Quantitative estimation of unexpected emissions in the atmosphere by stochastic inversion techniques

Inaugural – Dissertation
zur
Erlangung des Doktorgrades
der Mathematisch–Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Philipp Franke

aus Düsseldorf

Köln, 2018

Berichterstatter: PD Dr. H. Elbern
Prof. Dr. Y. Shao

Tag der letzten mündlichen Prüfung: 20.12.2017

Abstract

Unexpected emission events like volcanic eruptions pose a threat to humans, climate, economy, and aviation. For decision makers, the estimation of volcanic ash in the atmosphere and its uncertainty is essential. These information enables them to take safety precautions as accurate as possible. Currently, there exists no method for estimating the emissions of volcanic ash and its uncertainty for longer lasting eruptions, in which the eruption strength varies temporally in dependence on the emission height. Therefore, an analysis system, which is able to estimate the emission strength of a volcanic eruption in a highly temporal and vertical resolution, is established in this study. This analysis system is integrated into the atmospheric chemical part of the Ensemble for Stochastic Integration of Atmospheric Simulations (ESIAS-chem) that comprises a particle smoother in combination with a discrete-grid ensemble extension of the Nelder-Mead minimization method. The extended Nelder-Mead method makes the temporal and vertical resolution of the emission strength possible. The particle smoother reduces the variance in the analysis ensemble in order to provide an accurate and reliable estimate of the emission profile. The system validation addresses the special challenge of ash cloud height analyses in case of observations restricted to bulk column mass loading information. This reflects the typical case of geostationary satellite data like thus obtained by the Spinning Enhanced Visible and InfraRed Imager (SEVIRI), while height resolving measurements like lidar data are only sparsely available. The EURopean Air pollution Dispersion - Inverse Model (EURAD-IM), which was generalized to an ensemble system, is integrated into ESIAS-chem. The performance of ESIAS-chem is tested by identical twin experiments. These tests show that for both, strong and weak wind conditions, the model is able to retrieve the observed column mass loadings of volcanic ash. The potential of the method becomes especially evident under strong wind conditions at the volcano, for which the temporal and vertical varying volcanic emissions are analyzed up to an error of only 10 % by assimilating column mass loadings. For weak wind conditions, the analysis accuracy of the emission profile is limited, because the volcanic ash emitted at different times and heights is not sufficiently separated. Increasing the assimilation window length proves to be the only possibility to improve the analysis for these weak wind conditions. The dependence of the analysis on the wind conditions is also confirmed by a statistical analysis in the context of identical twin experiments. The analysis is biased because of weak winds at the volcano as well as an insufficiently short assimilation window. The investigation of this statistical analysis shows that an error correction of volcanic ash concentrations using the observed column mass loadings can achieve highly improved estimates of the volcanic ash concentrations. Thereafter, the mean error of the column mass loadings vanishes. The mean error in estimating the height of the maximum volcanic ash concentration is lower than the vertical resolution of the model. This indicates a successful estimation of the temporal and vertical distribution of volcanic ash by assimilating column mass loadings. The efficiency in the

calculation and the accuracy of the analysis of ESIAS-chem is demonstrated by a comparison with other methods for analyzing volcanic ash concentrations and its uncertainty. Especially for highly variable emissions, ESIAS-chem has more potential to estimate the emission profile and its uncertainty than currently existing methods. Besides volcanic eruptions, ESIAS-chem is applicable to various other emission scenarios. The special importance of risk assessment for unexpected emission scenarios makes ESIAS-chem most useful for these cases. The combination with a meteorological ensemble (e.g. ESIAS-met) allows for the consideration of the uncertainty of the meteorological variables.

Kurzzusammenfassung

Unerwartete Emissionen wie zum Beispiel durch Vulkaneruptionen, Waldbrände und Reaktorunfälle stellen eine Gefahr für den Menschen, das Klima, die Wirtschaft und den Flugverkehr dar. Eine Abschätzung der zu erwartenden Vulkanaschekonzentration sowie deren Unsicherheit ist für Entscheidungsträger von großer Bedeutung um mögliche Schutzvorkehrungen unter Berücksichtigung aller zur Verfügung stehenden Informationen zu ergreifen. Jedoch gibt es bisher keine Methode, die die Emissionen eines Vulkanausbruchs und deren Unsicherheit für einen längeren Ausbruch mit zeitlich variablen Emissionsstärken in Abhängigkeit der Emissionshöhe abschätzen kann. Daher wird in dieser Arbeit ein flexibles Analysesystem entwickelt, welches sowohl die Emissionsstärke in zeitlicher und vertikaler Auflösung als auch deren Unsicherheit ermittelt. Der hierfür neu entwickelte atmosphären-chemische Modellteil des Ensemble for Stochastic Integration of Atmospheric Simulations (ESIAS-chem) besteht aus einer Kombination aus einem "particle smoother" und einer ensemble-basierten Erweiterung des Nelder-Mead-Simplex-Minimierungsalgorithmus auf einem diskreten Gitter im Lösungsraum. Mit Hilfe des erweiterten Nelder-Mead-Algorithmus können die Quellstärken zeitlich und vertikal aufgelöst werden. Der particle smoother dient der zuverlässigen Abschätzung der Analyseunsicherheit. Das System ermöglicht eine Abschätzung der Höhe der Vulkanasche unter Benutzung der Beobachtungen von vertikal-integrierten Säulenmassen. Dies spiegelt die Assimilation von geostationären Satellitendaten, wie sie zum Beispiel durch den Spinning Enhanced Visible and InfraRed Imager (SEVIRI) bereitgestellt werden, wieder. Diese Daten sind in einer hohen zeitlichen Auflösung verfügbar, wohingegen höhenauflösende Messungen wie die von Lidar-Geräten nur weit verstreut verfügbar sind. In ESIAS-chem ist das EURopean Air pollution Dispersion - Inverse Model (EURAD-IM) integriert, welches in ein Ensemblesystem erweitert wurde. ESIAS-chem ist erfolgreich im Rahmen eines "identical twin experiments" getestet worden. Sowohl für starke als auch für schwache Winde konnte ESIAS-chem die vertikal-integrierte Säulenmasse der Beobachtungen reproduzieren. Bei starken Winden wird das Potential des Modells voll ausgeschöpft und die Abschätzung der Emissionen ist bis auf eine leichte vertikale und zeitliche Glättung der Quellstärke bis auf 10 % genau. Die Grenzen der Analysefähigkeit von ESIAS-chem ist bei zu schwachen Winden gegeben. In diesen Fällen ist die Separation der Vulkanasche, die zu unterschiedlichen Zeiten und in unterschiedlichen Höhen emittiert wurde, zu gering, sodass von den vertikal-integrierten Säulenmassen der Vulkanasche nicht exakt auf die zeitliche und vertikale Verteilung der Emissionen zurückgeschlossen werden kann. Eine Erweiterung des Assimilationsfensters zeigt hier eine Verbesserung der Analyse. Die Anbhängigkeit der Analyse von den Windbedingungen wurde durch eine Fehlerabschätzung mittels statistischer Kenngrößen bestätigt. Die Analyse zeigt einen systematischen Fehler, verursacht durch schwache Winde am Vulkan und ein zu kurzes Assimilationsfenster. Durch

eine Fehlerkorrektur der Vulkanaschekonzentration unter Berücksichtigung der assimilierten vertikal-integrierten Beobachtungen kann die Genauigkeit der Vorhersage der Vulkanaschekonzentration deutlich erhöht werden. Gemittelt über den Analysezeitraum zeigt nach der Fehlerkorrektur sowohl die abgeschätzte vertikal-integrierte Säulenmasse als auch die Vulkanaschekonzentration keinen Fehler bezüglich der Beobachtungen. Im Mittel liegt der Fehler bei der Abschätzung der Höhe der maximalen Vulkanaschekonzentration unterhalb der vertikalen Auflösung des Modells. Dies bestätigt die hohe Genauigkeit bei der Abschätzung der vertikalen und zeitlichen Verteilung der Vulkanaschekonzentration in der Atmosphäre. Ein Vergleich von ESIAS-chem mit weiteren Methoden zur Analyse von Vulkanaschekonzentrationen und deren Unsicherheit zeigt die Effizienz der Berechnung und Genauigkeit der Analyse durch ESIAS-chem. Gerade bei stark variablen Emissionen hat ESIAS-chem deutlich mehr Potenzial, das Emissionsprofile sowie dessen Unsicherheit abzuschätzen als zur Zeit existierende Methoden. Neben Vulkanausbrüchen ist das neu entwickelte ESIAS-chem System auf viele weitere Emissionsszenarien anwendbar. Besonders für unerwartete Emissionsereignisse aufgrund ihrer besonderen Bedeutung für die kurzfristige Risikobewertung ist die Benutzung von ESIAS-chem sinnvoll. Des Weiteren können Unsicherheiten in den meteorologischen Variablen durch Integration eines meteorologischen Ensemble (zum Beispiel von ESIAS-met) berücksichtigt werden.

Contents

Abstract **1**

Kurzzusammenfassung **3**

List of figures **7**

List of tables **9**

List of acronyms **11**

1 Introduction **13**

2 Data assimilation via particle filtering **19**
2.1 Bootstrap particle filter 19
2.2 Methods to overcome filter degeneracy 22

3 Minimization algorithm **27**
3.1 Ensemble of emission scenarios 27
3.2 Original Nelder-Mead method 29
3.3 Nelder-Mead method with adaptive parameters 34
3.4 Parallel Nelder-Mead method 41
3.5 Discrete Ensemble Nelder-Mead method 47

4 Model description **55**
4.1 EURAD-IM 55
4.2 Ensemble EURAD-IM 57
4.3 JUQUEEN 58

5 Ensemble for Stochastic Integration of Atmospheric Simulations (ESIAS) **61**
5.1 ESIAS-met 61
5.2 ESIAS-chem 63

6 Identical twin experiments **67**
6.1 Selection of test cases 67
6.2 Dependence on the assimilation window length 73

6.3 Analysis of the potential and limitations of ESIAS-chem 79
6.4 Statistical analysis of volcanic ash estimation 84

7 Comparison with literature **93**

8 Conclusion and outlook **99**

A A combined 4D-var / particle smoother for volcanic ash emission estimation **103**

B Localized particle filter **105**

Bibliography **107**

List of Figures

2.1 Schematic of the bootstrap particle filter 21

3.1 Conversion of the emission profile to a vector of emission factors for the minimization of an ensemble of emissions 28
3.2 Visualization of the Nelder-Mead minimization method 30
3.3 Comparison of ONM and ANM methods for cost function $f(\mathbf{x})$ 36
3.4 Percentage cost reduction of the ONM and ANM method for all experiments on cost function $g(\mathbf{x})$ 38
3.5 Comparison of the performance of the ONM and ANM methods. . . . 40
3.6 Centroid calculation in the PNM method 42
3.7 Comparison of the mean and standard deviation of the final cost reduction for S-PNM and A-PNM 45
3.8 Comparison of the relative frequency of the minimization steps for the S-PNM and A-PNM method for cost function $g(\mathbf{x})$ 46
3.9 Comparison of the relative frequency of the minimization steps for the S-PNM and A-PNM method for cost function $h(\mathbf{x})$ 46
3.10 Flowchart of the DENM method 49
3.11 Illustration of the DENM method using the two-dimensional Rosenbrock standard test function 51
3.12 Performance analysis of the DENM method for an ensemble of uniform initial simplices 53

4.1 Schematic of the parallelization of the ensEURAD-IM 58
4.2 Scaling performance of the ensEURAD-IM 60

5.1 Schematic of the ESIAS system 62

6.1 Hovmoeller plot of the nature run emission profile on April 15 and April 29, 2010 68
6.2 Meteorological conditions on April 15, 2010 69
6.3 Meteorological conditions on April 29, 2010 69
6.4 Nature run emission profile for the statistical analysis of ESIAS-chem 70
6.5 Frequency of the hypothetical lidar measurements taken at 33 Earlinet lidar stations 71
6.6 Domain of the ensEURAD-IM for the identical twin experiments . . . 72

6.7 Pattern correlation coefficient for different assimilation window lengths on April 15 and April 29, 2010 73
6.8 Mass mismatch for different assimilation window lengths on April 15 and April 29, 2010 75
6.9 Normalized distance of the centers of mass for different assimilation window lengths on April 15 and April 29, 2010 76
6.10 Relative mean absolute error of column mass loadings for different assimilation window lengths on April 15 and April 29, 2010 77
6.11 Profiles of the extinction coefficient at four hypothetical lidar stations on April 15-16, 2010 78
6.12 Relative mean absolute error of volcanic ash concentrations for different assimilation window lengths on April 15 and April 29, 2010 . . 78
6.13 Comparison of the column mass loadings of the nature run and ensemble mean on April 15, 2010, 00 UTC + 30 hours 80
6.14 Comparison of the vertical cross section of the nature run and ensemble mean volcanic ash concentrations on April 15, 2010, 00 UTC + 30 hours 81
6.15 Comparison of the emission profiles of the nature run and ensemble mean on April 15, 2010 82
6.16 Comparison of the emission profiles of the nature run and ensemble mean on April 29, 2010 83
6.17 Comparison of the absolute error of the ensemble mean and the ensemble spread of the column mass loadings at the 33 lidar stations . . 85
6.18 Daily relative mean error on April 14-18 and May 04-18, 2010 85
6.19 Wind profiles above the Eyjafjallajökull volcano on May 14-17, 2010 . 87
6.20 Wind profiles above the Eyjafjallajökull volcano on May 10-11, 2010 . 87
6.21 P-P plot for the statistical identical twin experiment 91

List of Tables

3.1 Codes for the initial simplices for the performance analysis of the Nelder-Mead methods 35
3.2 Percentage cost reduction of the ONM and ANM methods for cost function $g(\mathbf{x})$ 39

4.1 Initial median particle diameter and standard deviation of the modes in MADE 56
4.2 Speedup of the ensEURAD-IM normalized to 1024 cores 59
4.3 Comparison of I/O and computing time for EURAD-IM 60

6.1 Scale analysis for volcanic ash emission estimation 72
6.2 Statistics of the performance analysis of ESIAS-chem 89

7.1 Comparison of volcanic ash optimization methods 95

List of acronyms

3D-var	Three-Dimensional VARiational data assimilation
4D-var	Four-Dimensional VARiational data assimilation
ANM	Nelder-Mead method using adaptive parameters
A-PNM	parallel Nelder-Mead method using adaptive model parameters
BLUE	Best Linear Unbiased Estimate
BS	Brier Score
BSS	Brier Skill Score
CC	Correlation Coefficient
CDF	Cumulative Distribution Function
CTM	Chemistry-Transport-Model
CUT	Conjugate Unscent Transform
DENM	Discrete Ensemble Nelder-Mead
ECMWF	European Centre for Medium-range Weather Forecasts
ensEURAD-IM	ensemble EURAD-IM
EPS	Ensemble Prediction System
ESIAS	Ensemble for Stochastic Integration of Atmospheric Simulations
ESIAS-chem	chemical part of ESIAS
ESIAS-met	meteorological part of ESIAS
EURAD-IM	EURopean Air pollution Dispersion - Inverse Model
GEFS	Global Ensemble Forecast System
I/O	Input and Output

MACC	Modelling Atmospheric Composition and Climate
MADE	Modal Aerosol Dynamics module for Europe
ME	Mean Error
MER	Mass Emission Rate
MPI	Message Passing Interface
NCEP	National Centers for Environmental Prediction
ONM	original Nelder-Mead method
pcc	Pattern Correlation Coefficient
pCQ	polynomial Chaos Quadrature
PCQWE	polynomial Chaos Quadrature Weighted Estimate
PDF	Probability Density Function
PNM	parallel Nelder-Mead method
P-P plot	Probability-Probability plot
RMAE	Relative Mean Absolute Error
RME	Relative Mean Error
RMSE	Root Mean Square Error
ROC-A	Area under the Relative Operating Characteristics curve
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
SKEBS	Stochastic Kinetic Energy Backscatter Scheme
S-PNM	parallel Nelder-Mead method using the standard set of model parameters
SPPT	Stochastic Perturbed Parameterization Tendencies
WRF	Weather Research and Forecasting model
WRF-EDAS	Ensemble Data Assimilation System of the Weather Research and Forecasting model

Chapter 1

Introduction

Unexpected emission events can pose an enormous security threat to humans and environment. This type of incidents can be divided into natural emissions like volcanic eruptions and wild fires, and anthropogenic emissions owing to accidental releases of toxic gases, as in case of reactor failure or a gas pipeline explosion. From a simulation viewpoint, volcanic eruptions are challenging in many respects. Typically, they come as a sequence of emissions with highly varying ejection mass and height. This study, therefore, adopts volcanic eruptions as reference case for the analysis of accidental releases. It was shown, that in historic times 20-25 volcanic eruptions occurred on average in Iceland each century (*Thordarson and Larsen* [2007]). Taking into account that frequent wind systems are directed to transport volcanic ash from Iceland to continental Europe, eruptions of Icelandic volcanoes have a direct impact on life in that greater domain.

Volcanic eruptions may cause issues to humans health (*Baxter* [1999]), climate (*Robock and Oppenheimer* [2003]), environment (*Lecointre et al.* [2004]), and economy (*Guffanti et al.* [2009]). Additionally, volcanic ash particles may cause turbines of aircrafts to break down (*Casadevall* [1994]). This is the reason why it is necessary to consider closures of the air space after volcanic eruptions. Unfortunately, as this action has an exceptional high impact on affected economies, it needs to be traded off against risks. For this, volcanic ash concentrations and its uncertainty has to be known as accurate as possible.

As a recent example, the closure of the European air space as a consequence of the 2010 eruption of the Icelandic volcano Eyjafjallajökull forced more than 100,000 airplanes to stay on the ground worldwide with more than 4 billion passengers stranded. A direct economic damage of more than 1.3 billion Euros (*Oxford-Economics* [2010]) resulted. During the eruption, the European commission released thresholds of volcanic ash concentrations for which airplanes are allowed to fly only under certain restrictions (cf. *Kristiansen et al.* [2012] and references therein). However, at that time the estimation of volcanic ash concentrations could hardly be obtained by numerical predictions, which instigated efforts to develop methods for quantitative volcanic ash concentration forecasts along with their uncertainty. Hence, the prediction of skill or estimated uncertainty of a volcanic ash forecast is of fundamental

importance for cost - loss models, which may be used within the decision process.

The uncertainty by estimating volcanic ash in the atmosphere is due to manifold reasons. Firstly, quantitative observations of volcanic ash column mass loadings [g/m^2] by, for example, the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) on board the Meteosat Second Generation satellite, which are widely used for a quantitative estimation of volcanic ash, have large uncertainties of approximately 40 % (see *Wen and Rose* [1994]). Secondly, the location of the volcanic ash cloud depends on the atmospheric stratification at the volcano and the underlying wind field. Thirdly, the modeled volcanic ash concentration depends mainly on the emission parameters (e.g. exit velocity of the magma, gas content, and water availability) used to calculate the volcanic ash plume which are unknown and come with large uncertainties.

Models calculating the volcanic ash emissions can be subdivided into heuristic models and physical models. Heuristic models estimate the Mass Emission Rate (MER) of a volcanic eruption from measurements of the plume height through a statistical model based on historical eruptions (e.g. *Sparks et al.* [1997]; *Mastin et al.* [2009]). Due to the uncertainties of volcanic eruptions, which are mainly influenced by the crater geometry, gas content, temperature of the magma, exit velocity, and water availability, historical data have large variation. Hence, heuristic models have large uncertainties (*Mastin et al.* [2009]). Furthermore, the fraction of fine ash of MER, which is available for long range transport, is unknown and is an additional source of uncertainty.

Physical models compute the amount of emitted ash taking plume physics into account. Different models exist that calculate the volcanic ash plume rise with respect to meteorological conditions as wind speed, temperature, and humidity profiles of the atmosphere (see for example *Woodhouse et al.* [2013]; *Folch et al.* [2016]). Alternatively, *Suzuki* [1983] developed a theoretical model for plume rise and ash dispersion. Although more physical, these models need input parameters such as gas content, exit velocity and temperature of the emitted magma, which contribute to the emission strength and the type of a volcanic eruption. However, these are unknown and hardly measurable. Therefore, these parameters are insufficiently well known. In addition, these models assume that the emissions of ash occur mainly at the plume's top, not allowing for a secondary maximum below. As recent analyses show, emissions from a volcanic plume may occur with multiple maxima (cf. e.g. *Kristiansen et al.* [2015]), which renders physical models erroneous in these cases. These parameterizations are used in forecast models to improve the a priori knowledge of volcanic ash emissions.

As both, the observations and the forecast models, come with different sources of errors, the best estimate of the true atmospheric state is achieved by data assimilation methods, which combine all available information in order to improve the forecast of atmospheric dispersion. Data assimilation methods are basically divided into variational (e.g. Three-Dimensional VARiational data assimilation (3D-var), Four-Dimensional VARiational data assimilation (4D-var)) and sequential (e.g. Kal-

man filter) methods. In variational methods, the model state is varied by optimizing a cost function. In sequential methods, the model equations are adapted in order to account for all, the model state, the observations, and the errors, which come along with both. Both methodologies rely on the Best Linear Unbiased Estimate (BLUE) assumption. This means that the model state and the observations are unbiased and that the error statistics of both are Gaussian. Additionally, the models need to be linear or at least well represented with the tangent linear model.

As the computational power during the last decades has increased, the complexity of data assimilation algorithms increased likewise. Thus, in the last 20 years ensemble prediction systems entered the field of atmospheric sciences (e.g. *Molteni et al.* [1996]; *Toth and Kalnay* [1997]; *Marécal et al.* [2015]). Ensemble data assimilation systems are used to provide both, a reliable forecast/analysis and its uncertainty. Examples of ensemble analysis systems are the ensemble Kalman filter (*Evensen* [1994]) and its variants (e.g. ensemble transform Kalman filter proposed by *Bishop et al.* [2001] and the ensemble adjustment Kalman filter of *Anderson* [2001]), the ensemble of data assimilation operated at the European Centre for Medium-range Weather Forecasts (ECMWF) (*Bonavita et al.* [2012]), and the Ensemble Data Assimilation System of the Weather Research and Forecasting (WRF-EDAS) model (*Zupanski et al.* [2011]). Additionally, the development of an ensemble 4D-var system is a current research topic (e.g. *Zhang and Zhang* [2012]; *Clayton et al.* [2013]; *Yang et al.* [2015]) with the potential to overcome the need for an adjoint model (see *Liu et al.* [2008]).

In the aftermath of the 2010 Eyjafjallajökull eruption, several data assimilation methods were adapted in order to quantify the emission strength and the ensuing volcanic ash concentrations in the atmosphere. Some of these methods also applied an ensemble data assimilation system to provide estimations of the uncertainty of volcanic ash in the atmosphere. However, these methods make assumptions of the volcanic ash plume that need not to be applicable for individual volcanic eruptions. Most of these methods for analyzing volcanic ash emissions presented in the literature use satellite observations of volcanic ash column mass loading data to constrain volcanic ash emissions (e.g. *Lu et al.* [2016]; *Madankan et al.* [2014]; *Schmehl et al.* [2012]; *Kristiansen et al.* [2015]; *Wilkins et al.* [2016*a*] amongst others). However, the assimilation of column mass loadings lead to a highly ill-posed problem. This is due to the ambiguity of volcanic ash that is emitted in different heights and at different times resulting in similar volcanic ash patterns in the atmosphere at the analysis time (*Devenish et al.* [2012]; *Lu et al.* [2016]). The vertical wind shear may separate volcanic ash emitted at different heights/times, especially if the residence time of volcanic ash in the atmosphere is long enough. This would also reduce the correlation of the resulting column mass loadings. In contrast, an increased residence time of volcanic ash in the atmosphere leads to a vertical mixing and sedimentation of volcanic ash into lower heights, which in the end makes the distinction between volcanic ash emitted at different heights/times increasingly difficult and finally impossible. In an ideally observable atmosphere in which the volcanic ash emitted at

different heights/times is completely separated through the vertical wind shear, observations of column mass loadings could be used directly to constrain the volcanic ash emissions.

To quantify the amount of volcanic ash emitted by an eruption, *Stohl et al.* [2011] and later *Kristiansen et al.* [2012] and *Kristiansen et al.* [2015] used an ensemble of emission scenarios in the inversion technique of *Eckhardt et al.* [2008]. In this approach, each emission scenario consists of emissions in one model layer at one time instance (e.g. three hours). They assumed that the vertical wind shear is strong enough such that column mass loadings of volcanic ash emitted by different emission scenarios is distinguishable. For the inversion, a system of linear equations was solved taking the error of observations and the a priori emission profiles into account. In order to make the inversion robust, they applied a regularization that also vertically smooths the volcanic ash emission profile. However, multiple maxima of volcanic ash within the eruption column may be possible as become evident by images of volcanic ash plumes. Additionally, *Kristiansen et al.* [2012] suggested an ensemble of emission estimation in order to quantify the uncertainty of volcanic ash emissions as was indicated by sensitivity studies they performed.

Another approach to quantify volcanic ash emissions was proposed by *Schmehl et al.* [2012] in which the emission rate and the underlying wind field was estimated via a method they called "genetic algorithm variational approach". They started the analysis with a randomly chosen emission strength and wind field and used an ensemble of simulations to find the best estimate iteratively. Rather strong assumptions of the emission profile were made as they placed the emissions on a single model layer, which was expected to be the most sensitive one, and adjusted the winds only on this layer. Also, no variability of the emission strength with time was allowed and no information about the observation and a priori error was used in the minimization.

Wilkins et al. [2014] viewed the problem of quantifying volcanic ash concentrations from a different perspective. Instead of estimating the emission profile they developed the "data insertion" method to optimize the modeled ash concentrations with respect to observations of volcanic ash column mass loadings. They assumed a thickness of the ash cloud of 1 and 2 km with a normally distributed vertical ash profile. The height of the volcanic ash cloud as well as the concentrations were estimated by retrievals from column mass loadings according to the algorithm of *Francis et al.* [2012]. The algorithm was successfully applied to the eruptions of Eyjafjallajökull, Iceland, 2010 (*Wilkins et al.* [2016*a*]) and Grímsvötn, Iceland, 2011 (*Wilkins et al.* [2016*b*]). Nonetheless, no multi-layered volcanic ash cloud was considered and the insertion of volcanic ash into the model domain can imply inconsistencies with the modeled meteorology. For example, volcanic ash could be placed into grid boxes into which the modeled wind would not be in compliance with its volcanic origin. Also, the data insertion algorithm does not consider errors of observations and the a priori emissions.

An adjoint free ensemble version of the 4D-var method was developed by *Lu*

et al. [2016]. They decomposed the emission profiles into a mean emission profile and perturbations. Each emission profile is represented by an ensemble member. As a result, each ensemble member and, therefore, each perturbation of the emission profile is assigned a weight such that the resulting analysis emission profile is a weighted ensemble mean. The weights are computed by minimizing a cost function that compares the weighted ensemble mean with the observations of volcanic ash column mass loadings. However, this method provides the best estimate of the emission profile ignoring the uncertainty of the estimate, which is especially large if the observational data has large uncertainties. Although the method show a good performance in the context of identical twin experiments, it suffers from the lack of knowledge on the analysis' uncertainty.

The uncertainty estimation of volcanic ash in the atmosphere is addressed by the following methods. A general framework for calculating uncertainties of volcanic ash concentrations given any model and any observational data was proposed by *Denlinger et al.* [2012]. They started the analysis by setting an ensemble of emission parameters randomly. In order to find the best posterior estimate of the emission parameters, the L^1 norm of the mismatch between the observations and the corresponding ensemble values is maximized (e.g. using the quasi-Newton method) with respect to an ensemble of emission parameters. As they assumed a narrow ash cloud and a highly peaked probability distribution (i.e. high probability is concentrated in a small region in probability space), the uncertainty of the estimated emission parameters can be obtained through a saddle point approximation. The theory of the saddle point approximation implies that the Probability Density Function (PDF) is Gaussian. Now a sample from the posterior PDF is drawn to provide an estimate of the forecast uncertainty.

A method to provide uncertainty estimations of possible volcanic ash in the atmosphere before observations are available for assimilation was introduced by *Bursik et al.* [2012]. They developed a polynomial Chaos Quadrature Weighted Estimate (PCQWE) that weights the parameters of a plume model such that the moment errors of the PDF of volcanic ash in the atmosphere are minimized. Therefore, they used an ensemble of a coupled plume model with a transport and diffusion model and perturbed four parameters of the plume model, namely vent radius, vent velocity, and mean and variance of the grain size distribution. The system is constructed to provide all uncertainties contained in a volcanic eruption before observations of volcanic ash in the atmosphere are available. Nonetheless, all available observations of the volcanic ash plume can be used to constrain the emission parameters.

Stefanescu et al. [2014] extended this approach to take also uncertainties in the wind fields into account. This is done by first selecting the emission parameters using the Conjugate Unscent Transform (Conjugate Unscent Transform (CUT); *Adurthi et al.* [2012]), which allows to reduce the ensemble size significantly. Each set of emission parameters chosen is now propagated in time using the meteorological ensemble. Each meteorological ensemble member and emission parameter

is now assigned a weight according to the PCQWE methodology. Because of the computational complexity of the methods only a limited number of source parameters and meteorological ensemble members can be perturbed in order to generate a variance in the volcanic ash ensemble.

A polynomial Chaos Quadrature (pCQ) scheme was also used by *Madankan et al.* [2014] to generate hazard maps of volcanic ash in the atmosphere. They obtained the a priori representation of the emission parameters using the pCQ method in which the integrals are solved with the CUT approach. The a posteriori probability is then calculated by assimilating satellite observations in the context of a Kalman filter update (cf. *Kalnay* [2003]) in which the a priori estimate of the source parameters can be used to calculate the covariance matrix of the emission parameters. The use of the polynomial chaos theory in these three methods requires a large ensemble in order to represent all uncertainties in the perturbed parameters. Therefore, no temporal variations in the emission parameters can be involved in this method, which make the method less applicable for longer lasting eruptions with varying volcanic ash emissions.

More recently, another inversion technique for probabilistic volcanic ash estimates was proposed by *Zidikheri et al.* [2016] (cf. also *Zidikheri et al.* [2017*a*;*b*]). Here, the focus was placed on the horizontal distribution of volcanic ash column mass loadings rather than on the emission strength. This approach aims to provide fast first safety restrictions to aviation as it analyzes the horizontal position of volcanic ash. They assumed vertically uniform distributed volcanic ash emissions for which the plume top height, bottom height, and estimated eruption time is optimized. The analysis is done by systematic sampling of model parameters within a range of reasonably chosen bounds to generate an ensemble of possible volcanic ash cloud locations. For each ensemble member a (normalized) Pattern Correlation Coefficient is calculated in comparison with the observations. All ensemble members with a Pattern Correlation Coefficient larger than a certain threshold, e.g. 95 % as was used by *Zidikheri et al.* [2016], are taken in order to generate estimates on the forecast uncertainty. This uncertainty estimate is, therefore, dependent on the chosen threshold. It is noted that this method is not able to quantify volcanic ash concentrations but only the area, which is affected by volcanic ash.

No method so far considered both, vertical and temporal varying emission profiles and their uncertainty, which is likely to occur in longer lasting eruptions. Additionally, no model allows for multiple maxima in the vertical distribution of the volcanic ash emissions. As the volcanic emissions need to be discretized temporally, multiple maxima in the vertical distribution of the volcanic ash emissions are more likely because the eruption strength and type may have changed within the temporal resolution of the emission profile. Therefore, the objective of this work is to establish an inverse analysis system for volcanic ash emissions and its uncertainty that does not rely on assumptions about the emissions. This is achieved by allowing the system to dynamically adjust to the true volcanic ash emissions. Section 2 summarizes the particle filter methodology that is used to estimate the uncertainty in the

volcanic ash estimation. The particle filter methodology is extended to a particle smoother and combined with a minimization procedure to improve the estimation of the emission profile. This minimization procedure is developed in section 3, where also its applicability for estimating volcanic ash emissions is shown. Before the full analysis system is introduced in section 5, the used atmospheric models as well as the computer system are described in section 4. The performance of the resulting analysis system is investigated in section 6 in the context of identical twin experiments. Section 7 compares the analysis system with some of the existing methods described above. Conclusions and an outlook will be given in section 8.

Chapter 2

Data assimilation via particle filtering

One main part of the target data assimilation system for volcanic ash emission estimation is a particle smoother, which is a extension of the particle filtering methodology. The objective of data assimilation is to estimate the state of the atmosphere using all available information. These include the initial model state mostly obtained from a short-term forecast or the first guess of the emissions (a priori), the observations, and finally the prognostic model itself. All information sources are typically endowed with error estimates. One major disadvantage of the classical data assimilation methods as 4D-var and Kalman filtering is the underlying assumption that the error statistics are Gaussian and the model evolution is linear. As the complexity of numerical models in atmospheric sciences increase, more nonlinear processes are included and the grid resolution is increased as well. This leads to a multimodal a posteriori Probability Density Function (PDF, cf. *Ades and van Leeuwen* [2015]), such that the above mentioned assumptions become more and more obsolete. Therefore, new data assimilation methods are needed, which do not rely on the linearization and Gaussian assumptions. Therefore, a data assimilation technique is selected in the presented study, which do not rely on any assumptions of the model and error statistics. This technique is known as particle filtering and is introduced in the following. The nomenclature follows *Ide et al.* [1997] as far as possible.

2.1 Bootstrap particle filter

Originating in the early 1990s by the work of *Gordon et al.* [1993], particle filtering is a rather new technique compared to other data assimilation methods. First applications in atmospheric sciences followed around the turn of the millennium (*Krol et al.* [1998]; *van Leeuwen* [2003]). The particle filter base on a weighted mean of ensemble members (particles in this context), an idea which dates back to the mid of the 20th century (*Metropolis et al.* [1953]; *Hastings* [1970]) with the formulations

of the Metropolis-Hastings method for Markov Chain Monte Carlo models. The particle filter method develops from the Bayes' Theorem

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{\int p(\mathbf{y}|\mathbf{x})p(\mathbf{x})d\mathbf{x}}, \tag{2.1}$$

which states that the probability of the model state given the observations is proportional to the likelihood of the observations given the model state, times the a priori probability of the model state itself. The proportionality factor is the inverse of the probability of the observations $p(\mathbf{y})$. Instead of defining a first guess or background model state and its error covariances as it is done for other data assimilation techniques, particle filter represent the a priori PDF by an ensemble of N_{ens} model runs

$$p(\mathbf{x}) = \frac{1}{N_{ens}} \sum_{i=0}^{N_{ens}-1} \delta(\mathbf{x} - \mathbf{x}_i), \tag{2.2}$$

where $\delta(\cdot)$ denotes the Kronecker delta function and $\mathbf{x}_i$ is the model state of particle i.

The particle filter consists of two step, namely the prediction step and the filtering step. In the prediction step, the model state of each ensemble member is propagated forward to the next observation time t

$$p(\mathbf{x}_i^{(t)}|\mathbf{y}^{(t-1)}) = \int p(\mathbf{x}_i^{(t)}|\mathbf{x}_i^{(t-1)})p(\mathbf{x}_i^{(t-1)}|\mathbf{y}^{(t-1)})\mathbf{dx}_i^{(t-1)}. \tag{2.3}$$

It becomes apparent that the model is a Markov process, i.e. the model state at time step t only depends on the model state of the previous time step. Once each particle has calculated the model state at the observation time, the analysis is done (filtering step)

$$\begin{aligned} p(\mathbf{x}^{(t)}|\mathbf{y}^{(t)}) =& \frac{p(\mathbf{y}^{(t)}|\mathbf{x}^{(t)})p(\mathbf{x}^{(t)})}{\int p(\mathbf{y}^{(t)}|\mathbf{x}^{(t)})p(\mathbf{x}^{(t)})d\mathbf{x}} \\ =& \sum_{i=0}^{N_{ens}-1} \frac{p(\mathbf{y}^{(t)}|\mathbf{x}^{(t)})1/N_{ens}\delta(\mathbf{x} - \mathbf{x}_i^{(t)})}{\int p(\mathbf{y}^{(t)}|\mathbf{x}^{(t)}) \sum_{j=0}^{N_{ens}-1} 1/N_{ens}\delta(\mathbf{x} - \mathbf{x}_j^{(t)})d\mathbf{x}} \\ =& \sum_{i=0}^{N_{ens}-1} \frac{p(\mathbf{y}^{(t)}|\mathbf{x}_i^{(t)})}{\sum_{j=0}^{N_{ens}-1} p(\mathbf{y}^{(t)}|\mathbf{x}_j^{(t)})}\delta(\mathbf{x} - \mathbf{x}_i^{(t)}), \end{aligned} \tag{2.4}$$

where the ensemble representation of the a priori PDF as in (2.2) is used. It is noted that in the above derivation of the particle filter method no assumptions of the error statistics and the model were made. Hence, the particle filter is applicable to states with non-Gaussian error statistics and to non-linear models. The a posteriori PDF in (2.4) can be interpreted as a weighted mean of the ensemble members with weights defined by

$$w_i = \frac{p(\mathbf{y}|\mathbf{x}_i)}{\sum_{j=0}^{N_{ens}-1} p(\mathbf{y}|\mathbf{x}_j)}, \tag{2.5}$$

where from now on the superscript (t) indicating the time step is omitted unless necessary. As *Gordon et al.* [1993] proved, the ensemble representation of the a posteriori PDF $p(\mathbf{x}|\mathbf{y})$ tends asymptotically to the true distribution if N_{ens} increases to infinity. Therefore, an approximation of the statistics of the a posteriori PDF is also directly extractable from the ensemble, e.g. the mean $\overline{\mathbf{x}}$ of the distribution is defined as

$$\overline{\mathbf{x}} = \int \mathbf{x} p(\mathbf{x}|\mathbf{y}) d\mathbf{x} \approx \sum_{i=0}^{N_{ens}-1} w_i \mathbf{x}_i. \tag{2.6}$$

The advantages of the particle filter methodology are the simple theory, an easy implementation into existing data assimilation codes, its applicability to non-Gaussian error statistics, and its usability in non-linear models. Additionally, it is noted that no explicit formulation of the background error covariances is needed as it is represented by the spread of the ensemble members via the use of the a priori PDF $p(\mathbf{x})$. Therefore, the weights of the ensemble members depend solely on the likelihood $p(\mathbf{y}|\mathbf{x})$. However, particle filters are computationally demanding.

