Prediction of CO Concentrations in Monterrey, Mexico, by means of ARIMA Models

CLAUDIO GUARNACCIA*, SIMONA MANCINI, JOSEPH QUARTIERI
Applied Physics Lab, Department of Civil Engineering, University of Salerno
Via Giovanni Paolo II 132, Fisciano (SA)
ITALY
*corresponding author: cguarnaccia@unisa.it

JULIA GRISELDA CERON BRETON, ROSA MARIA CERON BRETON
Facultad de Química, Universidad Autónoma del Carmen
Calle 56 Num. 4, 24180 Ciudad del Carmen, Campeche
MEXICO

Abstract: - Physical and chemical pollutions are a key problem in populated urban areas. The long term monitoring of air pollutants concentrations is a very helpful aid for policy maker, to control the exposure and to keep track of the slope of the data. When and where measurements are not possible, predictive models can help in these issues. Among the several possible techniques, "AutoRegressive Integrated Moving Average" (ARIMA) models are a good choice when a sufficiently large database of measurements is available. In this paper, the authors use the CO concentrations measured in San Nicolas de Garza, in the Metropolitan Area of Monterrey, Mexico, to calibrate and implement two different models. Both the models will provide reliable predictions on a short time range, since they use in input the data measured in close past periods. For this reason, the ARIMA models presented here can provide predictions to maximum 24 hours forward the last measured data. 24, in fact, is the lag that maximizes the autocorrelation of the data and thus it is the seasonality implemented in the models. Finally, the authors will present a validation (comparison with data not used in the calibration) of the models in four different days along the year, showing that the models are not affected by overfitting effects and the results are good also on data not used during the model parameters tuning.

Key-Words: - Environment, ARIMA models, pollution forecast, CO concentration

1 Introduction

The need to monitor, control and possibly predict the slope of pollutants in proximity of human settlements is essential nowadays. Both chemical and physical agents affect the health of people living and working in areas with high concentrations of these pollutants.

One of the most important atmospheric oxidants in urban areas is tropospheric ozone ([1], [2]), that is recognized to be the primary source of OH radicals and the third most important greenhouse gas behind CO2 and CH4. Beside the positive effect of blocking high ultraviolet irradiation in the stratosphere, ozone is considered responsible of several adverse effects on human health, vegetation and materials ([3-9]). Among its precursors, CO is very important in urban areas, since it is related to vehicle exhaust emissions and to several human activities, such as power station burning coal, combustion of fossil fuels, etc.. In addition, CO is toxic and can be the reason of many health affections. In United States, it has been estimated that more than 40000 people per year seek medical attention for carbon monoxide poisoning [10]. Carbon monoxide (CO), is an odorless and colorless gas that is produced by the incomplete combustion of carbon compounds, so it can be considered as a tracer of mobile sources (automotive) and industrial sources of combustion on a smaller scale. Some natural sources of carbon monoxide emission include forest fires or their emission from the natural processes that take place in the oceans. From the point of view of health, intramural sources due to their accumulation in homes due to domestic processes and tobacco smoke deserve special attention. The main potential harmful effect of this contaminant is its affinity to combine with hemoglobin, resulting in a high level of carboxyhemoglobin formation and, as a consequence, the amount of oxyhemoglobin decreases and, therefore, the entry of oxygen into the human body. The risk of exposure to CO varies depending on the concentrations and is higher in individuals suffering from circulatory deficiencies (being particularly susceptible patients with angina pectoris, as well as those with arteriosclerosis), and acute poisoning can occur by inhalation of this pollutant at high concentrations in intramural

environments. For these reasons, measurement campaigns and predictive models development are very important issues to understand the behaviour and trends of CO concentrations, in order to provide suggestions and alerts to policy makers.

In this paper, the "Time Series Analysis" (TSA) models based on "AutoRegressive Integrated Moving Average" techniques, both Seasonal (SARIMA) or not (ARIMA), are considered to predict the slope of CO concentrations in the Metropolitan Area of Monterrey (Mexico). These techniques [11-14] have been developed and implemented in several domains, such as Economics, Management, Physics, etc. (see for instance [15-20]). In [21-25], the TSA deterministic decomposition and ARIMA models have been applied to acoustical noise in urban areas and in proximity of an airport, also in combination with non-homogeneous Poisson distribution [26]. In [27], some of the authors applied the TSA techniques to electric consumption by public transportation in Sofia, Bulgaria. As for air pollution, in [28] the authors applied the TSA deterministic decomposition model to the same CO dataset used in this paper, that has been collected in Nuevo Leon, Monterrey, Mexico. The deterministic model showed to be able to capture the average slope of the CO concentration (with a very low mean error) but failed to predict the local variations and oscillations of the pollutant. The main advantage of that technique is the possibility to extend the prediction to any time in the future, since it provides the deterministic forecasting function, with constant coefficients and parameters. On the contrary, the Seasonal ARIMA (SARIMA) models presented in this paper will give very precise results but with the limitation that the prediction is one step ahead in the calibration and can be extended maximum to 24 periods forward in the validation.

