Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN FISICA

CICLO XXIII

COORDINATORE Prof. Filippo Frontera

The Distribution of Atmospheric Pollutants in Europe: Optimal Use of Models and Observations with a Data Assimilation Approach

Settore Scientifico Disciplinare FIS/06

Dottorando

Dott. Messina Palmira

Tutore interno

Prof. Porcù Federico

Tutore esterno

Dott. Fierli Federico

Anni 2008/2010

Contents

Chapter 1

Introduction

Predicting air quality is of growing importance to society. The chemical composition of the atmosphere has been (and is being) significantly perturbed by emissions of trace gases and aerosols associated with a variety of anthropogenic activities. This changing of the chemical composition of the atmosphere has important implications for urban, regional and global air quality, and for climate change (Carmichael et al., 2008). Moreover, current health studies demonstrate that atmospheric pollutants are responsible for increasing breathing troubles in the population, with a non negligible impact on the morbidity and loss of life expectancy statistics. Studies carried out by the World Health Organization (WHO, 2006) clearly show that significant improvements are still needed to manage and control the impacts of air pollution on health. A comprehensive review of the state of the art of air quality research can be found in Monks (2000) showing that in the last three decades, atmospheric chemistry and air quality sciences have undergone increasing progress with the development of sophisticated *Chemical Transport Models* (CTMs) or more generally named *Air Quality Models* (AQMs) and with the growing extent of operational network. AQMs have become an

essential tool for providing science-based input into best alternatives for reducing urban pollution levels, for designing cost-effective emission control strategies, for the interpretation of observational data, and for assessments into how we have altered the chemistry of the global environment. The use of CTMs to produce air quality forecasts has become a new application area, providing important information to the public, decision makers and researchers.

There are different European initiatives aimed at (i) providing a fundamental understanding of the processes that control the distributions of chemical species in the atmosphere and their impact on global change and air quality, (ii) generating and publishing daily air quality forecasts and maps resulting from numerical simulations on different spatial scales, (iii) comparing daily different air quality models in order to better assess their capability in predicting chemical species using satellite and in-situ data; a non exhaustive list includes the "International Global Atmospheric Chemistry" (IGAC, http://igac.jisao.washington.edu/), "PREV'AIR system" (http://www.prevair.org/en/index.php, Honoré et al., 2008), "Global and regional Earth-system (Atmosphere) Monitoring" (GEMS, http://gems.ecmwf.int/about.jsp) .

Whereas significant advances in CTMs have taken place, predicting air quality remains a challenging problem due to the complex processes occurring at widely different scales and by their strong coupling across scales. A parallel increase of observations (i.e ground measurements, satellite data, lidar profiles) offers the possibility to both validate more extensively CTMs and to use observations to "correct" possible errors of CTMs to represent chemical species. Data assimilation (DA) techniques aimed at merging the informations brought by observations into models taking into account errors of both observation and modelled variables. DA has been used successfully in Numerical Weather Prediction (NPW) models and, since some decades, it has been implemented into several CTMs allowing a more

efficient use of measurements, thus possibly reducing model uncertainties (see for details section 2.3).

The research activity presented in this manuscript is belong to this research framework, implementing a methodology to merge in an optimal way atmospheric modelling and available observations at different spatial scales. In particular, we approach the problem of assimilation of ground measurements and satellite columnar data and how DA can improve CTMs and correct biases and errors in the chemical species forecast. The work focuses on tropospheric ozone and the species linked to its formation, since they play a crucial role in chemical processes during photochemical pollution events. Moreover, as they have a direct impact on human health, improvement of forecast accuracy and the quality of chemistry description in 3D models has a straightforward impact.

In the first part of the study we evaluate the improvement in the capability of regional models to reproduce the distribution of tropospheric pollutants, using the assimilation of surface chemical observations. Among the many causes of uncertainties of chemistry models simulations, a particular focus is given by uncertainties in emissions, that are known to be high, despite continuous improvement in compilation of emissions inventories. The scientific purpose is to analyse the efficacy of DA in correcting the biases due to perturbed emission. The work is performed using an Observing System Simulation Experiment (OSSE), which allowed the quantification of assimilation impact, through comparison with a reference state. Different sensitivity tests are carried out in order to identify how assimilation can correct perturbations on O_3, induced by NO_x emissions biased in flux intensity and time. Tests are performed assimilating different species, varying assimilation time window length and starting hour of assimilation. Emissions are biased quantitatively up to $\pm$ 50% and shifted temporally up to $\pm$ 2 hours.

The second part of PhD research activity deals with the evaluation of the

impact of assimilation of satellite NO_2 tropospheric columns on the distribution of pollutants at the ground level during photochemical pollution events at continental scale. In particular, we focus on the assimilation of observations from SCIAMACHY, on-board ENVISAT ESA satellite, and from OMI, on-board EOS Aura satellite, and its effect on ozone in the lowermost troposphere in Europe. For an effective improvement in assimilated fields it is particularly important the consistency between satellite and model resolution (Blond et al., 2007). SCIAMACHY and OMI have a considerable difference in spatial and temporal resolution, allowing to test the role of data resolution on the effectiveness of assimilation.

The study is carried out implementing and applying an Optimal Interpolation DA technique in the air quality model BOLCHEM and the chemical transport model CHIMERE. The OI routine is chosen because it has given satisfactory results in air quality modelling and because it is relatively simple and computationally inexpensive.

The thesis is organized as follow: Chapter 2 delineates the principal features of DA and the status of the art of the air quality models and their uncertainties. Chapter 3 presents the tool implemented in this work (AQMs and the Optimal Interpolation routine) together with observations used for validation and assimilation. Chapter 4 illustrates assimilation of ground measurements to evaluate to which extent DA can correct errors in pollutants simulations caused by biases in emissions. In Chapter 5 we focus on assimilation of tropospheric satellite data and the estimate of the impact on tropospheric ozone concentration. Conclusion will focus on the summary and discussion of the principal results and gives a short overview of research perspectives.

Chapter 2

Air quality models and data assimilation techniques

2.1 Analysis and predictability of air quality models and pollutants

Tropospheric chemical composition and hence the air quality is determined by an interplay of processes at different spatial and temporal scales. The concentration of pollutants in a peculiar region is a result of local (i.e. point emissions, turbolent mixing), regional (i.e. photochemistry, local scale circulation) and continental (long-range transboundary pollution) physics and chemistry, with an impact towards the global scale.

Figure 2.1 shows the NO_2 columns distribution measured by OMI. It can be taken as representative of photochemical pollution occurring in summer. The major European areas involved are, for instance, Po Valley, Ruhr region, Belgium, South England, Paris, Madrid, Athens and the other principal cities in Europe.

Within this frame, air quality models (AQMs) are key instrument for

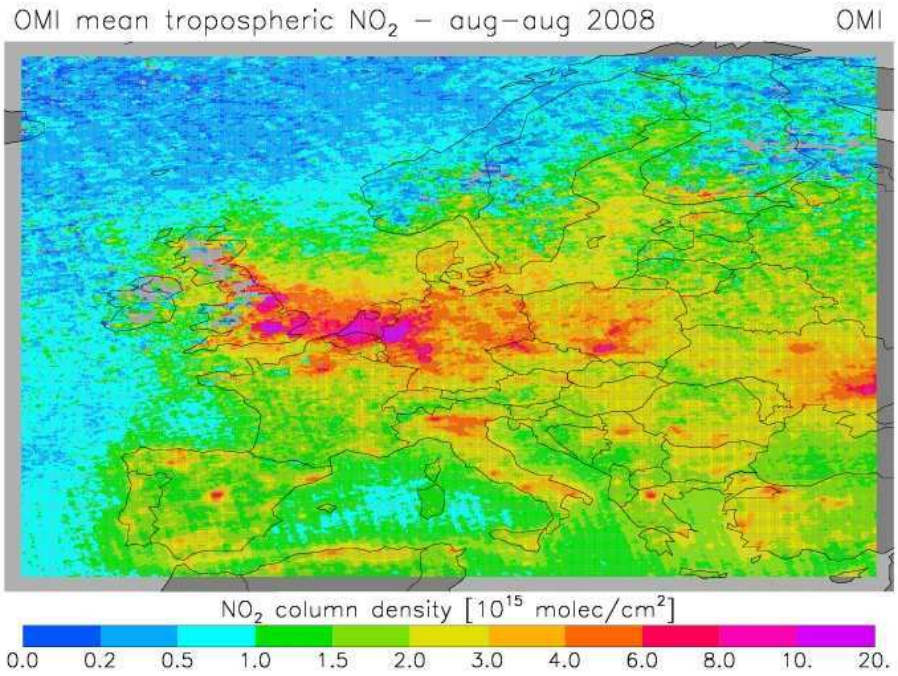

Figure 2.1: OMI mean tropospheric NO_2 in august 2008 (www.knmi.nl/research/climate_chemistry).

understanding and forecasting chemical pollution (see, for instance, Lawrence et al., 2003; Sportisse, 2007; Zhang, 2008), since they can be used in many applications ranging from the study of short-range dispersion of species (typically accidental release in the case of an industrial hazard) to atmospheric chemistry and upscaling to climate, as now widely recognised in the scientific community. Many chemical species are of interest for instance: ozone and volatile organic compounds (photochemistry), trace metals, mercury, methane, carbon monoxide, particulate matter (aerosols), radionuclides, biological species, etc.

In brief, AQMs simulate the time evolution of spatial fields for a set of chemical species (by extent of radionuclides and biological species). We refer to Figure 2.2 for the ensemble of processes generally described by AQM state of the art. Some primary species (for instance, nitrogen oxide or volatile organic compounds) are emitted either by anthropogenic or biogenic sources (surface emission or point emission). The species are then vertically diffused in the atmospheric boundary layer by turbulent eddies related to both mechanical forces (wind shear) and thermal forces (buoyancy), while the horizontal motion is due to wind advection. Gas-phase chemical reactions, mainly related to the oxidizing power of the atmosphere and to the radiative fluxes (through photolysis), lead to the production of secondary species (typically ozone, O_3). Mass transfer between gas phase, aqueous phase (cloud droplets) and particulate matter (solid or liquid particles in suspension) may also happen. Moreover, the evolution of aerosols is governed by microphysical processes such as nucleation (the formation of small clusters of gaseous molecules), coagulation (collision between particles), condensation/evaporation (mass transfer between the semi-volatile species and particles) or activation (the growth of aerosols to cloud droplets through condensation of water vapor). The loss processes from the atmosphere are dry deposition (when gases or particles impinges upon and stick to the surface)

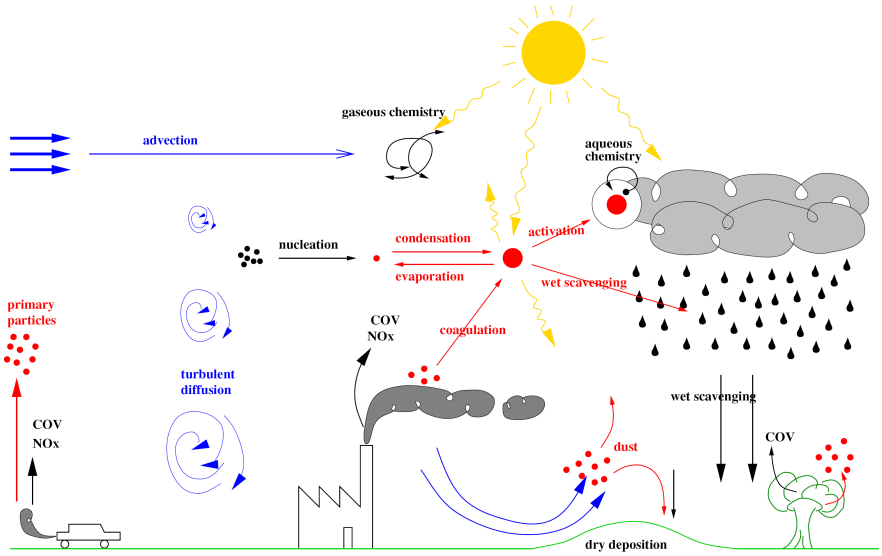

Figure 2.2: Scheme of processes described in a chemistry transport model (Sportisse, 2007).

and wet scavenging (washout by rains).

2.1.1 Air quality models uncertainties

Despite of the many developments, air quality modelling and simulations still suffer from many uncertainties (see for instance, Tilmes et al., 2002; Russell and Dennis, 2000) due to the large range of scale and high number of physical and chemical processes to be represented:

- Many input data are poorly known (for instance, ozone, NO_x and VOC initial conditions, initial conditions, NO_x /VOC emissions, biogenic emissions, rainfall amount, cloud liquid water content). Moreover, some forcing fields, such as meteorological fields, are computed with numerical models that may be uncertain.

- Physical and chemical parametrizations induce a high sensitivity to the way to represent process and to the values of parameter used to represent sub-grid processes.

- The numerical algorithms and the discretization also induce uncertainties, especially when resolution is coarse because of the computational burden, especially for aerosols.

Even if the models are "validated" (it is more rigorous to say that model-to-data comparisons are performed, when possible), one must keep in mind that there are a large amount of degrees of freedom (especially in parameterizations) and only a small number of model outputs can be measured. For instance, most of the existing CTMs have been extensively tuned to meet acceptable model-to-data error statistics for ozone peaks at ground. It does not ensure that the results are satisfactory for 3D fields and many other trace species. There is therefore a danger of using "overtuned" models, especially for impact studies or long-term scenario studies (Sportisse, 2007).

A particular remark has to be done to uncertainties in emissions (Sawyer et al., 2000; Simpson et al., 1999; Mendoza-Dominguez and Russell, 2001; Hanna et al., 2001, 1998). They in fact represent one of the major source of uncertainties in air quality models, since they continue to be high, despite many efforts made to provide a more accurate emissions inventory.

Emissions are supplied by annual total inventories, for instance EMEP Centre on Emission Inventories and Projections (CEIP) provide the main pollutants (NOx, SOx, CO, NH3 and NMVOCs), heavy metals, POP and PM, divided by production sectors and countries. For modelling purposes they are temporal and by-species disaggregated from annual total to hourly values profiles and to each activity (traffic, industry, energy extraction/production, residential, agriculture) and next

they are aggregated into model species. This whole procedure is based on a priori assumptions and can be greatly affected by errors (Tao et al., 2004; Placet et al., 2000; Tilmes et al., 2002; Menut et al., 2000). Hanna et al. (1998) and Mallet and Sportisse (2005) considered that NO_x emissions could be affected by an uncertainty of 30%-50%, while Beekmann and Derognat (2003) confirmed this assumption, considering an uncertainty of 40%; Hanna et al. (2001) considered NO_x biogenic emission uncertainties up to a factor of two.

Moreover, the impact of the temporal allocation of emissions on pollutant distribution is still under debate: Mallet and Sportisse (2005) found that ozone concentration can be sensitive to different temporal allocation of emissions, while Tao et al. (2004) found that daytime O_3 concentrations are slightly dependent on changes in the temporal allocation of NO_x emissions.

Some possible strategy for assessing the impact of uncertainties is to evaluate the sensitivity of some model outputs, with respect to uncertain input parameters. We refer for instance to for Carmichael et al. (1997); Sandu et al. (2003, 2005); He et al. (2000); Seigneur et al. (1982) a general presentation, Mallet and Sportisse (2005) for the sensitivity of ozone with respect to emissions, Zhang et al. (2005) for the sensitivity of ozone.

2.2 Data assimilation techniques

2.2.1 Historical notes and definition

In the late sixties, the development of satellite observing system, and the perspective that asynoptic observations, performed more or less continuously in time, would become more and more numerous in future, led to the notion that the dynamical evolution of the flow should be explicitly taken into account in the

very definition of the initial conditions of the forecast. The word *assimilation* was coined at that time denoting a process in which observations distributed in time are merged together with a dynamical numerical model of the flow in order to determine as accurately as possible the state of the atmosphere.

Since then, continuous in theory, in efficiency of numerical algorithms, as well as in available computing power, has led to a slow but steady progress in the method for assimilation. This progress, together with improvements to the quality of Numerical Weather Prediction (NPW) models and to a lesser extent, with the improvements in the observing system, has significantly contributed to the continuous increase observed in the last decades in the quality of numerical weather forecast. It is worth mentioning that the proportion of resources allocated to assimilation in the whole process of NWP has steady increase over time. This evolution has not resulted from a clearly stated voluntary choice, but it would be more appropriately described as a progressive "natural selection" process, during which increase of the proportion of resource allocated to assimilation repeatedly and consistently proved to be beneficial (Talagrand, 1997)

If assimilation of observations originated from the need of NWP, assimilations has already proven to be useful for other purpose than weather prediction see for example (Salstein and Rosen, 1986; Oort, 1989) and other fields of atmospheric modelling (climate modelling, air quality modelling, dynamical oceanography), (see Section 2.3).

An operational definition of data assimilation is "an analysis technique in which the observed information is accumulated into the model state by taking advantage of consistency constraints with laws of time evolution and physical properties". There are two basic approaches to data assimilation: sequential assimilation, that only considers observation made in the past until the time of analysis, which is the case of real-time assimilation systems, and non-sequential, or retrospective

assimilation, where observation from the future can be used, for instance in a reanalysis exercise. Another distinction can made between methods that are intermittent or continuous in time. In an intermittent method, observations can be processed in small batches, which is usually technically convenient. In a continuous method, observation batches over longer periods are considered, and the correction to the analysed state is smooth in time, which is physically more realistic. The four basic types of assimilation are depicted schematically in Figure 2.3. Compromises between these approaches are possible.

Both models and observations are not perfect and it is crucial in data assimilation formulation to refer to their uncertainties. This lead to a mathematical representation of errors through a statistical approach.

Many assimilation techniques have been developed for meteorology and oceanography (Figure 2.4). They differ in their numerical cost, their optimality, and in their suitability for real-time data assimilation. Some of them are explained in the sections hereafter.

Some basic references on assimilations techniques are (Daley, 1991; Lorenc, 1986; Ghil and Malanotte-Rizzoli, 1991; D'Isidoro, 2005).

2.2.2 Mathematical formalisation

The first step in the mathematical formalisation of the analysis problem is the definition of the work space. As in a forecast model, the collection of numbers needed to represent the atmospheric state of the model is collected as a column matrix called the *state vector* $\mathbf{x}$. How the vector components relate to the real state depend on the choice of discretization, which is mathematically equivalent to a choice of basis.

One must distinguish between reality itself (which is more complex than what

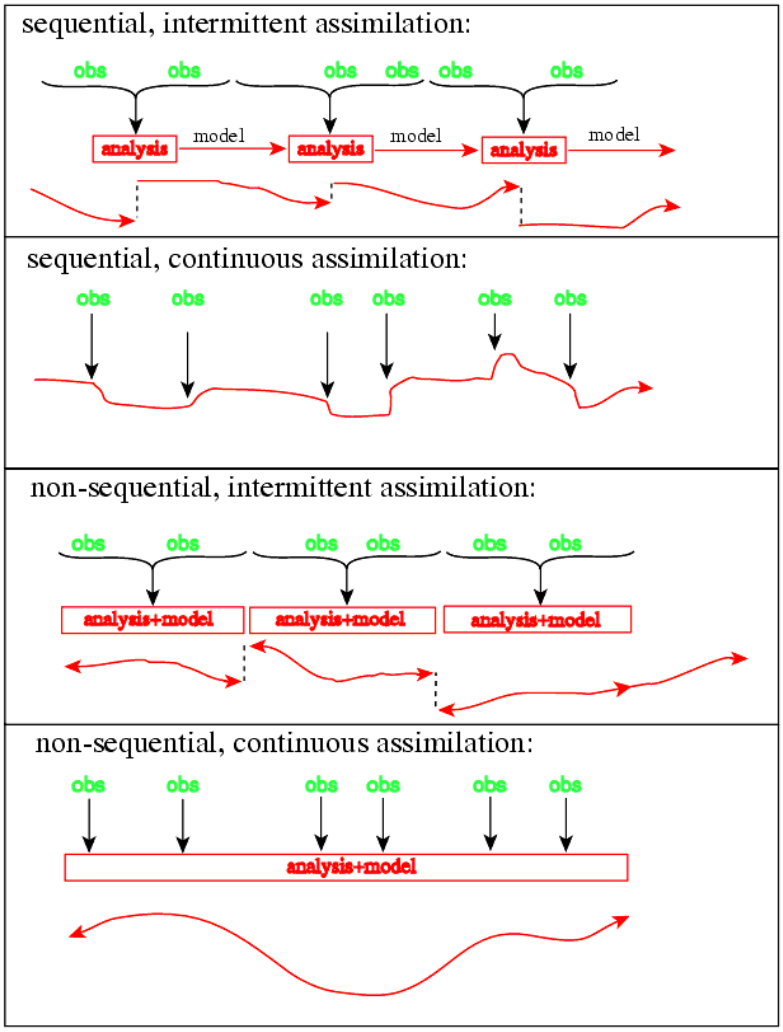

Figure 2.3: Representation of four basic strategies for data assimilation, as a function of time. The way the time distribution of observations ("obs") is processed to produce a time sequence of assimilated states (the lower curve in each panel) can be sequential and/or continuous (Bouttier and Courtier, 1999).

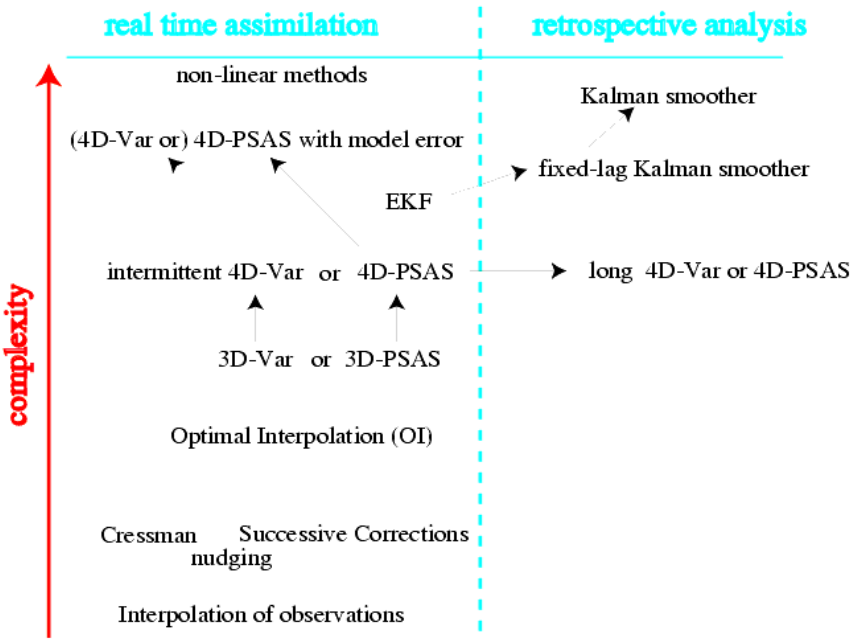

Figure 2.4: A summarized history of the main data assimilation algorithms used in meteorology and oceanography, roughly classified according to their complexity (and cost) of implementation, and their applicability to real-time problems. Currently, the most commonly used for operational applications are OI, 3D-Var and 4D-Var (Bouttier and Courtier, 1999).

can be represented as a state vector) and the best possible representation of reality as a state vector, which it shall be denoted $\mathbf{x_t}$, the true state at the time of the analysis. Another important value of the state vector is $\mathbf{x_b}$, the a priori or background estimate of the true state before the analysis is carried out, valid at the same time. Finally, the analysis is denoted $\mathbf{x_a}$, which is what we are looking for.

The analysis problem is to find a correction $\delta\mathbf{x}$ (or analysis increment) such that

$$\mathbf{x_a} = \mathbf{x_b} + \delta\mathbf{x} \tag{2.1}$$

is as close as possible to $\mathbf{x_t}$.

For a given analysis a number of observed values are used. They are gathered into an *observation vector* $\mathbf{y}$. To use them in the analysis procedure it is necessary to be able to compare them with the state vector. In general there are fewer observations than variables in the model and they are irregularly disposed, so that the only correct way to compare observations with the state vector is through the use of a function from model state space to observation space called an *observation operator* that we will denote by H. H is a collection of interpolation operators from the model discretization to the observation points, and conversions from model variables to the observed parameters.

The key to data analysis is the use of the discrepancies between observations and state vector. This is given by the vector of departures at the observation points:

$$\mathbf{y} - \mathbf{H}(\mathbf{x}) \tag{2.2}$$

When calculated with the background $\mathbf{x_b}$ it is called *innovations*, and with the *analysis* $\mathbf{x_a}$, analysis residuals. Their study provides important information about the quality of the assimilation procedure.

2.2.3 Sources of errors and their estimation

To represent the fact that there is some uncertainty in the background, in the observations and in the analysis we will assume some model of the errors between these vectors and their true counterparts. The correct way to do this is to assume some *probability density function (pdf)* of error.

Given a background field $\mathbf{x_b}$ just before doing an analysis, there is one and only one vector of errors that separates it from the true state:

$$\epsilon_\mathbf{b} = \mathbf{x_b} - \mathbf{x_t} \tag{2.3}$$

If each analysis experiment can be repeated a large number of times, under exactly the same conditions, but with different realizations of errors generated by unknown causes, $\epsilon_\mathbf{b}$ would be different each time. Statistics such as averages, variances and histograms of frequencies of $\epsilon_\mathbf{b}$ can be calculated. In the limit of a very large number of realizations, it is expected that the statistics to converge to values which depend only on the physical processes responsible for the errors, not on any particular realization of these errors. When we do another analysis under the same conditions, we do not expect to know what will be the error $\epsilon_\mathbf{b}$, but at least we will know its statistics. The probability density function of $\epsilon_\mathbf{b}$ is the best information about the distribution of $\epsilon_\mathbf{b}$ and from this function one can derive all statistics, including the average (or expectation) $\langle\epsilon_\mathbf{b}\rangle$ and the variances. A popular model of scalar pdf is the Gaussian function, which can be generalized to a multivariate pdf.

The errors in the background and in the observations are modelled as follows:

- **background errors**: $\epsilon_\mathbf{b} = \mathbf{x_b}$ - $\mathbf{x_t}$, of average $\langle\epsilon_\mathbf{b}\rangle$ and covariances $\mathbf{B} = \langle(\epsilon_\mathbf{b} - \langle\epsilon_\mathbf{b}\rangle)(\epsilon_\mathbf{b} - \langle\epsilon_\mathbf{b}\rangle)^\mathbf{T}\rangle$. They are the estimation errors of the background

state, i.e. the difference between the background state vector and its true value.

- **observation errors**: $\epsilon_{\mathbf{o}} = \mathbf{y} - \mathbf{H}(\mathbf{x_t})$, of average $\langle\epsilon_{\mathbf{o}}\rangle$ and covariances $\mathbf{R} = \langle(\epsilon_{\mathbf{o}} - \langle\epsilon_{\mathbf{o}}\rangle)(\epsilon_{\mathbf{o}} - \langle\epsilon_{\mathbf{o}}\rangle)^{\mathbf{T}}\rangle$. They contain errors in the observation process (instrumental errors, because the reported value is not a perfect image of reality), errors in the design of the operator H, and representativeness errors i.e. discretization errors which prevent $\mathbf{x_t}$ from being a perfect image of the true state.

- **analysis errors**: $\epsilon_{\mathbf{a}} = \mathbf{x_a}$ - $\mathbf{x_t}$, of average $\langle\epsilon_{\mathbf{a}}\rangle$. A measure of these errors is given by the trace of the analysis error covariance matrix $\mathbf{A} = \langle(\epsilon_{\mathbf{a}} - \langle\epsilon_{\mathbf{a}}\rangle)(\epsilon_{\mathbf{a}} - \langle\epsilon_{\mathbf{a}}\rangle)^{\mathbf{T}}\rangle$.
  They are the estimation errors of the analysis state, which is what we want to minimize.

Error covariances are more subtle and we will illustrate this with the background errors (all remarks apply to observation errors too). In a scalar system, the background error covariance is simply the variance, i.e. the root-mean-square (or r.m.s., or quadratic) average of departures from the mean:

$$\mathbf{B} = \mathbf{var}(\epsilon_{\mathbf{b}}) = \mathbf{var}(\langle(\epsilon_{\mathbf{b}} - \epsilon_{\mathbf{t}})^{\mathbf{2}}\rangle) \tag{2.4}$$

In a multidimensional system, the covariances are a square symmetric matrix. If the model state vector has dimension n, then the covariances are an $n \times n$ matrix. The diagonal of the matrix contain variances, for each variable of the model; the off-diagonal terms are cross-covariances between each pair of variables of the model. The matrix is positive. Unless some variances are zero, which happens only in the rather special case where one believes some features are perfect in the background,

the error covariance matrix is positive definite. For instance if the model state is tridimensional, and the background errors (minus their average) are denoted (e_1,e_2, e_3), then

$$B = \begin{bmatrix} var(e_1) & cov(e_1, e_2) & cov(e_1, e_3) \\ cov(e_2, e_1) & var(e_2) & cov(e_2, e_3) \\ cov(e_3, e_1) & cov(e_3, e_2) & var(e_3) \end{bmatrix} \tag{2.5}$$

The off-diagonal terms can be transformed into error correlations (if the corresponding variances are non zero):

$$\rho(e_i, e_j) = \frac{cov(e_i, e_j)}{\sqrt{var(e_i)var(e_j)}} \tag{2.6}$$

It is worth to underline that the error statistics are functions of the physical processes governing the meteorological situation and the observing network. They also depend on our a priori knowledge of the errors. Error variances in particular reflect our uncertainty in features of the background or the observations. In general, the only way to estimate statistics is to assume that they are stationary over a period of time and uniform over a domain so that one can take a number of error realizations and make empirical statistics. This is in a sense a climatology of errors. Another empirical way to specify error statistics is to take them to be a fraction of the climatological statistics of the fields themselves.

