



Characterisation of errors in an urban scale atmospheric dispersion model through clustering of performance metrics

Andrea L. Pineda Rojas¹ · Rafael Borge² · Emilio Kropff³

Received: 29 October 2021 / Accepted: 9 December 2021 / Published online: 16 January 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

This work introduces a methodology to perform a comprehensive analysis of the performance of atmospheric dispersion models, using the urban scale DAUMOD-GRS model as a testing ground. The estimation of NO₂ concentration of this model at an urban background site in the city of Buenos Aires is compared with long-term observed concentration series. Days are classified through clustering analysis according to their model performance level on a three-dimensional metric space (fractional bias, normalised mean square error and correlation coefficient), and four clusters are obtained. An assessment of whether or not classes occur under different typical meteorological conditions is then performed. The largest statistical differences amongst clusters occur for wind speed and air temperature. The method is also used to assess the performance of a modified version of DAUMOD-GRS. The modification leads to a general improvement, mostly due to greater accuracy during night-time, which improves the ability of the model to estimate the NO₂ peak concentration values, occurring at early morning and late evening hours. This is an example of how clustering techniques can be used to identify specific conditions under which a model underperforms, understand its causes and test potential model improvements.

Keywords Urban scale air quality modelling · Clustering analysis · DAUMOD-GRS · Performance evaluation

Introduction

Given the increasing value of air quality (AQ) models in research and management (e.g. Miranda et al. 2015; Pisoni et al. 2019; East et al. 2021), gaining confidence on their results is crucial. This is achieved through model performance evaluation (Hanna 1988; Chang and Hanna 2005; Dennis et al. 2010; Derwent et al. 2010). Ideally, this assessment uses large AQ datasets from many monitoring sites (to capture the large spatial variability associated to that of emissions) and long-term series (to cover all possible combinations of meteorological and emission conditions). In a

typical statistical evaluation, metrics are applied over the complete data set and a single set of measures is obtained for each monitoring site (e.g. Pineda Rojas and Venegas 2013). However, when only a few AQ monitoring sites are available, this kind of assessment may not allow identifying particular features as variations in model error under different input data conditions. An alternative is to compute performance metrics for different ranges of model input variables. This approach has the following two drawbacks: (i) results can be sensitive to the selection of ranges, and (ii) different performance levels may be due to the occurrence of a combination of conditions rather than to a single variable. When large AQ series are available, this can be overcome using big data techniques, such as clustering analysis.

The urban scale atmospheric dispersion model DAUMOD-GRS (Dispersión Atmosférica Urbana - MODElo coupled with the Generic Reaction Set) (Pineda Rojas and Venegas 2013) has been satisfactorily tested against observations of nitrogen dioxide (NO₂) and ozone (O₃) from several short-term (a few weeks) AQ monitoring campaigns carried out in different sites of the metropolitan area of Buenos Aires (MABA, 3830 km²) (Pineda Rojas 2014). Long-term (several years) measurements recently made available by the

✉ Andrea L. Pineda Rojas
pineda@cima.fcen.uba.ar

¹ CIMA, UMI-IFAECI/CNRS, FCEyN, Universidad de Buenos Aires - UBA/CONICET, C1428EGA Buenos Aires, Argentina

² Laboratorio de Modelización Ambiental, Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, Madrid, Spain

³ Leloir Institute - IIBBA/CONICET, C1405BWE Buenos Aires, Argentina

local environmental protection agency (APRA, ‘Agencia de Protección Ambiental’ in Spanish) allow a more detailed model evaluation under a wide range of conditions. In a previous work (Pineda Rojas and Borge 2019), the model was statistically evaluated against 4 years (2009–2012) of NO₂ observations from three monitoring sites in the city. When data were pooled together and a single set of metrics was used for the evaluation at each site, results showed that the general performance of the model was good, with the best performance occurring at the urban background (UB) station. Some overestimation was obtained at this site during nocturnal hours, possibly affecting the modelled peaks of NO₂. In this work, we focus on the UB station and expand the former analysis to better understand the ability of the DAUMOD-GRS model to estimate NO₂ concentrations under different input variable combinations. A simple methodology based on clustering analysis is presented to classify days by model performance levels. The objective is to assess whether model uncertainty is uniformly distributed or concentrated in particular groups of conditions which can give clues on model behaviour features and thus future improvement options.

To exemplify its applicability, the method is used to evaluate the impact of a previously proposed model change on its performance and to identify conditions under which the modified version outperforms the standard one. The modification, suggested in a previous work (Pineda Rojas et al. 2019), consists in removing the memory effect (a contribution to the modelled concentration of the residual from the previous hour) of the model.

Methodology

DAUMOD-GRS is an atmospheric dispersion model that results from coupling the DAUMOD model (Mazzeo and Venegas 1991) and the GRS simplified photochemical scheme developed by Azzi et al. (1992). DAUMOD is based on the two-dimensional equation of diffusion (Arya 1999) and assumes stationary conditions and that there is no transport of pollutant through the upper boundary of the plume. Originally, it was developed to estimate urban background concentrations of primary pollutants emitted to the atmosphere from area sources. The GRS allows to estimate the concentrations of nitrogen dioxide (NO₂) and ozone (O₃) resulting from NO_x and VOC emission sources with only seven reactions. A detailed description of the DAUMOD-GRS model can be found in Pineda Rojas and Venegas (2013). DAUMOD-GRS requires considerably less input data to be operated than complex multi-scale photochemical models (e.g. CMAQ (U.S. EPA 2014)), and it allows long-term (several years), high spatial (1 km²) and temporal (1 h) resolution simulations at low computational cost. Hence, a

large number of model results can be obtained and analysed to better understand model performance using a wide range of input data conditions.

Simulation conditions

Simulations are performed in an 85 km × 75 km modelling domain including the metropolitan area of Buenos Aires (MABA). Model input data consist of 4 years (2009–2012) of hourly surface meteorological information from the station Aeroparque located at the domestic airport (see Fig. 1) and area source emissions of NO_x and VOCs from the high resolution (1 km², 1 h) inventory developed for the MABA by Venegas et al. (2011). The emission inventory includes a typical hourly profile but does not consider weekend or monthly variations. Given that the MABA is surrounded by non-urban areas, the model assumes clean air concentration levels for NO_x and VOC boundary conditions (Pineda Rojas and Venegas 2013). Whilst the regional background O₃ concentration could present both temporal and spatial variations, due to the lack of observations in the MABA surroundings, we assume a constant value of 20 ppb following the results of Mazzeo et al. (2005).