2 Model Description

The ARIMA models implemented in this paper have been chosen according to the calibration dataset features, in order to optimize several variables and aspects, such as number of parameters, AIC and BIC, easiness of implementation, etc.. The two models presented in this section are seasonal, with a lag of 24 hours. The estimation of the coefficients is done in "*R*" software, in which also the validation can be run.

2.1 SARIMA(0,1,1)x(0,1,1)$_{24}$ model

The first model (Model 1) that we present and implement in this application is a Seasonal ARIMA model, without Auto Regressive coefficients, with seasonal lag equal to 24 hours (s = 24). According to the most used notation the model belongs to the SARIMA(0,1,1)x(0,1,1)$_{24}$ type. The model can be analytically formulated as follows:

$$Y_t = Y_{t-1} + Y_{t-24} - Y_{t-25} + \\ +\theta_1 e_{t-1} + \Theta_1 e_{t-24} + \theta_1\Theta_1 e_{t-25} + e_t$$

$$\hat{Y}_{t+1} = Y_t + Y_{t-23} - Y_{t-24} + \\ +\theta_1 \hat{e}_t + \Theta_1 \hat{e}_{t-23} + \theta_1\Theta_1 \hat{e}_{t-24}$$

where Y_t is the time series under study, $\hat{Y}_{t+1}$ is the prediction, θ and Θ are the coefficients related to the moving average, respectively with and without seasonality.

2.2 SARIMA(1,0,1)x(2,0,1)$_{24}$ model

The second model (Model 2) is again a Seasonal ARIMA model, with both Auto Regressive and Moving Average coefficients, with seasonal lag equal to 24 hours (s = 24). According to the standard notation, the model belongs to the SARIMA(1,0,1)x(2,0,1)$_{24}$ type. The model can be analytically formulated as follows:

$$Y_t = \phi_1(Y_{t-1} - \mu) + \Phi_1(Y_{t-24} - \mu) + \\ -\phi_1\Phi_1(Y_{t-25} - \mu) + \Phi_2(Y_{t-48} - \mu) + \\ -\phi_1\Phi_2(Y_{t-49} - \mu) + \mu + \theta_1 e_{t-1} + \Theta_1 e_{t-24} + \\ +\theta_1\Theta_1 e_{t-25} + e_t \quad (1)$$

$$\hat{Y}_{t+1} = \phi_1(Y_t - \mu) + \Phi_1(Y_{t-23} - \mu) + \\ -\phi_1\Phi_1(Y_{t-24} - \mu) + \Phi_2(Y_{t-47} - \mu) + \\ -\phi_1\Phi_2(Y_{t-48} - \mu) + \mu + \theta_1 e_t + \Theta_1 e_{t-23} + \\ +\theta_1\Theta_1 e_{t-24} \quad (2)$$

where Y_t , $\hat{Y}_{t+1}$, θ and Θ have the same meaning of above, ϕ and Φ are the coefficients related to the auto regression, respectively with and without seasonality, and μ is the intercept of the model (somehow related to the mean of the series).

2.3 Error metrics

In order to evaluate the performances of the models, both in the calibration and in the validation phase, error metrics can be calculated. In this paper, "Mean Percentage Error" (MPE) and "Coefficient of Variation of the Error" (CVE) are implemented and calculated, as well as the "Mean Absolute Scaled Error" (MASE) and the "Akaike's Information Criterion" (AIC).

The first quantitative metric gives a measurement of the error distortion, i.e. MPE is able to describe if the model overestimates or underestimates the actual data:

$$MPE = \frac{\sum_{t=1}^{n}\left(\frac{Y_t - \hat{Y}_t}{Y_t}\right)100}{n} \quad . \qquad (3)$$

CVE considers the variation from the reality in absolute value. In other words, it provides the error dispersion:

$$CVE = \frac{\sqrt{\frac{\sum_{t=1}^{n}(e_t)^2}{n-1}}}{\bar{Y}} \quad , \qquad (4)$$

where $\bar{Y}$ is the mean value of the actual data in the considered time range.

An effective measurement of forecast accuracy is also the "Mean Absolute Scaled Error" (MASE) [29]. The MASE for seasonal time series is computed according to the following formula:

$$MASE = \frac{1}{n}\sum_{t=1}^{n}\frac{|e_t|}{\frac{1}{n-k}\sum_{i=k+1}^{n}|Y_i - Y_{i-k}|} \quad . \qquad (8)$$

It is computed using a "naïve" model [30] in the denominator. In this application, the denominator is related to the difference between data observed at the period *t* and data observed *k* periods before, assuming that the period *t* can replicate the observed value at time *t-k*.

Considering that parameters of the models are estimated using the method of the likelihood maximization, also the Akaike's Information Criterion (AIC) is proposed to evaluate models performances.