2.2.4 Statistical interpolation with least squares estimation

In this section it is present the fundamental equation for linear analysis in a general algebraic form: *the least squares estimation*, also called *Best Linear Unbiased Estimator (BLUE)* (see for instance Bouttier and Courtier, 1999). This theorem is the basis of the most utilized data assimilation techniques as Optimal Interpolation, 3D-Var, Kalman Filtering, 4D-Var.

Given the dimension of the model state n and the dimension of the observation vector p, it is denoted:

- $\mathbf{x}_t$ true model state (dimension n)
- $\mathbf{x}_b$ background model state (dimension n)
- $\mathbf{x}_a$ analysis model state (dimension n)
- $\mathbf{y}$ vector of observations (dimension p)
- $\mathbf{H}$ observation operator (from dimension n to p)
- $\mathbf{B}$ covariance matrix of the background errors ($\mathbf{x}_b$ - $\mathbf{x}_t$) (dimension $n \times n$)
- $\mathbf{R}$ covariance matrix of observation errors ($\mathbf{y}$ - H [$\mathbf{x}_t$]) (dimension $p \times p$)
- $\mathbf{A}$ covariance matrix of the analysis errors ($\mathbf{x}_a$ - $\mathbf{x}_t$) (dimension $n \times n$)

The following hypotheses are assumed:

- **Linearized observation operator**: the variations of the observation operator in the vicinity of the background state are linear: for any $\mathbf{x}$ close enough to $\mathbf{x_b}$, $H(\mathbf{x})$ - H ($\mathbf{x_b}$) = $\mathbf{H}(\mathbf{x}$ - $\mathbf{x_b})$ where $\mathbf{H}$ is a linear operator.
- **Non-trivial errors**: $\mathbf{B}$ and $\mathbf{R}$ are positive definite matrices.
- **Unbiased errors**: the expectation of the background and observation errors is zero i.e. $\langle \mathbf{x_b} - \mathbf{x_t} \rangle = \langle \mathbf{y} - \mathbf{H}(\mathbf{x_t}) \rangle = 0$
- **Uncorrelated errors**: observation and background errors are mutually uncorrelated i.e. $\langle (\mathbf{x_b} - \mathbf{x_t})(\mathbf{y} - \mathbf{H}(\mathbf{x_t}))^{\mathbf{T}} \rangle = 0$
- **Linear analysis**: we look for an analysis defined by corrections to the background which depend linearly on background observation departures.

- **Optimal analysis**: we look for an analysis state which is as close as possible to the true state in an r.m.s. sense (i.e. it is a minimum variance estimate).

Under these hypothesis it can be demonstrated that:

(a) the *optimal least-squares estimator*, or *BLUE analysis*, is defined by the following interpolation equations:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - H[\mathbf{x}_b]) \tag{2.7}$$

$$\mathbf{K} = \mathbf{B}\mathbf{H}^{\mathbf{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathbf{T}} + \mathbf{R})^{-\mathbf{1}} \tag{2.8}$$

where the linear operator $\mathbf{K}$ is called the gain, or weight matrix, of the analysis.

(b) The *analysis error covariance matrix* is, for any $\mathbf{K}$:

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}(\mathbf{I} - \mathbf{K}\mathbf{H})^{\mathbf{T}} + \mathbf{K}\mathbf{R}\mathbf{K}^{\mathbf{T}} \tag{2.9}$$

If $\mathbf{K}$ is the optimal least-squares gain, the expression becomes

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B} \tag{2.10}$$

(c) The BLUE analysis is equivalently obtained as a solution to the *variational optimization problem*:

$$\mathbf{x}_a = Arg(minJ) \tag{2.11}$$

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + (\mathbf{y} - H[\mathbf{x}])^T\mathbf{R}^{-1}(\mathbf{y} - H[\mathbf{x}]) \tag{2.12}$$

$$= J_b(\mathbf{x}) + J_o(\mathbf{x})$$

where is called the *cost function* of the analysis (or *misfit*, or *penalty function*), J_b is the *background term*, J_o is the observation term.

(d) The analysis $\mathbf{x}_a$ is *optimal*: it is closest in an r.m.s. sense to the true state $\mathbf{x}_t$.

(e) If the background and observation error pdfs are Gaussian, then $\mathbf{x}_a$ is also the *maximum likelihood estimator* of $\mathbf{x}_t$.

In Figure 2.5 the dimensions of the matrix operators involved in computing the analysis is delineated.

2.2.5 Optimal Interpolation

The OI is an algebraic simplification of the computation of the weight $\mathbf{K}$ in the analysis equations (2.7) and (2.8). The equation (2.7) can be regarded as a list of scalar analysis equations, one per model variable in the vector $\mathbf{x}$.

For each model variable the analysis increment is given by the corresponding line of $\mathbf{K}$ times the vector of background departures $(\mathbf{y} - H[\mathbf{x}_b])$. The fundamental hypothesis in OI is: *For each model variable, only a few observations are important in determining the analysis increment.* It is implemented as follows:

- For each model variable $\mathbf{x}(i)$, select a small number of observations using empirical selection criteria.
- Form the corresponding list of background departures $(\mathbf{y} - H[\mathbf{x}_b])$, the background error covariances between the model variable $\mathbf{x}(i)$ and the model state interpolated at the p_i observation points (i.e. the relevant p_i coefficients of the i-th line of $\mathbf{BH^T}$), and the $p_i \times p_i$ background and observation error covariance submatrices formed by the restrictions of $\mathbf{HBH^T}$ and $\mathbf{R}$ to the selected observations.
- Invert the $p_i \times p_i$ positive definite matrix formed by the restriction of ($\mathbf{HBH^T}$ + $\mathbf{R}$) to the selected observations

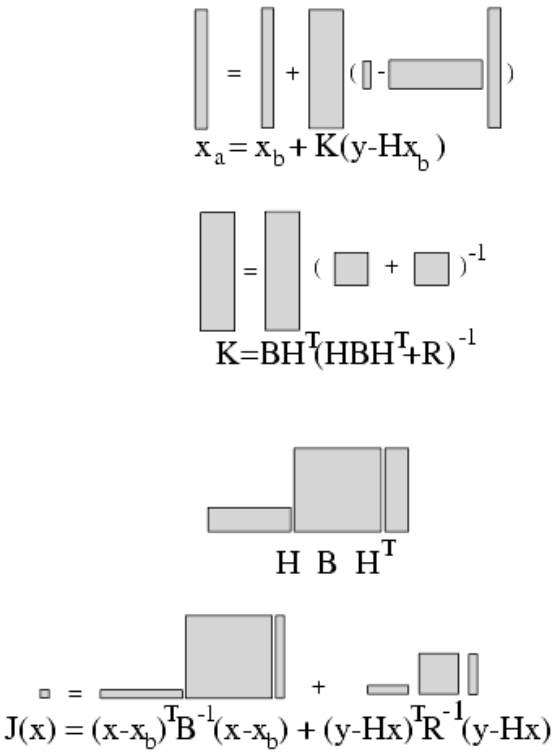

Figure 2.5: Sketches of the shapes of the matrices and vector dimensions involved in an usual analysis problem where there are many fewer observations than degrees of freedom in the model: from top to bottom, in the equations of the linear analysis, the computation of $\mathbf{K}$, of the $\mathbf{HBH^T}$ term, and the computation of the cost function J (Bouttier and Courtier, 1999).

- Multiply it by the i-th line of $\mathbf{BH^T}$ to get the necessary line of $\mathbf{K}$.

In the OI algorithm it is necessary to have the background error covariances $\mathbf{B}$ as a model which can easily be applied to pairs of model and observed variables, and to pairs of observed variables. This can be difficult to implement if the observation operators are complex. On the other hand, the $\mathbf{B}$ matrix needs not be specified globally, it can be specified in an "ad hoc" way for each model variable, as long as it remains locally positive definite. The specification of $\mathbf{B}$ usually relies on the design of empirical autocorrelation functions (e.g. Gaussian or Bessel functions and their derivatives), and on assumed amounts of balance constraints like hydrostatic balance or geostrophy (Bouttier and Courtier, 1999; Kalnay, 2003; D'Isidoro, 2005).

2.2.6 3D-Var

The principle of 3D-Var is to avoid the computation of the gain $\mathbf{K}$ completely by looking for the analysis as an approximate solution to the equivalent minimization problem defined by the cost function in 2.12. The solution is sought iteratively by performing several evaluations of the cost function

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + (\mathbf{y} - H[\mathbf{x}])^T \mathbf{R}^{-1} (\mathbf{y} - H[\mathbf{x}])$$

and of its gradient

$$\nabla J(\mathbf{x}) = 2\mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) - 2H^T \mathbf{R}^{-1} (\mathbf{y} - H[\mathbf{x}])$$

in order to approach the minimum using a suitable descent algorithm. The approximation lies in the fact that only a small number of iterations are performed. The minimization can be stopped by limiting artificially the number of iterations, or by requiring that the norm of the gradient $\|\nabla J(\mathbf{x})\|$ decreases by a predefined

amount during the minimization, which is an intrinsic measure of how much the analysis is closer to the optimum than the initial point of the minimization.

A significant difficulty with 3D-Var is the need to design a model for $\mathbf{B}$ that properly defines background error covariances for all pairs of model variables. The popularity of 3D-Var stems from its conceptual simplicity and from the ease with which complex observation operators can be used, since only the operators and the adjoints of their tangent linear need to be provided, whereas OI requires a background error covariance model between each observed variable and each model variable. Weakly non-linear observation operators can be used, with a small loss in the optimality of the result. As long as is strictly convex, there is still one and only one analysis (Bouttier and Courtier, 1999; Kalnay, 2003; Talagrand, 1997).

2.2.7 Extended Kalman Filter

Kalman filter (KF) is formally very similar to OI, but with one major difference: the forecast or background error covariance $\mathbf{P}^f(t_i)$ is advanced using the model itself, rather than estimating it as a constant covariance matrix $\mathbf{B}$. Following the notation of (Ide, 1997), let $\mathbf{x}^f(t_i) = M_{i-1}[\mathbf{x}^a(t_{i-1})]$ represent the (nonlinear) model forecast that advances from the previous analysis time $i-1$ to the current i. The model is imperfect. Therefore, it is assumed that for the true atmosphere

$$\mathbf{x}^t(t_i) = M_{i-1}[\mathbf{x}^t(t_{i-1})] + \eta(t_{i-1}) \tag{2.13}$$

where η is a noise process with zero mean and covariance matrix $\mathbf{Q}_{i-1} = \langle(\eta_{i-1})(\eta_{i-1}^T)\rangle$ that represent the model error.

In the *extended Kalman Filer*, the forecast error covariance is obtained linearizing the model about the non-linear trajectory of the model between $i-1$ e i, so that if we introduce a perturbation in the initial condition, the final perturbation

is given by

$$\mathbf{x}(t_i) + \delta\mathbf{x}(t_i) = M_{i-1}[\mathbf{x}(t_{i-1}) + \delta\mathbf{x}(t_{i-1})] = \tag{2.14}$$

$$M_{i-1}[\mathbf{x}(t_{i-1})] + \mathbf{M}_{i-1}\delta\mathbf{x}(t_{i-1}) + O(|\delta\mathbf{x}|^2) \tag{2.15}$$

The *linear tangent model* $\mathbf{M}_{i-1}$ is a matrix that transforms an initial perturbation at time t_{i-1} to the final perturbation at time t_i. The transpose of the linear tangent model or *adjoint model* makes the inverse i.e. transforms the perturbation from time t_i to time t_{i-1}.

If there are several steps in a time interval t_0 - t_i, the linear tangent model that advances a perturbation from t_0 to t_i is given by the product of the linear tangent model matrices that advance it over each step:

$$\mathbf{M}(t_0, t_i) = \prod_{j=i-1}^{0} \mathbf{M}(t_j, t_{j+1}) = \prod_{j=i-1}^{0} \mathbf{M}_j = \mathbf{M}_{i-1}\mathbf{M}_{i-2}...\mathbf{M}_0 \tag{2.16}$$

Therefore, the adjoint model is given by:

$$\mathbf{M}(t_i, t_0)^T = \prod_{0}^{j=i-1} \mathbf{M}(t_{j+1}, t_j)^T = \prod_{0}^{j=i-1} \mathbf{M}_j^T \tag{2.17}$$

Equation 2.17 shows that the adjoint model "advances" a perturbation backwards in time, from the final to the initial time. As done in OI and 3D-Var, observations are assumed to have a random errors with zero mean and an observational error covariance matrix $\mathbf{R}_i = \langle(\epsilon_i^o \epsilon_i^{o^T})\rangle$ and $\mathbf{y}_i^o = H(\mathbf{x}^t(t_i)) + \epsilon_i^o$ with observation operator H.

Note that the forecast error depends on the initial (analysis) error and on the error introduced by the forecast model during that period:

$$\epsilon_i^o = M_{i-1} + \eta_i - M_{i-1}(\mathbf{x}_{i-1}^a) = \tag{2.18}$$

$$M_{i-1}(\mathbf{x}_{i-1}^a + \mathbf{x}_{i-1}^t - \mathbf{x}_{i-1}^a) + \eta_i - M_{i-1}(\mathbf{x}_{i-1}^a) \approx \mathbf{M}_{i-1}\epsilon_{i-1}^a + \eta_i$$

The analysis and the forecast error covariance are defined from their corresponding errors at the appropriate time:

$$P_i = \langle(\epsilon_i \epsilon_i^T)\rangle$$

From these equations we can define extended Kalman filtering which consists of a "forecast step" that advances the forecast and the forecast error covariance, followed by an "analysis" or update step, a sequence analogous to OI. After the forecast step, an optimal weight matrix or *Kalman gain matrix* is calculated as in OI, and this matrix is used in the analysis step.

The forecast step is

$$\mathbf{x}^f(t_i) = M_{i-1}[\mathbf{x}^a(t_{i-1})] \tag{2.19}$$

$$\mathbf{P}^f(t_i) = \mathbf{M}_{i-1}\mathbf{P}^a(t_{i-1})\mathbf{M}_{i-1}^T + \mathbf{Q}(t_{i-1}) \tag{2.20}$$

The analysis step is written as in OI, with

$$\mathbf{x}^a(t_i) = \mathbf{x}^f(t_i) + \mathbf{K}_i\mathbf{d}_i \tag{2.21}$$

$$\mathbf{P}^a(t_i) = (\mathbf{I} - \mathbf{K}_i\mathbf{H}_i)\mathbf{P}^f(t_i) \tag{2.22}$$

where

$$\mathbf{d}_i = \mathbf{y}_i^o - H[\mathbf{x}^f(t_i)] \tag{2.23}$$

is the observational increment or innovation

The formula for the Kalman gain, computed after completing the forecast step, is obtained by minimizing the analysis error covariance $\mathbf{P}_i^a$. It is given by the same formula derived from OI, but with the constant background error covariance $\mathbf{B}$ replaced by the evolved forecast error covariance $\mathbf{P}^f(t_i)$:

$$\mathbf{K}_i = \mathbf{P}^f(t_i)\mathbf{H}_i^T[\mathbf{H}_i\mathbf{P}^f(t_i)\mathbf{H}_t^T + \mathbf{R}_i]^{-1} \tag{2.24}$$

The updating of the forecast error covariance matrix ensures that the analysis takes into account the forecast error. Unfortunately the extended Kalman filter is exceedingly expensive, since the linear model matrix $\mathbf{M}_{i-1}$ as size n, the number of degree of freedom of a modern model($n \approx 10^6$) and updating the error covariance is equivalent to performing $O(n)$ model integration. This method is simplified through some assumptions.

One promising simplification of Kalman filtering is *ensemble Kalman filtering*. In this approach, an ensemble of K data assimilation cycles is carried out simultaneously (Houtekamer and Mitchell, 2001; Houtekamer et al., 1996; Hamill et al., 2001). All the cycles assimilate the same real observations, but in order to maintain them realistically independent, different sets of random perturbations are added to the observations assimilated in each member of the ensemble data assimilations. This ensemble of data assimilation system can be used to estimate the forecast error covariance. After completing the ensemble of analysis at time t_{i-1}, and the K forecast $\mathbf{x}_k^f(t_i) = M_{i-1}^k[\mathbf{x}_k^a(t_{i-1})]$, one can obtain an estimate of the forecast error covariance from the K forecast $\mathbf{x}_k^f(t_i)$. For example one could assume:

$$\mathbf{P}^f \approx \frac{1}{K-1}\sum_{k=1}^{K}\left(\mathbf{x}_k^f - \overline{\mathbf{x}^f}\right)\left(\mathbf{x}_k^f - \overline{\mathbf{x}^f}\right)^T$$

where $\overline{\mathbf{x}^f}$ represents the ensemble average.

The ensemble Kalman filtering approach has several advantages: (a) K is of the order of 10-100, so that the computational cost (comapred with OI and 3D-Var) is increased by a factor 10-100. Although this increase cost may seem large, it is small compared to extended Kalman filtering, which requires a cost increase of order of number of degree of freedom of the model. (b) Ensemble Kalman filtering does not require the development of a linear and adjoint model. (c) It does not require the linearization of the evolution of the forecast error covariance. (d) It may provide useful initial perturbations for ensemble forecasting (Kalnay, 2003).

2.2.8 4D-Var

4D-Var is a simple generalization of 3D-Var for observations that are distributed in time. The equations are the same, provided the observation operators are generalized to include a forecast model that will allow a comparison between the model state and the observations at the appropriate time.

Over a given time interval, the analysis being at the initial time, and the observations being distributed among n times in the interval, we denote by the subscript i the quantities at any given observation time i. Hence, $\mathbf{y}_i$, $\mathbf{x}_i$ e $\mathbf{x}_i^t$ are the observations, the model and the true states at time i, and $\mathbf{R}_i$ is the error covariance matrix for the observation errors $\mathbf{y}_i - H_i(\mathbf{x}_i^t)$. The observation operator H_i at time i is linearized as $\mathbf{H}_i$. The background error covariance matrix $\mathbf{B}$ is only defined at initial time, the time of the background $\mathbf{x}_b$ and of the analysis $\mathbf{x}_a$.

In its general form, 4D-Var analysis is defined as the minimization of the following cost function:

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \sum_{i=0}^{n} (\mathbf{y}_i - H_i[\mathbf{x}_i])^T \mathbf{R}_i^{-1} (\mathbf{y}_i - H_i[\mathbf{x}_i]) \qquad (2.25)$$

which can be proven, like in the three-dimensional case detailed previously, to be equivalent to finding the maximum likelihood estimate of the analysis subject to the hypothesis of Gaussian errors. The 4D-Var analysis problem is by convention defined as the minimization problem of the (2.25) subject to the strong constraint that the sequence of model states $\mathbf{x}_i$ must be a solution of the model equations:

$$\forall i, \mathbf{x}_i = M_{0\to i}(\mathbf{x})$$

where $M_{0\to i}(\mathbf{x})$ is a predefined model forecast operator from the initial time to i. 4D-Var is thus a nonlinear constrained optimization problem which is very difficult to solve in the general case. Fortunately it can be greatly simplified with two hypotheses:

- **Causality**. The forecast model can be expressed as the product of intermediate forecast steps, which reflects the causality of nature. Usually it is the integration of a numerical prediction model starting with $\mathbf{x}$ as the initial condition, then by denoting M_i the forecast step from i to $i-1$, $\mathbf{x}_i$ is given by:

$$\mathbf{x}_i = M_i M_{i-1} ... M_i \mathbf{x}$$

- **Tangent linear hypothesis**. The cost function can be made quadratic by assuming, on top of H_i the linearization of M, that the operator can be linearized, i.e.

$$\mathbf{y}_i - H_i M_{0\to i}(\mathbf{x}) \approx \mathbf{y}_i - H_i M_{0\to i}(\mathbf{x}_b) - \mathbf{H}_i \mathbf{M}_{0\to i}(\mathbf{x} - \mathbf{x}_b)$$

where $\mathbf{M}$ is the **tangent linear model** as discussed in (2.2.7) i.e. the differential of M.

The two hypotheses above simplify the general minimization problem to an unconstrained quadratic one which is numerically much easier to solve. The first term J_b of the cost function is no more complicated than in 3D-Var, while the evaluation of the second term J_o would seem to require integrations of the forecast model from the analysis time to each of the observation times n, and even more for the computation of the gradient ∇J_o (Kalnay, 2003). It is given by

$$\nabla J_o = \left[\frac{\partial J_o}{\partial \mathbf{x}(t_0)}\right] = \sum_{i=0}^{N} \mathbf{M}_{i\to 0}^T \mathbf{H}_i^T \mathbf{R}_i^{-1} \left[H(\mathbf{x}_i - \mathbf{y}_i^o)\right] \tag{2.26}$$

where $\mathbf{x}_i = M_{(i-1)\to i}(\mathbf{x}_{i-1})$. The equation 2.26 shows that every iteration of the 4D-Var minimization of the gradient, i.e. computing the increments $[H(\mathbf{x}_i - \mathbf{y}_i^o)]$ at the observations times t_i during a forward integration, multiplying them by $\mathbf{H}_i^T \mathbf{R}_i^{-1}$ and integrated these weighted increments back to the initial time using the adjoint model $\mathbf{M}^T$ (Bouttier and Courtier, 1999; Kalnay, 2003).

For a more detailed description of the different data assimilation methods, it is suggested to refer to the mentioned bibliography in this Section.

2.3 Data Assimilation in Air quality models

Over the last decade, the great number of observations and advances in spatio-temporal data assimilation (DA) methods in atmospheric chemistry have allowed a more efficient use of measurements, thus possibly reducing model uncertainties. Nevertheless, the issue of determining an efficient technique for chemical DA remains challenging.

The Ensemble Kalman Filter is a promising method for chemical DA. Constantinescu et al. (2007b) assesses the performance of EnKF in an idealizing setting. In Constantinescu et al. (2007a) a new background error as a autoregressive

process is introduced as a key ingredient of data assimilation, obtaining a realistic estimation of background error distribution. In Constantinescu et al. (2007c) the performance of the EnKF method is compared with a state-of-the-art 4D-Var approach. Two Kalman filter algorithms, the reduced rank square root (RRSQRT) and the EnKF are implemented in Hanea et al. (2004, 2007). However, it is important to remark that, since air quality models are highly non-linear, EnKF implementation could be computationally expensive, as the number of ensemble members necessary to compensate the effects of non-linearity could be high (from 20 to 50 according to Hanea et al., 2007).

Four-dimensional variational techniques are often employed to study the sensitivity of chemical species in emissions and to improve their estimation at regional and continental scale, through the use of adjoint model (Mallet and Sportisse, 2005; Quélo et al., 2005). Menut et al. (2000) utilised the adjoint model of a simplified urban photochemical model to examine the sensitivity of ozone and nitrogen dioxide concentration to various parameter and in Menut (2003) a more comprehensive study, utilising adjoint modelling approach, to investigate sensitivity of O_3, O_x and NO_x concerning to reactions rates, emissions, boundary conditions, dry deposition, temperature and turbolent vertical diffusivity is presented. Whereas in Elbern and Schmidt (2001) a chemical 4D-Var system has been developed and applied for the study of an enhanced summery ozone level episode. Elbern et al. (2000) shows skill and limits of the 4D-Var technique applying to improving emissions rates of non-observed ozone precursors, when only ozone observations are are available. Further elaborations of this technique can be found in Elbern et al. (2007) and Chai et al. (2009). Despite the satisfactory results obtained, four-dimensional variational techniques require the formulation of the tangent linear model and its inverse, that are complex to build, given the intrinsic complexity of air quality models and require high computational costs and this

puts a limit to its comprehensive diffusion.

The optimal interpolation data assimilation technique is used by Jeuken et al. (1999) to combine total ozone columns with the three-dimensional tracer transport model TM3. Starting from a separable form of the forecast covariance matrix, the optimal interpolation equation could be rewritten into a horizontal and vertical analysis step. Ozone profiles from the assimilation appeared realistic and close to the ones observed. They were capable to describe dynamical features in the lower stratosphere. In Clerbaux et al. (2001) a sequential assimilation approach based on optimal interpolation is used to incorporate CO columnar dataset into a global three-dimensional chemistry-transport (MOZART). Surface CO mixing ratios computed after assimilation of total columns were compared with in situ measurements and a good agreement was found between the two data set. In Lamarque et al. (1999) observations of carbon monoxide columns were assimilated into a global three-dimensional chemistry transport model using the optimal interpolation technique. On the global scale the adjustment of the CO field resulting from the assimilation procedure was large at the beginning the assimilation suggesting discrepancies in the model initial conditions. Moreover the assimilation of CO significantly influenced the distribution of other chemical species, even over the limited time periods they analysed. Also in Riishojgaard et al. (2000) the optimal interpolation routine was chosen to an operative ozone observation assimilation, producing realistic total ozone fields, while a substantial bias with respect to observed profiles remains in the upper stratosphere.

Khattatov et al. (2000) presents a sequential data assimilation system that is based on suboptimal Kalman filter, that has the same formulation of the optimal interpolation, but with a parametrization of the temporal evolution of the $\mathbf{B}$ matrix. This sequential assimilation approach was developed in a general global chemistry transport model for assimilation of ozone observation, achieving

satisfactory results on global stratospheric ozone distribution. This technique was firstly presented and tested by Ménard and Chang (2000) for assimilating limb-sounding CH_4 observations of stratospheric chemical constituents into a tracer transport model. Same approach is successfully applied by Fierli et al. (2002) assimilating sequentially tracer measurements in isentropic chemistry-transport models (CTMs) of the stratosphere.

Wu et al. (2008) compared different algorithms for ozone forecasts (optimal interpolation, reduced-rank square root Kalman filter, ensemble Kalman filter, and 4D-variational assimilation), finding that the optimal interpolation algorithm (OI) and ensemble Kalman filter algorithm (EnKF) have the best performance, the former during assimilation periods, the latter during forecast.

So in this thesis work it was decided to implement an Optimal Interpolation routine, that has the advantage to be relatively easy to realize and to not imply high computational cost. Moreover, following the cited literature, in the contest of air quality modelling it is quite well performing with respect to more sophisticated data assimilation approach.

Chapter 3

BOLCHEM and CHIMERE and Optimal Interpolation

The first part of experimental work focused on implementing and testing an Optimal Interpolation routine in the air quality model BOLCHEM and the chemistry transport model CHIMERE. BOLCHEM is a model that consists of a mesoscale meterological model (BOLAM) coupled an on-line one way with a gas-phase chemistry processor. Whereas, CHIMERE does not explicitly calculates the meteorological part that is provided by the meteorological model BOLAM. Different assimilation experiments have been performed assimilating ground measurements and satellite tropospheric columns. In this chapter a brief description of the employed models, observations used for validation and assimilation and implemented Optimal Interpolation routine will be given.

3.1 BOLAM model

BOLAM (BOlogna Limited Area Model) is a meteorological model based on primitive equations in the hydrostatic approximation. It solves the prognostic equations for wind components u and v, potential temperature, specific humidity and surface pressure. Variables are defined on hybrid coordinates and are distributed on a non-uniformly spaced Lorenz grid. The horizontal discretization employs geographical coordinates, with latitudinal rotation on an Arakawa C-grid. The model implements a weighted average flux scheme for three-dimensional advection. The lateral boundary conditions are imposed by means of a relaxation scheme that minimizes wave energy reflection. The microphysical scheme has five prognostic variables (cloud water, cloud ice, rain, snow and graupel), as derived from the one proposed by Schultz (1995). Deep convection is parameterized with the scheme of Kain-Fritsch (Kain and Fritsch, 1990; Kain, 2004). The boundary layer scheme is based on the mixing length assumption and the explicit prediction of turbulent kinetic energy (Zampieri et al., 2005), while the surface turbulent fluxes are computed according to the Monin-Obukhov similarity theory. The parameterization of the effects of vegetation and soil processes (Pressman, 1994) is based on the water and energy balance in a four-layer soil model, and includes the diagnostic computation of skin temperature and humidity, seasonally dependent vegetation effects, evapo-transpiration and interception of precipitation. The radiation is computed with a combined application of the scheme from (Ritter and Geleyn, 1992) and the operational one from the ECMWF (Morcrette et al., 1998). Further details of the model are provided in Malguzzi et al. (2006) and Buzzi et al. (1994, 2003).

