Singular Vector Based Targeted Observations of Atmospheric Chemical Compounds

Inaugural–Dissertation

zur

Erlangung des Doktorgrades
der Mathematisch–Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von
Nadine Goris
aus Köln

Berichterstatter: Priv.-Doz. Dr. H. Elbern

Prof. Dr. A. Wahner

Tag der mündlichen Prüfung: 03.02.2011

Abstract

Measurements of the earth's environment provide only sparse snapshots of the state of the system due to their insufficient temporal and spatial density. In face of these limitations, the measurement configurations need to be optimized to get a best possible state estimate. One possibility to optimize the state estimate is provided by targeted observations of sensitive system states, where measurements are of great value for forecast improvements.

In the recent years, numerical weather prediction adapted singular vector analysis with respect to initial values as a novel method to identify sensitive locations. In the present work, this technique was transferred from meteorological to chemical forecast. Besides initial values, emissions were introduced as controlling variables. Since time-variant amounts of emissions continuously act on the chemical evolution, targeting observations of emissions is a challenging task. Alternatively, uncertainties in the amplitude of the diurnal profile of emissions are analyzed, yielding emission factors as target variables. Special operators were designed to address specific questions of atmospheric chemistry, like the VOC- versus NO_x-sensitivity of the ozone formation.

The concept of adaptive observations was studied on two levels of complexity: At first, targeted singular vectors were implemented in a chemistry box model which only treats chemical reaction kinetics. Due to the absence of spatial dimensions, the chemistry box model only provides ranking lists of measurement priorities for different compounds. In the second step, singular vector analysis was implemented in a chemical transport model to additionally determine optimal placements of measurements. Both models have been tested and evaluated by conducting a comprehensive set of case studies. Particular questions of specific interest for the chemical system were examined, where the newly designed operators were applied. Results show large differences in sensitivities of different compounds. Consequently, an optimal measurement configuration benefits from omitting measurements of compounds of low sensitivity. It is demonstrated how targeted observations of chemical compounds depend on the considered simulation interval, meteorological conditions, and the underlying chemical composition. Accomplished studies clearly identify strong differences between meteorological and chemical target areas. These differences reveal the importance of the chemical composition and emphasize the significance of chemical singular vectors for effective campaign planing.

Kurzzusammenfassung

Umweltbeobachtungen liefern aufgrund unzureichender zeitlicher und räumlicher Auflösung nur unvollständige Beschreibungen der Erdatmosphäre und ihrer Komponenten. Um trotz dieser Einschränkungen eine bestmögliche Zustandsabschätzung zu erhalten, ist eine Optimierung der Messkonfiguration erforderlich. Eine Möglichkeit der Optimierung liefern gezielte Beobachtungen von sensitiven Systemzuständen.

In der numerischen Wettervorhersage ist die Singulärwertanalyse hinsichtlich Anfangswerten eine neu eingeführte Methode zur Bestimmung sensitiver Systemzustände. In der vorliegenden Arbeit wurde diese Technik von meteorologischen auf atmosphärenchemische Simulationen übertragen. Neben Anfangswerten wurden Emissionen als Zielvariablen eingeführt. Da die zeitlich variierenden Emissionen kontinuierlich auf die chemische Entwicklung einwirken, sind gezielte Beobachtungen von Emissionen eine anspruchsvolle Aufgabe. Alternativ werden Unsicherheiten in der Amplitude des Emissionsratentagesganges analysiert, die neuen Optimierungsvariablen sind folglich Emissionsfaktoren. Um spezifische Fragestellungen der Atmosphärenchemie bearbeiten zu können, wurden neue Operatoren entwickelt.

Das Singulärwertanalyseverfahren wurde in ein chemisches Boxmodell implementiert, welches ausschließlich die chemische Reaktionskinetik betrachtet. Aufgrund der fehlenden räumlichen Dimension liefert dieses Modell eine Rangfolge der Messwichtigkeit verschiedener chemischer Komponenten. Zur zusätzlichen Identifizierung optimaler Messgebiete wurde die Singulärwertanalyse in ein chemisches Transportmodell eingebaut. Die beiden entwickelten Modelle wurden mit Hilfe umfangreicher Fallstudien getestet und evaluiert. Zur Lösung atmosphärenchemischer Fragestellungen wurden die neu entwickelten Operatoren eingesetzt. Bezüglich der gewählten Zielvorgabe wiesen die chemischen Komponenten erhebliche Sensitivitätsunterschiede auf. Adaptive Messungen von chemischen Verbindungen mit geringer Sensitivität liefern deshalb wenig Informationsgehalt. Eine optimierte Messkonfiguration kann durch Einsparung dieser Messungen bedeutende Ressourcengewinne erzielen. Es wurde gezeigt, dass optimale Messungen für chemische Verbindungen durch das betrachtete Simulationsintervall, meteorologische Bedingungen und das zugrunde liegende chemische Szenario bestimmt sind. Die durchgeführten Studien zeigen, dass sich die optimalen Messgebiete für chemische Verbindungen deutlich von den optimalen transportbestimmten Messgebieten unterscheiden. Diese Unterschiede offenbaren die Wichtigkeit der chemischen Zusammensetzung und somit die Signifikanz von chemischen Singulärvektoren für eine effektive Kampagnenplanung.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

It is a typical feature that measurements of the earth's environment have sparse temporal and spatial density and hence provide only incomplete snapshots of the state of the system. This applies to both in situ observations and retrievals from space borne sensors. Consequently, an optimized configuration of available observation capabilities has to be considered to improve the information content of our monitoring capabilities. Adaptive observations of selected parameters in well defined targeted areas can reduce uncertainty and decrease forecast errors (*Buizza et al.* [2007]). Target areas for most valuable observations are sensitive system states, where small variations of considered input parameters lead to a significant forecast change.

The optimal adaption of observations is a frequently investigated problem in numerical weather prediction. A classical topic are cases of explosive cyclogenesis at the North American east coast, which are often of highest relevance for European weather development and its forecast. Various strategies for targeting observations have been introduced, namely adjoint-sensitivity (*Buizza and Montani* [1999]), ensemble transformation (*Bishop and Toth* [1998]), statistical design (*Berliner et al.* [1998]), the breeding method (*Toth and Kalnay* [1993]), Lyapunov vectors (e.g., *Parker and O.Chua* [1989]) and singular vectors (*Buizza and Palmer* [1993]). In general, the denoted methods use a linear approach to evolve the uncertainties in time, even though the forecast involves nonlinear systems. Singular vectors of the tangent linear model identify the directions of fastest perturbation growth over a finite

time interval. Their application to numerical weather prediction was introduced by *Lorenz* [1965], who estimated the atmospheric predictability of an idealized model by computing the largest error growth. Because of the high computational expenditure, singular vector analyses were applied to realistic meteorological problems only in the late 1980's. Since the largest singular vectors contain the directions of fastest error growth (*Buizza and Palmer* [1993]), they are applied as reasonable tools to initialize ensemble forecasts. Their successful use in the ECMWF Ensemble Prediction System resulted in the first application of targeted singular vectors in a field campaign (*Buizza and Montani* [1999]). Several other field campaigns followed, including FASTEX (Fronts and Atlantic Storm-Track Experiment), NORPEX (North-Pacific Experiment), CALJET (California Land-falling JETs Experiment), the Winter Storm Reconnaissance Programs (WSR99 /WSR00) and NATReC (North Atlantic THORPEX Regional Campaign). *Buizza et al.* [2007] investigated the results of the latter campaigns and stated that targeted observations are more valuable than observations taken in random areas. However, the extend of the impact is strongly dependent on regions, seasons, static observing systems, and prevailing weather regimes. Since singular vector analysis is a well-established method within numerical weather prediction and is proven to be valuable, it is used as analysis method in this work.

The studies described above are dealing with perturbations of meteorological parameters. In atmospheric chemistry, studies attending targeted observations are rare. The earliest stimulus for analyzing uncertainties of the chemical composition was provided by *Khattatov et al.* [1999]. By investigation of the linearized model, Khattatov inferred, that a linear combination of 9 initial species' concentrations is sufficient to adequately forecast the concentrations of the complete set of 19 simulated species 4 days later. Since most instruments measure concentrations of individual species, the determination of linear combinations has only limited practical value. Yet, *Khattatov et al.* [1999] motivated to further examine the sensitivity of the initial chemical composition. *Sandu et al.* [2006] used singular vectors to estimate optimal adaptive measurements for chemical compounds. In this manner, application of the results to measurement strategies is feasible, as already demonstrated in the meteorological campaigns mentioned above. However, *Sandu et al.* [2006] especially focused on the optimal placement of observations, while the question which species are to be measured with priority remains mainly disregarded. In contrast, the intention of the present work goes beyond local optimization and furthermore addresses the problem of optimization with regard to species. Unfortunately, adaptive observations of initial values get less

valuable with growing simulation length. Meanwhile, the effect of emission rates on the final concentration increases. Therefore, singular vector analysis is not only applied with regard to initial values, but moreover with regard to emission rates.

In summary, the present work seeks to give insight into the impact of uncertainties in emission strengths and initial species concentrations. Its objective is the detection of sensitive locations and species for atmospheric chemistry transport models, i.e. to answer the following questions:

Which chemical species have to be measured with priority?

Where is the optimal placement for observations of these components?

In addition, the calculated directions of largest error growth can be utilized for sensitivity studies, to initialize ensemble-forecasts, and to form chemical covariances.

This study is organized as follows: The theory of singular vector analysis is presented in chapter 2, where the application on initial uncertainties and emission factors is described as well as newly introduced special operators. Singular vector analysis is implemented into a zero-dimensional model and into a 3-dimensional model. While the zero-dimensional model only takes the chemical kinetics into account, the 3-dimensional model additionally considers transport processes. In chapter 3 the setup of the adapted models is summarized. The zero-dimensional model is applied to analyze several tropospheric scenarios in chapter 4. Studies with the 3-dimensional model are described in chapter 5. Finally, the results of this work are summarized in chapter 6.

CHAPTER 2

Singular vector analysis

The singular vector analysis applied to a forecast model identifies sensitive system state modifications, where small variations of initial conditions lead to significant forecast changes. The leading singular vector reveals the direction of fastest perturbation growth during a finite time interval.

In this work the singular vector analysis was applied to atmospheric chemical modeling to study the influence of chemical initial concentrations and emissions on the temporal evolution of chemical compounds. In section 2.1 emphasis is placed on initial uncertainties, while in section 2.2 emission factors are addressed. These two parameters have been chosen since they both strongly determine the system's evolution. Meteorological fields, deposition velocities, and boundary conditions are other parameters, to which the evolution of chemical species is sensitive, but they go beyond the scope of this study.

2.1 Uncertainties of initial values

Deterministic chemical forecasts propagate the concentrations of chemical species $\boldsymbol{c} \in \mathbb{R}^n$ (denoted in mass mixing ratios) forward in time. With $\mathcal{M}_{t_I,t_F}$ denoting the model operator starting at initial time t_I and ending at final time t_F, the model solution reads:

$$\boldsymbol{c}(t_F) = \mathcal{M}_{t_I,t_F} \, \boldsymbol{c}(t_I). \tag{2.1}$$

Since chemistry-transport models rely on initial values, which do not exactly match the true chemical state, the initial values have initial errors or uncertainties $\delta\boldsymbol{c}(t_I)$. The problem of finding the most unstable initial uncertainty $\delta\boldsymbol{c}(t_I)$ can be envisaged as the search of the phase space direction $\delta\boldsymbol{c}(t_I)$ which results in maximum error growth.

By applying a first-order Taylor series approximation, the disturbed initial state evolves as follows:

$$
\begin{aligned}
\mathcal{M}_{t_I,t_F}\left[\boldsymbol{c}(t_I)+\delta\boldsymbol{c}(t_I)\right]=\\
\mathcal{M}_{t_I,t_F}\,\boldsymbol{c}(t_I)+\left.\frac{\partial\mathcal{M}_{t_I,t_F}}{\partial\boldsymbol{c}}\right|_{\boldsymbol{c}(t_I)}\delta\boldsymbol{c}(t_I)+\mathcal{O}\left[\delta\boldsymbol{c}(t_I)^2\right].
\end{aligned}
\tag{2.2}
$$

Due to the fact, that the term

$$
\mathbf{L}_{t_I,t_F}:=\left.\frac{\partial\mathcal{M}_{t_I,t_F}}{\partial\boldsymbol{c}}\right|_{\boldsymbol{c}(t_I)}
\tag{2.3}
$$

is linearized at the reference trajectory $\boldsymbol{c}(t)=\mathcal{M}_{t_I,t}\,\boldsymbol{c}(t_I)\ \forall\ t\in[t_I,t_F]$, $\mathbf{L}_{t_I,t_F}$ is termed the tangent-linear model. Considering initial errors sufficiently small to evolve linearly within a given time interval, terms of quadratic or higher order can be neglected:

$$
\mathcal{M}_{t_I,t_F}\left[\boldsymbol{c}(t_I)+\delta\boldsymbol{c}(t_I)\right]\approx\mathcal{M}_{t_I,t_F}\,\boldsymbol{c}(t_I)+\mathbf{L}_{t_I,t_F}\,\delta\boldsymbol{c}(t_I),
\tag{2.4}
$$

and the evolution of initial uncertainties can be described with the tangent linear model dynamics:

$$
\delta\boldsymbol{c}(t_F)=\mathbf{L}_{t_I,t_F}\,\delta\boldsymbol{c}(t_I).
\tag{2.5}
$$

For more details on the derivation of equation (2.5) see, for example, *Kalnay* [2003]. The ratio between perturbation magnitudes at final time t_F and initial time t_I can be used to define a measure of error growth $g(\delta\boldsymbol{c}(t_I))$:

$$
g(\delta\boldsymbol{c}(t_I)):=\frac{\|\delta\boldsymbol{c}(t_F)\|_2}{\|\delta\boldsymbol{c}(t_I)\|_2}
\tag{2.6}
$$

(see *Sandu et al.* [2006] for a comprehensive discussion). Substituting the definition of the Euclidean norm in (2.6) leads to

$$
g(\delta\boldsymbol{c}(t_I))=\sqrt{\frac{\delta\boldsymbol{c}(t_F)^T\delta\boldsymbol{c}(t_F)}{\delta\boldsymbol{c}(t_I)^T\delta\boldsymbol{c}(t_I)}}.
\tag{2.7}
$$

Maximizing this ratio with respect to the initial disturbance $\delta \boldsymbol{c}(t_I)$ provides the direction of maximal error growth $\delta \boldsymbol{c}(t_I)$. As $g(\delta \boldsymbol{c}(t_I)) \geq 0$, the initial perturbation $\delta \boldsymbol{c}(t_I)$ that maximizes the squared error growth $g^2(\delta \boldsymbol{c}(t_I))$, maximizes the error growth $g(\delta \boldsymbol{c}(t_I))$ as well. For convenience the squared error growth is treated henceforth:

$$\max_{\delta \boldsymbol{c}(t_I) \neq 0} g^2(\delta \boldsymbol{c}(t_I)) = \max_{\delta \boldsymbol{c}(t_I) \neq 0} \frac{\delta \boldsymbol{c}(t_F)^T \delta \boldsymbol{c}(t_F)}{\delta \boldsymbol{c}(t_I)^T \delta \boldsymbol{c}(t_I)}. \tag{2.8}$$

Using equation (2.5) the variable $\delta \boldsymbol{c}(t_F)$ may be eliminated from the maximization problem to leave

$$\max_{\delta \boldsymbol{c}(t_I) \neq 0} g^2(\delta \boldsymbol{c}(t_I)) = \max_{\delta \boldsymbol{c}(t_I) \neq 0} \frac{\delta \boldsymbol{c}(t_I)^T \mathbf{L}_{t_I,t_F}^T \, \mathbf{L}_{t_I,t_F} \, \delta \boldsymbol{c}(t_I)}{\delta \boldsymbol{c}(t_I)^T \delta \boldsymbol{c}(t_I)}. \tag{2.9}$$

The operator $\mathbf{L}_{t_I,t_F}^T$ denotes the adjoint model of the tangent-linear operator $\mathbf{L}_{t_I,t_F}$. The adjoint $\mathbf{K}^T$ of a real operator $\mathbf{K}$ is defined by the property

$$\langle \mathbf{K}\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \mathbf{K}^T \boldsymbol{y} \rangle \; \forall \; \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, \tag{2.10}$$

where $\mathbf{K}^T$ identifies the transpose operator and $\langle .,. \rangle$ denotes the canonical Euclidean scalar product (e.g., *Kalnay* [2003]). The operator $\mathbf{L}_{t_I,t_F}^T \, \mathbf{L}_{t_I,t_F}$ is also known as Oseledec operator. Obviously it is symmetric. Hence the ratio (2.9) is a Rayleigh quotient.

Rayleigh's principle

For a symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, a Rayleigh quotient $r^0(\mathbf{A}; \boldsymbol{x})$ with $\boldsymbol{x} \in \mathbb{R}^n$ is defined as

$$r^0(\mathbf{A}; \boldsymbol{x}) := \frac{\boldsymbol{x}^T \mathbf{A} \, \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}, \qquad \boldsymbol{x} \neq 0. \tag{2.11}$$

Let $\lambda_1^0 \geq \lambda_2^0 \geq ... \geq \lambda_n^0$ be the eigenvalues of $\mathbf{A}$ and $\boldsymbol{v}_i^0$ the associated eigenvectors. Rayleigh's principle states that a Rayleigh quotient $r^0(\mathbf{A}; \boldsymbol{x})$ reaches its maximum λ_1^0 when $\boldsymbol{v}_1^0$ is inserted (*Noble and Daniel* [1969]). Furthermore, for a symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ and a symmetric positive-definite matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$, a fraction of the type

$$r_g^0(\mathbf{A}, \mathbf{B}; \boldsymbol{x}) := \frac{\boldsymbol{x}^T \mathbf{A} \, \boldsymbol{x}}{\boldsymbol{x}^T \mathbf{B} \, \boldsymbol{x}}, \qquad \boldsymbol{x} \neq 0 \tag{2.12}$$

is referred to as generalized Rayleigh quotient $r_g^0(\mathbf{A}, \mathbf{B}; \boldsymbol{x})$. With the transformation $\mathbf{C} = \mathbf{B}^{-T/2} \mathbf{A} \mathbf{B}^{-1/2}$, $\boldsymbol{y} = \mathbf{B}^{1/2} \boldsymbol{x}$ the generalized Rayleigh quotient $r_g^0(\mathbf{A}, \mathbf{B}; \boldsymbol{x})$ can be reduced to a Rayleigh quotient $r^0 = (\mathbf{C}; \boldsymbol{y})$ and Rayleigh's

principle can be applied. Since various Rayleigh quotients will be considered in this section, notations (2.11) and (2.12) are adapted to

$$r(\mathbf{D}; \boldsymbol{x}) := \frac{\boldsymbol{x}^T \mathbf{D}^T \mathbf{D} \, \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}, \quad \boldsymbol{x} \neq 0 \tag{2.13}$$

and

$$r_g(\mathbf{D}, \mathbf{E}; \boldsymbol{x}) := \frac{\boldsymbol{x}^T \mathbf{D}^T \mathbf{D} \, \boldsymbol{x}}{\boldsymbol{x}^T \mathbf{E}^T \mathbf{E} \, \boldsymbol{x}}, \quad \boldsymbol{x} \neq 0 \tag{2.14}$$

to describe the subsequent problems more compactly. Thereby $\mathbf{D}^T \mathbf{D}$ and $\mathbf{E}^T \mathbf{E}$ have to be symmetric (which holds for every $\mathbf{D} \in \mathbb{R}^{n \times n}$, $\mathbf{E} \in \mathbb{R}^{n \times n}$) and $\mathbf{E}^T \mathbf{E}$, moreover, has to be positive-definite. For the new definitions, the transformation of the generalized Rayleigh quotient into a Rayleigh quotient changes to:

$$r_g(\mathbf{D}, \mathbf{E}; \boldsymbol{x}) = r(\mathbf{G}; \boldsymbol{y}), \quad \mathbf{G} = \mathbf{D} \mathbf{E}^{-1}, \quad \boldsymbol{y} = \mathbf{E} \, \boldsymbol{x}. \tag{2.15}$$

The maximal value of the Rayleigh quotient $r(\mathbf{D}; \boldsymbol{x})$ (2.14) is the largest eigenvalue λ_1 of the matrix $\mathbf{D}^T \mathbf{D}$. Accordingly, its assigned eigenvector $\boldsymbol{v}_1$ maximizes the fraction $r(\mathbf{D}; \boldsymbol{x})$.

Applying Rayleigh's principle to problem (2.9) results in searching for the largest eigenvalue λ_1 and the assigned eigenvector $\boldsymbol{v}_1(t_I)$ of the following eigenvalue problem:

$$\mathbf{L}_{t_I,t_F}^T \mathbf{L}_{t_I,t_F} \, \boldsymbol{v}(t_I) = \lambda \, \boldsymbol{v}(t_I). \tag{2.16}$$

Since the entire set of eigenvectors $\boldsymbol{v}_i(t_I)$ of $\mathbf{L}_{t_I,t_F}^T \mathbf{L}_{t_I,t_F}$ can be chosen to form an orthonormal basis in the n-dimensional tangent space of linear perturbations, the eigenvectors $\boldsymbol{v}_i(t_I)$, $i=2,...,n$ define secondary directions of instability. The amount of influence of eigenvector $\boldsymbol{v}_i(t_I)$ is defined by the magnitude of the square root of the associated eigenvalue λ_i.

The name *singular vector analysis* refers to the fact, that the square roots of the eigenvalues λ_i of $\mathbf{L}_{t_I,t_F}^T \mathbf{L}_{t_I,t_F}$ are the singular values σ_i of the tangent-linear model $\mathbf{L}_{t_I,t_F}$. The associated left and right singular vectors $\boldsymbol{u}_i(t_F) \in \mathbb{R}^n$ and $\boldsymbol{v}_i(t_I) \in \mathbb{R}^n$ of the operator $\mathbf{L}_{t_I,t_F}$ are defined satisfying the following conditions:

$$\mathbf{L}_{t_I,t_F} \, \boldsymbol{v}_i(t_I) = \sigma_i \, \boldsymbol{u}_i(t_F) \quad \text{and} \tag{2.17}$$

$$\mathbf{L}_{t_I,t_F}^T \, \boldsymbol{u}_i(t_F) = \sigma_i \, \boldsymbol{v}_i(t_I), \tag{2.18}$$

with $\|\boldsymbol{v}\|_2=1$ and $\|\boldsymbol{u}\|_2=1$. Combining these two equations

$$\mathbf{L}_{t_I,t_F}^T \mathbf{L}_{t_I,t_F} \, \boldsymbol{v}_i(t_I) = \sigma_i \, \mathbf{L}_{t_I,t_F}^T \, \boldsymbol{u}_i(t_F) = \sigma_i^2 \, \boldsymbol{v}_i(t_I) \tag{2.19}$$

reveals that the eigenvectors $\boldsymbol{v}_i(t_I)$ of the Oseledec operator are the right singular vectors of the tangent-linear operator $\mathbf{L}_{t_I, t_F}$. Hence, the right singular vector $\boldsymbol{v}_1(t_I)$ assigned to the largest singular value σ_1 of a chemistry-transport model characterizes the direction of maximum error growth over a finite time interval $[t_I, t_F]$. The singular value σ_1 is the maximum value of the original ratio (2.6) and defines the amount of error growth.

2.1.1 Relative error growth

Since concentrations of different species may vary by many orders of magnitude, perturbations of species with larger concentrations or higher reactivity are expected to dominate the error growth. To avoid this effect and gain a relative error growth, the absolute uncertainties $\delta\boldsymbol{c}(t) \ \forall\, t \in [t_I, t_F]$ are scaled by current concentrations $\boldsymbol{c}(t) \ \forall\, t \in [t_I, t_F]$. For this purpose, a weight matrix $\mathbf{W}_t \in \mathbb{R}^{n \times n}$,

$$\mathbf{W}_t := \operatorname{diag}\left(\boldsymbol{c}^{i,j,k,s}(t)\right)_{i,j,k,s} \quad \forall\, t \in [t_I, t_F] \tag{2.20}$$

is introduced, which assigns the concentration of chemical species s to each grid point (i, j, k). Its application provides the relative error $\delta\boldsymbol{c}_{\mathrm{r}} \in \mathbb{R}^n$

$$\delta\boldsymbol{c}_{\mathrm{r}}(t) := \mathbf{W}_t^{-1}\, \delta\boldsymbol{c}(t) \quad \forall\, t \in [t_I, t_F] \tag{2.21}$$

as well as the relative error growth g_{r} with

$$g_{\mathrm{r}}\left(\delta\boldsymbol{c}_{\mathrm{r}}(t_I)\right) := \frac{\|\delta\boldsymbol{c}_{\mathrm{r}}(t_F)\|_2}{\|\delta\boldsymbol{c}_{\mathrm{r}}(t_I)\|_2} = \frac{\|\mathbf{W}_{t_F}^{-1}\, \delta\boldsymbol{c}(t_F)\|_2}{\|\mathbf{W}_{t_I}^{-1}\, \delta\boldsymbol{c}(t_I)\|_2} \tag{2.22}$$

(*Sandu et al.* [2006]). Applying the squared measure and expressing the final perturbation in terms of the initial perturbation

$$\delta\boldsymbol{c}_{\mathrm{r}}(t_F) = \mathbf{W}_{t_F}^{-1}\, \delta\boldsymbol{c}(t_F) = \mathbf{W}_{t_F}^{-1}\, \mathbf{L}_{t_I, t_F} \delta\boldsymbol{c}(t_I) \tag{2.23}$$

leads to a generalized Rayleigh quotient with respect to $\delta\boldsymbol{c}(t_I)$:

$$\begin{aligned} g_{\mathrm{r}}^{2}\left(\delta\boldsymbol{c}_{\mathrm{r}}(t_I)\right) &= r_g(\mathbf{B}, \mathbf{W}_{t_I}^{-1}; \delta\boldsymbol{c}(t_I)), \\ \text{where} \quad \mathbf{B} &:= \mathbf{W}_{t_F}^{-1}\, \mathbf{L}_{t_I, t_F}. \end{aligned} \tag{2.24}$$

By formula (2.15) the generalized Rayleigh quotient (2.24) is transformed into a Rayleigh quotient

$$\begin{aligned} g_{\mathrm{r}}^{2}\left(\delta\boldsymbol{c}_{\mathrm{r}}(t_I)\right) &= r(\mathbf{B}_{\mathrm{r}}; \delta\boldsymbol{c}_{\mathrm{r}}(t_I)), \\ \text{where} \quad \mathbf{B}_{\mathrm{r}} &:= \mathbf{B}\, \mathbf{W}_{t_I} = \mathbf{W}_{t_F}^{-1}\, \mathbf{L}_{t_I, t_F}\, \mathbf{W}_{t_I}, \end{aligned} \tag{2.25}$$

and Rayleigh's principle can be applied. Thus the phase space direction $\delta \boldsymbol{c}_{\mathrm{r}}(t_I)$, for which the ratio (2.25) gains its maximal value, is the solution $\boldsymbol{v}_{\mathrm{r}1}(t_I)$ of the symmetric eigenvalue problem

$$
\begin{aligned}
\mathbf{B}_{\mathrm{r}}^T \, \mathbf{B}_{\mathrm{r}} \; \boldsymbol{v}_{\mathrm{r}}(t_I) &= \lambda_{\mathrm{r}} \, \boldsymbol{v}_{\mathrm{r}}(t_I) \\
\text{with} \quad \boldsymbol{v}_{\mathrm{r}}(t_I) :&= \mathbf{W}_{t_I}^{-1} \, \boldsymbol{v}(t_I)
\end{aligned}
\tag{2.26}
$$

assigned to the largest eigenvalue $\lambda_{\mathrm{r}1}$. Comparing problem (2.26) with the original problem (2.16), it is readily seen that the solution $\boldsymbol{v}_{\mathrm{r}1}(t_I) \in \mathbb{R}^n$ is the right singular vector of the operator $\mathbf{B}_{\mathrm{r}}$ and the square root of the eigenvalue $\lambda_{\mathrm{r}1}$ is the associated singular value $\sigma_{\mathrm{r}1}$. Due to this property the singular vector $\boldsymbol{v}_{\mathrm{r}1}(t_I)$ is called relative singular vector henceforth.

2.1.2   Projected error growth

Another central aim is to examine the error growth of a limited set of chemical species for limited geographical regions. To fulfill this requirement a projection operator $\mathbf{P}_t \in \mathbb{R}^{n \times n}$ is applied, which sets the entries of the perturbations to zero outside the focused species and regions (*Barkmeijer et al.* [1998]). Thus the projection operator is a diagonal matrix with binary entries, dependent on the spatial or chemical feature of interest. The projection operator reads

$$
\mathbf{P}_t := \mathrm{diag}\,(p_i)_i\,, \quad p_i = \begin{cases} 1 & \forall\, i \in \mathcal{P}(t) \\ 0 & \text{otherwise} \end{cases},
\tag{2.27}
$$

where $\mathcal{P}(t)$ denotes the set of selected chemical compounds and grid locations at time t. Using the example of $n{=}5$ and $\mathcal{P}(t){=}\{2,4,5\}$, the projection operator is given by

$$
\mathbf{P}_t = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.
$$

Note, that set $\mathcal{P}(t)$ and projection operator $\mathbf{P}_t$ are time-dependent. They may differ at initial and final time.

In order to consider the impact of a limited perturbation at initial time t_I on a limited perturbation at time t, the projected error $\delta \boldsymbol{c}_{\mathrm{p}} \in \mathbb{R}^n$ is defined as

$$
\delta \boldsymbol{c}_{\mathrm{p}}(t) := \mathbf{P}_t \, \mathbf{L}_{t_I,t} \, \mathbf{P}_{t_I} \, \delta \boldsymbol{c}(t_I).
\tag{2.28}
$$

In case, that there is no limitation of regions and species at time t, $t \in [t_I, t_F]$, the projection operator $\mathbf{P}_t$ equals the identity matrix. Since the projection operator is idempotent and $\mathbf{L}_{t,t} = \mathbf{I}$, the projected error at initial time reads

$$\delta c_{\mathrm{P}}(t_I) = \mathbf{P}_{t_I} \, \mathbf{L}_{t_I,t_I} \, \mathbf{P}_{t_I} \, \delta c(t_I) = \mathbf{P}_{t_I} \, \delta c(t_I). \tag{2.29}$$

After explicit use of equation (2.29), the projected error at final time becomes

$$\delta c_{\mathrm{P}}(t_F) = \mathbf{P}_{t_F} \, \mathbf{L}_{t_I,t_F} \, \delta c_{\mathrm{P}}(t_I). \tag{2.30}$$

Application of equation (2.29) and equation (2.30) amounts to the projected error growth g_{P}:

$$g_{\mathrm{P}} \left(\delta c_{\mathrm{P}}(t_I) \right) := \frac{\|\delta c_{\mathrm{P}}(t_F)\|_2}{\|\delta c_{\mathrm{P}}(t_I)\|_2} = \frac{\|\mathbf{P}_{t_F} \, \mathbf{L}_{t_I,t_F} \, \mathbf{P}_{t_I} \, \delta c(t_I)\|_2}{\|\mathbf{P}_{t_I} \, \delta c(t_I)\|_2}. \tag{2.31}$$

The squared projected error growth reduces to a Rayleigh quotient

$$\begin{aligned} g_{\mathrm{P}}{}^2 \left(\delta c_{\mathrm{P}}(t_I) \right) &= r(\mathbf{B}; \delta c_{\mathrm{P}}(t_I)), \\ \text{where} \quad \mathbf{B} :&= \mathbf{P}_{t_F} \, \mathbf{L}_{t_I,t_F}, \end{aligned} \tag{2.32}$$

subject to

$$[\delta c_{\mathrm{P}}(t_I)](j) = [\mathbf{P}_{t_I} \, \delta c(t_I)](j) = \begin{cases} [\delta c(t_I)](j) & \forall \, j \in \mathcal{P}_{t_I} \\ 0 & \text{otherwise.} \end{cases} \tag{2.33}$$

Here, $[\boldsymbol{x}](j)$ denotes the j-th component of a vector $\boldsymbol{x}$. According to Rayleigh's principle, the phase space direction that maximizes the Rayleigh quotient (2.43) is the solution $\boldsymbol{v}_{\mathrm{P}1}(t_I)$ of the symmetric eigenvalue problem

$$\mathbf{B}^T \mathbf{B} \; \boldsymbol{v}_{\mathrm{P}}(t_I) = \lambda_{\mathrm{P}} \, \boldsymbol{v}_{\mathrm{P}}(t_I) \tag{2.34}$$

assigned to largest eigenvalue $\lambda_{\mathrm{P}1}$. However, the solution $\boldsymbol{v}_{\mathrm{P}1}(t_I)$ does not necessarily ensure condition (2.33). In order to grant condition (2.33), the solution space has to be restricted. Therefore, equation (2.34) is left-multiplied with $\mathbf{P}_{t_I}$ at first:

$$\mathbf{P}_{t_I} \, \mathbf{B}^T \mathbf{B} \, \boldsymbol{v}_{\mathrm{P}}(t_I) = \lambda_{\mathrm{P}} \, \mathbf{P}_{t_I} \, \boldsymbol{v}_{\mathrm{P}}(t_I). \tag{2.35}$$

Application of

$$\mathbf{P}_{t_I} \, \boldsymbol{v}_{\mathrm{P}}(t_I) = \boldsymbol{v}_{\mathrm{P}}(t_I) \tag{2.36}$$

yields

$$\begin{aligned} \mathbf{B}_{\mathrm{P}}{}^T \mathbf{B}_{\mathrm{P}} \; \boldsymbol{v}_{\mathrm{P}}(t_I) &= \lambda_{\mathrm{P}} \; \boldsymbol{v}_{\mathrm{P}}(t_I), \\ \text{with} \quad \mathbf{B}_{\mathrm{P}} :&= \mathbf{B} \, \mathbf{P}_{t_I} = \mathbf{P}_{t_F} \, \mathbf{L}_{t_I,t_F} \, \mathbf{P}_{t_I}. \end{aligned} \tag{2.37}$$

Here, the solution $\boldsymbol{v}_{\mathrm{P}1}(t_I)$ assigned to the largest eigenvalue $\lambda_{\mathrm{P}1}$ holds the required restriction (2.33). The eigenvector $\boldsymbol{v}_{\mathrm{P}1}(t_I) \in \mathbb{R}^n$ is the right singular vector of the operator $\mathbf{B}_{\mathrm{P}}$ and therefore it is referred to as projected singular vector. The square root of the eigenvalue $\lambda_{\mathrm{P}1}$ is the associated projected singular value $\sigma_{\mathrm{P}1}$.

Furthermore, it is possible to combine the projected error (2.28) with the relative error (2.21), yielding the projected relative error $\delta\boldsymbol{c}_{\mathrm{pr}}(t)$:

$$\delta\boldsymbol{c}_{\mathrm{pr}}(t) := \mathbf{W}_t^{-1} \, \delta\boldsymbol{c}_{\mathrm{p}}(t). \tag{2.38}$$

Hence, the projected relative error growth $g_{\mathrm{pr}}\left(\delta\boldsymbol{c}_{\mathrm{pr}}(t_I)\right)$ is given by:

$$g_{\mathrm{pr}}\left(\delta\boldsymbol{c}_{\mathrm{pr}}(t_I)\right) := \frac{\|\delta\boldsymbol{c}_{\mathrm{pr}}(t_F)\|_2}{\|\delta\boldsymbol{c}_{\mathrm{pr}}(t_I)\|_2} = \frac{\|\mathbf{W}_{t_F}^{-1} \, \mathbf{P}_{t_F} \, \mathbf{L}_{t_I,t_F} \, \mathbf{P}_{t_I} \, \delta\boldsymbol{c}(t_I)\|_2}{\|\mathbf{W}_{t_I}^{-1} \, \mathbf{P}_{t_I} \, \delta\boldsymbol{c}(t_I)\|_2}. \tag{2.39}$$

After explicit use of

$$\delta\boldsymbol{c}_{\mathrm{pr}}(t_F) = \mathbf{W}_{t_F}^{-1} \, \mathbf{P}_{t_F} \, \mathbf{L}_{t_I,t_F} \, \delta\boldsymbol{c}_{\mathrm{p}}(t_I), \tag{2.40}$$

the squared projected relative error growth reads as a generalized Rayleigh quotient:

$$g_{\mathrm{pr}}{}^2(\delta\boldsymbol{c}_{\mathrm{pr}}(t_I)) = r_g(\mathbf{B}, \mathbf{W}_{t_I}^{-1}; \delta\boldsymbol{c}_{\mathrm{p}}(t_I)), \tag{2.41}$$

$$\text{where} \quad \mathbf{B} := \mathbf{W}_{t_F}^{-1} \, \mathbf{P}_{t_F} \, \mathbf{L}_{t_I,t_F}. \tag{2.42}$$

Use of transformation (2.15) reduces the generalized Rayleigh quotient into a Rayleigh quotient

$$g_{\mathrm{pr}}{}^2(\delta\boldsymbol{c}_{\mathrm{pr}}(t_I)) = r(\mathbf{B}\mathbf{W}_{t_I}; \delta\boldsymbol{c}_{\mathrm{pr}}(t_I)) \tag{2.43}$$

subject to

$$[\delta\boldsymbol{c}_{\mathrm{pr}}(t_I)](j) = [\mathbf{W}_{t_I}^{-1} \delta\boldsymbol{c}_{\mathrm{p}}(t_I)](j) = \begin{cases} [\frac{\delta\boldsymbol{c}(t_I)}{\boldsymbol{c}(t_I)}](j) & \forall \, j \in \mathcal{P}_{t_I} \\ 0 & \text{otherwise.} \end{cases} \tag{2.44}$$

The phase space direction that maximizes the Rayleigh quotient (2.43), is the solution $\boldsymbol{v}_{\mathrm{P}1}(t_I)$ of the symmetric eigenvalue problem

$$\begin{gathered} \mathbf{W}_{t_I}^T \, \mathbf{B}^T \, \mathbf{B} \, \mathbf{W}_{t_I} \, \boldsymbol{v}_{\mathrm{pr}}(t_I) = \lambda_{\mathrm{pr}} \, \boldsymbol{v}_{\mathrm{pr}}(t_I) \\ \text{with} \quad \boldsymbol{v}_{\mathrm{pr}}(t_I) := \mathbf{W}_{t_I}^{-1} \, \boldsymbol{v}_{\mathrm{p}}(t_I) \end{gathered} \tag{2.45}$$

assigned to the largest eigenvalue $\lambda_{\mathrm{pr}1}$. Since the solution $\boldsymbol{v}_{\mathrm{pr}1}(t)$ needs to hold condition (2.44), the solution space has to be restricted. Therefore, the eigenvalue problem is restated by multiplying with $\mathbf{P}_{t_I}$:

$$\mathbf{P}_{t_I} \, \mathbf{W}_{t_I}^T \, \mathbf{B}^T \, \mathbf{B} \, \mathbf{W}_{t_I} \, \boldsymbol{v}_{\mathrm{pr}}(t_I) = \lambda_{\mathrm{pr}} \, \mathbf{P}_{t_I} \, \boldsymbol{v}_{\mathrm{pr}}(t_I). \tag{2.46}$$

The new generalized eigenvalue problem ensures that condition (2.44) holds. In order to reduce the general eigenvalue problem to a symmetric eigenvalue problem (which is requested for the power method, see section 3.3.1) further modifications have to be made. Application of the fact, that diagonal matrices of the same dimension commute, yields

$$
\begin{aligned}
\mathbf{P}_{t_I}\, \boldsymbol{v}_{\mathrm{pr}}(t_I) &= \mathbf{P}_{t_I}\, \mathbf{W}_{t_I}\, \boldsymbol{v}_{\mathrm{p}}(t_I) \\
&= \mathbf{W}_{t_I}\, \mathbf{P}_{t_I}\, \boldsymbol{v}_{\mathrm{p}}(t_I) = \mathbf{W}_{t_I}\, \boldsymbol{v}_{\mathrm{p}}(t_I) = \boldsymbol{v}_{\mathrm{pr}}(t_I).
\end{aligned}
\tag{2.47}
$$

Substituting relation (2.47) into equation (2.46) gives a symmetric eigenvalue problem:

$$
\begin{aligned}
&\mathbf{B}_{\mathrm{pr}}{}^T\, \mathbf{B}_{\mathrm{pr}}\, \boldsymbol{v}_{\mathrm{pr}}(t_I) = \lambda_{\mathrm{pr}}\, \boldsymbol{v}_{\mathrm{pr}}(t_I), \\
&\text{where} \quad \mathbf{B}_{\mathrm{pr}} := \mathbf{W}_{t_F}^{-1}\, \mathbf{P}_{t_F}\, \mathbf{L}_{t_I, t_F}\, \mathbf{W}_{t_I}\, \mathbf{P}_{t_I}.
\end{aligned}
\tag{2.48}
$$

The solution $\boldsymbol{v}_{\mathrm{pr}_1}(t_I) \in \mathbb{R}^n$ assigned to the largest eigenvalue λ_{pr_1} of the final eigenvalue problem (2.48) is referred to as projected relative singular vector, since it is the right singular vector of the operator $\mathbf{B}_{\mathrm{pr}}$. The square root of the eigenvalue $\lambda_{\mathrm{pr}\,1}$ is the associated projected relative singular value $\sigma_{\mathrm{pr}\,1}$.

2.1.3 Grouped error growth

So far, the effect of each species was regarded independently. Instead of examining the initial disturbances individually, one can also decide to look at the influence of groups of chemical species, which chemically act in a similar manner. To the knowledge of the author this problem has not been addressed in the context of singular vector analysis previously. In order to implement this new approach, a family operator $\mathbf{F}_t \in \mathbb{R}^{n \times n}$ is introduced. Let $\mathcal{F}_k(t)$ represent one of $f(t)$ pairwise disjoint families with $m_k(t)$ members. Adopting the convention that $\mathbf{A}(i, j)$ refers to the entry that lies in the i-th row and the j-th column of a matrix $\mathbf{A}$, each entry of the family operator reads

$$
\mathbf{F}_t(i, j)_{i,j} = \begin{cases} 1 / m_k(t) & \forall\, i, j \in \mathcal{F}_k(t) \\ 0 & \text{otherwise.} \end{cases}
\tag{2.49}
$$

Note, that each chemical species is only allowed to belong to one family, i.e.

$$
\bigcap_{k=1}^{f(t)} \mathcal{F}_k(t) = \varnothing.
\tag{2.50}
$$

An example of a family operator at initial time t_I is

$$\mathbf{F}_{t_I} = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \end{pmatrix}, \tag{2.51}$$

where $n = 6$, $f(t_I) = 2$, $m_1(t_I) = 3$, $m_2(t_I) = 2$, $\mathcal{F}_1(t_I) = \{2, 4, 6\}$ and $\mathcal{F}_2(t_I) = \{1, 5\}$. If each family has only one member, the family operator equals the projection operator with $\mathcal{P}(t) = \bigcup_{k=1}^{f(t)} \mathcal{F}_k(t)$. The family operator is time dependent, hence the user can focus on distinct families at different times.

The influence of a set of chemical groups at initial time t_I on another set of chemical groups at time t is determined by the grouped error:

$$\delta\boldsymbol{c}_\mathrm{g}(t) := \mathbf{F}_t \, \mathbf{L}_{t_I,t} \, \mathbf{F}_{t_I} \delta\boldsymbol{c}(t_I). \tag{2.52}$$

If no chemical group is considered at time t, $t \in [t_I, t_F]$, the family operator $\mathbf{F}_t$ equals the identity matrix. With the aid of the idempotence of the family operator and $\mathbf{L}_{t,t} = \mathbf{I}$, the grouped error at initial time reduces to

$$\delta\boldsymbol{c}_\mathrm{g}(t_I) = \mathbf{F}_{t_I} \, \mathbf{L}_{t_I,t_I} \, \mathbf{F}_{t_I} \, \delta\boldsymbol{c}(t_I) = \mathbf{F}_{t_I} \, \delta\boldsymbol{c}(t_I). \tag{2.53}$$

Substituting the grouped error (2.52) into the measure of error growth gives

$$g_\mathrm{g}\left(\delta\boldsymbol{c}_\mathrm{g}(t_I)\right) := \frac{\|\delta\boldsymbol{c}_\mathrm{g}(t_F)\|_2}{\|\delta\boldsymbol{c}_\mathrm{g}(t_I)\|_2} = \frac{\|\mathbf{F}_{t_F} \, \mathbf{L}_{t_I,t_F} \, \mathbf{F}_{t_I}\delta\boldsymbol{c}(t_I)\|_2}{\|\mathbf{F}_{t_I} \, \delta\boldsymbol{c}(t_I)\|_2}. \tag{2.54}$$

After use of

$$\delta\boldsymbol{c}_\mathrm{g}(t_F) = \mathbf{F}_{t_F} \, \mathbf{L}_{t_I,t_F} \, \delta\boldsymbol{c}_\mathrm{g}(t_I), \tag{2.55}$$

the squared ratio becomes a Rayleigh quotient

$$\begin{aligned} g_\mathrm{g}{}^2(\delta\boldsymbol{c}_\mathrm{g}(t_I)) &= r(\mathbf{B}; \delta\boldsymbol{c}_\mathrm{g}(t_I)), \\ \text{with} \quad \mathbf{B} &:= \mathbf{F}_{t_F} \, \mathbf{L}_{t_I,t_F}, \end{aligned} \tag{2.56}$$

subject to

$$[\delta\boldsymbol{c}_\mathrm{g}(t_I)](j) = \begin{cases} \frac{1}{m_k} \sum_{i \in \mathcal{F}_k(t_I)} [\delta\boldsymbol{c}(t_I)](i) & \forall \, j \in \mathcal{F}_k(t_I) \\ 0 & \text{otherwise.} \end{cases} \tag{2.57}$$

The eigenvector $\boldsymbol{v}_{g1}(t_I)$ of the eigenvalue problem

$$\mathbf{B}^T \mathbf{B} \ \boldsymbol{v}_g(t_I) = \lambda_g \ \boldsymbol{v}_g(t_I) \tag{2.58}$$

assigned to the largest eigenvalue λ_{g1} is the phase space direction which results in a maximum Rayleigh quotient (2.56). In order to ensure, that condition (2.57) holds, the eigenvalue problem (2.58) needs to be restated. Multiplying equation (2.58) with $\mathbf{F}_{t_I}$ and applying

$$\mathbf{F}_{t_I} \ \boldsymbol{v}_g(t_I) = \boldsymbol{v}_g(t_I), \tag{2.59}$$

regroups equation (2.58) into the following equivalent eigenvalue problem:

$$\mathbf{B}_g{}^T \mathbf{B}_g \ \boldsymbol{v}_g(t_I) = \lambda_g \ \boldsymbol{v}_g(t_I),$$
$$\text{where} \quad \mathbf{B}_g := \mathbf{F}_{t_F} \ \mathbf{L}_{t_I, t_F} \ \mathbf{F}_{t_I}. \tag{2.60}$$

The solutions of the new eigenvalue problem hold condition (2.57). Henceforth, the eigenvector $\boldsymbol{v}_{g1}(t_I) \in \mathbb{R}^n$ assigned to largest eigenvalue λ_{g1} of equation (2.60) is denoted as grouped singular vector, since it is the right singular vector of the operator $\mathbf{B}_g$. Accordingly, the square root of the eigenvalue λ_{g1} is the associated grouped singular value σ_{g1}.

