutant for a known set of input data. By training with some known results an artificial neural network model (ANN), you can perform this task much more easily. To understand the capabilities of ANN, we first begin with a brief introduction. Later, we show a particular forecasting model whose use is relevant to the subject of air quality modelling. Artificial Neural Networks represent an advanced machine learning computing approach extensively used in the field of pattern recognition, dynamic system prediction, control, and optimisation (Zhang, Patuwo, 1998). The network is made by computational units (artificial neurons) through which information processing occurs by passing signals (often binary or real values) through links connecting network nodes. Each connection link has an associated weight representing its strength, and each node typically applies a nonlinear transformation, called an *activation function*, to its input to determine its output value. A neural network has an architecture describing the nodes' connection pattern, its connection weights' computational technique, and the activation function. The nodes are arranged in layers, namely *input* layer, *hidden* layer, and *output* layer forming the so called MLP (Multi Layer Perception) architecture. The input layer receives the input variables and provides information to the network. The output layer consists of the values predicted by the network. In a typical feed-forward neural network the nodes of one layer are connected to the nodes of adjacent layers but not to the nodes of the same layer. The number of hidden layers and hidden neurons are determined empirically. The output of each node depends on the output of the nodes of the previous layer connected to it and the corresponding weights. Each input to the network form a multi-dimensional vector $X = x_i$, $i = 1,..., m$, with the corresponding weights represented as a matrix $W = [w_{ij}]$, $i = 1,..., m$, $j = 1,..., n$. The network output is a vector $Y_j = f(X \bullet W_j - \theta_j)$, $j = 1,..., n$, where θ_j is the "threshold" value, X is the input vector, Wj is the weight vector for the j-th output node and f is the activation function. In a feed-forward ANN the most commonly used is the sigmoid function $f(x) = 1/(1+e-x)$. The choice of the sigmoid function is due to three important factors:

- it is a continuous function and therefore is everywhere differentiable;
- it is a nonlinear function;
- it is limited to [0, 1] and this translates into benefits in terms of learning.

In order to train and validate the ANN model, the data are divided into *training* and *validation* datasets. The suitability of the ANN model derives from its ability to reproduce the validation dataset with reasonable accuracy. Neural networks are widely used when you have no evident mathematical relationship between the variables. So, in the area of air quality forecasting the ANN model can be applied being not known a priori the connections between the operating parameters. In a neural network, determining the number of input and output neurons depends on the particular problem, so the number of hidden layers, whose choice significantly impacts on the network's performance. Besides, the training error can be influenced by the scaling of output values (Ojha, Singh, 2002). Too few nodes can lead to a poor approximation, while too many nodes can over fit the training data. So an empirical procedure is generally applied to decide on the optimal architecture.

The proposed ANN forecasting model

The neural network model chosen for the estimation of the pollutant emission into the atmosphere has a MLP feed-forward structure, with Levemberg-Marquardt learning algorithm and sigmoid activation function. The Levenberg-Marquardt training algorithm has been adopted because the classic steepest descent approach has a rather slow convergence to an absolute minimum due to its use of the gradient. This algorithm also uses the information on the error function Hessian without computing it explicitly, being therefore particularly fast when the number of inputs is not high. In the chosen neural network model the main features are:

- the number of hidden layers, set to one, as it has been shown that a neural network with one hidden layer is able to better approximate any continuous function;
- the number of hidden neurons, set to 20 for empirical reasons, to avoid over fitting;
- the number of input data used for the training dataset, for the resulting network to correctly "generalise" after the validation phase.

In particular, each input record consists of the last three forecasts (y), the related moving average (ma). Moreover, the CO_2 values are emissions in tons.