In this work, to exemplify the use of our analyses, we examine the performance of DAUMOD-GRS with a modification suggested in a previous study (Pineda Rojas et al. 2019). This modification consists in removing the memory effect of the model. When this feature is applied, the estimation of concentration at a given hour considers the influence from the previous hour. This element of the original model was later found to have a detrimental impact on the modelled peak O₃ concentrations in the MABA. Since O₃ and NO₂ are chemically coupled, this is expected to have an effect on predicted NO₂ concentrations as well. Results from Pineda Rojas et al. (2019) suggest that removing this “memory” from DAUMOD-GRS could improve night-time and early morning concentration predictions. To assess the impact of this modification on specific sets of input conditions, simulations with and without the memory effect are performed and their outcomes compared with observed data.

Air quality observations

The three APRA air quality stations are representative of urban background (Parque Centenario: CEN), urban traffic (Córdoba: COR) and residential industrial (La Boca: LB) sites. Pineda Rojas and Borge (2019) showed that the performance of the model to estimate hourly NO₂ concentration when considering the whole dataset is acceptable, and best performance metrics were obtained at CEN, as expected given that DAUMOD-GRS was developed to model urban background concentrations. The model presented a slight overestimation at CEN and some underestimation at the two

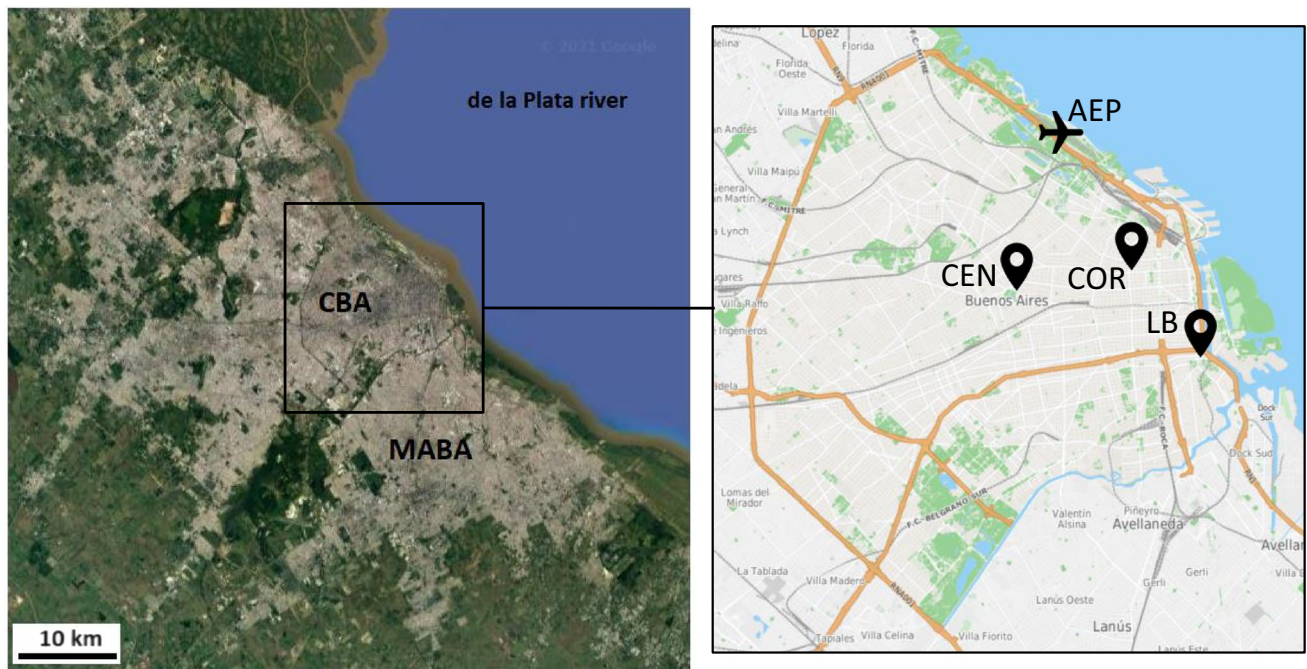


Fig. 1 Metropolitan area of Buenos Aires (MABA), including the city of Buenos Aires (CBA, 200 km²), the three air quality monitoring stations and the local airport. According to the air quality local authority (APRA), CEN (located close to a park in a commercial and residential area) is used as representative of urban background, COR

(located on one of the major traffic arteries of the city) as representative of urban traffic and LB (close to industries) as representative of residential/industrial areas. The local airport where meteorological data are obtained (AEP station) is indicated

other sites. An underestimation at COR is expected since this site is located in a street canyon where local effects that cannot be represented by the model make a contribution (Venegas et al. 2014). In turn, at LB, considerable underestimation when the wind comes from the N-ESE sector could be related to a non-negligible contribution from the power plants (point sources) that are located on the coast, which are not considered in the simulations (Pineda Rojas and Borge 2019). For these reasons, the present work focuses on the results obtained at the UB site (CEN: −34.60, −58.43) of the city of Buenos Aires.

Model performance metrics and cluster analysis

The fractional bias (FB), the normalised mean square error (NMSE) and the correlation coefficient (R) are widely used for statistical comparisons between modelled (C_m) and observed (C_o) concentrations (Chang and Hanna 2005) that can provide a fair representation of overall model performance. These metrics are computed for each day from C_m and C_o hourly values considering only days that have complete data (i.e. 24 hourly pairs of modelled and observed concentrations). The widely used unsupervised clustering algorithm *k*-means (MacQueen 1967) is used to find groups of days presenting similar model performance metrics (MATLAB function *k*-means). The silhouette criterion

(Rouseeuw 1987; Lleti et al. 2004; Kaufman and Rouseeuw 2009) is used to determine the optimal number (*k*) of clusters/groups by minimising the within-cluster vs between-cluster distance ratio (MATLAB function *silhouette*). As a reference, each cluster is assigned an index in increasing order according to

$$S = |\overline{FB}| + \overline{NMSE} + (1 - |\overline{R}|) \quad (1)$$

where the vertical bars denote absolute value and the over bar indicates the average overall members of the cluster. Roughly, this criterion orders clusters from better (cluster 1) to worse (cluster 4) performing. Although the indexing is somewhat arbitrary, it has no consequences on the results or the conclusions drawn from them.

