Atmospheric composition forecasting in Europe

L. Menut\textsuperscript{1} and B. Bessagnet\textsuperscript{2}
\textsuperscript{1}Institut Pierre-Simon Laplace, Laboratoire de Météorologie Dynamique, CNRS UMR 8539, Ecole Polytechnique, Palaiseau, France
\textsuperscript{2}INERIS, Institut National de l’Environnement Industriel et des Risques, Parc technologique ALATA, 60550 Verneuil en Halatte, France

Received: 25 August 2009 – Revised: 9 December 2009 – Accepted: 24 December 2009 – Published: 14 January 2010

Abstract. The atmospheric composition is a societal issue and, following new European directives, its forecast is now recommended to quantify the air quality. It concerns both gaseous and particles species, identified as potential problems for health. In Europe, numerical systems providing daily air quality forecasts are numerous and, mostly, operated by universities. Following recent European research projects (GEMS, PROMOTE), an organization of the air quality forecast is currently under development. But for the moment, many platforms exist, each of them with strengths and weaknesses. This overview paper presents all existing systems in Europe and try to identify the main remaining gaps in the air quality forecast knowledge. As modeling systems are now able to reasonably forecast gaseous species, and in a lesser extent aerosols, the future directions would concern the use of these systems with ensemble approaches and satellite data assimilation. If numerous improvements were recently done on emissions and chemistry knowledge, improvements are still needed especially concerning meteorology, which remains a weak point of forecast systems. Future directions will also concern the use of these forecast tools to better understand and quantify the air pollution impact on health.

Keywords. Atmospheric composition and structure (Pollution – urban and regional)

1 Introduction

The reduction of the air pollution impact on health may be achieved both with short and long term actions. The long term action is to globally improve the air quality by reducing anthropogenic emissions. This first option is a major goal of the 6th Environmental Action program and the Thematic Strategy on Air Pollution, adopted in September 2005. The technical project “Clean Air For Europe” (Amann et al., 2005) provided tools to assess the efficiency of the current legislation and a basis for its revision (Cuvelier et al., 2007). The short term action is to anticipate pollution events, a few days before, in order to warn the public in advance, particularly its most sensitive fraction (the elderly, young children and asthmatics), as well as to help the authorities to take efficient emergency control actions. Air quality forecasting is the tool that can help to deal with such objectives.

Two types of forecast systems exist, both based on the use of meteorological data and chemistry models. They are built on “statistical” and “deterministic” methods. Mainly for computer efficiency reasons, efforts were first done on the development and use of statistical models during the past decades. Easy to use and implement, simplified physics and chemistry equations systems were developed and tuned in order to estimate a probability of pollution event. This is achieved by using only meteorological parameters such as mean wind speed, solar radiation and temperature (Hrust et al., 2009). Using only few equations, these codes allowed a fast computation (a major constraint for forecast systems). They were widely used during many years and showed satisfactory results, at least for ozone daily maxima (Zeldin and Thomas, 1975; Simpson and Layton, 1983; Robeson and Steyn, 1990). On the other hand, very crude and uncertain (no boundary conditions for example), these systems were gradually less used even if many systems already exist and use sophisticated numerical approaches such as neural networks for air quality short term forecast (Kolehmainen et al., 2001; Chaloulakou et al., 2003; Ibarra-Berastegi et al., 2008; Pfeiffer et al., 2009) and health impacts (Schlink et al., 2006).

This paper deals with the “deterministic” atmospheric composition forecast in Europe. In the last decade, progresses in computing technologies allowed the use of simplified deterministic three-dimensional models, the so-called...
chemistry-transport models (CTM) such as Vautard et al. (2001); Tilmes (2002); Cope (2004); McHenry et al. (2004); Vaughan (2004); McKeen (2005). With respect to the statistical approaches, these models provide a more comprehensive picture of air quality for a given time period and geographical area, following the evolution of urban pollution plume. Moreover, they offer the possibility to assess the efficiency of a given emission scenario. In 2000, the state of the art of CTMs and described by Dabberdt and Miller (2000) was limited to a few set of models, each one having huge uncertainties both in their forcings and chemistry parameterizations. More recently, Ebel et al. (2005) presented a panorama of modeling systems in Europe: the situation was rather different with a significant evolution. The key points highlighted by the authors were the expectancy of better computer performances (for a better resolution as well as more complex chemical mechanisms), the chemical data assimilation and expected improvements in aerosol modeling. Today, the air quality modeling systems in Europe are much more numerous and complex than in 2005. They are able to model gaseous and aerosols species at various spatial scales. Most of them are used both for analyzes, scenarios and forecast.

