Atmos. Chem. Phys. Discuss., 11, 16745–16799, 2011
www.atmos-chem-phys-discuss.net/11/16745/2011/
doi:10.5194/acpd-11-16745-2011

Atmospheric Chemistry and Physics Discussions

Singular vector decomposition for sensitivity analyses of tropospheric chemical scenarios

N. Goris[1,2,*] **and H. Elbern**[1,2]

[1]Institute of Energy and Climate Research, IEK-8, Research Centre Jülich, Jülich, Germany
[2]Rhenish Institute for Environmental Research at the University of Cologne, Köln, Germany
[*]now at: Geophysical Institute at the University of Bergen, Bergen, Norway

Received: 13 May 2011 – Accepted: 23 May 2011 – Published: 16 June 2011

Correspondence to: N. Goris (nadine.goris@gfi.uib.no)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Observations of the chemical state of the atmosphere typically provide only sparse snapshots of the state of the system due to their insufficient temporal and spatial density. Therefore the measurement configurations need to be optimised to get a best possible state estimate. One possibility to optimise the state estimate is provided by observation targeting of sensitive system states, to identify measurement configurations of best value for forecast improvements. In recent years, numerical weather prediction adapted singular vector analysis with respect to initial values as a novel method to identify sensitive states. In the present work, this technique is transferred from meteorological to chemical forecast. Besides initial values, emissions are investigated as controlling variables. More precisely uncertainties in the amplitude of the diurnal profile of emissions are analysed, yielding emission factors as target variables. Singular vector analysis is extended to allow for projected target variables not only at final time but also at initial time. Further, special operators are introduced, which consider the combined influence of groups of chemical species.

As a preparation for targeted observation calculations, the concept of adaptive observations is studied with a chemistry box model. For a set of six different scenarios, the VOC versus NO_x limitation of the ozone formation is investigated. Results reveal, that the singular vectors are strongly dependent on start time and length of the simulation. As expected, singular vectors with initial values as target variables tend to be more sensitive to initial values, while emission factors as target variables are more sensitive to simulation length. Further, the particular importance of chemical compounds differs strongly between absolute and relative error growth.

ACPD

11, 16745–16799, 2011

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 optimised configuration of available observation capabilities has to be considered to improve the information content of our monitoring capabilities. In the realm of meteorology adaptive observations of sensitive areas can reduce uncertainty and decrease forecast errors (Buizza et al., 2007).

The optimal adaptation 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 Chua, 1989) and singular vectors (Buizza and Palmer, 1993). Most of these methods use a linearised approach to evolve the uncertainties of nonlinear systems in time.

Singular vectors of the tangent linear model can be applied to 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 idealised model by computing the largest error growth. Because of the high computational expenditure, singular vector analyses were applied to realistic meteorological problems not before the late 1980's. Since the singular vectors associated to largest singular values contain the directions of fastest error growth (Buizza and Palmer, 1993), they are applied as reasonable tools to initialise 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

ACPD
11, 16745–16799, 2011

(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.

Motivated by its success in meteorology singular vector analysis is used as analysis method for chemical modelling in this work. In atmospheric chemistry, studies attending targeted observations are rare. The earliest stimulus for analysing uncertainties of the chemical composition was provided by Khattatov et al. (1999). By investigation of the linearised 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 in meteorology the observation location of the standard parameter set wind, temperature, and humidity is of special interest, singular vector analysis for atmospheric chemistry differs in at least two ways:

1. The multitude of reactants enforces analyses which quantifies the species to be observed with preference, given limitations of the observational network and a well defined forecast time.

2. Initial values are not necessarily the only important parameters which are subject

ACPD
11, 16745–16799, 2011

to singular vector analysis. Rather, emissions are at least of like impact on forecasts. While adaptive observations of initial values become less important with growing simulation length, the effect of emission rates on the final concentration increases.

As a first step the present work seeks to give insight into the impact of uncertainties in emission strengths and initial species concentrations in a box model context. Its objective is the detection of sensitive species for atmospheric chemistry models, i.e. to answer the question, which chemical species have to be measured with priority. A follow up study will address the same problem with a full 3-dimensional model.

This paper is organised as follows: The theory of singular vector analysis is presented in Sect. 2, where the application on initial uncertainties and emission factors is described as well as newly introduced special operators. In Sect. 3 the setup of the adapted model is summarised. The zero-dimensional model is applied to analyse several tropospheric scenarios in Sect. 4. Finally, the results of this work are summarised in Sect. 5.

2 Singular vector analysis

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

In this work singular vector analysis is applied to atmospheric chemical modelling to study the influence of chemical initial concentrations and emissions on the temporal evolution of chemical compounds. These two parameters are 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.

ACPD
11, 16745–16799, 2011

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:

$$\boldsymbol{c}(t_F) = \mathcal{M}_{t_I,t_F}\,\boldsymbol{c}(t_I). \tag{1}$$

Chemical models rely on initial values with 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.

Assuming initial errors sufficiently small to evolve linearly within a given time interval, the evolution of initial uncertainties can be described with the tangent linear model $\mathbf{L}_{t_I,t_F}$:

$$\delta\boldsymbol{c}(t_F) = \mathbf{L}_{t_I,t_F}\,\delta\boldsymbol{c}(t_I). \tag{2}$$

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 $g(\delta\boldsymbol{c}(t_I))$:

$$g(\delta\boldsymbol{c}(t_I)) := \frac{\|\delta\boldsymbol{c}(t_F)\|_2}{\|\delta\boldsymbol{c}(t_I)\|_2} \tag{3}$$

(see Sandu et al. (2006) for a comprehensive discussion). Maximising this ratio with respect to the initial disturbance $\delta\boldsymbol{c}(t_I)$ provides the direction of maximal error growth $\delta\boldsymbol{c}_1(t_I)$. As $g(\delta\boldsymbol{c}(t_I)) \geq 0$, the initial perturbation $\delta\boldsymbol{c}_1(t_I)$ which maximises the squared error growth $g^2(\delta\boldsymbol{c}(t_I))$, maximises the error growth $g(\delta\boldsymbol{c}(t_I))$ as well. For convenience the squared error growth is henceforth maximised:

$$\max_{\delta\boldsymbol{c}(t_I)\neq 0} g^2(\delta\boldsymbol{c}(t_I)) = \max_{\delta\boldsymbol{c}(t_I)\neq 0} \frac{\delta\boldsymbol{c}(t_F)^T\,\delta\boldsymbol{c}(t_F)}{\delta\boldsymbol{c}(t_I)^T\,\delta\boldsymbol{c}(t_I)}. \tag{4}$$

ACPD
11, 16745–16799, 2011

Using Eq. (2) the variable $\delta \boldsymbol{c}(t_F)$ may be eliminated from the squared error growth to leave

$$g^2(\delta \boldsymbol{c}(t_I)) = \frac{\delta \boldsymbol{c}(t_I)^T \ \mathbf{L}_{t_I,t_F}^T \mathbf{L}_{t_I,t_F} \ \delta \boldsymbol{c}(t_I)}{\delta \boldsymbol{c}(t_I)^T \delta \boldsymbol{c}(t_I)}. \tag{5}$$

$\mathbf{L}_{t_I,t_F}^T$ denotes the adjoint model of the tangent-linear operator $\mathbf{L}_{t_I,t_F}$. 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 and ratio (5) is a Rayleigh quotient. Applying Rayleigh's principle to problem (4) results in searching for the largest eigenvalue λ_1 and the associated eigenvector $\boldsymbol{v}_1(t_I)$ of the following eigenvalue problem:

$$\mathbf{L}_{t_I,t_F}^T \mathbf{L}_{t_I,t_F} \ \boldsymbol{v}(t_I) = \lambda \boldsymbol{v}(t_I). \tag{6}$$

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 (Palmer, 1995). The influence of eigenvector $\boldsymbol{v}_i(t_I)$ is quantified by the magnitude of the square root of the associated eigenvalue λ_i. The term 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:

$$\mathbf{L}_{t_I,t_F} \boldsymbol{v}_i(t_I) = \sigma_i \boldsymbol{u}_i(t_F) \quad \text{and} \tag{7}$$

$$\mathbf{L}_{t_I,t_F}^T \boldsymbol{u}_i(t_F) = \sigma_i \boldsymbol{v}_i(t_I), \tag{8}$$

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

$$\mathbf{L}_{t_I,t_F}^T \mathbf{L}_{t_I,t_F} \boldsymbol{v}_i(t_I) = \sigma_i \mathbf{L}_{t_I,t_F}^T \boldsymbol{u}_i(t_F) = \sigma_i^2 \boldsymbol{v}_i(t_I) \tag{9}$$

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 $\boldsymbol{v}_1(t_I)$

ACPD
11, 16745–16799, 2011

assigned to the largest singular value σ_1 of a chemistry-transport model characterises 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 (3) and defines the amount of error growth.

2.1.1 Application of special operators

In order to address specific questions of atmospheric chemistry, special operators are applied to the initial and final perturbations, namely the weight matrix $\mathbf{W}_t$, the projection operator $\mathbf{P}_t$ and the family operator $\mathbf{F}_t$.