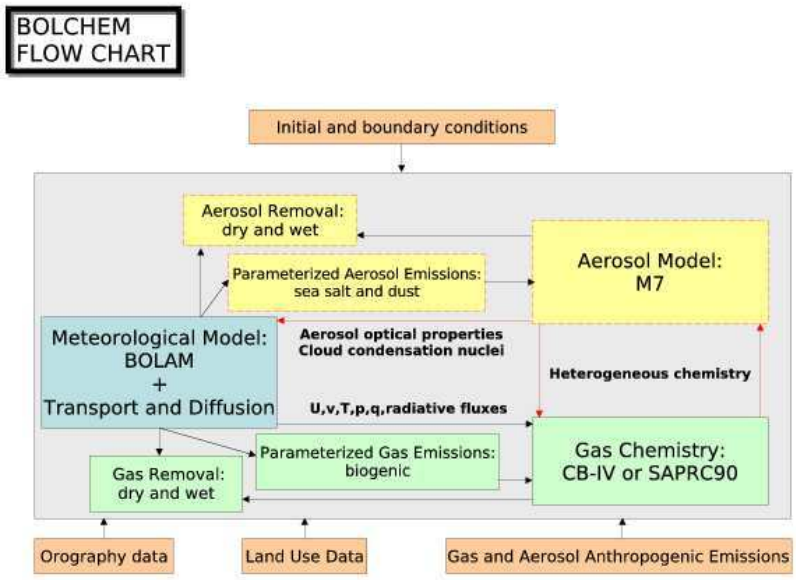

Figure 3.1: BOLCHEM flow chart (http://bolchem.isac.cnr.it/projects:bolchem.do).

3.2 BOLCHEM model

The air quality modelling system BOLCHEM is an on-line one-way coupling of the mesoscale meteorological model BOLAM with different types of gas-phase chemistry processor (see Figure 3.1). A detailed description of the model is presented in (Mircea et al., 2008). The model can be configured to run with two different gas-phase chemical mechanisms: SAPRC-90 (131 reactions with 35 chemical species (Carter, 1990)) and CB-IV (85 reactions and 30 chemical species (Gery et al., 1989)). BOLCHEM uses the same grid for meteorology and chemistry. The vertical coordinate system is terrain-following (σ), with variables distributed on a non-uniformly spaced staggered Lorenz grid, while the horizontal discretization uses geographical coordinates on an Arakawa C-grid. For each time step, the meteorological fields are computed and then used in the chemistry module in order to update the pollutant concentrations. The meteorological part is driven by BOLAM, which is based on hydrostatic primitive equations, with wind components u and v, potential temperature θ, specific humidity q and surface pressure P_s as dependent variables. The same 3-D WAF (Weighted Average Flux, Hubbard and Nikiforakis, 2003) mass conservative advection scheme is used both for meteorological quantities and pollutants. The dry deposition of gases is parametrized following a resistance analogy approach (Wesely, 1989), using the description of soil consistent with that of meteorology. The vertical diffusion scheme uses a parametrization dependent on turbulent kinetic energy.

Initial and lateral boundary conditions for meteorology were supplied by the European Centre for Medium-range Weather Forecasts (ECMWF) and interpolated over the BOLCHEM grid. Gas fields, as well as anthropogenic and biogenic emissions, can be provided by different databases (including climatology for 3-D fields).

3.3 CHIMERE model

CHIMERE is a three-dimensional (3-D) Eulerian chemistry-transport model. The CHIMERE multi-scale model is primarily designed to produce daily forecasts of ozone, aerosols and other pollutants and make long-term simulations for emission control scenarios. CHIMERE runs over a range of spatial scale from the regional scale (several thousand kilometers) (Schmidt and Martin, 2003; Schmidt et al., 2001) to the urban scale (100-200 Km) (Vautard et al., 2001; Beekmann and Derognat, 2003) with resolutions from 1-2 Km to 100 Km. CHIMERE offers the option to include different gas phase chemical mechanisms. The original, complete scheme called MELCHIOR1, describes more than 300 reactions of 80 gaseous species. The hydrocarbon degradation is fairly similar to the EMEP gas phase mechanism (Simpson, 1992). Adaptations are made in particular for low NO_x conditions and NO_x-nitrate chemistry. All rate constants are updated according to Atkinson et al. (1997) and DeMore et al. (1997). Heterogeneous formation of HONO from deposition of NO_2 on wet surfaces is now considered, using the formulation of Aumont et al. (2003). In order to reduce the computing time a reduced mechanism with less species and about 120 reactions is derived from MELCHIOR (Derognat et al., 2003), following the concept of "chemical operators" (Carter, 1990). This reduced mechanism is called MELCHIOR2.

The model formulation is based on the mass continuity equation for several species in every grid cell. The numerical method for the temporal solution of the stiff system of partial differential equation is adapted from the second-order TWOSTEP algorithm originally proposed by Verwer (1994). CHIMERE can run with several vertical resolutions, and with a wide range of complexity. It can run with several chemical mechanism configurations, simplified or more complete, with or without aerosols.

Vertical transport is assumed to balance horizontal mass divergence/ convergence. Horizontal turbulent fluxes are not considered. Vertical turbulent mixing takes place only in the boundary-layer and follows the parametrization of Troen and Mahrt (1986). Advection is performed either by a first upwind scheme, the Van Leer scheme or by the PPM (Piecewise Parabolic Method) 3d order scheme for slow species. Aerosol thermodynamic equilibrium is achieved using the ISORROPIA model.

In this work CHIMERE version V20050G1 over Europe is used. The meteorological fields are provided by the hydrostatic mesoscale model BOLAM (Bologna Limited Area Model) paragraph 3.1. The anthropogenic emissions are derived from the EMEP annual totals for 2002 (Vestreng et al., 2004) for NO_x, SO_2, CO, and non-methane volatile organic compounds. The boundaries concentrations for the continental set up are provided by MOZART CTM version 2 (Horowitz et al., 2003) with an horizontal resolution of about 0.5×0.5 and spatially interpolated from the EMEP grid onto the CHIMERE grid.

More details about CHIMERE model are given by (Schmidt et al., 2001; Vautard et al., 2003) and by the Laboratoire de Météorologie Dynamique on the web site http://euler.lmd.polytechnique.fr/chimere.

3.4 Ground based measurements

The first application (for details and results see chapter 4) of the implemented optimal interpolation routine consists in the assimilation of synthetic observations. To generate a realistic observations spatial distribution, we referred to a real observation network: the European Air quality dataBase (AirBase). Whereas in the second application (for details and results see chapter 5), pollutant concentrations of this database were utilised as *independent observations*, comparing them with the

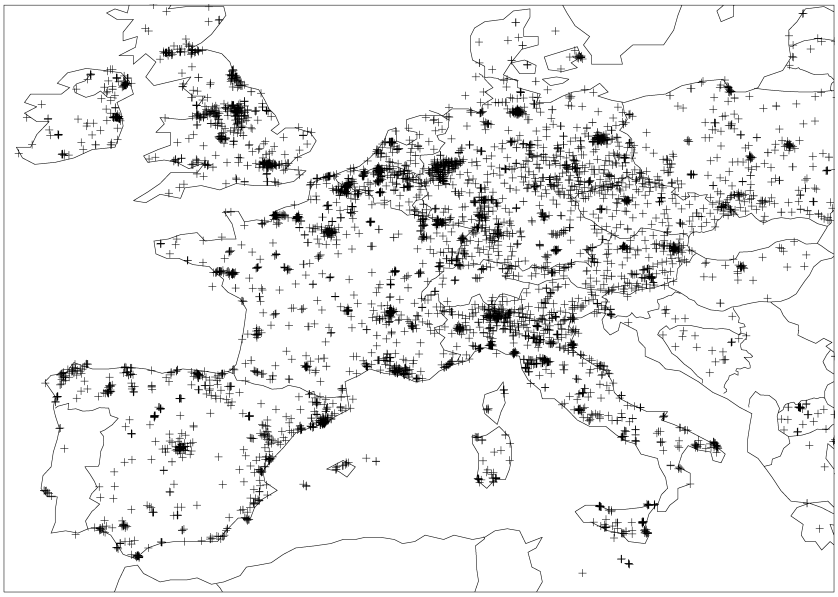

Figure 3.2: Air Base measurements network.

assimilated chemical species. AirBase is the public air quality database system of the European Environment Agency (EEA). It contains air quality monitoring data and information submitted by the participating countries throughout Europe. The air quality database consists of multi-annual time series of air quality measurement data and their statistics for a representative selection of stations and for a number of pollutants. It also contains meta-information on the involved monitoring networks, their stations and their measurements as type of station (traffic, industrial, background), type of area (urban, sub-urban, rural), characteristics of zone (residential, industrial, commercial, natural). The database covers geographically all countries from the European Union. In Figure 3.2 is depicted the measurement station that compose the Air Base network.

3.5 Satellite columns measurements

The trace gas sensor SCIAMACHY onboard European Space Agencys ENVISAT-1 satellite in sun-synchronous orbit provides a smooth continuation of measurements performed by its predecessor instrument, GOME, onboard the ERS-2 satellite. The enhanced capabilities of SCIAMACHY in measuring NO_2 include increased horizontal resolution (60 km across-track direction, 30 km along track direction) and a limb-nadir matching scan for obtaining the vertical structure of NO_2 as well as the NO_2 column amount. The satellite overpass time is about 10:00 local standard time (LST) and global coverage is achieved in 6 days at the equator (Boersma et al., 2004).

OMI is an instrument onboard NASAs Earth Observing System Aura satellite launched in July 2004. It is a nadir-viewing imaging spectrograph measuring direct and atmosphere-backscattered sunlight in the ultraviolet-visible (UV-VIS) range from 270 nm to 500 nm. Satellite has local overpass time at 13:30. The wide field of view of the nadir-pointing telescope (114°) gives OMI a swath width of 2600 km and provides daily global coverage with a high horizontal resolution. For the channel in which NO_2 is observed in global observation mode, the pixel size in the swath direction increases from 13 km x 24 km (along across track) at the exact nadir position to about 13 km $\times$ 150 km at the outermost swath angle (57°) (Boersma et al., 2006).

The NO_2 tropospheric column data sets utilized in this work are available on the TEMIS project web site (http://www.temis.nl). SCIAMACHY NO_2 tropospheric column is the result of a collaboration between the Belgian Institute for Space Aeronomy (BIRA-IASB) and the Royal Netherlands Meteorological Institute (KNMI) and the OMI NO_2 tropospheric column is a result of a collaboration between KNMI and NASA, in the framework of the ESA Data User Program

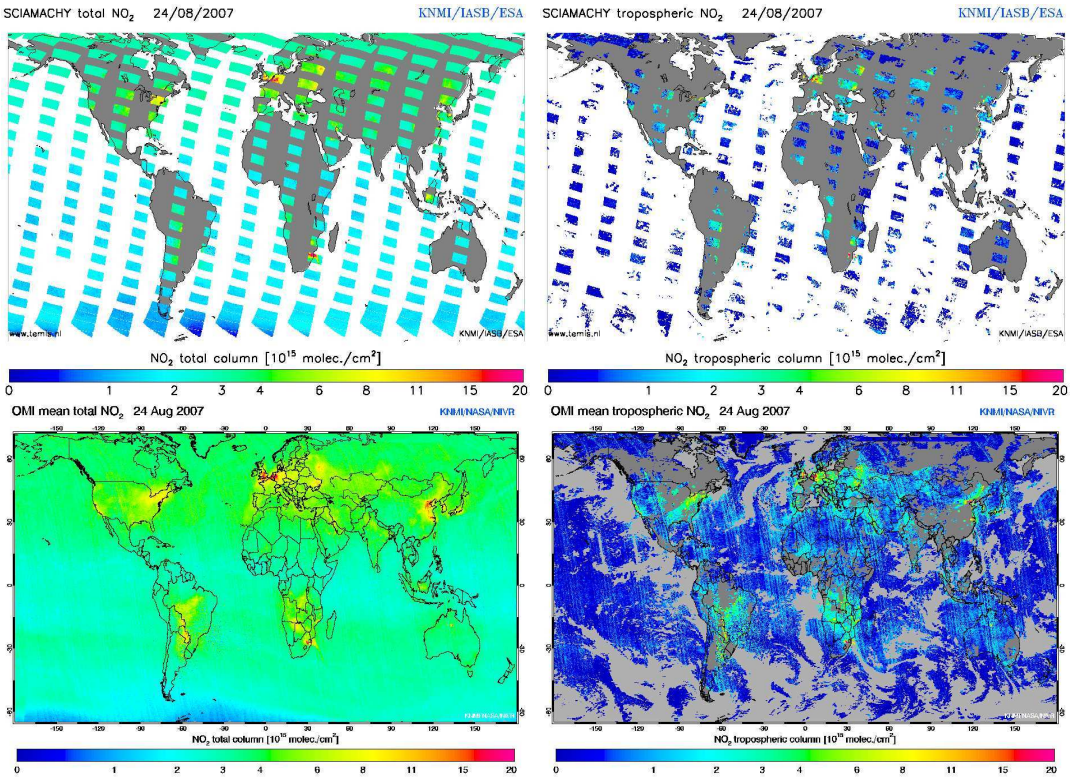

Figure 3.3: Total (upper left panel) and tropospheric (upper right panel) NO_2 column from SCHIAMACHY and total (lower left panel) and tropospheric (lower right panel) NO_2 column from OMI related to 24^{th} August 2007 (Ref.: http://www.temis.nl). NO_2 column are expressed in 10^{15}molecol/cm^2.

Tropospheric Emission Monitoring Internet Service (TEMIS) project. Description of the retrieval to obtain tropospheric columns method can be found in (Boersma et al., 2004) for SCIAMACHY and in (Boersma et al., 2006) for OMI.

In Figure 3.3 shows the total and tropospheric columns of NO_2 in 24^{th} August, it is evident the higher daily coverage of OMI with respect to SCIAMACHY. In the case taking into account, different polluted areas are detected by both satellites, as a vast area around Beijing, North-Central Europe, East and West cost of USA. Anyway finer resolution and better coverage of OMI makes it more suitable with respect to SCIAMACHY to provide daily picture of polluted area and to catch hot-spot area.

In the present work only the NO_2 tropospheric columns made with a scan angle below 55° and with a cloud fraction not exceeding 30% are utilised in the assimilation algorithm. First restriction was adopted in other to avoid pixel with a too large horizontal length with respect to other pixels and second restriction in order to avoid assimilation of columns with a too high overall error in vertical (Duncan et al., 2010; Boersma et al., 2006).

3.6 Optimal Interpolation algorithm

In the presented work, a sub-optimal Intepolation (OI) algorithm is implemented in both the BOLCHEM and CHIMERE models to assimilate surface observations and tropospheric columnar of gas constituents. The OI routine was chosen because it is relatively simple and computationally inexpensive. We refer to sections 2.2.5 and 2.2.4 and to (Talagrand, 1997; Kalnay, 2003) for a general formulation of assimilation problem.

The basic formulation of OI, with the hypothesis of unbiased errors for the background fields and uncorrelated observational errors, can be written:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T \left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1} (\mathbf{y}_o - \mathbf{y}_b) \tag{3.1}$$

where: $\mathbf{x}_a$ and $\mathbf{x}_b$ are the analysis and background (or first guess) vectors, respectively, dimensioned as the number of grid points I; $\mathbf{y}_o$ and $\mathbf{y}_b$ vectors, dimensioned as the number of observations K, are the observations and the background field interpolated over the observation sites, respectively; $\mathbf{B}$ is the covariance matrix of the background error dimensioned as $I \times I$ and $\mathbf{H}$ is the observation operator $(I \times K)$ and $\mathbf{R}$ is the $K \times K$ covariance matrix of observation errors. On the hypothesis of uncorrelated observation errors, $\mathbf{R}$ is diagonal. In the adopted procedure $\mathbf{H}$ and $\mathbf{B}$ matrices are not explicity calculated, but $\mathbf{G}$ ($\mathbf{B}\mathbf{H}^T$) and $\mathbf{S}$ ($\mathbf{H}\mathbf{B}\mathbf{H}^T$) matrices are computed. $\mathbf{G}$ is the $I \times K$ covariance matrix of background error between model grid points and observations locations; $\mathbf{S}$ is the $K \times K$ covariance matrix of the background error between observations points. A simplified implementation of covariance parametrization is adopted (Buzzi et al., 2003; Fierli et al., 2002; Khattatov et al., 2000) , choosing Gaussian shape functions factorised in horizontal and vertical components as follows:

$$f(i,j) = \sigma_b^2 \cdot e^{-\frac{1}{2}\left(\frac{d_h(i,j)}{L_h}\right)^2} \cdot e^{-\frac{1}{2}\left(\frac{z_i - z_j}{L_v}\right)^2} \tag{3.2}$$

where $d_h\,(i,j)$ and $z_i - z_j$ are the horizontal and vertical distances between points i and j and L_h and L_v are scale parameters defining the observation influence in the horizontal and vertical direction, respectively.

In the case of columnar observations $\mathbf{S}$ matrix does not depend on the vertical dimension while $\mathbf{G}$ matrix is divided in a part that weighs horizontally the covariance error with a correlation value of 30 km and that convolves vertically the analysis increments on the model grid proportionally to the model ozone profile. No temporal background error evolution is adopted because the NO_2 has

an high temporal variation, this approximation is particularly valid for assimilation of satellite data, as it is performed as a maximum once a day at the same location.

In the case of satellite observations, to obtain columnar values from model profile is often applied the Averaging Kernel (AK) (Eskes and Boersma, 2003; Boersma et al., 2004). Anyway, as underlined Blond et al. (2007) for SCIAMACHY data and Huijnen et al. (2010) for OMI data, the correction for NO_2 is in general very low, as NO_2 is trapped in low tropospheric layers and AK is not so varying in these layers. It was so decided to not apply AK in this work.

A further simplification is introduced by defining ϵ as the ratio between the observation error covariance and the background error covariance (σ_o^2/σ_b^2), assumed constant for every model assimilated variable. Thus, dividing all the covariance matrices by σ_b^2, the K diagonal elements of $\mathbf{R}$ are equal to ϵ, which becomes the single tuning parameter used to give more or less weight to the observations (Kalnay, 2003). Here, the values of L_v, L_h and ϵ are those chosen by (Buzzi et al., 2003) (for DA application in CTMs), to assimilate spatially non-uniform observations in the BOLAM model, with a horizontal resolution comparable to the one used here.

Chapter 4

Assimilated observations for improving tropospheric ozone

In this chapter is presented a first application of the data assimilation system aiming at evaluating the improvement in the capability of regional models to reproduce the distribution of tropospheric pollutants; for this we used surface chemical observations. Among the many causes of uncertainties of chemistry models simulations, a particular focus is given by uncertainties in emissions, that are known to be high, despite continuous improvement in compilation of emissions inventories. Thus we focused on evaluation of the efficacy of DA in correcting the error in ozone due to biases in emission scenarios. The study presented in this Chapter was carried out using the sequential Optimal Interpolation (OI) routine implemented in the air quality model BOLCHEM to perform ozone and nitrogen dioxide assimilation. This part of the work was performed using the Observing System Simulation Experiment (OSSE), which allows the quantification of assimilation impact, through comparison with a reference state. OSSE is typically designed to use data assimilation ideas to investigate the potential

impacts of prospective observing systems. In an OSSE, simulated rather than real observations are the input to a data assimilation system. This method was, for instance, successfully employed for satellite data assimilation to evaluate the expected performance of proposed observing strategies on improving CO in the lowermost troposphere (Edwards et al., 2009). Different sensitivity tests were carried out in order to identify how assimilation can correct perturbations on O_3, induced by NO_x emissions biased in both flux intensity and time. Tests were performed assimilating different species (only O_3, only NO_2 and both O_3 and NO_2), and varying assimilation window length (12 hours or 24 hours) and starting hour of assimilation (12 AM, 12 PM, 06 AM). Emissions were biased quantitatively up to $\pm$ 50% and shifted temporally up to $\pm$ 2hours. This part of the work dealt with assimilation of ground based observations with a realistic distribution derived from AirBase database network and was focused on Po Valley. However it can be easily extended in other polluted areas, characterized by a reasonable number of observational sites (order of 20-30 on 40000km area) as for instance European MEGACities and hot-spot areas.

4.1 Numerical experiments

4.1.1 Experimental set-up and case study description

Simulations were performed in a domain ranging between 38°N-48°N and 7°E-18°E, with an horizontal resolution of 0.2° $\times$ 0.2° and with 33 sigma-hybrid levels, the first level being about 40 meters above the ground. The chemical processes were calculated in the 16 lowermost model levels up to the mid-troposphere. The initial and boundary conditions for the meteorological variables were supplied by ECMWF, and updated every 6 hours with an horizontal resolution of 0.5° $\times$ 0.5°

and 60 sigma-hybrid levels. The emission fields for Italy were obtained from the national emission inventory of ARIA-NET S.r.l. in the frame of the MINNI project (National Integrated Modelling system for International Negotiation, Zanini et al., 2004), while for the non-Italian domain portions, they came from the EMEP inventory. They are both estimated from 1999 year inventories and they were regridded into BOLCHEM grid. The 3-hourly chemical boundary and initial conditions were also provided by EMEP model. The chemical scheme here used was SAPRC-90 (Mircea et al., 2008).

Simulations were performed in the Po Valley. The study focused on this region because is a highly urbanized and intensely industrialized area, which is one of the most polluted regions in Europe, and is often subjected to strong photochemical pollution episodes (Gabusi and Volta, 2005 and references herein), favored by frequent stagnant meteorological conditions associated with high insolation during summer (Dosio et al., 2002). The Po Valley can be primarily considered a NO_x sensitive region. Beekmann and Vautard (2009) found a mainly NO_x sensitive regime over the Mediterranean basin and Eastern Europe, with the exception of several large agglomerations (e.g. Barcelona, Milan). Moreover, the VOC sensitive chemical regime for European agglomerations is more pronounced in models with an urban scale resolution (Thunis et al., 2007).

It was chosen a four-day photochemical pollution event, lasting from July 20 to 24, 2004. This period, in fact, was characterized by high values of O_3 concentration at the ground throughout the Po Valley area. The synoptic situation was favorable to high photo-oxidant production, since high irradiance and temperature conditions were governed by a stationary high-pressure ridge at 500 hPa extending from North Africa to Southern Scandinavia (not reported here). The two upper panels in Figure 4.1 depict the ground level temperature (upper panel) and O_3 concentrations (middle panel), as simulated by the BOLCHEM model for July 22 at 15 UTC.

Simulated temperatures in the Po Valley between 12-18 UTC exceeded 30℃ on July 20 and 21, reaching 34℃ on July 22 and 23, and high temperatures were also estimated in large cities in Northern Italy. In Figure 4.1, middle panel, the ozone plume produced by photochemical pollution in the Po Valley is clearly visible. The ozone value significantly exceeds the information threshold (180 μg m^{-3}) and reaches the alert threshold (240 μg m^{-3}) in a few localized areas as Genoa, La Spezia, Trieste, Venice (European Commission directive 2002/3/EC). Figure 4.1, lower panel shows the daily cumulated NO_x emissions. The highest NO_x emissions are produced in the Milan and Genoa areas, this pattern remaining stable throughout the simulation.

4.1.2 Scenarios of biased NO_x emissions

In order to investigate the ability of DA to reduce the effect of an NO_x emissions bias on O_3 concentration, a set of Observing System Simulation Experiments (OSSE) (Arnold Jr and Dey, 1986 and Edwards et al., 2009, Constantinescu et al., 2007a for an application to tropospheric chemistry) were performed in a perfect model approach, assuming a simulation with unperturbed emissions as the reference atmosphere. Synthetic perfect observations of NO_2 and O_3 were extracted from the lowermost model level of the reference run in 28 spatial locations of AirBase observations (http://air-climate.eionet.europa.eu/databases/airbase/ airbasexml/) in the Po Valley. The selection was performed choosing the closest model grid point to each observational point and additional observations lying in the same grid cell were eliminated. The resulting spatial distribution of the synthetic observations is a not uniform distribution with a smaller number of observations (28) with respect to those available in AirBase database (48). The synthetic observational coverage is reported in the lower panel of Figure 4.1 (black crosses). The assimilation was

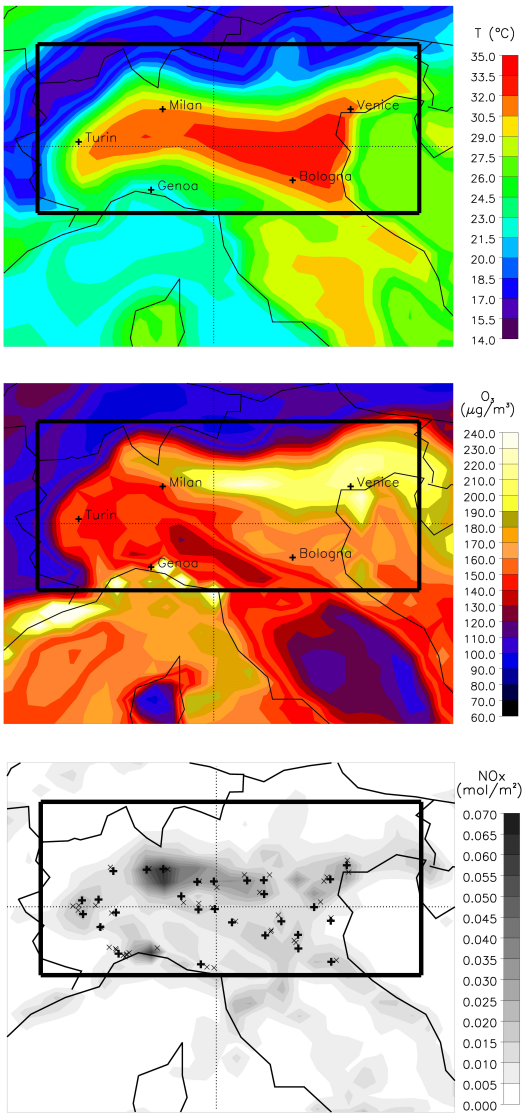

Figure 4.1: Upper and middle panels depict the ground level temperature (°C) and O_3 concentration (μg m^{-3}) simulated by the model, for July 22, 2004 at 15 UTC, respectively. The lower panel depicts the daily cumulative NO_x emissions (mol m^{-2}); the black crosses represent the locations of assimilated synthetic observations, while the thin "x" represent the real observation network used for setting up the spatial distribution of synthetic observations.

performed on the second day of the simulation (21 July), and was carried out every hour.

The OSSE approach has been chosen in order to separate and better identify the emissions bias effects and the role of assimilation in correction of the induced bias in O_3 field. Assimilation of actual observations instead of synthetic observations, should imply to assimilate also the reference simulation, as it could not be considered perfect any more. But this operation would shift chemical balance in reference case and it could be more difficult to distinguish the specific effect of assimilation in correcting biases in emissions and to study this issue at a methodological point of view. This increases the complexity of the problem, as it introduces the issue of the possible bias between the model chemical mechanism description and the "real" chemical mechanism description.

The characteristics of the assimilation experiments are summarized in Table 4.1. The first ensemble of experiments were defined (i) to assimilate O_3, NO_2 and both NO_2 and O_3, and (ii) to use 4 temporal assimilation windows: an entire day (DA_0024_O_3, DA_0024_NO_2 and DA_0024_NO_2_O_3 for assimilation of both species), the first 12 hours of the day (DA_0012_O_3, DA_0012_NO_2 and DA_0012_NO_2_O_3), the last 12 hours (DA_1224_O_3, DA_1224_NO_2 and DA_1224_NO_2_O_3), and during exposure to sunlight from 6 to 18 UTC (DA_0618_O_3, DA_0618_NO_2 and DA_0618_NO_2_O_3). Each assimilation experiment was composed of a set of 4 simulations with NO_x emissions biased by 25%, 50%, -25%, -50%, following the uncertainty estimates given in the introduction. The second ensemble of experiments was performed to evaluate the DA impact on the correction of the error caused by the temporal disaggregation of emissions. For this purpose, temporal shifts of ± 1 hour and ± 2 hour were applied to the emissions scenarios, performing 20 additional simulations based on DA_0024_NO_2, DA_0024_NO_2_O_3, DA_0618_NO_2 and DA_0618_NO_2_O_3

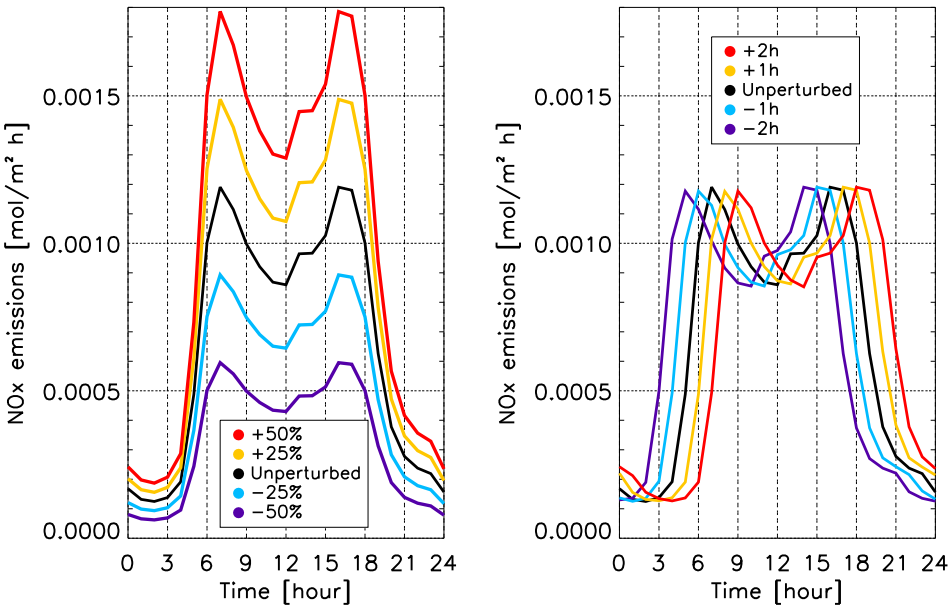

Figure 4.2: NO_x emissions expressed in mol m^{-2} hour $^{-1}$ and averaged in the area selected for analysis: the left panel shows the flux biased NO_x emissions scenarios, and the right panel shows the temporally biased ones.

configurations.