The atmospheric composition forecast is a recent activity in which research institutes are continuously involved, working on the model developments. These institutes are not dedicated to ensure daily results but much more to build the future forecast systems: the systems are thus called “experimental”. More recently, the first “operational systems” appeared but they remain scarce. In the air quality research community, the projects were first dedicated to model improvements mainly for the most “relevant” periods such as strong pollution episodes. Once the models were able to diagnose the peaks, mainly for ozone, the research turns to a better understanding of the whole “air pollution system”. Field campaigns and models were employed for longer periods and the goals turned to have accurate model results for all periods: with and without pollution episodes. With this point of view, the use of a model in a forecast mode is an objective tool to really evaluate its skills and weaknesses. In doing so, models run every day without a priori knowledge of the concentrations, certainly the best way to avoid model tuning and to improve the scientific knowledge. Knowing the fast improvements in numerics and network connections, a lot of model developers are testing their model by developing web forecast platforms. These platforms may be simple, but are all able to prognose pollutants concentrations up to three days in advance. This temporal limitation is primarily dependent on the weather forecast: a limit to predict meteorological fields enough accurate so that the cumulated uncertainty with emission calculations is acceptable.

The challenge is now to organize all these models in common frameworks and this part is being shifted from scientific to governmental activities communities. In parallel, and even if important scientific advancements were done the last years, the modeling systems are still evolving and must be considered as under development tools. Indeed, most of model scores (almost for particulate matter) are not sufficiently accurate to match the European directives on air quality forecasts accuracy. The needs are now a continuous models improvement and a better use and spreading of their outputs.

In this paper, the main principles of a deterministic atmospheric composition forecast system are described in the Sect. 2, the existing “experimental” and “operational” systems are described in Sect. 3. Future directions for models developments and improvements are discussed, first, about “processes” (Sect. 4) and, second, about “numerics” (Sect. 5).

2 The atmospheric composition modeling systems

Based on the same physical and chemical principles, all systems currently used in Europe and presented in Table 1 are different in their use and their results. Before calculating chemical concentrations, it is necessary to calculate the forcings of the system, i.e. the emissions and meteorology. For the systems described in this article, the principle remains the one-way nesting. Some models operate in two-ways nesting but are not currently used due to computational cost. The first set of differences is the meteorological drivers, including the parameterizations and the spatial and temporal resolutions. The second set of differences concerns the input data as the anthropogenic and biogenic emissions. In addition, some models, but not all, are able to prognose specific particle concentrations such as desert dust (i.e. mineral particles) or forest fires products.

2.1 The meteorology

Modeled concentrations are very sensitive to the physical parameters (wind, temperature, specific humidity) and diagnosed turbulent parameters (friction velocity, boundary layer height). A direct error on meteorological fields has a more or less direct and linear impact on concentrations (Menut, 2003; Minguzzi et al., 2005). In addition, the selected horizontal resolution of a data set has a large impact on results (Valari and Menut, 2008). During the 1990s, the global scale meteorological forecast systems outputs, such as NCEP (Kalnay et al., 1996) and ECMWF (Bechtold et al., 2008), were used as it, mainly using interpolations to regird the meso-scale data fields. In recent years, the forecast systems largely evolved: they now use a mesoscale model, forced by global meteorological fields, but more adapted to the fine resolutions and with more relevant landuses and turbulent parameterizations. In Europe, the MM5 (Grell et al., 1994) and WRF American models (Skamarock et al., 2007) are the most used, recognized for their ease of implementation and change. For forecast, new systems appear and are based on the Integrated Forecast System (IFS) operated in ECMWF.
Table 1. List of current operating forecast systems in Europe. Systems are sorted in alphabetical order. The operational systems are underlined. The MACC project is described at http://www.gmes-atmosphere.eu/.

