



# The assessment of potential observability for joint chemical states and emissions in atmospheric modelings

Xueran Wu<sup>1,2,3</sup> · Hendrik Elbern<sup>1,2</sup> · Birgit Jacob<sup>3</sup>

Accepted: 6 October 2021 / Published online: 28 October 2021  
© The Author(s) 2021

## Abstract

In predictive geophysical model systems, uncertain initial values and model parameters jointly influence the temporal evolution of the system. This renders initial-value-only optimization by traditional data assimilation methods as insufficient. However, blindly extending the optimization parameter set jeopardizes the validity of the resulting analysis because of the increase of the ill-posedness of the inversion task. Hence, it becomes important to assess the potential observability of measurement networks for model state and parameters in atmospheric modelings in advance of the optimization. In this paper, we novelly establish the dynamic model of emission rates and extend the transport-diffusion model extended by emission rates. Considering the Kalman smoother as underlying assimilation technique, we develop a quantitative assessment method to evaluate the potential observability and the sensitivity of observation networks to initial values and emission rates jointly. This benefits us to determine the optimizable parameters to observation configurations before the data assimilation procedure and make the optimization more efficiently. For high-dimensional models in practical applications, we derive an ensemble based version of the approach and give several elementary experiments for illustrations.

**Keywords** Data assimilation · Kalman smoother · Singular value decomposition · The degree of freedom for signal · Observability

## 1 Introduction

Climate change and air quality are influenced by fluxes of green house gases, reactive gas emissions and aerosols. The temporal evolution of reactive chemistry in the atmosphere is usually modeled by atmospheric chemistry transport models. Poorly known initial values, sources and sinks cause a serious problem for the quality of the simulation addressed by data assimilation and inverse modeling (e.g. Sandu and Chai (2011)). It is a typical situation in

practice data assimilation that the number of observations are markedly lower than the model degree of freedom (see Daley (1991)). In order to improve the quality of data assimilation and inverse modeling results, several aspects can be considered.

Firstly, the observation network can be optimized optimization problem subject to given external constraints, which has been addressed traditionally by Observation System Simulation Experiments [OSSEs, e.g. Daley (1991)]. The advanced concept of targeted observations has been popularized during the FASTEX campaign [e.g. Langland et al. (1999), Szunyogh et al. (1999)]. Theoretical studies are presented, for example by Berliner et al. (1999), or recently by Bellsky et al. (2014) for a case of study of highly nonlinear dynamics and Wu et al. (2016) for the optimal deployment of observations for time-varying system in a infinite dimensional domain within a finite-time interval. Secondly, the problem addressing the benefit assessment of individual observations or types of measurements has been investigated by Cardinali et al. (2004)

✉ Xueran Wu  
x.wu@fz-juelich.de

<sup>1</sup> IEK-8, Forschungszentrum Jülich, Wilhelm-Johnen-Straße, 52428 Jülich, Germany

<sup>2</sup> Rhenish Institute for Environmental Research at the University of Cologne, Pohligr. 3, 50969 Cologne, Germany

<sup>3</sup> Mathematics Department, University of Wuppertal, Gaußstraße 20, 42119 Wuppertal, Germany

and a sequence of related papers. Thirdly, the need to quantify the information provided by the observations can be satisfied by suitably selected measurements. Singular value decomposition (SVD) is a well-known tool applied to identify the priorities of observations by detecting the fastest growing uncertainties in meteorological models (e.g. Bousserez and Henze (2018), Buizza and Palmer (1995), Kang and Xu (2012), Khattatov et al. (1999), Liao et al. (2006), Lorenz (1965), Sandu et al. (2013), Singh et al. (2013), Spantini et al. (2015)). Besides, due to the importance of the sensitivity analysis of the model evolution with respect to the model errors and the observation network, the concept of the degree of freedom for signal (DFS) is also frequently applied to satellite retrieval problems, typically of significantly lower dimensions when compared to data assimilation [see for example Eyre (1990), Fisher (2003), Martynenko et al. (2010), Rabier et al. (2002), Rodgers (2000)]. Several methodologies have been proposed to account for model and estimation errors in both variational and ensemble data assimilation [e.g. Belsky et al. (2014), Daescu and Navon (2004), Gautam et al. (2021), Gautam et al. (2021), Li et al. (2009), Sharma et al. (2014), Smith et al. (2013), Zupanski et al. (2007)]. Navon (1997) outlined the perceptibility and stability in optimal parameter estimation in meteorology and oceanography. Cioaca and Sandu (2014b) introduced a general framework to optimize a set of parameters controlling the 4D-var data assimilation system and it was applied into shallow water model state and other parameter in Cioaca and Sandu (2014a).