The particle filter method is exemplarily illustrated in Fig. 2.1 in its basic form. Initially, the ensemble members are setup, ideally but not necessarily, according to the background error covariances. All ensemble members are assigned an initial weight equal to $1/N_{ens}$. Then, the ensemble of model runs is propagated to the first

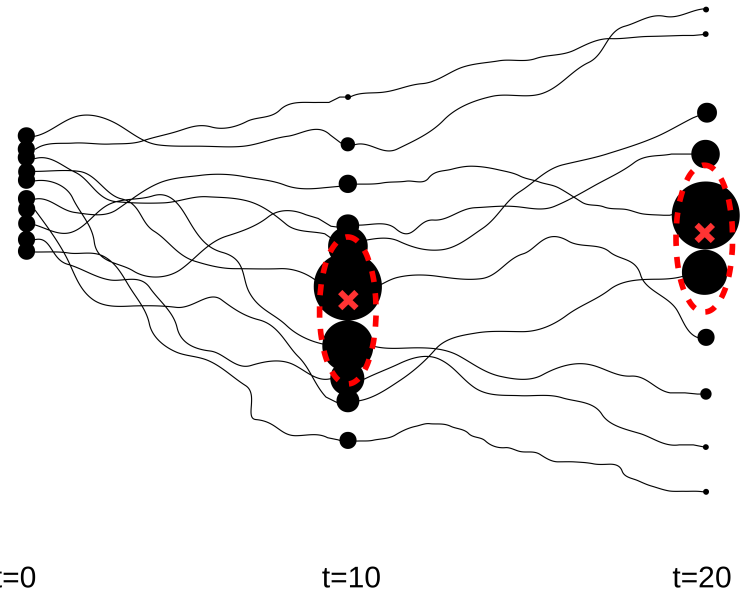

Figure 2.1: *Schematic of the bootstrap particle filter in phase space. The model state* $\mathbf{x}$ *varies along the vertical axis. Red crosses and dashed ellipses denote the observations and the high probability region around it, respectively. Black circles indicate the model equivalent of the individual ensemble members. The size of the circles represents the weight of the ensemble members. The prediction step is illustrated by the solid lines propagating the ensemble from one time step to the next.*

observation time ($t = 10$ in Fig. 2.1) calculating (2.3). At this time step the filtering step is executed by calculating the likelihood $p(\mathbf{y}|\mathbf{x}_i)$ and the weights w_i according to (2.5). Now, each ensemble member is assigned an individual weight representing the match of the ensemble members with the observations (illustrated by the circle size in Fig. 2.1). Note that the weights of the ensemble members can change significantly from one analysis time step to the other. Additionally, from Fig. 2.1 it can be deduced that if the ensemble is biased with respect to the observations, one ensemble member will get a weight close to 1 while all other ensemble members get vanishing small weights. Therefore, the choice of the ensemble members are ideally unbiased and the ensemble spread should be appropriate with respect to the background error covariances. Nonetheless, this requirement can be relaxed as methods exist, which implicitly reduce the bias of the ensemble and the observations (e.g. bias correction). After the filtering step, the statistics of the model state can be calculated. Then, the prediction step of the model again propagates the ensemble members to the next observation time.

From Fig. 2.1 it can be seen that after a certain time of integration, only a limited number of ensemble members may dominate, while the majority of the ensemble member weights vanish. This is a very likely behavior of particle filters, especially in high dimensions and is called filter degeneracy. This drawback and possible solutions to it are dealt with in the next section.

2.2 Methods to overcome filter degeneracy

The principle of particle filtering is promising in its applicability to non-linear models with non-Gaussian error statistics as it occurs in atmospheric dynamics and chemistry. Unfortunately, the poor divergence of the ensemble member weights (filter degeneracy) is the main drawback of particle filtering and is becoming more serious with increasing dimension. As the model state diverges, almost all weights tend to vanish with only one weight close to one gets all the probability mass. Therefore, the statistical value of the ensemble is lost once filter degeneracy occurs. *Bengtsson et al.* [2008] stated this fact as the "curse of dimensionality". In their analysis, they show for Gaussian and general independent and identically distributed kernels using weak assumptions that the increase of ensemble size must be super-exponential in the dimension to avoid collapse of the maximum weight to 1. *Bickel et al.* [2008] found same results but for slightly weaker assumptions on eigenvalues of the covariance matrix of the a priori PDF. In an additional analysis, *Snyder et al.* [2008] identified an effective dimension size, which depends on the state dimension, the prior distribution and the number and character of the observations. They argued that the ensemble size needs to be exponentially large in the effective dimension size to overcome filter degeneracy. Hence, a direct application of the bootstrap particle filter is not feasible because the state vector in atmospheric sciences is of the order of $\mathcal{O}(10^7 - 10^8)$.

Several techniques exist in atmospheric sciences, mostly in meteorology, which

aim to overcome this issue. The simplest technique is called resampling. The idea is to duplicate ensemble members with high weights and replace ensemble members with vanishing weights. In order to retrieve an ensemble size of N_{ens} again, the duplicates need to be perturbed. This reduces the ensemble variance in comparison to the bootstrap particle filter. Two effects emerge, which are competitive in their nature. Firstly, the reduction of the ensemble variance aims to avoid filter degeneracy. Secondly, the reduction of the variance reduces the diversity of the ensemble members such that the statistical meaning of the ensemble may get lost. This second effect is called filter impoverishment (*Li et al.* [2014]) and is a direct consequence of the resampling step.

Different resampling strategies exist in the literature (see *Douc et al.* [2005] and *Hol et al.* [2006] for a review). *Douc et al.* [2005] compared resampling algorithms and argued that stratified and residual resampling perform best. In stratified resampling the new ensemble is drawn from a multinomial distribution $Mult(N_{ens}; u_0, \ldots, u_{N_{ens}-1})$ with parameters $u_k = \frac{(k-1)+\tilde{u}_k}{N_{ens}}$, where $\tilde{u}_k$ is uniformly distributed according to $U[0,1)$ and $k = 0, 1, \ldots, N_{ens} - 1$. This choice reduces the ensemble variance compared to multinomial resampling, in which the parameters u_k of the multinomial distribution are given by the ensemble member weights ($u_k = w_k$; cf. *Hol et al.* [2006]). Because of its importance for the presented work, residual resampling is described in more detail. Residual resampling was introduced by *Liu and Chen* [1998]. Assuming that initially to the resampling step each particle has weight w_i. The duplication counts N_i, i.e. the number of duplicates for ensemble member i, is then calculated using

$$N_i = \lfloor N_{ens} w_i \rfloor + \overline{N}_i, \tag{2.7}$$

with $\lfloor \cdot \rfloor$ denoting the floor function. The residual $\overline{N}_i$ is chosen according to the multinomial distribution $Mult(N_{ens} - R; \overline{w}_0, \ldots, \overline{w}_{N_{ens}-1})$, with $R = \sum_j \lfloor N_{ens} w_j \rfloor$ and

$$\overline{w}_i = \frac{N_{ens} w_i - \lfloor N_{ens} w_i \rfloor}{N_{ens} - R}. \tag{2.8}$$

In practice, the residual $\overline{N}_i$ is calculated as follows:

First draw $N_{ens} - R$ random uniformly distributed variables $u_k \in [0, 1)$, with $k = 0, 1, \ldots, N_{ens} - R - 1$. The residual $\overline{N}_i$ is the number of variables u_k, which are in the interval $\{\sum_{j \leq i} \overline{w}_j, \sum_{j<i+1} \overline{w}_j\}$, with $i = 0, 1, \cdots, N_{ens} - 1$ and $\overline{w}_0 = 0$. As for stratified resampling, residual resampling is efficient in decreasing the ensemble variance (*Douc et al.* [2005]).

It was early recognized that resampling is not enough to overcome filter degeneracy in atmospheric sciences. The reason is that the dimension of the problems, in general, is still too large such that resampling has too little an effect. Different other particle filter extensions have been proposed in order to avoid filter degeneracy in high dimensions. Among those are, for example, the implicit particle filter (*Chorin and Tu* [2009]; *Chorin et al.* [2010]; *Morzfeld et al.* [2012]; *Atkins et al.* [2013]) and localized particle filters (e.g. *Penny and Miyoshi* [2016]; *Poterjoy* [2016]).

The implicit particle filter is a combination of the particle filter with the 4D-var method. First, the atmospheric state is approximated using 4D-var. Once the optimal model state is calculated, the ensemble is setup by solving an algebraic equation using the minimum of the cost function. This minimum is obtained by the 4D-var analysis. There exist different methods for the construction of the algebraic equation (cf. *Atkins et al.* [2013]). One example is the random map proposed by *Morzfeld et al.* [2012], where the algebraic equation is calculated by Gaussian random variables ξ as $J(\mathbf{x}) - \Phi_J = \frac{1}{2}\xi^T\xi$, with $J(\cdot)$ and Φ_J denoting the cost function and its minimum, respectively. The choice of $\xi \in [0, 1]$ ensures that the costs $J(\mathbf{x})$ of the ensemble members are close to the minimum Φ_J. It is noted that in the implicit particle filter first an approximation of the minimum of the cost function is determined and then an ensemble is setup around this optimal state. Therefore, the optimal state has to be close to the truth, otherwise the ensemble may be biased and not representative for the mean of the a posteriori PDF $p(\mathbf{x}|\mathbf{y})$. As stated above, this may be critical for volcanic ash emission estimation.

The localization in the localized particle filter proposed by *Poterjoy* [2016] aims mainly to reduce the dimension of the data assimilation problem. This is done by a radius of influence around the observations. Outside this radius of influence a global weight according to (2.5) is assigned to each ensemble member for all grid cells. At the observation location the weight is only controlled by the costs of this observation. Within the radius of influence a smooth transition between the weights is calculated such that atmospheric dynamics and correlations are not violated. The localized particle filter was tested for volcanic ash estimation. For convenience, it is described in more detail in the appendix.

Another set of approaches to avoid filter degeneracy uses the so called proposal density $q(\mathbf{x})$. It is noted that some of the above mentioned methods can also be interpreted in terms of proposal densities. The idea is to sample from the proposal density rather than from the original PDF $p(\mathbf{x})$. Therefore, the prediction density in (2.3) needs to be altered to

$$p(\mathbf{x}^{(t)}|\mathbf{y}^{(t-1)}) = \int \frac{p(\mathbf{x}^{(t)}|\mathbf{x}^{(t-1)})}{q(\mathbf{x}^{(t)}|\mathbf{x}^{(t-1)}, \mathbf{y}^{(t)})} q(\mathbf{x}^{(t)}|\mathbf{x}^{(t-1)}, \mathbf{y}^{(t)}) p(\mathbf{x}^{(t-1)}|\mathbf{y}^{(t-1)}) \mathbf{d}\mathbf{x}^{(t-1)}, \quad (2.9)$$

with $q(\mathbf{x}^{(t)}|\mathbf{x}^{(t-1)}, \mathbf{y}^{(t)})$ simulating information about the "future" observations. To sample from the proposal density provides two advantages. Firstly, the proposal density can be chosen arbitrarily, as far as the support is the same as for $p(\mathbf{x})$. For example, the proposal density can be chosen such that it is easier to sample from it than from the a priori PDF. Secondly, by including information of the "future" observations into the proposal density, it is expected that the analysis ensemble is closer to the observations. Hence, more ensemble members would get meaningful weights.

Particle filter methods using the proposal density are, for example, the auxiliary particle filter (*Pitt and Shephard* [1999]) and the equivalent weights particle filter (*van Leeuwen* [2010]; *Ades and van Leeuwen* [2013; 2015]). The auxiliary particle

filter consists of two stages. In the first stage a reduced ensemble (either reduced in ensemble size or complexity of the model) is propagated from the current time step t to the next filtering time step $t+1$. The full ensemble is resampled at time t using the weights gained at the filtering time step $t+1$ with the reduced ensemble. The resulting ensemble is propagated in full complexity to the filtering time step $t+1$. Formally, this is the sampling from the proposal density as the information from the "future" observations is taken. It is expected that the ensemble with full complexity is closer to the observations than it would be without the auxiliary ensemble. Hence, less computational power is lost by integrating model states, which get a low weight and, therefore, have low statistical value. At the filtering time step, the ensemble may be weighted again but a resampling step is not necessary since the resampling was done with the reduced ensemble in the first stage. The weights in the auxiliary particle filter are expected to be more similar and the ensemble is expected to be close to the observations because the first stage ensemble aims to provide a good approximation of the final ensemble distribution.

Besides using the information of "future" observations via the proposal density, the equivalent weights particle filter enforces the ensemble members to have equal weights. This is obtained by solving the quadratic equation

$$\begin{aligned} -\log w_i^{rest} + &\frac{1}{2}(\mathbf{x}_i^{(t)} - f(\mathbf{x}_i^{(t-1)}))^T \mathbf{B}^{-1}(\mathbf{x}_i^{(t)} - f(\mathbf{x}_i^{(t-1)})) \\ &+\frac{1}{2}(\mathbf{y}^{(t)} - h(\mathbf{x}_i^{(t)}))^T \mathbf{R}^{-1}(\mathbf{y}^{(t)} - h(\mathbf{x}_i^{(t)})) = C, \end{aligned} \tag{2.10}$$

where w_i^{rest} is the weight of former analysis time steps, $f(\cdot)$ is the (non-linear) model operator, $h(\cdot)$ is the (non-linear) observation operator, and C is chosen such that a certain amount of ensemble members (e.g. 80 %) can solve this equation. The ensemble members with too small weights are dumped and resampled after the analysis. This leads to the updated model equation

$$\mathbf{x}_i^{(t)} = f(x_i^{(t-1)}) + \alpha_i \mathbf{K}(\mathbf{y}^{(t)} - \mathbf{H} f(\mathbf{x}_i^{(t-1)})), \tag{2.11}$$

with

$$\begin{aligned} \alpha_i &= 1 \pm \sqrt{1 - b_i/a_i}, \\ a_i &= 0.5 \mathbf{d}_i^T \mathbf{R}^{-1} \mathbf{H} \mathbf{K} \mathbf{d}_i, \\ b_i &= 0.5 \mathbf{d}_i^T \mathbf{R}^{-1} \mathbf{d}_i - C - \log w_i^{rest}, \\ \mathbf{K} &= \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}, \\ \mathbf{d}_i &= \mathbf{y}^{(t)} - \mathbf{H} f(\mathbf{x}_i^{(t-1)}). \end{aligned} \tag{2.12}$$

In (2.12), $\mathbf{H}$ is the linearized observation operator, $\mathbf{K}$ is the Kalman gain matrix and $\mathbf{d}_i$ is the so called innovation vector, i.e. the difference between the observations and the model state. In *van Leeuwen and Ades* [2013] the equivalent weights particle filter is successfully applied to a model solving the barotropic vorticity equation on a 65,500 dimensional grid using only 32 ensemble members. Therefore, *van Leeuwen*

[2015*a*] stated that the equivalent weights particle filter is able to solve the curse of dimensionality. A detailed description of the multiple steps of the equivalent weights particle filter is given in *Ades and van Leeuwen* [2015].

For further reading, there exist various reviews on particle filter methods in the literature and the reader is referred to this for a more comprehensive view on this topic (e.g. *Doucet et al.* [2001]; *van Leeuwen* [2009]; *Andrieu et al.* [2010]; *van Leeuwen* [2015*b*]).

Chapter 3

Minimization algorithm

The objective of this work is to find the best estimate of volcanic ash concentrations in the atmosphere, including its uncertainty, by optimizing emission sources. Particle filtering provides a methodology to estimate the uncertainty of the model state if unbiased and reliable a priori information is given. However, only limited information is available for the estimation of the volcanic ash emission profiles such that the a priori PDF used in particle filtering is likely to be underrepresented by a small ensemble. In this section, a method is developed that is able to calculate an ensemble of analysis volcanic ash emissions that can be used as a priori information for particle filtering (see section 5.2 for the description of the full, combined analysis system). Therefore, an ensemble extension of the standard direct search method by *Nelder and Mead* [1965] using discrete solutions was developed. The extension aims to find the best representation of the cost function near its minimum in order to describe the uncertainty of the analysis by optimally combining pairwise distinct emission packages.

3.1 Ensemble of emission scenarios

The main concept of the assimilation system is to emit a set of N_{emis} pairwise distinct emission packages (emission scenarios as they were called by *Stohl et al.* [2011]) with a default unit ash amount such that the volcanic ash concentration resulting from each emission scenario is given through a forward model run. Each emission scenario is defined by an emission package for a single model layer and a single time frame (e.g. one hour). Since the observation operator for volcanic ash column mass loadings is linear, finding the optimal combination of the volcanic ash concentration patterns, which emerged from the pairwise distinct emission packages, is a reasonable approach. A similar approach for estimating the volcanic ash column mass loadings was used by *Stohl et al.* [2011] and *Kristiansen et al.* [2015] but only for the estimation of the optimal emission profile and not for its uncertainty.

The use of a linear combination of pairwise distinct emission packages as a priori information alters the cost function to be minimized. First, the observational part

of the cost function used in 4D-var is considered. Let $\mathbf{x}_t$ be the model state at time t and $\mathbf{y}_t$ the corresponding observations, then the observational part of the 4D-var cost function is

$$J(\mathbf{x}) = \sum_{t=0}^{L} (H\mathbf{M}_t(\mathbf{x}_0) - \mathbf{y}_t)^T \mathbf{R}^{-1} (H\mathbf{M}_t(\mathbf{x}_0) - \mathbf{y}_t) \tag{3.1}$$

where L is the length of the assimilation window, H is the (possibly non-linear) observation operator mapping the model state into observation space, $\mathbf{M}$ is the model operator integrating the model state from time 0 to t, and $\mathbf{R}$ is the observation error covariance matrix. The superscript T denotes the transpose of a matrix or vector. The use of emission packages with default emission strength changes the independent variable from the model state $\mathbf{x}$ to the emission factor $\mathbf{a}$. This emission factor scales the emission strength and the corresponding volcanic ash concentrations for each emission package and is, therefore, the variable to be optimized. Hence, the cost function changes to

$$J(\mathbf{a}) = \sum_{t=0}^{L} (H\tilde{M}_t[\mathbf{a}_t(\mathbf{z})\mathbf{e}_0] - \mathbf{y})^T \mathbf{R}^{-1} (H\tilde{M}_t[\mathbf{a}_t(\mathbf{z})\mathbf{e}_0] - \mathbf{y}), \tag{3.2}$$

with the source-receptor model $\tilde{M}_t$ mapping the unit emissions $\mathbf{e}_0$ onto the model state and transports it. Here, $\mathbf{a}_t(\mathbf{z})$ denotes the emission factor profile for a specific

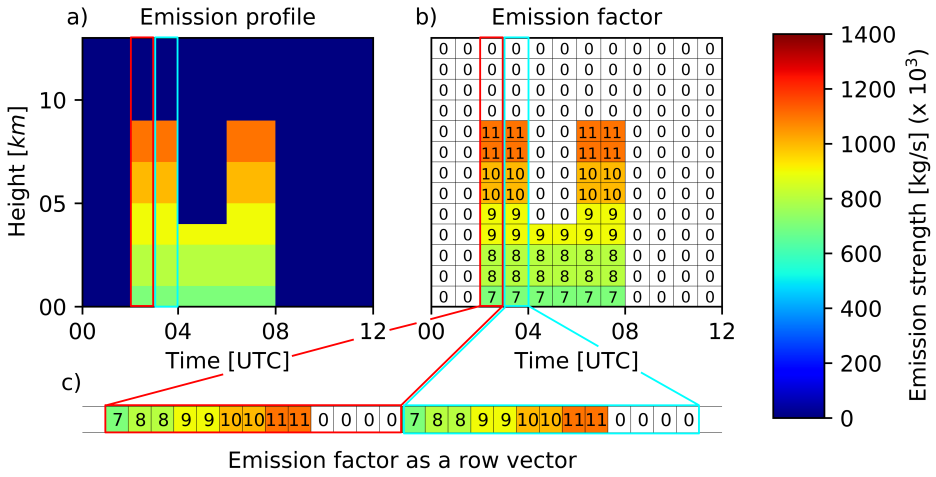

Figure 3.1: *Conversion of the emission profile to a row vector of emission factors as it is used by the DENM (see section 3.5) method. (a) Hovmoeller plot of the emission profile for each hour of the 12 hour analysis period. Colors indicate emission strengths of a hypothetical eruption (nature run emission profile). (b) Same as (a) but with emission factors for each emission package. In this case the emission factors are organized in a matrix. To be usable for the DENM method, the matrix in (b) is transformed into a vector (c).*

time for all heights **z**. This is exemplarily illustrated by the red and light blue rectangle in Fig. 3.1. A Hovmoeller plot (time vs. height) of the emission profiles of a hypothetical sub-plinian eruption with two explosive phases between 2-4 UTC and 6-8 UTC, respectively, is shown in Fig. 3.1a. The Hovmoeller plot shows in each column the emission profile of the volcanic eruption for a given time. The emission strength of the eruption range from 700 kg/s to 1100 kg/s, while the emissions for the bulk of the time-height combinations remain zero (blue color in Fig. 3.1a). The emission factors that correspond to a perfect analysis are summarized in Fig. 3.1b for this example. It is assumed that each ensemble member emits a default emission rate of volcanic ash of 100 kg/s for one hour in one model layer. For example, ensemble member 0 emits 100 kg/s from 0-1 UTC in model layer 1, ensemble member 1 emits 100 kg/s from 0-1 UTC in layer 2, and so on.
To be used in a minimization algorithm, the time-height matrix of Fig. 3.1b has to be transformed into the vector **a** of emission factors, wherein each entry corresponds to the emission factor for one emission package and, therefore, for one ensemble member. This transformation is illustrated in Fig. 3.1c. As can be seen from this graphic, the emission factors a_i that minimize the cost function are not smooth and the minimization has to deal with multiple discontinuities.

3.2 Original Nelder-Mead method

The choice made above for finding the optimal emission profile of volcanic ash by a combination of emission packages with (scaled) emission strengths leads to the need of a combinatorial minimization algorithm. The reason is that not all valid combinations of emission profiles can be calculated. Assuming N emission packages with M different possible emission strengths (i.e. emission factors) for each package, the number of combinations is $c = M^N = 10^{120}$ for, for example, 120 emission packages and 10 emission factors. This is not feasible to be calculated. Therefore, an algorithm is needed that finds the optimal combination with respect to the given cost function (3.2) within a limited number of steps. The direct search method proposed by *Nelder and Mead* [1965] is chosen, which is not restricted to only M different emission factors. This combinatorial minimization is not manageable for e.g. quasi-Newton methods as the limited memory Broyden-Flechter-Goldfarb-Shanno (L-BFGS) algorithm (*Liu and Nocedal* [1989]). Additionally, the Nelder-Mead algorithm needs relatively few function evaluations (mostly 1-2 per iteration, *Lagarias et al.* [1998]). As a direct search method, the Nelder-Mead minimization algorithm is expected to be robust, especially in cases where the function has discontinuities or the function values are noisy (see *McKinnon* [1998]).

The method proposed by *Nelder and Mead* [1965] is a popular direct search method (e.g. *Tseng* [1999]). The idea of the algorithm is to move a simplex on the surface of a N-dimensional function. A simplex is a hull in model space, which is defined by its $N+1$ vertices (or points). The simplex and the different moves of the Nelder-Mead algorithm are shown in Fig. 3.2, in which (a) the standard test function

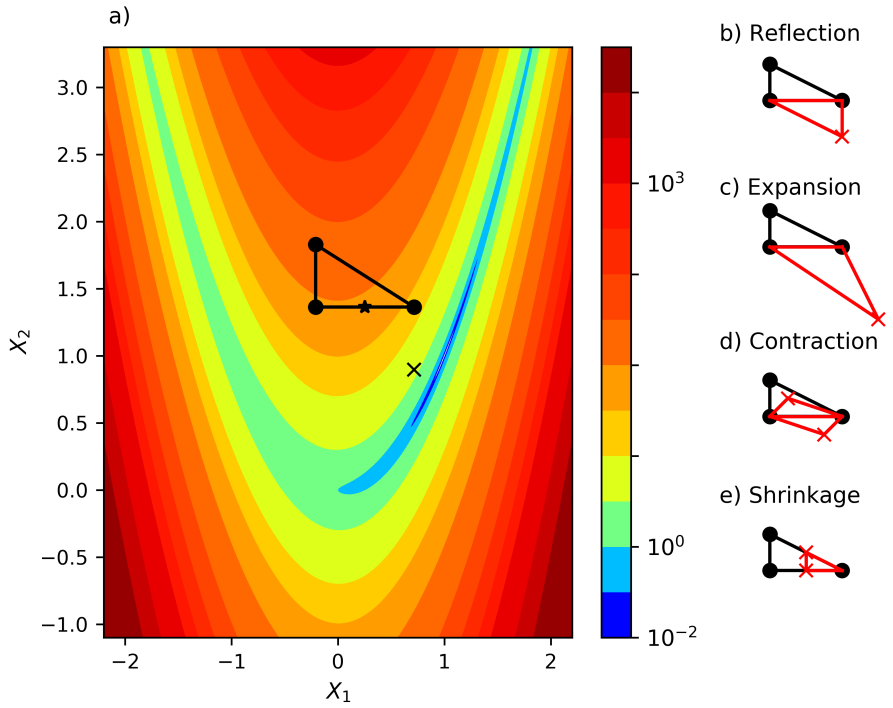

Figure 3.2: *(a) Example of a simplex on the standard test function Rosenbrock defined by* $f(\mathbf{x}) = (1 - x_1)^2 + 100 * (x_2 - x_1^2)^2$*. The black triangle indicates the simplex spanned by its vertices (black dots). The centroid is displayed as an asterisk, the reflection vertex as a cross. The simplex obtained through (b) reflection, (c) expansion, (d) contraction (two possibilities), and (e) shrinkage are overlaid in red over the initial simplex. Here the crosses denote the newly calculated vertices.*

Rosenbrock and a realization of a simplex (black triangle) is depicted. The black dots indicate the vertices that span the simplex. The centroid of the N best vertices and the reflection vertex are depicted as an asterisk and a cross, respectively. Additionally, the four possible moves of the simplex in the Nelder-Mead method, which are (b) reflection, (c) expansion, (d) inner and outer contraction, and (e) shrinkage, are shown on the right of Fig. 3.2. The simplex can take any shape on the cost functions surface and its inner angles can become arbitrarily small.

The initial simplex is constructed by choosing the first vertex reasonably well (background). Here, prior knowledge on the solution can be used. The other N vertices are generated by perturbing the initial vertex in each dimension. Hence, the simplex is able to move along every dimension. To exclude movement in one dimension, the value in this dimension can be kept fixed for all vertices. The initial simplex S is spanned by the initial vertices: $S = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{N-1}, \mathbf{x}_N\}$. It is noted that the vertices correspond to one vector of emission factors in the application to volcanic ash emission estimation. The movement of the simplex is performed by changing the position of the vertex with the highest costs according to the following steps (one iteration of the Nelder-Mead algorithm):

1. **Reflection** (cf. Fig. 3.2b): The first step is to calculate the reflected vertex. Without loss of generality the vertices are assumed to be ordered by costs, such that $f(\mathbf{x}_0) \leq f(\mathbf{x}_1) \leq \cdots \leq f(\mathbf{x}_{N-1}) \leq f(\mathbf{x}_N)$. In case of multiple vertices having the same function value, a proper tie-breaking rule as the one proposed by *Lagarias et al.* [1998] has to be applied in order to make the algorithm well-defined. To move the simplex towards the minimum, the worst vertex $\mathbf{x}_N$ is reflected at the centroid $\mathbf{c}$ of the N best vertices:

$$\mathbf{c} = \frac{1}{N} \sum_{i=0}^{N-1} \mathbf{x}_i,$$

$$\mathbf{x}^{(r)} = \mathbf{c} - \alpha(\mathbf{x}_N - \mathbf{c}),$$

where $\mathbf{x}^{(r)}$ is the reflection of $\mathbf{x}_N$ and $\alpha > 0$ is the reflection coefficient. The reflected simplex is: $S^{(r)} = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{N-1}, \mathbf{x}^{(r)}\}$. Once the reflection $\mathbf{x}^{(r)}$ is calculated, the choice of the remaining steps depends on its function value $f(\mathbf{x}^{(r)})$.

2. **Expansion** (cf. Fig. 3.2c): If $\mathbf{x}^{(r)}$ becomes the best vertex of the new simplex $S^{(r)}$, i.e. $f(\mathbf{x}^{(r)}) < f(\mathbf{x}_0)$, the reflected vertex is further expanded by

$$\mathbf{x}^{(e)} = \mathbf{c} + \beta(\mathbf{x}^{(r)} - \mathbf{c}),$$

where $\mathbf{x}^{(e)}$ is the expanded vertex and $\beta > 1$ is the expansion coefficient. The expanded simplex is: $S^{(e)} = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{N-1}, \mathbf{x}^{(e)}\}$. If the expanded vertex is a better estimate of the minimum than the reflected vertex (i.e. $f(\mathbf{x}^{(e)}) < f(\mathbf{x}^{(r)})$), S is replaced by $S^{(e)}$, otherwise by $S^{(r)}$. The next iteration is started.

3. **Contraction** (cf. Fig. 3.2d): If the reflected vertex is not an improvement over the best vertex (i.e. $f(\mathbf{x}^{(r)}) > f(\mathbf{x}_0)$) but it is an improvement over the worst vertex ($f(\mathbf{x}^{(r)}) < f(\mathbf{x}_N)$), the reflected vertex is moved towards the centroid (outer contraction)

$$\mathbf{x}^{(oc)} = \mathbf{c} + \gamma(\mathbf{x}^{(r)} - \mathbf{c}),$$

where $0 < \gamma < 1$ is the contraction coefficient. If the reflection is no improvement over the worst vertex ($f(\mathbf{x}^{(r)}) > f(\mathbf{x}_N)$), the worst vertex is moved towards the centroid (inner contraction)

$$\mathbf{x}^{(ic)} = \mathbf{c} - \gamma(\mathbf{x}^{(r)} - \mathbf{c}).$$

If the contracted vertex $\mathbf{x}^{(c)}$ (either $\mathbf{x}^{(oc)}$ or $\mathbf{x}^{(ic)}$) is an improvement over the worst vertex (i.e. $f(\mathbf{x}^{(c)}) < f(\mathbf{x}_N)$) the contracted vertex is accepted and the new simplex is $S^{(c)} = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{N-1}, \mathbf{x}^{(c)}\}$. The next iteration is started.

4. **Shrinkage** (cf. Fig. 3.2e): If the contraction vertex is no improvement over the worst vertex $\mathbf{x}^{(N)}$, non of the above steps lead to a vertex that is closer to the minimum than the worst vertex $\mathbf{x}_N$. In this case all vertices are shrank towards the best vertex

$$\mathbf{x}_i = \mathbf{x}_0 + \delta(\mathbf{x}_i - \mathbf{x}_0) \ , \forall i = 1, N.$$

Herein, $0 < \delta < 1$ is the shrinkage coefficient. The original simplex is replaced by the shrank one. The effect of the shrinkage for the simplex has an important consequence on the methods behavior. By shrinking the simplex towards the best vertex, the volume of the simplex is reduced. As a consequence, the step length for the other three steps of the algorithm (reflection, expansion, and contraction) is reduced as well, which is an advantage in the vicinity of the global minimum. If the step length is too large, even the reflection step overshoots such that it cannot find a new best vertex. In the vicinity of a local minimum, the shrinkage of the simplex may hinder the algorithm to overcome this local minimum. In this case the global minimum may not be found.

The above described iteration is terminated if a defined termination criterion is met. Different choices of termination criteria exist, for example

$$\sqrt{\frac{1}{N}\sum_{i=0}^{N}\left(f(\mathbf{x}_i) - \overline{f(\mathbf{x})}\right)^2} < \epsilon_1, \tag{3.3}$$

or

$$f(\mathbf{x}_0) < \epsilon_2, \tag{3.4}$$

where $\epsilon_1 \ll 1$ and $\epsilon_2 \ll 1$. The first criterion terminates the minimization if the simplex size has collapsed below a threshold. The second criterion terminates the minimization if the best vertex is close enough to the minimum. If one dimension is expected to be known more precisely, this dimension can be kept fixed for all initial vertices (*Nelder and Mead* [1965]) such that in this dimension no movement is possible anymore.
By first investigating the behavior of the minimization algorithm, *Nelder and Mead* [1965] found that the convergence rate depends on the size and orientation of the initial simplex. This can easily be understood by examining Fig. 3.2a. By increasing the volume of the simplex, the distance between the worst vertex, which is to be updated (or moved), and the centroid increases. Hence, the distance between the centroid and the reflection vertex increases as well. Also, if the simplex is turned clockwise around the centroid, the reflection vertex will turn clockwise, too. Likewise, the expansion, contraction, and shrinkage steps will also lead to different updates if the volume or orientation of the initial simplex is changed. This alters the performance of the minimization as was found by *Nelder and Mead* [1965].

Although the minimization algorithm is widely used (e.g. *Gao and Han* [2012]) and it shows satisfactory minimization behavior for most cost functions, not much

Algorithm 1 Pseudo-code of the original simplex method by *Nelder and Mead* [1965]. A detailed description of the algorithm can be found in the text.

Set initial vertex $\mathbf{x}_0$
generate simplex by perturbing initial vertex: $\mathbf{x}_i = \mathbf{x}_0 + \xi\mathbf{e}_i$, ($i = 1, \ldots, N; \xi$ arbitrarily, may depend on the problem)
while $\frac{1}{N+1}\sum_i \left(f(\mathbf{x}_i) - \overline{f}(\mathbf{x})\right)^2 > \epsilon^2$ **and** ITX < NITER **do**
  sort $\mathbf{x}_i$, $\{i = 0, \ldots, N\}$ such that $f(\mathbf{x}_0) = y_0 \leq y_1 \leq \cdots \leq y_N = f(\mathbf{x}_N)$
  compute centroid of N best vertices $\mathbf{c} = \frac{1}{N}\sum_{i=0}^{N-1} \mathbf{x}_i$
  compute reflection: $\mathbf{x}^{(r)} = \mathbf{c} - \alpha(\mathbf{x}_N - \mathbf{c})$
  if $f(\mathbf{x}^{(r)}) < y_0$ **then**
    compute extension: $\mathbf{x}^{(e)} = \mathbf{c} + \beta(\mathbf{x}^{(r)} - \mathbf{c})$
    if $f(\mathbf{x}^{(e)}) < f(\mathbf{x}^{(r)})$ **then**
      $\mathbf{x}_N = \mathbf{x}^{(e)}$
    else
      $\mathbf{x}_N = \mathbf{x}^{(r)}$
    end if
  else
    if $f(\mathbf{x}^{(r)}) < y_{N-1}$ **then**
      $\mathbf{x}_N = \mathbf{x}^{(r)}$
    else
      contract towards centroid
      if $f(\mathbf{x}^{(r)}) < y_N$ **then**
        outer contraction $\mathbf{x}^{(oc)} = \mathbf{c} + \gamma(\mathbf{x}^{(r)} - \mathbf{c})$
        $\mathbf{x}^{(c)} = \mathbf{x}^{(oc)}$
      else
        inner contraction $\mathbf{x}^{(ic)} = \mathbf{c} - \gamma(\mathbf{x}^{(r)} - \mathbf{c})$
        $\mathbf{x}^{(c)} = \mathbf{x}^{(ic)}$
      end if
      if $f(\mathbf{x}^{(c)}) < y_N$ **then**
        $\mathbf{x}_N = \mathbf{x}^{(c)}$
      else
        shrink to best vertex
        for i=1,N **do**
          $\mathbf{x}_i = \mathbf{x}_0 + \delta(\mathbf{x}_i - \mathbf{x}_0)$
        end for
      end if
    end if
  end if
  ITX += 1
end while

theoretical investigations about the minimization properties has been done. *Lagarias et al.* [1998] proofed the convergence of the algorithm in one dimensional problems for strictly convex functions. For two dimensional problems, they showed that the cost function values of the vertices converge to the same value and that the diameter of the simplex converges to zero. Nevertheless, they pointed out that this does not imply that the vertices converge to a single point. For higher dimensional cost functions no analysis of the behavior is available. For a set of strictly convex functions, *McKinnon* [1998] derived a family of initial simplices that converge to a non-stationary point. He showed that this is due to a non-stop sequence of inner contractions.