<table>
<thead>
<tr>
<th>System name (URL)</th>
<th>Country</th>
<th>Meteo/CTM</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CALIOPE (<a href="http://www.bsc.es/caliop">www.bsc.es/caliop</a>)</td>
<td>Spain</td>
<td>WRF/CMAQ</td>
<td>Baldasano et al. (2008)</td>
</tr>
<tr>
<td>CAMX-AMWFG (forecast.ua.org)</td>
<td>Greece</td>
<td>SKIRON/CAMx</td>
<td>Kallos et al. (2007)</td>
</tr>
<tr>
<td>CETEMPS (pumpkin.aquila.infn.it/forechem/)</td>
<td>Italy</td>
<td>MM5/CHIMERE</td>
<td>Cucci et al. (2008)</td>
</tr>
<tr>
<td>CHIMERE (MACC)</td>
<td>France</td>
<td>IFS/CHIMERE</td>
<td>Bessagnet et al. (2008)</td>
</tr>
<tr>
<td>EMEC-CWE (MACC)</td>
<td>Norway</td>
<td>IFS/EEMEP</td>
<td>Simpson et al. (2003)</td>
</tr>
<tr>
<td>EURAD (MACC) (<a href="http://www.eurad.uni-koeln.de">www.eurad.uni-koeln.de</a>)</td>
<td>Germany</td>
<td>MM5/EURAD-CTM</td>
<td>Elbern and Schmidt (2001)</td>
</tr>
<tr>
<td>FARM (<a href="http://www.aria-net.eu/QualeAria/">www.aria-net.eu/QualeAria/</a>)</td>
<td>Italy</td>
<td>RAMS/FARM</td>
<td>Zanini et al. (2005)</td>
</tr>
<tr>
<td>HIRLAM/MATCH (<a href="http://www.airviro.smhi.se/MATCH-AQ">www.airviro.smhi.se/MATCH-AQ</a>)</td>
<td>Sweden</td>
<td>HIRLAM/MATCH</td>
<td>Robertson et al. (1999)</td>
</tr>
<tr>
<td>LRTOS-EUROS (MACC)</td>
<td>Netherlands</td>
<td>LRTOS-EUROS</td>
<td>Schaap et al. (2008)</td>
</tr>
<tr>
<td>MATCH (MACC)</td>
<td>Sweden</td>
<td>IFS/MATCH</td>
<td>Robertson et al. (1999)</td>
</tr>
<tr>
<td>MOCAGE (MACC)</td>
<td>France</td>
<td>IFS/MOCAGE</td>
<td>Michou and Peuch (2002)</td>
</tr>
<tr>
<td>RCG (<a href="http://www.trumf.de">www.trumf.de</a>)</td>
<td>Germany</td>
<td>REM/CALGRID</td>
<td>Stern et al. (2003)</td>
</tr>
<tr>
<td>SILAM (MACC) (silam.fmi.fi)</td>
<td>Finland</td>
<td>HIRLAM/SILAM</td>
<td>Sofiev et al. (2006)</td>
</tr>
<tr>
<td>SKIRON/Dust (forecast.ua.org)</td>
<td>Greece</td>
<td>SKIRON</td>
<td>Kallos et al. (2007)</td>
</tr>
<tr>
<td>SMOGPROG (<a href="http://www.lml.rivm.nl/data/smogprog">www.lml.rivm.nl/data/smogprog</a>)</td>
<td>Netherlands</td>
<td>LRTOS-EUROS, CHIMERE</td>
<td>Schaap et al. (2008)</td>
</tr>
<tr>
<td>THOR (thor.dmu.dk)</td>
<td>Denmark</td>
<td>ETA/DEHM</td>
<td>Frohn and Brandt (2006)</td>
</tr>
<tr>
<td>UK AQ forecast (<a href="http://www.airquality.co.uk">www.airquality.co.uk</a>)</td>
<td>United Kingdom</td>
<td>ECMWF/NAME</td>
<td>Ryall and Maryon (1998)</td>
</tr>
<tr>
<td>WRF-CHIMERE (<a href="http://www.lmd.polytechnique.fr/cosy">www.lmd.polytechnique.fr/cosy</a>)</td>
<td>France</td>
<td>WRF/CHIMERE</td>
<td>Vautard et al. (2001)</td>
</tr>
<tr>
<td>ZAM (<a href="http://www.zamg.ac.at">www.zamg.ac.at</a>)</td>
<td>Austria</td>
<td>ALADIN/CAMx</td>
<td>Hirtl et al. (2007)</td>
</tr>
<tr>
<td>TAU (wind.tau.ac.il/dust8/dust.html)</td>
<td>Israel</td>
<td>MM5/DREAM</td>
<td>Kishcha et al. (2007)</td>
</tr>
</tbody>
</table>

2.2 The emissions

The other main forcing remains the surface emissions fluxes. Depending on species and their origin (anthropogenic or biogenic), they constitute the only pollutant sources and the most important unknown. Emissions are difficult to evaluate and are very sparse in space and variable in time. Emissions are often delivered to modelers after numerous “bottom-up” studies on cities (residential fluxes), industries, local measurements near sources (traffic). Results are classified per family species (NOx, VOCs, SO2, Particulate matter, etc.) and activity sectors (EMEP, Yttri et al., 2008; Vestreng, 2003, and EDGAR, Olivier et al., 1999, inventories). Data are often provided with poor time resolution, mainly monthly or annual masses. The modelers have to build their emission datasets for their own forecast modeling system. This increases the spread in emissions calculations and, consequently, in models performances. This effect was recently quantified by the way of models inter-comparisons exercises such as City-Delta (Cuvelier et al., 2007; Vautard et al., 2007) and GEMS (Simmons, 2004).