The evaluations of DA impact were performed over the Po Valley area, shown in Figure 4.1, identified as a 3-D box, of latitude 44.0°N-46.5°N, longitude 7.0°E-13.5°E, altitude below 600 m asl and model cells with less than 50% covered by sea. The chemical species selected for estimations were averaged over the model grid points where assimilation was not performed.

The different NO_x emission scenarios are shown in Figure 4.2, NO_x shows two maxima corresponding to traffic emissions in the morning and afternoon, the fluxes being equal for all days of the simulation.

Experiment name	DA window	Species assimilated	Perc. biases NO_x emiss.	Temp. biases NO_x emiss.
Reference	None	None	No	No
DA_NOASSIM	None	None	Yes	Yes
DA_0024_O_3	00 - 24 UTC	O_3	Yes	No
DA_0024_NO_2	00 - 24 UTC	NO_2	Yes	Yes
DA_0024_NO_2_O_3	00 - 24 UTC	O_3 NO_2	Yes	Yes
DA_0012_O_3	00 - 12 UTC	O_3	Yes	No
DA_0012_NO_2	00 - 12 UTC	NO_2	Yes	No
DA_0012_NO_2_O_3	00 - 12 UTC	O_3 NO_2	Yes	No
DA_1224_O_3	12 - 24 UTC	O_3	Yes	No
DA_1224_NO_2	12 - 24 UTC	NO_2	Yes	No
DA_1224_NO_2_O_3	12 - 24 UTC	O_3 NO_2	Yes	No
DA_0618_O_3	06 - 18 UTC	O_3	Yes	No
DA_0618_NO_2	06 - 18 UTC	NO_2	Yes	Yes
DA_0618_NO_2_O_3	06 - 18 UTC	O_3 NO_2	Yes	Yes

Table 4.1: Experiment reference names used in text, time window when DA is carried out, species assimilated, and indication of presence of percentage or temporal biases in NO_x emissions. Each experiment is composed of a set of 4 runs with $\pm$50%, $\pm$25% biased and/or $\pm$1hour and $\pm$2hours biased NO_x emissions.

4.2 Assimilation with variable NO_x emissions intensity

4.2.1 Sensitivity of model

The NO_2 concentration modeled at ground level was chosen as an indicator of the model's response to the NO_x emission perturbation. The model has a characteristic time to adjust the concentration of the species perturbed by the biased emissions. Figure 4.3 (panel a) depicts the relative difference of NO_2 between perturbed simulations and the reference, clearly showing that the model has a spin-up time of about 8 hours to adjust NO_2 concentrations to a stable perturbation. Therefore, the first 24 hours of the simulations were not taken in account, performing DA when emission perturbations led to a stable dependence with respect to the reference run.

Figure 4.3 (panel a) also shows that NO_2 concentrations have almost the same entity as the emission biases, whereas, in +50% NO_x emission biases, perturbation on NO_2 can reach 80%. This discrepancy can be explained by the fact that in the high NO_x state (when the NO_x emission rate is high Kleinman, 1994) insufficient concentrations of free radicals are produced to remove NO_x, which therefore accumulates more efficiently. The different emissions scenarios led to NO_x concentrations occurring in a range comparable with the transition between low NO_x and high NO_x state (Sillman, 1995; Kleinman, 1991; Kleinman et al., 1997; Lu and Chang, 1998).

Figure 4.3 (panel b) reports O_3 ground concentrations, showing that during the photochemically active period, NO_x emission biases significantly perturb ground-level O_3 concentrations with respect to the reference by up to 25 μg m^{-3}. The perturbation is smaller for simulations with positive biases in emissions. Such behaviour is attributed to the non-monotonic dependence of O_3 in a high NO_x

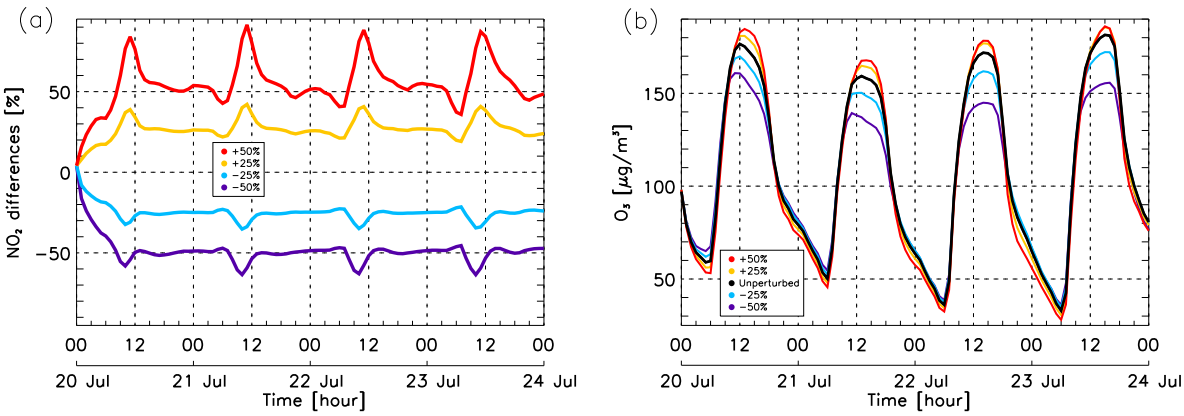

Figure 4.3: Panel (a) represents the relative difference of NO_2 at the model ground level, with respect to the reference for +50%, +25%, -25%, -50% NO_x emission biased simulations. Panel (b) shows the ground O_3 for the un-biased and biased simulations. Variables are averaged over the Po Valley area defined in Section 4.1.2.

concentration environment, with reduced O_3 production related to an increase in NO_x emissions (Logan et al., 1981; Sillman et al., 1990, 2003; Sillman and West, 2009; Poppe et al., 1993; Lin et al., 1988). In the case of low NO_x concentration environments, O_3 growth or decrease is directly proportional to NO_x emission variations (Wayne, 2000).

In the night-time regime the O_3 concentration is controlled by NO titration, ($NO + O_3 \rightarrow NO_2 + O_2$), leading to a reduction in O_3 with increasing NO (panel b, Figure 4.3) (Mircea et al., 2008; Hobbs, 2000; Wayne, 2000).

4.2.2 Model response to NO_2 and O_3 assimilation

Figure 4.4 reports the time series of ground O_3 concentrations and the relative differences for assimilated and non-assimilated simulations with respect to the reference for the DA_0024_NO_2 and DA_0024_O_3_NO_2 experiments (description in

Table 4.1).

The left column in Figure 4.4 shows that during the day positive biases in emissions induce differences in O_3 maxima with respect to reference between 4% and 5%, while negative biases induce differences between -15% and -5% (dashed lines). The NO_2 assimilation has a positive impact, bringing O_3 maxima close to the reference (within a maximum of 4%) (solid lines). This effect persists 36-40 hour after the assimilation ends on July 22.

During the night the effect of NO_2 assimilation is reversed, increasing the perturbation up to 10%-25%, bearing in mind that nocturnal O_3 corrections are around 20 μg m^{-3}. In the case of positive NO_x emission biases, during the photochemically active period, NO_2 assimilation acts to reduce the O_3 concentration. Therefore, with the activation of nocturnal chemistry, the ozone concentration is lower than in the DA_NOASSIM case. At night, the O_3 concentration further decreases with respect to the DA_NOASSIM case due to the effect of NO titration.

The same considerations hold in the case of the simulations relating to negative NO_x emission perturbations.

The right column in Figure 4.4 shows that, during daylight, the assimilation of both O_3 and NO_2 leads to similar results to those of the NO_2 assimilation. Regarding the DA_0024_O_3_NO_2 experiment (4.2, right column), significant differences with respect to the DA_0024_NO_2 one are only visible during night-time assimilation, when ozone concentrations are corrected close to the reference. However, a few hours after the end of assimilation, the ozone behaviour becomes similar to the case of NO_2 assimilation. This result is also supported by an experiment performed assimilating only O_3 Figure 4.5, highlighting the short persistence of the positive effects of O_3 assimilation (around 2-3 hours).

The persistence of the positive impact of NO_2 assimilation also beyond the

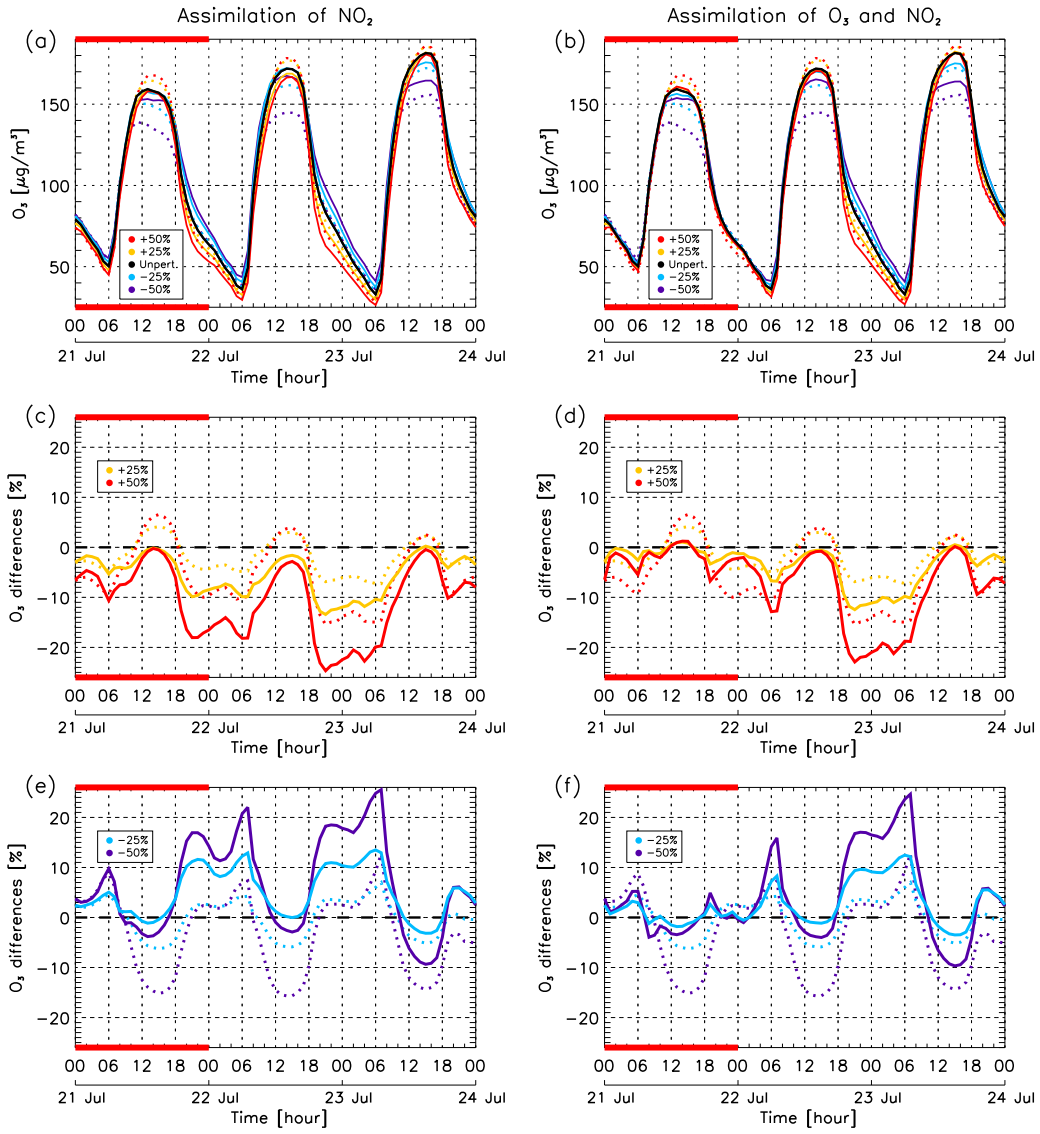

Figure 4.4: Time-series from second day of simulation of ground O_3 concentration for unbiased (black line) and biased (colored lines) NO_x emission simulations. Dashed lines represent the non-assimilated simulations. Solid lines in (a) and (b) represent the O_3 concentration obtained with NO_2 and with NO_2 and O_3 assimilations, respectively. Panels (c) and (e) represent the O_3 relative difference between DA_0024_NO_2 and reference for positive and negative NO_x emission perturbations, respectively. Panels (d) and (f) are the same as (c) and (e) but for DA_0024_NO_2_O_3.

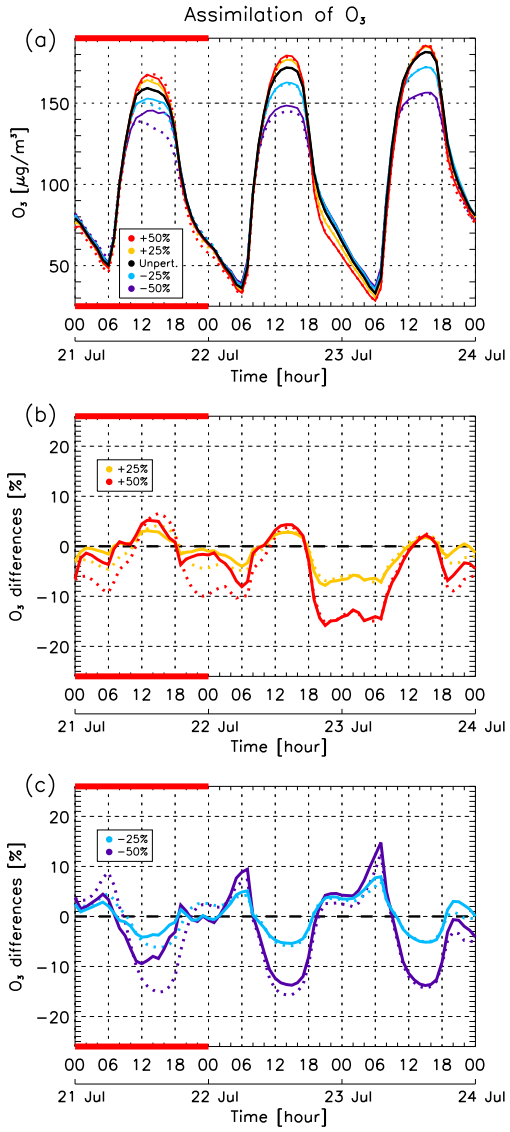

Figure 4.5: The same of Figure 4.4, but relative to DA_0024_O_3 experiment.

day of assimilation can be ascribed to the modification of the N_2O_5 and NO_3 night-time reservoirs. Figure 4.6 shows N_2O_5 and NO_3 concentrations at the model ground level for the DA_0024_NO_2 experiment. The NO_2 assimilation has a considerable impact on N_2O_5 and on NO_3 during the night of 21-22 July, increasing (or decreasing) them in the case of negative (or positive) NO_x emission biases. This impact on N_2O_5 and NO_3 persists to a lesser extent during the night of July 22-23. An increase (decrease) in N_2O_5 and NO_3 radicals, rapidly photolysed at the activation of photochemistry, can induce a larger (lesser) increase in NO_2 and NO concentrations. Consequently, assimilation can significantly modify NO_2 and ozone up to the photochemically active period of the following day.

Nevertheless, it should be remarked that the correction induced by NO_2 assimilation in N_2O_5 and NO_3 concentrations seems to be too large with respect to the reference run concentration. This behaviour is linked to a general problem of air quality models and assimilation procedure itself. At the state of the art, the chemistry mechanisms used in air quality models are designed and tested to represent correctly some major species, but they are not always guaranteed for the other ones, as they necessarily neglect some secondary species and some chemical reactions. Data assimilation technique that has been used, do not change the equations that compose the chemical mechanism, but it introduces a forcing that corrects concentrations of few species, shifting the related chemical balances. Therefore, it is expected that, for the principal species the model deals with, the chemical balance shifts have the same direction and similar entity of those in the real world, whereas it is not straightforward that "right" behaviour of minor species is guaranteed. If assimilation could be done for all species that interact each other probably a more a coherent correction could be assured, but in general this is not possible as the number of the species that are measured systematically and extensively are limited. This places restrictions to assimilation procedure.

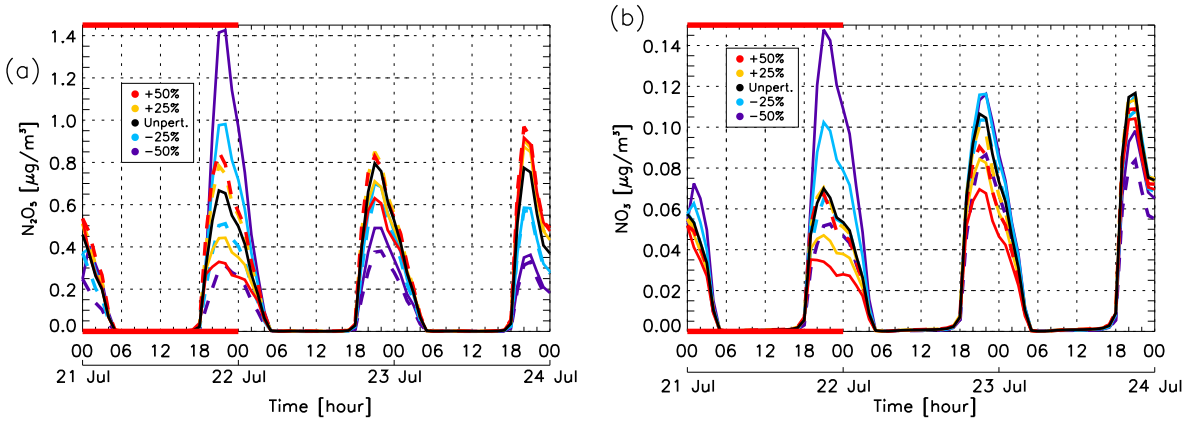

Figure 4.6: Time series of ground N_2O_5 (a) and NO_3 (b) concentrations for unbiased (black line) and biased (colored lines) NO_x emission simulations. Dashed lines represent the non-assimilated simulations and solid lines represent the simulations with NO_2 assimilation.

In order to enforce this consideration, in Figures 4.7 and 4.8 the all nitrogen species involved in the gas-photochemistry production of NO_2 and O_3 (expect N_2O_5 and NO_3 already presented in Figure 4.6) and calculated in the model chemical mechanism (Carter, 1990) are depicted. NO_y is obtained following the definition of given by Miyazaki et al. (2005) and computing it using the species explicitly calculated in the model chemistry mechanism (NO_2, NO, N_2O_5, NO_3, HO_2NO_2, HONO, RNO_3 and HNO_3). The principal species like NO, NO_2 and PAN showed a "right" behaviour, like so NO_y since its behaviour is dominated by NO_2, while like HO_2NO_2 and HNO_3 seems to have a response to biases in emissions and to assimilation more similar to those detected for N_2O_5 and NO_3. Whereas some other minor species, like HONO and RNO_3, does not seem to be strongly perturbed.

A second result of the analysis is that it is more effective to assimilate O_3 precursors (like NO_2) than O_3 itself in order to improve O_3 concentrations.

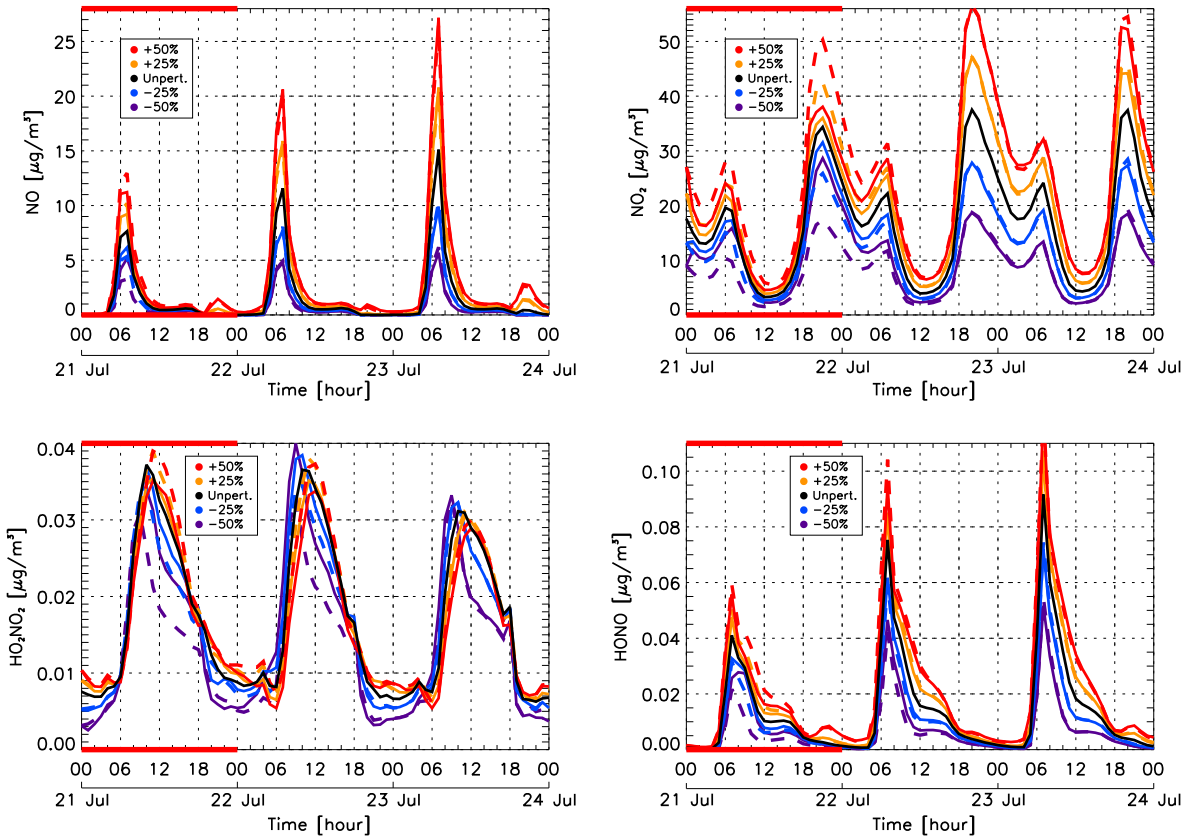

Figure 4.7: The same of Figure 4.6, but related to NO (uppper left panel), NO_2 (uppper right panel), HO_2NO_2 (lower left panel), HONO (lower right panel), respectively .

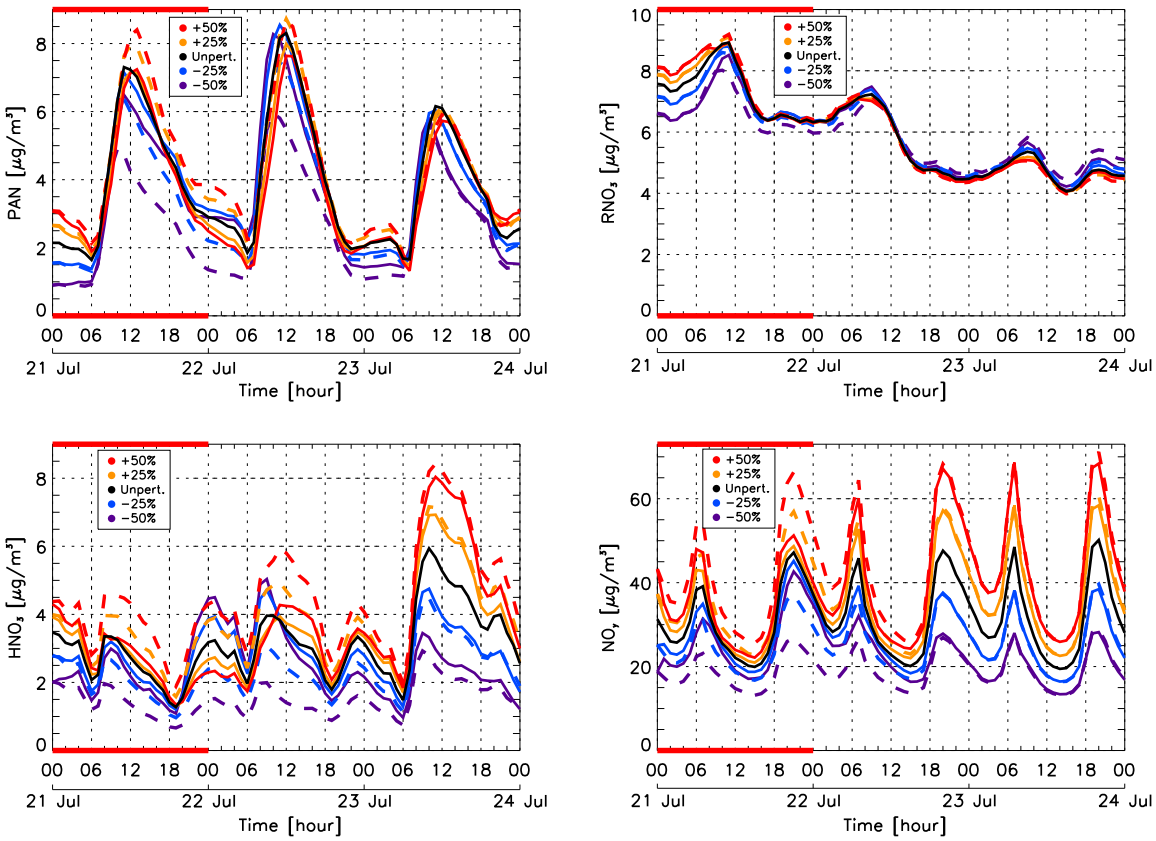

Figure 4.8: The same of Figure 4.6, but related to PAN (uppper left panel), RNO_3 (uppper right panel), HNO_3 (lower left panel), NO_y (lower right panel), respectively

.

4.2.3 Model response to different assimilation windows

The assimilation impact and timing may depend on the length of the assimilation window, in particular for quantities that can be highly variable in time, such as photolysis-driven chemical species. The analysis was restricted to NO_2 assimilation in view of the discussion in the previous section, and the assimilation window was varied in the four different ways described in Table 4.1. Figure 4.9 reports the absolute difference of ozone with respect to the reference for the DA_0024_NO_2, DA_1224_NO_2, DA_0012_NO_2 and DA_0618_NO_2 simulations.

Through the comparison between DA_0012_NO_2 and DA_1224_NO_2 simulations (green and light blue line, respectively), it is observed that it is more effective to assimilate in the first part of the day. In fact, for the DA_1224_NO_2 simulations the correction of the maximum O_3 value is rather poor during the assimilation window. In the case of -50% biased emissions (Figure 4.9, panel b) the correction is about 5 μg m^{-3}, whereas in the other experiments it is about 20 μg m^{-3}. This discrepancy can be ascribed to the fact that, in the DA_1224_NO_2 experiment, assimilation was activated when the photochemical processes had already started, so O_3 was already significantly more perturbed than in the other cases, and assimilation was not able to correct it efficiently.

DA_0618_NO_2 (blue lines), which covers the photochemically active period from 06 to 18 UTC, shows improvements in O_3 both during (up to 20 μ g m^{-3}) and after assimilation (up to 15 μg m^{-3}).

All experiments to some extent show a nocturnal perturbation of O_3 (a maximum of about 20 μg m^{-3} in DA_0024_NO_2), which could be explained by a change in titration equilibrium due to NO_2 and O_3 diurnal modification. It is observed that during night-time DA_0618_NO_2 and DA_1224_NO_2 are the closest to the reference.

DA_0024_NO_2 simulations (red lines) show that ozone maxima are slightly less well corrected in cases of positive biased NO_x emissions, with respect to the other assimilation windows, while ozone maxima are corrected better in the case of negative biased NO_x emissions. During the night, the increase in ozone perturbation is higher than in the other cases.

In DA_0024_NO_2 and DA_0012_NO_2, on the night of 21 - 22 July, O_3 exhibits similar perturbations, while at the activation of photochemistry (7 UTC 22 July), O_3 in DA_0012_NO_2 is closer to O_3 in DA_NOASSIM, in the case of both positive and negative biased emissions. Such behaviour is linked to the different perturbations of NO_2 concentration in these experiments. In fact, during the night of 21-22 July, in the DA_0024_NO_2 experiment, NO_2 is closer to reference due to the effect of assimilation, while in DA_0012_NO_2 NO_2 is closer to DA_NOASSIM, hence in the latter experiment, at the activation of phochemistry, O_3 is forced to be closer to DA_NOASSIM.

With regard to perturbation reduction in the ozone simulations, there are no marked discrepancies between the impacts of the 24h-assimilation and those of the 12h-assimilation. The assimilation tends to have a greater impact towards the start and finish, and is related to the first day after assimilation ends. All different assimilation windows tend to be quite similar on the last simulation day. In terms of computational cost reduction, the 12h-assimilation window is to be preferred.

