Geoscientific
Model Development

Singular vector-based targeted observations of chemical constituents: description and first application of the EURAD-IM-SVA v1.0

N. Goris[1,2,3] **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, Cologne, Germany
[3]Uni Research Climate, Bergen, Norway

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

Abstract. Measurements of the large-dimensional chemical state of the atmosphere provide only sparse snapshots of the state of the system due to their typically insufficient temporal and spatial density. In order to optimize the measurement configurations despite those limitations, the present work describes the identification of sensitive states of the chemical system as optimal target areas for adaptive observations. For this purpose, the technique of singular vector analysis (SVA), which has proven effective for targeted observations in numerical weather prediction, is implemented in the EURAD-IM (EURopean Air pollution and Dispersion – Inverse Model) chemical transport model, yielding the EURAD-IM-SVA v1.0. Besides initial values, emissions are investigated as critical simulation controlling targeting variables. For both variants, singular vectors are applied to determine the optimal placement for observations and moreover to quantify which chemical compounds have to be observed with preference. Based on measurements of the airship based ZEPTER-2 campaign, the EURAD-IM-SVA v1.0 has been evaluated by conducting a comprehensive set of model runs involving different initial states and simulation lengths. For the sake of brevity, we concentrate our attention on the following chemical compounds, O_3, NO, NO_2, $HCHO$, CO, $HONO$, and OH, and focus on their influence on selected O_3 profiles. Our analysis shows that the optimal placement for observations of chemical species is not entirely determined by mere transport and mixing processes. Rather, a combination of initial chemical concentrations, chemical conversions, and meteorological processes determines the influence of chemical compounds and regions. We furthermore demonstrate that the optimal placement of observations of emission strengths is highly dependent on the location of emission sources and that the benefit of including emissions as target variables outperforms the value of initial value optimization with growing simulation length. The obtained results confirm the benefit of considering both initial values and emission strengths as target variables and of applying the EURAD-IM-SVA v1.0 for measurement decision guidance with respect to chemical compounds.

1 Introduction

In meteorology and atmospheric chemistry, both data assimilation and inverse modelling seek to combine observations from a given observation network set-up with a model to reduce forecast errors. In contrast, the objective of targeted observations is to optimize the observation network for data assimilation and ensuing simulations applying a given model (e.g. Berliner et al., 1998; Daescu and Navon, 2004; Toth and Kalnay, 1993).

In numerical weather prediction, the optimal adaption of observations is a commonly investigated problem (e.g. Baker and Daley, 2000; Bishop and Toth, 1998; Palmer, 1995; Buizza and Palmer, 1993). It is typically studied to obtain a better estimate of initial values (Palmer, 1995). Events of explosive cyclogenesis at the North American eastern coast are often of highest relevance for European weather development and its forecast, and are therefore frequently taken

as study objects to obtain better configured observation sites and times. In order to find sensitive initial states, Lorenz (1965) introduced the application of singular vectors to numerical weather prediction by estimating the atmospheric predictability of an idealized model. Singular vectors determine the directions of fastest linear perturbation growth over a finite time interval and identify thereby sensitive system states, where small variations of considered input parameters lead to a significant forecast change. The identified sensitive system states are optimal target areas for adaptive observations, which help to optimize the information content of our monitoring capabilities and grant a better control of the dynamic system evolution by data assimilation. Likewise, this method can be effectively used for campaign planning (e.g. Gelaro et al., 1999; Langland et al., 1999; Kim et al., 2011). Buizza et al. (2007) investigated the results of field campaigns applying singular vector based targeted observations, including FASTEX (Fronts and Atlantic Storm-Track Experiment), NORPEX (North-Pacific Experiment), CAL-JET (California Land-falling JETs Experiment), the Winter Storm Reconnaissance Programs (WSR99/WSR00) and NA-TReC (North Atlantic THORPEX Regional Campaign), and stated that targeted observations are more valuable than observations taken in random areas. Yet, the extent of the impact is strongly dependent on regions, seasons, static observing systems, and prevailing weather regimes.

The successful application of singular vector analysis within numerical weather prediction motivated the transfer of this analysis method to chemical modelling, where studies addressing targeted observations are rare. Khattatov et al. (1999) gave the earliest stimulus for adaptive observations of chemical compounds. By investigation of the linearized model, Khattatov et al. 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. Hence, the problem of targeted observations of chemical compounds deals not only with the optimal placement of adaptive measurements, but also with the optimal set of chemical compounds to be measured. Daescu and Carmichael (2003) and Liao et al. (2006) introduced the application of an adjoint sensitivity method and of singular vector analysis, respectively, to chemical transport models (Lawrence et al., 2005). While Daescu and Carmichael (2003) and Liao et al. (2006) especially focused on the optimal placement of observations, a later study (Goris and Elbern, 2013) adapted singular vector analysis following the objective of Khattatov et al. (1999) and applied the theory to identify the optimal set of chemical compounds to be measured.

Initial values are not the only uncertainty when considering atmospheric chemical modelling. Errors in boundary conditions, emission rates, and meteorological fields add to the uncertainty of the chemical forecast (Liao et al., 2006). With progressing simulation time, the forecast solution is driven more by emission and less by initial values. While trace gas emissions are a forcing mechanism of prime importance for reactive chemistry simulations, they are not known exactly enough (e.g. Granier et al., 2011). This feature enforces the inclusion of emission rates in the data assimilation procedure (Elbern et al., 2007) and the need for targeting adaptive measurements for emission rates. In a first step, Goris and Elbern (2013) applied both emissions and initial values as target variables for singular vector analysis in a box-model context, yielding a relevance ranking of chemical compounds to be measured, while the optimal placement of those compounds is beyond the scope of zero-dimensional simulations.

In this work, the approach of Goris and Elbern (2013) was generalized for a three-dimensional chemistry transport model. The newly developed model set-up offers a comprehensive application of singular vector analysis by combining the idea of Goris and Elbern (2013) with the approach of Liao et al. (2006). Its objective is the detection of sensitive locations and species for atmospheric chemistry transport models. Specifically, the following questions are addressed: (i) which chemical species have to be measured with priority, and (ii) where is the optimal placement for observations of these components? Both questions are addressed with respect to emission strengths and initial species concentrations.

The present paper is organized as follows: the theory of singular vector analysis is presented in Sect. 2, where the application to initial concentration uncertainties and emission factors is described as well as the application of special operators. Singular vector analysis (SVA) is implemented in the EURAD-IM (EURopean Air pollution and Dispersion – Inverse Model, e.g. Elbern, 1997; Elbern and Schmidt, 1999; Elbern et al., 2007) three-dimensional chemical transport model, yielding the EURAD-IM-SVA v1.0, which is described in Sect. 3. In order to test and validate EURAD-IM-SVA v1.0, we focus on the model set-up of the ZEPTER-2 campaign (Zeppelin based tropospheric chemistry experiment, Part 2, Oebel et al., 2010; Wintel et al., 2013). The ZEPTER-2 campaign study configurations are described in Sect. 4. Results of singular vector analyses with respect to initial values and emission rates are presented in Sect. 5. Finally, the results of this work are summarized in Sect. 6.

2   Singular vector analysis for chemical models: theoretical background

The application of singular vector analysis to atmospheric chemical modelling allows for studying the influence of different kinds of uncertainties on the chemical forecast evolution. Within this work, we target the largest uncertainties in initial values and emissions, which both strongly determine the chemical system's evolution. A brief outline of the theoretical background of this application is presented in the following (see also Goris and Elbern, 2013, for a comprehensive discussion).

2.1 Initial values as target variables

A deterministic chemical forecast is processed by a typically nonlinear model operator, $\mathcal{M}_{t_I, t_F}$, propagating concentrations of a multitude of chemical species, $c \in \mathbf{R}^n$ (denoted in mass mixing ratios), forward in time:

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

with t_I initial time, and t_F final time.

For a three-dimensional transport model, the initial state of this equation is not entirely known, but has to be estimated relying on both former model results and assimilated observations. It is therefore subject to possible error growths. The evolution of an initial uncertainty or an initial error, $\delta c(t_I)$, which is sufficiently small to evolve linearly within a given limited time interval, can be modelled by the tangent linear model, $\mathbf{L}_{t_I, t_F}$ (Kalnay, 2002):

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

Our search for the most unstable initial uncertainty, $\delta c(t_I)$, can be described as the search for the phase space direction, which results in maximum error growth, $g(\delta c(t_I))$, at the end of the simulation:

$$\max_{\delta c(t_I) \neq 0} \left(g^2(\delta c(t_I)) = \frac{\|\delta c(t_F)\|_2^2}{\|\delta c(t_I)\|_2^2} \right) \tag{3}$$

$$= \max_{\delta c(t_I) \neq 0} \frac{\delta c(t_I)^T \, \mathbf{L}_{t_I, t_F}^T \mathbf{L}_{t_I, t_F} \, \delta c(t_I)}{\delta c(t_I)^T \delta c(t_I)},$$

where, for convenience, the squared error growth is maximized (Goris and Elbern, 2013). Here, $\mathbf{L}_{t_I, t_F}^T$ denotes the adjoint model and $\mathbf{L}_{t_I, t_F}^T \mathbf{L}_{t_I, t_F}$ the Oseledec operator. Since the Oseledec operator is symmetric, Rayleigh's principle can be applied (see, for example, Parlett, 1998). Accordingly, problem (3) can be solved by calculating the eigenvector $v_1(t_I)$ assigned to the largest eigenvalue λ_1 of the following eigenvalue problem:

$$\mathbf{L}_{t_I, t_F}^T \mathbf{L}_{t_I, t_F} \, v(t_I) = \lambda \, v(t_I). \tag{4}$$

The eigenvector, $v_1(t_I)$, of the Oseledec operator equals the right singular vector, $v_1(t_I)$, of the tangent-linear operator, $\mathbf{L}_{t_I, t_F}$. The singular value σ_1 equals the square root of the associated eigenvalue, λ_1, and is the maximum value of the error growth, $g(\delta c(t_I))$. It defines the amount of error growth at the end of integration time.