2.3 The chemistry-transport models

The two previous forcings, meteorology and emissions, are used to calculate the spatio-temporal concentrations of chemical species. Based on equations describing the transport, mixing, diffusion, deposition and chemistry of these species, these models have spatial resolutions of a few kilometers (∼5 km for the “city” scale domains) to a few tens of kilometers (∼50 km for the European domains). In general, the temporal integration step is always a few tens of seconds in order to respect the chemical mechanism reactions rates. The models most used in Europe are mainly CMAQ, CAMx and CHIMERE, as part of different forecasting systems including systems over cities with a higher horizontal resolution. Some of the models presented in the Table 1 are not designed for chemistry but only for transport of dust particles: CHIMERE-DUST (the mineral dust modeling part of CHIMERE), TAU and SKIRON/Dust.

Several choices are done by the model developers for their own platform. Some systems are using one model: for example, the PREVISAO-QAR in Portugal uses CHIMERE forced by the meteorological model MM5 (Dudhia, 1993). The two models are used with the same set of parameterizations but in nesting mode with several resolutions (at the
European scale and Portugal scale). The same methodology is used for the CALIOPE project, using the WRF-ARW (Skamarock et al., 2007) and CMAQ (Meng et al., 2007; Hakami et al., 2007) models. A second way is to use several models, each one modeling a specific area. For example, the SMOGPROG system uses the LOTOS-EUROS model for the European scale and drives the CHIMERE model for a zoom on the Netherlands.

3 Experimental and operational forecasting systems

Among the 23 systems identified (in the Table 1), 7 are operational: Both are derived from national initiatives (PREV’AIR and CALIOPE) and others are part of the MACC project (CHIMERE and MOCAGE, the same models than in PREV’AIR, EURAD, EMME-CWF, MATCH). The others are operated in research laboratories and are not obliged to deliver forecast results every days of the year. A flowchart describing the principle of air quality systems is displayed in Fig. 1.

Currently, the operational forecast systems in Europe are the PREV’AIR and the CALIOPE systems. The PREV’AIR system is a French national initiative and certainly the first step of an European system, under development in the framework of the European GEMS project. PREV’AIR (Rouil et al., 2009) is based on a partnership between french research institutes (CNRS and INERIS), the french government environmental agency (ADEME), meteorological center (Meteo-France), research ministry and the french ministry in charge of Ecology. In this system, engineers and researchers produce and provide forecast results every day of the year as in operational meteorological center. Hourly forecasts are performed for ozone, NOx, PM10 and Dust over France, Europe and at global scale with the CHIMERE and MOCAGE models. Scores for NO2, PM10 and O3 are calculated for daily peaks and mean values and statistics are displayed in terms of biases, RMSE (Root Mean Square Error), and correlations. For ozone daily peaks, correlations (in time and space) of 0.81 to 0.76 are reached respectively from the D+0 to D+2 lead time. NOx daily mean values are modeled with a correlation of 0.6 in summer and 0.5–0.7 in winter. For PM10, correlations are lower ranging from 0.55 to 0.35, stressing the gap of knowledge compared to gaseous species. For aerosols, biases are negative due to a lack in coarse particle concentrations other unquantified sources and badly modeled processes such as the secondary organic aerosols (SOA) formation (Honoré et al., 2008). For CALIOPE, its administrative structure merges universities and governmental authorities (the Barcelona Supercomputing Center (BSC), the CIEMAT, the Earth Sciences Institute “Jaume Almera” (IJA-CSIC) and the CEAM Foundation), moving it currently to an “operational” status.

The next step for “operational systems” will certainly be the results of the GMES project, going on with the FP7 European MACC project. More generally, the GEMS project is developing comprehensive data analysis and modeling systems for monitoring the global distributions of atmospheric constituents important for climate, air quality and UV radiation in Europe. A specific theme is dedicated to the regional air quality and is achieving its objective of producing daily regional forecasts of chemical species and air quality indicators based on an ensemble of models covering a common European domain. Ten modeling teams participate in this exercise. The models use a new emission dataset with a 5km resolution developed specially for GEMS and a common meteorological forcing from ECMWF operations. They take boundary conditions from the forecasts produced by the global GEMS system. Common archiving, display, and verification methods facilitate inter-comparison and forecast validation. Memoranda of Understanding negotiated with environment agencies cover their near-real-time supply of national air-quality measurements to GEMS for validation and their receipt of forecast data from GEMS. The regional models are also being run cooperatively to examine key periods from the extreme year 2003. The theme includes assessment of their products for human health protection. The forecast results are updated every day and available at http://gems.ecmwf.int/d/products/raq/.

4 Future directions in processes modeling

The forecast of European air quality is now acceptable for situations most commonly observed. But the systems always need scientific improvements mainly to add realistic variability in the results. This variability will depend on specific points: (i) the sub-grid scale processes such as the regional clouds and precipitations diagnostics, currently considered homogeneous over single grid cells, (ii) some poorly known chemical and physical processes: the SOA formation and the vertical mixing of PM in the surface layer, for example.