The weight matrix $\mathbf{W}_t \in \mathbb{R}^{n\times n}$ contains the concentrations of all considered chemical species s, $s = 1,\ldots,n$ at time t:

$$\mathbf{W}_t := \operatorname{diag}\left(\boldsymbol{c}^s(t)\right)_{s=1,\ldots,n} \quad \forall\, t \in [t_I, t_F]. \tag{10}$$

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. Application of the inverse weight matrix avoids this effect by scaling the absolute uncertainties $\delta\boldsymbol{c}(t)$ with current concentrations $\boldsymbol{c}(t)$. It provides the relative error $\delta\boldsymbol{c}_{\mathrm{r}} \in \mathbb{R}^n$

$$\delta\boldsymbol{c}_{\mathrm{r}}(t) := \mathbf{W}_t^{-1}\,\delta\boldsymbol{c}(t) \quad \forall\, t \in [t_I, t_F] \tag{11}$$

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

$$g_{\mathrm{r}}(\delta\boldsymbol{c}_{\mathrm{r}}(t_I)) := \frac{\|\delta\boldsymbol{c}_{\mathrm{r}}(t_F)\|_2}{\|\delta\boldsymbol{c}_{\mathrm{r}}(t_I)\|_2} \tag{12}$$

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

$$\delta\boldsymbol{c}_{\mathrm{r}}(t_F) = \mathbf{W}_{t_F}^{-1}\,\mathbf{L}_{t_I,t_F}\,\mathbf{W}_{t_I}\,\delta\boldsymbol{c}_{\mathrm{r}}(t_I) \tag{13}$$

leads to a Rayleigh quotient

ACPD
11, 16745–16799, 2011

$$g_r^2(\delta \boldsymbol{c}(t_l)) = \frac{\delta \boldsymbol{c}_r(t_l)^T \ \mathbf{B}_r^T \mathbf{B}_r \ \delta \boldsymbol{c}_r(t_l)}{\delta \boldsymbol{c}_r(t_l)^T \delta \boldsymbol{c}_r(t_l)}, \tag{14}$$

$$\text{with} \quad \mathbf{B}_r := \mathbf{W}_{t_F}^{-1} \ \mathbf{L}_{t_l,t_F} \ \mathbf{W}_{t_l}.$$

According to Rayleigh's principle the phase space direction $\delta \boldsymbol{c}_r(t_l)$, for which the ratio (14) gains its maximal value, is the solution $\boldsymbol{v}_{r_1}(t_l)$ of the symmetric eigenvalue problem

$$\mathbf{B}_r^T \mathbf{B}_r \ \boldsymbol{v}_r(t_l) = \lambda_r \boldsymbol{v}_r(t_l) \tag{15}$$

$$\text{with} \quad \boldsymbol{v}_r(t_l) := \mathbf{W}_{t_l}^{-1} \boldsymbol{v}(t_l)$$

assigned to the largest eigenvalue λ_{r_1}. Comparing problem (15) with the original problem (6), it is readily seen that the solution $\boldsymbol{v}_{r_1}(t_l) \in \mathbb{R}^n$ is the right singular vector of the operator $\mathbf{B}_r$ and the square root of the eigenvalue λ_{r_1} is the associated singular value σ_{r_1}. Due to this property the singular vector $\boldsymbol{v}_{r_1}(t_l)$ is called relative singular vector henceforth.

Another central aim is to examine the error growth of a limited set of chemical species. To fulfil this need 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 feature of interest (Barkmeijer et al., 1998). Thus the projection operator is a diagonal matrix with binary entries. In case of a chemical box model the projection operator reads

$$\mathbf{P}_t := \mathrm{diag}(p_i)_{i=1,\ldots,n}, \quad p_i = \begin{cases} 1 & \forall\, i \in \mathcal{P}(t) \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

where $\mathcal{P}(t)$ denotes the set of selected chemical compounds. Instead of examining the initial disturbances individually, one can also decide to look at the influence of groups of chemical species, which act chemically in a similar manner. To the knowledge of the authors 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 $\mathcal{F}_k(t), k = 1,\ldots,f(t)$ represent one of $f(t)$ pairwise disjoint families with

ACPD
11, 16745–16799, 2011

$m_k(t)$ 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

$$\mathbf{F}_t(i,j)_{i,j} = \begin{cases} 1/m_k(t) & \forall\, i,j \in \mathcal{F}_k(t) \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

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

$$\bigcap_{k=1}^{f(t)} \mathcal{F}_k(t) = \varnothing. \tag{18}$$

Since the family operator is time dependent, the user can focus on distinct families at different times. An example of a family operator at initial time t_l is

$$\mathbf{F}_{t_l} = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \end{pmatrix}, \tag{19}$$

where $n=6$, $f(t_l)=2$, $m_1(t_l)=3$, $m_2(t_l)=2$, $\mathcal{F}_1(t_l)=\{2,4,6\}$ and $\mathcal{F}_2(t_l)=\{1,5\}$. If each family has only one member, the family operator equals the projection operator with $\mathcal{P}(t)=\bigcup_{k=1}^{f(t)}\mathcal{F}_k(t)$. Therefore, the projected error can be regarded as a special case of the grouped error. For expository purposes we restrict attention to the application of the family operator, since applications of the projection operator proceed in the same manner.

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

$$\delta \boldsymbol{c}_{\mathrm{g}}(t) := \mathbf{F}_t\, \mathbf{L}_{t_l,t}\, \mathbf{F}_{t_l}\, \delta \boldsymbol{c}(t_l). \tag{20}$$

ACPD
11, 16745–16799, 2011

If no chemical group is considered at time $t \in [t_I, t_F]$, the family operator $\mathbf{F}_t$ equals the identity matrix $\mathbf{I}_n$. With the aid of the idempotence of the family operator (i.e. $\mathbf{F}_t^2 = \mathbf{F}_t$) and $\mathbf{L}_{t,t} = \mathbf{I}_n$, the grouped error reduces at initial time to

$$\delta \boldsymbol{c}_{\mathrm{g}}(t_I) = \mathbf{F}_{t_I}\, \delta \boldsymbol{c}(t_I). \tag{21}$$

Substituting the grouped error (20) into the measure of error growth gives

$$g_{\mathrm{g}}(\delta \boldsymbol{c}_{\mathrm{g}}(t_I)) := \frac{\|\delta \boldsymbol{c}_{\mathrm{g}}(t_F)\|_2}{\|\delta \boldsymbol{c}_{\mathrm{g}}(t_I)\|_2}. \tag{22}$$

After use of

$$\delta \boldsymbol{c}_{\mathrm{g}}(t_F) = \mathbf{F}_{t_F}\, \mathbf{L}_{t_I,t_F}\, \delta \boldsymbol{c}_{\mathrm{g}}(t_I), \tag{23}$$

the squared ratio takes on the form of a Rayleigh quotient

$$g_{\mathrm{g}}{}^2(\delta \boldsymbol{c}(t_I)) = \frac{\delta \boldsymbol{c}_{\mathrm{g}}(t_I)^T\ \mathbf{B}^T \mathbf{B}\ \delta \boldsymbol{c}_{\mathrm{g}}(t_I)}{\delta \boldsymbol{c}_{\mathrm{g}}(t_I)^T\, \delta \boldsymbol{c}_{\mathrm{g}}(t_I)}, \tag{24}$$

$$\text{with} \quad \mathbf{B} := \mathbf{F}_{t_F}\, \mathbf{L}_{t_I,t_F},$$

subject to condition

$$[\delta \boldsymbol{c}_{\mathrm{g}}(t_I)](j) = \begin{cases} \frac{1}{m_k} \sum_{i \in \mathcal{F}_k(t_I)} [\delta \boldsymbol{c}(t_I)](i) & \forall\, j \in \mathcal{F}_k(t) \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

Here, $[\boldsymbol{x}](j)$ denotes the j-th component of a vector $\boldsymbol{x}$. The eigenvector $\boldsymbol{v}_{\mathrm{g}_1}(t_I)$ of the eigenvalue problem

$$\mathbf{B}^T \mathbf{B}\, \boldsymbol{v}_{\mathrm{g}}(t_I) = \lambda_{\mathrm{g}}\, \boldsymbol{v}_{\mathrm{g}}(t_I) \tag{26}$$

assigned to the largest eigenvalue λ_{g_1} is the phase space direction which results in a maximum Rayleigh quotient (24). However, the solution $\boldsymbol{v}_{\mathrm{g}_1}(t_I)$ does not necessarily

ACPD
11, 16745–16799, 2011

ensure condition (25). In order to grant condition (25), the solution space has to be restricted. Multiplying Eq. (26) with $\mathbf{F}_{t_I}$ and applying

$$\mathbf{F}_{t_I}\, \boldsymbol{v}_{\mathrm{g}}(t_I) = \boldsymbol{v}_{\mathrm{g}}(t_I), \tag{27}$$

regroups Eq. (26) into the following equivalent eigenvalue problem:

$$\mathbf{B}_{\mathrm{g}}{}^T\, \mathbf{B}_{\mathrm{g}}\, \boldsymbol{v}_{\mathrm{g}}(t_I) = \lambda_{\mathrm{g}}\, \boldsymbol{v}_{\mathrm{g}}(t_I), \tag{28}$$

$$\text{with} \quad \mathbf{B}_{\mathrm{g}} := \mathbf{F}_{t_F}\, \mathbf{L}_{t_I,t_F}\, \mathbf{F}_{t_I}.$$

The solutions of the new eigenvalue problem hold condition (25). Henceforth, the eigenvector $\boldsymbol{v}_{\mathrm{g}_1}(t_I) \in \mathbb{R}^n$ assigned to largest eigenvalue λ_{g_1} of Eq. (28) is denoted as grouped singular vector, since it is the right singular vector of the operator $\mathbf{B}_{\mathrm{g}}$. Accordingly, the square root of the eigenvalue λ_{g_1} is the associated grouped singular value σ_{g_1}.

The grouped error (20) can be combined with the relative error (11), leading to the grouped relative error $\delta \boldsymbol{c}_{\mathrm{gr}}(t) \in \mathbb{R}^n$. 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

$$\delta \boldsymbol{c}_{\mathrm{gr}}(t) := \mathbf{F}_t\, \mathbf{W}_t^{-1}\, \mathbf{L}_{t_I,t}\, \mathbf{W}_{t_I}\, \mathbf{F}_{t_I}\, \mathbf{W}_{t_I}^{-1}\, \delta \boldsymbol{c}(t_I). \tag{29}$$

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

$$\delta \boldsymbol{c}_{\mathrm{gr}}(t_I) = \mathbf{F}_{t_I}\, \mathbf{W}_{t_I}^{-1}\, \delta \boldsymbol{c}(t_I). \tag{30}$$

Hence, the grouped relative error at final time t_F reads

$$\delta \boldsymbol{c}_{\mathrm{gr}}(t_F) = \mathbf{F}_{t_F}\, \mathbf{W}_{t_F}^{-1}\, \mathbf{L}_{t_I,t_F}\, \mathbf{W}_{t_I}\, \delta \boldsymbol{c}_{\mathrm{gr}}(t_I). \tag{31}$$

Inserting formula (31) into the squared grouped relative error growth

$$g_{\mathrm{gr}}{}^2(\delta \boldsymbol{c}_{\mathrm{gr}}(t_I)) := \frac{\|\delta \boldsymbol{c}_{\mathrm{gr}}(t_F)\|_2^2}{\|\delta \boldsymbol{c}_{\mathrm{gr}}(t_I)\|_2^2} \tag{32}$$

