Atmospheric Pollution Research 6 (2015) 928–938

Contents lists available at ScienceDirect

Atmospheric Pollution Research

journal homepage: http://www.journals.elsevier.com/locate/apr

Original article

Application of bias adjustment techniques to improve air quality forecasts

Camillo Silibello [a, *], Alessio D'Allura [a], Sandro Finardi [a], Andrea Bolignano [b], Roberto Sozzi [b]

[a] ARIANET Srl, via Gilino, 9, 20128 Milan, Italy
[b] ARPA Lazio, Via Garibaldi, 114, 02100 Rieti, Italy

ARTICLE INFO

Article history:
Received 1 December 2014
Received in revised form
10 April 2015
Accepted 11 April 2015
Available online 23 October 2015

Keywords:
Air quality forecast
Bias-correction
Kalman filter
Model evaluation
Score analysis

ABSTRACT

Two bias adjustment techniques, the hybrid forecast (HF) and the Kalman filter (KF), have been applied to investigate their capability to improve the accuracy of predictions supplied by an air quality forecast system (AQFS). The studied AQFS operationally predicts NO_2, ozone, particulate matter and other pollutants concentrations for the Lazio Region (Central Italy). A thorough evaluation of the AQFS and the two techniques has been performed through calculation and analysis of statistical parameters and skill scores. The evaluation performed considering all Lazio region monitoring sites evidenced better results for KF than for HF. RMSE scores were reduced by 43.8% (33.5% HF), 25.2% (13.2% HF) and 41.6% (39.7% HF) respectively for hourly averaged NO_2, hourly averaged O_3 and daily averaged PM10 concentrations. A further analysis performed clustering the monitoring stations per type showed a good performance of the AQFS for ozone for all the groups of stations ($r = 0.7$), while satisfactory results were obtained for PM_{10} and NO_2 at rural background ($r = 0.6$) and Rome background stations ($r = 0.7$). The skill scores confirmed the capability of the adopted techniques to improve the reproduction of exceedance events.

Software availability

The air quality forecast system is based on the chemical transport model FARM (Flexible Air quality Regional Model). FARM is an Open Source model available at: https://hpc-forge.cineca.it/. Further information can be found at the COST model inventory web site (https://www.mi.uni-hamburg.de/Model-Inventory.5554.0.html) or in the European Topic Centre for Air Pollution and Climate Change Mitigation (ETC/ACM) Model documentation system (http://acm.eionet.europa.eu/databases/MDS/index_html).

1. Introduction

During the last decade, deterministic air quality forecast systems (AQFS) have gradually been implemented to predict short-term air pollution episodes. Their development has been fostered to achieve compliance of short-term limit values imposed by the air quality legislation. Reliable air quality forecast, provided at least two days in advance, can in fact support the day-by-day adoption of mitigation measures, supply information to the public about health risks and give advices to reduce personal exposure. The use of AQFS is implicitly promoted by the "Clean Air for Europe" Directive 2008/50/EC, which states that "Member States shall ensure that timely information about actual or predicted exceedance of alert thresholds, and any information threshold is provided to the public".

Even if air quality models have proved to be capable to reproduce regional and urban atmospheric pollution phenomena and their reliability has been demonstrated by a number of single and multi-model evaluation studies, there are many sources of uncertainty in their use for operational applications (Kukkonen et al., 2012), the most important ones are related to uncertainties in emission data and meteorological predictions, and to the incomplete representation of the physical/chemical phenomena that determine pollutants concentrations. These uncertainties can determine model errors (Chang and Hanna, 2004; Borrego et al., 2008) and consequently failures in air quality predictions. Moreover, air quality models have space resolution limitations that do not permit to reproduce sub-grid scale features like urban hot-spot concentrations. In most applications, AQFS concentration fields can

* Corresponding author. Tel.: +39 0227007255; fax: +39 0225708084.
E-mail address: c.silibello@aria-net.it (C. Silibello).
Peer review under responsibility of Turkish National Committee for Air Pollution Research and Control.

http://dx.doi.org/10.1016/j.apr.2015.04.002
1309-1042/

be compared positively with rural and urban background air quality stations, while concentrations measured at traffic stations and in small towns can hardly be reconstructed. To improve the reproduction of measured concentration levels, different bias-correction techniques have been developed to remove systematic errors (Delle Monache et al., 2008; Kang et al., 2008, 2010; Borrego et al., 2011; Sicardi et al., 2012).

The PM_{10}, NO_2 and ozone predictions provided by the operational AQFS of the Lazio Region (Central Italy) during year 2012 have been used to verify the capability of bias-adjustment techniques to improve forecast accuracy. The existing AQFS provides satisfactory results for Rome city background stations (Finardi et al., 2010), while difficulties are encountered in reproducing pollutants concentration levels in smaller urban areas influenced by local scale phenomena not resolved by the model. To improve the AQFS prediction, we have post-processed the model results (raw model) with the so-called hybrid forecast (HF, Kang et al., 2008) and the Kalman filter (KF, Delle Monache et al., 2006) techniques. HF combines the latest observed concentration with model-predicted tendency for the subsequent forecast time period, while the KF is a linear, adaptive, recursive and optimal algorithm.