Overall, while no major differences exist among the various assimilation windows, it can be concluded that DA_0618_NO_2 yields better results.

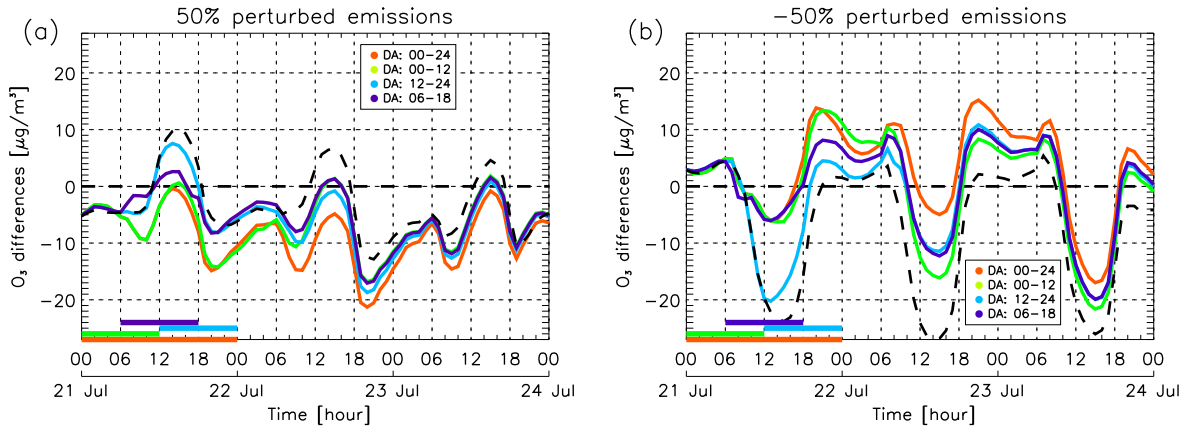

Figure 4.9: Time-series of the absolute differences with respect to reference of DA_0024_O_3_NO_2, DA_0012_O_3_NO_2, DA_1224_O_3_NO_2, DA_0618_O_3_NO_2 and DA_NOASSIM. The runs with 50%, -50% biased NO_x emissions are presented respectively in panels (a) and (b).

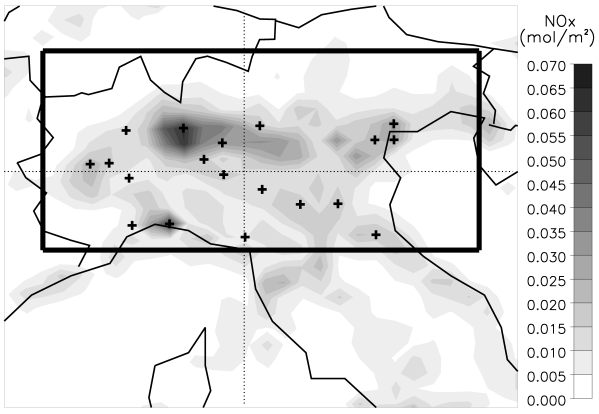

Figure 4.10: Daily cumulative NO_x emissions (mol m^{-2}); the black crosses represent the locations of assimilated synthetic observations related to the reduced number of monitoring stations (19).

4.2.4 Model response to reduction of assimilated observations

The actual number of assimilated points are 28 on an area of about 40000 km^2 (more than 100 model grid points). This is a reasonable distribution of measurements network representative of a number of observations sites in MEGAcities and/or hot-spot areas (as Po Valley, Ruhr, England). However, in order to further strengthen the analysis with respect to the number of assimilated observations, a test reducing the number of monitoring stations to 19 (see Figure 4.10) was performed. To do so the DA_0618_NO_2 experiments was chosen. It was not found a significative differences with respect to experiment with the not reducing number of monitoring stations (28), (see Figure 4.11 in comparison to Figure 4.12). The obtained result implies that the utilised technique, even if it was focused on Po Valley, can be easily extended in other polluted areas, characterized by a reasonable number of observational sites (order of 20-30 on 40000 km area), as for instance European MegaCities and hot-spot areas.

4.3 Assimilation with NO_x emission temporal biases

The temporal distribution of emission may play a role in ozone photochemistry, since a possible shift in peaks due to road traffic can modify the amount of NO_x during the photochemically active period. To analyse this effect, NO_x emissions are temporally shifted by +2, +1, -1, -2 hours. Assimilation of NO_2 is presented for experimentDA_0618_NO_2 only, in view of the discussion in the previous section.

Figure 4.13 reports the time series of ground O_3 concentration, showing that temporally biased emissions have an insignificant impact on ozone maximum values.

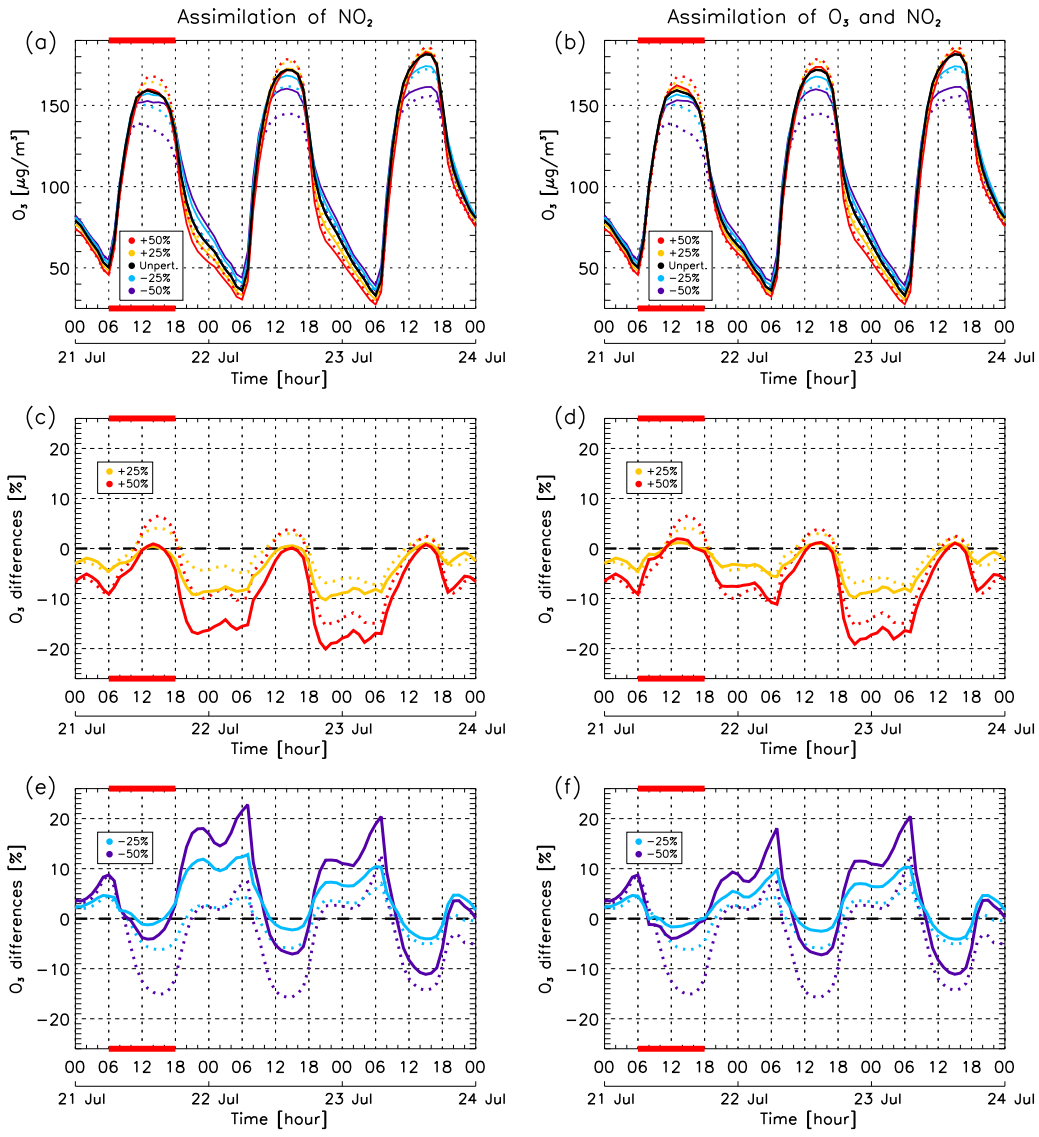

Figure 4.11: The same of Figure 4.4 but related to DA_0618_NO_2 experiment and with the reduced number of monitoring stations (19).

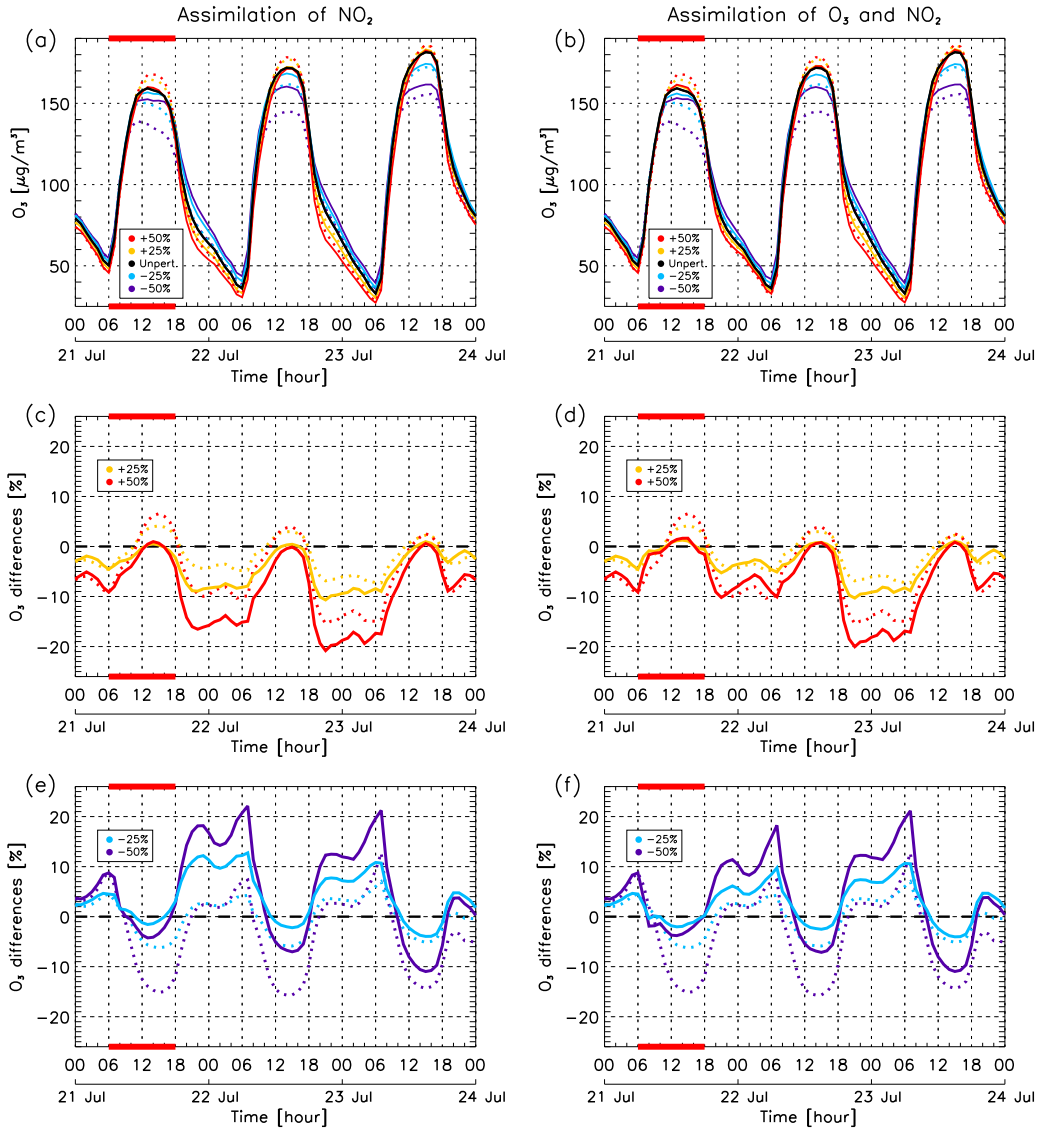

Figure 4.12: The same of Figure 4.11, but related to simulations with the not-reduced number of monitoring stations.

More significant effects are visible during night close to the transition periods between photochemical and nocturnal regime. It is observed that anticipated emissions induce a negative perturbation of O_3 concentration before the beginning of photochemistry and a positive perturbation after the beginning of the nocturnal regime. For postponed emissions (+1, +2 hours) the response of O_3 concentration is reversed. Such behaviour is strictly linked to the NO_x emission profile (Figure 4.2, right panel). Emission changes during the day are less than 30%, while a drastic transition between night and day regime can be observed, with a change of a factor of ten in a few hours. Consequently, in the central hours of the photochemically active period, biases in emission times do not significantly perturb O_3 maxima values. By contrast, a shift of $\pm$ 2 hours drastically changes NO_x emissions during night-time. A negative shift yields higher emissions at 5 UTC, while a positive shift provides higher emissions at 18 UTC, thus modifying the nocturnal chemistry. After the end of the photochemically active regime (18 UTC), the postponed NO_x emissions remain high, inducing negative O_3 perturbation (Figure 4.13) due to NO titration. Conversely, anticipated NO_x emissions are lower with respect to the reference case, leading to an accumulation of O_3. Similar considerations apply to anticipated emissions before the end of the nocturnal regime (5 UTC).

As the perturbations of O_3 maxima induced by biased emissions are very low, NO_2 assimilation does not improve them significantly. After the end of the assimilation window, the impact of assimilation is negligible.

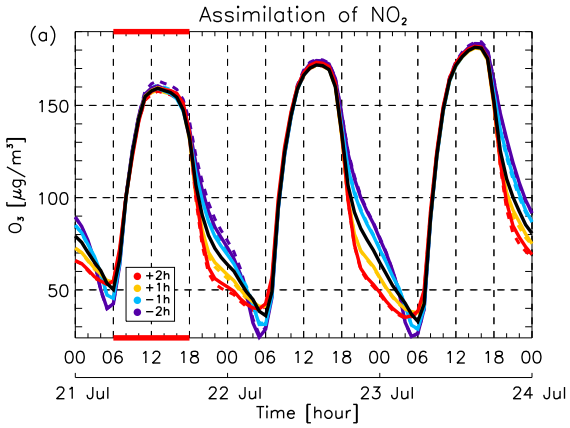

Figure 4.13: Time-series of ground O_3 concentration for unbiased NO_x emission simulation (black line) and biased NO_x emission simulations (colored lines); dashed lines represent the non-assimilated simulations and solid lines represent simulations with NO_2 assimilation in the 06-18 window.

4.4 Results summary

The objective of the study presented in this Chapter was to assess how perturbations in tropospheric O_3 concentrations caused by emission biases can be corrected by the assimilation of O_3 and NO_2 by means of an OI procedure. An OSSE approach is used to evaluate quantitatively the impact of DA, to be employed in the design of a simple assimilation procedure suitable for tropospheric ozone forecasts in polluted areas.

It was found that the air quality model BOLCHEM has a high sensitivity to intensity biases of NO_x emissions, which induce a change in the chemical compounds linked to photochemical processes (NO_2, O_3, N_2O_5, NO_3). A perturbation of up to 15% of the diurnal O_3 average in the whole domain is found.

NO_2 assimilation significantly improves O_3 maxima during the assimilation,

making it almost independent on different emission scenarios. The assimilation impact lasts up to 36-40 hours after the end of the assimilation window. This is a considerable result, especially when it is taken into account that DA generally yields significantly better forecasts in the 6-12 hours range, but improvements vanish afterwards (Elbern and Schmidt, 2001). The NO_2 night-time chemistry has the role of maintaining the correction of O_3 due to assimilation also in the following day. During the night, NO_2 assimilation increases the perturbations of O_3 with respect to the reference, deteriorating the quality of the analysis. This effect can be reduced by assimilating NO_2 and O_3 simultaneously, although the benefit lasts only a few hours after the end of the assimilation window.

Regarding the impact of different temporal assimilation windows, even if differences among them are not very marked, it was found that the initial and final assimilation times have a greater impact than the assimilation window length itself. The best results were achieved assimilating observations during the photochemically active period (06-18 UTC).

It was also observed that temporally biased NO_x emissions, during the strongest photochemical activity, slightly perturb the O_3 concentration, confirming the findings of Tao et al. (2004), while the perturbation is larger during night-time. Assimilation has a very low impact during the assimilation window and a negligible impact after its end.

In brief, even if based on an OSSE approach and focusing on one event only, this study showed that assimilating NO_2 (an O_3 precursor) can reduce the perturbations induced on O_3 forecasts by biased emissions. The positive effect remains up to 36-40 hours after the end of the assimilation, and the best result was obtained for assimilation during the photochemically active period.

Chapter 5

Different assimilated satellite geometry on tropospheric ozone

The aim of this part of the work was to evaluate the impact of assimilation of satellite NO_2 tropospheric columns on the distribution of pollutants at the ground level during photochemical pollution events. In particular, we focused on the assimilation of satellite NO_2 tropospheric columns from SCIAMACHY, on-board ENVISAT ESA satellite, and from OMI, on-board EOS Aura satellite, and its effect on ozone at low troposphere. SCIAMACHY and OMI have a considerable difference in spatial and temporal resolution and for an effective improvements in assimilate fields it is particularly important the consistency between satellite and model resolution. A second purpose was so to determine which was the more suitable satellite geometry for improving the model simulations of photochemical ozone. In this case the optimal interpolation routine was implemented in the three-dimensional Eulerian chemistry-transport model CHIMERE (see Chapter 3). The meteorological fields came from mesoscale hydrostatic model BOLAM. Two cases study were taken into account: in 2004 summer and in 2007 summer. Different

simulations were carried out, changing horizontal resolution (0.50° × 0.50° and 0.25° × 0.25°), assimilating NO_2 tropospheric columns from SCIAMACHY and from OMI. To evaluate more quantitatively and over time the impact of assimilation process and its effects on O_3 at local scale, sub-areas were selected, comparing the assimilated field with independent observations.

5.1 Numerical experiments

Two case studies of ozone tropospheric pollution above Europe are considered: (i) from 29^{th} July 2004 00 UTC till 4^{th} August 2004 00 UTC where NO_2 tropospheric columns from SCIAMACHY are assimilated; (ii) from 24^{th} August 00 UTC 2007 to 1^{st} September 00 UTC 2007 performing assimilation of NO_2 tropospheric columns from SCIAMACHY and from OMI. Five simulations are performed for these cases: two without assimilation (NOASSIM_2004_L, NOASSIM_2007_L) and the other ones activating the assimilation procedure (DA_2004_SCI_L, DA_2007_OMI_L, DA_2007_SCI_L). CHIMERE simulations are performed with an horizontal resolution of 0.5° × 0.5°, the vertical discretization consists in 8 sigma-hybrid pressure levels going from the surface up to 500hPa. Meteorological fields are provided by BOLAM model simulation at an horizontal grid of 0.4° × 0.4° and 33 sigma-hybrid pressure levels. For case study (ii) it is tested also the effect of changing horizontal resolution doubling it in CHIMERE (0.25° × 0.25°) and in BOLAM simulation (0.2° × 0.2°) (NOASSIM_2007_H, DA_2007_OMI_H). The characteristics of the performed simulation are summarized in Table 5.1.

Experiment name	Case Study	Satellite data utilized	Chemic. resolution	Meteo resolution
NOASSIM_2004_L	29-07/04-08 2004	None	0.5°×0.5°	0.4°×0.4°
DA_2004_SCI_L	29-07/04-08 2004	SCIAMACHY	0.5°×0.5°	0.4°×0.4°
NOASSIM_2007_L	24-08/01-09 2007	None	0.5°×0.5°	0.4°×0.4°
DA_2007_OMI_L	24-08/01-09 2007	OMI	0.5°×0.5°	0.4°×0.4°
DA_2007_SCI_L	24-08/01-09 2007	SCIAMACHY	0.5°×0.5°	0.4°×0.4°
NOASSIM_2007_H	24-08/01-09 2007	None	0.25°×0.25°	0.2°×0.2°
DA_2007_OMI_H	24-08/01-09 2007	OMI	0.25°×0.25°	0.2°×0.2°

Table 5.1: Experiment reference names used in text, case studies, satellite data used for assimilation, chemical and meteorological simulation resolutions.

5.2 SCHIAMACHY assimilation: 2004 event

The event occurring in summer 2004 presents favourable conditions to high photo-oxidant production and accumulation in many European regions, because of high irradiance and temperature and low wind intensities. Temperatures at the ground level simulated by BOLAM model (Figure 5.1, upper panel) between 12-18 UTC are around 30℃ in Po Valley, in many areas of Spain, in central France and between 25℃ - 28 ℃ in German and in Netherlands. In whole West-Europe the wind at the first model level from the ground is above 5 m/s. Figure 5.1, lower panel, also presents the typical pattern of NO emissions at 15 UTC adopted for the simulations and interpolated over the CHIMERE grid. Emissions are extracted from EMEP database. It is shown that high values are present in the principal European cities and most industrialized areas, in particular in Paris, Belgium and Ruhr. Combination of high emissions and favourable meteorological conditions causes to maximum ozone values to reach 150-180 μg m^{-3} in some areas and cities

as Po Valley, Paris, Belgium, Barcelona. In Figure 5.2 the observed ozone related to Belgium is showed. Concentrations come from Air Base database. The ozone average over this domain is also performed dividing urban, sub-urban and rural station. No significant differences are detected among the different averages. This can be ascribed to an uniform distribution of ozone over the whole region.

In Figure 5.3, upper panel, the NO_2 tropospheric columns from SCIAMACHY are reported. Observations show the elevated concentration of NO_2 corresponding to high polluted zones as Paris area, western Belgium, Barcelona area, Frankfurt region, ranging from 15 to 40 10^{15}molecol/cm^2.

Model innovations (Figure 5.3, lower panel) are a good indicator of the performance of data assimilation procedure and it is noticed that there is a direct correlation between high polluted areas and high innovations values. This means that model is not able to represent correctly NO_2 columns in these polluted areas. It is important to notice that the innovations are almost positive. This is in agreement with (Blond et al., 2007) who made a comprehensive comparison among SCHIAMACHY NO_2 column, CHIMERE simulations (with an horizontal resolution of 50km $\times$ 50km) and surface measurements, finding that CHIMERE systematically underestimates NO_2 concentration in the urban and sub-urban areas and underestimates SCIAMACHY NO_2 columns during spring-summer.

The perturbations to simulated NO_2 tropospheric columns, induced by SCIAMACHY assimilation, cause a direct modification in ozone concentrations (Figure 5.4). Comparing the left and right panel it is evident that ozone field at the ground without assimilation is spatially smoothed and in photochemical polluted areas it is lower with respect to the assimilated field.

For instance, Paris area is characterized by enhanced ozone values from 60-100 μg m^{-3} to 70-130 μg m^{-3}, in Belgium from 50-80 μg m^{-3} to 90-130 μg m^{-3}, in Barcelona area from 70-90 to 80-130, in Frankfurt region from 70-90 μg m^{-3} to

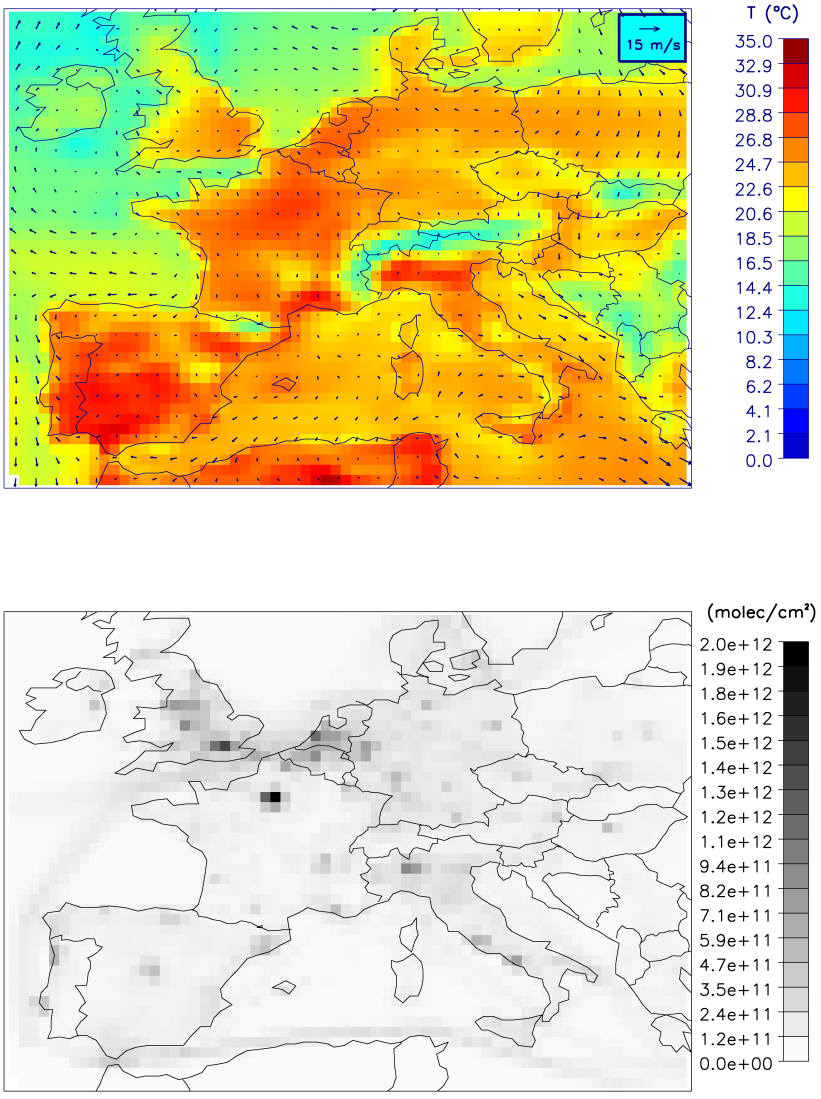

Figure 5.1: European temperature and wind (upper panel) modelled by BOLAM model and NO emissions distribution extracted from EMEP database (lower panel) at the first model level from the ground on 30^{th} July 2004 and interpolated over CHIMERE grid.

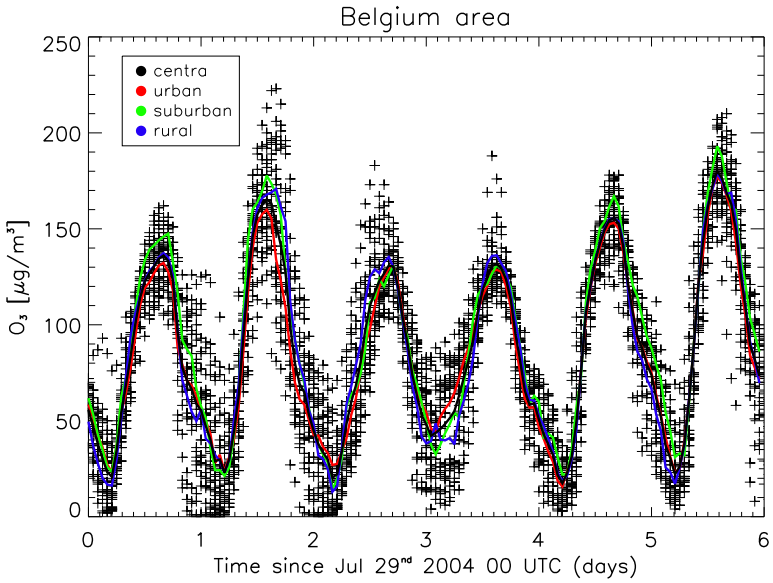

Figure 5.2: Observed ozone concentrations from Air Base network for Belgium. Black line represents the concentration averaged over the whole domain, red, green and blue line represent the average of the urban, sub-urban and rural stations, respectively. Black crosses are the values of the 26 stations (9 rural, 4 sub-urban, 13 urban).