The grouped error (2.52) can be combined with the relative error (2.21), leading to the grouped relative error $\delta\boldsymbol{c}_{gr}(t) \in \mathbb{R}^n$. For the latter it is of importance, that the individual errors of the chemical compounds are scaled *before* they are combined to the error of the group. Therefore, the grouped relative error is given by

$$\delta\boldsymbol{c}_{gr}(t) := \mathbf{F}_t \ \mathbf{W}_t^{-1} \ \mathbf{L}_{t_I, t} \ \mathbf{W}_{t_I} \ \mathbf{F}_{t_I} \ \mathbf{W}_{t_I}^{-1} \ \delta\boldsymbol{c}(t_I). \tag{2.61}$$

This rather complicated formula is caused by the fact, that the tangent linear model takes absolute uncertainties as input. At initial time t_I it reduces to

$$\delta\boldsymbol{c}_{gr}(t_I) = \mathbf{F}_{t_I} \ \mathbf{W}_{t_I}^{-1} \ \delta\boldsymbol{c}(t_I). \tag{2.62}$$

Hence, the grouped relative error at final time t_F reads

$$\delta\boldsymbol{c}_{gr}(t_F) = \mathbf{F}_{t_F} \ \mathbf{W}_{t_F}^{-1} \ \mathbf{L}_{t_I, t_F} \ \mathbf{W}_{t_I} \ \delta\boldsymbol{c}_{gr}(t_I). \tag{2.63}$$

Inserting formula (2.63) into the squared grouped relative error growth gives

$$g_{gr}{}^2(\delta\boldsymbol{c}_{gr}(t_I)) := \frac{\|\delta\boldsymbol{c}_{gr}(t_F)\|_2^2}{\|\delta\boldsymbol{c}_{gr}(t_I)\|_2^2} = \frac{\|\mathbf{F}_{t_F} \ \mathbf{W}_{t_F}^{-1} \ \mathbf{L}_{t_I, t_F} \ \mathbf{W}_{t_I} \ \delta\boldsymbol{c}_{gr}(t_I)\|_2^2}{\|\delta\boldsymbol{c}_{gr}(t_I)\|_2^2}. \tag{2.64}$$

It is straightforward to show that equation (2.64) reduces to a Rayleigh quotient

$$
\begin{aligned}
g_{\mathrm{gr}}{}^2 \left(\delta \boldsymbol{c}_{\mathrm{gr}}(t_I) \right) &= r(\mathbf{B}, \delta \boldsymbol{c}_{\mathrm{gr}}(t_I)), \\
\text{with} \quad \mathbf{B} &:= \mathbf{F}_{t_F} \, \mathbf{W}_{t_F}^{-1} \, \mathbf{L}_{t_I, t_F} \, \mathbf{W}_{t_I},
\end{aligned}
\tag{2.65}
$$

subject to

$$
[\delta \boldsymbol{c}_{\mathrm{gr}}(t_I)](j) := \begin{cases} \frac{1}{m_k} \sum_{i \in \mathcal{F}_k(t_I)} \frac{[\delta \boldsymbol{c}(t_I)](i)}{[\boldsymbol{c}(t_I)](i)} & \forall \, j \in \mathcal{F}_k(t_I) \\ 0 & \text{otherwise.} \end{cases}
\tag{2.66}
$$

According to Rayleigh's principle, the phase space direction that maximizes the Rayleigh quotient (2.65) is the solution $\boldsymbol{v}_{\mathrm{gr}_1}(t_I)$ of the symmetric eigenvalue problem

$$
\mathbf{B}^T \mathbf{B} \, \boldsymbol{v}_{\mathrm{gr}}(t_I) = \lambda_{\mathrm{gr}} \, \boldsymbol{v}_{\mathrm{gr}}(t_I)
\tag{2.67}
$$

which is assigned to the largest eigenvalue λ_{gr_1}. This solution does not necessarily hold condition (2.66). In order to grant condition (2.66), the eigenvalue problem (2.67) is rearranged. Making use of the idempotence of the family operator, it can readily be seen that

$$
\mathbf{F}_{t_I} \, \boldsymbol{v}_{\mathrm{gr}}(t_I) = \boldsymbol{v}_{\mathrm{gr}}(t_I).
\tag{2.68}
$$

Multiplying equation (2.67) with $\mathbf{F}_{t_I}$ and substituting relation (2.68) gives

$$
\begin{aligned}
\mathbf{B}_{\mathrm{gr}}{}^T \mathbf{B}_{\mathrm{gr}} \, \boldsymbol{v}_{\mathrm{gr}}(t_I) &= \lambda_{\mathrm{gr}} \, \boldsymbol{v}_{\mathrm{gr}}(t_I) \\
\text{where} \quad \mathbf{B}_{\mathrm{gr}} &:= \mathbf{F}_{t_F} \, \mathbf{W}_{t_F}^{-1} \, \mathbf{L}_{t_I, t_F} \, \mathbf{W}_{t_I} \, \mathbf{F}_{t_I}.
\end{aligned}
\tag{2.69}
$$

The new symmetric eigenvalue problem grants a feasible solution and holds condition (2.66). The vector $\boldsymbol{v}_{\mathrm{gr}_1}(t_I) \in \mathbb{R}^n$ assigned to the largest eigenvalue λ_{gr_1} of equation (2.69) is the right singular vector of the operator $\mathbf{B}_{\mathrm{gr}}$ and is denoted as grouped relative singular vector hereafter. Its associated grouped relative singular value is the square root of the eigenvalue λ_{gr_1}.

2.2 Uncertainties of emission factors

Since emissions are not sufficiently well known but appear to be at least as important as initial values (*Elbern et al.* [2007]), the concern of this section is to study the influence of uncertainties in emission rates on a chemistry-transport model. In contrast to initial concentrations, the total emitted amount of species s in grid cell (i, j, k) varies in time. Consequently, for a

forecast over a finite time interval $[t_I, t_F]$, the optimization problem of determining the most unstable emission uncertainties $\delta\boldsymbol{e}(t)$ has to be considered for each time step $t \in [t_I, t_F]$, resulting in high computational expenditure. By dealing with a time invariant vector of emission factors $\boldsymbol{e}_f$ instead of emission rates themselves, the set of optimization problems can be reduced to one optimization problem per time interval $[t_I, t_F]$. Here, the diurnal profile shape of emission rates is taken as strong constraint, such that only uncertainties in the amplitude are analyzed. This choice is reasonable, as the daily evolution of the emissions is far better known than the total emitted amount in a grid cell (*Elbern et al.* [2007]). The results of the singular vector analysis quantify for which species further emission strength ascertainment is most useful.

Emissions impact the final state $\boldsymbol{c}(t_F)$ according to the partial differential equations, which describe the chemical evolution:

$$\frac{d\boldsymbol{c}}{dt} = f(\boldsymbol{c}(t)) + \boldsymbol{e}(t). \tag{2.70}$$

Inserting the vector of emission factors $\boldsymbol{e}_f$ leads to

$$\frac{d\boldsymbol{c}}{dt} = f(\boldsymbol{c}(t)) + \mathbf{E}(t)\,\boldsymbol{e}_f, \tag{2.71}$$

where $\mathbf{E}(t)$ is a diagonal matrix with the vector of emission $\boldsymbol{e}(t)$ on its diagonal. Now let $\delta\boldsymbol{e}_f$ express uncertainties in the emissions. Then the tangent linear model integration of (2.70) describes the linear part of the evolution of perturbation $\delta\boldsymbol{c}(t_F)$ caused by uncertainties $\delta\boldsymbol{e}_f$ and $\delta\boldsymbol{c}(t_I)$:

$$\delta\boldsymbol{c}(t_F) = \widehat{\mathbf{L}}_{t_I,t_F}\,\delta\boldsymbol{z}(t_I), \tag{2.72}$$

where

$$\widehat{\mathbf{L}}_{t_I,t_F} := \left(\mathbf{L}_{t_I,t_F}, \mathbf{L}^{\boldsymbol{e}}_{t_I,t_F}\right), \qquad \delta\boldsymbol{z}(t_I) := (\delta\boldsymbol{c}(t_I), \delta\boldsymbol{e}_f)^T. \tag{2.73}$$

2.2.1   Error growth of emission factors

For the sake of convenience, emissions are disturbed first and uncertainties of initial concentrations are disregarded (i.e., $\delta\boldsymbol{c}(t_I) = 0$). Since uncertainties of emission factors already denote a relative disturbance, their relative impact is considered. The latter is expressed by the relative error at final time $\delta\tilde{\boldsymbol{c}}_{\mathrm{r}}(t_F)$:

$$\delta\tilde{\boldsymbol{c}}_{\mathrm{r}}(t_F) := \mathbf{W}^{-1}_{t_F}\,\delta\tilde{\boldsymbol{c}}(t_F). \tag{2.74}$$

Here, tilde denotes that the perturbation is merely induced by emissions uncertainties. Since the considered final uncertainty contains only perturbations in concentrations of species, the original weight matrix $\mathbf{W}_t$ (2.20) is applied. From definition (2.74), the relative error growth is obtained:

$$\tilde{g}_r(\delta \boldsymbol{e}_f) := \frac{\|\delta \tilde{\boldsymbol{c}}_r(t_F)\|_2}{\|\delta \boldsymbol{e}_f\|_2}. \tag{2.75}$$

Considering the squared ratio and applying equation (2.72) yields

$$\tilde{g}_r^{\,2}(\delta \boldsymbol{e}_f) = r(\tilde{\mathbf{B}}_r; \delta \boldsymbol{e}_f),$$
$$\text{where} \quad \tilde{\mathbf{B}}_r := \mathbf{W}_{t_F}^{-1} \, \mathbf{L}_{t_I, t_F}^e. \tag{2.76}$$

The phase space direction that maximizes the ratio (2.76) is the eigenvector $\tilde{\boldsymbol{v}}_{r1}$ of the eigenvalue problem

$$\tilde{\mathbf{B}}_r^T \tilde{\mathbf{B}}_r \, \tilde{\boldsymbol{v}}_r = \tilde{\lambda}_r \, \tilde{\boldsymbol{v}}_r \tag{2.77}$$

assigned to largest eigenvalue $\tilde{\lambda}_{r1}$ (again the tilde displays the fact, that only uncertainties in the emissions are considered). As the solution equals the right singular vector of the operator $\tilde{\mathbf{B}}_r$, it is denoted as relative singular vector with respect to emission uncertainties henceforth. Its associated singular value $\tilde{\sigma}_{r1}$ is the square root of $\tilde{\lambda}_{r1}$.

A special set of perturbation norms can be provided with the help of the projection operator $\mathbf{P}_t$ (2.27) and the family operator $\mathbf{F}_t$ (2.49). For the error growth of emission factor uncertainties, the projected relative singular vector as well as the grouped relative singular vector can be calculated following chapter 2.1.2 and 2.1.3.

2.2.2   Combined error growth

Consider now both uncertainties in initial values and emissions affecting the solution $\delta \hat{\boldsymbol{c}}(t_F)$ (where hat indicates that uncertainties of initial conditions and emissions were taken into account). Subject to the relative character of the emission factors $\boldsymbol{e}_f$, the singular vector analysis directly deals with the relative errors $\delta \boldsymbol{z}_r$ and $\delta \hat{\boldsymbol{c}}_r$:

$$\delta \hat{\boldsymbol{c}}_r(t_F) := \mathbf{W}_{t_F}^{-1} \, \delta \hat{\boldsymbol{c}}(t_F), \tag{2.78}$$

$$\delta \boldsymbol{z}_r(t_I) := \hat{\mathbf{W}}_{t_I}^{-1} \, \delta \boldsymbol{z}(t_I) = \begin{pmatrix} \mathbf{W}_{t_I}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \delta \boldsymbol{c}(t_I) \\ \delta \boldsymbol{e}_f \end{pmatrix}, \tag{2.79}$$

where $\mathbf{I}$ denotes the identity matrix and $\mathbf{W}_t$ the weight matrix (2.20). Then the measure of error growth is characterized by the ratio:

$$\hat{g}_{\mathrm{r}}\left(\delta \boldsymbol{z}_{\mathrm{r}}(t_I)\right) := \frac{\|\delta \hat{\boldsymbol{c}}_{\mathrm{r}}(t_F)\|_2}{\|\delta \boldsymbol{z}_{\mathrm{r}}(t_I)\|_2}. \tag{2.80}$$

Applying the squared ratio and substitution of equations (2.72) and (2.79) into equation (2.80) yields a Rayleigh quotient:

$$\hat{g}_{\mathrm{r}}^2\left(\delta \boldsymbol{z}_{\mathrm{r}}(t_I)\right) = r(\hat{\mathbf{B}}_{\mathrm{r}}; \delta \boldsymbol{z}_{\mathrm{r}}(t_I))$$
$$\text{with} \quad \hat{\mathbf{B}}_{\mathrm{r}} := \mathbf{W}_{t_F}^{-1}\, \hat{\mathbf{L}}_{t_I,t_F}\, \hat{\mathbf{W}}_{t_I}. \tag{2.81}$$

Consequently, the direction of largest relative error growth $\hat{\boldsymbol{z}}_{\mathrm{r}}(t_I)$ can be obtained directly via

$$\hat{\mathbf{B}}_{\mathrm{r}}^T\, \hat{\mathbf{B}}_{\mathrm{r}}\, \hat{\boldsymbol{v}}_{\mathrm{r}}(t_I) = \hat{\lambda}_{\mathrm{r}}\, \hat{\boldsymbol{v}}_{\mathrm{r}}(t_I)$$
$$\text{where} \quad \hat{\boldsymbol{v}}_{\mathrm{r}}(t_I) := \hat{\mathbf{W}}_{t_I}^{-1}\, \hat{\boldsymbol{v}}(t_I). \tag{2.82}$$

The phase space direction for which equation (2.80) reaches its maximum value is the singular vector $\hat{\boldsymbol{v}}_{\mathrm{r}1}(t_I)$ assigned to the largest eigenvalue $\hat{\lambda}_{\mathrm{r}1}$ of the operator $\hat{\mathbf{B}}_{\mathrm{r}}$. It is referred to as relative singular vector with respect to initial conditions and emission uncertainties. Its associated singular value $\hat{\sigma}_{\mathrm{r}1}$ is the square root of the eigenvalue $\hat{\lambda}_{\mathrm{r}1}$.

For the combined error growth of emission factors and initial condition uncertainties the projection operator $\mathbf{P}_t$ (2.27) and the family operator $\mathbf{F}_t$ (2.49) at initial time are extended to

$$\hat{\mathbf{P}}_{t_I} := \begin{pmatrix} \mathbf{P}_{t_I}^{\boldsymbol{c}} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{t_I}^{\boldsymbol{e}} \end{pmatrix} \quad \text{and} \quad \hat{\mathbf{F}}_{t_I} := \begin{pmatrix} \mathbf{F}_{t_I}^{\boldsymbol{c}} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{t_I}^{\boldsymbol{e}} \end{pmatrix}, \tag{2.83}$$

where the superscript $\boldsymbol{e}$ denotes the emission part and the superscript $\boldsymbol{c}$ the chemical part of the family and projection operator respectively. At final time the original operators are applied, since the considered final uncertainty contains only perturbations in concentrations of species. With definition (2.83) projected relative singular vectors as well as grouped relative singular vector can be calculated for the combined error growth.

CHAPTER 3

Model design

The evolution of chemical species in time and space is driven by transport, diffusion, deposition, emissions and chemical transformations. Since the specific goal of this work is the ranking of chemical constituents sorted by size of the associated singular vector entry, focus is placed on chemical reaction mechanisms. Therefore, chemical transformations were studied first utilizing a chemistry box model. This model simulates the chemical evolution of an air parcel without interferences with its neighborhood. The singular vector entries of the chemistry box model can be interpreted as *chemical sensitivities*. Details about the box model specifications for chemical sensitivities are given in sections 3.1, 3.2 and 3.3.

In order to observe complex interactions between emission sources, chemical transformation, transport, diffusion and deposition processes, a complete chemical transport model has been applied. Since singular vectors of chemical transport models identify not only sensitive chemical species but also chemically sensitive locations, they are denoted as *spatial chemical sensitivities*. Their implementation is explained in section 3.4.

3.1 Chemistry mechanisms

In this work, focus was placed on tropospheric gas-phase chemistry. Two different chemistry mechanism were adapted, which both satisfy the demand

to meet the computational limits on the one hand and to deliver a good representation of the tropospheric chemistry on the other hand. They are described in the following subsections.

3.1.1   RADM2

The first implemented chemistry mechanism is the second generation Regional Acid Deposition Model (RADM2, *Stockwell et al.* [1990]). It is an offspring of the RADM1 (*Chang et al.* [1987]), but offers a much more complete parametrization of many important species. To simplify the complex organic chemistry of the troposphere, the mechanism makes use of a reactivity lumped molecular approach. Therefore, hundreds of volatile organic compounds (VOC) are grouped together into 15 classes of reactive organic species based on similarity in oxidation reactivity and emission magnitudes.

Altogether RADM2 includes 63 species, 42 organic, and 21 inorganic compounds. The organic species are subclassified in 26 stable species and 16 peroxy radicals. Inorganic compounds include 14 stable species, 4 reactive intermediates, and 3 abundant stable species. The complete RADM2 chemical species list can be found in the appendix (Table A.1). The chemical compounds are coupled by 147 reactions that are divided into 21 photolysis reactions, 124 thermal reactions, 5 Troe reactions, 2 reactions calculated from equilibria, and 5 reactions with special rate expressions.

The RADM2 has been tested with more than 550 environmental chamber experiments performed in 4 different chambers by *Carter and Lurmann* [1990]. The results demonstrate that the RADM2 provides a good representation of the tropospheric chemistry.

3.1.2   RACM-MIM

The Regional Atmospheric Chemistry Mechanism (RACM, *Stockwell et al.* [1997]) is a revision of the RADM2 that provides a more detailed description of the chemistry of biogenic ozone precursors. Its reaction scheme considers 56 organic species and 21 inorganic species with 237 reactions including 33 photolysis reactions. Organic compounds are grouped into 32 stable and 24 intermediate compounds, while inorganic species are divided into 17 stable and 4 intermediates species.

Geiger et al. [2003] introduced an updated version of the RACM based on the Mainz Isoprene Mechanism (MIM, *U. Pöschl and Crutzen* [2000]), called RACM-MIM henceforth. The MIM itself is an enhanced condensed isoprene

degradation mechanism. It considers a number of products and reactions, which are not included in the original isoprene chemistry of RACM. In the course of the RACM update, the entire MIM was implemented into the RACM first. Performed sensitivity tests with the fully upgraded RACM identified redundant reactions, yielding to a reduced implementation of the MIM scheme. In comparison with the RACM, the RACM-MIM finally includes 7 further reactants as well as 7 new chemical reactions. Altogether, the RACM-MIM includes 221 chemical reactions, 23 photolysis reactions and 84 species. The complete RACM-MIM chemical species list can be found in the appendix (Table B.1).

3.2   The numerical solver

In order to calculate chemical sensitivities, the evolution of chemical constituents in time needs to be simulated. The Kinetic PreProcessor (KPP, *Sandu and Sander* [2006]) offers a set of stiff numerical integrators to efficiently simulate the chemical kinetic system. Furthermore, KPP includes the tangent linear and adjoint model with respect to initial conditions for some of the provided integrators.

As input, KPP needs a chemistry mechanism and a numerical integrator. Users may select from a given set of chemistry mechanisms or supply his or her own chemistry mechanism in terms of species, chemical reactions, reaction rates and initial values. A library of state of the art numerical integrators is provided. After setting up a chemistry mechanism and a numerical integrator, the KPP creates Fortran77, Fortran 90, Matlab or C routines for efficient numerical integration of the selected chemical kinetic system. In the following subsections the configuration applied in this work is described in detail. Since Fortran is the programming language of choice (due to conformance to the 3-dimensional chemistry-transport model, which is described in section 3.4.1), all newly written routines have been generated in Fortran77 or Fortran90.

3.2.1   Specification of the chemistry model

In the KPP library the chemistry mechanisms RADM2 and RACM are not implemented as standard. Therefore, their chemical compounds and reactions were specified for utilization within KPP. The reactions of both mechanisms can be classified into photolysis reactions, thermal reactions, falloff reactions, equilibrium reactions, and reactions with special rate expressions.

KPP offers reaction rate updates for thermal reactions in terms of Arrhenius and modified Arrhenius equations respectively, as well as for falloff reactions. New routines have been implemented to update equilibrium equations, reactions with special rate expressions, and photolysis reactions. In addition, molecular weights and Henry's law constants of chemical species are included, such that the user can control the compliance with mass conservation. Moreover, new routines have been created for emitted species to account for mass changes due to emissions.

3.2.2 The numerical integrator

KPP offers several Rosenbrock methods as well as Runge-Kutta methods and backward differentiation formulas. By carefully exploiting the sparsity structure of the Jacobian and the Hessian, KPP obtains efficiency for forward, tangent-linear and adjoint code (*Sandu and Sander* [2006]). The KPP version employed in this work (version 2.1) contains no adjoint model implementation for Runge-Kutta integrators and backward differentiation formulas, but only for Rosenbrock methods. Rosenbrock solvers have optimal stability properties and conserve the linear invariants of the system (*Hairer and Wanner* [1991]). Thus, they are well-suited for atmospheric chemistry applications. In this study a Rosenbrock method is applied.

The Rosenbrock tangent-linear models are generated utilizing the direct decoupled method, which has been shown to be cost-effective for low and medium accuracies (*Sandu et al.* [2003]). For Rosenbrock adjoint models KPP contains continuous and discrete adjoint models. Continuous adjoint models are evolved by forming the adjoint equation of the forward ODE first and discretizing the adjoint model equation afterwards (the name Rosenbrock adjoint model indicates that the continuous adjoint is integrated backward in time with the same numerical integrator as used in the forward integration). In contrast, discrete adjoint models are constructed by discretizing the problem in the first step and implementing the associated adjoint model subsequently. The continuous model has the advantage to be computationally less expensive, but the results are in general less accurate than the results of the discrete adjoint model. Furthermore, the discrete adjoint offers a much better consistency with the forward and tangent-linear model and, thus, it is our adjoint model of choice. In order to find the most appropriate Rosenbrock method for calculating the chemical sensitivities, the consistency of the adjoint model with the forward model has been tested for all Rosenbrock methods following *Chao and Chang* [1992]. All methods demonstrate a good performance. Since the order of accuracy as well as the computational work

per time step increases with the number of stages of Rosenbrock methods, a trade-off between precision and speed of computation has to be made. Due to time-consuming applications the second order Rosenbrock method (Ros-2) has been selected.

3.3 Solving the eigenvalue problems

Two methods have been implemented for solving the eigenvalue problems, namely the power method and the implicitly restarted Arnoldi method. Even though the extended power method computes only the dominant eigenpair with sufficient accuracy, it can be used to (roughly) confirm the eigenpairs computed with the Arnoldi process.

3.3.1 Extended power method

The power method (*Mises and Pollaczek-Geiringer* [1929]) is an iterative technique for computing the dominant eigenpair $(\lambda_1, \boldsymbol{v}_1)$ of a matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$. Here, only the case where $\mathbf{A}$ is symmetric and $\mathbf{A} \in \mathbb{R}^{n \times n}$ is considered. If the dominant eigenvalue λ_1 is strictly greater in absolute value than all other eigenvalues:

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|, \tag{3.1}$$

and if the start vector $\boldsymbol{v}^{(0)} \in \mathbb{R}^n$ is not orthogonal to the eigenspace of λ_1, then the sequences of vectors $\{\boldsymbol{v}^{(k)}\}_{k=1,\ldots s}$ and scalars $\{\lambda^{(k)}\}_{k=1,\ldots s}$ generated recursively by

$$\boldsymbol{v}^{(k)} = \mathbf{A}\,\boldsymbol{v}^{(k-1)} / \|\mathbf{A}\,\boldsymbol{v}^{(k-1)}\| \tag{3.2}$$

$$\lambda^{(k)} = \boldsymbol{v}^{(k)T}\mathbf{A}\,\boldsymbol{v}^{(k)} \tag{3.3}$$

will converge to the dominant eigenvector $\boldsymbol{v}_1$ and eigenvalue λ_1, respectively. The speed of convergence of the power method is proportional to $|\lambda_2/\lambda_1|$, which implies, that the rate of convergence is linear. The more the absolute values of the eigenvalues λ_1 and λ_2 differ, the faster the power method converges. For more details and a proof of the statements made in this section see *Deuflhard and Hohmann* [2003].

In order to find not only the leading eigenvalue and its associated eigenvector, but also the k largest eigenvalues and their eigenvectors an extension of the power method, called deflation method (*Acton* [1990]), can be applied. This approach is based on the idea to calculate the largest eigenpair $(\lambda_1, \boldsymbol{v}_1)$

of $\mathbf{A}$ first and simply 'remove' it from matrix $\mathbf{A}$ to gain a matrix $\mathbf{A}^{(2)}$, whose largest eigenpair is the second largest eigenpair $(\lambda_2, \boldsymbol{v}_2)$ of $\mathbf{A}$. Eigenpair $(\lambda_2, \boldsymbol{v}_2)$ can then be calculated by applying the power method on $\mathbf{A}^{(2)}$. For the third largest eigenpair of $\mathbf{A}$ a matrix $\mathbf{A}^{(3)}$ is considered, on which the eigenpairs $(\lambda_1, \boldsymbol{v}_1)$ and $(\lambda_2, \boldsymbol{v}_2)$ are 'removed' of $\mathbf{A}$ and so forth. In detail the sequence of matrices $\mathbf{A}^{(k)}$ is build using the following specification:

$$\mathbf{A}^{(k+1)} := \mathbf{A}^{(k)} - \lambda_k \, \boldsymbol{v}_k \, \boldsymbol{v}_k^T \quad \forall \quad k = 1, ..., n-2 \qquad (3.4)$$
$$\text{with} \quad \mathbf{A}^{(1)} := \mathbf{A} \quad \text{and} \quad \boldsymbol{v}_k^T \, \boldsymbol{v}_k = 1.$$

Obviously each matrix $\mathbf{A}^{(k+1)}$ has $n-k$ eigenpairs $(\lambda_i, \boldsymbol{v}_i), i = k+1, .., n$ identical to those of $\mathbf{A}$. For the remaining k eigenvalues of $\mathbf{A}^{(k+1)}$ the relation $\lambda_i = 0, i = 1, .., k$ holds.

Even though the deflation method is easy to realize, it is becoming rapidly less accurate with increasing number of eigenpairs to compute. This disadvantage is induced by the fact that each application of the power method to gain the largest eigenpairs of the considered matrix $\mathbf{A}^{(k)}$ results in an approximation of $(\lambda_k, \boldsymbol{v}_k)$ only. Hence the subsequent matrix $\mathbf{A}^{(k+1)}$ is imprecise, leading to an accumulation of errors for the approximation of $(\lambda_{k+1}, \boldsymbol{v}_{k+1})$. Since this feature degrades with each approximated eigenpair, the subsequently computed eigenpairs are only useful to roughly ensure the correctness of the eigenpairs computed with the Arnoldi package (see next section). In addition the deflation method is very cost intensive, since it produces a whole power-method-iteration-cycle for each designated eigenpair.

3.3.2   ARPACK

An appropriate method to find the k largest eigenvalues and their associated eigenvectors is implemented in the public software ARnoldi PACKage (ARPACK, *Lehoucq et al.* [1998], *Sorensen* [1996]). This software supplies a package of Fortran77 subroutines for solving large-scale eigenvalue problems. For that purpose it requires a number of subroutines from the Linear Algebra PACKage and the Basic Linear Algebra Subprograms (LAPACK and BLAS, both available at `http://www.netlib.org/`). ARPACK relies on the Lanczos and the Arnoldi process, dependent on the properties of the matrix $\mathbf{A}$. If $\mathbf{A}$ is symmetric, an algorithmic variant of the Implicitly Restarted Lanczos Method (IRLM) is used, otherwise a variant of the Implicitly Restarted Arnoldi Method (IRAM) is employed. These methods are presented in detail in *Sorensen* [1996]. Both have the ability to calculate the k largest eigenvalues and their associated eigenvectors in one iteration cycle. ARPACK has

the important advantage that it only needs the matrix-vector product instead of an explicit representation of the matrix **A**. Since the eigenvalue problems in this work include operators (see chapter 2), ARPACK is perfectly tailored to our needs.

3.4   Design of spatial chemical sensitivities

Spatial chemical sensitivities rely on a chemistry transport model, which propagates chemical species in time and space. The chemistry transport model employed in this work is described in section 3.4.1. For application of the singular vector analysis, the utilized model system had to be extended. New features and their implementation are addressed in section 3.4.2.

3.4.1   Chemistry transport model

Chemistry transport models simulate atmospheric chemistry by calculation of chemical transformation, meteorological transport, deposition and sedimentation processes and emission strengths. In this work, the EURopean Air pollution and Dispersion - Inverse Model (EURAD-IM, e.g., *Elbern et al.* [2007], *Elbern and Schmidt* [1999], and *Elbern* [1997]) is our model of choice. The EURAD-IM is an advanced Eulerian model. It operates from European down to local scale by applying a nesting technique with a horizontal solution down to 1 km. The horizontal grid design is based on Lambert conformal conic projections and employs the Arakawa C grid stencil (*Arakawa and Lamb* [1977]). The vertical grid structure of the EURAD-IM is defined by a terrain following σ-coordinate system. Due to the general focus on tropospheric applications in this work, the vertical limit is 100 hPa. 23 vertical model layers are implemented between surface and 100 hPa.

The EURAD-IM simulates the chemical development in time and space based on the following system of partial differential equations:

$$\frac{\partial c_i}{\partial t} = \underbrace{-\nabla(v c_i)}_{Advection} + \underbrace{\nabla(\rho K \nabla \frac{c_i}{\rho})}_{Diffusion} + \underbrace{A_i}_{Chemistry} + \underbrace{E_i}_{Emission} - \underbrace{\frac{\partial}{\partial z}(v_i^d c_i)}_{Deposition}, \qquad (3.5)$$

where $c_i, i = 1, ..., n$ denotes the mean mass mixing ratios of the chemical species, v the mean wind velocity, K the eddy diffusivity tensor, ρ the air density, A_i the chemical generation term for species c_i, E_i its emission rates and v_i^d the deposition velocity. The numerical solution of equation (3.5) employs

an operator splitting technique, which splits the partial differential equations into subproblems and treats them successively. In case of the EURAD-IM model, the dynamic procedures are splitted symmetrically, centered around the chemistry solver module $\mathbf{C}$:

$$c_i{}^{t+\Delta t} = \mathbf{T}_h^{\Delta t/2}\, \mathbf{T}_v^{\Delta t/2}\, \mathbf{D}_v^{\Delta t/2}\, \mathbf{C}^{\Delta t}\, \mathbf{D}_v^{\Delta t/2}\, \mathbf{T}_v^{\Delta t/2}\, \mathbf{T}_h^{\Delta t/2}\, c_i{}^t. \qquad (3.6)$$

Here, $\mathbf{T}_{h,v}$ and $\mathbf{D}_v$ denote the numerical solution for transport and diffusion with respect to horizontal (h) and vertical (v) direction. The subscript indicates whether the process is applied for the entire time-step (Δt) or for half of the time-step ($\Delta t/2$). The emission term is included in the chemistry solver module $\mathbf{C}$. In cases, where multiple solution-schemes are provided by EURAD-IM, the user can chose a suitable scheme. In this work, the upstream algorithm devised by *Bott* [1989] is chosen as advection scheme. Implemented are fourth order polynomials for the horizontal advection and second order polynomials for the vertical advection. For vertical diffusion the semi-implicit Thomas algorithm is applied. The chemical development is implemented with the software KPP, using an 2-stage Rosenbrock solver (see section 3.2).

3.4.2   Upgrading the EURAD-IM

Besides the possibility for operational forecast, the EURAD-IM provides variational data assimilation with initial value and emission rate optimization. In this study, the EURAD-IM was upgraded to enable the evaluation of spatial chemical sensitivities.

In order to realize the spatial chemical sensitivities, the singular vector analysis is applied (see chapter 2). The associated eigenvalue problems require the tangent linear and adjoint model of equation (3.5). Due to variational data assimilation, the EURAD-IM comprises adjoint modules for all considered processes. Furthermore, for the chemical evolution the tangent linear model with respect to initial conditions is provided by KPP (see section 3.2). For the remaining routines, the tangent linear model has been coded by hand. Furthermore, another version of the tangent linear routines was obtained by the online automatic differentiation engine TAPE-NADE (`http://tapenade.inria.fr:8080/tapenade/`). Both model variations were checked against each other and merged to a improved tangent linear model version.

Newly coded tangent linear routines have been checked for consistency with corresponding forward and adjoint modules. Tests were accomplished for

each module separately first, and then for the entire model. For consistency with the forward model, the gradient check ratio (*Navon et al.* [1992]) is applied, defined as

$$d = \frac{\texttt{FWD}(x + \alpha \delta x) - \texttt{FWD}(x)}{\texttt{TLM}(\alpha \delta x)}. \tag{3.7}$$

The abbreviations $\texttt{FWD}$ and $\texttt{TLM}$ denote the forward and tangent linear model, α is a scalar parameter. While α approaches zero, the ratio (3.7) should approach one. Note that round-off errors dominate in cases, where the magnitude of α is sufficiently small. Here, the limits of numerical precision are reached and results get worse. Within these limits, the new tangent linear routines demonstrate the required characteristics for equation (3.7) for considered test cases. Besides consistency with the forward model, the gradient ratio check indicates the accuracy of the tangent linear assumption. Application of the tangent linear model is only justified, if the considered perturbation is small enough to ensure $d \approx 1$. For each chemical scenario considered, the required accuracy has been achieved.

For consistency of tangent linear and adjoint model, *Navon et al.* [1992] suggest to inspect if the following equation holds:

$$(\texttt{TLM}(\delta x))^T (\texttt{TLM}(\delta x)) = \delta x^T \texttt{ADJ}(\texttt{TLM}(\delta x)), \tag{3.8}$$

where $\texttt{ADJ}$ denotes the adjoint model. The newly coded tangent linear routines satisfy equation (3.8) within the limits of numerical precision.

Employing the new tangent linear routines, one iteration of the eigenvalue problems can finally be calculated. For solving the eigenvalue problems, the extended power method (see section 3.3.1) and the software package PARPACK (Parallel ARPACK, for ARPACK see section 3.3.2) have been implemented and applied.

Chemical sensitivities for tropospheric chemistry scenarios

Poppe et al. [2001] introduced a set of six scenarios for modeling tropospheric chemistry. These scenarios are designed to cover conditions that are typical for the remote continental planetary boundary layer (briefly this scenario is called LAND), the ocean (scenario MARINE), the free troposphere (scenario FREE), the moderately polluted planetary boundary layer (scenario PLUME), the polluted planetary boundary layer (scenario URBAN) and an urban plume with biogenic impact (scenario URBAN/BIO). Since cases LAND, FREE, and MARINE represent rather clean air, these scenarios feature no emissions. In contrast, scenarios PLUME, URBAN, and URBAN/BIO include a varying burden of emissions. While cases PLUME and URBAN consider only anthropogenic emissions with different source strengths, scenario URBAN/BIO considers both anthropogenic and biogenic emissions. In detail, the first 60 hours of scenario URBAN/BIO are identical to those of scenario URBAN. For scenario URBAN/BIO the anthropogenic VOC emissions are switched off after 60 hours and biogenic emission (of isoprene) is switched on.

In order to verify, that the scenarios are unambiguously described, *Poppe et al.* [2001] simulated all cases with five different chemistry box models. The presented calculations demonstrate excellent agreement between the participating numerical solvers. Hence the definition of the scenarios is clear and complete and can be utilized for testing numerical solvers. Further, the scenarios can serve as well documented base cases for sensitivity studies. In this

work, both possibilities are employed. First, in section 4.1 the scenarios are applied to verify that the chemical models and solvers used in this work are suitable for the numerical solution of the balance equation. Second, in section 4.2 they are taken as base cases for sensitivity studies with main focus on the influence (in terms of initial concentrations and emission factors) of volatile organic compounds (VOC) and nitrogen oxides (NO_x) on the final concentration of ozone (O_3).

4.1   Description and forward simulation

The scenarios are defined via meteorological parameters (described in Table 4.1), chemical initial conditions (summarized in Table 4.2), photolysis frequencies (see appendix A.1.2) and emissions (see appendix A.1.3). The utilized chemistry mechanism is RADM2 (section 3.1.1), which considers tropospheric gas-phase chemistry only. All simulations start at July 1, local noon, and end five days (120 hours) later at July 6, local noon.

Table 4.1: *Meteorological parameters for scenarios defined by Poppe et al. [2001].*

	LAND, PLUME & MARINE	FREE	URBAN & URBAN/BIO
Altitude (km)	0	8	0
Temperature (K)	288.15	236.21	298.00
Pressure (hPa)	1013.25	356.50	1013.25
Air (1/cm^3)	$2.55 \cdot 10^{19}$	$1.09 \cdot 10^{19}$	$2.46 \cdot 10^{19}$

Table 4.2: *Initial mixing ratios for the gas-phase constituents (ppb) for scenarios defined by Poppe et al. [2001]. NMHC denotes non-methane hydrocarbons, other species are described in the RADM2 species list (Table A.1).*

	LAND	MAR.	FREE	PLUME	URBAN	URB./BIO
O_3	30	30	100	50	30	30
NO	0.1	0.1	0.05	0.2	0.1	0.1
NO_2	0.1	0.1	0.05	0.5	0.1	0.1
HNO_3	0.1	1.5	0.1	0.1	1.5	0.1

Continued next page

Table 4.2 – Continued

CO	100	100	100	200	100	100
CH$_4$	1700	1700	1700	1700	1700	1700
Isopr.	0	0	0	0	0	0
H$_2$	500	500	500	500	500	500
H$_2$O$_2$	2	2	2	2	2	2
HCHO	1	1	0	0	1	1
NMHC	0	0	0	0	0	0
SO$_2$	0	0	0	0	0	0

In order to test the numerical solver employed in this work, all scenarios were simulated. For the chemical constituents NO$_x$, O$_3$ and OH, the results of the simulations of all scenarios are depicted in Fig. 4.1 and Fig. 4.2. Since the depicted results show nearly perfect agreement with those of *Poppe et al.* [2001], the utilized configuration seems to be suitable to model the chemical evolution of the given scenarios.

4.2 VOC versus NO$_x$ limitation of the ozone formation

Ozone (O$_3$) formation is known to be dependent on two major classes of directly emitted precursors: nitrogen oxides (NO$_x$) and volatile organic compounds (VOC). The relation between O$_3$, NO$_x$ and VOC is driven by complex nonlinear photochemistry. An important indicator for the VOC versus NO$_x$ limitation of the O$_3$ formation is the VOC-to-NO$_x$ ratio. It can be expressed in terms of initial concentrations or emission rates. At high VOC-to-NO$_x$ ratios, O$_3$ increases with increasing NO$_x$ and changes little in response to increasing VOC. Therefore, regimes with high VOC-to-NO$_x$ ratios are considered to be NO$_x$ limited or NO$_x$ sensitive. In contrary, regimes with low VOC-to-NO$_x$ ratios are VOC limited. Here, O$_3$ decreases with increasing NO$_x$ and increases with increasing VOC. A detailed description of these mechanism can be found in *Seinfeld and Pandis* [1998]. In the following study, the issue of the VOC versus NO$_x$ limitation of the O$_3$ formation is applied to study singular vector analyses for atmospheric chemical modeling.

In case of the RADM2 mechanism, VOC consist of the model species ETH, HC3, HC5, HC8, OL2, OLT, OLI, ISO, TOL, CSL, XYL, HCHO, ALD,
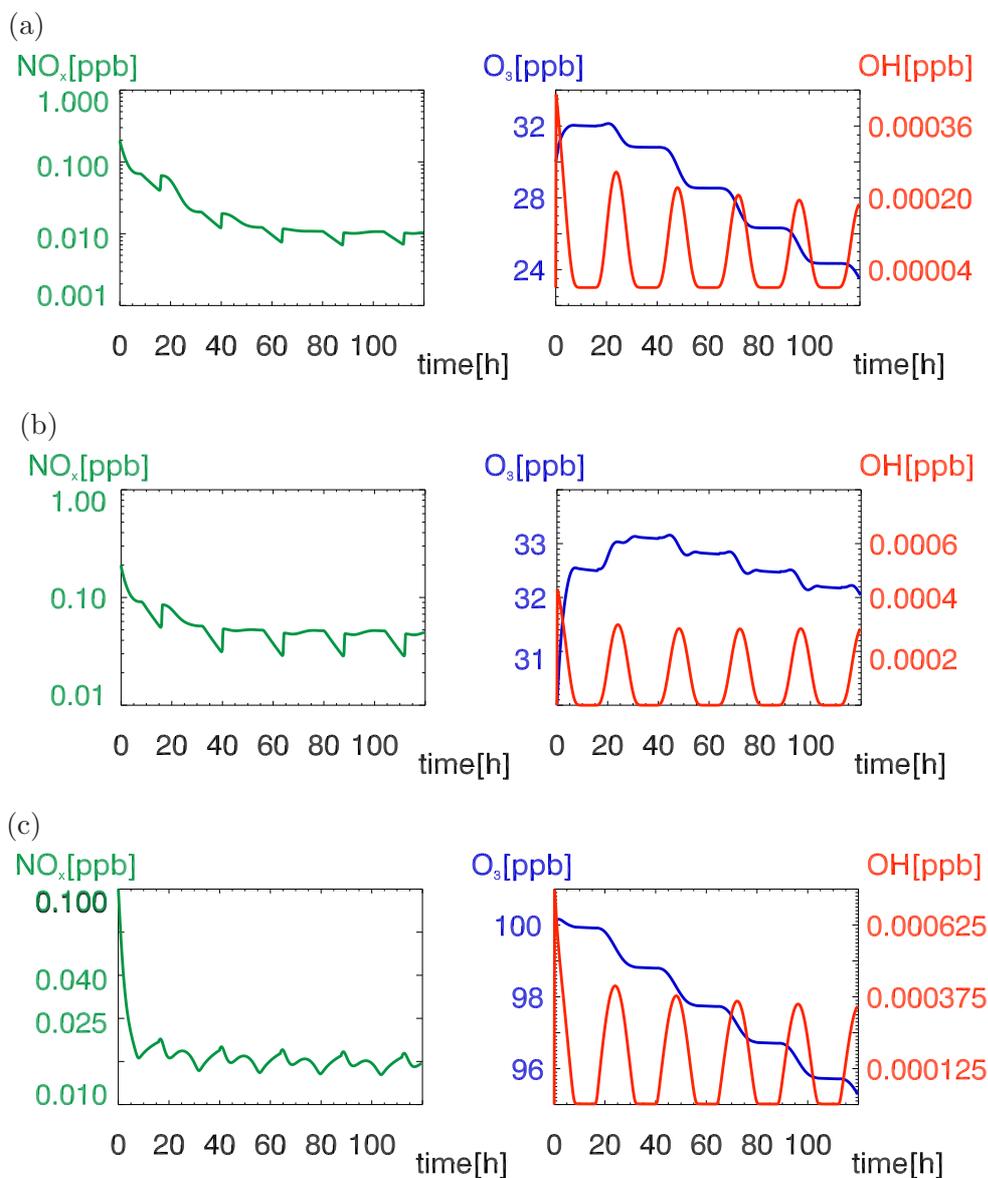
Figure 4.1: *Mixing ratios for NO_x, O_3 and OH as function of integration time for scenarios (a) LAND, (b) MARINE and (c) FREE.*

KET, GLY, MGLY, and DCB, whereas NO_x comprise the model species NO and NO_2. A short description of these acronyms can be found in the appendix (Table A.1). The sensitivity studies presented in the following consider these species with the objective to investigate the influence of initial values and emissions of volatile organic compounds (VOC) and nitrogen oxides (NO_x) on the final concentration of ozone (O_3). In order to examine the different er-
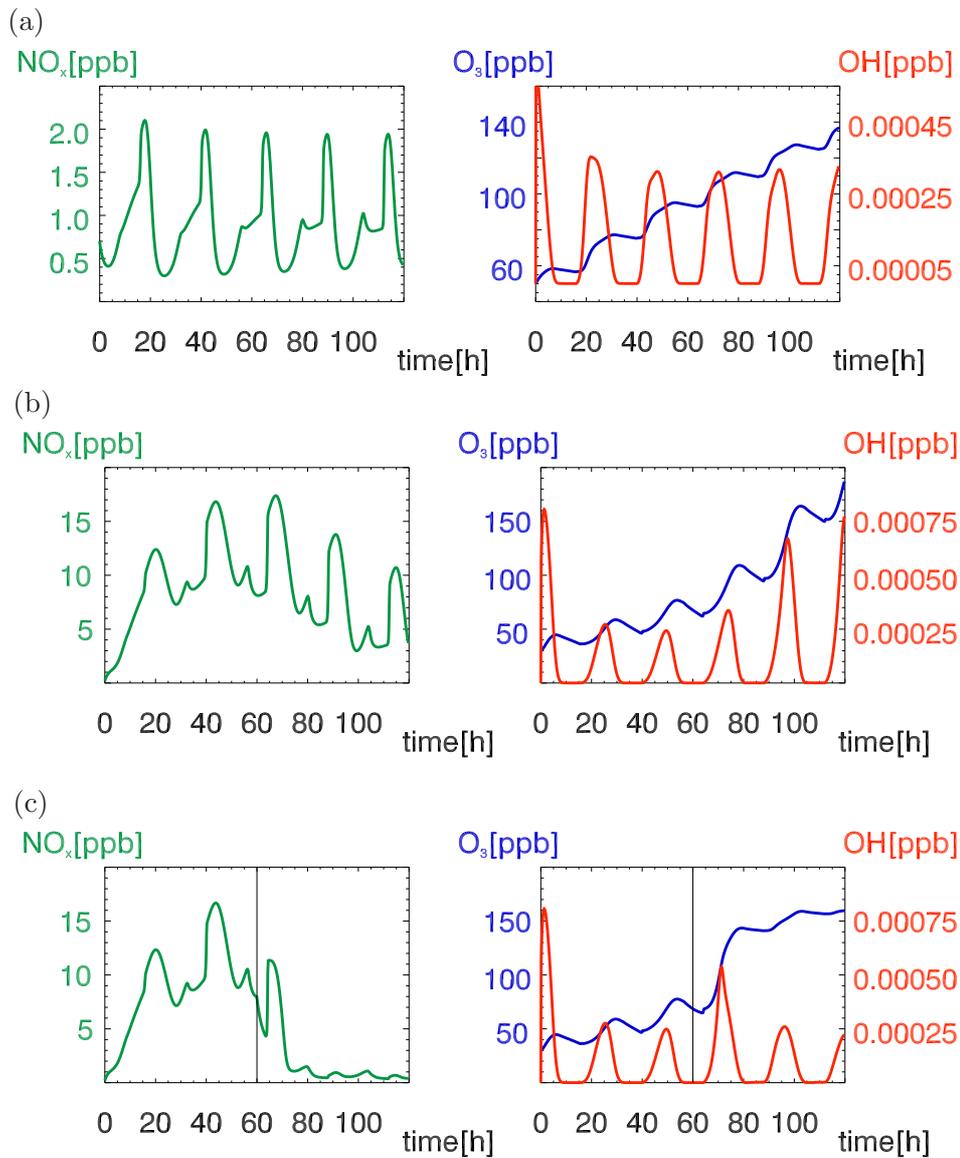
Figure 4.2: *Mixing ratios for NO$_x$, O$_3$ and OH as function of integration time for scenarios (a) PLUME, (b) URBAN and (c) URBAN/BIO. For scenario UR-BAN/BIO, the black continuous line marks the switch from anthropogenic to biogenic emissions.*

ror growths introduced in chapter 2, the singular vector analysis was applied with respect to the coupled error growth, the projected error growth, the coupled relative error growth as well as the projected relative error growth. Since sensitivities are dependent on integration time and aging, the time-dependence of the VOC versus NO$_x$ dominance is of special interest.

