Atmospheric Environment 43 (2009) 4050–4059

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Testing the capability of the chemistry transport model LOTOS-EUROS to forecast PM10 levels in the Netherlands

A.M.M. Manders [a,b,*], M. Schaap [a], R. Hoogerbrugge [b]

[a] TNO, Unit Environment, Health and Safety, P.O. Box 80015, 3508 TA Utrecht, The Netherlands
[b] National Institute for Public Health and the Environment (RIVM), P.O. Box 1, 3720 BA Bilthoven, The Netherlands

ARTICLE INFO

Article history:
Received 8 January 2009
Received in revised form
29 April 2009
Accepted 5 May 2009

Keywords:
Air quality forecasting
Particulate matter
Chemical transport models
Model evaluation

ABSTRACT

Since particulate matter has a direct and adverse impact on public health, a good air quality forecast is important. Several European countries presently use statistical forecasting models, which have their limitations, especially for PM10. An alternative approach is to use a chemistry transport model. Here, the ability of the chemical transport model LOTOS-EUROS to forecast PM10 concentrations in the Netherlands was investigated. LOTOS-EUROS models several PM10 components individually. For sulphate, nitrate and ammonium aerosol the evaluation against observations shows that the modelled annual mean concentrations are within 20% of the measured concentration and that the temporal correlation is reasonably good ($R > 0.6$). For sea salt the model tended to overestimate the measured concentrations. For elemental carbon the correspondence with black smoke observations was reasonable. However, total PM10 is seriously underestimated, due to unmodelled components (secondary organic aerosols, mineral dust) and missing sources. Therefore, a simple bias correction for four seasons was derived based on the years 2004–2006. The model was compared with the Dutch operational statistical model PROPART and ground-level observations. With bias correction, LOTOS-EUROS performed better than PROPART regarding the timing of events. The major flaw of LOTOS-EUROS was that high values ($>50\ \mu g\ m^{-3}$) were still underestimated. Another advantage of LOTOS-EUROS over the statistical model was the more detailed information in space and time, which facilitates communication of the forecast to the general public.

1. Introduction

While air quality in Europe has improved substantially over the past decades, air pollution still poses a significant threat to human health (EEA, 2007). Health effects of air pollution are dominated by particulate matter (PM), both PM2.5 and PM10. As the long-term exposure to PM is posed to cause a substantial reduction in life expectancy, it is thought to have major significance to public health (EEA, 2007). In addition, short-term exposure to PM has frequently been associated with inflammatory reactions in the lung, respiratory symptoms, adverse effects on the cardiovascular system and increases in hospital admissions and mortality (e.g., Brunekreef and Holgate, 2002). Children, senior citizens and people with heart and lung conditions are especially sensitive. Hence, the knowledge of current smog conditions and accurate short-term forecasts are important to plan their activities.

In the Netherlands, the National Institute for Public Health and the Environment (RIVM) monitors the air quality from day to day, and issues a smog forecast to civil authorities and the public. Currently, the PM10 prognosis is made using a statistical model named PROPART (Noordijk, 2003). Such a statistical approach is common practice in many European countries (e.g. Stadlober et al., 2008). Alternatively, neural networks are used (e.g. Kukkonen et al., 2003; Hooybergh et al., 2005), which are closely related and yield similar results. PROPART uses a limited number of input parameters (today's observed daily average concentration, temperature, rain, wind direction and wind speed in combination with the prognosis for the meteorological parameters for tomorrow) to generate a limited output (tomorrow's daily average concentrations at monitoring locations). The statistics are tailored to the Dutch monitoring locations. The accuracy of this relatively simple method is limited. Often it does not reproduce a large day-to-day change in PM10 concentration, whereas these changes are characteristic for this pollutant. Overall, the performance of PROPART is only slightly better than assuming persistence (Manders et al., 2008).

Given the impact of PM10 concentrations on public health, an improved forecast would be highly beneficial. A chemistry transport model (CTM) could be an alternative for the statistical models. Statistical PM forecasting models describe average relations

* Corresponding author. Tel.: +31 88 866 2260; fax: +31 88 866 2044.
E-mail address: astrid.manders@tno.nl (A.M.M. Manders).

1352-2310/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2009.05.006

between local PM concentrations and local atmospheric conditions. As a consequence, they average out subtle changes in air mass origin and history (other than local wind direction), and cannot capture the contribution of distant weather-dependent sources and circumstances favourable for the formation of secondary PM. In contrast, chemistry transport models provide a deterministic method to forecast PM levels. In the last decades a number of CTMs has been developed to model ambient PM concentrations (e.g. Schaap et al., 2004a,b; Bessagnet et al., 2004; Karydis et al., 2007). Although large uncertainties are still associated with the modelling of particulate matter, these systems are being used or set-up for air quality forecasting (e.g. McKeen et al., 2007; Hollingsworth et al., 2008). Compared to a statistical model, a CTM provides a means of incorporating a description of the air parcel history in the forecast.

In this study we explore the use of the CTM LOTOS-EUROS for forecasting PM concentrations for the Netherlands. The question that we like to answer is whether LOTOS-EUROS yields more reliable forecasts than PROPART. Using a CTM we restrict ourselves to regional background concentrations. Though most people live in cities, this appears justified as the temporal variability of PM10 inside urban areas is largely determined by the variability in the regional background concentrations. The absolute variability in local source contributions is much less. In Section 2 we describe the two models and the monitoring data used in this study. In Section 3 we present a validation of the LOTOS-EUROS model. We also present a methodology to account for the systematic underestimation of PM10. Such underestimations are found for nearly all current CTMs (e.g. Schaap et al., 2004b; Karydis et al., 2007; Stern et al., 2008). In Section 4 we compare LOTOS-EUROS and PROPART results. Finally, the results are discussed in Section 5.

2. Methodology

2.1. LOTOS-EUROS

We have investigated the capability of the regional air quality model LOTOS-EUROS (Schaap et al., 2008) to forecast PM concentrations and in particular episodes with high concentrations. The LOTOS-EUROS model is a 3D chemistry transport model aimed to simulate air pollution in the lower troposphere. The model has been used for the assessment of particulate air pollution in a number of studies directed to total PM10 (Denby et al., 2008; van Zelm et al., 2008), secondary inorganic components (Schaap et al., 2004a, Erisman et al., 1994; Barbu et al., 2008), primary carbonaceous components (Schaap et al., 2004b; Schaap and Denier van der Gon, 2007) and trace metals (Denier van der Gon et al., 2007). The model has participated frequently in international model comparisons addressing ozone (van Loon et al., 2007; Hass et al., 1997) and particulate matter (Cuvelier et al., 2007; Hass et al., 2003; Stern et al., 2008; Schaap et al., 2008). For a detailed description of the model we refer to these studies. Here, we describe the model simulation used in this study.

First, we performed a simulation on a European domain bound at 35° and 70° North and 10° West and 40° East. The grid resolution in this domain is 0.50° longitude × 0.25° latitude, which is approximately 25 × 25 km over the Netherlands. Using a one-way zoom option a high resolution simulation over the Netherlands and its direct surroundings with an increase in resolution of a factor 4 has been obtained. The model is forced using ECMWF meteorology and anthropogenic emissions (GEMS emission database, Visschedijk and Denier van der Gon, 2005). From the primary fine particles (PPM2.5) the amount of elemental carbon (EC) is subtracted following Schaap et al. (2004b) and EC is included in the model as a separate tracer. Natural sea salt emissions (SS) are calculated following Monahan et al. (1986). Hourly concentrations of all particulate components are stored. The PM10 concentrations were diagnosed as the sum of the separate model components:

$$\mathrm{PM10} = \mathrm{SO_4^{2-}} + \mathrm{NO_3^-} + \mathrm{NH_4^+} + \mathrm{PPM2.5} + \mathrm{PPMcoarse} + \mathrm{EC} + \mathrm{SS}$$

Crustal matter (CM) and secondary organic aerosols (SOA) are not yet incorporated in the current model version, due to a lack of solid knowledge on emission strengths for CM and formation routes for SOA.