Weight matrix and projection operator

To allow for the calculation of relative error growths and to place foci on limited sets of chemical compounds and limited areas, we extend the analysis above by applying two special operators, namely weight matrix, $\mathbf{W}_t \in \mathbf{R}^{n \times n}$, and projection operator, $\mathbf{P}_t \in \mathbf{R}^{n \times n}$:

$$\mathbf{W}_t := \mathrm{diag}\left(c^{i,j,k,s}(t) \right)_{i,j,k,s} \quad \text{and} \tag{5}$$

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

Since the weight matrix contains concentration of chemical species (here, s denotes the considered species, while (i, j, k) denotes the considered numerical grid point), application of the inverse weight matrix yields relative perturbations and prevents the uncertainties of species with larger concentrations from dominating the error growth.

The projection operator allows for analysis of a limited set, $\mathcal{P}(t)$, of chemical species and grid points by setting the entries of the perturbations to 0 when they are not within the chosen set of species and regions (Barkmeijer et al., 1998).

With the help of both projection operator and weight matrix, we can consider the relative impact of a limited set of perturbations at initial time, t_I, on a limited set of perturbation at time t:

$$\delta c_{\mathrm{pr}}(t) := \mathbf{W}_t^{-1} \, \mathbf{P}_t \, \mathbf{L}_{t_I, t} \, \mathbf{P}_{t_I} \, \delta c(t_I), \tag{6}$$

where $\delta c_{\mathrm{pr}} \in \mathbf{R}^n$ is denoted as the projected relative error. The associated squared projected relative error growth $g_{\mathrm{pr}}^2(\delta c_{\mathrm{pr}}(t_I))$ is given by

$$g_{\mathrm{pr}}^2(\delta c_{\mathrm{pr}}(t_I)) := \frac{|\delta c_{\mathrm{pr}}(t_F)|_2}{|\delta c_{\mathrm{pr}}(t_I)|_2} \tag{7}$$

$$= \frac{|\mathbf{W}_{t_F}^{-1} \, \mathbf{P}_{t_F} \, \mathbf{L}_{t_I, t_F} \mathbf{W}_{t_I} \, \delta c_{\mathrm{pr}}(t_I)|_2^2}{|\delta c_{\mathrm{pr}}(t_I)|_2^2}$$

subject to

$$[\delta c_{\mathrm{pr}}(t_I)](j) = \begin{cases} [\dfrac{\delta c(t_I)}{c(t_I)}](j) & \forall \, j \in \mathcal{P}_{t_I} \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Here, $[x](j)$ denotes the jth component of a vector x. The phase space direction that maximizes the Rayleigh quotient (7) and ensures condition (8) is the solution $v_{\mathrm{pr1}}(t_I) \in \mathbf{R}^n$ of the symmetric eigenvalue problem

$$\mathbf{B}_{\mathrm{pr}}^T \mathbf{B}_{\mathrm{pr}} \, v_{\mathrm{pr}}(t_I) = \lambda_{\mathrm{pr}} \, v_{\mathrm{pr}}(t_I), \tag{9}$$

where $\mathbf{B}_{\mathrm{pr}} := \mathbf{W}_{t_F}^{-1} \, \mathbf{P}_{t_F} \, \mathbf{L}_{t_I, t_F} \, \mathbf{W}_{t_I} \, \mathbf{P}_{t_I},$

assigned to the largest eigenvalue λ_{pr1} (see Goris and Elbern, 2013, for a derivation of the eigenvalue problem). We refer to the solution as a projected relative singular vector, since it is the right singular vector of the operator $\mathbf{B}_{\mathrm{pr}}$. The square root of the eigenvalue λ_{pr1} is the associated projected relative singular value σ_{pr1}.

2.2 Emissions as target variables

Emissions, $e(t)$, impact the final state, $c(t_F)$, according to the differential equations, which describe the chemical evolution:

$$\frac{dc}{dt} = f(c(t)) + e(t). \tag{10}$$

Here, the function $f(c(t))$ compromises all processes that influence the chemical evolution apart from emission sources (as those are added separately). For a chemical transport model, the function $f(c(t))$ describes advection and diffusion of chemical species as well as their chemical formation and destruction. Equation (10) differs from Eq. (1) as it describes the rate of change for each chemical species, while Eq. (1) combines the initial conditions with the rate of change to calculate the chemical concentration for another point in time.

Like initial values, emissions are subject to uncertainties or errors, since their estimate is dependent on imperfect models and observation. Yet, emissions vary in time, leading to uncertainties or errors, $\delta e(t)$, at each time step $t \in [t_I, t_F]$. Consequently, the associated directions of largest error growth differ for each time step and their identification results in one application of singular vector analysis per time step, $t \in [t_I, t_F]$. In order to reduce the degrees of freedom to keep ill-posedness of the optimization problem and computational expenditure under control, we define a time-invariant vector of emission factors, e_f, instead, representing the amplitude of a prescribed diurnal emission profile (Elbern et al., 2007). This is a reasonable constraint as the daily evolution of emissions is far better known than the total emitted amount in a grid cell. Furthermore, the application of e_f has the advantage of resulting in only one singular vector analysis per time interval, $[t_I, t_F]$. The associated results quantify for which grid cell and which chemical species further emission strength assessment is most beneficial.

Introducing the vector of emission factors, e_f, Eq. (10) reformulates to

$$\frac{dc}{dt} = f(c(t)) + \mathbf{E}(t) \, e_f, \tag{11}$$

where $\mathbf{E}(t)$ is a diagonal matrix with the vector of emissions $e(t)$ on its diagonal. Accordingly, we also implement the vector of emission factors in the forward model $\mathcal{M}_{t_I,t_F}$, leading to a forward model $\mathcal{M}_{t_I,t_F}^{e_f}$ (which is exactly the same model as in Eq. (1), only with a different expression for the emissions). In order to determine the evolution of emission factor uncertainties, we utilize $\mathcal{M}_{t_I,t_F}^{e_f}$ to calculate the tangent linear model with respect to emission factors, $\mathbf{L}_{t_I,t_F}^{e_f}$. The tangent linear model integration of Eq. (10) reads

$$\delta c^{e_f}(t_F) = \mathbf{L}_{t_I,t_F}^{e_f} \delta e_f. \tag{12}$$

Here, the superscript e_f denotes that the uncertainty at final time is solely caused by emission uncertainties. In contrast, the uncertainty $\delta c(t_F)$ as described in Eq. (2) is solely caused by initial value uncertainties.

As an analogue to Sect. 2.1, we further want to identify the most unstable emission factor, δe_f. The latter is achieved by calculating the phase space direction, which results in maximum error growth, $g_r^{e_f}(\delta e_f)$, at the end of the simulation. Since emission factors already denote a relative measure, we consider henceforth only the relative impact of their uncertainty:

$$\delta c_r^{e_f}(t_F) := \mathbf{W}_{t_F}^{-1} \, \delta c^{e_f}(t_F), \tag{13}$$

where $\mathbf{W}_{t_F}$ is the weight matrix as defined in Eq. (5). With these restrictions, the squared relative error growth with respect to emission factors, $g_r^{e_f \, 2}(\delta e_f)$, reads

$$
\begin{aligned}
g_r^{e_f \, 2}(\delta e_f) :&= \frac{|\delta c_r^{e_f}(t_F)|_2^2}{|\delta e_f|_2^2} \\
&= \frac{\delta e_f^T \, \mathbf{L}_{t_I,t_F}^{e_f \, T} \, \mathbf{W}_{t_F}^{-T} \, \mathbf{W}_{t_F}^{-1} \, \mathbf{L}_{t_I,t_F}^{e_f} \, \delta e_f}{\delta e_f^T \, \delta e_f}.
\end{aligned}
\tag{14}
$$

According to Rayleigh's principle, the phase space direction that maximizes the ratio (Eq. 14) is the eigenvector $v_{r1}^{e_f}$ of the eigenvalue problem

$$\mathbf{L}_{t_I,t_F}^{e_f \, T} \, \mathbf{W}_{t_F}^{-T} \, \mathbf{W}_{t_F}^{-1} \, \mathbf{L}_{t_I,t_F}^{e_f} \, v_r^{e_f} = \lambda_r^{e_f} \, v_r^{e_f} \tag{15}$$

assigned to the largest eigenvalue $\lambda_{r1}^{e_f}$. As the solution equals the right singular vector of the operator $\mathbf{W}_{t_F}^{-1} \, \mathbf{L}_{t_I,t_F}^{e_f}$, it is denoted as a relative singular vector with respect to emission uncertainties. Its associated singular value $\sigma_{r1}^{e_f}$ is the square root of $\lambda_{r1}^{e_f}$.