Some relevant processes for decision makers are missing in models, the terrigenous emissions in Europe is important for the surface particulate matter (PM) budget and exceedance calculations. Fire emissions is another issue and in directly impacts on radiation budget and surface PM load. Finally, the forecast systems will have to evolve to their primary goal: the short term impact on health, by developing extended models able to estimate the population exposure and thereafter the potential health impacts as displayed in Fig. 2.

From all forcings and mechanisms used in this type of chemistry-transport models, the main identified gaps first concern the direct forcings: meteorology, chemical boundary conditions and emissions as presented in Fig. 2. Some gaps are also identified in chemistry and some specific less studied topics such as the dust and fire emissions in Europe and the small scale vertical distribution of particles in the nocturnal boundary layer among others.
4.1 Meteorological forcings

Considering meteorology is not so complex and quite well forecasted, one can expect good scores for most of gaseous species peaks. More the domain sizes is reduced, more the local meteorology becomes sensitive to a mix of local and synoptic variability. The uncertainty due to the small scale meteorological processes becomes prevalent and the error on meteorology directly impacts the error on atmospheric pollution calculations.

The cloud modeling and their effects remains a key issue (Tao and Moncrieff, 2009). If clouds occurrences and lifetime is mainly diagnosed in global models, their physics have to be finely prognose in meso-scale models. Clouds location, liquid water content, convective and stratiform precipitations may drastically modify the surface pollutants budget (Sea- man, 2000). For convective clouds, an increased forced mixing in the boundary layer will have an impact on emissions vertical mixing (Zhang et al., 2003, 2007). The presence of clouds also affects the ozone production which depends on their radiative properties (photolysis rates) and their vertical distribution. Aerosol formation is largely affected by clouds as sulfate, nitrate chemistry. The particle scavenging depends on cloud water content and aerosol properties.

Another set of identified gaps concerns the transport, mixing, diffusion and deposition. For example, the vertical transport and deposition of gaseous species and aerosols from the free troposphere to the boundary layer remains uncertain. In this case, satellites could be able to give informations on the vertical structure of high concentration layers during long-range transport and to better understand fine stratified layers (Vuolo et al., 2009).

4.2 Emissions

The problem of surface emissions is one of the most complex because many problems accumulate. The chemical species involved are very different (gas and particles, their lifetimes, their different chemistry). Then, the emissions themselves are very difficult to measure and depend on estimated data from different sources such as laboratory measurements (for vehicles exhaust) or statistical data (annual masses) from industrial activities. In addition, these surface emissions data evolve a lot in time. Currently, at the European scale, the most consistent and updated emissions of the main gas and particles are those currently provided by EMEP (www.emep.int) with a 50 km resolution. In most of chemistry transport models, smooth seasonal, weekly and hourly profiles are generally used to get hourly emissions from annual totals. These profiles account for averaged meteorology and human activity changes.

4.2.1 Time updates of anthropogenic emissions

Surface emissions are also well known at various scales only as averaged fluxes in space and time. When extreme events are observed, the forecast systems suffered of a “non adaptability” to specific activities: currently, no system is able to be updated day by day to account for specific emissions such as industrial accidents or other unusual activities. Even if emissions are well prescribed, the issue of year-to-year update has also to be considered. For the surface emissions, satellites may deliver informations for the spatial distribution of main sources and day to day temporal evolution during many years (Konovalov et al., 2006). Depending on retrieved species (such as NO$_2$ and CO), some missing sources
may also be identified under low wind and low photochemistry conditions (during the night).

### 4.2.2 The ammonia emissions issue

Ammonia is an important PM precursor through ammonium nitrate formation. In a “real world”, ammonia emissions depend on the type of cattle, manures and fertilizers, spreading practices, meteorological and soil parameters. The time profiles for ammonia are not accurate enough to catch the real instantaneous emission when a parameter (temperature, soil humidity and spreading practices) largely deviates from its average value. Moreover, if annual quantities are not available for the studied year, the closest documented year is chosen assuming that inter-annual variations of emissions are small. In order to better estimate large ammonia emissions during specific meteorological conditions, new emission models allow to better account for ammonia emissions (Diaz Goebez et al., 2003) and the dynamical approach used in Skjøth et al. (2005) shows improved results in retrieving ammonia concentrations. Recently, Beuning et al. (2008) have developed ammonia emission models able to predict ammonia emissions peaks in Canada. In Europe, an inter-comparison exercise shows that ammonia emission models provided similar emission factors (Reidy, 2008). Therefore, a dynamical approach for the treatment of ammonia emissions needs to be implemented in chemistry transport models to obtain better model predictions of such high particulate matter episodes (Zhang et al., 2008).