A method that overcomes this issue was proposed by *Tseng* [1999], whose algorithm differs mainly in the descent criterion from the standard Nelder-Mead method. While *Nelder and Mead* [1965] require strict descend of the updated vertex with respect to the worst vertex of the simplex, *Tseng* [1999] demand a fortified-descent. This means that a new vertex is only accepted if the improvement over the worst vertex exceeds some threshold. Otherwise the iteration is repeated with the (non-improving) shrank simplex. Additionally, the fortified-descent minimization algorithm does not allow for a zero volume of the simplex and for zero angles at the vertices, a condition, which is not necessarily met in the original Nelder-Mead minimization algorithm. A first conclusion of *Tseng* [1999] indicates that the method settings closest to the Nelder-Mead algorithm produces the best results.

3.3 Nelder-Mead method with adaptive parameters

Although multiple examples exist in which the algorithm does not minimize correctly (cf. *Torczon* [1989]; *Wright* [1996]; *McKinnon* [1998]; *Han* [2000]), especially in higher dimensions, the Nelder-Mead method is used widely, also for larger dimensions. *Han and Neumann* [2006] analyzed the so called "effect of dimensionality" for the quadratic cost function $f(\mathbf{x}) = \mathbf{x}^T\mathbf{x}$. They showed that the performance of the method by *Nelder and Mead* [1965] decreases as the dimension increases. In the following, the issue introduced by higher dimensional problems is further addressed.

After the analysis of *Nelder and Mead* [1965], in which the performance of different values of the parameters α, β, γ, and δ (see algorithm 1) were compared, a set of standard choice of these parameters has established (cf. *Gao and Han* [2012])

$$\alpha = 1 \ , \ \beta = 2 \ , \ \gamma = 1/2 \ , \ \delta = 1/2. \tag{3.5}$$

The values in the literature may deviate because of different formulations of the four steps of the algorithm. *Gao and Han* [2012] proofed for uniformly convex functions that the expansion and contraction steps with these parameter values lose efficiency as the dimension increases. Further, they concluded that reducing the chances of reflection to occur and forcing the diameter of the simplex away from zero can

Table 3.1: *Codes for the initial simplices for the performance tests of the Nelder-Mead methods minimizing cost functions $f(\mathbf{x})$ and $g(\mathbf{x})$. The initial vertex $\mathbf{x}_0$ ranges from* **1** *to* **100** *(a-f; rows in the table). The perturbations σ (i-vii; columns in the table) of the N additional vertices ranges from 0.001 to 0.5.*

	perturbations σ						
$\mathbf{x}_0$	0.001	0.005	0.01	0.05	0.1	0.2	0.5
1	*a.i*	*a.ii*	*a.iii*	*a.iv*	*a.v*	*a.vi*	*a.vii*
5	*b.i*	*b.ii*	*b.iii*	*b.iv*	*b.v*	*b.vi*	*b.vii*
10	*c.i*	*c.ii*	*c.iii*	*c.iv*	*c.v*	*c.vi*	*c.vii*
20	*d.i*	*d.ii*	*d.iii*	*d.iv*	*d.v*	*d.vi*	*d.vii*
50	*e.i*	*e.ii*	*e.iii*	*e.iv*	*e.v*	*e.vi*	*e.vii*
100	*f.i*	*f.ii*	*f.iii*	*f.iv*	*f.v*	*f.vi*	*f.vii*

increase the performance of the algorithm. As a consequence, they adapted the parameters α, β, γ, and δ to the dimension of the problem such that

$$\alpha = 1 \ , \beta = 1 + \frac{2}{N} \ , \gamma = 0.75 - \frac{1}{2N} \ , \delta = 1 - \frac{1}{N}. \tag{3.6}$$

For increasing dimension N, the expansion, contraction, and shrinkage steps become smaller, hindering the volume of the simplex to decrease too rapidly. In an analysis, they demonstrated that the Nelder-Mead method using adaptive parameters (ANM) outperforms the standard version of the algorithm for various functions and dimensions, although in some instances more function evaluations were required. In the following, the behavior of the original Nelder-Mead method (ONM) and ANM is investigated.

The ANM method is compared with the original version of *Nelder and Mead* [1965] with special interest in the performance dependency on the choice of initial simplex and, thus, on its applicability to volcanic ash emission estimation. Therefore, for the N-dimensional spherical cost function

$$f(\mathbf{x}) = \sum_{i=0}^{N-1} x_i^2, \tag{3.7}$$

with the dimension of the problem $N = 120$, the minimization is performed for 100,000 iterations for various initial simplices. The initial simplices were constructed according to

$$\mathbf{x}_i = \mathbf{x}_0 + \delta_{ij}\sigma(\mathbf{e}_j^T\mathbf{x}_0)\mathbf{e}_j, \tag{3.8}$$

with δ_{ij} the Kronecker-Delta function and $\mathbf{e}_j$ the unit vector in dimension j. This means that first an initial vertex $\mathbf{x}_0$ is defined. All other vertices of the simplex are equal to that initial vertex except for one perturbed dimension i for vertex i. The values of the initial simplex and the perturbations are summarized in Tab. 3.1.

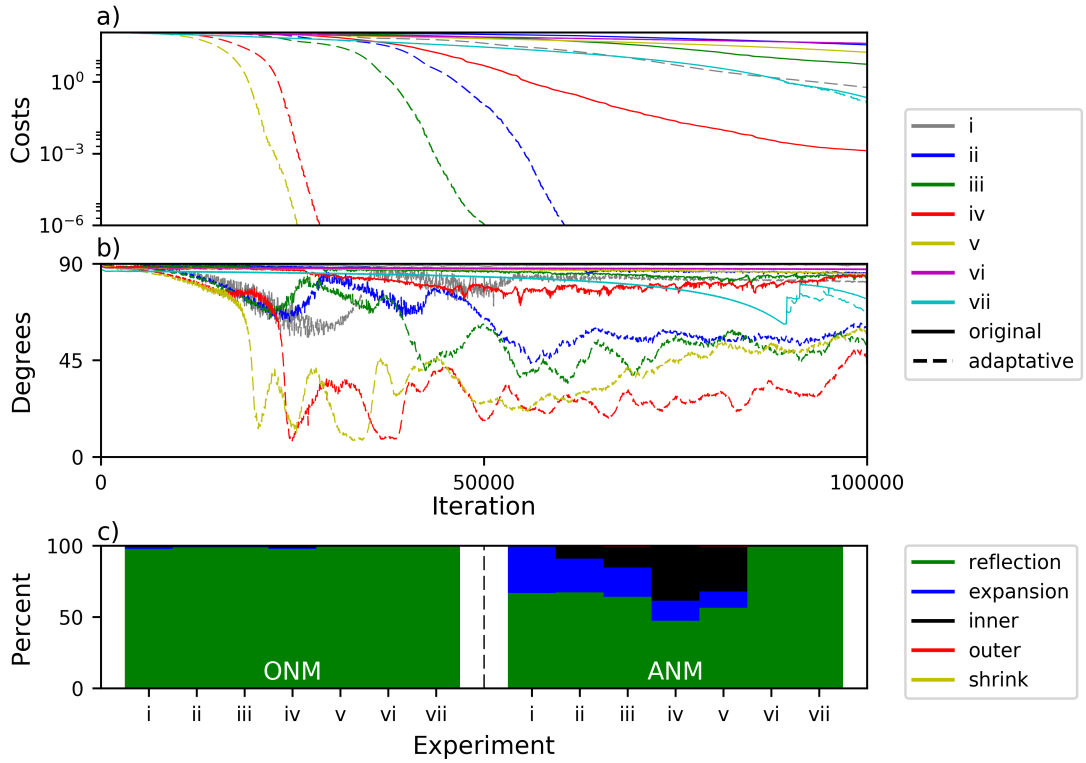

Figure 3.3: *Comparison of the performance of the ONM (solid lines) and the ANM (dashed lines) methods for the test cases with initial vertex* $\mathbf{x}_0 = \mathbf{1}$ *(experiment set a) minimizing cost function* $f(\mathbf{x})$ *(cf. also Tab. 3.1 for details on the experiment sets). (a) Cost evolution for the minimization; (b) angle* α *between the actual search direction and the optimal search direction (gradient for this cost function); (c) relative frequency of the five steps reflection (green), expansion (blue), inner contraction (black), outer contraction (red), and shrinkage (yellow). The abscissa in (c) indicates the perturbation* σ *to generate the initial simplex (cf. Tab. 3.1).*

For example, initial simplex *c.iv* consists of an initial vertex with $\mathbf{x}_0 = \mathbf{10}$ and N additional vertices with $\sigma = 0.05$.

The results for the different initial vertices are qualitatively similar, hence, only the test results with initial vertex $\mathbf{x}_0 = \mathbf{1}$ are shown in Fig. 3.3. The cost evolution for all seven perturbation tests are shown in Fig. 3.3a. For all perturbations, the ANM method (dashed lines) outperforms the ONM method (solid lines). The improvement stems from a reduced use of the reflection step, which was shown by *Gao and Han* [2012] to be less efficient than the extraction and contraction step for their analysis. This result is also depicted in Fig. 3.3c, where on the left side the relative counts of the five moves of the ONM method are shown, while the relative counts of the five steps of the ANM method are plotted on the right. The 7 columns in Fig. 3.3c correspond to the seven experiments using different perturbations of the initial vertex. For all perturbations except the largest two ($\sigma = 0.2$ and $\sigma =$

0.5), a clear reduction of the reflection counts is visible. For instance, while in the ONM method the minimization of the initial simplex spanned by 5 % perturbations (*a.iv*, red lines) is mainly obtained by reflection steps (98 %), the reflection step for experiment *a.iv* using the ANM method is used only 47 % of the time. For the smallest simplex *a.i*, the enhanced reduction rate of the costs is due to an increased usage of the expansion step (33 %, Fig. 3.3c). For the moderate simplex size in experiment *a.iv*, the improvement is mainly due to an increased usage of the inner contraction step (38 %, Fig. 3.3c).

Nonetheless, there are large differences visible for different choices of the initial perturbation. For the test function $f(\mathbf{x})$ in (3.7), best results were obtained by perturbing the initial vertex by 10 % or 5 % (yellow and red dashed lines in Fig. 3.3a, respectively). The reason for the large improvements in the minimization using the ANM method for the experiments *a.ii* - *a.v* (0.5 % - 10 % perturbation of the initial vertex) is the improved search direction opposed to the ONM method. In Fig. 3.3b, this is illustrated by the angle α between the actual search direction and the optimal search direction, which, for the given test function, is the gradient. It can be seen that the steeper the descend of the costs is, the smaller the angle α is. While the respective angle for the ANM method for the experiments *a.ii* - *a.v* falls below 45° in the iterations with the steepest descend, the respective angle for the ONM remains large (about 70° - 80°). This means that for the ONM method the search direction highly fluctuates in a zigzag course around the gradient. In contrast, in the ANM method in most cases the minimum is searched along steeper search directions, hence performing a much lower fluctuating zigzag course around the gradient. As can be seen by Fig. 3.3b, the iteration, at which the angle between the actual search direction and the gradient begins to reduce, is highly dependent on the initial simplex size. If the initial simplex is too small (e.g. 0.1 % perturbation, gray line in Fig. 3.3b), the respective angle does not reduce continuously. This indicates that the simplex is too small to generate a sequence of optimized search directions as it is in the cases for moderate sized initial simplices (e.g. 5 % and 10 % perturbations, red and yellow dashed lines in Fig. 3.3b, respectively). If the initial simplex is too large (20 % and 50 % perturbation, purple and cyan dashed lines), the angle between the actual search direction and the gradient remains large. Here, the step size is such that the reflection step provides the best updated vertex. As stated above, *Gao and Han* [2012] showed that the reflection step becomes ineffective in high dimensions. The reduction rate for these two experiments remained low leading to a failed minimization even for the ANM method.

To identify a benefit from the ANM method for the optimization of emission profiles of volcanic ash, the same tests with the initial simplices summarized in

Tab. 3.1 are performed using the spherical cost function

$$g(\mathbf{x}) = \sum_{i=0}^{N-1} (x_i - y_i)^2, \text{ with } y_i = \begin{cases} 4, & i < 33, \\ 0, & 33 \leq i < 50, \\ 1, & 50 \leq i < 72, \\ 2, & 72 \leq i < 73, \\ 0, & 73 \leq i < 110, \\ 3, & 110 \leq i < 119. \end{cases} \tag{3.9}$$

This cost function is comparable to the modified cost function in (3.2). For the similarity to the volcanic ash emission cost function see also the derivation of the vector of emission factors in Fig. 3.1. Although the cost function in (3.9) is irregular and non-smooth, it does not suffer from ill-posedness, such that the minimization of this cost function is expected to perform better than for volcanic ash emission estimation using the full atmospheric model.

By performing the minimization of the initial simplices defined in Tab. 3.1 for the cost function (3.9) it was seen that both, the performance of the minimization and the improvement of the ANM method over the ONM method, depend on the distance of the initial simplex to the minimum as well as the size of the initial simplex. For all initial simplices in Tab. 3.1, the percentage of cost reduction after 100,000 iterations for both, the ANM method and the ONM method, is depicted

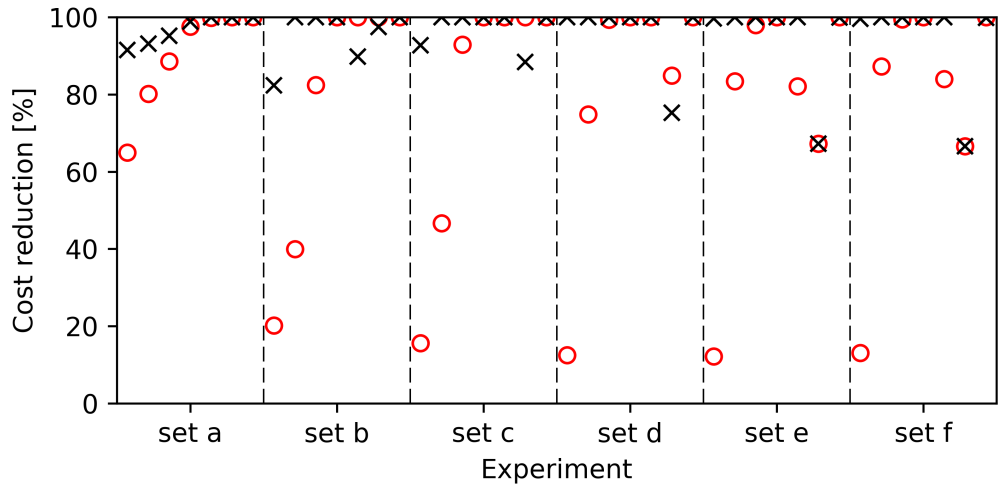

Figure 3.4: *Percentage cost reduction of the ONM (red circles) and ANM (black crosses) methods for the minimization of cost function $g(\mathbf{x})$ after 100,000 iterations using the initial simplices summarized in Tab. 3.1. For each set of initial vertices (e.g. set a corresponds to the experiments a.i - a.vii in Tab. 3.1) the seven different perturbations are shown (perturbation 'i' on the left to perturbation 'vii' on the right for each experiment set). The depicted data is summarized in Tab. 3.2 for convenience.*

in Fig. 3.4. The experiment sets *a-f* correspond to the different initial vertices in Tab. 3.1. Within each experiment set, the small-simplex-experiment (labeled *'i'* in Tab. 3.1) is on the left and the large-simplex-experiment (labeled *'vii'* in Tab. 3.1) is on the right of each column. The results shown in Fig. 3.4 are summarized in Tab. 3.2 for convenience. The tables layout corresponds to Tab. 3.1 for a better comparison.

For all experiments, it is found that the performance of both methods minimizing cost function $g(\mathbf{x})$ is poorer than minimizing cost function $f(\mathbf{x})$ (not shown). For the ONM method, the minimization of the smallest initial simplices with 0.1 % perturbations perform worst with a cost reduction between only 12 % (experiments *d.i* and *e.i*) and 64 % (experiment *a.i*). In contrast, the largest initial simplices with 50 % perturbations (experiments with label '*vii*' in Tab. 3.1) always reach the minimum for this cost function. For the remaining experiments, the cost reduction range from 40 % for experiment *b.ii* to 100 %. In total, a cost reduction of more than 97 %, which can be seen as a successful minimization, is reached for 22 out of 42 experiments using the ONM method. In this respect, the ANM method outperforms the ONM method for this set of experiments as it exceeds 80 % cost reduction for all but three choices of initial simplex (experiments *d.vi*, *e.vi*, and *f.vi*) and reaches more than 97 % cost reduction for 32 out of 42 experiments.

Nonetheless, large differences in the performance of the ANM method between

Table 3.2: *Percentage cost reduction of the ANM method (black) and ONM method (red) after 100,000 iterations for the experiments summarized in Tab. 3.1 for cost function $g(\mathbf{x})$. The initial vertex $\mathbf{x}_0$ is perturbed by σ according to (3.8). The data in the table is plotted in Fig. 3.4 with 'set a' denoting data in the first row, set b denoting data in the second row and all the rest of it.*

	perturbations σ						
$\mathbf{x}_0$	0.001	0.005	0.01	0.05	0.1	0.2	0.5
1	92% (64%)	93% (80%)	95% (89%)	99% (98%)	100% (100%)	100% (100%)	100% (100%)
5	82% (20%)	100% (40%)	100% (82%)	100% (100%)	90% (100%)	97% (100%)	100% (100%)
10	92% (16%)	100% (47%)	100% (93%)	100% (100%)	100% (100%)	88% (100%)	100% (100%)
20	100% (12%)	100% (75%)	100% (99%)	100% (100%)	100% (100%)	75% (85%)	100% (100%)
50	100% (12%)	100% (83%)	100% (98%)	100% (100%)	100% (82%)	67% (67%)	100% (100%)
100	100% (13%)	100% (87%)	100% (99%)	100% (100%)	100% (84%)	67% (67%)	100% (100%)

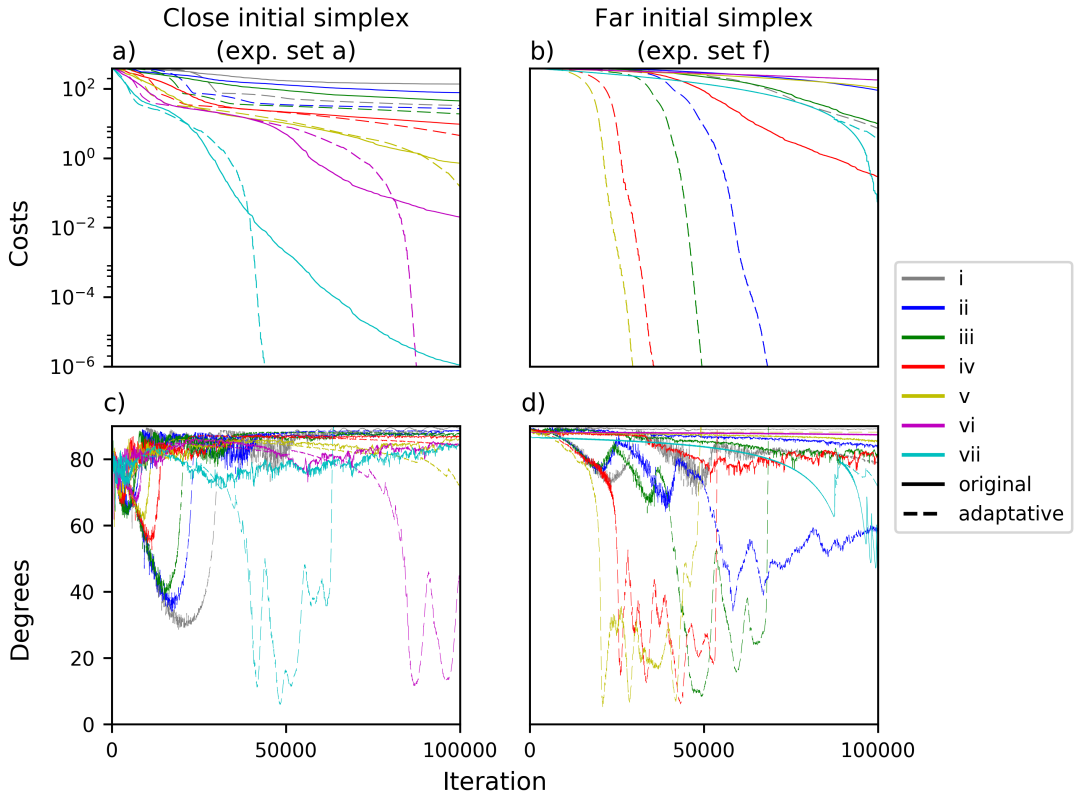

Figure 3.5: *Comparison of the performance for the ONM (solid lines) and the ANM (dashed lines) methods for experiment set a ((a) and (c)) and experiment set f ((b) and (d)) minimizing cost function $g(\mathbf{x})$. The cost evolution is depicted in panels (a) and (b). The angle between the actual search direction and the gradient is displayed in (c) and (d).*

the various initial simplices are visible. To investigate the reasons for this, Fig. 3.5 compares the evolution of the minimal costs over the iterations (Fig. 3.5a and 3.5b) and the angle α between the actual search direction and the gradient (Fig. 3.5c and 3.5d) for all simplex sizes with initial vertex $\mathbf{x}_0 = \mathbf{1}$ (Fig. 3.5a and 3.5c) and $\mathbf{x}_0 = \mathbf{100}$ (Fig. 3.5b and 3.5d). In the following, the experiments with initial vertex $\mathbf{x}_0 = \mathbf{1}$ are referred to as close initial simplices as the distance between the initial vertex and the minimum is small. Similarly, the experiments with initial vertex $\mathbf{x}_0 = \mathbf{100}$ are referred to as far initial simplices.

For the far initial simplex experiments (Fig. 3.5b and 3.5d), the evolution of the minimum costs and the angle α are qualitatively similar to the respective evolutions for cost function $f(\mathbf{x})$ (cf. Fig. 3.3). In contrast, the behavior of the close initial simplices differs during the minimization. Firstly, the differences between the ONM method and the ANM method are smaller for all initial simplex sizes (i.e. for all perturbations of the initial vertex). This is visible as the cost evolution of the two methods starts to deviate only after a certain number of iterations. For instance, the minimum costs for experiment *a.vii* are comparable for the first 40,000 itera-

tions for both methods and improves only afterwards for the ANM method over the ONM method. Secondly, for both methods, it can be seen that the larger the initial simplex is (i.e. the larger the perturbations are) the better is the performance of the minimization (see Fig. 3.5a), which is not true for the far initial simplices experiments (cf. Fig. 3.5b). Thirdly, although the minimum costs for the ANM method are reduced by 90 % to 95 % (cf. also Tab. 3.2), the minimization must be regarded as unsuccessful for the experiments *a.i*, *a.ii*, and *a.iii*. The same holds true for the ONM method, where the minimum costs of the simplices are reduced between 64 % and 89 % for the same experiments.

The reason for the different behavior of the minimization for the experiments shown in Fig. 3.5 can be understood by considering the displacement of the minimum of the cost function $g(\mathbf{x})$ with respect to cost function $f(\mathbf{x})$. While for some dimensions the initial vertex $\mathbf{x}_0 = \mathbf{1}$ meets the minimum, the distance varies between 1 and 3 for the other dimensions. Therefore, the distances of the initial vertex to the minimum for all dimensions is not uniform. By considering the movement of the simplex, which is shown in Fig. 3.2, it is obvious that moving the close initial simplex leads to an improvement in directions that do not match the minimum, while a degradation in dimensions, in which the initial simplex and the minimum coincide, occurs. Therefore, a longer adjustment period of the simplex towards optimal search directions with smaller angles α is needed (see Fig. 3.5a and 3.5c). Because of the large distance of the initial vertex $\mathbf{x}_0 = \mathbf{100}$ to the minimum, this non-uniformity of its distance to the minimum is less than for the close initial simplices. Therefore, the performance is comparable to the performance of the same set of initial simplices for cost function $f(\mathbf{x})$. Nonetheless, taking larger initial simplices with larger perturbations of the initial vertex, the ANM method is also for a displaced cost function an improvement of the minimization method by *Nelder and Mead* [1965] as can be seen by the higher cost reduction rate in Fig. 3.5a and 3.5b.

3.4 Parallel Nelder-Mead method

The results presented in the previous section show the ability of the ANM method to improve the performance of the minimization for a variety of initial simplices. Nonetheless, as is seen for the optimization of the displaced cost functions $g(\mathbf{x})$, the rate of convergence is low and depends on the chosen initial simplex. Therefore, a further improvement of the algorithm is needed. Another approach, proposed by *Lee and Wiswall* [2007], is implemented, which improves the rate of convergence by solving the minimization in parallel. The idea is to update the K worst vertices instead of only the worst vertex per iteration. By updating more than one vertex per iteration, the vertices used to calculate the centroid are expected to yield a better search direction because only the $N + 1 - K$ best vertices are taken into account for the calculation. Therefore, the algorithm minimizes much faster than the ONM method (cf. *Lee and Wiswall* [2007]). The parallel Nelder-Mead method (PNM) of *Lee and Wiswall* [2007] works as follows:

1. The initial simplex is created and the vertices $\mathbf{x}_i$ are sorted according to the individual costs.

2. The centroid of the $N+1-K$ best vertices is calculated

$$\mathbf{c} = \frac{1}{N+1-K} \sum_{i=0}^{N-K} \mathbf{x}_i. \tag{3.10}$$

3. For each of the K worst vertices the individual update is calculated, which is either the reflection ($\mathbf{x}^{(r)}$), expansion ($\mathbf{x}^{(e)}$), or contraction ($\mathbf{x}^{(c)}$) point. These steps are similar to the ONM method (cf. algorithm 1 and section 3.2 for more details).

4. If, for any of the K worst vertices, an update is available, no shrinkage is necessary because the simplex is improved with respect to the costs. If none of the K worst vertices has performed an update, the simplex is shrank towards the current best vertex $\mathbf{x}_0$. Again, this step does not deviate from the ONM method.

5. The next iteration starts with step 2 until the termination criterion is fulfilled.

It is obvious that the ONM method is included in this parallel extension of the method. Therefore, it is expected to perform at least as good as the ONM method.

Opposed to the description by *Lee and Wiswall* [2007], the calculation of the centroid (3.10) differs in the above summary of the method. This is a corrigendum of the PNM algorithm as described below. The main findings by *Lee and Wiswall* [2007] were the reduced number of iterations, since multiple vertices are corrected in one iteration, and the improved search direction as only the best $N+1-K$ vertices are contributing to the calculation of the centroid, as was previously explained. The proposed formula for the centroid by *Lee and Wiswall* [2007] is

$$\mathbf{c}' = \frac{1}{N} \sum_{i=1}^{(N+1)-K} \mathbf{x}_i. \tag{3.11}$$

This formula differs only by a factor from (3.10) but this has a large impact on the minimization performance. Fig. 3.6 demonstrates the behavior of both centroid calculations using (3.10) and (3.11). The initial simplex is defined by the vertices $\mathbf{x}_0$, $\mathbf{x}_1$, and $\mathbf{x}_2$. The different reflection vertices $\mathbf{x}_i^{(r)}$ are depicted, which define the search direction for the worst vertex $\mathbf{x}_2$ for the chosen number of updates K per iteration. $S_{K=2;right}$ and $S_{K=2;false}$ denote the search directions calculated with (3.10) and (3.11), respectively.

As can be seen in Fig. 3.6, the choice of $K = 1$ yields the same search direction (dashed line with vertices $\mathbf{x}_2$ and $\mathbf{x}_{K=1}^{(r)}$) as the ONM method. In contrast, by choosing $K = 2$, the search direction for the worst vertex $\mathbf{x}_2$ strongly deviates by using either of both formulas for the calculation of the centroid. By applying

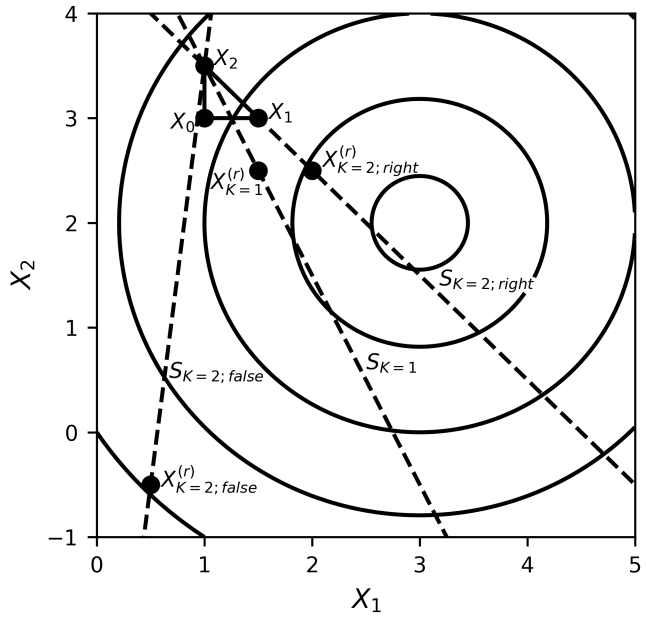

Figure 3.6: *Comparison of different search directions for a shifted spherical cost function (circular isolines) to the minimum coordinate (3,2). The simplex is defined by the vertices* $\mathbf{x}_0$, $\mathbf{x}_1$, *and* $\mathbf{x}_2$. *The reflection* $\mathbf{x}^{(r)}_{K=1}$ *is calculated according to the ONM method. The reflection vertex* $\mathbf{x}^{(r)}_{K=2;right}$ *is the corrected reflection vertex referred to Lee and Wiswall [2007] using (3.10). The reflection vertex* $\mathbf{x}^{(r)}_{K=2;false}$ *corresponds to the calculation from Lee and Wiswall [2007] using (3.11). The dashed lines represent the different search directions* S *with respect to the calculation of the centroid.*

(3.11) as proposed by *Lee and Wiswall* [2007], the search direction $S_{2;false}$ is highly nudged towards the origin of the coordinate system. For the cost functions used in the evaluation by *Lee and Wiswall* [2007], this is not an issue because the cost functions are all centered at the origin. Therefore, the methods performance is even increased by this formula. By applying the corrected formula for the calculation of the centroid (3.10), the search direction $S_{2;right}$ is nudged towards the minimum of the cost function, which illustrates well the potential of the PNM method to increase the rate of convergence.

A further extension of this PNM method subdivides the simplex into P subsimplices, which is beneficial especially if the dimension and the number of observations is large (*Klein and Neira* [2014]) and, hence, many and costly function evaluations are necessary. Each subsimplex comprises $N_{loc} = (N + 1)/P$ local vertices with $K_{loc} = K/P$ local updates. In the following, local variables refers to values, which are specific for each subsimplex. In contrast, global variables are equal for all subsimplices. The implementation follows mainly the procedure described above with

the following additions (cf. *Klein and Neira* [2014]):

1. The subsimplices are allocated on different processors;
2. A local centroid of each subsimplex j is calculated by

$$\tilde{\mathbf{c}}_j = \frac{1}{N+1-K} \sum_{i=0}^{N_{loc}-K_{loc}} \mathbf{x}_i; \tag{3.12}$$

3. The global centroid is calculated as the sum of the local centroids ($\mathbf{c} = \sum_{j=0}^{P-1} \tilde{\mathbf{c}}_j$);
4. The global minimum is determined and its corresponding vertex is distributed among the processors;
5. A flag is introduced that notifies if any vertex has been updated among the processors. If not, the subsimplices are shrank towards the best vertex with the global minimum costs.

With this implementation of the PNM method, the K vertices with highest costs are not equally distributed among the subsimplices. Hence, vertices with smaller costs than those K vertices will be updated. In contrast, some of the K vertices with the highest costs are used for the calculation of the centroid. This may lead to search directions that are less optimal compared to the PNM method by *Lee and Wiswall* [2007]. This will likely increase the number of iterations before the minimum is found. Nonetheless, *Klein and Neira* [2014] found that the above described distribution of the vertices among the subsimplices has only little effect on the algorithms performance, at least for their test functions.

The PNM method is tested for the cost function $g(\mathbf{x})$ defined by (3.9) and the set of initial simplices as described in the previous section (cf. Tab. 3.1). For each initial simplex, the minimization of cost function $g(\mathbf{x})$ is calculated for K = 1,2,5,10,20,40, and 80. However, in this analysis it is chosen that $P = 1$. The tests were conducted for both, using the PNM method with the standard set of model parameters defined by (3.5) and the set of adaptive parameters defined by (3.6). In the following, the former settings are referred to as S-PNM method and the latter settings are referred to as A-PNM method. Thus, in the A-PNM method the Nelder-Mead extensions proposed by *Gao and Han* [2012] and *Lee and Wiswall* [2007] are combined for the first time to the authors knowledge. This combination is expected to be most efficient. To analyze the results of these various tests, the mean and standard deviation of the final reduction of the minimum costs for all experiments for the different K-values are displayed in Fig. 3.7. Herein, the mean and standard deviation for both, the A-PNM and the S-PNM method, is calculated over 42 experiments. For the seven different K-values, this results in a total of 294 experiments. Out of this 294 experiments, using the S-PNM method only 64 reached minimum costs $f(\mathbf{x}_0) < 10^{-6}$, which is defined as a criterion for a successful

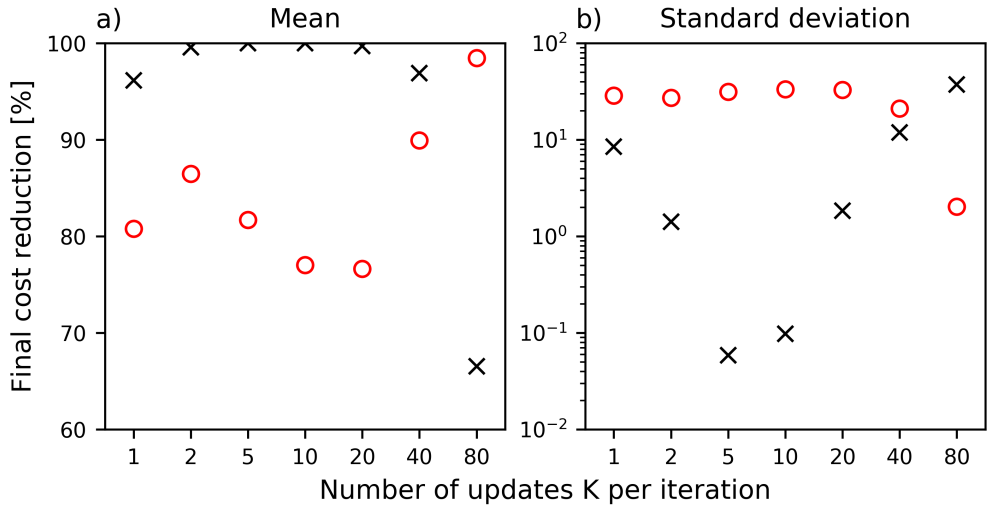

Figure 3.7: *Comparison of the mean (a) and standard deviation (b) of the final cost reduction for the S-PNM method (red circles) and A-PNM method (black crosses). The mean is calculated over the 42 experiments given by Tab 3.1 for seven different numbers of updates per iteration (K = 1, 2, 5, 10, 20, 40 and 80).*

minimization. In contrast, for the A-PNM method only 80 experiments did not reach this threshold. This is even more remarkable by taking into account that all experiments using $K = 80$ updates per iteration (42 experiments) fail in finding the minimum for both methods.