A focal set of initial and final perturbations can be examined with the help of the projection operator, $\mathbf{P}_t$ (defined in Eq. 5). The associated projected relative singular vector for the error growth of emission factor uncertainties can be calculated following Sect. 2.1, "Weight matrix and projection operator".

3 Model design

3.1 The inverse European air pollution and dispersion model (EURAD-IM)

For the design of a model enabling three-dimensional singular vector analysis of chemical species and their temporal evolution, we implement the theory as described in Sect. 2 in a chemistry transport model. Our chemistry model of choice is the EURopean Air pollution and Dispersion – Inverse Model (EURAD-IM; e.g. Elbern, 1997; Elbern and Schmidt, 1999; Elbern et al., 2007). EURAD-IM is an advanced Eulerian model operating from European down to local scale by applying a nesting technique with the smallest horizontal solution available being 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 upper boundary is 100 hPa. Between surface and 100 hPa, 23 vertical model layers are defined.

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

$$\frac{\partial \boldsymbol{c}_i}{\partial t} = \underbrace{-\nabla(\boldsymbol{v}\boldsymbol{c}_i) + \nabla(\rho K \nabla \frac{\boldsymbol{c}_i}{\rho}) - \frac{\partial}{\partial \boldsymbol{z}}(\boldsymbol{v}_i^d \boldsymbol{c}_i) + \mathbf{A}_i + \mathbf{e}_i}_{f(\boldsymbol{c}_i(t))}, \quad (16)$$

where $\boldsymbol{c}_i$, $i = 1, \ldots, n$ denotes the mean mass mixing ratio of the chemical species i, $\boldsymbol{v}$ is the mean wind velocity, K is the eddy diffusivity tensor, ρ the air density, $\mathbf{A}_i$ the chemical source term for species $\boldsymbol{c}_i$, $\mathbf{e}_i$ its emission rates, and $\boldsymbol{v}_i^d$ its deposition velocity. The first part of the right-hand side of Eq. (16) corresponds to the function $f(\boldsymbol{c}(t))$ as given in Eq. (10), but is presented here for individual species. The selected numerical solution of Eq. (16) employs a symmetrical operator splitting technique (Yanenko, 1971), which splits the differential equations into sub-problems and treats them successively, centred around the chemistry solver module. For each sub-problem, the EURAD-IM provides multiple solution schemes. Here, the upstream algorithm devised by Bott (1989) is chosen as advection scheme featuring fourth-order polynomials for the horizontal advection and second-order polynomials for the vertical advection. The vertical diffusion is discretized using the semi-implicit Crank–Nicholson scheme and solved with the Thomas algorithm (Lapidus and Finder, 1982). The chemical development is implemented with the Kinetic PreProcessor software package (KPP, Sandu and Sander, 2006) using a second-order Rosenbrock solver.

3.2 EURAD-IM-SVA v1.0: expansion of the EURAD-IM to allow for singular vector analysis

We augment the EURAD-IM to allow for the option of singular vector analysis (SVA), yielding the EURAD-IM-SVA v1.0. In order to calculate targeted singular vectors as described in Sect. 2, tangent linear as well as adjoint models with respect to initial values and emissions need to be provided. Since the EURAD-IM offers the possibility of variational data assimilation with initial value and emission rate optimization, it comprises adjoint modules for all considered processes already. Furthermore, KPP provides the tangent linear model with respect to initial conditions for the chemical evolution. The tangent linear models of the remaining routines have been coded by hand.

Newly coded tangent linear routines have been checked for consistency with corresponding forward and adjoint modules. For consistency with the forward model, the gradient

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

$$d = \frac{\mathrm{FWD}(\boldsymbol{x} + \alpha\delta\boldsymbol{x}) - \mathrm{FWD}(\boldsymbol{x})}{\mathrm{TLM}(\alpha\delta\boldsymbol{x})}. \quad (17)$$

The abbreviations FWD and TLM denote parts of the forward model and their associated tangent linear routines (allowing for piecewise code-checking); α is a scalar parameter. While α approaches 0, the ratio (Eq. 17) should converge towards 1 until the limits of numerical precision are reached and convergence falters. Within these limits, the new tangent linear routines demonstrate the required characteristics of Eq. (17) for considered test cases. 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 $d \approx 1$.

Consistency of tangent linear and adjoint model can be tested by inspecting the validity of the following equation:

$$(\mathrm{TLM}(\delta\boldsymbol{x}))^T (\mathrm{TLM}(\delta\boldsymbol{x})) = \delta\boldsymbol{x}^T \mathrm{ADJ}(\mathrm{TLM}(\delta\boldsymbol{x})) \quad (18)$$

(Navon et al., 1992), where ADJ denotes associated parts of the adjoint model. When testing Eq. (18) for the newly implemented tangent linear routines, single routines as well as the complete model demonstrate correctness.

The central task of the EURAD-IM-SVA v1.0 is the detection of singular vectors and their associated singular values. Two methods have been implemented for solving the eigenvalue problems: the power method (Mises and Pollaczek-Geiringer, 1929) and a distributed memory version of the implicitly restarted Arnoldi method (PARPACK, Maschho and Sorensen, 1996; Lehoucq et al., 1998; Sorensen, 1996). The EURAD-IM-SVA v1.0 offers both methods for singular vectors with respect to initial values. For singular vectors with respect to emission factors, however, only the power method is implemented in the current model version. While the power method converges iteratively to the dominant eigenpair $(\lambda_1, \boldsymbol{v}_1)$, PARPACK has the ability to calculate the k largest eigenvalues and their associated eigenvectors by one iteration cycle. PARPACK relies on the Lanczos and Arnoldi process, dependent on the properties of the considered 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. Specifically, we apply PARPACK routines "PSNAUPD" (features the computation of the matrix-vector product) and "PSNEUPD" (features the computation of the requested eigenvalues and eigenvectors). PARPACK has the important advantage that it only needs a matrix-vector product instead of an explicit representation of the matrix $\mathbf{A}$. Since the eigenvalue problems in this work include operators, PARPACK is perfectly tailored to our needs. For future versions of the EURAD-IM-SVA, we plan on providing PARPACK not only for the singular vector analysis with respect to initial values, but furthermore for emission factor uncertainties.

4 Case study: ZEPTER-2 measurement campaign

We apply the set-up of the ZEPTER-2 measurement campaign (Zeppelin based tropospheric chemistry experiment, Part 2, Oebel et al., 2010; Wintel et al., 2013) to test and validate the EURAD-IM-SVA v1.0.

ZEPTER-2 deployed the ZEPPELIN NT airship as a platform to measure the distribution of different trace gases, aerosols, and short-lived radicals in the planetary boundary layer. During the campaign, 25 flights were carried out within a 100 km radius of the home base at Friedrichshafen airport (FDH), southern Germany. Vertical profiles of trace gases were measured above different surface types, including Lake Constance and surrounding forests.

ZEPTER-2 was supported by daily 3D-var analyses and chemical forecasts modelled with the EURAD-IM. The ZEPTER-2 set-up of the EURAD-IM allows for a practical application of the theory of targeted observations. Here, we apply singular vector analysis to identify the most sensitive locations and chemical compounds with respect to their impact on the final concentration of ozone. This study is designed to give insight into example applications of singular vectors in future campaigns by answering the following questions.

Q_C: Which of the chemical compounds O_3, NO, NO_2, HCHO, CO, HONO, and OH has to be measured with priority to provide an improved forecast for given ozone profiles?

Q_L: Where is the optimal location for observations of these components?

Here Q_C denotes "question with regard to compounds", and Q_L "question with regard to location".