ACPD
11, 16745–16799, 2011

gives a Rayleigh quotient

$$g_{\mathrm{gr}}{}^2(\delta \boldsymbol{c}(t_l)) = \frac{\delta \boldsymbol{c}_{\mathrm{gr}}(t_l)^T \ \mathbf{B}^T \mathbf{B} \ \delta \boldsymbol{c}_{\mathrm{gr}}(t_l)}{\delta \boldsymbol{c}_{\mathrm{gr}}(t_l)^T \delta \boldsymbol{c}_{\mathrm{gr}}(t_l)}, \tag{33}$$

$$\text{with} \quad \mathbf{B} := \mathbf{F}_{t_F} \mathbf{W}_{t_F}^{-1} \mathbf{L}_{t_l,t_F} \mathbf{W}_{t_l},$$

subject to condition

$$[\delta \boldsymbol{c}_{\mathrm{g}}(t_l)](j) = \begin{cases} \frac{1}{m_k} \sum_{i \in \mathcal{F}_k(t_l)} \frac{[\delta \boldsymbol{c}(t_l)](i)}{[\boldsymbol{c}(t_l)](i)} & \forall\, j \in \mathcal{F}_k(t) \\ 0 & \text{otherwise.} \end{cases} \tag{34}$$

According to Rayleigh's principle, the phase space direction that maximises the Rayleigh quotient (33) is the solution $\boldsymbol{v}_{\mathrm{gr}_1}(t_l)$ of the symmetric eigenvalue problem

$$\mathbf{B}^T \mathbf{B} \ \boldsymbol{v}_{\mathrm{gr}}(t_l) = \lambda_{\mathrm{gr}} \boldsymbol{v}_{\mathrm{gr}}(t_l), \tag{35}$$

which is assigned to the largest eigenvalue λ_{gr_1}. This solution does not necessarily satisfy condition (34). In order to comply with condition (34), the eigenvalue problem (35) is rearranged. Making use of the idempotence of the family operator, it can readily be seen that

$$\mathbf{F}_{t_l} \boldsymbol{v}_{\mathrm{gr}}(t_l) = \boldsymbol{v}_{\mathrm{gr}}(t_l). \tag{36}$$

Multiplying Eq. (35) with $\mathbf{F}_{t_l}$ and substituting relation (36) gives

$$\mathbf{B}_{\mathrm{gr}}{}^T \mathbf{B}_{\mathrm{gr}} \ \boldsymbol{v}_{\mathrm{gr}}(t_l) = \lambda_{\mathrm{gr}} \ \boldsymbol{v}_{\mathrm{gr}}(t_l) \tag{37}$$

$$\text{with} \quad \mathbf{B}_{\mathrm{gr}} := \mathbf{F}_{t_F} \mathbf{W}_{t_F}^{-1} \mathbf{L}_{t_l,t_F} \mathbf{W}_{t_l} \mathbf{F}_{t_l}.$$

The new symmetric eigenvalue problem ensures a feasible solution and holds condition (34). The vector $\boldsymbol{v}_{\mathrm{gr}_1}(t_l) \in \mathbb{R}^n$ assigned to the largest eigenvalue λ_{gr_1} of Eq. (37) 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}.

ACPD
11, 16745–16799, 2011

For explanatory purposes, Fig. 1 gives an schematic example for grouped relative singular vectors. Here, the initial family operator focuses on families NO_x (Nitrogen Oxides) and VOC (Volatile Organic Compounds), while the final family operator focuses on O_3 alone. Assuming that each member of a family has the same initial relative perturbation, the associated grouped relative singular vector shows the direction of largest error. For clarification, the first singular vector is depicted on forward model runs with varying initial concentrations of VOC and NO_x, respectively.

2.2 Uncertainties of emissions

For skillful predictions, emissions are at least as important as initial values (Elbern et al., 2007). This section generalises the application of singular vector analysis to allow for sensitivities with respect to emission rates. In contrast to initial concentrations, the total emitted amount of species s varies in time. Due to this variation of emissions in time, the associated directions of largest error growth differ for each time step $t \in [t_I, t_F]$. Consequently, their identification results in one application of singular vector analysis per time step $t \in [t_I, t_F]$, yielding an ill-posed problem with high computational expenditure. By dealing with a time invariant emission factor e_{f_i} for species i, the set of singular vector analyses can be reduced to a single singular vector analysis 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 analysed. This choice is reasonable, as the relative 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 analyses 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:

$$\frac{d\boldsymbol{c}}{dt} = f(\boldsymbol{c}(t)) + \boldsymbol{e}(t). \tag{38}$$

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

ACPD
11, 16745–16799, 2011

$$\frac{d\boldsymbol{c}}{dt} = f(\boldsymbol{c}(t)) + \mathbf{E}(t)\boldsymbol{e}_f, \tag{39}$$

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 (38) describes the linear part of the evolution of perturbation $\delta\boldsymbol{c}(t_F)$ caused by uncertainties $\delta\boldsymbol{e}_f$ and $\delta\boldsymbol{c}(t_I)$:

$$\delta\boldsymbol{c}(t_F) = \tilde{\mathbf{L}}_{t_I,t_F}\,\delta z(t_I), \tag{40}$$

where

$$\tilde{\mathbf{L}}_{t_I,t_F} := \left(\mathbf{L}_{t_I,t_F}, \mathbf{L}^{\boldsymbol{e}}_{t_I,t_F}\right), \tag{41}$$
$$\delta z(t_I) := (\delta\boldsymbol{c}(t_I), \delta\boldsymbol{e}_f)^T.$$

For expository purposes, the perturbation is merely induced by emissions uncertainties ($\delta\boldsymbol{c}(t_I) = 0$). 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:

$$\delta\tilde{\boldsymbol{c}}_{\mathrm{r}}(t_F) := \mathbf{W}^{-1}_{t_F}\,\delta\tilde{\boldsymbol{c}}(t_F). \tag{42}$$

Here, tilde denotes that the perturbation is induced by emissions uncertainties. Since the considered final uncertainty contains only perturbations in concentrations of species, the original weight matrix $\mathbf{W}_t$ (10) is applied. From definition (42), the relative error growth is obtained:

$$\tilde{g}_{\mathrm{r}}(\delta\boldsymbol{e}_f) := \frac{\|\delta\tilde{\boldsymbol{c}}_{\mathrm{r}}(t_F)\|_2}{\|\delta\boldsymbol{e}_f\|_2}. \tag{43}$$

Considering the squared ratio and applying Eq. (40) yields

$$\tilde{g}_{\mathrm{r}}{}^2(\delta\boldsymbol{e}_f) = \frac{\delta\boldsymbol{e}_f^T\,\tilde{\mathbf{B}}_{\mathrm{r}}{}^T\tilde{\mathbf{B}}_{\mathrm{r}}\,\delta\boldsymbol{e}_f}{\delta\boldsymbol{e}_f^T\,\delta\boldsymbol{e}_f}, \tag{44}$$

ACPD
11, 16745–16799, 2011

with $\tilde{\mathbf{B}}_r := \mathbf{W}_{t_F}^{-1}\, \mathbf{L}^{e}_{t_I,t_F}$.

The phase space direction that maximises the ratio (44) is the eigenvector $\tilde{\boldsymbol{v}}_{r_1}$ of the eigenvalue problem

$$\tilde{\mathbf{B}}_r^T\, \tilde{\mathbf{B}}_r\ \tilde{\boldsymbol{v}}_r = \tilde{\lambda}_r\ \tilde{\boldsymbol{v}}_r, \tag{45}$$

associated to largest eigenvalue $\tilde{\lambda}_{r_1}$. As the solution gives the right singular vector of operator $\tilde{\mathbf{B}}_r$, it is henceforth denoted as relative singular vector with respect to emission uncertainties.

A special set of perturbation norms can be provided with the help of the projection operator $\mathbf{P}_t$ (16) and the family operator $\mathbf{F}_t$ (17). 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 Sect. 2.1.1.

3 Model design

In order to calculate singular vectors a chemical box model, the software package Kinetic PreProcessor (KPP, Sandu and Sander, 2006) is applied. KPP offers a set of numerical integrators, thereof the second order Rosenbrock method is chosen. Since Rosenbrock solvers have optimal stability properties and conserve the linear invariants of the system (Hairer and Wanner, 1991), they are well-suited for atmospheric chemistry applications. For Rosenbrock models KPP includes tangent linear and two different adjoint code generations (discrete and continuous) with respect to initial values. Here, the discrete adjoint model is our adjoint model of choice, since it offers a much better consistency with the forward and tangent-linear model than the continuous adjoint model. The tangent linear as well as the adjoint model with respect to emissions has been coded by hand.

Tangent linear routines have been checked for consistency with corresponding forward and adjoint modules. For consistency with the forward model, the gradient check

ACPD
11, 16745–16799, 2011

ratio (Navon et al., 1992) is applied, defined as

$$d = \frac{\mathrm{FWD}(x + \alpha\delta x) - \mathrm{FWD}(x)}{\mathrm{TLM}(\alpha\delta x)}. \quad (46)$$

The abbreviations FWD and TLM denote the forward and tangent linear model. Respectively α is a scalar parameter. While α approaches zero, ratio (46) should approach one. Note that rounding errors dominate in cases, where the magnitude of α is sufficiently small. Here, the limits of numerical precision are approached and results degrade. Within numerical limits, the new tangent linear routines demonstrate the required characteristics for Eq. (46) for considered test cases. Besides consistency with the forward model, the gradient ratio check indicates the accuracy of the tangent linear approximation. Application of the tangent linear model is only justified, if the considered perturbation is small enough to ensure $d \approx 1$.

For consistency of tangent linear and adjoint model, Navon et al. (1992) suggest to test whether the following equation holds:

$$(\mathrm{TLM}(\delta x))^T(\mathrm{TLM}(\delta x)) = \delta x^T \mathrm{ADJ}(\mathrm{TLM}(\delta x)), \quad (47)$$

where ADJ denotes the adjoint model. Again, the new routines demonstrate good performance.

Two methods have been implemented for solving the eigenvalue problems, namely the power method and the implicitly restarted Arnoldi 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 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

$$\boldsymbol{v}^{(k)} = \mathbf{A}\boldsymbol{v}^{(k-1)} / \|\mathbf{A}\boldsymbol{v}^{(k-1)}\| \quad (48)$$

$$\lambda^{(k)} = \boldsymbol{v}^{(k)T}\mathbf{A}\boldsymbol{v}^{(k)} \quad (49)$$

ACPD
11, 16745–16799, 2011

will converge to the dominant eigenpair $(\lambda_1, \boldsymbol{v}_1)$. The speed of convergence of the power method is proportional to $|\lambda_2/\lambda_1|$, hence 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.

An appropriate method to find the k largest eigenvalues and their associated eigenvectors is implemented in the open 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 processes, dependent on the properties of the matrix **A**. 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.