For the A-PNM method, this fact is clearly pointed out by Fig. 3.7. For $K = 80$ updates per iteration, the mean reduction of the minimum costs at the end of the minimization process amounts only $\sim$ 67 % with a standard deviation of $\sim$ 24 %. For all other K-values, the A-PNM method reaches a mean final reduction of the minimum costs of more than 95 %. The best performance of the A-PNM method is obtained by moderate K-values of 5 and 10 (approx. 10 % of the dimension size N), where all but 2 experiments are successful in minimizing the cost function. In contrast, the experiments using the S-PNM method show a slightly different behavior with respect to the K-values. Best results are obtained by using $K = 80$, considering the mean and standard deviation of the final cost reduction. Although no experiment finds the minimum for these experiments, all show a similar cost reduction of 98 %, hence, the standard deviation is small ($\sim$ 1 %). All other mean values range between 77 % and 90 % final cost reduction with a relatively high standard deviation of approx. 30 %. This indicates that for some experiments the S-PNM method is successful but for the majority of the experiments it fails to find the minimum. Therefore, the A-PNM method outperforms the S-PNM method in these experiments as it reaches in most cases the threshold of 10^{-6} for the minimum costs. In order to investigate the dependency of the performance on the number of updates K per iteration, Fig. 3.8 shows the relative frequency of the five steps of the

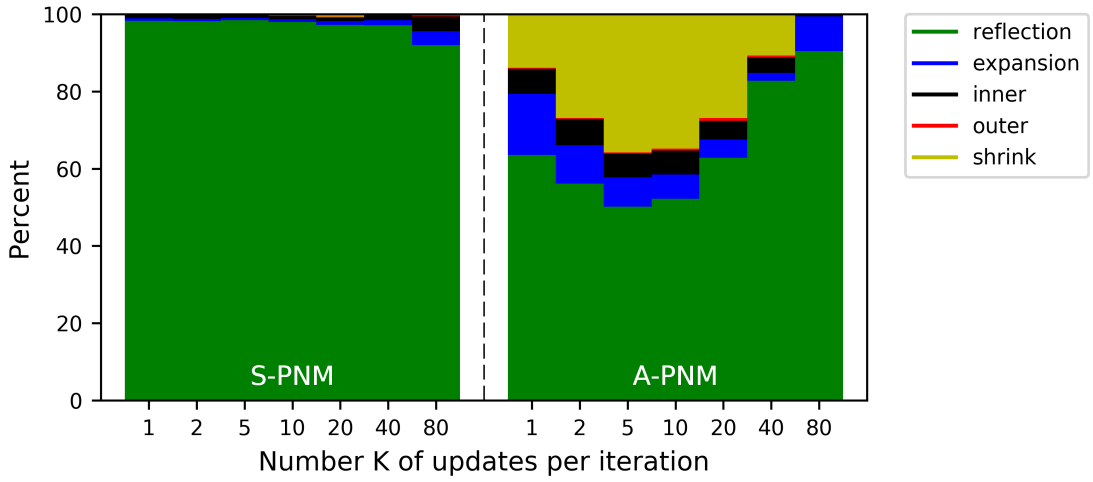

Figure 3.8: *Relative frequency of calls to the minimization steps of the Nelder-Mead minimization algorithm for cost function $g(\mathbf{x})$ for the S-PNM method (left) and the A-PNM method (right) for different number of updates K per iteration.*

Nelder-Mead algorithm for both, the S-PNM (left) and A-PNM (right) method. As stated above, *Gao and Han* [2012] found that decreasing the frequency of the reflection step improves the performance of the minimization. This finding is supported by the results of Fig. 3.8. For the K-values, which show a bad performance with respect to the mean final cost reduction, the frequency of performing a reflection step is above 90 %. This is the case for all K-values using the S-PNM method and for $K = 80$ using the A-PNM method. In contrast to the ANM method, where the reflection step is mainly replaced by expansion and contraction steps (cf. Fig. 3.3c), for the A-PNM method the reflection step is mainly replaced by the shrinkage step, especially for the K-values with the best performance. Hence, the benefits from the Nelder-Mead method with adaptive parameters and the parallel Nelder-Mead method are combined in the new A-PNM method.

To show the stability of the A-PNM method, the same tests are performed for cost function

$$h(\mathbf{x}) = \sum_{i=0}^{N-1} (x_i - y_i))^2, \tag{3.13}$$

where y_i is a randomly chosen integer between 0 and 100. This test shows the performance of the A-PNM method to find the minimum even if the distance between the initial simplex and the minimum is not uniform. For cost function $g(\mathbf{x})$, these are the experiments in which the ANM method is unsuccessful to find the minimum (see section 3.3). The mean and standard deviation of the final cost reduction are summarized in Fig. 3.9a and 3.9b, respectively.

The results are comparable with the analysis of cost function $g(\mathbf{x})$. For all but $K = 80$, the A-PNM method outperforms the S-PNM method in terms of mean final

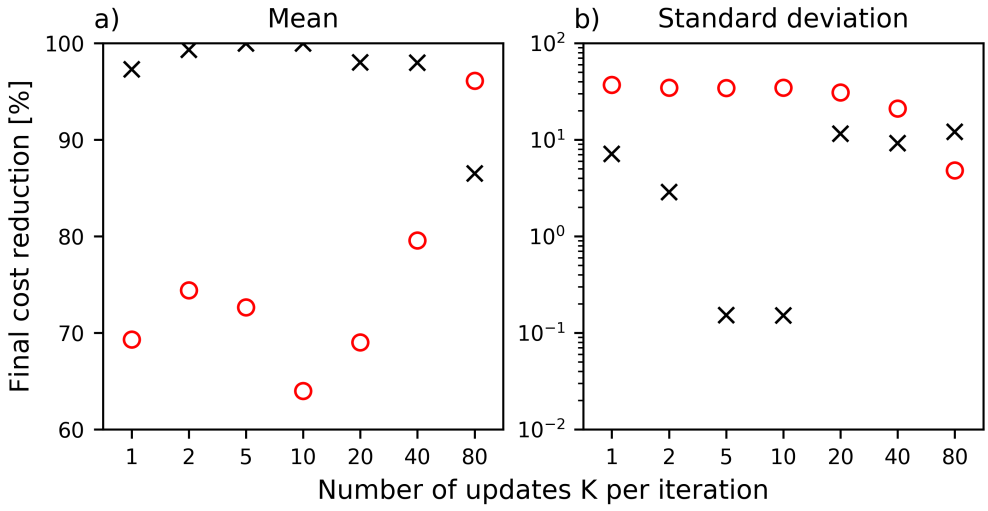

Figure 3.9: *Same as Fig. 3.7 but for cost function $h(\mathbf{x})$ in (3.13).*

cost reduction. Again 294 experiments are performed, out of which only four are successful in minimizing the cost function using the S-PNM method. In contrast, using the A-PNM method 211 experiments are successful in minimizing the cost function $h(\mathbf{x})$. For the S-PNM method, the mean final cost reduction ranges from $\sim$ 64 % for $K = 10$ to $\sim$ 80 % for $K = 40$ and reaches $\sim$ 96 % for $K = 80$, while the standard deviation is about $\sim$ 35 % for $K < 40$ and drops to $\sim$ 6 % for $K = 80$. For the A-PNM method, all mean final cost reduction lays beyond $\sim$ 97 % except for $K = 80$ with a mean value of $\sim$ 87 %. Again, the standard deviation is of the order of 10 and drops for moderate sized K-values of $5 - 10$ to approx. 10^{-1}. Hence, no degradation of the performance of the A-PNM method occurs by changing the cost function arbitrarily.

To conclude, the A-PNM method is successful in minimizing the cost function $h(\mathbf{x})$ for most initial simplices for moderate K-values. Nonetheless, the performance of the A-PNM method depends on the K-value chosen and for approx. 1/3 of the experiments the minimum is not found. Therefore, an extension of the A-PNM method is developed, which is introduced in the next section.

3.5 Discrete Ensemble Nelder-Mead method

One drawback of the Nelder-Mead method indicated by the performance analysis in the previous sections is that by choosing an initial simplex, it is not certain that the minimum is found. Anyway, the probability to find the minimum of the analyzed cost functions using the A-PNM method is higher for some initial simplices and K-values as summarized at the end of the previous section. For the application to uncertainty estimation of time and height dependent volcanic ash emission strengths, which is the focus of this study, the following extension is applied. In-

stead of performing the minimization with one initial simplex, an ensemble of initial simplices is chosen to minimize the cost function. For the uncertainty assessment, the outcome of this ensemble can be used as input for the particle smoother as will be described in section 5.2.

The main focus of an ensemble prediction system is to estimate not the minimum itself but the vicinity of the minimum and, therefore, its uncertainty. Using this fact, for volcanic ash emission estimation only discrete solutions are considered as vertices for the minimization steps, thus the method is called Discrete Ensemble Nelder-Mead (DENM) algorithm. The use of only discrete solutions in the DENM method increases the ability of the algorithm to come fast into the vicinity of the minimum but not limiting the performance of the minimization. As the A-PNM method, the DENM algorithm performs first a reflection or expansion step followed by a sequence of contraction and shrinkage steps. This is in line with a finding by *Gao and Han* [2012], who proofed for uniformly convex functions that the contraction step is also a "sufficient decent property" for the Nelder-Mead algorithm if the diameter of the simplex, which is proportional to its determinant, is not too small. The initial fast convergence rate to the minimum of the simplex is also supported by the fact that the DENM method generally may perform larger steps towards the minimum than the ONM method. This is because the minimum step size in one dimension is controlled by the resolution of the discrete emission strengths of, say 25 kg/s. Therefore, the minimization will not be trapped in increasingly small update steps as the size of the simplex reduces. In contrast, since there are only discrete vertices available as new estimates of the minimum, the method is likely to collapse on a single vertex. Therefore, the algorithm is extended by a restart step once the vertices are collapsed. This was also suggested by *Gao and Han* [2012], who argued that, since in larger simplices the expansion and contraction step have an increased reduction rate, a restart can improve the performance of the Nelder-Mead algorithm in large dimensions.

The DENM method deviates in only a few steps from the A-PNM method. This is depicted in Fig. 3.10, which displays a flowchart of the main steps of the algorithm. The respective parts from the different versions of the Nelder-Mead method are given in different colors, where black indicates the ONM method proposed by *Nelder and Mead* [1965], blue views the ANM method extensions proposed by *Gao and Han* [2012], and green highlights the PNM method extensions by *Lee and Wiswall* [2007]. The red parts in Fig. 3.10 indicate the newly introduced parts for the DENM method. The first newly included feature is the ensemble of N_{ens} initial simplices. As it was discussed in the previous section, the A-PNM method is able to find the minimum in many cases but there is no guarantee that the minimization succeeds. Therefore, the ensemble of initial simplices aims to overcome this issue. Additionally, as the assimilation of volcanic ash column observations is an ill-posed problem, the ensemble of initial simplices aims to find the N_{ens} best estimates of the solution. This will be needed by the particle smoother described in section 5.2.

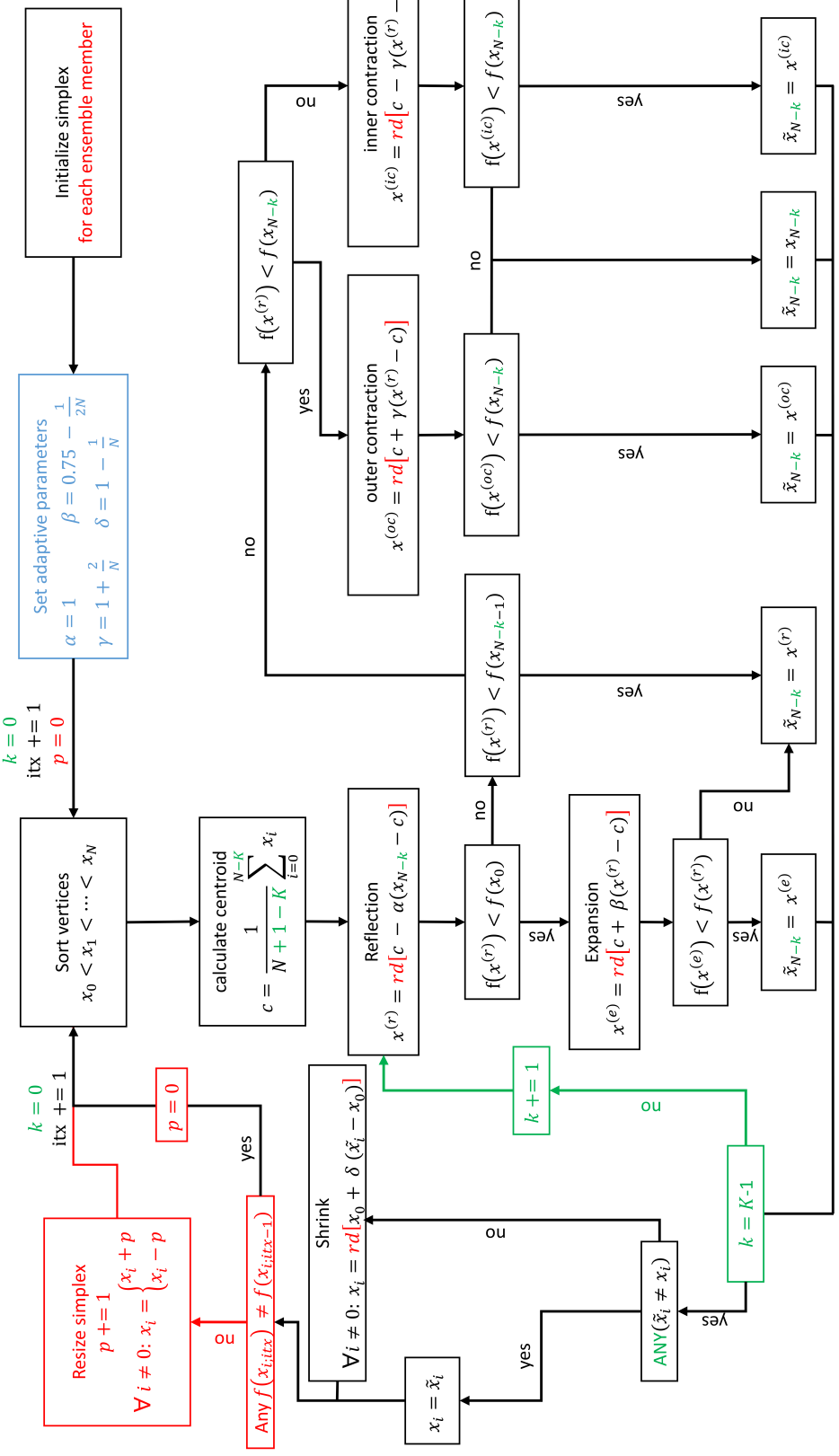

Figure 3.10: *Flowchart of the DENM method including adaptive model parameters (blue colored) and multiple updates per iteration (green colored). Additional extensions for the DENM method are colored red. See text for a further discussion.*

By applying the adaptive parameters α, β, γ, and δ from (3.6), the minimization is started for each ensemble member independently. Therefore, the following description concerns only one ensemble member. At first the centroid of the $N + 1 - K$ vertices with the lowest costs is calculated. Afterwards, the reflection vertex is calculated and, according to the costs of the reflection vertex $f(\mathbf{x}^{(r)})$, the expansion, inner, or outer contraction is computed as is described before. The new vertices $\mathbf{x}^{(r)}$, $\mathbf{x}^{(e)}$, $\mathbf{x}^{(ic)}$, and $\mathbf{x}^{(oc)}$ are rounded to the discrete value through function $rd[\cdot]$

$$\tilde{x}_i = rd[x_i] = \begin{cases} \lfloor x_i \rfloor, & \text{if } r \sim \mathrm{U}[0,1] \leq 0.5, \\ \lceil x_i \rceil, & \text{if } r \sim \mathrm{U}[0,1] > 0.5, \end{cases} \tag{3.14}$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denotes the floor and ceiling function, respectively, and $\mathrm{U}[0,1]$ is a uniform distribution on the interval [0,1]. The variable r is drawn randomly from an uniform distribution. This rounding procedures give probability to escape from a vertex even if the new vertices are closer to the starting vertex (e.g. if $\langle \mathbf{x}_{N-k}, \mathbf{x}^{(r)} \rangle < \langle rd[\mathbf{x}^{(r)}], \mathbf{x}^{(r)} \rangle$, $\langle \cdot \rangle$ denoting the dot product). Unfortunately, it is possible that the rounding function leads to higher costs than the actual vertex as it may round unprofitably.

If either of the above mentioned steps improves the $N - k$-th vertex, it is updated. This cycle starting with the reflection step is repeated K times. Afterwards, the algorithm checks if any of the K vertices with highest costs was updated during the iteration. If at least one vertex is improved with respect to the previous iteration, the next iteration is performed with the updated vertices (either $\mathbf{x}^{(r)}$, $\mathbf{x}^{(e)}$, $\mathbf{x}^{(c)}$, or $\mathbf{x}_{N-k}$ if there was no better approximation to the minimum found for this vertex). If no vertex was updated during the last iteration, all vertices except the best are shrank towards the best vertex $\mathbf{x}_0$. As only discrete solutions are a valid choice for the vertices, it is possible that the simplex shrinks towards a single vertex. If this occurs, the algorithm restarts the minimization by resizing the simplex through perturbations of the best vertex. Each vertex is perturbed by p in p_t dimensions. It is found that the best results are obtained if the perturbations p and the number of dimensions p_t to be perturbed increase if no better approximation to the minimum is found. If the simplex shrinks towards the same point, p is increased, otherwise it is set to 1. If p reaches a certain value $< N$ (e.g. $N/2$), i.e. if the perturbations are too large and no new best estimate is found, p is set to 1 and the minimization restarts by increasing p_t by 1. It is assumed that if the minimization is trapped in a local minimum there exist a simplex size (i.e. a perturbation p for a number of dimensions p_t) that ensures the algorithm to escape from being trapped. In contrast, if the minimization reaches the global minimum (or at least the closest point on the discrete grid) it will shrink towards this global minimum, no matter how large the simplex size is. For volcanic ash emission estimation, this expectation is relaxed. Considering that there are correlations between volcanic ash concentrations released by different emission packages as they where defined in section 3.1, the algorithm may be adapting the emission factor for the wrong emissions, which is correlated with the right ones. Therefore, the required step size for restarting the simplex may

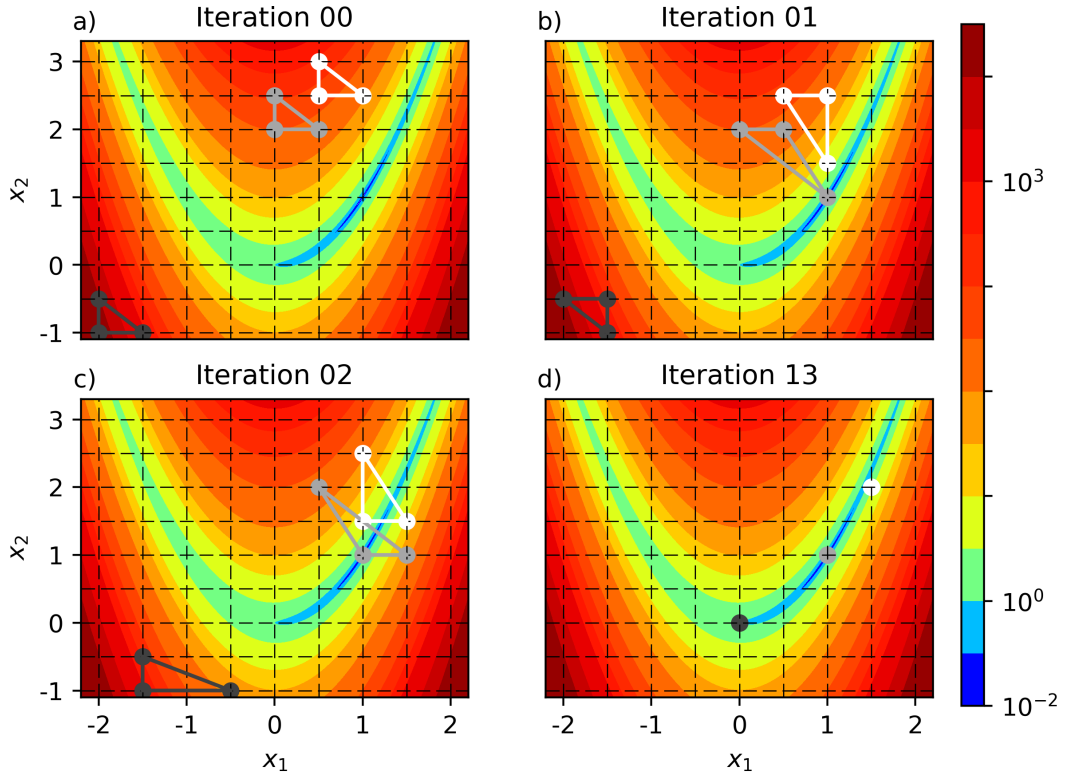

Figure 3.11: *Illustration of the DENM method using the two-dimensional Rosenbrock standard test function. The dashed grid shows the allowed points for vertices. The ensemble minimization is performed using an ensemble of size $N_{ens} = 3$ (indicated by different colors) and 13 iterations. The bullets depict vertices while each triangle is one simplex. Shown are (a) the initial simplices, (b) the simplices after iteration 1, (c) after iteration 2, and d) after final iteration 13.*

be large and leads to a better solution only in one direction. For the performance analysis of DENM method below, this correlation within the solution is not concerned but will be recalled in section 6.2. A drawback of restarting the simplex is the increased number of function evaluations, as for each restart N function evaluations have to be performed. Generally, in the application to volcanic ash emission estimation, the function evaluations are computationally cheap, which makes the restart within the DENM method feasible.

The idea of the DENM is illustrated by the Rosenbrock test function

$$f(\mathbf{x}) = (1 - x_1)^2 + 100 * (x_2 - x_1^2)^2 \tag{3.15}$$

in Fig. 3.11. The Rosenbrock test function is displayed in colors, blue indicating low values. In dashed lines the grid for the discrete solutions of the DENM method is overlaid. In the beginning, an ensemble of initial simplices is chosen (see Fig. 3.11a). Each simplex now performs the minimization as is illustrated by Fig. 3.11b and 3.11c. While in the first iterations (Fig. 3.11b) the white and light gray simplices perform

an extension step, the dark gray simplex is reduced by the rounding function $rd[\cdot]$ to the reflection vertex, although it actually performs an extension step, too. Because the restart step is omitted in this illustrative example, after 13 iterations all three simplices are reduced to their final vertex. A restart would lead all three simplices to the point (1,1) if and only if the restart by chance selects the global minimum vertex as for all other points the vertex to be updated has higher costs. Nonetheless, all three simplices come close to the global minimum and thus representing its uncertainty. It is noted that for ill-posed problems, as the assimilation of volcanic ash column observations, there are at least multiple physical solutions. Therefore, it is highly unlikely for all ensemble members to find the true solution (i.e. the true emission profile). The minimization is only expected to find the true solution if the vertical wind shear and the horizontal wind speed are large enough such that the volcanic ash emitted by different emission packages is well separated.

In the following, the rounding function $rd\,[\cdot]$ rounds to integers. The DENM method is tested with the cost function $g(\mathbf{x})$ defined by (3.9). The initial simplex is uniform in the way that the initial vertex $\mathbf{x}_0$ is constant in all dimensions (i.e. $x_0^i = \mathbf{c};\ i = 1, 2, \ldots, N; \mathbf{c}$ constant). Similarly, the perturbations are of the same magnitude for all remaining vertices but different for different ensemble members. Hence, this test is in line with the tests made in the previous sections, although the perturbations are larger in this case, as only integers are accepted as vertex values for each dimension (i.e. $x_j^i \in \mathbb{N}$). The results of the ensemble of initial simplices for different numbers of updates K per iteration are summarized in Fig. 3.12.

The boxplots in all graphics show the median as an orange line, the box indicates the boundaries of the 25 % - 75 % interval, while the error bars span approximately the 99 % interval of data points. It is noted that all ensemble members found the minimum of cost function $g(\mathbf{x})$. Compared to the A-PNM algorithm described above, this is a remarkably stable performance because the performance for the A-PNM method depends on the initial simplex. For the DENM method, no ensemble member needs more than 40,000 iterations for any number of updates per iteration. The median iterations performed before the minimum is found decreases from 10,364 iterations for $K = 1$ to 412 iterations for $K = 80$, which is a reduction of 96 %. Additionally, the spread of the ensemble members (excluding outliers) for the number of iterations decreases as the number of updates per iteration increases from the range 548 - 32,995 iterations for $K = 1$ to only 197 - 1,004 iterations for $K = 80$. Nonetheless, this reduction in the number of iterations needed to find the minimum of the ensemble median is not visible for the computing time (Fig. 3.12b). The reason is twofold. Firstly, by updating $K = 80$ vertices per iteration, it takes roughly 80 times more computing time per iteration than for $K = 1$ as in both cases the steps of the minimizer are the same except for the calculation of the centroid. This is calculated 80 times for 80 updates with $K = 1$ but only once for $K = 80$. Secondly, from Fig. 3.12d it can be seen, that the number of restarts (i.e. resizing the simplex if it has collapsed to a single point) and the number of function evaluations (Fig. 3.12c) do not reduce as much as the number of iterations in Fig. 3.12a. As the

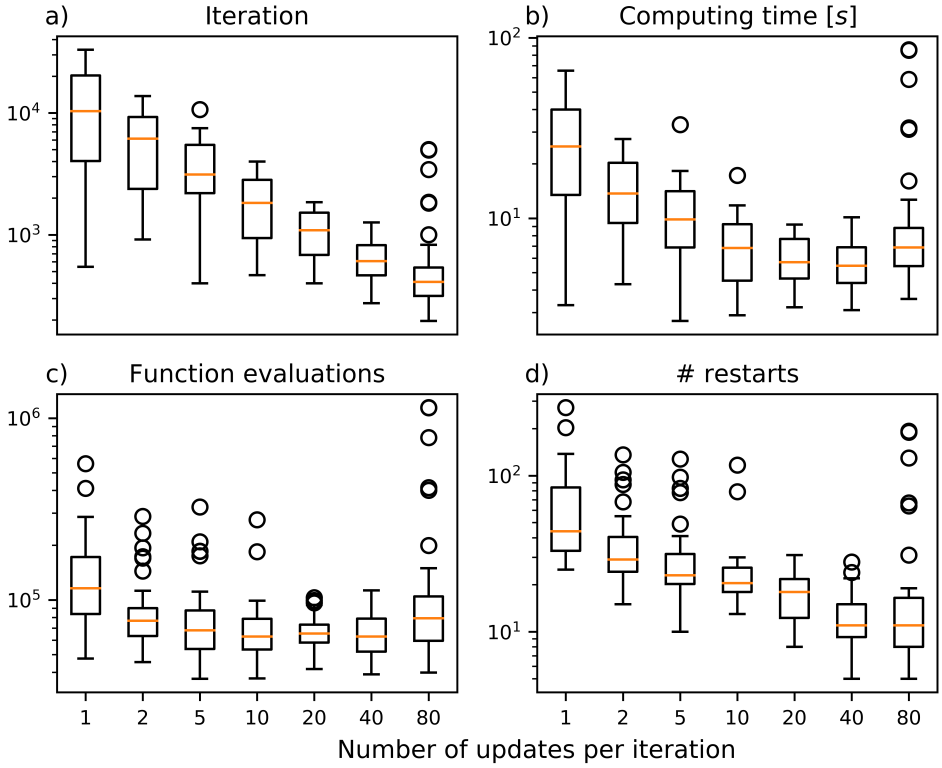

Figure 3.12: *Performance analysis of the DENM method for an ensemble of uniform initial simplices for different numbers of updates per iteration K for cost function $g(\mathbf{x})$. The box plots are defined by the orange line indicating the median, the box boundaries depicting the 25 % and 75 % percentile of the data, and the upper and lower error bar denoting approx. the 99 % percentile interval. Circles indicate outliers. (a) Number of iterations needed to find the minimum. (b) Computing time needed for the minimization. (c) Number of function evaluations. (d) Number of restarts as the simplex converges to a non-minimizing point.*

restart is computationally expensive, it equals the computing time for evaluations using different K-values.

The number of function evaluations does not change much by changing the number of updates per iteration (Fig. 3.12c). The median for the ensemble is of $\mathcal{O}(10^5)$ function evaluations. Especially for $K = 80$ outliers are visible with more than 10^6 function evaluations. These outliers correspond with the outliers in the number of restarts in Fig. 3.12d. Although the number of restarts are slightly higher for $K = 1$ than for $K = 80$, this is not visible in the number of function evaluations. The reason is again that in each iteration a different number of function evaluations is performed for different values of K, which leads to higher numbers of function evaluations for large K-values. It is considered that for $K = 80$ in each iteration a minimum number of 80 function evaluations have to be performed. For a restart, in these experiments 120 function evaluations need to be performed. Therefore, the

higher number of restarts for $K = 1$ does not increase the number of function evaluations in comparison to one iteration for $K = 80$. Hence, the number of function evaluations for this test case is almost equal.

The results presented above are also obtained by two different experiments, one with a randomly chosen initial simplex regarding its position and the perturbations and one with an additionally random displaced cost function (i.e. randomly chosen parameter y_i in (3.9) for all dimensions i). This leads to the conclusion that the higher the number of updates K per iteration is, the faster (and, therefore, computationally cheaper) is the minimization. For the performed tests, a critical number of updates per iteration of 20 - 40 is found after which the decrease in computing time stopped. If the dimension of the problem increases, it is expected that the critical number of updates per iteration will also increase. Nonetheless, the presented results indicate that even if the parameter K is not properly chosen, the minimization will lead to the correct solution although it may take more time and more iterations.

For an application to the minimization of the ill-posed problem induced by observations of column mass loadings of volcanic ash in the atmosphere, it is expected that the minimization will not lead to the absolute minimum but to local attractors close to the global minimum. This is due to the multiple solutions if the volcanic ash concentrations emitted by different emission packages are correlated such that they will lead to approximately similar column mass loadings. As the time difference between the release of volcanic ash and its observation increases, the vertical wind shear will likely separate the volcanic ash such that a better resolution for the analysis emission profile is expected. This will be analyzed in section 6.2. Additionally, the choice of discrete solutions of the minimization also prevents the algorithm to find the exact minimum. As mentioned above, this is not an issue because the ensemble system aims to find the uncertainty of the minimum rather than the exact value.

Chapter 4

Model description

For the uncertainty estimation of volcanic ash concentrations, the EURopean Air pollution Dispersion - Inverse Model (EURAD-IM) is used. The EURAD-IM comprises the adjoint of the chemistry and aerosol modules as well as of the transport modules (advection and diffusion). The adjoint modules are needed for initial value optimization as well as emission factor optimization in the framework of 4D-var. It was the first 4D-var system for atmospheric chemistry established in the mid 1990s (e.g. *Elbern et al.* [1997; 2000; 2007]). However, the work presented here uses only the so called forward run (forward time integration) because the adjoint code is not required by the analysis model used here (cf. section 5.2 for more details about this model).

The need for uncertainty assessment necessitates the extension of the EURAD-IM to an ensemble system. Therefore, a computer system is required that satisfies the demanding computational requirements for this ensemble system. Hence, the simulations were operated on the BlueGene/Q system (JUQUEEN) at the Jülich Supercomputing Centre. The meteorological fields needed by the EURAD-IM are calculated offline using the Weather Research and Forecasting model (WRF; *Skamarock et al.* [2008]). In the upcoming sections, the various modules of the model, which are essential for this analysis, are introduced and required changes to the system are shown.

4.1 EURAD-IM

The EURAD-IM is an Eulerian non-hydrostatic Chemistry-Transport-Model (CTM) solving the advection-diffusion-reaction equation

$$\frac{\partial c_i}{\partial t} = -\nabla(vc_i) + \nabla(\rho K \nabla \frac{c_i}{\rho}) + A_i + E_i - \frac{\partial}{\partial z}(v_i^d c_i), \tag{4.1}$$

for the concentrations of up to 109 gaseous species and 39 aerosols on a fixed grid. In (4.1) the variable c_i denotes the concentration of species i, v and v_i^d the wind and deposition velocity of species i, respectively, ρ the air density, K the eddy diffusivity tensor, A_i the chemical transformation, and E_i the emission of species i. The terms

Table 4.1: *Initial median diameter $D_{g,i}^{init}$ and standard deviations $\sigma_{g,i}$ of the three modes of the aerosol distribution in MADE.*

	$D_{g,i}^{init}$ $[\mu m]$	$\sigma_{g,i}$ $[\mu m]$
Aitken-mode	0.01	1.7
accumulation mode	0.07	2.0
coarse mode	1.0	2.2

on the right hand side of (4.1) represent advection, diffusion, chemical reactions, emissions, and deposition. As the operators for advection (hyperbolic) and diffusion (parabolic) are of different behavior, the transport is calculated using the operator splitting technique (e.g. *Yanenko* [1971]; *McRae et al.* [1982]). Hence, the discrete time integration of concentration c_i reads

$$c_i(t + \Delta t) = T_h T_v D_v M C D_v T_v T_h c_i(t), \tag{4.2}$$

with T and D denoting transport by advection and diffusion, respectively, in the horizontal (h) and vertical (v) direction, M representing the chemical transformation of aerosols by the Modal Aerosol Dynamics module for Europe (MADE) (*Ackermann* [1997]; *Ackermann et al.* [1998]) and C denoting the chemical transformation of gaseous species.

The meteorological conditions are supplied by WRF. For the advection, the monotone scheme by *Walcek* [2000] is used. The need for a monotone advection scheme stems from the purpose of volcanic ash clouds as they have sharp boundaries. Non-monotone advection schemes would lead to wiggles at the clouds boundaries, also known as Gibbs' phenomenon, entering unrealistic volcanic ash patterns or negative values into the simulation. The purpose of the advection scheme by *Walcek* [2000] to adjust the flux around local extrema reduces the numerical diffusion and preserves the positive definiteness.

The vertical coordinate is the terrain following sigma coordinate system

$$\sigma_k = \frac{p_k - p_{top}}{p_{bot} - p_{top}}, \tag{4.3}$$

where p_k is the pressure in model layer k, p_{top} is the pressure at the models top boundary, defined at 100 hPa, and p_{bot} is the pressure at the models surface. Hence, the vertical coordinate ranges from $\sigma_{top} = 0$ at the model top boundary at 100 hPa to $\sigma_{bot} = 1$ at the surface.

Aerosol properties such as size distributions and the physical and chemical transformations are calculated in MADE. The aerosols are represented by three log-normal distributions for the Aitken-, accumulation, and coarse modes, which have a fixed standard deviation $\sigma_{g,i}$ and an initial diameter $D_{g,i}^{init}$. The respective values for these variables are summarized in Tab. 4.1.

The initial median particle diameter serves also as minimum diameter for the

specific distribution. The evolution of the median particle diameter is calculated by

$$D_{g,i} = \sqrt[3]{\frac{M_{3,i}}{M_{0,i}} \exp\left(-\frac{9}{2}\ln^2 \sigma_{g,i}\right)}, \tag{4.4}$$

with i and g indexing the mode and aerosol species, respectively, and $M_{0,i}$ and $M_{3,i}$ representing the zeroth and third moment of the particle distributions, defined by the particle diameters D_p

$$\begin{aligned} M_{0,i} &= \int_0^\infty D_p^0 n_i(D_p) dD_p = N_i, \\ M_{3,i} &= \int_0^\infty D_p^3 n_i(D_p) dD_p = \frac{6}{\pi} V_i = \frac{6}{\pi} \sum_g \frac{c_{g,i}}{\rho_g}. \end{aligned} \tag{4.5}$$

In (4.5), N_i defines the number concentration in $[1/\mathrm{m}^3]$, V_i the volume concentration in $[\mathrm{m}^3/\mathrm{m}^3]$, and $c_{g,i}$ and ρ_g the mass concentration and the density of species g, respectively. For mode i, the log-normal distribution is defined by

$$n_i(\ln D_p) = \frac{N_i}{\sqrt{2\pi}\ln \sigma_{g,i}} \exp\left[-\frac{(\ln D_p - \ln D_{g,i})^2}{2\ln^2 \sigma_{g,i}}\right]. \tag{4.6}$$

The aerosol dynamics include nucleation, condensation, coagulation, deposition (dry and wet) and sedimentation. The chemical transformation of the aerosols is coupled with the gas phase chemistry module of EURAD-IM. The volcanic ash particles are only represented in the coarse mode. It is chemically inactive such that only transport and aerosol dynamics are relevant for the transformation of the median diameter of the distribution.

4.2 Ensemble EURAD-IM

For the analysis, EURAD-IM is extended to an ensemble system. In this context, the main changes concerned the model's parallelization using Message Passing Interface (MPI)[1]. The new parallelization of the ensemble system is illustrated in Fig. 4.1. On the left of Fig. 4.1, the parallelization of the original version of the EURAD-IM is shown. There is only one communicator MPI_COMM_WORLD with a subdivision of the MPI-processes into master and worker. For the ensemble EURAD-IM (ensEURAD-IM), the subdivision in master and worker is omitted and all processes act as workers. This has two main advantages. Firstly, all processes are involved in the simulation, which directly improves the performance of the model. Secondly, no process has to store the variables on the whole domain for Input and Output (I/O). Therefore, the model is more scalable and flexible in the application to larger domains. As the model domain increases, the total number of MPI-processes can be increased as well such that each process remains with the

[1] cf. *http://mpi-forum.org/docs/docs.html* for an overview

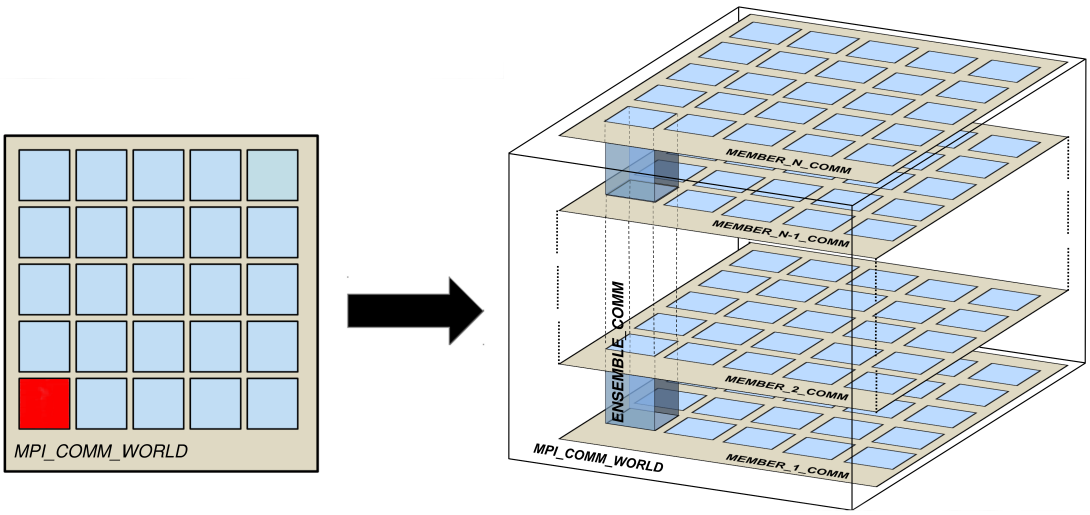

Figure 4.1: *Schematic of the parallelization of the ensEURAD-IM. The classical parallelization of the EURAD-IM with a subdivision of the MPI-processes in master (red) and worker (light blue) is illustrated on the left. The parallelization and different MPI-communicators of the ensEURAD-IM are illustrated on the right. Each gray layer indicate a single ensemble member.*

same amount of storage requirements as for smaller domains.

