Atmospheric Pollution Research 6 (2015) 928–938

HOSTED BY ELSEVIER

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

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

Table 1
Lazio Region monitoring network: stations names, coordinates (UTM: Zone 32), types and measured pollutants (PM_{10}, NO_2 and O_3). Blue cells associated to stations belonging to Rome domain.

Province	Station	UTM x [km]	UTM y [km]	Z a.s.l. [m]	Type	PM_{10}	NO_2	O_3
Frosinone	Fontechiari	889.2	4623.6	388	Rural background	V	V	V
	Alatri	860.9	4628.6	445	Urban background	V	V	
	Anagni	845.1	4630	401	Urban background	V	V	
	Cassino	903.3	4604.3	41	Urban traffic	V	V	
	Ceccano	861.7	4611	130	Urban traffic	V	V	
	Ferentino	853.7	4624.4	316	Urban traffic	V	V	
	Frosinone Mazzini	862.2	4618.9	153	Urban traffic	V	V	V
	Frosinone Scalo	860.8	4617.1	161	Urban traffic	V	V	
Latina	Aprilia	804.5	4611.3	83	Urban background	V	V	
	Latina Romag.	825.1	4598.4	23	Urban traffic	V	V	
	Latina Scalo	829.3	4605.3	18	Urban traffic	V	V	
	Latina Tasso	826.8	4597.7	21	Urban traffic	V	V	V
Rieti	Leonessa	825.2	4721	948	Rural background	V	V	V
	Rieti	817.5	4701.9	397	Urban traffic	V	V	V
Rome	*Castel di Guido*	*771*	*4642.7*	*61*	*Rural background*	*V*	*V*	*V*
	Cavaliere	*803.4*	*4648.6*	*48*	*Rural background*	*V*	*V*	*V*
	Villa Ada	*790.8*	*4648.3*	*50*	*Urban background*	*V*	*V*	*V*
	Arenula	*788.3*	*4643.8*	*31*	*Urban background*	*V*	*V*	*V*
	Bufalotta	*792.9*	*4650*	*41*	*Urban background*	*V*	*V*	*V*
	Cinecittà	*796.2*	*4640.1*	*53*	*Urban background*	*V*	*V*	*V*
	Cipro	*786*	*4645.1*	*31*	*Urban background*	*V*	*V*	*V*
	Preneste	*793.8*	*4643.9*	*37*	*Urban background*	*V*	*V*	*V*
	Malagrotta	*777.6*	*4641.3*	*50*	*Urban/Industrial*	*V*	*V*	*V*
	Guidonia	*808.7*	*4656*	*89*	*Urban traffic*	*V*	*V*	
	Ciampino	*799.7*	*4633.6*	*134*	*Urban traffic*	*V*	*V*	
	Fermi	*788*	*4640.5*	*26*	*Urban traffic*	*V*	*V*	
	Francia	*787.6*	*4649.8*	*43*	*Urban traffic*	*V*	*V*	
	Magna Grecia	*791.2*	*4642.7*	*49*	*Urban traffic*	*V*	*V*	
	Tiburtina	*794.3*	*4645.9*	*32*	*Urban traffic*	*V*	*V*	
	Colleferro Europa	833.5	4627	223	Urban/Industrial	V	V	
	Colleferro Oberdan	833.1	4627.6	219	Urban/Industrial	V	V	V
	Allumiere	740.3	4671.4	542	Rural background	V	V	V
	Civitavecchia	731.8	4663.8	26	Urban traffic	V	V	V
	Civitavecchia Porto	730.6	4664.3	6	Urban/Industrial	V		
	Civitavecchia Albani	731.4	4664.6	34	Urban traffic	V		
	Civitavecchia Morandi	732.1	4663.2	22	Urban traffic			
Viterbo	Acquapendente	735.5	4735.6	377	Rural background	V	V	V
	Civita Castellana	781.6	4687.5	139	Urban background	V	V	
	Viterbo	755.8	4701.3	338	Urban traffic	V	V	V

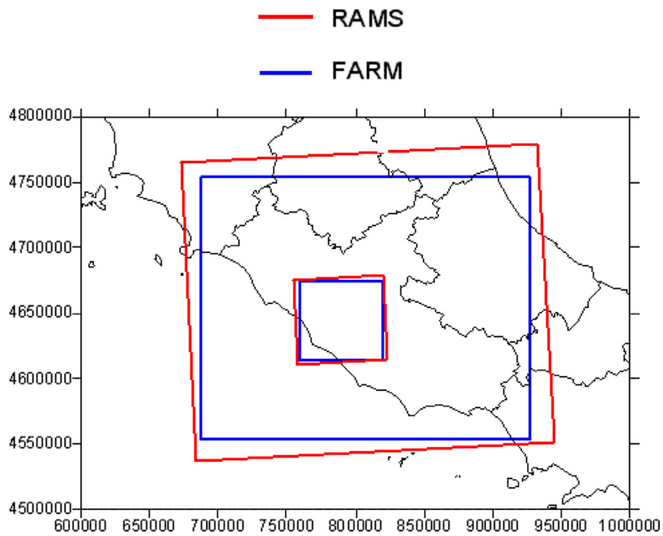

Fig. 1. Nested domains used by RAMS and FARM models over Lazio (Central Italy) and Rome conurbation (4 and 1 km horizontal resolutions).

from one day to the next are prevented. Fig. S1 provides the comparison between observed and predicted 24-hr averaged PM_{10}, O_3 and NO_2 concentrations at the "Villa Ada" urban background station. The monitoring station is located inside a large green area within the city of Rome (Villa Ada Park), about 200 m from the nearest street, and is considered representative of the background air quality within the Rome conurbation. The analysis of this Figure evidences that the daily/seasonal variations are quite well represented by the AQFS, although it shows a seasonal bias (see the overestimation of observed PM_{10} concentrations during colder periods probably determined by overestimations in residential heating emissions), which could be corrected by means of bias removal techniques. It is worth noting that the presented results refer to the coarse computational grid, even if the Rome area is covered by the fine model mesh which provides better results. This choice allows for a straightforward comparison of the bias adjustment techniques performances over the whole modelled area and highlights forecast improvements at critical locations (roadside and small town stations).