3 Application and Results

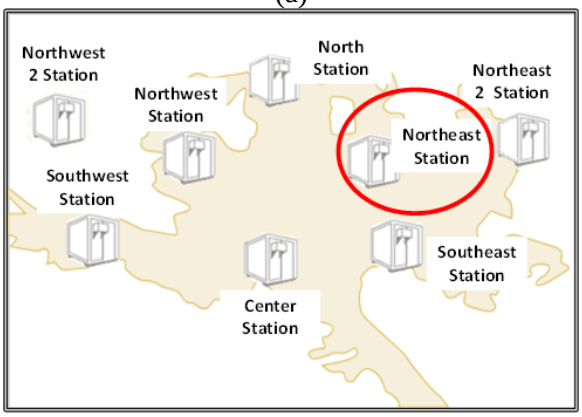
In this section, we present first the area under study, in which the field measurements have been performed, and the dataset used for the calibration. Then, the model implementation and parameters estimation are reported, together with the residuals evaluation. Finally, a validation on data not used in the calibration is performed.

3.1 Case study area and measurement description

The case study area is located in San Nicolas de Garza, one of the twelve municipalities of the Metropolitan Area of Monterrey (MAM), which constitutes the third largest urban area in Mexico (Figure 1). This city is located at 25°40'N and 100°18' W at 537 masl and it covers an area of 580.5 km^2. MAM has a population of about 4000000 habitants and it is an important area for several human activities. It is characterized by the presence of important education and research centers, business activities and industrial development. Road transportation and air sources (evaporative emissions from solvents, storage tanks, coatings, fuel marketing and other miscellaneous sources) are the dominant sources of O_3 precursors in MAM [31]. The specific sampling site was located within the facilities of Northeast Station of the SIMA, located in the Laboral Unity District in San Nicolas de los Garza, N.L. at 25° 43' 30 "N and 100° 18' 48" W at 500 m above sea level, within an area with high density of population.

More details on the area under study are reported in [28], in which meteorological features are resumed, together with the description of the field measurement campaign and instrumentation. The complete dataset, in fact, includes several pollutants such as O_3, NO_2, SO_2, PM_{10}, and meteorological conditions such as wind speed and direction, temperature, solar radiation and barometric pressure. In this paper, such as in [28], the authors focus on the time series of 8784 hourly CO concentrations observed in 2012, measured in ppm, that appeared to be the more interesting and the more suitable dataset for these applications.

(a)



(b)

Figure 1: (a) Area under study and (b) sampling site location in San Nicolas de Garza, Monterrey (Mexico) [28].

3.2 Models implementation and results

The models are calibrated on the complete dataset and, using the "*R*" software framework, the estimation of the parameters is performed. The estimated values of the models coefficients are reported in Table 1 and 2.

Table 1. Estimated value of the coefficients adopted by the SARIMA $(0,1,1)x(0,1,1)_{24}$ model (Model 1).

Coefficients	Estimated Value	Standard Error
MA1 (θ_1)	0.2104	0.0114
SMA1 (Θ_1)	-0.9453	0.0036
Log likelihood = -1449.45; AIC = 2904.91		

Table 2. Estimated values of the coefficients adopted by the SARIMA $(1,0,1)x(2,0,1)_{24}$ model (Model 2).

Coefficients	Estimated Value	Standard Error
AR1 (ϕ_1)	0.8151	0.0073
MA1 (θ_1)	0.2901	0.0114
SAR1 (Φ_1)	1.0381	0.0123
SAR2 (Φ_2)	-0.0426	0.0121
SMA1 (Θ_1)	-0.9415	0.0043
Intercept (μ)	0.6599	0.2017
Log likelihood = -1068.57; AIC = 2151.14		

Once the parameters are estimated, the analytical functions of the models can be implemented in any worksheet, to provide the results of the model in each period of the dataset. Let us underline that, since the models use in input some past data (see formulas above), the "predictions" can be obtained starting from a certain point of the dataset (period 26 for Model 1 and period 50 for Model 2) and can be extended in the future only if the inputs are available. For this reason, the "*R*" software extend the prediction outside the calibration dataset, for a maximum of 24 periods (i.e. the range fixed by the lag choice). These 24 data can be used for the "validation" of the model on data not used in the calibration process.

3.3 Calibration results

The main statistics of the observed data and of the models result in the calibration dataset are resumed in Table 3. It can be noticed that both the models achieve an excellent estimation of the mean, median and standard deviation of the dataset. Also skewness and kurtosis are very similar to the values of the calibration dataset, confirming that the predicted series well approximate the observed one.

These good results are confirmed calculating the residuals of the models, which are the difference between observed value and model predictions. The statistics of the residuals are reported in Table 4.

Figures 2 and 3 report the autocorrelation plot, the histogram and the Quantile-Quantile plot (Q-Q plot) for residuals calculated respectively with Model 1 and Model 2. Even if the two histograms are quite narrow and centred around zero residual, the two distributions differ significantly from the normal distribution, especially in the tails, as showed by the Q-Q plots. In addition, the autocorrelation plots show that Model 1 leaves a significant autocorrelation for a lag equal to 24 (and its multiples).