On the right of Fig. 4.1, the parallelization of the ensEURAD-IM is shown. Beside the change in the process hierarchy mentioned above, several new MPI-communicators are implemented. The ensemble is divided through the MPI-communicators MEMBER_1_COMM - MEMBER_N_COMM, which represent one ensemble member each. These ensemble member communicators are similar to the communicator of the classical EURAD-IM except for the transformation from the Master/worker to an all-worker parallel strategy. Another MPI-communicator is implemented for the communication of MPI-processes of the same rank of the ensemble member communicator. This ensures a most efficient communication between different ensemble members. To make the minimization more efficient, each ensemble member comprises a set of sub-ensemble members illustrated by the columns in each ensemble member communicator. The simplex for the minimization is subdivided such that each sub-ensemble member has to perform only a reduced number of updates per iteration. The cost calculation during the minimization is as well made in parallel within sub-ensemble members (number of rows in each ensemble member communicator), which leads to a most efficient minimization. Again, a MPI-communicator for the communication of MPI-processes at the same rank of the sub-ensemble members provide optimal exchange between the sub-ensemble members. Additional adaptations were necessary for the application of the ensEURAD-IM to a high performance computer. These are introduced in the next section.

4.3 JUQUEEN

The use of an ensemble system to calculate the analysis of volcanic ash dispersion demands high computational power. Therefore, the study was operated on the IBM BlueGene/Q system (JUQUEEN; *Jülich Supercomputing Centre* [2015]; *Stephan and Docter* [2015]) located at Jülich Supercomputing Centre. The system comprises 28 racks, each containing 32 nodes with 16 GB node memory. The nodes consist of 1.6 GHz PowerPC A2 processors with 16 cores/node and a maximum number of 64 threats/node using hyper-threading (*Brömmel et al.* [2015]).

The adaptation of the ensEURAD-IM to JUQUEEN required a large reorganization of the implemented MPI parallelization strategy. Originally, the EURAD-IM system was parallelized by a master/worker MPI parallelization (*Elbern et al.* [1997]). The bottleneck of this implementation is, firstly, the allocation of computing cores requires X-worker + 1 master, which on high performance computers results in the burden to allocate more cores than needed by the model. This is because on JUQUEEN only a fixed number of cores (or nodes to be more precise) of 2^y, $5 \leq y \leq 14$ can be allocated. Secondly, if the master only performs I/O operations, the master process is not in use for 99 % of the time and the required memory for storing the whole domain for all variables exceeds the capability of the core memory available on JUQUEEN. Therefore, the MPI parallelization strategy was transferred to an all-worker system, where one process (the former master process) only contains a little overload of I/O if parallel I/O is not required (e.g. for reading the namelist). To accomplish the transformation of the parallelization strategy, the main body of the code was reorganized, the I/O - processes were transformed to parallel netCDF, the interaction of the workers was reordered (as now (X+1)-workers are involved in the computation), and the domain subdivision had to be reorganized. Tab. 4.2 and Fig. 4.2 show the scaling behavior of the new ensEURAD-IM parallelization strategy normalized to 1024 cores. The data show the strong scaling of the ensEURAD-IM and, therefore, its applicability to high performance computers.

The saving of I/O time and computing time due to the new strategy and parallel netCDF is summarized in Tab. 4.3. For integration times of 6 h, 12 h, 24 h, and 48 h one forward run using the old (master/worker) and new (all worker) parallelization strategy is performed, respectively. The model runs are performed using 1024 cores. As can be seen directly from Tab. 4.3, both, I/O time and total

Table 4.2: *Speedup of the ensEURAD-IM normalized to 1024 cores.*

# cores	absolute timing in [s]	speedup
1024	1811	1.00
2048	986	1.84
4096	525	3.45
8192	308	5.87

Table 4.3: *Comparison of I/O and computing time for the EURAD-IM using the original and reorganized parallelization strategy and parallel netCDF. Shown are only forward runs for varying integration times running on 1024 cores. 'Master/worker' refers to the old parallelization strategy where the master process performs only I/O, and 'all worker' refers to the new parallelization strategy where the former master process performs only little more I/O.*

	I/O time		total comput. time forward run	
integration time	master/worker	All worker	master/worker	All worker
6 h	161.08 s	24.92 s	04:41 min	02:26 min
12 h	302.25 s	42.90 s	09:30 min	05:02 min
24 h	572.15 s	78.22 s	17:47 min	09:30 min
48 h	—	150.16 s	$>$ 30 min	19:08 min

computing time are reduced substantially for the new parallelization strategy. Remarkably, the I/O time of 48 h integration time using the new strategy is still less than for 6 h using the old strategy, which is an improvement of a factor of 8. The 48 h forward run using the old strategy required more than 30 minutes, which is the maximum wall-clock time on JUQUEEN for small jobs. Therefore, no I/O and total computing time for this configuration could be given. The two right columns of Tab. 4.3 show that for a given computing time the forecast length could almost be doubled using the model improvements. In addition to this computing time savings, the new parallelization strategy reduced the storage requirements. As is state above, the master process in the old parallelization strategy had to store the variables for the whole domain, which lead to a waste of computing power and

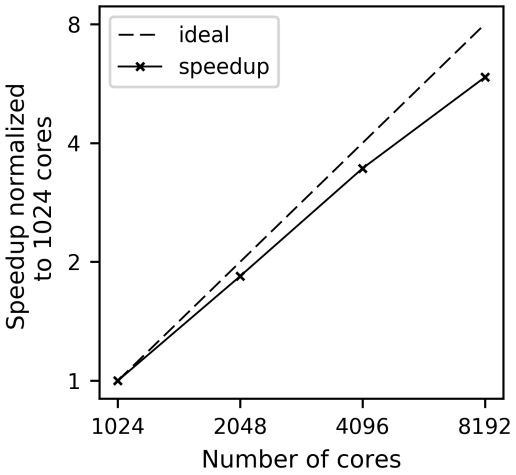

Figure 4.2: *Scaling performance of the ensEURAD-IM normalized to 1024 cores (solid line). The ideal scaling is defined by the dashed line.*

performance on the BlueGene/Q system for the old parallelization strategy. The new parallelization strategy allows for an application of the EURAD-IM to all horizontal resolutions and domain sizes.

Chapter 5

Ensemble for Stochastic Integration of Atmospheric Simulations (ESIAS)

The methods introduced in the previous sections are combined to an ensemble analysis system aiming to provide uncertainty forecasts of the atmospheric state. This Ensemble for Stochastic Integration of Atmospheric Simulations (ESIAS) comprises two model parts, which are an ensemble extension of the WRF model (ESIAS-met) and the previously introduced ensEURAD-IM (ESIAS-chem). In the following, the ESIAS system is described with a main focus on the chemical part of ESIAS (ESIAS-chem) model part as it was developed and used for the analysis presented here.

A schematic of ESIAS with the meteorological and the atmospheric chemical parts of the system is shown in Fig. 5.1. The system is coupled via the meteorological fields, which are generated by the meteorological part of ESIAS (ESIAS-met). For the analysis of volcanic ash, this meteorological ensemble provides the ability to represent the uncertainty of the driving meteorological fields. For this study, it is decided to use the same meteorology for both, the model run, which generates the artificial observations, and the analysis model runs. Thus, it is possible to analyze the performance of ESIAS-chem to estimate the true emission profile under idealized conditions.

5.1 ESIAS-met

ESIAS-met is an ensemble extension of the WRF model that is designed to operate efficiently ultra large ensembles with $\mathcal{O}(1000)$ ensemble members. ESIAS-met is initialized by two global ensemble systems, the Ensemble Prediction System (EPS) and Global Ensemble Forecast System (GEFS) operated by the ECMWF and National Centers for Environmental Prediction (NCEP), respectively. The global models also provide the boundary conditions. Model uncertainties are represented by stochastic

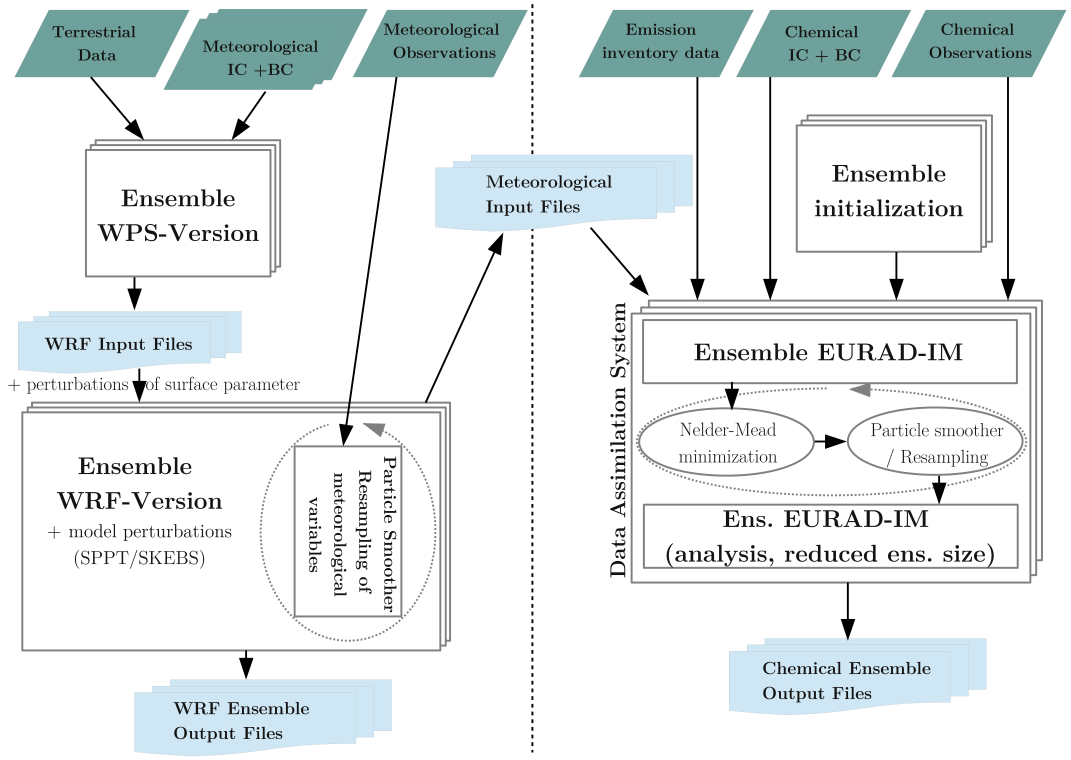

Figure 5.1: *Schematic of the ESIAS system including the meteorological (left) and atmospheric chemical (right) part. Staggered panels indicate ensemble versions. Green and light blue colors indicate model input and output, respectively. See the text for further information.*

perturbations of the parameterizations via the Stochastic Perturbed Parameterization Tendencies (SPPT; *Buizza et al.* [1999]) method as well as by stochastic perturbations of the energy budget via the Stochastic Kinetic Energy Backscatter Scheme (SKEBS, *Shutts* [2005]; *Berner et al.* [2009]). Additional uncertainty is introduced by perturbing the surface parameters and through different choices of model parameters (multi-parameter ensemble). The system is able to assimilate observations from land and sea surface, radiosondes, aircrafts and measurement towers. These data are assimilated via a particle smoother that uses all observations within a variable time window (e.g. from 00 - 06 UTC). At the end of the time window, the model parameters of the ensemble members are assigned a weight that is used for the generation of the next assimilation. The next assimilation extends the former time window (e.g. from 00 - 12 UTC) such that a smooth analysis trajectory is generated. The second analysis with the contribution of observations from 00 UTC to 12 UTC is expected to be closer to the observations as the model parameters were guided by the weights obtained in the first assimilation window. Further details on the ESIAS-met system as well as its application to wind energy analysis is given in *Berndt* [2018].

5.2 ESIAS-chem

The meteorological fields generated by ESIAS-met can be used as input for ESIAS-chem. Thus, the uncertainty of meteorological variables can be represented. This is especially important for wind, which is, besides the emission height, the major source of uncertainty for the location of the volcanic ash cloud in the atmosphere. However, because the focus of this study is placed on the estimate of emission profiles and the ensuing volcanic ash concentrations, the meteorological fields are assumed to be free of errors. For volcanic eruptions, the system aims to estimate the emissions using the separation of volcanic ash emitted at different times and heights through the vertical wind shear. This separation is usually larger for larger residence times in the atmosphere. Therefore, the main part of ESIAS-chem is a particle smoother with an increasing assimilation window. ESIAS-chem comprises the methods introduced in the previous sections. ESIAS-chem is flexible in integrating other modules and the used method is applicable to other atmospheric models as well.

Basically, the particle smoother in ESIAS-chem combines the ideas of two particle filter methods: the auxiliary particle filter and the implicit particle filter. In the former, first an approximation of the filter density with an auxiliary ensemble is calculated, before the full ensemble is used to analyze the atmospheric state. In the latter, first a representation of the minimum is estimated via the 4D-var method, before the ensemble is initialized around that minimum. ESIAS-chem is initialized with a large ensemble of pairwise distinct emission packages as is introduced in section 3.1. This ensemble of emission packages can be interpreted as an auxiliary ensemble in which not the ensemble size of the model is reduced but the emission profiles are reduced to single emission packages. Comparable to the auxiliary particle filter, in which the auxiliary ensemble is reduced in ensemble dimension and/or model complexity, the auxiliary ensemble in the presented particle smoother is reduced in its emission complexity. However, the dimension of the auxiliary ensemble may be larger than the analysis ensemble as it depends on the resolution of the emissions and the length of the assimilation window. This auxiliary ensemble is used to perform the minimization of the cost function using the DENM method.

Assume the first observations being available at $t = t_1$. The analysis is started by integrating the auxiliary ensemble from t_0 to the observation time t_1. Here, the DENM minimization algorithm is used to find the best ensemble representation of the atmospheric state. The DENM method searches for the minimum by combining the ash concentrations resulting from the different emission packages. As is introducted in section 3.2, this is done via geometric moves of the simplex on the hyperplane spanned by the cost function. The minimization step is comparable to the implicit particle filter in which the minimum of the cost function is calculated using 4D-var. In contrast to the implicit particle filter, in the new developed particle smoother not only a single estimate of the minimum of the cost function is

calculated but N_{ens} representations of the minimum of the cost function. However, there is no guarantee that the resulting ensemble is a good representation of the model state and its uncertainty because the variance of the ensemble might be too large. Beneath the well fitting ensemble members with low costs, there might exist ensemble members with high costs. Here, the filtering step of the particle smoother is applied by calculating weights w_i for each ensemble member

$$w_i = \frac{\tilde{w}_i}{\sum_j \tilde{w}_j}, \quad \tilde{w}_i = \left(\frac{N_{ens}}{J(\mathbf{y}|\mathbf{a})} \right)^{0.7}, \tag{5.1}$$

where the exponent was determined in a series of test runs with varying exponents and factors. The choice of this exponent makes the weights more equal, which is beneficial for omitting filter degeneracy. The cost function used here is

$$J(\mathbf{y}|\mathbf{a}) = -2\ln(p(\mathbf{y}|\mathbf{a})) = \sum_{t=0}^{L} (H\tilde{M}_t[\tilde{\mathbf{a}}_t(\mathbf{z})\mathbf{e}_0] - \mathbf{y})^T \mathbf{R}^{-1} (H\tilde{M}_t[\tilde{\mathbf{a}}_t(\mathbf{z})\mathbf{e}_0] - \mathbf{y}) + \mathbf{a}^T \mathbf{B}^{-1} \mathbf{a}, \tag{5.2}$$

where $\mathbf{B}$ is the emission factor error covariance matrix. The emission factor error covariance matrix aims to represent the uncertainty of the emission factors and is represented by a diagonal matrix

$$\mathbf{B} = diag(10). \tag{5.3}$$

This choice of emission factor error covariance matrix is a result of several test runs. More effort to find a reasonable emission factor error representation is required.

The reason for using (5.1) instead of (2.5) is that the cost function (5.2) is to be minimized. Therefore, low costs are equivalent to high weights. The application of (2.5) would give high weight to bad fitted ensemble members and vice versa. Within the DENM method, the distance between the model state and the observations is likely large, especially for the beginning of the minimization. The cost function needs to be capable of this fact. Hence, the likelihood $p(\mathbf{y}|\mathbf{a})$ is not applicable for the far distant model states with respect to the observations as its value vanishes in this case. Cost function (5.2) aims to overcome this issue.

In addition to the observation error covariance matrix, the observation error is represented by perturbed observations (cf. *Houtekamer and Mitchell* [1998]), which are assimilated by ESIAS-chem. Herein, the observations for each ensemble member are perturbed according to the observation error covariance matrix. For this purpose, the observation error is assumed to be Gaussian. Applying the perturbed observations for the assimilation increases the variance of the ensemble members.

The weights, which result from the filtering step, are applied to the emission profiles. Hence, the observations taken for the analysis are applied to update the emissions within the whole assimilation window. Once the weights are calculated, the ensemble of emission profiles is resampled using residual resampling such that the ensemble size N_{ens} is preserved.

With increasing time after the emission of volcanic ash, the amount of information obtained by the observations of column mass loadings is likely to increase. This is because more volcanic ash might be observed and the vertical wind shear is more likely to separate the volcanic ash emitted at different times and heights. Therefore, the assimilation is restarted from time t_0 with an increased assimilation window length $t_2 - t_0$, where t_2 is the time at which the next observations are available. In this way, the emission profiles are updated with as much information as possible.

Chapter 6

Identical twin experiments

Identical twin experiments test the assimilation system under idealized conditions. Artificial observations are generated by a model simulation, the so called nature run. Herein, a "true" emission profile is defined. The aim is to retrieve this true emission profile using ESIAS-chem. This setting assumes the model to be perfect. Additionally, the meteorological fields are assumed to be free of errors as well. Thus, deficiencies of the algorithm are analyzed as errors of the analysis only emerge from an erroneous algorithm. Identical twin experiments err on the optimistic side (*Daley* [1991]). However, they are indispensable to prove the concept and code correctness.

It is noted that all tests are performed without sedimentation. Thus, the volcanic ash is not removed from the atmosphere, which increases its residence time. Hence, the volcanic ash is amenable for height differentiation, which make the analysis more exact, especially in the statistical analysis. Tests showed, that not considering sedimentation has no effect on the performance of the assimilation algorithm (not shown). In the following, first the test cases, which are used for the performance analysis, are introduced. Thereafter, the dependency of the performance of ESIAS-chem on the length of the assimilation window as well as the potential and the limits of ESIAS-chem are shown. In the end, a statistical analysis of ESIAS-chem is performed.

6.1 Selection of test cases

To test the performance of ESIAS-chem for different assimilation window lengths, a hypothetical eruption of the Icelandic volcano Eyjafjallajökull is considered under real meteorological conditions. The Hovmoeller plot of the nature runs emission profile is shown in Fig. 6.1. It shows the emission rate (colored) at a given time (x-axis) and height above the volcano (y-axis). The sub-plinian eruption is characterized by two short explosive phases between 02-04 UTC and 06-08 UTC. The eruption is chosen to start at 02 UTC rather than at 00 UTC to test whether the analysis suffers from a temporal shift in the analyzed emission profile. Further, it is chosen to be rather strong with high emission rates especially at the top of the eruption column

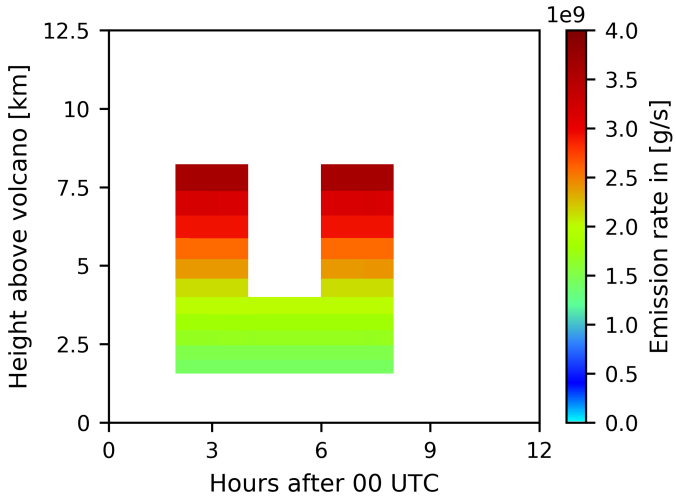

Figure 6.1: *Hovmoeller plot of the nature run emission profile for the performance analysis of ESIAS-chem in dependence on the length of the assimilation window on April 15 and April 29, 2010. Shown is the emission rate (colored) for a given time (x-axis) and height above the volcano (y-axis).*

at 8 km, where most of the volcanic ash is emitted from the eruption column.

The first test day to which ESIAS-chem is applied is April 15, 2010, which was characterized by strong winds over Iceland. This is illustrated by the wind profile at the volcano (Fig. 6.2a), the wind field over Europe in approx. 5 km height (Fig. 6.2b), and the geopotential height at 500 hPa and surface pressure (Fig. 6.2c), the latter all for 12 UTC on April 15. During this day, the polar front and, therefore, the polar jet stream is located over Iceland, forcing the volcanic ash to travel fast to the southeast. Thus, the volcanic ash is transported by strong winds of up to 60 m/s in heights of 5-8 km. Fig. 6.2 indicates, that the baroclinicity on April 15, 2010 over Iceland was low. Therefore, only a moderate vertical wind shear is expected. Hence, the volcanic ash column mass loadings resulting from the different emission packages that represent the a priori knowledge of the analysis system will likely separate only after some residence time in the atmosphere.

In addition to the April 15 scenario, a second analysis of the models performance is made for a hypothetical eruption of the Eyjafjallajökull on April 29, 2010. The same emission profile is taken (Fig. 6.1). This day is characterized by weak winds of approximately 10 m/s in the vicinity of the volcano, which is illustrated by Fig. 6.3. Thus, the emitted volcanic ash is only slowly transported.

To generate a larger database for a statistical analysis of the ESIAS-chem performance, an additional identical twin experiment is performed for the two eruptive phases of the 2010 Eyjafjallajökull eruption, which were on April 14-18 and May 04-18, 2010. In both eruptive phases the plume height was measured by radar observations to reach heights of 5-10 km (*Arason et al.* [2011]). From these measurements,

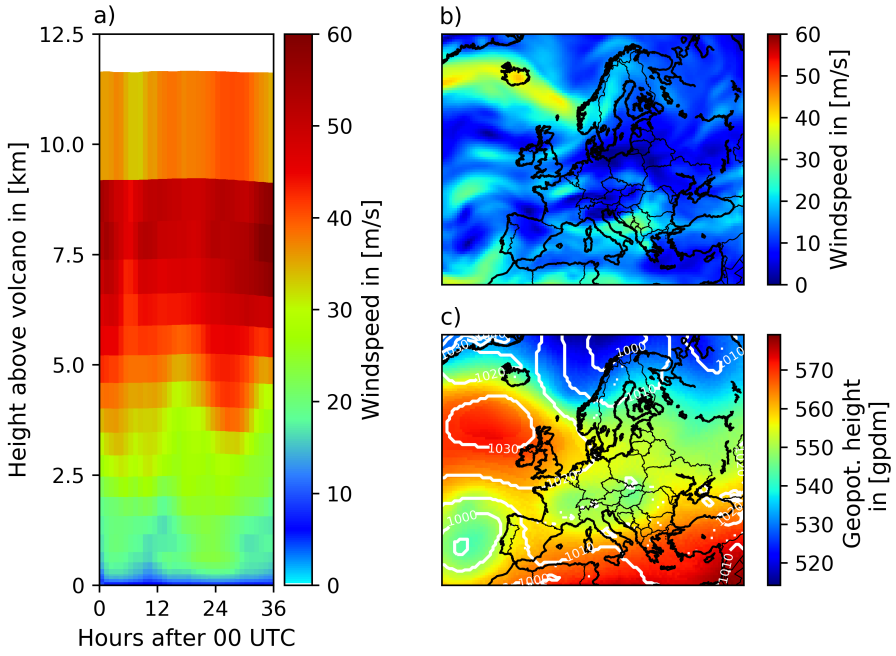

Figure 6.2: *Meteorological conditions on April 15, 2010. (a) Wind speed above the volcano. (b) Wind speed in approx. 5 km height at 12 UTC. (c) Geopotential height of 500 hPa and surface pressure (white contours) at 12 UTC.*

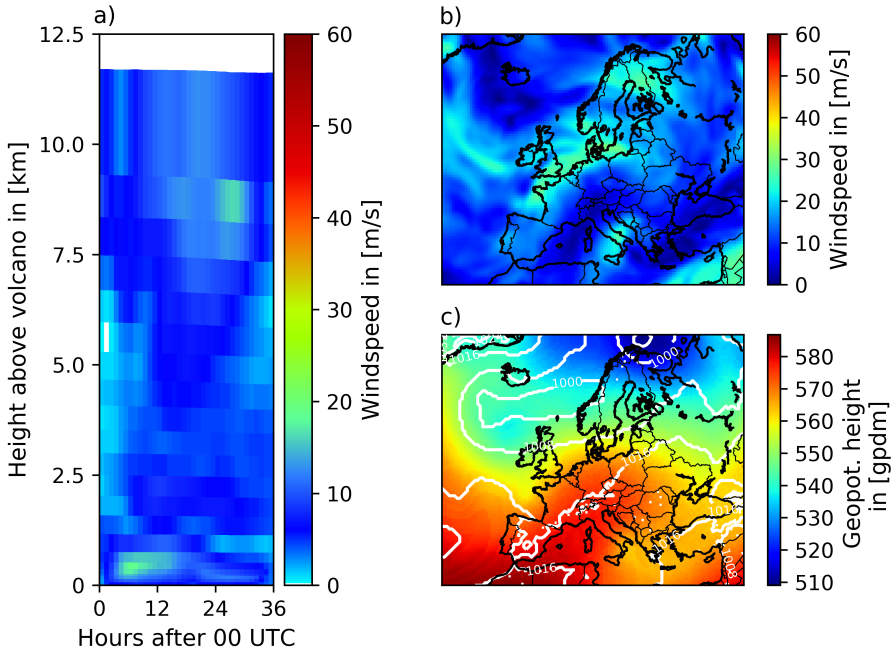

Figure 6.3: *Same as Fig. 6.2 but on April 29, 2010.*

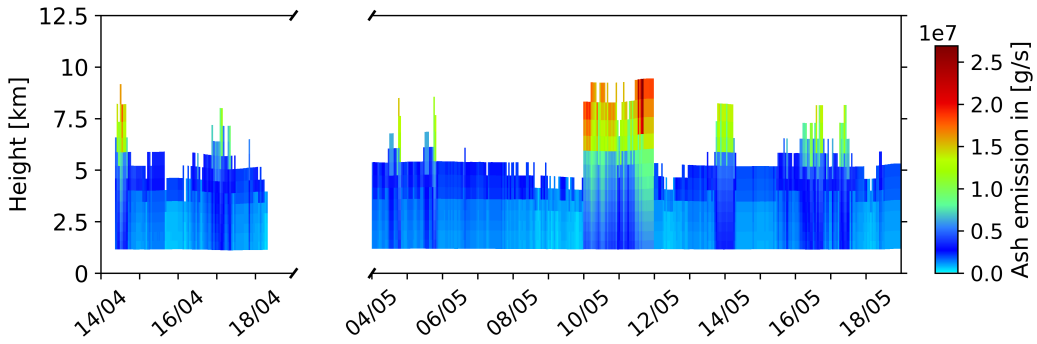

Figure 6.4: *Nature run emission profile for the identical twin experiments on April 14-18 and May 04-18, 2010. The heights of the eruption column are taken from Arason et al. [2011]. The data are averaged over one hour with data gaps filled by measurements from previous hours. On May 10-11, the measured heights of the eruption column are increased by 4 km.*

emission rates are calculated by applying the statistical eruption model of *Mastin et al.* [2009]. The magma flow rate M is defined by

$$M = \rho \left(\frac{H_T}{2} \right)^{\frac{1}{0.241}} f_{fa}, \tag{6.1}$$

with $\rho = 2,500$ kg/m^3 the magma density, H_T the plume top height measured by the radar in [km], and $f_{fa} = 0.1$ the fine ash fraction of the emitted magma. The data by *Arason et al.* [2011] are averaged over one hour, where missing data are replaced by the last valid measurement. In order to increase the variance of the nature run emissions, the measured plume-top heights on May 10-11, 2010, are increased by 4 km. The resulting nature run emission profiles are summarized in Fig. 6.4.

For the analysis of the identical twin experiment, the volcanic ash concentrations at the locations of 33 Earlinet stations (active and inactive[1]) are stored every model time step (10 minutes). Hence, a comparison with hypothetical lidar measurements can be performed. The locations of the lidar stations are shown in Fig. 6.5, where the different sizes of the red circles indicate the number of time steps at which volcanic ash is detected.

For all experiments, observations y_i of volcanic ash column mass loadings in [g/m^2] are taken from the nature run every six hours. Observed is the full domain, i.e. the grid cells containing volcanic ash and those without volcanic ash. This information is necessary in order to remove emissions at false times and heights. The observation error σ_{y_i} at location i is calculated by

$$\sigma_{y_i} = \max \left[\frac{(y_i * 0.4)^2}{\max[y_i * 0.4]}, 0.1 \right]. \tag{6.2}$$

[1]cf. *https://www.earlinet.org/index.php?id=105*; visited: September 26, 2017.

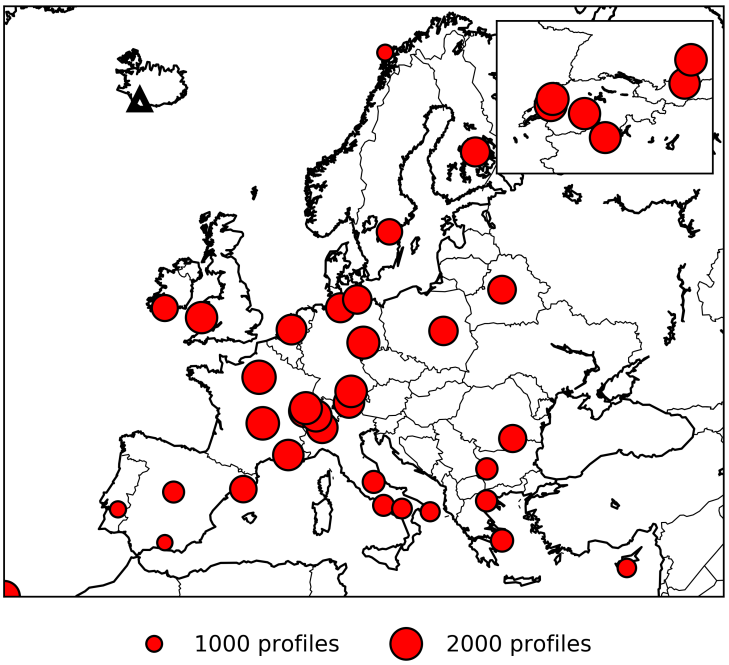

Figure 6.5: *Frequency of hypothetical lidar measurements taken at 33 Earlinet lidar stations. The size of the circles indicates the number of time steps for which volcanic ash is detected.*

The observation of the full domain includes locations with very small volcanic ash column mass loadings. This makes the limitation to $\sigma_{y_i} = 0.1$ necessary in order to reduce the relative influence of small observations on the cost function. Otherwise, changes of small column mass loadings would dominate the cost function value, which significantly reduces the performance of the analysis system.

To find an appropriate horizontal resolution for tests of the ensemble system, a scale analysis is done. The time t (in time step units) an information takes to be transported over one grid cell can be approximated by the CFL-ratio

$$t = \frac{\Delta x}{\overline{u}\Delta t}, \tag{6.3}$$

where Δx is the horizontal resolution (in m), $\overline{u}$ is the wind speed (in m/s), and Δt is the temporal resolution (in s) of the model domain. Tab. 6.1 summarizes the temporal resolution and the domain size to cover total Europe for different horizontal resolutions. It was decided that the optimal balance between temporal resolution of the emission data and the computing time can be achieved using a 45 km horizontal resolution for the identical twin experiments. The coarse domain was chosen in order to save available compute resources. In an application to real volcanic eruptions, a higher horizontal resolution of 15 km or less is appropriate. For the identical twin experiments, the horizontal resolution is of minor importance

Table 6.1: *Scale analysis for volcanic ash emission estimation using different horizontal (Δx) and temporal (Δt) model resolutions. For a wind speed of 10 m/s the temporal resolution is calculated by (6.3). The approximated domain size covering the whole European continent for the given horizontal resolution is also given.*

Δx [m]	Δt [s]	temporal resolution of emission data [h] (and time steps [ts])	approx. domain size in grid boxes to cover Europe
75,000	1200	2.10 (6.25 ts)	80 x 60
45,000	600	1.25 (7.50 ts)	130 x 100
15,000	300	0.41 (5.00 ts)	400 x 300
5,000	100	0.14 (5.00 ts)	1,200 x 900
1,000	40	0.03 (2.50 ts)	6,000 x 4,500

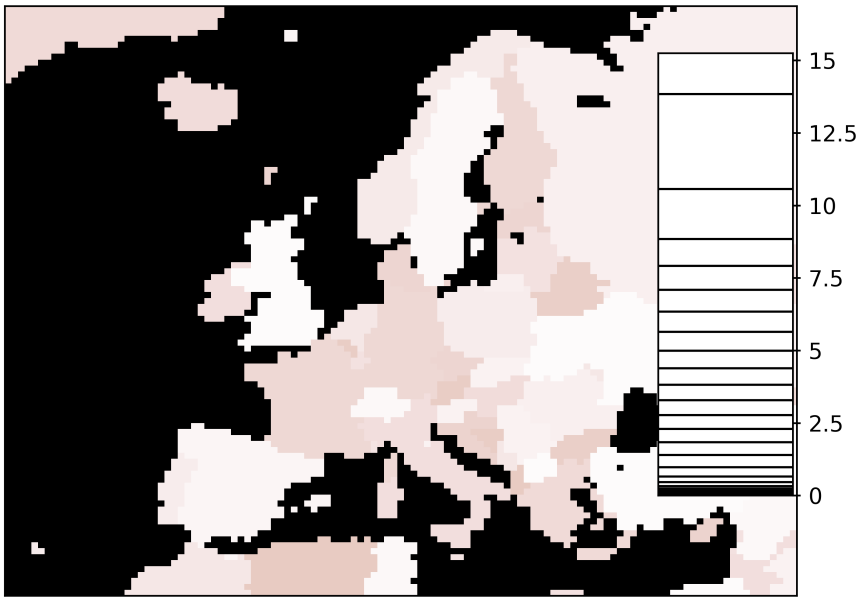

Figure 6.6: *Domain of the ensEURAD-IM for the identical twin experiments. The horizontal grid resolution is 45 km. In the vertical, the model is discretized in 23 layers, which are illustrated by the mean model layer height over the full domain (small subplot).*

as the observations are also generated with this coarse resolution. Ideally, the 45 km horizontal resolution results in a temporal resolution of the emission data of approx. one hour, which was decided to be the temporal resolution of the a priori emission packages (cf. Tab. 6.1). Nonetheless, horizontal and vertical diffusion and lower wind speeds will further reduce the actual temporal resolution of the analysis of emission profiles.

Fig 6.6 shows the coarse domain with 45 km horizontal resolution covering the

European continent from the east of Greenland in the northwest to the east of the Mediterranean Sea in the southeast. Vertically, the model is discretized in 23 layers up to 100 hPa (cf. scale on the right of Fig. 6.6). The vertical resolution in the height levels between 5 km and 10 km is approximately 1 km.

6.2 Dependence on the assimilation window length

The length of the assimilation window influences the performance of data assimilation algorithms. In this section, the performance of ESIAS-chem is tested for different assimilation windows. All assimilation windows start at 00 UTC for the specific day and last for 06-36 hours, where the assimilation window length increases by six hours. By increasing the assimilation window length, the observations include more information as the residence time of volcanic ash in the atmosphere is increased.