4 Tropospheric chemistry scenarios

Poppe et al. (2001) introduced a set of six scenarios for modelling 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, scenario URBAN/BIO features both anthropogenic and biogenic emissions. In detail, the first 60 h of scenario URBAN/BIO are identical to those of scenario URBAN. For scenario URBAN/BIO the anthropogenic VOC emissions are

ACPD
11, 16745–16799, 2011

switched off after 60 h and biogenic emission of isoprene is switched on. The emission strengths for scenario PLUME are specified in Table 1. For scenario URBAN and the first 60 h of scenario URBAN/BIO, the emission strength of NO of scenario PLUME has to be multiplied by 5, while the emission strengths of all other species remains unchanged.

The scenarios are defined via meteorological parameters (described in Table 2), chemical initial values (Table 3) and photolysis frequencies. The chemistry mechanism applied is the second generation Regional Acid Deposition Model (RADM2, Stockwell et al., 1990), which considers tropospheric gas-phase chemistry only. Poppe et al. (2001) simulated these scenarios with five different chemistry box models. All simulations started at 1 July, local noon, and ended five days (120 h) later at 6 July, local noon and demonstrated excellent agreement between the participating numerical solvers. Hence the scenarios are unambiguously described and can also be taken for testing numerical solvers. Simulations with the numerical solver employed in this work show nearly perfect agreement with those of Poppe et al. (2001). Consequently, the applied configuration can be regarded as suitable to model the chemical evolution of the given scenarios. Based thereupon, the scenarios 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).

Nitrogen oxides and volatile organic compounds are two major classes of directly emitted precursors of ozone formation. 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, KET, GLY, MGLY, and DCB, whereas NO_x comprise the model species NO and NO_2. 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-

ACPD
11, 16745–16799, 2011

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 order to investigate the influence of initial values and emissions of VOC and NO_x on the O_3-formation, singular vector analyses are applied. Singular vectors are not only dependent on the chemical scenario, but moreover time dependent on the Rayleigh quotient with all its metrics and operators. Since integration start and integration time are of similar importance, it can be expected to find significantly different singular vectors across a suite of simulations with different initial and integration times. For a comprehensive investigation of those effects, a temporal singular vector diagram (TSVD) is implemented. Each TSVD consists of a complete set of singular vectors, which is designed to cover the original time interval $[t_0, t_n]$. For this purpose the calculations are carried out in $n = \frac{t_n - t_0}{\Delta t}$ time rows TR(I), I=0,...,n-1, where Δt determines the minimal time interval chosen. Each time row TR(I) consists of n individual singular vector simulations, all starting at time $t_I = t_0 + I\Delta t$, 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. 2.

For spin up reasons, the starting point t_0 of the TSVD is 2 July, local noon. Further, t_n is determined as 6 July, 12 h and Δt is 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_0 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_n, whereas the end point of each individual singular vector calculation is denoted as final time t_F.

ACPD
11, 16745–16799, 2011

4.1 Error growth of initial uncertainties of VOC and NO_x family

For the first step of the present sensitivity study, focus is placed on changes in the final ozone concentration due to uncertainties of initial VOC and NO_x concentrations. The associated singular vector analyses answer the following question:

– Is the ozone evolution dominated by changes in the initial values of VOC or NO_x?

Therefore, the grouped error growth (22) is applied with final projection on ozone only and initial projections on the family VOC ($\mathcal{F}_1$) and the family NO_x ($\mathcal{F}_2$). Other species were not taken into account. Since the answer to the question above 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 these results dependent on the scenario selected and the time interval chosen?

In order to answer the last question, the TSVD of grouped singular vectors is applied to each scenario. Prior to analysing the results, the accuracy of the tangent linear assumption is tested for each grouped error growth calculated within the TSVD. On this account the associated first singular vector is inserted for δx into Eq. (46). Variations of scalar α reveal, that setting $\alpha = 0.25$ is sufficient to gain $|1.0 - d| \leq 0.01$. Reduction of α further improves the test-results (within numerical precision). Therefore, the ratios are close enough to 1 to suggest that the tangent linear assumption is sufficiently held.

In order to gain insight into the results of a complete TSVD, the singular vectors are visualised by considering the vector components corresponding to a specific chemical compound or family separately. Since the singular vectors are set to unit length, a vector component of 1 indicates that the ozone concentration at final time is solely influenced by this particular compound or family. Figure 3 presents the results of a complete TSVD for scenarios MARINE, FREE and PLUME. The illustration reveals a clear distinction between sensitivities with initial time t_I at day and sensitivities with

ACPD
11, 16745–16799, 2011

initial time t_I at night, which also occurs for non-displayed scenarios LAND, URBAN and URBAN/BIO. For all scenarios, main qualitative features are:

1. Initial time
   The importance of NO_x and VOC changes with initial time t_I at sunset or sunrise. The specific initial value at different day or night times does not seem to affect the results much.

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

Figure 3 only allows for a qualitative comparison between the importance of VOC and NO_x initial values. For a direct comparison, the results of the singular vector analyses (depicted in Fig. 3 for scenarios MARINE, FREE and PLUME) are categorised following the findings listed above. In order to present unambiguous results, a small amount of simulations are not considered for categorisation. First of all, simulations with initial time t_I at night and final time t_F before first sunrise are 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 are disregarded for categorisation. More precisely, hours with increasing insolation are defined to be between sunrise and 3 h after sunrise and hours with decreasing insolation are defined to be between 4 h before sunset and sunset. Thirdly, for scenario URBAN/BIO only the biogenic part of the scenario is considered, since the first 36 h equal those of scenario URBAN (remember the spin up run of 24 h). The biogenic part of the URBAN/BIO scenario is denoted as scenario BIO.

The results are categorised along the criteria "Initial time" and "Simulation length" listed above. Firstly, attention is paid to the first criterion (Initial time). Thus the results are categorised into results of calculations with initial time t_I at day (category C_a) and results of calculations with initial time t_I at night (category C_b). In order to obtain general statements about the VOC versus NO_x limitation of the ozone formation, some

ACPD
11, 16745–16799, 2011

statistical values are considered. For each category C_i a mean impact m_i is calculated following the equation

$$m_i(j) = \frac{1}{n_i} \sum_{i \in C_i} \left| \boldsymbol{v}_i^*(j) \right|, \tag{50}$$

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

$$c_i^+(j) = \max_i |\boldsymbol{v}_i^*(j)| \tag{51}$$

$$c_i^-(j) = \min_i |\boldsymbol{v}_i^*(j)| \tag{52}$$

as well as the standard deviation

$$s_i(j) = \sqrt{\frac{1}{n_i} \sum_{i \in C_i} \left(\left| \boldsymbol{v}_i^*(j) \right| - m_i(j) \right)^2}, \tag{53}$$

are calculated for each category C_i to examine the significance of the mean impact m_i.

Figure 4 displays an example of statistical results for categories C_a and C_b in case of the scenario MARINE. Generally, the standard deviations for the chosen categorisation 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. According to the second criterion (Simulation lengt), simulation length tends to be another influential feature. Therefore categories C_a and C_b are further subdivided depending on simulation length. Hence categories 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 k^{th} sunrise after initial time t_I. Sunrise 0 equals initial time t_I and sunrise 4 equals final time t_F, respectively.

Showcase results of the statistics of the new categories are illustrated for scenario MARINE in Fig. 5. For all scenarios, the subdivision in terms of simulation length

ACPD
11, 16745–16799, 2011

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. Tables 4 and 5 summarise the categorisation results in terms of mean impact and standard deviation for the NO_x-section. Since the singular vectors are normalised, the mean impact of the VOC-section can be derived directly. Notable findings of the categorisation are summarised in the following.

For simulations with initial time t_I at day, Table 4 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 5 exhibits that the results resemble for scenarios LAND, MARINE, PLUME and BIO. For the shortest time 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 slightly NO_x dominated to slightly 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 scenarios given and time interval chosen. According to the questions at the beginning of this section, the key findings are:

- For simulations with initial time t_I at day, scenarios with cleaner air are more NO_x sensitive.
- For simulations with initial time t_I at night, the VOC influence is typically decreasing

ACPD
11, 16745–16799, 2011

with simulation length. The decrement rate appears to be dependent on the degree of air pollution.

4.2 Error growth of initial uncertainties in VOC and NO_x species

The grouped error growth calculations carried out in the previous section are able to analyse the VOC versus NO_x dominance for a suite of time intervals and specific scenarios. The question to be faced in this section is:

– To which extent is each individual compound of NO_x or VOC responsible for the particular influence of the respective group?

Calculations with the projected error growth can deal with this question by focusing on the individual species belonging to VOC and NO_x at initial time t_I and on O_3 at final time t_F.

Figure 6 displays results of TSVDs of projected error growths for isoprene 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 categorisations and statistics introduced in Sect. 4.1 are applied. Results indicate that a categorisation into simulations with initial time t_I at day (C_a) and simulations with initial time t_I at night (C_b) is generally reasonable. For a ranking of species however, the projected singular vectors do not benefit much from a subcategorisation dependent on simulation length in all cases. A careful examination for each scenario is demanded to gain useful results. Figure 7 provides an example of both categorisations for scenario MARINE.

Results of all scenarios show, that the ranking within the NO_x group is independent of categorisation and can be summarised as follows:

– Within NO_x, NO and NO_2 show nearly equivalent influence (with NO_2 slightly more important).

The ranking for the species of VOC however is dependent on the chosen scenario, the initial time t_I of simulation (day or night) and the length of simulation.

ACPD
11, 16745–16799, 2011

For simulations with initial time t_I during daytime, 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 (except for scenario MARINE), OLI (except for scenarios PLUME and BIO), ISO, OLT, and HC5 (except for scenario PLUME) and OL2, HC3, and KET (+HC5 for scenario PLUME). Model species HCHO, GLY, and ETH are of least importance. 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 all these features.

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, but in terms of time-dependent behaviour. 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.

In summary the statistics for the VOC compounds for all scenarios indicate the following general results:

– 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.

4.3 Relative error growth of initial uncertainties of VOC and NO_x family

So far, investigation of the scenarios is carried out without any weighting of the perturbations. Hence existing results give insight into the influence of the kinetics without considering typical mixing ratios of different species. This property leads to domination

ACPD
11, 16745–16799, 2011

of error growth by more abundant species. In order to evaluate relative influences of chemical compounds, the problems considered in Sects. 4.1 and 4.2 are investigated with relative error growths. Unfortunately, a weighting with typical initial concentrations is not feasible for scenarios LAND, MARINE and FREE, due to the fact that the only reactive carbonaceous compounds are CO, CH_4 and HCHO. The remaining VOC have zero concentrations. Therefore, relative singular vector analyses are only applied to scenarios PLUME, URBAN and BIO. Recall that the notation BIO indicates that only the biogenic part of scenario URBAN/BIO is considered.