2.2. PROPART

PROPART was developed at RIVM and has been used since 1996 for the operational 1 day smog forecast in the Netherlands. PROPART is a statistical model, based on the idea that a forecast of PM10 concentrations for tomorrow can be constructed from today's observed concentration by multiplying it by a factor. This factor was constructed from measurements from the past and depends on today's observed concentration, the station type (traffic oriented, urban, and rural) and today's meteorological conditions and their forecast. The meteorological variables that are used in PROPART are daily mean wind speed and direction at 10 m above ground, daily minimum and maximum temperature at 2 m above ground, daily accumulated rain fall and duration. Each variable is subdivided into classes to model its impact on the forecast. In this way a decision tree is made, the contribution of each variable and subclass is determined on a statistical basis. For the meteorological variables, one value per meteorological variable per day is used for the whole country. This value represents an average over the country, which is especially problematic for rain since rain strongly influences the concentration but can be a highly discrete phenomenon.

Defining and building the statistics is the heart of the model and relatively time consuming. To build statistics, observations over several years are used and they only need to be updated after some years. Once statistics are available, the forecast can be made by looking up similar conditions in the past and applying the corresponding multiplication factor, which is a fast computation. PROPART provides forecasts of daily mean concentrations at the Dutch monitoring locations. Hence, the forecast has very limited spatial information. We refer to Noordijk (2003) for model details and Manders et al. (2008) for an evaluation.

2.3. Monitoring data

To verify the modelled concentrations in the Netherlands we use monitoring data for 2004–2006 from the national air quality network as operated by the RIVM (Beijk et al., 2007). For the evaluation of LOTOS-EUROS, only rural background stations were used, since the resolution of LOTOS-EUROS is representative for this scale. The network provides data for 16 rural stations for PM10 (see Fig. 1). The PM10 concentrations in the network were measured using an Anderson FH 62 I–N beta attenuation monitor. The measurement data are calibrated on the European reference method for PM10 measurements, resulting in a calibration factor of 1.17 for the slope and 2.7 μg m^{-3} for the intercept. For the calibrated data equivalence with the reference method has been demonstrated resulting in a measurement uncertainty for daily measurements of 17% (Beijk et al., 2008). All measured data were corrected for sampling losses using the standard procedures followed by RIVM. Furthermore, secondary inorganic aerosol concentrations are obtained for six rural locations. Finally, we used black smoke (BS) measurements to assess the model performance for EC. EC concentration data were estimated using the empirical relation between black smoke and EC as determined by Schaap and

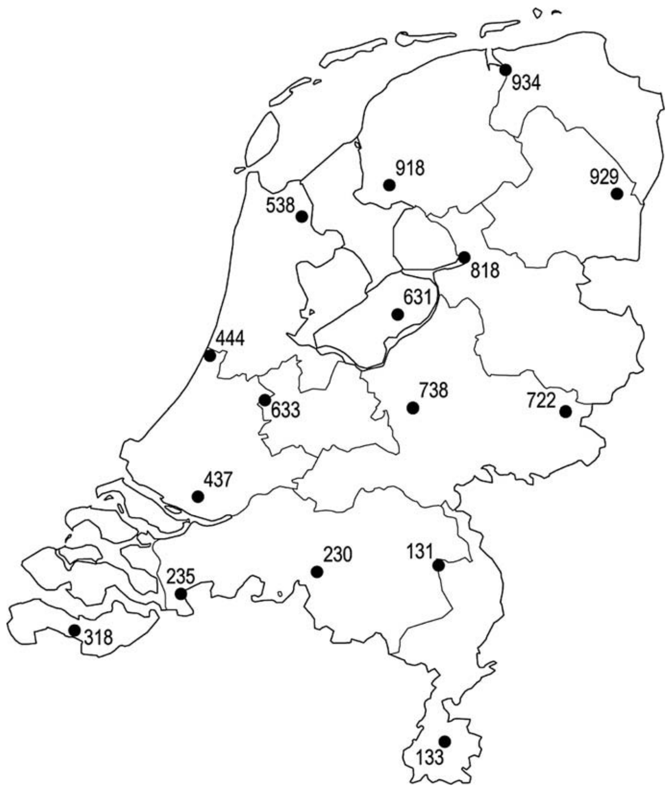

Fig. 1. Locations of the measurement sites used in this study. The numbers indicate the station code within the national air quality network (LML).

Denier van der Gon (2007). These authors found the relation based on simultaneous measurements of EC and BS at two regional sites (131 and 444) to be linear and highly correlated. As different analysis procedures for EC yield systematically different levels (up to a factor 2) (Ten Brink et al., 2004) the information is mainly used to assess the temporal performance of the model.

The results of the model at the abovementioned stations have been evaluated for PM10 using a range of statistical parameters. These are: average, root mean square error, residue, standard deviation of error, correlation, hit rate, and correct alarm (in the concentration interval 50–200 $\mu g\ m^{-3}$).

3. LOTOS-EUROS validation

3.1. Modelled spatial distributions

In Fig. 2 we present the annual mean modelled PM10 distribution for the European modelling domain and the Netherlands for 2005. The European domain shows maxima over the Benelux, Poland and the Po-Valley (Italy). In these areas the modelled concentrations exceed 15 $\mu g\ m^{-3}$. Similarly, large cities of agglomerations are recognizable with similar concentrations. Concentrations of 7.5–15 $\mu g\ m^{-3}$ are found for a band over north-western Europe, central and south-eastern Europe. In general, the modelled PM10 concentrations decline from central Europe to northern Scandinavia and towards the Iberian Peninsula.

The zoom domain provides a more detailed picture for the modelled concentrations over the Netherlands, which show levels between 10 and 20 $\mu g\ m^{-3}$. The lowest concentrations are modelled for the north east of the Country. Concentrations between 14 and 18 $\mu g\ m^{-3}$ are calculated for the populated and industrialised western part of the Netherlands and over the river area stretching towards the Ruhr area in Germany. Note that the concentrations over the Belgian cities are modelled to be higher than those in the Netherlands. Minimum concentrations over the zoom area are calculated over the forest regions of the Ardennes and Germany, south-east of the Ruhr-area.

3.2. Composition of PM

We have paired the observations available from the Dutch monitoring network to modell concentrations on a daily basis. Before we evaluate the total PM10 concentration we firstly examine its components, to the extent possible. Table 1 shows the mean and standard deviation of observed and modelled daily mean PM concentrations at monitoring locations, together with the correlation of observations and model results.

For the modelled secondary inorganic components the modelled annual mean concentrations are close to the observations. The contribution of nitrate (NO_3^-) is slightly overestimated by LOTOS-EUROS but the variability in the daily mean concentrations is quite comparable with the observations. Sulfate (SO_4^{2-}) is over-estimated or underestimated depending on site, whereas the modelled variability is slightly too small. Ammonium aerosol (NH_4^+) is comparable to or slightly overestimated with variability in accordance with the observations. The temporal correlation of nitrate is slightly better than for sulphate, which may be due to the generally higher concentrations of nitrate (and the associated lower amount of data below the detection limit).

The comparison for sea salt is hampered by the available data. The measurement strategy aimed at secondary inorganic aerosols also includes an analysis of the chloride (Cl^-) content. However, the measurement is not ideal as the available Cl^- observations do not cover the PM10 fraction but approximately the PM3 fraction. Moreover, Cl^- is lost due to chemical reactions in the atmosphere. This means that the sea salt estimate from chloride alone should be lower than the real and therefore the modelled concentrations, which is indeed the case. Lastly, many of the observed concentrations are below the detection limit. Consequently, we interpret the correlation coefficient ($R = 0.54$) given in Table 1 as a lower limit. Comparison with a single month of observed Na concentrations, which is available for April 2005, shows a fairly good absolute correspondence, although LOTOS-EUROS seems to overestimate concentrations in that month.

Primary PMcoarse and PM2.5 cannot be compared with observations as they are bulk components in emission databases. Nevertheless, the mean and variability are presented in Table 1 to complete the overview of modelled concentrations. The composition of these bulk components is a mixture of further unspecified species, except for elemental carbon which is estimated based on an earlier study (Schaap et al., 2004b). However, elemental carbon is not routinely measured in the monitoring network. Consequently, we have compared the elemental carbon model results with black smoke observations. Using the linear relation between black smoke (BS) and EC by Schaap and Denier van der Gon (2007) we have estimated the annual mean EC concentration (from BS = 6.12) at the Dutch sites to be $(0.056 * 6.12 + 0.12 =)$ 0.46. The modelled value is about a factor 2 higher, which is within the uncertainty of the measurements on which the relation was based (Ten Brink et al., 2004). The temporal correlation ($R = 0.75$) between the estimated EC or BS and the modelled values is, however, evident.