For a comprehensive investigation of the effects of initial concentrations and aging, a temporal singular vector diagram (TSVD) has been implemented. Here, a complete set of singular vectors with varying starting times and simulation intervals is calculated to cover the original interval $[t_I, t_F]$. The calculations are carried out in $n = \frac{t_F - t_I}{\Delta t}$ time rows, where Δt determines the chosen minimal simulation interval length. Each time row TR(i), i=0,...,n-1 consists of n individual singular vector simulations, all starting at time $t_i = t_I + i\Delta t$, but differing in terms of simulation length, which equals $(m + 1) \cdot \Delta t$ for calculation m, m=0,..,n-1. Altogether, each TSVD comprises n^2 singular vector analyses. A schematic overview of the TSVD is depicted in Fig. 4.3.

$$
\begin{array}{cccccc}
& & & & & [t_n, t_{2n} \quad] \\
& & & & [t_{n-1}, t_{2n-1}] & [t_n, t_{2n-1}] \\
& & \iddots & & & \\
& & \iddots & & \vdots & \vdots \\
& [t_1, t_{n+1}] & \cdots & \cdots & [t_{n-1}, t_{n+1}\ ] & [t_n, t_{n+1}\ ] \\
[t_0, t_n \quad] & [t_1, t_n \quad] & \cdots & \cdots & [t_{n-1}, t_n \quad] & \\
\vdots & \vdots & \iddots & & & \\
[t_0, t_2 \quad] & [t_1, t_2 \quad] & \iddots & & & \\
[t_0, t_1 \quad] & & & & & \\
\end{array}
$$

$$
\begin{array}{cccc}
\uparrow & \uparrow & \uparrow & \uparrow \\
\text{time row} & \text{time row} & \text{time row} & \text{time row} \\
\text{TR}(0) & \text{TR}(1) & \text{TR}(n\text{-}1) & \text{TR}(n)
\end{array}
$$

Figure 4.3: *Schematic presentation of the temporal singular vector diagram. Here, each interval $[t_i, t_f]$ represents one singular vector calculation starting at $t_i = t_I + i\Delta t, i = 1,.., n$ and ending at $t_f = t_I + f\Delta t, f = 1,.., 2n, f > i$.*

In the particular case of the following sensitivity study based on the scenarios by *Poppe et al.* [2001], each singular vector simulation had a spin up run of 24h to ensure chemical balance at initial time and to guarantee meaningful relative errors. Hence, the starting point t_I of the TSVD is July 2, local noon. Further, t_F is determined as July 6, 12h and Δt as one hour, leading to $n = 96$ and therefore 9216 singular vector calculations per TSVD.

For sake of clarity, the starting point of the TSVD is called starting time t_I henceforth, while the starting point of each individual singular vector calculation is called initial time t_i. Further, the end point of the TSVD is

denoted as end time t_F, whereas the end point of each individual singular vector calculation is denoted as final time t_f.

4.2.1 Uncertainties of initial conditions

4.2.1.1 Error growth of VOC and NO$_x$ families

The TSVD is applied on the grouped error growth (2.54) first. As the focus is placed on changes in the final ozone concentration due to uncertainties of initial VOC and NO$_x$ concentrations, the final projection operators $\mathbf{F}_{t_f}$ (2.49) target ozone only. The initial family operators $\mathbf{F}_{t_i}$ (2.49) contain the family VOC ($\mathcal{F}_1$) and the family NO$_x$ ($\mathcal{F}_2$). Other species were not taken into account. The associated singular vector analyses answer the following question:

Is the ozone development dominated by changes in the initial conditions of VOC or NO$_x$?

Since the answer to this question may differ for distinct scenarios and varied initial and final simulation times, it is of special interest to address the more in-depth issue:

To what extent are the results dependent on the given scenarios and the simulation interval chosen?

First, each grouped error growth calculated within the TSVD was considered independently. The assigned singular vectors were compared to forward model runs with varying initial concentrations of VOC and NO$_x$, respectively. By visual inspection of about 50 characteristic showcases per scenario, the accuracy of the tangent linear assumption and the correctness of the grouped singular vectors was proved. Fig. 4.4 gives an example for scenario LAND for the simulation interval $[t_0, t_{36}]$. All considered showcases illustrate that the grouped singular vectors correctly identify the direction of largest error growth and that the tangent linear approximation is accurate. Further, the accuracy of the tangent linear assumption was tested for each conducted singular vector calculation by inserting the first singular vector into equation (3.7). Results show, that setting $\alpha = 0.25$ is sufficient to gain $|1.0 - d| \leq 0.01$. Therefore, the ratios are close enough to 1, suggesting that the tangent linear assumption is sufficiently held.

The results of a complete TSVD for scenarios MARINE, FREE and PLUME are presented in Fig. 4.5. In order to visualize the singular vectors, the
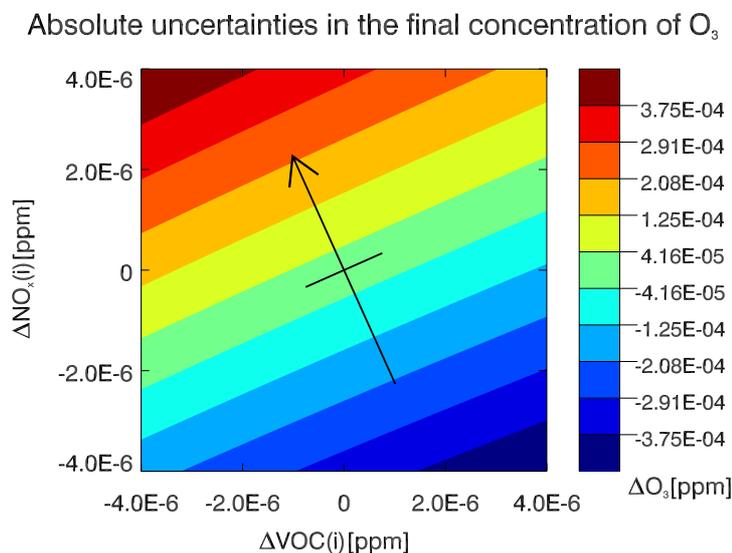
Figure 4.4: *Absolute uncertainties in the final concentration of O_3 for scenario LAND. Isopleths demonstrate the absolute change in the final concentration of O_3 resulting from variations in the initial concentrations of VOC and NO_x, respectively. The arrow indicates the direction of maximal grouped error growth. All calculations started at July 2, 12h and ended 36 hours later at July 3, 24h.*

vector entries corresponding to a specific chemical compound or family are considered separately. Further, the singular vectors were set to unit length. Therefore, an entry of 1 in a specific section of the singular vector indicates that the ozone concentration at final time is solely influenced by this particular section.

Fig. 4.5 reveals a clear distinction between sensitivities with initial time t_i at day and sensitivities with initial time t_i at night, which also occurs for the non-displayed scenarios LAND, URBAN and URBAN/BIO. For all scenarios, main qualitative features are:

i) *The importance of NO_x and VOC changes with initial time t_i at sunset or sunrise. The specific initial condition at different day or night times does not seem to affect the results much.*

ii) *With growing simulation length the initial influence of NO_x and VOC changes significantly. Hence simulation length appears to be another influential feature.*

Fig. 4.5 only allows for a qualitative comparison between the importance of VOC and NO_x initial values. For a direct comparison, the results of the

(a)

(b)

(c)
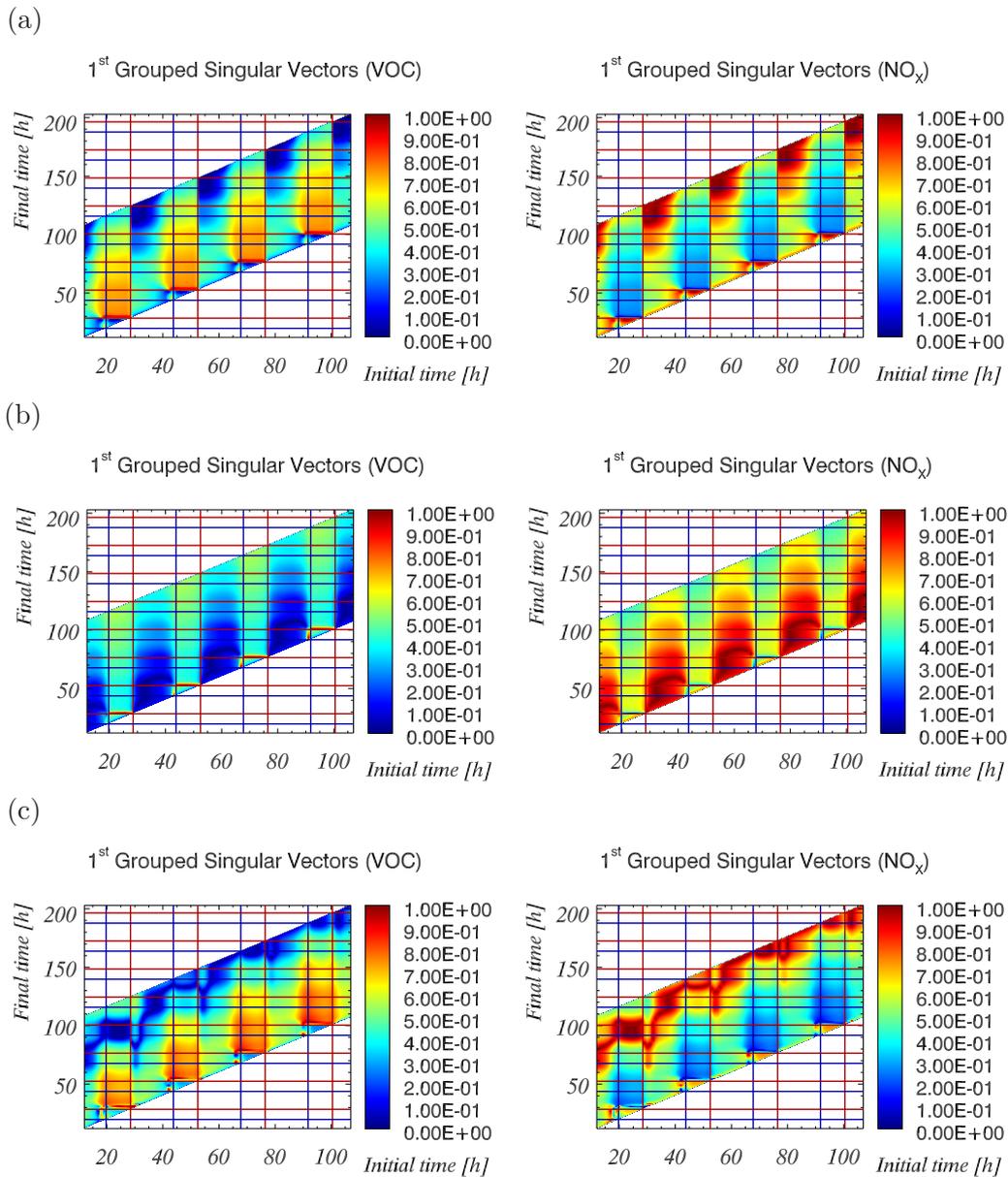
Figure 4.5: *TSVD of the optimal grouped singular vectors for scenarios (a) MA-RINE, (b) FREE and (c) PLUME. VOC-sections of the grouped singular vectors are depicted on the left panel column, NO$_x$-sections on the right panel column. The organization of the results follows the schematic diagram 4.3. Each color pixel indicates an individual singular vector calculation for interval $[t_i, t_f]$. To aid interpretation, singular vectors were set to unit length. Furthermore, sunrises are marked with red lines and sunsets with blue lines, respectively.*

singular vector analyses have been categorized following the findings **i)** and **ii)**. In order to grant unambiguous results, a small amount of the conducted (and for scenarios MARINE, FREE and PLUME already depicted) simulations were not considered for categorization. First of all, simulations with initial time t_i at night and final time t_f before first sunrise were not taken into account. In these special cases, the whole simulation takes place at nighttime when there is no photochemistry. Secondly, simulations with initial time t_i during hours with decreasing or increasing insolation were disregarded for categorization. More precisely, hours with decreasing insolation are defined to be between sunrise and 3 hours after sunrise and hours with decreasing insolation are defined to be between 4 hours before sunset and sunset. Thirdly, for scenario URBAN/BIO (Fig. 4.1(b)) only the biogenic part of the scenario is considered, since the first 36h equal those of scenario URBAN (remember the spin up run of 24h). For the remainder of this section, the term 'results' refers to the results remaining for categorization only. Furthermore, the biogenic part of the URBAN/BIO scenario is denoted as scenario BIO.

Paying attention to feature **i)** first, the results are categorized into results of calculations with initial time t_i at day (category $\mathcal{C}_a$) and results of calculations with initial time t_i at night (category $\mathcal{C}_b$). For each category $\mathcal{C}_i$ a mean impact m_i is calculated following the equation

$$m_i(j) = \frac{1}{n_i} \sum_{i \in \mathcal{C}_i} |\boldsymbol{v}_i^*(j)|, \tag{4.1}$$

where n_i denotes the number of calculations belonging to category $\mathcal{C}_i$ and $\boldsymbol{v}_i^*(j)$ the normalized singular vectors components of species j belonging to category $\mathcal{C}_i$. Furthermore, the absolute minimum and maximum values c_i^+ and c_i^-

$$c_i^+(j) = \max_i |\boldsymbol{v}_i^*(j)| \tag{4.2}$$

$$c_i^-(j) = \min_i |\boldsymbol{v}_i^*(j)| \tag{4.3}$$

as well as the standard deviation

$$s_i(j) = \sqrt{\frac{1}{n_i} \sum_{i \in \mathcal{C}_i} (|\boldsymbol{v}_i^*(j)| - m_i(j))^2}, \tag{4.4}$$

are calculated for each category $\mathcal{C}_i$ to examine the significance of the mean impact m_i.

Fig. 4.6 displays an example of statistical results for categories $\mathcal{C}_a$ and $\mathcal{C}_b$
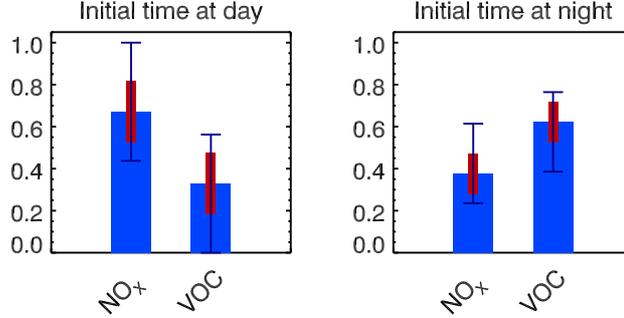
Figure 4.6: *Statistics of optimal grouped singular vectors for categories $\mathcal{C}_a$ and $\mathcal{C}_b$ for scenario MARINE. Depicted are mean impact (blue bars), minimum/maximum value (dark blue lines) and standard deviation (red bars).*

in case of the scenario MARINE. Generally, the standard deviations for the chosen categorization are relatively large. There are scenarios where none of both categories can definitely be assigned to be VOC or NO$_x$ dominated. Therefore, a more refined distinction appears to be advisable. Since simulation length tends to be another influential feature (see finding **ii)**), categories $\mathcal{C}_a$ and $\mathcal{C}_b$ are now further subdivided depending on simulation length. Hence categories $\mathcal{C}_{a_k/b_k}$, k=1, 2, 3, 4 represent results of calculations ending between sunrise k-1 and sunrise k. Thereby, sunrise k, k=1, 2, 3 specifies the kth sunrise after initial time t_i. Sunrise 0 equals time t_i and sunrise 4 equals t_f, respectively.

Showcase results of the statistics of the new categories are illustrated for
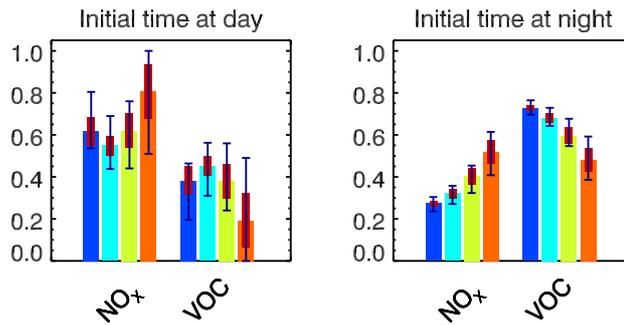
Figure 4.7: *Statistics of optimal grouped singular vectors for categories $\mathcal{C}_{a_{1/2/3/4}}$ and $\mathcal{C}_{b_{1/2/3/4}}$ for scenario MARINE. Depicted are mean impact ($\mathcal{C}_{a_1/b_1}$: blue bars, $\mathcal{C}_{a_2/b_2}$: turquoise bars, $\mathcal{C}_{a_3/b_3}$: green bars, $\mathcal{C}_{a_4/b_4}$: orange bars), minimum/maximum value (dark blue lines) and standard deviation (red bars).*

scenario MARINE in Fig. 4.7. For all scenarios, the subdivision in terms of simulation length leads to a reduction of the standard deviation. Not in

all cases the reduction is large enough to declare the mean impacts to be representative. Table 4.3 and Table 4.4 summarize the categorization results in terms of mean impact and standard deviation for the NO_x-section. Since the singular vectors are normalized, the mean impact of the VOC-section can be derived directly. Notable findings of the categorization are summarized below.

Table 4.3: *Mean impact and standard deviation of NO_x for categories $\mathcal{C}_{a_{1/2/3/4}}$. Considered are optimal grouped singular vectors.*

	LAND	MAR.	FREE	PLUME	URBAN	BIO
$\mathcal{C}_{a_1}$	$0.83^{\pm0.04}$	$0.62^{\pm0.07}$	$0.93^{\pm0.04}$	$0.49^{\pm0.06}$	$0.06^{\pm0.02}$	$0.49^{\pm0.02}$
$\mathcal{C}_{a_2}$	$0.72^{\pm0.09}$	$0.55^{\pm0.05}$	$0.90^{\pm0.05}$	$0.48^{\pm0.05}$	$0.05^{\pm0.02}$	$0.50^{\pm0.02}$
$\mathcal{C}_{a_3}$	$0.71^{\pm0.09}$	$0.62^{\pm0.08}$	$0.77^{\pm0.05}$	$0.60^{\pm0.11}$	$0.04^{\pm0.02}$	$0.51^{\pm0.02}$
$\mathcal{C}_{a_4}$	$0.75^{\pm0.09}$	$0.80^{\pm0.12}$	$0.66^{\pm0.04}$	$0.77^{\pm0.12}$	$0.04^{\pm0.02}$	$0.53^{\pm0.02}$

Table 4.4: *Mean impact and standard deviation of NO_x for categories $\mathcal{C}_{b_{1/2/3/4}}$. Considered are optimal grouped singular vectors.*

	LAND	MAR.	FREE	PLUME	URBAN	BIO
$\mathcal{C}_{b_1}$	$0.31^{\pm0.06}$	$0.27^{\pm0.01}$	$0.61^{\pm0.04}$	$0.24^{\pm0.04}$	$0.05^{\pm0.02}$	$0.37^{\pm0.10}$
$\mathcal{C}_{b_2}$	$0.35^{\pm0.06}$	$0.32^{\pm0.02}$	$0.57^{\pm0.04}$	$0.34^{\pm0.06}$	$0.04^{\pm0.02}$	$0.48^{\pm0.10}$
$\mathcal{C}_{b_3}$	$0.40^{\pm0.06}$	$0.40^{\pm0.04}$	$0.52^{\pm0.03}$	$0.56^{\pm0.18}$	$0.03^{\pm0.02}$	$0.60^{\pm0.14}$
$\mathcal{C}_{b_4}$	$0.45^{\pm0.07}$	$0.51^{\pm0.05}$	$0.48^{\pm0.02}$	$0.73^{\pm0.15}$	$0.03^{\pm0.02}$	$0.71^{\pm0.17}$

For simulations with initial time t_i at day, Table 4.3 indicates, that scenarios with rather clean air are in general more NO_x sensitive than scenarios with polluted air. The high NO_x values for case FREE (representing the cleanest air) and the low NO_x values for case URBAN (representing the most polluted air) are most remarkable. Accordingly, scenario BIO is nearly in VOC-NO_x balance. Further, simulation length tends to change the amount of the NO_x sensitivity, but no clear chains of cause and effect are identifiable.

For simulations with initial time t_i at night, Table 4.4 exhibits that the results resemble for scenarios LAND, MARINE, PLUME and BIO. For the shortest simulation interval, there is VOC dominance, which decreases with growing simulation length. Thereby the rate of decrease appears to be dependent on the degree of air pollution (the cleaner the air the lesser the rate of increase). Results for scenarios FREE and URBAN differ, both show increasing VOC

impact, whereas scenario URBAN remains clearly VOC controlled. In contrast, scenario FREE changes from NO$_x$ dominated to VOC dominated.

In summary there is no generally valid statement about the VOC versus NO$_x$ limitation of the ozone formation. As assumed, the results indicate that the VOC versus NO$_x$ dominance is highly dependent on given scenarios and the chosen simulation interval. According to the questions at the beginning of this section, the key findings are:

- *For simulations with initial time at day, scenarios with cleaner air are more NO$_x$ sensitive.*

- *For simulations with initial time at night, the VOC influence is typically decreasing with simulation length. The rate of decrease appears to be dependent on the degree of air pollution.*

4.2.1.2 Error growth of VOC and NO$_x$ species

The grouped error growth calculations carried out in the previous section were able to analyze the VOC versus NO$_x$ dominance of given situations. The consequential question to be faced in this section is:

To which extent is each individual compound of the NO$_x$ or VOC group responsible for the particular influence of the group?

Calculations with the projected error growth (2.31) can deal with this question by focusing on the individual species belonging to the VOC and NO$_x$ group at initial time t_i and on O$_3$ at final time t_f.

Fig. 4.8 displays results of projected error growth TSVDs for VOC compounds HC8 and ISO for scenario MARINE and gives an idea of the complexity of the underlying chemical processes. Since the concern of this section is a general ranking of the impact of the species, the categorizations and statistics introduced in section 4.2.1.1 were applied. Results indicate that a categorization into simulations with initial time t_i at day ($\mathcal{C}_a$) and simulations with initial time t_i at night ($\mathcal{C}_b$) is reasonable. For a ranking of species however, the projected singular vectors do not benefit much from a subcategorization dependent on simulation length in all cases, as shown in Fig. 4.9. Still there are other cases, where a subcategorization is necessary, since the standard deviation of the mean impact is quite large for categories $\mathcal{C}_a$ and $\mathcal{C}_b$. A careful examination for each scenario is demanded to gain useful results.

For the NO$_x$ compounds, the ranking of the compounds is independent of
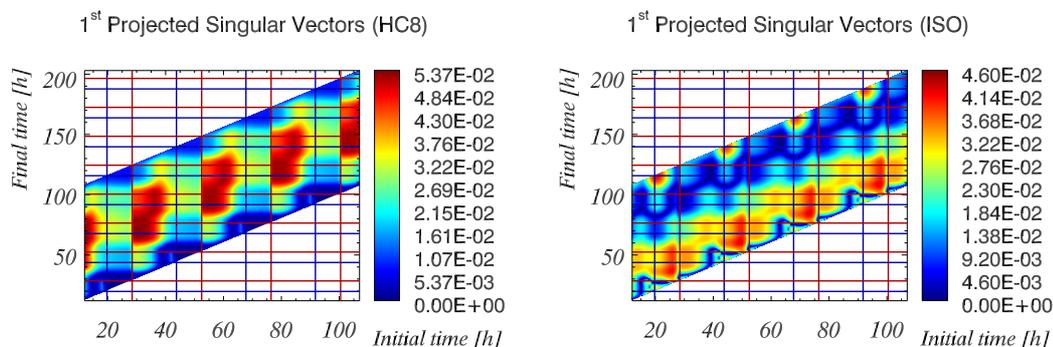
Figure 4.8: *TSVD of the optimal projected singular vectors for scenario MARI-NE. HC8-sections of the projected singular vectors are depicted on the left panel, ISO-sections on the right panel. Plotting conventions as in Fig. 4.5.*
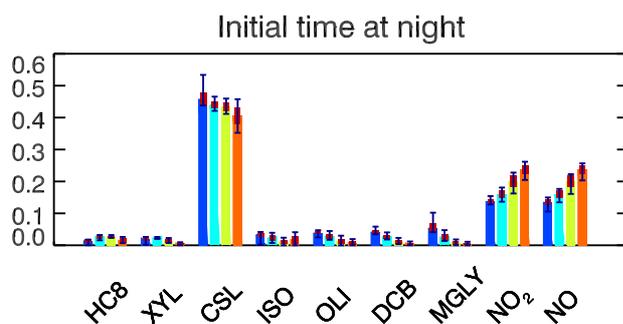
Figure 4.9: *Statistics of optimal projected singular vectors for categories $\mathcal{C}_{b_{1/2/3/4}}$ for scenario MARINE. The VOC and NO_x compounds with most influence have been chosen for presentation. Depicted are mean impact ($\mathcal{C}_{b_1}$: blue bars, $\mathcal{C}_{b_2}$: turquoise bars, $\mathcal{C}_{b_3}$: green bars, $\mathcal{C}_{b_4}$: orange bars), minimum/maximum value (dark blue lines) and standard deviation (red bars).*

categorization. Therefore, statistics for different categorizations are not presented here. The information content of those statistics can be summarized as follows:

  Within the NO_x group, NO and NO_2 show nearly equivalent influence (whereas NO_2 has little more importance).

Statistics for the VOC compounds for each scenario and all categories can be found in the appendix A.2.1. Examination of these statistics does not indicate a category-independent ranking for the VOC compounds. Nevertheless it can be stated:

  Within the VOC group, CSL, TOL, XYL, MGLY, DCB and ISO play a dominant role. In contrast, ETH and HCHO belong to the species with least influence.

Further, the statistics show, that the ranking order for each individual species of the VOC group is dependent on the chosen scenario, the initial time t_i of simulation (day or night) and the length of simulation. The most important findings will be summarized in the following.

Simulations with initial time at daytime

For simulations with initial time t_i at day, the particular influence of the VOC compounds are similar (in terms of ranking order) between scenarios LAND, MARINE, PLUME, and BIO. Here, model species CSL, XYL, TOL, MGLY, and DCB (+OLI for scenarios PLUME and BIO, +HC8 for scenario Marine) have strongest influence, followed by ALD, HC8 (not for scenario MARINE), OLI (not for scenarios PLUME and BIO), ISO, OLT, and HC5 (not for scenario PLUME) and OL2, HC3, and KET (+HC5 for scenario PLUME). Model species HCHO, GLY, and ETH are least important. For all scenarios the ranking is altered with changing simulation length. In general, TOL, HC8, HC5, HC3, and KET show increasing influence, while MGLY, DCB, and ALD show decreasing effect (ALD not for scenario URBAN) with growing simulation length. Scenario FREE, however, does not share these features.

Simulations with initial time at nighttime

For simulations with initial time t_i at night, CSL has remarkably higher influence than for simulations with initial time at day, yielding an outstanding impact for scenarios LAND, MARINE, PLUME, and BIO. Furthermore, the influence of model species ISO and OLI is ranked higher than for simulations with initial time at day for scenarios LAND, MARINE, FREE, and BIO. None of the scenarios shows further similarities in terms of ranking lists, but in terms of time-dependent behavior. For scenarios LAND, URBAN, and BIO, the influence of TOL, HC8, HC5, HC3, and KET tends to increase with growing simulation length, while the influence of MGLY, DCB, and ALD decreases.

So far, investigation of the scenarios has been carried out without any weighting of the perturbations. Hence existing results give insight into the influence of the kinetics without considering mixing ratios of different species. This property leads to domination of error growth by more abundant species. In order to evaluate relative influences of chemical compounds, the problems considered in sections 4.2.1.1 and 4.2.1.2 are investigated with the relative error growth. Unfortunately, a weighting with typical initial concentrations is

not feasible for scenarios LAND, MARINE and FREE, due to the fact that the only reactive carbon containing compounds are CO, CH_4 and HCHO. The remaining VOC have zero concentrations. Therefore, relative singular vector analyses were only applied to scenarios PLUME, URBAN and BIO. Again, the notation BIO indicates that only the biogenic part of scenario URBAN/BIO (Fig. 4.1(b)) is considered.

4.2.1.3   Relative error growth of VOC and NO_x families

Since grouped relative singular vectors weight the influence of chemical compounds by their typical concentrations, the main questions raised in subsection 4.2.1.1 are modified to:

> *Is the ozone development dominated by* relative *changes in the initial conditions of VOC or NO_x?*

> *To what extent are these results dependent on the given scenarios and the chosen simulation interval?*

In order to answer these questions, TSVDs of grouped relative singular vectors were calculated for scenarios PLUME, URBAN, and BIO.

For verification of the validity of the tangent linear assumption and of the accuracy of the grouped singular vectors, about 50 characteristic grouped relative singular vectors of each scenarios were regarded independently first. For those showcases the grouped relative singular vectors were depicted together with the results of simple forward model runs with varying initial concentrations of VOC and NO_x. For simulation interval $[t_{107}, t_{192}]$, Fig. 4.10 provides insight into the results for scenario PLUME. Similar to the depicted case, all samples demonstrate that the calculated grouped relative singular vectors are capable of accurately simulating the relative error growth and that the tangent linear approximation is adequate. In addition, the gradient check ratio introduced in equation (3.7) is applied to each conducted singular vector calculation. Here, the first singular vector is utilized as perturbation, while α is varied. Calculations prove that selecting $\alpha = 0.3$ is sufficient to gain $|1.0 - d| \leq 0.01$. In most of the cases, α can be increased without invalidating the approximation for d. The tangent linear assumption is sufficiently held.

TSVD-results for scenarios PLUME and URBAN are collected in Fig. 4.11. Scenario BIO is not displayed, but its TSVD-results are included in the interpretation of results. Remarkably, there is no similarity between the grouped error growth (section 4.2.1.1) and the grouped relative error growth. The

Relative uncertainties in the final concentration of O$_3$
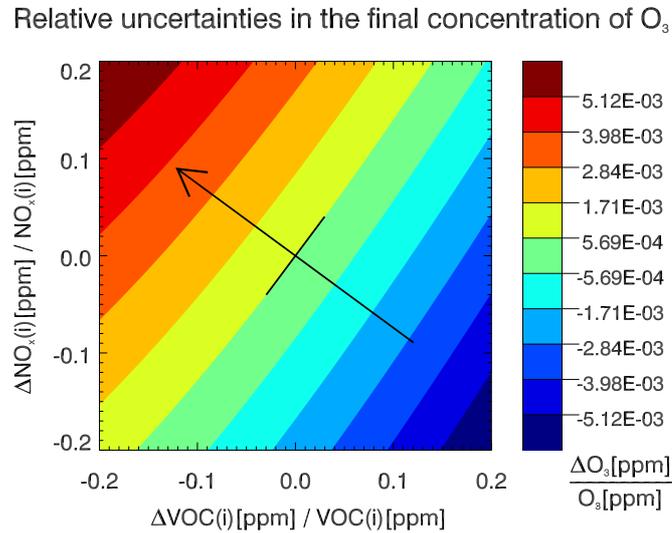
Figure 4.10: *Relative uncertainties in the final concentration of O_3 for scenario PLUME. Isopleths demonstrate the relative change in the final concentration of O_3 resulting from variations in the initial concentrations of VOC and NO_x, respectively. The arrow illustrates the assigned direction of largest grouped relative error growth. All calculations started at July 6, 23h and ended 85 hours later at July 10, 12h.*

singular vector entries of the grouped relative error growth do not show a periodic pattern for initial time t_i at day or night. Instead, each time row TR(i) (for explanation of time rows see Fig. 4.3) has its own pattern, which does not reiterate. Nevertheless, the evolutional behavior of the singular vectors seems to recur.

In order to estimate a more precise line of conduct for the grouped relative error, the evolutional behavior was categorized as introduced in section 4.2.1.1. But neither considering the deviation from the initial influence nor their modifications (taking the squared measure etc.) leads to standard deviations small enough to be meaningful. Another attempt to state the aging characteristic precisely is to apply the initial concentrations of the VOC and NO$_x$ components on the grouped relative error growth. Different applications were tested but none of them leads to precise characteristics. Using the same application to transform the grouped relative error into the grouped error growth did not succeed either. A straightforward pattern that helps to find meaningful time-dependent categorizations for the VOC or NO$_x$ dominance of the scenarios seems difficult to achieve.

Facing these difficulties, the scenarios are analyzed separately first. There-

(a)

(b)
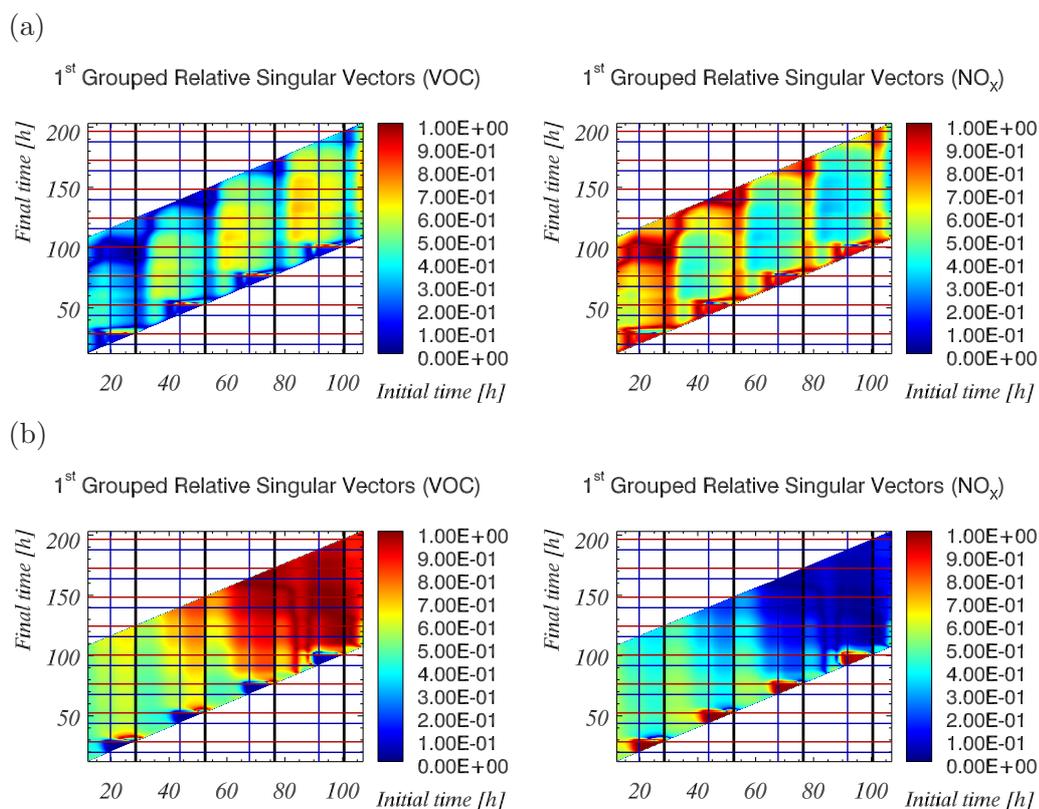
Figure 4.11: *TSVD of the optimal grouped singular vectors for scenarios (a) PLUME and (b) URBAN. VOC-sections of the grouped singular vectors are depicted on the left panel column, NO_x-sections on the right panel column. Plotting conventions as in Fig. 4.5. Here, in exception, sunrises are marked with red and black lines.*

fore, all simulations with initial time t_i within a fixed 24h-interval are denoted as $TSVD_{24}$-section henceforth. In detail, the fixed interval is determined to be $[S_R(i), S_R(i+1)], i = 1, 2, 3$, with $S_R(i)$ denoting the i-th sunrise after starting time t_I. Note, that this definition yields three $TSVD_{24}$-sections. The sunrises $S_R(i)$, that separate the different $TSVD_{24}$-sections, are indicated with black lines in Fig. 4.11. The evolutional behavior of $TSVD_{24}$-sections appears to roughly recur.

Scenario PLUME

For scenario PLUME, the evolutional behavior is well defined (Fig. 4.11(a)). For each time row $TR(i)$, the influence of VOC grows for simulations within the first half of the time row $TR(i)$ (i.e. simulations with shorter simulation lengths, see Fig. 4.3). In contrast, the VOC sensitivity decreases for simulations within the second half of the time row $TR(i)$. Thereby, the rate

of decrease is less for initial time t_i at night. Within each TSVD$_{24}$-section, time rows TR(i) with initial time t_i directly after sunrise have least VOC influence. Comparing the TSVD$_{24}$-sections among each other shows that the later the TSVD$_{24}$-section starts the more influence is gained by VOC.

Scenario URBAN

Scenario URBAN shows a plain evolutional behavior: With growing simulation length, the sensitivity of VOC increases. Similar to scenario PLUME, the evolutional behavior is less pronounced for initial time t_i at night. Depending on changing initial concentrations, the actual initial VOC sensitivity within each TSVD$_{24}$-sections increases the more, the later the associated simulations are starting. For comparison of the TSVD$_{24}$-sections, the same feature applies.

Scenario BIO

In contrary to the other scenarios, scenario BIO shows approximately constant influence within each time row TR(i) for simulations with initial time t_i at day. For initial time t_i at night, the effect of VOC decreases with increasing simulation length. Inside each TSVD$_{24}$-section the VOC influence is strongest for simulations starting after sunrise and reduces for simulations starting afterwards. In compliance with scenarios PLUME and URBAN, scenario BIO gains more VOC influence with growing distance between starting time t_I and start of the TSVD$_{24}$-section.

In summary, the key features of the grouped relative error growth are:

- *The evolutional behavior of the TSVD$_{24}$-sections is roughly periodic.*

- *The later the initial time of the simulations is, the more the influence of NO$_x$ per TSVD$_{24}$-sections decreases.*

- *The NO$_x$-sensitivity is larger for simulations starting during morning hours.*

4.2.1.4   Relative error growth of VOC and NO$_x$ species

For further understanding of the results of the grouped relative singular vectors it is of special interest to raise the following question:

To which extent is each individual compound of the NO$_x$ or VOC group responsible for the particular influence of the group?

The current section aims to answer this question by analyzing projected relative singular vectors. In contrast to projected singular vectors (section 4.2.1.2), projected relative singular vectors demonstrate the influence of *relative* changes in the initial concentrations of each individual compound.

Fig. 4.12 displays results of the TSVD of optimal projected relative singular
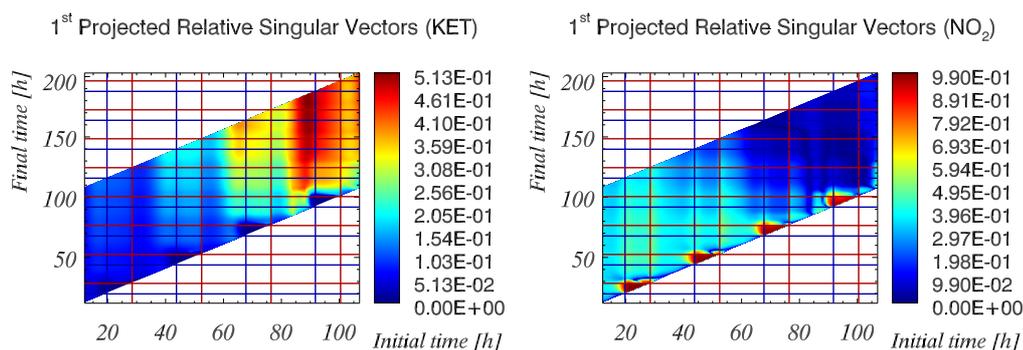
Figure 4.12: *TSVD of the optimal projected relative singular vectors for scenario URBAN. KET-sections of the projected relative singular vectors are depicted on the left panel, NO$_2$-sections on the right panel. Plotting conventions as in Fig. 4.5.*

vectors for compounds KET and NO$_2$ for scenario URBAN. These particular cases reveal at a glance, that the importance of the depicted compounds changes remarkably with growing distance between initial time t_i and starting time t_I. TSVD-results for other compounds and other scenarios are not depicted, but they confirm this finding. Therefore, the categorizations and statistics introduced in section 4.2.1.1 were applied for each day and night separately. Fig. 4.13 provides an example of the outcome for day 1 and night 1 of scenario URBAN. Here, day i contains all results of simulations starting between sunrise i-1 and sunset i, while night i comprises all results of simulations starting between sunset i and sunrise i. Sunrise i and sunset i, $i=1, 2, 3, 4$, specify the ith sunrise and i^{th} sunset after starting time t_I. Sunrise 0 equals starting time t_I.

For each scenario, the statistical results for all days and nights are depicted in the appendix (A.2.2). Their examination yields the following notable findings:

- *Within the VOC group, model species ALD, KET, HC3, HCHO, TOL and HC5 (and ISO for scenario BIO) play a dominant role. The influence of other species can be disregarded.*

- *The influence of all VOC compounds grows with increasing distance*
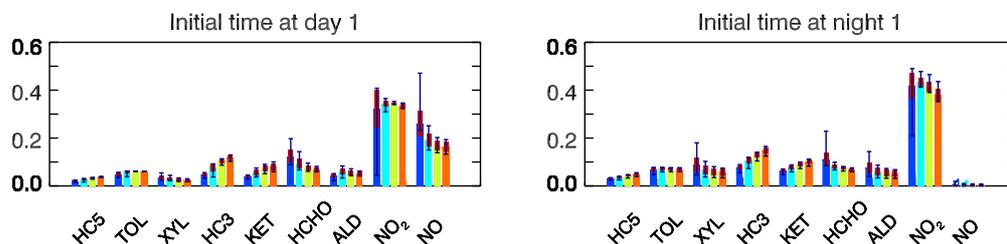
Figure 4.13: *Statistics of optimal projected relative singular vectors for day 1 (left panel) and night 1 (right panel) of scenario URBAN. Depicted are mean impact ($\mathcal{C}_{a_1/b_1}$: blue bars, $\mathcal{C}_{a_2/b_2}$: turquoise bars, $\mathcal{C}_{a_3/b_3}$: green bars, $\mathcal{C}_{a_4/b_4}$: orange bars), minimum/maximum values (dark blue lines) and standard deviation (red bars) for the nine most prominent VOC and NO$_x$ compounds for categories $\mathcal{C}_{a_{1/2/3/4}}$ (left panel) and $\mathcal{C}_{b_{1/2/3/4}}$ (right panel).*

> between starting time t_I and initial time t_i. Compatibly, the effect of the NO$_x$ compounds decreases in the meantime.

> • *Within the NO$_x$ group, NO is less important than NO$_2$. For initial time at night, NO has negligible impact.*

The exact amount of influence for each species of VOC and NO$_x$ differs scenario-wise. In the following, the most important findings per scenario are described shortly.

Scenario PLUME

For scenario PLUME, most of the influence of VOC is induced by ALD, HC3, and KET. Here, ALD gains influence with increasing simulation length, while HC3 and KET are losing effect at the same time. The influence of NO$_x$ is mainly determined by NO$_2$.

Scenario URBAN

Most prominent VOC compounds of scenario URBAN are HC3 and KET. Even though HCHO, ALD, HC5, and TOL have less influence, their effect is still considerable. Within the NO$_x$ group, NO$_2$ and NO have nearly the same amount of influence for initial time t_i at day 1. For the following days, the influence of both compounds declines the more the initial time t_i spreads away from starting time t_I. The rate of decrease, however, is larger for NO.

Scenario BIO

The VOC effect of scenario BIO is mainly determined by the compound ALD.

With growing simulation length, ALD loses influence. Most of the influence of NO_x is induced by NO_2.

4.2.1.5   Comparison of absolute and relative uncertainties

Even though both absolute and relative errors are apparently dependent on the chemical scenario, the simulation length and the initial time, their time-dependent behavior shows substantial differences. Based on recurring patterns, a categorization of the absolute error into simulations starting at daytime and simulations starting at nighttime is reasonable. With a further subcategorization into four different simulation length intervals, the development of the grouped singular vectors is adequately covered. Since relative errors do not show recurring patterns, a comparable categorization is not reasonable here. Grouped relative singular vectors are dependent on the initial concentration of the VOC and NO_x components respectively, with their evolutional behavior of $TSVD_{24}$-sections appearing to recur. Unfortunately, no categorization with adequate accuracy could be imposed for the evolutional behavior. Only a rough approximation of the further development of the relative influence of VOC and NO_x can be made. Hence, it can be stated that:

There is no simple relation between absolute and relative error growth.

Calculations with the projected error growth and the projected relative error growth stated the species that induced the impact of the VOC and NO_x group. For VOC, the species with most absolute influence were CSL, DCB, and MGLY, followed by the compounds HC8, TOL, XYL, ALD, OLI, and ISO. For relative influence this changed to ALD, KET, HC3, followed by HCHO, TOL, HC5, and ISO. For NO_x, a balanced absolute influence of NO and NO_2 changed to a NO_2 dominated relative influence. These differences can easily be explained by different initial concentrations. Differences in structural patterns for each species caused by chemical dynamics are quite complicated. Their explanation is beyond the scope of this work.

4.2.2   Uncertainties of emission factors

In the following, the VOC versus NO_x dependence of the ozone development is analyzed with respect to emissions. For the associated singular vector analyses, emission factors are considered instead of emissions (for reasons of practicability, see section 2.2). Since uncertainties in emission factors

already denote a relative disturbance, only relative error growths are applied here. Because the emission strength of chemical compounds NO$_2$, GLY, DCB, and MGLY is zero for scenarios defined by *Poppe et al.* [2001] (see appendix A.1.3), these compounds obviously have no emission impact within their family. Consequently, the NO$_x$ impact is solely determined by NO. Furthermore, only the scenarios PLUME, URBAN and BIO are analyzed, since other scenarios do not include emissions. Again, the notation BIO indicates that only the biogenic part of scenario URBAN/BIO (Fig. 4.1(b)) is considered.

4.2.2.1 Relative error growth of VOC and NO$_x$ families

The objective of this section is to evaluate the role of emission factors of VOC and NO$_x$ in the ozone development. More precisely, the main questions to be faced here are:

Is the ozone development dominated by changes in emissions of VOC or NO$_x$?

To what extent are these results dependent on the given scenarios and the chosen simulation interval?

In order to provide insight into these aspects, the TSVD is applied to the grouped relative singular vector analysis for scenarios PLUME, URBAN and BIO.