Since grouped relative singular vectors weight the influence of chemical compounds by their typical concentrations, the main questions raised in Subsect. 4.1 are modified to:

- Is the ozone evolution dominated by relative changes in the initial values of VOC or NO_x?
- To what extent are these results dependent on the scenario selected and the time interval chosen?

In order to answer these questions, TSVDs of grouped relative singular vectors are 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, the gradient check ratio introduced in Eq. (46) is applied to each conducted singular vector calculation. Here, the first singular vector is utilised as perturbation δx, while α is varied. Calculations prove that selecting $\alpha = 0.3$ is sufficient to gain $|1.0 - d| \leq 0.01$. Decreasing α leads to a better approximation (within the limits of numerical precision). The tangent linear assumption holds sufficiently well.

TSVD-results for scenarios PLUME and URBAN are collected in Fig. 8. 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 (Sect. 4.1) and the grouped relative error growth. The singular vector results of the grouped relative error growth do not show a periodic pattern for initial time t_I during day or night time.

ACPD
11, 16745–16799, 2011

Instead, each time row TR(*l*) (see Fig. 2) has its own pattern, which does not reiterate. Nevertheless, the behaviour of the deviation from the initial influence seems to recur. In order to estimate general quantitative features for the grouped relative error, the evolutional behaviour is categorised as introduced in Sect. 4.1. But neither considering the deviation from the initial influence nor their modifications (taking the squared measure etc.) leads to meaningful statistical results. Another attempt to state the ageing characteristic precisely is to apply the initial concentrations of the VOC and NO_x components on the grouped relative error growth. Different applications are tested but none of them leads to precise characteristics. Using the same application to transform the grouped relative error into the grouped error growth does not succeed either. A straightforward pattern that helps to find meaningful time-dependent categorisations for the VOC or NO_x dominance of the scenarios appears difficult to achieve.

Facing these difficulties, the qualitative features of the scenarios are analysed.

4.3.1 Scenario PLUME

For intermediate forecast lengths (i.e. simulation lengths shorter than $(t_n - t_0)/2$), the relative influence of VOC is increasing with increasing simulation length. For longer simulation lengths however (i.e. simulation lengths longer than $(t_n - t_0)/2$), the relative influence of VOC is decreasing with increasing simulation length. The relative importance of initial VOC-values is low during hours directly after sunrise. On a daily basis, the relative importance of VOC increases with ongoing ageing time between starting time t_0 and initial time t_l.

4.3.2 Scenario URBAN

Scenario URBAN shows a plain behaviour: With growing simulation length, the sensitivity of VOC increases. Similar to scenario PLUME, the evolutional behaviour is less pronounced for initial time t_l at night. The daily relative influence of VOC is increasing the later the simulation starts.

ACPD
11, 16745–16799, 2011

4.3.3 Scenario BIO

In contrast to the other scenarios, scenario BIO shows approximately constant VOC-influence with growing simulation length 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. The VOC influence is strongest for simulations starting after sunrise and reduces for simulations starting afterwards. In compliance with scenarios PLUME and URBAN, the daily relative influence of VOC increases with ongoing ageing time between starting time t_0 and initial time t_I.

In summary, the key features of the analysed scenarios are:

- The daily relative influence of NO_x decreases the later the initial time t_I of the simulations is.
- The NO_x-sensitivity is larger for simulations starting during morning hours.

4.4 Relative error growth of initial uncertainties 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 NO_x or VOC responsible for the particular influence of the respective group?

The current section aims to answer this question by analysing projected relative singular vectors. In contrast to projected singular vectors (Sect. 4.2), projected relative singular vectors demonstrate the influence of relative changes in the initial concentrations of each individual compound.

Figure 9 displays results of the TSVD of optimal projected relative singular vectors for compound KET for scenario PLUME. This particular case reveals at a glance, that the importance of the depicted compounds changes remarkably with growing distance between initial time t_I and starting time t_0 (recall, that the starting point of the TSVD is

ACPD
11, 16745–16799, 2011

called starting time t_0, while the starting point of each individual singular vector calculation is called initial time t_I). TSVD-results for other compounds and other scenarios are not depicted, but they confirm this finding. Therefore, the categorisations and statistics introduced in Sect. 4.1 are applied for each day and night separately.

Figure 10 provides an example of the results for day 1, day 4, night 1 and night 4 of scenario PLUME. 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 i-th sunrise and i-th sunset after starting time t_0. Sunrise 0 equals starting time t_0.

For all scenarios, examination of the statistical results for all days and nights 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 neglected.
- The influence of all VOC compounds grows with increasing distance between starting time t_0 and initial time t_I. Consistently, 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 only negligible impact.

The exact amount of influence for each species of VOC and NO_x differs by scenario. In the following, the most important findings per scenario are described shortly.

4.4.1 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.

ACPD
11, 16745–16799, 2011

4.4.2 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 differs from starting time t_0. The rate of decrease, however, is larger for NO.

4.4.3 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 impact of NO_x is induced by NO_2.

4.5 Relative error growth of emission factors of VOC and NO_x family

In the following, the VOC versus NO_x dependence of the ozone evolution is analysed with respect to emissions factors. Since uncertainties in emission factors are typically given as relative errors, 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), 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 analysed, since other scenarios do not include emissions. Again, the notation BIO indicates that only the biogenic part of scenario URBAN/BIO is considered.

The objective of this section is to evaluate the role of emission factors of VOC and NO_x in the ozone evolution. More precisely, the main questions to be faced here are:

– Is the ozone evolution dominated by changes in emissions of VOC or NO_x?

ACPD
11, 16745–16799, 2011

– To what extent are these results dependent on the scenario selected and the time interval chosen?

In order to provide insight into these aspects, the TSVD of grouped relative singular vector analysis is applied to scenarios PLUME, URBAN and BIO. In order to ensure that the tangent linear assumption is held for all TSVD-calculations, the gradient check ratio (46) is tested for the first singular vectors. 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 a reduction to 10 % perturbation of the emission factors ($\alpha = 0.1$) results in a sufficiently accurate tangent linear approximation ($|1.0 - d| \leq 0.01$). Decreasing α improves the approximation of the gradient check ratio (within the limits of numerical precision). 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. Again, 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 behaviour. Figure 11 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 include:

– The emission factors of VOC gain influence with growing distance between starting time t_0 and initial time t_I. An exception are simulations with final time t_F before and immediately after the first sunset in simulation time.

– The impact of the emission factors of VOC generally grows with increasing simulation length, except for simulations with final time t_F around sunrise, where the influence of VOC temporarily decreases.

ACPD
11, 16745–16799, 2011

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

4.6 Relative error growth of emission factors of VOC and NO_x species

In order to provide more detailed information about the VOC and NO_x limitation of the ozone evolution, 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. Figure 12 displays the results of the TSVD of scenario PLUME for OLI and 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 changes as Fig. 12 points out. Therefore, a categorisation for the ranking of species is not advisable.

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

ACPD
11, 16745–16799, 2011

- 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 analysed separately to provide a more comprehensive picture of the results.

4.6.1 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.

4.6.2 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, while ALD, ETH, and HCHO have least influence among the VOC species.

ACPD
11, 16745–16799, 2011

5 Summary and conclusions

In this work, singular vector analyses are applied to atmospheric chemical modelling in order to optimise the measurement configuration of chemical compounds. Initial chemical values and emissions are investigated as target variables. With the help of newly introduced operators, specific questions of atmospheric chemistry are addressed, like the VOC- versus NO_x-sensitivity of the ozone formation.

The application of emissions as target variables is a novel feature in the field of targeted observations. In order to reduce the amount of optimisations problems, the emission factor approach of Elbern et al. (2007) is adopted, where the amplitudes of the diurnal profile of emission rates are analysed. Since emission factors are time-invariant, only a single singular vector calculation has to be considered per time interval.

Special arrangements are made to adequately represent particular problems of atmospheric chemistry. Error growths are extended to allow for projected target variables not only at final time but also at initial time. Further, a family operator is 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 is applied to the RADM2-mechanism. For a set of six different scenarios, the VOC versus NO_x limitation of the ozone formation is investigated. The time-dependence of singular vectors is examined via differing initial and final simulation times.

Results reveal, that singular vectors are strongly dependent on start time and length of the simulation. Based on recurring patterns, a categorisation of the absolute error by simulations starting at daytime and simulations starting at nighttime is reasonable. With a further subcategorisation into four different simulation length intervals, the development of the grouped singular vectors can be adequately covered. Here, the importance of initial time is dependent on the time of the day and independent on the calender date

ACPD
11, 16745–16799, 2011

itself. For relative errors, however, the particular point in time is crucial. Grouped relative singular vectors are dependent on the initial concentration of the VOC and NO_x components respectively, therefore no categorisation with adequate accuracy can be imposed for the evolutional behaviour. Only a rough approximation of the development of the relative influence of VOC and NO_x can be made. The grouped error growth gives insight into the influence of the kinetics without considering mixing ratios of different species. In this manner it can be applied for a general investigation and verification of chemical reaction cycles. The grouped relative error growth however can identify groups of species which have to be measured with priority, given a well defined objective. A straightforward transformation from the grouped relative error into the grouped error growth is not possible.

Calculations with the projected error growth and the projected relative error growth identify the species which induced the impact of the VOC and NO_x group. For VOC, the species with most absolute influence are CSL, DCB, and MGLY, followed by the compounds HC8, TOL, XYL, ALD, OLI, and ISO. Considering relative influence this changes to ALD, KET, HC3, followed by HCHO, TOL, HC5, and ISO. For NO_x, a balanced absolute influence of NO and NO_2 changes to a NO_2 dominated relative influence. These differences can easily be explained by different initial concentrations.

As expected, singular vectors with initial values as target variables tend to be more sensitive to initial values, while emission factors as target variables are more sensitive to simulation length. This behaviour is due to the length of emission exposure on the given scenario. Species XYL and TOL induce the most emission factor impact of the VOC group, whereas the influence of HCHO, ETH, and ALD is negligible.

In summary it can be stated that singular vector analysis is a powerful tool, which identifies critical chemical species. It can be applied for investigation of chemical reaction cycles as well as for effective campaign planning. Further, the detected directions of largest error growth can be employed to initialise ensemble forecasts and to model covariances.

ACPD
11, 16745–16799, 2011

Acknowledgements. This study was funded by the virtual institute for Inverse Modelling of the Atmospheric Chemical COmposition (IMACCO) in the framework of the Helmholtz-Impuls- und Vernetzungsfonds under grant VH-VI-117. This support is gratefully acknowledged. All simulations utilised the Kinetic PreProcessor (KPP) developed by Valeriu Damian, Adrian Sandu, Mirela Damian, Florian A. Potra and Gregory R. Carmichael. Many thanks to Adrian Sandu for his technical assistance. We also thank Dirk Poppe for providing the tropospheric chemistry scenarios and his helpful advice on their implementation.