Let us underline that since the models have some negative terms, it can happen that a prediction is negative (see for instance Figure 4(a) below). If this is reasonable in cases in which the physical quantity under study can have negative values, in our case, for CO concentrations, this cannot be accepted. Anyway, Model 1 produced 98 negative forecasts (1.12% of the total calibration dataset), while Model 2 just 7 (0.08% of the total calibration dataset). The effect of these negative values on the residuals statistics is negligible. If we calculate mean and standard deviation removing negative values, the results are quite unchanged with respect to values reported in Table 4.

Table 5 reports the results of the error metrics defined in subsection 2.3. It can be noticed that Model 2 performs better than Model 1, except for the MPE.

Table 3. Summary statistics of the observed and predicted series.

	Mean [ppm]	Std dev [ppm]	Median [ppm]	Skew	Kurt
Observed	0.65	0.65	0.47	5.01	37.83
Model 1	0.64	0.68	0.47	4.08	34.17
Model 2	0.65	0.59	0.48	4.73	35.32

Table 4. Summary statistics of the residuals distribution evaluated on the calibration dataset for the two models.

	Mean [ppm]	Std dev [ppm]	Median [ppm]	Min [ppm]	Max [ppm]
Model 1	0.01	0.36	0.00	-3.86	5.47
Model 2	0.00	0.26	0.00	-3.15	4.47

Table 5. MPE, CVE, MASE (error metrics) and AIC values calculated in the calibration phase, for the two proposed models.

	MPE	CVE	MASE	AIC
Model 1	0.388	0.545	1.078	2904.91
Model 2	-4.283	0.402	0.880	2151.14

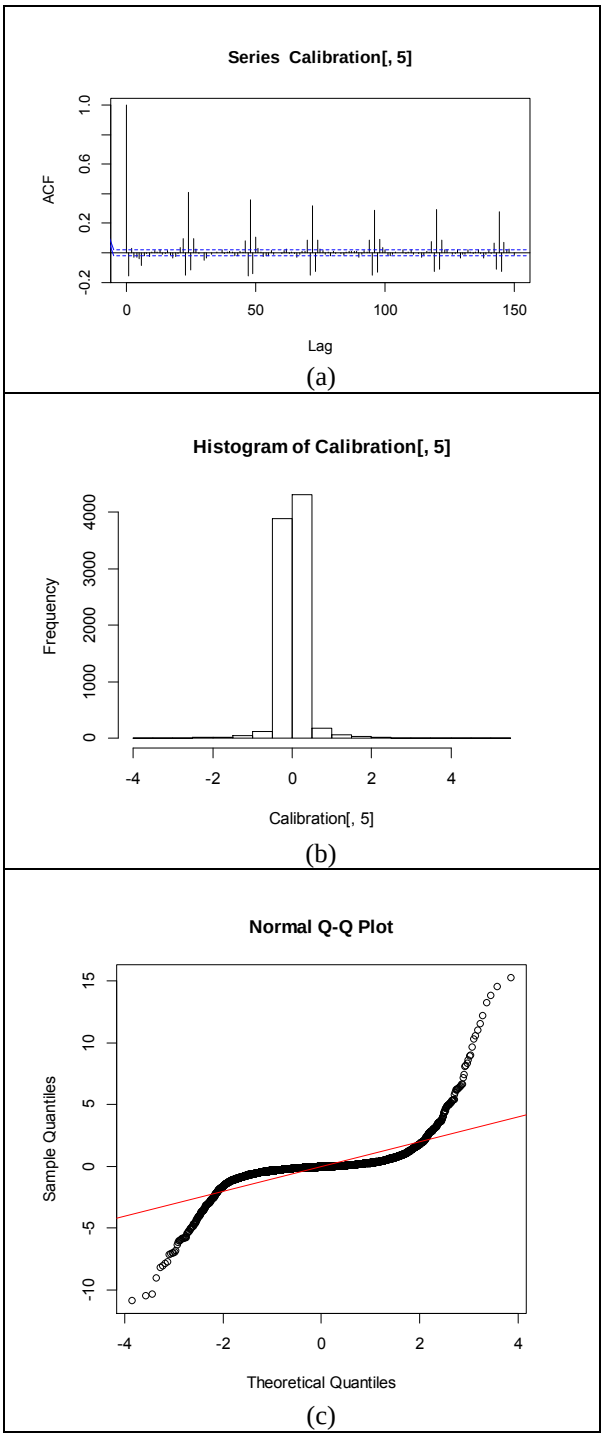


Figure 2: Residuals of the $(0,1,1)x(0,1,1)_{24}$ model applied to the calibration data: (a) correlogram plot; (b) histogram; (c) normal probability plot that describes the residuals behaviour compared to a normal distribution.