We choose all spatial projections to contain grid points with ZEPTER-2 measurements and all compound-wise projections to focus only on chemical compounds measured during the ZEPTER-2 campaign. In this manner, it is revealed how singular vector analyses can support the set-up of an optimal campaign design when the chemical compounds to be measured and an approximate measurement route are already set. At final time, we focus specifically on vertical measurement profiles, since measurement profiles grant a larger magnitude of the optimal initial perturbation than single ZEPTER-2 measurement points (the location of the vertical measurement profile at final time is denoted as "final profile $VP(t_F)$" henceforth). For local projection at initial time, it is not reasonable to focus on locations of measurements solely, since thereby (a) spatial optimization is omitted and (b) the dynamics of the system are very limited, resulting in nearly negligible eigenvalues. Hence, no local projection was chosen. Yet, the approximate measurement route is kept by considering only those final profiles $VP(t_F)$ that contain ZEPTER-2 measurements at initial time, in the centre of their backward wind plume. Since only hourly initial times can be
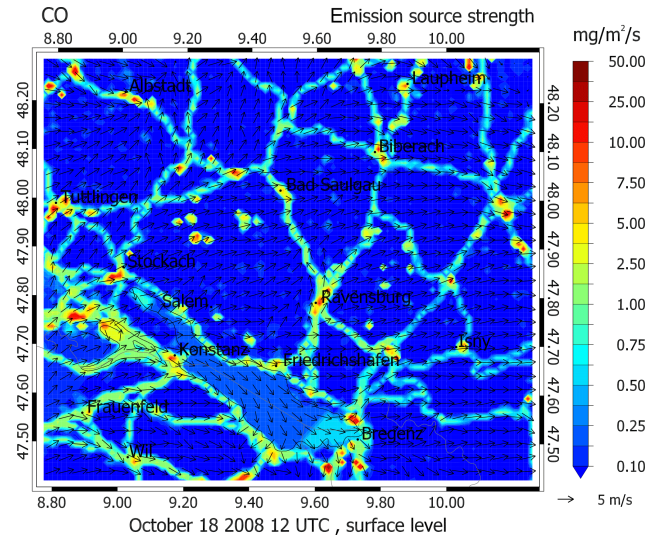

Figure 1. CO emission source strength ($\mathrm{mg\,m^{-2}\,s^{-1}}$) at the surface level of the ZPS grid for 18 October 2008, 12:00 UTC. Black arrows indicate the direction and strength of surface winds.

considered (due to the current EURAD-IM configuration), 17 simulation intervals meet the conditions described above. More details about the considered cases can be found in Table 1. Cases that share the same final profile $VP(t_F)$ are indicated with the same case number and subsequent distinctive letters.

EURAD-IM-SVA v1.0 configuration

The configuration of the EURAD-IM-SVA v1.0 applied in this study is based on the ZEPTER-2 set-up of the EURAD-IM. Here, RACM-MIM (Geiger et al., 2003) has been chosen as chemistry mechanism, while meteorological fields are provided by MM5 simulations (NCAR Mesoscale Meteorological Model, Grell et al., 1994). The ZEPTER-2 grid configuration of the EURAD-IM consists of a coarse European grid with a horizontal resolution of 45 km and a time step length of 600 s, and three nested grids with horizontal resolutions of 10, 5, and 1 km and time step lengths of 240, 120, and 60 s, 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 study. Due to its high horizontal resolution, the ZP3 grid provides a good representativeness of the measurements. In order to reduce the CPU time needed by singular vector calculations, the horizontal size of the ZP3 domain was reduced, resulting in a ZPS domain with $N_x = 111$, $N_y = 96$. Figure 1 illustrates the horizontal position of the ZPS domain. It was ensured that all flight trajectories remain within the ZPS grid. For a reference state in the centre of the ZPS domain, Table 2 lists the vertical grid structure in terms of height above ground.

Table 1. List of all singular vector simulations included in the ZEPTER-2 case study. Initial time (t_I) and final time (t_F) of simulation are given in UTC; the length of the simulation (time) is given in hours and minutes. $VP(t_F)$ denotes the location of the vertical measurement profile at final time, FDH designates Friedrichshafen airport, LC Lake Constance, FoA Forest of Altdorf, and Mengen denotes the city of Mengen.

Case	Flight	Date	t_I	t_F	Time	$VP(t_F)$
1a	02	18 Oct	12:00	13:30	1 h 30 min	LC
1b	02	18 Oct	13:00	13:30	0 h 30 min	LC
2a	02	18 Oct	11:00	14:00	3 h 00 min	FDH
2b	02	18 Oct	12:00	14:00	2 h 00 min	FDH
3	03	18 Oct	15:00	17:35	2 h 35 min	FDH
4a	04	19 Oct	09:00	12:15	3 h 15 min	FoA
4b	04	19 Oct	10:00	12:15	2 h 15 min	FoA
5a	05	19 Oct	14:00	15:20	1 h 20 min	FoA
5b	05	19 Oct	15:00	15:20	0 h 20 min	FoA
6	06	20 Oct	08:00	10:45	2 h 45 min	FDH
7a	07	20 Oct	13:00	14:45	1 h 45 min	LC
7b	07	20 Oct	14:00	14:45	0 h 45 min	LC
8a	08	24 Oct	16:00	18:00	2 h 00 min	FDH
8b	08	24 Oct	17:00	18:00	1 h 00 min	FDH
9a	21	7 Nov	10:00	11:25	1 h 25 min	Mengen
9b	21	7 Nov	11:00	11:25	0 h 25 min	Mengen
10	23	7 Nov	18:00	20:50	2 h 50 min	FDH

Emission estimates of the ZEPTER-2 set-up are provided by the EMEP (European Monitoring and Evaluation Programme) cooperative programme with a horizontal resolution of 50 km. The data consist 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 the EMEP emission data is not adequate for the considered ZPS grid, the horizontal resolution of the emission data sets was refined. For the refinement, land cover data sets of COoRdination of INformation on the Environment (CORINE) and of United States Geological Survey Global Land Cover Characterization (USGS-GLCC) were combined with data from GIS (Geographic Information Systems). In this manner of downscaling, emission data sets with a horizontal resolution of 1 km were generated, where consistency with the overlying EMEP emission data set is ensured. Emissions of small towns and busy roads are well resolved. An example of CO emissions on the ZPS grid can be found in Fig. 1.

Initial concentrations of all simulations are taken from 3D-var assimilation runs, conducted for the ZEPTER-2 campaign. Here, assimilation was accomplished every 4 h, starting at 02:00 UTC, and observational data of NO_2, NO, SO_2, O_3, CO, C_6H_6, $PM_{2.5}$, and PM_{10} were assimilated.

5 Results and discussion

In this section, elementary examples are demonstrated, illustrating performance and interpretation of singular vectors for observation targeting. The section is divided between initial value based singular vectors and those determined by emission rates. For both measures, we identify both optimal locations and optimal chemical compounds for additional measurements. Please note that the analysis of initial value uncertainties includes results of several leading singular vectors, while the analysis of emission factor uncertainties is only concerned with the leading singular vector. The latter is due to different implementations of eigenvalue problem solvers (see Sect. 3.2).

5.1 Singular vectors with respect to initial uncertainties

Singular vector calculations are based on the tangent linear model assuming that small perturbations evolve linearly within the simulation time. In order to grant meaningful results, this assumption has to be validated first. We apply Eq. (17) for validation and insert the chemical initial conditions of each simulation as x and the resulting singular vectors as perturbation, δx. 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 %) for each of the simulations. Hence, ratios are close enough to 1 to ensure that the tangent linear approximation is sufficiently accurate.

For initial uncertainties, we have calculated the five largest singular values for each of the considered cases using PARPACK (see Table 3). We find that the values of the singular vectors decrease relatively slowly. For 9 out of 17 cases, the fifth singular vector is still about half the value of the first singular value (see Table 3). The latter emphasizes the importance of all five leading singular vectors in our case study. For the sake of brevity, we restrict our identification of measure-
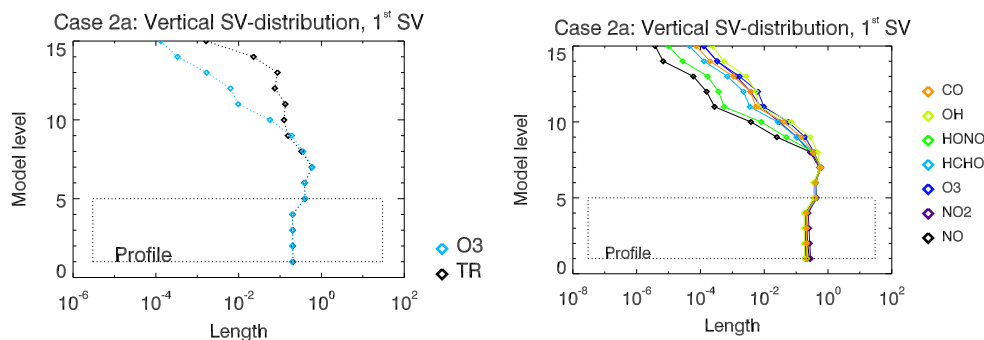
Figure 2. Vertical placement of the first singular vector with respect to initial value uncertainties for case 2a. Illustrated is the length of the vertical singular vector per model level for passive tracer and ozone (left panel) as well as for CO, OH, HONO, O_3, NO_2, and NO (right panel). Colour coding of each compound is denoted to the right of each panel. The black box indicates the height of the final profile VP(t_F).