The operational AQFS is presented in Section 2. The two bias-adjustment techniques used to improve forecast accuracy are described in Section 3. The analysis of the results obtained by the application of bias adjustment techniques is presented in Section 4.

2. AQFS description

The city of Rome and other urban areas in the Lazio Region are affected by high ambient concentrations of particulate matter, NO_2 and O_3. The air quality monitoring network (see Table 1) data show that the annual average concentration of NO_2 is above the European Union (EU) limit value for human health protection (40 $\mu g\ m^{-3}$) at most of the urban stations. The target value for the maximum daily value of the 8-h running mean of ozone concentration (120 $\mu g\ m^{-3}$) is exceeded several times a year, especially at regional background stations located in hilly and mountain areas. For PM_{10}, the annual mean limit value of 40 $\mu g\ m^{-3}$ is generally respected, with episodic violations recorded at urban traffic stations, while the maximum number (35) of exceedances of the daily average limit value for human health protection, fixed at 50 $\mu g\ m^{-3}$, is generally unattained at urban traffic stations. Because concentrations vary strongly among locations, air quality models can play an important role in assessing and predicting the spatial and temporal variation of the air quality across the region. The development and the use of air quality models is fostered by the EU Directive (EC, 2008) that encourages the combined use of monitoring data, emission inventories and modelling techniques; moreover, it requires the distribution of air quality information for the current day, together with trend and forecast for the next days, when concentrations are expected to exceed alert and information thresholds. To meet these requirements, the Lazio Region authorities supported the development of an AQFS covering the whole region and, with more detail, the Rome urban area (ARPA Lazio, 2010). The Lazio/Rome AQFS is built along the same lines as the atmospheric component of the "National Integrated Modelling system for International Negotiation on atmospheric pollution" (MINNI, 2008; D'Elia et al., 2009; Mircea et al., 2014). Operational since 2009, the AQFS considers two nested computational grids (Fig. 1): a background domain, covering a large portion of Central Italy (66 × 58 cells, with 4 by 4 km grid spacing) and a target domain including the Rome urban area (61 × 61 cells, with 1 by 1 km grid spacing). The "two-way nesting" approach used by the employed meteorological and air quality models allows for bi-directional information exchanges between the coarse (background) and the fine (Rome) computational grids.

The AQFS is based on the Flexible Air quality Regional Model (FARM; Gariazzo et al., 2007; Silibello et al., 2008, 2012) that employs the SAPRC-99 (Carter, 2000) chemical mechanism and the *aero3* modal aerosol scheme from CMAQ (Binkowski, 1999; Binkowski and Roselle, 2003).

2.1. Meteorology and natural emissions

The meteorological fields are produced by the prognostic and non-hydrostatic model RAMS version 6 (Cotton et al., 2003) using a two-way nested grid system. Further information required by FARM (gas-phase species deposition velocities, horizontal and vertical diffusivities) are calculated by an interface module (Finardi et al., 2005), as function of meteorological parameters (e.g. wind speed, solar radiation, temperature) and geographic characteristics (e.g. soil type and land cover). This module also includes algorithms to estimate natural emissions of aerosols (sea-salt and soil dust) driven by surface wind (Vautard et al., 2005; Zhang et al., 2005) and trace species from vegetation, depending on vegetation type and meteorological conditions (Guenther et al., 2006).

2.2. Anthropogenic emissions

Diffuse emissions are modelled from the latest available national emission inventory disaggregated at province level (year 2005), updated to the simulated year, using nationwide emissions trends available on a yearly basis for each pollutant and activity (ISPRA , 2015). The values from the national integrated assessment model GAINS-Italy (D'Elia et al., 2009) have been used to estimate PM10 and NMVOC emissions from wood and stubble burning (Caserini et al., 2007) and for two other sectors not included in the national inventory: construction and other combustion activities (barbecues, cigarettes smoke and fireworks). The largest industrial facilities are considered as point sources, with emission rates derived from stack measured data and owner declarations to local control authorities. A bottom-up methodology, based on the TREFIC model (Nanni and Radice, 2004), was adopted to estimate road traffic emissions from vehicles flows on the road network and fleet data (Gariazzo et al., 2007). TREFIC follows the COPERT IV approach and includes, for particulate matter, the emission factors developed by IIASA (IIASA, 2001) that consider both exhaust and non-exhaust (tyres, brakes, road coating) sources. For the Rome urban road network, limitation to the circulation of some categories of vehicles (e.g. non catalytic, EURO 1, etc.) in specific zones is also taken into account.

2.3. Boundary conditions

Boundary conditions for the 4-km FARM grid are provided by the "QualeAria" modelling system (QualeAria, 2015), developed within the COST Action ES0602 collaboration framework (COST ES0602, 2007; Kukkonen et al., 2012), which provides regional scale air quality forecast over the Italian peninsula starting from national and European emission inventories and synoptic scale weather forecast.

3. Bias adjustment methodology

3.1. Raw model

The dataset used in this work has been built using the first 24 h forecast (+24 h fcst) produced by the AQFS each day, for the whole year 2012, on the regional domain (4 km horizontal resolution). Since the "+24 h fcst" of day "*d*" uses previous day forecast as initial condition (e.g. hour 24 of "*d−1*" fcst), time series discontinuities