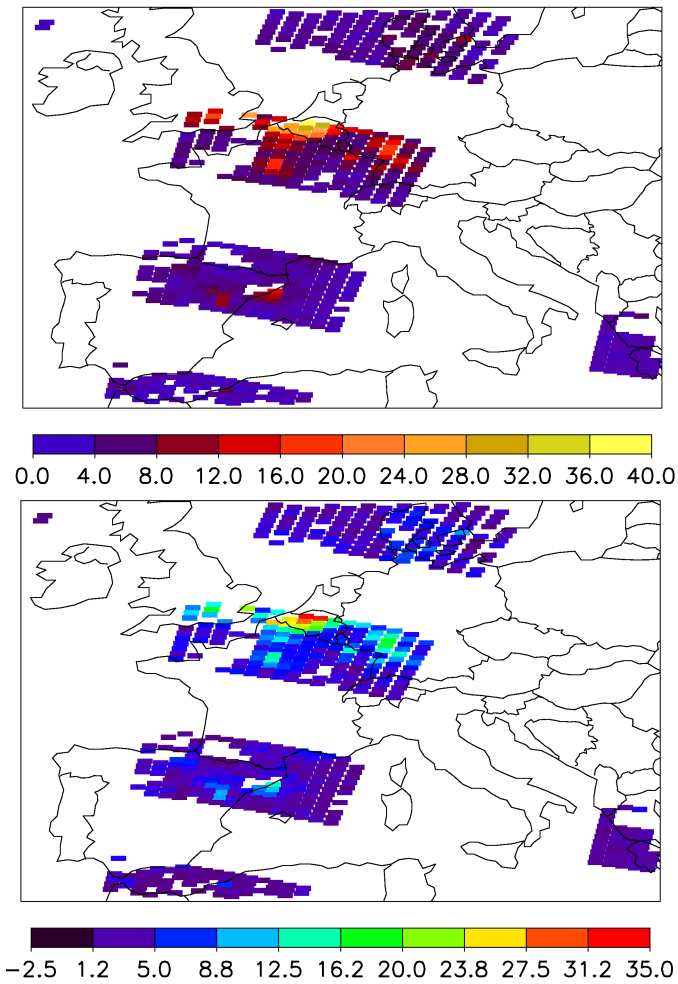

Figure 5.3: Upper panel shows the European distribution of NO_2 tropospheric columns from SCIAMACHY satellite on 30^{th} July 2004; lower panel shows the model innovations related to the same day. NO_2 tropospheric columns and innovations are expressed in 10^{15}molecol/cm^2.

80-95 μg m^{-3} and in Marseille from 80-90 μg m^{-3} to 100-105 μg m^{-3}.

Moreover, as depicted in Figure 5.4 right panel, the assimilation effect is localized to regions where the NO_2 tropospheric columns are high. This is caused to the choice in the assimilation procedure that has a short horizontal correlation in the background covariation matrix to be coherent to with the NO_2 field itself, that in general presents a structure that is highly spatially inhomogeneous.

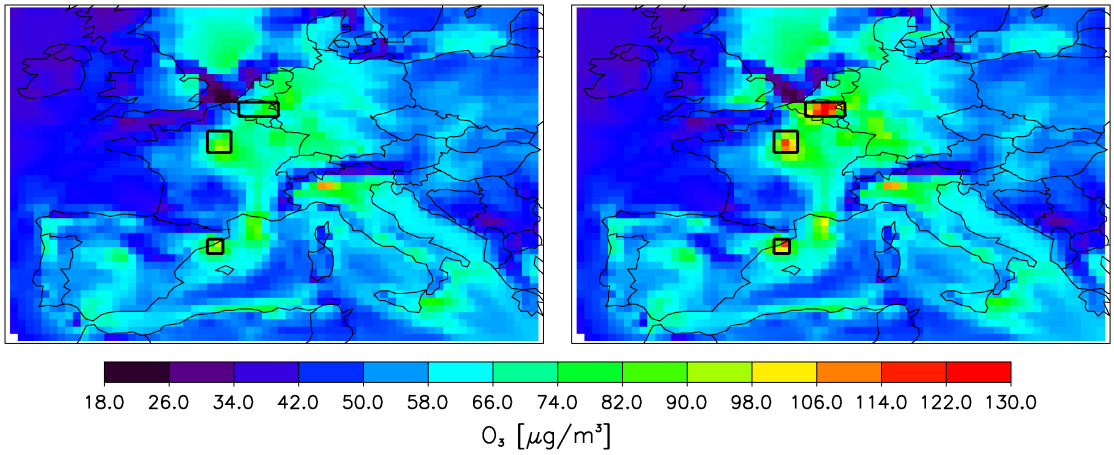

Figure 5.4: Ozone concentration at the ground level of the model (first model level has an height of about 50 m from the ground) on 30^{th} July 2004 15 UTC (after 5 hours the last assimilation). The left panel depicts NOASSIM_2004_L simulation and the right panel DA_2004_SCI_L one. Black boxes indicate the selected areas for following investigations.

5.2.1 Focus on specific sub-areas

As the effect of assimilation acts on specific regions, we focus on polluted sub-areas where the difference between ozone daily maxima of the assimilated simulation and not assimilated one is larger then 10%. The selected areas are: Paris (lat: 47.75°-49.25°, lon: 1.75°-3.25°), Belgium (lat: 50.25°-51.25°, lon: 3.75°-6.25°),Barcelona

(lat: 40.75°-41.75°, lon: 1.75°-2.75°) (reported as black boxes in Figure 5.4).

In Figure 5.5 we show the observations of NO_2 on the Belgium and Paris area, while ozone from NOASSIM_2004_L and DA_2004_SCI_L simulations and the difference between ozone from DA_2004_SCI_L and NOASSIM_2004_L are depicted in Figure 5.6. In Belgium, SCIAMACHY measures high NO_2 tropospheric values over a vast area (up to 40 10^{15}molecol/cm^2) and this brings to a high and broad correction in ozone field with a difference between the DA_2004_SCI_L and NOASSIM_2004_L up to 50 μg m^{-3}. In Paris area the measured columns are lower (up to 20 10^{15}molecol/cm^2) over a smaller area with respect to Belgium and consequently the correction on ozone field is lower and less extensive (from 10 μg m^{-3} to 35 μg m^{-3}).

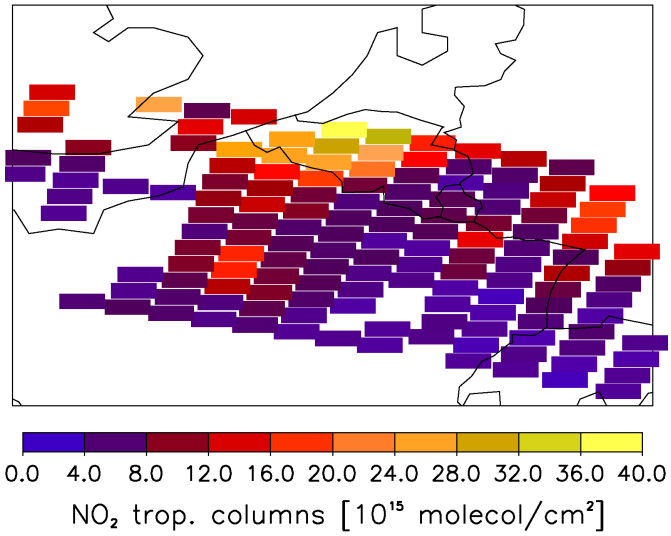

Figure 5.5: European distribution of NO_2 tropospheric columns from SCIAMACHY satellite on 30th July 2004 focused on Belgium and Paris area.

To evaluate more quantitatively the impact of assimilation process and the effects over the time, the ozone at ground level of the model is averaged in the three selected sub-areas for NOASSIM_2004_L and DA_2004_SCI_L simulations,

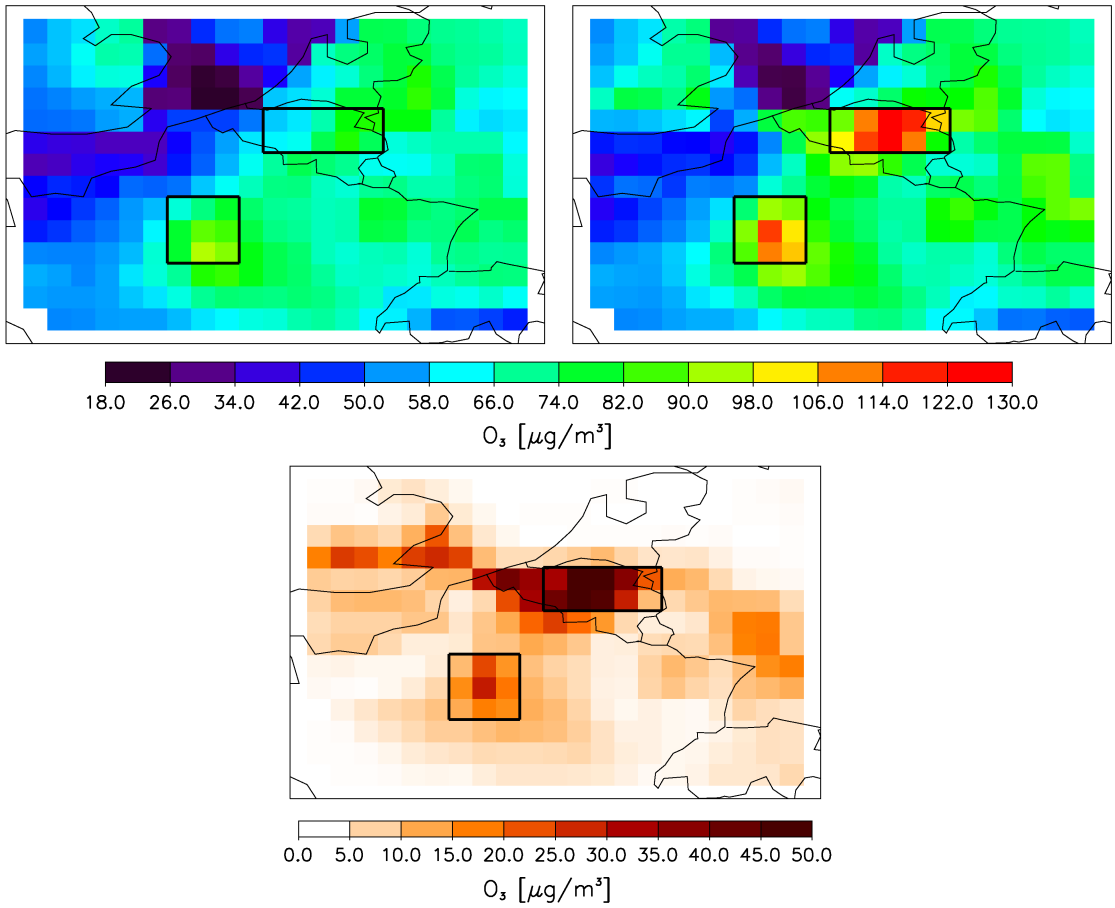

Figure 5.6: Focus on Belgium and Paris area of ozone concentration at the ground level of the model on 24^{th} at 15 UTC, related to NOASSIM_2004_L (upper left panel) and DA_2004_SCI_L (upper right panel) simulations and of the ozone differences between DA_2004_SCI_L and NOASSIM_2004_L simulations (lower panel).

respectively. The averaged values are compared to independent ground ozone observations coming from the AirBase Database (more details in Section 3.4). Firstly the observed ozone is divided in rural, urban, suburban stations and an average of each different typologies were performed, but not remarkable differences are found in all three sub-domains (see for instance 5.2 for Belgium area). Regarding Paris and Barcelona areas the stations are mostly collected in the city or they are very close to it, so even if some station are target as rural or sub-urban in case of photochemical pollution events, it is likely that they have a behaviour closer than the urban one. Regarding Belgium, the stations are quite well distributed over the whole sub-domain and as already explained in previous section, the reason of high similarity among urban, sub-urban and rural stations in the different averages could be ascribed to quasi-uniform distribution of ozone over the whole region. So, for the present comparison it was chosen to perform a simple average.

The ozone concentration averaged in Belgium area is represented in the upper panel of Figure 5.7. In that case the assimilation is done once at 10 UTC on 30^{th} July (shown as dashed red lines) and its effect is noticeable as ozone increase to 120 μg m^{-3} during the peak of photochemical pollution. Moreover, the perturbation on ozone simulation persists also at the re-activation of photochemistry in the following day. The assimilation effect lasts roughly 24-30 hours after the end of assimilation and this can be considered as a noticeable impact. A possible explanation is that NO_2 and O_3 perturbations are high enough to cause perturbations to NO_x nocturnal reservoirs like NO_3 and N_2O_5 as already described in Chapter 4.

Figure 5.7 (middle panel) shows the results for Paris area. The assimilation in that case takes place around 10 UTC on 29^{th} July and 30^{th} July. It can be noticed that assimilation improves the ozone field after 10 UTC during the photochemically active period, but the correction on ozone is drastically less than the difference with respect to observations. This could be ascribed to the fact that the observations

network gather in the city or very close to it and so they are characterized by an urban behaviour. SCIAMACHY measurement, having a resolution of 60 km $\times$ 30 km at nadir, is not able to catch the high variability of pollutants in this area, consequently assimilation effect on ozone is less marked and less close to observations. Moreover, it is observed an almost negligible impact of assimilation during night and at the re-activation of photochemistry in the following day. This indicates that model bias takes over with respect to assimilation during a time window of few hours.

In Figure 5.7 (lower panel) the ozone concentration averaged on Barcelona area is presented. Assimilation is activated at 10 UTC on 30^{th} July. In this case assimilation correction on O_3 shows the same entity of that in Paris area and there is no effect in the day following the assimilation. The degree of correction brought by assimilation is comparable to that in Paris area, anyway this is not an high photochemical pollution event based on ozone maxima and the correction brings the ozone rather close to ground observations.

5.3 OMI and SCIAMACHY assimilation: 2007 event

During the end of August 2007, the warm temperature and very low wind induce in the low troposphere the formation and accumulation of ozone, reaching high values in the principal cities of Europe. This is especially marked in South Europe and in Mediterranean basin where the temperatures are highest. BOLAM model (Figure 5.8) between 12-18 UTC shows that temperatures at the ground level are warm in Central-West Europe reaching 25℃. Regarding the emission pattern we refer to Figure 5.1 (lower panel), that is very similar to those used for simulations presented

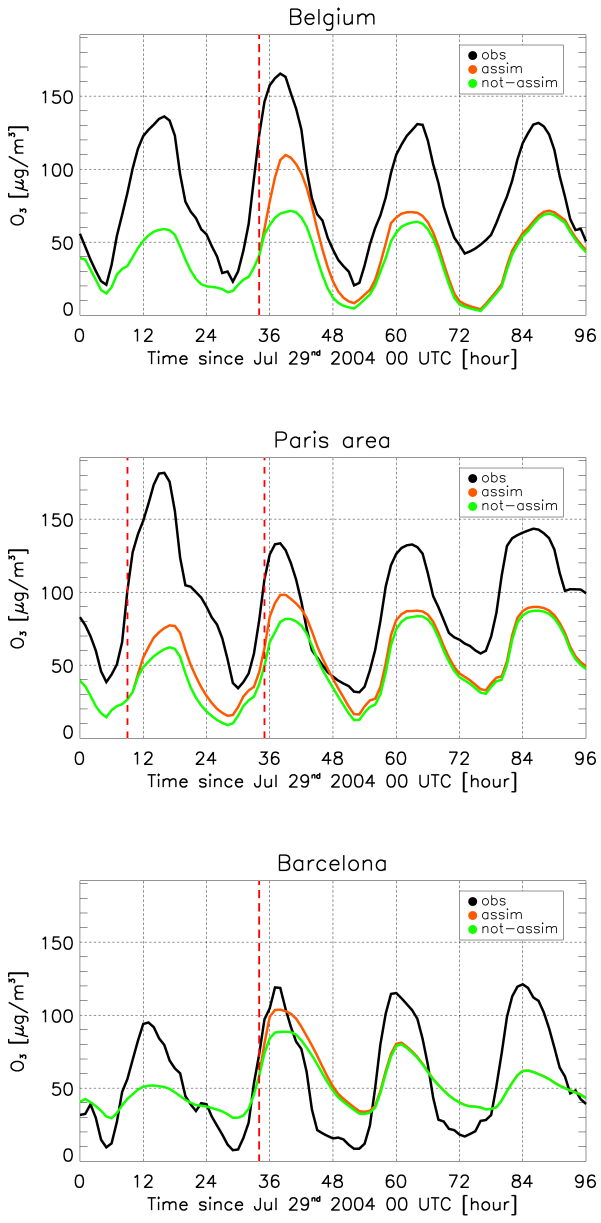

Figure 5.7: Averaged ozone in Belgium (upper panel), Paris (middle panel) and Barcelona (lower panel) area from 29^{th} July (first simulation day) till 1^{st} August (forth simulation day) for not-assimilated run (green line), assimilated one (red line) and for measurement stations (black line) respectively. Vertical red dashed lines indicate when assimilation occurs.

in the current Section. Ozone maxima concentrations from CHIMERE reach high values in Mediterranean area and South-East Europe (150-180 μg m^{-3}), while in Central Europe the values range around 100 μg m^{-3}. So we focus on less strong photochemical pollution event with respect to previous analysis. For instance, in Figure 5.9 the observed ozone in the Ruhr region is showed. Concentrations come from Air Base database. Also in this case, no significant differences are detected among urban, sub-urban and rural averages. Since the station are well distributed over this domain, also in this case we can suppose that ozone has a low variability over the whole region, with an average peak value of 115 μg m^{-3} on 25th August.

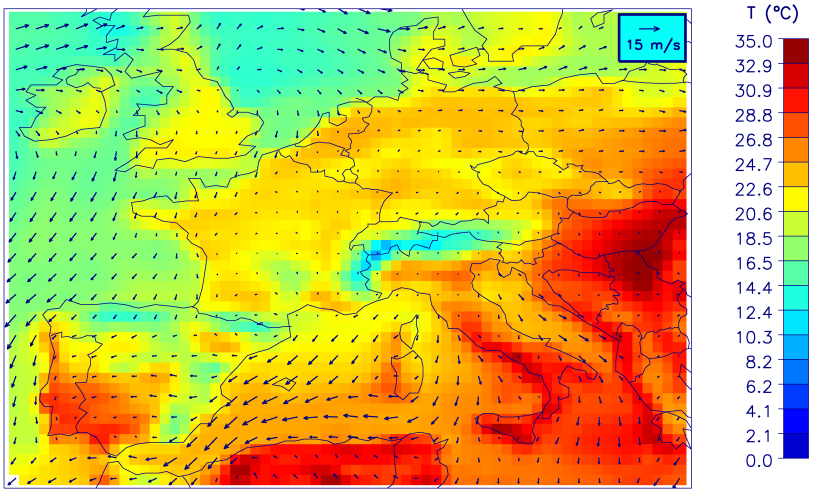

Figure 5.8: European temperature and wind distribution by BOLAM model at the ground level on 24th August 2007 and interpolated over CHIMERE grid.

Figure 5.10 depicts the NO_2 tropospheric column on 24th August 2007 for OMI and SCIAMACHY respectively. OMI provides daily global coverage with a higher horizontal resolution (13 km x 24 km) with respect to SCIAMACHY (30km x 60km). Due to the better coverage with respect to SCIAMACHY, OMI assimilation

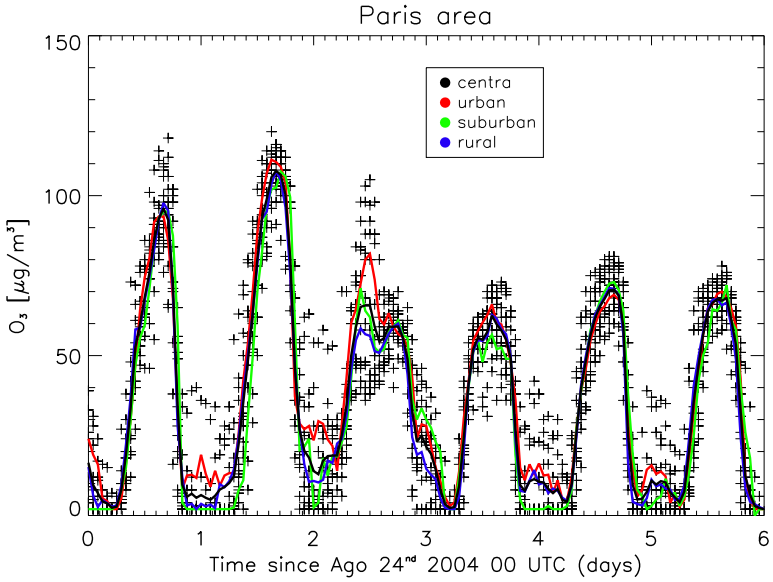

Figure 5.9: Observed ozone concentrations from Air Base network. Black line represents the concentration averaged over the whole domain, red, green and blue line represent the average of the urban, sub-urban and rural ones, respectively. Black crosses are the values of all 12 stations (7 rural, 1 sub-urban, 4 urban).

can be performed ideally once a day. High NO_2 concentrations are observed by both satellite in North-Central Europe. Nevertheless, values are lower with respect to 2004 event.

Instead of assimilating the individual OMI pixel, "super observation" were used. OMI observations lying in the same model grid were averaged and observations aggregated with the same resolution of model grid cells. The aim was to perform a DA experiment more coherent with model and observations resolutions. This choice was strengthened by the results of preliminary tests that aimed to evaluate the difference between assimilation using observations and super observation. It was obtained (not reported) that, except few localized areas, there were no significative differences in term of correction on ozone fields. This behaviour brings to the conclusion that it is not significative to assimilate different observations on the same grid cell. A remarkably lower computational cost can be achieved in the case of super observation assimilation.

Ozone concentration at the ground from NOASSIM_2007_L and DA_2007_OMI_L simulations on 24^{th} at 15 UTC is shown in Figure 5.11. As in case presented in section 5.2 by comparing the left and right panel is evident that assimilation causes an higher spatial variability in the ozone fields with respect to the case without assimilation and especially increases ozone maxima in polluted areas.

High value of ozone can be identified on the Ionian sea and South Tyrrhenian sea. It could be ascribed to the combined effect of high temperature and presumably to the intense forest fires events happened in that period in Greece (Turquety et al., 2009 and references therein). Figure 5.11 shows also that NO_2 assimilation increases ozone ground concentration in some polluted areas, especially in Ruhr region enhancing ozone from 50-70 μg m^{-3} to 70-100 μg m^{-3}, in Naples area increasing ozone more than 100 μg m^{-3} and in Palermo area rising it from 60-

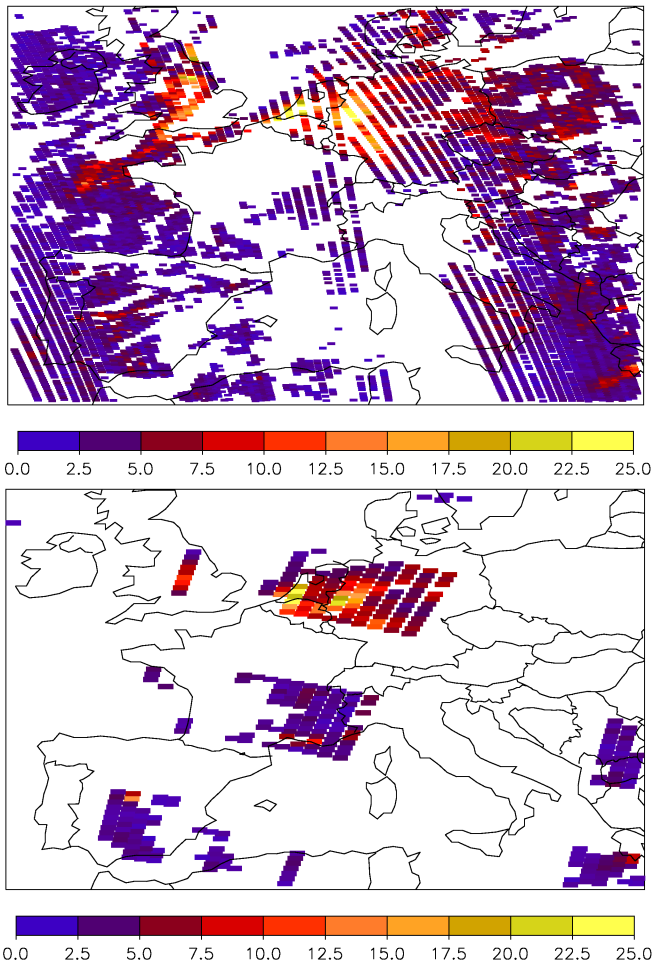

Figure 5.10: European distribution of NO_2 tropospheric columns from OMI (upper panel) and SCIAMACHY (lower panel) satellite on 24^{th}. Columns are expressed in 10^{15}molecol/cm^2.

70 μg m^{-3} to 90-100 μg m^{-3}.

In order to directly compare the performance of OMI and SCIAMACHY assimilation, a simulation with SCIAMACHY assimilation (DA_2007_SCI_L) is carried out in the period chosen for OMI assimilation experiments. Figure 5.12 shows the assimilated ozone on 24^{th} August. SCIAMACHY assimilation correction on ozone is less spatially extensive with respect to OMI one due to its lesser coverage. Anyway, in the areas where it is performed, the correction of ozone pattern is quite similar to that induced by OMI assimilation i.e. in Ruhr area, Ionian Sea. Anyway, except few zones, the correction degree is overall quite lower in areas where SCIAMACHY data are available (see Figure5.10); in Ruhr area O_3 goes from 50-70 μg m^{-3} to 60-80 μg m^{-3} and in Ionian Sea from 50-85 μg m^{-3} to 70-90 μg m^{-3}. This can be linked to the lower NO_2 columns values measured by SCIAMACHY with respect to OMI (Figure 5.10). No correction is applied in the South Tyrrhenian due to the absence of SCIAMACHY data.

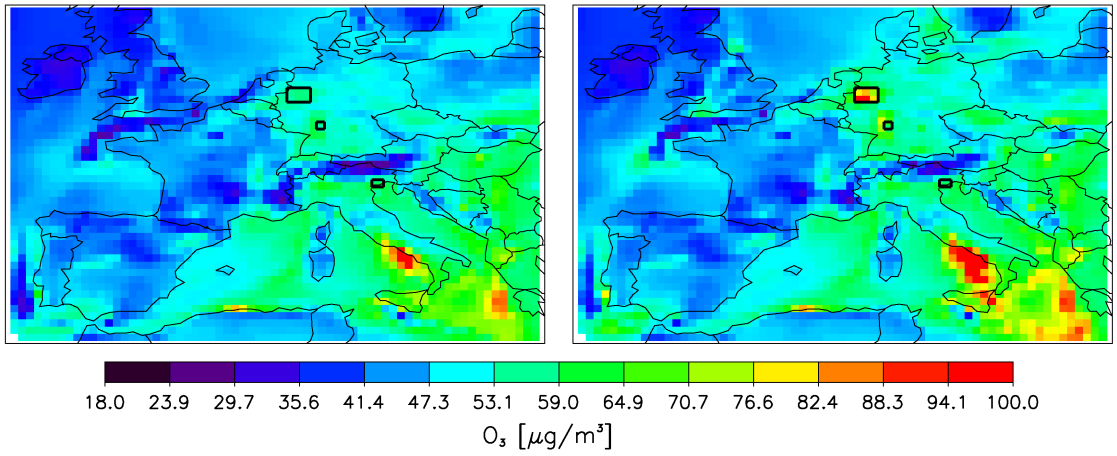

Figure 5.11: Ozone concentration at the ground level of the model on 24^{th} at 15 UTC (4 hours after the last assimilation), related to NOASSIM_2007_L (left panel) and DA_2007_OMI_L (right panel) simulations.

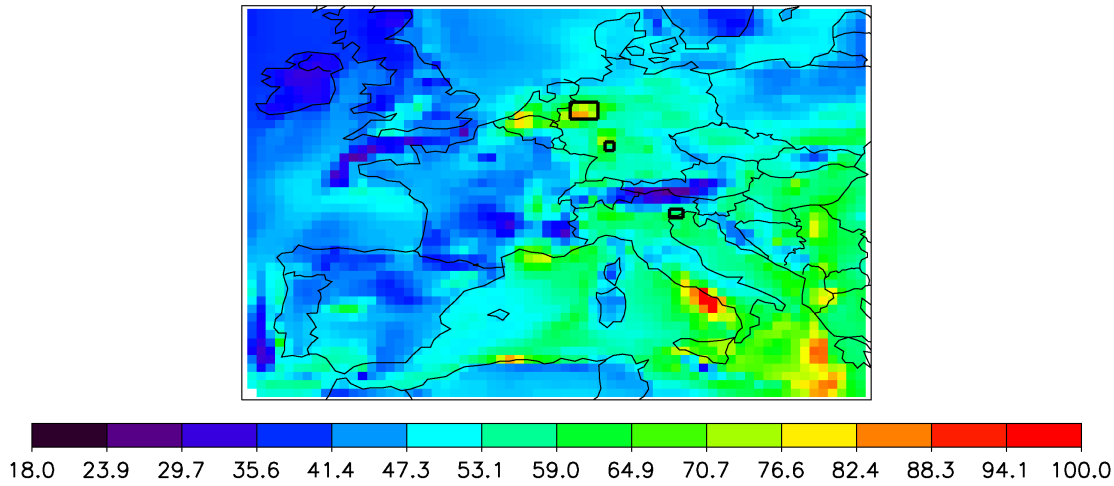

Figure 5.12: Ozone concentration at the ground level of the model on 24^{th} August 2007 at 15 UTC (4 hours after the last assimilation), related to DA_2007_SCI_L. Concentrations are expressed in μg m^{-3}.

5.3.1 Sensitivity to increased resolution

In this section the results obtained from the simulations with a finer resolution both in chemical model and meteorological one are presented. In this case the cell dimension is comparable to satellite pixel dimension and the aggregation of observation in super observations is not significant, anyway also in this case super observation are used, for a coherence with previous case.