References

Barkmeijer, J., Gijzen, M. V., and Bouttier, F.: Singular vectors and estimates of the analysis-error covariance metric, Q. J. Roy. Meteorol. Soc., 124, 1695–1713, 1998. 16753

Berliner, L. M., Lu, Z., and Snyder, C.: Statistical design for Adaptive Weather Observations, J. Atmos. Sci., 56, 2536–2552, 1998. 16747

Bishop, C. H. and Toth, Z.: Ensemble Transformation and Adaptive Observations, J. Atmos. Sci., 56, 1748–1765, 1998. 16747

Buizza, R. and Montani, A.: Targeting Observations Using Singular Vectors, J. Atmos. Sci., 56, 2965–2985, 1999. 16747

Buizza, R. and Palmer, T. N.: The Singular-Vector Structure of the Atmospheric Global Circulation, J. Atmos. Sci., 52, 1434–1456, 1993. 16747

Buizza, R., Cardinali, C., Kelly, G., and Thepaut, J. N.: The value of targeted observations, ECMWF Newsletter, 111, 11–20, 2007. 16747, 16748

Derwent, R. G. and Jenkins, M.-E.: Hydrocarbons and the long-range transport of ozone and PAN across Europe, Atmos. Environ., 25A, 1661–1678, 1991. 16783

Elbern, H., Strunk, A., Schmidt, H., and Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, Atmos. Chem. Phys., 7, 3749–3769, doi:10.5194/acp-7-3749-2007, 2007. 16758, 16779

Hairer, E. and Wanner, G.: Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems, Springer, Berlin, 1991. 16760

Khattatov, B. V., Gille, J., Lyjak, L., Brasseur, G., Dvortsov, V., Roche, A., and Waters, J.: Assimilation of photochemically active species and a case analysis of UARS data, J. Geophys. Res., 104, 18715–18738, 1999. 16748

ACPD

11, 16745–16799, 2011

Lehoucq, R. B., Sorensen, D. C., and Yang, C.: ARPACK Users' Guide: Solution of Large-scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods, SIAM, Philadelphia, 1998. 16762

Lorenz, E. N.: A study of the predictability of a 28 variable atmospheric model, Tellus, 17, 321–333, 1965. 16747

Mises, R. V. and Pollaczek-Geiringer, H.: Praktische Verfahren der Gleichungsauflösung, Z. Angew. Math. Mech., 9, 58–77, 1929. 16761

Navon, I. M., Zou, X., Derber, J., and Sela, J.: Variational data assimilation with an adiabatic version of the NMC spectral model, Mon. Weather Rev., 120, 1433–1446, 1992. 16761

Palmer, T. N.: Predictability of the atmosphere and oceans: from days to decades, in: Proceedings of the ECMWF seminar on Predictability - Volume I, p 83–141, 1995. 16751

Parker, T. S. and Chua, L. O.: Practical Numerical Algorithms for Chaotic Systems, Springer Verlag, New York, 1989. 16747

Poppe, D., Aumont, B., Ervens, B., Geiger, H., Herrmann, H., Röth, E.-P., Seidl, W., Stockwell, W. R., Vogel, B., Wagner, S., and Weise, D.: Scenarios for Modeling Multiphase Tropospheric Chemistry, J. Atmos. Chem., 40, 77–86, 2001. 16762, 16763, 16775, 16783, 16784, 16785

Sandu, A. and Sander, R.: Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1, Atmos. Chem. Phys., 6, 187–195, doi:10.5194/acp-6-187-2006, 2006. 16760

Sandu, A., Liao, W., Carmichael, G. R., and Chai, T.: Singular Vector Analysis for Atmospheric Chemical Transport Models, Month. Weather Rev., 134, 2443–2465, 2006. 16748, 16750, 16752

Seinfeld, J. H. and Pandis, S. N.: Atmospheric chemistry and physics, Wiley-Interscience, 1998. 16764

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. 16762

Stockwell, W. R., Middleton, P., and Chang, J. S.: The second generation regional acid deposition model chemical mechanism for regional air quality modeling, J. Geophys. Res., 95, 16343–16367, 1990. 16763, 16785

Toth, Z. and Kalnay, E.: Ensemble forecasting at NMC: The generation of perturbations, B. Am. Meteorol. Soc., 74, 2317–2330, 1993. 16747

ACPD

11, 16745–16799, 2011

Table 1. Emission strength as volume production rate in $cm^{-3}s^{-1}$ for scenarios defined by Poppe et al. (2001). The VOC-emission is further segregated according to Derwent and Jenkins (1991).

NO	NO_2	CO	CH_4	SO_2	VOC
1.1×10^6	0.0	2.4×10^6	0.0	2.2×10^5	3.0×10^6

ACPD
11, 16745–16799, 2011

Table 2. Meteorological parameters for scenarios defined by Poppe et al. (2001).

	LAND, MAR., PLUME	FREE	URBAN, URB./BIO
Alt. (km)	0	8	0
Temp. (K)	288.15	236.21	298.00
Pres. (hPa)	1013.25	356.50	1013.25
Air (#/cm^3)	$2.55 \cdot 10^{19}$	$1.09 \cdot 10^{19}$	$2.46 \cdot 10^{19}$

ACPD
11, 16745–16799, 2011

Table 3. Initial mixing ratios for the gas-phase constituents (ppb) of scenarios described by Poppe et al. (2001). NMHC denotes non-methane hydrocarbons, other species are described in the RADM2 species list (see Stockwell et al., 1990).

	LAND	MARINE	FREE
O_3	30	30	100
NO	0.1	0.1	0.05
NO_2	0.1	0.1	0.05
HNO_3	0.1	1.5	0.1
CO	100	100	100
CH_4	1700	1700	1700
Isopr.	0	0	0
H_2	500	500	500
H_2O_2	2	2	2
HCHO	1	1	0
NMHC	0	0	0
SO_2	0	0	0

	PLUME	URBAN	URBAN/BIO
O_3	50	30	30
NO	0.2	0.1	0.1
NO_2	0.5	0.1	0.1
HNO_3	0.1	1.5	0.1
CO	200	100	100
CH_4	1700	1700	1700
Isopr.	0	0	0
H_2	500	500	500
H_2O_2	2	2	2
HCHO	0	1	1
NMHC	0	0	0
SO_2	0	0	0

ACPD
11, 16745–16799, 2011

Table 4. Mean impact and standard deviation of NO_x for categories $C_{a_{1/2/3/4}}$ (initial time t_I at day). Considered are optimal grouped singular vectors.

	LAND	MARINE	FREE
C_{a_1}	0.83 ± 0.04	0.62 ± 0.07	0.93 ± 0.04
C_{a_2}	0.72 ± 0.09	0.55 ± 0.05	0.90 ± 0.05
C_{a_3}	0.71 ± 0.09	0.62 ± 0.08	0.77 ± 0.05
C_{a_4}	0.75 ± 0.09	0.80 ± 0.12	0.66 ± 0.04
	PLUME	URBAN	BIO
C_{a_1}	0.49 ± 0.06	0.06 ± 0.02	0.49 ± 0.02
C_{a_2}	0.48 ± 0.05	0.05 ± 0.02	0.50 ± 0.02
C_{a_3}	0.60 ± 0.11	0.04 ± 0.02	0.51 ± 0.02
C_{a_4}	0.77 ± 0.12	0.04 ± 0.02	0.53 ± 0.02

ACPD
11, 16745–16799, 2011

Table 5. Mean impact and standard deviation of NO_x for categories $\mathcal{C}_{b_{1/2/3/4}}$(initial time t_I at night) . Considered are optimal grouped singular vectors.

	LAND	MARINE	FREE
$\mathcal{C}_{b_1}$	0.31 ± 0.06	0.27 ± 0.01	0.61 ± 0.04
$\mathcal{C}_{b_2}$	0.35 ± 0.06	0.32 ± 0.02	0.57 ± 0.04
$\mathcal{C}_{b_3}$	0.40 ± 0.06	0.40 ± 0.04	0.52 ± 0.03
$\mathcal{C}_{b_4}$	0.45 ± 0.07	0.51 ± 0.05	0.48 ± 0.02
	PLUME	URBAN	BIO
$\mathcal{C}_{b_1}$	0.24 ± 0.04	0.05 ± 0.02	0.37 ± 0.10
$\mathcal{C}_{b_2}$	0.34 ± 0.06	0.04 ± 0.02	0.48 ± 0.10
$\mathcal{C}_{b_3}$	0.56 ± 0.18	0.03 ± 0.02	0.60 ± 0.14
$\mathcal{C}_{b_4}$	0.73 ± 0.15	0.03 ± 0.02	0.71 ± 0.17

ACPD
11, 16745–16799, 2011

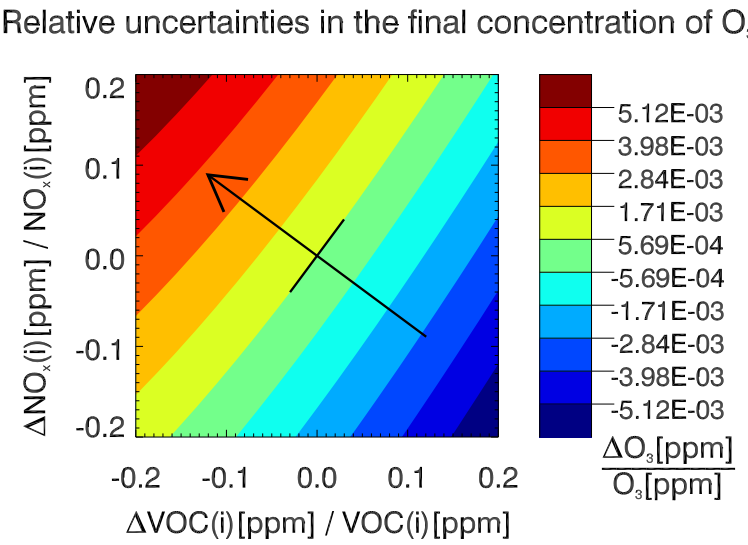

Fig. 1. Relative uncertainties in the final concentration of O_3. Isopleths demonstrate the relative change in the final concentration of O_3 resulting from variations in the initial concentrations of VOC and NO_x, respectively.