3.2. Hybrid forecast

The HF correction (Kang et al., 2008) is based on the simple assumption that the model is able to correctly predict the change of concentration over time (dC/dt). As an example, in Fig. S2 is presented the comparison between observed and predicted time variability (one day concentration increment/decrement) of PM_{10} concentrations at Villa Ada station. The Figure also reports the linear regression and the r^2 value for the linear regression (0.23) evidencing a tolerable capability of the model in reproducing PM_{10} concentration variation over time. According to the above hypothesis the forecast at a given monitoring location can be improved by combining the observed value at the previous time ("true" initial condition) with the forecasted change of concentration. So, the bias adjusted Hybrid Forecast (HF) for the future time $t + \Delta t$ ($HF_{t+\Delta t}$) is given by the following relationship:

$$HF_{t+\Delta t} = o_t + (f_{t+\Delta t} - f_t)$$

where o_t, f_t and $f_{t+\Delta t}$ are observations and forecasts at times t and $t + \Delta t$.

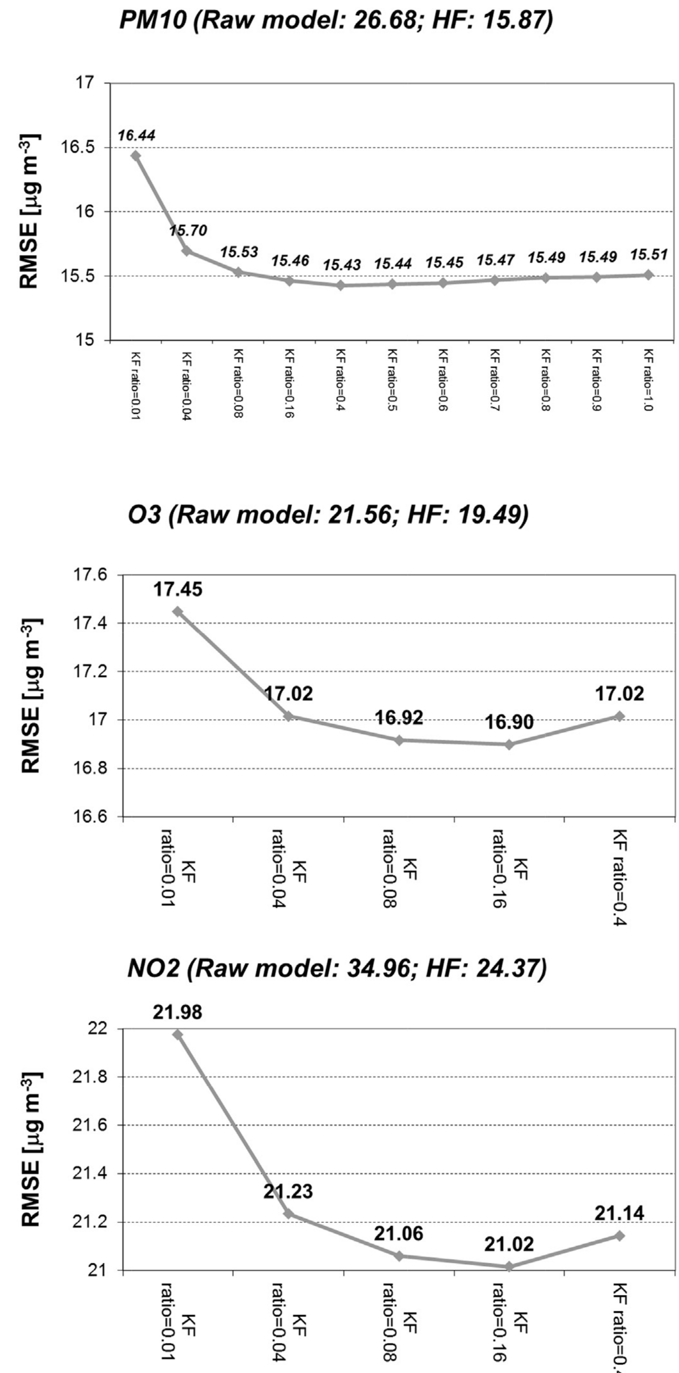

Fig. 2. RMSE analysis adopted to select the ratios *r* to be used by KF for the considered pollutants. Raw model and Hybrid Forecast (HF) RMSE values are also reported.

3.3. Kalman filter

The KF (Kalman, 1960; Delle Monache et al., 2006) is based on a recursive algorithm to estimate a signal from noisy measurements in which information from recent past forecasts and observations are used to revise the estimate of the current raw forecast. The main goal of KF is to estimate the forecast bias (x_t) between the forecast and the true (unobserved) concentration:

$$x_{t|t-\Delta t} = x_{t-\Delta t|t-2\Delta t} + \eta_{t-\Delta t}$$

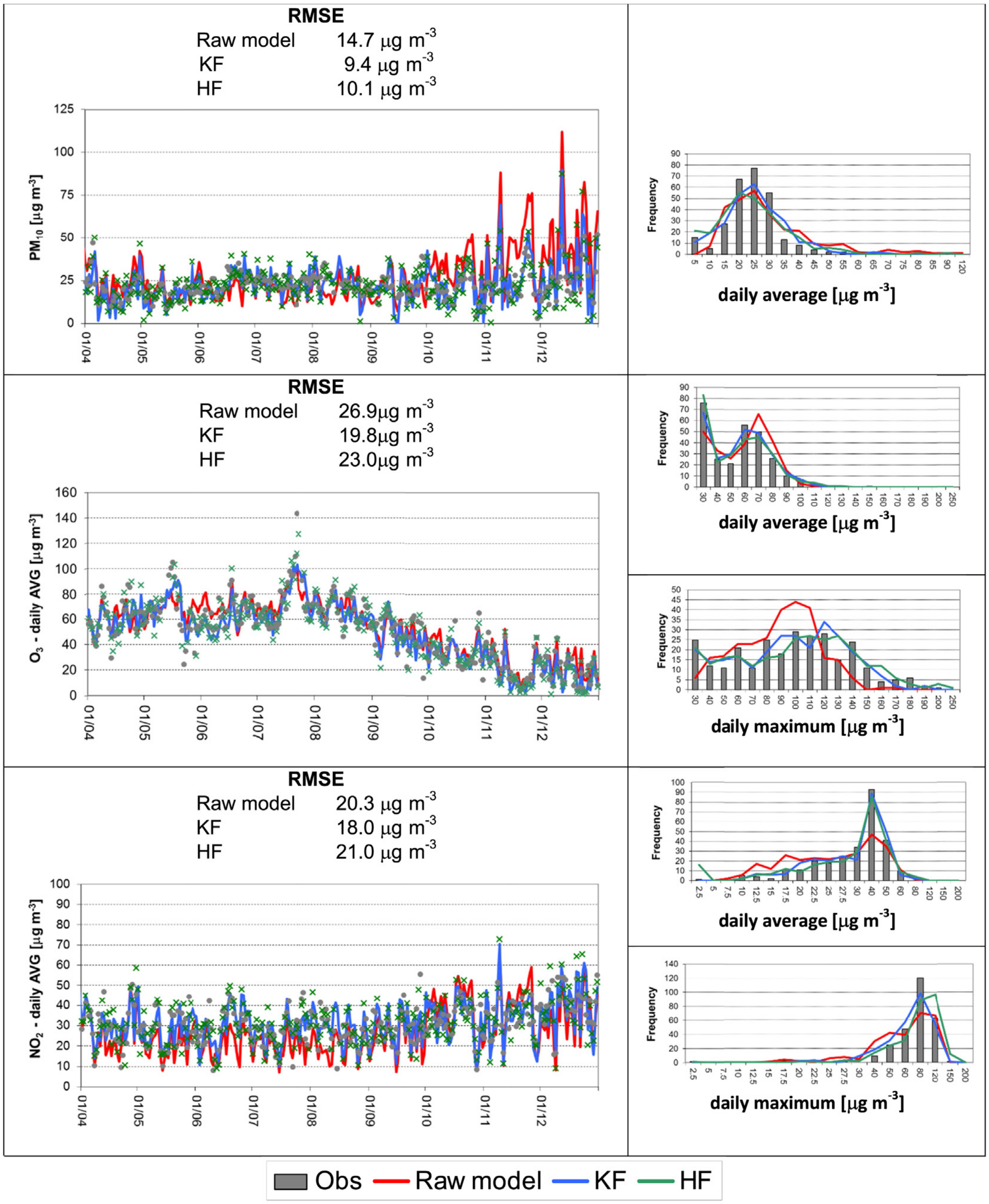

Fig. 3. Villa Ada station: time series for daily averaged PM10, NO_2 and O_3 [mg m^{-3}] observations, raw model, Kalman filter and Hybrid adjusted forecasts (left); frequency distributions of daily averaged and daily maximum (NO_2 and O_3) observed and predicted concentrations (right). RMSE values are also given.

where $t|t-\Delta t$ means that the value of the variable at time t depends on the value at time $t-\Delta t$; Δt is the time lag and η is called "*process noise*", uncorrelated in time, normally distributed with zero mean and variance σ^2_{η}. x_t is not observable but is related to the measurable bias y_t (the difference between forecasts and observations). Due to uncertainties in modelling, y_t is corrupted by a random error ε_t, called "*measurement noise*", again uncorrelated in time, normally distributed with zero mean and variance σ^2_{ε}:

$$y_t = f_t - o_t = x_t + \varepsilon_t = x_{t-\Delta t} + \eta_{t-\Delta t} + \varepsilon_t$$

The optimal recursive predictor of $\widehat{x}_t$ (e.g. the "best" estimate based on all previous measurements) can be written as

$$\widehat{x}_{t+\Delta t|t} = \widehat{x}_{t|t-\Delta t} + K_{t|t-\Delta t}\left(y_t - \widehat{x}_{t|t-\Delta t}\right)$$

where K, the "Kalman gain", is

$$K_{t|t-\Delta t} = \frac{p_{t-\Delta t|t-2\Delta t} + \sigma^2_{\eta}}{\left(p_{t-\Delta t|t-2\Delta t} + \sigma^2_{\eta} + \sigma^2_{\varepsilon}\right)}$$

and p, the expected mean square error, is

$$p_{t|t-\Delta t} = \left(p_{t-\Delta t|t-2\Delta t} + \sigma^2_{\eta}\right)\left(1 - K_{t|t-\Delta t}\right)$$

The error variances σ^2_η and σ^2_ε are not known a priori. They can be estimated by defining the new variable z_t as follows:

$$z_t = y_{t+\Delta t} - y_t = \ldots = \eta_t + \varepsilon_{t+\Delta t} - \varepsilon_t$$

that has variance $\sigma_z^2 = \sigma_\eta^2 + 2\sigma_\varepsilon^2 = (2+r)\sigma_\varepsilon^2$ where $r = \sigma_\eta^2/\sigma_\varepsilon^2$. The Kalman filter itself can be used to estimate σ^2_ε, and the process noise variance σ^2_η is then computed as $\sigma^2_\eta = r\,\sigma^2_\varepsilon$. As also pointed out by Delle Monache et al. (2006), the KF is sensitive to the ratio r: if the ratio is too high, the filter will put excessive confidence on the past forecast; vice versa, if the ratio is too low, the filter will be unable to respond to changes in bias.

Given the raw forecast and observation time series, estimates of σ^2_ε and the initial estimate of state $\hat{x}_{t_0|t_0}$ and $p_{t0|t0}$, the KF can recursively generate $\hat{x}_{t+\Delta t}$. Combined with the raw model forecast, it leads to the new KF forecast:

$$KF_{t+\Delta t} = f_{t+\Delta t} - \hat{x}_{t+\Delta t|t}$$

4. Bias-adjusted forecast

The above bias-adjustment techniques have been applied to hourly average data, using for each day only the measured and predicted values from previous day at the same hour (corresponding to a $\Delta t = 24$ h time delay). This approach takes into account the time-varying behaviour the bias may manifest at different times of the day (Delle Monache et al. 2006). An exception is made for PM_{10}: since its observations were available on a daily basis, KF and HF have been applied to daily average concentrations using a time lag $\Delta t = 1$ day (e.g. a given daily averaged concentration is corrected using only the daily averaged forecast and observation of the previous day). Following Delle Monache et al. (2006), KF today's forecast bias is therefore estimated using yesterday's bias, which in turn was estimated using the day-before-yesterday's bias, and so on. Delle Monache et al. (2006) used a value of 0.01 for r ratio taken from previous studies, where the KF was used to bias-correct weather forecasts (Roeger et al., 2003). This value was close to the optimal value (0.06) found by Homleid (1995), who tested three different values of the ratio (0.01, 0.06, and 0.16) for temperature forecast adjustment in Norway. Homleid (1995) also found that the sensitivity to the specification of the error ratio value was low with respect to bias reduction. Delle Monache et al. (2008) found an optimal value of 0.4 for eight Chemical Transport Models (CTM) ozone forecast simulations over north-eastern USA and southern Canada. Sicardi et al. (2012) investigated this issue to improve the simulated O_3 maximum concentration over Spain and used the following season dependent values: 0.4 (Winter), 0.2 (Spring), 0.15 (Summer) and 0.6 (Autumn). As also remarked by Delle Monache et al. (2008), since the KF is optimal in a least-square-error sense (i.e. it is designed to reduce RMSE), the “optimal” values for the ratio r should be selected by finding the values that minimize the RMSE metric considering observed and predicted concentrations at all monitoring sites.