In order to analyze the horizontal displacement of the volcanic ash cloud, the Pattern Correlation Coefficient (pcc) for the column mass loadings of the nature run (subscript y) and the ensemble mean (subscript x; according to (2.6)) are shown in Fig. 6.7. The Pattern Correlation Coefficient is defined by (cf. *Zidikheri et al.* [2016])

$$pcc = \frac{< \mathbf{va}'_{\mathbf{x}}, \mathbf{va}'_{\mathbf{y}} >}{|\mathbf{va}'_{\mathbf{x}}||\mathbf{va}'_{\mathbf{y}}|}, \tag{6.4}$$

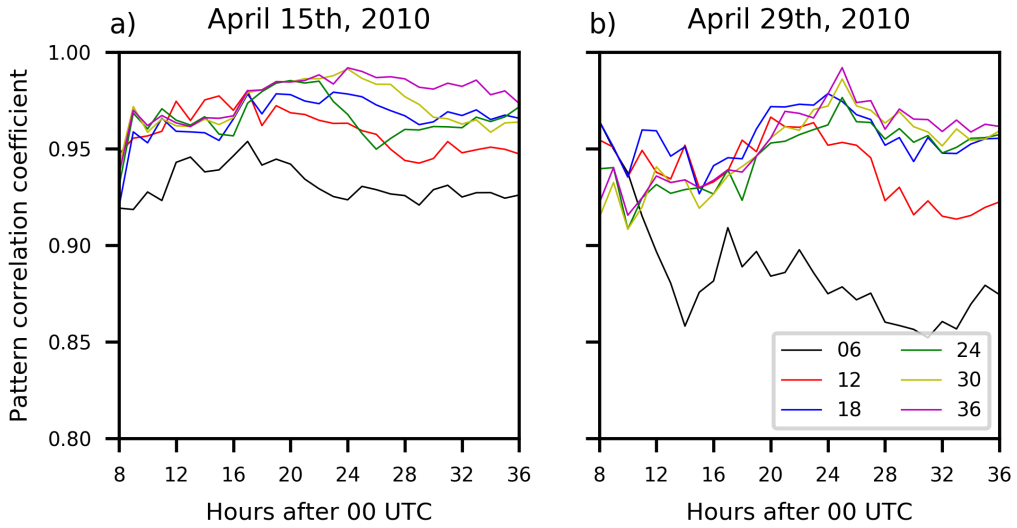

Figure 6.7: *Pattern correlation coefficient defined by (6.4) for the eruption on (a) April 15, 2010 and on (b) April 29, 2010. The different lines indicate different assimilation window lengths from 06 hours (gray) to 36 hours (magenta).*

with $\mathbf{va}' = \mathbf{va} - \overline{\mathbf{va}}$. Herein, the entries of the volcanic ash detection vector $\mathbf{va}$ are equal to 1 if the grid cell contains volcanic ash and 0 otherwise. The mean volcanic ash detection $\overline{\mathbf{va}}$ is calculated by

$$\overline{\mathbf{va}} = < \mathbf{1}, \mathbf{va} > / < \mathbf{1}, \mathbf{1} >, \tag{6.5}$$

where $\mathbf{1}$ denotes the unit vector with 1 on all entries and $< \cdot, \cdot >$ indicates the scalar product. The lines in Fig. 6.7 indicate different assimilation window lengths. The pcc ranges from 0 to 1, depending on the number of ash containing grid cells, in which nature run and ensemble mean do coincide. It is noted that for this model setup with perfect meteorology, pcc < 1 indicates that the analysis contains volcanic ash either in model layers or at times, where no volcanic ash is emitted in the nature run. In these cases, the volcanic ash of the ensemble members would be affected by different winds than the volcanic ash of the nature run. This is the only possibility to create different patterns of volcanic ash column mass loadings in this analysis setup. Furthermore, the pcc gives no information about the mass of volcanic ash that is falsely emitted.

Fig. 6.7 shows a large pcc > 0.85 after 08 UTC for both analysis days, April 15 and 29, 2010. The pcc is shown from 08 UTC on because this is the end of the eruption. The high pcc values indicate that the assimilation of column mass loadings to estimate volcanic ash emissions succeeds to retrieve the horizontal extent of the volcanic ash cloud. Hence, only in some cases the volcanic ash of the ensemble members is emitted on different layers or at different times than in the nature run.

The pcc on April 15 shows a smoother temporal evolution than on April 29. This can be understood by considering the different wind regimes on the two days. On April 15, the winds are strong such that volcanic ash is rapidly transported. If at one time instance the pcc is low, the strong winds transport the volcanic ash to different locations than in the nature run. On April 29, the winds are weak at all heights such that the volcanic ash emitted in false heights or at false times does frequently coincide with the nature run volcanic ash column mass loading cloud as it stays basically close to the volcano. Again, the pcc does not account for deviations in the strength of volcanic ash column mass loadings at locations in which the ensemble mean and the nature run differs in containing volcanic ash.

Increasing the assimilation window length (i.e. taking later observation into account) increases the pcc on both days. An exception is the pcc of the test case with 24 hours assimilation window on April 15, 2010, which decreases temporarily below 0.95. Thus, it has the second worst pcc among the different assimilation window lengths. Nonetheless, a pcc value of 0.95 is still a good result. Overall, the pcc is slightly larger on April 15 than on April 29. The lower pcc values do not necessarily imply that the performance of the analysis system is worse because it does not account for the actual mass of volcanic ash on the grid cells.

To investigate the mass on grid cells that contain volcanic ash erroneously,

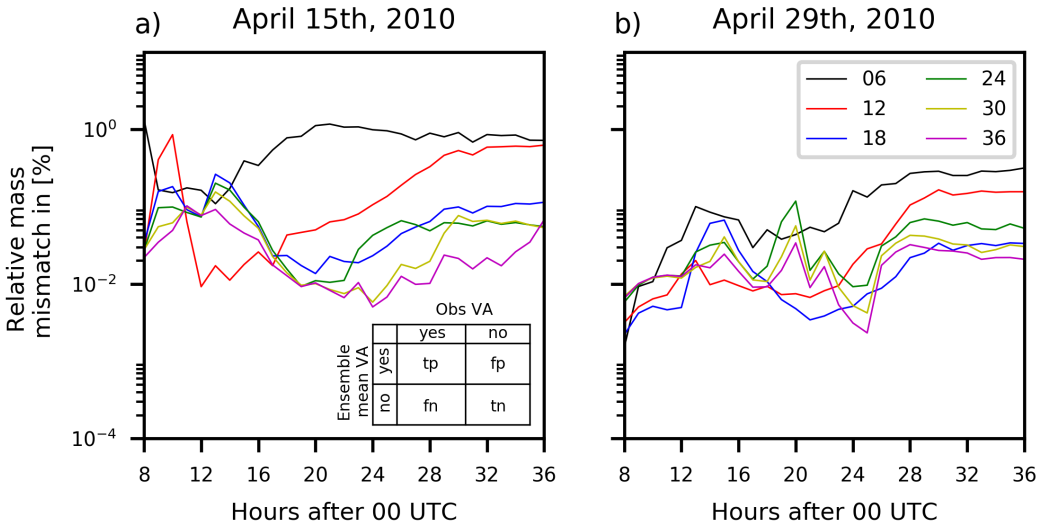

Figure 6.8: *Same as Fig. 6.7 but for the mass mismatch of the column mass loadings as defined by (6.6). The contingency table defines the phrases false positive, false negative, and true positive used in the text.*

Fig. 6.8 shows the relative amount of mass mismatch of volcanic ash

$$m = 100 \frac{\overline{m}_{fp} + m'_{fn}}{(\overline{m}_{tp} + m'_{tp})/2}, \tag{6.6}$$

with subscripts fp, fn, and tp indicate false positive, false negative, and true positive volcanic ash estimations (see the contingency table in Fig. 6.8), and $\overline{m}$ and m' denote the column mass loading of the ensemble mean and the nature run, respectively. The mass mismatch measures the mass of the ensemble mean that is at wrong locations (false positive) and the mass of the nature run at locations where it is not analyzed by the ensemble mean (false negative). This is normalized by the mean mass of the ensemble mean and the nature run, which is estimated at the same location (true positive).

Fig. 6.8 shows the mass mismatch for the identical twin experiments on April 15 and April 29, 2010, performed with different assimilation window length. Therein, less than 1 % of the intersecting mass of the nature run and the ensemble mean do not match. Especially for the last eight hours of the analysis, the reduction in the mass mismatch by increasing the assimilation window length is evident. During these hours, only the mass mismatch of the test case with 18 hours assimilation window on April 29 shows a contradiction to this finding as it is lower than the mass mismatch of the test case with 24 hours assimilation window. Nonetheless, for the test cases with assimilation window lengths longer than 18 hours, the mass mismatch ranges from $\sim 10^{-3}$ to $\sim 10^{-1}$%. This shows the good agreement of the column mass loadings from the nature run and the ensemble mean indicating that the analysis system is successful, especially for assimilation window lengths longer than 18 hours.

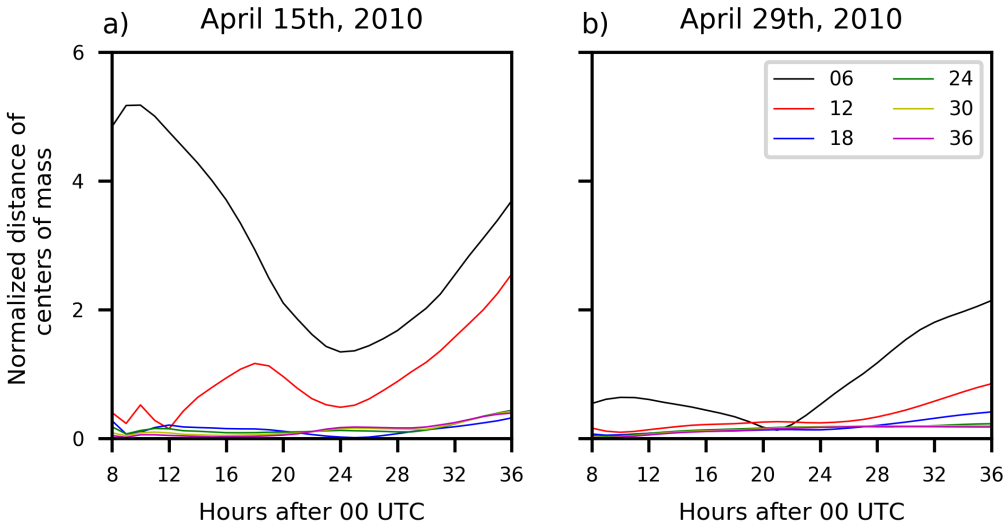

Figure 6.9: *Same as Fig. 6.7 but for the normalized distance of centers of mass of the column mass loadings defined by (6.9).*

Both, the pcc and the mass mismatch show the ability of the analysis system to reconstruct the lateral spread of the volcanic ash cloud in terms of column mass loadings. In addition, the normalized distance d between the centers of mass of the column mass loadings of the nature run and the ensemble mean is displayed in Fig. 6.9. The normalized distance of the centers of mass is defined by

$$d = \frac{1}{dx}|\mathbf{R}_y - \mathbf{R}_x|, \tag{6.7}$$

where dx is the horizontal resolution of the model, $\mathbf{R} = \frac{1}{\sum_j m_j}\sum_i m_i \mathbf{r}_i$ is the center of mass m of the nature run (subscript y) and ensemble mean (x), and r_i is the location of grid cell i. The normalized distance of the centers of mass is able to analyze the location and inner-cloud horizontal distribution of the volcanic ash. A value of $d < 1$ indicates that the distance of the centers of mass is smaller than the grid resolution of the model. Therefore, this can be seen as a criterion for a good analysis.

For both days analyzed, the normalized distance of the centers of mass are below 0.5 for assimilation windows longer than 18 hours. The increase of the values from approximately 28 hours after the simulation started results from differences in the emission profiles of nature run and ensemble mean. Especially on April 15, the vertical wind shear makes this difference evident: increasing the assimilation window length results in an increase in the distance of the centers of mass at the end of the analysis period. On April 29, the wind speed is low such that the volcanic ash is not transported far away from the volcano. Therefore, the distance of the centers of mass for assimilation window lengths is generally smaller than on April 15. For longer assimilation window lengths the distance of the centers of mass are small for both analysis days indicating the good performance of ESIAS-chem.

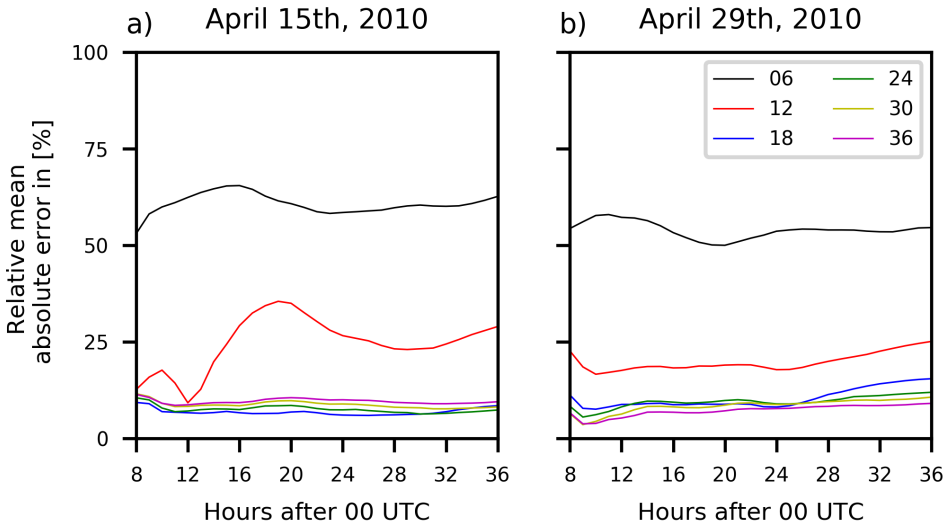

Figure 6.10: *Same as Fig. 6.7 but for the relative mean absolute error of column mass loadings defined by (6.8).*

Another criterion that analyzes the horizontal distribution of the volcanic ash within the cloud is the Relative Mean Absolute Error (RMAE) of the column mass loadings

$$RMAE = 100 \frac{1}{N_y} \sum_{j=1}^{N_y} \left| \frac{\overline{x}_j - y_j}{\sum_{i=1}^{N_y} y_i} \right|, \tag{6.8}$$

where N_y is the number of grid cells in which volcanic ash column mass loadings of the nature run y_j or the ensemble mean $\overline{x}_j$ are $\geq 10^{-3}$. Fig. 6.10 shows that the RMAE on both days is quite constant for the duration of the simulations. At the end of the simulation time at 36 hours, the RMAE for the test cases with assimilation windows greater than 18 hours is of the order of 10 % for both days. These low values show the good performance of the analysis for these assimilation window lengths with respect to the nature run.

In principle, Fig. 6.10 shows the same findings that are analyzed above, i.e. that by increasing the assimilation window length, the error of the ensemble mean decreases. One contradiction to this is the increase of the RMAE on April 15 for assimilation windows from 24 to 36 hours. This is a result of a convergence of volcanic ash in the upper troposphere south of Norway around 24 hours after the simulation has started. This convergence is illustrated in Fig. 6.11, which shows four profiles of the extinction coefficient that would be observed by lidar stations located south of Norway and Sweden at April 15-16, 2010, as obtained by the nature run.

The black arrows in Fig. 6.11 depict the arrival time of the upper tropospheric volcanic ash emitted by the two explosive eruptions (cf. Fig. 6.1). The extinction coefficient is calculated using the mass-extinction conversion factor proposed by

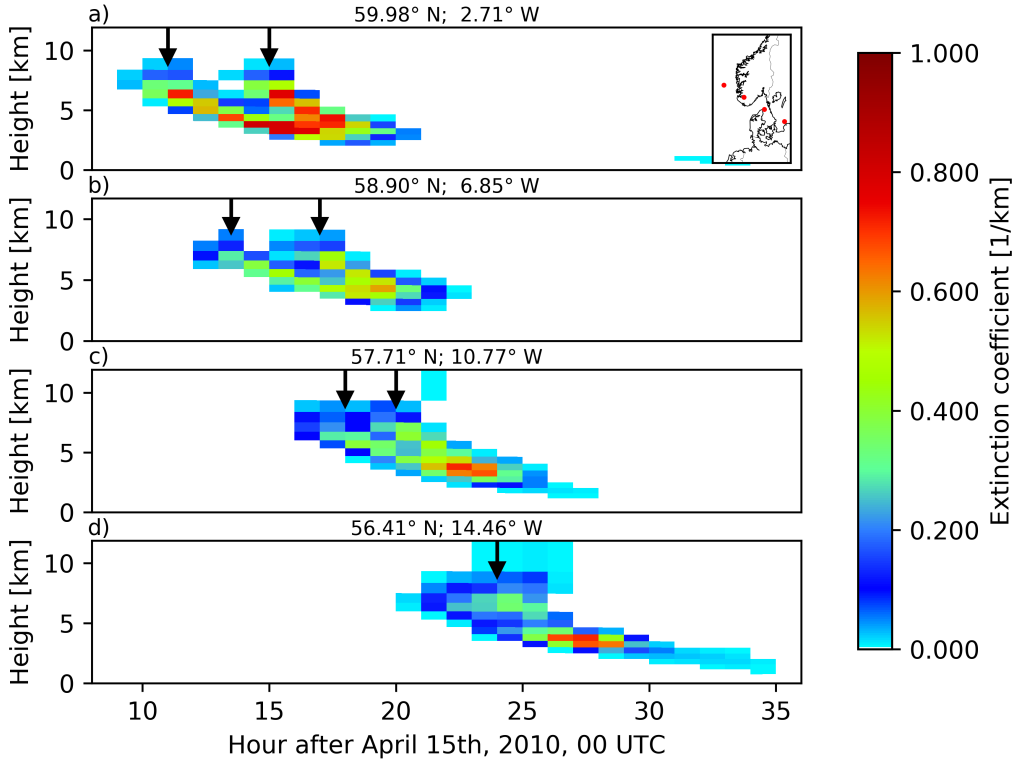

Figure 6.11: *Hypothetical lidar profiles of the extinction coefficient of volcanic ash calculated using (6.9) on April 15-16, 2010. The black arros indicate the maximum volcanic ash concentration in the upper troposphere. The locations of the hypothetical lidar instruments are shown as red marks on the map from the southwest of Norway (a) to the south of Sweden (d).*

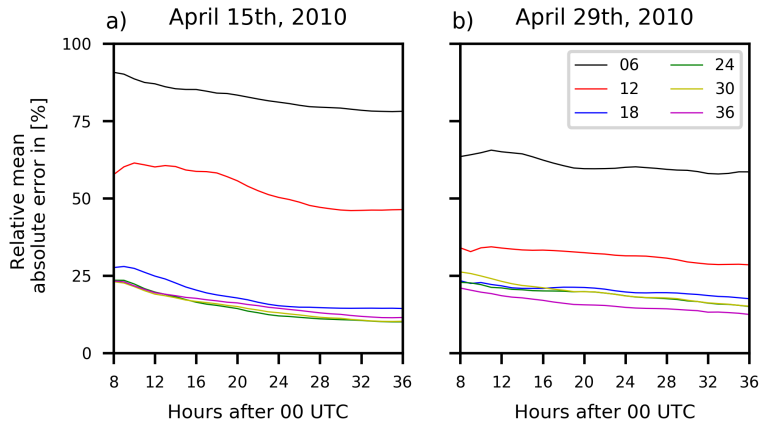

Figure 6.12: *Same as Fig. 6.7 but for the RMAE of the volcanic ash concentrations defined by (6.8).*

Gasteiger et al. [2011]

$$\eta = \frac{M}{\alpha} \approx 1.45 g/m^2, \tag{6.9}$$

with M and α denoting the volcanic ash concentration and extinction coefficient, respectively. Fig. 6.11 shows that as the volcanic ash is transported further south-east, the distinct signals of the extinction coefficient at 7-8 km converges until no distinction of the two signals is possible. This is when the volcanic ash cloud reaches the southern most lidar station (Fig. 6.11d). This convergence is finished at approximately April 16, 00 UTC, i.e. 24 hours after the simulation has started. Thereafter, observations of the volcanic ash column mass loadings do not allow to distinguish the two explosive eruptions.

All of the above analyses show that ESIAS-chem is able to reconstruct the column mass loadings of the nature run by assimilating vertically integrated values, that is, without information of the height distribution of volcanic ash. Fig. 6.12 shows the relative mean absolute error of the volcanic ash concentrations in order to investigate whether ESIAS-chem is able to retrieve the vertical distribution of the volcanic ash concentrations. The RMAE of the concentrations decreases by increasing the assimilation window length with the exception of the analysis using an assimilation window of 30 and 36 hours on April 15. Again, the reason is the convergence of the distinct ash clouds in the upper troposphere as shown in Fig. 6.11. Nonetheless, the RMAE of the volcanic ash concentrations for assimilation windows larger than 18 hours is of the order of 10 %, which shows the good performance of ESIAS-chem, not only to estimate the column mass loadings but also the vertical distribution of the volcanic ash in the atmosphere. This is further investigated and confirmed in the next section.

6.3 Analysis of the potential and limitations of ESIAS-chem

The former section shows the improvement in the analysis obtained by increasing the assimilation window length. In this section, the analysis results using an assimilation window of 24 hours are considered examplarily in more detail. The assimilation window length of 24 hours is chosen because it is a reasonable choice in applications to real volcanic eruptions and the results of the previous section indicate that the performance of the assimilation using an assimilation window of 24 hours is as good as longer assimilation window lengths for the selected analysis days. It is noted that the minimization starts with an arbitrarily chosen emission profile. Thus, the a priori volcanic ash concentrations are highly variable in time and space.

As an example, Fig. 6.13 shows the column mass loadings of (a) the nature run and (b) the ensemble mean at April 16, 2010, 06 UTC, which is 30 hours after the simulation has started and corresponds to a six hour forecast after the end of the assimilation window of 24 hours. Furthermore, Fig. 6.13c and Fig. 6.13d display the

relative error of the ensemble mean

$$RE = \frac{\overline{x} - y}{max(\overline{x})}, \tag{6.10}$$

and the relative ensemble standard deviation

$$\sigma_{rel} = \frac{\sigma_x}{max(\overline{x})}, \tag{6.11}$$

respectively. Herein, $\overline{x}$ and y are the ensemble mean and nature run column mass loading, respectively, and σ_x is the ensemble standard deviation. The column mass loadings at April 16, 06 UTC, show the good agreement of the ensemble mean and the nature run. Both show the correct location of the global maximum of the column mass loading pattern over the east Baltic Sea of up to 2.3 g/m^2 and the local maximum over Denmark of up 1.4 g/m^2. The relative error of the ensemble mean is less than 10 % and highest over the Baltic Sea, where large column mass loadings are present. In total, a small positive bias of the ensemble mean can

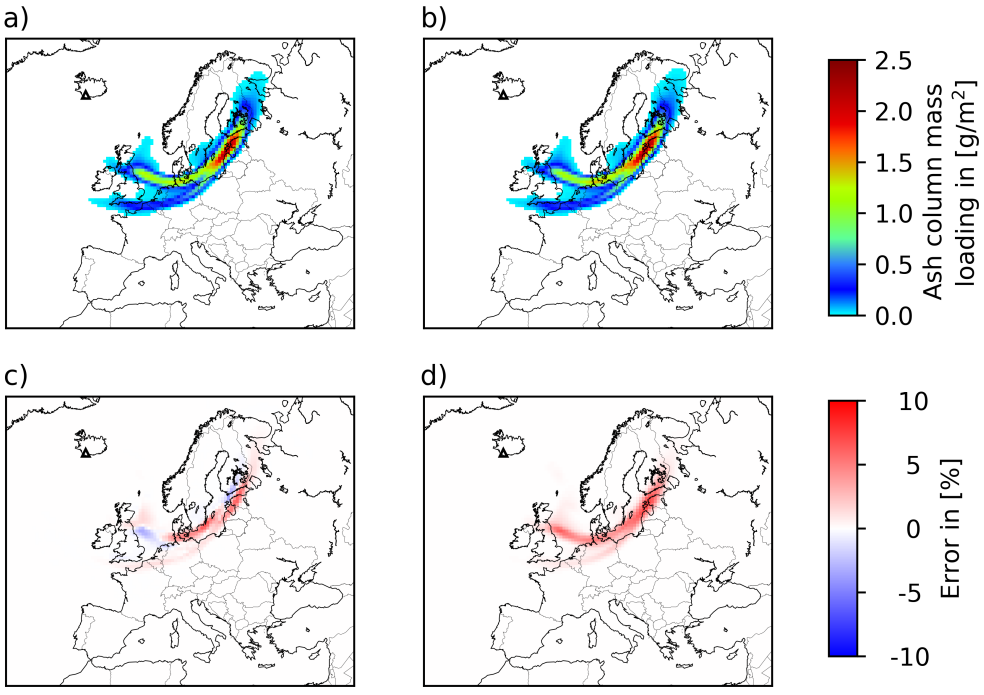

Figure 6.13: *Column mass loadings forecast of (a) nature run, (b) ensemble mean, (c) relative error of the ensemble mean according to (6.10), and (d) relative standard deviation of the ensemble according to (6.11) on April 15, 2010, 00 UTC + 30 hours using an assimilation window of 24 hours. The black triangle indicates the location of the erupting volcano.*

be seen by the relative error in Fig. 6.13c. The relative ensemble spread of the volcanic ash column mass loadings coincides with the relative error of the ensemble mean, which indicates again the applicability of ESIAS-chem to uncertainty quantification.

Fig. 6.14 depicts the vertical cross section of volcanic ash concentrations at April 15, 2010, 00 UTC + 30 hours along a hypothetical satellite path (black line on the map in Fig. 6.14a). Shown are the vertical cross section of (a) the nature run, (b) the ensemble mean, (c) the relative error of the ensemble mean, and (d) the relative ensemble standard deviation. Fig. 6.14 clearly shows the good agreement of the ensemble mean and the nature run volcanic ash concentrations along the satellite path. Regarding the vertical distribution of the volcanic ash, the analysis is able to reconstruct the second maximum along the satellite path in 8 km height at around grid box 60 beside the global maximum of volcanic ash concentrations in the vertical distribution in 3-4 km height. This is not possible for the most other existing analysis methods for volcanic ash (see the discussion in section 7 for more details).

The relative error of the ensemble mean (Fig. 6.14c) is of the order of 10 % in

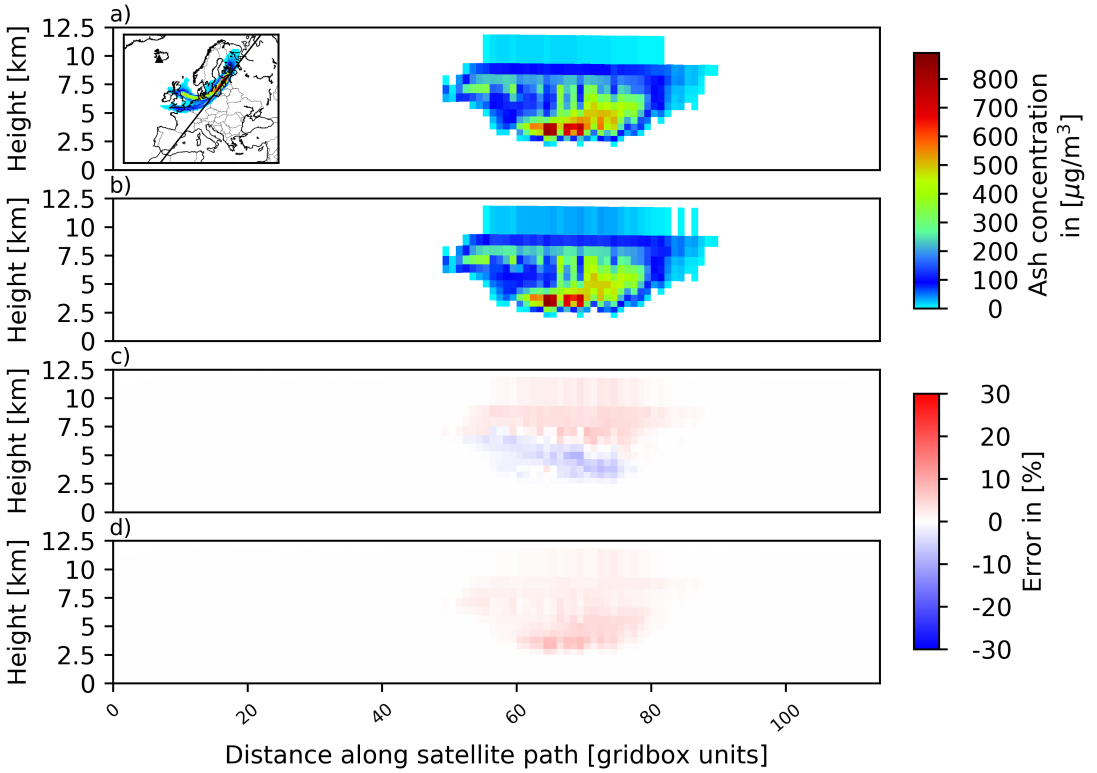

Figure 6.14: *Vertical cross section of (a) nature run, (b) ensemble mean, (c) relative error of the ensemble mean, and (d) relative standard deviation of the ensemble along a hypothetical satellite path (black line on the map) on April 15, 2010, 00 UTC + 30 hours.*

agreement with the RMAE of the volcanic ash concentrations in Fig. 6.12. The vertical distribution of the relative error of the ensemble mean is characteristic for the assimilation of vertically integrated column values like volcanic ash column mass loadings. As within the assimilation process only the vertically integrated column values are optimized, the vertical distribution of the volcanic ash is likely to deviate from the true vertical distribution. Only in an atmosphere with strong vertical wind shear, the vertically integrated column values can yield to the right vertical distribution of the volcanic ash. Therefore, the error of the ensemble mean in Fig. 6.14c shows an overestimation of volcanic ash of approximately 10 % in the upper troposphere and an underestimation of volcanic ash in the lower troposphere of the same order. Nonetheless, an error of only 10 % in estimating the vertical distributed volcanic ash concentrations is a remarkably good result of the assimilation process.

In a reliable analysis system, the ensemble standard deviation must be of the order of the error of the ensemble mean. In practice, the investigation of the reliability of the ensemble needs a large amount of data. Therefore, the relative ensemble standard deviation is only shown exemplarily without any statistical meaning. However, Fig. 6.14d shows that the relative ensemble standard deviation is of the order of the relative error of the ensemble mean for this vertical cross section. This indicates the potential of the analysis system to reliably predict the volcanic ash distribution in the atmosphere and its uncertainty.

The potential as well as the limits of ESIAS-chem can both be seen by the analyzed emission profiles. Fig. 6.15 and Fig. 6.16 display the profile of (a) the nature run emissions, (b) the ensemble mean emissions, (c) the relative error of the ensemble mean, and (d) the relative standard deviation of the ensemble mean for the identical twin experiments on April 15 and April 29, 2010, respectively. The total nature run emissions on both days ($4.25 \cdot 10^8$ tons and $4.30 \cdot 10^8$ tons on April 15 and April 29, respectively) are well captured by the analyzed total emissions ($4.58 \cdot 10^8$ tons and $4.10 \cdot 10^8$ tons, respectively). The error of the total emitted volcanic ash is 7.7 % and 4.7 %, respectively. On April 15, the analyzed emission profile of the ensemble mean shows the two explosive eruptions of the nature run emission profile with the correct height of the maximum emissions at the right time. Nonetheless, the ensemble mean shows a vertically and temporally diffused emission profile such that emissions are located at times and heights, at which no emissions are present in the nature run emission profile. However, these incorrect emissions are low with respect to the maximum emissions. The relative error of the ensemble mean emissions are of the order of 10 %-20 % for most emission times and heights (cf. Fig. 6.15c) and, therefore, the results are similar to the analysis made before. The relative ensemble standard deviation is of the same order as the relative error of the ensemble mean emissions. This test case can be seen as a demonstration of a successful analysis of the volcanic ash emissions at April 15, 2010.

In contrast, the emission profile of the ensemble mean on April 29, 2010,

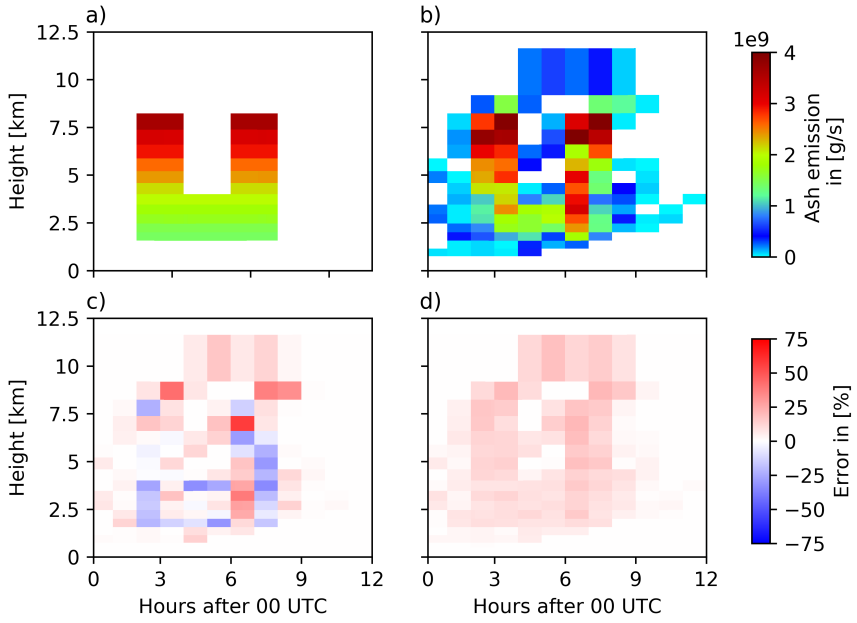

Figure 6.15: *Comparison of the emission profile of the nature run and ensemble mean on April 15, 2010. The figure shows (a) the nature run emission profile and (b) the ensemble mean emission profile. For comparison, (c) the relative error of the ensemble mean and (d) the relative ensemble standard deviation are illustrated.*

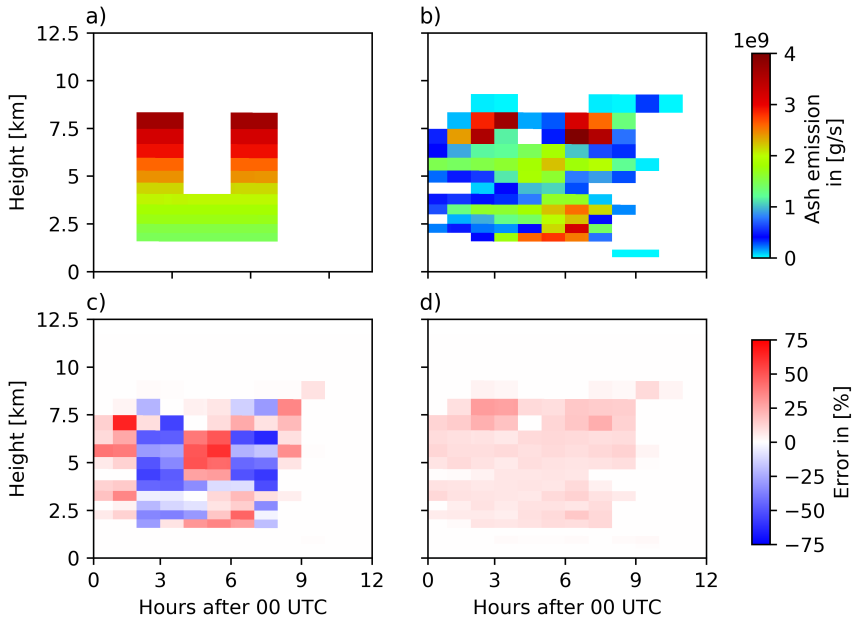

Figure 6.16: *Same as Fig. 6.15 but on April 29, 2010.*

(Fig. 6.16b) shows strong deviations from the nature run emission profile (Fig. 6.16a). Although the upper most emissions of the nature run emission profile in 8 km height are well captured by the ensemble mean, at lower levels no distinction between the two explosive eruptions is obtained. In comparison to the analyzed emissions on April 15, on April 29 the emissions of the ensemble mean are more uniform in time and height. Thus, large errors in both directions can be identified: negative errors during the explosive eruptions at around 03 UTC and 07 UTC indicating an underestimation of the emissions and positive errors outside the two explosive eruptions indicating an overestimation of the emissions.

The analyzed emission profile on April 29 shows the limits of ESIAS-chem. While the volcanic ash column mass loadings have only low errors as discussed in the previous section, the emission profile shows large deviations up to 60 % and more (Fig. 6.16c). The ensemble standard deviation of the emission profile (Fig. 6.16d) is lower than the relative error of the ensemble mean and ranges around 20 %. Although the analysis in the previous section shows a good agreement of the volcanic ash concentrations of the ensemble mean and the nature run, the large deviations of the analyzed emission profile may lead to larger errors if the forecast time is increased.

6.4 Statistical analysis of volcanic ash estimation

ESIAS-chem is a probabilistic forecast system. To evaluate probabilistic forecast systems, a larger data base is needed than based on a few days. In this section, the performance of ESIAS-chem is analyzed in the context of an identical twin experiment using a period of 20 analysis days for the volcanic eruption described in section 6.1. In total, there are 2,589 time steps, at which volcanic ash is detected at any of the 33 lidar stations (51,080 data pairs), considered for the analysis (see also Fig. 6.5).

Several analysis tools and scores exist to address the different characteristics of the forecast system. Among those, the following are considered in order to analyze the performance of ESIAS-chem:

- Relative mean error of the ensemble mean,
- Root mean square error of the ensemble mean,
- Mean error of the ensemble mean,
- correlation coefficient of the ensemble mean,
- Brier Score,
- Brier Skill Score,
- Area under the relative operating characteristics curve,

- Probability-Probability plot.

The quantities of the volcanic ash forecast that are used for the analysis at the 33 lidar stations are the column mass loadings, the maximum concentration, and the height of the maximum concentration in meter and σ-coordinates. These quantities are averaged over all occurrences and all lidar stations.

To investigate the reliability of ESIAS-chem, the relation of the ensemble spread (measured by the ensemble standard deviation) and the absolute error of the ensemble mean for the column mass loadings is shown in Fig. 6.17. The error of the ensemble mean exceeds the ensemble spread systematically. Thus, the ensemble spread does not reflect the error made by the forecast system. Especially, large errors of the ensemble mean column mass loadings forecast are underrepresented by the ensemble spread.