Once days are labelled (i.e. grouped according to their performance levels), the values of model input data variables (wind speed (WS), wind direction (WD), PGT atmospheric stability class (KST) (varying from 1-extremely unstable to 6-moderately stable), air temperature (T), sky cover (SC) and total solar radiation (TSR)) are analysed in order to identify whether different model performance levels are associated with distinct patterns of input data conditions. In this work, such analysis is performed by comparing the distribution of each variable across clusters through Kruskal–Wallis

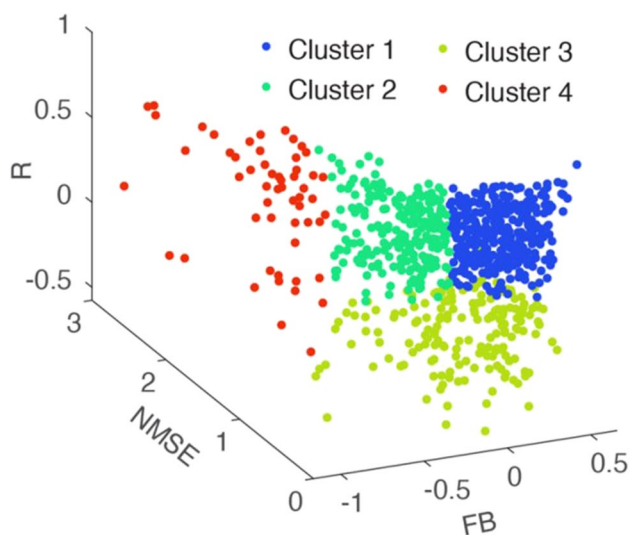


Fig. 2 Distributions of clustered days in the (3D) metric space. Clusters are ordered from best (#1, blue) to worst model performance (#4, red)

tests, Tukey–Kramer post-hoc multiple comparison analyses and bivariate polar plots (e.g. Carslaw 2018).

Finally, the impact of removing the memory effect from the model on the performance metrics is assessed using the clusters obtained from the standard simulation. By maintaining the ranking of days according to their performance in the standard run, a given labelled day (e.g. one belonging to cluster 4, coloured red) will have different metric values (or coordinates in metric space) in another simulation. In this way, the displacement of cluster points in the metric space (given by *FB*, *NMSE* and *R*) can be used to identify conditions under which a given model change improves performance.

A similar approach was developed for the three AQ sites in Pineda Rojas and Kropff (2021). Results from that work showed that this method produces different clusters at the three sites, suggesting significantly different outcomes on the conditions leading to worse model performance. In particular, the underestimation of concentrations for all clusters at the LB site, occurring specifically with ESE winds, supports the potential non-negligible contribution of power plants that were not included in the assessment. The present

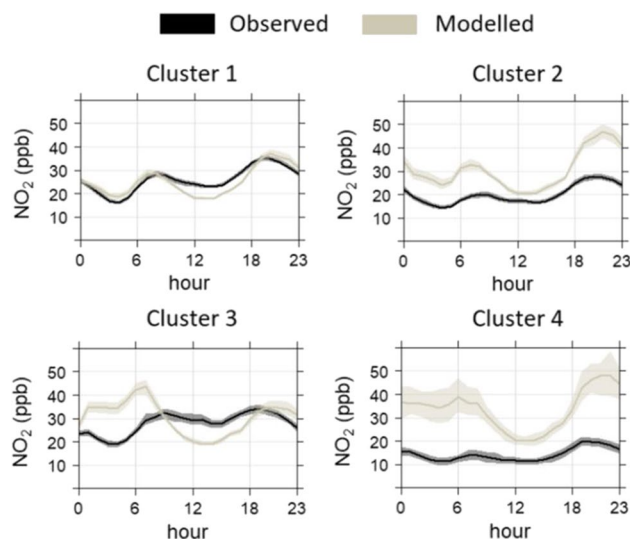


Fig. 3 Mean hourly variations of observed and modelled NO_2 concentrations by cluster. Shaded areas indicate 95% confidence interval in the mean

work focuses and expands the results obtained at the urban background site (CEN) where the model performs best.

Results

Clustering of model performance metrics

Applying the silhouette criterion as described in the section “[Model performance metrics and cluster analysis](#)”, an optimal value of $k=4$ is found. Figure 2 shows the distribution of days in the performance metric space given by *FB*, *NMSE* and *R*, and Table 1 their cluster-averaged values. Cluster 1 has the best performance for all metrics. Cluster 2 presents intermediate metric values, with better *R* but larger model-observed concentration differences (*FB* and *NMSE*) compared to the overall model performance. Whilst cluster 3 has lower performance levels exclusively in terms of *R*, and cluster 4 presents worse values of *FB* and *NMSE*, both of them perform similar to the average in terms of the other metrics (see Table 1).

