
TROPOSPHERIC CHEMICAL STATE ESTIMATION
BY FOUR-DIMENSIONAL VARIATIONAL
DATA ASSIMILATION ON NESTED GRIDS

I N A U G U R A L – D I S S E R T A T I O N
ZUR
ERLANGUNG DES DOKTORGRADES
DER MATHEMATISCH–NATURWISSENSCHAFTLICHEN FAKULTÄT
DER UNIVERSITÄT ZU KÖLN

vorgelegt von

Achim Strunk

aus Bad Marienberg

KÖLN, 2006

Berichterstatter: Prof. Dr. A. Ebel
Prof. Dr. J. Saur

Tag der mündlichen Prüfung: 12. Januar 2007

Abstract

The University of Cologne chemistry transport model EURAD and its four-dimensional variational data assimilation implementation is applied to a suite of measurement campaigns for analysing optimal chemical state evolution and flux estimates by inversion. In BERLIOZ and VERTIKO, interest is placed on atmospheric boundary layer processes, while for CONTRACE and SPURT upper troposphere and tropopause height levels are focussed. In order to achieve a high analysis skill, some new key features needed to be developed and added to the model setup. The spatial spreading of introduced observational information can now be conducted by means of a generalised background error covariance matrix. It has been made available as a flexible operator, allowing for anisotropic and inhomogeneous correlations. To estimate surface fluxes with high precision, the facility of emission rate optimisation using scaling factors is extended by a tailored error covariance matrix. Additionally, using these covariance matrices, a suitable preconditioning of the optimisation problem is made available. Furthermore, a module of adjoint nesting was developed and implemented, which enables the system to operate from the regional down to the local scale. The data flow from mother to daughter grid permits to accomplish nested simulations with both optimised boundary and initial values and emission rates. This facilitates to analyse constituents with strong spatial gradients, which have not been amenable to inversion yet. Finally, an observation operator is implemented to get to assimilate heterogeneous sources of information like ground-based measurements, airplane measuring data, Lidar and tethered balloon soundings, as well as retrieval products of satellite observations. In general, quality control of the assimilation procedure is obtained by comparison with independent observations. The case study analyses show considerable improvement of the forecast quality both by the joint optimisation of initial values and emission rates and by the increase of the horizontal resolutions by means of nesting. Moreover, simulation results for the two airplane campaigns exhibit outstanding characteristics of the assimilation system also in the middle and upper troposphere region.

Kurzzusammenfassung

Das Chemie-Transport-Modell EURAD der Universität zu Köln und sein vierdimensionales variationelles Datenassimilationsmodul wird zur Analyse optimaler chemischer Zustände und Flüsse auf eine Reihe unterschiedlicher Messkampagnen in der Troposphäre angewandt. Im Mittelpunkt der Simulationen für BERLIOZ und VERTIKO stehen Prozesse der atmosphärischen Grenzschicht, wohingegen im Rahmen von CONTRACE und SPURT die obere Troposphäre und die Tropopausenregion untersucht wird. Um bestmögliche chemische Zustandsanalysen zu erhalten, mussten verschiedene neue Schlüsselfunktionen entwickelt und dem Modellsystem hinzugefügt werden. Zum einen soll der Informationsgehalt weniger Messungen effizient genutzt werden, was durch die Parametrisierung der Hintergrundfehler-Kovarianzmatrix in Form eines Operators ermöglicht werden konnte. Zusätzlich können so flexibel anisotrope und inhomogene Korrelationen simuliert werden. Zur genauen Bestimmung von Flüssen wurde dem Optimierungsmodus für Emissionsraten mit skalierenden Faktoren eine angepasste Fehler-Kovarianzmatrix hinzugefügt. Die Kovarianzmatrizen werden ferner dazu verwendet, das Optimierungsproblem geschickt zu präkonditionieren. Eine fundamentale Neuentwicklung ist das Modul des adjungierten Nestens, wodurch das Assimilationssystem jetzt von der regionalen bis zur lokalen Skala operabel ist. Der Datenfluss von Mutter- zu Tochtergitter ermöglicht es, genestete Simulationen sowohl mit optimierten Rand- und Anfangswerten als auch mit verbesserten Emissionsraten durchzuführen. Durch diese Neuerung werden Spurengase mit hoher räumlicher Variabilität der Analyse zugänglich, die zuvor nicht analysiert werden konnten. Schließlich wurden leistungsfähige Beobachtungsoperatoren implementiert, die die bestmögliche, gemeinsame Nutzung unterschiedlichster Informationsquellen wie erdgestützte Messungen, Flugzeugbeobachtungen, Lidarsondierungen und Ballonaufstiege, sowie Erdbeobachtungen aus Satellitenmessungen ermöglicht. Die Qualität der bei den Fallstudien ermittelten Analysezustände wird generell durch den Vergleich mit unabhängigen Messungen ermittelt. Die Simulationen im Rahmen der Fallstudien zeigen eine starke Verbesserung der Vorhersagegüte sowohl durch die gemeinsame Optimierung von Anfangszuständen und Emissionsfaktoren als auch durch die Erhöhung der horizontalen Gitterauflösung mittels Nestens. Darüberhinaus stellen die Simulationen für die obere Troposphäre die herausragenden Fähigkeiten des Assimilationssystems unter Beweis.

Contents

1	Introduction	1
2	Data Assimilation	7
2.1	Objective Analysis using Least-Squares Estimation	8
2.1.1	Definition of errors	8
2.1.2	Minimum variance and maximum likelihood	10
2.1.3	Conditional probabilities	13
2.2	Optimal Interpolation and Kalman Filtering	14
2.3	4D-Var Data Assimilation	16
2.4	4D-Var vs. KF	22
3	The EURAD 4D-Var Data Assimilation System	23
3.1	The EURAD-CTM and its Adjoint	25
3.2	Minimisation Algorithm: L-BFGS	29
3.3	The Rosenbrock solver	30
3.4	Optimisation of Initial Values and Emission Factors	31
3.5	Adjoint Nesting Technique	36
3.6	Observation Operator	40

4	Covariance Modelling	43
4.1	Incremental formulation	45
4.2	Initial Value Background Error Covariance Matrix	46
4.3	Emission Factor Background Error Covariance Matrix	50
4.4	Observation Error Covariance Matrix	53
5	Campaign Simulations	57
5.1	BERLIOZ Campaign	59
5.1.1	Model design and observational data basis	59
5.1.2	Simulation Results	62
5.2	VERTIKO Campaign	74
5.2.1	Model design and observational data basis	74
5.2.2	Simulation Results	76
5.3	CONTRACE I Campaign	95
5.3.1	Model design and observational data basis	95
5.3.2	Simulation Results	100
5.4	SPURT IOP 2 Campaign	107
5.4.1	Model design and observational data basis	107
5.4.2	Simulation Results	108
6	Summary and Discussion of Results	115
A	Vertical Grid Structure	119
B	Error Specifications	123
C	Definition of Statistical Measures	127
	Bibliography	129
	Acknowledgements	137

List of Figures

2.1	Example for a minimum variance/maximum likelihood estimate.	12
2.2	Principle of sequential data assimilation algorithms.	15
2.3	Principle of 4d-var data assimilation algorithm.	21
3.1	Flowchart of the EURAD 4d-var data assimilation system. . .	24
3.2	Domain decomposition strategy.	28
3.3	Steepest descent and Newton minimisation method.	29
3.4	Controllability of model parameters in tropospheric chemical data assimilation.	31
3.5	The principle benefit of joint optimising initial values and emission rates.	33
3.6	Definition of emission factors subject to optimisation.	34
3.7	Portrayal of the nesting strategy in the EURAD 4d-var data assimilation system.	38
3.8	Initial ozone mixing ratios on a nested domain, according to the initialisation strategy.	39
3.9	Strategy to calculate a model equivalent for an observation coming with a footprint information.	41
3.10	Examples of satellite based measurements coming with footprint information.	42

4.1	Parametrisation of horizontal influence radii and their impact on the analysis result.	49
4.2	Background emission factors error correlation matrix.	52
4.3	Impact of averaging high resolving air-borne measurements.	54
5.1	BERLIOZ: Nested model grid sequence.	60
5.2	BERLIOZ CG: Footprints of NNORSY GOME ozone profile retrievals.	63
5.3	BERLIOZ CG: Assimilation results for GOME NO ₂ tropospheric columns.	64
5.4	BERLIOZ CG: Normalised cost function values for ground based observations with respect to assimilated species and optimisation mode.	66
5.5	BERLIOZ CG: Observations and model realisations in grid cell (43,39,1).	67
5.6	BERLIOZ N1: Observations and model realisations in grid cells (40,40,1) and (40,41,1).	69
5.7	BERLIOZ N2: Observations and model realisation for nitrogen oxides at selected stations.	70
5.8	BERLIOZ N2: Observations and model realisation for hydrogen peroxide, sulphur dioxide and formaldehyde at selected stations.	71
5.9	BERLIOZ N2: Optimised emission factors.	72
5.10	BERLIOZ N2: Surface ozone mixing ratios on July 20.	73
5.11	VERTIKO: Nested model grid sequence.	75
5.12	VERTIKO CG: Partial costs by assimilated GOME O ₃ profile retrievals.	77
5.13	VERTIKO CG: Comparison of NO ₂ column retrievals from GOME and SCIAMACHY with model equivalents.	78
5.14	VERTIKO CG: Temporal evolution of the cost function for ground based observations.	82
5.15	VERTIKO N1: Temporal evolution of cost function for ground based observations.	83
5.16	VERTIKO: Hourly averaged bias and rms.	86
5.17	VERTIKO N2: Bias and rms for NH ₃ campaign data from June 15 to 19, 2003.	93

5.18	VERTIKO N2: Optimised NH ₃ emission factors for June 18, 2003.	94
5.19	Satellite image of Europe and the North-Atlantic on November 14 and 15, 2001.	96
5.20	CONTRACE: Model grid sequence.	97
5.21	CONTRACE CG: Partial costs by GOME O ₃ profile retrievals.	98
5.22	CONTRACE CG: O ₃ profile retrievals by NNORSY from GOME observations on November 13, 2001.	99
5.23	CONTRACE CG: CO mixing ratios during November 14 and 15, 2001.	101
5.24	CONTRACE CG: NO mixing ratios on November 14 and 15, 2001.	102
5.25	CONTRACE N1: Observational campaign data and MOZAIC data for O ₃ and CO on November 14, 2001.	103
5.26	CONTRACE N1: Observational campaign data and MOZAIC data for NO, NO _y and H ₂ O ₂ on November 14, 2001.	104
5.27	CONTRACE N1: Partial costs by the campaign data.	105
5.28	SPURT: Flight paths of the IOP2 campaign data and available MOZAIC data.	108
5.29	SPURT: Averaged observations and model simulation results for January 18, 2002.	109
5.30	SPURT: Averaged observations and model simulation results for January 19, 2002.	110
5.31	SPURT: Averaged MOZAIC observations and model simulation results.	111
5.32	SPURT: Impact of campaign data on model equivalents for GOME O ₃ profile retrievals.	112

List of Tables

2.1	Definition of model variable notation.	9
4.1	Standard deviations of the transformed emission factors. . .	50
5.1	BERLIOZ: Simulation setup.	59
5.2	VERTIKO: Simulation setup.	74
5.3	VERTIKO: Bias and rms for satellite column retrievals. . . .	77
5.4	VERTIKO: Bias and rms statistics for ground based stations.	85
5.5	CONTRACE: Simulation setup.	95
5.6	CONTRACE: Bias and rms statistics.	98
5.7	SPURT: Simulation setup.	107
5.8	SPURT: Bias and rms error for GOME O ₃ satellite profiles.	113
A.1	Vertical grid structure for the 23 layer design.	119
A.2	Vertical grid structure for the 26 layer design.	120
B.1	Minimum estimates for the absolute background errors. . . .	123
B.2	Absolute and relative measurement errors for ground based stations.	124
B.3	Absolute representativeness error estimates at surface level. .	125
B.4	Representativeness lengths for ground based observations. . .	125

List of Acronyms

4d-var	Four-Dimensional Variational
AK	Averaging Kernel
BERLIOZ	Berliner Ozon Experiment
BLUE	Best Linear Unbiased Estimate
CONTRACE	Convective Transport of Trace Gases into the Upper Troposphere over Europe: Budget and Impact on Chemistry
CTM	Chemistry Transport Model
DLR	Deutsches Zentrum für Luft- und Raumfahrt
DOAS	Differential Optical Absorption Spectroscopy
ECMWF	European Centre for Medium-Range Weather Forecasts
EEA	European Environment Agency
EEM	EURAD Emission Model
EMEP	Co-operative programme for monitoring and evaluation of the long-range transmissions of air pollutants in Europe
EURAD	European Air Pollution and Dispersion [model]
GMES	Global and regional Earth-system Monitoring using Satellite and in-situ data
GOME	Global Ozone Mapping Experiment
IOP	Intensive Observational Period

KF	Kalman Filter
KNMI	Koninklijk Nederlands Meteorologisch Instituut
KPP	Kinetic PreProcessor
L-BFGS	Limited-Memory Broyden Fletcher Goldfarb Shanno
MM5	Fifth-Generation NCAR / Penn State Mesoscale Model
MOZAIC	Measurements of Ozone and Water Vapour by in-service Airbus Aircraft
NCAR	National Center for Atmospheric Research
NNORSY	Neural Network Ozone Retrieval System
PROMOTE	Protocol Monitoring for the GMES Service Element
RACM	Regional Atmospheric Chemistry Mechanism
RADM	Regional Acid Deposition Model
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography
SPURT	Trace Gas Transport in the Tropopause Region
UTC	Coordinated Universal Time
UT/LS	Upper Troposphere/Lower Stratosphere
VERTIKO	Vertical Transport of Energy and Trace Gases at Anchor Stations and Their Spatial/Temporal Extrapolation under Complex Natural Conditions
VOC	Volatile Organic Compounds

CHAPTER 1

Introduction

Studying atmospheric processes has always been motivated by the desire to predict future system states. *Bjerknes* [1911] formulated such a forecast as a deterministic initial value problem: "The solution of a system of differential equations with concrete initial values can lead to a weather forecast. But as a matter of fact, a good prediction always needs an exact guess of the atmosphere's actual state." Until the 70ies of the last century, observations scattered in time have been used to draw synoptical maps. Not before the advent of numerical models and their computational efficiency, weather forecasts relied on subjective analyses of the current physical state of the atmosphere. When suitable error characterisations or probability density functions of the available information have been specified, this allows for an objective analysis in terms of the most probable system state. The technique of *statistical interpolation* (*Eliassen* [1954], *Gandin* [1963]) was more and more used in all fields of geophysical interest. It is, *e. g.*, known as *Gauss-Markov* algorithm in oceanography and as *Kriging* method in applied geophysics. Subsequently, objective analysis by *data assimilation* gained acceptance in meteorology (*Daley* [1991]).

The observational basis in meteorology as well as in atmospheric chemistry changed substantially within the last decades. In addition to enhanced ground based observational networks with a denser set of measurement stations, satellite based remote sensing data is now available at high resolution levels and with remarkable global coverage.

Hence, new observation techniques favoured the use of spatio-temporal techniques as the four-dimensional variational (4d-var) algorithm (*e. g.*, Courtier [1997]) and the Kalman Filter (Kalman [1960], Kalman and Bucy [1962]), in order to fully use the information contained in the heterogeneous observation types. In contrast to sequentially employed spatial algorithms, which do not use physical or chemical constraints to join the scattered information, spatio-temporal algorithms allow the fusion of both heterogeneous observations and the physical and chemical laws coded in numerical models. These algorithms are used to retrieve analyses, that lead to consistent system evolutions and that are available on regular grids. In the field of meteorology, the 4d-var method entered operational weather forecasting in the nineties. The most prominent example is the first operational 4d-var implementation at the ECMWF in 1997 (Rabier *et al.* [2000]), while all meteorological centres thereafter experienced substantial benefits from going to 4d-var (Rabier [2005]).

During the last decade much progress has been attained in developing comprehensive chemistry transport models for both the stratosphere and the troposphere. In concert with rapidly increasing computer power and laboratory results on new chemical reaction paths, spatially highly resolved model simulations with sophisticated chemistry kernels and promising forecast behaviour over longer periods have become available. In addition to user-oriented air quality forecast services¹, long-term model simulations are used for episodic assessments of suitable political regulations to enhance the air quality in Europe (Memmesheimer *et al.* [1996], Memmesheimer *et al.* [2004b])².

Moreover, spatio-temporal data assimilation algorithms have been introduced in the field of atmospheric chemistry, exhibiting a development comparable to the development in meteorology, with 20 years delay, however. The usefulness of the variational approach for atmospheric chemistry has been shown by Fisher and Lary [1995] with box model simulations for the stratosphere. The spatio-temporal extension with the 4d-var technique for the full tropospheric chemistry kernel RADM2 has been applied by Elbern *et al.* [1997]. Elbern and Schmidt [1999] extended the regional EURAD model to the first available complete adjoint chemistry transport model for the troposphere, including parallelisation, pre-conditioning and off-diagonal covariance modelling. Further, the real world applicability for an ozone case study has been shown in Elbern and Schmidt [2001].

The application of advanced data assimilation techniques using complex

¹Visit, *e. g.*, www.gse-promote.org/services/airquality.html for a short overview.

²For a list of German clean air plans surf to www.umweltbundesamt.de/luft/index.htm.

chemistry transport models has only become feasible since the last decade, due to several reasons: Firstly, realistic chemical systems include at least ten times more control parameters (chemical species, see *Stockwell et al.* [1990] or *Sander et al.* [2005]) than those used in meteorology or oceanography. This influences the complexity of statistics concerning covariance models as well as the demands in computer power. Secondly, there is a huge imbalance between available measurements and model parameters. The number of observations is usually far below the number of parameters in a chemistry transport model, such that the system is grossly underdetermined (it may as well be locally overdetermined). Additionally, there is not necessarily a straight forward model equivalent to each kind of observation due to the often lumped formulation of the chemical reaction schemes or due to some methods of measuring, where classes of species are observed as single quantity.

Another reason is the more complicated way in which the forecast skill of the atmospheric chemistry is influenced by initial values. At least in the troposphere, there might be several other processes which have a significant impact on the temporal evolution of the system like emission rates and deposition velocities. Especially emission rates have a strong influence on the trace gas concentrations, at most in the vicinity of the sources. Moreover, hourly emission rates are calculated on the basis of annually emitted amounts, given on a presumably different inventory grid, and thus are subject to huge uncertainties. However, due to a possibly strong exchange between the atmospheric boundary layer and the free troposphere, the emissions can strongly influence large model areas by long-range transport. Hence, a number of source and sink estimation assessments have been conducted: *Kaminski et al.* [1999] estimated CO₂ fluxes using a 4d-var implementation in the TM2 model and global observations. The first 4d-var implementation for a complete set of emitted species was introduced by *Elbern et al.* [2000], who additionally showed the applicability with identical twin experiments for the EURAD model. *Quélo et al.* [2005] used a one year comparison of observations and simulations with the POLAIR3D model to retrieve NO_x emission time distribution, which exhibits to be robust and which significantly improved ozone forecast skills. The combination of emission rate and chemical state optimisation of precursor trace gases with 4d-var by observations of product species and a systematic assessment of improved forecast quality by observations withheld from the assimilation procedure is shown in *Elbern et al.* [2006] for an ozone episode. Hence, best estimates also for surface fluxes are indispensable for good predictions of the chemical composition.

However, despite the attained complexity and the sophistication level of the 4d-var data assimilation algorithm, the following problems still prevent the achievement of high quality tropospheric model simulations:

(i) Fine model resolutions are needed in order to predict primary emitted species like nitrogen oxides, which exhibit strong spatial gradients. Otherwise, for example, the ozone chemistry simulations may reside in a wrong chemical regime. Moreover, enhanced model resolution will significantly reduce biases in the representativeness of assimilated observations for the grid cells. The favourable option to increase the model resolution are nested sub-domains, which, as a by-product, lead to optimised boundary values in terms of analysed model states from the respective mother domain.

To this end, an adjoint nesting approach has been developed in this work, allowing to telescope from regional scale down to local scale. The benefit of this facility for forecast skills and the analysis quality is addressed in boundary layer campaign simulations.

(ii) One of the most crucial aspects of spatio-temporal data assimilation is the treatment of background error covariances. Both the estimated variances of the background fields as well as their correlations manifest the quality of an analysis, at most in data void regions. By reason of the size of one model state vector, the background error covariance matrix for comprehensive chemistry transport models cannot be treated numerically in a straight forward way. One option is to provide it as an operator instead of its matrix representation. Further, due to surface textures and corresponding correlation features in the background state, an inhomogeneous and anisotropic design is eligible.

To this end, a diffusion operator is employed in this work to construct the correlations in the initial value background error covariance matrix. This enables to flexibly model covariances in the background fields and can be easily extended to represent anisotropic and inhomogeneous correlations. Additionally, for the optimisation of emission rates, a background error covariance matrix has been coded from scratch, allowing to model reasonable correlations between different emitted trace gases and to reduce the degrees of freedom in the analysis procedure. Both covariance matrices now admit to simultaneously optimise the initial model state and emission rates in a balanced and suitably preconditioned way.

(iii) To allow for the use of all available information during a data assimilation procedure in the best possible manner, a suitable observation operator needed to be constructed. This facilitates the confrontation of a model simulation with remote sensing data like tropospheric column and profile retrievals from satellite observations, lidar and tethered balloon soundings,

as well as air-borne and ground based in-situ data. The use of air-borne and satellite measurements is discussed for a data assimilation suite of two air-borne campaigns, focussing on the free and upper troposphere.

Hence, the main intention of this work is to demonstrate the feasibility of sophisticated spatio-temporal data assimilation for the modelling of tropospheric chemistry from mesoscale down to local scale and its favourable impact on the forecast skill. The chosen algorithm is the 4d-var method, which is introduced and shortly compared to the Kalman Filter in Chapter 2. The extensions and improvements of the EURAD 4d-var data assimilation system developed in this thesis — namely the adjoint nesting technique, the technique of a balanced joint estimation of initial chemical state and emission rates, an improved chemistry solver with an updated mechanism, and the observation operator — will be shown in Chapter 3 in the framework of a concise system description. The modelling of the variances and covariance matrices involved in the assimilation procedure can be found in Chapter 4. The complete model system is then applied to four campaign data sets, in slightly different configuration with varying subject matter. The results of which are detailed in Chapter 5. Finally, Chapter 6 summarises and discusses the results of this thesis.

CHAPTER 2

Data Assimilation

The main objective of data assimilation is to provide an as accurate as possible and consistent image of a system's state at a given time (the optimal guess) by taking into account all available information about the system. This information may consist of

1. observations providing samples of the system's current state,
2. the intrinsic laws of the system's evolution in space and time, coded in a numerical prognostic model, and
3. a previous estimate of the true state (encompassing all information gathered in the past), usually referred to as the *background knowledge*.

Combining all these types of information, the observed knowledge may propagate to all state variables of the model, while the model provides for the constraints on the propagation within the system. Hence, the innovation can accumulate in time (*Bouttier and Courtier* [1999]), being the core principle of all model based data assimilation techniques. It is however affected by errors of often poorly known quantities (*e. g.*, biases, discretisation errors). Observations usually vary in nature and accuracy and are strongly irregular distributed in space and time. Concerning atmospheric chemistry, only a few system's instances are observed (*i. e.*, chemical species) and often there is a lack of validated error statistics. So none of this available information is exactly true. We cannot "trust" completely neither the model,

nor the observations, nor the a-priori information (which may consist of climatological statistics, the result of a previous analysis, a model forecast or combinations of those), since they are affected by uncertainties. This leads to the formal requirement to take into account the data errors (their probability distribution functions) such that the resulting analysis error is minimal. In some suitable sense this will allow us to write the analysis problem as an optimisation problem. The following section will present this in the framework of least-squares estimation.

For a detailed classifications of data assimilation techniques and the history of data assimilation in meteorology see *Lorenc* [1986], *Ghil* [1989], *Daley* [1991] and *Talagrand* [1997].

2.1 Objective Analysis using Least-Squares Estimation

To derive the most common data assimilation techniques we need first to assume some properties of the error variances of the input parameters. The error estimation and the generalisation to covariances will evolve throughout this work, namely when introducing the data assimilation algorithms in the next sections, in Chapter 4 and Appendix B. The notation given in Table 2.1 will be used, following as close as possible the suggestions given in *Ide et al.* [1997].

2.1.1 Definition of errors

The background errors or forecast errors

$$\boldsymbol{\varepsilon}^b = \boldsymbol{x}^b - \boldsymbol{x}^t \quad (2.1)$$

are the only terms that separate the background knowledge from the true state (*Bouttier and Courtier* [1999]). They are the estimation errors of the background state, which is the best model based guess of the system's current state. The errors represent both the uncertainties in the up-to-now knowledge and the forecast errors, if a model integration was used to produce the background model state.

The observation errors

$$\begin{aligned} \boldsymbol{\varepsilon}^o &= \underbrace{(\boldsymbol{y}^t - H(\boldsymbol{x}^t))}_{\boldsymbol{\varepsilon}^r} + \underbrace{(\boldsymbol{y} - \boldsymbol{y}^t)}_{\boldsymbol{\varepsilon}^m} \\ &= \boldsymbol{y} - H(\boldsymbol{x}^t) \end{aligned} \quad (2.2)$$

include the errors due to the statistical performances of the measuring instruments ($\boldsymbol{\varepsilon}^m$) and an estimate of representativeness error ($\boldsymbol{\varepsilon}^r$). H is the *observation operator* or *forward interpolation operator* that generates the *model equivalents* to the observations in the case that both the observations and the model state vector were true and that there are no forecast errors. In practise H is a collection of interpolation operators from the model discretisation to the observational points and conversions from model variables to the observed parameters. Some remarks on the properties of H as implemented in the EURAD 4d-var system are given in Section 3.6.

The representativeness error reflects the degree to which an observation is able to represent the optimal model equivalent $H(\boldsymbol{x}^t)$, basing on volume-average model state values. Representativeness errors are usually taken to be unbiased, and their magnitudes may depend on the volume size, the sensor's time average, the volume size which one wants the observation to represent, and the chemical state of the atmosphere. This error has to be included in the observational error (Daley [1991]).

Table 2.1: Definition of model variable notation. The number of observations is denoted by p , while a vector of an initial model state and the emission factors is of size n and m , respectively.

Parameter	Definition	Dimension
\boldsymbol{x}^t	true model state	n
\boldsymbol{x}^b	background model state	n
\boldsymbol{x}^a	analysis state	n
\boldsymbol{y}	observations	p
\boldsymbol{y}^t	true observational state	p
H	observation operator	—
\mathbf{H}	linearised observation operator	$p \times n$
\mathbf{B}	background error covariance matrix (initial values)	$n \times n$
\mathbf{K}	background error covariance matrix (emission factors)	$m \times m$
\mathbf{O}	measurement error covariance matrix	$p \times p$
\mathbf{F}	representativeness error covariance matrix	$p \times p$
$\mathbf{R} = \mathbf{O} + \mathbf{F}$	observation error covariance matrix	$p \times p$
\mathbf{A}	analysis error covariance matrix	$n \times n$

The terms $[\mathbf{y} - H(\mathbf{x})]$ are called *departures* or the *innovation vector* and are the key to data analysis based on the discrepancies between observations and model states (*Parrish and Derber [1992]*).

We assume those errors to be unbiased,

$$\langle \boldsymbol{\varepsilon}^b \rangle = 0 \quad \text{and} \quad \langle \boldsymbol{\varepsilon}^o \rangle = 0, \quad (2.3)$$

where $\langle \rangle$ is the expectation operator. Additionally, we consider the observational and the background errors to be mutually uncorrelated, *i. e.*

$$\langle \boldsymbol{\varepsilon}^b \boldsymbol{\varepsilon}^o \rangle = 0. \quad (2.4)$$

The analysis errors

$$\boldsymbol{\varepsilon}^a = \mathbf{x}^a - \mathbf{x}^t \quad (2.5)$$

is what we want to find, such that the analysis is *optimal* in the sense that it is as close as possible to the true model state \mathbf{x}^t .

2.1.2 Minimum variance and maximum likelihood

What we are looking for is an optimal estimation for the analysis \mathbf{x}^a . Consider the problem of estimating a scalar parameter x with x^t being the true value, having p measurements y_i , $1 \leq i \leq p$ with measurement errors ε_i (see *Daley [1991]*). The expected error variances are denoted $\sigma_i^2 = \langle \varepsilon_i^2 \rangle$ and suppose the errors are unbiased and uncorrelated (as described above). The estimate x^e should be a linear composition of the observations:

$$x^e = \sum_i c_i y_i,$$

with $c_i \geq 0$ for $i = 1, \dots, p$ being the weights to be specified. ε^e denotes the error of the estimate. The bias then reads

$$\langle \varepsilon^e \rangle = \langle x^e \rangle - \langle x^t \rangle = \sum_i c_i \langle \varepsilon_i \rangle - \langle x^t \rangle \left[1 - \sum_i c_i \right].$$

We attain an unbiased linear estimate $\langle \varepsilon^e \rangle = 0$ if $\sum_i c_i = 1$. For the expected error variance we yield

$$\begin{aligned}
 \langle (\varepsilon^e)^2 \rangle &= \langle (x^e - x^t)^2 \rangle \\
 &= \left\langle \left(\sum_i c_i y_i - \sum_i c_i x^t \right)^2 \right\rangle \\
 &= \left\langle \left(\sum_i c_i \varepsilon_i \right)^2 \right\rangle \\
 &= \sum_i c_i^2 \sigma_i^2
 \end{aligned} \tag{2.6}$$

The *minimum variance estimate* is now defined to be the estimate x^a that minimises (2.6). This may be solved by the method of Lagrange's undetermined multipliers. The solution is characterised as follows: The minimum variance estimate or the optimal linear unbiased estimate (also called *best linear unbiased estimate* or **BLUE**) is given by

$$x^a = \frac{\sum_i \sigma_i^{-2} y_i}{\sum_i \sigma_i^{-2}}. \tag{2.7}$$

The minimum error variance reads

$$\langle (\varepsilon^a)^2 \rangle = \sum_i c_i^2 \sigma_i^2 = \frac{1}{\sum_i \sigma_i^{-2}}$$

or

$$\frac{1}{(\sigma^a)^2} = \sum_i \frac{1}{\sigma_i^2}.$$

Each additional observation thus reduces the expected error variance $\langle (\varepsilon^a)^2 \rangle$ of the minimum variance estimate. Therefore the expected error variance is always smaller than the least expected error variance of \mathbf{y} (see Fig. 2.1 for an example of $p = 2$ and Gaussian distributed errors).

For Gaussian distributed error probabilities

$$P(\varepsilon_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp \left[-\frac{\varepsilon_i^2}{2\sigma_i^2} \right]$$

the joint error probability of all p observations is given by

$$\begin{aligned}
 P(\varepsilon_1)P(\varepsilon_2) \cdots P(\varepsilon_p) &= \prod_i \frac{1}{\sigma_i \sqrt{2\pi}} \exp \left(-\frac{\varepsilon_i^2}{2\sigma_i^2} \right) \\
 &= \left[\prod_i \frac{1}{\sigma_i \sqrt{2\pi}} \right] \exp \left[-\sum_i \frac{(y_i - x^t)^2}{2\sigma_i^2} \right].
 \end{aligned} \tag{2.8}$$

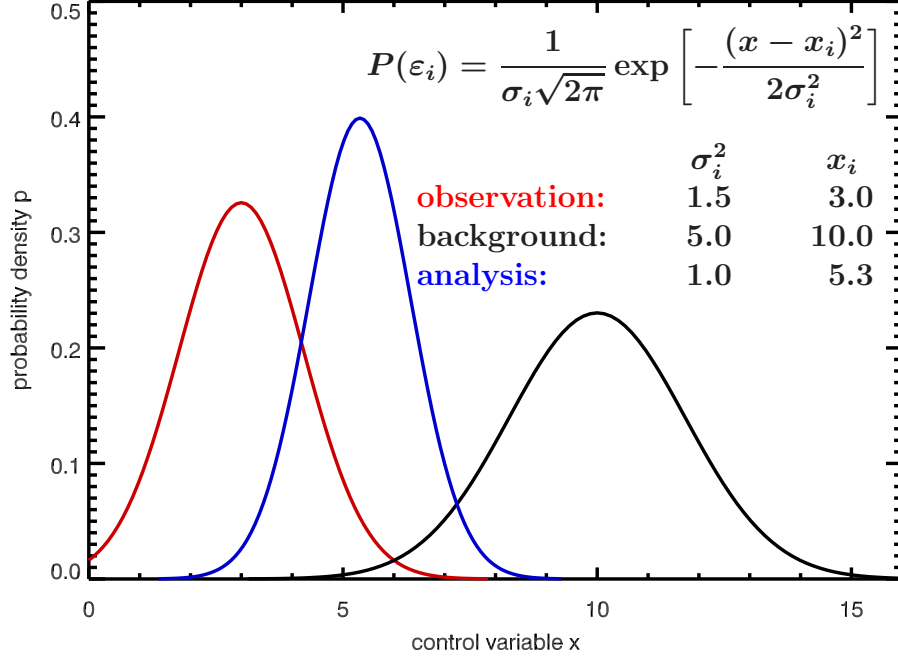


Figure 2.1: Example for a minimum variance/maximum likelihood estimate for $p = 2$ (the second "observation" is for future reference called background) and Gaussian probability distribution functions. The analysis distribution is the best linear unbiased estimate **BLUE**, which has a higher probability of the expectation value and a smaller variance than both observations (a sharper pdf).

The most probable value x^a is that one which maximises (2.8). This occurs when the argument to the exponential function reaches a minimum. The *maximum likelihood estimate* must minimise

$$J = \frac{1}{2} \sum_i \sigma_i^{-2} (x^a - y_i)^2. \quad (2.9)$$

Differentiating with respect to x^a and setting the result to zero leads to

$$x^a = \frac{\sum_i \sigma_i^{-2} y_i}{\sum_i \sigma_i^{-2}},$$

which is the same as (2.7). When error distribution is Gaussian, minimum variance and maximum likelihood estimates thus lead to the same optimal linear unbiased estimate (or **BLUE**).

The following sections discussing advanced data assimilation techniques will generalise (2.7) and (2.9) under different constraints.

2.1.3 Conditional probabilities

The analysis problem can be described in terms of conditional or *Bayesian* probabilities (Lorenz [1986]). These probabilities are based on the Theorem of Bayes defining the joint pdf of two events x and y

$$P(x|y)P(y) = P(y|x)P(x), \quad (2.10)$$

where $P(x)$ denotes the probability of an event x to be true and $P(x|y)$ the probability of the event x under the condition of event y to be true. This theorem is useful for cases in which there is an event y that can be observed while there is an event x that has to be determined.

Let us consider a set of observations $Y_j = \{y_0, y_1, \dots, y_j\}$ gathered in a time window $[0, t_j]$, $j = 0, \dots, N$ and a model state \mathbf{x} to be determined, conditional to the information Y_j . We then yield (Cohn [1997])

$$\begin{aligned} P(x|Y_j) &= P(x|y_j, Y_{j-1}) = \frac{P(x, y_j, Y_{j-1})}{P(y_j, Y_{j-1})} \\ &= \frac{P(y_j|x, Y_{j-1})P(x, Y_{j-1})}{P(y_j, Y_{j-1})} \\ &= \frac{P(y_j|x, Y_{j-1})P(x|Y_{j-1})P(Y_{j-1})}{P(y_j|Y_{j-1})P(Y_{j-1})} \\ &= \frac{P(y_j|x, Y_{j-1})P(x|Y_{j-1})}{P(y_j|Y_{j-1})}, \end{aligned} \quad (2.11)$$

which is only another version of Bayes' rule. This very general result can be used to derive the data assimilation techniques presented below under different constraints and assumptions, because it delivers a method to derive $P(x|Y_j)$ from $P(x|Y_{j-1})$; *i. e.*, the analysis update using a newly introduced information y_j and the background information pdf $P(x|Y_{j-1})$. It is both the basis of conditional mean estimation, which is identical to the minimum variance estimator, as well as of conditional mode estimation, resulting in the maximum likelihood estimator. In the former case, the result for x will be the expectation value of $P(x|Y_j)$ (under the assumption of its existence), where in the latter case the numerator of (2.11) is evaluated for a maximum, resulting in the most probable value of x . This will be done in Section 2.3 to derive the objective function describing the 4d-var method.

This fundamental probabilistic formalism of the analysis problem may be generalised to arbitrary probability distribution functions leading to the possibility to account for control variables which have strongly non-Gaussian error pdfs, like log-normal distributions. In some recent works this is accounted for (*e. g.*, Fletcher and Zupanski [2006]).

2.2 Optimal Interpolation and Kalman Filtering

This section very briefly introduces two important data assimilation methods that attempt to minimise analysis errors by minimum variance (or conditional mean estimator) algorithms.

Optimal interpolation — often referred to as statistical interpolation — is still widely used in atmospheric data assimilation. The method reads as follows:

With the notation defined in Section 2.1 we write the analysis increment ($\mathbf{x}^a - \mathbf{x}^b$) as a linear combination of the observation departures (or innovations),

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y} - H(\mathbf{x}^b)), \quad (2.12)$$

where the matrix \mathbf{K} contains the weights to be determined. Here the observations may be displaced to the model grid points, so that we need to calculate the model equivalent of the observations by applying the observation operator H to the model state. For convenience we consider the observation errors and the background errors to be unbiased and mutually uncorrelated (Eqs. (2.3)-(2.4)).

Under the condition, that the optimal weights \mathbf{K} are encountered, when the analysis error is minimal, the result can be shown to be (*e. g.*, Daley [1991])

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T [\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T]^{-1}, \quad (2.13)$$

where \mathbf{K} is referred to as the *gain matrix* and \mathbf{H} is a linear observation operator, linearised in the vicinity of the background model state \mathbf{x}^b , such that

$$H(\mathbf{x}) - H(\mathbf{x}^b) = \mathbf{H}(\mathbf{x} - \mathbf{x}^b) + \mathcal{O}[(\mathbf{x} - \mathbf{x}^b)^2]$$

is satisfied (Bouttier and Courtier [1999]). The algorithm (2.12) is only optimal if the covariance matrices \mathbf{B} and \mathbf{R} and the forward operator \mathbf{H} are correct.

Until the late nineties Optimal Interpolation was the most common used data assimilation technique — at least in operational weather forecast and chemistry transport models (Talagrand [1997]). It was well suited for the demands of operational forecasting because the implementation of (2.12) only needed the inversion of $(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)$. The algorithm can be performed in a *sequential* mode, which means that the analysis procedure is performed each time a new observation is available. The result of the analysis step is taken as initial state for a model forecast until the next analysis time. This

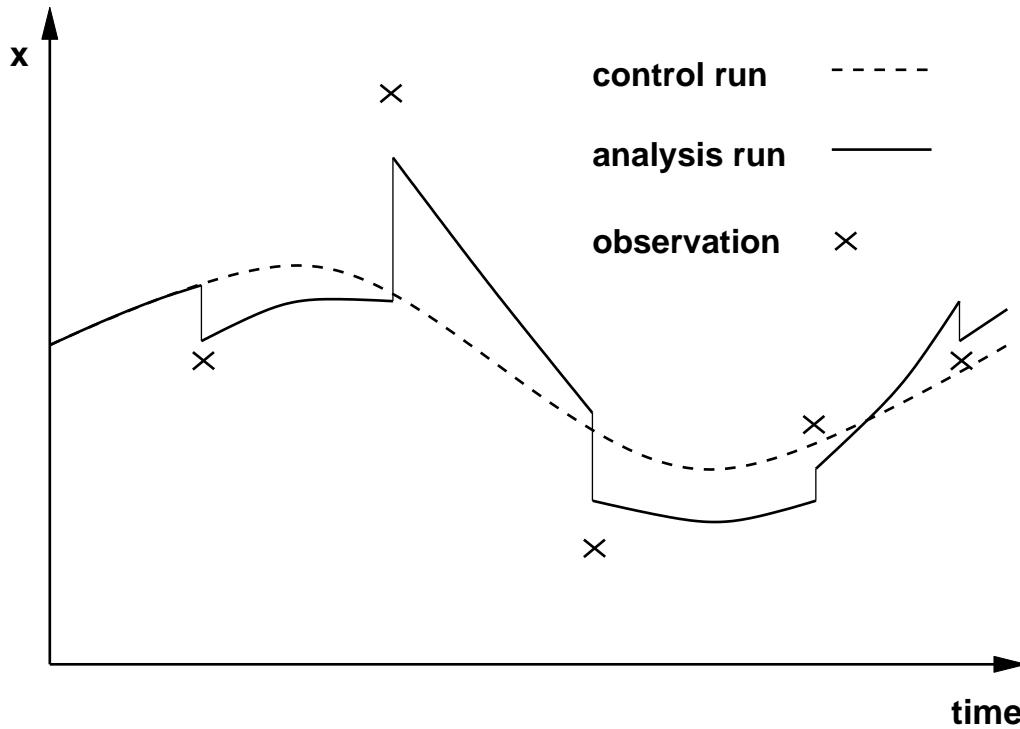


Figure 2.2: Principle of sequential data assimilation algorithms, shown for one single model parameter. At each analysis instance the model state is updated by the newly introduced observations. If there are only a few instances of the system measured, the successive model integration forward in time relaxes the usually distorted model state towards the control run, due to the chemically coupled evolution of the whole system.

scheme is depicted in Fig. 2.2 for a single model variable. This primitive sketch already shows some of the properties of statistical interpolation: At each analysis instance (here the observational times) the model state is corrected to the **BLUE** analysis state given by the observation and the a-priori model state. This usually leads to an analysis model state which is (chemically) distorted and the correction often vanishes after a short successive forecast period — *i. e.*, the model trajectory relaxes towards the background state.

Let us now introduce a prognostic model description by

$$\frac{d\mathbf{x}}{dt} = \mathbf{M}(\mathbf{x}) + \boldsymbol{\eta} \quad (2.14)$$

where \mathbf{M} is a linear model describing the time dependence of the model state, being itself independent from model state \mathbf{x} and having errors $\boldsymbol{\eta}$

with covariances \mathbf{Q} . Having an analysis state \mathbf{x}_{j-1}^a at time $t = t_{j-1}$ for $j = 1, \dots, N$, we can write the forecast step $t_{j-1} \rightarrow t_j$ as

$$\mathbf{x}_j^f = \mathbf{M}_j \mathbf{x}_{j-1}^a + \boldsymbol{\eta}, \quad (2.15)$$

where \mathbf{M}_j denotes the model operator propagating the model state from \mathbf{x}_{j-1} to \mathbf{x}_j . With the same assumptions as in (2.12) we can derive the related forecast error covariance matrix \mathbf{P}_j^f to (see *Daley [1991]*)

$$\mathbf{P}_j^f = \mathbf{M}_j \mathbf{P}_{j-1}^a \mathbf{M}_j^T + \mathbf{Q}_{j-1}. \quad (2.16)$$

The analysed variables evolve from the forecast by the following formulas

$$\mathbf{x}_j^a = \mathbf{x}_j^f + \mathbf{K}_j \mathbf{d}_j \quad (2.17)$$

$$\mathbf{P}_j^a = (\mathbf{I} - \mathbf{K}_j \mathbf{H}_j) \mathbf{P}_j^f, \quad (2.18)$$

where $\mathbf{d}_j = (\mathbf{y} - \mathbf{H}\mathbf{x}_j^f)$ denote the innovation vector and

$$\mathbf{K}_j = \mathbf{P}_j^f \mathbf{H}_j^T (\mathbf{H}_j \mathbf{P}_j^f \mathbf{H}_j^T + \mathbf{R}_j)^{-1} \quad (2.19)$$

is the *Kalman gain matrix*. Eqs. (2.15-2.18) define the discrete *Kalman Filter* (*Kalman [1960]*, *Kalman and Bucy [1962]*, *Jazwinski [1970]*), which is the minimum variance solution to the analysis problem under the condition that both the model and the observation operator are linear. This algorithm exploits all available information and successively updates the analysis error covariances (and thus delivers the background error covariances for the next step). This part of the Kalman Filter is the by far most demanding procedure, because (2.16) involves large matrix multiplications. Therefore, the application of the Kalman Filter in concert with large numerical models is still impossible (*Loon and Heemink [1997]*, *Bouttier and Courtier [1999]*, *Constantinescu et al. [2006]*). It is mostly applied in sub-optimal versions, *e.g.* the Ensemble Kalman Filter (EnKF, *Houtekamer et al. [2005]*) or the Reduced Rank Square Root Kalman Filter (RRSQRT, *Heemink et al. [2001]*). For an elementary comparison of the Kalman Filter and the 4d-var data assimilation technique see Section 2.4.