In this perspective, a calibration period of three months (e.g. from January to March 2012) has been considered to train the KF and to find the values of r that minimize RMSE. Fig. 2 presents the RMSE values computed considering all stations and by varying the r ratio from 0.01 to 0.4, for NO_2 and O_3, and from 0.01 to 1, for PM_{10}. The RMSE curves show minimum values for the following ratios: 0.16 for NO_2 and O_3 and 0.4 for PM_{10}. Ratios computed at different station types (traffic, background, etc.) are generally coherent with these overall values. Even if the optimal error ratios vary in space, Kang et al. (2008) found that the impact on the bias-adjusted forecast of using different values over the model domain was insignificant when compared to employ a unique representative value. The Figure reports also the RMSE values computed for the raw model and the HF method. The lower RMSE values obtained using the KF bias adjustment technique point out its better potential to reduce air quality forecast errors. This has been verified by applying bias correction, for all the available monitoring stations, over the subsequent time period (i.e. from April to December 2012). As an example, Fig. 3 shows the comparison between observed, raw model, KF and HF adjusted forecasts concentration time series for PM_{10}, NO_2 and O_3 at the urban background site of Villa Ada. This Figure also reports the comparison of the frequency distributions of daily averaged and daily maximum (daily maximum of 1-hr average concentrations – for NO_2 and ozone -) observed and predicted concentrations and the RMSE values. Both bias removal predictors improve forecast skills: the KF performs better than the HF in terms of RMSE, while HF better reproduces the peaks even if it shows a tendency to over-prediction. As suggested by Delle Monache et al. (2006), the lower capability of KF to reproduce short-time pollution events could be ascribed to the KF inability “to predict a large bias when all biases for the past few days have been small”.

4.1. Forecast evaluation and skill scores analysis

A thorough assessment of forecast quality requires the computation of statistical parameters and skill scores in order to evaluate how the two bias correction techniques improve the AQFS capability to forecast high concentration episodes and possible exceedances of air quality limits over Lazio Region. The analysis of threshold exceedance events is performed considering the following scores that are based on the so-called “Contingency Table” (Table 2) and are commonly adopted to evaluate meteorological and air quality forecasts:

$$\textit{Bias score}: B = \frac{a+b}{a+c}\cdot 100,\ B = 100\ \textit{unbiased},\quad B < 100\ \textit{underforecast}\ \ B > 100\ \textit{overforecast}$$

$$\textit{Probability of detection}: POD = \frac{a}{a+c}\cdot 100,\ \ 0 \le POD \le 100,\ \textit{best score}: POD = 100,\ \textit{best score} \ne \textit{perfect fcst}$$

$$\textit{False alarm ratio}: FAR = \frac{b}{a+b}\cdot 100,\ \ 0 \le FAR \le 100,\ \textit{best score}: FAR = 0,\ \textit{best score} \ne \textit{perfect fcst}$$

$$\textit{Accuracy}: ACC = \frac{a+d}{N}\cdot 100,\ \ 0 \le ACC \le 100,\ \ \textit{best score}: ACC = 100,\ \ \textit{best score} = \textit{perfect fcst}$$

where a, b, c and d represent respectively the number of “hits”, “false alarms”, “misses forecast” and “no misses/correct rejects forecast”. These scores have the following meaning: ***B*** indicates whether the forecast overestimate or underestimate the number of exceedances, ***POD*** gives an idea of the fraction of exceedances actually forecasted by the system, ***ACC*** measures the percentage of simulations that correctly reproduce exceedance and non-exceedance events and ***FAR*** is the fraction of forecasted exceedances that did not occur.

Table 2
Contingency table.

		observations		
		yes	*no*	
forecasts	*yes*	*a* *hits*	*b* *false alarms*	*a+b* *yes fcsts*
	no	*c* *misses*	*d* *correct rejects*	*c+d* *no fcsts*
		a+c *yes obs*	*b+d* *no obs*	*N* *total fcsts,obs*

The short-term limit values imposed by the European Directive 2008/50/EC on air quality (EC, 2008) for PM_{10} (50 μg m^{-3} daily average that should not be exceeded more than 35 times per year), NO_2 (200 μg m^{-3} for hourly average concentration) and ozone (180 μg m^{-3} for hourly average concentration) cannot be used to identify the exceedance events, because they are very rarely reached at the Lazio Region stations during the investigated period (April–December 2012). According to Pay et al. (2014) we have used the 75th percentile of the observed concentrations for each species.

Table 3 shows the statistical parameters (see Appendix A for their definition) and the scores computed for the AQFS (raw model) and the bias corrected forecasts (KF and HF) for all Lazio region stations. The evaluation of the results should also account for the fact that measurements are not the absolute truth and have some degree of uncertainty. The analysis of the statistical parameters reported in Table 3 confirms the capability of the two bias-adjustment techniques to improve the raw model forecast, as well as the better results obtained using KF (i.e. statistical parameter closer to their ideal values). In particular, 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 PM_{10} concentrations. As for the raw model, the statistical parameters are comparable with those provided by similar applications (Pay et al., 2014), indicating a good performance of the AQFS particularly for ozone and PM_{10}. The lower performance of the raw

Table 3
NO_2, O_3 and PM_{10} forecast evaluation and skill scores analysis for the AQFS (raw model) and the two bias correction techniques (KF and HF) considering all Lazio region stations. Investigated period: April–December 2012.