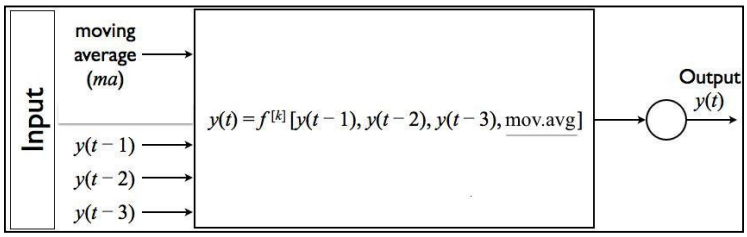
In this work, these parameters have been defined experimentally, by trial-and-error after various tests by minimisation of the mean square error indicator:

$$\mathrm{MSE}^{[k]} = \frac{1}{n}\sum_{j=1}^{n} (s_j^{[k]} - y_j^{[k]})^2 \quad (4)$$

where $s_j^{[k]}$ represents the desired j−th output value and $y_j^{[k]}$ is the obtained j−th output value for the k−th input pattern given to the network for training.

Results and discussion

In this work we have adopted a neural network model of "nonlinear auto-regressive with exogenous inputs moving average" type (see Fig. 1) in order to make predictions (Zhang, 1998) in the time domain of CO_2 pollutant inferred by the energy consumption of the companies under study. The reference database is composed of 220 cereal Apulian companies that are introducing, or have introduced, on part of the cultivation area less polluting methods with low energy content. We simulated different scenarios that could draw on emissions in cereal sector as a result of this transformation. The results of the most significant experiment follow in detail. The underlying mathematical model is shown in Fig. 1. The experimental results are detailed in Table 1, while the correlation between achieved outputs and desired targets are illustrated in Fig. 2.



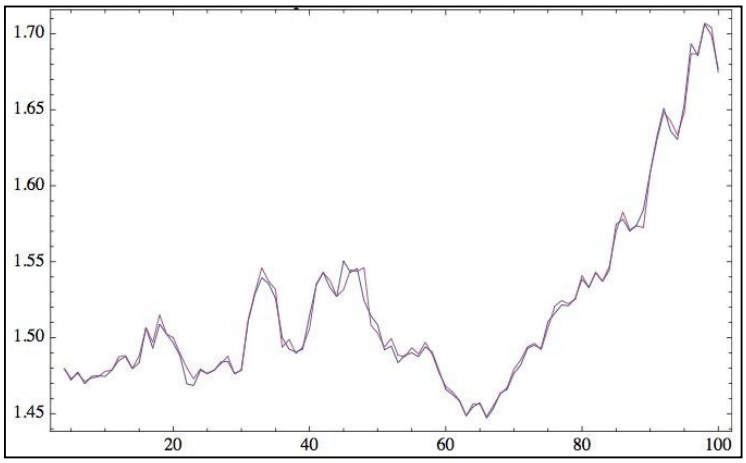
Source: our processing

Figure 1: The block diagram of the chosen network architecture.

Input	$CO_2(t-1)$, $CO_2(t-2)$, CO2($t-3$), ma[$CO_2(t-3)$, $CO_2(t-2)$, $CO_2(t-1)$]		
Target	$CO_2(t)$		
	%	# of samples	MSE
Training	70	154	3.433333333
Validation	30	66	3.841666667

Source: our processing

Table 1: Experimental results after the application of the prediction model.



Source: our processing

Figure 2: The network model's performance as a function of iteration.

Conclusions

Air quality control shows that pollution is mostly caused by anthropogenic factors. In this paper we introduce a model based on artificial neural networks to predict pollution's short-term rate for supporting environmental policy decisions. The adopted forecasting model is of non-linear auto-regressive type with exogenous input, and has demonstrated its validity for short-term previsions. The model provides good results (as to the MSE indicator) for estimating CO_2 pollution given by a sample of Apulian companies in the cereal sector. Its application shows that it is possible to make predictions up to three time periods. It is not possible to predict at a longer time period given the low correlation that exists between the outputs obtained and the desired targets. In addition, the adopted model is stable or time-independent, in the sense that the analysed phenomenon does not depend on the time but on a combination of meteorological factors present in that place and both the chemical and physical processes acting at that precise moment and in the moments before. The use of specific software designed and implemented for the management of cereal farms in all its aspects will monitor both the effects of optimisation of inputs in companies' life and in their financial statements and the environmental effects of the activities of these businesses. Future research trends will aim to investigate more thoroughly towards the interactions between environmental (exogenous) and company-related (endogenous) factors to better estimate the correlation between CO_2 emission and companies' size and operating parameters.