2.3 4D-Var Data Assimilation

Using variational data assimilation techniques, the analysis problem is formulated as a minimisation problem of the variational calculus (see, *e.g.*, *Courant and Hilbert [1953]*). A *objective* or *cost function* is defined, measuring the distance between a model simulation and the observations (see

Eq. (2.9)). To derive the cost function from (2.11), we need to define both the properties of the mapping between the model phase space and the observational phase space (*i. e.*, the forward model H) and of the prognostic model M itself. We want to use the model as a strong constraint to the optimisation problem. Therefore the model is regarded to be perfect and the error $\boldsymbol{\eta}$ vanishes in (2.14). Further, we assume that there are no stochastic processes influencing the model development in time, which makes the model deterministic. This leads to

$$\forall j, \quad \boldsymbol{x}_j^t = M_j \boldsymbol{x}_{j-1}^t, \quad (2.20)$$

where \boldsymbol{x}_j^t is the true model state at time $t = t_j$ and M_j a predefined non-linear model operator performing the integration from time t_{j-1} to t_j . Hence we can describe the model evolution from the initial state \boldsymbol{x}_0 to the state \boldsymbol{x}_j by recurrence,

$$\boldsymbol{x}_j = M_j M_{j-1} \dots M_2 M_1 \boldsymbol{x}_0. \quad (2.21)$$

Under the assumption, that a current observation y_j is independent from the set of previous observations Y_{j-1} and with the properties of the prognostic model given above, we can rewrite (2.11) as

$$P(x|Y_j) \propto P(x|Y_{j-1})P(y_j|x), \quad (2.22)$$

where $P(x|Y_{j-1})$ is the prior pdf, $P(y_j|x)$ is the likelihood of observation y_j and $P(x|Y_j)$ is the posterior pdf.

In Section 3.4 there will be a brief introduction to the kind of model parameter that is amenable to variational data assimilation under a certain observational basis. Moreover, the formulation of the cost function is devised for the combination of initial values and emission rates as control parameters. For the sake of convenience we will now assume, that the initial state of the model integration \boldsymbol{x}_0 is the only set of parameters which is subject to the minimisation procedure — *i. e.*, the set of parameters to be analysed. The background knowledge is denoted \boldsymbol{x}^b and \boldsymbol{y} is a set of observations, being distributed in space and time and vector valued for each time instance. Considering the first term on the right side of (2.22) constituting the background model state based on the information available before the time instance for which the new information y_j is introduced, and assuming that Gaussian distribution functions sufficiently describe the properties of

both the background errors and the observation errors, we yield

$$P(\mathbf{x}_0|\mathbf{Y}) \propto \exp \left\{ -\frac{1}{2} [\mathbf{x}_0 - \mathbf{x}^b]^T \mathbf{B}^{-1} [\mathbf{x}_0 - \mathbf{x}^b] \right\} \times \exp \left\{ -\frac{1}{2} \sum_{j=0}^N [\mathbf{y}_j - G_j(\mathbf{x}_0)]^T \mathbf{R}^{-1} [\mathbf{y}_j - G_j(\mathbf{x}_0)] \right\}, \quad (2.23)$$

where G_j is the combination of forward operator H_j and model operators M , reading

$$G_j := \{H_j M_j M_{j-1} \cdots M_1\} \quad \text{and} \quad M_0 = \mathbf{I}$$

generating the model equivalent of the observations at time t_j , when applied to the initial state. The covariance matrix of the observation errors is given by \mathbf{R} , and N denotes the time index t_N — *i. e.*, the end of the assimilation window.

The principle of 4d-var is defined by finding the maximum likelihood estimation of \mathbf{x}_0 by iteratively maximise its conditional probability (2.23), which is identical to the problem of minimising the scalar valued *objective function* or *cost function* J ,

$$\begin{aligned} J(\mathbf{x}_0) &= -\log [P(\mathbf{x}_0|\mathbf{Y})] \\ &= J_{\text{iv}} + J_{\text{obs}} \\ &= \frac{1}{2} [\mathbf{x}_0 - \mathbf{x}^b]^T \mathbf{B}^{-1} [\mathbf{x}_0 - \mathbf{x}^b] + \\ &\quad \frac{1}{2} \sum_{j=0}^N [\mathbf{y}_j - G_j(\mathbf{x}_0)]^T \mathbf{R}^{-1} [\mathbf{y}_j - G_j(\mathbf{x}_0)]. \end{aligned} \quad (2.24)$$

The minimisation is usually achieved by calculating the gradient of the cost function with respect to the control parameters \mathbf{x}_0 , given to a standard gradient descent minimisation algorithm performing successive iterations. This gradient can be obtained numerically in different ways. The most efficient way for high dimensional problems like in chemical data assimilation uses principles of the theory of adjoint equations and can be described as follows:

The application of the prognostic model to small variations $\delta\mathbf{x}_0$ and linearising the result leads to

$$\delta\mathbf{x}_j = \mathbf{M}_j \mathbf{M}_{j-1} \cdots \mathbf{M}_1 \delta\mathbf{x}_0, \quad (2.25)$$

where the Jacobians with respect to the state variable

$$\mathbf{M}_j := \left. \frac{\partial M_j}{\partial \mathbf{x}_j} \right|_{\mathbf{x}_j} \quad (2.26)$$

are linearised model operators, referred to as the *tangent linear* model. Under the assumption of both a linear model and a linear forward operator — G_j is linear in \mathbf{x}_j , the cost function (2.24) is a quadratic form and has a unique minimum. Hence, the minimisation will lead to an optimal guess of the initial values. On the validity of the *tangent linear approximation* (*i. e.*, the linearisation of both the model and observation operator still leads to a nearly optimal analysis) the reader may refer to *Courtier et al.* [1994], *Talagrand* [1997], *Cohn* [1997] and *Bouttier and Courtier* [1999].

We want to derive a formula for the gradient of the cost function with respect to the initial values. Therefore, using the canonical scalar product $\langle \mathbf{a}, \mathbf{b} \rangle = \sum_i a_i^T b_i$ (note the difference to $\langle \cdot \rangle$ denoting the expectation value), we first describe small variations of the scalar valued function $J(\mathbf{x})$ as a result of small variations $\delta \mathbf{x}$

$$\delta J \approx \langle \nabla_{\mathbf{x}} J, \delta \mathbf{x} \rangle .$$

Hence, the following holds for the observational term J_{obs} in the cost function (2.24):

$$\delta J_{\text{obs}} = \sum_{j=0}^N \langle \nabla_{\mathbf{x}_j} J_{\text{obs}}, \delta \mathbf{x}_j \rangle$$

Introducing the tangent linear model equation (2.25), small variations in the cost function are obtained as a result of small variations in initial conditions:

$$\delta J_{\text{obs}} = \sum_{j=0}^N \langle \nabla_{\mathbf{x}_j} J_{\text{obs}}, \mathbf{M}_j \mathbf{M}_{j-1} \cdots \mathbf{M}_1 \delta \mathbf{x}_0 \rangle . \quad (2.27)$$

Using $\langle \mathbf{a}, \mathbf{Qb} \rangle = \langle \mathbf{Q}^T \mathbf{a}, \mathbf{b} \rangle$, where the superscript T denotes transposition, we can rewrite (2.27) to

$$\delta J_{\text{obs}} = \sum_{j=0}^N \langle \mathbf{M}_1^T \mathbf{M}_2^T \cdots \mathbf{M}_j^T \nabla_{\mathbf{x}_j} J_{\text{obs}}, \delta \mathbf{x}_0 \rangle ,$$

where \mathbf{M}_j^T are the *adjoint* operators to the tangent linear model (2.25). Thus the gradient of the cost function with respect to the initial conditions is given by

$$\begin{aligned} \nabla_{\mathbf{x}_0} J &= \nabla_{\mathbf{x}_0} (J_{\text{iv}} + J_{\text{obs}}) \\ &= \mathbf{B}^{-1} [\mathbf{x}_0 - \mathbf{x}^{\text{b}}] + \sum_{j=0}^N \tilde{\mathbf{M}}_j^T \mathbf{H}_j^T \mathbf{R}^{-1} [\mathbf{y}_j - G_j(\mathbf{x}_0)] , \end{aligned} \quad (2.28)$$

where J_{iv} is the cost function with respect to the updated initial values. \mathbf{H}^T is the adjoint operator of the linearised forward operator and

$$\tilde{\mathbf{M}}_j^T := \prod_{i=j}^0 \mathbf{M}_i^T \quad (2.29)$$

is the adjoint model performing the integration from $t = t_j$ to $t = t_0$.

Consequently, the gradient of the cost function with respect to the initial values can be obtained by integrating the direct model forward in time followed by an adjoint model integration backward in time. At the end of the forward integration, the adjoint model state, here denoted by \mathbf{x}^* , is initialised with $\mathbf{x}^* = 0$. This adjoint model state is integrated backward in time using the adjoint model operators. At each time instance t_i , for which a new set of observations \mathbf{y}_j is available, the state of the adjoint integration is updated by the observational forcing term $\mathbf{y}_j - G_j(\mathbf{x}_0) = \mathbf{y}_j - H_j(\mathbf{x}_j)$. At the end of the integration, the gradient of the cost function with respect to the initial values is given by the final state of the adjoint model run. The main advantage of this algorithm is that it allows to calculate a complete and accurate gradient of the cost function with one direct (forward) integration and one successive adjoint (backward) integration of the model, which is numerically much cheaper than a perturbation method. The disadvantages are at hand: First, the adjoint model operators have to be coded and, second, the forward model states have to be completely available prior to the adjoint integration.

The former problem means, that a new model has to be developed as complex as the forward model. This can be achieved in various ways: One possibility is to adjoin the system of model equations with its own formulation, discretisation and linearisation (if needed). Another possibility is to derive the tangent linear of a piece of code and to adjoin line by line (*Talagrand* [1991], *Schmidt* [1996], *Giering and Kaminski* [1998]), which preserves the forward formulation and discretisation. Nowadays there are *adjoint* compilers that help creating adjoint pieces of code, for example TAMC (*Giering* [1999]) and *Odyssée* (*Faure and Papegay* [1998]).

The second problem is due to the fact that the forward model trajectory explicitly appears in the observational forcing terms and that it additionally needs to be available in order to get the right adjoint operator of the tangent linear solution (2.25), in the case of a non-linear forward model M and/or observation operator H . This leads to storage demands, which are still unfeasible using a model with a phase space size of order 10^6 , which is the case for the EURAD chemistry transport model (see, *e. g.*, *Schmidt* [1999]).

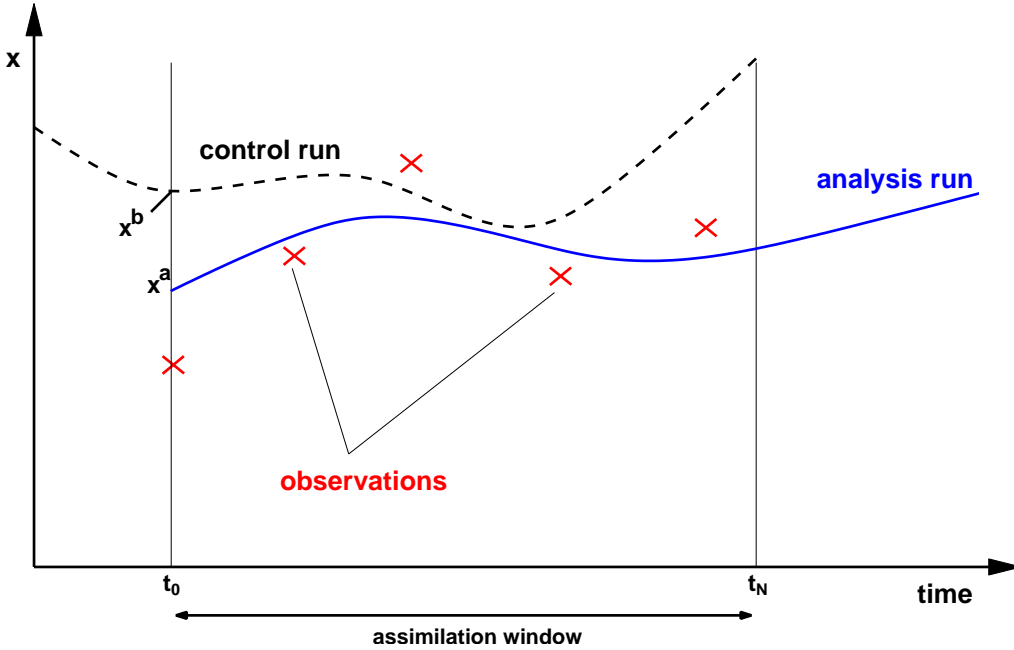


Figure 2.3: Principle of 4d-var data assimilation algorithm, here for one single model parameter. Because 4d-var fits a model trajectory to a set of observations within the predefined assimilation window, it is a smoother algorithm. The result is a **BLUE** for each time step.

The solution to these problems as implemented in the current EURAD 4d-var data assimilation system will be shown in Section 3.1.

Minimising the cost function (2.24) under the nonlinear constraint (2.20) — provided errors are Gaussian — will produce a **BLUE** of the system under consideration at any time $t_j \in [t_0, t_N]$, if the tangent linear assumption holds in the neighbourhood of \mathbf{x}^b and the error statistics are unbiased and correct (Lorenz [1986]). Due to the fact that it propagates information forward and backward in time, 4d-var is regarded a *smoother*, fitting a model simulation to a set of observations distributed in a predefined time window (see Fig. 2.3). This is in contrast to the principle of filtering, propagating information forward in time only (cf. Section 2.2). Many applications of 4d-var to real data experiments in meteorology, oceanography and atmospheric chemistry exhibit the capability of minimising efficiently the penalty function even under non-linear model constraints (see, e.g., Talagrand [1997] and references therein). Due to its formulation, 4d-var is able to produce physically and chemically consistent model simulations within the whole assimilation window and therefore to analyse even constituents which are not observed or to propagate information to data void regions.

2.4 4D-Var vs. KF

The 4d-var method and the Kalman Filter (KF) are spatio-temporal algorithms, which exploit the available information about the system state on the highest sophistication level. Under the assumption of a perfect model, that is that the model error $\boldsymbol{\eta}$ in (2.14) vanishes, both methods lead to the same analysis at the end of the 4d-var assimilation window. However, "the ambitious and elusive goal of data assimilation [...] to provide a dynamically consistent motion picture of the atmosphere and oceans, in three space dimensions, with known error bars" (*Ghil and Malanotte-Rizzoli [1991]*), can only be provided by the Kalman Filter without recourse to features of the minimisation solver. It has the certain appeal of integrating the initial background error covariance matrix in time, leading to forecast and analysis error covariance matrices, fulfilling the *elusive goal* mentioned above. An estimate of the analysis error covariance matrix by the 4d-var method can be yielded by, for example, singular value decomposition, which is used in a limited memory version by the current model version as minimisation routine (see Section 3.2). However, this method is of about the same complexity as the 4d-var method itself (*Elbern [2001]*).

Kalman Filtering comes with huge computational demands in the framework of large numerical models. This allows its application in atmospheric chemistry only for models of a strongly reduced complexity. Another method to supersede the unfeasibility is to reduce the complexity of the Kalman Filter itself (see Section 2.2), leading to very promising results in the stratosphere as well as in the troposphere. The advantage of those algorithms is the easy implementation, compared to the need to code the adjoint model version in the framework of 4d-var. While the KF serves for propagated covariance matrices, the analysis by 4d-var is physically and chemically consistent over the whole assimilation window without intermittent breaks at observation times.

Besides real data comparisons of 3d-var and the Ensemble Kalman Filter (EnKF) in the operational meteorological framework (*Houtekamer et al. [2005]*), first parallel implementations of EnKF and 4d-var in state-of-the-art chemistry transport models show small predominance of 4d-var during the assimilation window, while comparable results are achieved for both methods in subsequent forecasts (*Constantinescu et al. [2006]*). Possible future developments will take into account and exploit both methods.

CHAPTER 3

The EURAD 4D-Var Data Assimilation System

The EURAD (EUROpean Air pollution and Dispersion) data assimilation system consists of four main parts, the EURAD model system, the adjoint chemistry transport model (ADCTM), the minimisation procedure L-BFGS¹ including a preconditioning module, and the observational data preprocessor PREP, collecting available observations and providing them in a standardised file.

The EURAD model system is an advanced numerical model system able to simulate the physical and chemical evolution of trace gases in the troposphere and lower stratosphere (*Hass [1991]*, *Hass et al. [1995]*, *Petry et al. [1997]*). It consists of the meteorological driver MM5² (*Grell et al. [1994]*), the EURAD emission model EEM (*Memmesheimer et al. [1991]*) and the EURAD chemistry transport model (EURAD-CTM). The EURAD-CTM is a mesoscale chemistry transport model involving transport, diffusion, chemical transformation, wet and dry deposition of tropospheric trace gases (*Hass [1991]*, *Memmesheimer et al. [1997]*). It is used for long term simulations focussing on emission directives and their influence on air quality (*Memmesheimer et al. [2000]*, *Memmesheimer et al. [2004]*) as well as for operational forecast from hemispheric down to local scale using nesting technique *Jakobs et al. [1995]* and episodic scenarios focussing on special aspects

¹limited memory **B**royden **F**letcher **G**oldfarb **S**hanno

²developed at the National Center for Atmospheric Research, Boulder, USA, in cooperation with The Penn State University and the University Corporation for Atmospheric Research. www.mmm.ucar.edu/mm5

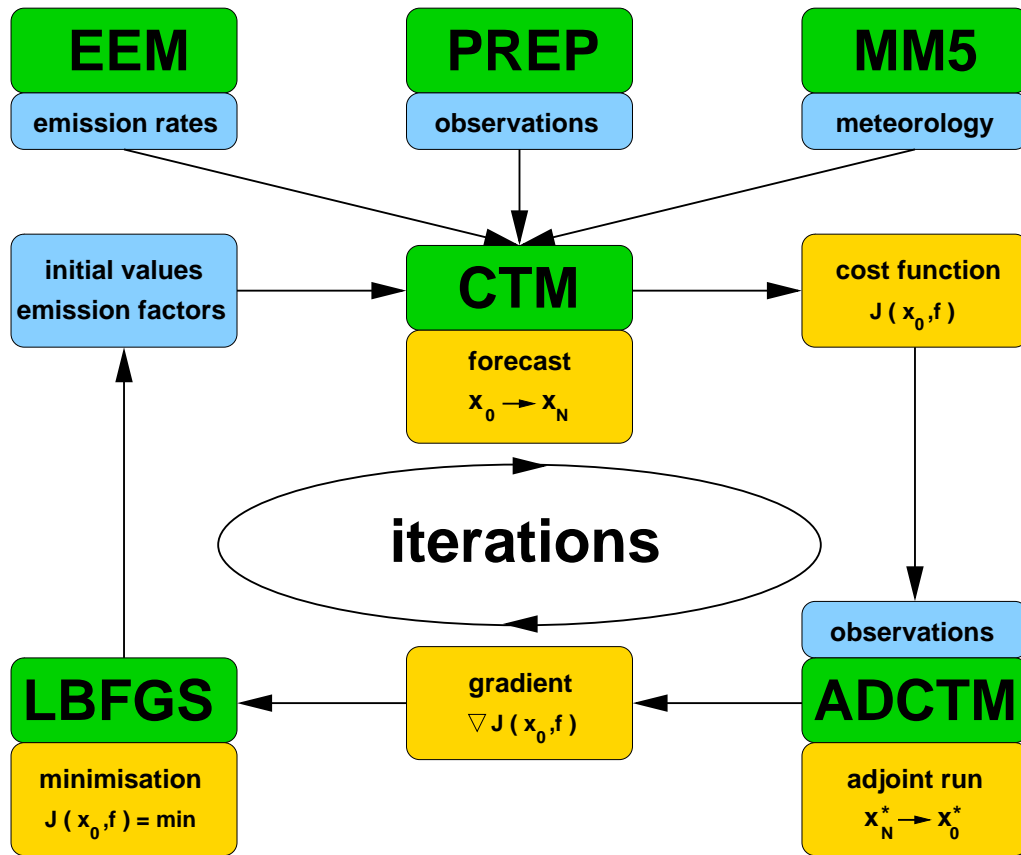


Figure 3.1: Flowchart of the EURAD 4d-var data assimilation system.

of air quality simulations.³

The resulting data flow of the data assimilation system, which is depicted in Fig. 3.1, can be described as follows: Given emission rates and meteorological parameters, the CTM integrates the model trajectory forward in time until the end of the assimilation window. Additionally, the model state is compared with the available observations and the cost function is calculated. The ADCTM performs a subsequent adjoint model integration backward in time, resulting in the gradient of the cost function with respect to the control parameters (here \mathbf{x}_0 for initial values and \mathbf{f} for emission factors). Using this gradient information, the minimisation procedure L-BFGS delivers new (improved) control parameters. Due to the tangent linear approximation of the gradient (see Section 2.3), one minimisation step is not sufficient to find the global minimum of the cost function. So the system

³The daily operational forecasts and its validation with observations is available and documented at www.eurad.uni-koeln.de.

is applied iteratively until the decrease of the cost function from one iteration to the next is lower than a certain threshold or a maximum number of iterations has been performed.

This chapter is structured as follows: First the CTM and its adjoint version ADCTM, which are building the kernel of the data assimilation system, as well as the minimisation procedure L-BFGS will be discussed. Sections 3.3-3.6 describe in detail the parts of the data assimilation system which have been newly developed or newly implemented for this work, namely the new chemistry kernel and its solver, the formulation of the joint optimisation of emission factors and initial values, the adjoint nesting technique, and the observation operator. The description of the new modelling of the error covariance matrices, which is one of the most crucial aspects of data assimilation, is dedicated an own chapter and can therefore be found in Chapter 4.

3.1 The EURAD-CTM and its Adjoint

The EURAD-CTM, which is an offspring of the Regional Acid Deposition Model RADM (*Chang et al.* [1987]), solves a set of partial differential equations for the tendency of M different species in an Eulerian framework (*Hass* [1991], *Seinfeld and Pandis* [1998]):

$$\frac{\partial c_i}{\partial t} = -\nabla(\mathbf{u}c_i) + \nabla(\rho\mathbf{K}\nabla\frac{c_i}{\rho}) + A_i + E_i - S_i, \quad (3.1)$$

where $c_i, i = 1, \dots, M$ are mean mass mixing ratios, \mathbf{u} are mean wind velocities, \mathbf{K} is the eddy diffusivity tensor, ρ is air density, A_i is the chemical generation term for species i , E_i and $S_i = \tilde{v}_i^d c_i$ its emission and removal fluxes, respectively. \tilde{v}_i^d denotes the deposition velocity for species c_i .

Let P_i and L_i be the production rate and the loss term of species i , respectively, R the number of chemical reactions, $k(r)$ the reaction rate of reaction r , and $s_i(r_{+,-})$ the stoichiometric coefficient of species i in reaction r , either productive (r_+) or destructive (r_-). Following *Finlayson-Pitts and Pitts* [1986], the chemical transformation term A_i can then be written as

$$A_i = P_i + L_i = \sum_{r=1}^R \left(k(r) [s_i(r_+) - s_i(r_-)] \prod_{j=1}^M c_j^{s_j(r_-)} \right).$$

The different processes are treated in an operator splitting scheme (*McRae et al.* [1982]), where a symmetric splitting of the dynamic procedures is

performed, encompassing the chemistry solver module C (see *Hass* [1991]),

$$c_i^{t+\Delta t} = T_h T_v D_v C D_v T_v T_h c_i^t, \quad (3.2)$$

where $T_{h,v}$ and D_v denote transport and diffusion operators, respectively, in horizontal (h) and vertical (v) direction.

The EURAD-CTM is a mesoscale- α model operating on integration domains with a boundary length of about 5000 km. In this work the horizontal grid designs are based on Lambert conformal conic projections with horizontal grid spacings ranging from 125 km down to 5 km. This is achieved by applying the nesting technique which is described in *Jakobs et al.* [1995] for the EURAD-CTM. The nesting procedure implemented in the data assimilation system will be outlined in Section 3.5. The gridded definition of the state variables follows the Arakawa-C-grid stencil (*Arakawa and Lamb* [1977]). In the vertical a terrain following sigma coordinate is used which is defined as

$$\sigma_k = \frac{p_k - p_{\text{top}}}{p_{\text{bot}} - p_{\text{top}}},$$

where p_{top} and p_{bot} is the model pressure at model top and model bottom, respectively. p_k is the pressure of the k^{th} model layer. In this work the number of vertical layers has been extended to 23 and 26 layers (see Appendix A), depending on the target area of the measurement campaign for which the simulations have been performed. The model top pressure has been limited to 100 hPa for all assimilation experiments, which roughly corresponds to 16 km height. The advection scheme is an upstream algorithm, devised by *Bott* [1989], fourth order in the horizontal and second order in the vertical. The vertical diffusion is an semi-implicit approach, which is solved by the Thomas algorithm.

The data assimilation procedure works as follows: At the beginning of a campaign simulation, initial values for each species are obtained by interpolating an externally provided vertical profile. This profile has only a seasonal variation, so that the initial condition of such a species is zonally uniform. Only for a few species like nitrogen oxides and ozone the user can chose to interpolate vertical profiles with a meridional and seasonal variation. Additionally, one can initialise ozone and hydrogen peroxide by correlation with potential vorticity and dew point temperature, respectively. A more detailed listing of the climatological information can be found in *Schell* [1996]. Generally a spin-up period with current meteorology and emissions is performed prior to assimilation experiments to force the model state to current conditions. Performing sequences of assimilation procedures, the analysis state based forecast of the previous experiment is used as initial condition

of the current day. Since better knowledge is missing, the boundary values in inflow regions are set to the initial conditions which can have a strong influence on the forecast, if the lateral boundary is not far enough from the region of interest. The model top is treated as a rigid lid, while the fluxes through the model bottom are given by deposition and emissions. The cross-dependency of initial and lateral boundary values with the nesting technique and the duration of the assimilation window is discussed in Section 3.5.

Elbern et al. [1997] provided the adjoint of the EURAD-CTM chemistry kernel built by the Second Generation Regional Acid Deposition Model RADM2 (*Stockwell et al.* [1990]), which consists of 63 species with a set of 158 gas phase and photolysis reactions. The set of stiff ordinary differential equations was solved by a semi-implicit and quasi steady state approximation (QSSA) method as derived by *Hesstvedt et al.* [1978]. The calculation of the time step followed *McRae et al.* [1982], applying a minimal time step of 1/100 min to maintain computational efficiency. In this work, a new chemistry solver has been implemented for another chemical reaction scheme, which will both be shortly discussed in Section 3.3. The radiative transfer equation has to be solved to retrieve the actinic fluxes of the photolysis frequencies. This is done by the radiation scheme of *Madronich* [1987].

Applying the rules of variational calculus (see, *e. g.*, *Courant and Hilbert* [1953]) to the chemistry transport equation (3.1) one can derive the adjoint equation (*Schmidt* [1999], *Elbern* [2001]):

$$\begin{aligned} \frac{\partial \delta c_i^*}{\partial t} = & -\mathbf{u} \nabla \delta c_i^* - \frac{1}{\rho} \nabla (\rho \mathbf{K} \nabla \delta c_i^*) + \tilde{v}_i^d c_i^* \\ & + \sum_{r=1}^R \left(k(r) \frac{s_i(r_-)}{c_i} \prod_{j=1}^U c_j^{s_j(r_-)} \sum_{n=1}^U [s_n(r_+) - s_n(r_-)] \delta c_n^* \right), \end{aligned} \quad (3.3)$$

where c_i^* is the adjoint variable to c_i . Emission strength vanishes due to its independence from c_i . The adjoint model could be elaborated from (3.3), but — as described in *Elbern and Schmidt* [1999] — the adjoint version of the operators in (3.2) have been coded resulting in the backward application of (3.2) with the operators substituted by their adjoint counterparts. The code generation strategy (line by line of the forward code as described by *Talagrand* [1991] or using the Adjoint Model Compiler, *Giering* [1999]), the necessary algorithmic simplifications and the practical implementation issues are discussed in detail in *Elbern et al.* [1997], *Schmidt* [1999], *Elbern et al.* [2000] and *Elbern* [2001].

As outlined in Section 2.3, the backward integration of the adjoint model

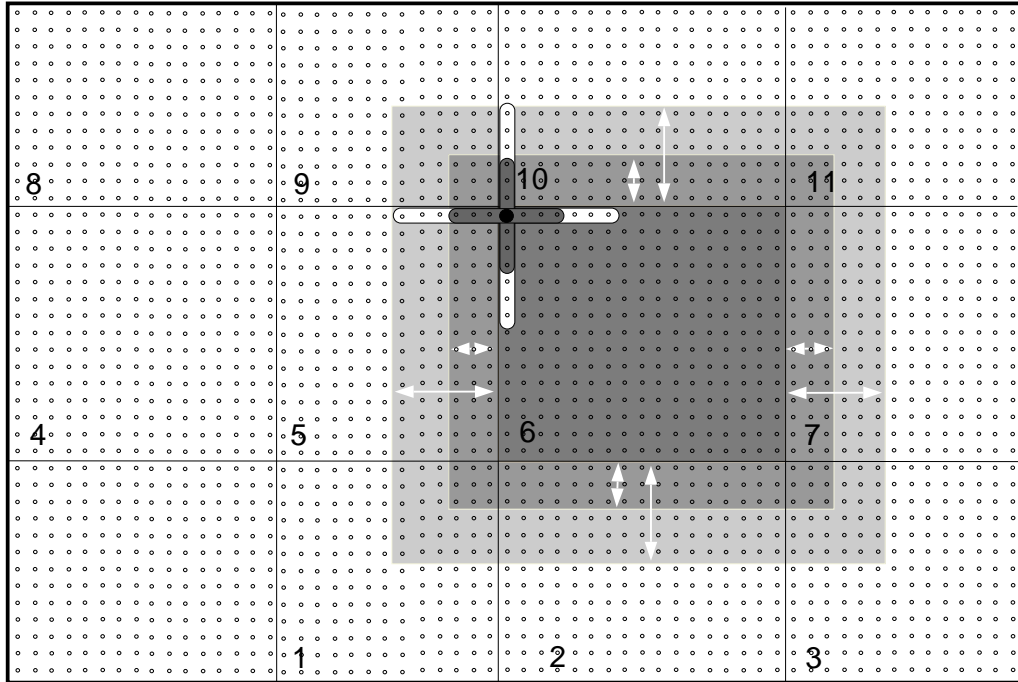


Figure 3.2: Domain decomposition strategy of the parallel implementation. Each processor hosts its own connected private sub-domain (the dark grey area for processor 6). Each unfilled circle denotes a grid box. Lighter shaded areas are overlap regions needed for the horizontal transport (three boxes for Bott's fourth order scheme and six for its adjoint version).

needs the complete model trajectory of the forward integration. A 24 hour forward run of a regional chemistry transport model with an advanced chemistry kernel leads to a storage demand of about $40 \cdot 10^{12}$ Bytes (see *Elbern et al.* [1999]), which is still far beyond feasible capacities. Hence storage must be traded with recalculations. The implemented solution follows the *dynamic stepwise* approach, that leads to three forward integrations per iteration cycle. Taking into account that an adjoint piece of code needs about twice the time of its respective forward part, the complexity in terms of forward integrations is about five, meaning that the runtime of one iteration cycle is about five times higher than that one of a single forward integration. As mentioned above, the EURAD 4d-var data assimilation system is completely parallelised, using a domain decomposition method (see Fig. 3.2) and the Message Passing Interface Environment for communication on distributed memory machines, in order to cope with these computational requirements (for a detailed description of these implementation issues the

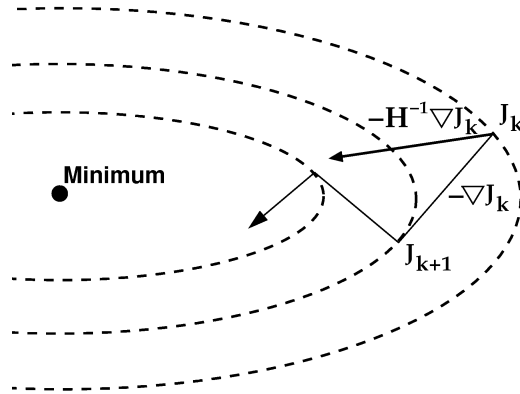


Figure 3.3: Illustration of steepest descent and Newton minimisation method. The cost function isopleths are forming a narrow valley, where the gradient of the cost function does not point towards the minimum. Employing the Hessian results in the exact search direction, which is done in Newton algorithms. After Daley [1991].

reader may refer to *Elbern* [2001] and the references therein).

All simulations shown in this work have been carried out on the massively parallel IBM eServer J_U^{MP} at the Forschungszentrum Jülich.

3.2 Minimisation Algorithm: L-BFGS

As mentioned in Section 2.3, the 4d-var algorithm rests on finding the minimum of the cost function (2.24). We will assume here, that the cost function can be locally approximated by a smooth quadratic function. The minimum will be found by evaluating the gradient of the cost function with respect to the control variables. An update to the model state by the information inherent in the gradient only is called *steepest descent*. This method is very inefficient due to the problem that isopleths of the cost function do not necessarily need to be spherical, resulting in (elongated) ellipsoids (see Fig. 3.3). The local gradient thus does not point towards the minimum, requiring many iterations to minimise the cost function (see, *e. g.*, *Press et al.* [1992]). The classical Newton method

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{H}^{-1}(\mathbf{x}_k)\mathbf{g}_k, \quad (3.4)$$

where the k^{th} iteration state is updated by the gradient $\mathbf{g}_k = \nabla_z J$ and the inverse of the Hessian \mathbf{H} , points towards the minimum of the gradient in one iteration, if the function is exactly quadratic and the gradient and the

Hessian are exact as well. The calculation of the Hessian and its inverse in high-dimensional cases as presented here, is computationally too demanding (*Wang et al.* [1992]). Favourable techniques are *conjugate gradient* and *quasi-Newton* methods.

The quasi-Newton BFGS algorithm is devised to approximate the Hessian iteratively. This method requires the storage of matrices of size n^2 , where n is the number of control variables, making the application of the original BFGS impossible. To this end, the L-BFGS (*Nocedal* [1980], *Liu and Nocedal* [1989]) is used in the current EURAD 4d-var system. Here the number of approximations q can be controlled by the user, resulting in $2q$ states of size n that have to be stored. For a detailed description of the L-BFGS in the EURAD 4d-var data assimilation system the reader may refer to *Schmidt* [1999].

The main aspect of the cost function which affects the minimisation performance is its *condition number* — *i. e.*, the ratio of the largest and the lowest eigenvalue in the Hessian. The larger the number is, the more ill-conditioned is the problem. The idea of *pre-conditioning* is not to minimise the original cost function and its gradient but a problem that is better conditioned and from where the analysis state can easily be obtained. Following (3.4), the perfect preconditioner is the Hessian itself. Chapter 4 will give details how preconditioning is implemented in the current model version, using the background error covariance matrices.

3.3 The Rosenbrock solver

To allow for more flexibility in adopting new chemistry mechanisms, a Rosenbrock solver and its adjoint have been implemented, which will be introduced shortly.

The system of gas phase reactions is given in the non-autonomous form for a set of chemical constituents y :

$$\dot{y} = f(y, t) \quad \text{with} \quad y_{t=t_0} = y_0. \quad (3.5)$$

Explicit integration methods like QSSA schemes tend to choose small time steps to preserve stability in chemically active regions (like during dawn in polluted areas). In contrast, implicit schemes have infinite stability regions and evaluating the step size only needs to fulfil the user's accuracy requirements (*Sandu et al.* [1996]).

The newly implemented solver is a stage-2 Rosenbrock method. Rosenbrock methods are also known as linearly implicit Runge-Kutta methods,

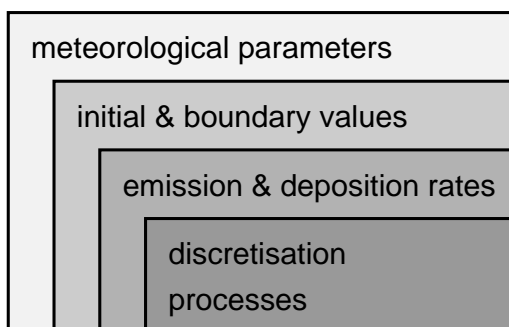


Figure 3.4: Controllability of model parameters in tropospheric chemical data assimilation. From light to grey / top down: higher resolving observational data is needed to analyse parameters. Process treatment and discretisation issues are only subject to optimisation when analysis error statistics are available. After Elbern [2001].

since they avoid nonlinear systems by replacing them by a sequence of linear systems.

The solver has been implemented for the Regional Atmospheric Chemistry Mechanism RACM (*Stockwell et al.* [1997]), which is based upon the earlier RADM2 and the more detailed Euro-RADM mechanisms. It includes 237 reactions of about 80 species. Besides other improvements, a new condensed reaction mechanism has been included for the biogenic compounds isoprene, α -pinene and d-limonene. The mechanism is created to be fully capable to simulate all tropospheric conditions, remote to polluted, and all height levels up to the upper troposphere.

The code of the chemistry solver has been constructed using the Kinetic Pre-Processor KPP (*Damian-Iordache* [1996]) Version 2.1 by *Sandu and Sander* [2006]. The main advantage of creating the Fortran code using KPP is the powerful exploitation of the sparsity of the Jacobian by the current version. Moreover, KPP is able to generate the tangent-linear and the adjoint model version of the solver as well. The original KPP version has been modified to directly produce library routines suitable for the EURAD 4d-var data assimilation system (*E. Friese*, personal information). The correctness of the program code has been tested using the method of *Chao and Chang* [1992]. All simulations shown in this thesis have been performed with the Rosenbrock solver for the RACM gas phase chemistry mechanism.

3.4 Optimisation of Initial Values and Emission Factors

Chemical weather forecasting in the troposphere is governed by several physical and chemical processes and controlled by parameters which are not known precisely enough. In section 2.3 the theoretical background of 4d-var has been introduced briefly, the system's initial values being the only set

of parameters subject to optimisation. This is due to the fact, that 4d-var has first been devised in the framework of meteorological systems, where initial conditions strongly determine the system's evolution. In tropospheric chemistry, this predominance is weakened by the strong influence of emissions, at least for emitted constituents and their products in the planetary boundary layer.

However, 4d-var admits for the optimisation of further parameters than initial values only. But which set of parameters should be chosen? A comprehensive answer to this question can only be given for a special system under consideration. Generally, those parameters should be optimised which are poorly known and to which the system's evolution in time is very sensitive. Small variations in these parameters lead to a significant change in model trajectory. Doing tropospheric chemical data assimilation, Fig. 3.4 gives an introductory classification of those parameters that are amenable to quality control under different observational information densities.

Elbern [2001] showed the basic feasibility of the adjoint modelling technique for emission rate assessment by identical twin experiments. The current implementation of the EURAD 4d-var data assimilation system allows to optimise both initial values and emission rates, and — as a key feature presented in the simulation results in Chapter 5 — the joint optimisation of both parameter sets. The motivation for jointly analysing initial values and emission rates is given in Fig. 3.5, where the time evolution of one single model parameter's mixing ratio is given for the different optimisation modes. Having a biased first guess model trajectory due to wrong background emission estimates, initial value optimisation will lead to an analysis that compensates for wrong emissions by wrong initial values. The same applies to emission rate optimisation, where corrected emission estimates lead to an opposite bias later on. Only jointly optimising both initial values and emission rates enables to fit the model trajectory to the truth. Moreover, combining the adjoint nesting technique, that has been newly developed and that will be introduced in the next section, with a successive application of assimilation experiments, the gap between initial and boundary conditions is filled and the model setup is able to optimise the set of initial values, boundary values and emission rates (see also Section. 3.5).

Emission estimates used in this work have been provided by the Co-operative programme for monitoring and evaluation of the long-range transmissions of air pollutants in Europe (EMEP, www.emep.int) as yearly emitted masses. These estimates have then been further processed by the EURAD emission model EEM, primarily adapting the data to the appropriate projection combined with infra-structural information and assuming some detailed time

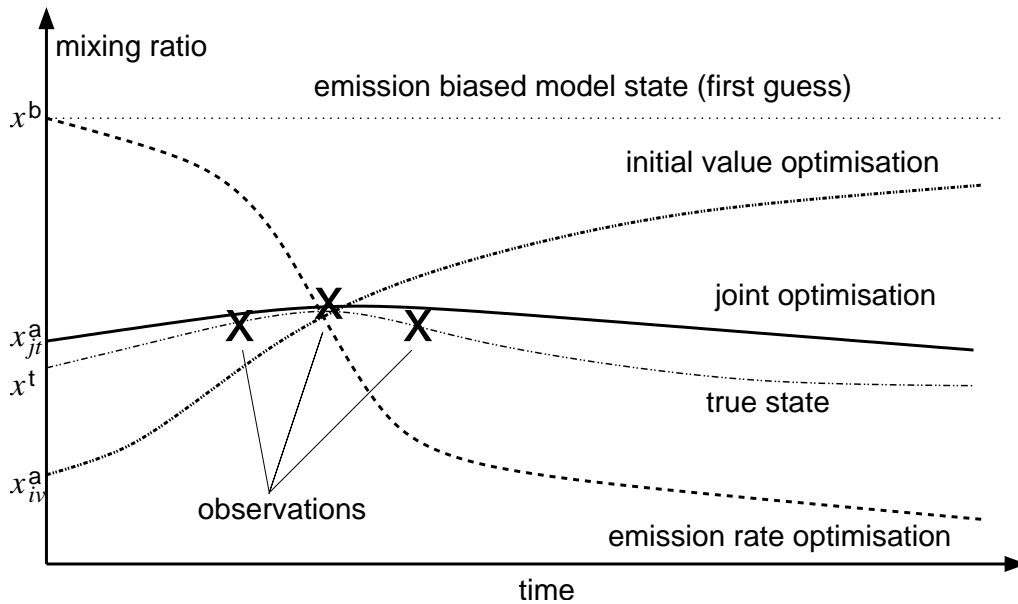


Figure 3.5: The principle benefit of joint optimising initial values and emission rates. Dash-dot-dotted line: true state. Crosses: observations. Dotted line: first guess model trajectory. Bold dash-dot-dotted line: initial value optimisation model trajectory. Dashed line: emission rate optimisation model trajectory. Solid line: joint optimisation model trajectory. Subscripts "iv" and "jt" denote initial value and joint optimisation, respectively.

functions to distribute the yearly emitted amount down to hourly values. For a more detailed description see Memmesheimer *et al.* [1995]. In the first study concerning emission rate optimisation, Elbern *et al.* [2000] showed that even species which are not observed are principally feasible for analysis when chemically linked to observed ones. But the authors emphasised, that a proper regularisation to emission estimates is needed to introduce further constraints to the optimisation process in order to reduce the number of degrees of freedom.

The emission rates are optimised as follows: The daily cycle of each emitted species in each grid cell is preserved, while the total amount of emitted mass is scaled by a time invariant *emission factor* $f = f(\text{BoxIndex}, \text{Species})$, which is to be optimised (see Fig. 3.6). This kind of regularisation is called *temporal regularisation* and seems to be well suited for this kind of problem, since the daily evolution of the emissions is far better known than the total emitted amount in a grid cell (M. Memmesheimer, personal communication). Moreover, one prerequisite of the optimality of the 4d-var method

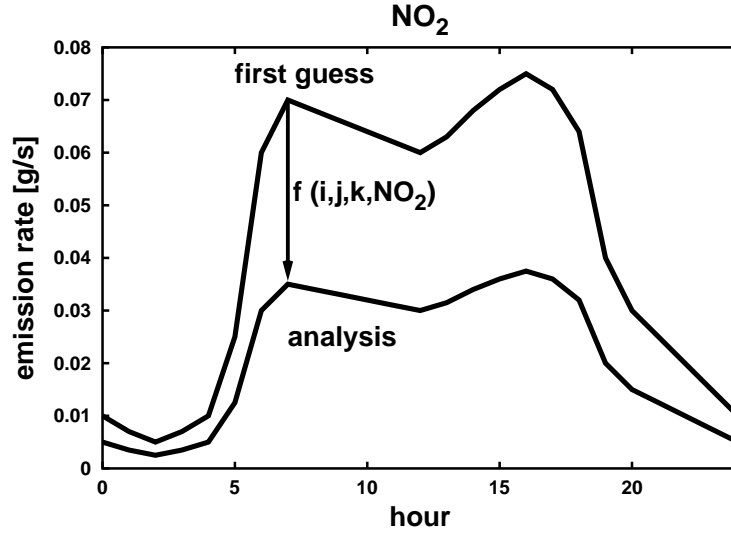


Figure 3.6: Definition of emission factors (here for NO_2 , (i,j,k) are the spatial box indices) being control parameters and thus subject to optimisation. The diurnal cycle of emission rates is preserved while the total emitted amount is scaled by the factor.