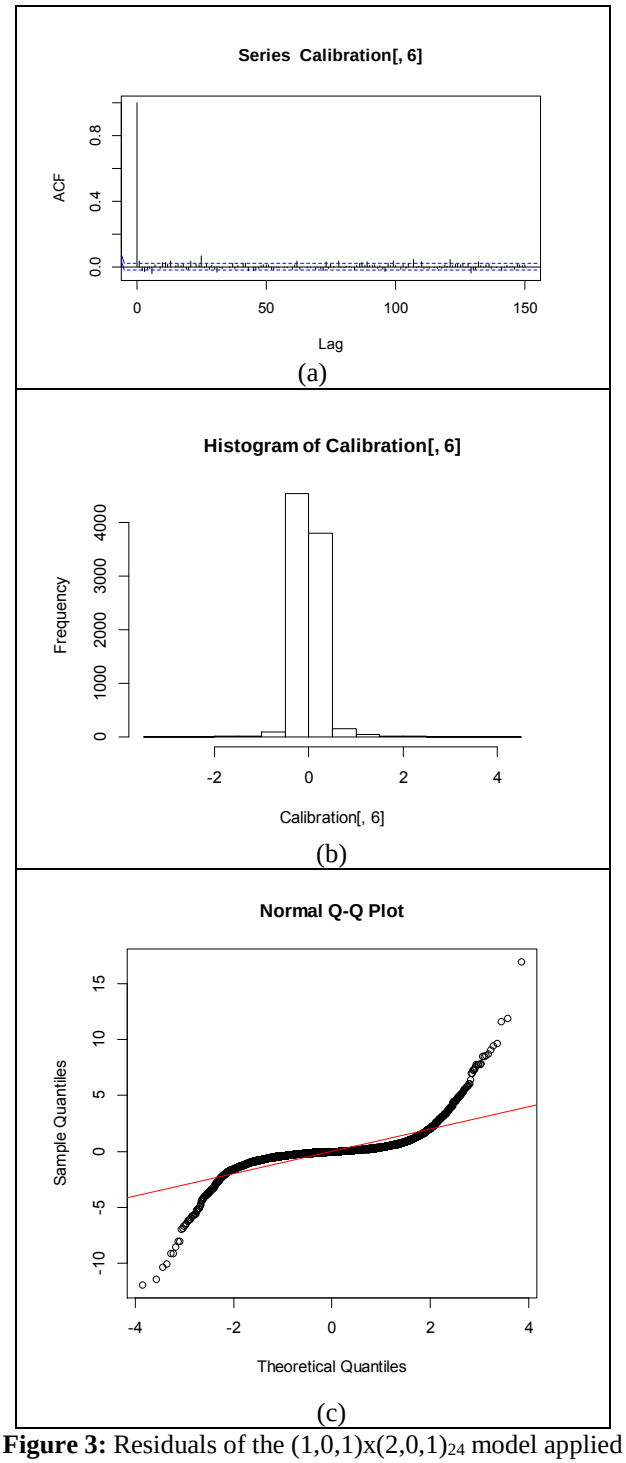


Figure 3: Residuals of the $(1,0,1)x(2,0,1)_{24}$ model applied to the calibration data: (a) correlogram plot; (b) histogram; (c) normal probability plot that describes the residuals behaviour compared to a normal distribution.

The superimposition of observed values and models results is reported in Fig. 4, for 24 hours in 4 different days of the calibration dataset (February 9, May 9, August 9 and November 9 2012). The good agreement between forecasts and observed values is evident. It can be observed that in some time ranges there is a kind of "one period delay", that is common when dealing with ARIMA models (see for instance [32]). This is due to the structure of the model, which takes one time step to adapt to changes in the series and, thus, can sometimes miss the sudden peaks of the physical quantity under study.

The choice to do not present the comparison in the entire calibration dataset was due to the difficulties in highlighting small variations in a so large dataset.