First, about 50 grouped relative singular vectors of each scenarios were considered to verify the validity of the tangent linear assumption and the accuracy of the grouped relative singular vectors. For those characteristic showcases, the grouped relative singular vectors were depicted on the results of simple forward model runs with varying emission factors for each member of VOC and NO$_x$. Fig. 4.14 is a case in point for simulation interval $[t_{72}, t_{137}]$ of scenario URBAN. It captures the main feature, that the tangent linear approximation for emission factors is less accurate than the tangent linear approximation for initial uncertainties in case of grouped relative singular vectors. Nevertheless, in each of the showcases a reduction to 10% perturbation of the emission factors results in a sufficient accurate tangent linear approximation. All samples demonstrate that the calculated grouped relative singular vectors are capable of accurately simulating the relative error growth. In order to ensure that the tangent linear assumption is held for all time series calculations, the gradient check ratio (3.7) is tested for the first singular vectors. It can be proven that setting $\alpha = 0.1$ is sufficient to gain $|1.0 - d| \leq 0.01$. The shorter the simulation interval, the more α can

Relative uncertainties in the final concentration of O₃
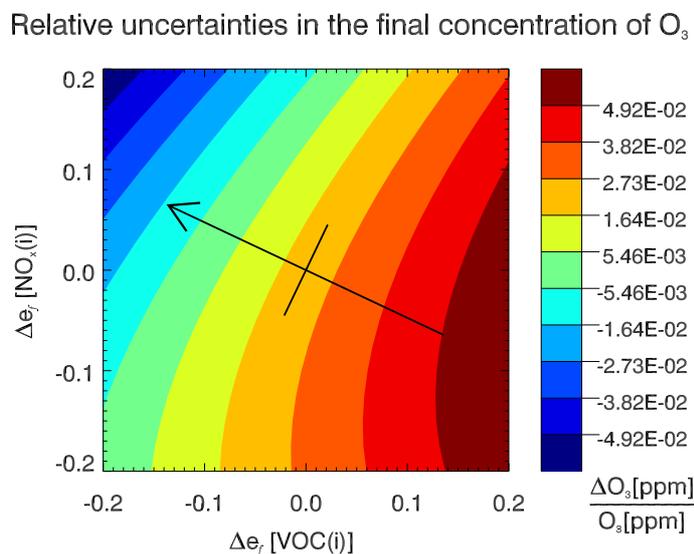
Figure 4.14: *Relative uncertainties in the final concentration of O_3 for scenario URBAN. Isopleths demonstrate the relative change in the final concentration of O_3 resulting from variations in the emission factors of VOC and NO_x, respectively. The arrow illustrates the assigned direction of largest grouped relative error growth. All calculations started at July 5, 12h and ended 65 hours later at July 8, 9h.*

be increased without violating the approximation for d. The tangent linear assumption is sufficiently valid.

Results of the TSVD calculations show substantial differences between scenarios PLUME and URBAN on the one hand and scenario BIO on the other hand. Specifically, *scenario BIO is solely influenced by VOC*. This effect results from the fact, that scenario BIO only includes biogenic emission of isoprene. In contrast, the grouped relative singular vectors of scenarios PLUME and URBAN show a time-dependent behavior. Fig. 4.15 provides insight into the results for the complete TSVD for both scenarios. Note that, in contrast to other TSVD illustrations, for both scenarios only the VOC-entries of the singular vectors are depicted. Furthermore, for scenario PLUME the sensitivity-scale of the VOC entries reaches only 0-20% impact.

Remarkable common features of scenarios PLUME and URBAN are:

- *The emission factors of VOC gain influence with growing distance between starting time t_I and initial time t_i. An exception are simulations with final time t_f before and immediately after the first sunset in simulation time.*
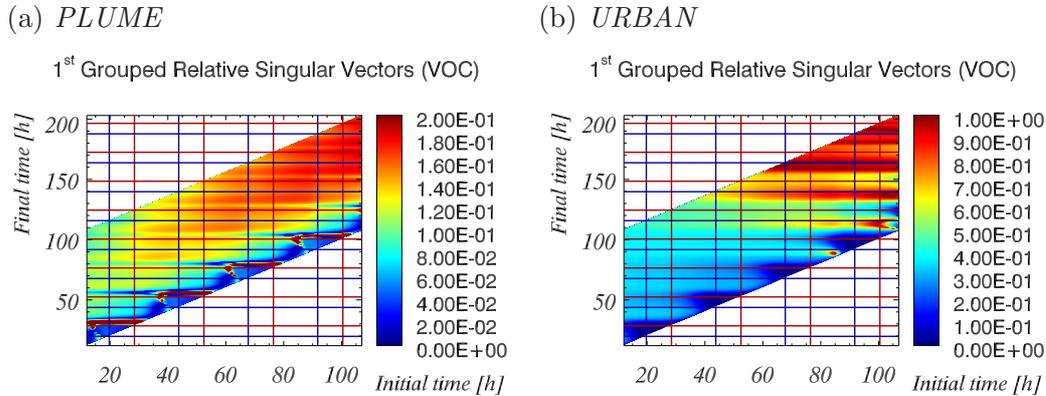
(a) *PLUME*

(b) *URBAN*

Figure 4.15: *TSVD of the optimal grouped relative singular vectors for scenarios (a) PLUME and (b) URBAN. Both panels depict VOC-sections of the grouped relative singular vectors. Plotting conventions as in Fig. 4.5.*

- *The impact of the emission factors of VOC generally grows with increasing simulation length. Simulations with final time t_f around sunrise are an exception, where the influence of VOC temporarily decreases.*

A categorization of the emission factor impact is not advisable since gradual changes of influence appear for both initial times t_i and simulation lengths. This behavior reflects the length of emission exposure on the given scenario. Thereby the VOC sensitivities of scenario URBAN gain influence from 35% (for simulation interval $[t_0, t_{35}]$) to 100% (for simulation interval $[t_{96}, t_{190}]$). For scenario PLUME the described behavior is less pronounced with 9% VOC impact for simulation interval $[t_0, t_{35}]$ and 19% VOC impact for simulation interval $[t_{96}, t_{190}]$. The different rates of increase of influence appear to be associated with the different amounts of NO-emissions applied for scenario PLUME and URBAN.

4.2.2.2 Relative error growth of VOC and NO$_x$ species

In order to provide more detailed information about the VOC and NO$_x$ limitation of the ozone development, the following question is addressed:

To which extent is each individual compound of the NO$_x$ or VOC group responsible for the particular influence of the group?

This question was investigated with TSVDs of projected relative singular vectors. As for the grouped relative errors, results show substantial differences between scenarios PLUME and URBAN on the one hand and scenario

BIO on the other hand. Since scenario BIO only includes biogenic emission of isoprene, the VOC influence of scenario BIO is solely determined by model species ISO.

Fig. 4.16 displays the results of the TSVD of scenario PLUME for OLI and
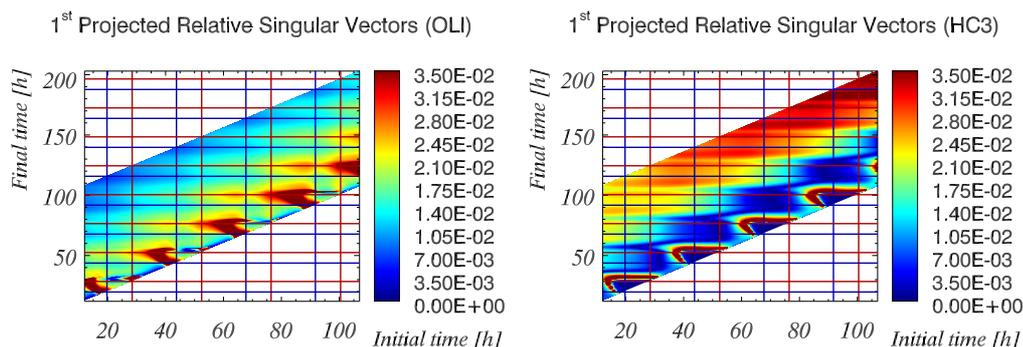
Figure 4.16: *TSVD of the optimal projected relative singular vectors for scenario PLUME. OLI-sections of the projected relative singular vectors are depicted on the left panel, HC3-sections on the right panel. Plotting conventions as in Fig. 4.5.*

HC3. Results for other species and scenario URBAN are not depicted, but nevertheless considered for the investigation. The projected relative singular vectors of scenario PLUME and URBAN show gradual changes of influence for different initial times t_i and different simulation lengths. Furthermore, the influence of different species has distinct gradual changes as Fig. 4.16 points out. Therefore, a categorization for the ranking of species is not advisable.

Visual inspection of the TSVDs for scenario PLUME and scenario URBAN exhibit little similarity. Most notable features shared are:

- *XYL and TOL play a dominant role within the effect of VOC, whereas the influence of HCHO, ETH, and ALD is negligible.*

- *Species ISO, CSL, GLY, MGLY, DCB, and NO_2 are not emitted and therefore have no emission factor impact.*

In the following, scenarios PLUME and URBAN are analyzed separately to provide a more comprehensive picture of the results.

Scenario PLUME

Within scenario PLUME the impact of VOC is almost entirely determined by chemical compounds TOL and XYL. The structural pattern of the influence

of TOL and KET is similar to the structural pattern of the VOC-influence. KET, OLI, HC3, OLI, OLT, HC5, and HC8 show medium impact within the VOC group. Compared to the influence of TOL and XYL, the maximum impact of some of the species with medium influence is quite large. Nevertheless, their mean influence is substantially below their maximum impact. Chemical compounds ETH, ALD, and HCHO show relatively large differences between maximum impact and mean influence as well. They have the least influence within the VOC compounds.

Scenario URBAN

For scenario URBAN, the most dominant species are XYL, HC3, TOL, HC5, and HC8 (sorted by decreasing mean influence). Since the structural pattern of TOL and XYL is less pronounced than the structural pattern of HC3, HC5, and HC8, the order of maximum impact does not match the order of mean influence. Compounds with medium impact within the VOC group are OL2, OLT, OLI, and KET. ALD, ETH, and HCHO have least influence among the VOC species.

Spatial chemical sensitivities for the ZEPTER-2 campaign

The second ZEPpelin based Tropospheric photochemical chemistry expER-iment (ZEPTER-2-campaign, e.g. *Häseler et al.* [2009]) took place between October 17, 2008 and November 8, 2008, employing the unique capabilities of a modified Zeppelin NT ('Neue Technologie', German for new technology) airship as an airborne measurement platform. The Zeppelin was equipped with instruments to conduct in-situ measurements of concentrations of OH, HO_2, O_3, HONO, HCHO, NO_x, CO and several other compounds as well as particular photolysis frequencies. During the campaign 25 flights were carried out within a 100 km radius of the home base Friedrichshafen airport (FDH). Vertical profiles of trace gases were measured above different surfaces, including Lake Constance and surrounding forests. Table 5.1 summarizes accomplished ZEPTER-2-flights and associated vertical profile measurements.

Table 5.1: *List of ZEPTER-2-flights. All flights started and ended at Friedrichshafen airport (FDH). Start and end time are given in UTC, # VP denotes number of vertical profiles.*

Flight	Date	Start	End	Flight-Location	# VP
01	Oct 17	17:20	19:32	Lake Constance	3
02	Oct 18	09:45	14:08	Lake Constance	4
03	Oct 18	14:11	17:45	Lake Constance	2
04	Oct 19	08:45	13:25	Ravensburg, Forest of Altdorf	2

Continued next page

Table 5.1 – Continued

05	Oct 19	13:42	17:36	Ravensburg, Forest of Altdorf	2
06	Oct 20	04:43	10:50	Lake Constance	3
07	Oct 20	11:26	17:18	Lake Constance	3
08	Oct 24	14:38	18:08	Lake Constance	2
09	Oct 25	13:20	16:44	Lake Constance	2
10	Oct 26	12:40	17:01	Forest of Altdorf	3
11	Oct 26	17:40	20:45	Forest of Altdorf	1
12	Oct 27	06:20	06:58	Forest of Altdorf	0
13	Oct 31	15:04	17:54	Lake Constance	1
14	Nov 02	11:02	14:31	Forest of Tettnang	1
15	Nov 02	15:15	17:57	Forest of Tettnang	1
16	Nov 03	10:07	13:40	Ravensburg, Forest of Altdorf	3
17	Nov 03	14:10	17:37	Ravensburg, Forest of Altdorf	2
18	Nov 03	18:01	20:59	Ravensburg, Forest of Altdorf	1
19	Nov 05	10:44	14:31	Hinterland of Lake Constance	4
20	Nov 05	15:50	20:28	Lake Constance	1
21	Nov 07	09:09	13:25	Hinterland of Lake Constance	6
22	Nov 07	15:07	16:40	Hinterland of Lake Constance	1
23	Nov 07	17:27	20:55	Hinterland of Lake Constance	2
24	Nov 08	11:09	14:14	Hinterland of Lake Constance	3
25	Nov 08	14:33	17:19	Lake Constance	1

The collected data of the ZEPTER-2-campaign allows a practical application of the theory of targeted observations. In detail, the singular vector analysis was applied to find out the most sensitive locations and chemical species with respect to their impact on the final concentration of ozone. For subsequent studies, ZEPTER-2-measurements offer the perspective to verify the evaluated sensitivities of chemical species by data assimilation.

5.1 Design of sensitivity experiments

Spatial chemical sensitivities are not only influenced by chemical transformation but also by meteorological processes. Therefore, they do not only determine the measurement priority of each chemical species, but also the optimal placement for each measurement. Accordingly, singular vectors can

help to fulfill the request of effective campaign design. With specific operators (see chapter 2), singular vector analysis can be adjusted to answer campaign specific questions (in terms of targeted observations). In the following, the adjustment of singular vector analysis to the ZEPTER-2-campaign is described. According to chapter 2, the starting time of the considered simulations is denoted as initial time t_I, while the associated end time is denoted as final time t_F.

Since the ZEPTER-2-campaign has already taken place, the singular vector analysis cannot be employed for actual campaign design here. Nevertheless, this study gives an insight in how singular vectors can be applied in future campaigns. Furthermore, accomplished ZEPTER-2-measurements can help to confirm the calculated relevance ranking of measurements of chemical compounds. Therefore, the projection operators $\mathbf{P}_{t_I}$ and $\mathbf{P}_{t_F}$ (2.27) were adapted to focus on ZEPTER-2-measurements. This design provides the possibility to verify the calculated relevance ranking of measurements with data assimilation at a later date. In the process of verification, the measurements at initial time are assimilated in order of sensitivity. Subsequently, the obtained ozone concentrations at final time are compared with the ZEPTER-2-measurements at the same time. The improvement of the forecast of ozone should reflect the relevance ranking of the sensitivities.

The design of the following sensitivity experiments is defined by the chosen projection operators. As mentioned above, the latter focus on ZEPTER-2-measurements. In detail, they are set up to result in answering the following guiding questions for singular vector analysis:

$\mathbf{q}_1$) *Which of the species O_3, NO, NO_2, HCHO, CO, HONO and OH has to be measured with priority to provide an enhanced forecast of locally predetermined ozone profiles?*

$\mathbf{q}_2$) *Where is the optimal location for observations of these components (subject to the above specified aim)?*

According to these questions, the chemical projection of the final projection operator $\mathbf{P}_{t_F}$ focuses on O_3. The local projection at final time is predetermined by ZEPTER-2 vertical profile measurements. Compared to final projections on a single measurement point, final projections on a measurement profile grant a larger magnitude of the optimal initial perturbation. For notational convenience, the location of the vertical measurement profile at final time is denoted as 'final profile $\mathrm{VP}(t_F)$' in the following.

The initial projection operator $\mathbf{P}_{t_I}$ focuses on species O_3, NO, NO_2, HCHO, CO, HONO and OH. The relatively large amount of considered species within

the initial projection set $\mathcal{P}(t_I)$ (2.27) serves the purpose to obtain relevance rankings of species. For local projection at initial time it is not reasonable to focus on locations of measurements solely, since thereby a) the local optimization is omitted, b) different initial air masses are not considered, and c) the dynamics of the system are very limited, resulting in nearly negligible eigenvalues. Hence, no local projection was chosen for initial time. Nevertheless, the possibility of a subsequent data assimilation application is preserved because only profiles and simulation intervals $[t_I, t_F]$ were regarded, where the center of the backward plume of the final profile $\mathrm{VP}(t_F)$ contains ZEPTER-2-measurements. Since only hourly initial times can be considered (due to the current EURAD-IM configuration), 17 simulation intervals meet the conditions for the backward plume of the final profile $\mathrm{VP}(t_F)$. More details about the considered cases can be found in Table 5.2. Cases that share the same final profile $\mathrm{VP}(t_F)$ are indicated with the same case number and subsequent distinct letters.

Table 5.2: *List of all ZEPTER-2 singular vector simulations. Initial time t_I and final time t_F of simulation are given in UTC. VP(t_F) denotes the location of the vertical measurement profile at final time, FDH designates Friedrichshafen airport.*

Case	Flight	Date	Profile	t_I	t_F	VP(t_F)-Location
1a	02	Oct 18	02	12:00	13:30	Lake Constance
1b	02	Oct 18	02	13:00	13:30	Lake Constance
2a	02	Oct 18	03	11:00	14:00	FDH
2b	02	Oct 18	03	12:00	14:00	FDH
3	03	Oct 18	07	15:00	17:35	FDH
4a	04	Oct 19	08	09:00	12:15	Forest of Altdorf
4b	04	Oct 19	08	10:00	12:15	Forest of Altdorf
5a	05	Oct 19	10	14:00	15:20	Forest of Altdorf
5b	05	Oct 19	10	15:00	15:20	Forest of Altdorf
6	06	Oct 20	14	08:00	10:45	FDH
7a	07	Oct 20	16	13:00	14:45	Lake Constance
7b	07	Oct 20	16	14:00	14:45	Lake Constance
8a	08	Oct 24	19	16:00	18:00	FDH
8b	08	Oct 24	19	17:00	18:00	FDH
9a	21	Nov 07	32	10:00	11:25	Mengen
9b	21	Nov 07	32	11:00	11:25	Mengen
10	23	Nov 07	34	18:00	20:50	FDH

For all cases the projected *relative* error growth (2.39) is evaluated for initial value and emission uncertainties. The results will be presented in section 5.2 and section 5.3, after the description of the applied EURAD-IM configuration is presented.

5.1.1 EURAD-IM configuration

For the case studies the upgraded EURAD-IM model (see section 3.4.1 and section 3.4.2) was employed. The applied configuration is based on the ZEP-TER-2-EURAD-IM-setup, which was designed to deliver forecasts and 3D-var analyses for the ZEPTER-2 campaign. As chemistry mechanism, RACM-MIM (section 3.1.2) has been chosen, while meteorological fields are provided by MM5 (NCAR Mesoscale Meteorological Model) simulations. The original ZEPTER-2-EURAD-IM-grid configuration consists of a coarse European grid with a horizontal resolution of 45 km and a time step length of 600 sec, and three nested grids with horizontal resolutions of 15 km, 5 km, and 1 km and time step lengths of 240 sec, 120 sec, and 60 sec, respectively. The finest grid (ZP3) covers the region of Lake Constance. Since all flight trajectories are located within the ZP3-grid, the ZP3-domain is sufficient for the considered case studies. Due to its high horizontal resolution, the ZP3-grid provides a good representativeness of the measurements. Unfortunately, its large amount of grid cells ($N_x \times N_y \times N_z$ grid cells with N_x=236, N_y=146, N_z=23) leads to excessive computational requirements. In order to reduce the CPU time needed by the singular vector calculations, the horizontal size of the ZP3-domain was reduced resulting in a new ZPS-domain with N_x=111, N_y=96. Fig. 5.1 illustrates the horizontal position of the ZPS-domain. It was assured, that all flight trajectories remain within the ZPS-grid. For a reference state in the center of the ZPS-domain, Table 5.3 lists the vertical grid structure in terms of height above ground.

Table 5.3: *Vertical grid structure of EURAD-IM for the reference state (latitude: 47.85336, longitude: 9.500639). Given are model level (ML) and height above ground (HT). The superscripts $^+$ and $^-$ indicate upper and lower boundary of the associated layer.*

ML	HT$^-$ (m)	HT$^+$ (m)
23	10937.5000	14009.1934
22	8766.0986	10937.5000
21	7060.0742	8766.0986
20	5643.5728	7060.0742
19	4426.4463	5643.5728

Continued next page

Table 5.3 – Continued

18	3355.8394	4426.4463
17	2397.9038	3355.8394
16	2040.8547	2397.9038
15	1696.9287	2040.8547
14	1446.9795	1696.9287
13	1203.4612	1446.9795
12	1005.1839	1203.4612
11	810.9417	1005.1839
10	658.3343	810.9417
9	508.1113	658.3343
8	396.9637	508.1113
7	287.0776	396.9637
6	214.5087	287.0776
5	142.4796	214.5087
4	106.6640	142.4796
3	70.9805	106.6640
2	35.4280	70.9805
1	0.0042	35.4280

Emission estimates are provided by the cooperative program EMEP (European Monitoring and Evaluation Programme) with a horizontal resolution of 50 km. The data consists of annual emissions of CO, SO_2, NO_x, NH_3, VOC and particulates ($PM_{2.5}$, PM_{10}) provided for 11 anthropogenic source-sectors. Since the horizontal resolution of all EMEP emission data is not adequate for the considered ZPS-grid, the horizontal resolution of the emission was refined. For the refinement, data sets of COoRdination of INformation on the Environment (CORINE) Land Cover and of United States Geological Survey Global Land Cover Characterization (USGS-GLCC) were combined with data from GIS (Geographic Information Systems). In this manner, emission data sets with a horizontal resolution of 1 km were generated, which are consistent with the overlying EMEP emission data set. Emissions of small towns and busy roads are located correctly. An example for CO-emissions on the ZPS-grid can be found in Fig. 5.1.

Initial concentrations are taken from previously conducted 3D-var assimilation runs. Here, assimilation was accomplished every four hours, starting at 02 UTC. Observational data of NO_2, NO, SO_2 O_3, CO, C_6H_6, $PM_{2.5}$ and
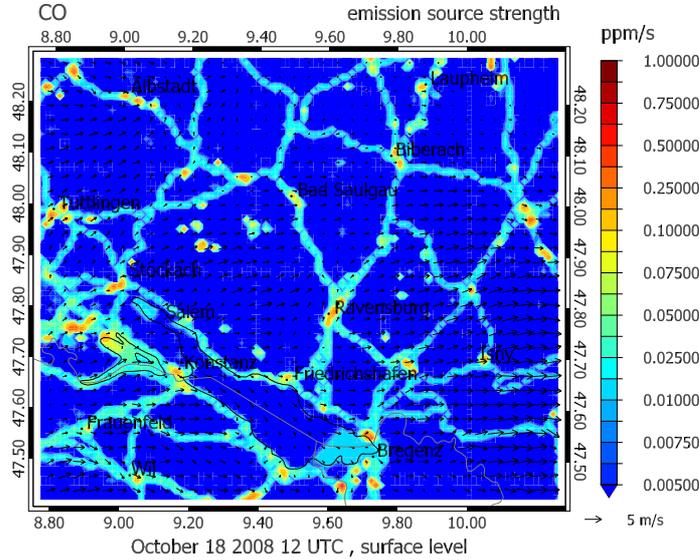
Figure 5.1: *CO emission source strength at surface level of the ZPS-grid.*

PM_{10} were assimilated.

5.2 Singular vectors with respect to initial uncertainties

Singular vector calculations are based on the assumption, that small perturbations evolve linearly. In order to grant meaningful singular vectors, this assumption has to be tested first. Therefore, the calculated singular vectors are inserted as perturbation in equation (3.7). Results demonstrate that $|1.0 - d| \leq 0.001$ is achieved by reducing α to 0.1 (which equals a relative initial disturbance of 10%) in every case. Hence, ratios are close enough to 1 to ensure that the tangent linear approximation is sufficiently accurate.

Singular vectors comprise sensitivities of all species s and grid points (i, j, k) within the initial projection set $\mathcal{P}(t_I)$ (2.27). This set of species and grid points can be analyzed in terms of a) optimal placement of observations and b) measurement priority of species. In detail, each singular vector entry $v(i, j, k, s)$ refers to the sensitivity of species s at grid point (i, j, k). Here, i and j indicate horizontal grid coordinates, while k denotes the considered model level. For each species s, the magnitudes of the singular vector entries $\{v(i, j, k, s)|i, j, k \text{ variable}, s \text{ fix}\}$ determine the relevance of measurements at each grid point (i, j, k). Hence, the grid points with largest magnitudes

define the optimal placement for the considered species. Furthermore, the measurement priority of species is achieved by arranging the singular vector entries $\{\mathrm{v}(i,j,k,s)|i,j,k \text{ fix}, s \text{ variable}\}$ according to magnitude. Note, that the relevance ranking of species may differ for each considered grid point (i,j,k).

In the subsequent exposition, sensitive locations are analyzed first (associated to section 5.1, question **q2)**), followed by sensitive chemical compounds (associated to section 5.1, question **q1)**).

5.2.1   Optimal placement of observations

In order to assess the role of meteorology in the distribution of optimal placement of measurements of chemical compounds, sensitive regions of a passive tracer are compared to sensitive regions of chemical compounds in the next section. Subsequent, the optimal placement of observations of each chemical compound is considered in section 5.2.1.2.

5.2.1.1   Influence of meteorological conditions

In terms of optimal placement of adaptive chemical observations, meteorological conditions obviously play an important role. Indeed, for different singular vector calculations with shared final profiles $\mathrm{VP}(t_F)$, the optimal placement drifts further away from the final profile for longer simulation intervals. At the same time, the areas of maximum error growth are expanding.

In order to analyze the influence of meteorological conditions, the optimal placement of observation of chemical compounds is compared to the optimal placement of observation of a passive tracer. The passive tracer and chemical species are both driven by transport processes, but the passive tracer is not effected by chemical processes (chemical transformation, emission and deposition). Based on this condition, the optimal observation placement of the passive tracer is found by running the singular vector analysis with initial and final species projection on the tracer itself. Before investigating the optimal placement of observations of passive tracer and chemical species, some notations are introduced. First, the term 'placement' is substituted for 'placement of observations'. Furthermore, 'placement of observation of the passive tracer' is shortened to 'meteorological placement' and 'placement of observation of chemical compounds' is shortened to 'chemical placement'.

Results for placements of passive tracer and chemical species can be considered in terms of vertical and horizontal optimal placement. The horizontal

placement distribution disregards effects of the vertical distribution and of different species magnitudes. This is accomplished by the following modification of the singular vector entries $v(i, j, k, s)$:

$$v_h(i, j, k, s) = \frac{v(i, j, k, s)}{|v(k, s)|}. \tag{5.1}$$

Here, each section of the singular vector $\boldsymbol{v}$ with fixed level k and fixed species s (consequently these are 'horizontal sections') is scaled by its length $|v(k, s)|$, where

$$|v(k, s)| := \sqrt{\sum_{i=1}^{i_{max}} \sum_{j=1}^{j_{max}} v(i, j, k, s)^2}. \tag{5.2}$$

The modified singular vector $\boldsymbol{v_h}$ with entries $v_h(i, j, k, s)$ is called horizontal singular vector henceforth. Since $|v(k, s)|$ determines the length of the optimal perturbation of model level k and species s, it reveals the height dependent relevance of each species. In order to disregard effects of species magnitudes, the length $|v(k, s)|$ is scaled by the length of the entire species perturbation:

$$v_v(k, s) = \frac{|v(k, s)|}{|v(s)|}, \tag{5.3}$$

where

$$|v(s)| := \sqrt{\sum_{i=1}^{i_{max}} \sum_{j=1}^{j_{max}} \sum_{k=1}^{k_{max}} v(i, j, k, s)^2}.$$

The vector $\boldsymbol{v_v}$ with entries $v_v(k, s)$ is defined as vertical singular vector. Both vertical and horizontal singular vectors allow for direct comparison of local sensitivities of different species (in terms of horizontal and vertical placement).

In this study, the optimal meteorological placement is compared to the optimal placement of ozone only. The placement of ozone is a reasonable representative for all chemical placements as the comparison of the individual species placement points out (see section 5.2.1.2).

Horizontal placements of ozone and passive tracer for all cases can be found in the appendix (see B.2.1). Fig. 5.2 gives an idea of different outcomes for the horizontal distribution. Keep in mind, that the horizontal placement disregards effects of the vertical placement distribution and of different species magnitudes. For an comprehensive picture of the placement distribution, vertical and optimal placement have to be combined. Fig. 5.2 illustrates,
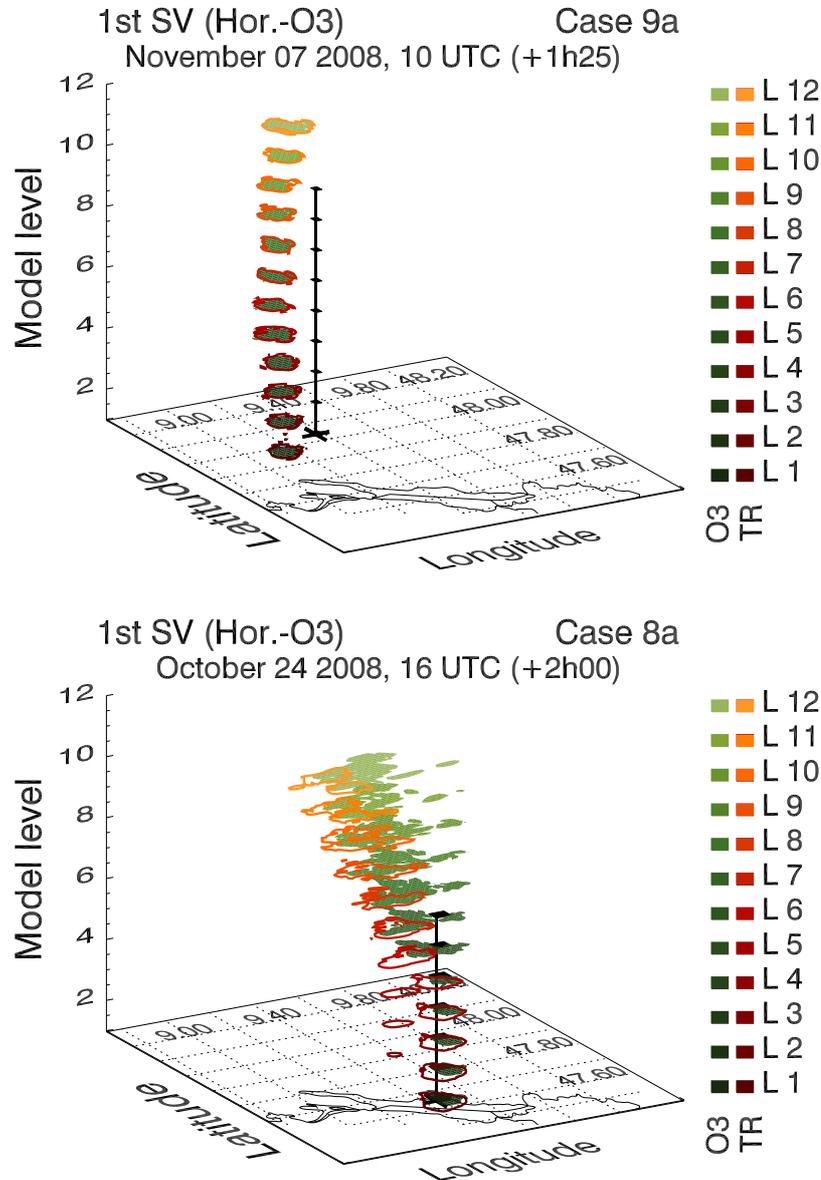
Figure 5.2: *Optimal horizontal placement of measurements for case 9a (top panel) and case 8a (bottom panel). The horizontal placement distribution disregards effects of the vertical distribution and of different species magnitudes. Shown are 0.01-isopleths of the horizontal singular vector for passive tracer (red framed shading) and ozone (green filled shading). The optimization criterion are the ozone values of the final profile $VP(t_F)$(marked with a black line, the black cross indicates its horizontal position). Case numbers and simulation intervals are given on top of each panel.*

that the optimal horizontal placement for case 9a is completely driven by meteorology. Examination of all cases reveals, that meteorological conditions do not only determine the horizontal placement for case 9a, but also for case 9b (Fig. B.16). Interestingly, the location of the final profile $\mathrm{VP}(t_F)$ of cases 9a and 9b is clearly different from the locations of other considered final profiles $\mathrm{VP}(t_F)$ (see Table 5.2). All remaining cases show differences between horizontal placement of passive tracer and ozone. Fig. 5.2 shows subtle differences between meteorological and chemical placement for case 8a. Nevertheless, the overlap of chemical and meteorological horizontal placement is relatively large for all cases. This behavior confirms the important role of meteorology for the optimal chemical placement. Recall that the considered cases share relative short simulation intervals (the longest simulation interval lasts 3h15) and a local projection on the final profile $\mathrm{VP}(t_F)$. Both features restrict the dynamics of the system. Presumably, meteorological and chemical placement differs more in case of longer simulation intervals and/or a less restrictive final local projection. Further placement examination reveals that *meteorological and chemical horizontal placement generally disperse more in higher model levels.*

Fig. 5.3 provides an overview of the vertical placement of passive tracer and ozone. The results for the selected cases reveal at a glance that *chemical placement is less important than meteorological placement in higher model levels.* This feature is also confirmed by the cases not visualized here.

In summary, it can be stated that:

Meteorology conditions are an important factor for the optimal observation placement of chemical compounds, but their consideration is not sufficient to determine the optimal placement.

5.2.1.2 Influence of chemical compounds

To assess the species dependence of the optimal placement, the horizontal and vertical placement for each species is examined.

Horizontal placement

General results for the horizontal placement reveal, that the considered species tend to have approximately the same optimal horizontal placement. Remarkable differences within the chemical placement are discovered for case 6, case 7a, case 8b and case 10 only. In each of these cases, the differing horizontal placements are locally within the optimal placement area of ozone (the optimal placement of ozone for all cases is depicted in appendix B.2.1). Therefore,
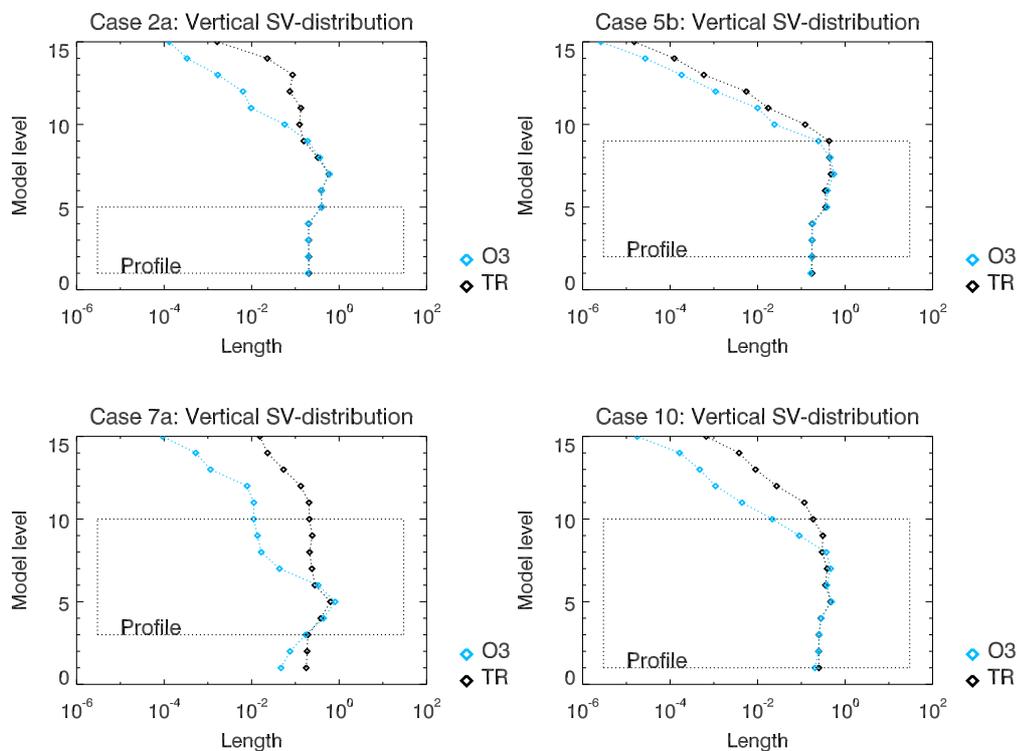
Figure 5.3: *Optimal vertical placement for case 2a (top left), case 5b (top right), case 7a (bottom left) and case 10 (bottom right). The vertical placement distribution disregards effects of different species magnitudes. Shown is the length of the vertical singular vector per model level for passive tracer (black diamonds) and ozone (blue diamonds). The optimization criterion are the ozone values of the final profile $VP(t_F)$. The black box indicates the extent of the height of the final profile $VP(t_F)$.*

horizontal placement differences are considered with regard to the horizontal placement of ozone. The most important findings for the placement differences of case 6, case 7a, case 8b and case 10 are summarized in the following.

In case 6 and case 7, the distinct horizontal placement occurs only at lower model levels. For case 6, horizontal placement differences (with regard to the horizontal placement of ozone) are detected for NO_2, NO and HONO in model levels 1-4. In case 7a, the horizontal placement of all considered species differs from the horizontal placement of ozone at surface level. A closer look reveals, that these horizontal placement differences appear to be initial value related. Fig. 5.4 displays the initial value dependence of the horizontal placement for case 6.
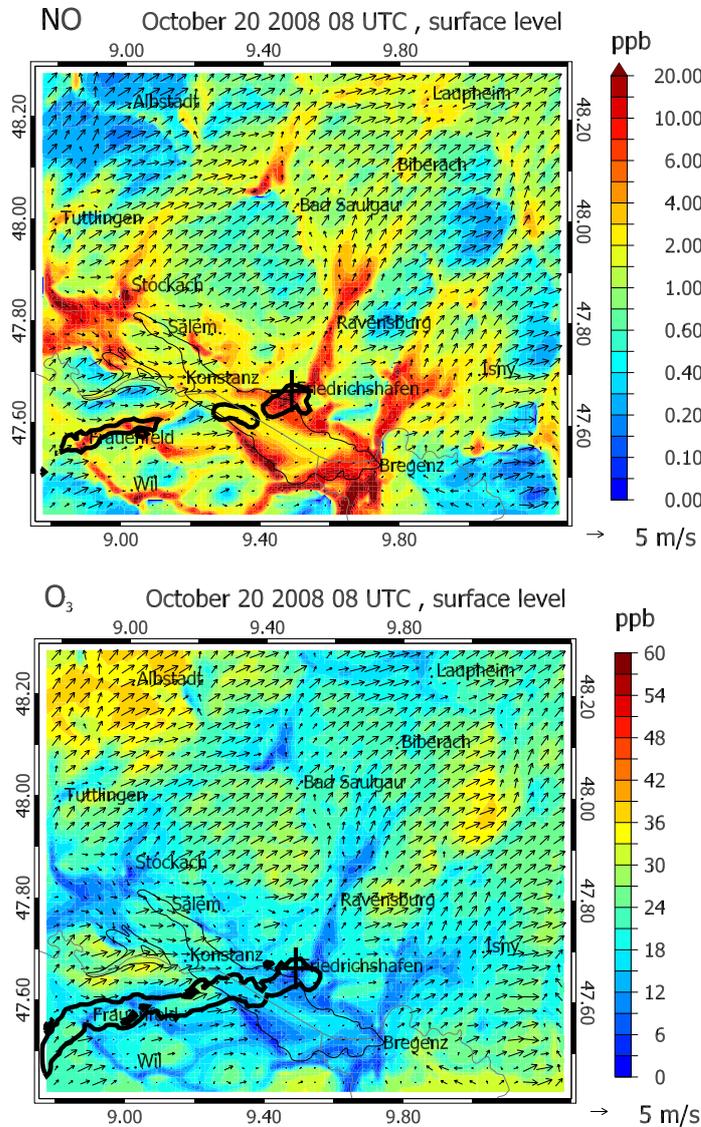
Figure 5.4: *Initial concentrations and optimal horizontal placement of NO (top panel) and O_3 (bottom panel) at surface level for case 6. The 0.01-isopleths of the optimal horizontal placement are indicated with black lines. The optimization criterion for the horizontal placement are the ozone values of the final profile $VP(t_F)$ (the horizontal position of the final profile $VP(t_F)$ is marked with a black cross). Date and time are denoted above each panel.*

In case 8b and case 10, the horizontal placement area of NO is remarkably more compact than the horizontal placement area of ozone. Both cases share relatively late initial times t_I (see Table 5.2). The relative error growth of

NO appears to be dependent on the initial time t_I, presumably due to the associated initial concentrations of NO.

Vertical placement

Fig. 5.5 provides an example of the optimal vertical placement for case 1a. Results for other cases can be found in the appendix (see B.2.1). Keep in
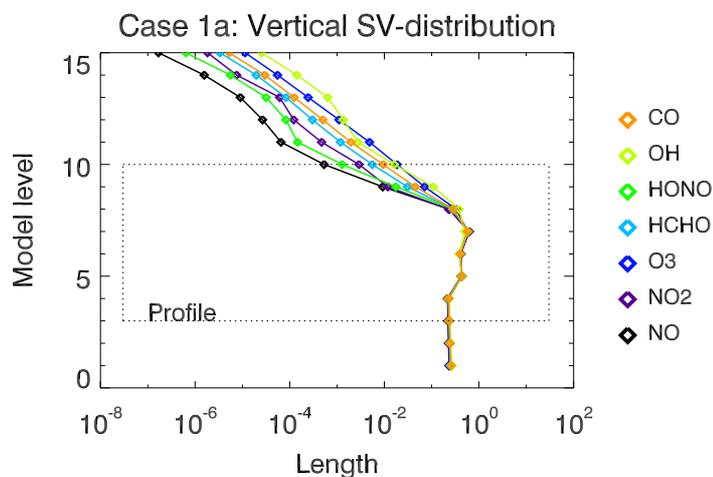
Figure 5.5: *Optimal vertical placement of chemical compounds for case 1a. The vertical placement distribution disregards effects of different species magnitudes. Shown is the length of the vertical singular vector per model level for CO, OH, HONO, O_3, NO_2 and NO (color code of the species is denoted to the right of the panel). The optimization criterion are the ozone values of the final profile $VP(t_F)$. The black box indicates the extent of the height of the final profile $VP(t_F)$.*

mind, that the vertical placement disregards effects of species magnitudes. Therefore, Fig. 5.5 does not illustrate a priority of measurement with respect to species, but only a height dependent relevance ranking of measurements for each individual species. Investigation of the vertical length distribution exhibits, that the relevance of measurements of all species decreases in higher model levels. Remarkably, the model level, at which the decrease becomes apparent, is not indicated by the height of the final profile $VP(t_F)$. The rate of decrease differs among the chemical compounds. The relevance of measurements of O_3 and CO decreases slower than the relevance of measurements of NO and HCHO for all considered cases. This behavior is not directly depending on initial time t_I or simulation length. Presumably, the different rates of decrease are based on different species concentrations in higher model levels.

5.2.2 Relevance ranking of chemical compounds

According to question $\mathbf{q_1}$), the different influences of species O_3, NO, NO_2, HCHO, CO, HONO and OH on the ozone values of the final profile $VP(t_F)$ are examined in this section. A relevance ranking of the associated chemical compounds can be established for each grid point (i, j, k) by arranging the associated singular vector entries $v(i, j, k, s)$ according to magnitude. The most sensitive species at grid point (i, j, k) is denoted to be ranked first, the second most sensitive species is ranked second and so forth. Since seven species are considered for optimization (O_3, NO, NO_2, HCHO, CO, HONO and OH), each species can be ranked m^{th}, m=1,...,7. The relevance ranking of species s may differ for each considered grid point (i, j, k). In order to gain a spatially more comprehensive pattern, the rankings are generalized for the area, that comprises all grid points (i, j, k) with $\sqrt{\sum_s v(i, j, k, s)^2} > 10^{-4}$. Since the selected area is large enough to contain different air masses, a more general picture is obtained.

Within the considered area, a relative ranking $rk(k, s)$ for each species s and each model level k is established, indicating a general relevance ranking for the selected area. Each relative ranking $rk(k, s)$ comprises the relative ranks $rk^m(k, s)$, m=1,...,7 (due to the amount of considered species). The relative rank $rk^m(k, s)$ simply counts how often species s is ranked m^{th} within the considered area of level k and then divides this number by the number of considered grid points:

$$rk^m(k, s) := \frac{\sum_i \sum_j p(i, j) \cdot r(i, j)}{\sum_i \sum_j p(i, j)}, \tag{5.4}$$

where

$$p(i, j) = \begin{cases} 1 & \sqrt{\sum_s v(i, j, k, s)^2} > 10^{-4} \\ 0 & \text{elsewhere,} \end{cases}$$

$$r(i, j) = \begin{cases} 1 & \text{if species } s \text{ is ranked } m^{th} \text{ at grid point } (i, j, k) \\ 0 & \text{elsewhere.} \end{cases} \tag{5.5}$$

Fig. 5.6 provides an example of the relative ranking of O_3 and CO for model level 1 and model level 9. Results for all species and levels can be found in the appendix B.2.2. Note, that the number of grid points (i, j, k) with $\sqrt{\sum_s v(i, j, k, s)^2} > 10^{-4}$ decreases with growing height. If a case is not depicted for a particular level, then the number of grid point (i, j, k) which
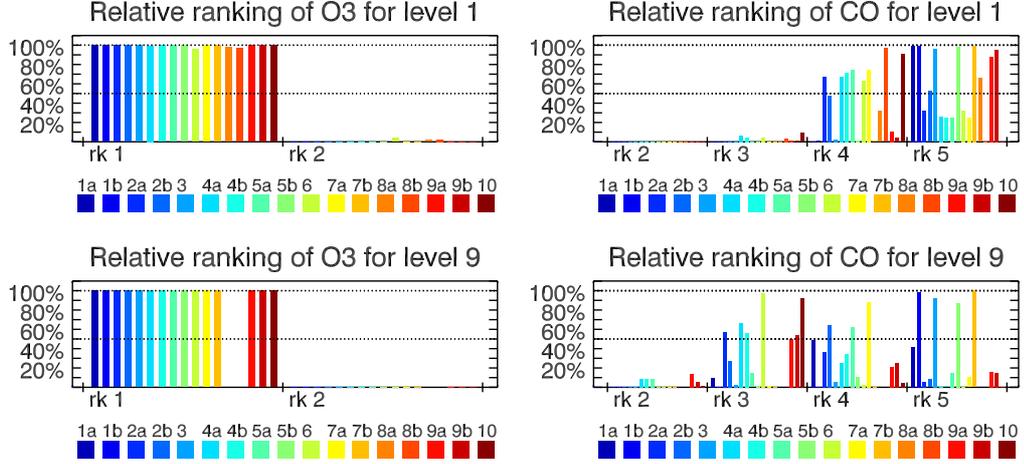
Figure 5.6: *Relative ranking of the effect of initial uncertainties of O_3 (left panel column) and CO (right panel column) for model level 1 (top panel row) and level 9 (bottom panel row) for all 17 case studies. The optimization criterion are the ozone values of the final profile $VP(t_F)$. Results are sorted by relative ranks (denoted below each bar plot). Rank m is only depicted, if the associated species is ranked m^{th} for at least one considered grid point. Color code of cases is indicated below each panel.*

hold $\sqrt{\sum_s \mathrm{v}(i,j,k,s)^2} > 10^{-4}$ is simply zero. The ranking of ozone is not depicted in the appendix, since O_3 is ranked first in nearly every level and every case (see Fig. 5.6). For less than 5% of the considered grid points, O_3 is listed as the second most important species in surface level. None of the other species reveals such a plain behavior in all cases. Examination of all results reveals that in terms of initial uncertainties

- O_3 *has most relevance among the considered species,*

- *NO, NO_2, CO and HCHO show average influence and*

- *OH and HONO have least effect.*

In most cases, the relevance of OH is ranked 7th, while the effect of HCHO is ranked 6th. Furthermore, the relevance of CO increases with growing height (see Fig. 5.6).

In order to analyze, whether the relative ranking is dependent on simulation length or initial time t_I, the cases have been sorted according to both factors. This approach reveals that the influence of initial uncertainties of NO is depending on initial time t_I. The effect of NO is ranked higher for initial time

t_I during midday hours and reduces rank for initial time t_I in the morning or in the afternoon/evening. This feature appears to be related to the initial concentration of NO. The influence of the other species is not depending on initial time t_I or simulation length. Presumably, their differing rankings are based on different air masses.