Comparing the ozone in NOASSIM_2007_H simulation (Figure 5.13, left panel) with that in NOASSIM_2007_L (Figure 5.11, left panel), the ozone field spatial structure is more detailed and that there are more regions with high values of ozone, as for example, the principal cities of Balkans area, in South and Central Italy. These regions cover an area of few model grids, so it can be assumed that combination of higher resolution emission and meteorological parameter (as temperature, humidity, wind) plays a key role in changing simulated concentration

of atmospheric pollutants at local scale.

In Figure 5.13, right panel, ground ozone from DA_2007_OMI_H simulation is reported. We notice again that OMI assimilation enhances the ozone concentration in some polluted areas and increases its spatial variability. In particular, in Ruhr region assimilation enhances ozone from 50-70 μg m^{-3} to 75-100 μg m^{-3}, in Naples area makes broader the polluted area with ozone higher than 100 μg m^{-3}, in Palermo area assimilation rises ozone from 60-70 μg m^{-3} to 90-100 μg m^{-3}, and in Frankfurt area enhances it from 70-80 μg m^{-3} to 75-100 μg m^{-3}. Moreover, comparing these results to those showed in previous paragraph and related to DA_2007_OMI_L simulation, it is observed that in some areas, like Frankfurt on 24^{th} August and gulf of Venice on 25^{th} August (not reported), ozone DA_2007_OMI_H is significantly changed, while ozone in DA_2007_OMI_L simulation is not modified. The reason of this correction could be ascribed to the higher resolution of DA_2007_OMI_H simulation and the highly spatial variability of NO_2 and O_3 fields that is better described by a finer resolution.

5.3.2 Focus on specific sub-areas

In this section an analysis of sub-areas involved by significative correction of ozone assimilation is done through the comparison of the performance at local scale of the DA_2007_OMI_L, DA_2007_SCI_L DA_2007_OMI_H simulations. The criterion to select the sub-domains is the same adopted in Section 5.2.1 adding the condition of contemporaneous presence of OMI and SCIAMACHY measurements. The selected zones embraces the Ruhr (lat: 51.25° - 52.25°; lon: 6.75° - 8.75°) and include two smaller areas centred on Venice (lat: 45.375° - 45.875°;lon: 12.125° - 12.875°) and Frankfurt (lat: 49.375° - 49.875°; lon: 8.625° - 9.125°).

In Figure 5.14 and 5.15 NO_2 satellite super observation (that is what is

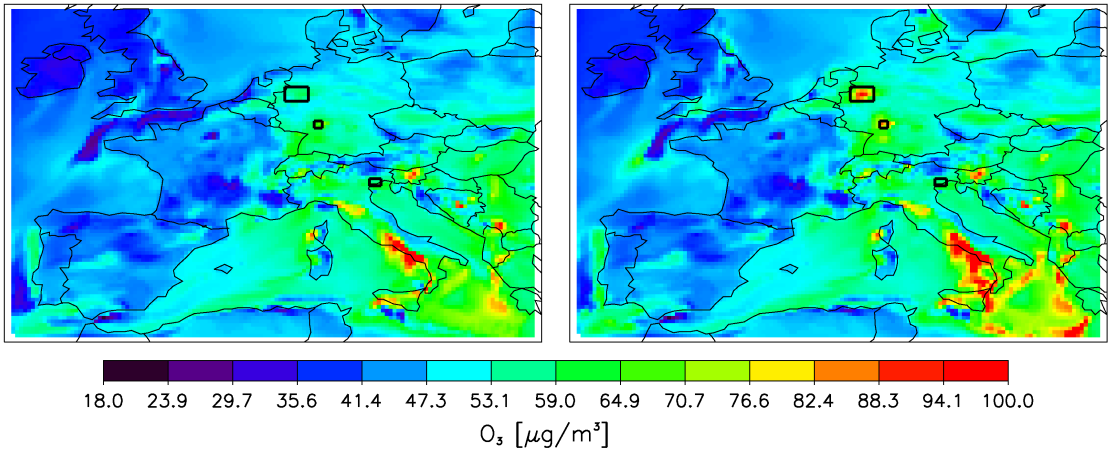

Figure 5.13: Ozone concentration at the ground level of the model on 24^{th} August 2007 15 UTC (4 hours after the last assimilation), related to NOASSIM_2007_H (left panel) and DA_2007_OMI_H (right panel).

effectively assimilated) and ozone from NOASSIM_2007_L and DA_2007_OMI_L simulations on the Ruhr and Frankfurt area are depicted. In Ruhr area OMI measures high NO_2 tropospheric values (up to 25 10^{15}molecol/cm^2) and this brings to a high and broad correction in ozone field with a difference between the DA_2007_OMI_L and NOASSIM_2007_L from 20 to 50 μg m^{-3}. Regarding the Frankfurt area the correction is almost negligible.

In Figure 5.16 and 5.17 NO_2 satellite data assimilated in DA_2007_SCI_L simulation and ozone from DA_2007_SCI_L are shown. As in the 2004 event, SCIAMACHY columns values are lower than OMI ones (up to 20 10^{15}molecol/cm^2) and this brings in Ruhr area a lower correction in ozone field with a difference between the DA_2007_SCI_L and NOASSIM_2007_L from 15 to 40 μg m^{-3}.

Figure 5.18 depicts the NO_2 satellite super-observations that has a 25km x 25km resolution and that were used in DA_2007_OMI_H experiments. Respect to those in DA_2007_OMI_L experiment (Figure 5.14) they have a finer resolution.

Thus, the assimilation in the Ruhr makes the correction to ozone more spatially variable and more localized with a difference between the DA_2007_OMI_L and NOASSIM_2007_H from 10 to 50 μg m^{-3}. Only in DA_2007_OMI_H simulation the assimilation is able to bring a correction in Frankfurt area. This is a direct impact of assimilation at higher resolution (25km), since the run at 25km without DA (NOASSIM_2007_H) does not show differences with respect to the low resolution (NOASSIM_2007_L).

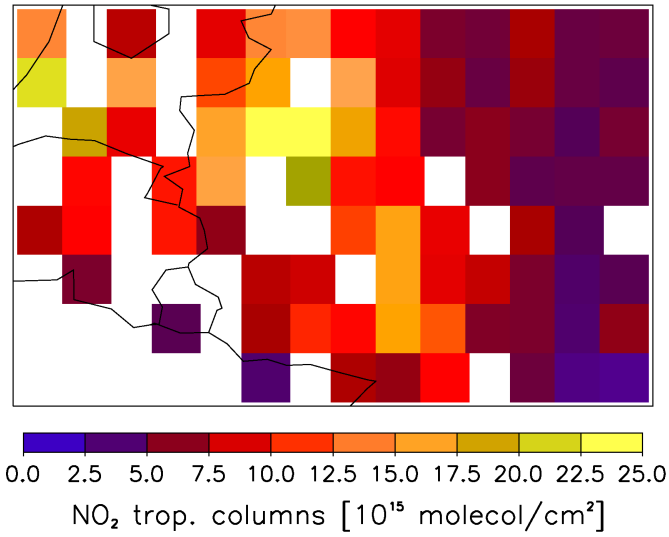

Figure 5.14: European distribution of NO_2 tropospheric columns from OMI satellite super observations at a 50km x 50km resolution on 24^{th} August 2007 focused on Ruhr and Frankfurt area.

To evaluate more quantitatively the impact of assimilation process and the effects over the time, the ozone at ground level of the model is averaged in the three selected sub-areas (Ruhr, Frankfurt, Venice) for NOASSIM_2007_L, DA_2007_OMI_L, DA_2007_SCI_L, NOASSIM_2007_H and DA_2007_OMI_H simulations, respectively. Ozone values are compared to the ground ozone observations coming from the AirBase Database. OMI assimilation

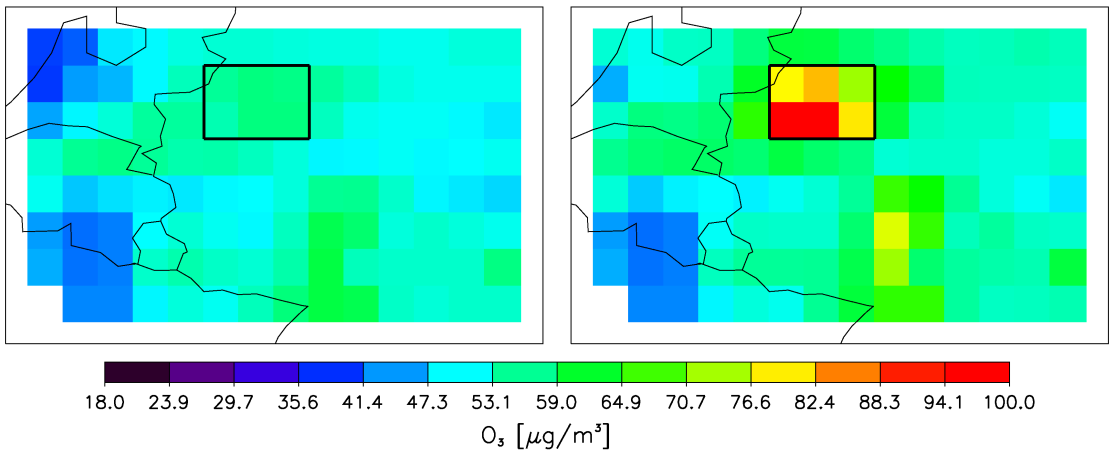

Figure 5.15: Focus on Ruhr and Frankfurt area of ozone concentration at the ground level of the model on 24^{th} at 15 UTC, related to NOASSIM_2007_L (left panel) and DA_2007_OMI_L (right panel) simulations.

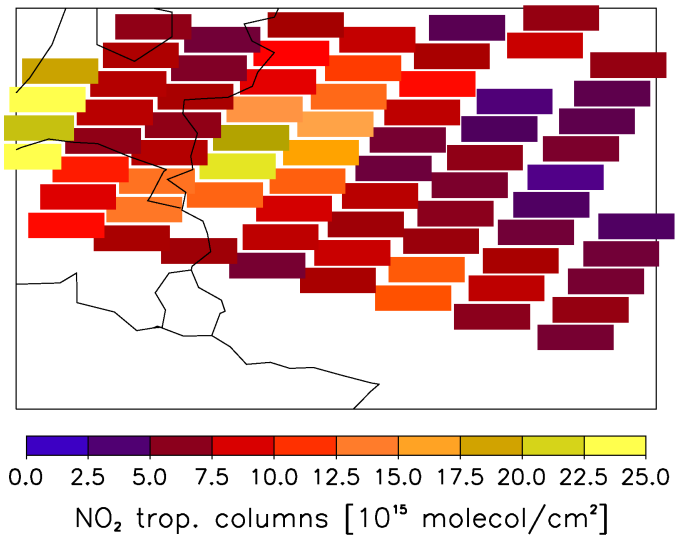

Figure 5.16: European distribution of NO_2 tropospheric columns from SCIAMACHY satellite on 24^{th} August 2007 focused on Ruhr and Frankfurt area.

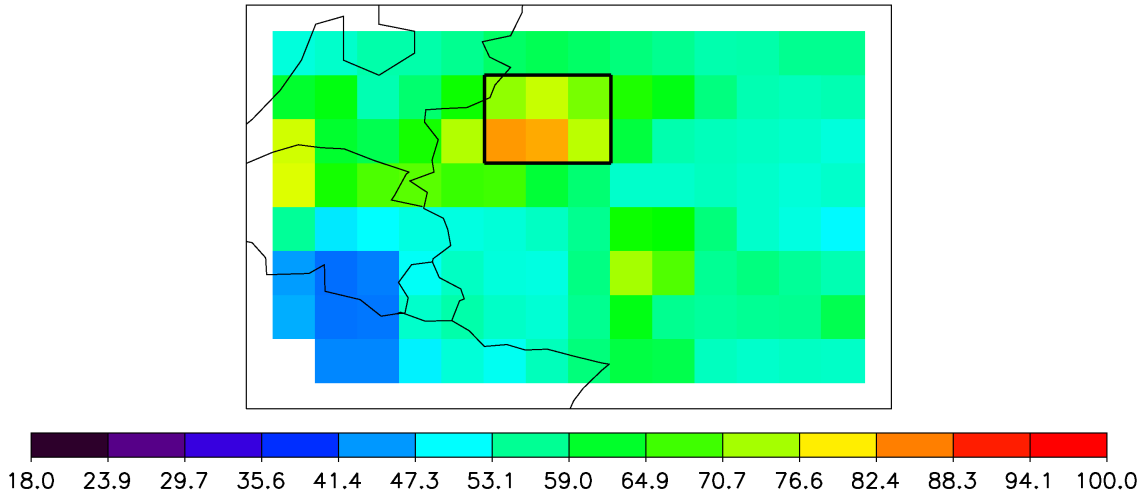

Figure 5.17: Focus on Ruhr and Frankfurt area of ozone concentration at the ground level of the model on 24^{th} at 15 UTC, related to DA_2007_SCI_L. .

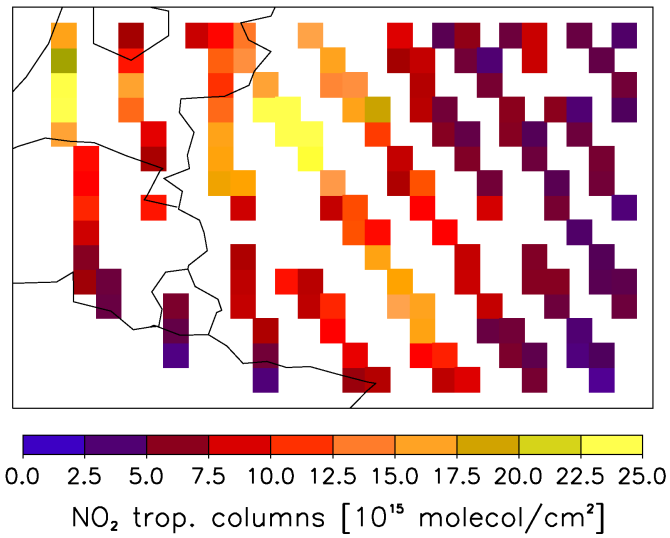

Figure 5.18: European distribution of NO_2 tropospheric columns from OMI satellite super observations at a 25km x 25km resolution on 24^{th} August 2007 focused on Ruhr and Frankfurt area.

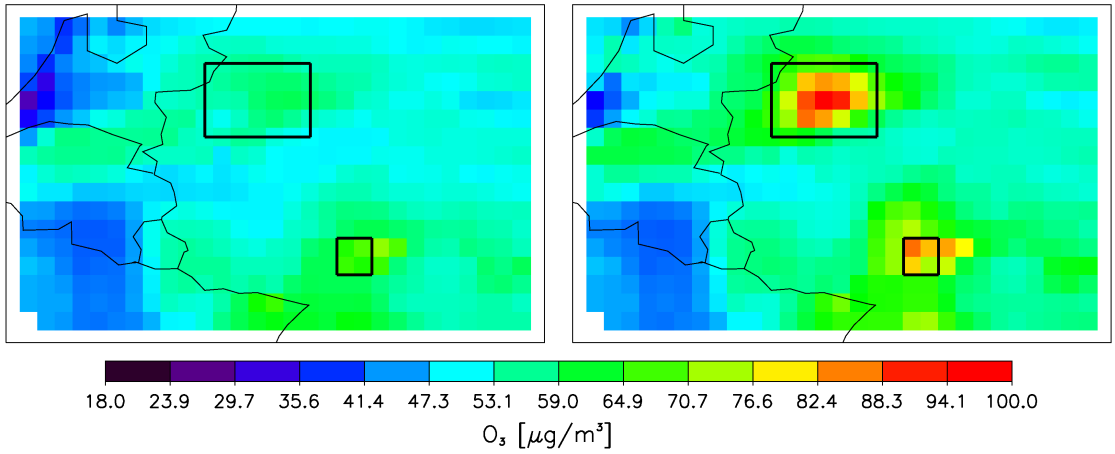

Figure 5.19: Same of Figure 5.15, but related to NOASSIM_2007_H (left panel) and DA_2007_OMI_H (right panel) simulations

takes place almost every day around 11 UTC, while SCIAMACHY at least twice in each areas.

In all cases, both SCIAMACHY and OMI assimilation take place during the more photochemical active period (around 10 UTC for SCIAMACHY and around 11 UTC for OMI) bringing to a correction of maximum ozone values.

Regarding the Ruhr area, all assimilation configurations (Figure 5.20) bring to partial correction of the ozone maxima values and they show significant improvement on ozone concentration especially in those days when the satellite coverage is more widespread (first and second day). In case of OMI data, assimilation seems to be strong enough to last more than 24 hours. Coarser OMI assimilation perform better respect to finer one. This behaviour is link to assimilation set up and OMI data distribution. In this case OMI data does not cover the whole area and super observations in 50km x 50km and in 25km x 25km resolution are quite different in terms of spatial coverage (see Figure 5.14 and 5.18), causing a more spatially extensive correction on ozone field in case

of lower resolution with respect to finer resolution case. The difference between SCIAMACHY and OMI assimilation at coarse resolution is ascribed to the lower SCIAMACHY data with respect to OMI ones.

Both Frankfurt (Figure 5.21) and Venice (Figure 5.22) areas are less extensive polluted zones then Ruhr one and present a higher spatial variability. Correction on ozone maxima, in these two cases, shows a different behaviour with respect to that in Figure 5.20. It is observed that OMI assimilation in the finer resolution simulation brings the ground ozone more closer to observations with respect to OMI assimilation in the coarse resolution simulation and SCIAMACHY assimilation.

5.4 Results summary

The experiments carried out in this chapter aimed to assess if the assimilation of NO_2 satellite tropospheric columns improves ozone field at the ground level. We focused on OMI and SCIAMACHY data assimilation, comparing their performance in two case studies.

It was that the difference between modelled and observed NO_2 tropospheric columns (innovations) were almost positive. Strongest differences were found in the most polluted areas in Europe (Ruhr, Belgium, Paris). We derived that the model has the tendency to underestimate NO_2 concentration especially in the polluted areas.

The perturbation on NO_2 field due to assimilation causes a modification on ozone field that appears more spatially variable and higher in some photochemical polluted areas. Similar effects are detected both for SCIAMACHY and OMI assimilation. Significative effects of assimilation on ozone can be appreciate at local scale.

Focusing on specific sub-domains, the effect of assimilation and its perturbations

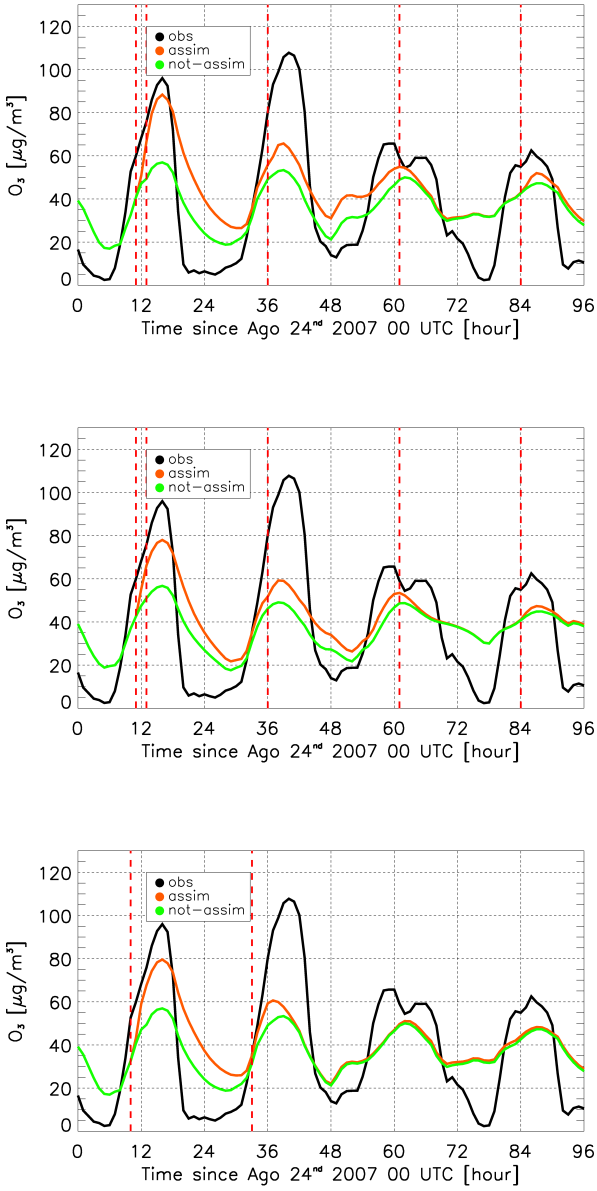

Figure 5.20: Averaged ozone in the Ruhr area, from 24^{th} 00 UTC to 28^{th} August 00 UTC, for DA_2007_OMI_L (upper panel), for DA_2007_OMI_L (middle panel) and for DA_2007_SCI_L (lower panel). Green line represents the mean from not-assimilated run, red line the mean from assimilated one and black line the mean from measurement stations (Airbase database). Vertical red dashed lines indicate when assimilation occurs.

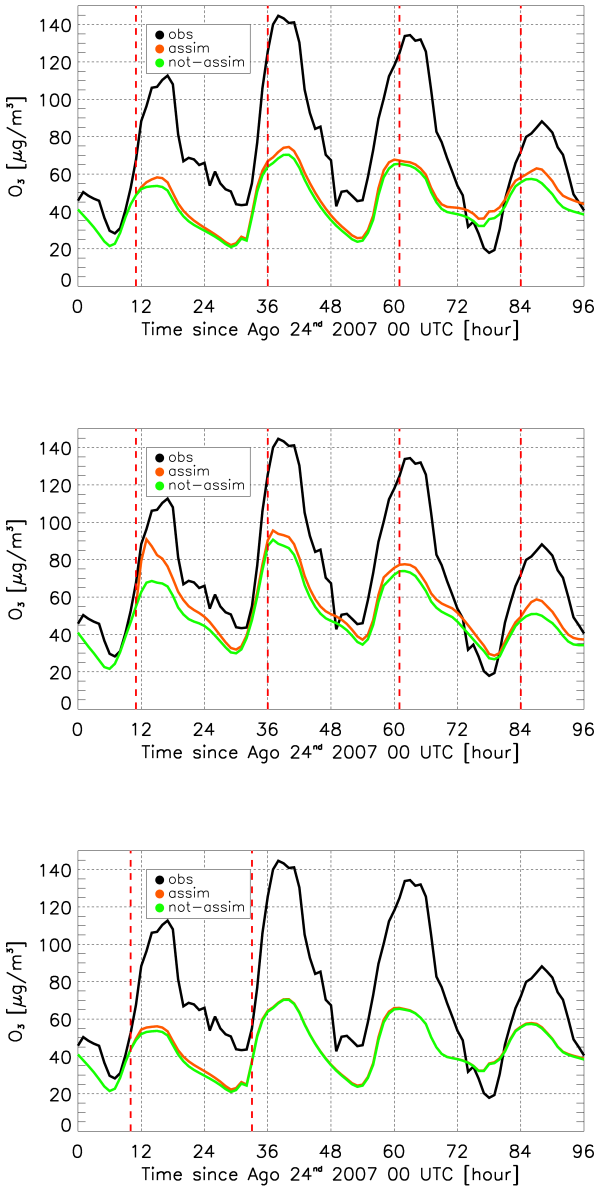

Figure 5.21: Same as 5.20, but for Frankfurt area, from 24^{th} 00 UTC to 28^{th} August 00 UTC.

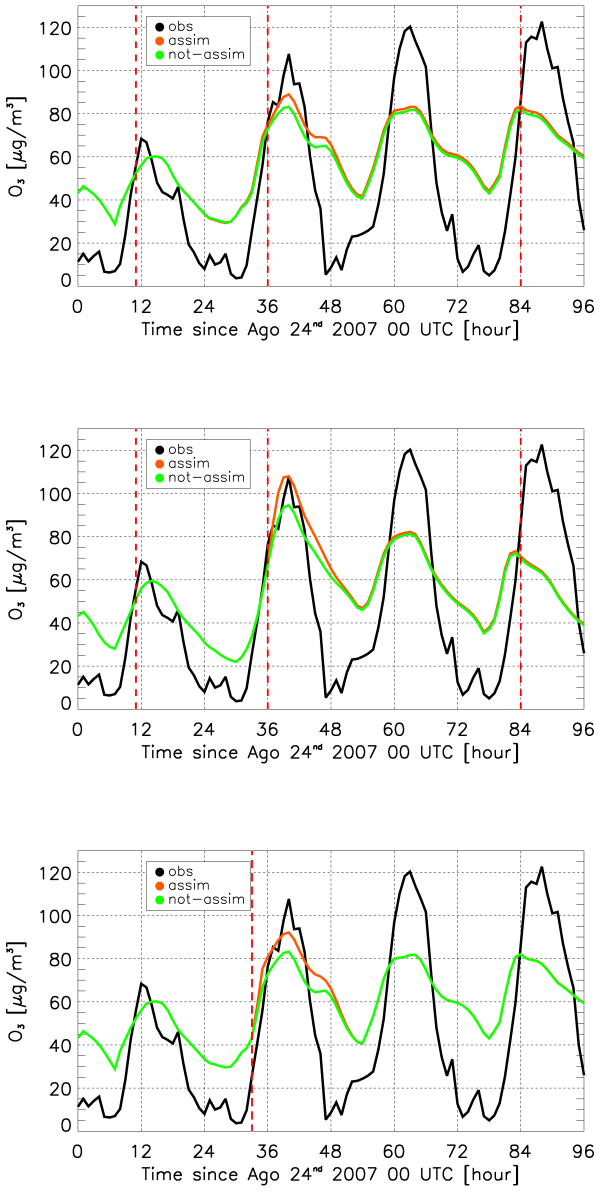

Figure 5.22: Same as 5.20, but for Venice area, from 24^{th} 00 UTC to 28^{th} August 00 UTC.

on ozone fields over the time is well visible. It was found that the effect of assimilation in general lasts 8 hours (till to the end of the active photochemical period) and in few cases until the reactivation of active photochemical period in the following day (around 30 hours after the assimilation end). This is a strong impact, considering that assimilation is performed at most once a day and it is probably linked to the model underestimation of ozone and its precursors in polluted areas with respect to those measured by SCIAMACHY and OMI.

In wide and highly polluted areas (as Belgium and Ruhr) assimilation achieves satisfactory results, comparing simulated ground ozone with independent ground measurements. In Ruhr region where OMI assimilation in the coarse resolution simulation, SCIAMACHY assimilation and OMI assimilation in the finer resolution simulation are confronted, we can conclude that these different assimilation set-up are almost similar. OMI assimilation in the coarse resolution simulation is slightly most performing and this results can be ascribe to specific characteristics of assimilated columns. In particular, with respect to SCIAMACHY assimilation it is because SCIAMACHY columns are in this case lower than OMI ones and with respect to OMI assimilation in the finer resolution simulation it is because OMI data does not cover the whole area and super observations in 50km x 50km has a larger spatial coverage with respect to 25km x 25km resolution, consequently ozone field in case of lower resolution has more spatially extensive correction with respect to finer resolution case.

Whereas, in not extensive polluted areas OMI assimilation in the finer resolution simulation performs better with respect to OMI assimilation in the coarse resolution simulation and SCIAMACHY assimilation.

Regarding the temporal distribution, OMI assimilation potentially performs better, because assimilation can be activated ideally once a day.

The efficacy of assimilation is strictly linked to good representation of pollutants

distribution of assimilated satellite data. For instance SCIAMACHY measurement, having a resolution of 60 km $\times$ 30 km at nadir, is not able to catch the high variability of pollutants in some areas, consequently assimilation effect on ozone is less marked and hence less close to ground observations. Same restrictions arise from a too coarse model resolution.

Conclusions

The present Phd thesis dealt with the assimilation of ground measurements and satellite columnar data and how this could improve and correct biases and errors in the chemical species forecast. The work focalised on tropospheric ozone and the species linked to its formation as NO_2.

The study was carried on implementing and applying an Optimal Interpolation DA technique in the air quality model BOLCHEM and the chemical transport model CHIMERE. The OI routine was chosen because it has given satisfactory results in air quality modelling and because it is relatively simple and computationally inexpensive.

Emissions represent one of the major source of uncertainties in air quality models, since they continue to be high, despite many efforts made to provide a more accurate emissions inventory. The objective of the first part of the PhD work was to assess how perturbations in tropospheric O_3 concentrations caused by emission biases could be corrected by the assimilation of O_3 and NO_2 by means of DA. An OSSE approach was used to evaluate quantitatively the impact of DA, to be employed in the design of a simple assimilation procedure suitable for tropospheric ozone forecasts in polluted areas. We assimilated ground based synthetic observations with a realistic distribution derived from AirBase database network. We focused on Po Valley, however the work can be easily extended in

other polluted areas, characterized by a reasonable number of observational sites (order of 20-30 on 40000km area) as for instance European MegaCities and hot-spot areas.