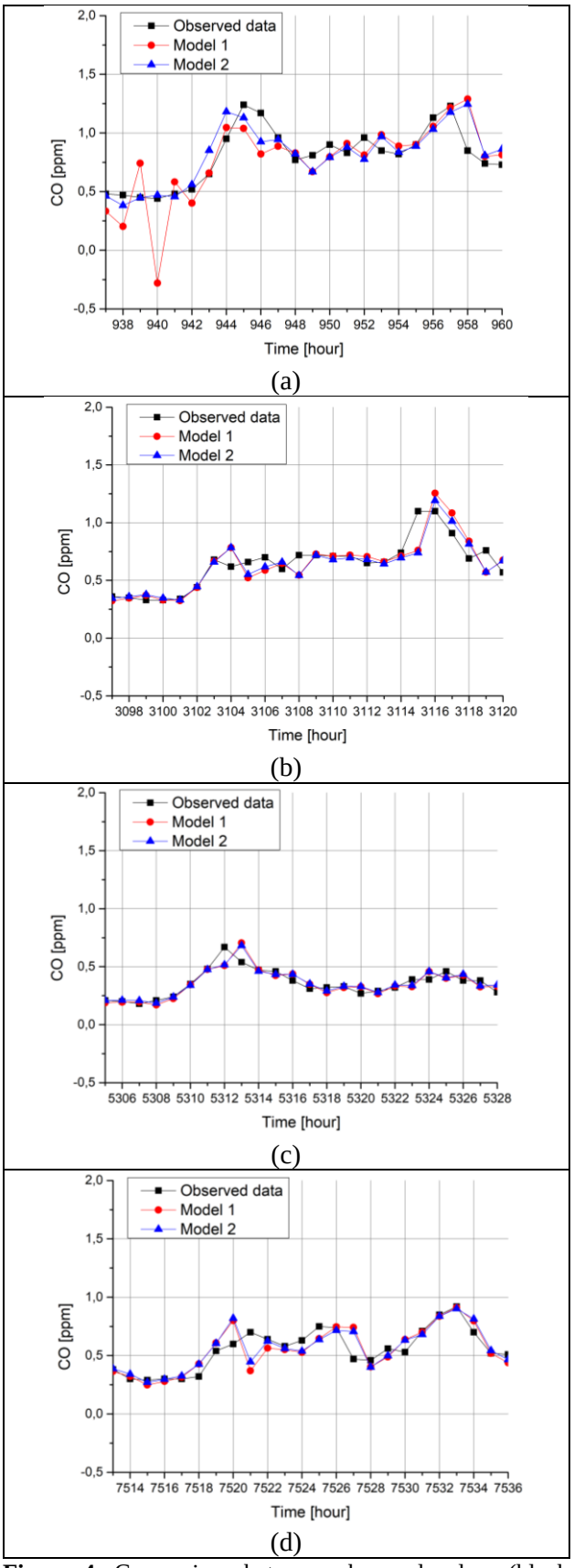


Figure 4: Comparison between observed values (black line and squares), Model 1 (red line and circles) and Model 2 (blue line and triangles) performed on (a) February 9 2012, (b) May 9 2012, (c) August 9 2012 and (d) November 9 2012.

3.4 Validation results

The validation of the models on a single day, i.e. a 24 hours measurement range not used in the calibration, is presented in this section, together with a validation on a larger period, i.e. one month, in two different parts of the year.

Since the previous calibration dataset ends on the 31st of December 2012, the validation should be done on January 1, 2013. Anyway, this procedure should be barely general, since the first day of the year is a special day, with celebration, fireworks, etc., that could affect the CO concentration.

For this reason, in order to perform a validation on an unbiased day, a new calibration has been performed on the first 5112 periods (up to the end of July 2012) for both the models. The models built with the new parameters have been tested on the 24 hours of August 1, 2012, that was a Wednesday. Results of the comparison are reported in Figure 5 and in Table 6.

It can be noticed that the average slope is well depicted by both the models. Again, as in [28], the single peaks are barely reproduced, since the parameters of the models are fixed at the last calibration period and cannot follow the very short term variations. In the calibration phase, this phenomenon was reduced because each period prediction uses the previous data, allowing to follow the possible high variability of the dataset.

The statistics of the errors (difference between measured and predicted values in the validation phase) reported in Table 6 and the error metrics, reported in Table 7, suggest that in this validation dataset, made of 24 periods, Model 1 performs slightly better than Model 2.

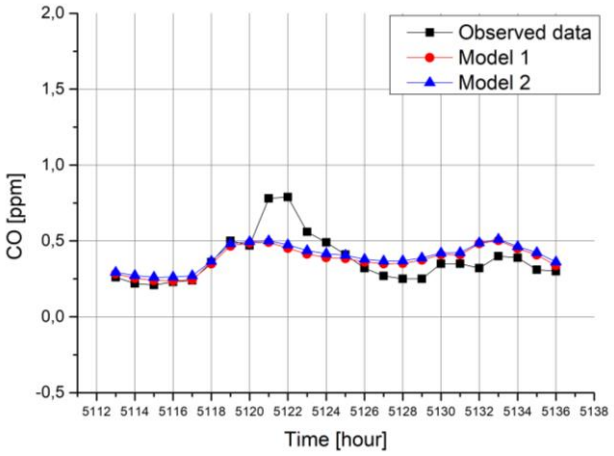


Figure 5: Comparison between observed values (black line and squares), Model 1 (red line and circles) and Model 2 (blue line and triangles) performed on August 1, 2012, after the new calibration on the first 5112 periods.

Table 6. Summary statistics of the errors evaluated on the 24 hourly data of the 1st of August for the two models.

	Mean [ppm]	Std dev [ppm]	Median [ppm]	Min [ppm]	Max [ppm]
Model 1	-0,004	0.119	-0.03	-0.16	0.34
Model 2	-0.021	0.117	-0.05	-0.17	0.32

Table 7. MPE, CVE, MASE (error metrics) calculated on the 24 hourly data of the 1st of August, for the two proposed models.

	MPE	CVE	MASE
Model 1	-8.593	0.317	1.235
Model 2	-13.718	0.317	1.323

A similar validation but on a larger dataset can be performed, in order to avoid local fluctuations of the measurement. In Table 8, the mean and the standard deviation of the errors evaluated in the entire months of August and December 2012 are reported. Of course, the model parameters are frozen at the end of July, i.e. at the end of the calibration dataset.