(the **BLUE** property) is the description of error pdfs as Gaussians (see section 2.3). Emission estimates, being bounded from below as non-negative parameters, generally do not have Gaussian error distributions (the errors due to truncation are in fact small, when the expectation value is large compared to the variances). However, the idea is to logarithmically transform the emission factors resulting in the equivalent partial costs for the emission factors 2.0 and 0.5, for example. Related to emission rate estimation, the control variable \mathbf{u} is thus given by

$$\mathbf{u} = \ln(\mathbf{f}) \quad \text{and} \quad \mathbf{u}^b = \ln(\mathbf{f}^b), \quad (3.6)$$

with associated covariance matrix \mathbf{K} . Consequently, the evolution of the model state \mathbf{x} is now not only predefined by the initial state \mathbf{x}_0 , but also by the emission estimate \mathbf{E} , here encoded in the model operator M :

$$\mathbf{x}_j = M_j(\mathbf{E}_j) \mathbf{x}_{j-1}$$

Hence, the model trajectory is uniquely defined by the initial conditions \mathbf{x}_0 and the emission rates for all $t \in [t_0, t_j]$. In the following, the subscripts ef , iv and jt denote parameters linked to emission factor estimation, initial value estimation and joint estimation of both sets, respectively. The cost

function for emission rate optimisation then reads

$$\begin{aligned}
J &= J_{\text{ef}} + J_{\text{obs}} \\
&= \frac{1}{2} [\mathbf{u} - \mathbf{u}^{\text{b}}]^T \mathbf{K}^{-1} [\mathbf{u} - \mathbf{u}^{\text{b}}] + \\
&\quad \frac{1}{2} \sum_{j=0}^N [\mathbf{y}_j - G_{j,\text{ef}}(\mathbf{x}_0)]^T \mathbf{R}^{-1} [\mathbf{y}_j - G_{j,\text{ef}}(\mathbf{x}_0)],
\end{aligned} \tag{3.7}$$

where $G_{j,\text{ef}}$ is the resolvent depending on the emission rates. If we describe the emission rates as the background knowledge of the emission rates scaled by the emission factor (see above),

$$\mathbf{E}_j = \mathbf{f} \cdot \mathbf{E}_j^{\text{b}} = \exp(\mathbf{u}) \cdot \mathbf{E}_j^{\text{b}}, \tag{3.8}$$

the model operator M is independent from the background emission rate \mathbf{E}_j^{b} and the tangent linear model can be written as

$$\mathbf{M}_{j,\text{ef}} = \left. \frac{\partial M_j}{\partial \mathbf{u}} \right|_{\mathbf{x}_j} = \left. \frac{\partial M_j}{\partial \ln(\mathbf{f})} \right|_{\mathbf{x}_j}. \tag{3.9}$$

In equivalence to (2.28), the gradient of the cost function with respect to \mathbf{u} is given by

$$\begin{aligned}
\nabla_{\mathbf{u}} J &= \nabla_{\mathbf{u}} (J_{\text{ef}} + J_{\text{obs}}) \\
&= \mathbf{K}^{-1} [\mathbf{u} - \mathbf{u}^{\text{b}}] + \sum_{j=0}^N \tilde{\mathbf{M}}_{j,\text{ef}}^T \mathbf{H}_j^T \mathbf{R}^{-1} [\mathbf{y}_j - G_{j,\text{ef}}(\mathbf{x}_0)],
\end{aligned} \tag{3.10}$$

where $\tilde{\mathbf{M}}_{j,\text{ef}}^T$ is the adjoint model operator, defined in the same way as $\tilde{\mathbf{M}}_j^T$ in (2.28). The current implementation first calculates $\nabla_{\mathbf{f}} J_{\text{obs}}$ and then transforms to $\nabla_{\mathbf{u}} J_{\text{obs}} = \mathbf{f} \cdot \nabla_{\mathbf{f}} J_{\text{obs}}$, the equivalence of which is easily shown.

To derive the algorithm for jointly optimise initial values and emission factors, we define the merged set of control variables as

$$\mathbf{z} := (\mathbf{x}_0, \mathbf{u})^T$$

with joint covariance matrix $\begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{K} \end{pmatrix}$. The block diagonal formulation of the covariance matrix is due to the assumption, that errors of the initial values and the emission factors are mutually uncorrelated. This is a crude simplification, at least for the application of joint optimisation in an operational mode, but the required statistical information needed to quantify the covariances are not yet available.

The complete cost function for joint optimisation can then be written as the sum of the background costs of both initial values and emission factors and the observational costs, reading

$$\begin{aligned}
J &= J_{\text{iv}} + J_{\text{ef}} + J_{\text{obs}} \\
&= \frac{1}{2} [\mathbf{x}_0 - \mathbf{x}^{\text{b}}]^T \mathbf{B}^{-1} [\mathbf{x}_0 - \mathbf{x}^{\text{b}}] + \\
&\quad \frac{1}{2} [\mathbf{u} - \mathbf{u}^{\text{b}}]^T \mathbf{K}^{-1} [\mathbf{u} - \mathbf{u}^{\text{b}}] + \\
&\quad \frac{1}{2} \sum_{j=0}^N [\mathbf{y}_j - G_j(\mathbf{x}_0)]^T \mathbf{R}^{-1} [\mathbf{y}_j - G_j(\mathbf{x}_0)]
\end{aligned} \tag{3.11}$$

and the gradient of this cost function with respect to the joint control parameters is given in a concise matrix notation by

$$\begin{aligned}
\nabla_{\mathbf{z}} J &= - \begin{pmatrix} \mathbf{B}^{-1} & 0 \\ 0 & \mathbf{K}^{-1} \end{pmatrix} \begin{bmatrix} \mathbf{x}_0 - \mathbf{x}^{\text{b}} \\ \mathbf{u} - \mathbf{u}^{\text{b}} \end{bmatrix} \\
&\quad - \sum_{j=0}^N \begin{bmatrix} \tilde{\mathbf{M}}_{j,\text{iv}}^T \\ \tilde{\mathbf{M}}_{j,\text{ef}}^T \end{bmatrix} \mathbf{H}_j^T \mathbf{R}^{-1} [\mathbf{y}_j - G_j(\mathbf{x}_0)],
\end{aligned} \tag{3.12}$$

with $\tilde{\mathbf{M}}_{j,\text{iv}}^T$ defined in Eq. (2.29). The adjoint model operators $\tilde{\mathbf{M}}_{\text{iv}}$ and $\tilde{\mathbf{M}}_{\text{ef}}$ are identical to those used in the single initial value and emission rate optimisation only during the first backward integration. Later the tangent linear model evolution is different due to the different \mathbf{x}_i in the local partial derivatives (see Eqs. (2.26) and (3.9)).

The method of jointly optimising initial values and emission factors is a key feature of the EURAD 4d-var system used for the simulations described in Chapter 5, where the predominance over simple state estimation will be shown. The covariance matrices, which are used to precondition and balance the minimisation problem, are described in detail in Chapter 4.

3.5 Adjoint Nesting Technique

The widespread nesting technique is used for successive grid refinements while each nested model domain is a horizontal sub-domain of the respective mother domain. This facilitates to simulate large-scale features together with rather detailed structures by telescoping down to smaller scales in the region of interest. This technique is employed in two different modes, namely one-way and two-way nesting. The difference of these modes is the

way the simulations on the nested domains are performed. Taking two-way nesting, the simulations on the different grids are carried out simultaneously, with information flow in both directions: from mother domain to daughter domain and vice versa. If a model sequence uses one-way nesting, the information flow is only from mother domain to daughter domain, thus the simulations on the daughter domain do not feedback the results on the mother domain. Therefore, one-way nesting can be performed one after another. In both modes, the coarser grid serves for the initial and boundary values for the finer grid. This is achieved by interpolation, which is one of the crucial properties of the refinement approach, since it needs to preserve the system's shape and monotonicity properties. Applying two-way nesting, additionally the model state of the coarser domain is updated at each time step by averaged values of the finer grid.

The EURAD-CTM admits to be run in a multiple, one-way nested mode following a scheme devised by *Pleim et al.* [1991]. It has been first introduced by *Jakobs et al.* [1995] and in the following been employed in various model studies (see, *e. g.*, *Memmesheimer et al.* [2004] and citations therein).

In this work, this technique has been adapted to the 4d-var data assimilation system. The default configuration of the informational flow between the different nested grids and successive episodic assimilation experiments is shown in Fig. 3.7. On each nest level (here coarse grid CG and nest N1), the background knowledge of the initial values is retrieved by a forecast F on the same nest level (CG \rightarrow CG, N1 \rightarrow N1, etc.) using the analysed model parameters from the previous day ($F[\mathbf{x}_{n-1}^a, \mathbf{u}_{n-1}^a] \rightarrow \mathbf{x}_n^b$, called "analysis run" hereafter). The background knowledge of the emission factors (if optimised) is equal to the analysis of the previous day ($\mathbf{u}_{n-1}^a \rightarrow \mathbf{u}_n^b$). n denotes the day number.

Dynamical boundary values are provided by the analysis run of the respective mother domain. This scheme of inheritance of information is chosen for two reasons: 1) The analysis on the CG for the current day already comprises the knowledge about the system coded in the observations from the same day and, thus, is a step beyond any forecast resting on an analysis from the previous day. However, following the principle of nesting, 2) the forecast on the nested domain can contain physical and chemical features which are not resolved by a coarse grid simulation and thus are not destroyed by taking the initial values from a model simulation on the same nest level. Inflow regions of the nested domain will always gain the information already contained in the mother domain. Fig. 3.8 shows surface ozone concentrations obtained by interpolation of a mother domain's analysis state and nested forecast using the analysis of the previous day. Both

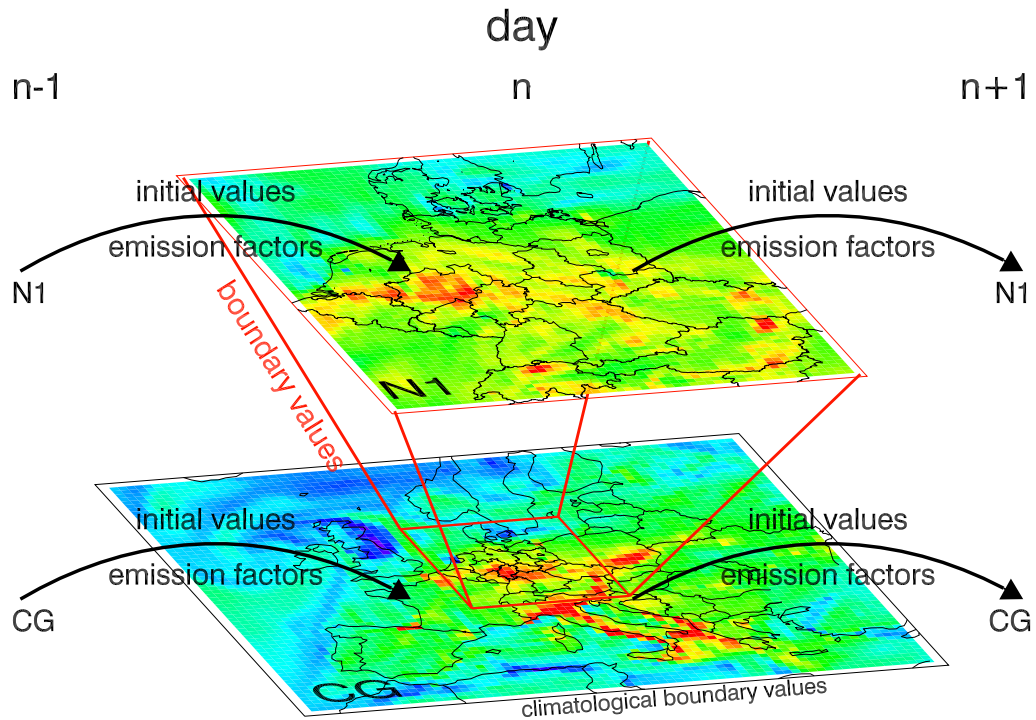


Figure 3.7: Simplified portrayal of the default nesting strategy in the EURAD 4d-var data assimilation system. Plots show surface ozone distribution at noon on July 19, 1998, as simulated by the control run for the BERLIOZ campaign simulation.

ozone distributions exhibit the same main features, while the nested forecast already contains sharp local structures, like motorway signatures by low ozone mixing ratios due to high nitrogen oxides emissions (*e. g.*, central Bavaria). The model simulation, starting with interpolated distributions, will need a certain spin-up time for developing these local patterns.

As in meteorological data assimilation, the main stimulus of initial value optimisation is, that model integrations based on optimised initial model states will lead to improved forecast skills. To put the strategy — introduced in Section 2.3 — simple, the information of an observation is integrated backward in time and the model state is changed at the resulting position. If this location lies outside the model domain, there is no direct possibility to use this observation for adjusting the initial model state. This applies for all parts of a model domain, which are influenced by boundary conditions during the forward integration. Consequently, this will happen even more often on a much smaller, nested domain. One direct conclusion is, that initial value optimisation will lose its information gain under strong

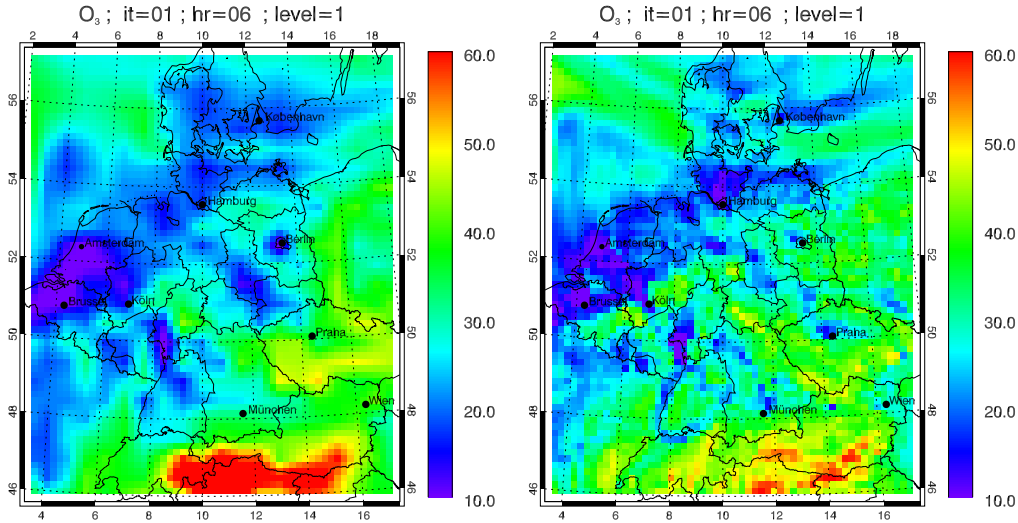


Figure 3.8: Initial ozone mixing ratios (in ppbV) on July 20, 1998, 06 UTC as used by the BERLIOZ campaign simulations for Nest 1, according to the initialisation strategy. Left: Interpolated values of the mother domain's analysis state. Right: Forecast on the same nestlevel, based on the analysis of the previous episode. Ozone initial values.

inflow conditions, when the domain is too small compared to the length of the assimilation window. This will lead to a stronger focus on emission factor optimisation on nested domains.

Future developments should therefore focus on the model design and to assure that the main region of interest is always sufficiently far enough from lateral boundaries. This can be achieved, e.g. by analysing the current dynamics on the coarse grid and to decide the size and location of a nested domain afterwards. This will lead to a more flexible model design, allowing to optimally exploit the adjoint nesting technique at same numerical expenses.

Concerning possible benefits of nesting in data assimilation it has to be stressed here, that the observational errors (2.2) comprise measurement errors as well as representativeness errors (see Section 4.4), which are due to the limitations given by $(H\mathbf{x}^t)$ generating the model equivalence to \mathbf{y}^t . The telescoping facilities of the nesting technique allow to successively enhance the horizontal resolution of a model simulation. Due to the finer discretisation in \mathbf{x} , this will lead to a step-by-step reduction of the error of $H\mathbf{x}$. The benefits of the nesting technique in simulating trace gases with a high spatial variability like nitrogen oxides has been shown in various model simulations and will be expanded to data assimilation in Chapter 5.

In addition to the possible benefits of nested forecasting, the application of the nesting technique in an advanced chemical data assimilation system supplies the feasibility of high-quality spatio-temporal model state analyses at an almost arbitrary horizontal resolution. This possibly enables the variational calculus to identify wrong process treatment or inappropriate discretisation. With increasing computer power and high resolving campaign data supported by satellite based earth observations, these issues will be obvious (see Fig. 3.4).

3.6 Observation Operator

The observation operator and its adjoint are the interfaces between the observation space and the model space. They have to be tailored for each observation type anew. The operator is needed to get the model equivalent to each type of measurement, yielding the possibility to compare the model state to various kinds of observations. The design of this operator is the key to a powerful data assimilation implementation, namely the ability to introduce knowledge to the model simulation coded in various kinds of observation. Additionally, a proper design may reduce the representativeness error of observations (see section 4.4). Hence, the observation operator must

- provide a mapping from model parameters to observed quantities,
- interpolate gridded information to measurement locations and
- integrate or average sub-domains,

or even combine all of those. Further, for the adjoint integration in (2.28) the adjoint of the observation operator must be available. Therefore, all facilities need to be coded in a forward and an adjoint version. This section will give a short overview of two implemented features dealing with earth observations by satellites. The descriptions are given for the operator H , while both the operator itself as well as its adjoint H^T have been coded.

Satellite based measurements are radiances. Usually a radiative transfer model and its inverse are employed to retrieve direct information about a trace gas distribution by these observed radiances. The retrieval model uses some assumptions about the trace gas distribution, for example a-priori information like a climatological profile, and some smoothness constraints are applied. The retrieval result is a compromise between the information

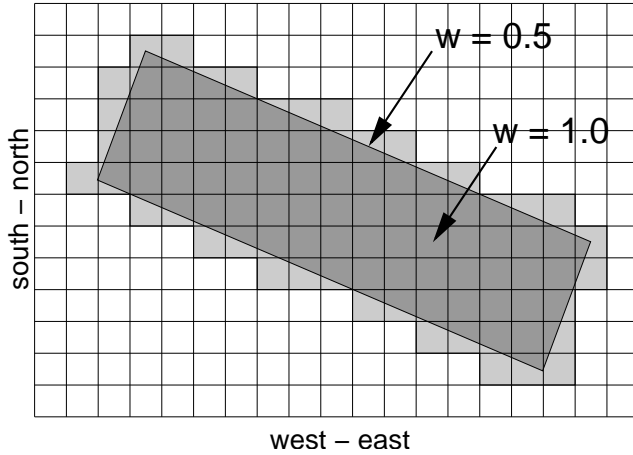


Figure 3.9: Strategy to calculate a model equivalent for an observation coming with a footprint information (dark shaded area). The fractional coverage w of all grid cells involved (light shaded boxes) is calculated and used for a weighted average. The weights are indicated for two grid cells.

contained in the observations and some other information sources. One way to deal with this limitation inherent in satellite based retrievals is to use averaging kernel information if available. The averaging kernel (AK) is a matrix giving the sensitivity of the retrieval result \hat{x} with respect to small changes in the modelled trace gas distribution \mathbf{x} , reading

$$\hat{x} - \hat{x}_a = \mathbf{A}(\mathbf{x} - \mathbf{x}_a), \quad (3.13)$$

where \mathbf{A} is the AK and \mathbf{x} a vertical profile, both defined on the retrieval grid layers, and subscript a denotes a-priori information. Using AK in the observation operator H of a data assimilation system enables to calculate the forecast by the retrieval procedure, given the data assimilation system's model state. For a comprehensive discussion of satellite based measurements and retrieval procedures see, *e. g.*, Rodgers [2000].

In this work, retrievals of total tropospheric DOAS NO_2 columns from the Koninklijk Nederlands Meteorologisch Instituut (KNMI) have been assimilated using AK information. Following Eskes and Boersma [2003], a comparison of DOAS retrieved total columns of optically thin absorbers like NO_2 with any model profile \mathbf{x} is independent of the a-priori profile which is used in the retrieval procedure. Therefore Eq. (3.13) reduces to

$$\hat{x} = A(\mathbf{x}),$$

where \mathbf{x} is now an arbitrary model profile and \hat{x} is a total NO_2 column. The implemented observation operator involves interpolation from the vertical CTM grid to the retrieval grid layers and summation from bottom to top, weighted by the AK.

For all satellite-borne data which has been assimilated in this work, the retrieval pixels have been available as *footprints*, comprising latitude and

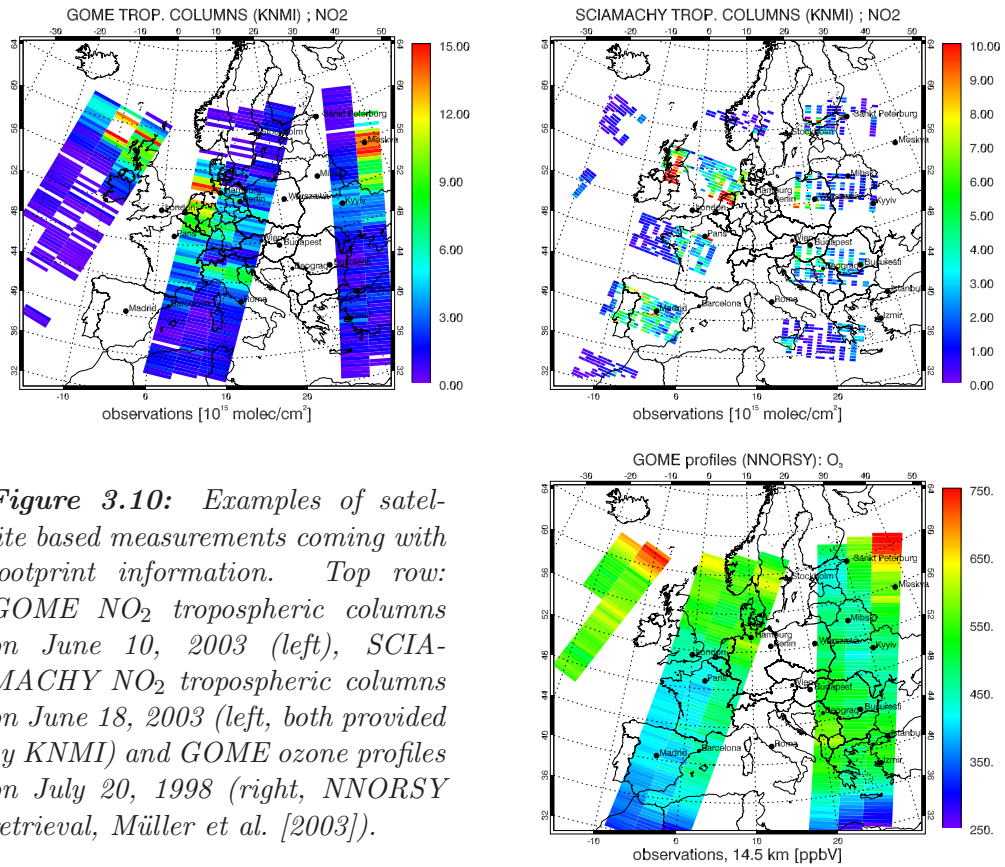


Figure 3.10: Examples of satellite based measurements coming with footprint information. Top row: GOME NO₂ tropospheric columns on June 10, 2003 (left), SCIAMACHY NO₂ tropospheric columns on June 18, 2003 (left, both provided by KNMI) and GOME ozone profiles on July 20, 1998 (right, NNORSY retrieval, Müller *et al.* [2003]).

longitude values of the four pixel corners, properly coding the representative area. To calculate the model equivalent for such an observation, the fractional coverage of all grid cells by the footprint area is calculated and a weighted average over those grid cells, which are affected by the footprint area. This scheme, which is sketched out in Fig. 3.9, allows for a much more sophisticated and realistic treatment of such kind of observations than simply retrieving the model equivalent by that grid cell, in which the centre of the observation resides. Three retrieval examples, NO₂ column retrievals (provided by KNMI) from Global Ozone Mapping Experiment (GOME) and SCIAMACHY (Scanning Imaging Absorption Spectrometer for Atmospheric Chartography), and O₃ retrievals for a height of 14.5 km, being part of tropospheric O₃ profiles (NNORSY, Müller *et al.* [2003]), are shown in Fig. 3.10.

CHAPTER 4

Covariance Modelling

Error covariances are symmetric and positive definite matrices, where the diagonal elements contain the variances and the off-diagonal elements are covariances. As introduced in Section 3.4, correlations between initial values, emission factors and observations are neglected. Thus there are three covariance matrices in the implemented model version, namely the observation error covariance matrix \mathbf{R} and two background error covariance matrices \mathbf{B} and \mathbf{K} for initial values and emission factors, respectively.

The weight given an observation at a certain location is mainly determined by its a-priori known error statistics and the background error at this location (or, more precisely, at the location of the backward trajectory at analysis time). Hence, one needs to specify both standard deviations correctly to produce an optimal analysis. Moreover, the background fields usually are results of a previous forecast and thus governed by characteristic dynamical and chemical processes. Any knowledge about the influence of these processes on the background errors can be coded into the correlation matrix, which is inherent in a covariance matrix. The minimum of the cost function will then not only be encountered by adjusting the background value for the observed constituent at the observational location, but as well all model parameters it is correlated with. Thus, the information of a single observation can be exploited to spread information in the model state. This enables to efficiently introduce information to the model even by using observations with sparse spatial coverage, *e. g.*, air-borne data. Correlations

between different model locations for a given constituent are building the univariate part of the background error covariance matrices, while the cross-correlations of different trace gases are known as multivariate correlations. In the case of the initial value background error covariance matrix \mathbf{B} , the multivariate part is of less importance in chemical data assimilation compared to meteorological applications. A chemical imbalance at initial time will relax towards a chemical equilibrium while a physical imbalance in meteorology can cause problems associated with gravity waves (see, *e. g.*, Daley [1991]).

Further, the Hessian of the cost function can be expressed by the model operator and the covariance matrices involved, while its computation is numerically too expensive. The first term in the Hessian, defined by the second derivative of the background term J^b in the cost function with respect to the control parameters, is given by

$$\nabla_z (\nabla_z J^b) = \begin{pmatrix} \mathbf{B}^{-1} & 0 \\ 0 & \mathbf{K}^{-1} \end{pmatrix} + \dots$$

As the inverse of the Hessian matrix, which is proportional to the analysis error covariance matrix, is the perfect preconditioner of the minimisation problem, a first order approximation to the conditioning problem is available by the background error covariance matrices \mathbf{B} and \mathbf{K} .

Now, for a modelling system like the EURAD system, the number of state variables n is of order $\mathcal{O}(10^6)$. Therefore, the storage of the initial value background error covariance matrix \mathbf{B} , which is of size $(n^2 + n)/2$, is not anymore feasible. Moreover, a proper guess of the entries in \mathbf{B} is difficult to obtain. Since the true model state is not available, one has to compare an enormous number of realisations of background fields, to sample the error distribution function and to derive by that the correlations encoded in \mathbf{B} .

An applicable data assimilation system thus needs to truncate the information in \mathbf{B} and to circumvent these problems. One possibility is to provide \mathbf{B} by some leading singular vectors only, as demonstrated by *Elbern and Schmidt* [2001]. A more flexible technique is to replace the univariate part in \mathbf{B} by an operator as devised by *Weaver and Courtier* [2001], which will be introduced in Section 4.2. As a prerequisite, we will first reformulate the cost function in an incremental way.

Then there will be a description of the multivariate background error covariance matrix \mathbf{K} , which is substantially easier to obtain, and the properties of the observational error covariance matrix \mathbf{R} .

4.1 Incremental formulation

The incremental version of the cost function (3.11) can be obtained by translating the control variables by their respective background values,

$$\delta \mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}^b \quad \text{and} \quad \delta \mathbf{u} = \mathbf{u} - \mathbf{u}^b, \quad (4.1)$$

where $\delta \mathbf{x}_0$ and $\delta \mathbf{u}$ will be called *increments* hereafter. Additionally, we use the square root of the inverse background error covariance matrices to obtain new initial state variables \mathbf{v} and \mathbf{w} by

$$\mathbf{v} = \mathbf{B}^{-1/2} \delta \mathbf{x}_0 \quad \text{and} \quad \mathbf{w} = \mathbf{K}^{-1/2} \delta \mathbf{u}. \quad (4.2)$$

Let $\mathbf{d}_j := \mathbf{y}_j - G_j(\mathbf{x}^b, \mathbf{u}^b)$ denote the innovation vector at time $t = t_j$ and the difference in the trajectories of the model equivalent with respect to the initial values be given by

$$\delta \hat{\mathbf{y}}_j = G_j(\mathbf{x}_0, \mathbf{u}) - G_j(\mathbf{x}^b, \mathbf{u}^b) \approx \mathbf{G}_j(\delta \mathbf{x}_0, \delta \mathbf{u}).$$

The cost function (3.11) can then be rewritten as an incremental version by

$$J(\mathbf{v}, \mathbf{w}) = \frac{1}{2} \mathbf{v}^T \mathbf{v} + \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{2} \sum_{j=0}^N [\delta \hat{\mathbf{y}}_j - \mathbf{d}_j]^T \mathbf{R}^{-1} [\delta \hat{\mathbf{y}}_j - \mathbf{d}_j]. \quad (4.3)$$

With this transformation, the gradient with respect to the state variables $\tilde{\mathbf{z}} = (\mathbf{v}, \mathbf{w})^T$ reads

$$\nabla_{\tilde{\mathbf{z}}} J = - \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} - \begin{pmatrix} \mathbf{B}^{T/2} & 0 \\ 0 & \mathbf{K}^{T/2} \end{pmatrix} \sum_{j=0}^N \begin{bmatrix} \tilde{\mathbf{M}}_{j,\text{iv}}^T \\ \tilde{\mathbf{M}}_{j,\text{ef}}^T \end{bmatrix} \mathbf{H}_j^T \mathbf{R}^{-1} [\delta \hat{\mathbf{y}}_j - \mathbf{d}_j]. \quad (4.4)$$

The benefit of this formulation is, that only $\mathbf{B}^{T/2}$ and $\mathbf{K}^{T/2}$ are needed for the transformation of the gradient, and — upon return from the minimisation procedure — $\mathbf{B}^{1/2}$ and $\mathbf{K}^{1/2}$ for the transformation of \mathbf{v} and \mathbf{w} back to the increments $\delta \mathbf{x}_0$ and $\delta \mathbf{u}$, leading to the optimised state variables

$$\mathbf{x}_0 = \mathbf{B}^{1/2} \mathbf{v} + \mathbf{x}^b \quad \text{and} \quad \mathbf{u} = \mathbf{K}^{1/2} \mathbf{w} + \mathbf{u}^b. \quad (4.5)$$

Provided initial values and background fields coincide at optimisation start, the variables \mathbf{v} and \mathbf{w} equal zero and, upon saving for later use at all subsequent iterations, the inverse square roots of the covariance matrices in (4.2) are never needed. Hence, having a suitable square root decomposition of the background error covariance matrices, a sophisticated preconditioning of the minimisation problem is available and a balanced minimisation can be performed.

4.2 Initial Value Background Error Covariance Matrix

Since the background error covariance matrix \mathbf{B} is symmetric, it can be decomposed in a diagonal matrix $\mathbf{\Sigma}$ and a correlation matrix \mathbf{C} , such that

$$\mathbf{B} = \mathbf{\Sigma} \mathbf{C} \mathbf{\Sigma}$$

holds. The main diagonal of $\mathbf{\Sigma}$ contains estimates of the background error standard deviations, which are given for the different constituents in Appendix B. Therefore, providing \mathbf{B} can be split in the specification of standard deviations of the model parameters and by finding a suitable factorisation $\mathbf{C} = \mathbf{C}^{1/2} \mathbf{C}^{T/2}$ of the correlation matrix, such that

$$\begin{aligned} \mathbf{B} &= \mathbf{B}^{1/2} \mathbf{B}^{T/2} \\ &= \mathbf{\Sigma} \mathbf{C}^{1/2} \mathbf{C}^{T/2} \mathbf{\Sigma}. \end{aligned} \tag{4.6}$$

Following *Weaver and Courtier* [2001], a Gaussian function, which only depends on the Euclidean distance $r = |z - z'|$ between two points, is homogeneous and isotropic and it defines a valid positive definite correlation function. Therefore, the classical diffusion equation, the solution of which is given by the convolution of the functional value at the origin with a Gaussian, serves for a valid operator generating the correlation function. The authors generalise the classical diffusion equation to a larger class of *generalised diffusion equations* (GDE), allowing even to account for negative correlations on the sphere, featuring, for example, radially fading pressure waves often encountered in meteorological conditions.

The advantage of using a diffusion approach to provide the correlation matrix as an operator is, that the factorisation by the square root and its transpose can be easily obtained by adjusting the characteristic parameters of the diffusion process, that is the length scales, given by the diffusion coefficient and the integration time. *Weaver and Courtier* [2001] emphasise, that the square root is yielded when applying a diffusion operator only half the integration time. Moreover, the Laplacian, encountered in the diffusion equation, is self adjoint. This fundamental property can be exploited to identify a factorisation $\mathbf{C}^{1/2} \mathbf{C}^{T/2}$ of the correlation matrix:

Let \mathbf{L} denote a discretised solver to the diffusion equation in an explicit forward differencing form,

$$\mathbf{L} = \{\mathbf{I} + \kappa \Delta t \mathbf{D}\}^M,$$

such that $\eta(t_M) = \mathbf{L} \eta(t_0)$ holds for any arbitrary scalar field η . The total number of integration steps is given by M , hence $T \equiv t_M = M \Delta t$, and \mathbf{D} is

a matrix representation of the discretised Laplacian, being self-adjoint with respect to a diagonal matrix \mathbf{W} , which contains the area elements of the grid. The symmetric representation of the diffusion operator can be written as (*Weaver and Courtier [2001]*)

$$\mathbf{LW}^{-1} = \mathbf{L}^{1/2} \mathbf{W}^{-1} \mathbf{L}^{T/2}. \quad (4.7)$$

with \mathbf{L} and thus $\mathbf{L}^{1/2}$ being self-adjoint with respect to matrix \mathbf{W} . Hence, the correlation matrix is obtained by multiplying the scalar field with the inverse metric \mathbf{W}^{-1} and applying the discretised diffusion operator over M time steps, representing the desired integration time while taking into account stability properties of the diffusion solver. Without loss of generality be M an even number of time steps. Then the square root of the diffusion operator is obtained by

$$\mathbf{L}^{1/2} = \{\mathbf{I} + \kappa \Delta t \mathbf{D}\}^{M/2}$$

Finally, the symmetric factorisation of \mathbf{B} reads

$$\begin{aligned} \mathbf{B} &= \mathbf{\Sigma} \mathbf{C} \mathbf{\Sigma} \\ &= \mathbf{\Sigma} \mathbf{C}^{1/2} \mathbf{C}^{T/2} \mathbf{\Sigma} \\ &= (\mathbf{\Sigma} \mathbf{\Lambda} \mathbf{L}^{1/2} \mathbf{W}^{-1/2}) (\mathbf{W}^{-1/2} \mathbf{L}^{T/2} \mathbf{\Lambda} \mathbf{\Sigma}) \\ &= \mathbf{B}^{1/2} \mathbf{B}^{T/2} \end{aligned} \quad (4.8)$$

where $\mathbf{\Lambda}$ denotes a diagonal normalisation matrix — here introduced on both sides to maintain symmetry. This ensures that the correlation matrix \mathbf{C} has unit variance and therefore the main diagonal of \mathbf{B} consists of the variances of the model state variables, as encoded in $\mathbf{\Sigma}$. In an unlimited domain, the normalisation factor can be derived analytically for the isotropic case. For a regional model with boundaries in both vertical and horizontal directions and, moreover, using an inhomogeneous diffusion operator, which will be introduced below, the normalisation matrix has to be rendered numerically. The normalisation matrix $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$ is yielded by applying the complete filter to unit vectors $\mathbf{e}_l = (0, \dots, 0, 1, 0, \dots, 0)^T$, the entry equal to one being located at the l^{th} grid point. The normalisation coefficients are then defined by

$$\lambda_l = \frac{1}{\sqrt{t_l}} \quad \text{with} \quad t_l = \mathbf{e}_l^T \mathbf{LW}^{-1} \mathbf{e}_l.$$

If different spatial correlations are specified for the different constituents, this method has to be performed for all species, resulting in a very expensive algorithm if performed sequentially, since the number of state variables

exceeds 10^6 in the EURAD model system. *Weaver and Courtier* [2001] propose another method to estimate the variances of the filter \mathbf{LW}^{-1} approximately. This algorithm bases on a randomisation method. Therefore, the unnormalised correlation matrix \mathbf{LW}^{-1} is rewritten to

$$\mathbf{LW}^{-1} \approx \frac{1}{Q} \sum_{q=1}^Q \tilde{\mathbf{v}}_q \tilde{\mathbf{v}}_q^T = \frac{1}{Q} \sum_{q=1}^Q (\mathbf{L}^{1/2} \mathbf{W}^{-1/2} \mathbf{v}_q) (\mathbf{L}^{1/2} \mathbf{W}^{-1/2} \mathbf{v}_q)^T,$$

where $\tilde{\mathbf{v}} = \mathbf{L}^{1/2} \mathbf{W}^{-1/2} \mathbf{v}$, with \mathbf{v} being a Gaussian random vector, drawn from a population having zero mean and unit variance. The diagonal elements converge to the true variances (which are the inverse of the searched normalisation factors) when Q increases. Under the assumption of Gaussian pdf the standard error in the estimated variations is $1/(2Q)^{1/2}$.

Due to its feasibility on massive parallel machines like the IBM Regatta, the exact computation of the normalisation matrix, as discussed above, has been parallelised and implemented for the simulations shown in this thesis.

The diffusion length scale L , which is often referred to as *influence radius* in the framework of background error covariance modelling, is the characteristic parameter of the diffusion approach and given by the formula

$$L = \sqrt{2\kappa T}.$$

It specifies the distance of two model grid points, having a mutual correlation of $c_{r=L} = \exp(-1/2)$. To deduce the implementation in the EURAD 4d-var data assimilation system, Fig. 4.1 shows the SO₂ emission data base for eastern Germany, where a very noisy pattern can be seen, and we can be sure that the near surface SO₂ concentrations will show a comparable pattern. In contrast, O₃ profile retrievals for a height of 15 km, which is also shown, exhibit completely different length scales. In order to reflect the governing processes producing the correlations in the background errors, the chosen horizontal influence radii weakly increase from 1.5 km to 10 km within the boundary layer, while a correlation length of 250 km is assumed for the model top at 100 hPa (see Fig. 4.1). The vertical correlation length is determined by the vertical diffusion coefficient as provided by MM5, and hence the parametrisation uses the accessible knowledge about the true diffusivity of the atmosphere as a first approximation of valid influence radii. Applying the filter with these properties leads to the concentration increments shown in Fig. 4.1. The noisy character of surface SO₂ is conserved while the information contained in the satellite ozone tracks is smoothed on a length-scale of about 250 km.

In order to simulate spatial error correlations in three dimensions, *Weaver*

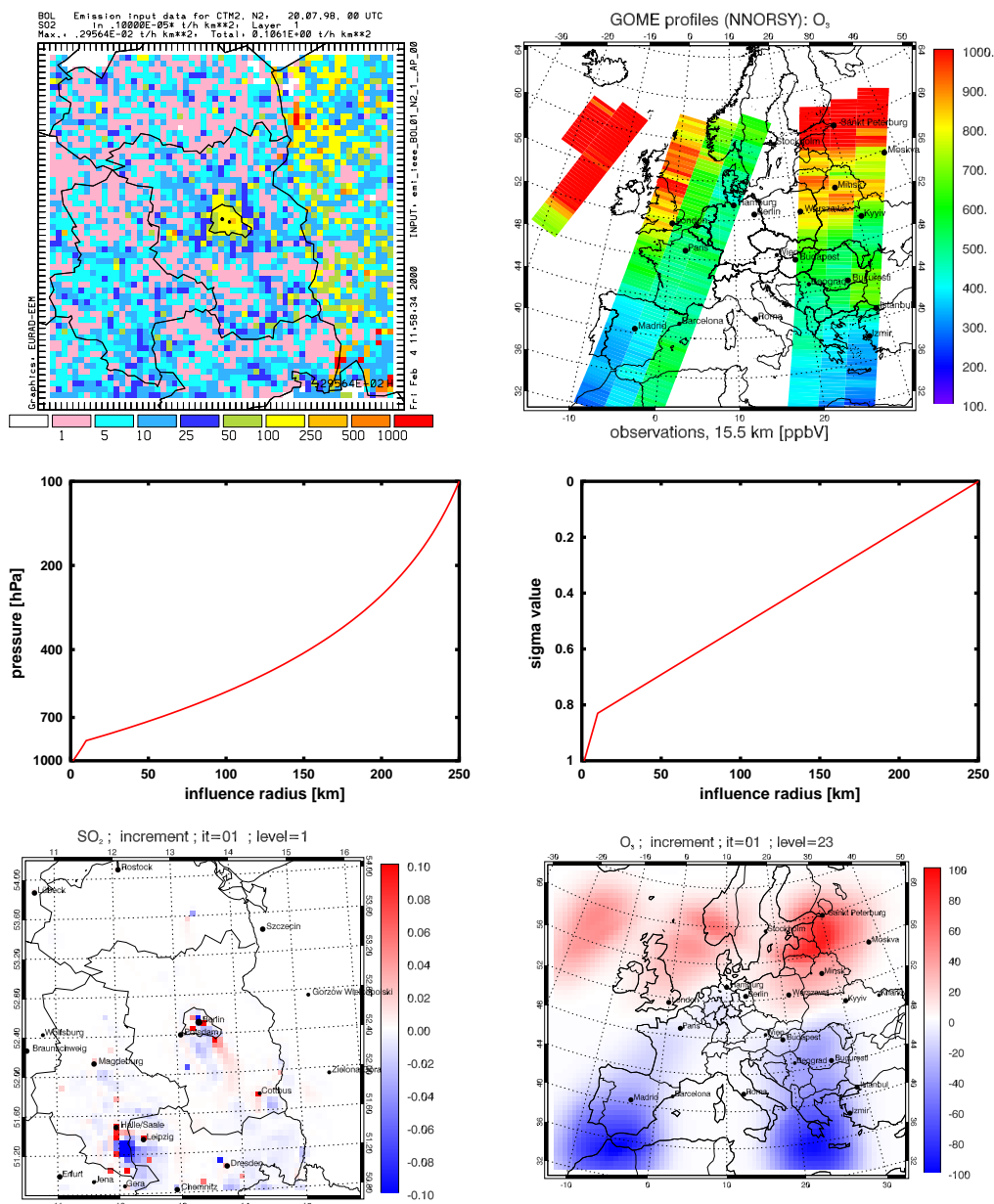


Figure 4.1: Parametrisation of horizontal influence radii and their impact on the analysis result. Top row: SO_2 emissions in $\text{t}/(\text{h km}^2)$ in the eastern Germany (left) and O_3 mixing ratios in 15.5 km height by satellite profile retrievals on January 18, 2002 (right). Middle row: Correlation length estimates with respect to model pressure (left) and sigma level value (right). Bottom row: Analysis increments derived from given correlation length for surface SO_2 (left) and ozone at ~15 km height (right).

and Courtier [2001] split both the diffusion operator \mathbf{L} and the metric \mathbf{W} in a horizontal and a vertical component,

$$\mathbf{L}\mathbf{W}^{-1} = \mathbf{L}_v\mathbf{L}_h\mathbf{W}_v^{-1}\mathbf{W}_h^{-1}.$$

However, due to both the vertical variations of the horizontal influence radii and the horizontal variations of the vertical influence radii, the three dimensional diffusion operator does not remain separable into symmetric square roots

$$\mathbf{L} = \mathbf{L}_h^{1/2}\mathbf{L}_v^{1/2}\mathbf{L}_v^{1/2}\mathbf{L}_h^{1/2}$$

The matrices \mathbf{L}_v and \mathbf{L}_h do not commute, since the spatial derivatives of the diffusion coefficients no longer vanish. However, *Weaver and Courtier* [2001] show, that a non-separable approximation of the square root of the three-dimensional diffusion operator can be achieved by adjusting the diffusion coefficients $\kappa_{v,h}$, the integration time steps $\Delta t_{v,h}$ and the number of diffusion steps $M_{v,h}$ in a suitable way, while preserving the desired influence radii $L_{v,h}$. The skill of the approximation can be improved by increasing the number of integration time steps $M_{v,h}$.

For future developments, the diffusion approach also enables to simulate flow-dependent, *i. e.*, anisotropic and inhomogeneous due to known dynamic features, correlations in three dimensions, using coordinate stretching and rotation.

4.3 Emission Factor Background Error Covariance Matrix

As shown by *Elbern et al.* [2000], the EURAD 4d-var data assimilation system allows for emission rate optimisation. A suitable preconditioning of the minimisation problem as well as a limitation of degrees of freedom to the optimisation problem can be achieved by a proper emission factor error covariance matrix. Compared to \mathbf{B} , the emission factor background error

Table 4.1: *Estimated standard deviations of the transformed emission factors (personal information by Michael Memmesheimer). The others section comprises all volatile organic compounds (VOCs) as well as NO₂, CO and SO₄.*

Species	Standard Deviation
SO ₂	$\ln(1.7) = 0.53$
NH ₃	$\ln(1.7) = 0.53$
NO	$\ln(1.3) = 0.26$
others	$\ln(2.0) = 0.69$

covariance matrix \mathbf{K} is substantially easier to obtain under the assumption, that spatial correlations between emission rates can be neglected. Moreover, we will assume, that the remaining multivariate correlations between emitted species in one grid cell do not vary over the whole domain. Hence, the correlations to be coded reduce to a symmetric square matrix of size $m \times m$, when m is the number of emitted species, provided by the emission model EEM.