5.3 Singular vectors with respect to emission factors

Prior to analyzing the singular vectors with respect to emission factors, the linearity assumption is tested. For this purpose, the calculated perturbations of largest error growth are inserted in equation (3.7). Reducing α to 0.1 (which equals an emission factor disturbance of 10%) ensures $|1.0 - d| \leq 0.01$ for each considered case. Note, that in most cases even $|1.0 - d| \leq 0.001$ is achieved. Therefore, the tangent linear approximation is considered to be sufficiently accurate.

Singular vectors with respect to emission factors can be analyzed in terms of optimal placements (with regard to question **q1**) and optimal chemical composition (with regard to question **q2**) for adaptive observations. Only the surface level is considered for examination of emission factors. Furthermore, the singular vector entries of species O_3, HONO and OH are not considered for analysis, since O_3, HONO and OH are not emitted. Under these conditions, sensitive areas are analyzed in section 5.3.1, while sensitive chemical compounds are addressed in section 5.3.2.

5.3.1 Optimal placement of observations

In contrast to initial values, emission factors are not assigned to a fixed point in time but to the entire simulation interval. Hence, the area, where placements can have an impact is provided by the backward plume of the entire simulation interval, originated at the final profile $VP(t_F)$. This plume expands with (backward) time, but does not drift away from the final profile $VP(t_F)$. It is obvious, that the optimal placement of emissions will strongly be influenced by locations of emission sources within this plume.

In order to inspect the optimal placement of emissions, the optimal horizontal placement of emissions was compared to the optimal horizontal placement of initial values. Fig. 5.7 gives an example of this comparison for case 5a. Here, other cases and species are not given, but their results confirm that:
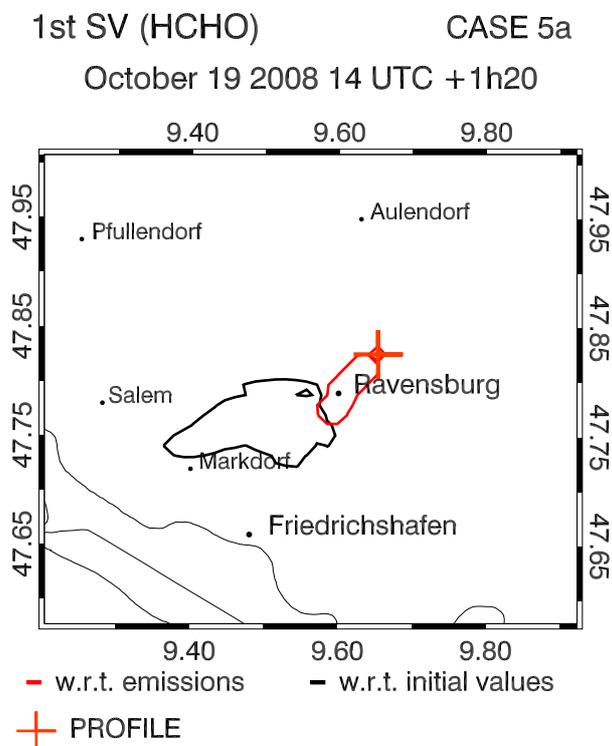
Figure 5.7: *Optimal horizontal placement of emissions and initial values for HCHO at surface level for case 5a. The horizontal placement distribution disregards effects of the vertical distribution and of different species magnitudes. The 0.01-isopleths of the optimal horizontal placement are indicated with black lines (for initial values) and red lines (for emissions). The optimization criterion are the ozone values of the final profile VP(t_F). The horizontal position of the final profile VP(t_F) is indicated with a red cross.*

- *placement of emissions is usually located closer to the final profile $VP(t_F)$ than placement of initial values,*

- *placement of emissions is closely related to emission source locations.*

Even though the area where placements can have an impact is larger for emissions than for initial values, the actual horizontal placement area of emissions is smaller. Since the horizontal placement distribution disregards effects of the vertical distribution and of different species magnitudes, the horizontal singular vector sections for a fixed species and a fixed model level have unit length. Therefore, the small placement areas of emissions lead to the conclusion, that the relevance of placements of emissions is relatively high (compared to local sensitivities of initial values) at few grid points and decreases fast for the surrounding grid points. Further, the placement of

emissions of different species differs quite a lot in some cases, while in other cases all species show the same placements. This behavior occurs due to different emission source strengths for different species and will be explained in more detail at the end of the next section.

5.3.2   Relevance ranking of chemical compounds

According to question $\mathbf{q}_1$), a relevance ranking of the emission influences of NO, NO$_2$, HCHO and CO is assessed in this section. Note, that species O$_3$, OH and HONO are not considered since they are not emitted. Following the approach of section 5.2.2, an area is encompassed which comprises all grid points (i, j, k) with $\sqrt{\sum_s \mathrm{v}(i, j, k, s)^2} > 10^{-4}$. Within this area, a relevance ranking of emission influence is calculated for each grid point. For each species s and each model level k these rankings are subsumed to relative rankings $\mathrm{rk}(k, s)$, which comprise the relative ranks $\mathrm{rk}^m(k, s)$. Since only four species are considered for optimization (NO, NO$_2$, HCHO and CO), m takes value between one and four.

Results for all considered levels and species are depicted in Fig. 5.8. Based
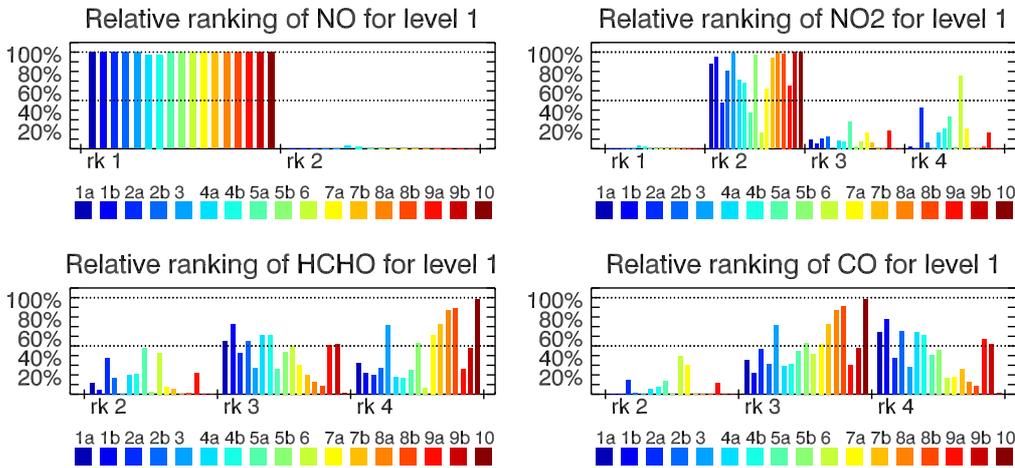
Figure 5.8: *Relative ranking of the effect of emission uncertainties of NO (top left), NO$_2$ (top right), HCHO (bottom left) and CO (bottom right) for surface level for all 17 case studies. Plotting conventions as in Fig. 5.6.*

on these results, it can be stated that:

- *The influence of emissions of NO is most important.*

- *The influences of emissions of NO_2, CO and HCHO are less important than the effect of emissions of NO. Emissions of NO_2 tend to have second most influence, while emissions of CO and HCHO generally alternate between third and fourth rank.*

In order to understand the cause for differing rankings, the cases have been sorted according to initial time t_I and simulation length, but no clear dependencies could be assigned. Nevertheless, locations are assumed to influence emission factors, since emission sources are clearly locally delimited. Hence, the location dependence of the rankings is examined. Fig. 5.9 and Fig. 5.10
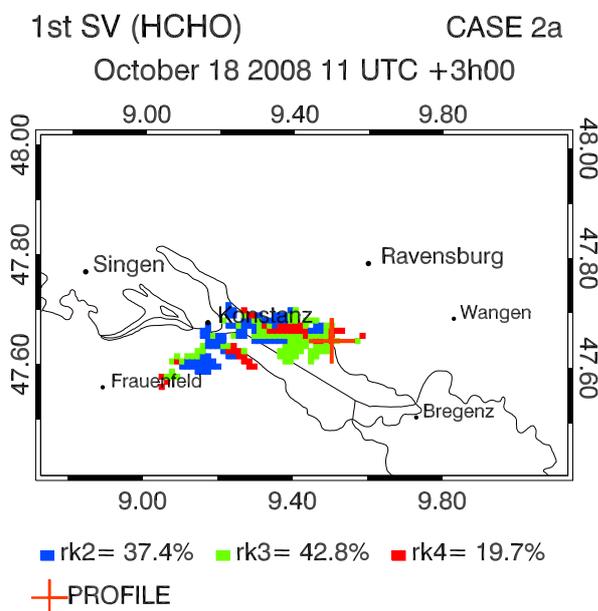
Figure 5.9: *Location dependent relevance rankings of emission influences of HCHO at surface level for case 2a. Color code of ranks is indicated below the panel. Each grid point (i,j,k) with $\sqrt{\sum_s v(i,j,k,s)^2} > 10^{-4}$ is evaluated. The optimization criterion are the ozone values of the final profile $VP(t_F)$. Its horizontal position is indicated with a red cross.*

serve to give an idea about the location dependence of rankings for emission influences of HCHO and CO for case 2a. Results for HCHO and CO for all cases can be found in appendix B.3. Based on the analyses of all 17 cases, the following conclusions can be drawn:

- *The importance of emissions of HCHO tends to increase in urban plumes at the expense of the influence of emissions of CO and NO.*
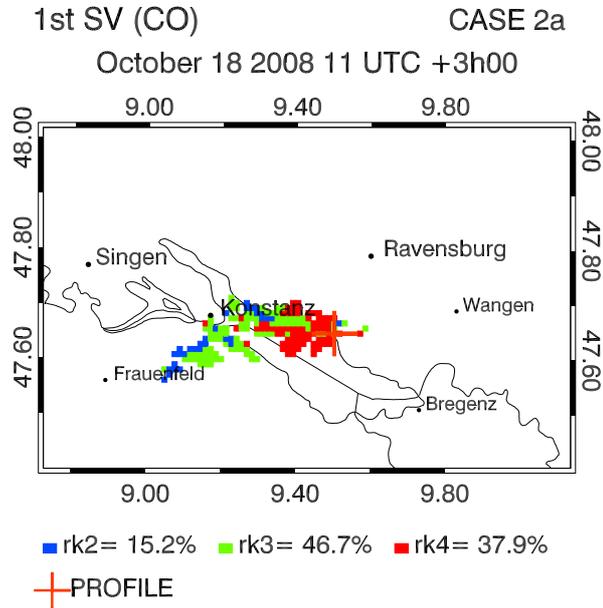
Figure 5.10: *Location dependent relevance rankings of emission influences of CO at surface level for case 2a. Plotting conventions as in Fig. 5.9.*

- *The influence of emissions of CO tends to increase at busy roads. As compensation, the influence of emissions of HCHO and NO decreases.*

Both findings can be utilized for optimal placement as well.

5.4 Magnitudes of singular values

The singular values associated with the optimal singular vectors analyzed in section 5.2 and section 5.3 determine the relative error growths $g_{\mathrm{pr}}\left(\delta \boldsymbol{c}_{\mathrm{pr}}(t_I)\right)$ (2.39) and $\tilde{g}_{\mathrm{r}}\left(\delta \boldsymbol{e}_f\right)$ (2.75), respectively. Table 5.4 captures the singular values for the ZEPTER-2 calculations for both target variables (initial values and emissions). Here, only singular values of simulations with a shared final profile $\mathrm{VP}(t_F)$ are listed.

Singular values of both target variables are shown in Table 5.4 to be depending on simulation length. The influence of initial values decreases with growing simulation length, whereas the effect of emissions increases. Nevertheless, the rate of decrease and increase, respectively, is strongly depending on the case considered and the selected target variable. The same dependence applies for the absolute amount of the singular values.

Table 5.4: *Singular values (SV) for initial values (IV) and emissions (EM) as target variables. VP(t_F) denotes the considered final profile and ML the associated model levels. (a) marks the longer simulation interval and (b) the shorter simulation interval, $t(a)$ and $t(b)$ are the associated simulation lengths.*

VP(t_F)	t(a)	t(b)	SV$_{IV}$ (a)	<	SV$_{IV}$ (b)	SV$_{EM}$(a)	>	SV$_{EM}$(b)
1, ML: 3-10	1h30	0h30	0.3376	<	0.6218	0.0271	>	0.0100
2, ML: 1- 5	3h00	2h00	0.2388	<	0.3294	0.0960	>	0.0931
4, ML: 3- 9	3h15	2h15	0.2770	<	0.3506	0.0724	>	0.0550
5, ML: 2- 9	1h20	0h20	0.5239	<	1.0064	0.1123	>	0.0592
7, ML: 3-10	1h45	0h45	0.4215	<	0.6127	0.0463	>	0.0340
8, ML: 1- 7	2h00	1h00	1.5171	<	1.6142	2.7603	>	1.3248
9, ML: 1- 9	1h25	0h25	0.6886	<	0.8065	0.0353	>	0.0384

Table 5.4 furthermore points out that for most of the calculated ZEPTER-2-cases the magnitude of the singular values is smaller than 1. This means, that the final perturbation is smaller in magnitude than the perturbation of initial values or emission rates (the perturbation of initial values or emission rates is scaled to unit length). However, it should be considered, that the final projection focuses on the ozone perturbation of the final profile VP(t_F). It seems obvious, that the magnitude of the ozone perturbation of the final profile VP(t_F) is smaller in amount than the magnitude of the locally not focused initial value perturbation. For emission rates, three features restrict the dynamics of the system. Firstly, the final species projection is on ozone, but ozone itself is not emitted. Secondly, the final local projection is on a vertical profile, whose vertical extensions range between model level 1 and model level 10. Since the emissions do neither influence the entire vertical profile nor the concentration of ozone directly, some integration time is needed before the effect of emissions on the final perturbation becomes apparent. Remarkably, case 8a and case 8b show singular value magnitudes over 1 (for initial value optimization case 5b as well) despite those restriction for the influence of initial values and emissions. In these cases, targeted observations are especially valuable.

CHAPTER 6

Summary and conclusion

In this work, singular vector analysis was applied to atmospheric chemical modeling in order to optimize the measurement configuration of chemical compounds. Initial chemical values and emissions were investigated as target variables. With the help of newly introduced operators, specific questions of atmospheric chemistry were addressed.

The application of emissions as target variables is a newly introduced feature within the field of targeted observations. Singular vector analysis considers target variables at a particular point in time. Formally, targeting observations of emissions for a given simulation interval hence results in as many singular vector calculations as time steps considered. In order to reduce the amount of optimizations problems, the emission factor approach of *Elbern et al.* [2007] was adopted. Instead of emission strengths, the amplitudes of the diurnal profile of emission rates are analyzed. Due to the fact, that the daily evolution of emissions is far better known than the total emitted amount in a grid cell (*Elbern et al.* [2007]), this choice is reasonable. Since emission factors are time-invariant, only a single singular vector calculation has to be considered per simulation interval.

Special arrangements were made to adequately represent particular problems of atmospheric chemistry. Error growths were extended to allow for projected target variables not only at final time but also at initial time. Further, a family operator was introduced, which considers the combined influence of groups of chemical species. Since it is possible to choose between different

operators and, further, between different projections and families, a flexible algorithm is at disposition.

The adapted theory of targeted singular vectors was first applied to the RADM2-mechanism only. For a set of six different scenarios, the VOC versus NO_x limitation of the ozone formation was investigated. In this process, all introduced operators were applied. Upon investigation, the time-dependence of singular vectors was examined via differing initial and final simulation times. Results reveal, that the singular vectors are strongly dependent on start time and length of the simulation. For the absolute error growth, recurring patterns made a categorization dependent on initial and final time possible. Here, the importance of initial time is dependent on the time of the day and independent on the calender date itself. For relative errors, however, the particular point in time is crucial. Therefore, a categorization is not helpful. As expected, singular vectors with initial values as target variables tend to be more dependent on initial values, while emission factors as target variables are more dependent on simulation length. The particular importance of individual chemical compounds differs strongly between absolute and relative error growth. In both cases, results revealed a wide range of chemical sensitivities, indicating that species of low sensitivity are insignificant. Their measurement is negligible.

Optimal placements of measurements of chemical compounds are indicated by singular vectors of a full chemistry transport model. This application has been implemented into the EURAD-IM. The extended model was tested and evaluated by conducting a comprehensive set of case studies based on the accomplished ZEPTER-2 campaign. Particular questions were examined, which allow for verification of the results with data assimilation at a later date. The optimal placement of measurements of chemical compounds was compared to the optimal placement of measurements of transport and diffusion. Analyses demonstrate, that these placements differ significantly. Different chemical compounds showed a similar optimal placement of measurements of initial values. Occurring differences appear to be related to initial concentrations or time of simulation start. Optimal placement of measurements of emission factors is strongly dependent on the location of emission sources. For campaign planning, the application of singular vector analysis is advantageous, since the optimal placement of measurements is dependent on several processes and is not straightforward to predict. In terms of a relevance ranking of chemical species, the measurement priority of different species is location dependent. A general categorization in low, medium and high measurement relevance of species is nevertheless possible. Due to a wide range of sensitivity values, measurements of compounds of low sensitiv-

ity can be omitted. Both optimal placement of measurements and relevance ranking of chemical compounds confirm the benefit of singular vectors for measurement decision guidance with respect to chemical compounds.

Singular vector analysis therefore is a powerful tool, which identifies critical chemical species and chemical locations. It can be applied for effective campaign-planning. Further, the detected directions of largest error growth can be employed to initialize ensemble forecasts and to model covariances.

APPENDIX A

Specifications and results
for tropospheric chemistry scenarios

This chapter is designed to give further insight into the scenario study introduced in chapter 4. First, in section A.1 the RADM2 model species are introduced. Further, the applied update-mechanisms for photolysis frequencies and emissions are described. Detailed results of the conducted calculations are presented in section A.2.

A.1 Model specifications

A.1.1 RADM2 species

Table A.1: RADM2 species list according to Stockwell et al. [1990]. PR denotes peroxy radicals.

Species	Definition
	Stable Inorganic Compounds
Nitrogen	
NO	Nitric oxide
NO_2	Nitrogen dioxide

Continued next page

Table A.1 – Continued

HONO	Nitrous acid
NO_3	Nitrogen trioxide
N_2O_5	Nitrogen pentoxide
HNO_4	Pernitric acid
HNO3	Nitric acid
Oxidants	
O_3	Ozone
H_2O_2	Hydrogen peroxide
Sulfur	
SO_2	Sulfur dioxide
SULF	Sulfuric acid
Carbon oxides	
CO	Carbon monoxide
CO_2	Carbon dioxide (product only)
H_2	Hydrogen (product only)
	Inorganic Short-Lived Intermediates
Atomic species	
O^3P	Ground state oxygen atom
O^1D	Excited state oxygen atom
Odd hydrogen	
HO	Hydroxy radical
HO_2	Hydroperoxy radical
	Abundant Stable Species
O_2	Oxygen
N_2	Nitrogen
H_2O	Water
	Stable Organic Compounds
Alkanes	
CH_4	Methane
ETH	Ethane
HC3	Alkanes with HO rate constant (298 K, 1 atm) between 2.7×10^{-13} cm^{3}s^{-1} and 3.4×10^{-12} cm^{3}s^{-1}
HC5	Alkanes with HO rate constant (298 K, 1 atm) between 3.4×10^{-12} cm^{3}s^{-1} and 6.8×10^{-12} cm^{3}s^{-1}

Continued next page

Table A.1 – Continued

HC8	Alkanes with HO rate constant (298 K, 1 atm) greater than 6.8×10^{-12} cm^3s^{-1}
Alkenes	
OL2	Ethene
OLT	Terminal alkenes
OLI	Internal alkenes
ISO	Isoprene
Aromatics	
TOL	Toluene and less reactive aromatics
CSL	Cresol and other hydroxy substituted aromatics
XYL	Xylene and more reactive aromatics
Carbonyls	
HCHO	Formaldehyde
ALD	Acetaldehyde and higher aldehydes
KET	Ketones
GLY	Glyoxal
MGLY	Methylglyoxal
DCB	Unsaturated Dicarbonyl
Organic nitrogen	
PAN	Peroxyacetyl nitrate and higher PANs
TPAN	H(CO)CH=CHCO3NO2
ONIT	Organic nitrate
Organic peroxides	
OP1	Methyl hydrogen peroxide
OP2	Higher organic peroxides
PAA	Peroxyacetic acid
Organic acids	
ORA1	Formic acid
ORA2	Acetic acid and higher acids
	Organic Short-Lived Intermediates
PR from alkanes	
MO2	Methyl peroxy radical
ETHP	Peroxy radical formed from alkane, ETH
HC3P	Peroxy radical formed from alkane, HC3
HC5P	Peroxy radical formed from alkane, HC5

Continued next page

Table A.1 – Continued

HC8P	Peroxy radical formed from alkane, HC8
PR from alkenes	
OL2P	Peroxy radical formed from alkene, OL2
OLTP	Peroxy radical formed from alkene, OLT
OLIP	Peroxy radical formed from alkene, OLI
PR from aromatics	
TOLP	Peroxy radical formed from aromatics, TOL
XYLP	Peroxy radical formed from aromatics, XYL
PR with carbonyl groups	
ACO_3	Acetylperoxy radical
KETP	Peroxy radical formed from ketone, KET
TCO_3	$H(CO)CH=CHCO_3$
PR involving nitrogen	
OLN	NO3-alkene adduct
XNO_2	Accounts for additional organic nitrate formation affected by the lumped organic species
XO_2	Accounts for additional NO to NO_2 conversions affected by the lumped organic species

A.1.2   Photolysis frequencies

The scenario protocol prescribes photolysis frequencies based on the radiation transfer model by *Roeth* [1992]. The latter is adapted to the photolysis processes of RADM2. In conclusion, clear sky photolysis rates are calculated based on the empirical formula:

$$\text{PHUX(a,b,c)} = a \cdot \exp^{b(1-\sec(c\theta))}, \tag{A.1}$$

where θ is the solar zenith angle, a is the photolysis rate at an overhead sun ($\theta = 0$) and b and c are empirical coefficients. Concrete values for the considered scenarios are specified in Tables A.2 and A.3.

Table A.2: *Photolysis-Parameter according to equation (A.1) for scenarios LAND, MARINE, PLUME, URBAN, and URBAN/BIO.*

Reaction	a	b	c
$NO_2 \rightarrow O^3P + NO$	1.03^{-2}	0.96180	0.84671
$O_3 \rightarrow O^1D + O_2$	5.00^{-5}	3.29332	0.80782

Continued next page

Table A.2 – Continued

	a	b	c
$O_3 \rightarrow O^3P + O_2$	5.11^{-4}	0.37195	0.92289
$HONO \rightarrow HO + NO$	2.36^{-3}	1.06560	0.83644
$HNO_3 \rightarrow HO + NO_2$	8.07^{-7}	2.30845	0.81364
$HNO_4 \rightarrow HO_2 + NO_2$	4.88^{-6}	2.08052	0.81320
$NO_3 \rightarrow NO + O_2$	2.59^{-2}	0.29618	0.93748
$NO_3 \rightarrow NO_2 + O^3P$	2.30^{-1}	0.33518	0.93059
$H_2O_2 \rightarrow 2\ HO$	1.18^{-5}	1.65050	0.81606
$HCHO \rightarrow H_2 + CO$	5.12^{-5}	1.44263	0.81851
$HCHO \rightarrow 2\ HO_2 + CO$	4.51^{-5}	1.81238	0.81930
$ALD \rightarrow MO2 + HO_2 + CO$	7.49^{-6}	2.20021	0.81543
$OP1 \rightarrow HCHO + HO_2 + HO$	6.81^{-6}	1.60212	0.81688
$OP2 \rightarrow ALD + HO_2 + HO$	6.81^{-6}	1.60212	0.81688
$PAA \rightarrow MO2 + CO_2 + HO$	1.28^{-8}	7.94062	0.74435
$KET \rightarrow ACO_3 + ETHP$	6.46^{-7}	2.99467	0.80969
$GLY \rightarrow 0.13\ HCHO + 1.87$ CO	2.89^{-3}	0.57643	0.89043
$GLY \rightarrow 0.45\ HCHO + 1.55$ $CO + 0.80\ HO_2$	2.89^{-3}	0.57643	0.89043
$MGLY \rightarrow ACO_3 + HO_2 + CO$	3.15^{-3}	0.61557	0.88505
$DCB \rightarrow 0.98\ HO_2 + TCO_3 +$ $0.02\ ACO_3$	6.30^{-4}	1.27788	0.82502
$ONIT \rightarrow 0.20\ ALD + 0.80$ $KET + HO_2 + NO_2$	1.50^{-7}	7.85847	0.74473

Table A.3: *Photolysis-Parameter according to equation (A.1) for scenario FREE.*

Reaction	a	b	c
$NO_2 \rightarrow O^3P + NO$	1.35^{-02}	0.34016	0.93483
$O_3 \rightarrow O^1D + O_2$	6.64^{-05}	2.36935	0.86368
$O_3 \rightarrow O^3P + O_2$	5.68^{-04}	0.24033	0.91557
$HONO \rightarrow HO + NO$	3.16^{-03}	0.37063	0.93207
$HNO_3 \rightarrow HO + NO_2$	1.05^{-06}	1.33763	0.88651
$HNO_4 \rightarrow HO_2 + NO_2$	7.58^{-06}	1.09416	0.89213
$NO_3 \rightarrow NO + O_2$	2.73^{-02}	0.15360	0.93979
$NO_3 \rightarrow NO_2 + O^3P$	2.48^{-01}	0.15452	0.94442
$H_2O_2 \rightarrow 2\ HO$	1.75^{-05}	0.74993	0.90478
$HCHO \rightarrow H_2 + CO$	8.25^{-05}	0.57493	0.91546

Continued next page

Table A.3 – Continued

HCHO $\rightarrow$ 2 HO$_2$ + CO	6.83^{-05}	0.87898	0.90341
ALD $\rightarrow$ MO2 + HO$_2$ + CO	1.17^{-05}	1.19741	0.89116
OP1 $\rightarrow$ HCHO + HO$_2$ + HO	1.01^{-05}	0.72054	0.90563
OP2 $\rightarrow$ ALD + HO$_2$ + HO	1.01^{-05}	0.72054	0.90563
PAA $\rightarrow$ MO2 + CO$_2$ + HO	2.44^{-08}	6.67910	0.78023
KET $\rightarrow$ ACO$_3$ + ETHP	1.06^{-06}	1.87408	0.87448
GLY $\rightarrow$ 0.13 HCHO + 1.87 CO	3.42^{-03}	0.20636	0.94732
GLY $\rightarrow$ 0.45 HCHO + 1.55 CO + 0.80 HO$_2$	3.42^{-03}	0.20636	0.94732
MGLY $\rightarrow$ ACO3 + HO$_2$ + CO	3.76^{-03}	0.21624	0.94721
DCB $\rightarrow$ 0.98 HO$_2$ + TCO$_3$ + 0.02 ACO$_3$	8.85^{-04}	0.49137	0.92075
ONIT $\rightarrow$ 0.20 ALD + 0.80 KET + HO$_2$ + NO$_2$	2.83^{-07}	6.58637	0.78094

A.1.3   Emission rates

The utilized emission strengths in chapter 4 are given as volume production rate in cm^{-3}s^{-1}. For scenarios PLUME, URBAN and the urban part of scenario URBAN/BIO the emission data is specified in Table A.4.

Table A.4: *Emission strength as volume production rate in cm^{-3}s^{-1} for scenarios defined by Poppe et al. [2001].*

	PLUME	URBAN
NO	$1.1 \cdot 10^6$	$5.5 \cdot 10^6$
NO$_2$	0.0	0.0
CO	$2.4 \cdot 10^6$	$2.4 \cdot 10^6$
CH$_4$	0.0	0.0
SO$_2$	$2.2 \cdot 10^5$	$2.2 \cdot 10^5$
VOC	$3.0 \cdot 10^6$	$3.0 \cdot 10^6$

The total VOC-emission for all three scenarios is segregated into species and categories according to *Derwent and Jenkins* [1991]. The outcome is presented in Table A.5.

Table A.5: *Emitted VOC species, associated categories C and weighting factors W_f for scenarios PLUME, URBAN and URBAN/BIO.*

Specie	Formula	C	W_f
Ethane	C_2H_6	2	0.0340956
Propane	C_3H_8	3	0.00554352
n-Butane	$n\text{-}C_4H_{10}$	4	0.0800388
i-Butane	$i\text{-}C_4H_{10}$	4	0.0434109
2,2-Dimethylbutane	$CH_3CH_3C_4H_8$	4	0.00155535
n-Pentane	$n\text{-}C_5H_{12}$	5	0.0240418
i-Pentane	$i\text{-}C_5H_{12}$	5	0.0371555
n-Hexane	$n\text{-}C_6H_{14}$	5	0.0137236
2-Methylpentane	$CH_3C_5H_{11}$	5	0.0155535
3-Methylpentane	$CH_3C_5H_{11}$	5	0.0109789
2,3-Dimethylbutane	$CH_3CH_3C_4H_8$	5	0.00494051
n-Heptane	$n\text{-}C_7H_{16}$	6	0.00621589
2-Methylhexane	$CH_3C_6H_{13}$	6	0.00519302
3-Methylhexane	$CH_3C_6H_{13}$	6	0.00448488
n-Octane	$n\text{-}C_8H_{18}$	6	0.00400313
Methylheptanes	$CH_3C_7H_{15}$	6	0.0138039
n-Nonane	$n\text{-}C_9H_{20}$	6	0.00737645
Methyloctanes	$H_3C_8H_{17}$	6	0.00301205
n-Decane	$n\text{-}C_{10}H_{22}$	6	0.005541
Methylnonanes	$CH_3C_9H_{19}$	6	0.00354624
n-Undecane	$n\text{-}C_{11}H_{24}$	6	0.00706122
n-Duodecane	$n\text{-}C_{12}H_{26}$	7	0.00647971
Ethylene	C_2H_4	9	0.0646318
Propylene	C_3H_6	10	0.0155491
But-1-ene	C_4H_8	11	0.0073062
Pent-1-ene	C_5H_{10}	11	0.00314729
2-Methylbut-1-ene	$CH_3C_4H_7$	11	0.00179845
3-Methylbut-1-ene	$CH_3C_4H_7$	11	0.00213566
But-2-ene	C_4H_8	12	0.0134884
Pent-2-ene	C_5H_{10}	12	0.00809302
2-Methylbut-2-ene	$CH_3C_4H_7$	12	0.00393411
Butylene	C_4H_8	13	0.00210756
Benzene	C_6H_6	14	0.0171487
Toluene	$CH_3C_6H_5$	15	0.0624326

Continued next page

Table A.5 – Continued

Ethylbenzene	$C_2H_5C_6H_5$	15	0.010392
n-Propylbenzene	$n\text{-}C_3H_7C_6H_5$	15	0.00216376
i-Propylbenzene	$i\text{-}C_3H_7C_6H_5$	15	0.00111466
o-Xylene	$o\text{-}C_6H_4(CH_3)CH_3$	16	0.0126188
m-Xylene	$p\text{-}C_6H_4(CH_3)CH_3$	16	0.0163303
p-Xylene	$m\text{-}C_6H_4(CH_3)CH_3$	16	0.0163303
1,2,3-Trimethylbenzene	$CH_3CH_3CH_3C_6H_3$	16	0.00255717
1,2,4-Trimethylbenzene	$CH_3CH_3CH_3C_6H_3$	16	0.00983527
1,3,5-Trimethylbenzene	$CH_3CH_3CH_3C_6H_3$	16	0.00367183
o-Ethyltoluene	$o\text{-}CH_3C_6H_4C_2H_5$	16	0.00334399
m-Ethyltoluene	$CH_3C_6H_4CH_3$	16	0.00467639
p-Ethyltoluene	$p\text{-}CH_3C_6H_4C_2H_5$	16	0.00413081
Formaldehyde	$HCHO$	19	0.0196705
Acetaldeyde	CH_3CHO	20	0.00178823
Proprionaldehyde	C_2H_5CHO	20	0.00149225
Butyraldehyde	$n\text{-}C_3H_7CHO$	20	0.00079775
i-Butyraldehyde	$(CH_3)_2CHCHO$	20	0.000710325
Valeraldehyde	$n\text{-}C_4H_9CHO$	20	3.65964^{-05}
Benzaldehyde	C_6H_5CHO	20	0.000296914
Acetone	CH_3COCH_3	21	0.023062
Methyl-ethyl ketone	$CH_3COC_2H_5$	22	0.0316914
Methyl-isobutyl ketone	$CH_3COC_4H_9$	22	0.00778954
Acetylene	C_2H_2	24	0.0393411
Tetrachloroethylene	C_2Cl_4	25	0.00671677
Methanol	CH_3OH	27	0.0209
Methyl acetate	$CH_3CO_2CH_3$	27	0.00170124
Ethyl acetate	$C_2H_5CO_2CH_3$	27	0.00554352
Ethanol	C_2H_5OH	28	0.174469
i-Propyl acetate	$i\text{-}C_3H_7CH_3CO_2$	28	0.0021599
n-Butyl acetate	$n\text{-}C_4H_9CH_3CO_2$	29	0.00244186
i-Butyl acetate	$s\text{-}C_4H_9CH_3CO_2$	29	0.00305233

With these specifications the emission strength of each VOC model species can be calculated according to the following formulas:

$$E_{\text{ETH}} = E_{\text{VOC}}/\text{Air} \cdot (1.000 \cdot sW_f(\ 2)),$$
$$E_{\text{HC3}} = E_{\text{VOC}}/\text{Air} \cdot (0.519 \cdot sW_f(\ 3) + 0.964 \cdot sW_f(\ 4)$$

$$
\begin{aligned}
&\hphantom{E_{HC5}} + 0.343 \cdot sW_f(24) + 0.078 \cdot sW_f(25)\\
&\hphantom{E_{HC5}} + 0.404 \cdot sW_f(27) + 1.215 \cdot sW_f(28)),\\
E_{HC5} &= E_{VOC}/Air \cdot (0.956 \cdot sW_f(\ 5) + 1.075 \cdot sW_f(29)),\\
E_{HC8} &= E_{VOC}/Air \cdot (0.945 \cdot sW_f(\ 6) + 1.141 \cdot sW_f(\ 7),\\
&\hphantom{E_{HC5}} + 1.011 \cdot sW_f(30)),\\
E_{OL2} &= E_{VOC}/Air \cdot (1.000 \cdot sW_f(\ 9)),\\
E_{OLT} &= E_{VOC}/Air \cdot (1.000 \cdot sW_f(10) + 1.000 \cdot sW_f(11)\\
&\hphantom{E_{HC5}} + 0.500 \cdot sW_f(13) + 1.000 \cdot sW_f(18)),\\
E_{OLI} &= E_{VOC}/Air \cdot (1.000 \cdot sW_f(12) + 0.500 \cdot sW_f(13)),\\
E_{TOL} &= E_{VOC}/Air \cdot (0.293 \cdot sW_f(14) + 1.000 \cdot sW_f(15)\\
&\hphantom{E_{HC5}} + 1.000 \cdot sW_f(18)),\\
E_{XYL} &= E_{VOC}/Air \cdot (1.000 \cdot sW_f(16),\\
E_{CSL} &= E_{VOC}/Air \cdot (1.000 \cdot sW_f(17)),\\
E_{HCHO} &= E_{VOC}/Air \cdot (1.000 \cdot sW_f(19)),\\
E_{ALD} &= E_{VOC}/Air \cdot (1.000 \cdot sW_f(20)),\\
E_{KET} &= E_{VOC}/Air \cdot (0.253 \cdot sW_f(21) + 1.000 \cdot sW_f(22)).
\end{aligned}
$$

Here, E_s is the emission strength of species s, Air is the air density of the utilized scenario (see Table 4.1) and $sW_f(i)$ are the summed up weighting factors of category i.

For the biogenic part of the scenario URBAN/BIO, only isoprene is emitted with a volume source strength of $E_{ISO} = 4.0 \cdot 10^{-11} \cdot J(NO_2) \cdot Air$. Here, $J(NO_2)$ is the photolysis frequency of NO_2 (calculated according to Table A.2).

A.2   Results

A.2.1   Error growth of VOC and NO_x species

Table A.6: *Mean impact and standard deviation of VOC compounds for category $\mathcal{C}_a$ for scenarios LAND, MARINE, and FREE. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	LAND	MARINE	FREE
CSL	$(2.60 \pm 1.71)_{10}\text{-}\mathbf{1}$	$(1.60 \pm 0.94)_{10}\text{-}\mathbf{1}$	$(1.27 \pm 0.42)_{10}\text{-}\mathbf{1}$
TOL	$(1.45 \pm 0.98)_{10}\text{-}\mathbf{1}$	$(1.18 \pm 0.63)_{10}\text{-}\mathbf{1}$	$(3.72 \pm 1.73)_{10}\text{-}\mathbf{2}$

Continued next page

Table A.6 – Continued

XYL	$(1.11 \pm 0.42)_{10}\text{-}\mathbf{1}$	$(8.46 \pm 3.15)_{10}\text{-}\mathbf{2}$	$(6.53 \pm 2.41)_{10}\text{-}\mathbf{2}$
MGLY	$(0.89 \pm 1.13)_{10}\text{-}\mathbf{1}$	$(8.28 \pm 8.55)_{10}\text{-}\mathbf{2}$	$(1.91 \pm 0.56)_{10}\text{-}\mathbf{1}$
DCB	$(8.13 \pm 8.12)_{10}\text{-}\mathbf{2}$	$(8.54 \pm 6.14)_{10}\text{-}\mathbf{2}$	$(1.47 \pm 0.44)_{10}\text{-}\mathbf{1}$
ALD	$(5.20 \pm 4.41)_{10}\text{-}\mathbf{2}$	$(6.14 \pm 3.66)_{10}\text{-}\mathbf{2}$	$(1.02 \pm 0.45)_{10}\text{-}\mathbf{1}$
HC8	$(7.30 \pm 4.12)_{10}\text{-}\mathbf{2}$	$(9.87 \pm 4.69)_{10}\text{-}\mathbf{2}$	$(4.05 \pm 1.78)_{10}\text{-}\mathbf{2}$
OLI	$(4.19 \pm 1.93)_{10}\text{-}\mathbf{2}$	$(5.57 \pm 2.52)_{10}\text{-}\mathbf{2}$	$(6.32 \pm 1.99)_{10}\text{-}\mathbf{2}$
ISO	$(3.15 \pm 1.44)_{10}\text{-}\mathbf{2}$	$(4.53 \pm 2.03)_{10}\text{-}\mathbf{2}$	$(5.99 \pm 3.65)_{10}\text{-}\mathbf{2}$
OLT	$(3.11 \pm 1.47)_{10}\text{-}\mathbf{2}$	$(4.39 \pm 2.01)_{10}\text{-}\mathbf{2}$	$(5.47 \pm 2.90)_{10}\text{-}\mathbf{2}$
HC5	$(3.08 \pm 2.31)_{10}\text{-}\mathbf{2}$	$(5.91 \pm 3.82)_{10}\text{-}\mathbf{2}$	$(1.89 \pm 0.99)_{10}\text{-}\mathbf{2}$
HC3	$(1.58 \pm 1.26)_{10}\text{-}\mathbf{2}$	$(4.16 \pm 2.98)_{10}\text{-}\mathbf{2}$	$(7.69 \pm 3.69)_{10}\text{-}\mathbf{2}$
OL2	$(2.04 \pm 1.37)_{10}\text{-}\mathbf{2}$	$(2.48 \pm 1.57)_{10}\text{-}\mathbf{2}$	$(3.47 \pm 2.07)_{10}\text{-}\mathbf{2}$
KET	$(9.84 \pm 7.60)_{10}\text{-}\mathbf{3}$	$(3.00 \pm 2.28)_{10}\text{-}\mathbf{2}$	$(5.82 \pm 3.63)_{10}\text{-}\mathbf{3}$
HCHO	$(2.76 \pm 1.81)_{10}\text{-}\mathbf{3}$	$(3.93 \pm 3.56)_{10}\text{-}\mathbf{4}$	$(2.39 \pm 1.39)_{10}\text{-}\mathbf{2}$
GLY	$(2.62 \pm 2.36)_{10}\text{-}\mathbf{3}$	$(2.11 \pm 3.03)_{10}\text{-}\mathbf{4}$	$(2.06 \pm 1.26)_{10}\text{-}\mathbf{2}$
ETH	$(2.09 \pm 1.98)_{10}\text{-}\mathbf{3}$	$(8.79 \pm 8.21)_{10}\text{-}\mathbf{3}$	$(7.95 \pm 4.98)_{10}\text{-}\mathbf{4}$

Table A.7: *Mean impact and standard deviation of VOC compounds for category $\mathcal{C}_a$ for scenarios PLUME, URBAN, and BIO. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	PLUME	URBAN	BIO
CSL	$(2.37 \pm 0.82)_{10}\text{-}\mathbf{1}$	$(5.30 \pm 1.73)_{10}\text{-}\mathbf{2}$	$(4.96 \pm 0.45)_{10}\text{-}\mathbf{1}$
TOL	$(9.57 \pm 5.43)_{10}\text{-}\mathbf{2}$	$(6.72 \pm 1.28)_{10}\text{-}\mathbf{2}$	$(1.03 \pm 0.45)_{10}\text{-}\mathbf{1}$
XYL	$(8.26 \pm 2.96)_{10}\text{-}\mathbf{2}$	$(1.29 \pm 0.17)_{10}\text{-}\mathbf{1}$	$(9.31 \pm 1.75)_{10}\text{-}\mathbf{2}$
MGLY	$(9.93 \pm 6.03)_{10}\text{-}\mathbf{2}$	$(1.17 \pm 0.38)_{10}\text{-}\mathbf{1}$	$(4.74 \pm 3.91)_{10}\text{-}\mathbf{2}$
DCB	$(1.02 \pm 0.49)_{10}\text{-}\mathbf{1}$	$(1.20 \pm 0.31)_{10}\text{-}\mathbf{1}$	$(4.07 \pm 3.02)_{10}\text{-}\mathbf{2}$
ALD	$(7.36 \pm 3.35)_{10}\text{-}\mathbf{2}$	$(3.80 \pm 0.82)_{10}\text{-}\mathbf{2}$	$(3.73 \pm 1.29)_{10}\text{-}\mathbf{2}$
HC8	$(3.79 \pm 2.18)_{10}\text{-}\mathbf{2}$	$(6.17 \pm 2.79)_{10}\text{-}\mathbf{2}$	$(3.34 \pm 1.57)_{10}\text{-}\mathbf{2}$
OLI	$(6.99 \pm 2.76)_{10}\text{-}\mathbf{2}$	$(7.30 \pm 0.79)_{10}\text{-}\mathbf{2}$	$(5.06 \pm 1.43)_{10}\text{-}\mathbf{2}$

Continued next page

Table A.7 – Continued

ISO	$(4.48 \pm 3.07)_{10}$-**2**	$(7.67 \pm 0.79)_{10}$-**2**	$(2.45 \pm 0.82)_{10}$-**2**
OLT	$(4.26 \pm 2.53)_{10}$-**2**	$(6.87 \pm 0.64)_{10}$-**2**	$(2.53 \pm 0.74)_{10}$-**2**
HC5	$(1.60 \pm 0.94)_{10}$-**2**	$(3.96 \pm 2.06)_{10}$-**2**	$(1.31 \pm 0.74)_{10}$-**2**
HC3	$(2.31 \pm 1.78)_{10}$-**2**	$(2.73 \pm 1.54)_{10}$-**2**	$(8.93 \pm 6.26)_{10}$-**3**
OL2	$(2.36 \pm 1.98)_{10}$-**2**	$(4.06 \pm 0.69)_{10}$-**2**	$(1.57 \pm 0.69)_{10}$-**2**
KET	$(2.96 \pm 2.21)_{10}$-**2**	$(3.35 \pm 1.46)_{10}$-**2**	$(4.57 \pm 2.92)_{10}$-**3**
HCHO	$(7.40 \pm 5.30)_{10}$-**3**	$(2.27 \pm 0.71)_{10}$-**2**	$(3.54 \pm 1.79)_{10}$-**3**
GLY	$(8.12 \pm 7.46)_{10}$-**3**	$(2.55 \pm 0.93)_{10}$-**2**	$(2.22 \pm 1.28)_{10}$-**3**
ETH	$(6.44 \pm 6.18)_{10}$-**3**	$(6.59 \pm 5.94)_{10}$-**3**	$(1.06 \pm 0.85)_{10}$-**3**

Table A.8: *Mean impact and standard deviation of VOC compounds of categories $\mathcal{C}_{a_{1/2/3/4}}$ for scenario LAND. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{a_1}$	$\mathcal{C}_{a_2}$	$\mathcal{C}_{a_3}$	$\mathcal{C}_{a_4}$
CSL	$(1.35\pm0.4)_{10}$-**1**	$(2.85\pm1.7)_{10}$-**1**	$(2.91\pm1.7)_{10}$-**1**	$(2.73\pm1.8)_{10}$-**1**
TOL	$(1.10\pm0.7)_{10}$-**2**	$(9.78\pm4.4)_{10}$-**2**	$(1.69\pm0.6)_{10}$-**1**	$(2.32\pm0.8)_{10}$-**1**
XYL	$(3.66\pm2.0)_{10}$-**2**	$(1.15\pm0.3)_{10}$-**1**	$(1.34\pm0.2)_{10}$-**1**	$(1.29\pm0.2)_{10}$-**1**
MGLY	$(3.03\pm0.8)_{10}$-**1**	$(9.66\pm7.0)_{10}$-**2**	$(2.59\pm1.5)_{10}$-**2**	$(1.75\pm2.2)_{10}$-**2**
DCB	$(2.34\pm0.4)_{10}$-**1**	$(9.05\pm5.2)_{10}$-**2**	$(4.29\pm1.6)_{10}$-**2**	$(2.49\pm1.2)_{10}$-**2**
HC8	$(2.57\pm1.7)_{10}$-**2**	$(5.91\pm3.3)_{10}$-**2**	$(8.95\pm3.6)_{10}$-**2**	$(9.92\pm3.4)_{10}$-**2**
ALD	$(1.21\pm0.3)_{10}$-**1**	$(7.35\pm2.8)_{10}$-**2**	$(3.27\pm1.5)_{10}$-**2**	$(1.47\pm1.3)_{10}$-**2**
OLI	$(5.78\pm1.6)_{10}$-**2**	$(5.44\pm1.5)_{10}$-**2**	$(4.23\pm1.2)_{10}$-**2**	$(2.56\pm1.1)_{10}$-**2**
ISO	$(3.11\pm1.7)_{10}$-**2**	$(3.93\pm1.4)_{10}$-**2**	$(3.68\pm1.0)_{10}$-**2**	$(2.32\pm1.0)_{10}$-**2**
OLT	$(1.85\pm1.2)_{10}$-**2**	$(3.74\pm1.5)_{10}$-**2**	$(4.00\pm1.1)_{10}$-**2**	$(2.80\pm1.1)_{10}$-**2**
HC5	$(5.56\pm3.4)_{10}$-**3**	$(1.79\pm1.2)_{10}$-**2**	$(3.60\pm1.8)_{10}$-**2**	$(5.06\pm2.1)_{10}$-**2**
OL2	$(5.19\pm2.9)_{10}$-**3**	$(1.36\pm1.0)_{10}$-**2**	$(2.62\pm1.2)_{10}$-**2**	$(3.03\pm1.1)_{10}$-**2**
HC3	$(2.56\pm1.7)_{10}$-**3**	$(8.53\pm5.9)_{10}$-**3**	$(1.76\pm0.9)_{10}$-**2**	$(2.68\pm1.2)_{10}$-**2**
KET	$(2.62\pm1.7)_{10}$-**3**	$(5.56\pm3.3)_{10}$-**3**	$(1.03\pm0.5)_{10}$-**2**	$(1.65\pm0.8)_{10}$-**2**

