Chapter 84
On the Spatial Support of Time Series of Monitoring Data for Model Evaluation

Efisio Solazzo and Stefano Galmarini

Abstract We use time series of hourly records of ozone for a whole year (2006) collected by the European AirBase network to analyse the area of representativeness of monitoring stations and find, for similar class of stations (urban, suburban, rural), large heterogeneity and high sensitivity to the density of the network and to the noise of the signal. This suggests the mere station classification to be not a suitable method to help select the pool of stations used in model evaluation. Therefore a novel, more robust technique is developed consisting in studying the spatial properties of the associativity of the spectral component of the ozone time series, in an attempt to determine the level of homogeneity.

84.1 Introduction

This study investigates the use and limitations of monitoring data of relevance for the correct assessment of air quality modelling systems, with focus on the spatial representativeness and the homogeneity of signals. The aim is to provide modellers a methodology for selecting homogeneous spatially distributed signal which also carries information about its spatial support. We introduce a method for the estimation of the area of representativeness of monitoring receptors and the limitations it imposes on its use for model evaluation. We then propose a novel methodology consisting in studying the associativity of the spectral decomposition of the pollutants time series rather than the raw data. The criterion beyond such choice consists in assuming that components pertaining to different scales may show different levels of associativity and still be usable selectively to evaluate their counterpart in model data. Should that be the case only the components of similar associativity should be compared with the modelled counterpart.

E. Solazzo (✉) · S. Galmarini
European Commission, Joint Research Centre, Institute for Environment
and Sustainability, Air and Climate Unit, Ispra, Italy
e-mail: efisio.solazzo@jrc.ec.europa.eu

84.2   The Area of Representativeness

We initially estimate the distance at which the influence of the signal of each receptor is felt by the other receptors of the AirBase network. For this we apply variogram analysis complemented with a fitting procedure (Solazzo et al. 2013). Results show that the range (the distance after which the memory of the signal of the receptor can be considered lost) is approximately the same for all three categories of stations (urban, suburban, rural), thus not informative for selecting stations to be used to evaluate regional scale air quality models (Fig. 84.1). Refer to Solazzo and Galmarini (2015) for details.

84.3   Stations Associativity and Implications for Model Evaluation

Hourly time series of ozone concentrations have been first decomposed using an iterative moving average approach known as Kolmogorov-Zurbenko (kz) low-pass filter. The components of interest for the analysis are the intra-day (ID) component (periods less than 12 h), the diurnal (DU) component (periods of half-day to ~ 2.5 days), the synoptic (SY) component (between ~ 2.5 and ~ 21 days), the long-term (LT) component (between ~ 21 and ~ 90 days) and the seasonal (SE) component (base line containing periods longer than 90 days). Hierarchical clustering was applied to the spectral components of the hourly-ozone monitored time series in Europe. By examining the clusters of receptors produced at different
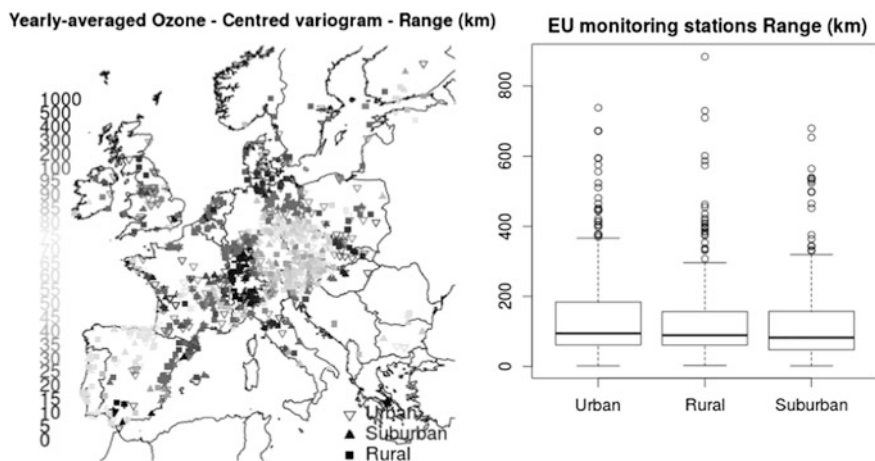

Fig. 84.1 Range (*left*) calculated using the point-centred variogram method at the AirBase monitoring sites. Percentile distribution (*right*) of the range (up to 1000 km and nugget < 50 ppb^2) by area-type. *Tick black line* is the median
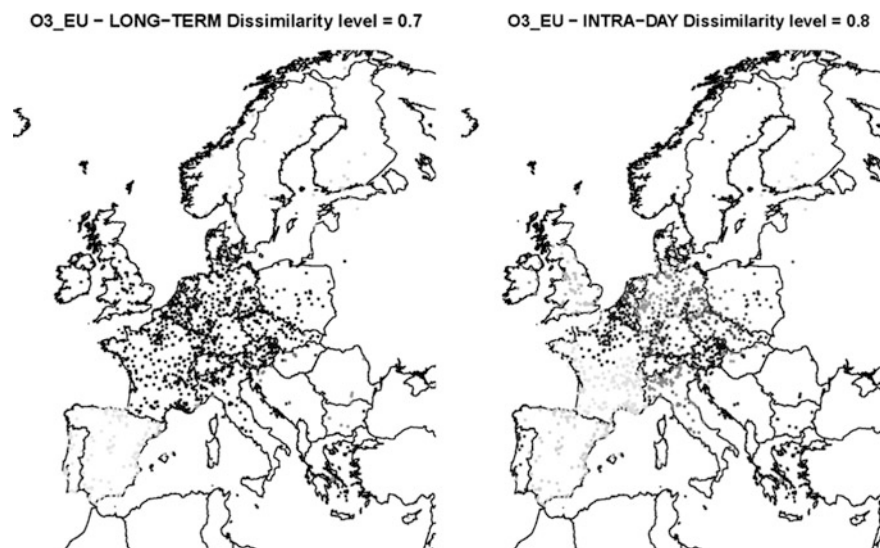
Fig. 84.2 Spatial associativity for time series components of ozone