ACPD
11, 16745–16799, 2011

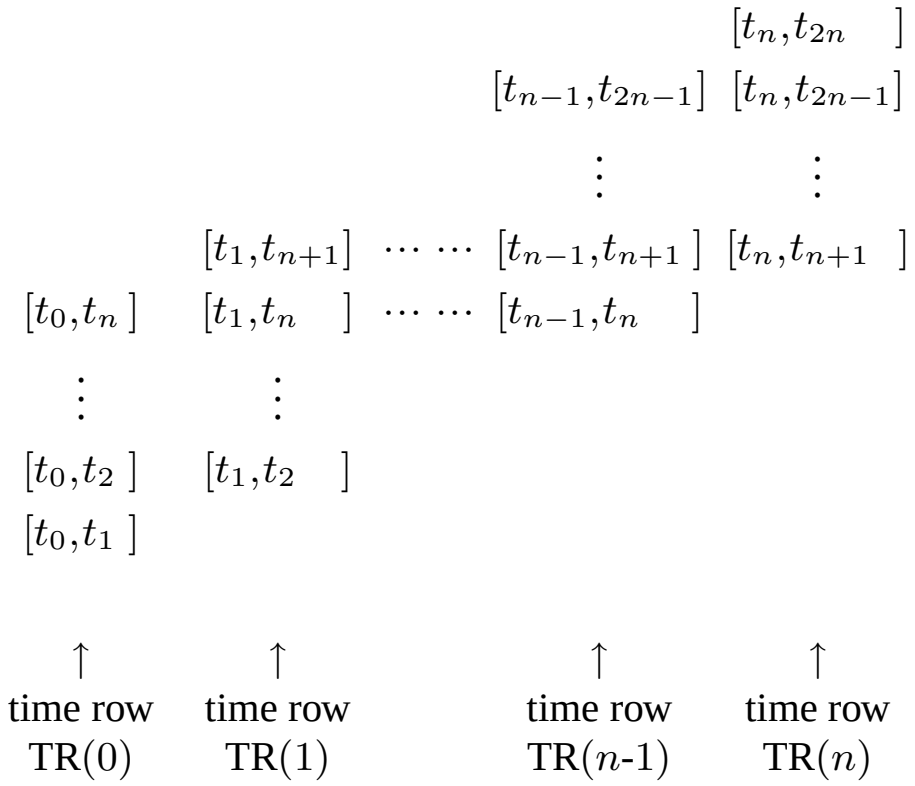

Fig. 2. Schematic presentation of the temporal singular vector diagram. Here, each interval $[t_I, t_F]$ represents one singular vector calculation starting at initial time $t_I = t_0 + I\Delta t$, $I = 0,..,n-1$ and ending at final time $t_F = t_0 + F\Delta t$, $F = 1,..,2n$, $F > I$.

ACPD
11, 16745–16799, 2011

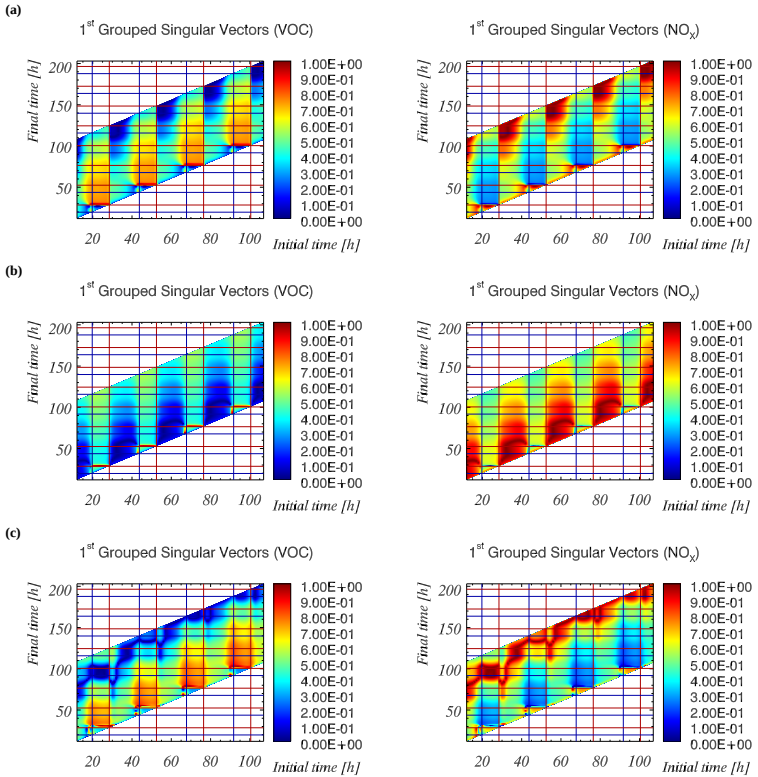

Fig. 3. TSVD of the optimal grouped singular vectors for scenarios **(a)** MARINE, **(b)** FREE and **(c)** PLUME. VOC-sections of the grouped singular vectors are depicted on the left panel column, NO_x-sections on the right panel column. The organisation of the results follows the schematic diagram displayed in Fig. 2. Each colour pixel indicates an individual singular vector calculation with initial time t_I and final time 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.

ACPD
11, 16745–16799, 2011

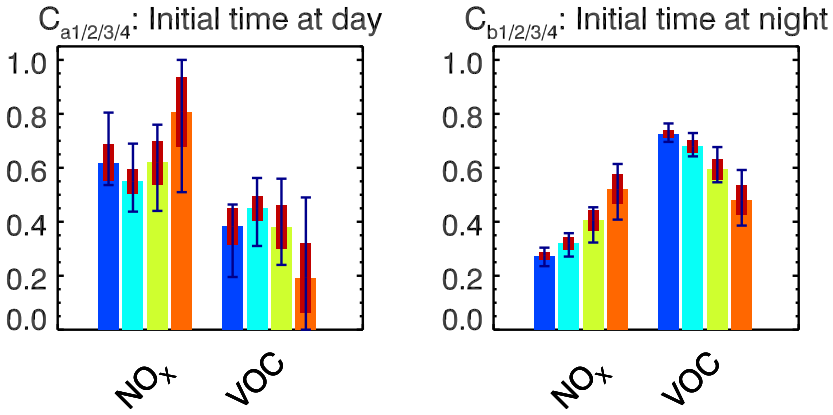

Fig. 4. Statistics of optimal grouped singular vectors for categories C_a and C_b for scenario MARINE. Depicted are mean impact (blue bars), minimum/maximum value (dark blue lines) and standard deviation (red bars).

ACPD
11, 16745–16799, 2011

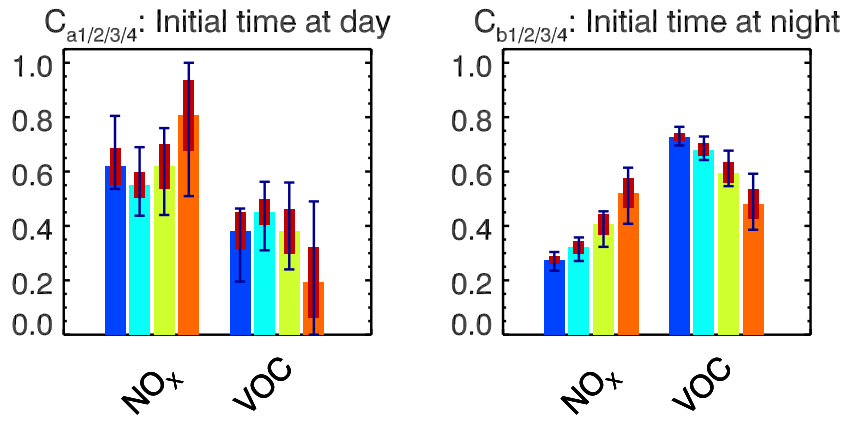

Fig. 5. Statistics of optimal grouped singular vectors for categories $C_{a_{1/2/3/4}}$ and $C_{b_{1/2/3/4}}$ for scenario MARINE. Depicted are mean impact (C_{a_1/b_1}: blue bars, C_{a_2/b_2}: turquoise bars, C_{a_3/b_3}: green bars, C_{a_4/b_4}: orange bars), minimum/maximum value (dark blue lines) and standard deviation (red bars).

ACPD
11, 16745–16799, 2011

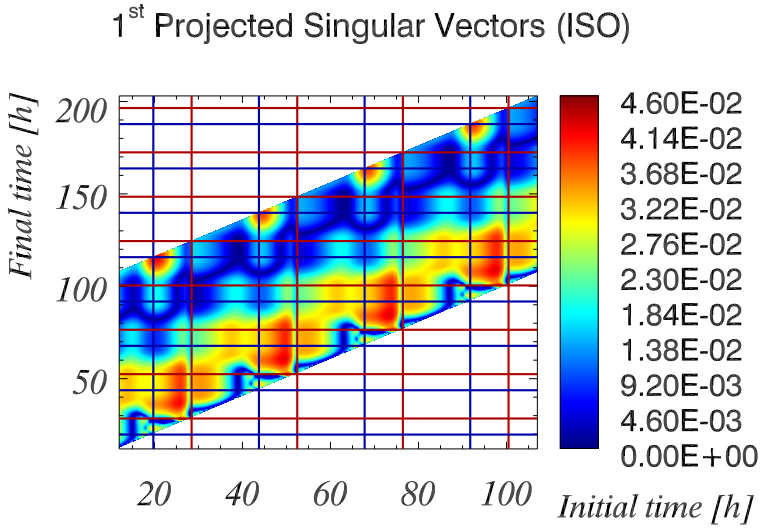

Fig. 6. ISO-section of TSVD of the optimal projected singular vectors for scenario MARINE. Plotting conventions as in Fig. 3.

ACPD
11, 16745–16799, 2011

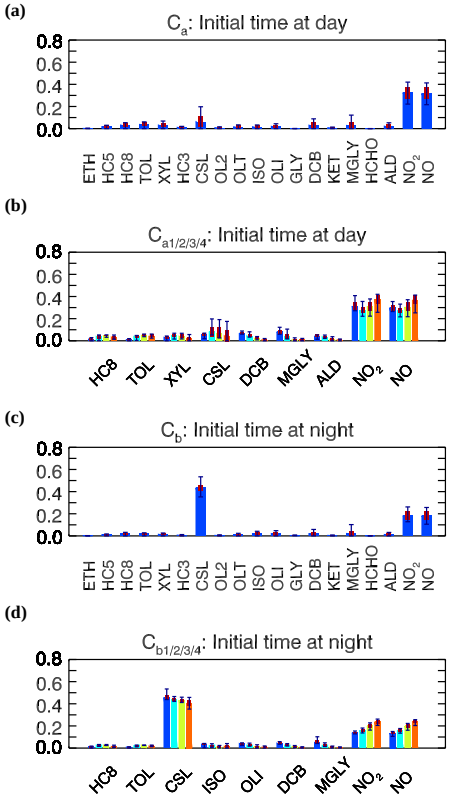

Fig. 7. Statistics of optimal projected singular vectors with respect to initial uncertainties for scenario MARINE. Panels **(a)** and **(c)** depict categories $\mathcal{C}_a$ and $\mathcal{C}_b$ respectively for all VOC/NO_x compounds. Panels **(b)** and **(d)** illustrate categories $\mathcal{C}_{a_{1/2/3/4}}$ and $\mathcal{C}_{b_{1/2/3/4}}$ respectively for the VOC/NO_x compounds with most influence. Both panels depict mean impact ($\mathcal{C}_{a/b}$: blue bars, $\mathcal{C}_{a_1/b_1}$: 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).