In accordance with the specification of the initial value background error covariance matrix and as needed by the incremental formulation of the cost function, we write \mathbf{K} as a factorisation

$$\mathbf{K} = \mathbf{\Gamma} \mathbf{D}^{1/2} \mathbf{D}^{T/2} \mathbf{\Gamma}, \quad (4.9)$$

where $\mathbf{\Gamma}$ is a diagonal matrix containing the standard deviations of the emission factors. Equivalent to the initial values, we can use $\mathbf{D}^{T/2}\mathbf{\Gamma}$ for the transformation of the gradient in (4.4) and $\mathbf{\Gamma}\mathbf{D}^{1/2}$ for the transformation (4.5) back to the emission factors.

The standard deviations coded in $\mathbf{\Gamma}$ apply to the transformed state variable \mathbf{u} . Therefore we need to transform the standard deviations as well, leading to the needed Gaussian error distributions. The current values in $\mathbf{\Gamma}$ are given in Table 4.1. Besides SO_2 , NH_3 and NO , which are thought of to be slightly better estimated, the error bar on an emission factor of $f = 1.0$ ranges from 0.5 to 2.0, resulting in a transformed emission factor $u = 0 \pm 0.69$ (all estimates by personal communication with M. Memmesheimer).

To estimate the background error correlation matrix \mathbf{D} , annual amounts of emitted NO_x , SO_x and volatile organic compounds (VOCs) divided into different polluter groups have been analysed. The statistical data has been provided by EMEP (for further information the reader may refer to www.emep.int, www.iiasa.ac.at and www.rivm.nl/edgar) and used during model simulations for the AZUR project (*Memmesheimer et al.* [2004b]). Starting from scratch, the block diagonal matrix $\mathbf{D}^{1/2}$ has been encoded in such a way, that the significant multivariate correlations between the emitted constituents are represented in \mathbf{D} . This method ensures a positive definite correlation matrix \mathbf{D} and provides directly the factorisation $\mathbf{D} = \mathbf{D}^{1/2}\mathbf{D}^{T/2}$. The upper triangle sub-matrix of \mathbf{D} is given in Fig. 4.2. Correlations greater than 0.1 are encountered, for example between NO_x and CO , SO_x and NO_x , and between species, which are chemically related to each other, like SO_2 and SO_4 , NO and NO_2 or families within the VOCs (alkanes, alkenes, carbonyl compounds, aromatic hydrocarbons). Additionally, there are reasonable correlations between those constituents which are emitted in mixed (deciduous and coniferous) woodland, namely OLT (ter-

	SO2	SO4	NO2	NO	ALD	HCHO	NH3	HC3	HC5	HC8	ETH	CO	OL2	OLT	OLI	TOL	XYL	KET	ISO
SO2	100	40	12	12	0.4	0.4	0.4	0.4	0.4	0.4	0.4	9	0.4	0.4	0.4	0.4	0.4	0.4	0
SO4		100	12	12	0.4	0.4	0.4	0.4	0.4	0.4	0.4	9	0.4	0.4	0.4	0.4	0.4	0.4	0
NO2			100	26	8	8	7	8	8	8	8	12	8	8	8	7	7	8	0.4
NO				100	8	8	7	8	8	8	8	12	8	8	8	7	7	8	0.4
ALD					100	26	5	6	6	6	6	5	10	6	6	5	5	26	0.3
HCHO						100	5	6	6	6	6	5	10	6	6	5	5	26	0.3
NH3							100	5	5	5	5	5	5	5	5	5	5	5	0.3
HC3								100	23	23	23	5	18	15	15	6	6	6	0.3
HC5									100	23	23	5	18	15	15	6	6	6	0.8
HC8										100	23	5	18	15	15	6	6	6	0.8
ETH											100	5	18	15	15	6	6	6	0.8
CO												100	6	5	5	5	5	5	0.3
OL2													100	23	23	6	6	10	1
OLT														100	27	5	5	6	15
OLI															100	5	5	6	15
TOL																100	24	5	0.3
XYL																	100	5	0.3
KET																		100	0.3
ISO																			100

> 10 %
 > 1 %
 < 1 %

Figure 4.2: The upper sub-matrix of the symmetric emission factors background error correlation matrix. Values are given in percent. The different colour codes indicate different levels of importance. The shortcuts of emitted species or species groups are according to the chemistry mechanisms RADM2 (Stockwell et al. [1990]) and RACM (Stockwell et al. [1997]).

minal olefins), OLI (olefins with internal double bond), and ISO (isoprene and other alkenes with conjugated double bonds).

The described method of determining emission factor correlations implicitly underlies the assumption, that the fraction of a special emitted constituent due to a special pollutant group is homogeneous over the whole domain. Most probably, a better parametrisation would be available, if the emission data would already contain the fractional parts of the emitted masses according

to the different pollutant groups. Hence, the emission factor background error covariance matrix could be extended to include spatial variations.

4.4 Observation Error Covariance Matrix

The observation error covariance matrix is the sum of two other covariance matrices, namely the measurement error covariance matrix, containing the estimates of the statistical performance of the measuring instrument, and the representativeness error covariance matrix.

The representativeness error reflects the degree to which an observation represents the volume-average value. Representativeness errors are usually unbiased, and their magnitudes may depend on the volume size, the sensor's time average, the volume size which one wants the observation to represent, and the chemical weather state. In this work, the representativeness error covariance matrix is assumed to be diagonal, *i. e.*, the estimates are uncorrelated. This assumption usually holds, when the observing system is representative for the true observational state. Otherwise, the representativeness errors are correlated, but the information needed to specify cross-correlations in the covariance matrix is not yet available.

The specification of representativeness errors for ground based observations uses the information about the environment in which the measurement station is located. *Tilmes* [1999] used a classification consisting of five different location types: *remote*, *rural*, *suburban*, *urban* and *traffic*. Due to an increasing amount of available observational data and missing information about the stations, this scheme needed to be extended by the class *unknown*. A characteristic representativeness length has been assigned to each of these classes. This representativeness length is then compared to the current horizontal grid spacing and a grid dependent scaling factor is obtained. This factor is used to scale a species' characteristic error to retrieve an estimate of the representativeness error. For measurements like balloon soundings or lidar observations, rough estimates about the characteristic length scales are provided and the representativeness error is yielded in the same way as for ground based observations. Observations based on retrievals of satellite measurements coming with a footprint information are taken to perfectly represent the model equivalent, since footprint dimensions are larger or comparable to horizontal grid sizes. However, satellite profile observations, having a certain vertical resolution, are assigned a small representativeness error due to the usually coarser model grid in the vertical.

The estimates of the representativeness lengths and the resulting representativeness errors are given in Appendix B.

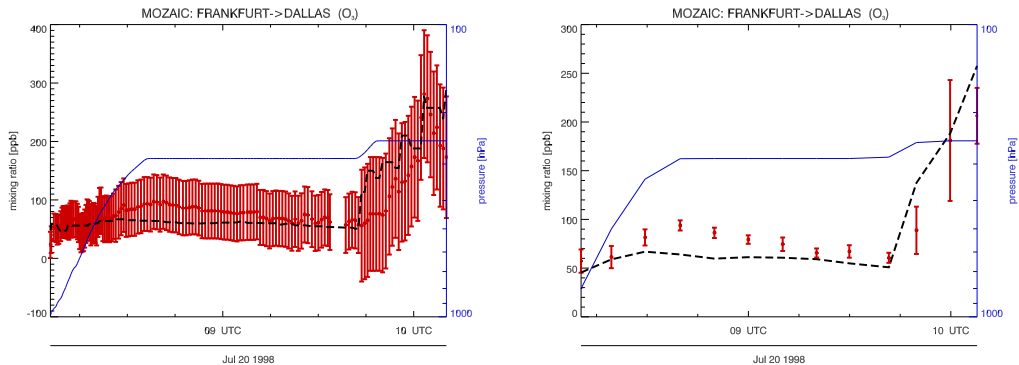


Figure 4.3: Impact of averaging high resolving air-borne measurements, here for ozone measured on a MOZAIC flight from Frankfurt to Dallas on July 20, 1998. Left: Original measurement data. Right: Averaged data with representativeness error included, derived from variance within averaged intervals. Black dashed line: Model realisation. Blue solid line: Flight level.

The same difficulties apply to the specification of measurement error cross-correlations. The assumption of mutual independency holds in most cases, since different measurement devices are used to observe different species. However, there may be time dependent auto-correlations which are difficult to account for. Anyway, the specification of a valid and suitable measurement error covariance matrix is up to institutions which collect and process observations. Often ground based observations miss a specification of the standard deviation. For these observations, minimum absolute and relative errors are defined following a proposal by *Mohnen* [1999]. These errors are used to assign uncertainties of ground based in-situ measurements and are given in Appendix B.

Another problem arising with unfeasible covariance statistics, is the unrealistic weight given to high frequency observations in the assimilation process, compared to the weight given to background knowledge. This is most crucial for air-borne observations, where measurement periods of 1 s can be encountered and no further informations on auto-correlations are available. To circumvent unrealistic analysis states inferred by such observations, an averaging approach has been implemented to construct superobservations. All observations of a certain species and measurement type for one dynamical model time step are averaged and the variance of the encountered observations is used to derive an estimate of the representativeness error for the averaged observation. The impact on the measurement density and the resulting error bars is shown in Fig. 4.3, where original air-borne ozone data from a MOZAIC flight is compared to the sparser superobservations.

Regions with a high variability in the measured ozone mixing ratios are assigned large error bars for the superobservations, due to the poor representativeness of the data for the specific grid cell.

CHAPTER 5

Campaign Simulations

In this chapter the main results of the data assimilation experiments for the selected campaigns will be shown. Further to some insights into the campaign goals, the characteristics of the chosen system configuration, including the extensions developed for this work, will be presented and evaluated. The first campaign simulation for BERLIOZ is used to introduce the main features of the EURAD 4d-var system, using the newly developed adjoint nesting technique and applying all optimisation modes, while focussing on boundary layer processes. Moreover, the observation basis is the richest in all simulations shown in this thesis, and therefore both a proper characteristics of the system performance for various constituents can be derived and a comparison of the optimisation modes can be achieved. Using the VERTIKO campaign data focussing on the boundary layer too, the data assimilation procedure is then applied over ten days allowing for a more comprehensive statistics and a deeper insight in the performance of the system over a longer period. The air-borne data campaigns CONTRACE and SPURT will then be used to investigate the model performance in the free and upper troposphere, combining in-situ air-borne observations and retrievals from satellite observations.

Validation of an assimilation result can only be given by comparison with independent observations, either withheld from the assimilation procedure and used for quality control within the assimilation window, or lying in the future and used to examine the forecast skill. Here, both methods will be

presented. For boundary layer campaign simulations with a relatively dense set of ground based stations, the a-posteriori validation focuses on ameliorated forecasts, because keeping back an arbitrary set of measurement stations could force the set of remaining observations towards an unfavourable representation for both the involved grid cells and the complete set of observations. However, owing to the mostly sparse spatial coverage of in-situ air-borne data and satellite observations on a daily basis, there is a probable poor influence on the forecast for another model region with observations during the following days. Therefore, the success of the simulations for the CONTRACE and SPURT campaigns will additionally be investigated by information withheld from the assimilation procedure.

Besides the cost function value, the conventional statistical measures to quantify an improved model simulation are biases and root mean square errors. In those cases, where a greater number of observations (with different confidence intervals) lies within one grid cell, the bias and rms will both be presented, normalised by the error estimations of the observations (for definition see Appendix C). In the case of several observations in one larger grid cell, a reduced bias does not necessarily mean a better estimate. Moreover, in most cases the cost function values will be normalised by the control run value in order to allow for a comparison of the model performance of different days.

Due to possible confusion about background knowledge on different nested grids, a short nomenclature related to the simulation runs will be given: Hereafter, all legacy simulations — *i. e.*, first guess runs without doing any data assimilation — will be called *control runs* (cr). On coarse grids, *first guess runs* (fgr) base on the best knowledge about the model system, having assimilated all observations *prior* to initial time. On nested grids, the first guess run uses background values provided by an analysis run on the respective mother domain, while the initial state is still based only on observations prior to the initial time (see also Section 3.5). An *analysis run* (ar) is the simulation having taken into account all available information and thus resting on the best knowledge about the system after data assimilation. This implies optimised boundary and initial values as well as improved emission factors, if subject to optimisation.

The different optimisation modes, namely emission factor (EF), initial value (IV) and joint emission factors-initial value (JT) optimisation, will always be denoted.

Further, all simulations performed in this work were carried out in episodic mode, such that meteorological analysis products by operational weather services were available and used by MM5 as initial conditions.



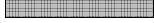






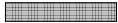






5.1 BERLIOZ Campaign

The main stimulus of the "Berliner Ozone Experiment" (BERLIOZ) campaign, which was part of the "Troposphärenforschungsschwerpunkt" (TFS) funded by the German Ministry of Science and Research, was to analyse the transport and chemical processes responsible for photo-oxidant formation and the contribution of groups of precursor substances to the ozone formation in the Berlin area and so to enhance process understanding of photochemical smog. To this end a measurement campaign — including many speciated hydrocarbons and carbonyl compounds besides common air quality species like ozone, nitrogen compounds, sulphur dioxide and carbon monoxide — has been designed and carried out from July 5 to August 7, 1998 in order to trace the Berlin plume up to a distance of about 100 km. Using the results of the measurement campaign should enable chemistry transport models to estimate the influence of emission regulation on air quality in densely populated regions significantly better, especially during summer smog episodes.

5.1.1 Model design and observational data basis

The simulations for the BERLIOZ campaign using the first application of the adjoint nesting technique, include a coarse grid (CG) with a horizontal grid spacing of 54 km encompassing Europe and parts of North-Africa.

Table 5.1: BERLIOZ: Simulation setup. Assimilation experiments are indicated by black solid bars (EF: emission factors, IV: initial values, JT: joint EF+IV), while forecasts are given by dashed areas.

		July 1998			
Nestlevel	Episode 1		Episode 2		
	18	19	20	21	
CG	 		 		
N1			 		
N2			 		

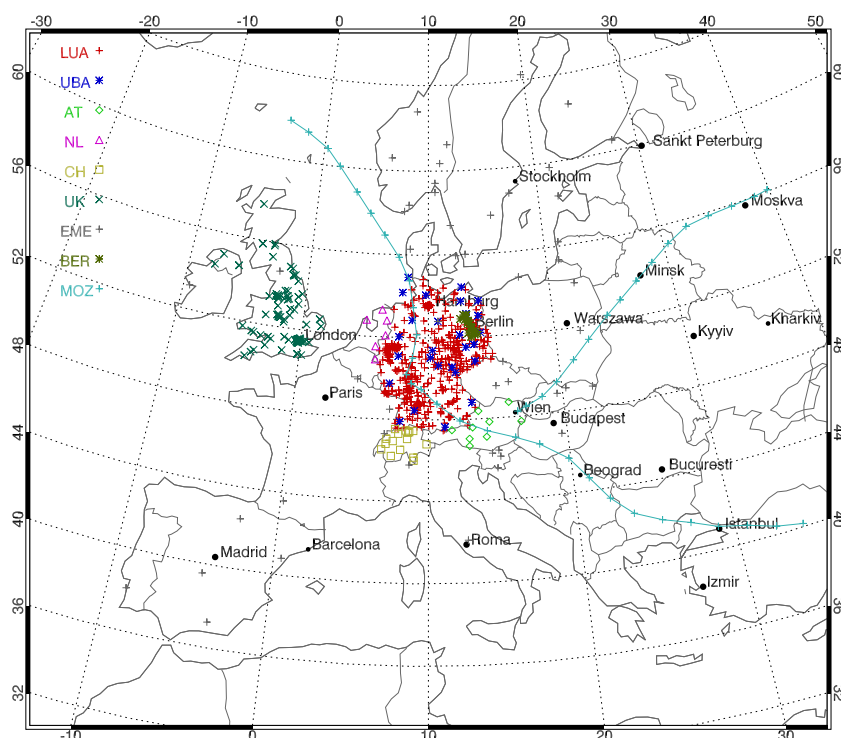
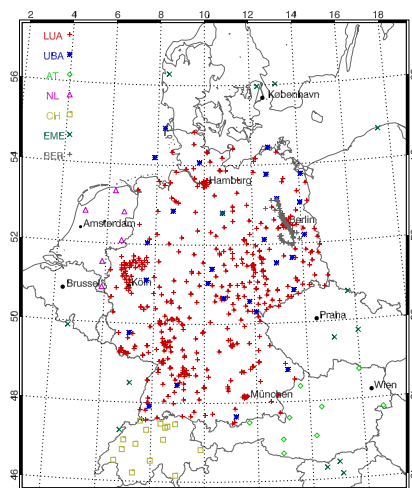
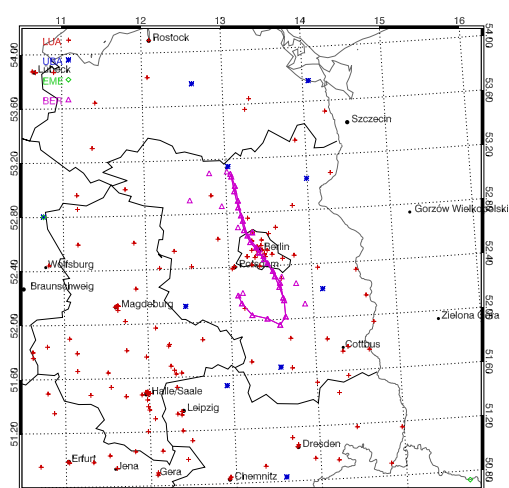
(a) Coarse Grid, $\Delta x = 54$ km(b) Nest 1, $\Delta x = 18$ km(c) Nest 2, $\Delta x = 6$ km

Figure 5.1: Nested model grid sequence for the BERLIOZ campaign simulations. The locations of in-situ measurements and air plane flight paths are indicated. The grid spacings are 54-18-6 km.

Two domains have been nested with a nesting ratio of three, resulting in a sequence of 54-18-6 km, telescoping from regional scale down to urban scale. The Nest 2 covers large parts of Eastern Germany with the greater Berlin area, being the region of main interest. The nesting sequence is shown in Fig. 5.1 together with the locations of in-situ measurements. The BERLIOZ campaign stations were located in a SE-NW line, due to a high probability of this wind direction during July and thus the possibility to trace the plume. However, the meteorological situation during BERLIOZ was exceptionally dominated by westerly wind regimes (*Volz-Thomas et al.* [2003]), and intensive observational periods (IOPs) took place only on July 20 and 21 and August 4 and 5, due to the detectability of the plume by the campaign stations. The days analysed by the EURAD 4d-var data assimilation system are July 18 to July 21, 1998.

The data assimilation procedure, which is sketched out in Table 5.1, was applied as follows: For July 18 (first day of episode 1), after a spin up run of six hours, observations from 06 UTC to 20 UTC have been assimilated in three modes, namely initial value (IV), emission factor (EF) and joint optimisation of initial values and emission factors (JT). Subsequent forecasts were then initiated for 48 hours on all three grids, based on the analysis result on the CG, delivering new initial values for July 20, 06 UTC. On July 20 (first day of episode 2), the data assimilation procedure has now been applied on all three domains and for all assimilation modes, again for 14 hours from 06 UTC to 20 UTC. A subsequent forecast for 42 hours then serves as quality control by means of improved forecast skill. For comparison, control runs have been performed without data assimilation.

To ensure a sufficient vertical resolution in the boundary layer, the number of vertical layers has been increased to 23, with about fifteen layers in the boundary layer and a representative height of about 19 meters in the lowest layer (see Appendix A).

The simulation results for the BERLIOZ campaign have been obtained by the newly implemented Rosenbrock solver of the RACM chemistry kernel, introduced in Section 3.3.

The observational basis consists of ground based measurements, tethered balloon and Lidar soundings, and air-borne data as well as retrievals from satellite based measurements. In-situ observations of sulphur dioxide (SO_2), nitrogen dioxide (NO_2), nitrogen oxide (NO), ozone (O_3), carbon monoxide (CO) and benzene (C_6H_6) were available as routinely gathered European measurements, provided by national institutions and the Co-operative programme for monitoring and evaluation of the long-range transmissions of air pollutants in Europe (EMEP). The BERLIOZ observational data has

been available as timely high-resolving measurements, comprising nitric acid (HNO_3), hydrogen peroxide (H_2O_2), formaldehyde (HCHO), peroxyacetyl-nitrate (PAN) and nitrogen oxides (NO_x) in addition to the constituents mentioned above.

Air-borne data by courtesy of the MOZAIC (Measurements of OZone and water vapour by in-service AIRbus airCRAFT) project and GOME (The Global Ozone Monitoring Experiment) partial satellite NO_2 columns provided by KNMI (Koninklijk Nederlands Meteorologisch Instituut) and GOME O_3 profiles retrieved by the NNORSY (Neural Network Ozone Retrieval System, Müller *et al.* [2003]) have been used additionally for the assimilation procedure on the Coarse Grid.

5.1.2 Simulation Results

The main focus of the assimilation experiments lies on the results on July 20 and a possible improvement of the forecast skill during Episode 2 (July 20 and 21), which can be interpreted as a-posteriori validation of the assimilation success.

Firstly, the impact of assimilated satellite retrievals from July 18 on the first guess forecast skill on Episode 2 will be shortly outlined. To this end, the averaged NNORSY ozone profile retrievals from GOME are shown in Fig. 5.2. The retrieved profiles from July 20 have been averaged for each observation height and been divided into four latitudinal regions, in order to take into account the northward decrease of the tropopause height and the associated increase of ozone concentrations at equivalent heights. Besides the remarkably good analysis quality on July 20, the first guess shows small to moderate improvements compared to the control run, most notably in the middle and upper troposphere. This is due to the information introduced in the model by assimilating the observations of July 18 and spreading with an influence radius. The positive effect is the stronger, the weaker the daily and lateral variability in the observations is and the higher the coverage of the model domain by the measurements' footprints is. The standard deviation according to the variance of the retrievals has been drawn as the error bars on the averaged observations. Due to the higher spatial coverage, the quality of the first guess is increasing with increasing latitude. The analysis gain by assimilated tropospheric O_3 profile retrievals from July 18 therefore significantly remained in the model over a period of at least 48 hours.

Fig. 5.3 shows joint optimisation results for tropospheric NO_2 columns on July 20, retrieved from GOME observations by KNMI. The observational data, which comes with averaging kernel (AK) and footprint information,

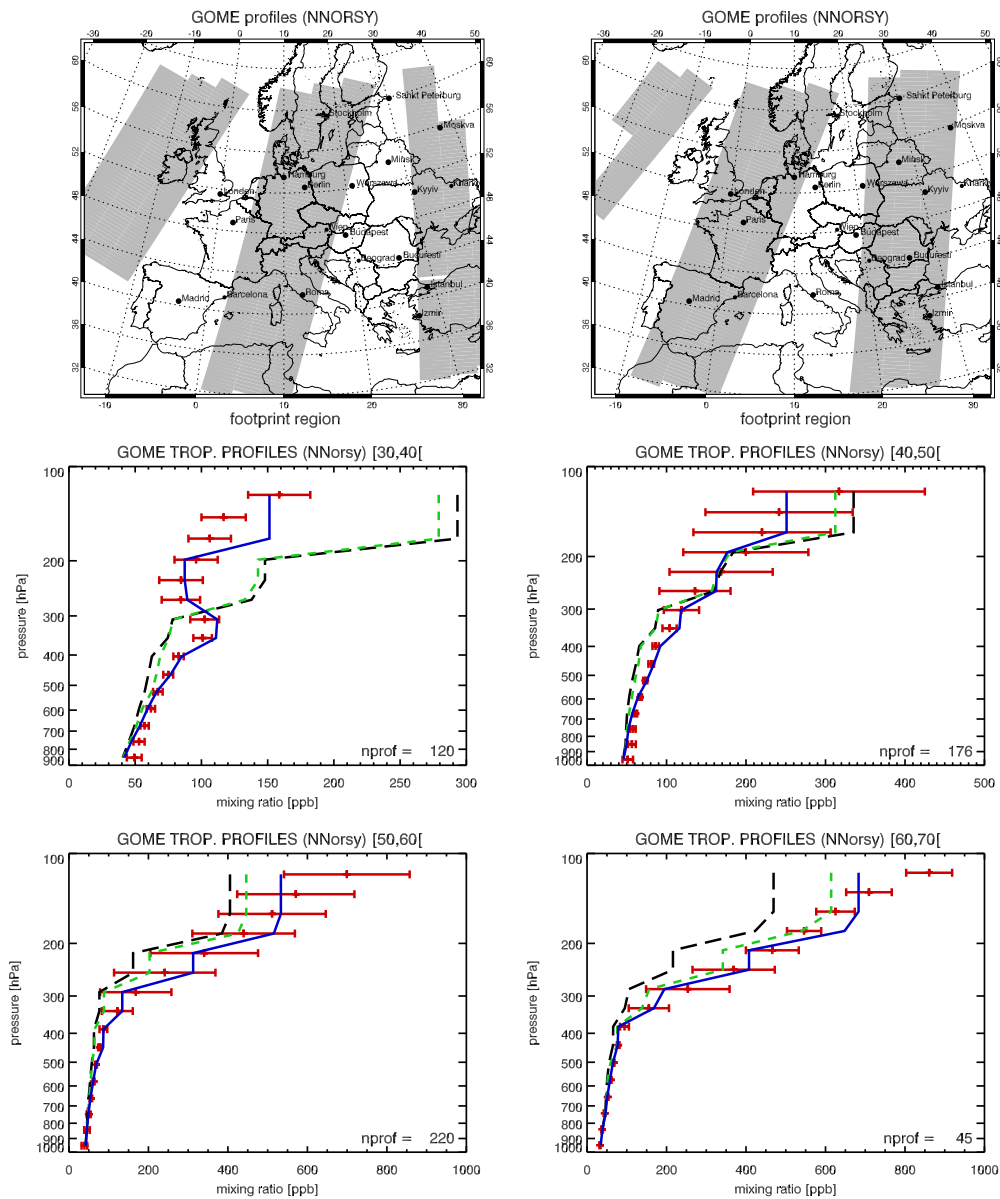


Figure 5.2: Coarse Grid: Footprints of GOME ozone profile retrievals on July 18 (upper left) and July 20 (upper right), 1998. Lower panels: Averaged ozone profile retrievals (NNORSY) together with the averaged model equivalents for July 20 (Episode 2), broken down by their latitudinal positions (upper left: 30° - 40° N, upper right: 40° - 50° N, lower left: 50° - 60° N, lower right: 60° - 70° N). The error bars on the observations (red) are given by the square root of the variance of the averaged profile values at respective height. Dashed black line: Control run. Dashed green line: First guess run. Solid blue line: Analysis run.

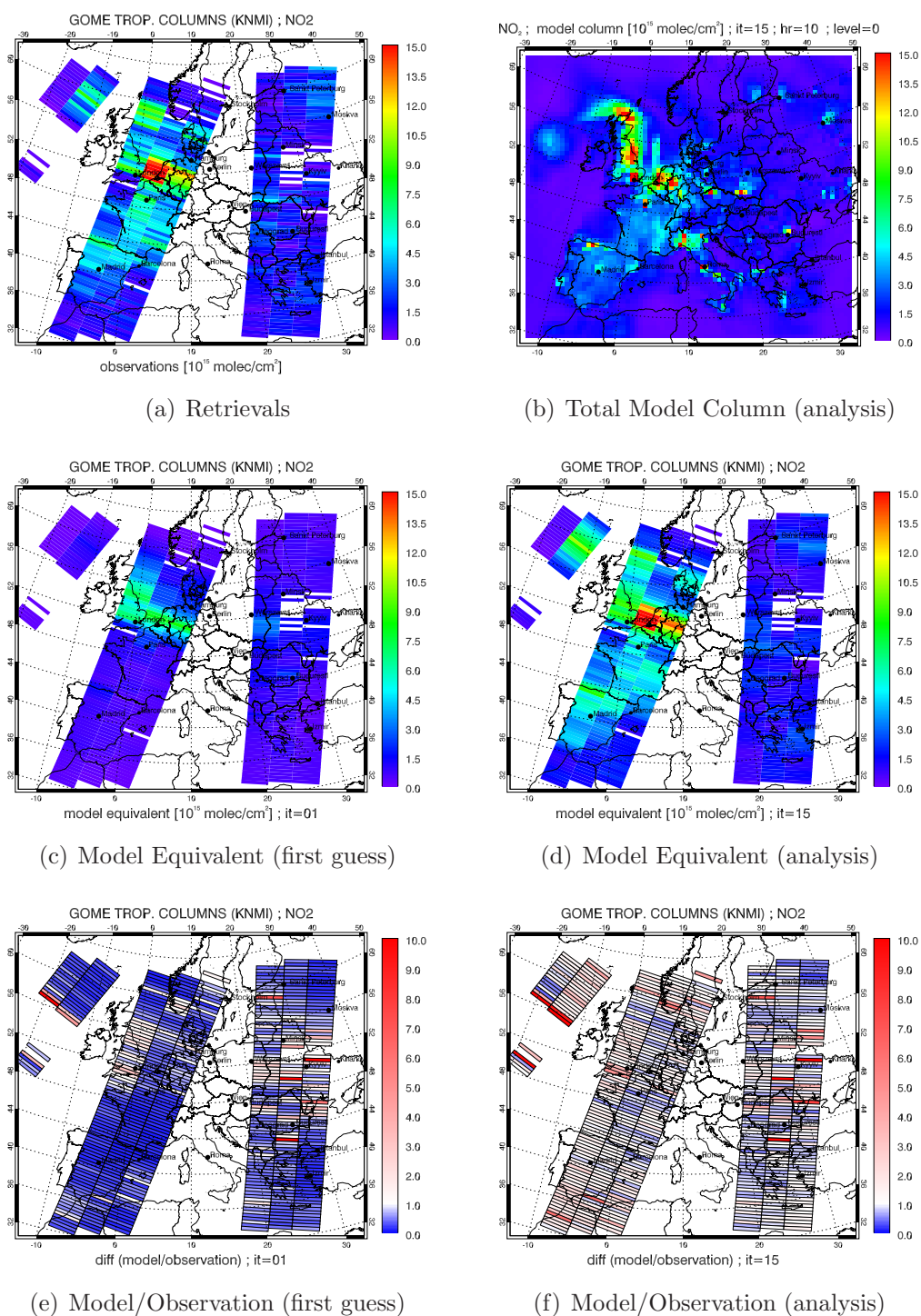


Figure 5.3: BERLIOZ CG: Assimilation results for GOME NO_2 tropospheric columns on July 20, 1998. Column densities (a-d) are in $10^{15} \text{ molec/cm}^2$.

shows features with elevated NO_2 column densities coincidently with higher expected NO_2 surface concentrations, namely the area from London over the Channel and Netherlands/Belgium to the German Rhein-Ruhr area as well as Paris and the Spanish north-coast (Fig. 5.3a). Although underestimated, the higher NO_2 column densities over the Channel and the adjacent areas already become apparent in the model equivalents by the first guess model simulation, while the simulation lacks a satisfying forecast for all other areas (Fig. 5.3c). This is obvious in the model equivalent to observation ratios by the blue values (Fig. 5.3e). However, the structures in the observed NO_2 column densities can be reproduced by the analysis run in a remarkable manner (Fig. 5.3d), leading to model equivalent to observation ratios of about 1.0 over the whole model domain (Fig. 5.3f). The remaining discrepancies coincide with very low observed values, which the model does not reproduce. To emphasise the kind of information contained in the observed column densities including footprint information, the total tropospheric model columns as produced by the analysis run is given in Fig. 5.3b, too. Here, the locations with high NO_2 surface concentrations also push through in the tropospheric column density, but concentrated ashore, where congested areas reside. The higher NO_2 column densities over Scotland and the Atlantic Ocean are linked to a cyclonic low, transporting polluted air masses, which originate from surface levels, over long distances.

A comparison of the data assimilation skill on the coarse grid with respect to the optimisation modes is pointed out in Fig. 5.4 for the ground based observations. The analysis run costs, normalised by the control run costs, are given for each assimilated species and the three different optimisation modes. The cost function values have been normalised to assure independence from the number of observations and thus to yield comparability of the different constituents. Obviously, there is no general characterisation concerning the different optimisation modes. A satisfactory cost reduction for SO_2 , O_3 , HNO_3 , HCHO , PAN and CO can be stated, with different sensitivities of the species to the optimisation modes. It can be stated, that here the joint optimisation mode joins the benefits of both initial value and emission rate optimisation and thus is superior to the two other modes. Note that there are very few observations for HNO_3 , H_2O_2 , HCHO , PAN and NO_y , so that very little "focus" is given to those species in the minimisation procedure and therefore the statistics is very crude. Time series for these species with sparsely measurements will be shown below using a nested model simulation. Further, it turns out that the gain is negligible for NO , NO_2 and benzene irrespective of the optimisation mode. This is thought of

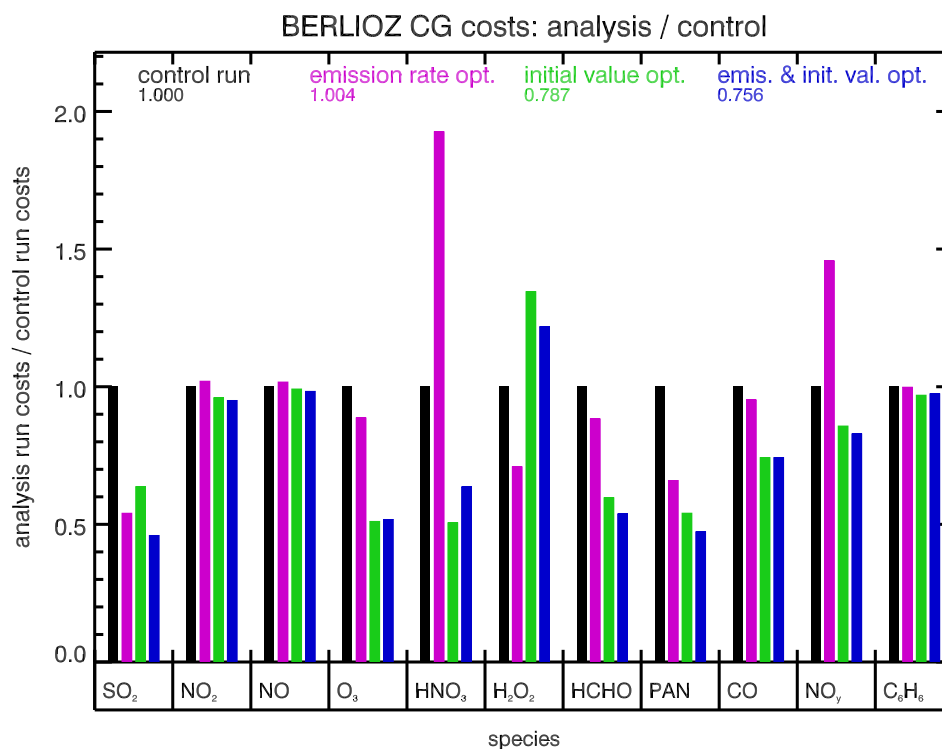


Figure 5.4: *BERLIOZ CG: Cost function ratios $J_{\text{ar}}/J_{\text{cr}}$ of the analysis runs normalised by the respective costs of the control run, given for ground based observations with respect to the assimilated species and the different optimisation modes. Numbers in the figure key indicate values averaged over all species.*

to be mainly due to the bad representativeness of the measurements for a grid cell having coarse model resolution. Most of the measurement devices are located in densely populated areas, mostly in the vicinity of transportation routes. Therefore, having coarse resolution, these situations cannot be sufficiently modelled. To emphasise this point, the available observations of NO_x species, NO_y , SO_2 , O_3 and CO for grid cell $(i, j, k) = (43, 39, 1)$ are plotted in Fig. 5.5, with (i, j, k) being the spatial indices in x , y and z direction, respectively. This box covers the south-eastern part of Berlin and part of its surroundings. While a large number of ozone observations are available and the analysis run forecasts for initial value and joint optimisation are already satisfactory, the observations for NO and NO_2 exhibit that very different characteristics reside in the same grid cell, while the mixing ratios are generally strongly underestimated. It must be stated, that, consequently, the representativeness errors for some observations within that grid

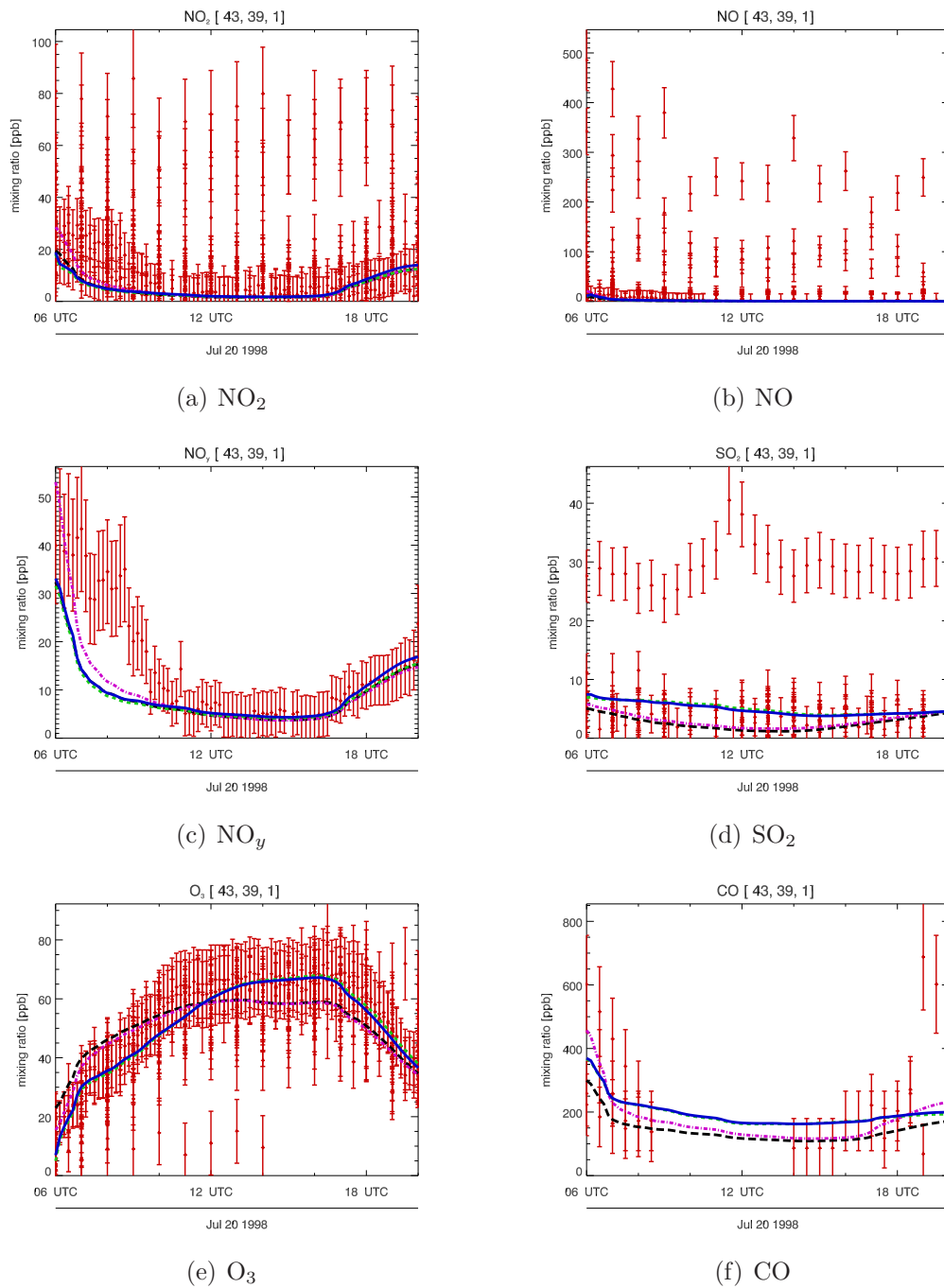


Figure 5.5: BERLIOZ CG: Observations (red) and model realisation in grid cell (43,39,1), partly covering the south-eastern part of Berlin. Black dashed line: Control run. Magenta dashed line: Analysis run, emission factor optimisation. Green dashed line: Analysis run, initial value optimisation. Solid blue line: Analysis run, joint optimisation.

cell have been estimated too low, resulting in the obvious outlier situation. Hence, one stimulus of nested data assimilation is to further increase the grid resolution such that higher resolving simulations will result in better representativeness of species like NO and NO₂ as well as SO₂, which have a strong local variability and, additionally for the NO_x species, a strong chemical coupling with a variety of other species due to a large number of reaction paths. Therefore higher resolved simulations by nesting are required, to enable improved forecast skills even for those species.

For the sake of a concise discussion, the nested simulations will be shown hereafter only for simulations applying joint optimisation, which shaped out to be superior to the single initial value and emission factor optimisation. It must be noted, that the time series, which are shown in the following, encompass the whole Episode 2, including the forecast for quality control, while only observations from the first 14 hours (from 06 to 20 UTC) have been assimilated (grey shaded areas in the time series plots).

Fig. 5.6 presents SO₂, NO₂ and NO observations in the grid cells (40, 40, 1) and (40, 41, 1) of the first nested domain N1. These two surface level boxes are both nested grid cells of the CG box shown in Fig. 5.5 and thus covering the south-eastern part of Berlin, too. The NO and NO₂ observations in grid cell (40, 40, 1) still exhibit a strong variety, especially for NO, where a huge range of different observed values in one grid cell remain. However, there is no strong outlier left for the SO₂ measurements in both grid cells, while the overestimation of the mixing ratio is efficiently reduced, leading to very improved forecast behaviour. Moreover, the representativeness of the NO and NO₂ observations in grid box (40, 41, 1) could be significantly advanced. While a slight underestimation of the observations remain, the nested simulation notably enhances the forecast skill. Moreover, the forecast skill is also improved for the second day. The analysis procedure therefore achieves to keep information in the model state (by means of the initial values and the emission factors) to produce a better forecast. To emphasise this property, Figs. 5.7-5.8 show observations and model results for selected stations on the second nestlevel N2. The forecast skill is very good to excellent, not only for the assimilation window but also for the subsequent forecast. Concentration peaks, which have not been predicted by the control run, can be reproduced in a fully satisfying way by the analysis run. Hence, observed time series of species which have a strong local variability have become sufficiently predicted by optimised model simulations with a spatial resolution of 6 km. Consequently, besides the facility of improving the horizontal resolution by nesting, the feature of joint initial value and emission rate optimisation appears to be prerequisite for skillful forecast of emitted

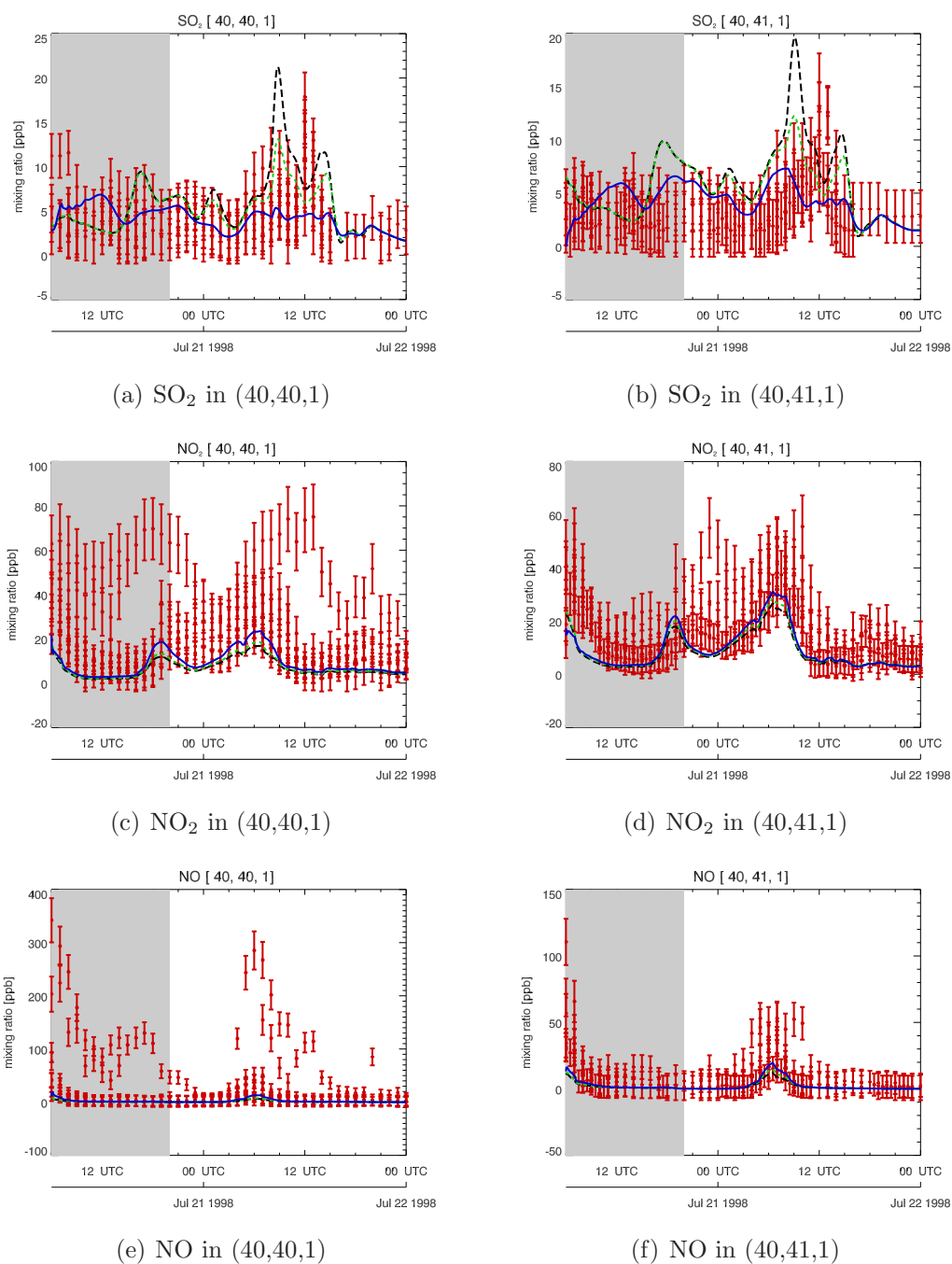


Figure 5.6: BERLIOZ N1: Observations (red) and model realisations for SO_2 , NO_2 and NO in grid cells (40,40,1) (left panels) and (40,41,1) (right panels) during episode 2. Black dashed line: Control run. Green dashed line: First guess run. Solid blue line: Analysis run. Assimilated observations: 06-20 UTC on July 20 (grey shaded).