The performance of the LOTOS-EUROS system is satisfactory for secondary inorganic aerosol. For both sea salt and black carbon we can only conclude that the temporal variation is well represented. Strong conclusions on the absolute concentrations are not possible besides that the values are not unrealistic. The present findings are

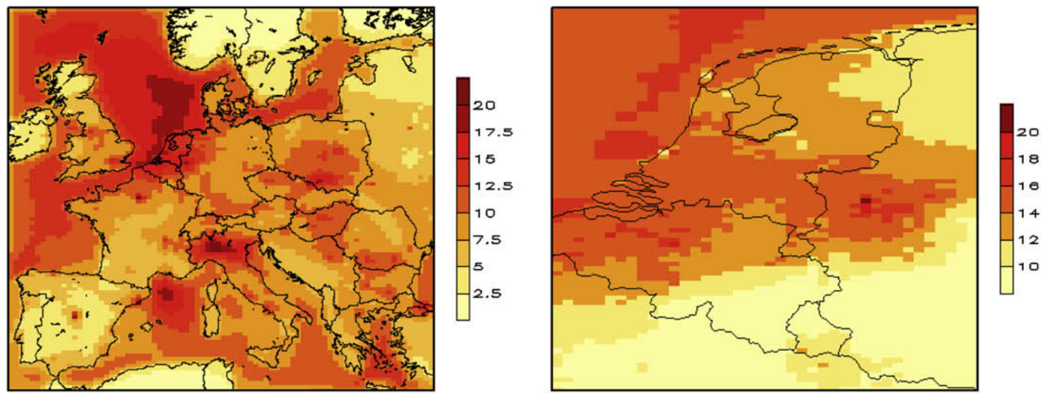

Fig. 2. Annual mean modelled PM10 concentration (μg m^{-3}) for the European domain (left) and the Netherlands (right) for 2005.

in line with earlier comparisons against European data for secondary inorganic aerosols and elemental carbon (Schaap et al., 2004b, 2008).

3.3. Variability of composition

Fig. 3 shows the contribution of several components to total PM10 as a function of total PM10 concentration. There is approximately a linear relationship between total PM10 and the SO_4^{2-} and NH_4^{2-} aerosol, both in the observations and in the LOTOS-EUROS model results. For NO_3^- a slightly more variable contribution is found with enhanced concentrations at the high end of the PM10 range. This results in a nearly constant relative contribution of about 33% for these components. In contrast, sea salt, approximated by using modelled Na^+ and observed Cl^- concentrations, declines with increasing PM10 concentrations. This behaviour is explained by the high sea salt loads in periods with fast transport from the Atlantic and North Sea, whereas high PM10 concentrations are associated with continental flows and periods with low wind speeds. The relative contribution of the primary elemental carbon concentrations tends to decline slightly with increasing PM10 concentrations, which is both observed and modelled. These findings indicate that the species as modelled by LOTOS-EUROS show a behaviour that is very similar to that of the observations over the full range of total PM10 concentrations.

One should note that the measured components do not sum up to the observed total PM10 concentrations. We know that sea salt calculated from Cl^- is an underestimation of the sea salt concentration. But the most important reason is that not all PM10 components were measured. Organic carbon (OC) and elemental composition (tracers for dust, metals) were not available. OC and dust may both contribute 15–20% to the total mass (Querol et al., 2004). In addition, the total PM10 mass measurements used in this study have an uncertainty of 17%. Furthermore, differences in sampling characteristics between the PM monitor, the reference methodology and the separate methodologies for the composition may contribute to mismatches. Most studies find about 20% of unaccounted mass, even tough they do sample such that a complete characterization is possible (e.g. Putaud et al., 2004).

Table 1
Statistical comparison between modelled and observed concentrations of PM10 and its components for the year 2005. We compare modelled and observed mean concentration and standard deviation (st dev) and provide root mean square error (RMSE) and correlation coefficient. All statistical parameters were determined based on daily data for the individual stations and then averaged. Sea salt observations are not well constrained and only the correlation coefficient is given. The same applies for black carbon, for which the observations are estimated from black smoke observations (see text).

Species	Observed		LOTOS-EUROS			
	Mean	St dev	Mean	St dev	RMSE	Correlation
Total PM10	25.58	10.63	13.81	9.68	15.39	0.68
NO_3^-	3.20	3.02	3.70	3.12	2.48	0.70
NH_4^+	1.59	1.29	1.90	1.33	1.05	0.71
SO_4^{2-}	1.98	1.63	2.20	1.39	1.43	0.59
Sea salt	*1.42*	*1.19*	2.81	2.84		*0.54*
Elemental carbon	0.46	0.43	0.89	0.44		0.75
Prim PMcoarse			0.91	0.55		
Prim PM2.5			2.39	1.23		

3.4. PM10 bias characterisation

The modelled PM10 concentration is the sum of the individual model components. Since the model does not include all PM sources and components the model underestimates observed PM10 levels. This feature is a common feature of present day CTMs (Yu et al., 2008; Stern et al., 2008). Like for the modelled components, the temporal correlation ($R = 0.68$) for total PM10 is quite reasonable (see Table 1). To use LOTOS-EUROS to assess the variability in total PM10, the missing fraction must be compensated for. On average, LOTOS-EUROS underestimates the observed PM10 by 46% or 11.7 μg m^{-3}. In Fig. 4 we explore the bias as function of season. In the figure we compare all observations for all sites with the modelled value over the years 2004–2006. The scatter plots indicate that the behaviour is different for summer than for the other three seasons. In summer the variability in PM10 is lower than in the other seasons and the model explains a lower percentage of the observed variability. This is partly explained by a too high variability in modelled ammonium nitrate in summer, which is related to the sensitivity of the ammonium nitrate formation to temperature and precursor emissions. In addition, in summer chemistry is more effective yielding a higher ratio secondary to primary components. Furthermore, the impact of missing primary sources would be less, as the mixing layer is much deeper. Consequently, the model is slightly closer to the observed concentrations as there is no significant bias in the secondary inorganic components. Hence, we have used the seasonal fit parameters to estimate the bias corrected values:

$$\begin{aligned}
PM10_{biascor} &= 1.54{*}PM10 + 8.1 \quad R^2 = 0.52 \quad \text{Winter (DJF)} \\
PM10_{biascor} &= 1.42{*}PM10 + 7.5 \quad R^2 = 0.47 \quad \text{Spring (MAM)} \\
PM10_{biascor} &= 0.76{*}PM10 + 13.5 \quad R^2 = 0.26 \quad \text{Summer (JJA)} \\
PM10_{biascor} &= 1.31{*}PM10 + 9.1 \quad R^2 = 0.51 \quad \text{Fall (SON)}
\end{aligned}$$

We have neglected the variation between stations and parts of the country to keep the procedure simple and transparent.