Continued next page

Table A.8 – Continued

HCHO	$(5.32\pm2.6)_{10}$-**3**	$(2.40\pm1.3)_{10}$-**3**	$(1.89\pm0.5)_{10}$-**3**	$(2.32\pm0.8)_{10}$-**3**
GLY	$(5.72\pm4.0)_{10}$-**3**	$(2.12\pm1.2)_{10}$-**3**	$(1.69\pm0.7)_{10}$-**3**	$(2.09\pm1.0)_{10}$-**3**
ETH	$(2.40\pm1.2)_{10}$-**4**	$(8.57\pm6.7)_{10}$-**4**	$(2.14\pm1.2)_{10}$-**3**	$(3.92\pm2.0)_{10}$-**3**

Table A.9: *Mean impact and standard deviation of VOC compounds for categories $\mathcal{C}_{a_{1/2/3/4}}$ for scenario MARINE. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{a_1}$	$\mathcal{C}_{a_2}$	$\mathcal{C}_{a_3}$	$\mathcal{C}_{a_4}$
CSL	$(1.53\pm0.1)_{10}$-**1**	$(1.79\pm0.7)_{10}$-**1**	$(1.71\pm0.8)_{10}$-**1**	$(1.37\pm1.3)_{10}$-**1**
TOL	$(2.41\pm1.3)_{10}$-**2**	$(8.02\pm2.6)_{10}$-**2**	$(1.41\pm0.3)_{10}$-**1**	$(1.81\pm0.3)_{10}$-**1**
HC8	$(3.35\pm2.0)_{10}$-**2**	$(8.26\pm3.2)_{10}$-**2**	$(1.33\pm0.3)_{10}$-**1**	$(1.30\pm0.3)_{10}$-**1**
DCB	$(1.90\pm0.5)_{10}$-**1**	$(9.95\pm3.5)_{10}$-**2**	$(5.69\pm1.4)_{10}$-**2**	$(3.61\pm2.3)_{10}$-**2**
XYL	$(7.31\pm2.3)_{10}$-**2**	$(1.04\pm0.1)_{10}$-**1**	$(1.01\pm0.2)_{10}$-**1**	$(6.67\pm3.8)_{10}$-**2**
MGLY	$(2.27\pm0.9)_{10}$-**1**	$(8.35\pm5.4)_{10}$-**2**	$(1.92\pm1.6)_{10}$-**2**	$(4.94\pm3.1)_{10}$-**2**
ALD	$(1.14\pm0.1)_{10}$-**1**	$(7.85\pm2.0)_{10}$-**2**	$(3.14\pm1.9)_{10}$-**2**	$(3.85\pm2.7)_{10}$-**2**
HC5	$(9.22\pm6.7)_{10}$-**3**	$(3.34\pm1.8)_{10}$-**2**	$(7.27\pm2.6)_{10}$-**2**	$(9.85\pm2.1)_{10}$-**2**
OLI	$(7.31\pm2.2)_{10}$-**2**	$(7.75\pm0.9)_{10}$-**2**	$(5.23\pm1.4)_{10}$-**2**	$(2.97\pm1.8)_{10}$-**2**
ISO	$(4.82\pm2.2)_{10}$-**2**	$(6.10\pm1.1)_{10}$-**2**	$(4.40\pm1.2)_{10}$-**2**	$(2.94\pm2.0)_{10}$-**2**
OLT	$(3.59\pm2.0)_{10}$-**2**	$(6.09\pm1.2)_{10}$-**2**	$(5.16\pm1.0)_{10}$-**2**	$(2.79\pm1.6)_{10}$-**2**
HC3	$(5.55\pm4.1)_{10}$-**3**	$(2.02\pm1.1)_{10}$-**2**	$(4.68\pm1.8)_{10}$-**2**	$(7.41\pm1.9)_{10}$-**2**
KET	$(5.60\pm2.7)_{10}$-**3**	$(1.34\pm0.6)_{10}$-**2**	$(3.03\pm1.3)_{10}$-**2**	$(5.54\pm1.7)_{10}$-**2**
OL2	$(6.65\pm4.0)_{10}$-**3**	$(2.31\pm1.2)_{10}$-**2**	$(3.99\pm0.9)_{10}$-**2**	$(2.79\pm1.4)_{10}$-**2**
ETH	$(6.18\pm4.9)_{10}$-**4**	$(2.76\pm1.6)_{10}$-**3**	$(8.11\pm4.0)_{10}$-**3**	$(1.79\pm0.7)_{10}$-**2**
HCHO	$(2.70\pm2.6)_{10}$-**4**	$(2.71\pm2.3)_{10}$-**4**	$(3.34\pm2.5)_{10}$-**4**	$(5.47\pm3.9)_{10}$-**4**
GLY	$(3.66\pm6.1)_{10}$-**4**	$(1.01\pm1.4)_{10}$-**4**	$(1.26\pm1.0)_{10}$-**4**	$(2.68\pm1.8)_{10}$-**4**

Table A.10: *Mean impact and standard deviation of VOC compounds for categories $\mathcal{C}_{a_{1/2/3/4}}$ for scenario FREE. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{a_1}$	$\mathcal{C}_{a_2}$	$\mathcal{C}_{a_3}$	$\mathcal{C}_{a_4}$
MGLY	$(1.58\pm0.4)_{10}\text{-}\mathbf{1}$	$(2.51\pm0.5)_{10}\text{-}\mathbf{1}$	$(2.12\pm0.3)_{10}\text{-}\mathbf{1}$	$(1.52\pm0.2)_{10}\text{-}\mathbf{1}$
DCB	$(8.80\pm3.3)_{10}\text{-}\mathbf{2}$	$(1.74\pm0.5)_{10}\text{-}\mathbf{1}$	$(1.72\pm0.2)_{10}\text{-}\mathbf{1}$	$(1.39\pm0.1)_{10}\text{-}\mathbf{1}$
CSL	$(1.19\pm0.3)_{10}\text{-}\mathbf{1}$	$(8.43\pm3.7)_{10}\text{-}\mathbf{2}$	$(1.43\pm0.3)_{10}\text{-}\mathbf{1}$	$(1.51\pm0.2)_{10}\text{-}\mathbf{1}$
ALD	$(2.71\pm1.6)_{10}\text{-}\mathbf{2}$	$(1.13\pm5.0)_{10}\text{-}\mathbf{1}$	$(1.33\pm0.1)_{10}\text{-}\mathbf{1}$	$(1.10\pm0.1)_{10}\text{-}\mathbf{1}$
XYL	$(7.86\pm0.9)_{10}\text{-}\mathbf{2}$	$(3.63\pm2.4)_{10}\text{-}\mathbf{2}$	$(6.30\pm1.8)_{10}\text{-}\mathbf{2}$	$(8.14\pm0.4)_{10}\text{-}\mathbf{2}$
OLI	$(7.03\pm2.2)_{10}\text{-}\mathbf{2}$	$(4.02\pm1.9)_{10}\text{-}\mathbf{2}$	$(6.61\pm1.2)_{10}\text{-}\mathbf{2}$	$(7.49\pm0.1)_{10}\text{-}\mathbf{2}$
ISO	$(1.16\pm0.3)_{10}\text{-}\mathbf{1}$	$(4.12\pm3.7)_{10}\text{-}\mathbf{2}$	$(4.09\pm1.6)_{10}\text{-}\mathbf{2}$	$(5.88\pm0.5)_{10}\text{-}\mathbf{2}$
OLT	$(9.70\pm1.1)_{10}\text{-}\mathbf{2}$	$(4.34\pm3.6)_{10}\text{-}\mathbf{2}$	$(3.59\pm1.5)_{10}\text{-}\mathbf{2}$	$(5.45\pm0.6)_{10}\text{-}\mathbf{2}$
HC8	$(3.93\pm0.6)_{10}\text{-}\mathbf{2}$	$(2.91\pm1.8)_{10}\text{-}\mathbf{2}$	$(3.15\pm1.5)_{10}\text{-}\mathbf{2}$	$(5.51\pm1.1)_{10}\text{-}\mathbf{2}$
TOL	$(3.99\pm0.9)_{10}\text{-}\mathbf{2}$	$(2.16\pm1.5)_{10}\text{-}\mathbf{2}$	$(2.96\pm1.4)_{10}\text{-}\mathbf{2}$	$(5.20\pm0.7)_{10}\text{-}\mathbf{2}$
OL2	$(5.43\pm0.9)_{10}\text{-}\mathbf{2}$	$(5.39\pm1.6)_{10}\text{-}\mathbf{2}$	$(1.70\pm1.5)_{10}\text{-}\mathbf{2}$	$(2.20\pm0.8)_{10}\text{-}\mathbf{2}$
HCHO	$(3.97\pm0.3)_{10}\text{-}\mathbf{2}$	$(3.74\pm0.5)_{10}\text{-}\mathbf{2}$	$(1.93\pm0.7)_{10}\text{-}\mathbf{2}$	$(9.32\pm2.5)_{10}\text{-}\mathbf{3}$
GLY	$(3.77\pm0.6)_{10}\text{-}\mathbf{2}$	$(3.09\pm0.4)_{10}\text{-}\mathbf{2}$	$(1.60\pm0.6)_{10}\text{-}\mathbf{2}$	$(7.84\pm2.0)_{10}\text{-}\mathbf{3}$
HC5	$(2.26\pm0.5)_{10}\text{-}\mathbf{2}$	$(2.88\pm0.6)_{10}\text{-}\mathbf{2}$	$(1.04\pm0.9)_{10}\text{-}\mathbf{2}$	$(1.46\pm0.7)_{10}\text{-}\mathbf{2}$
HC3	$(8.45\pm2.2)_{10}\text{-}\mathbf{3}$	$(1.08\pm0.2)_{10}\text{-}\mathbf{2}$	$(4.16\pm3.4)_{10}\text{-}\mathbf{3}$	$(6.97\pm3.0)_{10}\text{-}\mathbf{3}$
KET	$(3.44\pm0.6)_{10}\text{-}\mathbf{3}$	$(2.46\pm1.3)_{10}\text{-}\mathbf{3}$	$(5.27\pm2.6)_{10}\text{-}\mathbf{3}$	$(9.69\pm2.1)_{10}\text{-}\mathbf{3}$
ETH	$(8.84\pm2.8)_{10}\text{-}\mathbf{4}$	$(1.40\pm0.2)_{10}\text{-}\mathbf{3}$	$(6.07\pm5.0)_{10}\text{-}\mathbf{4}$	$(3.84\pm2.1)_{10}\text{-}\mathbf{4}$

Table A.11: *Mean impact and standard deviation of VOC compounds for categories $\mathcal{C}_{a_{1/2/3/4}}$ for scenario PLUME. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{a_1}$	$\mathcal{C}_{a_2}$	$\mathcal{C}_{a_3}$	$\mathcal{C}_{a_4}$
CSL	$(2.11\pm0.3)_{10}\text{-}\mathbf{1}$	$(2.18\pm0.3)_{10}\text{-}\mathbf{1}$	$(2.50\pm0.5)_{10}\text{-}\mathbf{1}$	$(2.64\pm1.2)_{10}\text{-}\mathbf{1}$
DCB	$(1.69\pm0.3)_{10}\text{-}\mathbf{1}$	$(1.21\pm0.2)_{10}\text{-}\mathbf{1}$	$(8.40\pm2.7)_{10}\text{-}\mathbf{2}$	$(6.37\pm3.8)_{10}\text{-}\mathbf{2}$
MGLY	$(1.91\pm0.5)_{10}\text{-}\mathbf{1}$	$(1.19\pm0.3)_{10}\text{-}\mathbf{1}$	$(7.34\pm2.7)_{10}\text{-}\mathbf{2}$	$(5.57\pm3.9)_{10}\text{-}\mathbf{2}$

Continued next page

Table A.11 – Continued

TOL	$(2.63\pm1.6)_{10}$-**2**	$(7.42\pm2.4)_{10}$-**2**	$(1.24\pm0.2)_{10}$-**1**	$(1.32\pm0.6)_{10}$-**1**
XYL	$(7.70\pm2.8)_{10}$-**2**	$(1.00\pm0.1)_{10}$-**1**	$(8.89\pm2.2)_{10}$-**2**	$(6.74\pm3.3)_{10}$-**2**
ALD	$(1.09\pm0.2)_{10}$-**1**	$(9.85\pm1.2)_{10}$-**2**	$(6.56\pm1.9)_{10}$-**2**	$(4.01\pm2.5)_{10}$-**2**
OLI	$(8.20\pm2.2)_{10}$-**2**	$(9.05\pm0.7)_{10}$-**2**	$(6.63\pm2.1)_{10}$-**2**	$(4.98\pm2.8)_{10}$-**2**
ISO	$(3.12\pm2.0)_{10}$-**2**	$(4.02\pm1.1)_{10}$-**2**	$(3.42\pm1.7)_{10}$-**2**	$(5.76\pm4.3)_{10}$-**2**
OLT	$(2.44\pm1.2)_{10}$-**2**	$(4.68\pm1.4)_{10}$-**2**	$(3.95\pm1.5)_{10}$-**2**	$(4.72\pm3.5)_{10}$-**2**
HC8	$(1.53\pm1.0)_{10}$-**2**	$(3.03\pm1.7)_{10}$-**2**	$(5.67\pm1.5)_{10}$-**2**	$(4.43\pm2.0)_{10}$-**2**
KET	$(5.24\pm3.4)_{10}$-**3**	$(1.49\pm0.7)_{10}$-**2**	$(3.35\pm1.2)_{10}$-**2**	$(5.11\pm1.8)_{10}$-**2**
OL2	$(2.82\pm0.5)_{10}$-**2**	$(1.65\pm1.0)_{10}$-**2**	$(1.65\pm2.1)_{10}$-**2**	$(3.05\pm2.6)_{10}$-**2**
HC3	$(3.22\pm1.9)_{10}$-**3**	$(1.11\pm0.7)_{10}$-**2**	$(2.92\pm0.9)_{10}$-**2**	$(3.89\pm1.6)_{10}$-**2**
HC5	$(1.24\pm0.4)_{10}$-**2**	$(8.20\pm4.9)_{10}$-**3**	$(1.85\pm0.8)_{10}$-**2**	$(2.27\pm1.0)_{10}$-**2**
GLY	$(4.99\pm5.6)_{10}$-**3**	$(3.46\pm3.0)_{10}$-**3**	$(7.09\pm7.3)_{10}$-**3**	$(1.39\pm0.7)_{10}$-**2**
HCHO	$(9.70\pm2.9)_{10}$-**3**	$(5.98\pm4.1)_{10}$-**3**	$(6.22\pm6.4)_{10}$-**3**	$(8.34\pm6.0)_{10}$-**3**
ETH	$(4.35\pm2.5)_{10}$-**4**	$(1.87\pm1.2)_{10}$-**3**	$(6.25\pm2.7)_{10}$-**3**	$(1.27\pm0.5)_{10}$-**2**

Table A.12: *Mean impact and standard deviation of VOC compounds for categories $\mathcal{C}_{a_{1/2/3/4}}$ for scenario URBAN. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{a_1}$	$\mathcal{C}_{a_2}$	$\mathcal{C}_{a_3}$	$\mathcal{C}_{a_4}$
XYL	$(1.44\pm0.1)_{10}$-**1**	$(1.37\pm0.1)_{10}$-**1**	$(1.25\pm0.1)_{10}$-**1**	$(1.17\pm0.1)_{10}$-**1**
DCB	$(1.59\pm0.3)_{10}$-**1**	$(1.29\pm0.2)_{10}$-**1**	$(1.11\pm0.2)_{10}$-**1**	$(1.01\pm0.2)_{10}$-**1**
MGLY	$(1.66\pm0.4)_{10}$-**1**	$(1.26\pm0.3)_{10}$-**1**	$(1.05\pm0.2)_{10}$-**1**	$(9.37\pm2.6)_{10}$-**2**
ISO	$(8.48\pm0.6)_{10}$-**2**	$(8.02\pm0.5)_{10}$-**2**	$(7.52\pm0.6)_{10}$-**2**	$(7.13\pm0.7)_{10}$-**2**
OLI	$(7.09\pm0.8)_{10}$-**2**	$(7.64\pm0.7)_{10}$-**2**	$(7.42\pm0.7)_{10}$-**2**	$(7.08\pm0.8)_{10}$-**2**
OLT	$(6.66\pm0.8)_{10}$-**2**	$(7.20\pm0.5)_{10}$-**2**	$(6.99\pm0.5)_{10}$-**2**	$(6.66\pm0.6)_{10}$-**2**
TOL	$(4.74\pm1.1)_{10}$-**2**	$(6.53\pm1.0)_{10}$-**2**	$(7.31\pm0.6)_{10}$-**2**	$(7.46\pm0.5)_{10}$-**2**
HC8	$(3.63\pm1.8)_{10}$-**2**	$(5.07\pm2.1)_{10}$-**2**	$(6.65\pm2.4)_{10}$-**2**	$(7.88\pm2.6)_{10}$-**2**
CSL	$(3.87\pm2.3)_{10}$-**2**	$(4.75\pm1.2)_{10}$-**2**	$(5.66\pm1.2)_{10}$-**2**	$(6.18\pm1.4)_{10}$-**2**

Continued next page

Table A.12 – Continued

OL2	$(4.02\pm1.4)_{10}\text{-}\mathbf{2}$	$(4.18\pm0.6)_{10}\text{-}\mathbf{2}$	$(4.13\pm0.2)_{10}\text{-}\mathbf{2}$	$(3.97\pm0.3)_{10}\text{-}\mathbf{2}$
HC5	$(2.04\pm1.1)_{10}\text{-}\mathbf{2}$	$(3.14\pm1.5)_{10}\text{-}\mathbf{2}$	$(4.31\pm1.8)_{10}\text{-}\mathbf{2}$	$(5.24\pm1.9)_{10}\text{-}\mathbf{2}$
ALD	$(2.64\pm1.1)_{10}\text{-}\mathbf{2}$	$(4.01\pm0.7)_{10}\text{-}\mathbf{2}$	$(4.15\pm0.4)_{10}\text{-}\mathbf{2}$	$(3.99\pm0.4)_{10}\text{-}\mathbf{2}$
KET	$(1.65\pm0.6)_{10}\text{-}\mathbf{2}$	$(2.67\pm0.9)_{10}\text{-}\mathbf{2}$	$(3.68\pm1.0)_{10}\text{-}\mathbf{2}$	$(4.45\pm1.2)_{10}\text{-}\mathbf{2}$
HC3	$(1.09\pm0.6)_{10}\text{-}\mathbf{2}$	$(2.09\pm1.0)_{10}\text{-}\mathbf{2}$	$(3.07\pm1.3)_{10}\text{-}\mathbf{2}$	$(3.78\pm1.4)_{10}\text{-}\mathbf{2}$
GLY	$(3.90\pm1.0)_{10}\text{-}\mathbf{2}$	$(2.68\pm0.6)_{10}\text{-}\mathbf{2}$	$(2.22\pm0.5)_{10}\text{-}\mathbf{2}$	$(2.02\pm0.5)_{10}\text{-}\mathbf{2}$
HCHO	$(3.12\pm0.3)_{10}\text{-}\mathbf{2}$	$(2.41\pm0.6)_{10}\text{-}\mathbf{2}$	$(2.06\pm0.6)_{10}\text{-}\mathbf{2}$	$(1.88\pm0.6)_{10}\text{-}\mathbf{2}$
ETH	$(1.51\pm0.9)_{10}\text{-}\mathbf{3}$	$(3.65\pm2.4)_{10}\text{-}\mathbf{3}$	$(6.85\pm4.3)_{10}\text{-}\mathbf{3}$	$(1.09\pm0.7)_{10}\text{-}\mathbf{2}$

Table A.13: *Mean impact and standard deviation of VOC compounds for categories $\mathcal{C}_{a_{1/2/3/4}}$ for scenario BIO. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{a_1}$	$\mathcal{C}_{a_2}$	$\mathcal{C}_{a_3}$	$\mathcal{C}_{a_4}$
CSL	$(4.89\pm0.6)_{10}\text{-}\mathbf{1}$	$(5.25\pm0.3)_{10}\text{-}\mathbf{1}$	$(5.03\pm0.4)_{10}\text{-}\mathbf{1}$	$(4.75\pm0.4)_{10}\text{-}\mathbf{1}$
TOL	$(3.38\pm1.7)_{10}\text{-}\mathbf{2}$	$(7.86\pm2.3)_{10}\text{-}\mathbf{2}$	$(1.18\pm0.1)_{10}\text{-}\mathbf{1}$	$(1.45\pm0.1)_{10}\text{-}\mathbf{1}$
XYL	$(6.78\pm2.5)_{10}\text{-}\mathbf{2}$	$(9.53\pm1.3)_{10}\text{-}\mathbf{2}$	$(1.02\pm0.1)_{10}\text{-}\mathbf{1}$	$(9.90\pm0.4)_{10}\text{-}\mathbf{2}$
OLI	$(7.16\pm0.9)_{10}\text{-}\mathbf{2}$	$(5.81\pm0.9)_{10}\text{-}\mathbf{2}$	$(4.49\pm0.5)_{10}\text{-}\mathbf{2}$	$(3.79\pm0.4)_{10}\text{-}\mathbf{2}$
MGLY	$(1.09\pm0.4)_{10}\text{-}\mathbf{1}$	$(5.63\pm2.3)_{10}\text{-}\mathbf{2}$	$(3.08\pm0.9)_{10}\text{-}\mathbf{2}$	$(2.00\pm0.7)_{10}\text{-}\mathbf{2}$
DCB	$(8.99\pm3.0)_{10}\text{-}\mathbf{2}$	$(4.52\pm1.7)_{10}\text{-}\mathbf{2}$	$(2.76\pm1.1)_{10}\text{-}\mathbf{2}$	$(2.07\pm1.0)_{10}\text{-}\mathbf{2}$
ALD	$(5.41\pm0.9)_{10}\text{-}\mathbf{2}$	$(4.57\pm0.8)_{10}\text{-}\mathbf{2}$	$(3.30\pm0.5)_{10}\text{-}\mathbf{2}$	$(2.56\pm0.4)_{10}\text{-}\mathbf{2}$
HC8	$(1.62\pm1.0)_{10}\text{-}\mathbf{2}$	$(2.04\pm1.0)_{10}\text{-}\mathbf{2}$	$(3.87\pm0.8)_{10}\text{-}\mathbf{2}$	$(4.84\pm0.3)_{10}\text{-}\mathbf{2}$
OLT	$(1.41\pm0.7)_{10}\text{-}\mathbf{2}$	$(2.76\pm0.6)_{10}\text{-}\mathbf{2}$	$(2.92\pm0.3)_{10}\text{-}\mathbf{2}$	$(2.71\pm0.4)_{10}\text{-}\mathbf{2}$
ISO	$(2.05\pm1.6)_{10}\text{-}\mathbf{2}$	$(2.32\pm0.6)_{10}\text{-}\mathbf{2}$	$(2.67\pm0.4)_{10}\text{-}\mathbf{2}$	$(2.60\pm0.4)_{10}\text{-}\mathbf{2}$
OL2	$(1.13\pm0.5)_{10}\text{-}\mathbf{2}$	$(8.70\pm4.3)_{10}\text{-}\mathbf{3}$	$(1.70\pm0.4)_{10}\text{-}\mathbf{2}$	$(2.19\pm0.3)_{10}\text{-}\mathbf{2}$
HC5	$(1.24\pm0.3)_{10}\text{-}\mathbf{2}$	$(5.85\pm4.4)_{10}\text{-}\mathbf{3}$	$(1.01\pm0.5)_{10}\text{-}\mathbf{2}$	$(2.03\pm0.4)_{10}\text{-}\mathbf{2}$
HC3	$(3.95\pm1.5)_{10}\text{-}\mathbf{3}$	$(3.06\pm1.7)_{10}\text{-}\mathbf{3}$	$(8.42\pm3.0)_{10}\text{-}\mathbf{3}$	$(1.58\pm0.3)_{10}\text{-}\mathbf{2}$
KET	$(1.21\pm0.8)_{10}\text{-}\mathbf{3}$	$(2.48\pm1.0)_{10}\text{-}\mathbf{3}$	$(4.61\pm1.2)_{10}\text{-}\mathbf{3}$	$(7.65\pm1.5)_{10}\text{-}\mathbf{3}$
HCHO	$(2.24\pm1.4)_{10}\text{-}\mathbf{3}$	$(2.85\pm1.7)_{10}\text{-}\mathbf{3}$	$(3.78\pm1.5)_{10}\text{-}\mathbf{3}$	$(4.43\pm1.5)_{10}\text{-}\mathbf{3}$

Continued next page

Table A.13 – Continued

GLY	$(2.22\pm1.9)_{10}$-**3**	$(1.53\pm0.9)_{10}$-**3**	$(1.92\pm0.9)_{10}$-**3**	$(2.82\pm0.8)_{10}$-**3**
ETH	$(6.63\pm1.6)_{10}$-**4**	$(3.19\pm2.3)_{10}$-**4**	$(7.43\pm3.9)_{10}$-**4**	$(1.96\pm0.6)_{10}$-**3**

Table A.14: *Mean impact and standard deviation of VOC compounds for category $\mathcal{C}_b$ for scenarios LAND, MARINE, and FREE. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	LAND	MARINE	FREE
CSL	$(7.77 \pm 0.41)_{10}$-**1**	$(6.94 \pm 0.64)_{10}$-**1**	$(1.93 \pm 0.56)_{10}$-**1**
MGLY	$(2.41 \pm 2.62)_{10}$-**2**	$(3.19 \pm 2.93)_{10}$-**2**	$(3.58 \pm 0.55)_{10}$-**1**
DCB	$(1.81 \pm 1.49)_{10}$-**2**	$(3.10 \pm 1.92)_{10}$-**2**	$(1.66 \pm 0.18)_{10}$-**1**
OLI	$(4.37 \pm 1.31)_{10}$-**2**	$(3.10 \pm 1.86)_{10}$-**2**	$(5.64 \pm 0.83)_{10}$-**2**
ISO	$(3.79 \pm 1.62)_{10}$-**2**	$(3.05 \pm 1.74)_{10}$-**2**	$(2.62 \pm 1.93)_{10}$-**2**
HC8	$(1.71 \pm 0.97)_{10}$-**2**	$(3.23 \pm 1.36)_{10}$-**2**	$(1.82 \pm 1.20)_{10}$-**2**
TOL	$(2.67 \pm 2.11)_{10}$-**2**	$(3.05 \pm 1.31)_{10}$-**2**	$(1.60 \pm 1.04)_{10}$-**2**
XYL	$(1.33 \pm 0.67)_{10}$-**2**	$(2.33 \pm 1.03)_{10}$-**2**	$(2.73 \pm 1.43)_{10}$-**2**
ALD	$(9.97 \pm 7.44)_{10}$-**3**	$(1.97 \pm 1.35)_{10}$-**2**	$(5.10 \pm 1.84)_{10}$-**2**
OLT	$(9.75 \pm 3.31)_{10}$-**3**	$(1.83 \pm 0.90)_{10}$-**2**	$(2.12 \pm 1.04)_{10}$-**2**
HC5	$(7.32 \pm 5.33)_{10}$-**3**	$(1.84 \pm 0.90)_{10}$-**2**	$(7.25 \pm 3.52)_{10}$-**3**
OL2	$(4.79 \pm 3.24)_{10}$-**3**	$(8.42 \pm 5.02)_{10}$-**3**	$(1.25 \pm 0.60)_{10}$-**2**
HC3	$(3.74 \pm 2.88)_{10}$-**3**	$(1.26 \pm 0.65)_{10}$-**2**	$(3.08 \pm 1.70)_{10}$-**3**
GLY	$(3.20 \pm 1.29)_{10}$-**3**	$(5.54 \pm 1.62)_{10}$-**3**	$(3.05 \pm 1.62)_{10}$-**2**
KET	$(2.33 \pm 1.75)_{10}$-**3**	$(8.98 \pm 4.78)_{10}$-**3**	$(2.89 \pm 2.40)_{10}$-**3**
HCHO	$(7.67 \pm 2.52)_{10}$-**4**	$(9.06 \pm 5.25)_{10}$-**4**	$(9.71 \pm 3.85)_{10}$-**3**
ETH	$(4.98 \pm 4.56)_{10}$-**4**	$(2.54 \pm 1.78)_{10}$-**3**	$(2.82 \pm 1.50)_{10}$-**4**

Table A.15: *Mean impact and standard deviation of VOC compounds for category $\mathcal{C}_b$ for scenarios PLUME, URBAN, and BIO. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	PLUME	URBAN	BIO
CSL	$(5.24 \pm 1.05)_{10}\text{-}\mathbf{1}$	$(1.08 \pm 0.21)_{10}\text{-}\mathbf{1}$	$(4.11 \pm 1.58)_{10}\text{-}\mathbf{1}$
MGLY	$(6.98 \pm 3.97)_{10}\text{-}\mathbf{2}$	$(1.02 \pm 0.33)_{10}\text{-}\mathbf{1}$	$(3.97 \pm 3.29)_{10}\text{-}\mathbf{2}$
DCB	$(5.89 \pm 3.31)_{10}\text{-}\mathbf{2}$	$(1.19 \pm 0.34)_{10}\text{-}\mathbf{1}$	$(2.93 \pm 2.54)_{10}\text{-}\mathbf{2}$
OLI	$(4.18 \pm 2.78)_{10}\text{-}\mathbf{2}$	$(6.18 \pm 0.81)_{10}\text{-}\mathbf{2}$	$(8.56 \pm 5.39)_{10}\text{-}\mathbf{2}$
ISO	$(4.55 \pm 3.62)_{10}\text{-}\mathbf{2}$	$(6.23 \pm 0.79)_{10}\text{-}\mathbf{2}$	$(1.06 \pm 0.60)_{10}\text{-}\mathbf{1}$
HC8	$(1.58 \pm 0.98)_{10}\text{-}\mathbf{2}$	$(5.97 \pm 2.59)_{10}\text{-}\mathbf{2}$	$(2.57 \pm 1.35)_{10}\text{-}\mathbf{2}$
TOL	$(4.68 \pm 2.24)_{10}\text{-}\mathbf{2}$	$(6.55 \pm 1.11)_{10}\text{-}\mathbf{2}$	$(8.55 \pm 3.35)_{10}\text{-}\mathbf{2}$
XYL	$(4.76 \pm 2.45)_{10}\text{-}\mathbf{2}$	$(1.28 \pm 0.15)_{10}\text{-}\mathbf{1}$	$(9.32 \pm 2.43)_{10}\text{-}\mathbf{2}$
ALD	$(4.56 \pm 2.32)_{10}\text{-}\mathbf{2}$	$(5.23 \pm 1.07)_{10}\text{-}\mathbf{2}$	$(3.23 \pm 1.26)_{10}\text{-}\mathbf{2}$
OLT	$(2.78 \pm 1.49)_{10}\text{-}\mathbf{2}$	$(5.78 \pm 0.59)_{10}\text{-}\mathbf{2}$	$(3.60 \pm 3.16)_{10}\text{-}\mathbf{2}$
HC5	$(7.30 \pm 3.88)_{10}\text{-}\mathbf{3}$	$(3.82 \pm 1.89)_{10}\text{-}\mathbf{2}$	$(9.46 \pm 5.38)_{10}\text{-}\mathbf{3}$
OL2	$(1.71 \pm 0.74)_{10}\text{-}\mathbf{2}$	$(3.77 \pm 0.49)_{10}\text{-}\mathbf{2}$	$(1.02 \pm 0.59)_{10}\text{-}\mathbf{2}$
HC3	$(1.04 \pm 0.69)_{10}\text{-}\mathbf{2}$	$(2.63 \pm 1.41)_{10}\text{-}\mathbf{2}$	$(6.43 \pm 4.75)_{10}\text{-}\mathbf{3}$
GLY	$(1.93 \pm 0.81)_{10}\text{-}\mathbf{2}$	$(1.99 \pm 0.83)_{10}\text{-}\mathbf{2}$	$(1.83 \pm 0.74)_{10}\text{-}\mathbf{2}$
KET	$(1.41 \pm 0.72)_{10}\text{-}\mathbf{2}$	$(3.25 \pm 1.33)_{10}\text{-}\mathbf{2}$	$(3.69 \pm 2.00)_{10}\text{-}\mathbf{3}$
HCHO	$(5.16 \pm 3.69)_{10}\text{-}\mathbf{3}$	$(2.11 \pm 0.74)_{10}\text{-}\mathbf{2}$	$(7.07 \pm 3.55)_{10}\text{-}\mathbf{3}$
ETH	$(2.85 \pm 2.07)_{10}\text{-}\mathbf{3}$	$(6.40 \pm 5.58)_{10}\text{-}\mathbf{3}$	$(7.73 \pm 6.27)_{10}\text{-}\mathbf{4}$

Table A.16: *Mean impact and standard deviation of VOC compounds for categories $\mathcal{C}_{b_{1/2/3/4}}$ for scenario LAND. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{b_1}$	$\mathcal{C}_{b_2}$	$\mathcal{C}_{b_3}$	$\mathcal{C}_{b_4}$
CSL	$(7.48 \pm 0.4)_{10}\text{-}\mathbf{1}$	$(7.70 \pm 0.3)_{10}\text{-}\mathbf{1}$	$(7.87 \pm 0.3)_{10}\text{-}\mathbf{1}$	$(8.04 \pm 0.4)_{10}\text{-}\mathbf{1}$
OLI	$(5.65 \pm 0.9)_{10}\text{-}\mathbf{2}$	$(4.90 \pm 0.8)_{10}\text{-}\mathbf{2}$	$(3.94 \pm 0.8)_{10}\text{-}\mathbf{2}$	$(2.90 \pm 0.9)_{10}\text{-}\mathbf{2}$
ISO	$(4.91 \pm 1.6)_{10}\text{-}\mathbf{2}$	$(4.51 \pm 1.3)_{10}\text{-}\mathbf{2}$	$(3.48 \pm 1.0)_{10}\text{-}\mathbf{2}$	$(2.12 \pm 1.0)_{10}\text{-}\mathbf{2}$

Continued next page

Table A.16 – Continued

TOL	$(2.60\pm1.7)_{10}$-**3**	$(1.87\pm1.0)_{10}$-**2**	$(3.65\pm1.3)_{10}$-**2**	$(4.99\pm1.8)_{10}$-**2**
MGLY	$(5.97\pm2.7)_{10}$-**2**	$(2.49\pm1.4)_{10}$-**2**	$(7.57\pm5.2)_{10}$-**3**	$(5.44\pm5.7)_{10}$-**3**
DCB	$(3.75\pm1.4)_{10}$-**2**	$(2.00\pm0.9)_{10}$-**2**	$(9.74\pm5.1)_{10}$-**3**	$(5.45\pm3.8)_{10}$-**3**
HC8	$(5.83\pm3.7)_{10}$-**3**	$(1.58\pm0.7)_{10}$-**2**	$(2.28\pm0.8)_{10}$-**2**	$(2.36\pm0.8)_{10}$-**2**
XYL	$(7.35\pm4.1)_{10}$-**3**	$(1.45\pm0.6)_{10}$-**2**	$(1.67\pm0.6)_{10}$-**2**	$(1.42\pm0.7)_{10}$-**2**
ALD	$(1.72\pm0.7)_{10}$-**2**	$(1.34\pm0.5)_{10}$-**2**	$(5.82\pm3.7)_{10}$-**3**	$(3.23\pm3.2)_{10}$-**3**
OLT	$(8.55\pm3.4)_{10}$-**3**	$(1.15\pm0.3)_{10}$-**2**	$(1.09\pm0.2)_{10}$-**2**	$(7.54\pm2.9)_{10}$-**3**
HC5	$(1.35\pm0.9)_{10}$-**3**	$(5.33\pm2.8)_{10}$-**3**	$(9.94\pm3.7)_{10}$-**3**	$(1.28\pm0.4)_{10}$-**2**
OL2	$(8.91\pm5.7)_{10}$-**4**	$(4.09\pm2.2)_{10}$-**3**	$(6.89\pm2.3)_{10}$-**3**	$(7.20\pm2.5)_{10}$-**3**
HC3	$(6.59\pm4.5)_{10}$-**4**	$(2.54\pm1.3)_{10}$-**3**	$(4.96\pm1.9)_{10}$-**3**	$(6.97\pm2.5)_{10}$-**3**
GLY	$(3.51\pm1.9)_{10}$-**3**	$(2.78\pm0.8)_{10}$-**3**	$(3.05\pm0.9)_{10}$-**3**	$(3.55\pm1.2)_{10}$-**3**
KET	$(5.72\pm3.6)_{10}$-**4**	$(1.55\pm0.7)_{10}$-**3**	$(2.92\pm1.1)_{10}$-**3**	$(4.37\pm1.7)_{10}$-**3**
HCHO	$(7.56\pm2.7)_{10}$-**4**	$(6.65\pm1.7)_{10}$-**4**	$(7.59\pm2.2)_{10}$-**4**	$(9.12\pm2.8)_{10}$-**4**
ETH	$(5.40\pm3.6)_{10}$-**5**	$(2.72\pm1.6)_{10}$-**4**	$(6.37\pm2.7)_{10}$-**4**	$(1.07\pm0.4)_{10}$-**3**

Table A.17: *Mean impact and standard deviation of VOC compounds for categories* $\mathcal{C}_{b_{1/2/3/4}}$ *for scenario MARINE. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{b_1}$	$\mathcal{C}_{b_2}$	$\mathcal{C}_{b_3}$	$\mathcal{C}_{b_4}$
CSL	$(6.31\pm0.2)_{10}$-**1**	$(6.48\pm0.1)_{10}$-**1**	$(7.27\pm0.4)_{10}$-**1**	$(7.77\pm0.3)_{10}$-**1**
HC8	$(1.69\pm0.8)_{10}$-**2**	$(3.89\pm0.9)_{10}$-**2**	$(4.46\pm0.3)_{10}$-**2**	$(2.59\pm1.1)_{10}$-**2**
MGLY	$(7.50\pm2.0)_{10}$-**2**	$(3.53\pm1.4)_{10}$-**2**	$(1.08\pm0.8)_{10}$-**2**	$(7.69\pm4.7)_{10}$-**3**
DCB	$(5.78\pm0.8)_{10}$-**2**	$(3.73\pm0.8)_{10}$-**2**	$(1.81\pm0.8)_{10}$-**2**	$(1.09\pm0.5)_{10}$-**2**
OLI	$(5.34\pm0.6)_{10}$-**2**	$(3.94\pm0.9)_{10}$-**2**	$(1.75\pm1.2)_{10}$-**2**	$(1.36\pm1.1)_{10}$-**2**
TOL	$(1.18\pm0.6)_{10}$-**2**	$(3.18\pm0.9)_{10}$-**2**	$(4.32\pm0.3)_{10}$-**2**	$(3.34\pm0.8)_{10}$-**2**
ISO	$(4.61\pm0.8)_{10}$-**2**	$(3.11\pm1.1)_{10}$-**2**	$(1.56\pm0.9)_{10}$-**2**	$(3.10\pm2.3)_{10}$-**2**
XYL	$(2.80\pm0.9)_{10}$-**2**	$(3.27\pm0.2)_{10}$-**2**	$(2.05\pm0.7)_{10}$-**2**	$(1.06\pm0.5)_{10}$-**2**
ALD	$(3.57\pm0.8)_{10}$-**2**	$(2.60\pm0.7)_{10}$-**2**	$(7.40\pm7.0)_{10}$-**3**	$(1.02\pm0.6)_{10}$-**2**

Continued next page

Table A.17 – Continued

HC5	$(5.54\pm3.1)_{10}$-**3**	$(1.77\pm0.6)_{10}$-**2**	$(2.69\pm0.3)_{10}$-**2**	$(2.23\pm0.4)_{10}$-**2**
OLT	$(2.34\pm0.7)_{10}$-**2**	$(2.61\pm0.3)_{10}$-**2**	$(1.55\pm0.6)_{10}$-**2**	$(7.11\pm3.9)_{10}$-**3**
HC3	$(3.38\pm1.9)_{10}$-**3**	$(1.08\pm0.4)_{10}$-**2**	$(1.80\pm0.3)_{10}$-**2**	$(1.80\pm0.2)_{10}$-**2**
KET	$(2.68\pm1.2)_{10}$-**3**	$(6.92\pm2.1)_{10}$-**3**	$(1.21\pm0.2)_{10}$-**2**	$(1.43\pm0.1)_{10}$-**2**
OL2	$(3.73\pm2.2)_{10}$-**3**	$(1.15\pm0.4)_{10}$-**2**	$(1.27\pm1.1)_{10}$-**2**	$(4.47\pm4.3)_{10}$-**3**
GLY	$(4.52\pm1.3)_{10}$-**3**	$(4.55\pm0.8)_{10}$-**3**	$(5.96\pm1.2)_{10}$-**3**	$(7.30\pm1.3)_{10}$-**3**
ETH	$(4.06\pm2.3)_{10}$-**4**	$(1.59\pm0.6)_{10}$-**3**	$(3.42\pm0.9)_{10}$-**3**	$(4.88\pm0.6)_{10}$-**3**
HCHO	$(7.03\pm3.4)_{10}$-**4**	$(7.54\pm4.0)_{10}$-**4**	$(1.01\pm0.5)_{10}$-**3**	$(1.18\pm0.7)_{10}$-**3**

Table A.18: *Mean impact and standard deviation of VOC compounds for categories $\mathcal{C}_{b_{1/2/3/4}}$ for scenario FREE. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{b_1}$	$\mathcal{C}_{b_2}$	$\mathcal{C}_{b_3}$	$\mathcal{C}_{b_4}$
MGLY	$(3.99\pm0.4)_{10}$-**1**	$(3.97\pm0.4)_{10}$-**1**	$(3.43\pm0.3)_{10}$-**1**	$(2.88\pm0.2)_{10}$-**1**
CSL	$(1.53\pm0.5)_{10}$-**1**	$(2.00\pm0.6)_{10}$-**1**	$(2.14\pm0.5)_{10}$-**1**	$(2.03\pm0.4)_{10}$-**1**
DCB	$(1.54\pm0.2)_{10}$-**1**	$(1.78\pm0.2)_{10}$-**1**	$(1.72\pm0.1)_{10}$-**1**	$(1.57\pm0.1)_{10}$-**1**
OLI	$(5.28\pm1.5)_{10}$-**2**	$(5.62\pm0.6)_{10}$-**2**	$(5.88\pm0.2)_{10}$-**2**	$(5.75\pm0.1)_{10}$-**2**
ALD	$(2.33\pm1.1)_{10}$-**2**	$(5.08\pm1.1)_{10}$-**2**	$(6.34\pm0.5)_{10}$-**2**	$(6.53\pm0.3)_{10}$-**2**
GLY	$(5.29\pm1.3)_{10}$-**2**	$(3.36\pm0.6)_{10}$-**2**	$(2.13\pm0.4)_{10}$-**2**	$(1.44\pm0.2)_{10}$-**2**
XYL	$(1.79\pm0.5)_{10}$-**2**	$(1.36\pm0.6)_{10}$-**2**	$(3.32\pm0.8)_{10}$-**2**	$(4.67\pm0.5)_{10}$-**2**
ISO	$(5.12\pm1.8)_{10}$-**2**	$(1.75\pm1.4)_{10}$-**2**	$(1.35\pm0.8)_{10}$-**2**	$(2.58\pm0.8)_{10}$-**2**
OLT	$(2.64\pm0.7)_{10}$-**2**	$(9.18\pm6.1)_{10}$-**3**	$(1.98\pm0.7)_{10}$-**2**	$(3.16\pm0.4)_{10}$-**2**
HC8	$(1.18\pm0.1)_{10}$-**2**	$(6.88\pm3.2)_{10}$-**3**	$(2.05\pm0.8)_{10}$-**2**	$(3.59\pm0.6)_{10}$-**2**
TOL	$(1.23\pm0.2)_{10}$-**2**	$(5.73\pm3.9)_{10}$-**3**	$(1.71\pm0.7)_{10}$-**2**	$(3.11\pm0.5)_{10}$-**2**
OL2	$(1.85\pm0.2)_{10}$-**2**	$(1.09\pm0.5)_{10}$-**2**	$(6.53\pm2.9)_{10}$-**3**	$(1.52\pm0.4)_{10}$-**2**
HCHO	$(1.49\pm0.2)_{10}$-**2**	$(1.11\pm0.2)_{10}$-**2**	$(7.50\pm1.4)_{10}$-**3**	$(5.27\pm0.8)_{10}$-**3**
HC5	$(8.14\pm1.5)_{10}$-**3**	$(6.12\pm2.4)_{10}$-**3**	$(4.29\pm1.8)_{10}$-**3**	$(1.12\pm0.4)_{10}$-**2**
HC3	$(3.07\pm0.6)_{10}$-**3**	$(2.30\pm1.0)_{10}$-**3**	$(1.98\pm0.9)_{10}$-**3**	$(5.35\pm1.7)_{10}$-**3**

Continued next page

Table A.18 – Continued

KET	$(9.42\pm1.3)_{10}\text{-}4$	$(9.90\pm4.0)_{10}\text{-}4$	$(3.50\pm1.3)_{10}\text{-}3$	$(6.49\pm1.3)_{10}\text{-}3$
ETH	$(3.42\pm0.9)_{10}\text{-}4$	$(3.62\pm0.6)_{10}\text{-}3$	$(1.24\pm1.1)_{10}\text{-}4$	$(3.13\pm1.7)_{10}\text{-}4$