The results are very interesting, showing that the performances of the models decrease when moving from August to December, due to the distance between the calibration dataset and the validation range. Anyway, the distributions of the errors are still centred close to zero, even if with a standard deviation that is about 5 times the standard deviation observed in the validation performed on the August dataset.

Table 8. Mean and standard deviations of the errors evaluated on two validation datasets of 744 periods (one month) for both the models.

	August 2012 (data from 5113 to 5856)		December 2012 (data from 8041 to 8784)	
	Mean [ppm]	Std dev [ppm]	Mean [ppm]	Std dev [ppm]
Model 1	0.000	0.104	0.005	0.581
Model 2	0.000	0.097	0.037	0.561

4 Conclusions

In this paper, the authors present the application of two Seasonal AutoRegressive Integrated Moving Average (SARIMA) models are to CO concentrations in Metropolitan Area of Monterray, Mexico. The two models have been chosen according to the minimization of the AIC index. Both of them have been calibrated on a large dataset, made of 8784 hourly CO concentrations observed in 2012, in one of the monitoring stations of the site.

Once the parameters have been evaluated, the residuals (difference between observed value and models results, i.e. a simple predictive error metric) have been calculated and analysed. The results are very good for the selected models, both in terms of residuals mean and standard deviation. Also the graphical comparison showed very good predictive performances by both the models, suggesting the suitability of this predictive technique for CO concentration forecasts.

Then, in order to perform a validation, i.e. comparison of models results with data not used in the parameters tuning, a new calibration has been performed using the data from January to July 2012. This has been done to avoid a validation on a special day, such as the 1st of January 2013. Thus, the comparison between models predictions and observed CO concentrations in the 24 hours of the 1st of August 2012, has been performed. The results, in terms of mean error and standard deviation, are even better than in the calibration phase.

Anyway, the single day validation can be affected by local variabilities. For this reason, a validation on one-month data has been performed in the last part of the paper. The results showed a better average results of the mean error and standard deviation, but an expected worsening of the performances of the models moving from August to December, i.e. moving away from the calibration dataset.

Acknowledgments

This paper reports one of the last research topics pursued by Carmine Tepedino, who passed away on November 30, 2017, at the early age of 34. All the authors deeply miss his scientific and human great qualities. This paper is dedicated to his memory.

The authors are grateful to the local Government of San Nicolas de los Garza, Nuevo Leon, especially to the Integrated System of Environmental Monitoring (SIMA), for providing the air pollutants levels data measured in MAM used in this paper.

References:

[1] Van Eijkeren J.C., Freijer J.I., Van Bree L., A model for the effect on health of repeated exposure to ozone. *Environmental Modelling and Software*, Vol. 17, 2002, pp. 553-562.

[2] Xu J., Zhu Y., Some characteristics of ozone concentrations and their relations with meteorological factors in Shangai, *Atmospheric Environment*, Vol. 20, 1994, pp. 3387-3392.

[3] Lee D.S., Holland M.R., Falla W., The potential impact of ozone on materials in the U.K.,

Atmospheric Environment, Vol. 30, N. 7, 1996, pp. 1053-1065.

[4] Cass G.R., Nazarof W.W., Tiller C., Whitmore P.M., Protection of works of art from damage due to the atmospheric ozone, *Atmospheric Environment*, Vol. 25A, 1991, pp. 441-451.

[5] Wang S.W., Georgopoulos P.G., *Observational and mechanistic studies of tropospheric studies of ozone, precursor relations: photochemical models performance evaluation with case study*, Technical Report ORC-TR99-03, 2001.

[6] Cerón-Bretón J.G., Cerón-Bretón R.M., Guerra-Santos J.J., Córdova-Quiroz A.V., Vargas-Cáliz C., Aguilar-Bencomo L.G., Rodriguez-Heredia K., Bedolla-Zavala E., Pérez-Alonso J., Effects of simulated tropospheric ozone on soluble proteins and photosynthetic pigments levels of four woody species typical from the Mexican Humid Tropic, *WSEAS Transactions on Environment and Development*, Vol. 6, N. 5, 2010, pp. 335-344.

[7] Cerón-Bretón J.G., Cerón-Bretón R.M., Rangel-Marrón M., Vargas-Cáliz C., Aguilar-Bencomo L.G., Muriel-García M., Effects of simulated tropospheric ozone on foliar nutrients levels (Ca2+, Mn2+, Mg2+, and K+) of three woody species of high commercial value typical from Campeche, Mexico, *WSEAS Transactions on Environment and Development*, Vol. 6, N. 11, 2010, pp. 731-743.

[8] Aris R.M., Christian D., Hearne P.Q., Kerr K., Finkbeiner W.E., Balmes J.R., Ozone-induced airway inflammation in human subjects as determined by airway lavage and biopsy, *American Review of Respiratory Disease*, Vol. 148, 1993, pp. 1363-1372.

[9] Coleridge J.C., Coleridge H.M., Schelegle E.S., Green J.F., Acute inhalation of ozone stimulates bronchial C-fibers and rapidly adapting receptors in dogs, *Journal of Applied Physiology*, Vol. 74, 1993, pp. 2345-2352.