Corresponding author:
Mariantonietta Fiore
Department of Economics, University of Foggia, Largo Papa Giovanni Paolo II, 1 71121 Foggia, Italy
Phone: +39.0881.781806, fax: +39.0881.753711, E-mail: mariantonietta.fiore@unifg.it

References

[1] Beychok, M. R., Fundamentals of stack gas dispersion. Self-published. 2005, ISBN 0-9644588-0-2.

[2] Comrie, A. C. Comparing neural networks and regression models for ozone forecasting. Journal of the Air and Waste Management Association. 1997, 47, No.6, p. 653 – 663, ISSN 1047-3289.

[3] Contò, F., Prota, F. Politiche ambientali e incentivi per l'innovazione tecnologica. Rivista di diritto tribut intern. 2004, 2-3, p. 451 – 463, ISSN 1593-8441.

[4] Dentener, F. et al.. The global atmospheric environment for the next generation. Environmental Science & Technology, 2006, 40, p. 3586 – 3594, doi: 10.1021/es0523845.

[5] Esty, D. C., Levy, M., Srebotnjak, T., De Sherbinbin, A. Environmental Sustainability Index: Benchmarking National Environmental Stewardship. New Haven: Yale Center for Environmental Law & Policy, 2005.

[6] Gallo, C., Contò, F., La Sala, P., Antonazzo, A. P. A Neural Network Model for Classifying Olive Farms. Procedia Technology, 2013, 8, p. 593 – 599, ISSN 2212-0173.

[7] Haykin, S. Neural Networks: a comprehensive foundation. 2001, London: Prentice Hall. ISBN 0132733501.

[8] Holmes, N. S., Morawska, L. A review of dispersion modeling and its application to the dispersion of particles: An overview of different dispersion models available. Atmospheric Environment, 2006, 40, p. 5902 – 5928. ISSN 1352-2310.

[9] Hornik, K., Stinchcombe, M., White, H. Multilayer feedforward networks are universal approximators. Neural Networks, 1989, No. 2, p. 359 – 366. ISSN 0893-6080.

[10] Hussain A.K.M.F. and Reynolds W.C.. Measurements in fully developed turbulent channel flow. ASME. Journal of Fluids Engineering, Transactions of the ASME. 1975, 97, p. 568 – 580. ISSN 0098-2202.

[11] Krzyzanowski, M. WHO Air Quality Guidelines for Europe. Journal of Toxicology and Environmental Health, 2008, A, No.71, p. 47 – 50. ISSN 1528-7394.

[12] Mori, A. Integration of plume and puff diffusion models/application of CFD. Atmospheric Environment, 2000, 34, p. 45 – 49, ISSN 1352-2310.

[13] Nema, A., Tare, V., Atmospheric dispersion under non-homogeneous and unsteady conditions. Atmospheric Environment, 1989, 23, No.4, p. 851 – 856. ISSN 1352-2310.

[14] O'Neill, M. S., Kinney, P. L., Cohen, A. J. Environmental equity in air quality management: Local and international implications for human health and climate change. Journal of Toxicology and Environmental Health, 2008, A, 71, 9, p. 570 – 577. ISSN 1093-7404.

[15] Ojha, C. S. P., Singh, V. P. A chapter on ANN Modeling in Watershed Hydrology. In: Singh V. P., Frevert D. K., Editors. Mathematical Models of Water Shed Hydrology. Water Resources Publishers, Colorado, USA. ISSN 1084-0699.

[16] Ramanathan, V. Feng, Y. On avoiding dangerous anthropogenic interference with the climate system, Formidable challenges ahead, 2008,105, p. 14245 – 14250, doi 10.1073/pnas.0803838105.

[17] Stanley, H. E. Turbulence: The legacy of A.N. Kolmogorov. Journal of Statistical Physics, 1997, 88, No.1, p. 521 – 523, ISSN 0022-4715.

[18] Zhang, G. P., Patuwo, M. Y. H. Forecasting with artificial neural networks. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 1998,1, No.14, p. 35 – 62, doi:10.5194/isprsarchives-XL-5-W1-43-2013.