All	NO_2			O_3			PM_{10}		
stations	39			21			38		
Mean (O) [μg m^{-3}]	3.8			56.9			25.5		
Threshold [μg m^{-3}]	48			83			30		
Set	Raw model	KF	HF	Raw model	KF	HF	Raw model	KF	HF
Mean (P) [μg m^{-3}]	14.3	33.6	34.4	59.3	57.4	58.8	17.8	25.5	25.6
r	0.45	0.76	0.70	0.70	0.85	0.81	0.37	0.76	0.76
RMSE [μg m^{-3}]	30.7	17.3	20.5	25.7	19.2	22.3	16.1	9.4	9.7
IA	0.59	0.86	0.83	0.81	0.92	0.90	0.57	0.87	0.87
FAC2 [%]	32.5	79.5	73.6	75.8	79.9	77.4	70.7	93.5	92.5
MFB [%]	−82.7	6.4	−0.8	17.4	8.8	2.1	−35.9	−0.6	−2.2
a	9028	37,106	35,191	15,709	23,268	22,916	593	1626	1618
b	4606	17,494	22,242	6695	6445	8812	408	1051	1049
c	48,985	20,907	20,111	14,989	7430	7293	1800	767	722
d	171,038	158,150	152,403	88,401	88,648	82,026	7072	6429	6285
BIAS [%]	23.5	94.1	103.9	73.0	96.8	105.0	41.8	111.9	114.0
POD [%]	15.6	64.0	63.6	51.2	75.8	75.9	24.8	67.9	69.1
FAR [%]	33.8	32.0	38.7	29.9	21.7	27.8	40.8	39.3	39.3
ACC [%]	77.1	83.6	81.6	82.8	89.0	86.7	77.6	81.6	81.7

Note: Mean (O) and Mean (P) indicate respectively the observed and predicted mean concentrations; Threshold is the threshold [μg m^{-3}] applied in the score evaluation calculated as the 75° percentile of the observed concentrations. *a*, *b*, *c* and *d* represent the number of "hits", "false alarms", "misses forecast" and "no misses/correct rejects forecast".

model for NO_2 is due to the inclusion of roadside stations in the verification dataset. Traffic station measurements are underestimated due to the model coarse resolution, but the application of KF and HF allows to verify the improvement obtained by systematic bias removal. This approach is conceptually analogous to the use of KF to estimate local minimum/maximum temperatures from low resolution weather forecast (Libonati et al., 2007). As far as skill scores are concerned, the results shown in Table 3 confirm the capability of KF and HF to capture exceedance events not revealed by the AQFS, with generally better scores using KF.

To provide further insight into the capability of the AQFS and the bias-corrected forecasts at the Lazio region monitoring sites, Fig. 4 shows the RMSE results for the raw model and the bias-adjusted forecasts. The analysis of this Figure evidences that KF and HF methods better fit PM_{10}, NO_2 and O_3 measurements, with generally lower RMSE obtained using KF. It is worth noting that the RMSE improvement is larger for stations displaying a poor raw model performance. For Rome city stations (left side of Fig. 4), where the raw model provides already a good performance, the RMSE decrease is limited.

The capability of the adopted bias-adjustment techniques to improve the air quality forecasts has also been checked over a longer time period: Fig. S3 presents the 48-hr NO_2 forecasts for 14–15 December 2012 at several urban background/traffic and rural background stations. The analysis of this Figure evidences that the KF and the HF are able to improve the simulated concentrations and that the correction is relevant for the urban traffic stations and particularly for those located outside the Rome domain.

To highlight the performance of the forecasting system at different locations of the target domain, a further analysis has been performed by clustering the monitoring station per type. Tables S1, S2 and S3 show the statistical and scores analysis computed respectively for NO_2, O_3 and PM_{10} for the following subgroups: rural background (RB), background stations (rural, suburban and urban) belonging to the Rome domain (Rome) and urban traffic (UT). These Tables also report the "performance goal" (defined as the level of accuracy that is considered to be close to the best one that a model can be expected to achieve) and the "performance criteria" (defined as the level of accuracy that is considered to be acceptable for regulatory applications) as summarized by Thunis et al. (2011).

4.1.1. NO_2

The analysis of Table S1 evidences a better performance of the AQFS for rural background and Rome background stations. As expected, the statistical parameters and the skill scores for the raw model are generally better for background than for urban traffic sites (UT). This result is due to several factors: the calculated concentrations are averaged within the lowest model layer (20 m) for grid cells having an horizontal resolution of 4 km, while the observed levels are affected by sub-grid scale phenomena and by local emissions that can hardly be estimated (the emission inventory is originally available at province level on an annual basis). The comparison between raw model FAC2 and proposed performance criteria and goal evidences an unsatisfactory performance. Nevertheless, if PM_{10} performance criteria and goals are considered (e.g. [0.40–0.48] for r; [0.60–0.72] for IA and [60–30%] for |MFB|, see Table S3) the performance could be considered acceptable (except for MFB at urban traffic stations). The analysis of skill scores confirms the better capability of the AQFS to forecast the NO_2 exceedances for rural background and Rome background stations. This capability is clearly improved using the bias-adjustment techniques. The statistical parameters and the skill scores for the KF and HF methods are closer to the ideal values and the results at urban traffic stations are significantly improved. As evidenced by Table S1, the adoption of these techniques leads to an increase of the number of exceedances correctly forecasted by the system, but also of the "false alarms" and consequently of the FAR score. The very high number of "correct rejects" leads to similar ACC values (close to 1) for the raw model and the two correction techniques. The comparison between KF and HF statistical parameters and skill scores evidences a better performance of KF (except for the MFB parameter) and a similar capability to forecast NO_2 exceedances (similar POD scores).

4.1.2. O_3

The analysis of Table S2 shows a good performance of the AQFS with respect to this pollutant. The statistical parameters meet the proposed performance criteria and for some indicators and groups of stations also satisfy the performance goals. This result is mainly due to the ubiquitous nature of this pollutant that is less influenced by uncertainties in local emission estimations. The analysis of skill scores confirms the capability of the AQFS to forecast the O_3 exceedances as indicated by the high POD values (except for RB stations due to the underestimation of ozone concentrations at Fontechiari and Leonessa rural background sites; see the higher RMSE presented in Fig. 4). The use of bias-adjustment methods further improves the forecast performance and the detection of ozone exceedances (the POD score increases above 70% for the three station types). The comparison between KF and HF computed statistical parameters and skill scores indicates a similar capability to forecast ozone exceedances and a lower number of "false alarms" using KF. As for NO_2, the very high number of "correct rejects" leads to similar ACC values (close to 1). A further increase of the ACC values obtained using KF and HF is due to both the increase of "hits" and the diminution of "correct rejects".

4.1.3. PM_{10}

The analysis of Table S3 evidences, as in the case of NO_2, a better performance of the AQFS at rural background and Rome background stations. For those two stations subgroups, the statistical parameters satisfy the performance goal, while for urban traffic sites the performance criteria are satisfied only for the FAC2 and MFB.