### 4.2.3 Ship emissions

Ships with huge engines running on bunker fuel without emission controls, thousands of diesel trucks per day, diesel locomotives, and other polluting equipment and activities at modern seaports cause an array of environmental impacts that can seriously affect local communities and marine and land-based ecosystems throughout a region. These impacts range from increased cancer risk in nearby communities and increased regional smog, to contamination of water bodies, the introduction of destructive foreign species and aesthetic effects on local communities and public lands (Bailey and Solomon, 2004).

A study in the port city of Vancouver, British Columbia (Lu et al., 2006) showed that several pollution episodes were caused by fresh ship plumes. In terms of exposure, a recent study (Andersson et al., 2009) showed that international shipping was estimated to contribute 5% to the total Primary PM\(_{2.5}\) population weighted average concentration (PWC) and 9% to the secondary inorganic aerosol PWC. Nevertheless, Hellebust et al. (2009) found that fresh ship plumes were not found to make a significant contribution to primary PM\(_{2.5-0.1}\) concentrations adjacent to a shipping channel. However, this was partially attributed to the ultrafine nature of ship emissions and the majority of the toxic metal content was attributed to emissions associated with heavy oil combustion sources, which include ship engines. The issue of shipping emissions is still open and there is a
urgent need of better inventories particularly for PM species, Moldanov et al. (2009) showed that there is not a consensus on gas and PM emission factors. Moreover, the use of coarse resolution inventories need shipping routes database to better refine these inventories for regional modeling applications.

4.2.4 Mineral dust emissions

Dust are commonly identified as natural but in some cases human is responsible for soil erodability as in the former Republics of the Soviet Union, then the origin can be partly anthropogenic in those cases. Missing natural particles sources were recently identified in Europe: ignored or badly modeled, further developments are necessary. During some dust outbreaks inside Europe, the transport of these huge masses of aerosols may induce pollution alerts falsely attributed to anthropogenic pollution then it is natural. These events remain poorly represented in models and a recent study was devoted to the Chernozemic soils explaining particles peaks in remote and urbanized areas (Bessagnet et al., 2008). But this latter event remains exceptional: in average, the dust emissions inside Europe are moderate and their transport is local.

Terrigenous matter may come from other continents and impact European particulate matter budget. The Saharan dust outbreaks are sporadic and intense. The forecast of their emissions and transport aggregate a large number of uncertainties. Their predictability is very uncertain as estimated in Menut et al. (2009). To improve the system, the first step will be to merge already existing large-scale modeling system (such as hemispheric dust models) and regional CTMs. A second challenge would be to avoid risks of false air quality alerts due to large overestimated intense dust production events. For a large part of current CTMs used for forecast, boundary conditions are global monthly climatologies from model outputs. This may be sufficient for some aerosols species, relatively constant in time and space. In the case of mineral dust, nor the mean nor median values are able to accurately describe the temporal variability. An example is displayed in Fig. 3: the time series of monthly climatology of dust (modeled by the GOCART model, Ginoux et al., 2004) versus hourly modeled (with CHIMERE-DUST) shows that day to day regional forecast scores could be improved if the hourly frequency variability is taken into account.

4.2.5 Fires products emissions

Another missing or badly estimated sources are vegetation fires. As shown by Langmann et al. (2009), they represent an important input of gaseous and aerosol compounds in the atmosphere. Whether domestic wood burning emissions can be included in usual inventories, vegetation fires emissions (as prescribed and wild fires) are difficult (i) to estimate for past year simulations and (ii) a fortiori to predict in a forecast mode. Since vegetation fires strongly depend on meteorological factors and the biomass availability, they have a great inter-annual variability (Langenfelds et al., 2002). Usually, fire emissions for a specific species are calculated as the product of burned area, fuel load, combustion completeness and specific emission factors. In Europe, the Mediterranean Basin and the Portugal (Miranda et al., 1994; Hodzic et al., 2007) is often submitted to fires in summertime. During the 2003 fire outbreaks the modeled wildfire emissions caused an increase in average PM$_{2.5}$ ground concentrations from 20 to 200% (Hodzic et al., 2007). Burned areas are detected and calculated thank to satellite observations like the MODIS Rapid Response System (http://rapidfire.sci.gsfc.nasa.gov) that provides near real time products. If such data have been already used in air quality models (Roy et al., 2007; Wiedinmyer et al., 2006), these data could be used in a forecast framework by trying to elaborate a more or less complex relationship between the fire duration and the meteorological parameter forecast (wind speed, soil moisture, temperature, etc.).

4.2.6 The issue of the injection height of emissions

The issue of emission injection height concerns several types of emissions: (i) Industrial and Forest fires emissions, (ii) Industrial point source emissions and (iii) Volcanoes emissions. The injection height of smoke plumes from forest fires is a large source of uncertainty in transport models used to predict the effect of emissions from fires on air quality (Mazzoni et al., 2007). It is well known that crown fires generate sufficient energy to loft smoke plumes above the boundary layer (Lavoué et al., 2000), facilitating long-range transport
of gases and particulate matter. A large fraction of smoke aerosols remains in the near-surface boundary layer, and emissions that rapidly escape the boundary layer are more likely to contribute to long distance transport. Mazzoni et al. (2007) have shown how data mining methods applied to imagery and higher level data products from MISR and MODIS on Terra are capable of generating partially automated retrievals of smoke plume injection height.