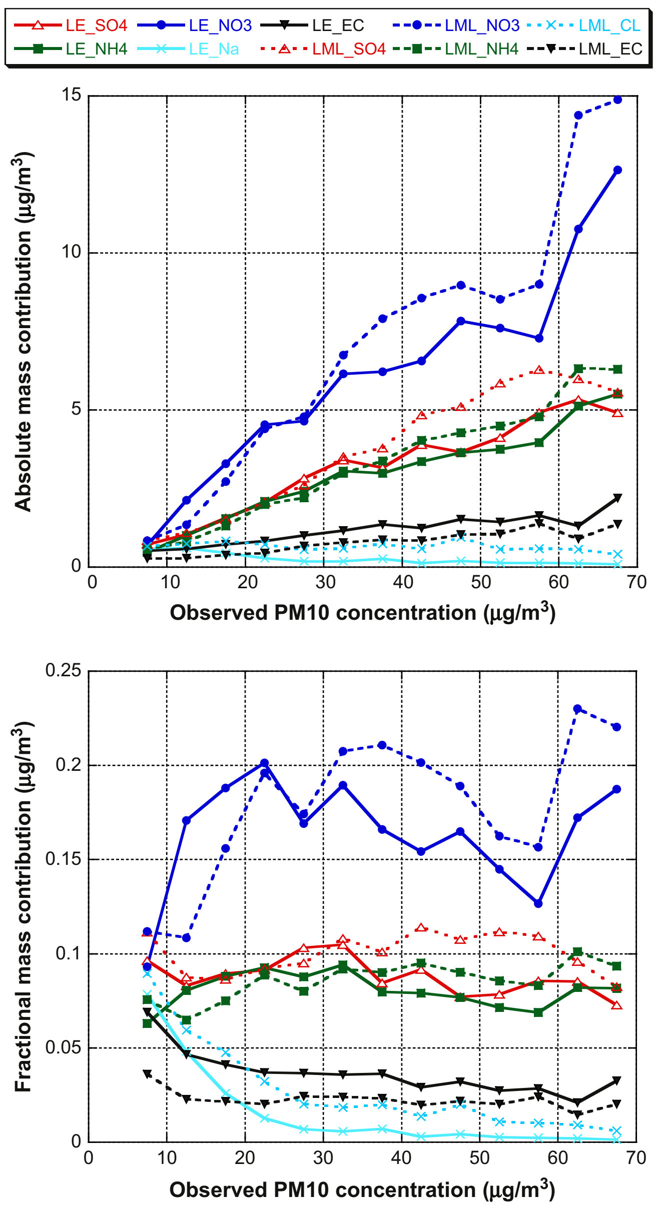

Fig. 3. Modelled (LE) and observed (LML) absolute (top) and relative (bottom) contribution of SO_4^{2-}, NO_3^-, NH_4^+, EC and sea salt in Vredepeel (131) versus observed total PM10. Vredepeel is shown as it is the only Dutch EMEP site with both PM10 and all speciation data available. Note that the observed chloride (Cl^-) and modelled sodium (Na^+) sea salt tracers were not transformed to a sea salt equivalent and are meant to illustrate their behaviour. Observed elemental carbon (EC) concentrations are derived from black smoke measurements.

3.5. Variability with meteorological conditions

Besides emissions, meteorology is the driving force for air quality. In Fig. 5 the modelled concentrations of LOTOS-EUROS are compared with observations as a function of meteorological parameters. Results are averaged over classes and all stations are averaged to focus on average statistical relationships. One should keep in mind that the spread of the individual results within a class is large: the relative standard deviation is around 42% for the observations, PROPART and LOTOS-EUROS and around 31% for LOTOS-EUROS with bias correction, with very little difference between the meteorological variables. For all models and the observations, main features are the clear decrease in concentration with increasing wind speed, the increased concentrations for easterly winds, the decrease in concentration when rain cleans the atmosphere and the relatively high concentrations for both very cold and very warm days. Very cold and very warm days are generally associated with stagnant weather conditions with (very) weak easterly and southerly winds. Note that these relations are statistical and by no means meant to be explanatory as meteorology has a complex interaction with air quality in which none of the abovementioned variables plays an independent role. The bias corrected model captures the general magnitude and variability of the PM10 concentrations as a function of meteorology, albeit with some exceptions. These are the underestimations for easterly winds (slightly overestimated by PROPART) and most prominently the underestimation for days with high temperatures (again slightly overestimated by PROPART). For the summer months, the uncorrected LOTOS-EUROS results hardly increase with increasing temperature, a shortcoming that cannot be repaired by a simple concentration-dependent bias correction. We must conclude that the bias correction effectively compensates for the missing PM10 fraction in most, but not all cases.

4. Model intercomparison

In this section we compare the performance of the statistical PROPART model and LOTOS-EUROS for predicting PM10 levels. As a benchmark we include the persistence model, i.e. using today's concentration as the prediction for tomorrow. Here, we limit ourselves to a 1 day forecast using analysed ECMWF meteorological data to minimize the impact of the uncertainty in the meteorological forecast. For the same reason, the PROPART forecasts were produced using observed meteorology instead of forecasted meteorology. The period 2004–2006 is now studied. First, time series results are shown and interpreted and then statistical parameters are presented and discussed.

4.1. Time series

In Fig. 6 we compare the day by day predictions of the models for a 3 month period in spring, 2005. In this period there are three episodes with concentrations above the daily limit value (50 µg m^{-3}), the first at the 6th , the second at the 28th of March and the third at the 17th of April. All three episodes are characterised by weak southerly or easterly winds. Fig. 7 displays the FLEXTRA back trajectories (http://tarantula.nilu.no/trajectories) for these episodes illustrating two episode types with: (1) stagnant conditions over the Netherlands (6th); and (2) slow transport and build up of air pollutants from continental areas to the south (28th) or east (17th) of the Netherlands. This makes them classical smog episodes. In the first and major episode temperatures are around zero degrees. Both PROPART and LOTOS-EUROS represent the high values around these episodes reasonably well, albeit PROPART tends to be 1 day behind the observed maximum value. LOTOS-EUROS correctly times the day with maximum value but underestimates the high concentrations significantly. In general, LOTOS-EUROS correlates well with the observations but underestimates the variability, i.e. underestimates the highest values and slightly overestimates the lowest values. In contrast, predicted maximum values by PROPART are often 1 day behind the observations but represents the observed variability well. This is caused by the use of today's average for the forecast of tomorrow, which gives PROPART some characteristics of a persistence model. Hence, fast changes in PM10 levels associated with a shift in the large scale synoptic situation are not captured well by PROPART.

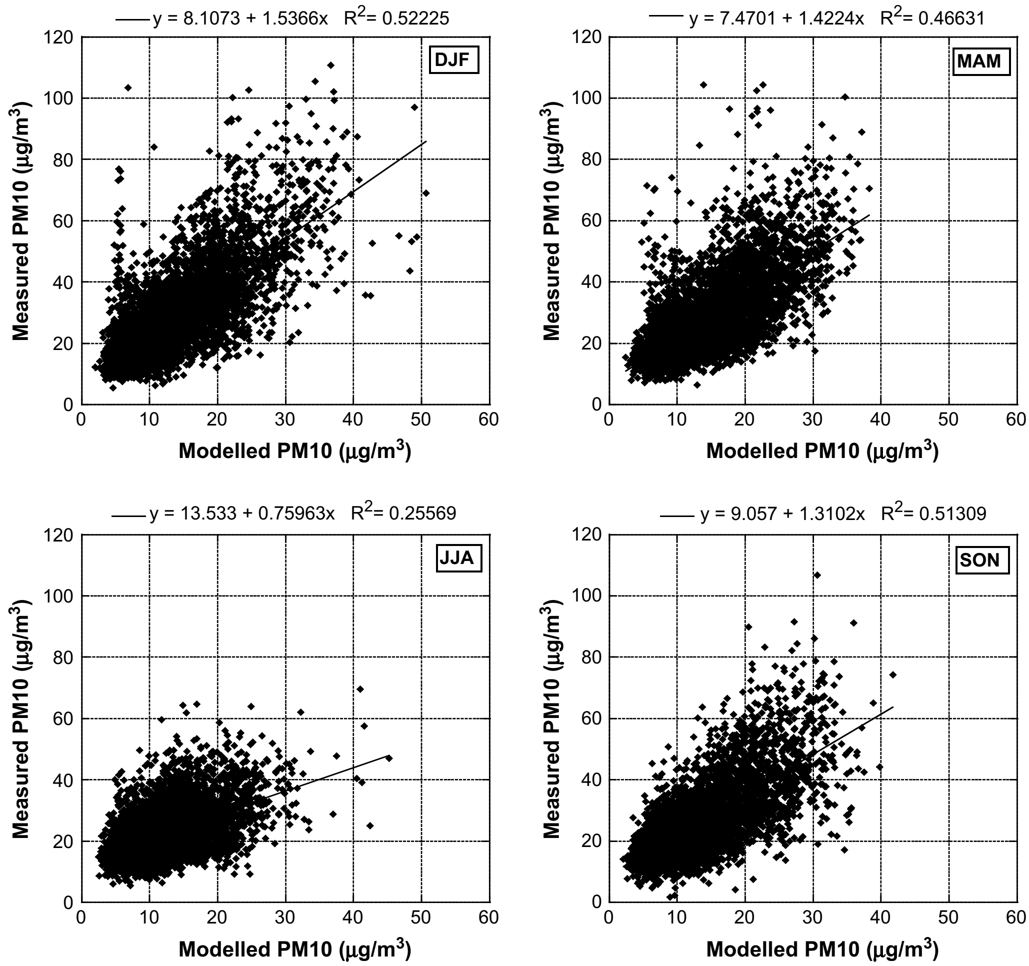

Fig. 4. Comparison between modelled and measured PM10 concentrations for all LML locations. The comparison is made for every season: winter (DJF), spring (MAM), summer (JJA) and fall (SON).