Table 2. Vertical grid structure of the EURAD-IM-SVA v1.0 for the reference state 47.85° N, 9.50° E. Given are model level (ML) and height above ground (HT) in metres (m). The superscripts $^+$ and $^-$ indicate the upper and lower boundaries of the associated layer.

ML	HT$^-$ (m)	HT$^+$ (m)
23	10 937.50	14 009.19
22	8766.10	10 937.50
21	7060.07	8766.10
20	5643.57	7060.07
19	4426.45	5643.57
18	3355.84	4426.45
17	2397.90	3355.84
16	2040.85	2397.90
15	1696.93	2040.85
14	1446.98	1696.93
13	1203.46	1446.98
12	1005.18	1203.46
11	810.94	1005.18
10	658.33	810.94
9	508.11	658.33
8	396.96	508.11
7	287.08	396.96
6	214.51	287.08
5	142.48	214.51
4	106.66	142.48
3	70.98	106.66
2	35.43	70.98
1	0.00	35.43

ment priorities to the results of the first and second singular vectors.

### 5.1.1 Optimal placement of observations

An evident point of interest for chemistry is the relation between singular vectors resulting from passive tracer advection–diffusion, as merely controlled by meteorological parameters, and those which are also affected by reactive chemistry. Their differences can be visualized via horizontal and vertical placement (for a definition of horizontal and vertical placement, see Appendix A1). In the case of the latter, the left panel of Fig. 2 displays the vertical profile of the horizontal placement for the leading singular vector, broken down for the lower 15 model levels for a passive tracer "ozone" and reactive ozone for case 2a. It can be seen that up to a height limit of approximately 450 m (level 8), initial values of both passive and reactive chemistry demonstrate a similar influence per height level. The faster levelling of the reactive chemistry profile above level 8 indicates that initial values of higher levels are first transported into lower air masses before chemical production processes take place. The same pattern is seen for all considered cases and all considered chemical compounds (right panel, Fig. 2) with varying lower height limits for the faster levelling of reactive chemistry. These results can be expected as ozone production is initiated by chemical production processes at lower elevation or, in the case of ozone itself, ozone decomposition at lower elevation. Concerning differences in the levelling of different chemical compounds, we find that the relevance of measurements of O_3 and CO decreases slower than the relevance of measurements of NO and HCHO, independent of initial time t_I or simulation length (see Fig. 2). It can be assumed that this feature is linked to differing vertical profiles.

We find the same properties to be true for the vertical profile of the second singular vector. The left panel of Fig. 3 illustrates the vertical placement for the first and second singular vectors for reactive ozone for case 2a. It can be seen that the vertical profiles of first and second singular vectors are relatively similar to the second singular vector exhibiting slightly smaller values in lower air masses and higher values in higher air masses. Yet, compared to the passive tracer "ozone" (left panel, Fig. 2), the reactive chemistry profile of the second singular vector also exhibits a faster decrease with height for all considered cases and all considered chemical compounds (right panel, Fig. 3). Again, the relevance of mea-

Table 3. Five largest singular values (SV) with respect to initial value uncertainties for all 17 case studies. Case numbers are denoted according to Table 1.

Case	SV_1	SV_2	SV_3	SV_4	SV_5
1a	0.33756	0.21116	0.15000	0.12025	0.09680
1b	0.62180	0.43528	0.39816	0.36516	0.32796
2a	0.23881	0.08695	0.05089	0.01897	0.01732
2b	0.32939	0.15073	0.09439	0.04336	0.03302
3	0.20785	0.12149	0.08432	0.06091	0.05030
4a	0.27697	0.13624	0.06797	0.04604	0.02720
4b	0.35056	0.22871	0.10714	0.09889	0.05292
5a	0.52395	0.34937	0.31069	0.23216	0.22084
5b	1.00638	0.86925	0.82216	0.73719	0.70424
6	0.05874	0.01023	0.00872	0.00183	0.00132
7a	0.42151	0.24298	0.17263	0.13601	0.12783
7b	0.62200	0.43488	0.37958	0.35852	0.32628
8a	1.51770	1.18979	1.04014	0.92703	0.79162
8b	1.61465	1.24563	1.23831	1.07596	1.02942
9a	0.68862	0.60123	0.44726	0.35885	0.34969
9b	0.80649	0.77847	0.64214	0.58633	0.55604
10	0.28409	0.25807	0.23173	0.17787	0.15934
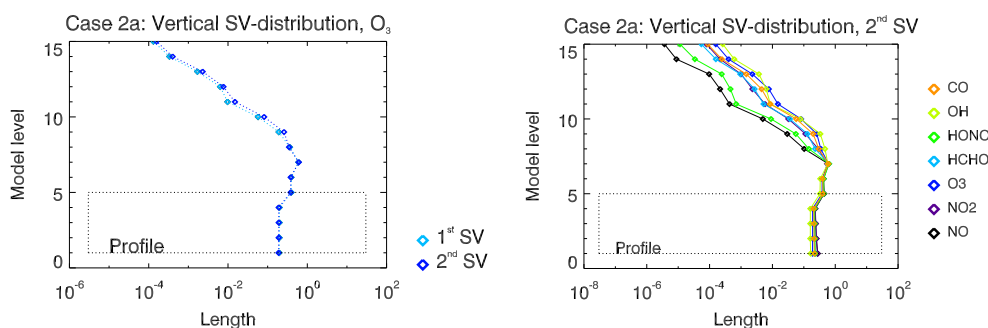
Figure 3. Vertical placement of the first and second singular vectors with respect to initial value uncertainties for case 2a. Illustrated is the length of the first and second vertical singular vectors per model level for ozone (left panel) as well as the length of the second vertical singular vector for CO, OH, HONO, O_3, NO_2, and NO (right panel). Colour coding of each compound is denoted to the right of each panel. The black box indicates the height of the final profile $VP(t_F)$.

surements of O_3 and CO decreases slower than the relevance of measurements of NO and HCHO.

Examination of the horizontal placement (for a definition of horizontal placement, see Appendix A1) of the first and second singular vectors for all cases confirms that the placement of passive tracer and ozone generally diverges more in higher model levels (as seen in the left panel of Fig. 4 for case 8a). Since the horizontal placement disregards effects of the vertical placement distribution and of different species magnitudes, a broader 0.01 isopleth in higher model levels (as seen in the left panel of Fig. 4) means that neighbouring grid cells show only small differences in placement importance. In comparison to passive tracer ozone, the reactive ozone of both first and second singular vectors reveals smaller isopleths at lower elevation and broader isopleths in higher model levels. The latter indicates varying chemical concentrations in lower air masses driven by locations of production sources and photochemical lifetimes. Even though ozone itself is not emitted into the atmosphere, its precursors are strongly influenced by emissions, leading to a highly variable distribution of ozone in lower levels of the troposphere, while it is relatively uniform in higher model levels. Due to this feature, placement differences between first and second singular vectors are less pronounced in lower air masses and most pronounced in higher model levels.

Results reveal furthermore that the horizontal placement of all considered chemical compounds usually coincides. Remarkable differences within the chemical placement are only discovered for cases 6, 7a, 8b, and 10, and can be explained by varying initial concentrations within the otherwise advection-controlled placement area. The horizontal distribution of the first and second singular vectors at the lowest
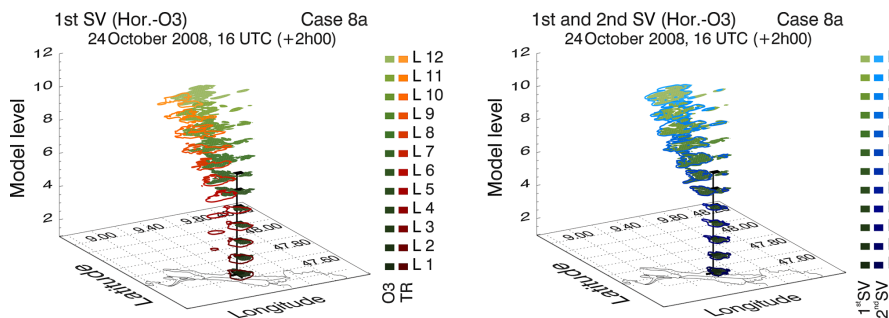
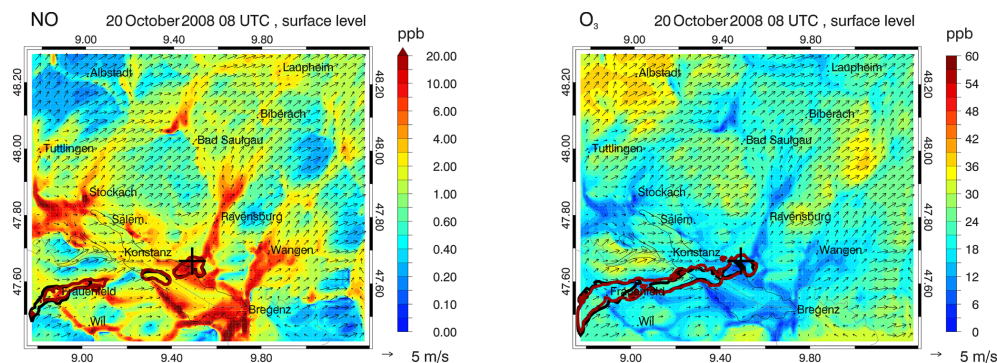**Figure 4.** Horizontal placement of the first and second singular vectors with respect to initial value uncertainties for case 8a. Left panel: 0.01 isopleths of the first horizontal singular vector for passive tracer (red framed shading) and ozone (green filled shading). Right panel: 0.01 isopleths of the first (green filled shading) and second (blue framed shading) horizontal singular vectors for ozone. In both figures, the final profile $VP(t_F)$ is marked with a black line and the black cross indicates its horizontal position. Case numbers and simulation intervals are given on top of each panel.