Table A.19: *Mean impact and standard deviation of VOC compounds for categories $\mathcal{C}_{b_{1/2/3/4}}$ for scenario PLUME. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{b_1}$	$\mathcal{C}_{b_2}$	$\mathcal{C}_{b_3}$	$\mathcal{C}_{b_4}$
CSL	$(4.97\pm0.6)_{10}\text{-}1$	$(5.70\pm0.9)_{10}\text{-}1$	$(5.72\pm0.8)_{10}\text{-}1$	$(4.41\pm1.2)_{10}\text{-}1$
MGLY	$(1.10\pm0.3)_{10}\text{-}1$	$(6.43\pm3.3)_{10}\text{-}2$	$(4.25\pm2.8)_{10}\text{-}2$	$(6.69\pm3.8)_{10}\text{-}2$
DCB	$(7.54\pm2.2)_{10}\text{-}2$	$(4.80\pm2.3)_{10}\text{-}2$	$(4.15\pm3.0)_{10}\text{-}2$	$(7.56\pm3.9)_{10}\text{-}2$
XYL	$(5.48\pm1.5)_{10}\text{-}2$	$(4.82\pm2.2)_{10}\text{-}2$	$(3.73\pm2.1)_{10}\text{-}2$	$(5.17\pm3.4)_{10}\text{-}2$
TOL	$(2.71\pm1.0)_{10}\text{-}2$	$(5.64\pm1.2)_{10}\text{-}2$	$(6.18\pm2.0)_{10}\text{-}2$	$(3.77\pm2.4)_{10}\text{-}2$
ALD	$(6.93\pm1.0)_{10}\text{-}2$	$(5.48\pm1.6)_{10}\text{-}2$	$(2.98\pm1.7)_{10}\text{-}2$	$(2.85\pm1.8)_{10}\text{-}2$
ISO	$(2.34\pm1.0)_{10}\text{-}2$	$(1.68\pm1.0)_{10}\text{-}2$	$(5.24\pm2.6)_{10}\text{-}2$	$(9.49\pm2.8)_{10}\text{-}2$
OLI	$(4.02\pm1.8)_{10}\text{-}2$	$(2.40\pm2.0)_{10}\text{-}2$	$(3.68\pm2.1)_{10}\text{-}2$	$(7.07\pm2.9)_{10}\text{-}2$
OLT	$(3.50\pm1.0)_{10}\text{-}2$	$(2.66\pm1.0)_{10}\text{-}2$	$(1.52\pm1.1)_{10}\text{-}2$	$(3.68\pm1.7)_{10}\text{-}2$
GLY	$(1.99\pm0.5)_{10}\text{-}2$	$(2.09\pm0.5)_{10}\text{-}2$	$(2.02\pm0.9)_{10}\text{-}2$	$(1.58\pm1.1)_{10}\text{-}2$
OL2	$(2.03\pm0.4)_{10}\text{-}2$	$(1.55\pm0.8)_{10}\text{-}2$	$(1.43\pm0.8)_{10}\text{-}2$	$(1.89\pm0.7)_{10}\text{-}2$
HC8	$(7.41\pm6.5)_{10}\text{-}3$	$(1.88\pm0.9)_{10}\text{-}2$	$(2.11\pm0.9)_{10}\text{-}2$	$(1.46\pm0.9)_{10}\text{-}2$
KET	$(5.41\pm2.1)_{10}\text{-}3$	$(1.34\pm0.4)_{10}\text{-}2$	$(1.99\pm0.4)_{10}\text{-}2$	$(1.69\pm0.8)_{10}\text{-}2$
HC3	$(2.15\pm1.3)_{10}\text{-}3$	$(1.04\pm0.4)_{10}\text{-}2$	$(1.66\pm0.4)_{10}\text{-}2$	$(1.18\pm0.7)_{10}\text{-}2$
HC5	$(7.79\pm2.1)_{10}\text{-}3$	$(5.16\pm1.7)_{10}\text{-}3$	$(8.94\pm4.4)_{10}\text{-}3$	$(7.40\pm5.2)_{10}\text{-}3$
HCHO	$(4.71\pm4.5)_{10}\text{-}3$	$(4.54\pm2.9)_{10}\text{-}3$	$(5.64\pm3.5)_{10}\text{-}3$	$(5.82\pm3.8)_{10}\text{-}3$
ETH	$(3.55\pm2.0)_{10}\text{-}4$	$(1.99\pm0.8)_{10}\text{-}3$	$(4.16\pm1.0)_{10}\text{-}3$	$(4.91\pm1.8)_{10}\text{-}3$

Table A.20: *Mean impact and standard deviation of VOC compounds for categories* $\mathcal{C}_{b_{1/2/3/4}}$ *for scenario URBAN. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{b_1}$	$\mathcal{C}_{b_2}$	$\mathcal{C}_{b_3}$	$\mathcal{C}_{b_4}$
XYL	$(1.46\pm0.1)_{10}\text{-}\mathbf{1}$	$(1.33\pm0.1)_{10}\text{-}\mathbf{1}$	$(1.21\pm0.1)_{10}\text{-}\mathbf{1}$	$(1.14\pm0.1)_{10}\text{-}\mathbf{1}$
DCB	$(1.57\pm0.4)_{10}\text{-}\mathbf{1}$	$(1.20\pm0.2)_{10}\text{-}\mathbf{1}$	$(1.06\pm0.2)_{10}\text{-}\mathbf{1}$	$(9.59\pm2.0)_{10}\text{-}\mathbf{2}$
CSL	$(8.21\pm2.4)_{10}\text{-}\mathbf{2}$	$(1.08\pm0.9)_{10}\text{-}\mathbf{1}$	$(1.19\pm0.1)_{10}\text{-}\mathbf{1}$	$(1.23\pm0.1)_{10}\text{-}\mathbf{1}$
MGLY	$(1.36\pm0.4)_{10}\text{-}\mathbf{1}$	$(1.04\pm0.2)_{10}\text{-}\mathbf{1}$	$(8.99\pm2.2)_{10}\text{-}\mathbf{2}$	$(8.03\pm2.4)_{10}\text{-}\mathbf{2}$
TOL	$(5.23\pm1.3)_{10}\text{-}\mathbf{2}$	$(6.67\pm0.7)_{10}\text{-}\mathbf{2}$	$(7.07\pm0.6)_{10}\text{-}\mathbf{2}$	$(7.16\pm0.5)_{10}\text{-}\mathbf{2}$
ISO	$(6.51\pm1.0)_{10}\text{-}\mathbf{2}$	$(6.40\pm0.7)_{10}\text{-}\mathbf{2}$	$(6.13\pm0.6)_{10}\text{-}\mathbf{2}$	$(5.88\pm0.6)_{10}\text{-}\mathbf{2}$
OLI	$(6.44\pm1.0)_{10}\text{-}\mathbf{2}$	$(6.36\pm0.7)_{10}\text{-}\mathbf{2}$	$(6.08\pm0.7)_{10}\text{-}\mathbf{2}$	$(5.83\pm0.6)_{10}\text{-}\mathbf{2}$
HC8	$(3.74\pm1.5)_{10}\text{-}\mathbf{2}$	$(5.38\pm1.9)_{10}\text{-}\mathbf{2}$	$(6.83\pm2.3)_{10}\text{-}\mathbf{2}$	$(7.98\pm2.5)_{10}\text{-}\mathbf{2}$
OLT	$(5.91\pm0.8)_{10}\text{-}\mathbf{2}$	$(5.95\pm0.5)_{10}\text{-}\mathbf{2}$	$(5.71\pm0.5)_{10}\text{-}\mathbf{2}$	$(5.52\pm0.5)_{10}\text{-}\mathbf{2}$
ALD	$(5.82\pm1.4)_{10}\text{-}\mathbf{2}$	$(5.53\pm0.9)_{10}\text{-}\mathbf{2}$	$(4.99\pm0.7)_{10}\text{-}\mathbf{2}$	$(4.53\pm0.7)_{10}\text{-}\mathbf{2}$
HC5	$(2.14\pm1.0)_{10}\text{-}\mathbf{2}$	$(3.39\pm1.4)_{10}\text{-}\mathbf{2}$	$(4.47\pm1.7)_{10}\text{-}\mathbf{2}$	$(5.32\pm1.8)_{10}\text{-}\mathbf{2}$
OL2	$(3.77\pm0.9)_{10}\text{-}\mathbf{2}$	$(3.92\pm0.2)_{10}\text{-}\mathbf{2}$	$(3.75\pm0.2)_{10}\text{-}\mathbf{2}$	$(3.63\pm0.3)_{10}\text{-}\mathbf{2}$
KET	$(1.82\pm0.6)_{10}\text{-}\mathbf{2}$	$(2.91\pm0.8)_{10}\text{-}\mathbf{2}$	$(3.80\pm1.0)_{10}\text{-}\mathbf{2}$	$(4.50\pm1.2)_{10}\text{-}\mathbf{2}$
HC3	$(1.24\pm0.6)_{10}\text{-}\mathbf{2}$	$(2.31\pm1.0)_{10}\text{-}\mathbf{2}$	$(3.19\pm1.2)_{10}\text{-}\mathbf{2}$	$(3.82\pm1.3)_{10}\text{-}\mathbf{2}$
HCHO	$(2.64\pm0.6)_{10}\text{-}\mathbf{2}$	$(2.19\pm0.6)_{10}\text{-}\mathbf{2}$	$(1.90\pm0.7)_{10}\text{-}\mathbf{2}$	$(1.70\pm0.7)_{10}\text{-}\mathbf{2}$
GLY	$(2.59\pm0.9)_{10}\text{-}\mathbf{2}$	$(2.01\pm0.7)_{10}\text{-}\mathbf{2}$	$(1.77\pm0.7)_{10}\text{-}\mathbf{2}$	$(1.62\pm0.7)_{10}\text{-}\mathbf{2}$
ETH	$(1.79\pm1.0)_{10}\text{-}\mathbf{3}$	$(4.36\pm2.5)_{10}\text{-}\mathbf{3}$	$(7.82\pm4.5)_{10}\text{-}\mathbf{3}$	$(1.20\pm0.7)_{10}\text{-}\mathbf{2}$

Table A.21: *Mean impact and standard deviation of VOC compounds for categories* $\mathcal{C}_{b_{1/2/3/4}}$ *for scenario BIO. Considered are optimal projected singular vectors with respect to initial uncertainties.*

	$\mathcal{C}_{2a}$	$\mathcal{C}_{2b}$	$\mathcal{C}_{2c}$	$\mathcal{C}_{2d}$
CSL	$(5.00\pm0.9)_{10}\text{-}\mathbf{1}$	$(4.68\pm1.1)_{10}\text{-}\mathbf{1}$	$(3.73\pm1.5)_{10}\text{-}\mathbf{1}$	$(2.94\pm1.9)_{10}\text{-}\mathbf{1}$
ISO	$(5.09\pm2.4)_{10}\text{-}\mathbf{2}$	$(8.84\pm3.5)_{10}\text{-}\mathbf{2}$	$(1.27\pm0.5)_{10}\text{-}\mathbf{1}$	$(1.60\pm0.6)_{10}\text{-}\mathbf{1}$
XYL	$(8.81\pm3.0)_{10}\text{-}\mathbf{2}$	$(9.96\pm2.2)_{10}\text{-}\mathbf{2}$	$(9.58\pm2.2)_{10}\text{-}\mathbf{2}$	$(8.77\pm2.1)_{10}\text{-}\mathbf{2}$

Continued next page

Table A.21 – Continued

OLI	$(4.76\pm2.3)_{10}\text{-}\mathbf{2}$	$(7.10\pm3.1)_{10}\text{-}\mathbf{2}$	$(9.83\pm5.0)_{10}\text{-}\mathbf{2}$	$(1.28\pm0.7)_{10}\text{-}\mathbf{1}$
TOL	$(4.66\pm2.1)_{10}\text{-}\mathbf{2}$	$(8.01\pm2.1)_{10}\text{-}\mathbf{2}$	$(1.02\pm0.2)_{10}\text{-}\mathbf{1}$	$(1.12\pm0.2)_{10}\text{-}\mathbf{1}$
MGLY	$(7.97\pm4.4)_{10}\text{-}\mathbf{2}$	$(3.90\pm1.3)_{10}\text{-}\mathbf{2}$	$(2.41\pm0.8)_{10}\text{-}\mathbf{2}$	$(1.71\pm0.7)_{10}\text{-}\mathbf{2}$
OLT	$(2.22\pm1.7)_{10}\text{-}\mathbf{2}$	$(2.33\pm1.7)_{10}\text{-}\mathbf{2}$	$(4.25\pm3.0)_{10}\text{-}\mathbf{2}$	$(5.79\pm4.2)_{10}\text{-}\mathbf{2}$
ALD	$(4.66\pm1.1)_{10}\text{-}\mathbf{2}$	$(3.55\pm0.9)_{10}\text{-}\mathbf{2}$	$(2.63\pm0.6)_{10}\text{-}\mathbf{2}$	$(2.08\pm0.6)_{10}\text{-}\mathbf{2}$
DCB	$(6.06\pm3.2)_{10}\text{-}\mathbf{2}$	$(2.83\pm1.1)_{10}\text{-}\mathbf{2}$	$(1.69\pm1.0)_{10}\text{-}\mathbf{2}$	$(1.27\pm0.9)_{10}\text{-}\mathbf{2}$
HC8	$(1.01\pm0.9)_{10}\text{-}\mathbf{2}$	$(2.21\pm1.0)_{10}\text{-}\mathbf{2}$	$(3.43\pm0.8)_{10}\text{-}\mathbf{2}$	$(3.60\pm0.9)_{10}\text{-}\mathbf{2}$
GLY	$(2.36\pm1.0)_{10}\text{-}\mathbf{2}$	$(1.79\pm0.6)_{10}\text{-}\mathbf{2}$	$(1.64\pm0.5)_{10}\text{-}\mathbf{2}$	$(1.56\pm0.5)_{10}\text{-}\mathbf{2}$
OL2	$(5.32\pm5.3)_{10}\text{-}\mathbf{3}$	$(8.72\pm5.0)_{10}\text{-}\mathbf{3}$	$(1.28\pm0.5)_{10}\text{-}\mathbf{2}$	$(1.41\pm0.4)_{10}\text{-}\mathbf{2}$
HC5	$(8.09\pm3.3)_{10}\text{-}\mathbf{3}$	$(4.23\pm2.0)_{10}\text{-}\mathbf{3}$	$(1.04\pm0.4)_{10}\text{-}\mathbf{2}$	$(1.60\pm0.4)_{10}\text{-}\mathbf{2}$
HCHO	$(6.86\pm3.7)_{10}\text{-}\mathbf{3}$	$(7.08\pm3.6)_{10}\text{-}\mathbf{3}$	$(7.25\pm3.5)_{10}\text{-}\mathbf{3}$	$(7.04\pm3.4)_{10}\text{-}\mathbf{3}$
HC3	$(1.98\pm1.7)_{10}\text{-}\mathbf{3}$	$(3.34\pm2.0)_{10}\text{-}\mathbf{3}$	$(8.29\pm2.7)_{10}\text{-}\mathbf{3}$	$(1.26\pm0.3)_{10}\text{-}\mathbf{2}$
KET	$(1.65\pm0.9)_{10}\text{-}\mathbf{3}$	$(2.79\pm0.8)_{10}\text{-}\mathbf{3}$	$(4.31\pm1.2)_{10}\text{-}\mathbf{3}$	$(6.16\pm1.6)_{10}\text{-}\mathbf{3}$
ETH	$(4.18\pm1.7)_{10}\text{-}\mathbf{4}$	$(2.57\pm1.3)_{10}\text{-}\mathbf{4}$	$(8.44\pm3.6)_{10}\text{-}\mathbf{4}$	$(1.68\pm0.5)_{10}\text{-}\mathbf{3}$

A.2.2   Relative error growth of VOC and NO_x species

Figure A.1: *Statistics of optimal projected relative singular vectors with respect to initial uncertainties for categories $\mathcal{C}_{a_{1/2/3/4}}$ and $\mathcal{C}_{b_{1/2/3/4}}$ for scenario PLUME. The VOC and NO_x compounds with most influence have been chosen for presentation. Depicted are mean impact ($\mathcal{C}_{a_1/b_1}$: blue bars, $\mathcal{C}_{a_2/b_2}$: turquoise bars, $\mathcal{C}_{a_3/b_3}$: green bars, $\mathcal{C}_{a_4/b_4}$: orange bars), minimum/maximum value (dark blue lines) and standard deviation (red bars).*
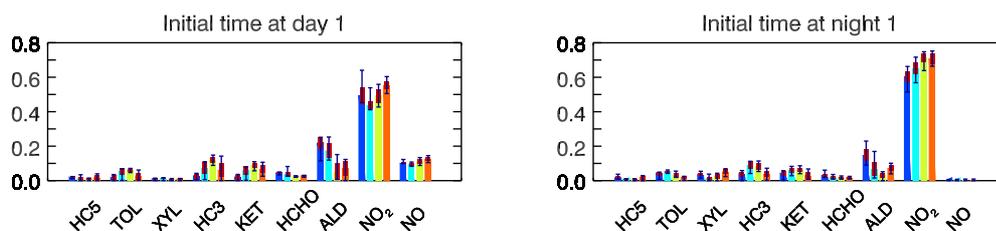
Figure A.1: *-Continued.*
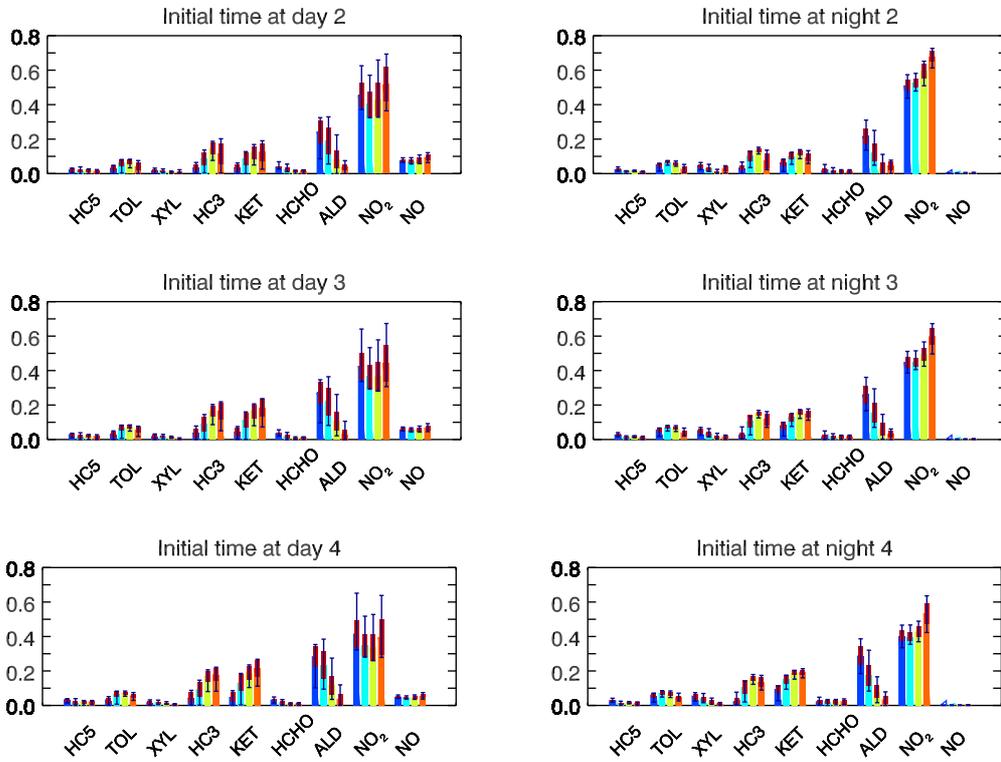
Figure A.2: *Statistics of optimal projected relative singular vectors with respect to initial uncertainties for categories $C_{a_{1/2/3/4}}$ and $C_{b_{1/2/3/4}}$ for scenario URBAN. Plotting conventions as in Fig. A.1.*
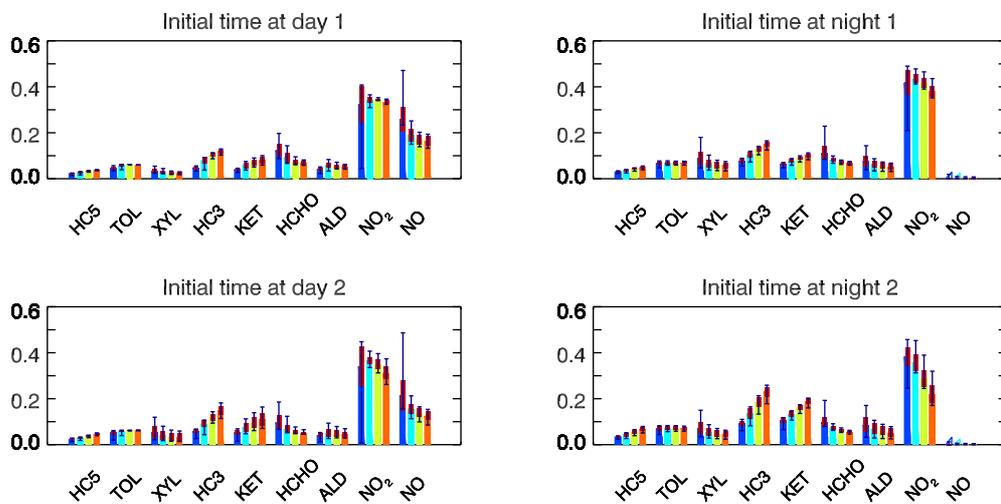
Figure A.2: -Continued.
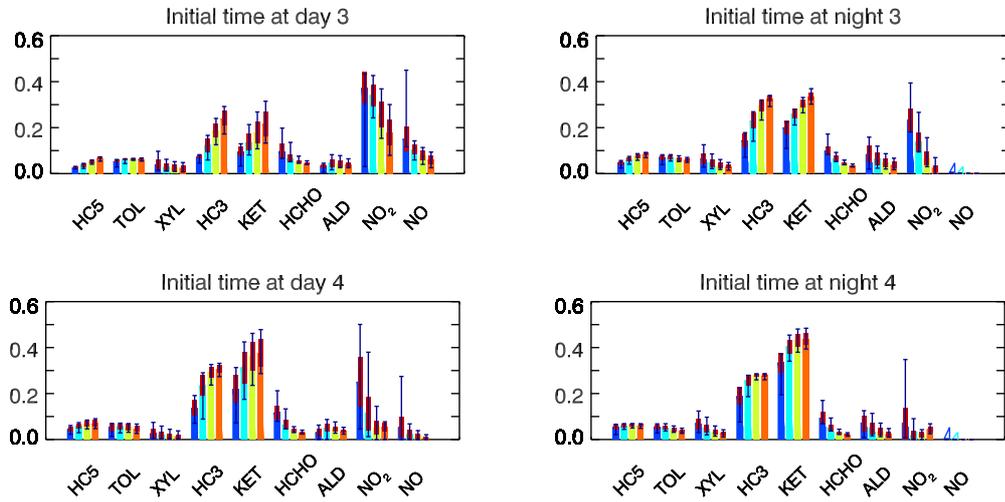
Figure A.3: Statistics of optimal projected relative singular vectors with respect to initial uncertainties for categories $\mathcal{C}_{a_{1/2/3/4}}$ and $\mathcal{C}_{b_{1/2/3/4}}$ for scenario BIO. Plotting conventions as in Fig. A.1.
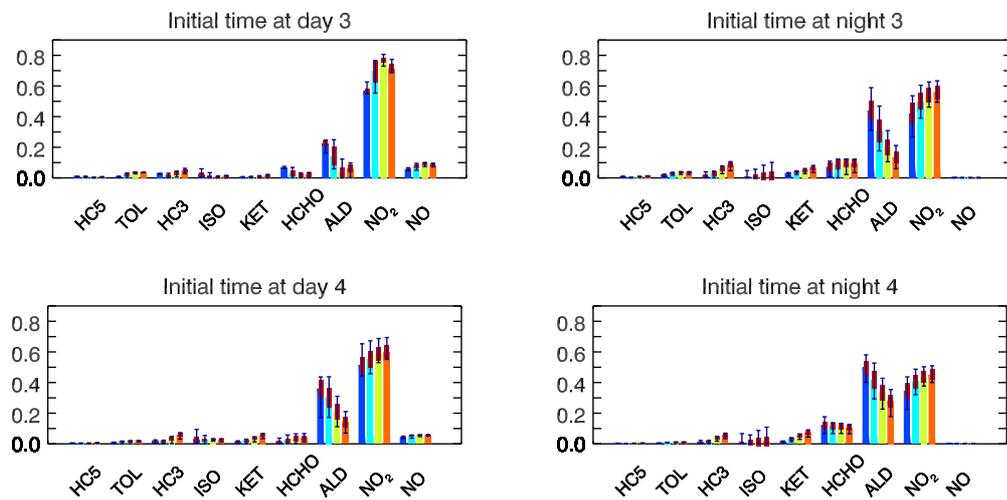
Results for the ZEPTER-2 campaign

In the following, specifications and results for the considered ZEPTER-2-case study are depicted. In order to ease orientation, this chapter is organized according to chapter 5.

B.1 Design of sensitivity experiments

B.1.1 RACM-MIM species

Table B.1: *RACM-MIM species list according to Stockwell et al. [1997] and Geiger et al. [2003]. RACM species are presented first, additional MIM-reactants are denoted subsequent. PR indicates peroxy radicals.*

Species	Definition
Stable Inorganic Compounds	
Oxidants	
O_3	Ozone
H_2O_2	Hydrogen peroxide
Nitrogenous comp.	
NO	Nitric oxide
NO_2	Nitrogen dioxide

Continued next page

Table B.1 – Continued

NO_3	Nitrogen trioxide
N_2O_5	Dinitrogen pentoxide
HONO	Nitrous acid
HNO_3	Nitric acid
HNO_4	Pernitric acid
Sulfur compounds	
SO_2	Sulfur dioxide
SULF	Sulfuric acid
Carbon oxides	
CO	Carbon monoxide
CO_2	Carbon dioxide
	Abundant Stable Species
N_2	Nitrogen
O_2	Oxygen
H_2O	Water
H_2	Hydrogen
	Inorganic Short-Lived Intermediates
Atomic species	
O^3P	Ground state oxygen atom
O^1D	Excited state oxygen atom
Odd hydrogen	
HO	Hydroxy radical
HO_2	Hydroperoxy radical
	Stable Organic Compounds
Alkanes	
CH_4	Methane
ETH	Ethane
HC3	Alkanes, alcohols, esters and alkynes with HO rate constant (298 K, 1 atm) less than 3.4×10^{-12} cm^{3}s^{-1}
HC5	Alkanes, alcohols, esters and alkynes with HO rate constant (298 K, 1 atm) between 3.4×10^{-12} cm^{3}s^{-1} and 6.8×10^{-12} cm^{3}s^{-1}
HC8	Alkanes, alcohols, esters and alkynes with HO rate constant (298 K, 1 atm) greater than 6.8×10^{-12} cm^{3}s^{-1}
Alkenes	

Continued next page

Table B.1 – Continued

ETE	Ethene
OLT	Terminal alkenes
OLI	Internal alkenes
DIEN	Butadiene and other anthropogenic dienes
Stable biogenic alkenes	
ISO	Isoprene
API	α-pinene and other cyclic terpenes with one double bond
LIM	d-limonene and other cyclic diene-terpenes
Aromatics	
TOL	Toluene and less reactive aromatics
XYL	Xylene and more reactive aromatics
CSL	Cresol and other hydroxy substituted aromatics
Carbonyls	
HCHO	Formaldehyde
ALD	Acetaldehyde and higher aldehydes
KET	Ketones
GLY	Glyoxal
MGLY	Methylglyoxal and other α-carbonyl aldehydes
DCB	Unsaturated Dicarbonyls
MACR	Methacrolein and other unsaturated monoaldehydes
UDD	Unsaturated dihydroxy dicarbonyl
HKET	hydroxy ketone
Organic nitrogen	
ONIT	Organic nitrate
PAN	Peroxyacetyl nitrate and higher saturated PANs
TPAN	Unsaturated PANs
Organic peroxides	
OP1	Methyl hydrogen peroxide
OP2	Higher organic peroxides
PAA	Peroxyacetic acid and higher analogs
Organic acids	
ORA1	Formic acid
ORA2	Acetic acid and higher acids

Continued next page

Table B.1 – Continued

	Organic Short-Lived Intermediates
PR from alkanes	
MO2	Methyl peroxy radical
ETHP	Peroxy radical formed from ETH
HC3P	Peroxy radical formed from HC3
HC5P	Peroxy radical formed from HC5
HC8P	Peroxy radical formed from HC8
PR from alkenes	
ETEP	Peroxy radicals formed from ETE
OLTP	Peroxy radicals formed from OLT
OLIP	Peroxy radicals formed from OLI
PR from biog. alkenes	
ISOP	Peroxy radicals formed from ISO and DIEN
APIP	Peroxy radicals formed from API
LIMP	Peroxy radicals formed from LIM
PR from aromatics	
TOLP	Peroxy radicals formed from TOL
XYLP	Peroxy radicals formed from XYL
CSLP	Peroxy radicals formed from CSL
PR with carbonyl gr.	
ACO_3	Acetyl peroxy and higher saturated acyl peroxy radicals
TCO_3	Unsaturated acyl peroxy radicals
KETP	Peroxy radicals formed from KET
Other PR	
OLNN	NO_3-alkene adduct reacting to form carbonitrates + HO_2
OLND	NO_3-alkene adduct reacting via decomposition
XO_2	Accounts for additional NO to NO_2 conversions
	Additional MIM-reactants
HACE	Hydroxyacetone and other C_3 ketones
ISHP	β-hydroxy hydroperoxides from ISOP+HO_2
ISON	β-hydroxyalkylnitrates from ISOP+NO and alkylnitrates from ISO+NO_3
MACP	Peroxy radicals from MACR+OH
MAHP	Hydroperoxides from MACP+HO2

Continued next page

Table B.1 – Continued

MPAN	Peroxymethacryloylnitrate and other higher peroxya-cylnitrates from isoprene oxidation
NALD	Nitrooxyacetaldehyde

B.2 Singular vectors with respect to initial uncertainties

B.2.1 Optimal placement of observations

Figure B.1: *Optimal placement for case 1a. The optimization criterion are the ozone values of the final profile VP(t_F).*

The optimal horizontal placement of measurement is illustrated in the left panel. Shown are 0.01-isopleths of the horizontal singular vector for passive tracer (red shading) and ozone (green shading, isopleths are given as color fill). The final profile VP(t_F) is marked with black dots. To aid interpretation, its horizontal position is marked with a black cross on surface level. Case numbers, simulation starts and simulation lengths are denoted on top of the panel.

The optimal vertical placement of measurement is illustrated in the right panel. Shown is the length of the vertical singular vector per model level for CO, OH, HONO, O$_3$, NO$_2$ and NO (color code of the species is denoted to the right of the panel). The black box indicates the extent of the height of the final profile VP(t_F). The considered case is denoted on top of the panel.
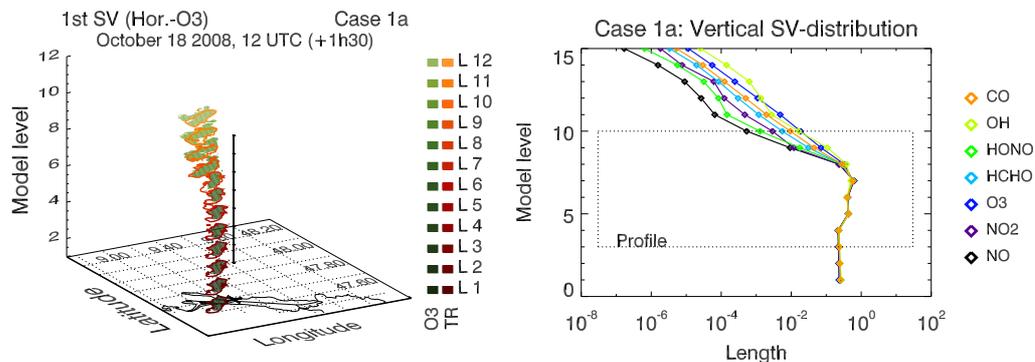
Figure B.2: *Optimal placement for case 1b. Plotting conventions as in Fig. B.1.*
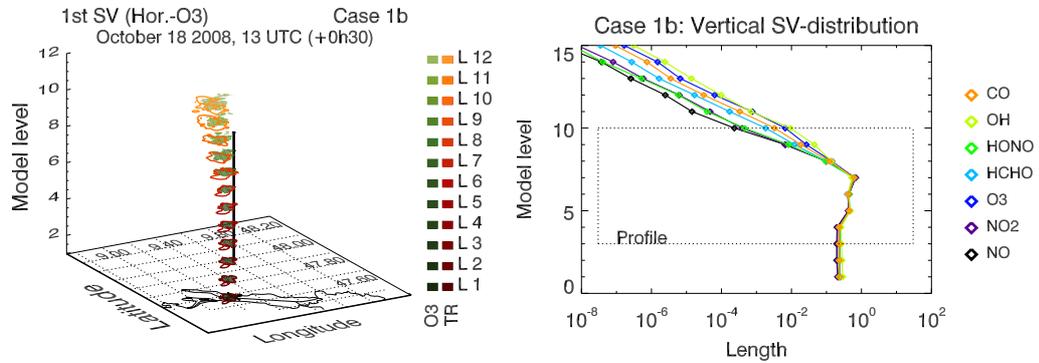
Figure B.3: *Optimal placement for case 2a. Plotting conventions as in Fig. B.1.*
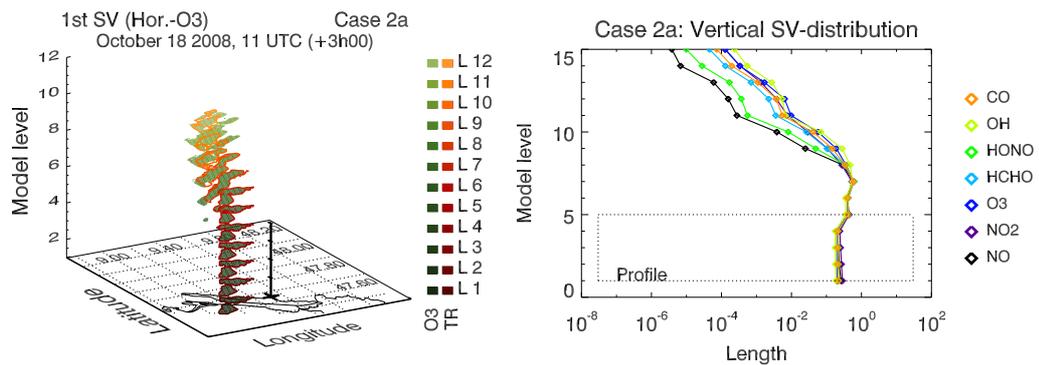
Figure B.4: *Optimal placement for case 2b. Plotting conventions as in Fig. B.1.*
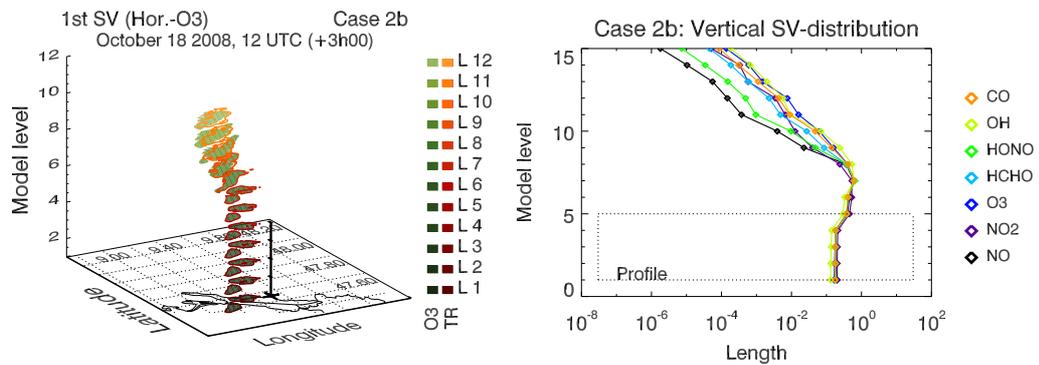
Figure B.5: *Optimal placement for case 3. Plotting conventions as in Fig. B.1.*
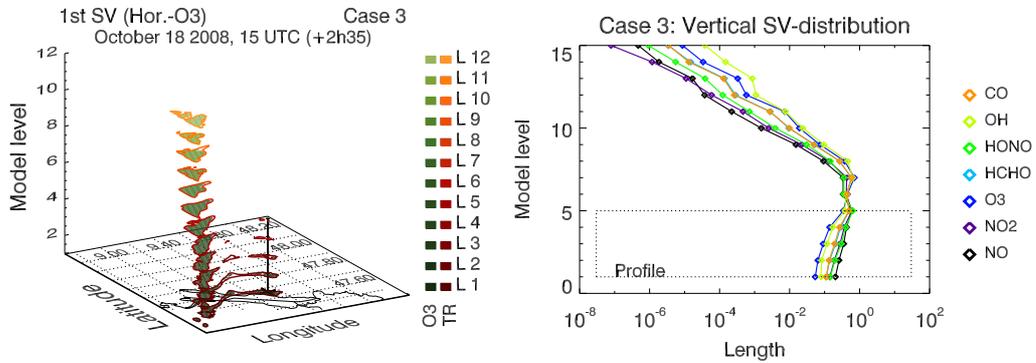
Figure B.6: *Optimal placement for case 4a. Plotting conventions as in Fig. B.1.*
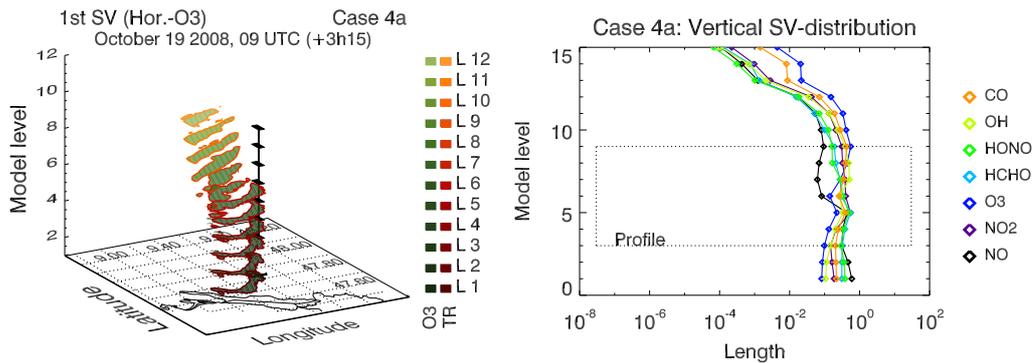
Figure B.7: *Optimal placement for case 4b. Plotting conventions as in Fig. B.1.*
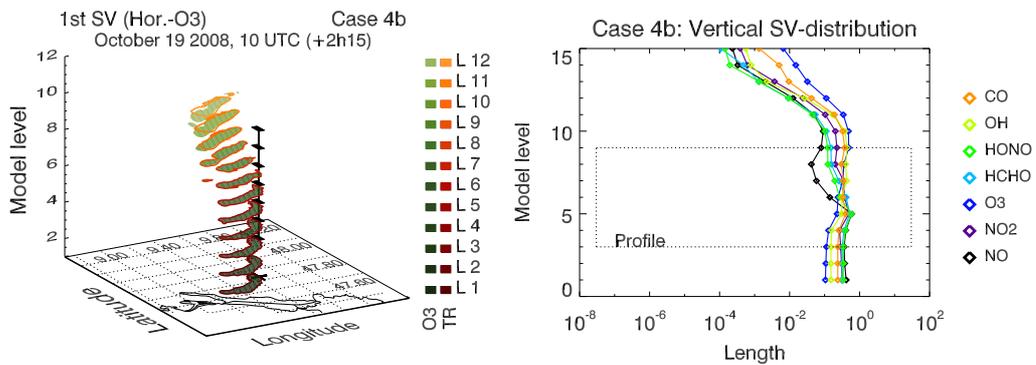
Figure B.8: *Optimal placement for case 5a. Plotting conventions as in Fig. B.1.*
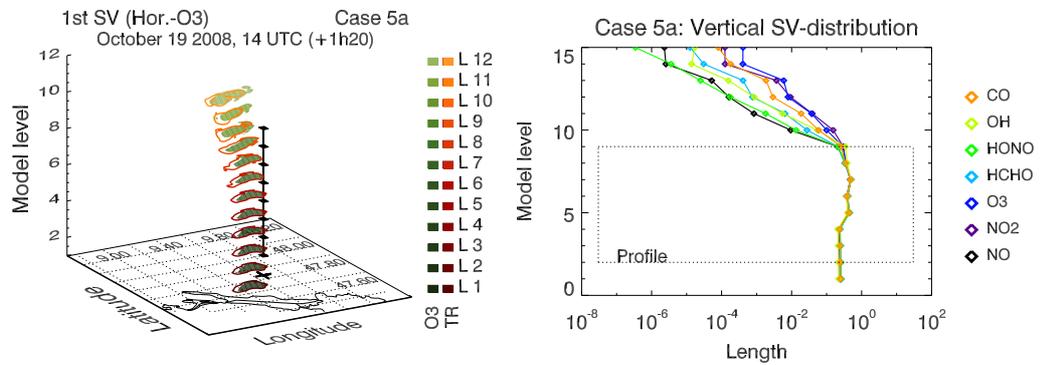
Figure B.9: *Optimal placement for case 5b. Plotting conventions as in Fig. B.1.*
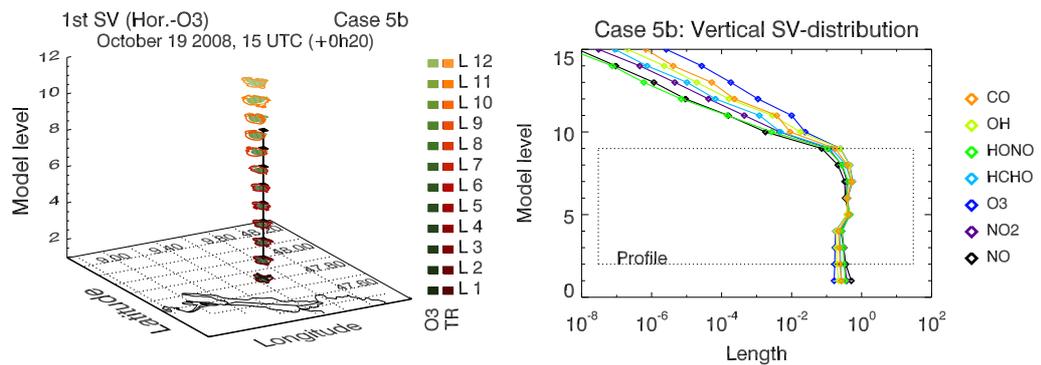
Figure B.10: *Optimal placement for case 6. Plotting conventions as in Fig. B.1.*
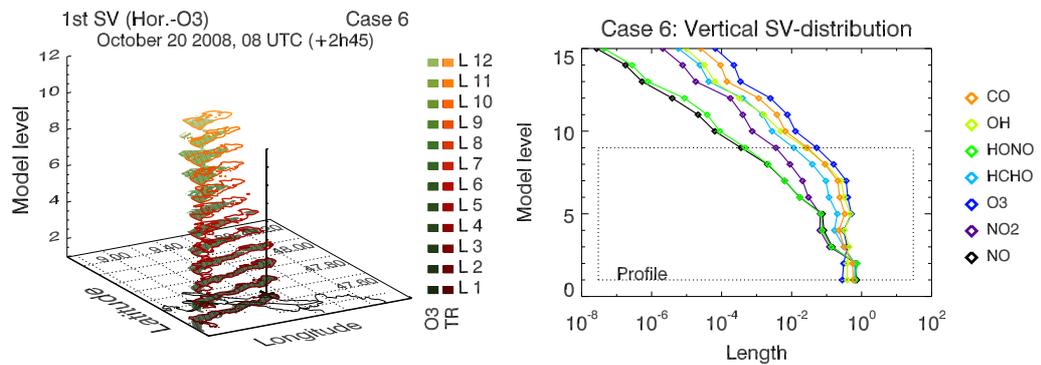
Figure B.11: *Optimal placement for case 7a. Plotting conventions as in Fig. B.1.*
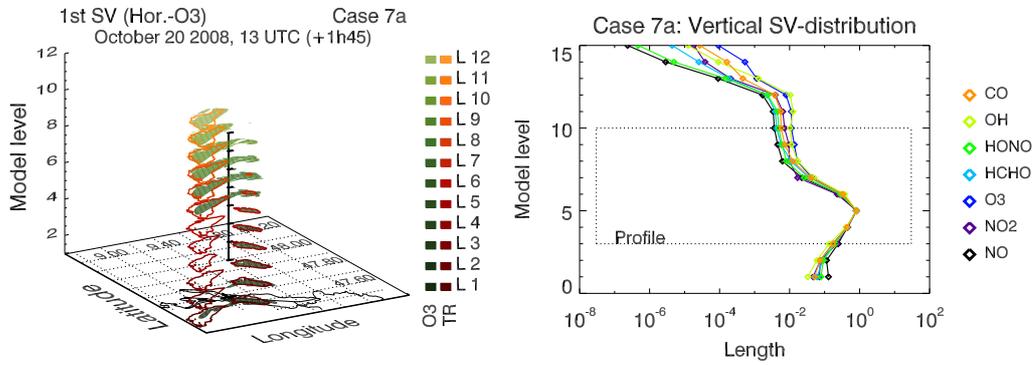
Figure B.12: *Optimal placement for case 7b. Plotting conventions as in Fig. B.1.*
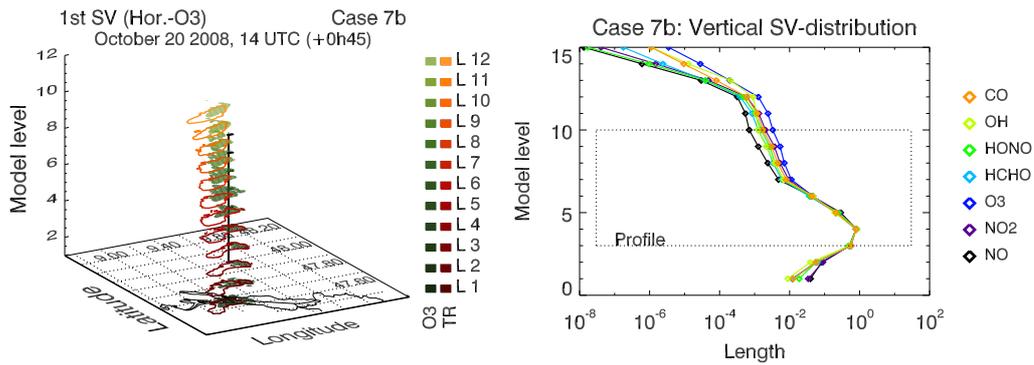
Figure B.13: *Optimal placement for case 8a. Plotting conventions as in Fig. B.1.*
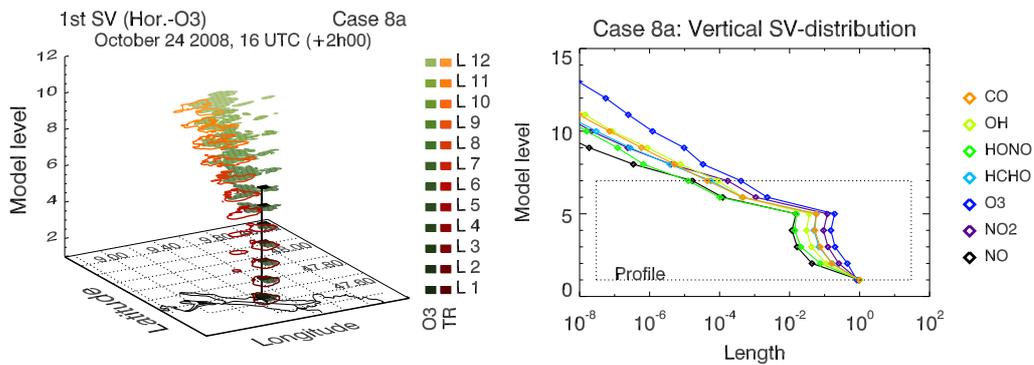
Figure B.14: *Optimal placement for case 8b. Plotting conventions as in Fig. B.1.*
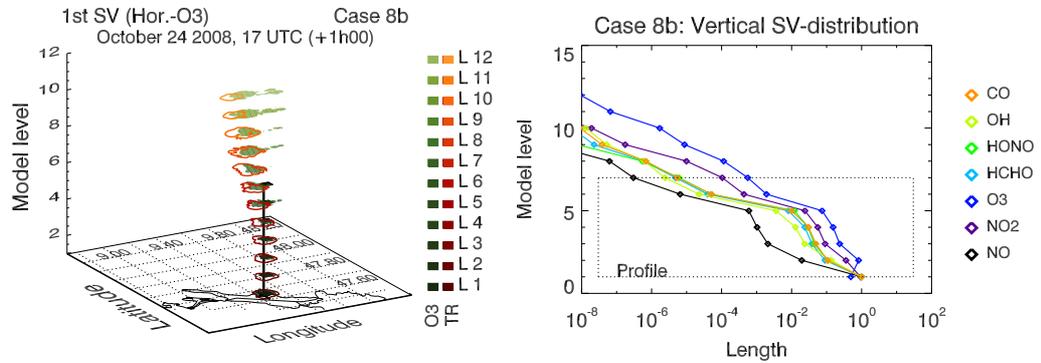
Figure B.15: *Optimal placement for case 9a. Plotting conventions as in Fig. B.1.*
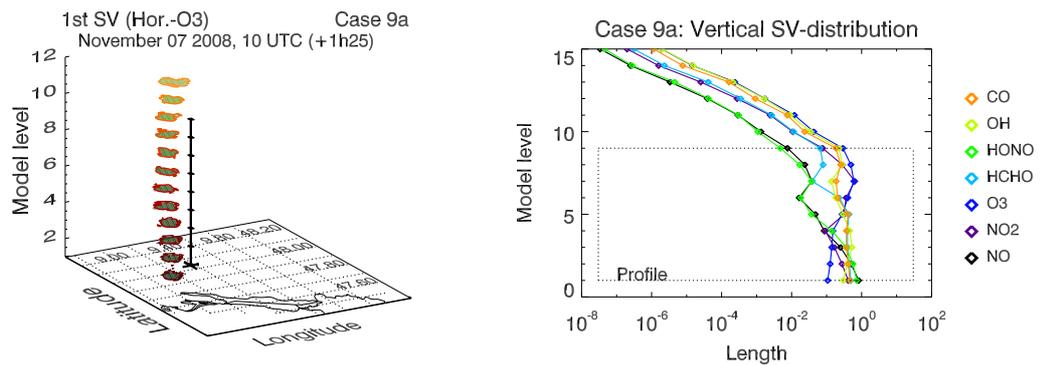
Figure B.16: *Optimal placement for case 9b. Plotting conventions as in Fig. B.1.*
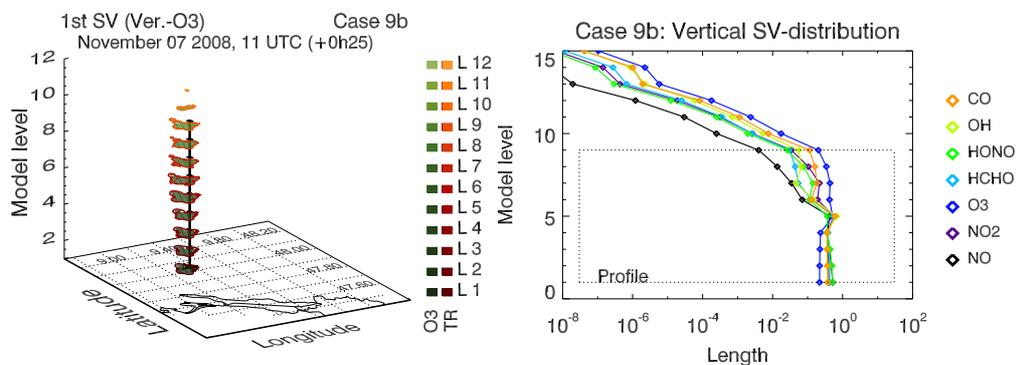
Figure B.17: *Optimal placement for case 10. Plotting conventions as in Fig. B.1.*
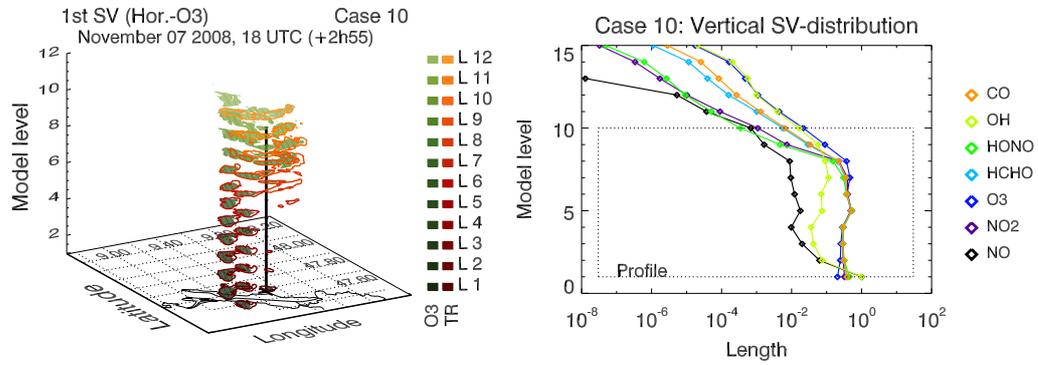
B.2.2   Relevance ranking of chemical compounds