4.2. Statistics

Table 2 contains the values of the most important statistical parameters which are generally used to evaluate the performance of models. In the Appendix, the statistics for individual stations are presented. The mean of the bias corrected LOTOS-EUROS simulations are very close to the observed mean (equal to the mean of persistence), as it should be because of the bias correction. The absolute bias at individual stations varies between 0 and 4 $\mu g\ m^{-3}$. The standard deviation of error, a measure for the non-systematic part of the RMS error, is comparable for the three models. The ability to model the variability of the observations correctly (skill variance) is by definition perfect for persistence. In this respect PROPART is also very good. The bias correction has improved the skill variance of LOTOS-EUROS considerably from 0.5 to 0.7. However, it still means that the bias corrected LOTOS-EUROS underestimates the ambient variability. This is clearly seen for Vredepeel during spring where also with the bias correction one does not capture the high concentrations. Regarding temporal correlation, LOTOS-EUROS performs better than PROPART and persistence. This means that LOTOS-EUROS better predicts changes in concentrations. The bias corrected LOTOS-EUROS also has the smallest root mean squared error value. The hit rate was determined, in this case based on the ability of the model to be within 20% of the observed value. This represents the accuracy of the observed concentrations (Beijk et al., 2007). None of the models comes close to the ideal 100% and differences are small, except for the uncorrected LOTOS-EUROS, which is always too low resulting in a poor hit rate.

For smog forecasting, it is important to predict exceedance values correctly. Therefore, the skill score for values larger than the threshold value of 50 $\mu g\ m^{-3}$ but smaller than 200 $\mu g\ m^{-3}$ was determined. Values larger than 200 $\mu g\ m^{-3}$ rarely occur and were therefore not considered. Within the 3 year period about 1000 values were registered within this interval summed over all the monitoring sites in the Netherlands. The persistence model predicts the total number of exceedance (per definition) correctly as the persistence model is the measured PM time series with 1 day delay. From these predictions 42% was an exceedance day indicating that 58% of the exceedances are single day events. PROPART predicts more exceedances but has a lower percentage of correct predictions than persistence. Hence, in this respect PROPART does not add information. LOTOS-EUROS underpredicts the number of exceedance days by more than a factor 2. On the other hand, the percentage of correct predictions is substantially higher than that of the other models. The reason is associated with the underestimation of the highest peak values by LOTOS-EUROS, which is also expressed in the underestimation of the observed variability (skvar = 0.7). The number of exceedances will therefore be low, but the chance for a correct prediction is higher as one tends to under predict the measurements. Below we investigate an alternative bias correction aimed to optimize the variability and address the skill for predicting exceedance values.

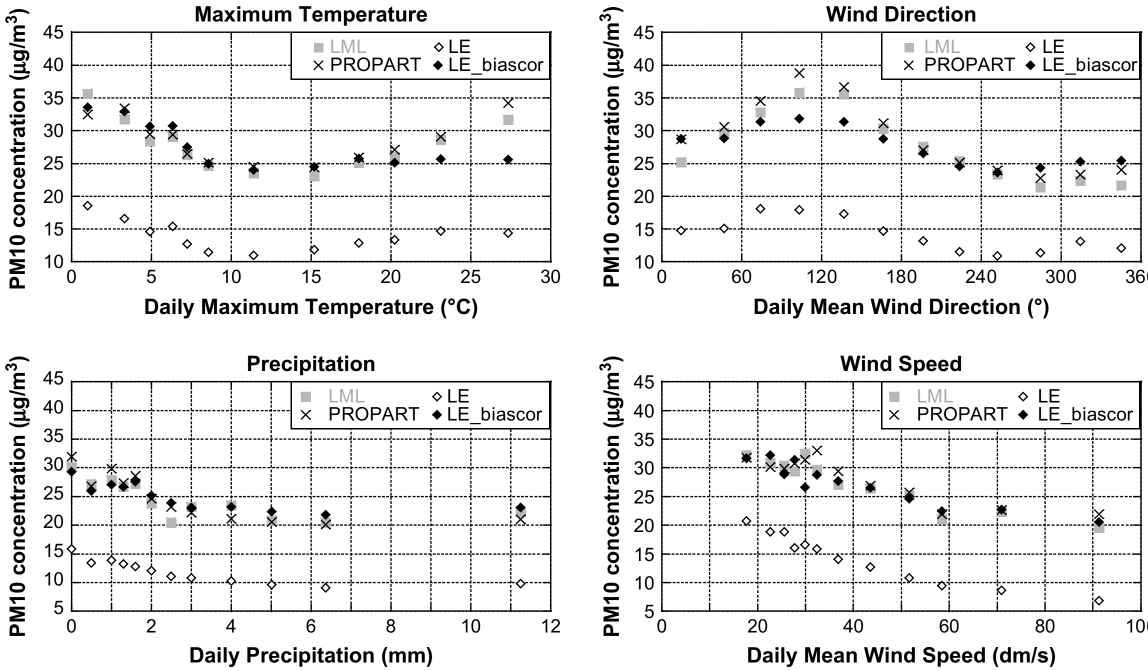

Fig. 5. Observed and modelled PM10 concentration as function of daily maximum temperature (upper left), wind direction (upper right), rain fall (lower left) and wind speed (lower right). All data represent average values for classes over all measurement locations. Average relative standard deviations for the different models were 43% for the observations (LML), 42% for PROPART, 44% for LOTOS-EUROS and 31% for LOTOS-EUROS with bias correction.

4.3. Alternative bias correction

The bias correction presented above provides the correction with the best statistical performance over all cases. Inspection of the scatter plots (Fig. 5) shows that the linear fits are relatively flat and underestimate the concentrations at the high end. This is especially the case for the fall, winter and spring seasons. A simple way to amplify the variability in the bias corrected model for these seasons is to force the fit through zero. The resulting relations are:

$$\begin{array}{lllll} PM10_{biascor} = 2.00{*}PM10 & R^2 = 0.46 & \text{Winter} & (DJF) \\ PM10_{biascor} = 1.85{*}PM10 & R^2 = 0.42 & \text{Spring} & (MAM) \\ PM10_{biascor} = 1.86{*}PM10 & R^2 = 0.40 & \text{Fall} & (SON) \end{array}$$

In Table 2 we present the statistical performance of this bias correction, in which we keep the summer to the old simulation as the fit though zero for the summer yields a very low correlation coefficient. The observed variability in the PM10 measurements is better represented by the model (skvar = 0.89). The slightly higher RMSE and residual for the alternative bias correction indicate that this is at the expense of the general performance. This is caused by the slightly lower correlation coefficient of this bias correction. On the other hand, the alternative bias correction still yields a better performance than PROPART and persistence. Focussing on the exceedance days the alternative bias correction yields twice as many days (725) with a correct prediction percentage of 50%. Hence, the performance for predicting exceedance days can be improved at the cost of a slight decrease in the general performance of the bias corrected model.