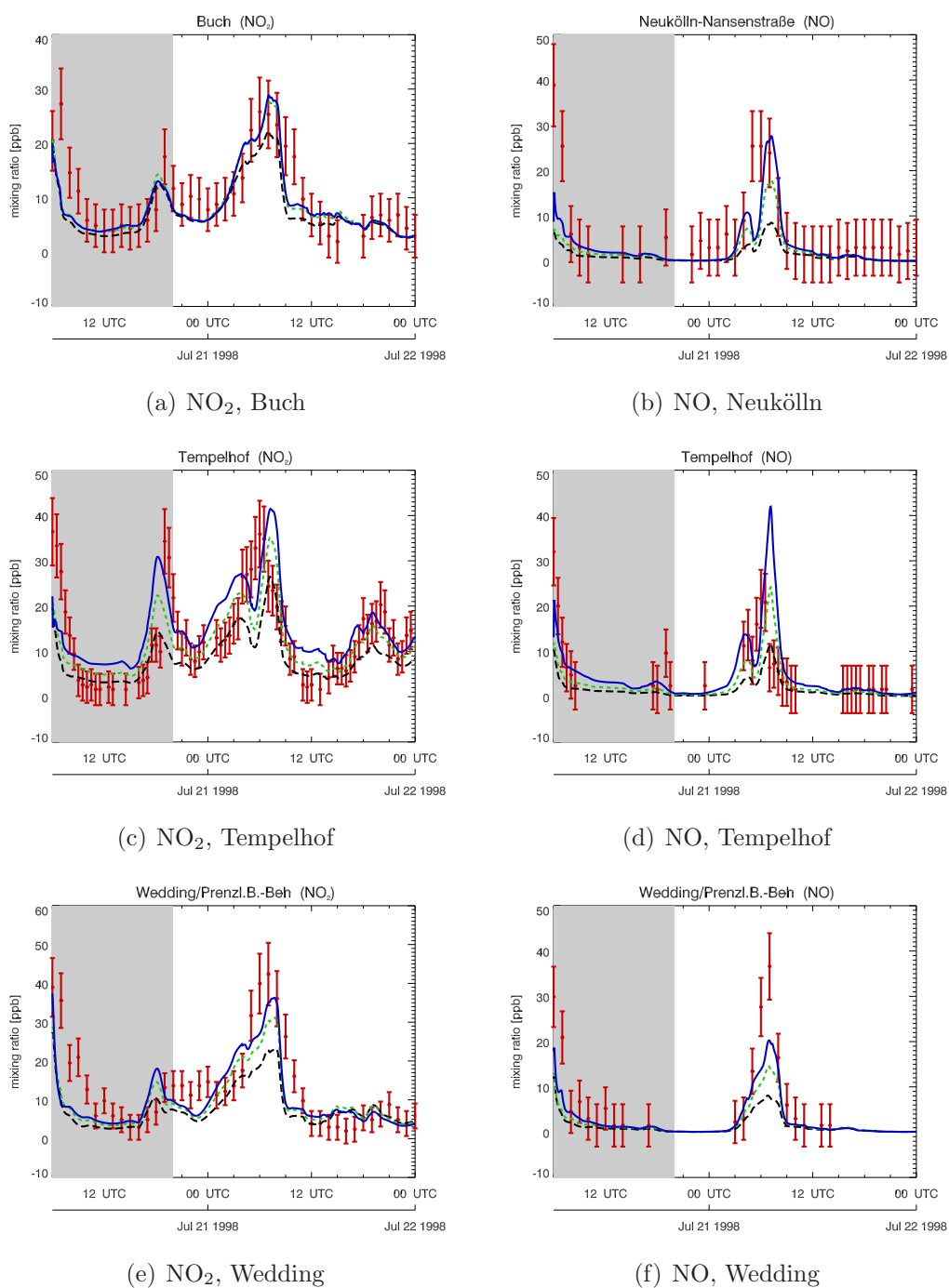


Figure 5.7: BERLIOZ N2: Observations (red) and model realisations for nitrogen oxides at selected stations during episode 2. Black dashed line: Control run. Green dashed line: First guess run. Solid blue line: Analysis run. Assimilated observations: 06-20 UTC on July 20 (grey shaded).

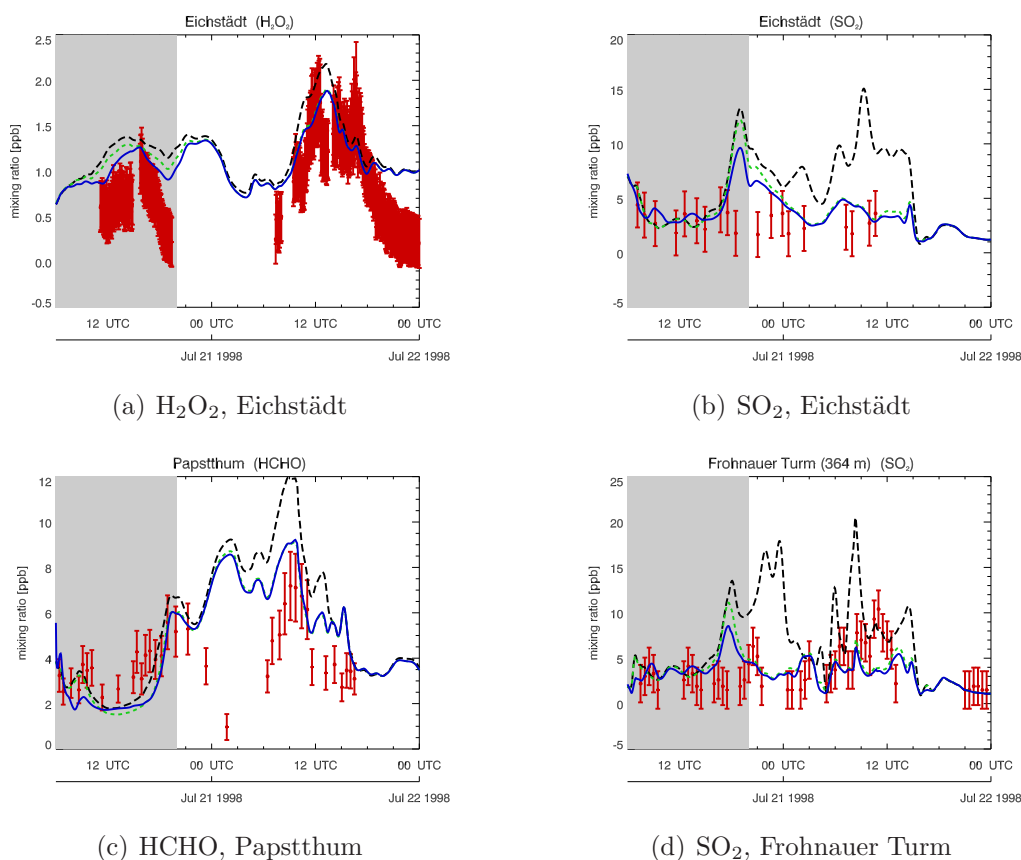


Figure 5.8: BERLIOZ N2: Observations (red) and model realisations for hydrogen peroxide, sulphur dioxide and formaldehyde at selected stations during episode 2. Black dashed line: Control run. Green dashed line: First guess run. Solid blue line: Analysis run. Assimilated observations: 06-20 UTC on July 20 (grey shaded).

species. These results will be investigated in a more comprehensive way and on a broader statistical basis in the course of the VERTIKO simulations in Section 5.2.

As a result of the assimilation procedure, the analysed emission factors for sulphur dioxide, nitrogen oxide, carbon monoxide and xylene are given in Fig. 5.9. The different optimisation stages in terms of boxes of Coarse Grid and Nests 1 and 2 are obvious. Each nestlevel refines the previous determined scaling factors. The resulting emission factors show moderate values, thus the analysis runs do not enforce a vigorous change in emission inventory to reproduce the observational information. These optimised emission rates are a direct estimate of improved surface fluxes.

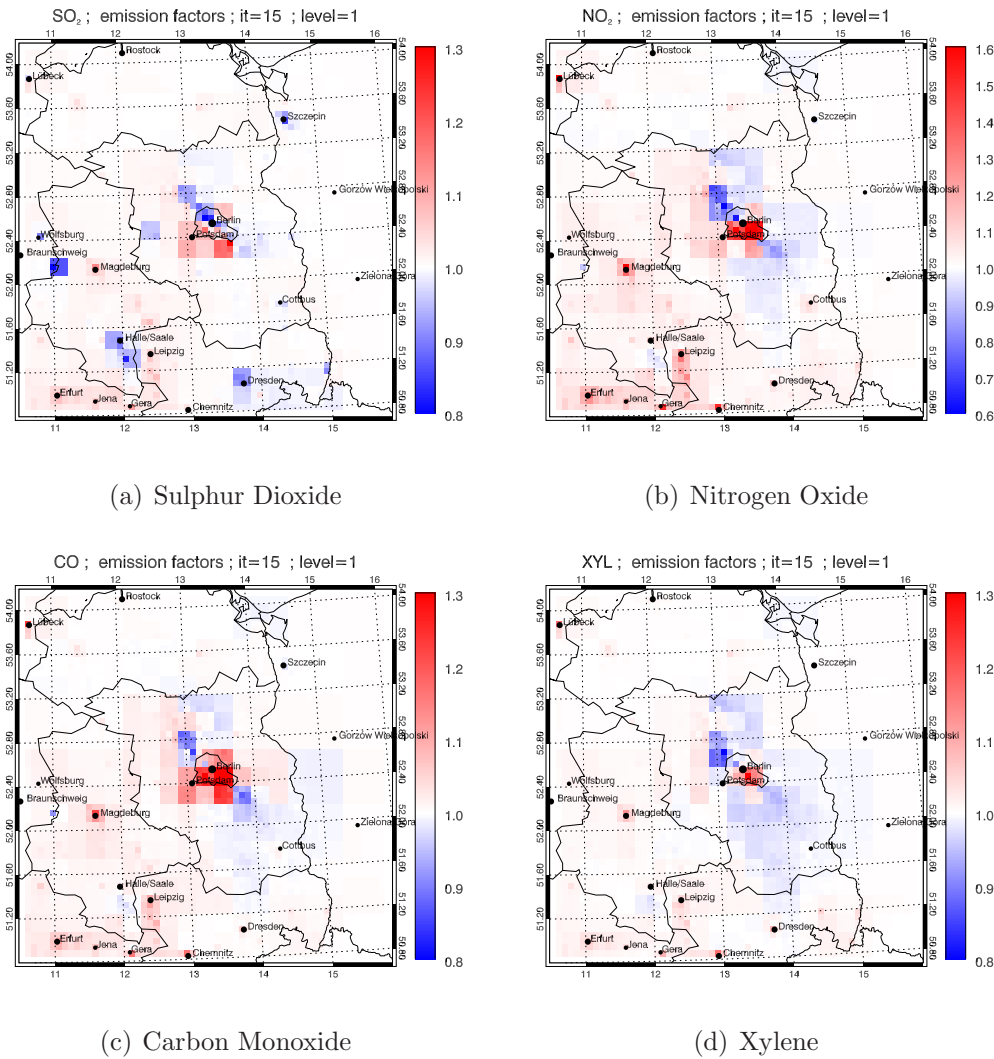
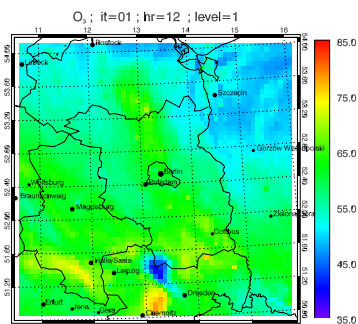
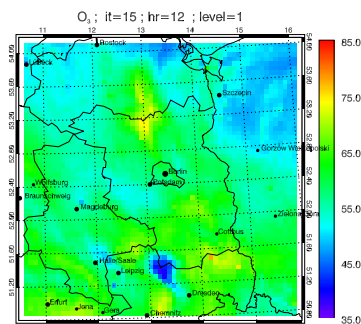


Figure 5.9: BERLIOZ N2: Optimised emission factors for SO₂, NO, CO and xylene as obtained by joint optimisation on episode 2.

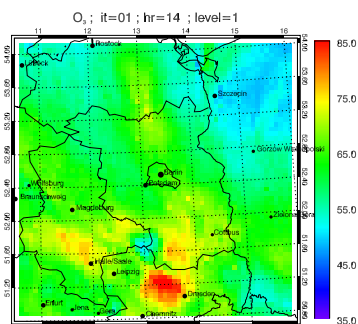
One core objective of the BERLIOZ campaign was to trace the Berlin ozone plume. The implication of the assimilation procedure on the simulation of the ozone plume is shown in Fig. 5.10, where ozone surface mixing ratios during the early afternoon on July 20, 1998, are presented. The analysis run significantly strengthens the ozone mixing ratios in the north or leeward of Berlin. The gained insight in the involved processes needed for the production of the plume and its transformation are therefore most qualified answers on the questions asked by the BERLIOZ project.



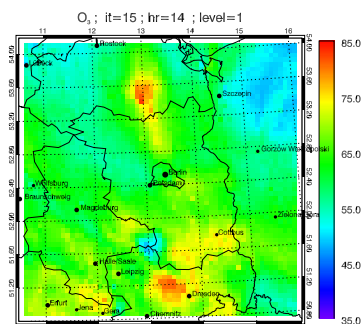
(a) Control run, 12 UTC



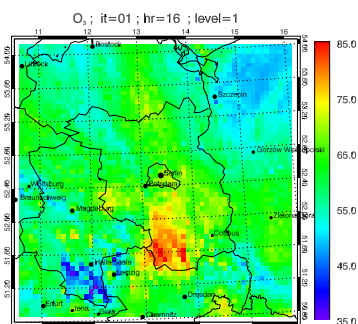
(b) Analysis run, 12 UTC



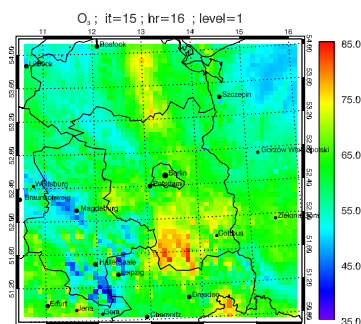
(c) Control run, 14 UTC



(d) Analysis run, 14 UTC



(e) Control run, 16 UTC



(f) Analysis run, 16 UTC

Figure 5.10: BERLIOZ N2: Surface ozone mixing ratios in ppbV during the afternoon on July 20. Left panel: Control run. Right panel: Analysis run.

5.2 VERTIKO Campaign

The VERTIKO (Vertical Transport of Energy and Trace Gases at Anchor Stations and Their Spatial/Temporal Extrapolation under Complex Natural Conditions) network, which is a co-operation of 13 institutions, aimed at measuring and modelling land-surface-atmosphere interactions by means of momentum flux and transport of energy and trace gases (especially water vapour, carbon dioxide and nitrogen species) over inhomogeneous surfaces. In addition to measurement campaigns, continuous measurements have been performed at several anchor stations, located in woods and fields. The target area encompassed the region between the mountain range Erzgebirge in the South and the Oder-Spree lake district in the North, comprising a variety of natural small scale variability. A detailed project summary is given in *Bernhofer and Köstner [2005]*.

5.2.1 Model design and observational data basis

As a special campaign study of the VERTIKO project, the MORE II campaign was selected for analysis. The newly available adjoint nesting technique, which has already been introduced in the BERLIOZ section, has been applied to telescope down from regional scale to local scale, focussing on the MORE II target area. To this end, the model design includes three grids, an European coarse grid and two nested domains with a sequence of

Table 5.2: VERTIKO: Simulation setup. Assimilation experiments are indicated by black solid bars. The assimilation window lasts from 00 UTC to 24 UTC. The days without data assimilation on the nested grids have been used as spin up time.

		June 2003								
Nestl.	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed
	9	10	11	12	13	14	15	16	17	18
CG	■	■	■	■	■	■	■	■	■	■
N1			■	■	■	■	■	■	■	■
N2						■	■	■	■	■

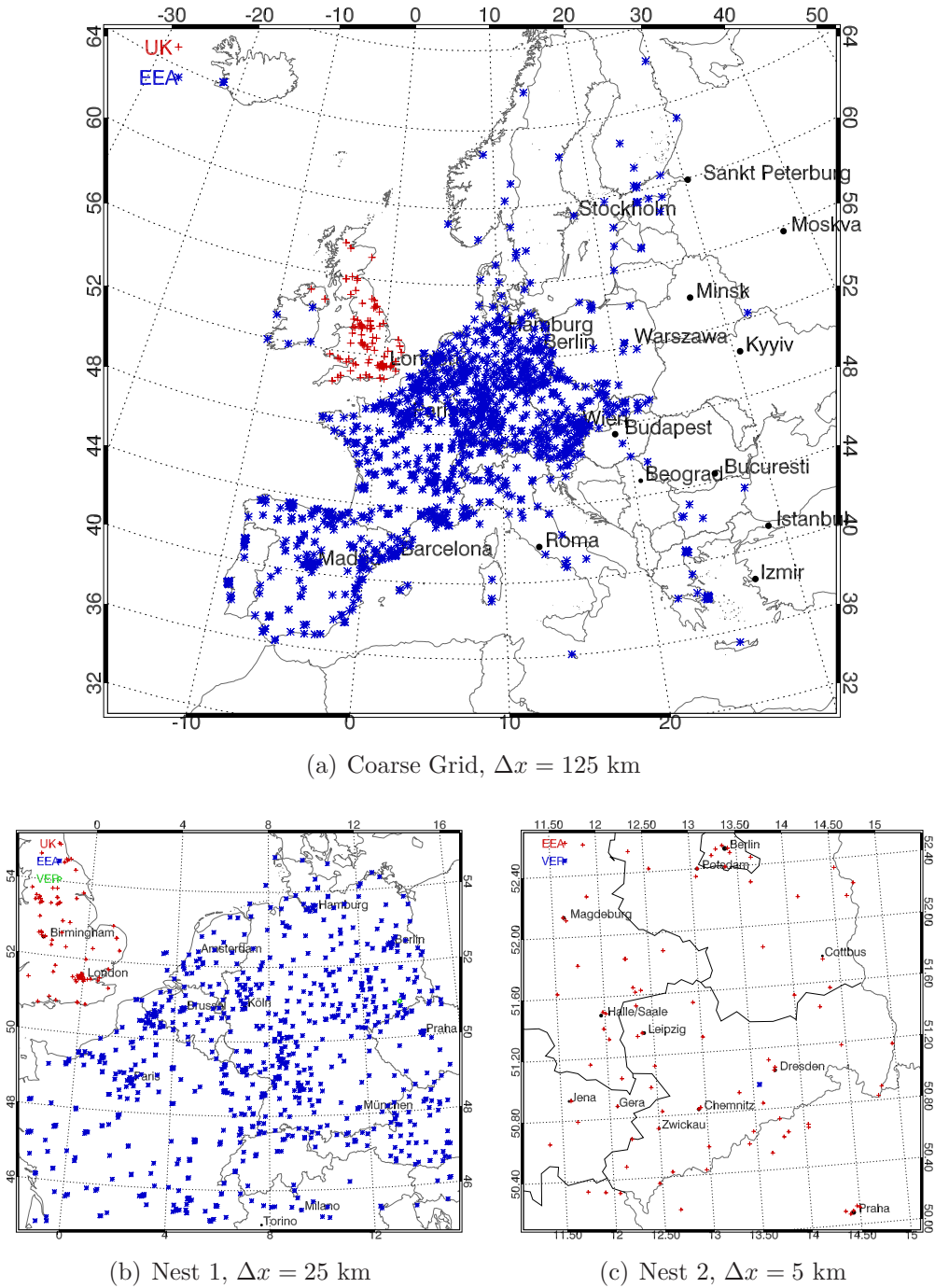


Figure 5.11: VERTIKO: Nested model grid sequence. The position of in-situ measurements are indicated. The grid resolutions are 125-25-5 km.

horizontal grid resolutions of 125-25-5 km. The finest nested domain covers the Saxonian area, in which the VERTIKO sub-campaign MORE II took place (see Fig. 5.11). Like for the BERLIOZ simulations, the number of vertical layers is 23, with about 15 layers within the boundary layer and a representative height of about 19 metres in the lowermost layer.

Ten days in June, 2003, namely June 09 to 18, have been selected as time window for the data assimilation cycles. The suite of assimilation experiments on the three grids is shown in Table 5.2: During the whole period, observational data has been assimilated on the CG. After a spin up period of two days using optimised boundary values by means of coarse grid results, the data assimilation started on Nest 1. From June 14 to 18, optimisation has been performed on all three domains. Further, the data assimilation procedure has been applied for all three optimisation modes.

The selected period has been chosen due to the richest observational basis obtained. The MORE II data comprise NO, NO₂, O₃ and NH₃ as high frequency routine observations at reference height, and SO₂, NO₂, HNO₃ and HNO₂ at two elevated levels, measured about twice a day. In addition to the campaign data, European in-situ observations as provided by the European Environment Agency (EEA) have been used (see Fig. 5.11). This data includes half-hourly values of about 2500 measurement devices for SO₂, NO₂, NO, O₃, NH₃, CO and C₆H₆. Available data for the United Kingdom has been used as provided by the *Air Quality Archive* (www.airquality.co.uk) due to its partly higher resolving information compared to the respective data from EEA.

Fortunately, during the selected period, both NO₂ retrievals of GOME and SCIAMACHY (Scanning Imaging Absorption Spectrometer for Atmospheric Chartography) have been available, provided by KNMI as tropospheric columns. In order to analyse whether there is a positive impact from GOME observations on the model forecast for SCIAMACHY, the SCIAMACHY observations have been withheld from assimilation progress and used as quality control of which only. Further, NNORSY O₃ profiles retrieved from GOME measurements have been assimilated too. Satellite based measurements, again, have only been used for the CG simulations.

5.2.2 Simulation Results

As already demonstrated by the analysis of BERLIOZ, the EURAD 4d-var data assimilation system allows to use retrievals from satellite measurements for the optimisation process. The VERTIKO simulation now offers the opportunity to investigate the model performance over a much longer

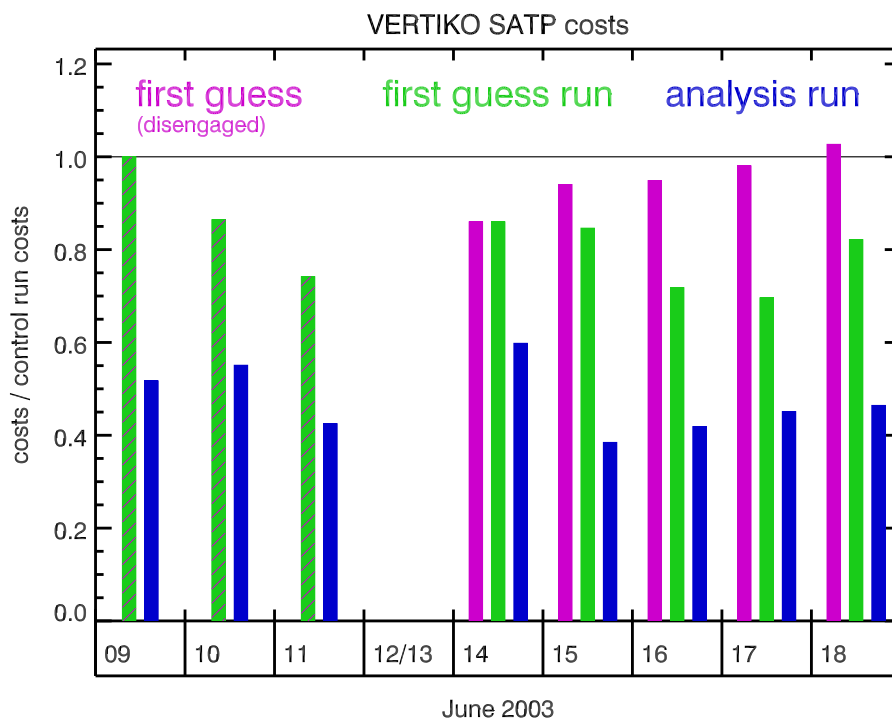


Figure 5.12: VERTIKO CG: Partial costs by the assimilated GOME O_3 profile data for the whole modelling period, normalised by the control run costs, here shown for initial value optimisation. For June 12-13 there are no GOME profile observations available. The disengaged first guess shows the evolution of costs, when no new observations are introduced after June 11, 2003.

Table 5.3: VERTIKO CG: Bias (observation minus model) and rms error for tropospheric NO_2 column retrievals from GOME and SCIAMACHY observations on June 18, 2003. Values are given for control (cr), first guess (fgr) and analysis run (ar) in 10^{15} molec/cm².

VERTIKO CG: NO_2 satellite column statistics						
Statistical Measure [10^{15} molec/cm ²]	GOME			SCIAMACHY		
	cr	fgr	ar	cr	fgr	ar
bias	0.36	-0.15	-0.19	1.44	1.07	0.67
rms error	1.03	1.03	0.88	2.71	2.52	2.36

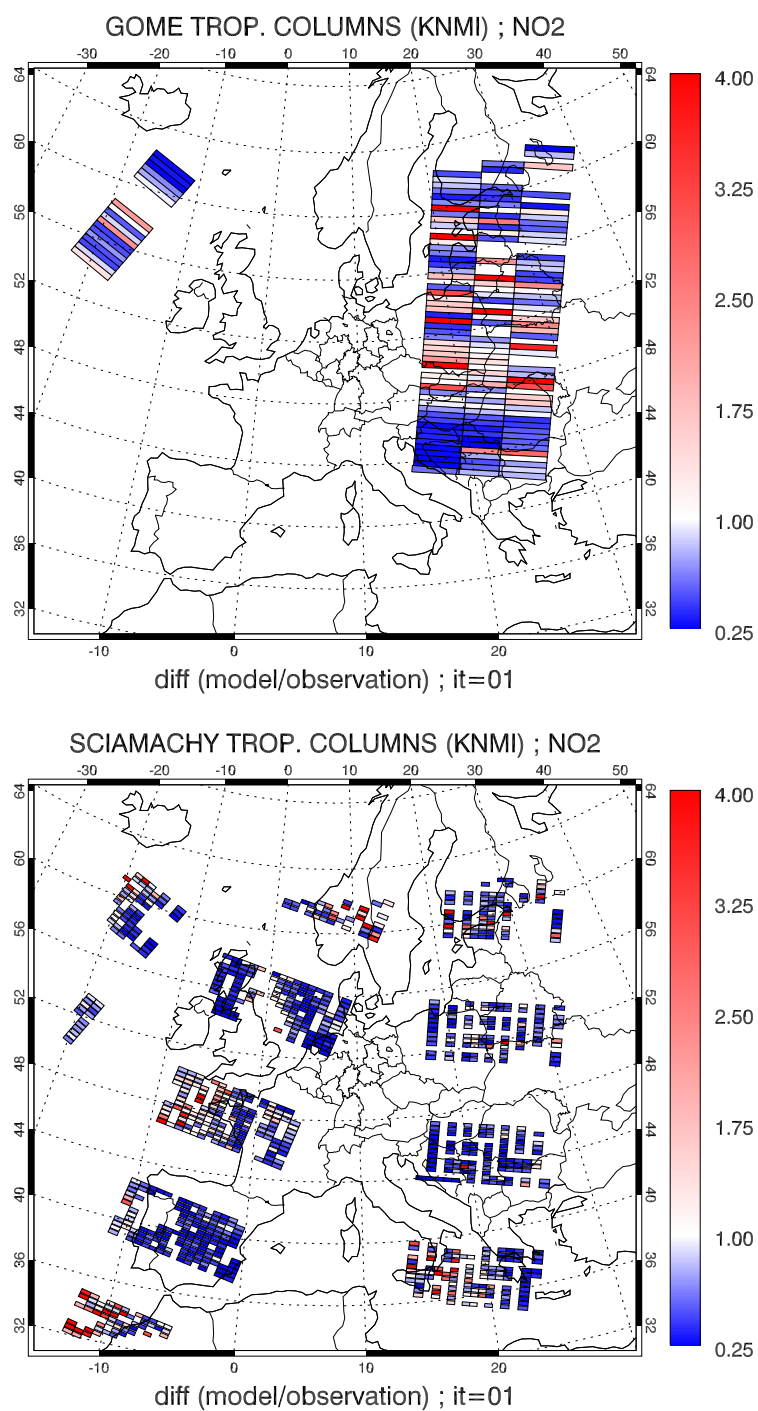


Figure 5.13: VERTIKO CG: Comparison of tropospheric NO₂ columns from GOME (top) and SCIAMACHY (bottom) with model equivalents from the control run, given as ratio (model equivalent/observation) on June 18, 2003.

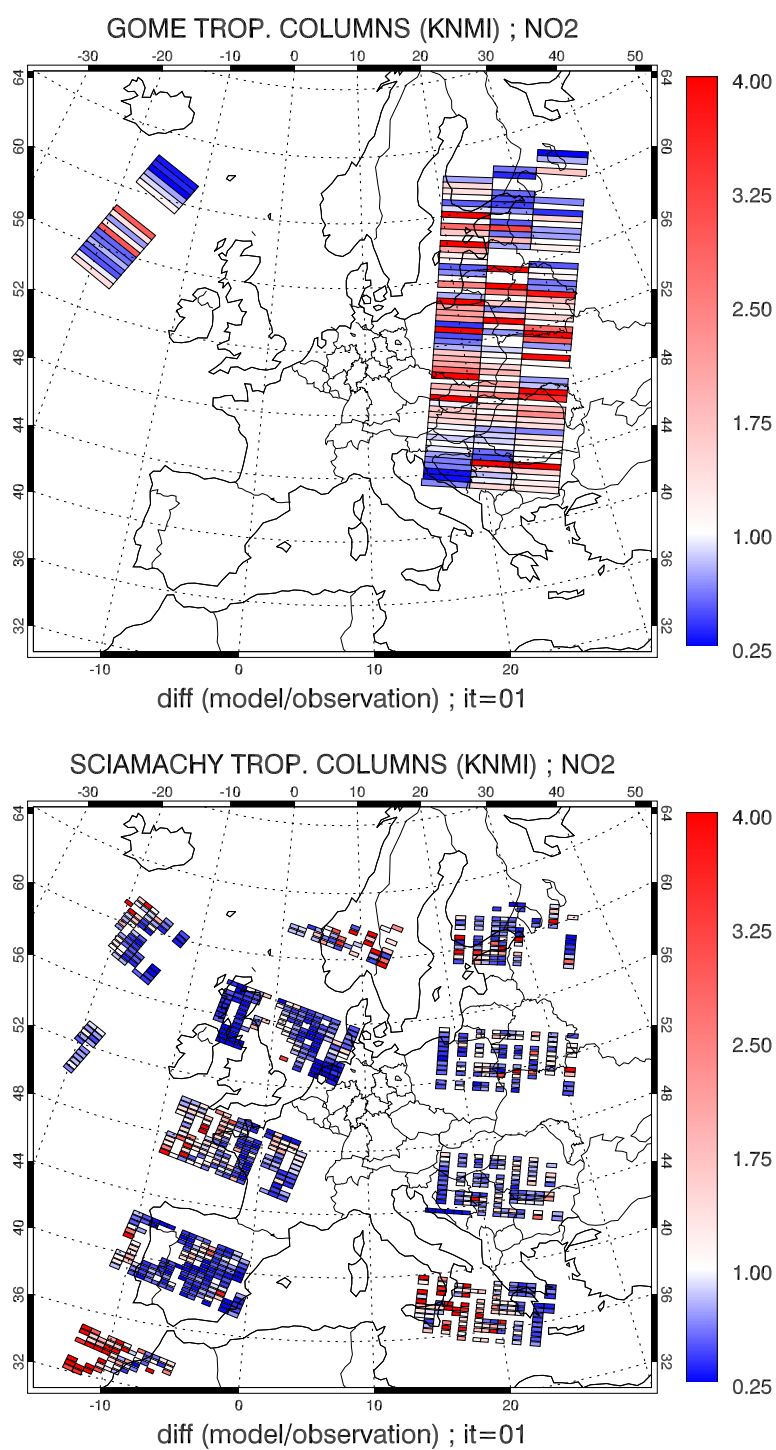


Figure 5.13: (continued) VERTIKO CG: Model equivalents from first guess run.

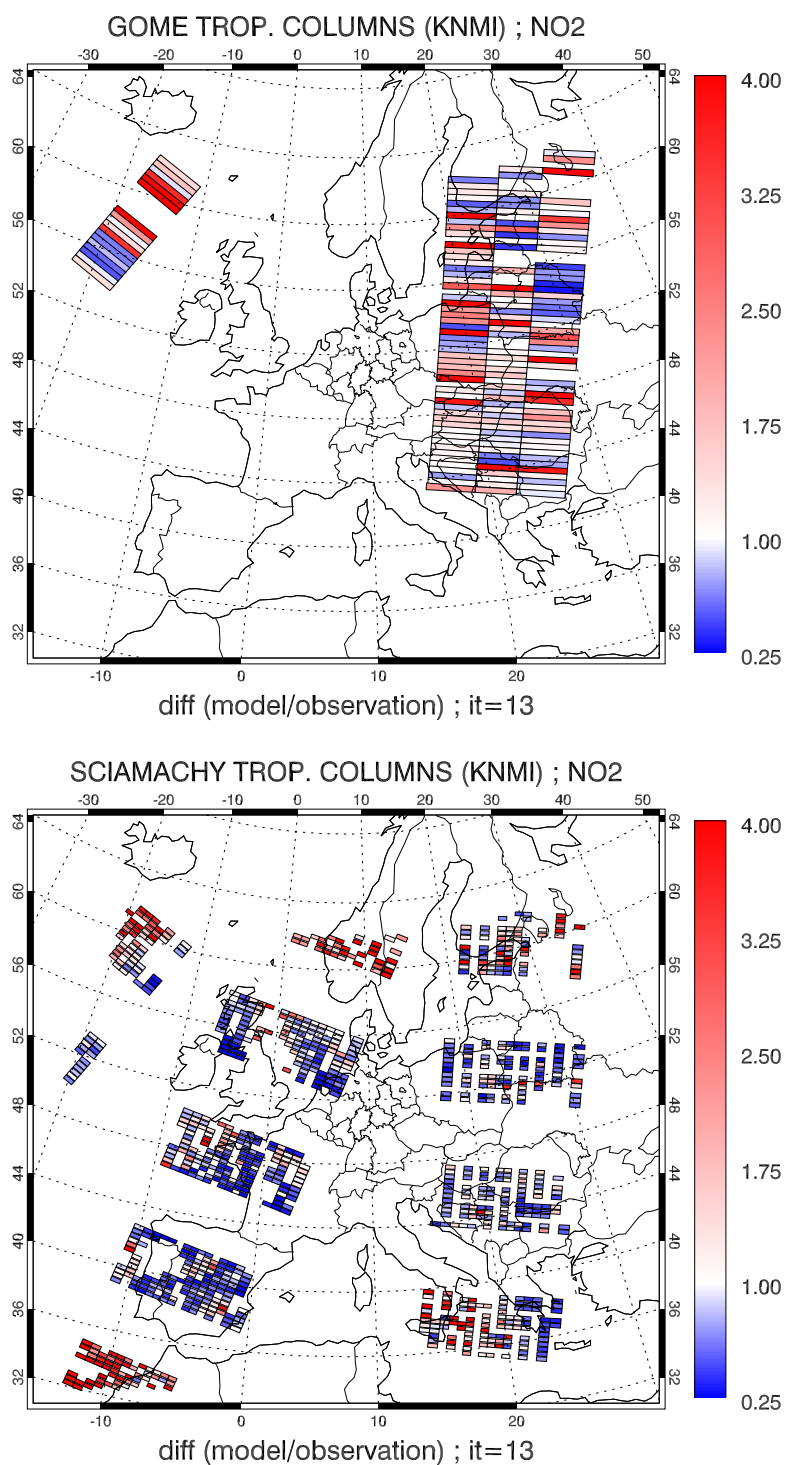


Figure 5.13: (continued) VERTIKO CG: Model equivalents from analysis run.

period, for 10 days here. To this end, Fig. 5.12 shows the costs of the different model realisations when doing initial value optimisation only. This is sufficient when surveying ozone profile assimilation, due to the presumably negligible influence of surface emission rates on the profiles with the lowest observation at 500 m above ground, compared to the retrieval errors for these heights.

During the first three days, the first guess values (green bars) are significantly reduced, equivalent to a successively improved forecast skill for the following day. For June 12 and 13 there were no satellite observations available. Subsequently, an additional forecast has been initiated, without further assimilating any data, in order to check the relaxation process (denoted *disengaged* first guess, magenta bars). As can be seen in Fig. 5.12, the increasing forecast costs reach the control run costs (given by the line denoting 1.0) on June 18, one week after the last analysis procedure for this simulation on June 11. Hence, this can be seen as the time span, in which the whole domain is influenced by boundary values and therefore the previous analysis gain for tropospheric ozone mixing ratios has vanished. However, the analysis procedure, resulting in a cost reduction of about 60 %, leads to first guess forecast costs on the following days of less than 80 % of the values encountered for the control run.

While the previous data assimilation experiment only used tropospheric NO₂ column retrievals by GOME observations, SCIAMACHY measurements have been used here to cross-validate the analysis results obtained by GOME observations. The GOME data has been assimilated during the whole simulation period. Fig. 5.13 shows the control run, first guess and analysis run results for both GOME and SCIAMACHY observations on June 18, given as ratio (model equivalent / observations). Compared to the control run, the first guess run exhibits slight changes in the model state due to GOME observations, especially over the Balkan region and Finland. This can be noticed as well in the model comparison with SCIAMACHY data. Moreover, the regions in which the model equivalent is forced towards the observations, the discrepancies in the SCIAMACHY equivalents slightly vanish, too. In GOME NO₂ data void regions, there is also strong improvement in the model equivalent for SCIAMACHY data (for example over Scotland and the North Sea), which can only be due to other assimilated information, viz GOME ozone profile retrievals being available for this day and covering this region (not shown), and ground based in-situ observations.

The quantification of these qualitative statements is given in Table 5.3 in terms of bias and rms error for GOME and SCIAMACHY observations on

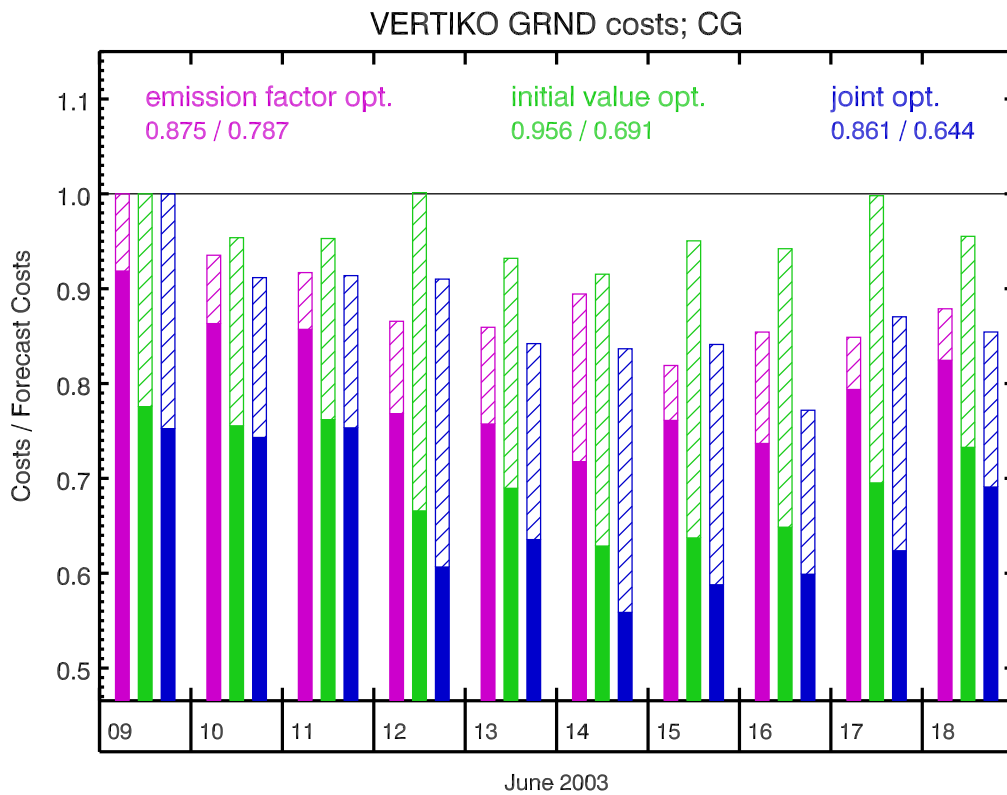


Figure 5.14: VERTIKO CG: Temporal evolution of the cost function for ground based observations with respect to the optimisation mode, normalised by the respective control run costs. Solid areas denote the normalised analysis costs, while line-filled bars indicate the first guess value before optimisation. Numbers in figure key give averaged values for first guess and analysis run, respectively.

June 18. Compared to the control runs, the improved bias and rms error values for both first guess runs, is owing to the assimilation procedures accomplished during the previous days. While present for the SCIAMACHY data, the missing reduction of the rms error for the GOME data is presumably due to the missing data for the middle track covering central Europe, being the area with the highest data covering and thus having most improved model states. The obtained error reduction for the SCIAMACHY observations both for the first guess and analysis run allows to state a successful cross-validation of the assimilation procedure, by means of independent measurements. However, the NO_2 data sets from KNMI base on the same measurement technique and are retrieved by a similar algorithm, which may slightly influence this result.