Table 1 Model performance metrics (dimensionless) and their standard deviation values (in brackets) obtained at CEN station, considering the whole dataset (2009–2012) and the classification of days shown in Fig. 2

	All data	Cluster			
		1	2	3	4
<i>N</i>	18,912	7,800	5,544	4,248	1,320
<i>FB</i>	−0.198 (0.312)	0.019 (0.177)	−0.426 (0.127)	−0.102 (0.257)	−0.821 (0.135)
<i>NMSE</i>	0.374 (0.374)	0.144 (0.089)	0.457 (0.214)	0.363 (0.222)	1.413 (0.435)
<i>R</i>	0.476 (0.322)	0.638 (0.166)	0.638 (0.183)	0.002 (0.207)	0.526 (0.320)

N number of hours, *FB* fractional bias, *NMSE* normalised mean square error and *R* correlation coefficient

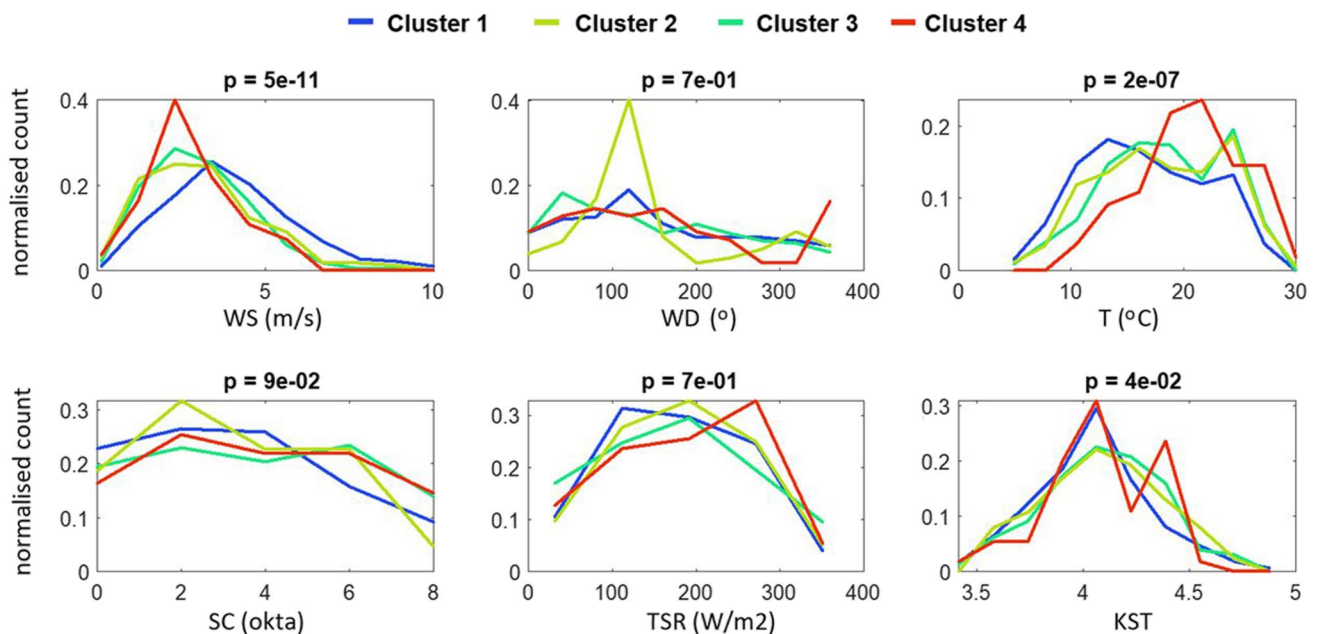


Fig. 4 Distributions of daily mean meteorological variables by cluster. The largest statistical difference amongst clusters is indicated with the p -value (Kruskal–Wallis)

The hourly mean observed and modelled NO_2 concentrations in each cluster are shown in Fig. 3. As expected, days included in cluster 1 show the best representation of the observed daily profile. Differences between observed and modelled values in days belonging to clusters 2 and 4 are larger at night. Figure 3 also shows that these two clusters present relatively lower observed NO_2 concentration levels.

Whilst clusters are defined based on multidimensional metric values, an assessment of whether or not this results in a separation of clusters in the space of input data conditions is performed. This unbiased approach is important because it has the potential to highlight conditions under which the model underperforms in a stereotypic way.

Differences between clusters in meteorological conditions

The distributions of daily mean values of meteorological variables for each cluster are shown in Fig. 4. To understand if observed differences are significant, a Kruskal–Wallis test ($\alpha: 0.01$) is used for each meteorological variable, against the null hypothesis that the median value of the variable for all clusters is the same. Due to its circular nature, for WD, a multi-sample test for equal median directions is used instead (Berens 2009). Significant differences are obtained exclusively for wind speed (WS) and air temperature (T). For WS, a Tukey–Kramer post-hoc multicomparison test indicates that the effect is due to the fact that the median for cluster 1 is significantly higher than the medians for all other clusters.

The same procedure applied to T indicates that significant differences exist between all pairs of clusters with the exception of the comparison between clusters 2 and 3. In general, worse performing days occur with relatively lower WS and higher T. Whilst the poorer performance under low WS has been observed using traditional analyses; errors associated to high T are seen with this method for the first time.

Next, hourly modelled (C_m) and observed (C_o) concentrations are compared by plotting the mean discrimination index $(C_m - C_o)/(C_m + C_o)$ (with values between -1 and 1) versus combinations of meteorological variables in bivariate polar plots for each cluster (Fig. 5). In these polar plots, the angle represents the direction of the wind and the radius represents different variables: WS, T, SC, TSR and H (hour of the day). Red areas indicate conditions for which the model tends to overestimate concentrations, whilst black areas indicate a tendency to underestimate them. The largest positive discrimination indexes are concentrated in clusters 2 and 4, where overestimation of observed NO_2 concentrations occurs for almost all combinations of meteorological variables. This suggests that NO_x emissions could be overestimated during days belonging to those clusters. An analysis (not shown) of the distribution of days by season and day of week for each cluster confirms that these clusters present larger fractions of summer and weekend days than clusters 1 and 3. In particular, cluster 4 has the largest frequency of both groups of days (35% and 51%, respectively). Since the inventory of emissions does not include corrections for weekends and summer holidays, this could in part explain