5. Discussion and conclusions

We have explored the use of the CTM LOTOS-EUROS for forecasting PM concentrations for the Netherlands in comparison to

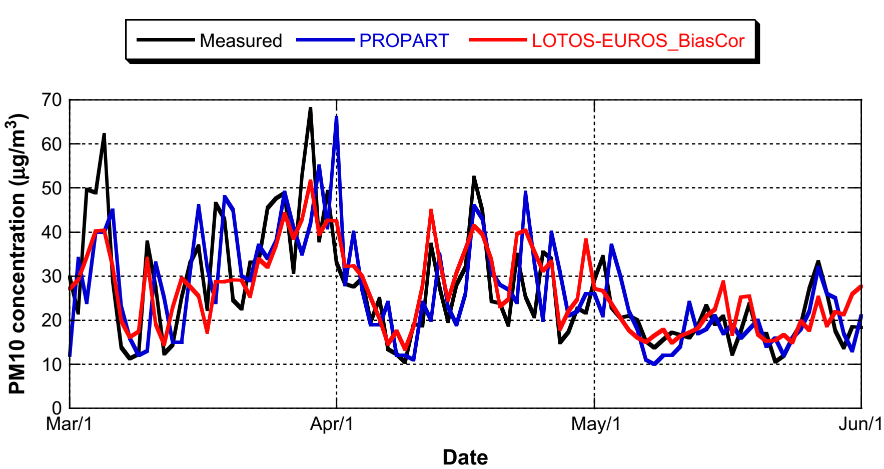

Fig. 6. Observed and modelled daily average PM10 concentration for Vredepeel, spring 2005.

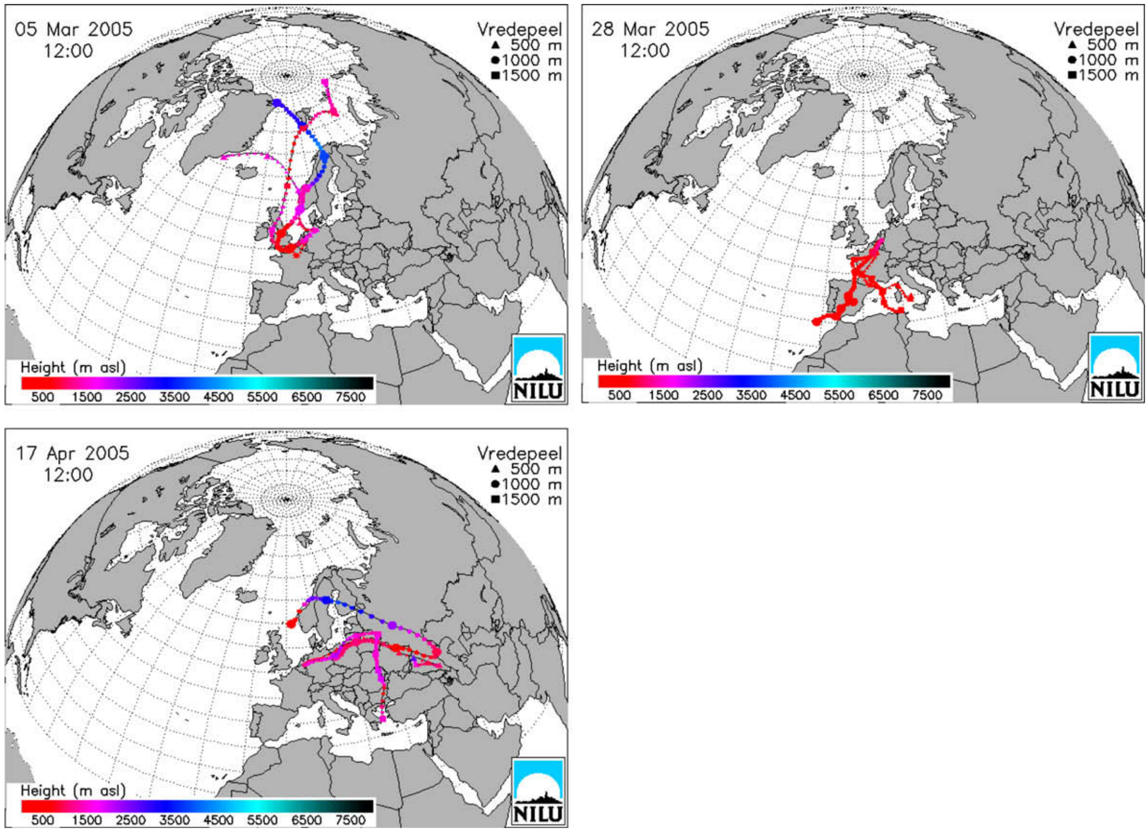

Fig. 7. Trajectories for Vredepeel on the 3 days with high concentrations in both observations and LOTOS-EUROS.

a traditional statistical approach. LOTOS-EUROS is capable of modelling the temporal behaviour of PM10 concentrations rather well. However, the absolute concentration is not captured and a systematic bias exists between the model and the measured concentrations. This bias is a common feature of CTMs (Stern et al., 2008). On average, the underestimation by LOTOS-EUROS is 13.2 μg m^{-3} at regional sites in the Netherlands for the period 2004–2006. Our comparison with observed data indicated that the LOTOS-EUROS model does not show a large systematic bias for the contribution of the species that can be validated (SO_4^{2-}, NO_3^-, NH_4^+, EC and sea salt) over the full range of PM10 concentrations. This means that the missing mass is largely associated with the non-modelled components and the components for which no measurement data are available. An important part of the missing mass is associated with crustal matter and organic carbon. Crustal matter (CM) and secondary organic aerosols (SOA) were not yet incorporated in the model version due to a lack of solid knowledge on emission strengths for CM and formation routes for SOA. The emissions of organic material (POM), though in principle included in the emission data for PPM2.5, may be significantly underestimated (Schaap and Denier van der Gon, 2007) and its behaviour after emission is uncertain (Robinson et al., 2007). In addition, besides the common issues concerning monitoring PM10 it has been posed that PM10 measurements include a portion of water, which might contribute to the systematic underestimation of the observed levels (Tsyro, 2005). In general, uncertainties in (the timing of) emissions, chemistry, meteorology and approximations with respect to the boundary layer behaviour impact the performance of the model (e.g. McKeen et al., 2007; Stern et al., 2008).

The systematic underestimation of the observed PM10 concentrations forces to apply a bias correction to the predictions of LOTOS-EUROS. Though the bias correction is an empirical approach, the LOTOS-EUROS predictions yield promising results compared to the operational model PROPART. The PROPART results follow the observations quite well, albeit that the forecasts peak 1 day later than the observations, which is due to the use of the today's average for the forecast of tomorrow. Hence, the statistical model is not able to provide an accurate forecast when the PM10 levels rise or drop sharply. These phenomena are associated with a shift in the large scale synoptic meteorological situation. LOTOS-EUROS does represent the sharp concentration changes and the peaks are often predicted at the correct day. The reason for the better timing is that the model explicitly includes the transport, formation and emissions of PM inside and upwind from the Netherlands. In this way fast changes in PM levels can be better modelled than using a statistical approach based on daily averages. The major flaw of the bias corrected LOTOS-EUROS model is that the highest concentrations are still underestimated. We have shown that the performance for predicting exceedance days can be improved by an alternative bias correction at the cost of a slight decrease in the general performance of the bias corrected model.

A CTM like LOTOS-EUROS is an approximation of reality and therefore describes average conditions (through emissions, parameterizations, etc.). Observations, on the other hand, are often influenced by contributions derived from events. They may derive from special activities, such as Easter fires or fire works. Also, long range transport of forest fires plumes or dust storms may occasionally reach the Netherlands (Hodzic et al., 2006; Birmili et al., 2008; Bruckmann et al., 2008). At present these sources are not represented in the emission module and can therefore not be represented by LOTOS-EUROS. In addition, these events cannot be captured by PROPART and persistence either. But these models have

Table 2
Comparison of the statistical performance of LOTOS-EUROS (LE), PROPART, the bias corrected LOTOS-EUROS (LE_biascor) and LOTOS-EUROS with an alternative bias correction (LE_alternative). The data represent average statistical properties over daily PM10 values covering the full years 2004–2006 and stemming from 16 rural stations, resulting in a total of 17,536 data points. Statistical parameters were determined for the individual stations and then averaged.