The analysis of skill scores confirms the better capability of the AQFS to forecast the PM_{10} exceedances at rural background and Rome background stations. This capability is improved using the bias-adjustment techniques. As for NO_2, the statistical parameters and the skill scores for the KF and HF methods are closer to the ideal values and the results for the urban traffic stations are significantly improved. As shown by Table S3, the adoption of these techniques leads to an increase of the exceedances actually forecasted by the system but also of the "false alarms" and consequently of the FAR score. The comparison between KF and HF statistical parameters and skill scores evidences a better performance of KF. The two techniques exhibit a similar capability to forecast PM_{10} exceedances.

5. Summary and conclusions

The potential of two bias adjustment techniques to improve forecast accuracy has been tested on the operational forecasts produced over the whole year 2012 by the Lazio Region AQFS. The techniques are the so-called hybrid forecast (HF), that combines the most recent observed concentration with model-predicted tendency for the subsequent forecast time period, and the Kalman filter (KF), that takes into account the temporal variation of forecast errors at specific locations. A key parameter in the KF approach is the error ratio, which determines the relative weighting of

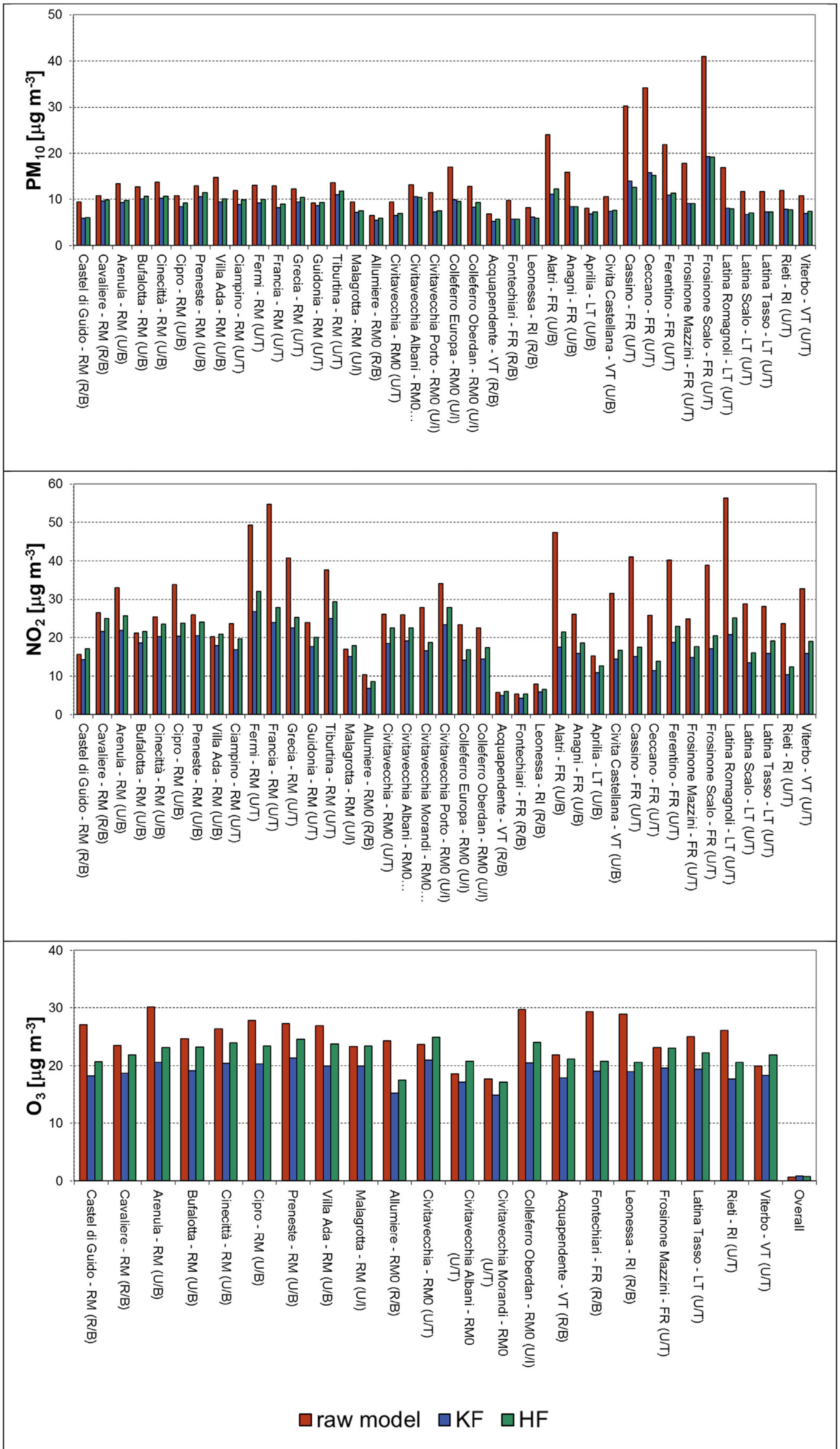

Fig. 4. RMSE values between observed and predicted PM$_{10}$ (1-day average), NO$_2$ (1-hr average) and O$_3$ (1-hr average) concentrations for Lazio Region monitoring network. RM0 indicates Rome province stations outside the Rome domain. Following station classification is used: U/T (Urban traffic); U/I (Urban/Industrial), U/B (Urban background) and R/B (Rural background).

observed and predicted pollutants concentrations. We selected the "optimal" values for this parameter searching for the values that minimize the RMSE metric, considering all monitoring sites and a calibration period of three months lasting from January to March 2012. The following "optimal" ratios have been found: 0.4 for PM_{10} and 0.16 for NO_2 and O_3. The analysis of bias correction results has been then performed over the period subsequent the calibration time window (i.e. from April to December 2012) for all available monitoring stations. A thorough evaluation of the AQFS and of the two bias-adjustment techniques has been performed by means of statistical parameters and skill scores.