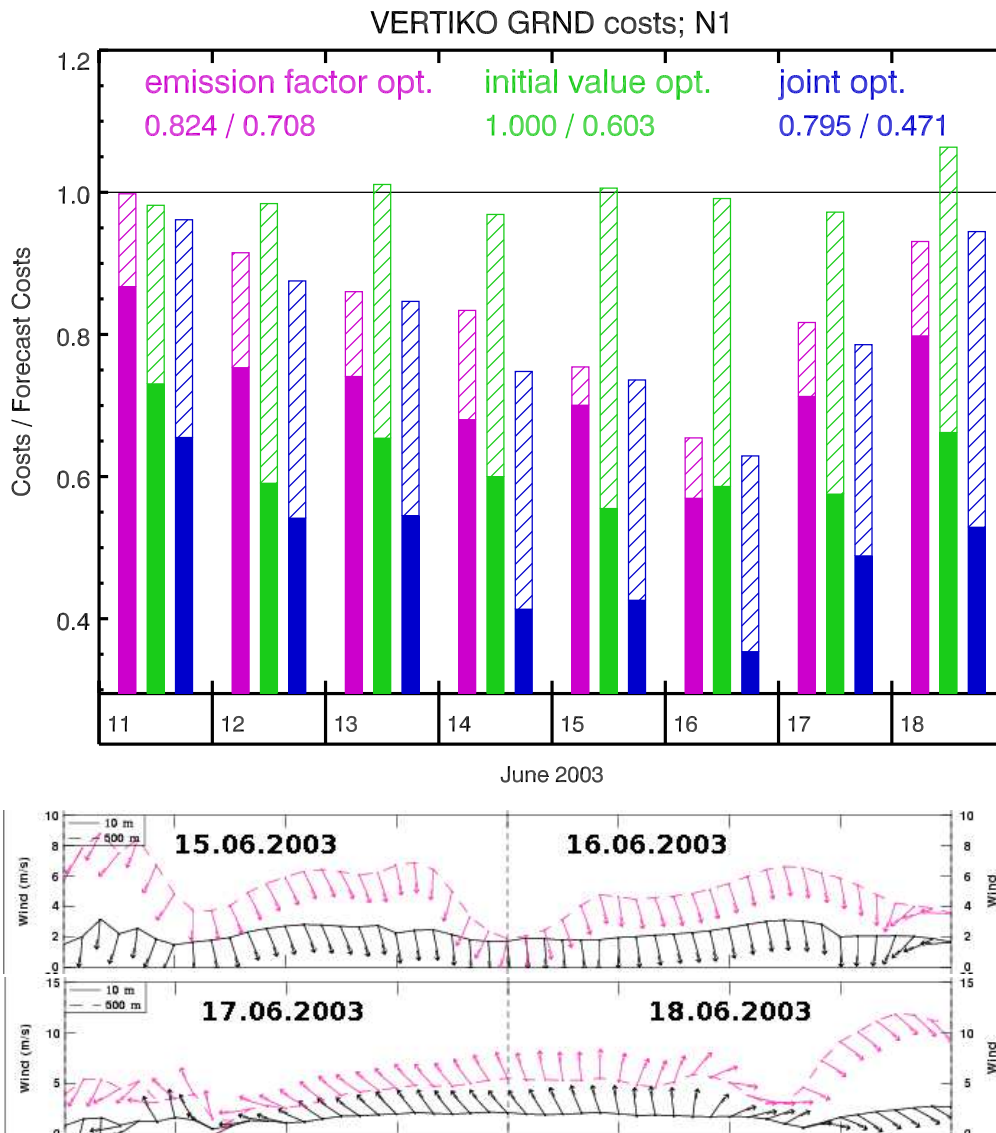


Figure 5.15: VERTIKO N1: Temporal evolution of cost function for ground based observations with respect to the optimisation mode, normalised by the respective control run costs (top). Solid areas denote the normalised analysis costs, while line-filled bars indicate the first guess value before optimisation. Numbers in figure key give averaged values for first guess and analysis run. Bottom: Horizontal wind barbs for Dresden as forecasted by MM5 from June 15 to June 18. Black: at 10 m, magenta: at 500 m height.

The ground based observational data provided by EEA offers the opportunity to assimilate measurements European-wide with an almost equal density. In the following we will refine to local scale emphasising the results obtained on the different nesting stages. Fig. 5.14 shows the temporal evolution of the costs with respect to the optimisation mode and the analysis stage. The different optimisation modes exhibit their characteristics as expected: Compared to the first guess at the same day, initial value optimisation shows a strong gain in the analysis, but the gain is more or less completely lost at the following day. This is mainly due to the fact that, in the boundary layer, initial values lose their influence on the current state after at last one day due to emissions governing the further evolution. Therefore, emission factor optimisation leads to moderate improvements when comparing analysis and first guess run costs at the same day, but analysis gain is preserved by means of the emission factors, and the following forecast is still significantly improved. Finally, joint optimisation of emission rates and initial values combines the advantages of both strategies, leading to the best analysis and the best forecast skill for the following day (according to the results obtained for the BERLIOZ simulations). The same plot is shown for Nest 1 in Fig. 5.15, exhibiting the same features as discussed for the CG. However, to stress the properties of the assimilation modes, it must be noted, that, on June 17, there is a strong increase in first guess costs for both emission factor optimisation and joint optimisation, while initial value optimisation performance stays unchanged. This is due to a strong change in wind direction from June 16 to June 17 (see bottom panel of Fig. 5.15). Therefore the previously windward analysed emission factors lose their influence on the observing locations and emission factors have to be estimated anew. In general, the deterioration after the change of meteorological conditions regime may have two reasons: 1) The emission optimisation has been incomplete during the day prior of June 17, as only those areas can be improved which are upwind of observation sites, and 2) there has been an over-adaptation for the very situation prior to change. Presumably a combination of both cases occur.

To investigate the nesting success, Fig. 5.16 shows bias and rms values for the three domains. The respective hourly values for bias and rms, encountered for the observed constituents, have been averaged over available simulation days. The time series are given for the analysis runs of the three optimisation modes and the control run. Together with the complete averages for bias and rms values given in Tab. 5.4 for joint optimisation, the following results can be stated: SO_2 , NH_3 and C_6H_6 show strong enhancement in bias and rms when nested grids are employed. The daily trend in

Table 5.4: VERTIKO: Bias (observation minus model) and rms error statistics for joint optimisation analyses obtained at the different domains for ground based measurement stations. Values are given in ppbV for the control runs (cr), first guess runs (fgr) and analysis runs (ar).

VERTIKO ground statistics									
Species	bias [ppbV]								
	CG			N1			N2		
	cr	fgr	ar	cr	fgr	ar	cr	fgr	ar
SO ₂	0.7	0.4	0.7	-1.6	0.6	0.4	-5.7	-0.9	-0.8
NO ₂	7.8	5.9	6.9	6.0	4.6	5.0	6.0	4.9	5.0
NO	9.7	8.5	8.8	7.7	7.2	7.3	8.0	7.6	7.5
O ₃	-0.8	0.1	1.3	1.3	1.2	1.9	2.1	2.3	2.0
NH ₃	-4.3	7.0	6.2	-5.9	2.9	2.5	-5.3	-2.7	-0.9
CO	197.3	145.7	118.1	103.0	80.6	65.9	113.1	78.3	62.5
C ₆ H ₆	0.42	0.39	0.30	0.30	0.25	0.16	0.17	0.22	0.15
Species	rms error [ppbV]								
	CG			N1			N2		
	cr	fgr	ar	cr	fgr	ar	cr	fgr	ar
SO ₂	6.7	6.2	6.0	6.8	4.3	3.9	9.3	3.5	3.0
NO ₂	12.4	11.3	11.6	10.6	9.4	9.3	9.3	8.3	8.2
NO	20.4	19.6	19.6	18.2	17.8	17.6	14.8	14.4	14.4
O ₃	15.3	14.1	12.1	13.8	13.3	9.9	12.5	12.1	9.5
NH ₃	15.2	13.0	12.6	13.6	8.4	7.0	6.3	3.3	2.3
CO	365.1	332.3	315.0	272.5	227.4	208.7	225.1	188.9	170.2
C ₆ H ₆	0.63	0.60	0.59	0.43	0.39	0.33	0.61	0.40	0.32

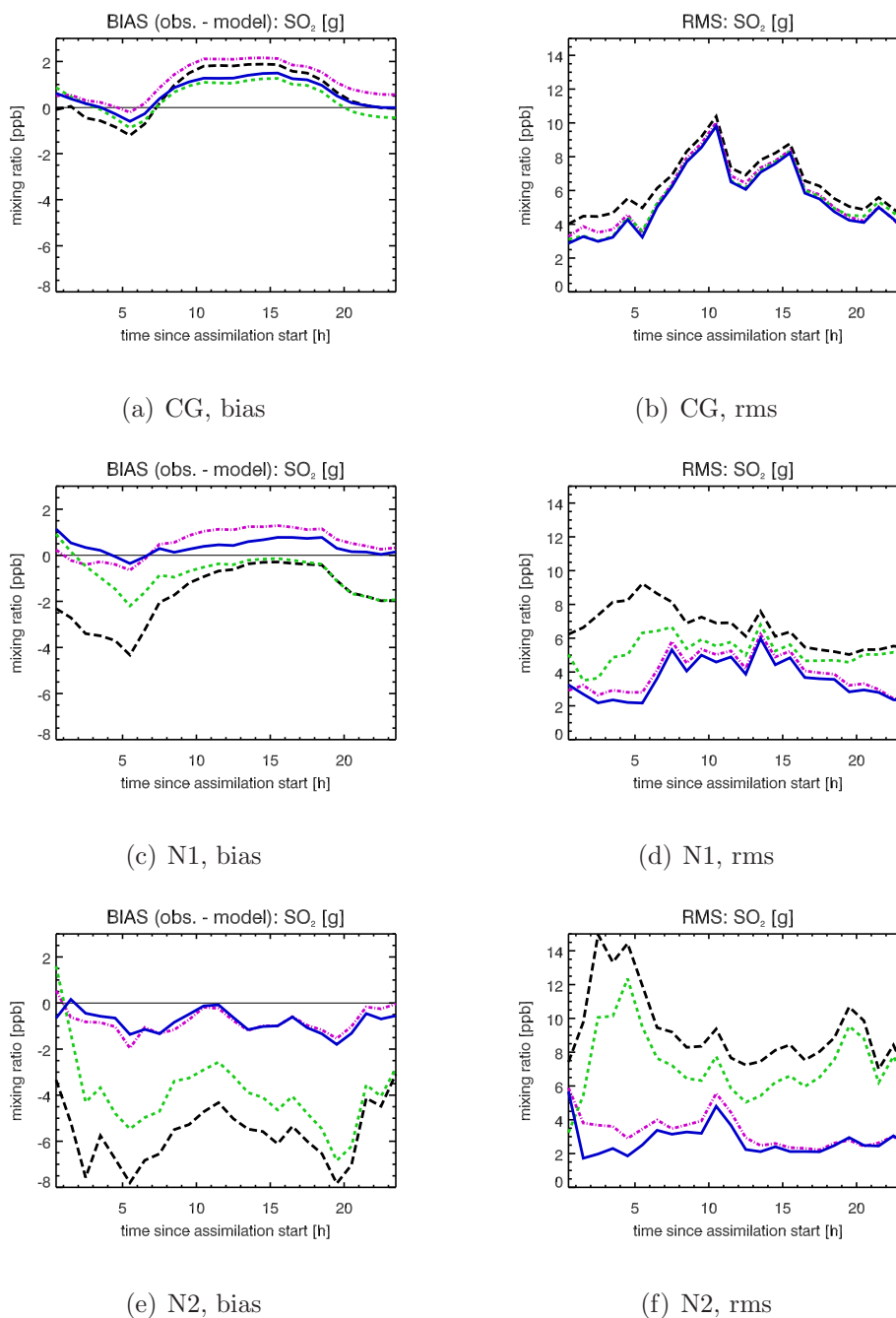
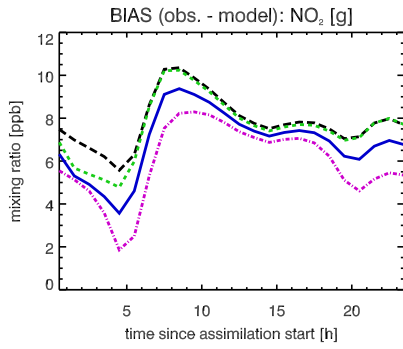
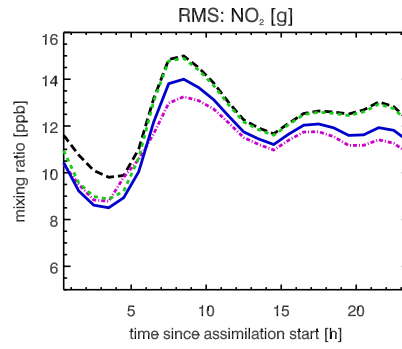


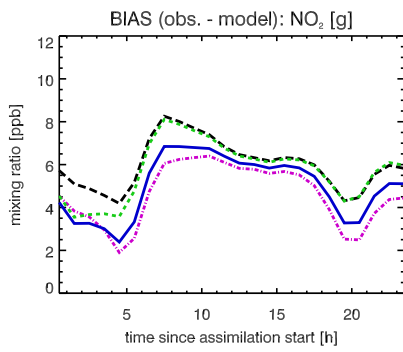
Figure 5.16: VERTIKO: Hourly averaged bias and rms error for ground based SO_2 observations with respect to hour after assimilation start and optimisation mode. Values have been averaged over the available days. Dashed black line: Control run. Dash-dotted magenta line: Analysis run (EF). Dashed green line: Analysis run (IV). Solid blue line: Analysis run (JT).



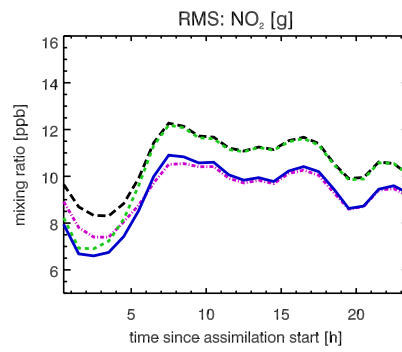
(a) CG, bias



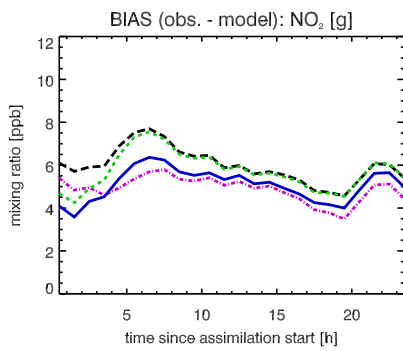
(b) CG, rms



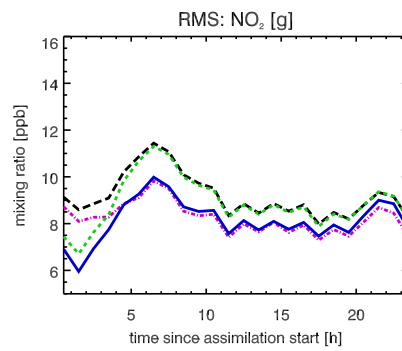
(c) N1, bias



(d) N1, rms

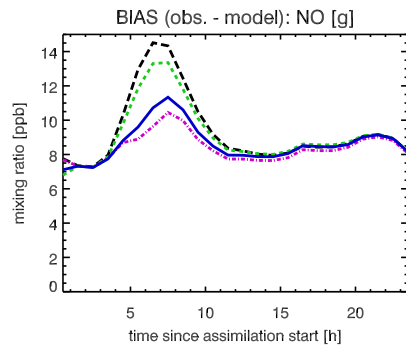


(e) N2, bias

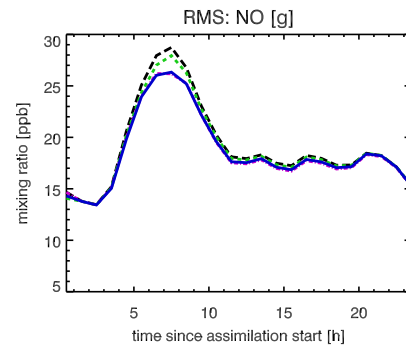


(f) N2, rms

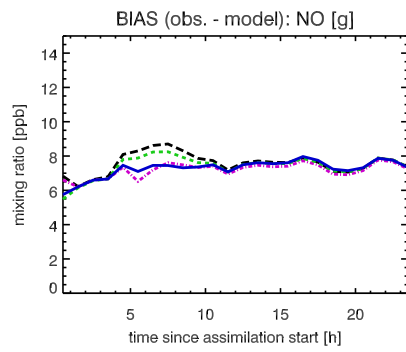
Figure 5.16: (continued) VERTIKO: NO₂.



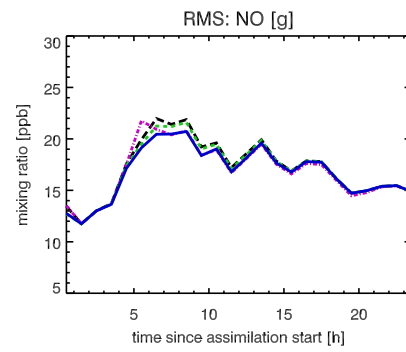
(a) CG, bias



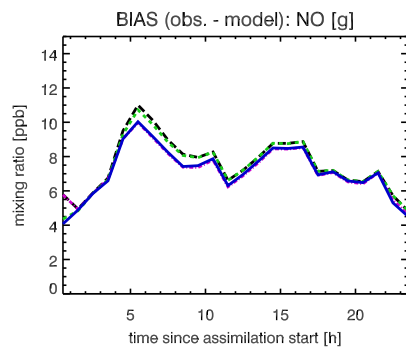
(b) CG, rms



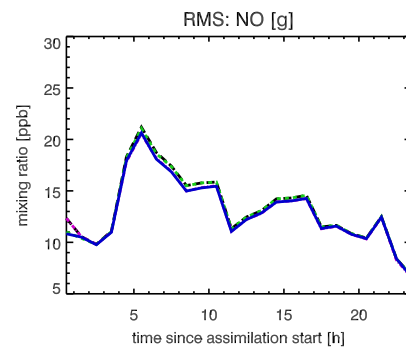
(c) N1, bias



(d) N1, rms



(e) N2, bias



(f) N2, rms

Figure 5.16: (continued) VERTIKO: NO.

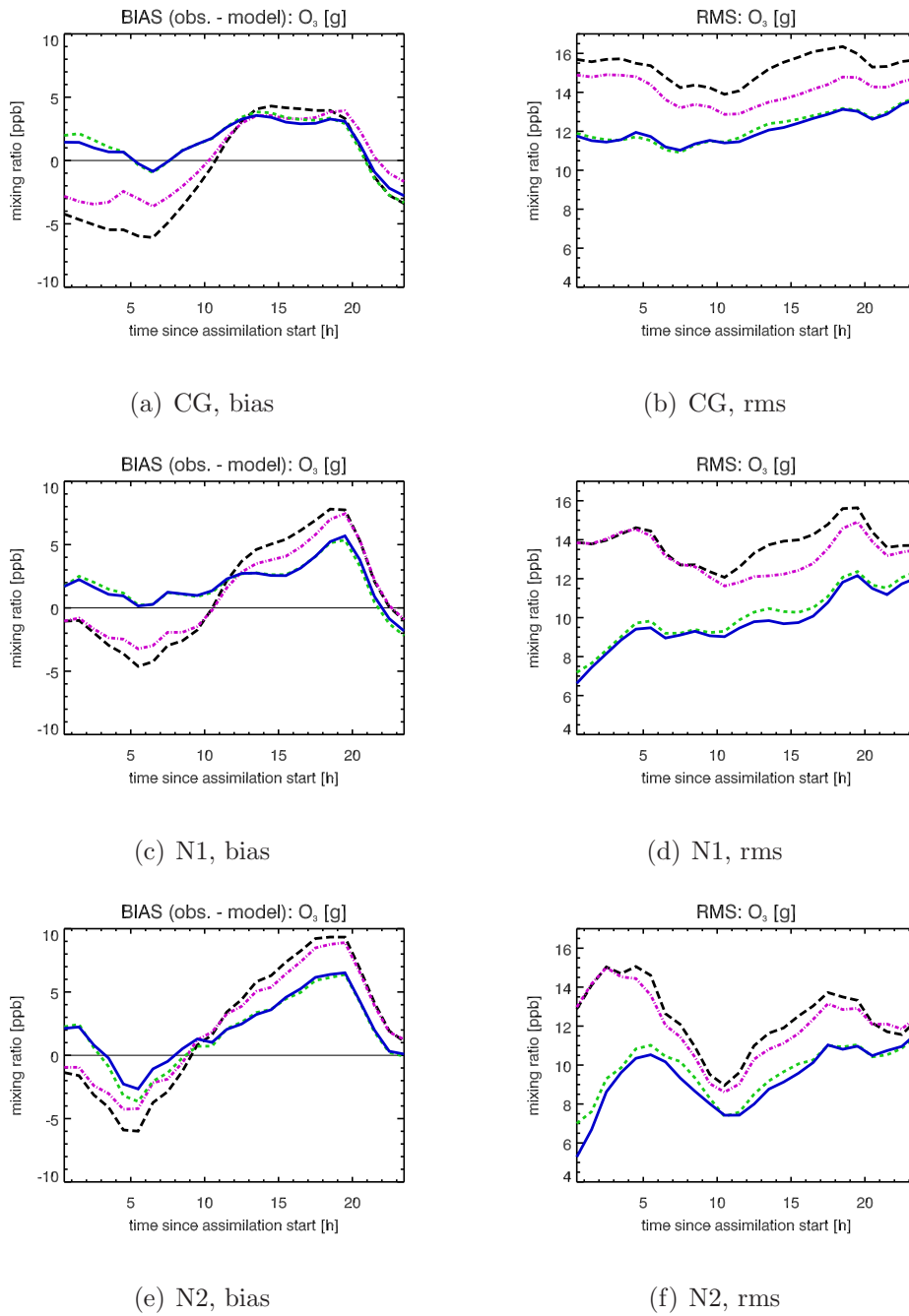


Figure 5.16: (continued) VERTIKO: O_3 .

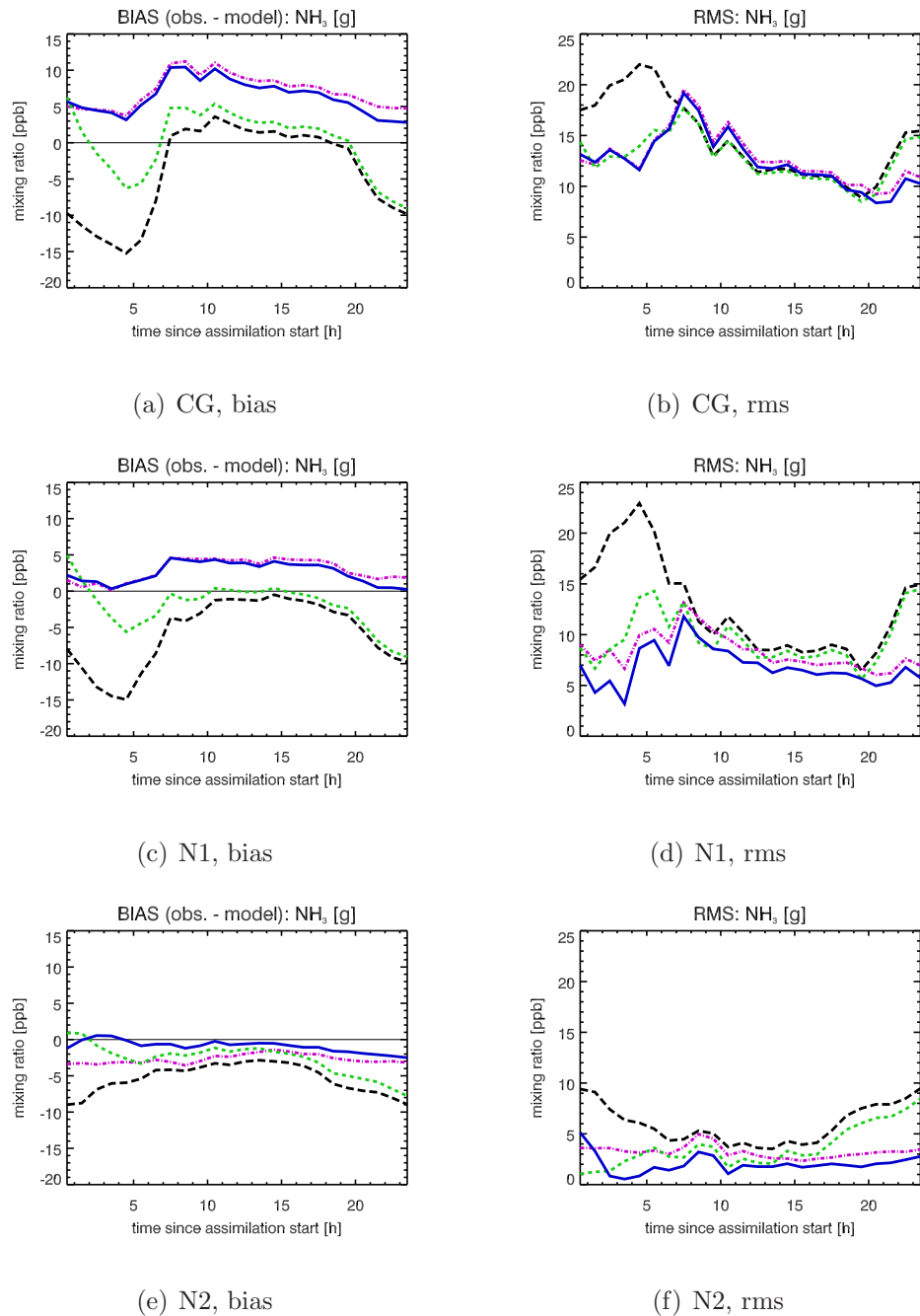
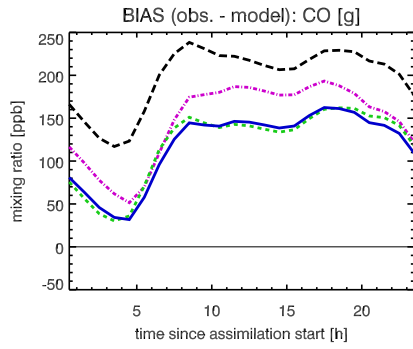
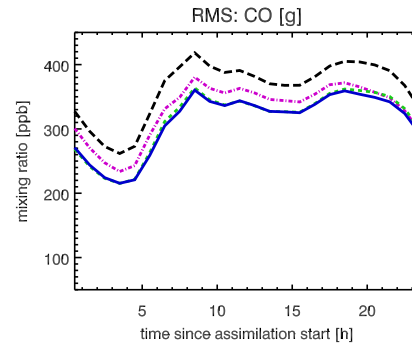


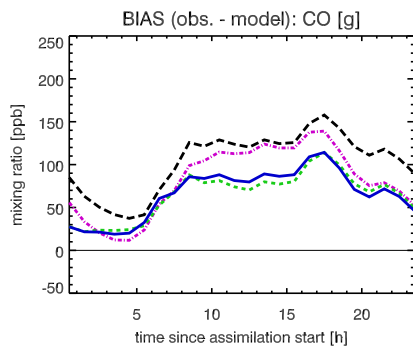
Figure 5.16: (continued) VERTIKO: NH_3 .



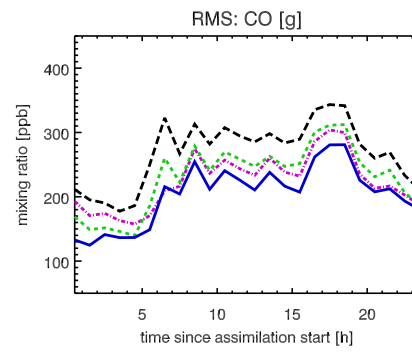
(a) CG, bias



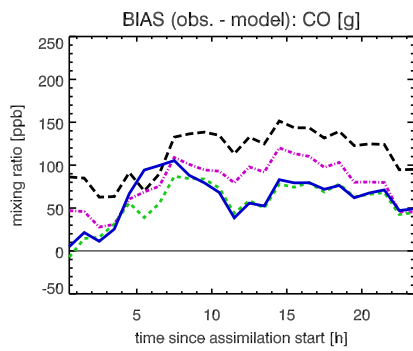
(b) CG, rms



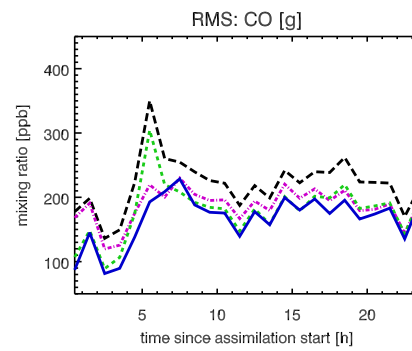
(c) N1, bias



(d) N1, rms



(e) N2, bias



(f) N2, rms

Figure 5.16: (continued) VERTIKO: CO.

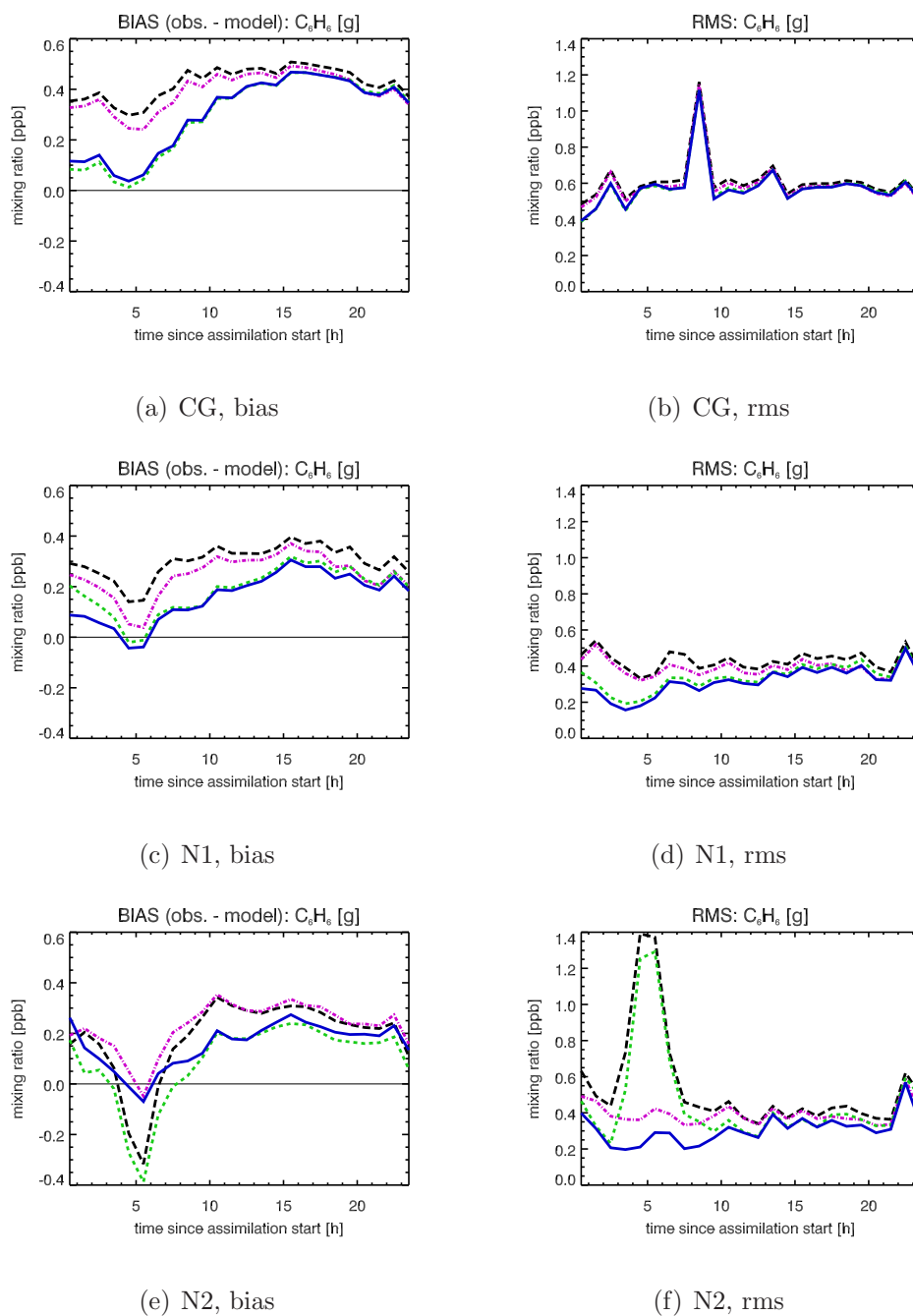


Figure 5.16: (continued) VERTIKO: C_6H_6 .

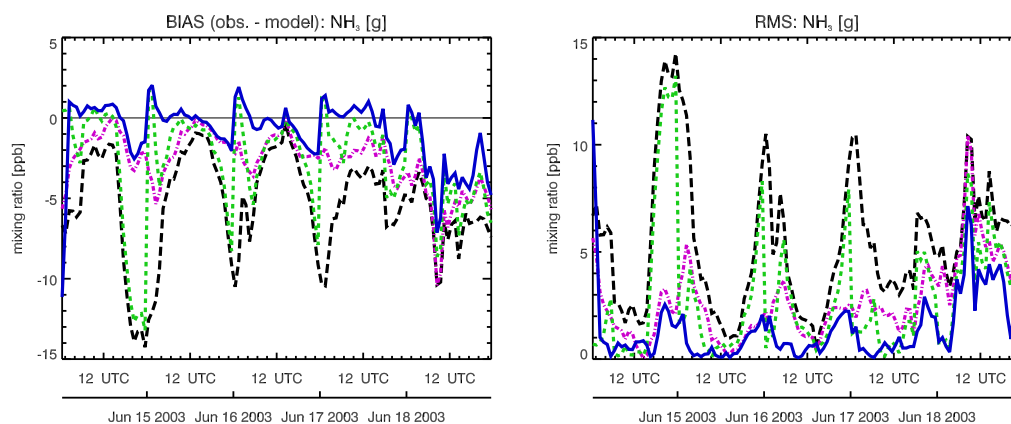


Figure 5.17: VERTIKO N2: Time series of bias (left) and rms (right) for NH_3 campaign data from June 15, 2003, 00 UTC to June 19, 2003, 00 UTC. Black dashed line: Control run. Magenta dash-dotted line: Analysis run (EF). Green dashed line: Analysis run (IV). Solid blue line: Analysis run (JT).

bias and rms of SO_2 and NH_3 has completely been removed, while that of NO_2 and C_6H_6 ameliorate. O_3 and CO show moderate improvements from Nest 1 to Nest 2, the main analysis benefit is already obtained by CG and N1 assimilations, reflecting the characteristic length of representativeness of these two species. The model equivalent for NO and NO_2 only slightly depend on the nesting procedure. This is thought of to be mainly due to the strong local variability, mainly of NO . Here, even simulations with a grid spacing of 5 km are not sufficient to efficiently analyse these constituents. Strictly wrong assumptions about the partial fracture of NO and NO_2 and the daily cycle in the original NO_x emissions may prohibit the analysis using time independent, correlated emission factors.

However, in general, a significant reduction of rms values have been obtained for all species, when model resolution is increased. Concerning the optimisation mode, the different species exhibit different response. Neither emission factor nor initial value optimisation is obviously superior to the other mode. Only jointly optimising both emission factors and initial values reliably shows the best performance of the three optimisation modes, which is in accordance to the simulation results already obtained for the BERLIOZ campaign.

The VERTIKO campaign data comprises timely high resolving NH_3 observations. The bias and rms time series for this data is plotted from June 15 00 UTC to June 19 00 UTC in Fig. 5.17 for the different optimisation modes. Obviously, the analysis procedure works successful in reducing both

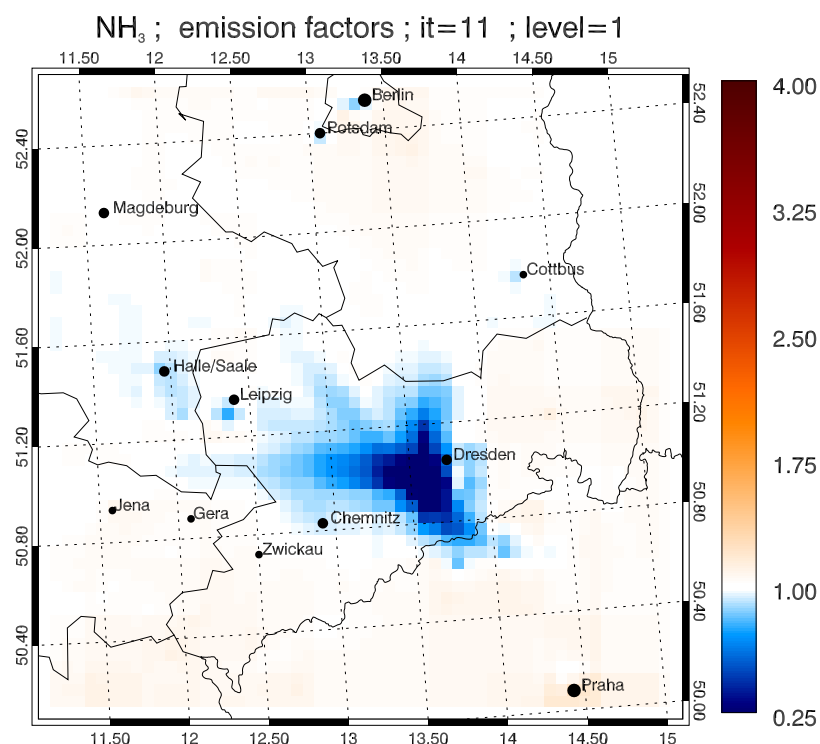


Figure 5.18: VERTIKO N2: Optimised emission factors for NH_3 as analysed by joint optimisation for June 18.

bias and rms, most notably when joint optimising initial values and emission factors. After June 17, both bias and rms exhibit a strong increase towards the control run values when emission factors are optimised. This is due to the changed main horizontal wind direction (see Fig. 5.15) and emission rates in the newly windward direction have not yet been analysed. The emission factors for NH_3 on June 19 as retrieved by joint optimisation are given in Fig. 5.18. Following the analysis, the emission rate estimates of NH_3 must be reduced throughout northwestern Saxony to enable the model to be in accordance with the campaign data. The emission factors unequal to one, *e.g.* in central Berlin, where no NH_3 measurements are available, is due to the correlation of the NH_3 emission factor with another observed and emitted trace gas, coded in the emission factor background error covariance matrix.

5.3 CONTRACE I Campaign

The CONTRACE project (Convective Transport of Trace Gases into the Upper Troposphere over Europe: Budget and Impact on Chemistry) focused on the long-range transport of trace gases from the boundary layer into the upper troposphere (UT) by isolated convective clouds, frontal or large-scale uplifting. The relative importance of convective clouds, frontal systems and long-range advection in transporting nitrogen oxides and other chemical species from the boundary layer into the free and upper troposphere should be investigated, and the transport of NO_x and other ozone precursors should be quantified.

The project aimed to achieve these objectives by a suitable combination of chemical and meteorological in-situ measurements, accomplished with the DLR research aircraft Falcon, analysis of satellite data and chemical transport modelling. It was planned to make flights over strongly polluted areas and less polluted areas during unstable weather situations with focus on the upper troposphere (7-12 km height).

5.3.1 Model design and observational data basis

The first CONTRACE episode with a special flight on November 14, 2001 was selected for upper tropospheric data assimilation, with warm conveyor belt features induced by cyclone dynamics and strong uplifting due to cumulonimbus clouds over the north-western Mediterranean (see Fig. 5.19).

Table 5.5: *Simulation setup for the CONTRACE campaign. Assimilation experiments are indicated by black solid bars. The assimilation window lasts from 00 UTC to 24 UTC on the CG, while on the N1 observations from 08 UTC to 18 UTC have been assimilated. Days without data assimilation have been used as spin up period or quality control.*

	November 2001					
Nestl.	Sat 10	Sun 11	Mon 12	Tue 13	Wed 14	Thu 15
CG		■	■	■	■	
N1					■	

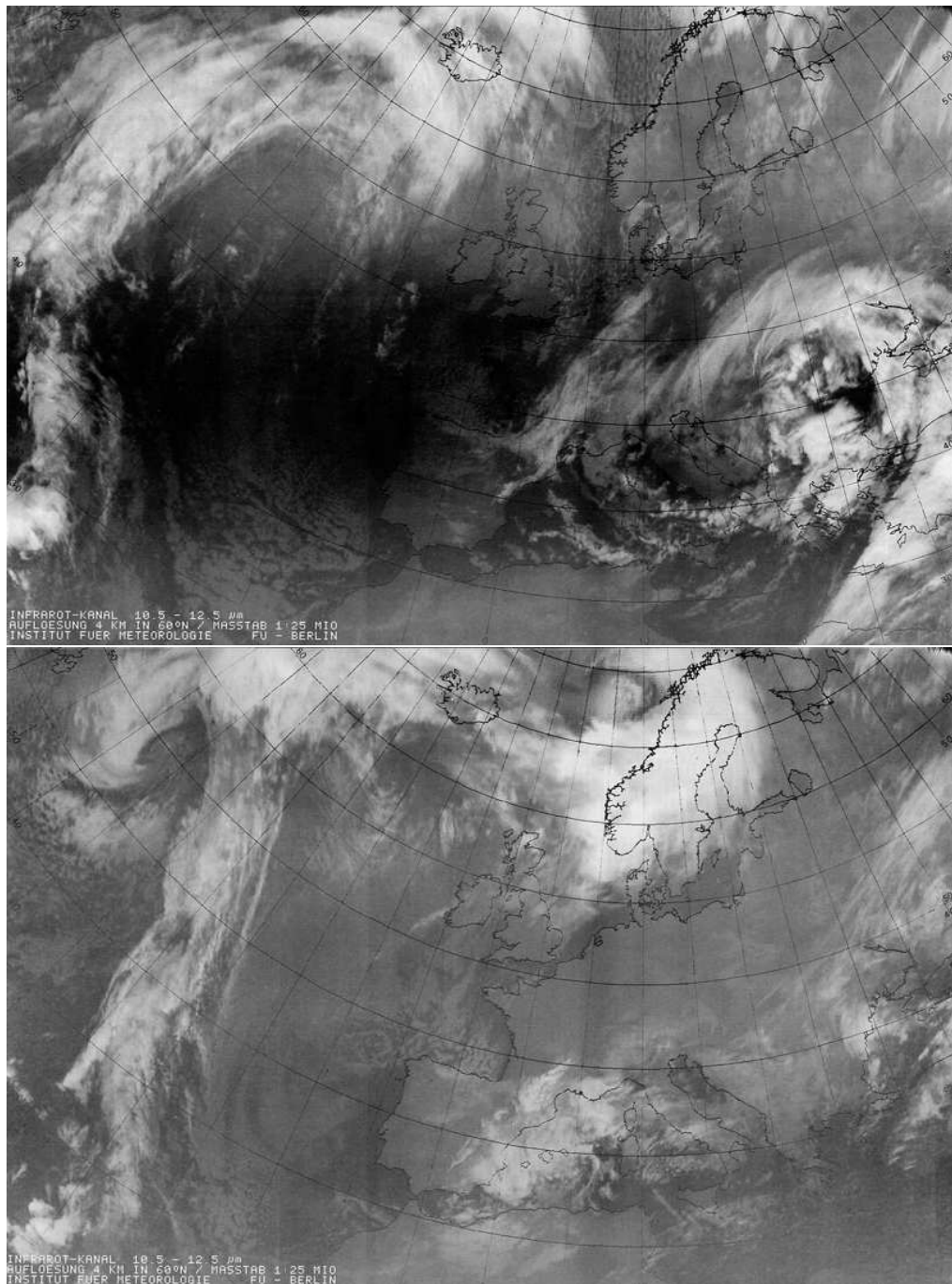


Figure 5.19: Satellite image (Meteosat 7, IR channel) of Europe and the North-Atlantic on November 14, (top) and November 15 (bottom), 2001, 00 GMT. The region of interest is the north-western Mediterranean where thunderstorms evolve with lightning during the second half of the day.

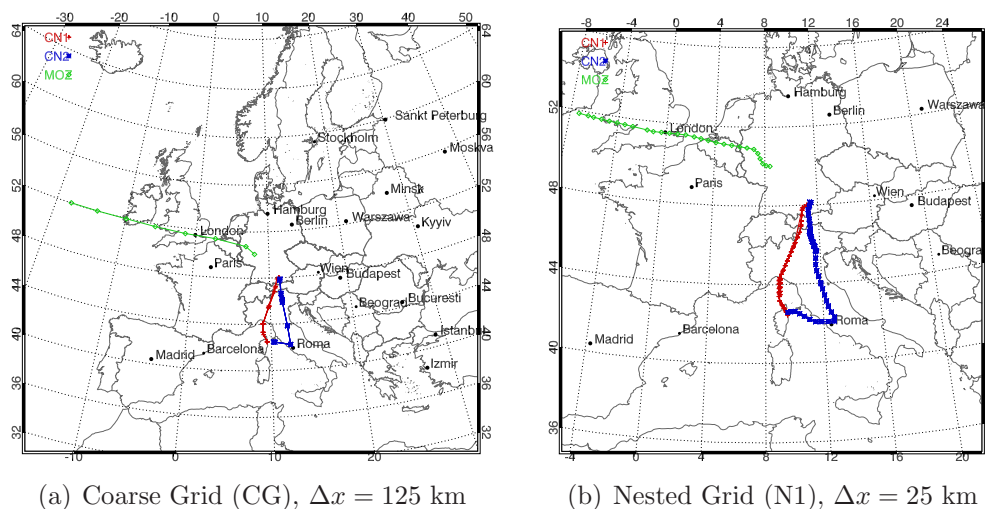


Figure 5.20: *CONTRACE*: Domain areas for the two employed grids having horizontal resolutions of 125 and 25 km, respectively. The flight paths of the assimilated air-borne observations from November 14, 2001, are overlaid.