ACPD
11, 16745–16799, 2011

Singular vector analyses of tropospheric chemical scenarios

N. Goris and H. Elbern

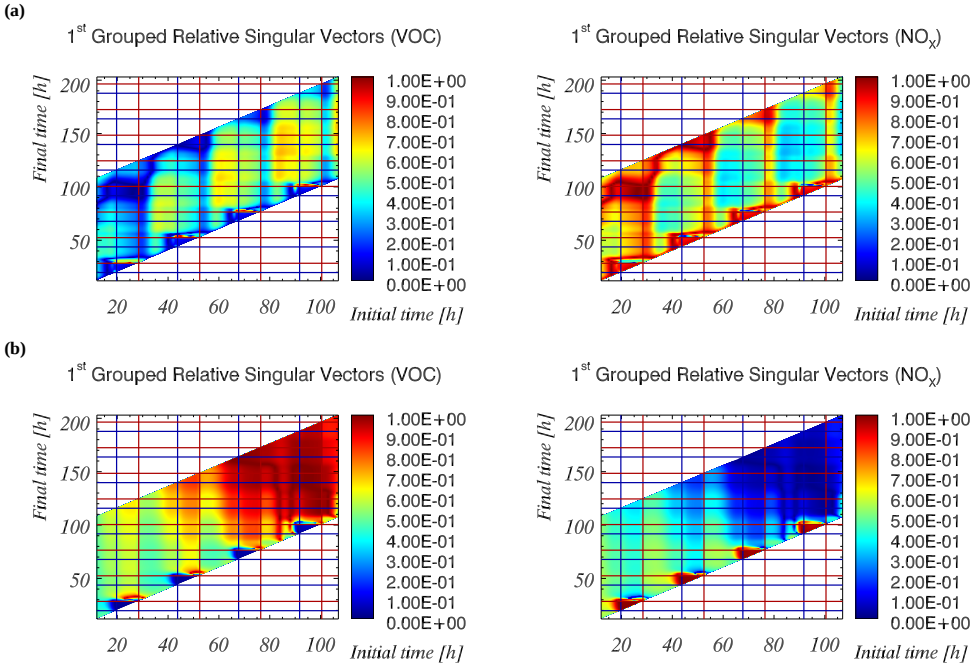

Fig. 8. TSVD of the optimal grouped singular vectors for scenarios **(a)** PLUME and **(b)** URBAN. VOC-sections of the grouped singular vectors are depicted on the left panel column, NO_x-sections on the right panel column. Plotting conventions as in Fig. 3.

ACPD
11, 16745–16799, 2011

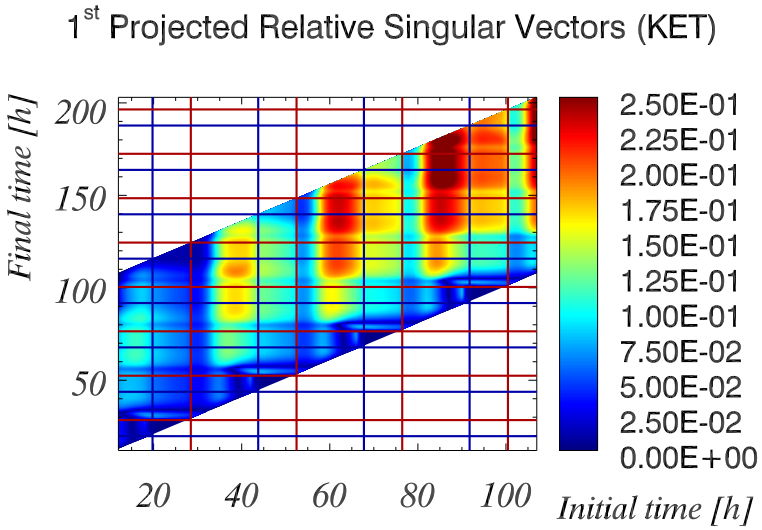

Fig. 9. KET-sections of TSVD of optimal projected relative singular vectors for scenario PLUME. Plotting conventions as in Fig. 3.

ACPD
11, 16745–16799, 2011

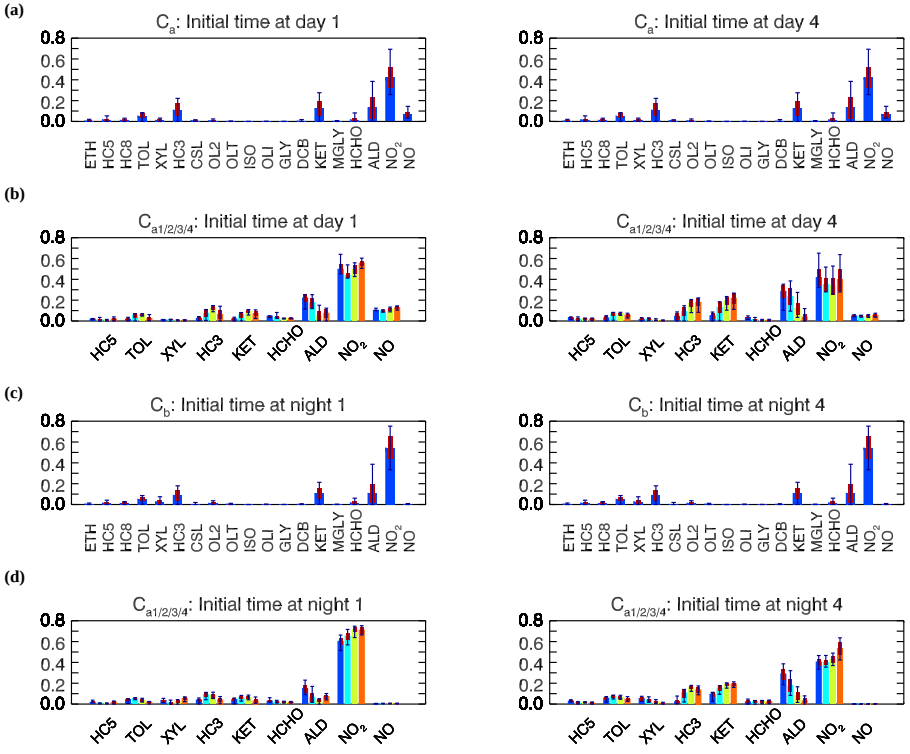

Fig. 10. Statistics of optimal projected relative singular vectors with respect to initial uncertainties for scenario PLUME. Panels **(a)** and **(c)** depict categories $\mathcal{C}_a$ and $\mathcal{C}_b$, respectively for all VOC/NO_x compounds. Panels **(b)** and **(d)** illustrate categories $\mathcal{C}_{a_{1/2/3/4}}$ and $\mathcal{C}_{b_{1/2/3/4}}$, respectively for the VOC/NO_x compounds with most influence. Both panels depict mean impact ($\mathcal{C}_{a/b}$: blue bars, $\mathcal{C}_{a_1/b_1}$: 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).

ACPD
11, 16745–16799, 2011

Singular vector analyses of tropospheric chemical scenarios

N. Goris and H. Elbern

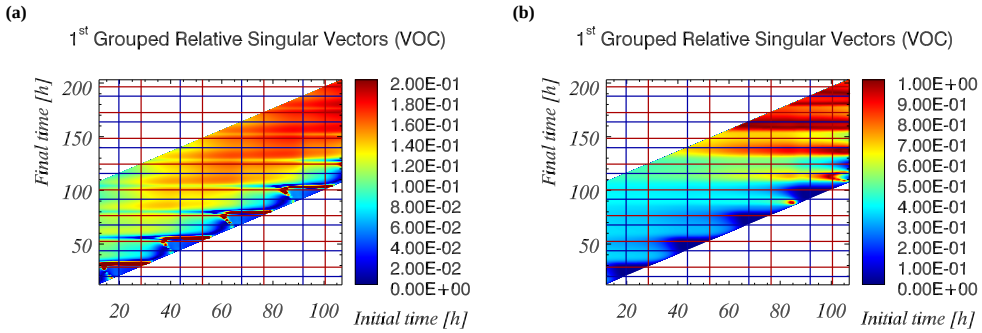

Fig. 11. TSVD of the optimal grouped relative singular vectors for scenarios **(a)** PLUME and **(b)** URBAN. Both panels depict VOC-sections of the grouped relative singular vectors. Plotting conventions as in Fig. 3.

ACPD
11, 16745–16799, 2011

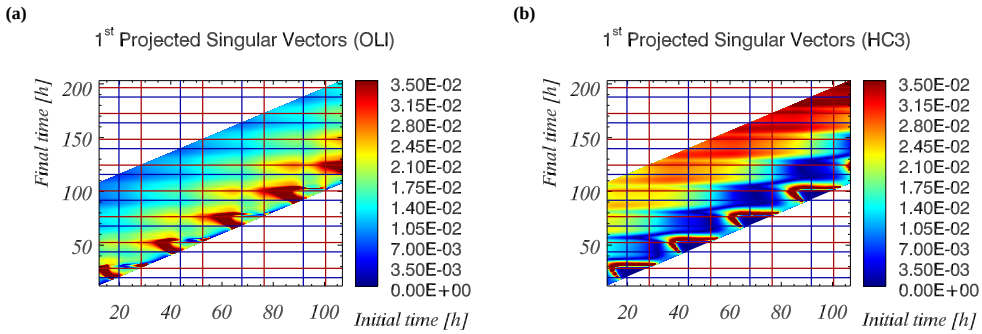

Fig. 12. TSVD of the optimal projected relative singular vectors for scenario PLUME. OLI-sections of the projected relative singular vectors are depicted on the left panel, HC3-sections on the right panel. Plotting conventions as in Fig. 3.

ACPD
11, 16745–16799, 2011