As a baseline, the analysis performed considering all Lazio region monitoring sites evidenced a performance of the AQFS raw forecast comparable with that provided by similar systems, with better results particularly for ozone and PM_{10}; in addition to that, the analysis of bias-adjusted forecasts pointed out to the capability of KF and HF to capture exceedances events not revealed by the AQFS, with better scores obtained using KF. For what concerns the behaviour in different parts of the region, the RMSE analysis confirmed the good performance of the AQFS at monitoring sites inside the Rome urban area, and the capability of KF and HF methods to improve the fit with PM_{10}, NO_2 and O_3 measurements. A further analysis has been performed by clustering the monitoring station as follows: all rural background stations, background stations (rural, suburban and urban) belonging to Rome domain, and all urban traffic stations. A better performance of the AQFS has been obtained for NO_2 and PM_{10} at rural background and Rome background stations, while for ozone a good performance of the AQFS has been obtained for all considered station subgroups. The use of bias-adjustment methods improved the forecast performance and the detection of exceedances for the three considered pollutants. The easy implementation of these techniques encourages their application for areas evidencing a poor performance of the adopted AQFS; vice versa, for areas showing a good performance of the AQFS, their application should be checked in advance to avoid possible overestimation of the exceedances. In this work we considered a 3-month training period to select the optimal values for the KF error ratio. Since this parameter can vary with season its determination for each season and for each pollutant could be recommended for practical implementations of this techniques. Once the optimal error ratios have been estimated a short training period, lasting some weeks, is actually needed since the KF approach adapts its coefficients at each iteration (Delle Monache et al., 2006).

The difficulties in simulating NO_2 and PM_{10} levels at urban traffic stations can be ascribed to the horizontal resolution adopted by the model and to uncertainties in emissions estimation. These results evidence the need to improve the emission inventory, particularly for the areas outside Rome city, and to increase the horizontal resolution for the Lazio region domain, to possibly better capture local scale phenomena not reproduced at present spatial resolution.

The promising results obtained using KF and HF convinced the Regional Air Quality Agency to implement these bias-adjustment techniques in the operational AQFS. The KF is now operational and the combined use of both bias-adjusted techniques is also foreseen. During the last year, further activities have been performed in order to improve the regional emission inventory, to apply a different prognostic meteorological model (WRF) and to include two other inner domains having a horizontal resolution of 1 km. The results obtained by these activities will provide useful information to further investigate current AQFS uncertainties and to improve its forecast skill.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

Acknowledgements

The development of the operational AQFS and the related research, including the one reported in this paper, were funded by Lazio Region Government. The authors are grateful to Luca Delle Monache and Roland Stull for providing their original Kalman filter source code used in this work. The authors wish to thanks Lorenzo Mauri (Gexcon AS, formerly Arianet S.r.l.) for his support in implementing Kalman filter in the AQFS and to Paola Radice (Arianet S.r.l.) for preparing emission data needed by QualeAria and Lazio Region forecast systems. The authors would also like to thank the anonymous reviewers for their valuable comments and suggestions.

Appendix A. Statistical parameters

Table A1
Performance metrics.

Parameter	Formula
RMSE [0,+∞)	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(f_i - o_i)^2}$
FAC2 [0,+1]	fraction of data that satisfy $0.5 \le f_i/o_i \le 2$
r [0,+1]	$\frac{\frac{1}{N}\sum_{i=1}^{N}(f_i-\bar{f})(o_i-\bar{o})}{\left[\frac{1}{N}\sum_{i=1}^{N}(f_i-\bar{f})^2\right]^{1/2}\left[\frac{1}{N}\sum_{i=1}^{N}(o_i-\bar{o})^2\right]^{1/2}}$
IA [0,+1]	$1-\frac{\sum_{i=1}^{N}(f_i-o_i)^2}{\sum_{i=1}^{N}(\lvert f_i-\bar{o}\rvert+\lvert o_i-\bar{o}\rvert)^2}$
MFB [-2,+2]	$\frac{2}{N}\sum_{i=1}^{N}\frac{(f_i-o_i)}{(f_i+o_i)}$

Note: f_i = simulation, o_i = observation, N = monitoring sites. A perfect model would have RMSE = 0, FAC2 = 1; r = 1; IA = 1 and MFB = 0.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.apr.2015.04.002.

References

ARPA Lazio (Lazio Region Environmental Protection Agency), 2010. http://www.arpalazio.net/main/aria/sci, accessed in April 2015.
Binkowski, F.S., 1999. The aerosol portion of models-3 CMAQ. In: Byun, D.W., Ching, J.K.S. (Eds.), Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System, pp. 1–23. EPA-600/R-99/030.
Binkowski, F.S., Roselle, S.J., 2003. Models-3 community multiscale air quality (CMAQ) model aerosol component 1. Model description. J. Geophys. Res. 108 (D6), 4183.
Borrego, C., Monteiro, A., Ferreira, J., Miranda, A.I., Costa, A.M., Carvalho, A.C., Lopes, M., 2008. Procedures for estimation of modelling uncertainty in air quality assessment. Environ. Int. 34, 613–620.
Borrego, C., Monteiro, A., Pay, M.T., Ribeiro, I., Miranda, A.I., Basart, S., Baldasano, J.M., 2011. How bias-correction can improve air quality forecasts over Portugal. Atmos. Environ. 45, 6629–6641.
Carter, W.P.L., 2000. Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment. Final Report to California Air Resources Board, Contract 92–329 and 95–308. SAPRC, University of California, Riverside, CA.
Caserini, S., Fraccaroli, A., Monguzzi, A.M., Moretti, M., Angelino, E., 2007. New insight into the role of wood combustion as key PM source in Italy and in Lombardy region. In: Proc. of 16th Annual International Emissions Inventory Conference "Emission Inventories: Integration, Analysis, and Communications" Raleigh, North Carolina, May 14 – 17, 2007.
Chang, J.C., Hanna, S.R., 2004. Air quality model performance evaluation. Meteorol. Atmos. Phys. 87, 167–196.
COST ES0602, 2007. Action ENCWF: towards a European Network on Chemical Weather Forecasting and Information Systems. http://www.chemicalweather.eu/. accessed in April 2015.
Cotton, W.R., Pielke, R.A., Walko, R.L., Liston, G.E., Tremback, C.J., Jiang, H., McAnelly, R.L., Harrington, J.Y., Nicholls, M.E., Carrio, G.G., McFadden, J.P., 2003. RAMS 2001: current status and future directions. Meteorol. Atmos. Phys. 82, 5–29.
Delle Monache, L., Nipen, T., Deng, X., Zhou, Y., Stull, R.B., 2006. Ozone ensemble forecasts: 2. A Kalman filter predictor bias-correction. J. Geophys. Res. 111, D05308.
Delle Monache, L., Wilczak, J., Mckeen, S., Grell, G., Pagowski, M., Peckham, S., Stull, R., Mchenry, J., McQueen, J., 2008. A Kalman filter bias correction method