For a further investigation, the Relative Mean Error (RME) of the column mass loadings forecast is shown in Fig. 6.18. The RME is defined by

$$RME = \frac{\sum_{i=1}^{N}(\overline{x}_i - y_i)}{\sum_{i=1}^{N} y_i}, \tag{6.12}$$

with N the number of observations y_i, and $\overline{x}_i$ the corresponding ensemble mean of the column mass loadings for each observation time averaged over all stations. The RME shows a strong dependence on the forecast day. In the beginning of both analysis periods, i.e. on April 14, and May 04, 2010, the RME is of the order of 10 %, which corresponds to the error analyzed in the previous sections. On the subsequent days, the RME increases. At the end of the second analysis period, the

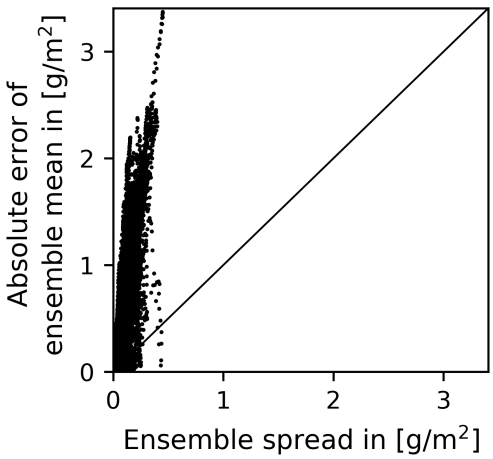

Figure 6.17: *Comparison of the absolute error of the ensemble mean and the ensemble spread of column mass loadings aggregated over the 33 lidar stations for April 14-18 and May 04-18, 2010. The black diagonal indicates a perfect forecast system.*

RME reaches values of 60 %-75 %.

To test the influence of the wind speed on the RME, Fig. 6.19 shows exemplarily the wind profile at the volcano for May 14-17. These days with large RMEs are accompanied by weak winds at the volcano. Although in section 6.2 it is shown that the error of the column mass loadings analysis is of the order of 10 % regardless of the wind conditions, in the subsequent forecast the error induced by the insufficiently represented volcanic ash emissions increases. This is due to the bad representation of the vertical distribution of the emissions for weak winds as is discussed in the previous sections. As argued before, these weak wind conditions are the limit of ESIAS-chem, which can only be overcome by increasing the assimilation window length. The optimal assimilation window length depends on the specific wind field and emission profile. Hence, it is not possible to give a fixed rule on the optimal assimilation window length.

In contrast to this weak wind period, Fig. 6.20 shows exemplarily the wind profile above the volcano on May 10-11, 2010. The wind speed on these days ranges from 20 m/s at the beginning of May 10 to approx. 50 m/s on May 11 in the upper troposphere. The corresponding RME reduces from more than 40 % on May 09 to approx. 20 % on May 10-11. This displays the ability of ESIAS-chem to provide a good forecast of volcanic ash if the winds are sufficiently strong.

The high RME for the later days of the analysis periods does not mean that the analysis system fails. The problems induced by weak wind conditions can partly be overcome by performing an error correction on the volcanic ash concentrations. Therefore, for the following investigation of the performance of ESIAS-chem, an error correction is applied to the volcanic ash concentrations after the forecast is made. It is noted that this error correction can, in principle, be applied to the emissions within the analysis process such that no postprocessing is needed. Because of the lack of computing time, this online error correction could not been tested. As the RME shown in Fig. 6.18 is given in terms of column mass loadings, which are the observed quantity, a multiplicative error correction is used in order to correct the

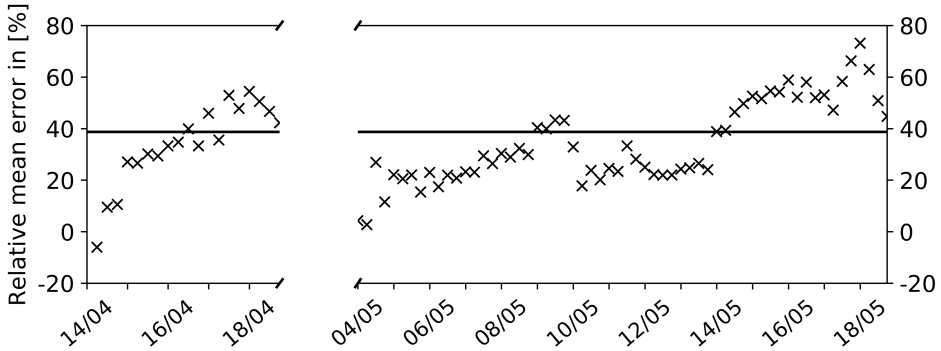

Figure 6.18: *Daily relative mean error for the identical twin experiment analysis on April 14-18 and May 04-18, 2010, for each observation time (crosses). The averaged RME is also displayed (solid line).*

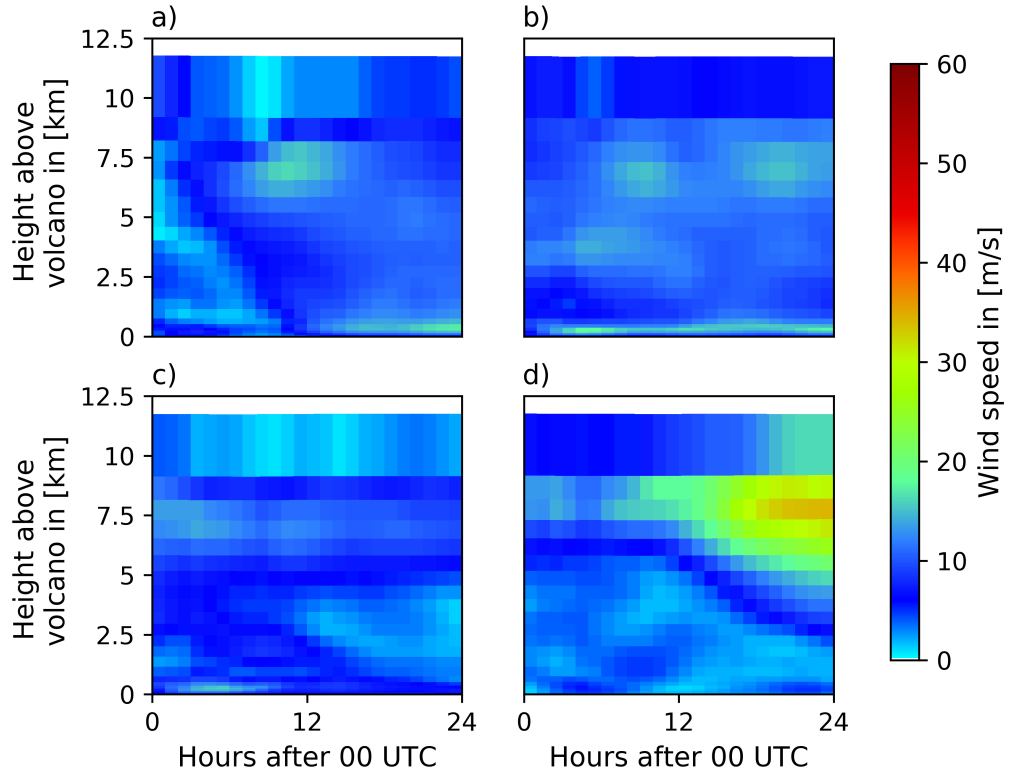

Figure 6.19: *Wind profile above the Eyjafjallajökull volcano on (a) May 14, (b) May 15, (c) May 16, and (d) May 17, 2010.*

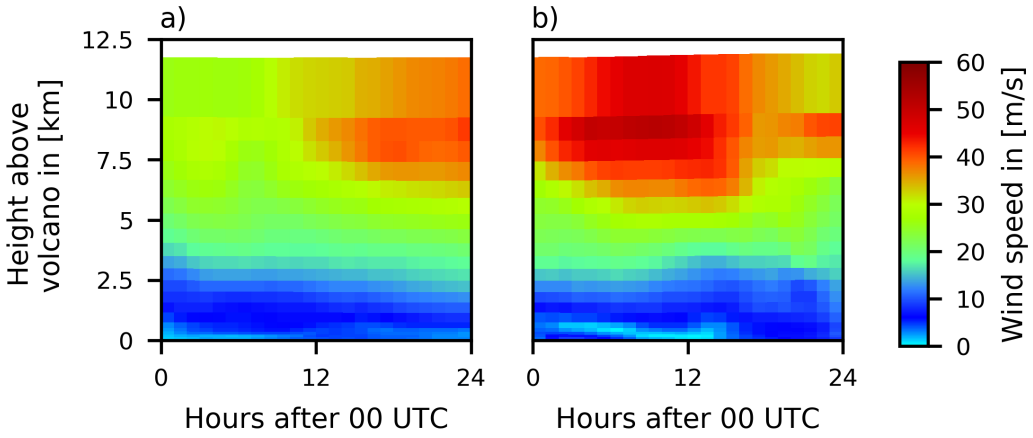

Figure 6.20: *Wind profile above the Eyjafjallajökull volcano on (a) May 10 and (b) May 11, 2010.*

volcanic ash concentration $\mathbf{x}$

$$\mathbf{x}_{ec} = c\mathbf{x} = \frac{1}{1+RME}\mathbf{x}, \tag{6.13}$$

where $\mathbf{x}_{ec}$ is the volcanic ash concentration after multiplication with the error correction factor c. The derivation of the error correction factor c starts with the assumption that the corrected ensemble mean volcanic ash concentration $\overline{\mathbf{x}}_{\mathbf{ec}}$ is on average unbiased

$$\mathcal{E}\left(H(\overline{\mathbf{x}}) - \mathbf{y}\right) = \mathcal{E}\left(\sum_{i=1}^{N_z} \overline{\mathbf{x}}_{ec}^{i} \Delta z^i - \mathbf{y}\right) = \mathcal{E}\left(c\sum_{i=1}^{N_z} \overline{\mathbf{x}}^{i} \Delta z^i - \mathbf{y}\right) \stackrel{!}{=} 0, \tag{6.14}$$

with $\mathcal{E}(\cdot)$ the expectation, and N_z the number of vertical model layers of thickness Δz^i. Using (6.12) the formula can be transformed to

$$\begin{aligned} c\mathcal{E}\left([\mathbf{y} + RME\mathbf{y}]\right) - \mathcal{E}\left(\mathbf{y}\right) =& 0 \\ c = \frac{\mathcal{E}\left(\mathbf{y}\right)}{\mathcal{E}\left(\mathbf{y}\right) + RME\mathcal{E}\left(\mathbf{y}\right)} =& \frac{1}{1+RME}. \end{aligned} \tag{6.15}$$

The error correction factor applied here is

$$c = \frac{1}{1+0.388}.$$

In an application to a real volcanic eruption, the error correction factor can be computed for each assimilation by the observed values of the column mass loadings.

Tab. 6.2 compares several statistics for the column mass loadings, maximum ash concentration, and height of the maximum ash concentration at the 33 lidar stations averaged over all time steps before and after the error correction is applied. The height of the maximum ash concentration is given in both, meter above ground and σ-coordinates. Compared are the RME, Root Mean Square Error (RMSE), Mean Error (ME), Correlation Coefficient (CC), Brier Score (BS), Brier Skill Score (BSS) and the Area under the Relative Operating Characteristics curve (ROC-A). The BS and BSS are calculated for the event of column mass loadings CML > 1 g/m^2 and maximum volcanic ash concentration MAX > 0.5 g/m^3. The amount of data is 51,080 for RME, RMSE, ME, and CC, 9,443 for BS and BSS, and 7,541 for ROC-A. Tab. 6.2 shows the improvement of the analysis by applying the error correction clearly. The RME reduces from 36.5 % to 1.7 % for the column mass loadings (CML) and from 30.0 % to -0.1 % for the maximum volcanic ash concentration (MAX). The RMSE and the ME for both quantities reduces similarly. Remarkably, the ME of the column mass loadings after the error correction (CML$_{ec}$) vanishes. This is a clear indication that the multiplicative error correction is successful and a reasonable choice for the analysis method. Because the error correction only alters the strength of the volcanic ash concentrations, it has almost no effect on the height in meter above the ground (HGT^m) of the maximum volcanic ash concentrations. The RME

Table 6.2: *Statistics of the performance analysis of ESIAS-chem aggregated over all 33 lidar stations and analysis time steps before and after the error correction (subscript "ec" for the latter). Shown are the relative mean error (RME), root mean square error (RMSE), correlation coefficient (CC), Brier Score (BS), Brier Skill Score (BSS), and the area under the relative operating characteristics curve (ROC-A). The statistics are calculated for the column mass loadings (CML), the maximum volcanic ash concentration (MAX) and the height of the maximum volcanic ash concentration above the lidar both, in meter (HGT^m) and σ-coordinates (HGT^σ). The RME, RMSE, and CC are aggregated over all volcanic ash values (51,080 data pairs). The BS, BSS, and ROC-A are aggregated for the column mass loadings for which the nature run values are larger than 1 gm^{-2} (9,443 data pairs) and for the maximum volcanic ash concentrations larger than 0.5 gm^{-3} (7,541 data pairs).*

	RME	RMSE	ME	CC	BS	BSS	ROC-A
CML	36.5 %	0.33 gm^{-2}	0.2 gm^{-2}	0.98	0.19	-0.01	0.69
CML_{ec}	1.7 %	0.14 gm^{-2}	0.0 gm^{-2}	0.98	0.10	0.48	0.93
MAX	30.0 %	142.7 μgm^{-3}	64.6 μgm^{-3}	0.97	0.18	-1.34	0.70
MAX_{ec}	-0.1 %	80.8 μgm^{-3}	-14.9 μgm^{-3}	0.96	0.14	0.23	0.89
HGT^m	8.0 %	1373 m	336 m	0.82	-	-	-
HGT^m_{ec}	7.3 %	1325 m	305 m	0.83	-	-	-
HGT^σ	3.7 %	2.2	0.5	0.84	-	-	-
HGT^σ_{ec}	3.5 %	2.1	0.5	0.84	-	-	-

and RMSE reduces only from 8.0 % to 7.3 % and from 1373 m to 1325 m, respectively. The same is true for the height of the maximum volcanic ash concentrations in σ-coordinates (HGT^σ). The ME of HGT^σ shows that on average the maximum volcanic ash layer is displaced by 0.5 vertical layers. Thus, the averaged variation of the layer containing the maximum volcanic ash concentration is smaller than the vertical resolution of the model. This is a remarkable result, as column mass loadings of volcanic ash are assimilated. Thus, ESIAS-chem is able to retrieve the vertical distribution of volcanic ash by assimilating these vertically integrated values.

In order to investigate possible phase errors of the volcanic ash concentrations, the CC is given in Tab. 6.2 for the quantities mentioned above. The CC range between 0.96 and 0.98 for the maximum volcanic ash concentrations and column mass loadings and from 0.82 to 0.84 for the height of the maximum volcanic ash concentrations in meter and σ-coordinates. The estimation of the height of the maximum volcanic ash concentrations is more uncertain, leading to larger variations in the estimate and, therefore, to a slightly lower CC. However, all CC are high, depicting the ability of ESIAS-chem to estimate on average the correct temporal evolution of the volcanic ash concentrations at the lidar stations. Therefore, the volcanic ash analysis is free of any phase error.

The statistics BS, BSS, and ROC-A value the forecast system with respect to an event. For the column mass loadings forecast, this event is CML > 1 g/m^2. For the

forecast of the maximum volcanic ash concentrations, the event is MAX > 0.5 g/m^3. Although the value of 0.5 g/m^3 for the maximum volcanic ash concentrations is no official threshold for air safety considerations, it is chosen in order to have a larger data base than for higher thresholds.

The accuracy of the probabilistic forecast system and its skill is measured by the Brier Score and the Brier Skill Score. The Brier Score is defined as the mean square error of a probabilistic forecast

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2, \tag{6.16}$$

where N is the number of all occurrences, p_i denotes the predicted likelihood of an event to be occur, and $o_i = 0$ if the event is not observed and $o_i = 1$ if the event is observed. A good forecasting system predicts a high probability p_i of the occurrence of an event if the event occurs ($o_i = 1$) and vice versa. Thus, a Brier Score close to zero is characteristic for a good forecast system. The Brier Skill Score measures the skill of the probabilistic forecast according to a reference forecast and is defined by

$$BSS = 1 - \frac{BS}{BS_{ref}}, \tag{6.17}$$

where BS$_{ref}$ is the Brier Score of the reference forecast. In this study, the sample climatology is taken as reference forecast. The BSS ranges from $-\infty$ to 1, with BSS=0 indicating no skill above the reference and BSS=1 indicating a perfect skill. Tab. 6.2 summarizes the BS for the column mass loadings to be larger than 1 g/m^2 and maximum volcanic ash concentrations to be larger than 0.5 g/m^3 at the 33 lidar stations. Both quantities are reduced by the error correction from 0.19 to 0.10 for the column mass loadings and from 0.18 to 0.14 for the maximum volcanic ash concentration forecast. The BSS is increased from -0.01 to 0.48 for the column mass loadings and from -1.34 to 0.23 for the maximum of the volcanic ash concentrations. This displays the benefit gained by the error correction. In principle, the BS and BSS indicate an accurate forecast system at least for the column mass loadings. The accuracy of the maximum volcanic ash concentrations at the 33 lidar stations is reduced in comparison to the accuracy of the column mass loadings forecast, which basically depicts the issue of weak winds within the analysis period and, therefore, the insufficiently estimated volcanic ash emissions.

So far, the considered measures of ESIAS-chem analyze the overall quality either of the deterministic ensemble mean forecast (RME, RMSE, ME) or the probabilistic ensemble forecast (BS, BSS). The ROC-A is a score that measures the ability of the forecasting system to discriminate between the occurrence and non-occurrence of an event. A ROC-A value of 0.5 is equal to no skill of the forecasting system. The relative operating characteristics curve compares the hit rate and the false alarm rate for a given event. The hit rate is the probability to which an event is truly forecasted. In contrast, the false alarm rate is the probability to which an event, that does not occur, is forecasted. To construct the relative operating characteristics

curve, the forecasts are binned into probability categories for which the hit rate and the false alarm rate are computed. Hence, a ROC-A value larger than 0.5 reflects that, given a probability threshold, the forecast system is more likely to predict the occurrence of the event than the non-occurrence (*Casati et al.* [2008]). The score is equal to 1 if the forecast system is able to discriminate the events occurrence from non-occurrence.

The ROC-A in Tab. 6.2 shows the improvement of the analysis due to the error correction. For the column mass loadings forecast, the ROC-A increases from 0.69 to 0.93. For the maximum volcanic ash concentrations, the ROC-A increases from 0.70 to 0.89. These ROC-A values reflect the ability of ESIAS-chem to correctly discriminate between the occurrence and non-occurrence of whether the column mass loadings exceed 1 g/m^2 and the maximum volcanic ash concentrations exceed 0.5 g/m^3.

In addition to the statistics summarized in Tab. 6.2, Fig. 6.21 shows the Probability-Probability plot (P-P plot) of the column mass loadings aggregated over the 33 lidar stations. A P-P plot compares the Cumulative Distribution Function (CDF) of two data sets, here, the observations and the probabilistic forecast. For a given probability value, the P-P plot measures the percentage of the forecasts and observations that fall at or below this value. Thus, the closer the P-P plot is to the diagonal the more do both CDFs agree. Fig. 6.21 indicates a strong agreement of the forecast CDF with the CDF of the observations for the column mass loadings aggregated over all data at the 33 lidar stations. It shows an underprediction for

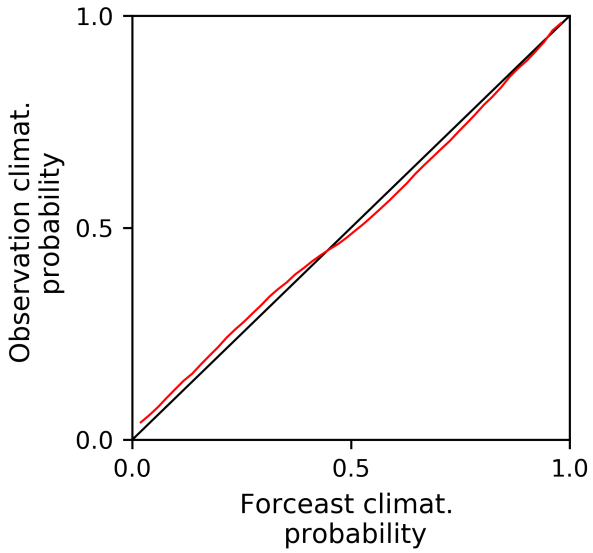

Figure 6.21: *Probability-Probability plot (P-P plot) aggregated over the 33 lidar stations for April 14-18 and May 04-18, 2010. The black diagonal indicates a perfect forecast system.*

low forecast probabilities (red curve is above the diagonal) and an overprediction for high forecast probabilities (red curve is below the diagonal). Additionally, the P-P plot supports the finding that the forecast is on average free of systematic errors after error correction.

Chapter 7

Comparison with literature

The identical twin experiments described in the previous sections confirm that ESIAS-chem is able to estimate the vertical distribution of volcanic ash concentrations on average adequately by assimilating vertically integrated volcanic ash column mass loadings. This strength of ESIAS-chem is more pronounced under strong wind conditions. Under weak wind conditions, the assimilation window needs to be elongated in order to give the volcanic ash emitted at different heights the chance to separate. Then, the assimilation of volcanic ash column mass loadings can yield sufficiently good estimates of the vertical distribution of volcanic ash as well. Nonetheless, the actual length of the assimilation window depends strongly on the wind field and the eruption style.

In order to analyze the benefit gained from ESIAS-chem, Tab. 7.1 summarizes the main features of the various methods proposed in the literature for estimating volcanic ash concentrations and its uncertainty. Beneficial features for analyzing volcanic ash concentrations are marked in bold. It is not the purpose of this study to value the results of the different methods, which are explained shortly in the introduction. All methods are shown to perform reasonably well in their field of application (cf. references given in Tab. 7.1). Here, only the complexity of the emission profiles, the applicability to real volcanic ash forecasts, and their computational costs are compared. It is noted that all methods assimilating column mass loadings of volcanic ash suffer from the same dependence on the wind conditions as ESIAS-chem does.

The most important positive property of an assimilation system for volcanic ash concentrations in the atmosphere is the ability to include temporal varying emissions by making as few as possible assumptions of the eruption style. Most assimilation systems summarized in Tab. 7.1 make assumptions of the emission profile. They assume volcanic ash emitted on a predefined height and time (*Schmehl et al.* [2012]), a Gaussian volcanic ash layer in the atmosphere (*Wilkins et al.* [2014]) or a Gaussian emission profile (*Denlinger et al.* [2012]), or they model volcanic ash emissions using perturbed parameters (*Bursik et al.* [2012]; *Madankan et al.* [2014]). Only a few assimilation systems make almost no assumptions of the emission profile (*Stohl et al.* [2011]; *Lu et al.* [2016], ESIAS-chem). In the method by *Stohl et al.* [2011] and in

ESIAS-chem, the flexible temporal and vertical resolution of the volcanic ash emissions stems from the pairwise distinct emission packages used for the optimization of the emission profile. The temporal and vertical resolution in these methods can be adapted such that it is most applicable with respect to compute resources and the expected variability of the true emissions. However, ESIAS-chem allows only for discrete emission strength, which is not critical as the discretization may be refined if the required compute resources are available. In contrast, the method proposed by *Lu et al.* [2016] gives an a priori estimate of the emissions and optimizes the emissions according to the adjoint-free ensemble 4D-var method. The analysis does depend on the a priori estimate of the emission profiles but this is not referred to as an assumption of the emission profile in this discussion.

In addition to the representation of the emissions, the methods differ in the number of simulations needed to generate the analysis. This is examined together with the estimated parallel section of the model simulations. It is noted that the actual information about the parallel section of the method is missing in the literature. Hence, it is only estimated and may, therefore, vary in the specific application. Only two methods require few model simulations, i.e. one simulation in the method by *Wilkins et al.* [2014] and $\mathcal{O}(100)$ simulations in the method by *Lu et al.* [2016]. In contrast, three methods need $\mathcal{O}(10,000)$ simulations (*Schmehl et al.* [2012]; *Bursik et al.* [2012]; *Zidikheri et al.* [2016]). The other methods, as well as ESIAS-chem need moderate sized ensembles of $\mathcal{O}(200)$-$\mathcal{O}(500)$. In *Zidikheri et al.* [2016], all simulations can be run in parallel. Therefore, this method is almost as fast as running one single model run. Nonetheless, the method by *Zidikheri et al.* [2016] makes the strongest assumptions of the emission profile (i.e. constant emissions for all times and heights) and analyzes only the horizontal dispersion of the volcanic ash and its uncertainty without any information on the volcanic ash concentrations. For the most other methods, the percentage of simulations that can be run in parallel is of the order of 98 %. For these methods, most simulations are run in parallel (e.g. the simulations that generate the a priori estimate). The optimization can only be performed if the a priori simulations are generated, which reduces the parallel part of the methods. Possibly, after the optimization, additional simulations are necessary in order to calculate the analysis, which further reduces the parallel section. The parallel section of the method by *Denlinger et al.* [2012] is estimated to be 50 % because it estimates the probability distribution iteratively. It is assumed that at least some of this method can be run in parallel. The method by *Lu et al.* [2016] is only parallel to approx. 20 % because of its application of an ensemble 4D-var. Herein, the iterations needed to calculate the optimal emission profile are to be calculated sequentially, while the ensemble in each iteration can be run in parallel. The method by *Schmehl et al.* [2012] uses the genetic algorithm for optimizing volcanic ash emissions. The genetic algorithm can be run in parallel such that each generation (i.e. a set of model runs) can be rapidly calculated. However, the expensive mutation and cross over steps still need to be run sequentially. ESIAS-chem requires $\mathcal{O}(500)$ model simulations to calculate the analysis out of which 98 % can be run in

Reference	Optimization method	Emission profile	# Simulations	Est. parallel section	Uncert. quantification	Error representation
Stohl et al. [2011]	least squares inversion	**free; time variable**	$\mathcal{O}(500)$	**98**%	-	**Obs+BG**
Schmehl et al. [2012]	genetic algorithm	fixed; time invariant	$\mathcal{O}(10{,}000)$[1]	60%	-	-
Wilkins et al. [2014]	data insertion	Gaussian[2]; **time variable**	**1**	-	-	-
Lu et al. [2016]	ensemble 4D var	**free; time variable**	$\mathcal{O}$**(100)**[1]	20%	-	**Obs+BG**
Denlinger et al. [2012]	Quasi-Newton / saddle pt. approx.	Gaussian, time invariant	$\mathcal{O}(500)$[1]	**50**%	**yes**	**Obs+BG**
Bursik et al. [2012]	PCQWE + CUT	Paramet.; time invariant	$\mathcal{O}(10{,}000)$	**98**%	**yes**	-
Madankan et al. [2014]	gPC + CUT / Kalman filter	Paramet.; time invariant	$\mathcal{O}(200)$	**98**%	**yes**	**Obs+BG**
Zidikheri et al. [2016]	pattern correlation	constant; **time variable**	$\mathcal{O}(10{,}000)$	**100**%	**yes**	-
ESIAS-chem	particle filter & DENM	**free; time variable**	$\mathcal{O}(500)$	**98**%	**yes**	**Obs+BG**
PF/DA	particle filter & 4D var	**free; time variable**	$\mathcal{O}$**(200)**	20%	**yes**	**Obs(+BG)**
LPF	localized particle filter	constant[2]; **time variable**	$\mathcal{O}$**(50)**	**100**%	**yes**	**Obs(+BG)**

[1] Values are estimated because no explicit information is given in the literature.
[2] No optimization of emission profiles but of volcanic ash concentrations in the atmosphere.

Table 7.1: *Comparison of volcanic ash optimization methods. Positive properties are written in bold. See text for further details.*

parallel. Therefore, ESIAS-chem is fast in calculating the analysis of volcanic ash emission.

Another difference among the various methods is the representation of observation and background errors. Most methods account for both, observation and background errors (*Stohl et al.* [2011]; *Lu et al.* [2016]; *Denlinger et al.* [2012]; *Madankan et al.* [2014], and ESIAS-chem). Four methods do not account for neither observation nor background errors (*Schmehl et al.* [2012]; *Wilkins et al.* [2014]; *Bursik et al.* [2012]; *Zidikheri et al.* [2016]) within the optimization procedure, which overrates the accuracy of the observations.

In the comparison above, the benefits of ESIAS-chem become evident. Although all other methods perform equally well in analyzing volcanic ash concentrations in the atmosphere, they suffer more or less from specific limitations. The most severe limitation is that 4 out of 9 methods (*Schmehl et al.* [2012]; *Denlinger et al.* [2012]; *Bursik et al.* [2012]; *Madankan et al.* [2014]) do not allow for temporal varying volcanic ash emissions. This means that in the analysis averaged volcanic ash emissions are estimated, which is especially for longer lasting volcanic eruptions a large source of uncertainty. Although ESIAS-chem allows only for discrete emission rates, it makes no assumptions of the style of the volcanic eruption. Thus, the variability of volcanic ash emissions can be represented by the method. Only two other methods do not make any assumptions of the eruption style either (*Stohl et al.* [2011]; *Lu et al.* [2016]) but both methods do not forecast uncertainties of volcanic ash predictions, which is a highly desired feature of the analysis system to quantify possible risks imposed by volcanic ash in the atmosphere as was also concluded in the literature (cf. *Kristiansen et al.* [2012]). Making no restricting assumptions of the eruption style implies that the methods are able to represent multi-layered volcanic ash emissions and, hence, concentrations in the atmosphere, which might be possible because of the wind profile and the thermal stratification at the volcano.

Although the method by *Wilkins et al.* [2014] is the most efficient method as it needs only one model simulation to estimate the volcanic ash concentrations, it is not able to quantify the uncertainty of that estimate. Besides, it does not account for model and observation errors. In fact, in *Wilkins et al.* [2014] the observations are taken as truth, which overrates the observations and neglects both, measurement and representativity errors. Observation errors are represented in some of the methods that quantify the uncertainty of the volcanic ash forecast (*Denlinger et al.* [2012]; *Madankan et al.* [2014], ESIAS-chem). Both is needed for the estimation of volcanic ash concentrations. However, the different methods represent different sources of uncertainty. In *Denlinger et al.* [2012], *Bursik et al.* [2012], and *Madankan et al.* [2014], the uncertainty with respect to four emission parameters (vent radius, vent velocity, mean and variance of the grain size distribution) is analyzed. In *Zidikheri et al.* [2016], the uncertainty in the horizontal dispersion of the volcanic ash cloud is considered. Alternatively, ESIAS-chem analyzes the uncertainty induced by the vertical and temporal distribution of the volcanic ash emissions as well as the emission strength. Thus, ESIAS-chem provides an estimate of the uncertainty of both,

the emission source and the horizontal distribution of the volcanic ash, except for the grain size distribution of the volcanic ash. Further, all these methods that are able to analyze the uncertainty of the volcanic ash forecast need comparable computing time except for the method by *Denlinger et al.* [2012] because of the limited amount of parallelism. However, the methods by *Bursik et al.* [2012] and *Zidikheri et al.* [2016] need approx. a factor of 10 more simulations to calculate the analysis. Remarkably, among these methods only ESIAS-chem makes no assumptions of the emissions, thus providing a highly flexible approach, which is also applicable to longer lasting volcanic eruptions and to eruptions with multi-layered emission profiles and highly variable emission strength.

In addition to ESIAS-chem, two other promising analysis methods are tested for the estimation of volcanic ash concentrations and its uncertainty. The first analysis method, which is newly developed, is a combination of a particle smoother and 4D-var (PF/DA in Tab. 7.1; see appendix A for further details on the method). In this combination, the gradient of the cost function of the 4D-var iterations is used in order to calculate time and height dependent weights for the emissions. The second method is a localized particle filter proposed by *Poterjoy* [2016] (LPF in Tab. 7.1; see appendix B for a short description of this method). The LPF method does not estimate the emission profile but it adjusts the volcanic ash concentrations of the ensemble members according to local weights in the vicinity of observations. Generally, this locality of the analysis makes it possible to investigate temporal variable emissions.

Both methods can be used for uncertainty quantification. The LPF methods requires less simulations in order to calculate the analysis and is fully parallel in running the simulations. Thus, it requires only limited compute resources and provides early analyses in the application to a real volcanic eruption. As the PF/DA method solves the optimization iteratively, the parallel section of the simulations is only of the order of 20 %. Hence, the PF/DA method needs more time to calculate the analysis, which limits the applicability to real volcanic eruptions. The two methods PF/DA and LPF represent the observation error explicitly, but the background error is only represented by the a priori ensemble spread. Both methods are able to analyze temporal variable emissions but only the PF/DA method is able to retrieve the vertical distribution of the emissions. In contrast, in the LPF method, for each ensemble member an emission profile is defined prior to the analysis. This emission profile is not altered as observations are assimilated. Thus, the a priori representation of the emissions is crucial for this method.

For the PF/DA method, the diffusion of the gradient of the cost function limits the performance of the analysis. This diffusion reduces the variance in the signals each ensemble member receives in the gradient, which leads to a false weighting of the ensemble members. Hence, the method suffers from a biased estimate of the volcanic ash concentrations.

Although the PF/DA and LPF methods show the general applicability to volcanic eruptions, ESIAS-chem proved to outperform both methods in estimating the

volcanic ash concentrations and its uncertainty. As is suggested in the literature (e.g. *Durand and Alliot* [1999]; *Chelouah and Siarry* [2003]), it was tested to initialize the Nelder-Mead minimization by a genetic algorithm. It was recognized, that the extension of the Nelder-Mead method to the discrete ensemble method makes the use of the genetic algorithm obsolete for this application. The use of discrete solutions enables the rapid convergence of the variable, which is to be optimized, to the minimum.

Chapter 8

Conclusion and outlook

In this study, a new method for estimating volcanic ash emissions and its uncertainty from column mass loadings observations is developed. On the one hand, the method comprises an ensemble-based particle smoother, which extends the assimilation window if new observations are available. Thus, the latest observations are taken for the estimation of the emission profile. On the other hand, the Discrete Ensemble Nelder-Mead (DENM) method is developed in order to achieve an efficient a priori ensemble representation of the optimum of the cost function. The DENM combines the parallel version of the Nelder-Mead method with adaptive parameters and, additionally, extends the pure combination of these Nelder-mead methods to an ensemble formulation with discrete solutions. This extension is shown to increase the performance of the optimization tremendously because the solution space is reduced by the discretization. The ensemble representation increases the diversity of the solution set, which serves as input for the particle smoother.

The use of the DENM minimization method enables the optimization of the emission profiles by a linear combination of pairwise distinct emission packages. Thus, measurements of volcanic ash column mass loadings can be used to optimize the emission profile in a high temporal and vertical resolution. Therefore, the emission profile is separated into single emission packages, each for a single height and unit duration. The emitted volcanic ash packages are transported to the observation time, where the impact of each package on the observations is measured. The DENM minimization combines the volcanic ash packages in order to minimize the distance to the observations. The weight for each volcanic ash package is applied to the corresponding emission package for the generation of the analysis. When the latest observations become available, the assimilation window is extended. Thus, the latest observations, which have the largest information content with respect to the separation of volcanic ash that results from emissions on different heights and at different times, are added to the minimization. It is shown that, in general, the longer the assimilation window, the better is the analysis of volcanic ash. The method is integrated into the new ensemble version of the EURAD-IM, forming the chemical part of ESIAS (ESIAS-chem). ESIAS-chem is flexible in integrating other modules and is applicable to other atmospheric models as well.

The performance of ESIAS-chem is tested by identical twin experiments. These are performed with a coarse grid of 45 km horizontal resolution and 23 vertical σ-layers in order to reduce the computational complexity due to restricted computer facilities. For applications to real volcanic eruptions, a higher horizontal and vertical resolution is recommended to be able to resolve fine volcanic ash structures in the atmosphere. For the identical twin experiments, the horizontal and vertical resolution is of minor importance.

The dependence of the analysis on the wind field for varying assimilation window lengths is investigated in the context of identical twin experiments. On two days with strong and weak winds over the volcano, a nature run emission profile containing two short explosive eruptions are analyzed by ESIAS-chem with assimilation window lengths of 06-36 hours. Thus, the influence of assimilating later observations depending on the wind conditions is analyzed. It is shown that the analysis of the column mass loadings improves if later observations are assimilated, i.e. if the time between observations and emissions increases. For the strong wind case, it is found that once the volcanic ash is mixed in the atmosphere, the analysis cannot be further improved by later observations. In the strong wind case, it is also possible to retrieve the vertical distribution of the volcanic ash from the assimilated column mass loadings. This includes the detection of a second maximum in the vertical distribution of the volcanic ash concentrations, which is not possible for most other existing analysis methods for volcanic ash concentrations and its uncertainty. The relative mean absolute error of the analyzed volcanic ash concentrations for this test case is shown to be of the order of 15 % – 20 %. For the weak wind case, the analyzed column mass loadings are in good agreement with the observations as well. The relative mean absolute error of the volcanic ash concentrations of the ensemble mean to the nature run volcanic ash concentrations is larger than for the strong wind case but still of the order of 20 %.