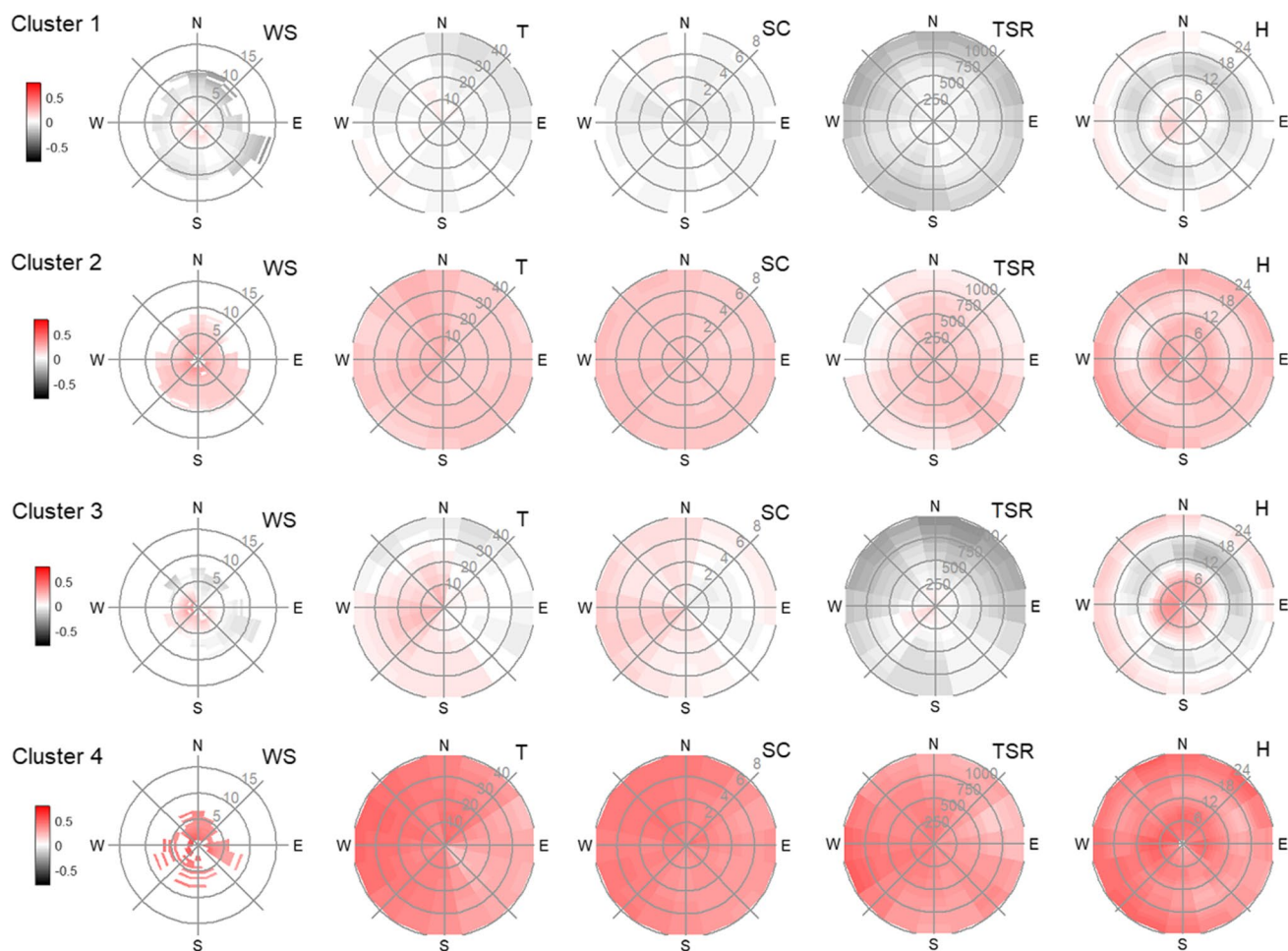


Fig. 5 Bivariate polar plots of the mean discrimination index $(C_m - C_o)/(C_m + C_o)$ for each cluster (variables in the radial axis: wind speed (WS, m/s), temperature (T, °C), sky cover (SC, okta), total solar radiation (TSR, W/m^2) and hour (H))

the association between poor performance and high temperature shown in Fig. 4. When removing days belonging to cluster 4, only a slight performance improvement is obtained in global FB (changing its value from -0.198 to -0.151) and NMSE (from 0.374 to 0.296). This relatively small impact on the overall model performance is probably due to the fact that cluster 4 includes fewer days than other clusters (see Table 1) and highlights the importance of fine-grain analyses to understand possible causes of specific types of model underperformance. Cluster 3 shows a mild variation of the index with WD. Overestimation tends to occur during night hours with low intensity winds from the 3rd and 4th quadrants, whilst some underestimation is observed for large total solar radiation (TSR) values and winds from the 1st and 4th quadrants. This could be due to reasons such as variations in the regional background ozone concentration, which in the model is assumed to be constant. However, the overall differences between C_m and C_o are small and tend

to disappear under winds from the ESE sector, where more than half of the data points in this cluster lie (Fig. 4).

Impact of proposed model change on model performance

In this section, the performance of the DAUMOD-GRS model when removing the memory effect is assessed using the same clustering (classification of days) previously discussed. This allows to identify whether this model version brings about performance changes under specific conditions. Figure 6 shows the distributions of days (colour points) by cluster in different projections of the multidimensional metric space. This is done for the following two simulations: (a) the standard run and (b) a new simulation without the memory effect (i.e. information from the residual pollutant concentration from the previous hour is no longer available). By keeping the classification of the