Figure 5. Initial concentrations and horizontal placement of the first and second singular vectors with respect to initial value uncertainties for case 6. Illustrated are results for NO (left panel) and O_3 (right panel) at surface level. The 0.01 isopleths of the first and second horizontal singular vectors are indicated with red and black lines, respectively, and the horizontal position of the final profile $VP(t_F)$ is marked with a black cross. Date and time are denoted above each panel.

level for case 6 is displayed in Fig. 5 for NO (left panel) and ozone (right panel). The westward orientation of the influence area displays the upwind domain of the Friedrichshafen target location, and shows a fairly evenly distributed domain for possible ozone measurements. It can be assumed that this area is mostly controlled by transport and diffusion processes. In contrast, the areas of sensitivity for NO cover three or four (depending on the singular vector considered) disconnected sub-domains enclosed by the ozone sensitivity area. These patches are associated with NO emission areas, and indicate the sensitivity of the ozone evolution to direct interaction with NO in the nearby area of Friedrichshafen, and also to indirect interaction (via NO_2) for the longer distance area at the westerly map border. Figure 5 furthermore confirms that the 0.01 isopleth of the horizontal placement of first and second singular vectors are fairly similar.

The analysed ZEPTER-2 cases share a relative short simulation interval (the longest simulation interval lasts 3 h 15 min) and a local projection on the final profile $VP(t_F)$. Both features restrict the dynamics of the system. It can be expected that the chemical placements are likely to differ

more when choosing longer simulation intervals (as is the case in simulations done by Liao et al., 2006).

5.1.2 Measurement priority of chemical compounds

Optimal compounds for additional measurements can be determined via the relative ranking defined in Appendix A2. Here, we consider the influence of compounds O_3, NO, NO_2, HCHO, CO, HONO, and OH on the ozone evolution.

Figure 6 provides an example of the relative ranking of the first and second singular vectors for O_3 and CO at model level 1 (ground level). Note that if a case is not depicted for a particular level, then the number of grid points (i, j, k) that hold $\sqrt{\sum_s v(i, j, k, s)^2} > 10^{-4}$ equals 0. Results of all cases reveal that O_3 is ranked first for more than 95 % of the considered grid points for all cases. None of the other species reveals such a distinct behaviour. Yet, it is possible to come to the following conclusions: (1) O_3 has most relevance among the considered chemical compounds, (2) NO, NO_2, HCHO, and CO show medium relevance, and (3) OH and HONO have least relevance. In most cases, the relevance
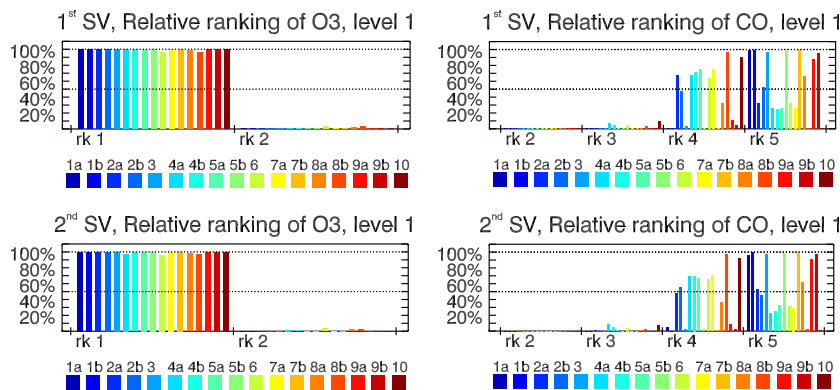
Figure 6. Relative ranking of the first (upper panel) and second (lower panel) singular vectors with respect to initial value uncertainties. Illustrated are results for O_3 (left panel column) and CO (right panel column) at surface level for all 17 case studies. Relative ranks are denoted below each bar plot. A rank m is only depicted if the associated chemical compound is ranked mth for at least one considered grid point. The colour coding of each case is denoted below each panel.

of OH is ranked seventh, while HONO is ranked sixth. In lower air masses, NO and NO_2 tend to be ranked second or third, while HCHO tends to be ranked third or fourth and CO fourth or fifth. This general ranking applies for both the first and second singular vectors. The revealed measurement priority meets our expectations as NO_x, CO, and volatile organic compounds are important precursors of ozone (Seinfeld and Pandis, 1998). Here, the considered cases are in general NO_x sensitive (see also Goris and Elbern, 2013).

We also find that the measurement priority of NO is higher for simulations starting during noon hours, while it is lower for simulations starting in the morning or in afternoon/evening time frames. This feature is related to the initial mixing ratio of NO which is close to 0 during night-time (Seinfeld and Pandis, 1998).

5.2 Singular vectors with respect to emission uncertainties

Prior to analysing the singular vectors with respect to emission factors, the linearity assumption is tested by inserting the calculated perturbations of largest error growth into Eq. (17). 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.

The optimization of observational networks with respect to measurements of emissions themselves is somewhat artificial, as only for very special cases are flux tower observations of CO_2 and, even more sparsely, other greenhouse gases, available. Nevertheless, formally it can be applied in very much the same way as for initial values and, for reactive emission sources under conditions with sufficiently large Damköhler numbers and small background concentrations,

traditional observations in emitting areas can serve as a supplement.

The subsequent analysis in Sect. 5.2.1 and 5.2.2 discusses only results for the first singular vector as further singular vectors are not available (see Sect. 3.2). Furthermore, we concentrate only on results for surface level and for chemical compounds NO, NO_2, HCHO, and CO. This is due to the fact that O_3, HONO, and OH are not emitted and, in the case of the ZEPTER-2 configuration, emissions are only included for surface level.

5.2.1 Optimal placement of observations

Figure 7 exhibits an example for formaldehyde (HCHO), which is both emitted into and produced in the atmosphere. Correspondingly, a spatial comparison between singular vectors of initial values and emission rate optimization will reveal spatial differences. It can be seen from the map that, influenced by the spatial distribution of the emission fields, the area for optimal observations of emissions is close to the final profile, while the area of optimal observations of initial values is in a larger distance. This outcome is valid for all cases and can be explained by the fact that the target area for emissions is the result of an optimization over the entire simulation interval. The target area of initial values can only be located within the area of the backward plume at its initial time, yet the target area of emissions can be any point within the entire advection trace area of the backward plume. Hence, the optimal placement of observations of emissions is strongly influenced by locations of emission sources within this plume (Fig. 7). The importance of emission sources is confirmed by the smaller extent of the target area of emissions, in comparison to initial values. Since the horizontal singular vector sections have unit length for a fixed compound and a fixed model level, a small extent of the target area shows that the additional value of observations is rela-
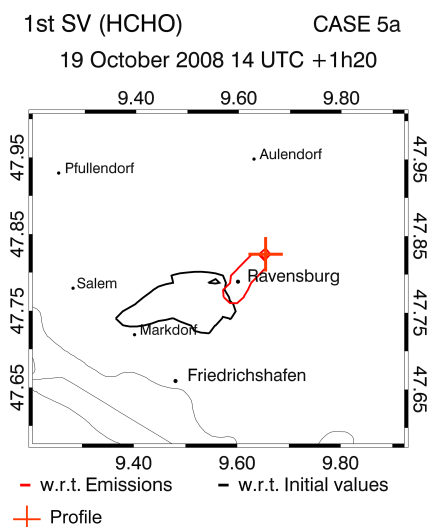
Figure 7. Optimal horizontal placement of emissions and initial values for HCHO at surface level for case 5a. The 0.01 isopleths of the optimal horizontal placement are indicated with a black line (initial values) and a red line (emissions). The horizontal position of the final profile $VP(t_F)$ is indicated with a red cross.

tively high at few grid points and decreases sharply for the surrounding grid points.

Comparing the target area of emissions for different compounds, we find that the target areas differ quite substantially in some cases. This feature occurs due to different emission source strengths for different compounds and will be explained in more detail at the end of the next section.

5.2.2 Relevance ranking of chemical compounds

In response to question Q_C, a relevance ranking for the emission influences of NO, NO_2, HCHO, and CO is assessed in this section (see Appendix A2). Note that species O_3, OH, and HONO are not emitted and are therefore not to be taken into account.