The vertical grid structure has been refined for the free troposphere region, having now 26 layers with suitable layer thicknesses (see Appendix A). The simulations have been splitted into two parts (see Tab. 5.5): 1) The period from November 10-14 has been simulated on a Coarse Grid with a horizontal grid spacing of 125 km (see Fig. 5.20a). Taking November 10 as spin up run, available satellite observations and air-borne data from November 11-14 have been assimilated within an assimilation window of 24 hours. Initial values have been optimised for midnight (00:00 UTC) to ensure a daily consistent chemical model state on the Coarse Grid during the whole period. 2) To facilitate high quality analysis fields describing the processes of interest in a spatially sufficient manner, an additional grid has been nested into the Coarse Grid, applying a nesting ratio of 5. The resulting grid is shown in Fig. 5.20b, having a horizontal grid size of 25 km. It encompasses central Europe and parts of North-Africa, while the nested domain focuses on the region of interest. The initial and boundary values have been provided by coarse simulations in 1). Due to the huge numerical demands, the assimilation window has been shortened to be from 08 UTC to 18 UTC on November 14, suitably enclosing the *CONTRACE* flight schedules.

Due to the favourable optimisation properties obtained in the previous sections, the results of the data assimilation experiments shown here have been performed in joint optimisation mode only, and — as in the previous sections, too — the Rosenbrock solver to the RACM chemistry kernel has been

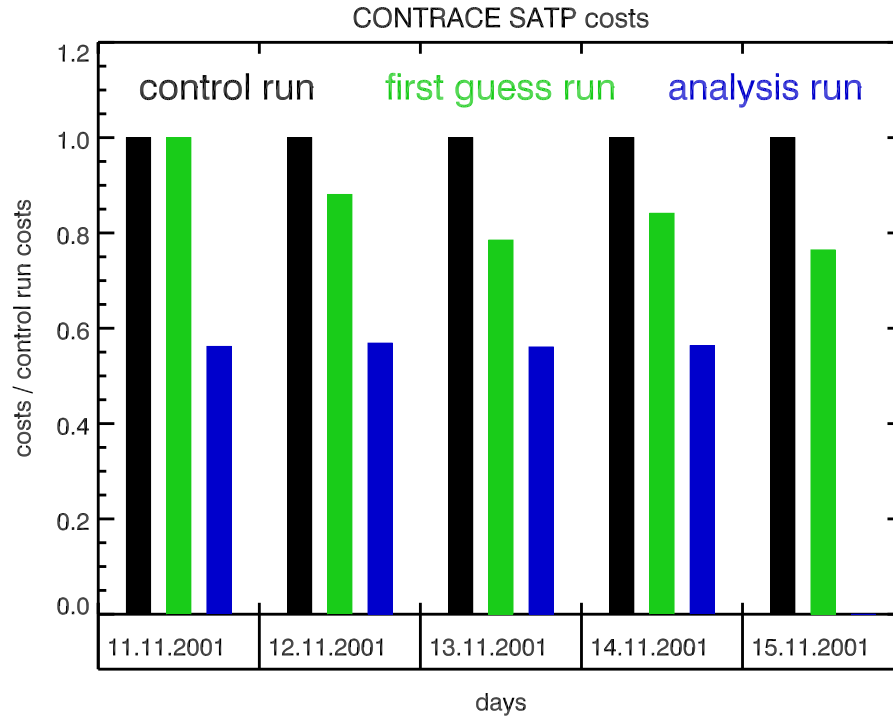


Figure 5.21: *CONTRACE CG: Partial costs by the assimilated GOME O₃ profile data for the whole modelling period, normalised by the control run costs. On November 15 only forecast verification has been conducted.*

Table 5.6: *CONTRACE: Normalised bias (observation minus model) and rms error (see Eqs. (C.2)-(C.4)) statistics by GOME O₃ observations obtained above 5000 m for control, first guess and analysis runs.*

	Normalised bias					
	11.11.	12.11.	13.11.	14.11.	15.11.	
control run	-0.21	-0.28	-0.15	-0.37	-1.13	
first guess run	-0.21	-0.18	0.06	-0.12	-0.77	
analysis run	0.18	0.23	0.17	0.19	—	
	Normalised rms error					
	control run	1.53	1.50	1.48	1.67	2.02
	first guess run	1.53	1.39	1.26	1.51	1.73
analysis run	1.04	0.99	0.98	1.16	—	

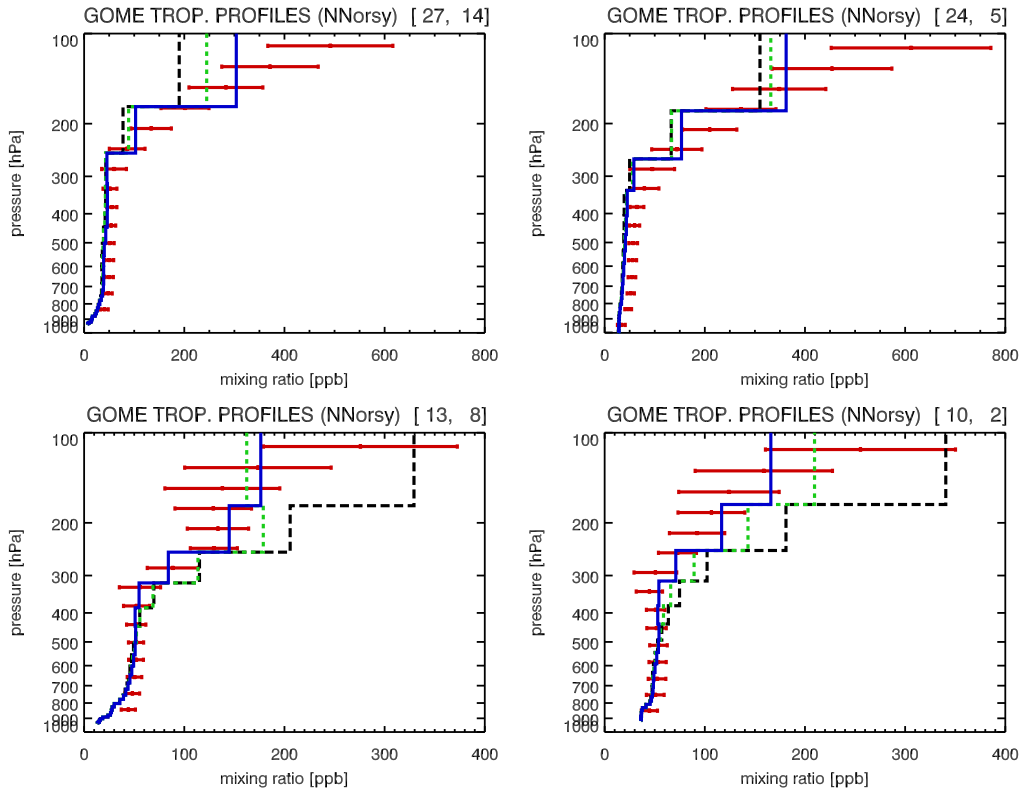


Figure 5.22: CONTRACE CG: O₃ profile retrieval examples by NNORSY from GOME observations (red crosses) on November 13, 2001, and the respective model columns. Black dashed line: Control run. Green dashed line: First guess run. Solid blue line: Analysis run.

employed.

The observational basis consists of air-borne data and available tropospheric satellite retrievals. The former is given by the CONTRACE I campaign data for November 14 and observations from the MOZAIC project. Assimilated species from air-borne data are O₃, NO₂, H₂O₂, CO and NO_y. The observations of HCHO have been skipped, due to most probable bad data quality (H. Huntrieser, personal communication). The latter data comprises tropospheric NO₂ column retrievals provided by KNMI and NNORSY O₃ profile retrievals (Müller *et al.* [2003]), both from GOME measurements. Satellite based data has been assimilated on the Coarse Grid only, while air-borne data was used on both grids.

5.3.2 Simulation Results

Starting on November 10 with a 24 hour spin up run, the first data assimilation experiment has been undertaken on November 11. The suite of partial costs for the GOME O₃ profiles during the whole simulation period is shown in Fig. 5.21. The costs are normalised by the control run, to ensure comparability between the different days. Since data assimilation process started on November 11, the control and first guess run are identical. During the following days, the first guess costs successively reduce to about 80 %, while the analysis run produces stable results with a cost reduction of more than 40 %, which are comparable results to those obtained in the VERTIKO simulations. This fact represents the procedural benefit introduced in the model by the observations and the conservation of the knowledge over, at least, the forecast period (see also Fig. 5.12). To emphasise the forecast improvement by comparison with independent observations, the values of the normalised bias and rms error have been calculated for the observations above 5000 m height, where the strongest discrepancies between the model forecast and the observations can be encountered. The results are given in Table 5.6 for the control, first guess, and analysis runs. Moderate biases occur in the control runs with overestimation of the observations by the model realisations. These biases are significantly reduced in the first guess runs, while the rms errors show more or less the same behaviour as the cost function, shown in Fig. 5.21. The sign change of the bias in the analysis runs is due to the tendency to underestimate the ozone mixing ratios in the lower model domain while the overestimation in the upper model domain is corrected. As an illustration, Fig. 5.22 shows four examples of GOME based O₃ profile retrievals on November 13, 2001, and the respective model columns, calculated with the different model runs, exhibiting the properties of the assimilation procedure discussed above. While the lower part of the O₃ profiles shows the tendency of a positive observation-minus-model bias, discrepancies in the upper model region are successively corrected and the first guess model state profits from the previously assimilated observations. To focus on the uplifting of boundary layer air masses into the free troposphere, Fig. 5.23 shows the temporal evolution of CO from November 14, 00 UTC until November 15, 12 UTC. The mixing ratios are given for model levels 10, 16 and 21, corresponding to a height of about 800, 2300 and 5500 m, respectively. Carbon monoxide is shown because of its relatively long lifetime and its almost negligible daily variations, which enables to use it as a tracer over time spans of two days. Due to frontal processes during the previous days, the near surface air over the northern part of Europe is slightly less polluted than over southern Europe. During these

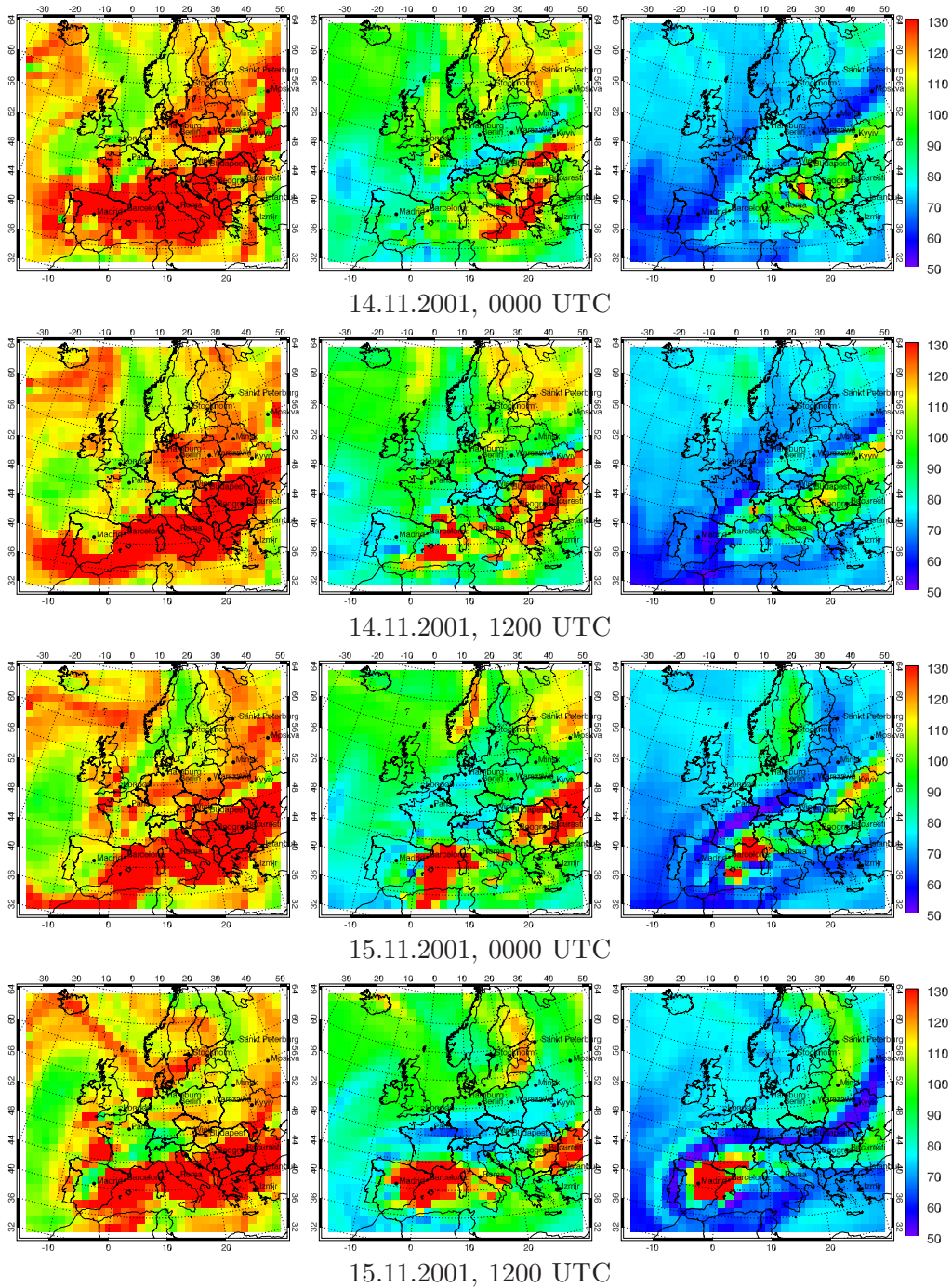


Figure 5.23: CONTRACE CG: Carbon monoxide mixing ratios [ppbV] during November 14 and 15, 2001. Left panel: Layer 10 (~ 800 m). Middle panel: Layer 16 (~ 2300 m). Right panel: Layer 21 (~ 5500 m).

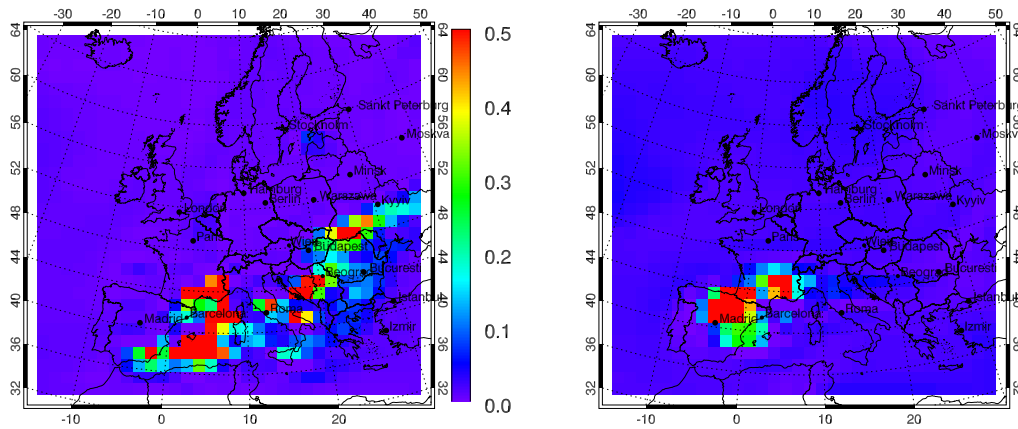


Figure 5.24: *CONTRACE CG: Nitrogen oxide mixing ratios [ppbV]. Left: Layer 16 (~ 2300 m) on November 14, 1200 UTC. Right: Layer 23 (~ 7700 m) on November 15, 1200 UTC.*

two days shown, a low pressure system develops over Corsica and slowly moves westward. As mentioned above, thunderstorms in concert with lightning occurred along the Mediterranean coast as well as strong uplifting of boundary layer air in clouds during November 14. This evolution is sketched out by the CO increase in the free troposphere due to rapid and strong uplifting, resulting in high CO mixing ratios. Another meteorological feature is found over the Balkan and eastern Europe. According to the model simulations, strong uplift along a frontal system occurred, too, showing characteristic aspects of convective transport by a north-easterly warm conveyor belt with raised air masses connected to cloud occurrence (see also Fig. 5.19). The mixing ratios of nitrogen oxide as one key target species of the CONTRACE campaign are shown in Fig. 5.24 for November 14, 1200 UTC in about 2300 m height and for November 15, 1200 UTC in about 7700 m height. On November 14, both uplift features can be seen in the NO mixing ratios. Over the Iberian peninsula and southern France, the highest NO mixing ratios in the upper troposphere occur on November 15, the sources of which may be strong uplifting in convective clouds as well as nitrogen oxide production by lightning. However, the chemical composition in this region is significantly changed and the air masses will presumably be transported over long distances.

While having shown only coarse simulations yet, the data assimilation results for the CONTRACE campaign data and the MOZAIC flights on the nested domain with a horizontal resolution of 25 km are shown in Fig. 5.25

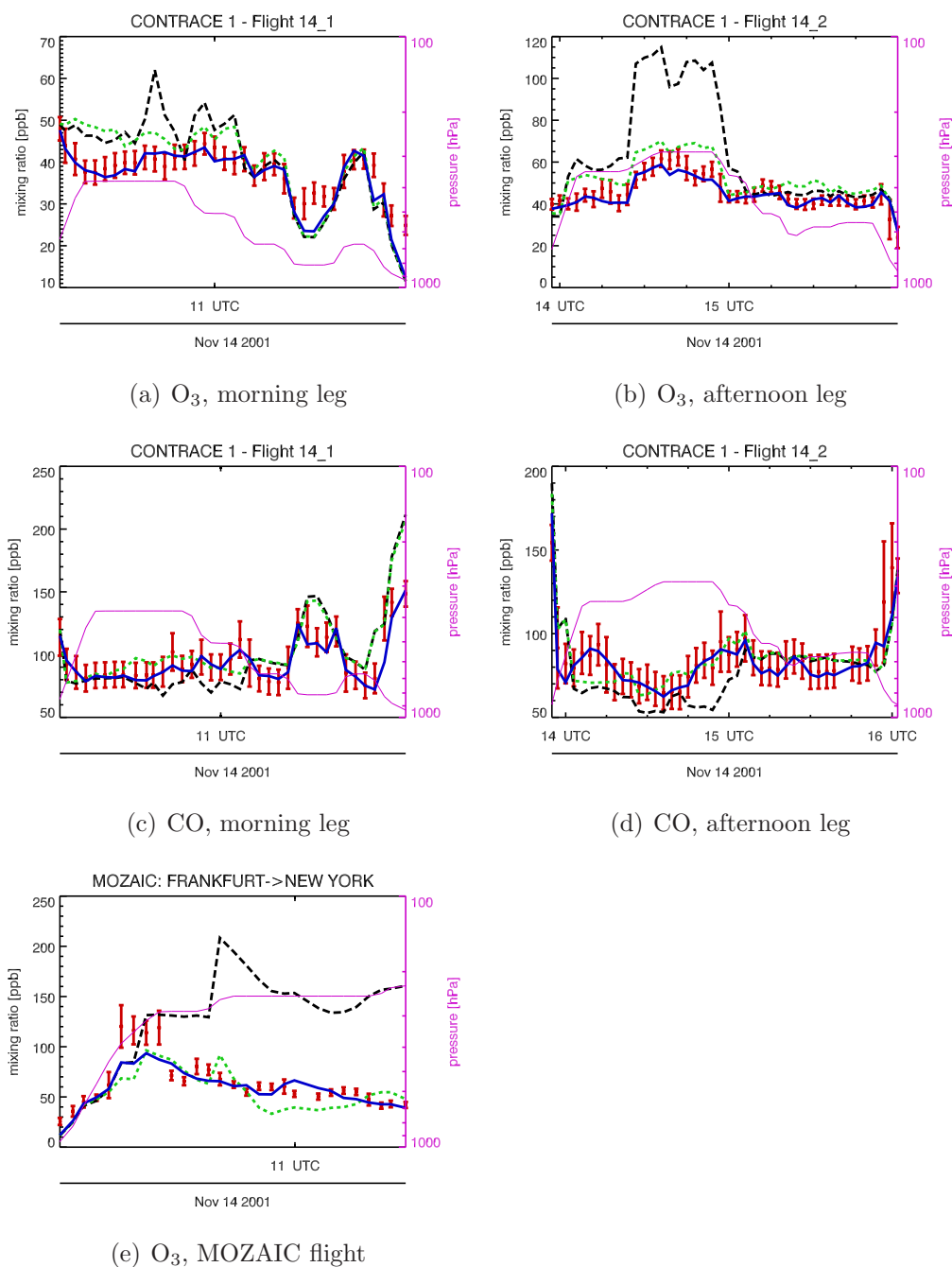


Figure 5.25: *CONTRACE N1: Observational campaign data (a-d) and MOZAIC data (e) for ozone (a,b,e) and carbon monoxide (c,d) on November 14, 2001. Red crosses: Observations. Dashed black line: Control run. Dashed green line: First guess run. Solid blue line: Analysis run. Solid magenta line: Flight altitude as pressure [hPa].*

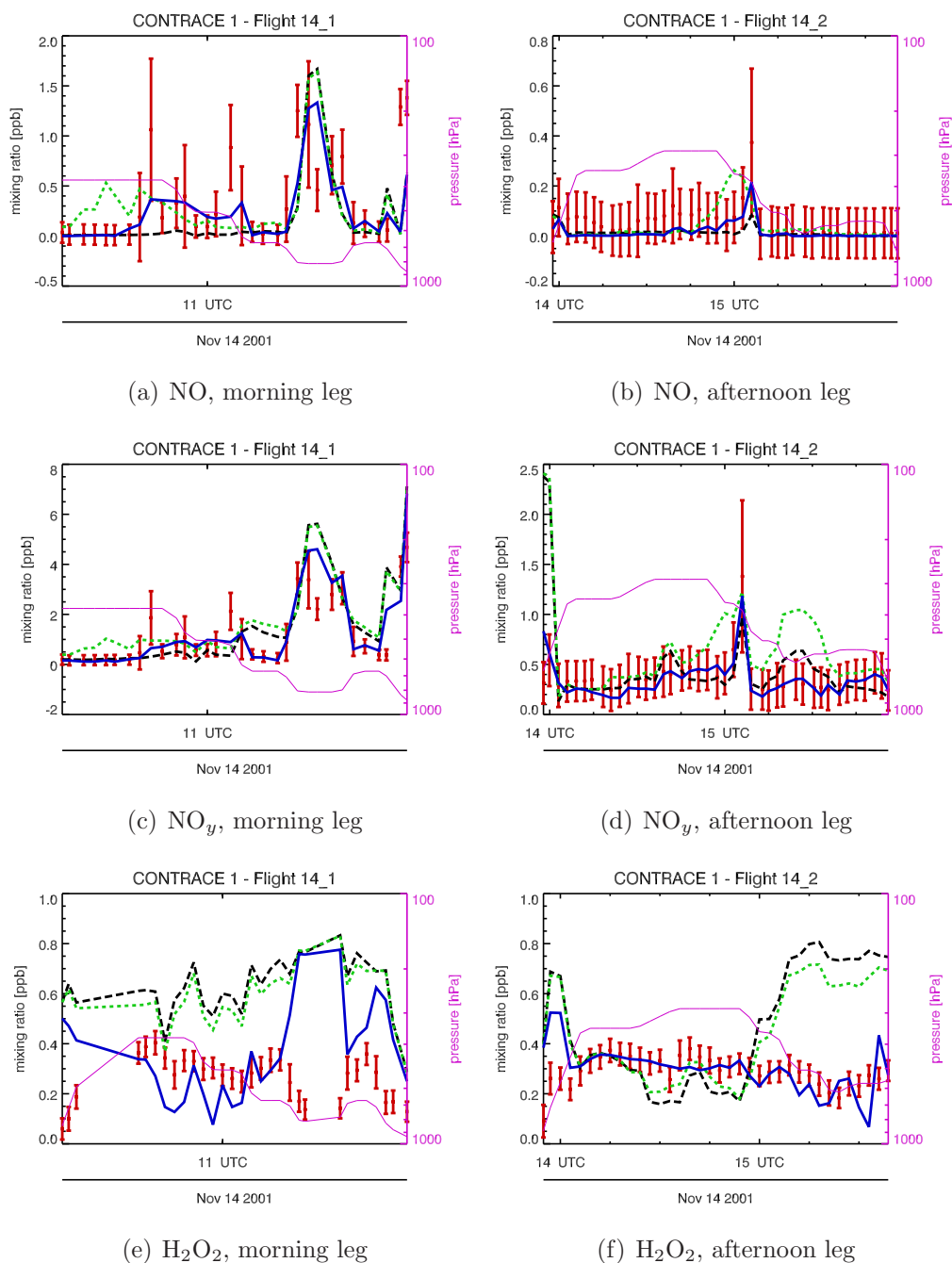


Figure 5.26: *CONTRACE N1: Observational campaign data and MOZIC data for nitrogen oxide (a,b), nitrogen compounds (c,d) and hydrogen peroxide (e,f). Red crosses: Observations. Dashed black line: Control run. Dashed green line: First guess run. Solid blue line: Analysis run. Solid magenta line: Flight altitude as pressure [hPa].*

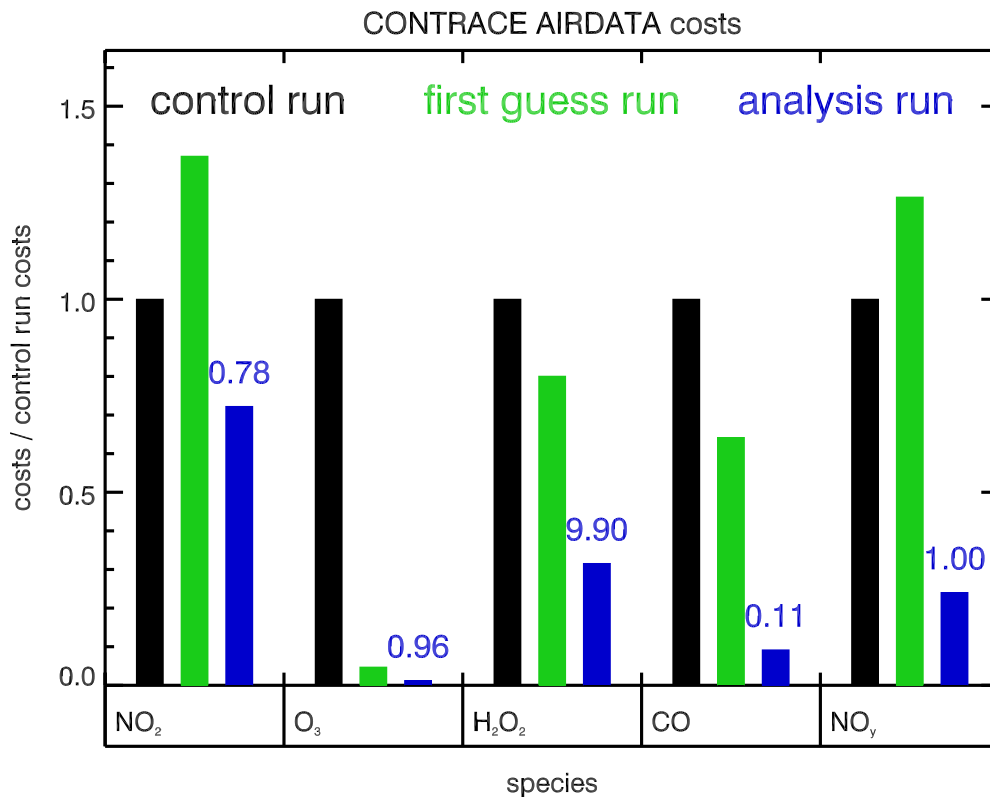


Figure 5.27: *CONTRACE N1: Partial costs by the campaign data, gathered on November 14, 2001, for the different model runs, normalised by the costs of the control run. Blue numbers give the value of cost per observation for the analysis run.*

for ozone and carbon monoxide and in Fig. 5.26 for nitrogen oxide, nitrogen compounds and hydrogen peroxide. The timely high resolving observations have been averaged as described in Section 4.4, in order to retrieve a suitable representativeness error. The moderate discrepancies for O₃ and CO in the control run are mostly corrected by the data assimilation procedures undertaken on the previous days. Thus the first guess run already produces good estimates for these two constituents in the free troposphere. The data assimilation procedure succeeds in fitting the model forecast excellently to the observations, made possible by large background errors compared to measurement errors.

The morning observations for NO and NO_y on the southward flight leg show limited areas with strongly varying mixing ratios, noticeable by outliers and their large error bars. The possible reasons of these high nitrogen oxide mixing ratios are those mentioned above, namely uplifted air from the

boundary layer as well as NO_x production by lightning. Although lightning is not parameterised in the current model version, NO and NO_y are satisfyingly simulated by the control run as well as by the first guess run. The analysis procedure significantly reduces the discrepancies — while weakly limited by smoothness constraints inferred by background error covariances (see Chapter 4). The hydrogen peroxide forecast is of moderate to bad quality and is not significantly improved in the first guess simulation. Only during the afternoon the analysis run produces good mixing ratios for H_2O_2 . The cost statistics for all campaign data constituents is shown in Fig. 5.27. The degraded forecast skill of the first guess run for NO and NO_y is probably due to the assimilation of satellite based NO_2 column retrievals, which are not shown here. Even by using additional averaging kernel for the column retrieval information as described in Section 3.6, the model fit to this kind of data inherits strong ambiguity that may produce the slightly worse first guess state encountered here, compared to the control simulation. Nevertheless, the analysis procedure is successful in correcting NO_2 and NO_y mixing ratios, leading to very good model predictions. Additionally, the cost value per observation is given for the analysis run, according to the observed species. Under the assumption of uncorrelated background and observation errors, one can find the expectation value of the cost function to be $p/2$, when p is the number of observations. Therefore in the optimal case, each observation would contribute with a partial cost of 0.5 to the cost function. The resulting partial costs for CO indicate over-fitting, while H_2O_2 cannot be reproduced sufficiently by the model simulation, mainly due to the time series of the morning flight.

A proper assimilation performance has been obtained for the campaign data, especially for ozone and carbon monoxide. The forecast skill for nitrogen compounds has not been substantially improved by assimilating satellite observations on the previous days, hampered by the ambiguity of introduced information and possible strong source processes that are not parameterised in the model. Nevertheless, the analysis procedure succeeds in correcting discrepancies by adjusting initial values only, leading to chemical consistent simulations on a daily basis. Generally, the feasibility of free troposphere simulations with the EURAD 4d-var data assimilation system can thus be stated.

5.4 SPURT IOP 2 Campaign

One aim of the SPURT (Trace Gas Transport in the Tropopause Region) project is to analyse the chemical state of the upper troposphere/lower stratosphere (UT/LS) region and its seasonal variations from the subtropics to the Arctic region. The project planned to help understanding the processes in the UT/LS region and the governing budgets, namely, *e. g.*, tropopause-stratosphere exchange by isentropic transport and diabatic descent of vortex air. To this end, air-borne in-situ measurements of H_2O , O_3 , NO_y , CO , N_2O , CH_4 , CO_2 , SF_6 , CFC-11 and CFC-12 were collected during eight sub-campaigns.

5.4.1 Model design and observational data basis

The SPURT intensive operational period (IOP) 2 has been selected for the analysis procedure conducted for the upper troposphere/lower stratosphere region. The suite of flights carried out from January 17 to January 19, 2002, provide observations at UT/LS heights on a south-north cross section over Europe (see Fig. 5.28). On January 17, the flight track headed from northern Germany to Portugal and further to Casablanca, with the home bound flight during the following day. On January 19 the measurements have been obtained on a flight to northern Norway and back on the same day. O_3 , CO and NO_y have been assimilated from the set of observed trace gases. Further, MOZAIC campaign data has been available for the selected period. In-situ air-borne observations have been treated during the assimilation as described in Section 4.4. In addition to air-borne data, GOME measurements were available as tropospheric NO_2 column retrievals (courtesy of KNMI) and NNORSY O_3 profile retrievals (*Müller et al.* [2003]). The experiments involve coarse grid data assimilation only, where the grid design is identical to the coarse grid used for the CONTRACE simulations (see Fig. 5.20) using the same vertical refinement in the UT/LS region. An enhanced horizontal grid resolution by a nested sub-domain, like for

	January 2002				
Nestl.	Tue 15	Wed 16	Thu 17	Fri 18	Sat 19
CG		■	■	■	■

Table 5.7: SPURT: Campaign simulation setup. Assimilation experiments are indicated by black solid bars with an assimilation window of 24 hours. January 15 has been used as spin up run.

the CONTRACE simulations, has not been undertaken due to the coarse structures observable at heights above 10 km. Due to the height of the observations, initial values have been the only control parameter to the optimisation procedure.

The main objective of the simulations performed here was to investigate the possible benefit of air-borne data on the forecast for satellite based observations. To this end, the satellite information has been assimilated on January 16 only, while, during the following days, the data has been withheld from the assimilation procedure. So, independent observations were available from January 17-19 in order to validate the forecast improvement by the data assimilation procedure.

5.4.2 Simulation Results

The air-borne time series of O_3 , CO and NO_y as collected by the SPURT campaign are given in Figs. 5.29 and 5.30 together with the model forecasts. Compared to the observations obtained by the CONTRACE campaign, the time series show less variations due to the significantly higher flight levels. Compared to the control run, the forecast quality of the first guess run has not been much improved by the previous assimilation of GOME O_3 profile and NO_2 column retrievals as well as MOZAIC air-borne data for January 16 and 17 due to the limited coverage of the assimilated data. However, the analysis procedure finds initial values at 00 UTC facilitating a model simulation of good, partly excellent agreement with the observations. The same applies for the MOZAIC data, which is shown in Fig. 5.31 exhibiting the partly excellent performance of the data assimilation algorithm.

This allows to state a sound performance of the chemistry transport model

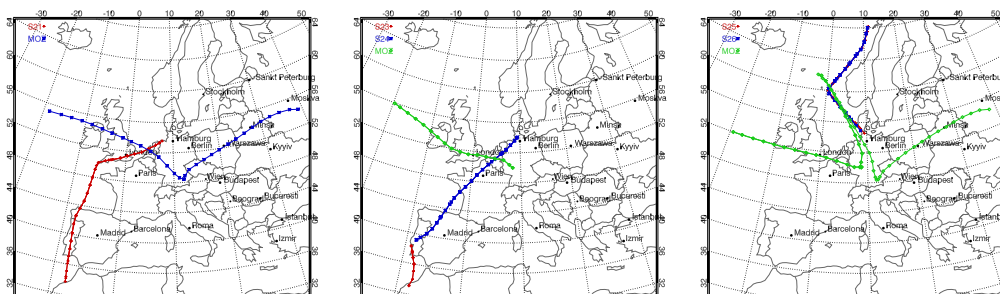


Figure 5.28: SPURT: Flight paths of the IOP2 campaign flights and the available MOZAIC flight tracks. From left to right: January 17-19.

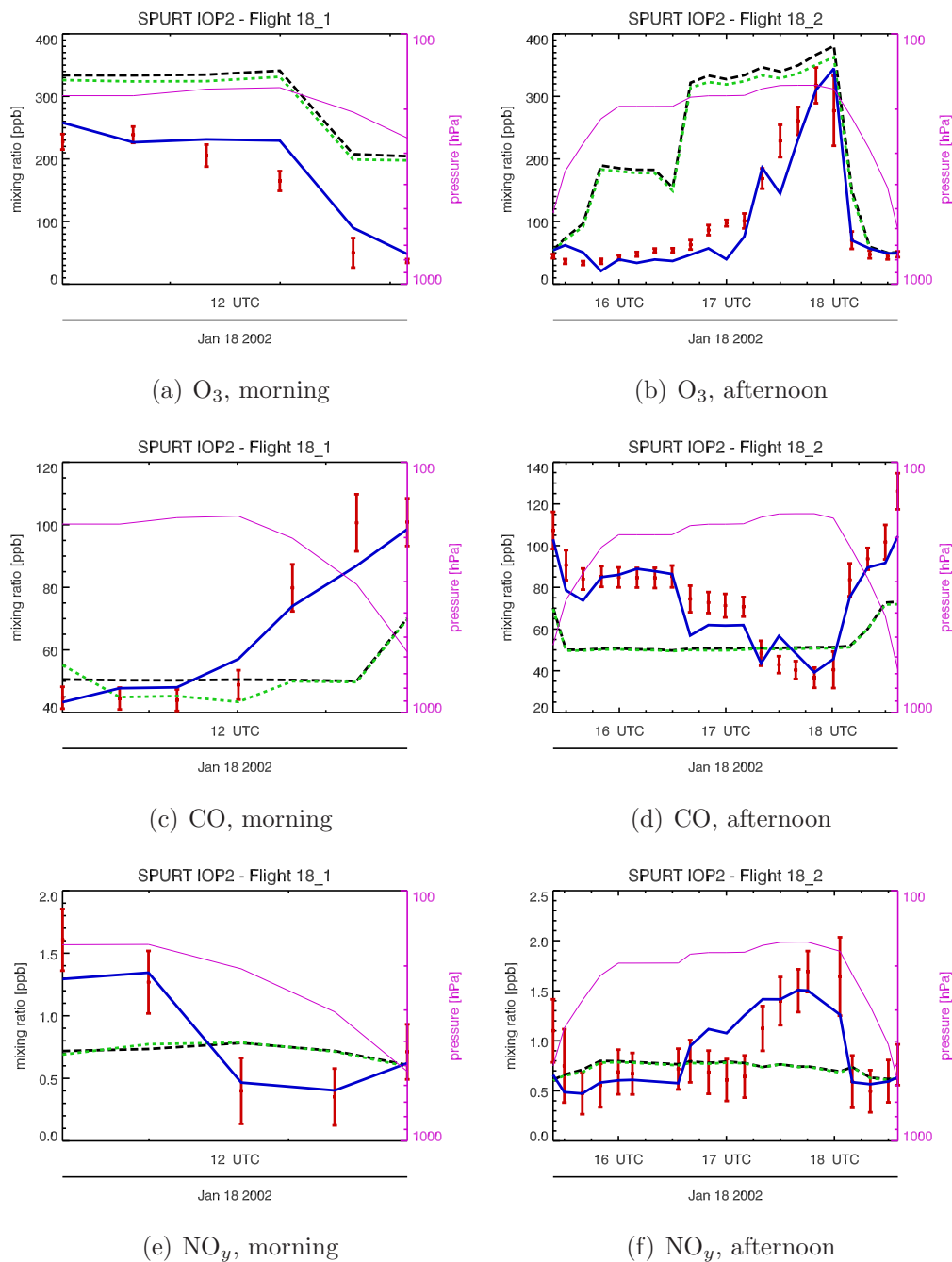


Figure 5.29: SPURT: Averaged IOP2 observations for January 18, 2002 and the model simulation results. Red crosses: Observations. Black dashed line: Control run. Green dashed line: First guess run. Blue solid line: Analysis run. Magenta line: Flight height.

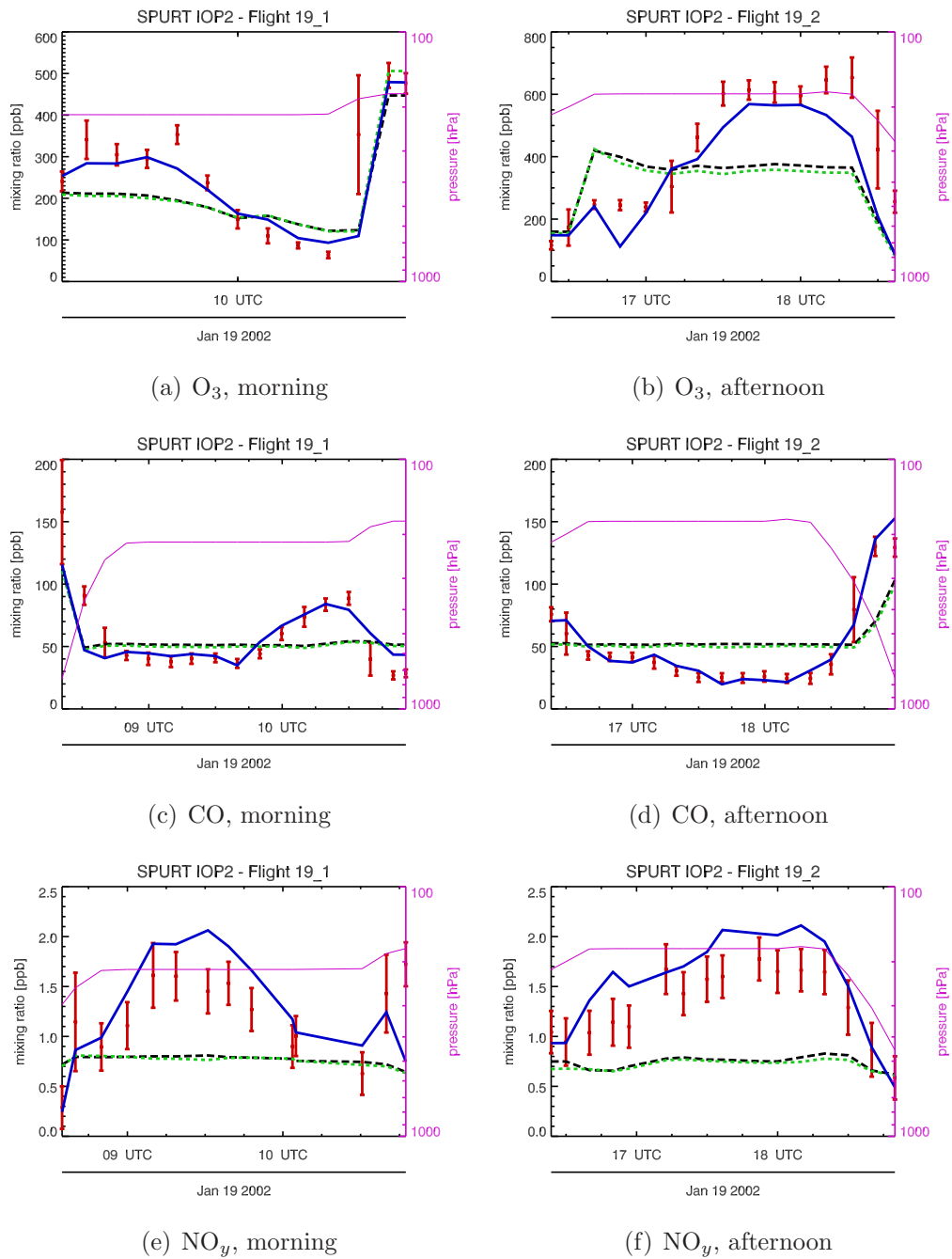


Figure 5.30: SPURT: Averaged IOP2 observations for January 19, 2002 and the model simulation results. Red crosses: Observations. Black dashed line: Control run. Green dashed line: First guess run. Blue solid line: Analysis run. Magenta line: Flight height.

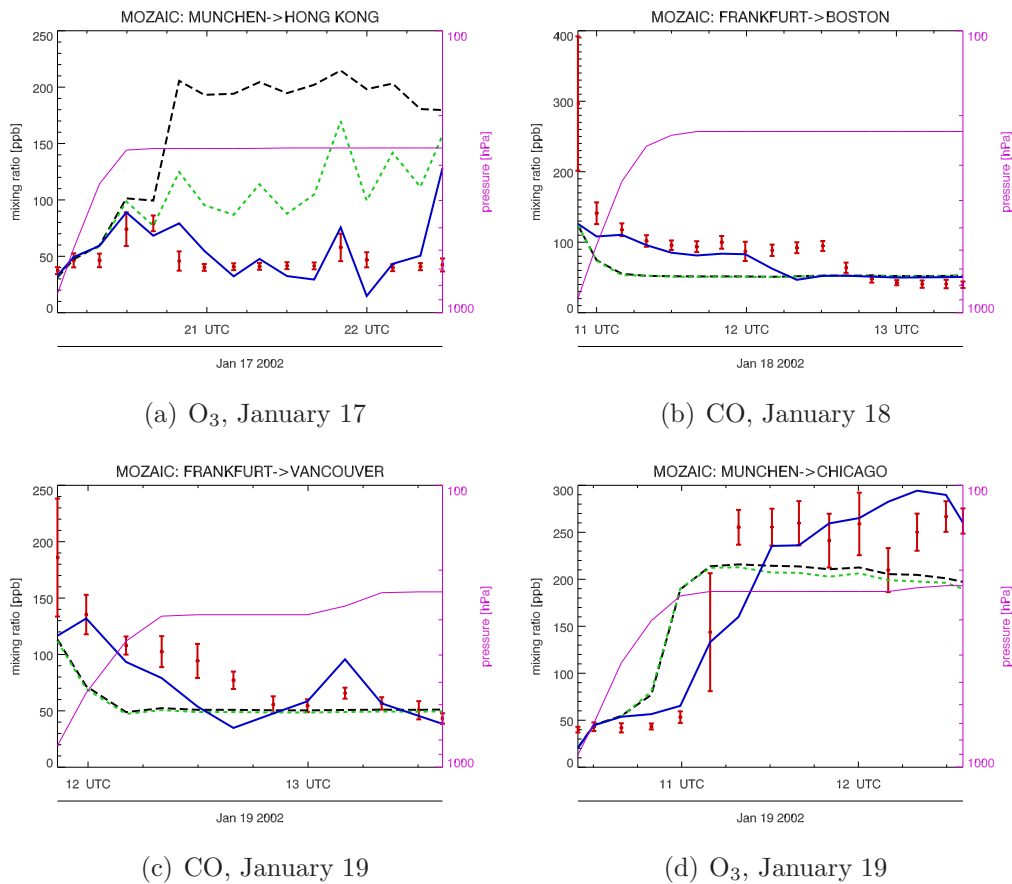


Figure 5.31: SPURT: Averaged MOZIAIC observations and model simulation results. Dates and constituents are given by sub-captions. Red crosses: Observations. Black dashed line: Control run. Green dashed line: First guess run. Blue solid line: Analysis run. Magenta line: Flight height.

in simulating the chemical processes in the UT/LS region, at least over periods of about some days. Even in the tropopause region the EURAD 4d-var data assimilation system thus enables to produce chemically consistent estimates of the state of the atmosphere.