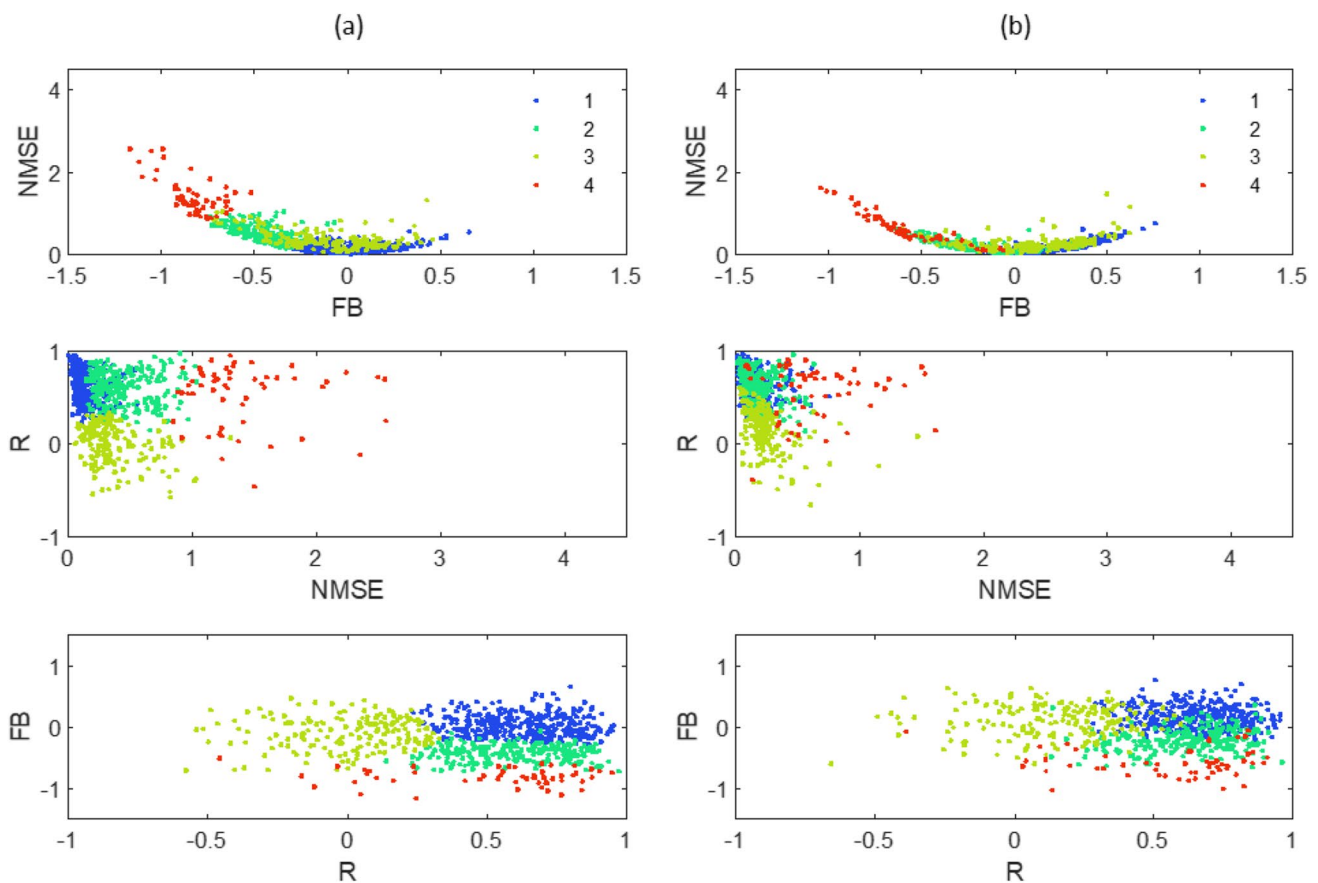


Fig. 6 Distributions of clustered days in the (2D) metric's planes for: **a** the standard run and **b** without the memory effect of the model, using the same classification of days as in **a** (i.e. each cluster groups the same days in the two figures **a** and **b**)

standard simulation, it is possible to observe the displacement of each cluster due to the proposed model change. A small general improvement for all data points is observed as a result of the modification. Box plots of the distribution of metric values for each cluster (Fig. 7) show that the largest improvement occurs in days belonging to clusters 2 and 4 (i.e. those with the largest overestimation). This is also evident in the average hourly NO_2 concentration profiles (Fig. 8) where the night-time overestimation is greatly reduced with little impact on the diurnal values. This result is consistent with those from Pineda Rojas et al. (2019) showing that the memory effect has a larger impact on modelled ozone concentrations under stable conditions. However, considerable differences between modelled and observed values still persist in cluster 4 which presents the lowest observed NO_2 concentration levels. This reinforces the need of improving the emission estimates by including monthly and weekly variations. Finally, Fig. 9 shows that the scatter plot of modelled vs observed daily maximum NO_2 concentrations around the line $C_o = C_m$ is largely reduced. The fraction of values within a factor two

improves from 0.739 to 0.931. Other metrics also improve, for example, $NMSE = 0.490$ and $FB = -0.428$ under the standard run and $NMSE = 0.160$ and $FB = 0.076$ when the memory effect is removed from the model.

Conclusions

A comprehensive study of the performance of the DAU-MOD-GRS model to estimate the nitrogen dioxide (NO_2) concentration using the first available long-term (4 years) air quality record, at the urban background site of the city of Buenos Aires, is performed. We present a novel methodology to study whether and how model errors vary with input data conditions. Applying a simple clustering analysis over three performance metrics that are computed daily, we assess differences between groups of days sharing similar model performance levels. The main advantage of this methodology is that it allows grouping data to study patterns in input data that may be associated to model errors.