Figure B.18: *Relative ranking of the effect of initial uncertainties of NO (left panel column) and NO_2 (right panel column) for level 1, level 3, level 5, level 7 and level 9 for all 17 cases studies. The optimization criterion are the ozone values of the final profile $VP(t_F)$. Results are sorted by relative ranks (denoted below each bar plot). Rank m is only depicted, if the associated species is ranked m^{th} for at least one considered grid point. Color code of cases is indicated below each panel.*
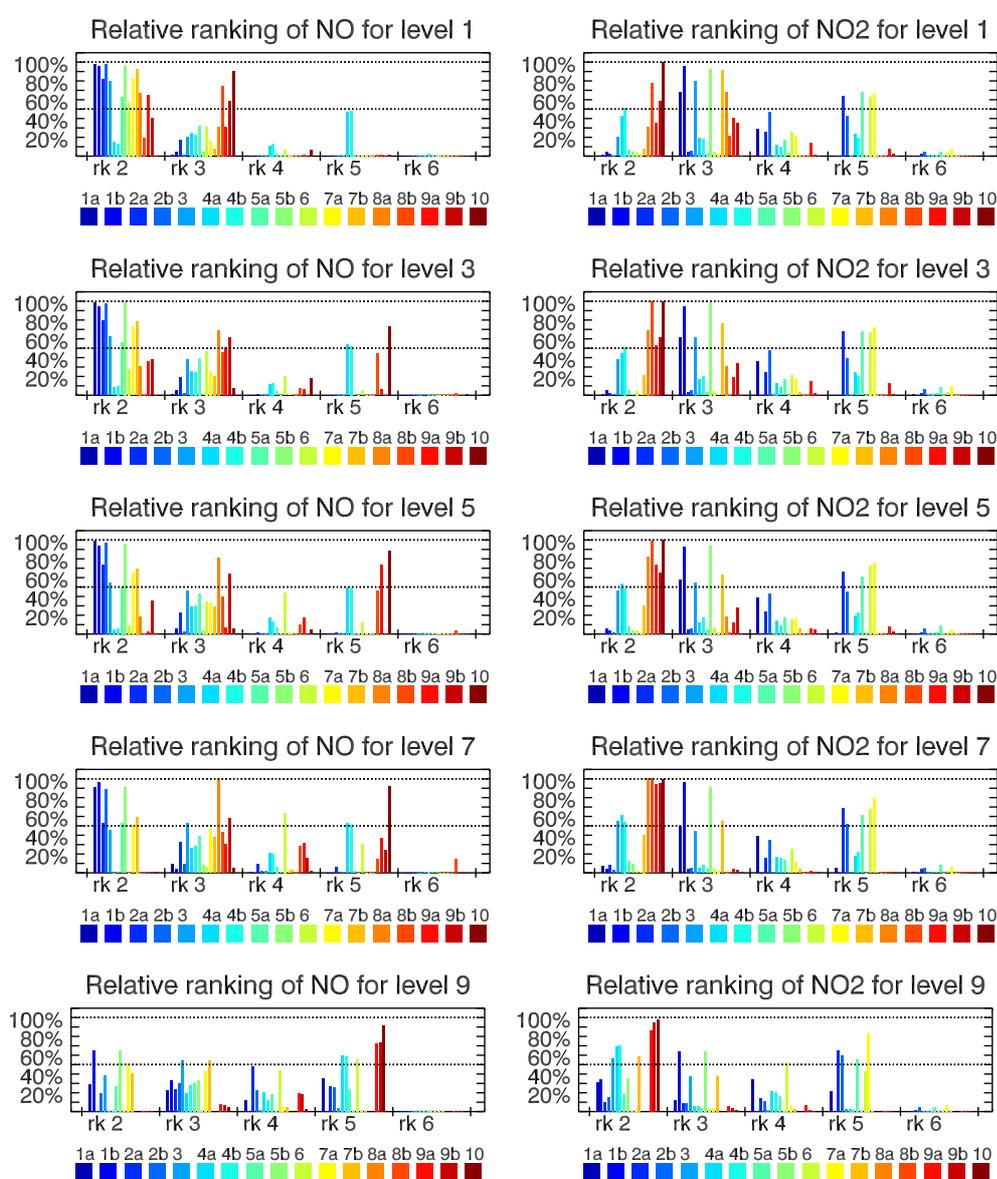
Figure B.19: *Relative ranking of the effect of initial uncertainties of HCHO (left panel column) and CO (right panel column) for level 1, level 3, level 5, level 7 and level 9 for all 17 cases. Plotting conventions as in Fig. B.18.*
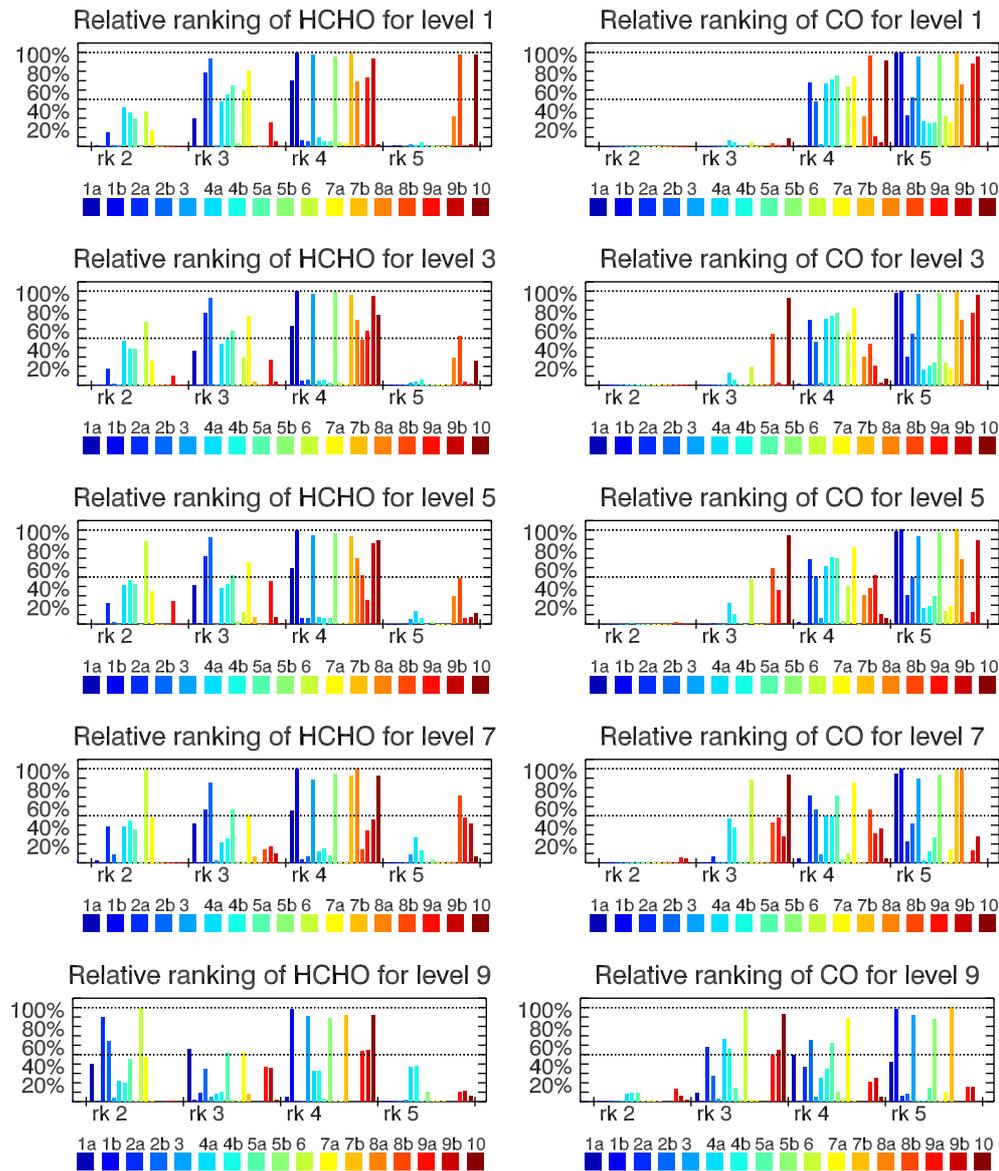
Figure B.20: *Relative ranking of the effect of initial uncertainties of HONO (left panel column) and OH (right panel column) for level 1, level 3, level 5, level 7 and level 9 for all 17 cases. Plotting conventions as in Fig. B.18.*
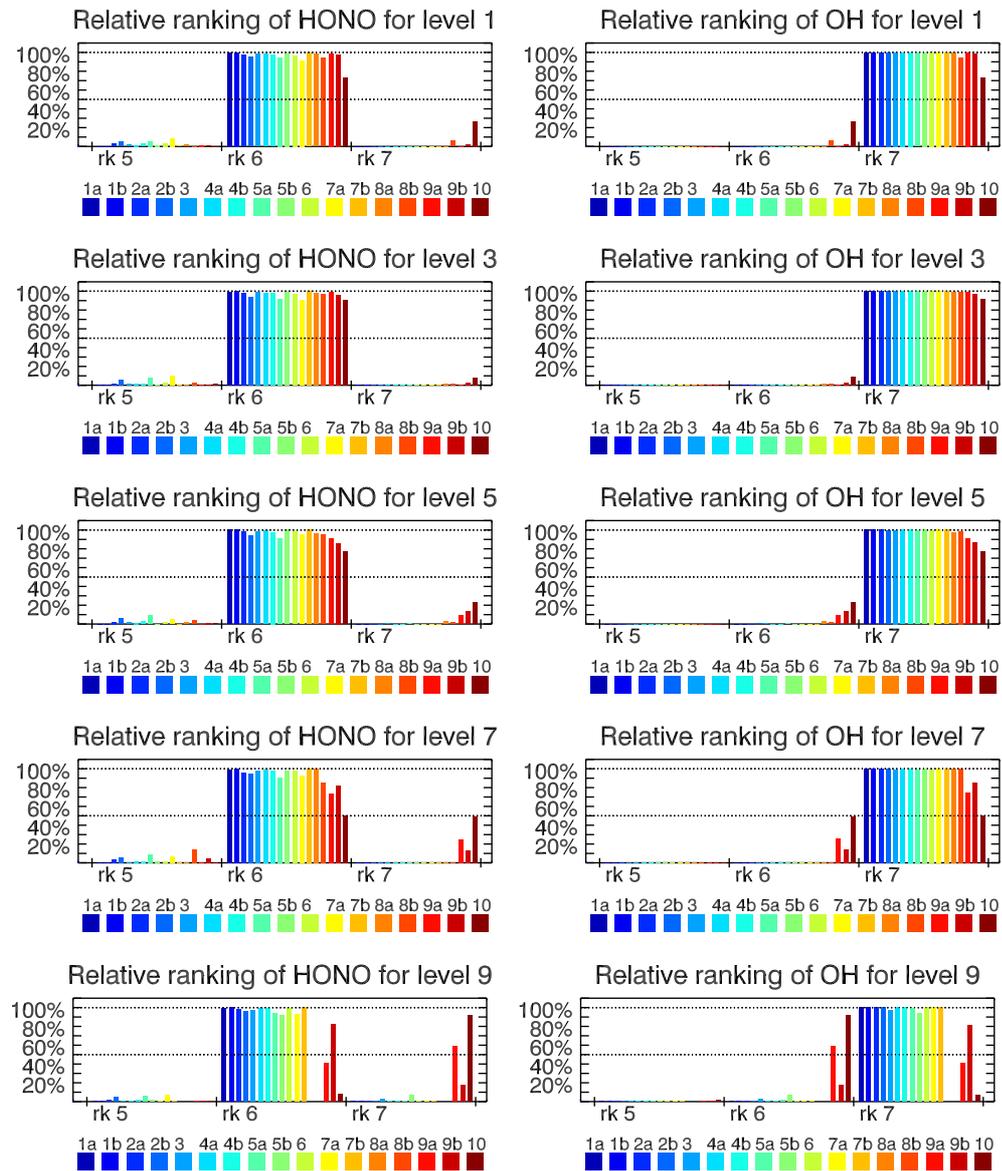
B.3   Singular vectors with respect to emission factors

Figure B.21: *Location dependent relevance rankings of emission influences of HCHO (left column) and CO (right column) at surface level for case 1a and case 1b. Color code of ranks is indicated below each panel. Each grid point (i, j, k) with $\sqrt{\sum_s v(i, j, k, s)^2} > 10^{-4}$ is evaluated. The optimization criterion are the ozone values of the final profile $VP(t_F)$. Its horizontal position is indicated with a red cross. Case number are denoted on top of each panel.*
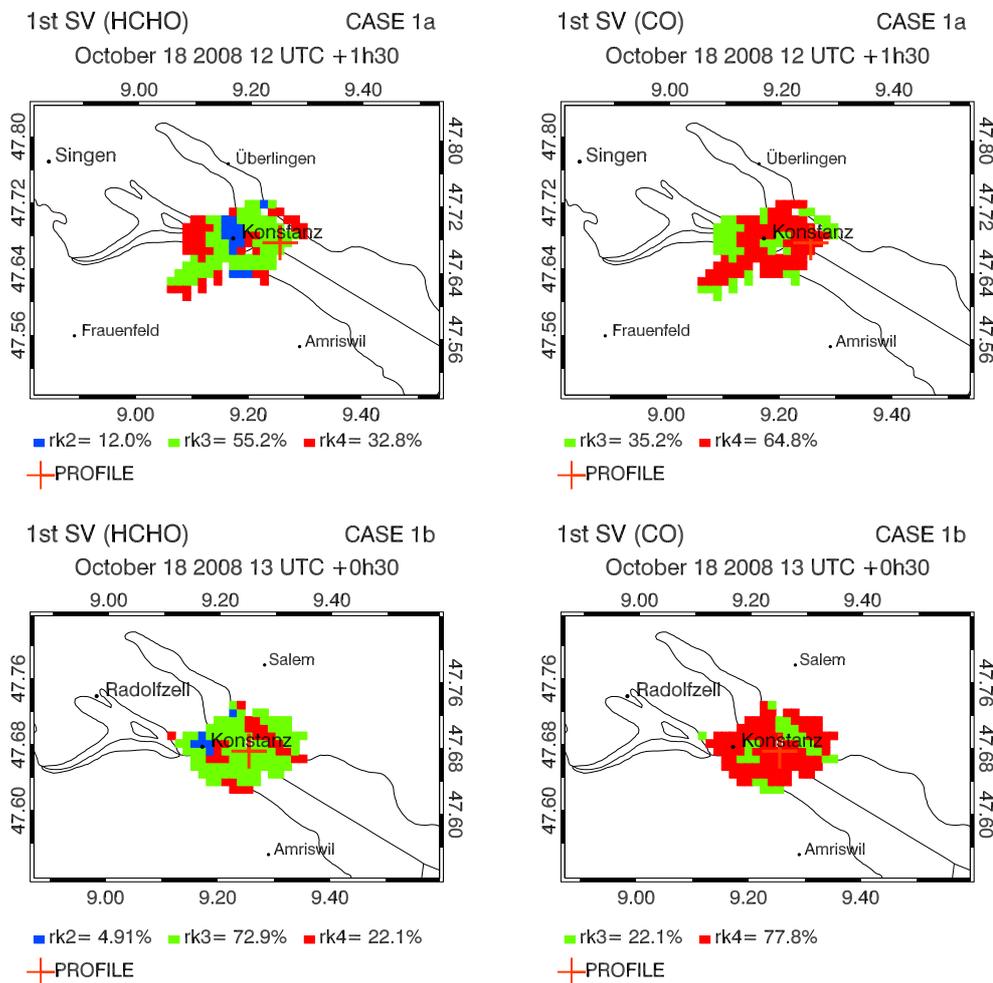
Figure B.22: *Location dependent relevance rankings of emission influences of HCHO (left column) and CO (right column) at surface level for case 2a, case 3 and case 4a. Plotting conventions as in Fig. B.21.*
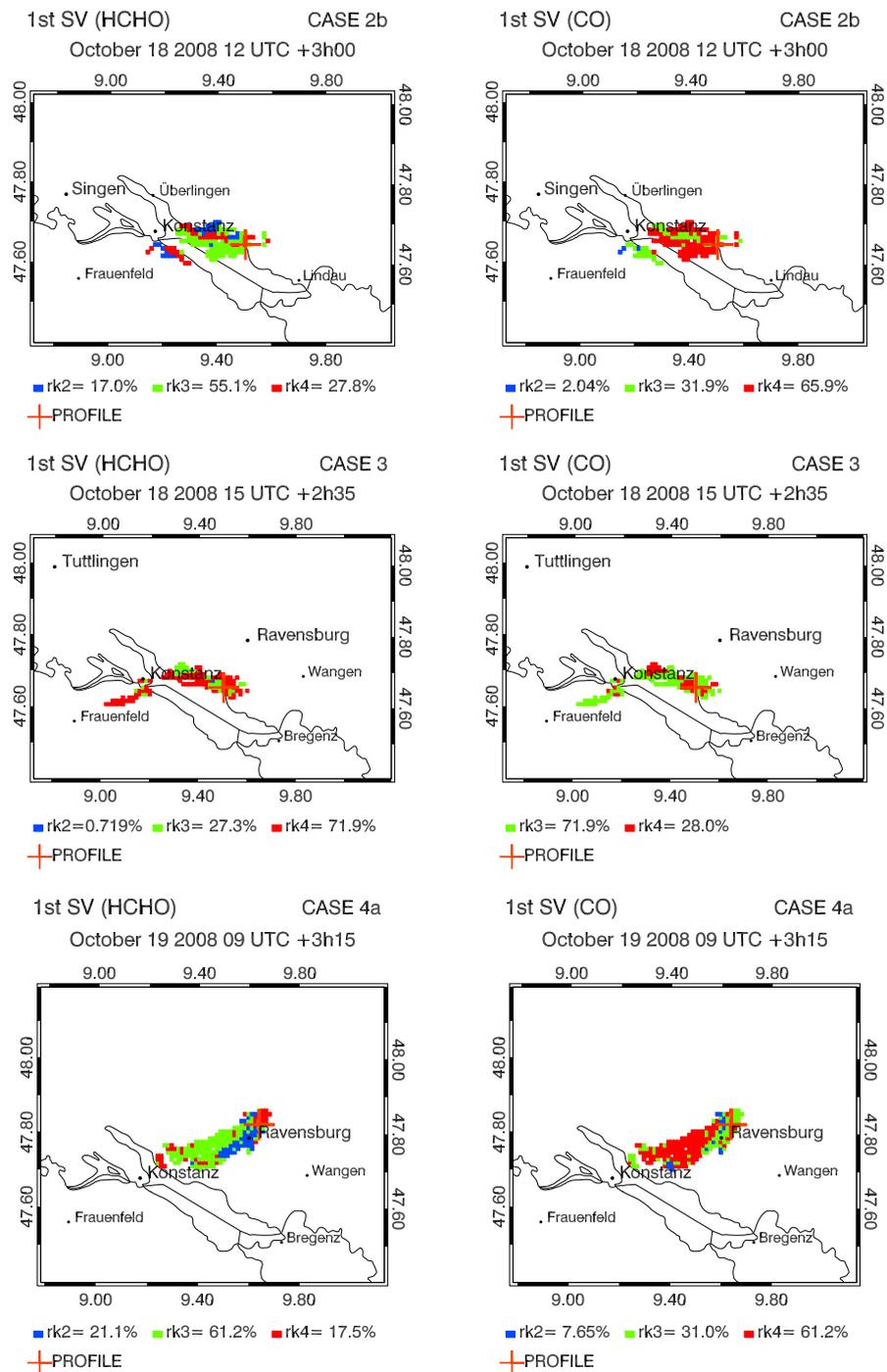
Figure B.23: *Location dependent relevance rankings of emission influences of HCHO (left column) and CO (right column) at surface level for case 4b, case 5a and case 5b. Plotting conventions as in Fig. B.21.*
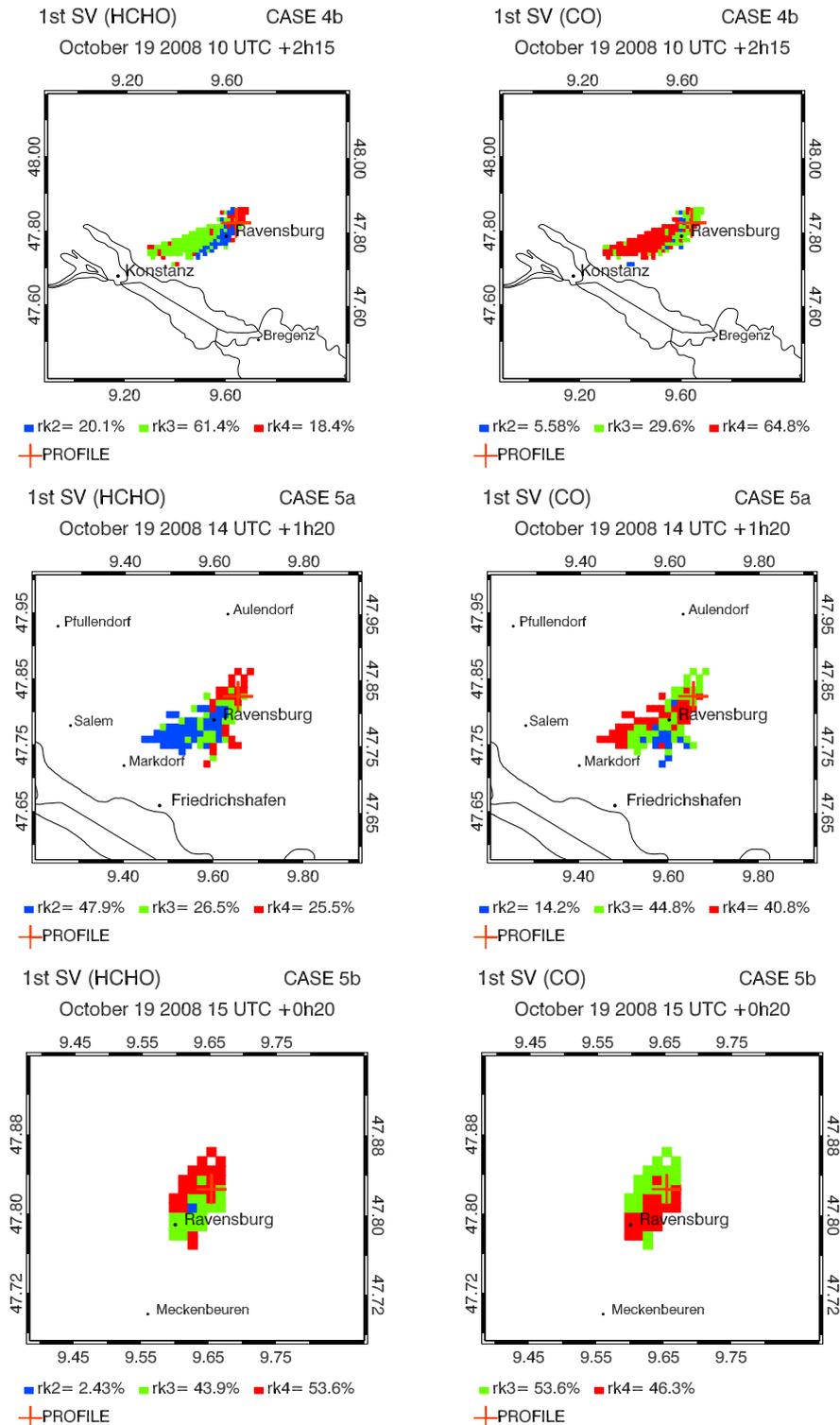
Figure B.24: *Location dependent relevance rankings of emission influences of HCHO (left column) and CO (right column) at surface level for case 6, case 7a and case 7b. Plotting conventions as in Fig. B.21.*
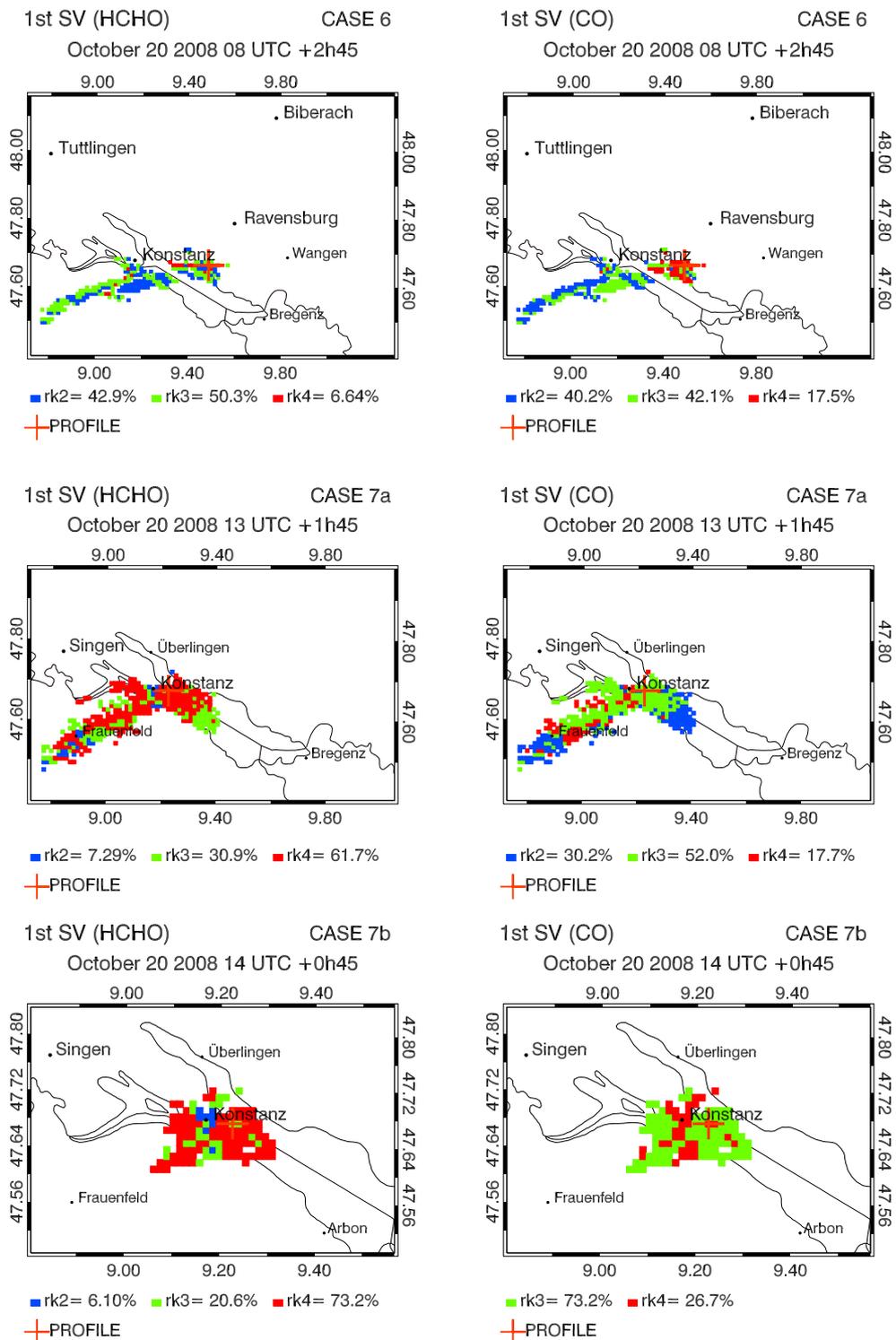
Figure B.25: *Location dependent relevance rankings of emission influences of HCHO (left column) and CO (right column) at surface level for case 8a, case 8b and case 9a. Plotting conventions as in Fig. B.21.*
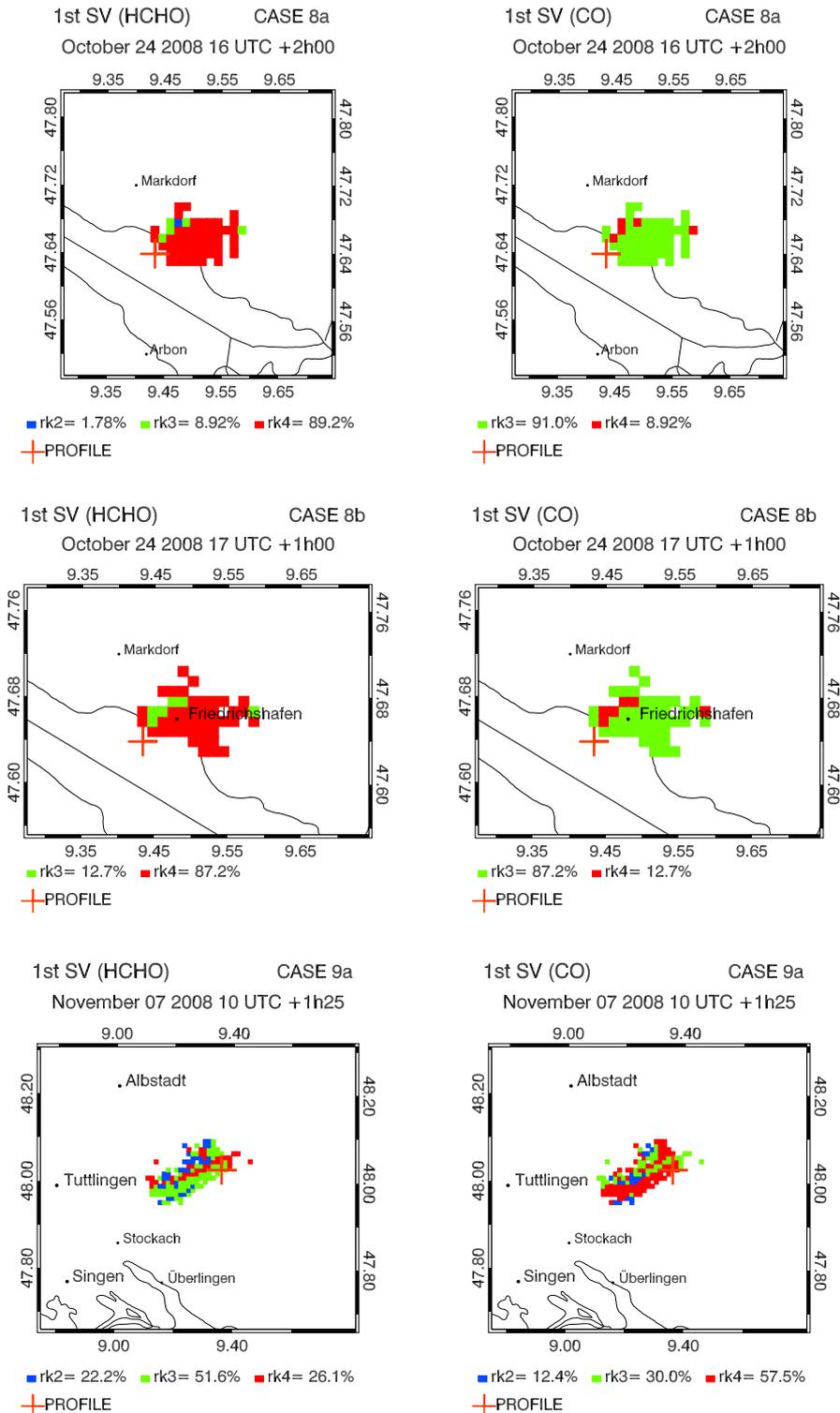
Figure B.26: *Location dependent relevance rankings of emission influences of HCHO (left column) and CO (right column) at surface level for case 9b and case 10. Plotting conventions as in Fig. B.21.*
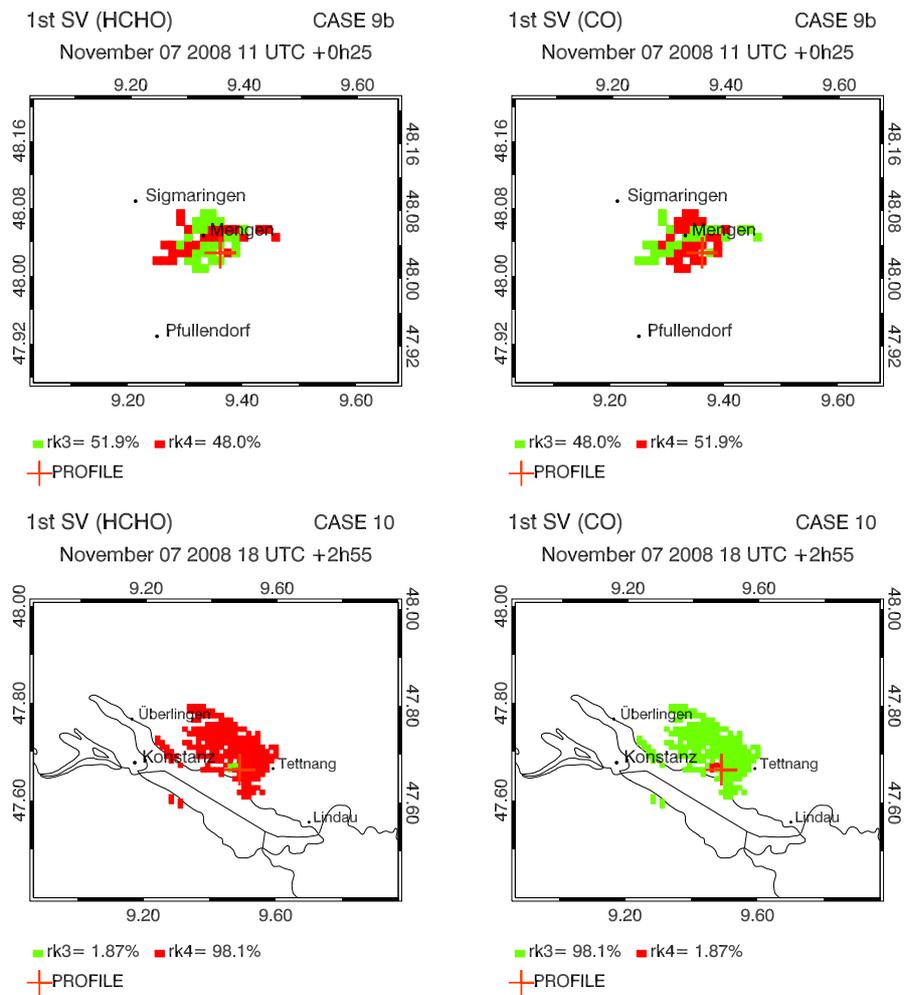
Bibliography

Acton, F. S., *Numerical Methods That Work*, The Mathematical Association of America, 1990.

Arakawa, A. and V. R. Lamb, Computational design for the basic dynamical processes of UCLA general circulation model, *Meth. Comp. Phys.*, *17*, 174–264, 1977.

Barkmeijer, J., M. V. Gijzen and F. Bouttier, Singular vectors and estimates of the analysis-error covariance metric, *Q. J. R. Meteorol. Soc.*, *124*, 1695–1713, 1998.

Berliner, L. M., Z. Lu and C. Snyder, Statistical design for adaptive weather observations, *J. Atmosph. Sci.*, *56*, 2536–2552, 1998.

Bishop, C. H. and Z. Toth, Ensemble transformation and adaptive observations, *J. Atmosph. Sci.*, *56*, 1748–1765, 1998.

Bott, A., A positive definit advection scheme obtained by nonlinear renormalization of the advective fluxes, *Month. Weath. Rev.*, *117*, (5), 1006–1015, 1989.

Buizza, R. and A. Montani, Targeting observations using singular vectors, *J. Atmosph. Sci.*, *56*, (17), 2965–2985, 1999.

Buizza, R. and T. N. Palmer, The singular-vector structure of the atmospheric global circulation, *J. Atmosph. Sci.*, *52*, (9), 1434–1456, 1993.

Buizza, R., C. Cardinali, G. Kelly and J. N. Thepaut, The value of targeted observations, *ECMWF Newsletter*, *111*, 11–20, 2007.

Carter, W. P. L. and F. W. Lurmann, Evaluation of the RADM gas-phase chemical mechanism, Final Report EPA-600/3-90-001, EPA, 1990.

Chang, J. S., R. A. Brost, I. S. A. Isaksen, S. Madronich, P. Middleton, W. R. Stockwell and C. J. Walcek, A three-dimensional Eulerian acid deposition model: physical concepts and formulation, *J. Geophys. Res.*, *92*, (14), 618–700, 1987.

Chao, W. C. and L. P. Chang, Development of a four-dimensional variational analysis system using the adjoint method at GLA Part I: Dynamics, *Month. Weath. Rev.*, *120*, (8), 1661–1673, 1992.

Derwent, R. G. and M.-E. Jenkins, Hydrocarbons and the long-range transport of ozone and PAN across europe, *Atmos. Environ.*, *25A*, 1661–1678, 1991.

Deuflhard, P. and A. Hohmann, *Numerical Analysis in Modern Scientific Computing: An Introduction*, Springer, New York, 2003.

Elbern, H., Parallelization and load balancing of a comprehensive atmospheric chemistry transport model, *Atmos. Environ.*, *13*, 3561 – 3574, 1997.

Elbern, H. and H. Schmidt, A four–dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling, *J. Geophys. Res.*, *104*, (D15), 18583–18598, 1999.

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.

Geiger, H., I. Barnes, I. Bejan, T. Benter and M. Spittler, The tropospheric degradation of isoprene: an updated module for the regional atmospheric chemistry mechanism, *Atmos. Environ.*, *37*, 1503 – 1519, 2003.

Hairer, E. and G. Wanner, *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*, Springer, Berlin, 1991.

Häseler, R., T. Brauers, F. Holland and A. Wahner, Development and application of a new mobile LOPAP instrument for the measurement of HONO altitude profiles in the planetary boundary layer, *Atmos. Meas. Tech. Discuss.*, *2*, 2027 – 2054, 2009.

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

Khattatov, B. V., J. Gille, L. Lyjak, G. Brasseur, V. Dvortsov, A. Roche and J. Waters, Assimilation of photochemically active species and a case analysis of UARS data, *J. Geophys. Res.*, *104*, (D15), 18715–18738, 1999.

Lehoucq, R. B., D. C. Sorensen and C. Yang, *ARPACK Users' Guide: Solution of Large-scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*, SIAM, Philadelphia, 1998.

Lorenz, E. N., A study of the predictability of a 28 variable atmospheric model, *Tellus*, *17*, 321–333, 1965.

Mises, R. V. and H. Pollaczek-Geiringer, Praktische Verfahren der Gleichungsauflösung, *Z. Angew. Math. Mech.*, *9*, 58–77, 1929.

Navon, I. M., X. Zou, J. Derber and J. Sela, Variational data assimilation with an adiabatic version of the NMC spectral model, *Month. Weath. Rev.*, *120*, 1433–1446, 1992.

Noble, B. and J. Daniel, *Applied Linear Algebra*, Prentice Hall, 1. Auflage, 1969.

Parker, T. S. and L. O.Chua, *Practical Numerical Algorithms for Chaotic Systems*, Springer Verlag, 1989.

Poppe, D., B. Aumont, B. Ervens, H. Geiger, H. Herrmann, E.-P. Röth, W. Seidl, W. R. Stockwell, B. Vogel, S. Wagner and D. Weise, Scenarios for modeling multiphase tropospheric chemistry, *J. Atmos. Chem.*, *40*, (1), 77–86, 2001.

Roeth, E. P., Description of the anisotropic radiation transfer model ART to determine photodissociation coefficients, Berichte des Forschungszentrums Jülich JUEL-3960, Jülich, Germany, 1992.

Sandu, A. and R. Sander, Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1, *Atmos. Chem. Phys.*, *6*, 187–195, 2006.

Sandu, A., D. N. Daescu and G. R. Carmichael, Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: Part I – Theory and software tools, *Atmos. Environ.*, *37*, 5083–5096, 2003.

Sandu, A., W. Liao, G. R. Carmichael and T. Chai, Singular vector analysis for atmospheric chemical transport models, *Month. Weath. Rev.*, *134*, (9), 2443–2465, 2006.

Seinfeld, J. H. and S. N. Pandis, *Atmospheric chemistry and physics*, Wiley-Interscience, 1998.

Sorensen, D. C., Implicitly restarted Arnoldi/Lanczos methods for large scale eigenvalue calculations, ICASE Report No. 96-40, Institute for Computer Applications in Science and Engineering, Houston, Texas, 1996.

Stockwell, W. R., P. Middleton and J. S. Chang, The second generation regional acid deposition model chemical mechanism for regional air quality modeling, *J. Geophys. Res.*, *95*, (D10), 16343–16367, 1990.

Stockwell, W. R., F. Kirchner, M. Kuhn and S. Seefeld, A new mechanism for regional atmospheric chemistry modeling, *J. Geophys. Res.*, *102*, (D12), 25847–25879, 1997.

Toth, Z. and E. Kalnay, Ensemble forecasting at NMC: The generation of pertubationsraman study of sulfuric acid at low temperatures, *Bull. Amer. Meteor. Soc.*, *74*, 2317–2330, 1993.

U. Pöschl, R. von Kuhlmann, N. P. and P. J. Crutzen, Development and intercomparison of condensed isoprene oxidation mechanisms for global atmospheric modeling, *J. Atmos. Chem.*, *37*, 29–52, 2000.

Acknowledgements

Erklärung

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

Teilpublikationen:

A. Strunk, A. Ebel, H. Elbern, E. Friese, N. Goris und L. Nieradzik,
Four-dimensional Variational Assimilation of Atmospheric Chemical Data –
Application to Regional Modelling of Air Quality,
in 'Large-Scale Scientific Computing: 7th International Conference', I. Lirkov, S. Margenov, J. Wasniewski (Hrsg.),
4.-8. Juni 2009, Sozopol, Bulgarien, ISBN 978-3-642-12534-8, Springer 2010

H. Elbern, A. Strunk, E. Friese, N. Goris, K. Kasradze und L. Nieradzik,
IMACCO: Inverse Modelling of the Atmospheric Chemical Composition,
in 'NIC Symposium 2010', G. Münster, D. Wolf, M. Kremer (Hrsg.),
24.-25. Februar 2010, Jülich, Germany, IAS Series Volume 3, Jülich, 2010

Lebenslauf

Nadine Goris

Aachener Str. 239

50931 Köln

geb. am 28.10.1979 in Köln

Staatsangehörigkeit: Deutsch

Schulausbildung

1986 – 1990	Grundschule Berchum, Hagen
1990 – 1993	Gymnasium Hohenlimburg, Hagen
1993 – 1999	Gymnasium Kreuztal, Abschluss: Abitur

Hochschulstudium

10/1999 – 09/2000	*Gerhard-Mercator-Universität Duisburg*
	Studiengang: Lehramt Sekundarstufe 2, Fächer: Mathematik, Psychologie
10/2000 – 09/2004	*Westfälische Wilhelms-Universität Münster*
	Studiengang: Mathematik Diplom, Nebenfach Psychologie
09/2004 – 02/2005	*Université Nice Sophia Antipolis*
	Erasmus-Aufenthalt in Nizza, Frankreich
02/2005	Diplom in Mathematik,
	Diplomarbeit: Hinreichende Optimalitätsbedingungen für optimale Steuerprozesse mit Zustandsbeschränkungen und linear auftretender Steuerung: Beispiele aus Medizin und Physik
seit 04/2005	*Universität zu Köln*
	Studiengang: Promotion Meteorologie

Berufliche Tätigkeiten

seit 04/2005	Wissenschaftliche Mitarbeiterin am Rheinischen Institut für Umweltforschung
	Arbeitsgebiet: Sensitivitätsanalyse für chemische Konstituenten