The analysis brought to the conclusions that NO_2 assimilation significantly improves O_3 maxima during the assimilation, making it almost independent on different emission scenarios. The assimilation impact lasts up to 36-40 hours after the end of the assimilation window. This is a considerable result, especially when it is taken into account that DA generally yields significantly better forecasts in the 6-12 hours range, but improvements vanish afterwards (Elbern and Schmidt, 2001). The NO_2 night-time chemistry has the role of maintaining the correction of O_3 due to assimilation also in the following day. It was observed that during night, NO_2 assimilation increases the perturbations of O_3 with respect to the reference, deteriorating the quality of the analysis. This effect can be reduced by assimilating NO_2 and O_3 simultaneously, although the benefit lasts only a few hours after the end of the assimilation window. Regarding the impact of different temporal assimilation windows, the best results are achieved assimilating observations during the photochemically active period (06-18 UTC). It was also found that temporally biased NO_x emissions only slightly perturb O_3 concentration during the photochemically active regime, confirming the findings of Tao et al. (2004), while the perturbation is larger during night-time. Assimilation has a very low impact during the assimilation window and a negligible impact after its end.

The second item was to assess if assimilation of NO_2 satellite tropospheric columns improved ozone field at the ground level. The choice of NO_2 satellite assimilation was enforced by the results obtained in the first part that confirmed the efficacy of assimilate NO_2 for correcting ozone field. We focused on OMI and SCIAMACHY data assimilation, comparing their performance in two case studies. The role of data resolution on the effectiveness of assimilation was investigated

using two different dataset and model resolutions.

It was found the perturbation on NO_2 field due to assimilation causes a modification on ozone field that appears more spatially variable and higher in some photochemical polluted areas. Similar effects are detected both for SCIAMACHY and OMI assimilation. Significative effects of assimilation on ozone can be appreciate in polluted areas at local scale.

Focusing on specific sub-domains, it was found that the effect of assimilation lasts, in general, 8 hours (till to the end of the active photochemical period) and in few cases until the reactivation of active photochemical period in the following day (around 30 hours after the assimilation end). This is a strong impact, considering that assimilation is performed at most once a day and it is probably linked to the model underestimate of ozone and its precursors in polluted areas with respect to those measured by SCIAMACHY and OMI.

In wide and highly polluted areas assimilation achieves satisfactory results, comparing simulated ground ozone with independent ground measurements. In that region where OMI assimilation in the coarse and fine resolution simulations and SCIAMACHY are confronted, we can conclude that these different assimilation set-up are almost similar. Whereas, in more localised polluted areas (i.e. comparable to model and satellite resolution), OMI assimilation in the finer resolution simulation performs better with respect to OMI assimilation in the coarse resolution simulation and SCIAMACHY assimilation. Regarding the temporal distribution, OMI assimilation potentially performs better, because assimilation can be activated ideally once a day.

As a general conclusive statement, assimilation can be an important tool to make the spatial and temporal distribution of pollutants more realistic and closer to the specific local differences with the caveat of horizontal resolution of the assimilated columns and model simulations. We have in fact shown that results are

sensitive to the resolution when the latter is comparable with the typical extension of the polluted area under exam.

We consider that this work has three logical extensions:

[1] The OSSE approach should be implemented in different events and regions to further analyse the impact of DA and especially its duration in time. Moreover OSSE approach can be used to evaluate any model bias; so a straightforward application is the evaluation of the role of meteorological parametrization that have an impact on modelled local mixing, fluxes and temperatures.

[2] The analysis of satellite observations gives initial interesting findings that deserve further studies. In particular, it would be necessary to both extent the analysis on larger time periods (i.e. a whole summer season) and to perform additional experiments to quantitatively evaluate the limits of observations and model resolution on pollutants concentrations correction. For this it might be useful to adopt an OSSE approach as demonstrated here for ground based data.

[3] A last point, that has not been fully addressed in this manuscript is the potential of DA in correcting the continental O_3 field from unknown or not suitably represented sources as forest fires. Chapter 5 shows O_3 was increased by DA during the intense biomass burning event in Greece in late August 2007; moreover this event had a large impact on whole southern Europe (see Cristofanelli et al., 2009). So, it would be important to extend the analysis presented the in paper above-mentioned, testing to which extent NO_2 and possibly CO (from the Infrared Atmospheric Sounding Interferometer, IASI) satellite observations improve the description of transport and chemistry during such continental scale events.

Appendix A

List of acronyms

AK	Averaging Kernel
AQM	Air Quality Model
BIRA-IASB	Belgian Institute for Space Aeronomy
BLUE	Best Linear Unbiased Estimator
BOLAM	BOlogna Limited Area Model
BOLCHEM	BOlogna Limited area model for meteorology and CHEMistry
CEIP	Centre on Emission Inventories and Projections
CTM	Chemical Trasport Model
DA	Data Assimilation
ECWF	European Centre for Medium-range Weather Forecast
EEA	European Environment Agency
EMEP	European Monitoring and Evaluation Programme
EnKF	Ensemble Kalman Filter
ESA	European Space Agency
GEMS	Global and regional Earth-system Monitoring

GOME	Global Ozone Monitoring Experiment
IASI	Infrared Atmospheric Sounding Interferometer
IGAC	International Global Atmospheric Chemistry
KNMI	Koninklijk Nederlands Meteorologisch Instituut
MINNI	Integrated National Model system for International Negotiation
NPW	Numerical Weather Prediction
OI	Optimal Interpolation
OMI	Ozone Monitoring Instrument
OSSE	Observing System Simulation Experiment
Prev'Air	Prévisions et observations de la qualité de l'air
RRSQRT	Reduced Rank SQuare RooT
TEMIS	Tropospheric Emission Monitoring Internet Service
SCIAMACHY	SCanning Imaging Absorption SpectroMeter for Atmospheric CartograpHY
VOC	Volatile Organic Compound
WHO	World Health Organization
3D-Var	Three-Dimensional VARiational assimilation
4D-Var	Four-Dimensional VARiational assimilation

Bibliography

Arnold Jr, C. and C. Dey, 1986: Observing-systems simulation experiments: Past, present, and future. *Bulletin of the American Meteorological Society*, **67**, 687–695.

Atkinson, R., D. L. Baulsch, R. A. Cox, R. F. Hampton, J. A. Kerr, M. J. Rossi, and J. Troe, 1997: Evaluated kinetics, photochemical and heterogeneous data. *J. Phys. Chem.*, **26**, 521 – 1012.

Aumont, B., F. Chervier, and S. Laval, 2003: Contribution of HONO sources to the NOx/HOx/O3 chemistry in the polluted boundary layer. *Atmospheric Environment*, **37**, 487–498.

Beekmann, M. and C. Derognat, 2003: Monte Carlo uncertainty analysis of a regional-scale transport chemistry model constrained by measurements from the Atmospheric Pollution Over the Paris Area (ESQUIF) campaign. *Journal of Geophysical Research*, **108**, 8559.

Beekmann, M. and R. Vautard, 2009: A modelling study of photochemical regimes over Europe: robustness and variability. *Atmospheric Chemistry and Physics Discussions*, **9**, 1521–1560.

Blond, N., K. Boersma, H. Eskes, R. Van der A, M. Van Roozendael, I. De Smedt,

G. Bergamatti, and R. Vautard, 2007: Intercomparison of SCIAMACHY nitrogen dioxide observations, in situ measurements and air quality modeling results over Western Europe. *J. Geophys. Res*, **112**.

Boersma, K., H. Eskes, and E. Brinksma, 2004: Error analysis for tropospheric NO2 retrieval from space. *Journal of Geophysical Research*, **109**, D04311.

Boersma, K., H. Eskes, J. Veefkind, E. Brinksma, M. Sneep, G. Van Der Oord, P. Levelt, P. Stammes, J. Gleason, and E. Bucsela, 2006: Near-real time retrieval of tropospheric NO2 from OMI. *Atmospheric Chemistry and Physics*, **6**, 12301–12345.

Bouttier, F. and P. Courtier, 1999: Data assimilation concepts and methods. *MeteorologicalMeteorological Training Course Lecture Series*, 1–59.

Buzzi, A., M. D'Isidoro, and S. Davolio, 2003: A case-study of an orographic cyclone south of the alps during the map sop. *Quarterly Journal of the Royal Meteorological Society*, **129**, 1795–1818.

Buzzi, A., M. Fantini, P. Malguzzi, and P. Nerozzi, 1994: Validation of a limited area model in cases of mediterranean cyclogenesis: surface fields and precipitation scores. *Meteorology and Atmospheric Physics*, **53**, 137–153.

Carmichael, G., A. Sandu, and A. Potra, 1997: Sensitivity analysis for atmospheric chemistry models via automatic differentiation. *Atmospheric Environment*, **31**, 475–489.

Carmichael, G. R., A. Sandu, T. Chai, D. N. Daescu, E. M. Constantinescu, and Y. Tang, 2008: Predicting air quality: Improvements through advanced methods to integrate models and measurements. *J. Comput. Phys.*, **227**, 3540–3571.

Carter, W., 1990: A detailed mechanism for the gas-phase atmospheric reactions of organic compounds. *Atmospheric Environment*, **24**, 481–518.

Chai, T., G. Carmichael, Y. Tang, A. Sandu, A. Heckel, A. Richter, and J. Burrows, 2009: Regional NOx emission inversion through a four-dimensional variational approach using SCIAMACHY tropospheric NO_2 column observations. *Atmospheric Environment*, **43**, 5046–5055.

Clerbaux, C., J. Hadji-Lazaro, D. Hauglustaine, G. Mégie, B. Khattov, and J. Lamarque, 2001: Assimilation of carbon monoxide measured from satellite in a three-dimensional chemistry-transport model. *Journal of Geophysical Research*, **106**, 15–385.

Constantinescu, E., T. Chai, A. Sandu, and G. Carmichael, 2007a: Autoregressive models of background errors for chemical data assimilation. *Journal of Geophysical Research*, **112**.

Constantinescu, E., A. Sandu, T. Chai, and G. Carmichael, 2007b: Assessment of ensemble-based chemical data assimilation in an idealized setting. *Atmospheric Environment*, **41**, 18–36.

Constantinescu, E., A. Sandu, T. Chai, and G. Carmichael, 2007c: Ensemble-based chemical data assimilation. I: General approach. *Quarterly Journal of the Royal Meteorological Society*, **133**, 1229–1244.

Cristofanelli, P., A. Marinoni, J. Arduini, U. Bonafè, F. Calzolari, T. Colombo, S. Decesari, R. Duchi, M. Facchini, F. Fierli, et al., 2009: Significant variations of trace gas composition and aerosol properties at Mt. Cimone during air mass transport from North Africa–contributions from wildfire emissions and mineral dust. *Atmospheric Chemistry & Physics Discussions*, **9**, 7825–7872.

Daley, R., 1991: *Atmospheric Data Analysis*, Cambridge University Press.

DeMore, W., S. Sander, D. Golden, R. Hampson, M. Kurylo, C. Howard, A. Ravishankara, C. Kolb, and M. Molina, 1997: Chemical kinetics and photochemical data for use in stratospheric modeling. Evaluation number 12. *JPL Publication*, **97**, 1–266.

Derognat, C., M. Beekmann, M. Baeumle, D. Martin, and H. Schmidt, 2003: Effect of biogenic volatile organic compound emissions on tropospheric chemistry during the Atmospheric Pollution Over the Paris Area (ESQUIF) campaign in the Ile-de-France region. *Journal of Geophysical Research*, **108**, 8560.

D'Isidoro, M., 2005: Valutazione dellImpatto di Dati Osservativi sulle Previsioni Numeriche a Mesoscala, Ph.D. thesis, Ferrara University.

Dosio, A., S. Galmarini, and G. Graziani, 2002: Simulation of the circulation and related photochemical ozone dispersion in the Po plains (northern Italy): comparison with the observations of a measuring campaign. *Journal of Geophysical Research*, **107**, 8189.

Duncan, B., Y. Yoshida, J. Olson, S. Sillman, R. Martin, L. Lamsal, Y. Hu, K. Pickering, C. Retscher, D. Allen, et al., 2010: Application of OMI observations to a space-based indicator of NOx and VOC controls on surface ozone formation. *Atmospheric Environment*.

Edwards, D., A. Arellano Jr, and M. Deeter, 2009: A satellite observation system simulation experiment for carbon monoxide in the lowermost troposphere. *Journal of Geophysical Research*, **114**, D14304.

Elbern, H. and H. Schmidt, 2001: Ozone episode analysis by four-dimensional

variational chemistry data assimilation. *Journal of Geophysical Research*, **106**, 3569–3590.

Elbern, H., H. Schmidt, O. Talagrand, and A. Ebel, 2000: 4D-variational data assimilation with an adjoint air quality model for emission analysis. *Environmental Modelling and Software*, **15**, 539–548.

Elbern, H., A. Strunk, H. Schmidt, and O. Talagrand, 2007: Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmospheric Chemistry and Physics*, **7**, 3749–3769.

Eskes, H. and K. Boersma, 2003: Averaging kernels for DOAS total-column satellite retrievals. *Atmospheric Chemistry and Physics Discussions*, **3**, 895–910.

Fierli, F., A. Hauchecorne, S. Bekki, B. Théodore, et al., 2002: Data assimilation of stratospheric ozone using a high-resolution transport model. *Geophysical Research Letters*, **29**, 1381.

Gabusi, V. and M. Volta, 2005: Seasonal modelling assessment of ozone sensitivity to precursors in northern Italy. *Atmospheric Environment*, **39**, 2795–2804.

Gery, M., G. Witten, J. Killus, and M. Dodge, 1989: A photochemical kinetics mechanism for urban and regional scale computer modeling. *Journal of Geophysical Research*, **94**, 12925–12956.

Ghil, M. and P. Malanotte-Rizzoli, 1991: Data assimilation in meteorology and oceanography. *Adv. Geophys*, **33**, 141–266.

Hamill, T., J. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, **129**, 2776–2790.

Hanea, R., G. Velders, and A. Heemink, 2004: Data assimilation of ground-level ozone in Europe with a Kalman filter and chemistry transport model. *Journal of Geophysical Research*, **109**, 1–19.

Hanea, R., G. Velders, A. Segers, M. Verlaan, and A. Heemink, 2007: A hybrid Kalman filter algorithm for large-scale atmospheric chemistry data assimilation. *Monthly Weather Review*, **135**, 140–151.

Hanna, S., J. Chang, and M. Fernau, 1998: Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables. *Atmospheric Environment*, **32**, 3619–3628.

Hanna, S., Z. Lu, H. Christopher Frey, N. Wheeler, J. Vukovich, S. Arunachalam, M. Fernau, and D. Alan Hansen, 2001: Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. *Atmospheric Environment*, **35**, 891–903.

He, S., G. Carmichael, A. Sandu, B. Hotchkiss, and V. Damian-Iordache, 2000: Application of ADIFOR for air pollution model sensitivity studies. *Environmental Modelling & Software*, **15**, 549–557.

Hobbs, P., 2000: *Introduction to atmospheric chemistry*, Cambridge University Press.

Honoré, C., L. Rou

"ıl, R. Vautard, M. Beekmann, B. Bessagnet, A. Dufour, C. Elichegaray, J. Flaud, L. Malherbe, F. Meleux, et al., 2008: Predictability of European air quality: Assessment of 3 years of operational forecasts and analyses by the PREVAIR system. *J. Geophys. Res*, **113**.

Horowitz, L., S. Walters, D. Mauzerall, L. Emmons, P. Rasch, C. Granier, X. Tie, J. Lamarque, M. Schultz, G. Tyndall, et al., 2003: A global simulation of tropospheric ozone and related tracers: Description and evaluation of MOZART, version 2. *J. Geophys. Res*, **108**, 4784.

Houtekamer, P., L. Lefaivre, J. Derome, H. Ritchie, and H. Mitchell, 1996: A system simulation approach to ensemble prediction. *Monthly Weather Review*, **124**, 1225–1242.

Houtekamer, P. and H. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, **129**, 123–137.

Hubbard, M. and N. Nikiforakis, 2003: A three-dimensional, adaptive, godunov-type model for global atmospheric flows. *Monthly Weather Review*, **131**, 1848–1864.

Huijnen, V., H. Eskes, A. Poupkou, H. Elbern, K. Boersma, G. Foret, M. Sofiev, A. Valdebenito, J. Flemming, O. Stein, et al., 2010: Comparison of OMI NO2 tropospheric columns with an ensemble of global and European regional air quality models. *Atmos. Chem. Phys*, **10**, 3273–3296.

Ide, K., 1997: P. Courtier, M. Ghil, and AC Lorenc, 1997: Unified notation for data assimilation: Operational, sequential, and variational. *J. Meteor. Soc. Japan*, **75**, 181–189.

Jeuken, A., H. Eskes, P. Van Velthoven, H. Kelder, and E. Hólm, 1999: Assimilation of total ozone satellite measurements in a three-dimensional tracer transport model. *Journal of Geophysical Research*, **104**, 5551–5563.

Kain, J., 2004: The kain-fritsch convective parameterization: an update. *Journal of Applied Meteorology*, **43**, 170–181.

Kain, J. and J. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *Journal of the Atmospheric Sciences*, **47**, 2784–2802.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press.

Khattatov, B., J. Lamarque, L. Lyjak, R. Menard, P. Levelt, X. Tie, G. Brasseur, and J. Gille, 2000: Assimilation of satellite observations of long-lived chemical species in global chemistry transport models. *Journal of Geophysical Research*, **105**, 29135.

Kleinman, L., 1991: Seasonal dependence of boundary layer peroxide concentration: The low and high NOx regimes. *Journal of Geophysical Research*, **96**.

Kleinman, L., 1994: Low and high NOx tropospheric photochemistry. *Journal of Geophysical Research*, **99**.

Kleinman, L., P. Daum, J. Lee, Y. Lee, L. Nunnermacker, S. Springston, L. Newman, J. Weinstein-Lloyd, and S. Sillman, 1997: Dependence of ozone production on NO and hydrocarbons in the troposphere. *Geophysical Research Letters*, **24**.

Lamarque, J., B. Khattatov, J. Gille, and G. Brasseur, 1999: Assimilation of Measurement of Air Pollution from Space (MAPS) CO in a global three-dimensional model. *Journal of Geophysical Research*, **104**, 26209.

Lawrence, M., P. Rasch, R. Von Kuhlmann, J. Williams, H. Fischer, M. De Reus, J. Lelieveld, P. Crutzen, M. Schultz, P. Stier, et al., 2003: Global chemical weather forecasts for field campaign planning: predictions and observations of

large-scale features during MINOS, CONTRACE, and INDOEX. *Atmospheric Chemistry and Physics*, **3**, 267–289.

Lin, X., M. Trainer, and S. Liu, 1988: On the nonlinearity of the tropospheric ozone production. *Journal of Geophysical Research*, **93**.

Logan, J., M. Prather, S. Wofsy, and M. McElroy, 1981: Tropospheric chemistry: A global perspective. *Journal of Geophysical Research*, **86**, 7210–7254.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Q. J. Roy. Meteor. Soc.*, **112**, 1177–1190.

Lu, C. and J. Chang, 1998: On the indicator-based approach to assess ozone sensitivities and emissions features. *Journal of Geophysical Research*, **103**.

Malguzzi, P., G. Grossi, A. Buzzi, R. Ranzi, and R. Buizza, 2006: Climate and Dynamics-D24106-The 1966" century" flood in Italy: A meteorological and hydrological revisitation (DOI 10.1029/2006JD007111). *Journal of Geophysical Research-Part D-Atmospheres*, **111**.

Mallet, V. and B. Sportisse, 2005: A comprehensive study of ozone sensitivity with respect to emissions over Europe with a chemistry-transport model. *Journal of Geophysical Research*, **110**.

Ménard, R. and L. Chang, 2000: Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part II: χ 2-validated results and analysis of variance and correlation dynamics. *Monthly Weather Review*, **128**, 2672–2686.

Mendoza-Dominguez, A. and A. Russell, 2001: Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. *Atmospheric Environment*, **35**, 2879–2894.

Menut, L., 2003: Adjoint modeling for atmospheric pollution process sensitivity at regional scale. *Journal of Geophysical Research*, **108**, 8562.

Menut, L., R. Vautard, M. Beekmann, and C. Honoré, 2000: Sensitivity of photochemical pollution using the adjoint of a simplified chemistry-transport model. *Journal of Geophysical Research*, **105**, 15.

Mircea, M., M. D'Isidoro, A. Maurizi, L. Vitali, F. Monforti, G. Zanini, and F. Tampieri, 2008: A comprehensive performance evaluation of the air quality model BOLCHEM to reproduce the ozone concentrations over Italy. *Atmospheric Environment*, **42**, 1169–1185.

Miyazaki, Y., Y. Kondo, N. Takegawa, R. Weber, M. Koike, K. Kita, M. Fukuda, Y. Ma, A. Clarke, V. Kapustin, et al., 2005: Contribution of particulate nitrate to airborne measurements of total reactive nitrogen. *Journal of Geophysical Research*, **110**, D15304.

Monks, P., 2000: A review of the observations and origins of the spring ozone maximum. *Atmospheric Environment*, **34**, 3545–3561.

Morcrette, J., S. Clough, E. Mlawer, and M. Iacono, 1998: Impact of a validated radiative transfer scheme, RRTM, on the ECMWF model climate and 10-day forecasts, ECMWF. *Reading, UK, technical memorandum*, **252**.

Oort, A., 1989: Angular momentum cycle in the atmosphere-ocean-solid earth system. *Bulletin of the American Meteorological Society*, **70**, 1231–1242.

Placet, M., C. Mann, R. Gilbert, and M. Niefer, 2000: Emissions of ozone precursors from stationary sources: a critical review. *Atmospheric Environment*, **34**, 2183–2204.

Poppe, D., M. Wallasch, and J. Zimmermann, 1993: The dependence of the concentration of OH on its precursors under moderately polluted conditions: a model study. *Journal of Atmospheric Chemistry*, **16**, 61–78.

Pressman, D., 1994: Chislennaja model'gidrotermicheskikh processov v pochve kak chast'skhemy mezomasshtabnogo prognoza (A numerical model of hydrothermical processes in soil as a part of a mesoscale weather forecast scheme). *Meteorologiya i gidrologiya (Meteorology and Hydrology)*, **11**, 62–75.

Quélo, D., V. Mallet, and B. Sportisse, 2005: Inverse modeling of NOx emissions at regional scale over Northern France. Preliminary investigation of the second-order sensitivity. *Journal of Geophysical Research*, **110**.

Riishojgaard, L. P., I. Stajner, and G. P. Lou, 2000: The geos ozone data assimilation system. *Advances in Space Research*, **25**, 1063 – 1072.

Ritter, B. and J. Geleyn, 1992: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Monthly weather review*, **120**, 303–325.

Russell, A. and R. Dennis, 2000: NARSTO critical review of photochemical models and modeling. *Atmospheric Environment*, **34**, 2283–2324.

Salstein, D. and R. Rosen, 1986: Earth Rotation as a Proxy for Interannual Variability in Atmospheric Circulation, 1860-Present. *Journal of Applied Meteorology*, **25**, 1870–1877.

Sandu, A., D. Daescu, and G. Carmichael, 2003: Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: Part I–theory and software tools. *Atmospheric Environment*, **37**, 5083–5096.

Sandu, A., D. Daescu, G. Carmichael, and T. Chai, 2005: Adjoint sensitivity analysis of regional air quality models. *Journal of Computational Physics*, **204**, 222–252.

Sawyer, R., R. Harley, S. Cadle, J. Norbeck, R. Slott, and H. Bravo, 2000: Mobile sources critical review: 1998 NARSTO assessment. *Atmospheric Environment*, **34**, 2161–2181.

Schmidt, H., C. Derognat, R. Vautard, and M. Beekmann, 2001: A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in western europe. *Atmospheric Environment*, **35**, 6277 – 6297.

Schmidt, H. and D. Martin, 2003: Continental scale adjoint sensitivity of ozone concentrations in the Paris region during the ESQUIF episodes. *Journal of Geophysical Research*, **108**, 4313.

Seigneur, C., G. Stephanopoulos, and R. Carr Jr, 1982: Dynamic sensitivity analysis of chemical reaction systems:: A variational method. *Chemical Engineering Science*, **37**, 845–853.

Sillman, S., 1995: The use of NOy, H_2O_2, and HNO_3 as indicators for ozone-NOx-hydrocarbon sensitivity in urban locations. *Journal of Geophysical Research*, **100**, 14175–14188.

Sillman, S., J. Logan, and S. Wofsy, 1990: The sensitivity of ozone to nitrogen oxides and hydrocarbons in regional ozone episodes. *Journal of Geophysical Research*, **95**, 1837–1851.

Sillman, S., R. Vautard, L. Menut, and D. Kley, 2003: O_3-NOx-VOC sensitivity and NOx-VOC indicators in Paris: Results from models and Atmospheric Pollution

Over the Paris Area (ESQUIF) measurements. *Journal of Geophysical Research*, **108**, 8563.

Sillman, S. and J. J. West, 2009: Reactive nitrogen in Mexico City and its relation to ozone-precursor sensitivity: results from photochemical models. *Atmospheric Chemistry and Physics*, **9**, 3477–3489.

Simpson, D., 1992: Long period modeling of photochemical oxidants in Europe, calculations for July 1985. *Atmospheric Environment*, **26**, 1609 – 1634.

Simpson, D., W. Winiwarter, G. Börjesson, S. Cinderby, A. Ferreiro, A. Guenther, C. Hewitt, R. Janson, M. Khalil, S. Owen, et al., 1999: Inventorying emissions from nature in Europe. *Journal of Geophysical Research*, **104**.

Sportisse, B., 2007: A review of current issues in air pollution modeling and simulation. *Computers & Geosciences*, **11**, 159–181.

Talagrand, O., 1997: Assimilation of observations, an introduction. *Journal of the Meteorological Society of Japan*, **75**, 191–209.

Tao, Z., S. Larson, A. Williams, M. Caughey, and D. Wuebbles, 2004: Sensitivity of regional ozone concentrations to temporal distribution of emissions. *Atmospheric Environment*, **38**, 6279–6285.

Thunis, P., L. Rouil, C. Cuvelier, R. Stern, A. Kerschbaumer, B. Bessagnet, M. Schaap, P. Builtjes, L. Tarrason, J. Douros, et al., 2007: Analysis of model responses to emission-reduction scenarios within the CityDelta project. *Atmospheric Environment*, **41**, 208–220.

Tilmes, S., J. Brandt, F. FlatØy, R. Bergström, J. Flemming, J. Langner, J. Christensen, L. Frohn, Ø. Hov, I. Jacobsen, et al., 2002: Comparison of five

Eulerian air pollution forecasting systems for the summer of 1999 using the German ozone monitoring data. *Journal of Atmospheric Chemistry*, **42**, 91–121.

Troen, I. and L. Mahrt, 1986: A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. *Boundary-Layer Meteorology*, **37**, 129–148.

Turquety, S., D. Hurtmans, J. Hadji-Lazaro, P. Coheur, C. Clerbaux, D. Josset, and C. Tsamalis, 2009: Tracking the emission and transport of pollution from wildfires using the IASI CO retrievals: analysis of the summer 2007 Greek fires. *Atmospheric Chemistry & Physics*, **9**, 4897–4913.

Vautard, R., M. Beekmann, J. Roux, and D. Gombert, 2001: Validation of a hybrid forecasting system for the ozone concentrations over the Paris area. *Atmospheric Environment*, **35**, 2449–2461.

Vautard, R., D. Martin, M. Beekmann, P. Drobinski, R. Friedrich, A. Jaubertie, D. Kley, M. Lattuati, P. Moral, B. Neininger, et al., 2003: Paris emission inventory diagnostics from ESQUIF airborne measurements and a chemistry transport model. *Journal of Geophysical Research*, **108**, 8564.

Verwer, J., 1994: Gauss-seidel iterations for stiff odes from chemical kinetics. *SIAM Journal of Scientific Computing*, **15**, 1243 – 1250.

Vestreng, V., M. Adams, and J. Goodwin, 2004: *Inventory Review 2004, Emission Data reported to CLRTAP and under the NEC Directive.*

Wayne, R., 2000: *Chemistry of atmospheres*, Oxford University Press.

Wesely, M., 1989: Parameterization of surface resistances to gaseous dry deposition in regional scale numerical models. *Atmospheric Environment*, **23**, 1293–1304.

WHO, 2006: Health risks of particulate matter from long-range transboundary pollution. *Regional Office for Europe, World Health Org., Copenhagen.*

Wu, L., V. Mallet, M. Bocquet, and B. Sportisse, 2008: A comparison study of data assimilation algorithms for ozone forecasts. *Journal of Geophysical Research*, **113**.

Zampieri, M., P. Malguzzi, and A. Buzzi, 2005: Sensitivity of quantitative precipitation forecasts to boundary layer parameterization: a flash flood case study in the Western Mediterranean. *Natural Hazards and Earth System Sciences*, **5**, 603–612.

Zanini, G., F. Monforti, P. Ornelli, G. Vialetto, G. Brusasca, G. Calori, G. Finardi, and C. Silibello, 2004: The MINNI project. *9th Conference on Harmonization within Atmospheric Dispersion Modelling for Regulatory Purposes, 1-4 June 2004, Garmisch-Partenkirchen, Germany.*

Zhang, Y., 2008: Online-coupled meteorology and chemistry models: history, current status, and outlook. *Atmospheric Chemistry and Physics*, **8**, 2895–2932.

Zhang, Y., C. Bischof, R. Easter, and P. Wu, 2005: Sensitivity analysis of photochemical indicators for O3 chemistry using automatic differentiation. *Journal of Atmospheric Chemistry*, **51**, 1–41.