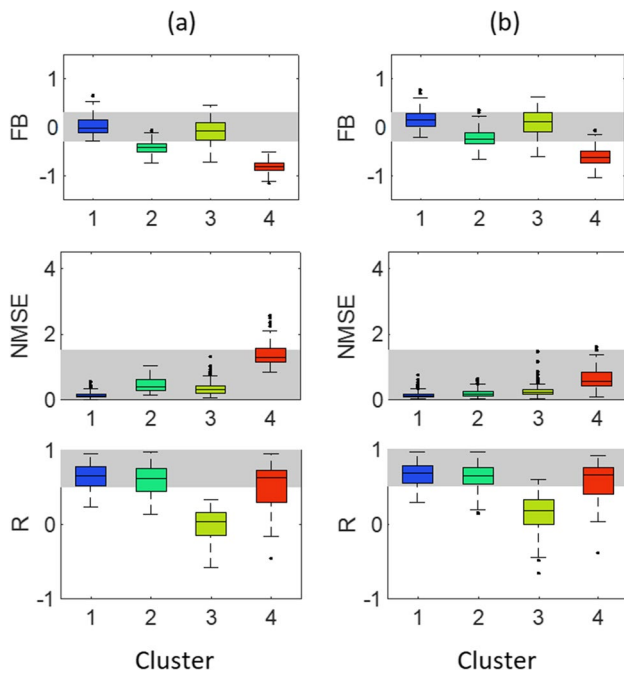
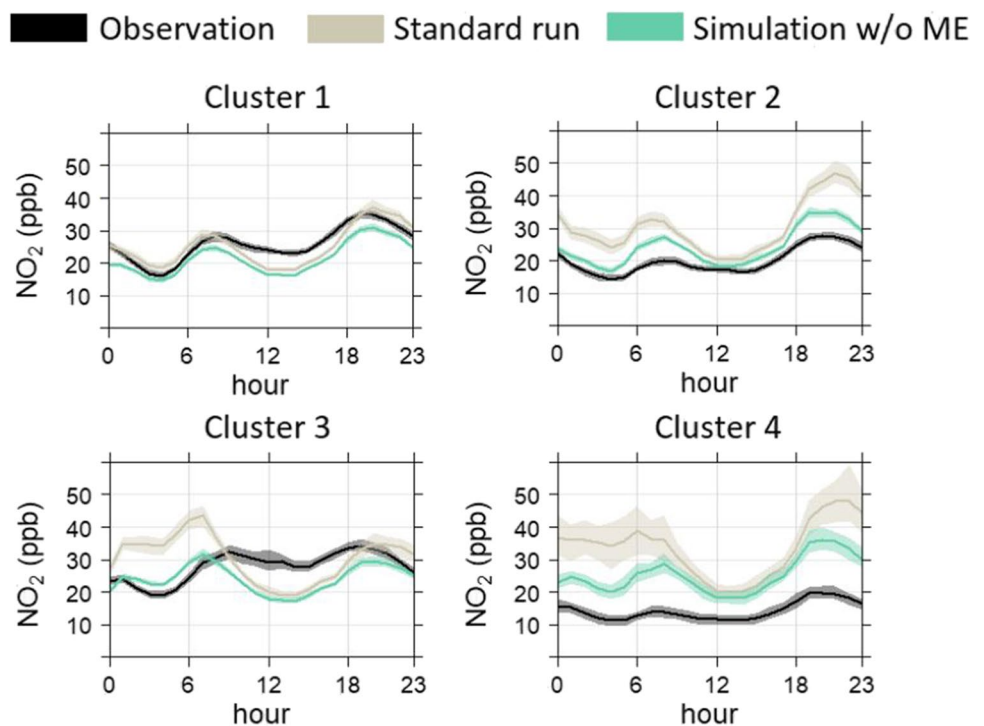


Fig. 7 Box plots of the three metrics by cluster, for two simulations: **a** standard and **b** without the memory effect of the model, considering the same classification used in **a** (i.e. each cluster groups the same days in the two figures **a** and **b**). Grey areas indicate better performance ranges (FB in $[-0.3, 0.3]$, NMSE $[0, 1.5]$ and R $[0.5, 1]$)

Four clusters of model performance metrics are found and these are ordered from best (cluster 1) to worst (cluster 4) performing days. Statistical differences between

Fig. 8 Mean hourly variations of observed and modelled NO_2 concentrations under the standard simulation and that without memory effect (ME), by cluster. Shaded areas indicate 95% confidence interval in the mean



clustered daily mean meteorological input variables are significant only for wind speed (WS) and air temperature (T), indicating relatively worse model performance under conditions of lower WS and higher T. This was not noticed using a traditional analysis of performance metrics by ranges of model input variables. When using the classification in combination with bivariate polar plots, groups of days presenting almost uniform overestimation were isolated from those presenting variability relative to the wind direction. An analysis of the cluster distributions of days by seasons and day of week reveals that cluster 4 presents a much larger proportion of weekend and summer days (i.e. when the emission inventory may be overestimated) compared to other clusters.

Finally, the proposed classification of days based on model performance metrics is used to assess the impact of removing the memory effect from the model (i.e. the residual pollutant concentration from the previous hour) on its ability to estimate the hourly concentrations of NO_2 at this site. In general, model performance improves over most conditions. Most of the improvement is achieved during nighttime, early morning and late evening hours, resulting in a better estimation of the peak concentrations. However, some differences still persist in the cluster of “worse performing days” highlighting the need to include a more realistic temporal allocation of emissions in our modelling system.

Overall, the proposed methodology allows for the identification of conditions mostly influencing model performance. Understanding whether model uncertainty is uniformly distributed or concentrated in particular

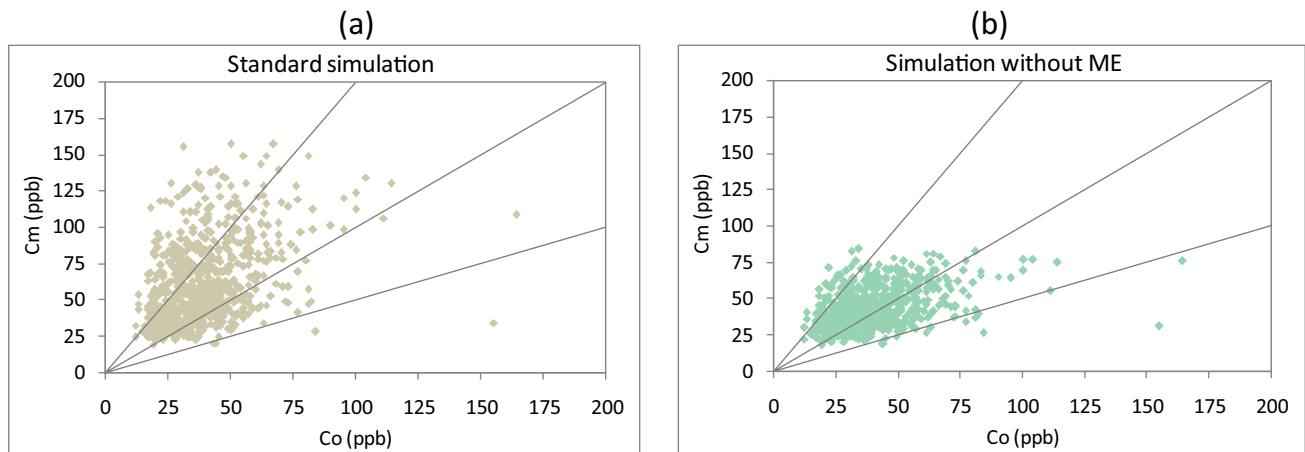


Fig. 9 Modelled (C_m) vs observed (C_o) daily maximum NO_2 concentrations at CEN, under **a** standard conditions and **b** without the memory effect (ME) of the model

conditions can help to identify aspects of the model that require further attention. The method can also be useful to gauge performance improvements related to specific model parameter or option changes.