[10] Hampson N.B., Emergency Department visits for carbon monoxide poisoning, *Journal of Emergency Medicine*, Vol. 16, 1998, pp. 695-698.

[11] Box G. E. P., Jenkins G., *Time Series Analysis: Forecasting and Control*, Holden-Day, 1976.

[12] Chatfield C., *The Analysis of Time Series: an Introduction*, Chapman&Hall, New York, 1975.

[13] Box G. E. P., Pierce D. A., Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models, *Journal of the American Statistical Association*, Vol. 65, 1970, pp. 1509–1526.

[14] Cryer P.D., Chan K., *Time Series Analysis, with applications in R*, II Edition, Springer, 2008.

[15] Manganelli B., Tajani F., Macroeconomic variables and real estate in Italy and USA, *Italian Journal of Regional Science*, Vol. 14, N. 3, 2015, pp. 31-48.

[16] Manganelli B., Morano P., Tajani F., House Prices and Rents. The Italian experience, *WSEAS Transactions on Business and Economics*, Vol. 11, 2014, pp. 219-226.

[17] Morano P., Tajani F., Di Liddo F., Guarnaccia C., The Value Of the Energy Retrofit in the Italian Housing Market: Two Case-Studies Compared, *WSEAS Transactions on Business and Economics*, Vol. 15, 2018, pp. 249-258.

[18] Milanato D., *Demand Planning. Processi, metodologie e modelli matematici per la gestione della domanda commerciale*, Springer, Milano, 2008, in Italian.

[19] Chase R. B., Aquilano N. J., *Operations Management for Competitive Advantage*, Irwin Professional Pub, 10th edition, 2004.

[20] Di Matteo T., Aste T., Dacorogna M.M., Scaling behaviors in differently developed markets, *Physica A: Statistical Mechanics and its Applications*, Vol. 324, 2003, pp. 183-188.

[21] Guarnaccia C., Quartieri J., Mastorakis N.E., Tepedino C., Development and Application of a Time Series Predictive Model to Acoustical Noise Levels, *WSEAS Transactions on Systems*, Vol. 13, 2014, pp. 745-756.

[22] Guarnaccia C., Quartieri J., Rodrigues E.R., Tepedino C., Acoustical Noise Analysis and Prediction by means of Multiple Seasonality Time Series Model, *International Journal of Mathematical Models and Methods in Applied Sciences*, Vol. 8, 2014, pp 384-393.

[23] Guarnaccia C., Quartieri J., Tepedino C., *A Hybrid Predictive Model for Acoustic Noise in Urban Areas Based on Time Series Analysis and Artificial Neural Network*, in Proc. of the Int. Conf. on Applied Mathematics and Computer Science, AIP Conference Proceedings 1836, 020069, pp. 1-8, 2017.

[24] Guarnaccia C., Mastorakis N.E., Quartieri J., Tepedino C., Kaminaris S.D., *Development of Seasonal ARIMA Models for Traffic Noise Forecasting*, in Proc. Of the 21st Int. Conf. on Circuits, Systems, Communications and Computers, CSCC 2017, Heraklion, Crete, Greece. *MATEC Web of Conferences*, Vol. 125, 2017.

[25] Guarnaccia C., Quartieri J., Tepedino C., *Deterministic decomposition and seasonal ARIMA time series models applied to airport*

noise forecasting, in Proc. of the Int. Conf. on Applied Mathematics and Computer Science, AIP Conference Proceedings 1836, 020079, pp. 1-7, 2017.

[26] Guarnaccia C., Quartieri J., Tepedino C., Rodrigues E. R., A time series analysis and a non-homogeneous Poisson model with multiple change-points applied to acoustic data, *Applied Acoustics*, Vol. 114, 2016, pp. 203-212.

[27] Tepedino C., Guarnaccia C., Iliev S., Popova S., Quartieri J., A Forecasting Model Based on Time Series Analysis Applied to Electrical Energy Consumption, *International Journal of Mathematical Models and Methods in Applied Sciences*, Vol. 9, 2015, pp 432-445.

[28] Guarnaccia C., Cerón Bretón J.G., Quartieri J., Tepedino C., Cerón Bretón R.M., An Application of Time Series Analysis for Forecasting and Control of Carbon Monoxide Concentrations, *International Journal of Mathematical Models and Methods in Applied Sciences*, Vol. 8, 2014, pp 505-515.

[29] P.H. Franses, A note on the Mean Absolute Scaled Error, *International Journal of Forecasting*, Vol. 32, 2016, pp. 20–22.

[30] R.J. Hyndman, G. Athanasopoulos, Forecasting: principles and practice, OTexts, 2013.

[31] SEMARNAT-INE. 2005. Emissions Inventory for the Metropolitan Area of Monterrey.

[32] Cadenas E., Rivera W., Campos-Amezcua R., Heard C., Wind Speed Prediction Using a Univariate ARIMA Model and a Multivariate NARX Model, *Energies*, Vol. 9, 2016, pp. 1-15.