applied to deterministic, ensemble averaged, and probabilistic forecasts of surface ozone. Tellus, B 60, 238–249.
D'Elia, I., Bencardino, M., Ciancarella, L., Contaldi, M., Vialetto, G., 2009. Technical and non technical measures for air pollution emission reduction: the integrated assessment of the regional air quality management plans through the Italian national model. Atmos. Environ. 43, 6182–6189.
European Commission, 2008. Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on Ambient Air Quality and Cleaner Air for Europe. Off. J. Eur. Comm. Technical Report 2008/50/EC, L152.
Baklanov, A., Clappier, A., Fay, B., Joffre, S., Karppinen, A., Ødegard, V., Slørdal, L.H., Sofiev, M., Sokhi, R.S., Stein, A., 2005. In: Finardi, S. (Ed.), Improved Interfaces and Meteorological Pre-processors for Urban Air Pollution Models, p. 55. FUMAPEX Report D5.2e3, Milan, Italy. http://fumapex.dmi.dk.
Finardi, S., D'Allura, A., Silibello, C., Radice, P., Stocchi, P., Bolignano, A., Morelli, M., Sozzi, R., 2010. Forecast and near real time air quality modeling systems for Rome metropolitan Area: description and preliminary performance evaluation. In: 13th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Paris, France, 1–4 June 2010.
Gariazzo, C., Silibello, C., Finardi, S., Radice, P., Piersanti, A., Calori, G., Cucinato, A., Perrino, C., Nussio, F., Cagnoli, M., Pelliccioni, A., Gobbi, G.P., Di Filippo, P., 2007. A gas/aerosol air pollutants study over the urban area of Rome using a comprehensive chemical transport model. Atmos. Environ. 41, 7286–7303.
Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P.I., Geron, C., 2006. Estimates of global terrestrial isoprene emissions using MEGAN (Model of emissions of gases and aerosols from nature). Atmos. Chem. Phys. 6, 3181–3210.
Homleid, M., 1995. Diurnal corrections of short-term surface temperature forecasts using the Kalman filter. Weather Forecast. 10, 689–707.
ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale), 2015. SINAnet: Rete del Sistema Informativo Nazionale Ambientale. http://www.sinanet.isprambiente.it/it/sia-ispra/serie-storiche-emissioni/. accessed in April 2015.
IIASA (International Institute for Applied Systems Analysis), 2001. RAINS-Europe. http://www.iiasa.ac.at/~rains/home.html. accessed in April 2015.
Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. J. Basic Eng. 82, 35–45.
Kang, D., Mathur, R., Rao, S.T., Yu, S., 2008. Bias adjustment techniques for improving ozone air quality forecasts. J. Geophys. Res. 113, D23308.
Kang, D., Mathur, R., Trivikrama Rao, S., 2010. Assessment of bias-adjusted PM2.5 air quality forecasts over the continental United States during 2007. Geosci. Model Dev. 3, 309–320.
Kukkonen, J., Olsson, T., Schultz, D.M., Baklanov, A., Klein, T., Miranda, A.I., Monteiro, A., Hirtl, M., Tarvainen, V., Boy, M., Peuch, V.-H., Poupkou, A., Kioutsioukis, I., Finardi, S., Sofiev, M., Sokhi, R., Lehtinen, K.E.J., Karatzas, K., San José, R., Astit, M., 2012. A review of operational, regional-scale, chemical weather forecasting models in Europe. Atmos. Chem. Phys. 12, 1–87.
Libonati, R., Trigo, I., Da Camara, C.C., 2007. Correction of 2 m-temperature forecasts using Kalman filtering technique. Atmos. Res. 183–197.
MINNI (Modello Integrato Nazionale a supporto della Negoziazione Internazionale sui temi dell'inquinamento atmosferico), 2008. http://www.minni.org, accessed in April 2015.
Mircea, M., Ciancarella, L., Briganti, G., Calori, G., Cappelletti, A., Cionni, I., Costa, M., Cremona, G., D'Isidoro, M., Finardi, S., Pace, G., Piersanti, A., Righini, G., Silibello, C., Vitali, L., Zanini, G., 2014. Assessment of the AMS-MINNI system capabilities to predict air quality over Italy for the calendar year 2005. Atmos. Environ. 84, 178–188.
Nanni, A., Radice, P., 2004. Sensitivity analysis of three EF methodologies for PM10 in use with climatological dispersion modelling in urban Italian study cases. In: Proc. of 9th Int. Conf. on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, 1–4 June 2004, Garmisch-Partenkirchen (Germany), vol. 1, pp. 309–314.
Pay, M.T., Martínez, F., Guevara, M., Baldasano, J.M., 2014. Air quality forecasts on a kilometer-scale grid over complex Spanish terrains. Geosci. Model Develop. 7, 1979–1999.
QualeAria, 2015. Sistema per la Previsione della Qualità dell'Aria. http://www.qualearia.it. accessed in April 2015.
Roeger, C., Stull, R.B., McClung, D., Hacker, J., Deng, X., Modzelewski, H., 2003. Verification of mesoscale numerical weather forecast in mountainous terrain for application to avalanche prediction. Weather Forecast. 18, 1140–1160.
Sicardi, V., Ortiz, J., Rincón, A., Jorba, O., Pay, M.T., Gassó, S., Baldasano, J.M., 2012. Assessment of Kalman filter bias-adjustment technique to improve the simulation of ground-level ozone over Spain. Sci. Total Environ. 416, 329–342.
Silibello, C., Calori, G., Brusasca, G., Giudici, A., Angelino, E., Fossati, G., Peroni, E., Buganza, E., 2008. Modelling of PM10 concentrations over Milan urban area using two aerosol modules. Environ. Model. Softw. 23, 333–343.
Silibello, C., Calori, G., Costa, M.P., Dirodi, M., Mircea, M., Radice, P., Vitali, L., Zanini, G., 2012. Benzo[a]pyrene modelling over Italy: comparison with experimental data and source apportionment. Atmos. Pollut. Res. 3, 399–407.
Thunis, P., Georgieva, E., Pederzoli, A., 2011. The DELTA Tool and Benchmarking Report Template. Concepts and User's Guide. Joint Research Centre, Ispra. http://fairmode.ew.eea.europa.eu/models-benchmarking-sg4/delta_tool_concepts_userguide.pdf.
Vautard, R., Bessagnet, B., Chin, M., Menut, L., 2005. On the contribution of natural Aeolian sources to particulate matter concentrations in Europe: testing hypotheses with a modelling approach. Atmos. Environ. 39, 3291–3303.
Zhang, K.M., Knipping, E.M., Wexler, A.S., Bhave, P.V., Tonnesen, G.S., 2005. Size distribution of sea-salt emissions as a function of relative humidity. Atmos. Environ. 39, 3373–3379.