Parameter	Persistence	PROPART	LE	LE_biascor	LE_alternative
Mean	26.51	27.22	13.25	26.45	25.21
Bias	0.02	0.82	−13.23	−0.03	−1.28
Standard deviation of error	10.99	10.50	9.60	9.06	9.47
Skill variance	1.00	0.99	0.50	0.70	0.89
Correlation	0.63	0.66	0.68	0.70	0.70
Residue	7.64	7.53	13.40	6.49	6.94
RMSE	10.99	10.54	16.34	9.06	9.56
Hit rate	46.32	45.43	8.30	49.21	44.47
# predicted 50–200	991	1047	1	380	725
# observed 50–200	983	949	1003	1003	1003
% correct 50–200	42	40	100	61	50
% datacoverage	94	94	96	96	96

the most primitive form of data-assimilation since they always start from the right initial concentrations. In case of a prolonged event, they are able to pick up such signals and follow episodes, albeit with 1 day delay. CTMs provide a means to incorporate these events using routines for wind blown dust (e.g. Bessagnet et al., 2004) or satellite derived fire counts in combination with an emission module (e.g. Saarikosi et al., 2007). The magnitude of the emissions and the specific conditions during such events are difficult to capture satisfactorily. Data-assimilation of both in-situ and remote sensed observations upwind from the Netherlands may therefore improve the air quality analysis. LOTOS-EUROS is equipped with a data-assimilation system (e.g. van Loon et al., 2000; Barbu et al., 2008; Denby et al., 2008). The challenges regarding data-assimilation are many but its use for air quality forecasting should be investigated.

Based on the results of this study we conclude that the LOTOS-EUROS model has many advantages over PROPART for air quality forecasting for PM10. LOTOS-EUROS already outperforms the statistical models in most respects, in particular with respect to timing of events. Furthermore, it provides maps and time evolution which enhances the possibility of communicating the results to the general public, and more detailed knowledge on PM composition and origin. LOTOS-EUROS can be further improved on by including better routines for dust and data-assimilation.

Table A1
Performance for individual stations of the bias corrected LOTOS-EUROS over 2004–2006.

Station	Observed mean	LE-bc mean	Bias	Skill variance	RMSE	Correlation
LML131	25.76	26.86	1.09	0.72	8.22	0.79
LML133	25.30	26.28	0.97	0.92	7.51	0.74
LML230	30.63	27.33	−3.30	0.65	10.00	0.76
LML235	29.40	29.08	−0.32	0.77	9.32	0.72
LML318	26.02	28.16	2.14	0.74	9.29	0.76
LML437	26.35	27.30	0.94	0.72	8.62	0.75
LML444	27.09	27.16	0.07	0.78	8.58	0.65
LML538	23.34	24.72	1.38	0.74	7.57	0.69
LML631	24.91	24.98	0.07	0.62	9.03	0.72
LML633	23.31	28.20	4.90	0.75	9.71	0.72
LML722	26.80	26.76	−0.04	0.70	8.04	0.78
LML738	27.24	25.20	−2.04	0.64	9.41	0.72
LML818	27.61	25.6	−2.02	0.62	9.8	0.67
LML918	24.22	25.32	1.10	0.66	8.77	0.65
LML929	25.76	23.83	−1.93	0.56	9.43	0.73
LML934	26.58	24.69	−1.89	0.59	9.13	0.67

Appendix. Performance for individual stations

The performance of the models was evaluated for the individual stations. The results for the bias corrected LOTOS-EUROS are shown in Table A1. Results for the other models are not shown for conciseness. LOTOS-EUROS captures the gradient over the Netherlands with higher concentrations in the south and east of the country and lower concentrations in the north and west. The spread of the results is underestimated, especially in the north. Bias, RMSE and correlation do not show a regional trend. Results for the traffic oriented station in Rotterdam (433) and two urban background sites (441, Dordrecht, and 520, Amsterdam) do not deviate significantly from other stations according to these statistics, although in the observations the concentrations are somewhat higher than their neighbouring rural stations. However, these differences are smaller than the differences between the rural stations.

References

Barbu, A., Segers, A., Schaap, M., Heemink, A., Builtjes, P.J.H., 2008. A multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe. Atmospheric Environment 43, 1622–1631.

Beijk, R., Mooibroek, D., Hoogerbrugge, R., 2007. Air Quality in the Netherlands, 2003–2006 (Jaaroverzicht Luchtkwaliteit, in Dutch). RIVM report 680704002. Bilthoven, the Netherlands.

Beijk, R., Mooibroek, D., Kassteele, Jvan de, Hoogerbrugge, R., 2008. PM10: Equivalence study 2006: Demonstration of Equivalence for the Automatic PM10 Measurements in the Dutch National Air Quality Monitoring Network: a Technical Background Report, RIVM Report 680708002 Bilthoven, the Netherlands.

Bessagnet, B., Hodzic, A., Vautard, R., Beekmann, M., Cheinet, S., Honoré, C., Liousse, C., Rouil, L., 2004. Aerosol modeling with CHIMERE: preliminary evaluation at the continental scale. Atmospheric Environment 38, 2803–2817.

Birmili, W., Schepanski, K., Ansmann, A., Spindler, G., Tegen, I., Wehner, B., Nowak, A., Reimer, E., Mattis, I., Müller, K., Brüggemann, E., Gnauk, T., Herrmann, H., Wiedensohler, A., Althausen, D., Schladitz, A., Tuch, T., Löschau, G., 2008. A case of extreme particulate matter concentrations over Central Europe caused by dust emitted over the southern Ukraine. Atmospheric Chemistry and Physics 8 (4), 997–1016.

Bruckmann, P., Niemann, K., Wurzler, S., Engel, I., Hoogerbrugge, R., Garus, G., Olschewski, A., 2008. High PM_{10} levels covering large parts of Europe caused by long range transport. Gefahrstoffe Reinhaltung der Luft 68 (5), 189–195.

Brunekreef, B., Holgate, S.T., 2002. Air pollution and health. The Lancet 360, 1233–1242.

Cuvelier, C., Thunis, P., Vautard, R., Amann, M., Bessagnet, B., Bedogni, M., Berkowicz, R., Brandt, J., Brocheton, F., Builtjes, P., Coppalle, A., Denby, B., Douros, G., Graf, A., Hellmuth, O., Honoré, C., Hodzic, A., Jonson, J., Kerschbaumer, A., de Leeuw, F., Minguzzi, E., Moussiopoulos, N., Pertot, C., Pirovano, G., Rouil, L., Schaap, M., Stern, R., Tarrason, L., Vignati, E., Volta, M., White, L., Wind, P., Zuber, A., 2007. CityDelta: a model intercomparison study to explore the impact of emission reductions in European cities in 2010. Atmospheric Environment 41, 189–207.

Denby, B., Schaap, M., Segers, A., Builtjes, P., Horalek, J., 2008. Comparison of two data assimilation methods for assessing PM10 exceedances on the European scale. Atmospheric Environment 42, 7122–7134.

Denier van der Gon, H.A.C., Hulskotte, J.H.J., Visschedijk, A.J.H., Schaap, M., 2007. A revised estimate of copper emissions from road transport in UNECE Europe and its impact on predicted copper concentrations. Atmospheric Environment 41, 8697–8710.

EEA, 2007. Europe's Environment. The Fourth Assessment. EEA, Copenhagen.

Erisman, J.W., van Pul, A., Wyers, P., 1994. Parameterization of surface-resistance for the quantification of atmospheric deposition of acidifying pollutants and ozone. Atmos. Environ. Vol. 28, 2595–2607.

Hass, H., Builtjes, P.J.H., Simpson, D., Stern, R., 1997. Comparison of model results obtained with several European regional air quality models. Atmospheric Environment 31, 3259–3279.

Hass, H., van Loon, M., Kessler, C., Stern, R., Matthijsen, J., Sauter, F., Zlatev, Z., Langner, J., Foltescu, V., Schaap, M., 2003. Aerosol Modelling: Results and Intercomparison from 15 European Regional-scale Modelling Systems. EUROTRAC-2 Special report. Eurotrac-ISS, Garmisch Partenkirchen, Germany.

Hodzic, A., Vautard, R., Chepfer, H., Goloub, P., Menut, L., Chazette, P., Deuzeć, J.L., Apituley, A., Couvert, P., 2006. Evolution of aerosol optical thickness over Europe during the August 2003 heat wave as seen from CHIMERE model simulations and POLDER data. Atmospheric Chemistry and Physics 6, 1853–1864.