Acknowledgements The authors are thankful to the ‘Agencia de Protección Ambiental (APRA)’ of the city of Buenos Aires and the ‘Servicio Meteorológico Nacional (SMN)’ of Argentina for providing the air quality and meteorological data, respectively.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Andrea L. Pineda Rojas and Emilio Kropff. The first draft of the manuscript was written by Andrea L. Pineda Rojas and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work has been funded by Project PICT2018-1624 from the National Agency for the Promotion of Science and Technology (ANPCyT) of Argentina.

Data availability The data sets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. The air quality and meteorological data used in this work are public and freely available from APRA and SMN, respectively.

Declarations

Competing interests The authors declare no competing interests.

References

- Arya SP (1999) Air pollution meteorology and dispersion. Oxford University Press, New York, p 310
- Azzi M, Johnson G, Cope M (1992) An introduction to the generic reaction set photochemical smog model. Proc. 11th Int. Clean Air Conf, pp 451–462. Available at: <http://hdl.handle.net/102.100.100/247315?index=1>. Accessed 5-10 Jul 1992
- Berens P (2009) CircStat: A Matlab Toolbox for circular statistics. J Stat Softw 31 (10). <http://www.jstatsoft.org/v31/i10>. Accessed 23 Sept 2009
- Carslaw DC (2018) The openair manual — open-source tools for analysing air pollution data. Manual for version 2.6–0, University of York
- Chang JC, Hanna SR (2005) Air quality model performance evaluation. Meteorol Atmos Phys 87:167–196
- Dennis R, Fox T, Fuentes M, Gilliland A, Hanna S, Hogrefe C, Irwin J, Rao ST, Scheffe R, Schere K, Steyn D, Venkatram A (2010) A framework for evaluating regional-scale numerical photochemical modeling systems. Environ Fluid Mech 10:471–489
- Derwent D, Fraser A, Abbott J et al. (2010) Evaluating the performance of air quality models. DEFRA Report. Issue 3/June 2010
- East J, Montealegre JS, Pachon JE, Garcia-Menendez F (2021) Air quality modeling to inform pollution mitigation strategies in a Latin American megacity. Sci Total Environ. 776:145894
- Hanna SR (1988) Air quality model evaluation and uncertainty. JAPCA 38(4):406–412. <https://doi.org/10.1080/08940630.1988.10466390>
- Kaufman L, Rousseeuw PJ (2009) Finding groups in data: an introduction to cluster analysis. John Wiley & Sons, New York, pp 344
- Lleti R, Ortiz MC, Sarabia LA, Sánchez MS (2004) Selecting variables for k-means cluster analysis by using a genetic algorithm that optimises the silhouettes. Anal Chim Acta 515:87–100
- Mazzeo NA, Venegas LE, Choren H (2005) Analysis of NO , NO_2 , O_3 and NO_x concentrations measured at a green area of Buenos Aires City during wintertime. Atmos Environ 39:3055–3068
- Mazzeo NA, Venegas LE (1991) Air pollution model for an urban area. Atmos Res 26:165–179
- MacQueen J (1967) Some methods for classification and analysis of multivariate observations. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, pp 281–297. Available at: https://digitalassets.lib.berkeley.edu/math/ucb/text/math_s5_v1_article-17.pdf. Accessed June 21–July 18, 1965
- Miranda A, Silveira C, Ferreira J et al (2015) Current air quality plans in Europe designed to support air quality management policies. Atmos Pollut Res 6:434–443
- Pineda Rojas AL (2014) Simple atmospheric dispersion model to estimate hourly ground-level nitrogen dioxide and ozone concentrations at urban scale. Environ Model Softw 59:127–134

- Pineda Rojas AL, Borge R (2019) Statistical evaluation of the urban atmospheric dispersion model DAUMOD-GRS to estimate NO₂ concentrations using new available data from Buenos Aires. Proceedings of the '19th international conference on harmonisation within atmospheric dispersion modelling for regulatory purposes (Harmo19)', 3–6 June 2019, Bruges, Belgium, H19–061, 5pp. Available at: <https://www.harmo.org/conference.php?id=19>. Accessed Jun 2019
- Pineda Rojas AL, Kropff E (2021) Exploring error types and performance of an air quality model through clustering analysis. Proceedings of the '20th international conference on harmonisation within atmospheric dispersion modelling for regulatory purposes (Harmo20)', 14–18 June 2021, Tartu, Estonia, H20–107, 5pp. Available at: <https://www.harmo.org/conference.php?id=20>. Accessed Jun 2021
- Pineda Rojas AL, Leloup JA, Kropff E (2019) Spatial patterns of conditions leading to peak O₃ concentrations revealed by clustering analysis of modeled data. *Air Qual Atmos Health* 12(6):743–754
- Pineda Rojas AL, Venegas LE (2013) Upgrade of the DAUMOD atmospheric dispersion model to estimate urban background NO₂ concentrations. *Atmos Res* 120–121:147–154
- Pisoni E, Guerreiro C, Lopez-Aparicio S, Guevara M, Tarrason L, Janssen S, Thunis P et al (2019) Supporting the improvement of air quality management practices: the “FAIRMODE pilot” activity. *J Environ Manage* 245:122–130
- Rouseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 20(1):53–65
- U.S. EPA (2014) CMAQv5.0.2 (version 5.0.2). Office of research and development, U.S. Environmental Protection Agency, Research Triangle Park, NC. Zenodo. 10.5281/zenodo.1079898
- Venegas LE, Mazzeo NA, Dezzutti MC (2014) A simple model for calculating air pollution within street canyons. *Atmos Environ* 87:77–86
- Venegas LE, Mazzeo NA, Pineda Rojas AL (2011) Chapter 14: Evaluation of an emission inventory and air pollution in the metropolitan area of Buenos Aires. En: D. Popovic (ed.) *Air quality-models and applications*, Editorial In-Tech, 261–288

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.