Results for all considered levels and species are depicted in Fig. 8. It is found that (1) the influence of NO emissions is most important, and (2) emissions of NO_2 tend to have the second-most influence, while (3) in the majority of cases, the importance of emissions of CO and HCHO alternates between the third and fourth rank. This result is to be expected, as NO_x, CO, and volatile organic compounds are the most important precursors of the ozone production. Dependent on the existing mixing ratio, the ozone production is NO_x or VOC sensitive (Seinfeld and Pandis, 1998; Goris and Elbern, 2013). Here, the considered cases are all NO_x sensitive.

Figure 9 serves to give an idea about the location dependence of the ranking of emission influences of HCHO and CO for case 2a. Based on the analyses of all 17 cases, the following conclusions can be drawn: (1) the importance of emissions of HCHO tends to increase in urban plumes at

the expense of the influence of emissions of CO and NO, and (2) the influence of emissions of CO tends to increase at busy roads. As compensation, the influence of emissions of HCHO and NO decreases. These findings are consistent with the modelled strength of different emission sources per compound.

5.3 Magnitudes of the leading singular values

The singular values of our calculations determine the relative error growths of uncertainties in initial values and emissions, respectively. Table 4 captures the leading singular values for the ZEPTER-2 calculations for both target variables (initial values and emissions) for simulations with a shared final profile $VP(t_F)$.

We find that the influence of singular values with respect to initial values decreases with growing simulation length, whereas the influence of singular values with respect to emissions increases (Table 4). This behaviour is expected since continuous emissions and their uncertainties affect the chemical evolution at every time step. Therefore, the emission sensitivity increases with each added time step. Uncertainties in initial values, on the other hand, influence the forecast mostly at initial time, with declining importance with time.

Furthermore, Table 4 reveals that, for most of the calculated cases, the magnitude of the singular values is smaller than 1, meaning that the final perturbation is smaller in magnitude than the perturbation of initial values or emission rates. Considering that we apply singular vector analyses to find the initial and emission uncertainties that cause the largest error growth, a small error growth seemingly suggests that the benefit of singular vector analysis is small. However, it should be considered that we analyse only very restricted cases. Due to the focus on vertical profiles, the final projections cover only 5 to 10 grid points, and it can be expected that the magnitude of the final ozone perturbation will be smaller in amount than the magnitude of the locally unfocused initial value perturbation. For emission rates, the dynamics of the system is mainly limited by two features. 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 influence neither 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. Despite those restrictions, case 8a and case 8b (and case 5b for initial value optimization) show singular values greater than 1, proving the value of singular vector analysis even in the case of strongly restricted dynamics.

Table 4. Singular values (SV) with respect to initial values (iv) and emissions (em). $VP(t_F)$ denotes the considered final profile (numbers according to Table 1) and ML the associated model levels. Only simulations with a shared final profile $VP(t_F)$ are listed; "a" marks the simulation with the longer simulation interval and "b" the simulation with the shorter simulation interval. $t(a)$ and $t(b)$ are the associated simulation lengths.

$VP(t_F)$	ML	$t(b)$	$t(a)$	$SV_{iv}(b)$	$SV_{iv}(a)$	$SV_{em}(b)$	$SV_{em}(a)$
1	3–10	0 h 30 min	1 h 30 min	0.622	0.338	0.010	0.027
2	1–5	2 h 00 min	3 h 00 min	0.329	0.239	0.093	0.096
4	3–9	2 h 15 min	3 h 15 min	0.351	0.277	0.055	0.072
5	2–9	0 h 20 min	1 h 20 min	1.006	0.524	0.059	0.112
7	3–10	0 h 45 min	1 h 45 min	0.613	0.422	0.034	0.046
8	1–7	1 h 00 min	2 h 00 min	1.614	1.517	1.325	2.760
9	1–9	0 h 25 min	1 h 25 min	0.807	0.689	0.035	0.038
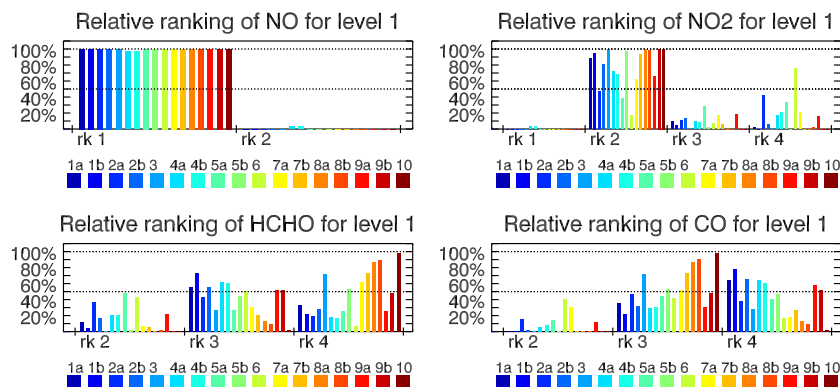
Figure 8. Relative ranking of the first singular vector with respect to emission uncertainties. Illustrated are results for NO (top left), NO_2 (top right), HCHO (bottom left), and CO (bottom right) at surface level for all 17 case studies. Plotting conventions as in Fig. 6.

6 Summary and conclusions

The EURAD-IM has been augmented to allow for singular vector analysis (SVA), resulting in the new EURAD-IM-SVA v1.0 model. The purpose of the EURAD-IM-SVA v1.0 is the calculation of the most sensitive chemical configuration with respect to initial values and emissions. The calculated sensitive configurations can be utilized to stabilize the chemical forecast by targeting sensitive system states for additional measurements. In this manner, the new tool can especially be applied for effective campaign planning.

In the framework of the model augmentation, newly coded or embedded routines are tested for accuracy. Within the limits of numerical precision, single routines as well as the complete model demonstrate correctness. Subsequently, the EURAD-IM-SVA v1.0 is evaluated by conducting a set of case studies based on the accomplished ZEPTER-2 campaign. Here, we evaluate the importance of measurements with regards to their ability to improve the forecast for locally predetermined ozone profiles. We investigate the influence of additional measurements of O_3, NO, NO_2, HCHO, CO, HONO, and OH. Since the considered simulation cases focus on the chemistry of ozone production and advection–diffusion dynamics in selected areas, they allow for a retrac-

ing of the results and a confirmation of their correctness. Elementary examples are presented, illustrating performance and interpretation of singular vectors for observation targeting.

Results of the singular vector decomposition with respect to initial values reveal that the optimal placement for additional observations is linked to height, with observations being more important at lower elevation where most of the chemical production of ozone takes place. Here, optimal target areas are controlled by mixing ratios of ozone precursors and their photochemical lifetimes, as well as transport and diffusion processes. In terms of a relevance ranking of chemical species, the measurement priority of species differs location-wise, dependent on initial concentrations and the importance of the precursor in the chemical formation of ozone. Overall, O_3 has most relevance among the considered species, while NO, NO_2, CO, and HCHO show medium relevance, and OH and HONO have least relevance. The revealed measurement priority meets our expectations as NO_x, CO, and volatile organic compounds are important precursors of ozone (Seinfeld and Pandis, 1998).

The singular vector decomposition with respect to emissions shows that the optimal placement of measurements of emission factors is strongly dependent on the location of
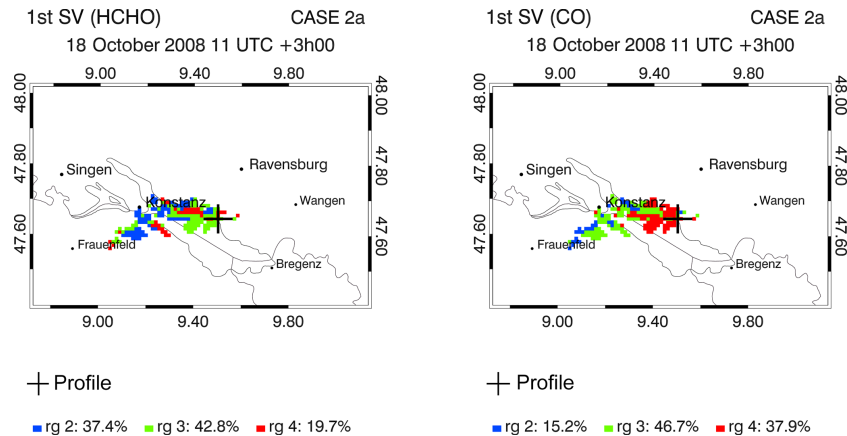
Figure 9. Spatially dependent measurement priorities of the first singular vector with respect to emission uncertainties. Illustrated are results for HCHO (left) and CO (right) at surface level for case 2a. Please note that the ranking is only depicted within the area of the relevance ranking. For each panel, the horizontal position of the final profile VP(t_F) is indicated with a black cross and the colour coding of each rank is denoted below each panel.

emission sources. When considering the relevance ranking of considered emitted species, we find that, for most cases, the influence of emissions of NO is most important, followed by emissions of NO_2, which of course are chemically closely linked. In these cases, a choice between both compounds for measurement network design may follow practical considerations. The importance of emissions of CO and HCHO, in the majority of cases, alternates between third and fourth rank.