level of similarity we find the spatial associativity of Fig. 84.2. Grouping of seasonal component confirms a strong correlation between all stations, up to a level of correlation of 0.75. Below the value of 0.25 some heterogeneity appears. When moving to examine the LT component at $d_0 = 0.7$ a strong spatial structure is detected (Fig. 84.2a). The result is a net trans-boundary grouping, most notably between Spain and France, and between the Scandinavian countries and the rest of the continent. The signal recorded at these stations has a higher correlation with stations of the same countries (even if far apart) than with stations across the border (although sited within their immediate vicinity). The SY component at $d_0 = 0.4$ (similarity of 0.6) and the ID component (Fig. 84.2b) confirms the same marked country-wise grouping, starting from high level of dissimilarity (~ 0.85), suggesting this latter component to be removed prior to carry out a model evaluation study. When repeating the analysis using model results from AQMEII, the detected cross-border associativity patterns are not captured by the modeling systems, supporting the argument that lack of harmonization among countries and networks can originate noise that deteriorates the information of the monitoring data for use in model evaluation and development.

When the portions of the signal that are not representative are removed from both measured and modelled data and the remaining components compared, the models performance improve. In fact the mean square error between the fifteen regional models participating to AQMEII and the hourly ozone concentration, over all the AirBase receptors is reduced on average of over 2.5 % (ranging from 1.1 to 6.2 %, with the ensemble mean improved by 2 %) when the ID component is removed from both the modelled and the observed time-series, for the all year.

Similar range of improvement for the mean squared error is achieved when considering the maximum daily concentration. When the analysis is restricted to the summer months (June–August) the noise suppression results in improved accuracy of 5 % on average (ranging from 1.4 to 9.6 %, with the ensemble mean improved by 4.3 %).

84.4   Conclusions

Analysis of spatial representativeness has been carried out on the AirBase stations and by using gridded model outputs as proxy of concentration. While the categorization of stations contained in the meta-data is not a viable indicator of the spatial extension of the area of influence of a receptor, we find that variogram methods in conjunction to spatially uniform proxy fields might provide a valid alternative. These methods, however, have a large degree of dependence on the quality of the field and on the fitting procedure itself and are probably better suited for isolated receptors. On the other hand, the method of searching for spatial grouping of stations based on signal frequency, with the aim of identifying homogeneous regions to facilitate model evaluation, has revealed to be more robust, allowing i) to capture the 'portion' of desired signal detected by each stations, even those located in heavily polluted urban areas, and ii) to discard the unwanted part. This feature is of high significance in a context of dynamic and diagnostic model evaluation (full details are provided in Solazzo and Galmarini, 2015). More specifically we find:

- Only the spectral components with high signal-to-noise ratio (SE, DU) exhibit grouping properties across the European continent as a whole, showing regions where the signal of these components is homogeneous. Using only the stations of these homogeneous regions will facilitate the model evaluation activity and interpretation of the results. The identification of these regions also indicates a novel strategy for sensitivity analysis and process-evaluation of models. For instance, clustering of the DU components can reveal which stations are most sensitive to the ozone diurnal cycle and use only those to diagnose models capability to reproduce it.
- The other components (LT, SY, ID) have surprisingly shown country-sensitive grouping properties: the signal is more strongly associated among stations sited in the same country but hundreds of kilometres apart than among stations sited on either side of the country border and only a few kilometres apart. This behaviour belongs to all European countries. Trans-boundary effects of fast fluctuating components cannot be captured by the modelling systems and should be removed prior of any model-to-data comparison.
- By filtering-out the ID component, the model accuracy (in terms of mean square error) improves significantly by up to 5 % on average over fifteen modelling systems for the months of June–August and on average of 2.5 % for the whole year and for the daily maximum concentration.

Question and Answer

Questioner: P. Makar

Question: Is there any way to distinguish between low correlations due to instrumental error versus due to stations representativeness?

Answer: In case of a dense network, the signal recorded by a faulty station is likely to be spotted more easily than in the case of isolate stations. In fact, if the correlation among a close group of stations (cluster) is uniform and only for one of them the correlation is lower, then that station requires closer inspection and the error would be diagnosed. On the other hand, if the station is isolated or close to a highly localized emission source the faulty signal would be harder to detect.

Questioner: J. Bieser

Question: How large is, on average, the fraction of the error that cannot be reduced on regional domains due to the effect being on a sub-grid scale?

Answer: We have found that the intra-day (ID) portion of the signal (for ozone concentration) recorded by the stations is weakly representative and models with grid spacing exceeding ~ 10–15 km might show an increase of the mean square error of ~ 5 % due to the limited representativeness of the stations for the ID share of the signal. The error is indeed due to sub-grid processes that cannot be picked up by the model.

References

Solazzo E et al (2013) Report of 2013 WG2/SG1 activity on spatial representativity. European Commission EUR 26539 EN—Joint Research Centre, Institute for the Environment and Sustainability, Luxembourg, 22 p. doi:10.2788/17619

Solazzo E, Galmarini S (2015) Comparing apples with apples: Using spatially distributed time series of monitoring data for model evaluation. Atmos Environ 112:234–245