For point source emissions (PSE), two cases are considered: first, the PSE inventory is separated from the other surface emissions, then a specific module can compute a plume rise based on Brigg’s formula (Briggs, 1971) for instance that account for thermodynamics and dynamics of emission fluxes and meteorological conditions. Or, secondly, the PSE are not separated in the emission database as in EMEP emissions in Europe (Vestreng, 2003), then vertical profiles per activity sectors are applied to distribute the emissions in the model layers.

Using German industrial point source information a comparison of data for 1996 and 2004 performed by Pregger and Friedrich (2009) shows changes in parameter values which are different depending on source category and parameter, reflecting changes within a certain plant inventory such as shutdown or retrofitting of old plants, implementation of new plants, energy saving and emission abatement measures. Some variations are influenced by data corrections and a changing scope of the reporting obligations as a result of control act amendments. The study stated that these temporal changes may be significant and concluded that an update of stack information or resulting effective emission heights was recommended if emission data and the model base year are updated.

In Europe, only volcanoes in the vicinity of Sicily are still active and contribute to pollutant emissions. The EMEP (Vestreng, 2003) database only consider SO₂ emissions. The way to inject volcanoes emissions in CTMs is quite uncertain because it depends on volcano activity. In addition, when a volcano is erected in a flat environment it is difficult to identify the model layer that fit to the top of the volcano.

### 4.3 Chemistry

All models used in forecasting use chemical gas mechanisms which are recognized as fairly stable (Seinfeld and Pandis, 1998). The main issues for the next years concern the chemistry of aerosols, and more specifically the poorly known processes of SOA formation. SOA modeling remains highly uncertain and too simple, they need to be improved by adding aqueous and heterogeneous pathways, as well as taking into account the multi-step oxidation processes and their dependence on the NO₃ regime (Pun and Seigneur, 2007). One important weakness was identified to be the estimate of biogenic precursor emissions. A first step towards improving our SOA knowledge is improvements of the biogenic emission inventories (Simpson et al., 2007; Arneth et al., 2008).

Another improvement is the quality of meteorological data feeding the model. For instance, an underestimate of temperatures implies three kind of issues: (i) an underestimate of total SOA formation because of kinetic rates favored by high temperature, (ii) an underestimate of SOA concentrations because of lower biogenic emissions, and (iii) an overestimate of gas to particle transfer of semi-volatile SOA. A recent work by Shrivastava et al. (2006) suggests that particle emission factors could be underestimated because they are calculated at given temperature, relative humidity and dilution ratio in order to avoid artefacts during the sampling process. In so doing, a non negligible fraction, the so-called SVOCs (semi-volatile VOCs), may have not been taken into account in current gas and particle emission inventories when these data are applied for modeling “real” atmospheres. Moreover, recent studies pointed out the possible role of primary SVOCs in SOA formation (Schauer et al., 2002; Donahue et al., 2006; Robinson et al., 2007). These precursors evaporate during the emission dilution process and could be converted into the particulate phase after oxidation.

### 4.4 The boundary conditions

For short term forecast and long-term trends, a specific attention must be devoted to the spatial and temporal resolution of chemical boundary conditions as highlighted in Szopa et al. (2009). This concerns all chemical species involved in the pollution budget and becomes critical when polluted air masses are stagnant or recirculate over a region. For short term events with a fast transport, too averaged boundary conditions data are not able to catch intense polluted plumes. In addition, the use of different chemical schemes between the coarse and the nested simulation raises two issues. First, the family species can largely differ between the two schemes and require strong assumption for the reallocation from the coarse to the nest. Second, the chemical regime can also be different implying boundary effects for the nested domain.

Moreover, the use of measurements remains a critical issue: if data such as ozonesondes, satellite data, could be of great help, there is currently no system able to assimilate their outputs quickly and each day before to realize a forecast.