Considering the error growth of uncertainties in initial values and emission strength, we find that the influence of singular values with respect to initial values decreases with growing simulation length, whereas the influence of singular values with respect to emissions increases. Due to short simulation intervals and focus on selected ozone profiles at the end of the simulation, the error growth is smaller than 1 in most of

the cases, meaning that the final uncertainty is smaller in percentage than the initial uncertainty. Yet, there are also cases that show singular values greater than 1, proving the value of singular vector analysis even in the case of strongly restricted dynamics.

Altogether, our case study shows that the newly designed EURAD-IM-SVA v1.0 is a powerful tool which identifies critical chemical species and chemical locations with respect to initial values and emissions. Both optimal placement of measurements and relevance ranking of chemical compounds confirm the benefit of singular vectors for measurement selection guidance. This can be applied for effective campaign planning. Furthermore, the detected directions of largest error growth can be employed to initialize ensemble forecasts and to model covariances.

Appendix A: Usage of singular vectors for determining targeted observations

For three-dimensional chemical transport models, a singular vector $\boldsymbol{v}$ comprises vector entries $v(i, j, k, s)$ for each chemical species s and each grid point (i, j, k) (i and j indicate horizontal grid coordinates, while k denotes the considered vertical model level), referring to each species' local sensitivity to perturbations of initial values or emissions. This set of vector entries can be analysed in terms of (a) optimal placement of observations and (b) measurement priority of considered species.

A1 Horizontal and vertical placement

The optimal observation location for a given species s is determined by the magnitudes of the singular vector entries $v(i, j, k, s)$ with i, j, k variable and s fixed. Accordingly, the grid point with largest magnitude defines the optimal placement for a considered species s.

We analyse the optimal placement in terms of vertical and horizontal optimal placement. The *horizontal placement* disregards effects of the vertical distribution and of different species' magnitudes, answering the question of optimal placement in a given horizontal plane:

$$v_{\mathrm{h}}(i, j, k, s) = \frac{v(i, j, k, s)}{|v(k, s)|}, \tag{A1}$$

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

Here, each horizontal section of the singular vector $\boldsymbol{v}$ with fixed level k and fixed species s is scaled by its length $|v(k, s)|$. In this manner, the combined singular vector entries of each horizontal plane of a given species have unit length and allow for a horizontal placement comparison between species. The modified singular vector $\boldsymbol{v}_{\mathrm{h}}$ with entries $v_{\mathrm{h}}(i, j, k, s)$ is referred to as a horizontal singular vector.

Likewise, for the *vertical placement*, we want to yield placement priorities with respect to vertical levels. 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 all perturbations associated with species s:

$$v_{\mathrm{v}}(k, s) = \frac{|v(k, s)|}{|v(s)|}, \tag{A2}$$

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

The vector $\boldsymbol{v}_{\mathrm{v}}$ with entries $v_{\mathrm{v}}(k, s)$ is defined as a vertical singular vector. In terms of optimal placement, both vertical and horizontal singular vectors allow for direct comparison of local sensitivities of different species.

A2 Relative rankings of chemical compounds

A measurement priority 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.

Since the measurement priority of species s may differ for each considered grid point (i, j, k), we are interested in gaining a picture representative of a specific height level. Accordingly, we select an area that is large enough to contain different air masses (here: all grid points with $\sqrt{\sum_s v(i, j, k, s)^2} > 10^{-4}$). Within the considered area, we establish a *relative ranking* $\mathrm{rk}(k, s)$ for each species s and each model level k. Each relative ranking $\mathrm{rk}(k, s)$ comprises the relative ranks $\mathrm{rk}^m(k, s)$, $m = 1, \ldots, n$ (where n is the number of considered species). The relative rank $\mathrm{rk}^m(k, s)$ simply counts how often the measurement priority of species s is ranked mth within the considered area of level k and then divides this number by the number of considered grid points:

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

$$p(i, j, k, s) := \begin{cases} 1, & \text{if } \sqrt{\sum_s v(i, j, k, s)^2} > 10^{-4} \\ 0, & \text{elsewhere,} \end{cases} \tag{A3}$$

$$r(i, j, k, s) := \begin{cases} 1, & \text{if } s \text{ is ranked } m\text{th in } (i, j, k) \\ 0, & \text{elsewhere.} \end{cases}$$

In this manner, a general measurement priority is provided for the selected area.

Code availability

References

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

Baker, N. L. and Daley, R.: Observation and background adjoint sensitivity in the adaptive observation-targeting problem, Q. J. Roy. Meteor. Soc., 126, 1431–1454, 2000.

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

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

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

Bott, A.: A positive definit advection scheme obtained by nonlinear renormalization of the advective fluxes, Mon. Weather Rev., 117, 1006–1015, 1989.

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

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

Daescu, D. N. and Carmichael, G. R.: An Adjoint Sensitivity Method for the Adaptive Location of the Observations in Air Quality Modeling, J. Atmos. Sci., 60, 434–450, 2003.

Daescu, D. N. and Navon, I. M.: Adaptive observations in the context of 4D-Var data assimilation, Meteorol. Atmos. Phys., 85, 205–226, 2004.

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

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

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.

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

Gelaro, R., Langland, R., Rohaly, G. D., and Rosmond, T. E.: An assessment of the singular vector approach to targeted observations using the FASTEX data set, Q. J. Roy. Meteor. Soc., 125, 3299–3327, 1999.

Goris, N. and Elbern, H.: Singular vector decomposition for sensitivity analyses of tropospheric chemical scenarios, Atmos. Chem. Phys., 13, 5063–5087, doi:10.5194/acp-13-5063-2013, 2013.

Granier, C., Bessagnet, B., Bond, T., D'Angiola, A., Denier van der Gon, H., Frost, G. J., Heil, A., Kaiser, J. W., Kinne, S., Klimont, Z., Kloster, S., Lamarque, J.-F., Liousse, C., Masui, T., Meleux, F., Mieville, A., Ohara, T., Raut, J.-C., Riahi, K., Schultz, M. G., Smith, S. J., Thompson, A., van Aardenne, J., van der Werf, G. R., and van Vuuren, D. P.: Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional scales during the 1980–2010 period, Climatic Change, 109, 163–190, 2011.

Grell, G. A., Dudhia, J., and Stauffer, D. R.: A description of the fifth-generation Penn State/ NCAR mesoscale model (MM5), NCAR Technical Note, NCAR/TN-398+STR, Boulder, 117 pp., 1994.

Kalnay, E.: Atmospheric Modeling, Data Assimilation, and Predictability, Cambridge University Press, Cambridge, 2002.

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.

Kim, H. M., Kim, S.-M., and Jung, B.-J.: Real-Time Adaptive Observation Guidance Using Singular Vectors for Typhoon Jangmi (200815) in T-PARC 2008, Weather Forecast., 26, 634–649, 2011.

Langland, R. H., Toth, Z., Gelaro, R., Szunyogh, I., Shapiro, M. A., Majumdar, S. J., Morss, R. E., Rohaly, G. D., Velden, C., Bond, N., and Bishop, C. H.: The North Pacific Experi-

ment (NORPEX-98): Targeted Observations for Improved North American Weather Forecasts, B. Am. Meteorol. Soc., 80, 1363–1384, 1999.

Lapidus, L. and Finder, G. F.: Numerical Solution of Partial Differential Equations in Science and Engineering, John Wiley and Sons, New York, New York, 1982.

Lawrence, M. G., Hov, O., Beekman, M., Brandt, J., Elbern, H., Eskes, H., Feichter, H., and Takigawa, M.: The Chemical Weather, Environ. Chem., 2, 1–3, doi:10.1071/EN05014, 2005.

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.

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

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

Maschho, K. J. and Sorensen, D. C.: A portable implementation of ARPACK for distributed memory parallel architectures, Proceedings of the Copper Mountain Conference on Iterative Methods, 1996.

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

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.

Oebel, A., Broch, S., Raak, D., Bohn, B., Rohrer, F., Hofzumahaus, A., Holland, F., and Wahner, A.: In-situ measurements of vertical profiles of chemical tracers in the PBL using the airship Zeppelin NT, Proceedings of the 15th International Symposium for the Advancement of Boundary Layer Remote Sensing (ISARS 2010), Paris, France, 2010.

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

Parlett, B. N.: The Symmetric Eigenvalue Problem, 2nd Edn., SIAM, Philadelphia, 1998.

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.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric chemistry and physics, Wiley-Interscience, John Wiley and Sons, Inc., Hoboken, New Jersey, 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.

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

Wintel, J., Hösen, E., Koppmann, R., Krebsbach, M., Hofzumahaus, A., and Rohrer, F.: Stable carbon isotope ratios of toluene in the boundary layer and the lower free troposphere, Atmos. Chem. Phys., 13, 11059–11071, doi:10.5194/acp-13-11059-2013, 2013.

Yanenko, N. N.: The method of fractional steps, Springer-Verlag, New York, 1971.