As mentioned above, the benefit impact of the assimilated air-borne campaign data on the forecast skill for satellite based measurements should be investigated. To this end, Fig. 5.32a shows the footprint area of the available but not assimilated satellite data on January 18, the blue shaded track being the track of special interest. This track, measured at about 0930 UTC, is crossed by the SPURT flight on the same day some hours later, between 14 and 17 UTC in the afternoon (compare Figs. 5.28 and

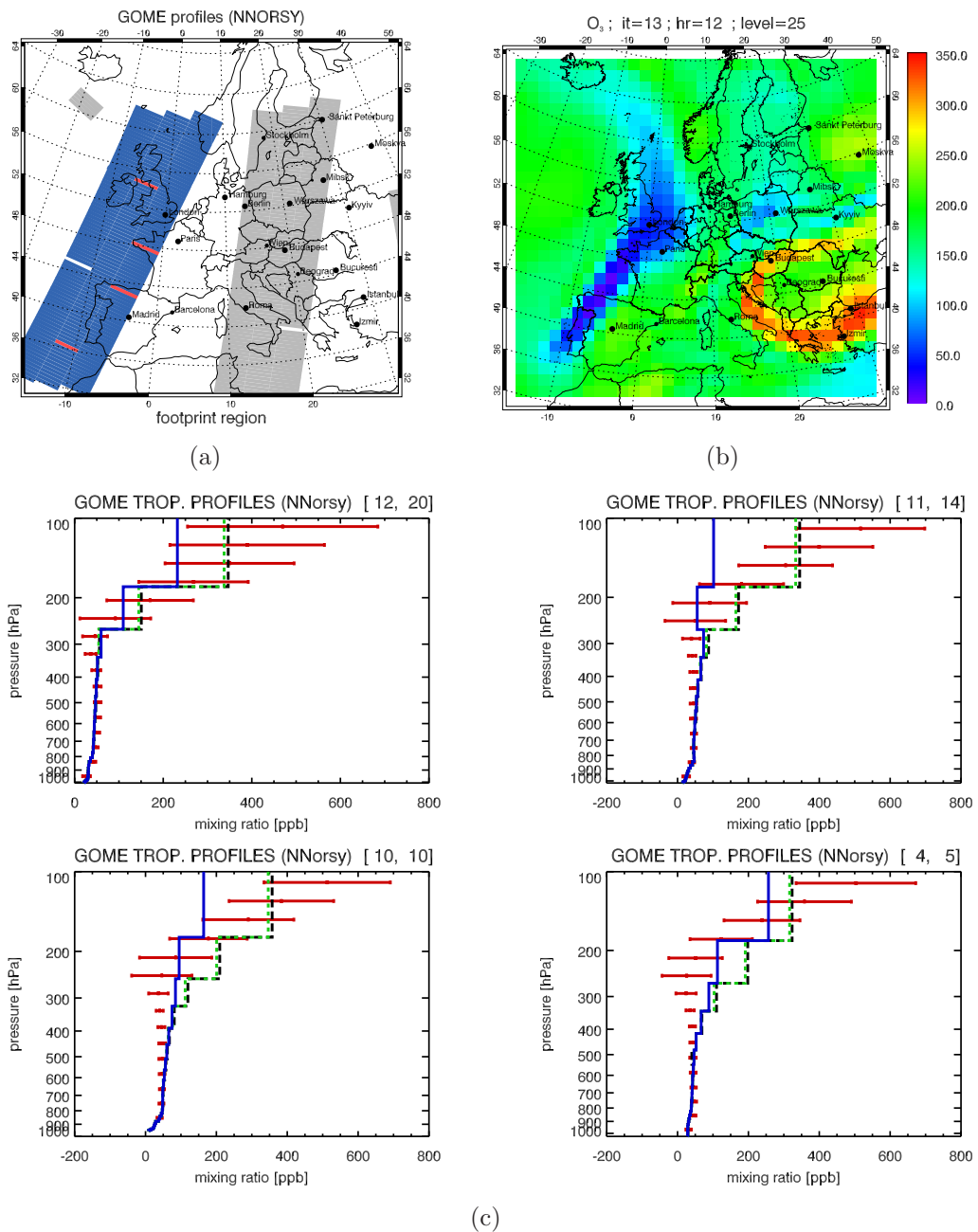


Figure 5.32: SPURT: Impact of campaign data on model equivalents for GOME O₃ profile retrievals. (a) GOME observational footprint areas for January 18. The footprints of the profiles in (c) are filled in red, while the blue shaded data has been used to calculate bias and rms error of the GOME observations, shown in Table 5.8. (b) Assimilation result for O₃ at ≈ 11 km height and 1200 UTC. (c) Example O₃ profile retrievals (red) and model equivalents. Black dashed line: Control run. Green dashed line: First guess run. Blue solid line: Analysis run.

5.29). The analysis result for ozone at model level 25 ($\sim 10\text{-}12$ km height) is presented in Fig. 5.32b. The flight path from the Iberian Peninsula to Germany is obvious by means of reduced ozone mixing ratios. Additionally, four representative ozone profiles of the middle satellite track are given in Fig. 5.32c, clearly showing an analysis state different to the first guess and the control run in the uppermost model layers (the profile locations are indicated by red footprints in Fig. 5.32a). The analysis procedure reduces the discrepancies between the forecast and the satellite observations, at least at observational heights (*i. e.*, the flight level). A quantification of the forecast improvement is given in Table 5.8, where the bias (observation minus model) and rms error with respect to the GOME data for the middle satellite track and observation heights between 150 and 400 hPa are given. The assimilation procedure on January 16 and 17 already leads to a small reduction of both the overestimation of the ozone mixing ratio and the rms error. The introduction of the air-borne data — gathered some hours later that day on January 18, reduces the bias to almost 25 % of its control run value, while the rms error is about 15 % lower. Hence, slight to moderate improvements in the model state become apparent when compared to the O_3 profile retrievals, such that the information inherent in both data sources mainly agree and a successful cross-validation of information sources can be stated. The weak deterioration of the model equivalent to the satellite data in the uppermost layer is due to the insufficient vertical resolution at that height. As a first solution, a vertical interpolation algorithm in the observation operator would presumably weaken this effect. However, a significantly higher vertical resolution would allow the air-borne data to improve the model state at flight height, while layers above could stay unchanged and thus reflect the strong vertical and presumably horizontal ozone gradient in the UT/LS region.

Table 5.8: *SPURT: Bias (observation minus model) and rms error for GOME O_3 satellite profiles on January 18, 2002, observed between 150 and 400 hPa and westward of 8°E .*

	bias [ppbV]	rms error [ppbV]
control run (cr)	-48.4	88.4
first guess run (fgr)	-42.7	85.8
analysis run (ar)	-12.8	74.8

CHAPTER 6

Summary and Discussion of Results

Previous work clearly demonstrated the potential of 4d-var data assimilation. Nevertheless, a variety of features remained missing, hampering the full exploitation of key features of both chemistry transport modelling and advanced data assimilation. In order to further enhance the usage of the assimilation system for comprehensive case study analyses, a number of developments was due. The current version of the EURAD 4d-var data assimilation system has been developed by a complete revision of the precursor version. In the course of this work many features have been added to the assimilation procedure in order to enhance the model performance. Firstly, a sophisticated modelling of the background error covariance matrix following *Weaver and Courtier* [2001] has been implemented. Additionally, using pollutant group based statistics, a suitable emission factor error covariance matrix has been coded, introducing correlations between different emitted constituents. Hence, the number of degrees of freedom has substantially been reduced, forcing the scaling emission factors to stay in realistic ranges. Further, the covariance matrices are used to precondition the minimisation problem.

The first adjoint nesting technique in chemical 4d-var data assimilation has been developed for the EURAD assimilation system. This enables the chemistry transport model and its adjoint to telescope from regional scale down to urban scale, making arbitrary horizontal resolutions available. A systematic informational data transfer from mother domains to nested grids is estab-

lished, facilitating model simulations on nested grids with both optimised initial and boundary values. Moreover, optimised emission rate estimates are passed from one assimilation experiment to the following, allowing for a continual improvement by long term simulations.

Facing a wide variety of different measurement types and retrieval products, a sophisticated observation operator has been implemented in order to avoid any restriction on possible information sources. This allows assimilation of complete data sets from measurement campaigns with different scientific aims. Testing the boundary layer performance of the assimilation system, simulations using BERLIOZ and VERTIKO data have been performed, while the campaigns CONTRACE and SPURT focussed on lower/free troposphere and upper troposphere/lower stratosphere properties, respectively.

Besides assessing the cost function evolution, quality control of the forecast behaviour for the various assimilation experiments has been performed using the statistical measures bias and rms error. The simulations for the boundary layer campaigns BERLIOZ and VERTIKO used — as key features — the adjoint nesting technique and the joint optimisation of initial values and emission rates. The joint optimisation has been successfully employed, exhibiting to be generally superior to the single initial value and emission factor optimisations. Constituents like ozone, carbon monoxide and benzene prove to be analysed well to excellently by initial value optimisation, while sulphur dioxide, nitrogen compounds and volatile organic compounds (VOCs) are quite more sensitive to emission factor inversion. However, the joint optimisation mode effectively joins the benefits of both methods, thus, in most cases, leading to the best analysis for the trace gases considered in this thesis. Improving the horizontal grid resolution on the nested domains resulted in strongly reduced representativeness errors of the observations. While the bias and rms error values of all species are significantly reduced by enhanced grid resolution, the analysis of species with high spatial variability in urban areas, which have previously not been amenable to optimisation, is facilitated. However, the highest resolution of 5 km grid spacing turns out to still be insufficient for a perceptible optimisation of nitrogen oxide. The reason of which may be a still too poor representation of nitrogen oxide observations for this model resolution due to the proximity of measurement devices to sources, *e. g.* roads. Moreover, even a too low vertical resolution may have reasonable impact on the modelled near surface concentrations, due to the direct "dispersion" of emission strength over the whole grid volume. Another reason may be essentially too low emission estimates for nitrogen oxides in strongly polluted areas, which bars the

optimisation procedure from a considerable improvement of forecasted nitrogen oxide mixing ratios. Nevertheless, the optimisation procedure shows strongly improved model forecast behaviour, because of analysed emission factors, which are direct estimates of enhanced surface fluxes.

Model simulations for a selected flight of the CONTRACE campaign show a sound performance of the model in the free troposphere. Optimising only initial values at 00 UTC during a model suite of a few days leads to chemical consistent simulations — on a daily basis — with good to excellent forecast skills for ozone and carbon monoxide. A very good performance for nitrogen compounds is hampered presumably by the lack of lightning induced production of nitrogen oxides in the model. Nevertheless, the analysis for observed nitrogen compounds is considerably improved compared to the control run. Another key result is, that strong uplift of near surface air into the free troposphere by clouds, being a target of the CONTRACE campaign, could be identified in the model simulations. Focussing on the upper troposphere/lower stratosphere (UT/LS) region, the simulations for the SPURT IOP 2 campaign exhibit a similar performance. The analysis procedure finds drastically improved initial values allowing to reproduce the measured mixing ratios for ozone, carbon monoxide and nitrogen compounds in the UT/LS region.

In addition to in-situ observations, satellite-borne earth observations have been assimilated in all four campaign simulations. The beneficial impact on the quality of ozone prediction in the free troposphere has been shown by both forecast improvement as well as by observations withheld from the assimilation procedure. It has been found, that — due to the influence of horizontal boundary conditions by advection — the improved predictive skill for ozone in the middle and upper troposphere is lost after one week without application of the 4D-var procedure. Further, GOME ozone profile retrievals withheld from the assimilation procedure, have been used for quality control, showing the beneficial impact of SPURT air-borne data on the forecasted ozone values. Moreover, retained column retrievals from SCIAMACHY observations have been analysed by introducing GOME satellite retrievals. Hence, cross validations of different measurement types can be demonstrated.

Finally, the system demonstrated the ability to provide improved chemical consistent simulation results throughout the troposphere, from polluted urban environments up to the upper troposphere. It turns out, that the features of the EURAD 4d-var data assimilation system combine to form a powerful tool for analysing tropospheric trace gas distributions. The combination of the model facilities makes this system unique in the world. Future

applications include long-term simulations with emission factor optimisation, offering the potential of reliable emission rate estimates obtained in a rigorous way as well as nested simulations in a hemispheric mother domain¹. Moreover, the operator approach for the background error covariances allows to flexibly take into account anisotropic and inhomogeneous correlations. This will enable to model correlation features in the background error covariance matrix reflecting coherent areas of homogeneous surface textures. The new features described in this study are adopted as key contributions of the two presently developed projects, namely the European Space Agency project PROMOTE and GEMS², funded by the European Union, and is scheduled for operational use.

¹For first simulation results on the northern hemisphere with the EURAD model see db.eurad.uni-koeln.de/prognose/chem.html

²Please visit www.gse-promote.org and www.ecmwf.int/research/EU_projects/GEMS

APPENDIX A

Vertical Grid Structure

The vertical Lorentz type grid structure used in this work is given in the following tables. The number of layers has been extended from 15 (*Elbern et al.* [1997]) to 23 for the BERLIOZ and VERTIKO simulations, with refinements in the boundary layer. The CONTRACE and SPURT simulations have been performed on 26 layers, where the additional layers increased the resolution in the free troposphere.

Table A.1: *The vertical grid structure for the 23 layer design, given by the terrain following σ coordinate. Pressures are given for a surface pressure of 1013.25 hPa, and heights are valid for the U.S. standard atmosphere.*

Layerindex	σ value	Pressure [hPa]	Height [m]
1	1.0 - 0.995	1013.25 - 1008.68	0 - 38
2	0.995 - 0.99	1008.68 - 1004.12	38 - 76
3	0.99 - 0.985	1004.12 - 999.55	76 - 115
4	0.985 - 0.98	999.55 - 994.99	115 - 153
5	0.98 - 0.97	994.99 - 985.85	153 - 231
6	0.97 - 0.96	985.85 - 976.72	231 - 309
7	0.96 - 0.945	976.72 - 963.02	309 - 427
8	0.945 - 0.93	963.02 - 949.32	427 - 546
9	0.93 - 0.91	949.32 - 931.06	546 - 708
10	0.91 - 0.89	931.06 - 912.79	708 - 872

Layerindex	σ value	Pressure [hPa]	Height [m]
11	0.89 - 0.865	912.79 - 889.96	872 - 1081
12	0.865 - 0.84	889.96 - 867.13	1081 - 1294
13	0.84 - 0.81	867.13 - 839.73	1294 - 1556
14	0.81 - 0.78	839.73 - 812.34	1556 - 1825
15	0.78 - 0.74	812.34 - 775.81	1825 - 2196
16	0.74 - 0.70	775.81 - 739.28	2196 - 2581
17	0.70 - 0.60	739.28 - 647.95	2581 - 3615
18	0.60 - 0.50	647.95 - 556.63	3615 - 4775
19	0.50 - 0.40	556.63 - 465.30	4775 - 6101
20	0.40 - 0.30	465.30 - 373.98	6101 - 7658
21	0.30 - 0.20	373.98 - 282.65	7658 - 9560
22	0.20 - 0.10	282.65 - 191.33	9560 - 12064
23	0.10 - 0.0	191.33 - 100.00	12064 - 16179

Table A.2: The vertical grid structure for the 26 layer design, given by the terrain following σ coordinate. Pressures are given for a surface pressure of 1013.25 hPa, and heights are valid for the U.S. standard atmosphere.

Layerindex	σ value	Pressure [hPa]	Height [m]
1	1.0 - 0.995	1013.25 - 1008.68	0 - 38
2	0.995 - 0.99	1008.68 - 1004.11	38 - 76
3	0.99 - 0.985	1004.11 - 999.55	76 - 115
4	0.985 - 0.98	999.55 - 994.99	115 - 153
5	0.98 - 0.97	994.99 - 985.85	153 - 231
6	0.97 - 0.96	985.85 - 976.72	231 - 309
7	0.96 - 0.945	976.72 - 963.02	309 - 427
8	0.945 - 0.93	963.02 - 949.32	427 - 546
9	0.93 - 0.91	949.32 - 931.06	546 - 708
10	0.91 - 0.89	931.06 - 912.79	708 - 872
11	0.89 - 0.865	912.79 - 889.96	872 - 1081
12	0.865 - 0.84	889.96 - 867.13	1081 - 1294
13	0.84 - 0.81	867.13 - 839.73	1294 - 1556
14	0.81 - 0.78	839.73 - 812.33	1556 - 1825
15	0.78 - 0.74	812.33 - 775.81	1825 - 2196
16	0.74 - 0.70	775.81 - 739.28	2196 - 2581
17	0.70 - 0.65	739.28 - 693.61	2581 - 3084
18	0.65 - 0.60	693.61 - 647.95	3084 - 3615

Layerindex	σ value	Pressure [hPa]	Height [m]
19	0.60 - 0.54	647.95 - 593.16	3615 - 4294
20	0.54 - 0.48	593.16 - 538.36	4294 - 5025
21	0.48 - 0.41	538.36 - 474.43	5025 - 5959
22	0.41 - 0.34	474.43 - 410.51	5959 - 7001
23	0.34 - 0.26	410.51 - 337.45	7001 - 8368
24	0.26 - 0.18	337.45 - 264.39	8368 - 9999
25	0.18 - 0.09	264.39 - 182.19	9999 - 12374
26	0.09 - 0.0	182.19 - 100.00	12374 - 16179

APPENDIX B

Error Specifications

Initial Value Background Errors

A crude estimate of the errors in the initial values can be obtained by taking a fix but reasonable fraction of the fields themselves. Reflecting the lack of direct observations and thus a reduced analysis skill in the upper model domain, a height dependence has been introduced, while preserving a lower

Table B.1: Minimum estimates for the absolute background errors. The others block contains all constituents not mentioned.

Species	e^{iv} [ppbV]	Species	e^{iv} [ppbV]	Species	e^{iv} [ppbV]
SO ₂	20	HCHO	10	OL2/ETE	5
SULF	5	NH ₃	10	OLT	10
NO ₂	5	HC3	5	OLI	5
NO	5	HC5	5	TOL	5
O ₃	15	HC8	5	XYL	5
HNO ₃	5	ETH	5	ISO	5
ALD	10	CO	400	others	1

bound for the absolute values. Let i, j , and k denote the spatial indices of the background field, and let l denote the species index. The estimates of the background errors are given by

$$\varepsilon_{i,j,k,l}^b = \max \left(x_{i,j,k,l}^b e_k^{\text{rel,iv}}, e_l^{\text{iv}} \right),$$

where the relative error $e_k^{\text{rel,iv}}$ is given by

$$e_k^{\text{rel,iv}} = 0.5 \exp \left[0.69 \frac{k-1}{k_{\text{max}}-1} \right], \quad k = 1, \dots, k_{\text{max}}$$

leading to a minimum relative error of 50% for the surface layer ($k = 1$) and 100% for the uppermost layer ($k = k_{\text{max}}$). The minimum absolute errors e^{iv} are given in Table B.1.

Observation Errors

The observation errors are the sum of measurement and representativeness errors,

$$\varepsilon_i^o = \varepsilon_i^m + \varepsilon_i^r, \quad i = 1, \dots, p \quad (\text{B.1})$$

Table B.2: Minimal absolute and relative measurement errors for ground based measurement stations, following the suggestions by Mohnen [1999].

Species	e^m [ppbV]	$e^{\text{rel,m}}$ [%]	Species	e^m [ppbV]	$e^{\text{rel,m}}$ [%]
SO ₂	1.5	10	NO ₃ [*]	0.01	20
NO ₂	1.5	15	PAN [*]	0.1	20
NO	1.0	10	CO	30.0	15
O ₃	2.0	10	HONO [*]	0.01	20
HNO ₃ [*]	0.3	20	HNO ₄ [*]	0.1	20
H ₂ O ₂ [*]	0.2	10	NO _y	2.5	25
HCHO [*]	0.5	10	NO _x	1.5	15
NH ₃ [*]	0.3	30	C ₆ H ₆ [*]	0.2	20
N ₂ O ₅ [*]	0.1	20			

^{*} No reference has been available and the given value has been estimated for the simulations performed in this thesis.

Table B.3: Absolute representativeness error estimates at surface level. To retrieve the representativeness error of a ground based observation, the given values are scaled by a factor determined by the grid resolution and the station characteristics.

Species	e_{surface}^r [ppbV]	Species	e_{surface}^r [ppbV]	Species	e_{surface}^r [ppbV]
SO ₂	0.4	HCHO	0.05	HONO	0.05
NO ₂	1.4	NH ₃	0.2	HNO ₄	0.3
NO	3.0	N ₂ O ₅	0.2	NO _y	0.5
O ₃	1.2	NO ₃	0.1	NO _x	0.5
HNO ₃	0.05	PAN	0.02	C ₆ H ₆	0.2
H ₂ O ₂	0.05	CO	15.0		

with p being the number of observations.

For a ground based observations lacking error information, an error estimate is obtained by

$$\varepsilon^m = \max(e^m, e^{\text{rel},m} \cdot y).$$

The values for e^m and $e^{\text{rel},m}$ follow suggestions devised by *Mohnen* [1999] and are given in Table B.2.

In this work the representativeness errors are retrieved by defining a characteristic absolute error e^f for each species and level, scaled by a factor d determined by the grid resolution Δx and a characteristic representativeness length L^f , which depends on the observation type and — if it is a ground based observation — on the station characteristics. The error is defined to

Table B.4: Representativeness lengths for ground based observations, broken down by the station characteristics.

Station Type	L^f [km]	Station Type	L^f [km]	Station Type	L^f [km]
Remote	20.0	Suburban	4.0	Traffic	1.0
Rural	10.0	Urban	2.0	Unknown	3.0

be

$$\begin{aligned}\varepsilon^r &= d \cdot e^r \\ &= \sqrt{\frac{L^r}{\Delta x}} \cdot e^r.\end{aligned}\tag{B.2}$$

The estimates of the absolute errors e^r are given for the surface level in Table B.3, and the characteristic representativeness lengths are shown in Table B.4. For Lidar observations and tethered balloon soundings, the representativeness length is estimated by $L^r = 250$ km, resulting in a small representativeness error for the grid resolutions encountered in this work. For observations by satellite retrievals, the representativeness error is assumed to depend not on the horizontal grid resolution, since the usual pixel size is larger than the grid spacing. Representativeness errors on satellite column retrievals are therefore neglected. For satellite profile retrievals the representativeness error is taken to be $\varepsilon^r = 0.05 \cdot e^r$ with respect to the possible coarser vertical grid resolution than that of the profile retrieval. As mentioned in Section 4.4 the air-borne measurements are averaged and the encountered variance is used to determine the representativeness error of the averaged observation.

APPENDIX C

Definition of Statistical Measures

In this thesis the following statistical measures have been used (\mathbf{y} and $\boldsymbol{\sigma}$ denote observations and their standard deviations, respectively, and $\hat{\mathbf{x}}$ are the model equivalents, all of size N):

- bias (observation minus model)

$$\langle \mathbf{y} - \hat{\mathbf{x}} \rangle = \frac{1}{N} \sum_{j=1}^N y_j - \hat{x}_j \quad (\text{C.1})$$

- normalised bias (observation minus model)

$$\langle \widehat{\mathbf{y} - \hat{\mathbf{x}}} \rangle = \frac{1}{N} \sum_{j=1}^N \frac{y_j - \hat{x}_j}{\sigma_j} \quad (\text{C.2})$$

- root mean square error (rms)

$$\text{rms} = \sqrt{\frac{1}{N} \sum_{j=1}^N (y_j - \hat{x}_j)^2} \quad (\text{C.3})$$

- normalised rms

$$\widehat{\text{rms}} = \sqrt{\frac{1}{N} \sum_{j=1}^N \left(\frac{y_j - \hat{x}_j}{\sigma_j} \right)^2} \quad (\text{C.4})$$

Bibliography

- Arakawa, A. and V. R. Lamb**, Computational design for the basic dynamical processes of UCLA general circulation model, *Meth. Comp. Phys.*, 17, 174–264, 1977.
- C. Bernhofer and B. Köstner (Ed.), *Vertical Transport of Energy and Trace Gases at Anchor Stations and Their Spatial/Temporal Extrapolation under Complex Natural Conditions (VERTIKO): Project Summary*, vol 12, Tharandter Klimaprotokolle, Technische Universität Dresden, 2005.
- Bjerknes, V.**, Dynamic meteorology and hydrography, in *Kinematics*, Carnegie Institute, Gibson Bros., New York, 1911.
- Bott, A.**, A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes, *Month. Weath. Rev.*, 117, (5), 1006–1015, 1989.
- Bouttier, F. and P. Courtier**, Data assimilation concepts and methods, ECMWF Meteorological Training Course Lecture Series, 1999.
- Chang, J. S., R. A. Brost, S. A. Isaksen, S. Madronich, P. Middleton, W. R. Stockwell and C. J. Walcek**, A three dimensional Eulerian acid deposition model: Physical concepts and formulation, *J. Geophys. Res.*, 92, (D12), 14,681–14,700, 1987.
- Chao, W. C. and L. P. Chang**, Development of a four-dimensional variational analysis system using the adjoint method at GLA Part I: Dynamics, *Month. Weath. Rev.*, 120, (8), 1661–1673, 1992.

- Cohn, S. E.**, An introduction to estimation theory, *J. Meteor. Soc. Japan*, 75, 257–288, 1997.
- Constantinescu, E. M., A. Sandu, T. Chai and G. R. Carmichael**, Ensemble-based chemical data assimilation II: Real observations, Technical Report TR-06-07, Computer Science, Virginia Polytechnic Institute and State University, 2006.
- Courant, R. and D. Hilbert**, *Methods of Mathematical Physics*, vol 1, Wiley-Interscience, 1953.
- Courtier, P.**, Variational methods, *J. Meteor. Soc. Japan*, 75, (1B), 211–218, 1997.
- Courtier, P., J. N. Thépaut and A. Hollingsworth**, A strategy for operational implementation of 4D-Var, using an incremental approach, *Q. J. R. Meteorol. Soc.*, 120, (519), 1367–1387, 1994.
- Daley, R.**, *Atmospheric Data Analysis*, Cambridge Univ. Press, 1991.
- Damian-Iordache, V.**, *KPP - Chemistry Simulation Development Environment*, Ph.D. Thesis, The University of Iowa, Dezember 1996.
- Elbern, H.**, Inversion of a complex chemistry transport model, Habilitationsschrift vorgelegt der Mathematisch-Naturwissenschaftlichen Fakultät der Universität zu Köln, Januar 2001.
- Elbern, H. and H. Schmidt**, A four-dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling, *J. Geophys. Res.*, 104, (D15), 18,583–18,598, 1999.
- Elbern, H. and H. Schmidt**, Ozone episode analysis by four-dimensional variational chemistry data assimilation, *J. Geophys. Res.*, D4, (106), 3569–3590, 2001.
- Elbern, H., H. Schmidt and A. Ebel**, Variational data assimilation for tropospheric chemistry modeling, *J. Geophys. Res.*, 102, (D13), 15,967–15,985, 1997.
- Elbern, H., H. Schmidt and A. Ebel**, 4D-variational data assimilation for an Eulerian chemistry transport model. proceedings of the NATO advanced research workshop., in *Large-Scale Computations in Air Pollution Modelling*, 151–160, Kluwer Academic, 1999.

- Elbern, H., H. Schmidt, O. Talagrand and A. Ebel**, 4D-variational data assimilation with an adjoint air quality model for emission analysis, *Environ. Model. and Software*, 15, 539–548, 2000.
- Elbern, H., A. Strunk, H. Schmidt and O. Talagrand**, Emission rate and chemical state estimation by 4-dimensional variational inversion, in preparation, 2006.
- Eliassen, A.**, Provisional report on calculation of spatial covariance and autocorrelation of the pressure field, technical Report 5, Videnskaps-Akademiet Institut for Vaer og Klimaforskning, Oslo, 1954.
- Eskes, H. J. and K. F. Boersma**, Averaging kernels for DOAS total-column satellite retrievals, *Atmos. Chem. Phys.*, 3, 1285–1291, 2003.
- Faure, C. and Y. Papegay**, Odyssee Version 1.7 Users Guide, Technical Report 0224, Institut National de Recherche en Informatique et en Automatique, September 1998.
- Finlayson-Pitts, B. J. and J. N. Pitts**, *Atmospheric Chemistry*, Wiley-Interscience, 1986.
- Fisher, M. and D. Lary**, Lagrangian four-dimensional variational data assimilation of chemical species, *Q. J. R. Meteorol. Soc.*, 121, (527), 1681–1704, 1995.
- Fletcher, S. J. and M. Zupanski**, A data assimilation method for log-normal variables and errors, submitted to *Q. J. R. Meteorol. Soc.*, 2006.
- Gandin, L.**, Objective analysis of meteorological fields, Leningrad: Gridromet, 1963.
- Ghil, M.**, Meteorological data assimilation for oceanographers. Part I: Description and theoretical framework, *Dyn. Atmos. Oceans*, 13, 171–218, 1989.
- Ghil, M. and P. Malanotte-Rizzoli**, Data assimilation in meteorology and oceanography, *Adv. Geophys.*, 33, 141–266, 1991.
- Giering, R.**, *Tangent linear and Adjoint Model Compiler, Users manual TAMC Version 5.2*, September 1999.
- Giering, R. and T. Kaminski**, Recipes for adjoint code construction, *ACM Trans. Mathematical Software*, 24, (4), 437–474, 1998.

- Grell, G. A., J. Dudhia and D. R. Stauffer**, A description of the fifth-generation Penn State/NCAR mesoscale model MM5, technical report, National Center for Atmospheric Research, 1994.
- Hass, H.**, *Description of the EURAD Chemistry-Transport-Model version2 (CTM2)*, vol 83, Mitteilungen aus dem Institut für Geophysik und Meteorologie der Universität zu Köln, 1991.
- Hass, H., H. J. Jakobs and M. Memmesheimer**, Analysis of a regional model (EURAD) near surface gas concentration predictions using observations from networks, *Met. Atmos. Phys.*, 57, 173–200, 1995.
- Heemink, A., M. Verlaan and A. Segers**, Variance reduced ensemble Kalman filtering, *Month. Weath. Rev.*, 129, 1718–1728, 2001.
- Hesstvedt, E., Ø. Hov and I. S. A. Isaksen**, Quasi-steady state approximation in air pollution modeling: Comparison of two numerical schemes for oxidant prediction, *Int. J. Chem. Kinet.*, 10, 971–994, 1978.
- Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek and B. Hansen**, Atmospheric data assimilation with the ensemble Kalman filter: Results with real observations, *Month. Weath. Rev.*, 133, (3), 604–620, 2005.
- Ide, K., P. Courtier, M. Ghil and A. C. Lorenc**, Unified notation for data assimilation: operational, sequential and variational, *J. Meteor. Soc. Japan*, 75, (1B), 181–189, 1997.
- Jakobs, H. J., H. Feldmann, H. Hass and M. Memmesheimer**, The use of nested models for air pollution studies: An application of the EURAD model to a SANA episode, *J. Appl. Meteor.*, 34, (6), 1301–1319, 1995.
- Jazwinski, A. H.**, *Stochastic Processes and Filtering Theory*, Academic Press, 1970.
- Kalman, R. E.**, A new approach to linear filtering and prediction problems, *Trans. ASME, J. Basic Eng.*, 82D, 35–45, 1960.
- Kalman, R. E. and R. S. Bucy**, New results in linear filtering and prediction theory, *Trans. ASME, J. Basic Eng.*, 83D, 95–108, 1962.

- Kaminski, T., M. Heimann and R. Giering**, A coarse grid three-dimensional global inverse model of the atmospheric transport — 1. adjoint model and jacobian matrix, *J. Geophys. Res.*, *104*, (D15), 18535–18553, 1999.
- Liu, D. C. and J. Nocedal**, On the limited memory BFGS method for large scale optimization, *Math. Programming*, *45*, 503–528, 1989.
- Loon, M. V. and A. Heemink**, Kalman filtering for nonlinear atmospheric chemistry models: first experiences, technical report, CWI, MAS-R9711, Amsterdam, 1997.
- Lorenc, A. C.**, Analysis methods for numerical weather prediction, *Q. J. R. Meteorol. Soc.*, *112*, 1177–1194, 1986.
- Madronich, S.**, Photodissociation in the atmosphere. 1. actinic flux and the effects of ground reflections and clouds, *J. Geophys. Res.*, *92*, 9740–9752, 1987.
- McRae, G. J., W. R. Goodin and J. H. Seinfeld**, Numerical solution of the atmospheric diffusion equation for chemically reacting flows, *J. Comp. Phys.*, *45*, 1–42, 1982.
- Memmesheimer, M., J. Tippke, A. Ebel, H. Hass, H. J. Jakobs and M. Laube**, On the use of EMEP emission inventories for European scale air pollution modelling with the EURAD model, in *Proceedings of the EMEP workshop on photooxidant modelling for long range transport in relation to abatement strategies*, edited by J. Pankrath, 307–324, UBA, Berlin, 1991.
- Memmesheimer, M., H. Hass, J. Tippke and A. Ebel**, Modeling of episodic emission data for Europe with the EURAD Emission Model EEM, in *Proceedings of the International Speciality Conference "Regional Photochemical Measurement and modeling studies"*, vol 2, San Diego, CA, USA, AWMA, 1995.
- Memmesheimer, M., H. J. Jacobs, G. Piekorz, A. Ebel and H. Hass**, *Großräumige Ozon-Modellrechnungen mit verschiedenen Emissionsminderungsszenarien mit Hilfe des Modellsystems EURAD*, Abschlussbericht zum Teilvorhaben A5: Großräumige Ozon-Ausbreitungsrechnungen im Rahmen des FuE Aktionsprogramms und Massnahmenplanes, 1996.

- Memmesheimer, M., M. Roemer and A. Ebel**, Budget calculations for ozone and its precursors: Seasonal and episodic features based on model simulations, *J. Atmos. Chem.*, *28*, (1-3), 283–317, 1997.
- Memmesheimer, M., H. J. Jakobs, H. Feldmann, G. Piekorz, C. Kessler, E. Friese and A. Ebel**, Computergestützte Langzeitsimulationen zur Bewertung von Strategien zur Luftreinhaltung, Abschlussbericht zum Vorhaben COSIMA im Auftrag des Landesumweltamtes Nordrhein-Westfalen, 2000.
- Memmesheimer, M., E. Friese, A. Ebel, H. J. Jakobs, H. Feldmann, C. Kessler and G. Piekorz**, Long-term simulations of particulate matter in Europe on different scales using sequential nesting of a regional model, *Int. J. Environm. and Pollution*, *22*, (1-2), 108–132, 2004.
- Memmesheimer, M., E. Friese, H. J. Jakobs, C. Kessler, H. Feldmann, G. Piekorz and A. Ebel**, Ausbreitungsrechnungen zur zukünftigen Entwicklung der Luftqualität in Nordrhein-Westfalen: Bewertung und Maßnahmenplanung zur Luftreinhaltung, Abschlussbericht zum LuE-Vorhaben AZUR im Auftrag des Landesumweltamtes Nordrhein-Westfalen, November 2004b.
- Mohnen, V.**, Data quality assessment - an overview for the TRACT 16./17. september 1992 field intensive, Report to the coordinator of TFS-LT1, IFU Garmisch-Partenkirchen, 1999.
- Müller, M. D., A. K. Kaifel, M. Weber, S. Tellmann, J. P. Burrows and D. Loyola**, Ozone profile retrieval from Global Ozone Monitoring Experiment (GOME) data using a neural network approach (Neural Network Ozone Retrieval System (NNORSY)), *J. Geophys. Res.*, *108*, (D16), doi:10.1029/2002JD002784, 2003.
- Nocedal, J.**, Updating quasi-newton matrices with limited storage, *Math. Comput.*, *35*, (151), 773–782, 1980.
- Parrish, D. and J. Derber**, The national meteorological centre's spectral statistical-interpolation analysis system, *Month. Weath. Rev.*, *120*, 1747–1763, 1992.
- Petry, H., A. Ebel, V. Franzkowiak, J. Hendricks, E. Lippert and M. Möllhoff**, Impact of aircraft exhaust on the atmosphere: Box model studies and 3-d mesoscale numerical case studies of seasonal differences, in *Proceedings of Impact of aircraft emissions upon the atmosphere*, 241–246, Office National d'Etudes et de Recherche Aérospatiale, 1997.

- Pleim, J. E., J. S. Chang and K. Zhang**, A nested grid mesoscale atmospheric chemistry model, *J. Geophys. Res.*, *96*, 3065–3084, 1991.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery**, *Numerical Recipes*, Cambridge Univ. Press, 1992.
- Quélo, D., V. Mallet and B. Sportisse**, Inverse modeling of nox emissions at regional scale over northern france. preliminary investigation of the second-order sensitivity, *J. Geophys. Res.*, *110*, (D24), D24310, 2005.
- Rabier, F.**, Overview of global data assimilation developments in numerical weather-prediction centres, *Q. J. R. Meteorol. Soc.*, *131*, (613), 3215–3233, 2005.
- Rabier, F., H. Jarvinen, E. Klinker, J. F. Mahfouf and A. Simmons**, The ecmwf operational implementation of four-dimensional variational assimilation. i: Experimental results with simplified physics, *Q. J. R. Meteorol. Soc.*, *126*, (564), 1143–1170, 2000.
- Rodgers, C. D.**, *Inverse Methods for Atmospheric Sounding*, World Scientific, 2000.
- Sander, R., A. Kerkweg, P. Jöckel and J. Lelieveld**, Technical note: The new comprehensive atmospheric chemistry module MECCA, *Atmos. Chem. Phys.*, *5*, 445–450, 2005.
- Sandu, A. and R. Sander**, Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1, *Atmos. Chem. Phys.*, *6*, 187–195, 2006.
- Sandu, A., F. A. Potra, G. R. Carmichael and V. Damian**, Efficient implementation of fully implicit methods for atmospheric chemical kinetics, *J. Comp. Phys.*, *129*, 101–110, 1996.
- Schell, B.**, Zu den Rand- und Anfangswertbedingungen des EURAD Chemie-Transport-Modells Version2 (CTM2): Erstellung eines Szenarios auf der Basis von Messdaten, Master Thesis, Institut für Geophysik und Meteorologie der Universität zu Köln, 1996.
- Schmidt, H.**, Variationelle Datenassimilation mit Hilfe eines adjungierten Modells am Beispiel des RADM2-Gasphasenmechanismus, Master Thesis, Institut für Geophysik und Meteorologie der Universität zu Köln, 1996.

- Schmidt, H.**, *Vierdimensionale Datenassimilation nach der Variationsmethode für ein mesoskaliges Chemietransportmodell*, Ph.D. Thesis, Institut für Geophysik und Meteorologie der Universität zu Köln, 1999.
- Seinfeld, J. H. and S. N. Pandis**, *Atmospheric chemistry and physics*, Wiley-Interscience, 1998.
- Stockwell, W. R., P. Middleton and J. S. Chang**, The second generation regional acid deposition model chemical mechanism for regional air quality modeling, *J. Geophys. Res.*, *95*, (D10), 16343–16367, 1990.
- Stockwell, W. R., F. Kirchner, M. Kuhn and S. Seefeld**, A new mechanism for regional atmospheric chemistry modeling, *J. Geophys. Res.*, *102*, (D12), 25847–25879, 1997.
- Talagrand, O.**, The use of adjoint equations in numerical modelling of the atmospheric circulation, in *Proceedings of Workshop on Automatic differentiation of algorithms: Theory, implementation and application*, edited by A. Griewand and G. G. Corliss, 1991.
- Talagrand, O.**, Assimilation of observations, an introduction, *J. Meteor. Soc. Japan*, *75*, (1B), 191–209, 1997.
- Tilmes, S.**, *Verfahren zur Analyse von Messungen atmosphärischer Spurengase mit dem Ziel der Assimilation in Chemie-Transportmodellen*, Ph.D. Thesis, Berichte des Deutschen Wetterdienstes, DWD Offenbach, 1999.
- Volz-Thomas, A., H. Geiss, A. Hofzumahaus and K. H. Becker**, Introduction to special section: Photochemistry experiment in berlioz, *J. Geophys. Res.*, *108*, (D4), 8252, 2003.
- Wang, Z., I. M. Navon, F. X. L. Dimet and X. Zou**, The second order adjoint analysis, theory and applications, *Met. Atmos. Phys.*, *50*, 3–20, 1992.
- Weaver, A. and P. Courtier**, Correlation modelling on the sphere using a generalized diffusion equation, *Q. J. R. Meteorol. Soc.*, *127*, 1815–1846, 2001.

Acknowledgements

This work has been supported directly and indirectly by many institutions and single persons. Without their contributions — no matter to what extent — it would not be the same.

I am very grateful to Prof. Dr. Adolf Ebel for his support and commitment. Thanks to his dedication for his work and RIU, I had the great opportunity to join his group and to learn lots of things.

I kindly thank Prof. Dr. Joachim Saur for joining the doctoral committee, acting as a referee for this thesis.

I gratefully acknowledge the assistance, generosity, and advice I received from PD Dr. Hendrik Elbern. Special thanks for never rejecting a discussion, for being always available and long-sighted, for his suggestions, and his great support, also in difficult times. By the way, many thanks for his flowery language, even when discussing most difficult topics. Finally, I acknowledge the essential meliorations of this manuscript by his proof-reading.

I extend many thanks to Dr. Michael Memmesheimer for the countless discussions and his profound explanations about emissions and chemistry and their modelling aspects.

The *data* in data assimilation has been provided by various campaign consortia. I therefore owe a special note of gratitude to the persons and groups involved in BERLIOZ and TFS, MOZAIC, CONTRACE, SPURT and VERTIKO for the unfettered access to their measurement data. In addition, the routinely gathered data has been obtained by various European institutions, which I hereby acknowledge. The satellite column retrievals used in this work have been supplied by Dr. Henk Eskes, while the NNORSY ozone

profiles have been processed and provided by Anton Kaifel. Meteorological analyses have been used by courtesy of NCEP and ECMWF. The chemistry library routines have been provided and validated by Elmar Friese, who adapted the KPP to the needs of the EURAD system. Thanks to Dr. Hermann Jakobs, who bailed me out with the MM5.

Due to the immense computational demands, the simulations in this work have been carried out at the IBM p690 eServer _{JüMP} at the Forschungszentrum Jülich. I thank for the support by the Zentrum für Angewandte Mathematik and the John-von-Neumann Institut for Computing. I am very grateful for the in-house service accomplished by Georg Piekorz, who reliably found the trouble spots and solved all problems.

My friendly thanks are devoted to Jörg Schwinger, the great companion throughout these years. Without him, I definitely would have never enjoyed the work so much. Thank you, Jörg, "you'll never walk alone"... Thanks to Hendrik Feldmann for the very enjoyable time in our office. Moreover, I would like to thank the RIU staff for the exceptional social skills, making it quite impossible not to feel good being part of you. And, very important, I like to express my gratitude to the Mensacombo, thanks for the great time! "FLAAAIIIIISCH"!

Many thanks to Nils Alhäuser for persistently working through this thesis and his helpful remarks on the perceivability. Moreover, he always kept me aware that even the Borg never succeeded in assimilating (Mr.) data.

Finally, I sincerely thank Astrid Leive, my mother Hannelore and my sisters Petra and Angela for their love, and their endless support and patience with the "Schwachkopf".

This work has been funded under FKZ 07ATF03 in the framework of the atmospheric research programme AFO2000 by the German Federal Ministry of Education and Research.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. A. Ebel betreut worden.

Teilpublikationen

Elbern, H., A. Strunk, Tracer analyses by mesoscale experiments with 4D-var data assimilation, in *Results of the German atmospheric research programme AFO2000 final publication*, Editor GSF, 2005.