### 5 Conclusions

The previous sections have identified all of the processes involved in calculating the atmospheric composition for the forecast. Other routes of improvements are possible and relate to methods for using these models (the ensemble methods) and use more completely measurements to better initialize the forecast calculations (data assimilation). These methods are still very new in this context and many projects (such as FP7/MACC) are currently underway to evaluate their performances.
5.1 Ensemble forecasting

The recent studies of Galmarini et al. (2004b) and Galmarini et al. (2004a) used several models in Europe to analyze trends of ozone and particles concentrations. All chemical concentrations fields are combined to constitute an “ensemble”. The idea is to have a more realistic result by combining all the results statistically. For now, these methods have mainly been applied in “analysis” mode to quantify the behavior of these systems. Van Loon et al. (2007) shows a wide dispersion between the results and this is due both to different model parameterizations but also to their horizontal resolution. Concerning surface ozone concentrations, the authors showed that ozone daily maxima are better simulated than daily averages, and summertime concentrations are better simulated than wintertime concentrations. They also showed that the use of many models in an ensemble approach may add informations for daily ozone modeling. More recently, Vautard et al. (2009) showed that ozone forecast remains very sensitive to long-range boundary conditions transport. In the same time, NO2 ensemble is not enough representative due to missing real high values by all models. If the forecast ensemble seems to be an approach to continue, it is nevertheless necessary to continue to improve the models to reduce their uncertainties.

5.2 Data assimilation benefits

One key issue regarding atmospheric composition modeling performances could be the data assimilation (Constantinescu et al., 2007; Davakis et al., 2007). Data assimilation has been well developed for weather forecasting. The nature of the problem (an “initial conditions” problem), highly non-linear, helped to significantly improve the forecast scores. But the benefit of assimilation for atmospheric composition is less obvious: the problem is more a problem of boundary conditions (emissions, transport of pollutants has long distance) (Honoré et al., 2000; Elbernt and Schmidt, 2001; Blond and Vautard, 2004). If it can be expected to somewhat improve the forecasts, the assimilated data, such as satellites, will mainly benefit to more easily update the variability of emissions, for example. These limitations would be more precisely quantified in the next years in order to optimize the use of existing (surface, sondes) and future data (satellite products).

5.3 Model resolutions and scores

Exposure studies suggest that the future needs for air quality would be to increase the horizontal and vertical resolution until a “human” scale. This is not realistic for the moment to imagine a CTM with a few meters spatial resolution and a few seconds of temporal resolution. Even if this would be possible and accurate, the amount of data would be impossible to analyze. Thus, an optimum resolution must be found.

The current consensus is to approach the “turbulence cute” spatial resolution, i.e. one kilometer. Below this limit, the eddies must be explicitly solved and the actual meteorological meso-scale models have to be replaced by LES models (Large Eddy Simulation).

The most relevant resolution is the one giving the best scores (measurements versus modeled concentrations) for the species of interest and every day, but exposure studies require the retrieval of urban scale pollutants concentrations (Baklanov et al., 2007). Model may be used until a one kilometer resolution, but the results are improved mainly for areas with well-marked orography and during very stable conditions (Fay and Neunhauserer, 2006) and not everywhere and everytime. This is mainly due to the input parameters accuracy such as the orography description, the urban roughness and the surface exchanges fluxes (Baklanov et al., 2006). But, even with the finest meteorology scale, the modeled concentrations are not necessary better. The finest resolution will increase the wind speed and the boundary layer height variability, for example: the forecast error may thus increase with the resolution and the optimum is not always the finest scale (Valari and Menut, 2008). This paradoxal result is mainly due to the fact that by increasing emissions and meteorology spatial resolutions, uncertainties also increase with the risk of model errors.

5.4 Retroactions modelling

The challenge is now to accurately model sporadic processes, such as mineral dust and fire transport, together with small urban scale processes (locations where a major part of people is living). Even if numerical capabilities are growing fast, this is not currently possible to model the whole Northern Hemisphere with a complete set of gaseous and particle species and with a one kilometer horizontal resolution. A first option is to build a forecast air quality system with multi domains and nesting (Jacobson, 2001). The nesting can be
one-way (from the coarse to the nested domain only) or two-way. A second option is to use new grid meshes such as the one presented in Fig. 4 and currently under development for the CHIMERE model (in this case, only the CTM has this type of mesh; the emissions and meteorological forcings remain in their original regular grids). The zoom represents the target for the forecast e.g. the France with a resolution of 5 km when the largest cells are 1° × 1° over the Atlantic Ocean. A more powerful approach could be the development of unstructured meshes for CTMs as currently used in computational fluid dynamics or oceanography. Currently, all models used in forecast mode do not have two-ways meteorology/CTM retroactions due to a too large numeric cost. These models exist and remain research tools mainly for climate feedbacks evaluations. Finally, retroactions from concentrations fields to meteorology are not taken into account in forecast modeling systems. The potential impact of large plumes of mineral dust or fires products on radiation will be a challenge and certainly fields of innovative research for the next years.

Acknowledgements. The COST 602 and 728 european projects and all their members are acknowledged for inputs and fruitful discussions.

Topical Editor F. D’Andrea thanks two anonymous referees for their help in evaluating this paper.

The publication of this article is financed by CNRS-INSU.

References


http://www.atmos-chem-phys.net/7/2199/2007/.