The analyzed emission profile differs between the two test cases. Besides some temporal and vertical diffusion, the emission profile for the strong wind case is comparable to the nature run emission profile. In contrast, for the weak wind case the two explosive eruptions in the nature run emission profile cannot be resolved by ESIAS-chem. However, the total mass emitted by the volcanic eruption for both test cases is similar to the nature run emissions with less than 10 % error. Thus, the strong wind case shows the potential of ESIAS-chem to analyze the emission profile from observations of column mass loadings. For weak wind conditions, the analysis of the emissions from column mass loadings is limited but can partly be improved by increasing the assimilation window length.

The performance of ESIAS-chem is investigated by an identical twin experiment comprising 20 days of successive volcanic eruptions on April 14-18 and May 04-18, 2010. The assimilation takes observations over a 24 hour period into account and starts every 12 hours, which is a reasonable choice for real volcanic eruptions. The observations are column mass loadings over the full domain, including zero values, every six hours. The observations taken for the performance analysis are profiles of

volcanic ash concentrations at 33 hypothetical lidar stations located at the Earlinet sites.

The error of the ensemble mean is shown to be larger than the ensemble spread, hence, the ensemble is underdispersive. This indicates that the analysis system is not able to represent the forecast uncertainty to full extend. An analysis of the temporal distribution of the relative mean error proves that for strong winds the system is able to analyze the volcanic ash concentrations. If the relative mean error is increased, the winds at the volcano are weaker. This shows that the assimilation window length of 24 hours is too short for a reliable forecast for this case. Longer assimilation windows are favorable. The optimal assimilation window length depends on the actual wind field (e.g. wind speed and areas of convergence) and the emission profile (e.g. temporal and vertical distribution of the emissions). Therefore, no general rule for the optimal length of the assimilation window can be provided. However, an error correction using observations of column mass loadings is applied to the volcanic ash concentrations, which is proved by several statistics to achieve on average reliable and accurate forecasts of volcanic ash concentrations from observations of vertically integrated volcanic ash column mass loadings. Generally, the error correction can be applied during the calculation of the analysis with minimum efforts because it corrects for the relative mean error in the column mass loadings estimate. The uncertainty of estimating the vertical distribution of the volcanic ash, which is measured by the height of the maximum volcanic ash concentration at the 33 analyzed lidar stations, is smaller than the vertical resolution of ESIAS-chem. Hence, the vertical distribution of the volcanic ash can be retrieved by column mass loadings observations using ESIAS-chem.

ESIAS-chem is compared to other methods for estimating volcanic ash and its uncertainty given in the literature. It is shown that ESIAS-chem combines the positive features of several other methods. It is able to estimate temporal and vertical variable emissions, it provides estimates of the emissions uncertainty, and it represents the observation and background error covariances within the minimization procedure. Besides this, it is highly parallelized such that the main of the moderate number of simulations needed within the assimilation can be run in parallel. In contrast to ESIAS-chem, two other methods are shown to suffer from limitations in the estimation of volcanic ash concentrations. The first method is a localized particle filter proposed by *Poterjoy* [2016], which can estimate the vertical distribution of volcanic ash in the atmosphere only if the initial choice of the emission profiles of the ensemble is appropriate. The second method is a new developed combination of a particle smoother and 4D-var data assimilation, which uses the profile of the gradient of the cost function at the volcano to calculate the ensemble member weights. This method suffers from a temporal and vertical diffusion of the signal in the gradient of the cost function yielding to incorrect weights of the ensemble members.

In a next step, the model will be tested for a real eruption of the Eyjafjallajökull volcano in 2010. Therefore, an additional consideration of the uncertainty in the

wind field will be applied. This uncertainty is provided by the meteorological part of the ESIAS system and can easily be applied to ESIAS-chem. By considering uncertain wind fields, the model is able to give estimates on the uncertainty of the horizontal displacement of the volcanic ash. Additionally, the representation of the emission factor error covariances will be investigated and optimized. This will further improve the reliability of the probabilistic forecast of volcanic ash concentrations as the ensemble spread will increase, especially for weak wind conditions.

Besides volcanic ash eruptions, ESIAS-chem is applicable to a variety of emission scenarios, especially unexpected emission events. Therefore, it provides a fast and efficient model for source term estimation including uncertainty representation. In principle, the method can be adapted to multi-source emission scenarios. The enhanced need for compute resources can partly be absorbed by a reduced resolution of the emission profile and will be in the focus of future work.

Appendix A

A combined 4D-var / particle smoother for volcanic ash emission estimation

In Four-Dimensional VARiational data assimilation (4D-var), the gradient of the cost function (cf. 3.1 for the observational part) is used to calculate the best estimate of the model state iteratively

$$\nabla J = \mathbf{B}^{-1}\left[\mathbf{x} - \mathbf{x}_b\right] + \sum_{i=1}^{N} \mathbf{M}^T \mathbf{H}^T \mathbf{R}^{-1}\left[\mathbf{H}(\mathbf{M}_i(\mathbf{x}_0)) - y_i\right], \tag{A.1}$$

with $\mathbf{H}^T$ and $\mathbf{M}^T$ denoting the transposed, or adjoint, observation and model operator, respectively. The adjoint observation operator $\mathbf{H}^T$ maps the innovation vector $[\mathbf{H}(\mathbf{M}_i(\mathbf{x}_0)) - y_i]$, normalized by the observation error covariance matrix, into model space. This includes a redistribution of the information obtained by the column mass loading measurements on vertical model layers. In this application, this is done by setting a Gaussian with standard deviation $\sigma = 0.5$ onto the model layers, which contain volcanic ash in the a priori model state, and weighting the innovation vector with this set of Gaussians.

In this combination of 4D-var with the particle filter method, the profile of the gradient (A.1) of the cost function at the volcano location is used for the generation of the weights in (2.5). The idea is that, similarly to the forward model integration, the adjoint model will transport the gradient backward in time and space to the volcano. The profile of the gradient of the cost function at the volcano can be taken for emission optimization. This gradient, which is transformed into a time/height field at the volcano location, contains information of times and height at which the volcanic ash emissions are too high/low. If the model state is close to the observations, the gradient of the cost function at the corresponding emission time and height in the profile at the volcano is small. Therefore, small cost function values contribute to large weights and vice versa.

In practice, one iteration of 4D-var is performed for an ensemble of N_{ens} model

runs. Instead of running the minimization procedure, each ensemble member is assigned a profile of weights that correspond to the profile of the gradient of the cost function at the volcano. These weights are used in the resampling step. It was recognized that running this process for several iterations, the results were best. Therefore, in the resampling step the emissions are not replaced by the resampled ensemble members but are a weighted mean of the emissions of the replaced and better fitted ensemble members. Initially, each ensemble member contains constant emissions in time and height, where the emission strength differs among the ensemble members.

Appendix B

Localized particle filter

The localized particle filter proposed by *Poterjoy* [2016] aims to reduce the dimension of the data assimilation problem as the weights of the ensemble members are calculated locally. The following description follows *Poterjoy* [2016]. The update of the model state $\mathbf{x}_i^{(y_t)}$ of ensemble member i is equal to the update of the bootstrap particle filter $\mathbf{x}_{k_i}^{(y_{t-1})}$ (cf. section 2.1) distant to the observations. Herein, the index k_i indicates that the model state is possibly resampled after the filtering step, but not perturbed. Therefore, multiple copies of $\mathbf{x}_{k_i}^{(y_{t-1})}$ may exist. Close to the observations, the update of the model state is equal to the model state $\mathbf{x}_i^{(y_{t-1})}$ prior to the assimilation of the current observation to which the local weights are applied. This results in the update equation

$$\mathbf{x}_i^{(y_t)} = \overline{\mathbf{x}}^{(y_t)} + \mathbf{r}_1 \circ (\mathbf{x}_{k_i}^{(y_{t-1})} - \overline{\mathbf{x}}^{(y_t)}) + \mathbf{r}_2 \circ (\mathbf{x}_i^{(y_{t-1})} - \overline{\mathbf{x}}^{(y_t)}), \tag{B.1}$$

with $\circ$ indicating the Schur product and

$$\begin{aligned} r_{1,j} =& \frac{\sigma_j^{(y_t)}}{\sqrt{\frac{1}{N_{ens}-1}\sum_{i=0}^{N_{ens}-1}\left[x_{k_i,j}^{(y_{t-1})} - \overline{x}_j{}^{(y_t)} + c_j\left(x_{i,j}^{(y_{t-1})} - \overline{x}_j{}^{(y_t)}\right)\right]^2}}, \\ r_{2,j} =& c_j r_{1,j}, \\ c_j =& \frac{N_{ens}\left(1 - l[x_j, y_t, r]\right)}{l[x_j, y_t, r]\tilde{W}}. \end{aligned} \tag{B.2}$$

Herein, the localization function is defined by (cf. (4.10) in *Gaspari and Cohn* [1999])

$$l[x_j, y_t, r] = \begin{cases} -0.25d_j^5 + 0.5d_j^4 + \frac{5}{8}d_j^3 - \frac{5}{3}d_j^2 + 1, & 0 \leq d_j < 1, \\ \frac{1}{12}d_j^5 - 0.5d_j^4 + \frac{5}{8}d_j^3 + \frac{5}{3}d_j^2 - 5d_j + 4 - \frac{2}{3d_j}, & 1 \leq d_j < 2, \end{cases} \tag{B.3}$$

with d_j being the Euclidean distance of the location of model state x_i to observation y_t. In (B.2) the a posteriori mean and standard deviation are given by

$$\begin{aligned} \overline{x}_j^{(y_t)} &= \sum_{i=0}^{N_{ens}-1} \frac{\omega_{i,j}^{(y_t)}}{\Omega_j^{(y_t)}} x_{i,j}^{(y_0)}, \\ \sigma_j &= \sum_{i=0}^{N_{ens}-1} \frac{\omega_{i,j}^{(y_t)}}{\Omega_j^{(y_t)}} \left[\mathbf{x}_{i,j}^{(y_0)} - \overline{x}_j^{(y_t)} \right], \end{aligned} \tag{B.4}$$

with local weights calculated according to

$$\omega_{i,j}^{(y_t)} = \omega_{i,j}^{(y_{t-1})} \left(\left[p(y_t | x_i^{(y_0)}) - 1 \right] l[x_j, y_t, r]\alpha + 1 \right), \tag{B.5}$$

and

$$\Omega_j^{(y_t)} = \sum_{n=0}^{N_{ens}-1} \omega_{i,j}^{(y_t)}. \tag{B.6}$$

The inflation factor $\alpha < 1$ controls the ability of the algorithm to account for bias in the ensemble by giving more weights to the tails of the probability density function (cf. *Poterjoy and Anderson* [2016]).

Bibliography

Ackermann, I. J., *MADE: Entwicklung und Anwendung eines Aerosol-Dynamikmodells für dreidimensionale Chemie-Transport-Simulationen in der Troposphäre*, Ph.D. Thesis, Institut für Geophysik und Meteorologie der Universität zu Köln, 1997.

Ackermann, I. J., H. Hass, M. Memmesheimer, A. Ebel, F. S. Binkowski and U. Shankar, Modal aerosol dynamics model for Europe: Development and first applications, *Atmos. Environ.*, *32*, 2981–2999, 1998.

Ades, M. and P. J. van Leeuwen, An exploration of the equivalent weights particle filter, *Q. J. R. Meteorol. Soc.*, *139*, (672), 820–840, 2013.

Ades, M. and P. J. van Leeuwen, The equivalent-weights particle filter in a high-dimensional system, *Q. J. R. Meteorol. Soc.*, *141*, (687), 484–503, 2015.

Adurthi, N., P. Singla and T. Singh, The conjugate unscented transform - An approach to evaluate multi-dimensional expectation integrals, in *2012 American Control Conference (ACC)*, 5556–5561, IEEE, Montreal, QC, Canada, 2012.

Anderson, J. L., An ensemble adjustment Kalman filter for data assimilation, *Month. Weath. Rev.*, *129*, (12), 2884–2903, 2001.

Andrieu, C., A. Doucet and R. Holenstein, Particle Markov chain Monte Carlo methods, *J. R. Statist. Soc. B*, *72*, (3), 269–342, 2010.

Arason, P., G. N. Petersen and H. Bjornsson, Altitude estimates of the volcanic plume-top measured by C-band weather radar, 9–17, PANGEA, 2011.

Atkins, E., M. Morzfeld and A. J. Chorin, Implicit particle methods and their connection with variational data assimilation, *Month. Weath. Rev.*, *141*, (6), 1786–1803, 2013.

Baxter, P. J., Impacts of eruptions on human health, in *Encyclopaedia of Volcanoes*, edited by H. Siggurdson, 1035–1043, Academic Press, San Diego, CA, USA, 1999.

Bengtsson, T., P. Bickel and B. Li, Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems, in *Probability and Statistics:*

Essays in Honor of David A. Freedman, edited by D. Nolan and T. Speed, vol 2 of *Collections*, 316–334, Institute of Mathematical Statistics, Beachwood, Ohio, USA, 2008.

Berndt, J., *On the predictability of exceptional error events in wind power forecasting — an ultra large ensemble approach —*, Ph.D. Thesis, University of Cologne, Cologne, 2018.

Berner, J., G. J. Shutts, M. Leutbecher and T. N. Palmer, A spectral stochastic kinetic energy backscatter scheme and its impact on flow–dependent predictability in the ECMWF Ensemble Prediction System, *J. Atmosph. Sci.*, *66*, 603–626, 2009.

Bickel, P., B. Li and T. Bengtsson, Sharp failure rates for the bootstrap particle filter in high dimensions, in *Pushing the limits of contemporary statistics: Contributions in honor of Jayanta K. Ghosh*, vol 3, 318–329, Institute of Mathematical Statistics, 2008.

Bishop, C. H., B. J. Etherton and S. J. Majumdar, Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects, *Month. Weath. Rev.*, *129*, (3), 420–436, 2001.

Bonavita, M., L. Isaksen and E. Hólm, On the use of EDA background error variances in the ECMWF 4D-var, *Q. J. R. Meteorol. Soc.*, *138*, 1540–1559, 2012.

Brömmel, D., W. Frings and B. J. N. Wylie, Extreme-scaling Applications 24/7 on JUQUEEN BlueGene/Q, in *Parallel computing: On the road to exascale, Proceedings of the international conference on parallel computing, ParCo 2015, 1-4 September 2015, Edinburgh, Scotland, UK*, 817–826, Amsterdam, 2015.

Buizza, R., M. Miller and T. N. Palmer, Stochastic representation of model uncertainties in the ECMWF ensemble prediction system, *Q. J. R. Meteorol. Soc.*, *125*, 2887–2908, 1999.

Bursik, M., M. Jones, S. Carn, K. Dean, A. Patra, M. Pavolonis, E. B. Pitman, T. Singh, P. Singla, P. Webley, H. Bjornsson and M. Ripepe, Estimation and propagation of volcanic source parameter uncertainty in an ash transport and dispersal model: Application to the Eyjafjallajökull plume of 14-16 April 2010, *Bull Volc.*, *74*, 2321–2338, 2012.

Casadevall, T. J., The 1989-1990 eruption of Redoubt volcano, Alaska: Impacts on aircraft operations, *J. Volc. Geotherm. Res.*, *62*, (1-4), 301–316, 1994.

Casati, B., L. J. Wilson, D. B. Stephenson, P. Nurmi, A. Ghelli, M. Pocernich, U. Damrath, E. E. Ebert, B. G. Brown and S. Mason, Forecast verification: Current status and future directions, *Meteorol. Appl.*, *15*, (1), 3–18, 2008.

Chelouah, R. and P. Siarry, Genetic and Nelder-Mead algorithms hybridized for a more accurate global optimization of continuous multiminima functions, *Europ. J. Operat. Research*, *148*, 335–348, 2003.

Chorin, A. J. and X. Tu, Implicit sampling for particle filters, in *Proceedings of the National Academy of Sciences of the United States of America*, vol 106, 17249–17254, Nat. Acad. Sci., 2009.

Chorin, A., M. Morzfeld and X. Tu, Implicit particle filters for data assimilation, *Comm. Appl. Math. and Comp. Sci.*, *5*, (2), 221–240, 2010.

Clayton, A. M., A. C. Lorenc and D. M. Barker, Operational implementation of a hybrid ensemble/4d-var global data assimilation system at the MetOffice, *Q. J. R. Meteorol. Soc.*, *139*, (675), 1445–1461, 2013.

Daley, R., *Atmospheric Data Analysis*, Cambridge Univ. Press, 1991.

Denlinger, R. P., M. Pavolonis and J. Sieglaff, A robust method to forecast volcanic ash clouds, *J. Geophys. Res.*, *117*, (D13208), 2012.

Devenish, B. J., P. N. Francis, B. T. Johnson, R. S. J. Sparks and D. J. Thomson, Sensitivity analysis of dispersion modeling of volcanic ash from Eyjafjallajökull in May 2010, *J. Geophys. Res.*, *117*, (D00U21), 2012.

Douc, R., O. Cappé and E. Moulines, Comparison of resampling schemes for particle filtering, in *ISPA 2005. Proceedings of the 4th International Symposium on Image and Signal Processung and Analysis*, 64–69, Zagreb, 2005.

Doucet, A., N. de Freitas and N. Gordon, *Sequential Monte-Carlo methods in practice*, Springer, 2001.

Durand, N. and J.-M. Alliot, A combined Nelder-Mead simplex and genetic algorithm, in *Proceedings of the genetic and evolutionary computation conference GECCO 1999*, edited by W. Banzhaf, J. Daida, A. E. Eiben, M. H. Garzon, V. Honavar, M. Jakiela, and R. E. Smith, Orlando, FL, USA, 1999.

Eckhardt, S., A. J. Prata, P. Seibert, K. Stebel and A. Stohl, Estimation of the vertical profile of sulfure dioxide injection into the atmosphere by a volcanic eruption using satellite column meansurements and inverse transport modeling, *Atmos. Chem. Phys.*, *8*, 3881–3897, 2008.

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, 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, *Atmos. Chem. Phys.*, *7*, 1–59, 2007.

Evensen, G., Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*, (C5), 10143–10162, 1994.

Folch, A., A. Costa and G. Macedonio, FPLUME-1.0: An integral volcanic plume model accounting for ash aggregation, *Geosci. Model Dev.*, *9*, (1), 431–450, 2016.

Francis, P. N., M. C. Cooke and R. W. Saunders, Retrieval of physical properties of volcanic ash using Meteosat: A case study from the 2010 Eyjafjallajökull eruption, *J. Geophys. Res.*, *117*, (D00U09), 2012.

Gao, F. and L. Han, Implementing the Nelder-Mead simplex algorithm with adaptive parameters, *Comput. Optim. Appl.*, *51*, (1), 259–277, 2012.

Gaspari, G. and S. E. Cohn, Construction of correlation functions in two and three dimensions, *Q. J. R. Meteorol. Soc.*, *125*, 723–757, 1999.

Gasteiger, J., S. Groß, V. Freudenthaler and M. Wiegner, Volcanic ash from Iceland over Munich: Mass concentration retrieved from ground-based remote sensing measurements, *Atmos. Chem. Phys.*, *11*, 2209–2223, 2011.

Gordon, N. J., D. J. Salmond and A. F. M. Smith, Novel approach to nonlinear/non-gaussian bayesian state estimation, in *IEE Proceedings-F (Radar and Signal Processing)*, vol 140, 107–113, 1993.

Guffanti, M., G. C. Mayberry, T. J. Casadevall and R. Wunderman, Volcanic hazards to airports, *Natural Hazards*, *51*, (2), 287–302, 2009.

Han, L., *Algorithms for unconstrained optimization*, Ph.D. Thesis, University of Connecticut, 2000.

Han, L. and M. Neumann, Effect of dimensionality on the Nelder-Mead simplex method, *Optim. Meth. Software*, *21*, (1), 1–16, 2006.

Hastings, W., Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, *57*, (1), 97–109, 1970.

Hol, J. D., T. B. Schön and F. Gustafsson, On resampling algorithms for particle filters, in *2006 IEEE Nonlinear statistical signal processing workshop*, 79–82, Sept. 2006.

Houtekamer, P. L. and H. L. Mitchell, Data assimilation using an ensemble Kalman filter technique, *Month. Weath. Rev.*, *126*, 796–811, 1998.

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.

Jülich Supercomputing Centre, JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre, *Journal of large-scale research facilities*, *1*, (A1), 2015.

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

Klein, K. and J. Neira, Nelder-Mead simplex optimization routine for large-scale problems: A distributed memory implementation, *Comput. Econ.*, *43*, (4), 447–461, 2014.

Kristiansen, N. I., A. Stohl, A. J. Prata, N. Bukowiecki, H. Dacre, S. Eckhardt, S. Henne, M. C. Hort, B. T. Johnson, F. Marenco, B. Neininger, O. Reitebuch, P. Seibert, D. J. Thomson, H. N. Webster and B. Weinzierl, Performance assessment of a volcanic ash transport model mini-ensemble used for inverse modeling of the 2010 Eyjafjallajökull eruption, *J. Geophys. Res.*, *117*, (D00U11), 2012.

Kristiansen, N. I., A. J. Prata, A. Stohl and S. A. Carn, Stratospheric volcanic ash emissions from the 13 February 2014 Kelut eruption, *Geophys. Res. Lett.*, *42*, 588–596, 2015.

Krol, M., P. J. van Leeuwen and J. Lelieveld, Global OH trend inferred from methylchloroform measurements, *J. Geophys. Res.*, *103*, (D9), 10697–10711, 1998.

Lagarias, J. C., J. A. Reeds, M. H. Wright and P. E. Wright, Convergence properties of the Nelder-Mead simplex method in low dimensions, *SIAM J. Optim.*, *9*, (1), 112–147, 1998.

Lecointre, J., K. Hodgson, V. Neall and S. Cronin, Lahar-triggering mechanisms and hazard at Ruapehu volcano, New Zealand, *Natural Hazards*, *31*, (1), 85–109, 2004.

Lee, D. and M. Wiswall, A parallel implementation of the simplex function minimization routine, *Comput. Econ.*, *30*, 171–187, 2007.

Li, T., S. Sun, T. P. Sattar and J. M. Corchado, Fighting sample degeneracy and impoverishment in particle filters: A review of intelligent approaches, *Expert Sys. w. Appl*, *41*, 3944–3954, 2014.

Liu, J. S. and R. Chen, Sequential Monte Carlo methods for dynamic systems, *J. Americ. Stat. Assoc.*, *93*, (443), 1032–1044, 1998.

Liu, D. C. and J. Nocedal, On the limited memory BFGS method for large scale optimization, *Math. Programming*, *45*, 503–528, 1989.

Liu, C., Q. Xiao and B. Wang, An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test, *Month. Weath. Rev.*, *136*, (9), 3363–3373, 2008.

Lu, S., H. X. Lin, A. W. Heemink, G. Fu and A. J. Segers, Estimation of volcanic ash emissions using trajectory-based 4d-var data assimilation, *Month. Weath. Rev.*, *144*, (2), 575–589, 2016.

Madankan, R., S. Pouget, P. Singla, M. Bursik, J. Dehn, M. Jones, A. Patra, M. Pavolonis, E. B. Pitman, T. Singh and P. Webley, Computation of probabilistic hazard maps and source parameter estimation for volcanic ash transport and dispersion, *J. Comp. Phys.*, *271*, 39–59, 2014.

Marécal, V., V.-H. Peuch, C. Andersson, S. Andersson, J. Arteta, M. Beekmann, A. Benedictow, R. Bergström, B. Bessagnet, A. Cansado, F. Chéroux, A. Colette, A. Coman, R. L. Curier, H. A. C. D. van der Gon, A. Drouin, H. Elbern, E. Emili, R. J. Engelen, H. J. Eskes, G. Foret, E. Friese, M. Gauss, C. Giannaros, J. Guth, M. Joly, E. Jaumouillé, B. Josse, N. Kadygrov, J. W. Kaiser, K. Krajsek, J. Kuenen, U. Kumar, N. Liora, E. López, L. Malherbe, I. M. Marco, D. Melas, F. Meleux, L. Menut, P. Moinat, T. M. Morín, J. Parmentier, A. Piacentini, M. Plu, A. Poupkou, S. Queguiner, L. Robertson, L. Rouïl, M. Schaap, A. Segers, M. Sofiev, L. Tarasson, M. Thomas, R. Timmermans, A. Valdebenito, P. van Velthoven, R. van Versendaal, J. Vira and A. Ung, A regional air quality forecasting system over Europe: the MACC-II daily ensemble production, *Geosci. Model Dev.*, *8*, (9), 2777–2813, 2015.

Mastin, L. G., M. Guffanti, R. Servranckx, P. Webley, S. Barsotti, K. Dean, A. Durant, J. W. Ewert, A. Neri, W. I. Rose, D. Schneider, L. Siebert, B. Stunder, G. Swanson, A. Tupper, A. Volentik and C. F. Waythomas, A multidisciplinary effort to assign realistic source parameters to models of volcanic ash–cloud transport and dispersion during eruptions, *J. Volc. Geotherm. Res.*, *186*, (1-2), 10–21, 2009.

McKinnon, K. I. M., Convergence of the Nelder-Mead simplex method to a nonstationary point, *SIAM J. Optim.*, *9*, (1), 148–158, 1998.

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), 1–42, 1982.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth and A. H. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.*, *21*, 1087–1091, 1953.

Molteni, F., R. Buizza, T. N. Palmer and T. Petroliagis, The ECMWF ensemble prediction system: Methodology and validation, *Q. J. R. Meteorol. Soc.*, *122*, (529), 73–119, 1996.

Morzfeld, M., X. Tu, E. Atkins and A. J. Chorin, A random map implementation of implicit filters, *J. Comp. Phys.*, *231*, (4), 2049–2066, 2012.

Nelder, J. A. and R. Mead, A simplex method for function minimization, *Comp. J.*, *7*, (4), 308–313, 1965.

Oxford-Economics, The economic impacts of air travel restrictions due to volcanic ash, 2010.

Penny, S. G. and T. Miyoshi, A local particle filter for high-dimensional geophysical systems, *Nonlinear Proc. in Geophys.*, *23*, (6), 391–405, 2016.

Pitt, M. K. and N. Shephard, Filtering via simulation: Auxiliary particle filters, *J. Americ. Stat. Assoc.*, *94*, (446), 590–599, 1999.

Poterjoy, J., A localized particle filter for high-dimensional nonlinear systems, *Month. Weath. Rev.*, *144*, (1), 59–76, 2016.

Poterjoy, J. and J. L. Anderson, Efficient assimilation of simulated observations in a high-dimensional geophysical system using a localized particle filter, *Month. Weath. Rev.*, *144*, (5), 2007–2020, 2016.

Robock, A. and C. Oppenheimer, *Volcanism and the earth's atmosphere*, American Geophysical Union, Geophysical Monograph, 2003.

Schmehl, K. J., S. E. Haupt and M. J. Pavolonis, A genetic algorithm variational approach to data assimilation and application to volcanic emissions, *Pure Appl. Geophys.*, *169*, (3), 519–537, 2012.

Shutts, G., A kinetic energy backscatter algorithm for use in ensemble prediction systems, *Q. J. R. Meteorol. Soc.*, *131*, 3079–3102, 2005.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X.-Y. Huang, W. Wang and J. G. Powers, A description of the advanced research WRF version 3, technical report, NCAR Technical note NCAR/TN-475+STR, June 2008.

Snyder, C., T. Bengtsson, P. Bickel and J. Anderson, Obstacles to high-dimensional particle filtering, *Month. Weath. Rev.*, *136*, (12), 4629–4640, 2008.

Sparks, R. S. J., M. I. Bursik, S. N. Carey, J. Gilbert, L. S. Glaze, H. Sigurdsson and A. Woods, *Volcanic plumes*, John Wiley, Chichester, N. Y., USA, 1997.

Stefanescu, E. R., A. K. Patra, M. I. Bursik, R. Madankan, S. Pouget, M. jones, P. Singla, T. Singh, E. B. Pitman, M. Pavolonis, D. Morton, P. Webley and J. Dehn, Temporal, probabilistic mapping of ash clouds using wind field stochstic variability and uncertain eruption source parameters: Example of the 14 April 2010 Eyjafjallajökull eruption, *J. Adv. Model. Earth Syst.*, *6*, (4), 1173–1184, 2014.

Stephan, M. and J. Docter, JUQUEEN: IBM BlueGene/q supercomputer system at the Jülich Supercomputing Centre, *J. Large-Scale Res. Facil.*, *1*, (A1), 2015.

Stohl, A., A. J. Prata, S. Eckhardt, L. Clarisse, A. Durant, S. Henne, N. I. Kristiansen, A. Minikin, U. Schumann, P. Seibert, K. Stebel, H. E. Thomas, T. Thorsteinsson, K. Tørseth and B. Weinzierl, Determination of time- and height-resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: the 2010 Eyjafjallajökull eruption, *Atmos. Chem. Phys.*, *11*, (9), 4333–4351, 2011.

Suzuki, T., *Arc volcanism: Physics and tectonics*, chapter A Theoretical model for dispersion of tephra, 95–113, Terra Scientific Publishing Company (TERRAPUB), Tokyo, 1983.

Thordarson, T. and G. Larsen, Volcanism in Iceland in historcal time: Volcano types, eruption styles and eruptive history, *J. Geodyn.*, *43*, (1), 118–152, 2007.

Torczon, V., *Multi-directional search: A direct search algorithm for parallel machines*, Ph.D. Thesis, Rice University, 1989.

Toth, Z. and E. Kalnay, Ensemble forecasting at NCEP and the breeding method, *Month. Weath. Rev.*, *125*, 3297–3319, 1997.

Tseng, P., Fortified-descent simplicial search method: A general appraoch, *SIAM J. Optim.*, *10*, (1), 269–288, 1999.

van Leeuwen, P. J., A variance-minimizing filter for large-scale applications, *Month. Weath. Rev.*, *131*, (9), 2071–2084, 2003.

van Leeuwen, P. J., Particle filtering in geophysical systems, *Month. Weath. Rev.*, *137*, (12), 4089–4114, 2009.

van Leeuwen, P. J., Nonlinear data assimilation in geosciences: an extremely efficient particle filter, *Q. J. R. Meteorol. Soc.*, *136*, (653), 1991–1999, 2010.

van Leeuwen, P. J., vol 8964, chapter Aspects of particle filtering in high-dimensional spaces, Springer, Cham, 2015*a*.

van Leeuwen, P. J., vol 2, chapter Nonlinear data assimilation for high-dimensional systems, Springer, Cham, 2015*b*.

van Leeuwen, P. J. and M. Ades, Efficient fully nonlinear data assimilation for geophysical fluid dynamics, *Comp. Geosci.*, *55*, 16–27, 2013.

Walcek, C. J., Minor flux adjustment near mixing ratio extremes for simplified yet highly accurate monotonic calculation of tracer advection, *J. Geophys. Res.*, *105*, (D7), 9335–9348, 2000.

Wen, S. and W. I. Rose, Retrieval of sizes and total masses of particles in volcanic clouds using AVHRR bands 4 and 5, *J. Geophys. Res.*, *99*, (D3), 5421–5431, 1994.

Wilkins, K. L., S. Mackie, I. M. Watson, H. N. Webster, D. J. Thomson and H. F. Dacre, Data insertion in volcanic ash cloud forecasting, *Ann. Geophys.*, *57*, 1–6, 2014.

Wilkins, K. L., I. M. Watson, N. I. Kristiansen, H. N. Webster, D. J. Thomson, H. F. Dacre and A. J. Prata, Using data insertion with the NAME model to simulate the 8 May 2010 Eyjafjallajökull volcanic ash cloud, *J. Geophys. Res.: Atmos.*, *121*, (1), 306–323, 2016*a*.

Wilkins, K. L., L. M. Western and I. M. Watson, Simulating atmospheric transport of the 2011 Grímsvötn ash cloud using a data insertion update scheme, *Atmos. Environ.*, *141*, 48–59, 2016*b*.

Woodhouse, M. J., A. J. Hogg, J. C. Phillips and R. S. J. Sparks, Interaction between volcanic plumes and wind during the 2010 Eyjafjallajökull eruption, Iceland, *J. Geophys. Res.: Sol. Ea.*, *118*, (1), 92–109, 2013.

Wright, M. H., Direct search methods: Once scorned, now respectable, in *Numerical analysis 1995: Proceedings of the 1995 Dundee biennial conference in numerical analysis*, edited by D. F. Griffiths and G. A. Watson, Harlow U. K.: Addison Wesley Longman, 1996.

Yanenko, N. N., *The method of fractional steps: solution of problems of mathematical physics in several variables*, Springer, 1971.

Yang, Y., C.Robinson, D. Heitz and E. Mémin, Enhanced ensemble-based 4D-var scheme for data assimilation, *Computers and Fluids*, *115*, 201–210, 2015.

Zhang, M. and F. Zhang, E4DVar: Coupling an ensemble Kalman filter with four-dimensional variational data assimilation in a limited-area weather prediction model, *Month. Weath. Rev.*, *140*, (2), 587–600, 2012.

Zidikheri, M. J., R. J. Potts and C. Lucas, A probabilistic inverse method for volcanic ash dispersion modelling, in *Proceedings of the 17th Biennial Computational Techniques and Applications Conference, CTAC-2014*, edited by J. Sharples and J. Bunder, vol 56, C194–C209, 2016.

Zidikheri, M. J., C. Lucas and R. J. Potts, Estimation of optimal dispersion model source parameters using satellite detections of volcanic ash, *J. Geophys. Res.: Atmos.*, *122*, (15), 8207–8232, 2017*a*.

Zidikheri, M. J., C. Lucas and R. J. Potts, Toward quantitative forecasts of volcanic ash dispersial: Using satellite retrievals for optimal estimation of source terms, *J. Geophys. Res.: Atmos.*, *122*, 8187–8206, 2017*b*.

Zupanski, D., S. Q. Zhang, M. Zupanski, A. Y. Hou and S. H. Cheung, A prototype WRF-based ensemble data assimilation system for dynamically downscaling satellite precipitation observations, *J. Hydrometeor.*, *12*, (1), 118–134, 2011.

Acknowledgements/Danksagung

The author gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUQUEEN at Jülich Supercomputing Centre (JSC).

Im Folgenden, möchte ich in meiner Muttersprache an alle Beteiligten meinen tiefsten Dank richten. Zuallererst möchte ich denjenigen danken, die eine direkte Unterstützung zu dieser Arbeit geleistet haben. Zum Einen ist Sebastian Lührs zu nennen, der die Ensembleversion des EURAD-IM initialisiert hat, zum Anderen danke ich Olaf Stein vom SimLab des Jülich Supercomputing Center für seine technische Unterstützung bei der Umsetzung der neuen MPI-Strategie.

Mein ganzer Dank geht an meinen Betreuer Hendrik Elbern für seine Ideen und Hinweise, wenn ich mich einmal gedanklich in einer Sackgasse befunden habe. Des Weiteren möchte ich Yaping Shao für seine Bereitschaft danken, als zweiter Gutachter meiner Arbeit zur Verfügung zu stehen.

Ein ganz besonderer Dank geht an alle ehemaligen und derzeitigen Mitarbeiter des RIU, die mich herzlich Willkommen hießen und bei allen Fragen beratend zur Seite standen. Vor allem außerhalb des Beruflichen habt ihr die Zeit im Institut deutlich verschönert. Besonders hervorheben möchte ich Georg Piekorz, mit dem ich so manchen einsamen frühen Morgen im Institut verbracht habe, Elmar Friese und Lars Nieradzik für ihre Unterstützung beim Kennenlernen des Modells, Anne Caroline Lange, weil sie seit nun mehr sieben Jahren meine Zimmernachbarin ist und die noch nicht einmal auf meine Fragen antworten musste um diese zu beantworten, sowie Johannes Klimpt für seine hilfreichen Ideen und tollen Gespräche über Kultur (Thüringer Klöße) und den Alltag. Des Weiteren geht mein Dank an Jonas Berndt, mit dem ich interessante Diskussionen über die verwendeten Methoden hatte und der mir oft einen anderen Blickwinkel auf meine Probleme gegeben hat. Danke auch an das “Frauenzimmer” im RIU für die tollen Gespräche über Essen und Sprachen. Bei euch im Zimmer war ich immer ein bisschen im Urlaub.

Ohne meine Eltern und meine Schwester wäre ich heute nicht an dem Punkt, den ich erreicht habe. Sie haben immer an mich geglaubt, mir alles ermöglicht und vor allem geholfen, mein Leben so zu gestalten, dass ich mich in jeder Hinsicht zu den Privilegierten zählen darf.

Zum Schluss kommt das wichtigste im Leben, meine kleine Familie.

Liebe Susanne, danke für deine Hilfe, Zuversicht, Unterstützung an schweren Tagen, Geduld und Rücksicht, für deine Fragen, ob “der Wert null ist”, deine Eiskreationen, die mir so manchen Abend gerettet haben und dein Vertrauen in meine Fähigkeiten. Lieber Jakob, danke für den schönsten Moment des Tages, wenn wir uns nach einem langen Tag wiedersehen. Darauf freue ich mich schon morgens. Lieber Eliah, nun bist du erst so kurz bei uns und musstest schon so viel zurückstecken. Ich verspreche dir, in Zukunft mehr Zeit zu haben. Ich danke dir dafür, dass du immer wieder den Fokus auf die wirklich wichtigen Dinge im Leben gelenkt hast. Ohne euch wäre ich nicht der, der ich bin. Danke.

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit einschlielich 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 Herrn PD. Dr. Hendrik Elbern betreut worden.

Es liegen keine Teilpublikationen vor.

_____

Philipp Franke