Hollingsworth, A., Engelen, R.J., Textor, C., Benedetti, A., Boucher, O., Chevallier, F., Dethof, A., Elbern, H., Eskes, H., Flemming, J., Granier, C., Kaiser, J.W., Morcrette, J.-J., Rayner, P., Peuch, V.-H., Rouil, L., Schultz, M.G., Simmonsa, A.J., 2008. Toward a monitoring and forecasting system for atmospheric

composition GEMS project. Journal of Atmospheric and Oceanic Technology 89 (8), 1147–1164. doi:10.1175/2008BAMS2355.1.
Hooyberghs, J., Mensink, C., Dumont, G., Fierens, F., Brasseur, O., 2005. A neural network forecast for daily average PM10 concentrations in Belgium. Atmospheric Environment 39, 3279–3289.
Karydis, V.A., Tsimpidi, A.P., Pandis, S.N., 2007. Evaluation of a three-dimensional chemical transport model (PMCAMx) in the eastern United States for all four seasons. Journal of Geophysical Research D: Atmospheres 112 (14), D14211. doi:10.1029/2006JD007890.
Kukkonen, J., Partanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, M., Niska, H., Dorling, S., Chatterton, T., Foxall, R., Cawley, G., 2003. Extensive evaluation of neural network models for the prediction of NO2 and PM10 concentrations, compared with a deterministic modelling system and measurements in central Helsinki. Atmospheric Environment 37, 4539–4550.
van Loon, M., Builtjes, P.J.H., Segers, A., 2000. Data assimilation of ozone in the atmospheric transport model LOTOS. Modeling Software 15, 603–609.
van Loon, M., Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P.J.H., Christensen, J.H., Cuvelier, K., Graf, A., Jonson, J.E., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R., Tarrasón, L., Thunis, P., Vignati, E., White, L., Wind, P., 2007. Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble average. Atmospheric Environment 41, 2083–2097.
Manders, A., Nguyen, L., Hoogerbrugge, R., 2008. Evaluatie van RIVM-modellen voor de ozon-en fijnstofverwachting RIVM report no. 680704004, Bilthoven, The Netherlands.
McKeen, S.A., Chung, S.H., Wilczak, J., Grell, G., Djalalova, I., Peckham, S., Gong, W., Bouchet, V., Moffet, R., Tang, Y., Carmichael, G.R., Mathur, R., Yu, S., 2007. Evaluation of several PM2.5 forecast models using data collected during the ICARTT/NEAQS 2004 field study. Journal of Geophysics Research D: Atmospheres 112, D10S20. doi:10.1029/2006JD007608.
Monahan, E.C., Spiel, D.E., Davidson, K.L., 1986. A model of marine aerosol generation via whitecaps and wave disruption. In: Monahan, E.C., MacNiocaill, G. (Eds.), Oceanic Whitecaps and Their Role in Air-sea Exchange.
Noordijk, H., 2003. PROPART en PROPART; statistische modellen voor smogprognose RIVM report no. 725301012, Bilthoven, The Netherlands.
Putaud, J.-P., Raes, F., Van Dingenen, R., Brüggemann, E., Facchini, M.-C., Decesari, S., Fuzzi, S., Gehrig, R., Hüglin, C., Laj, P., Lorbeer, G., Maenhaut, W., Mihalopoulos, N., Müller, K., Querol, X., Rodriguez, S., Schneider, J., Spindler, G., Ten Brink, H., Tørseth, K., Wiedensohler, A., 2004. A European aerosol phenomenology – 2: chemical characteristics of particulate matter at kerbside, urban, rural and background sites in Europe. Atmospheric Environment 38, 2579–2595.
Querol, X., Alastuey, A., Ruiz, C.R., Artiñano, B., Hansson, H.C., Harrison, R.M., Buringh, E., ten Brink, H.M., Lutz, M., Bruckmann, P., Straehl, P., Schneider, J., 2004. Speciation and origin of PM10 and PM2.5 in selected European cities. Atmospheric Environment 38, 6547–6555.
Robinson, A.L., Donahue, N.M., Shrivastava, M.K., Weitkamp, E.A., Sage, A.M., Grieshop, A.P., Lane, T.E., Pierce, J.R., Pandis, S.N., 2007. Rethinking organic aerosols: semivolatile emissions and photochemical aging. Science 315, 1259–1262. doi:10.1126/science.1133061.
Saarikosi, S., Sillanpää, M., Sofiev, M., Timonen, H., Saarnio, K., Teinilä, K., Karppinen, A., Kukkonen, J., Hillamo, R., 2007. Chemical composition of aerosols during a major biomass burning episode over northern Europe in spring 2006: experimental and modelling assessments. Atmospheric Environment 41, 3577–3589.
Schaap, M., Denier van der Gon, H.A.C., 2007. On the variability of black smoke and carbonaceous aerosols in the Netherlands. Atmos. Env. 41, 5908–5920.
Schaap, M., van Loon, M., ten Brink, H.M., Dentener, F.D., Builtjes, P.J.H., 2004a. Secondary inorganic aerosol simulations for Europe with special attention to nitrate. Atmospheric Chemistry and Physics 4, 857–874.
Schaap, M., Denier Van Der Gon, H.A.C., Dentener, F.J., Visschedijk, A.J.H., van Loon, M., Ten Brink, H.M., Putaud, J.-P., Guillaume, B., Liousse, C., Builtjes, P.J.H., 2004b. Anthropogenic black carbon and fine aerosol distribution over Europe. Journal of Geophysical Research 109. doi:10.1029/2003JD004330 D18201.
Schaap, M., Timmermans, R.M.A., Sauter, F.J., Roemer, M., Velders, G.J.M., Boersen, G.A.C., Beck, J.P., Builtjes, P.J.H., 2008. The LOTOS-EUROS model: description, validation and latest developments. International Journal of Environment and Pollution 32 (2), 270–289.
Stadlober, E., Hörmann, S., Pfeiler, B., 2008. Quality and performance of a PM10 daily forecasting model. Atmospheric Environment 42, 1098–1109.
Stern, R., Builtjes, P., Schaap, M., Timmermans, R., Vautard, R., Hodzic, A., Memmesheimer, M., Feldmann, H., Renner, E., Wolke, R., Kerschbaumer, A., 2008. A model inter-comparison study focussing on episodes with elevated PM10 concentrations. Atmospheric Environment 42, 4567–4588.
Ten Brink, H., Maenhaut, W., Hitzenberger, R., Gnauk, T., Spindler, G., Even, A., Chi, X., Bauer, H., Puxbaum, H., Putaud, J.-P., Tursic, J., Berner, A., 2004. INTERCOMP2000: the comparability of methods in use in Europe for measuring the carbon content of aerosol. Atmospheric Environment 38 (38), 6507–6519.
Tsyro, S.G., 2005. To what extent can aerosol water explain the discrepancy between model calculated and gravimetric PM10 and PM2.5? Atmospheric Chemistry and Physics 5 (2), 515–532.
Visschedijk, A.J.H., Denier van der Gon, H.A.C., 2005. Gridded European anthropogenic emission data for NOx, SOx, NMVOC, NH_3, CO, PPM10, PPM2.5 and CH_4 for the year 2000, TNO-Report B&O-A R 2005/106.
Yu, S., Mathur, R., Schere, K., Kang, D., Pleim, J., Young, J., Tong, D., Pouliot, G., McKeen, S.A., Rao, S.T., 2008. Evaluation of real-time PM2.5 forecasts and process analysis for PM2.5 formation over the eastern United States using the Eta-CMAQ forecast model during the 2004 ICARTT study. Journal of Geophysical Research D: Atmospheres 113 (6). doi:10.1029/2007JD009226 D06204.
van Zelm, R., Huijbregts, M.A.J., den Hollander, H.A., van Jaarsveld, H.A., Sauter, F.J., Struijs, J., van Wijnen, H.J., van de Meent, D., 2008. European characterization factors for human health damage of PM10 and ozone in life cycle impact assessment. Atmospheric Environment 42 (3), 441–453. doi:10.1016/j.atmosenv.2007.09.072.