Most studies cited above were based on the classical data assimilation problem, of which the initial value or the prognostic state variable is the only parameter to be optimized. However, for chemistry transport or greenhouse gas models driven by emissions in the troposphere, the optimization of emission rates is at least as important as the initial value in the data assimilation. In order to get better analysis from combining the model with observations, efforts of joint optimization have been made by adding emission rates to concentrations in amount of manuscripts, such as Bocquet and Sakov (2013), Elbern and Schmidt (1999), Elbern et al. (2000), Elbern et al. (2007), Goris and Elbern (2013), Goris and Elbern (2015), Miyazaki et al. (2012), Winiarek et al. (2014). However, the lack of ability to observe and estimate surface emission fluxes directly is still a major roadblock hampering the progress in predictive skills of climate and atmospheric chemistry models. For example studies with focus on urban area emissions with interactions of longer range transport need to discriminate between initial values driven by external emissions, distant emissions like biomass burning, and local emissions [e.g. Duarte et al. (2021), Kaskaoutis et al. (2014), Kumar et al. (2019)]. In all cases, the question of

model and emission uncertainty is crucial, as it blurs the model based capacities to discriminate between initial value and emission rate controlled simulations. This is especially challenging in case of biogenic emissions, which require related knowledge on plant and soil properties [e.g. Vogel and Elbern (2021a), Vogel and Elbern (2021b)].

Therefore, in this paper we novelly establish the dynamic model of emission rates and extended the traditional chemistry transport model. It provide the initial value and emission rates the same importance in the optimization. Based on the extended model, we investigate an approach to identify and assess the potential observability and sensitivity of any given observation configuration to the initial value and emission rates individually for atmospheric transport diffusion models by considering the (ensemble) Kalman smoother as underlying data assimilation method. Further, through the sensitivities of the initial value and emission rates, we have the opportunity to quantitatively balance weights between the initial value and emission rates. Those can help us determining the sensitive parameters and quantitatively build the initial variance matrices for both concentrations and emission rates in advance of the data assimilation process so that the computation resource can be saved and the accuracy of optimization can be improved.

The rest of paper is organized as follows. In Sect. 2, we establish the dynamic model for emissions and extend the original atmospheric transport model by emission rates in a novel way. In Sect. 3, based on the Kalman smoother, we present the specific approach to determine the degree of freedom for signal for both initial values and emission rates. In Sect. 4, we develop the ensemble approach to evaluate the degree of freedom for signal of the initial value and emission rates based on the non-singularity of the background covariance matrix. In Sect. 5, we identify the sensitive directions of the initial values and emission rates separately through maximizing the ratio of magnitudes of observation perturbation and the initial perturbation. It also provides us a possibility to estimate sensitivities of the initial value and emission rates by few leading singular values and singular vectors. In Sect. 6, we extend a 3D advection-diffusion equation with the dynamic model of the emission rate and give several elementary experiments to verify and demonstrate the ensemble approach. Finally, in Sect. 7, we conclude main contributions of this paper and discuss possible extensions.

## 2 Atmospheric inverse modeling extended by emission rates

We usually describe the concentration change rate by the following prognostic atmospheric transport model

$$\frac{dc(t)}{dt} = \mathcal{A}(c) + e(t), \tag{1}$$

where  $\mathcal{A}$  is a nonlinear model operator,  $c(t)$  and  $e(t)$  are the state vector of chemical constituents and emission rates at time  $t$ , respectively.

The prior estimate of the state vector of concentrations  $c(t)$  is given and denoted by  $c_b(t)$ , termed as the background state. The prior estimate of emission rates, usually taken from emission inventories, is denoted by  $e_b(t)$ .

Let  $\mathbf{A}$  be the tangent linear operator of  $\mathcal{A}$ ,  $\delta c(t_0) = c(t_0) - c_b(t_0)$  and  $\delta e(t) = e(t) - e_b(t)$ . The linear evolution of the perturbation of  $c(t)$  follows the tangent linear model as

$$\frac{d\delta c}{dt} = \mathbf{A}\delta c + \delta e(t). \tag{2}$$

By the discretization of the tangent linear model in space, it is straightforward to obtain the linear solution of (2) discretized in space and continuous in time as

$$\delta c(t) = M(t, t_0)\delta c(t_0) + \int_{t_0}^t M(t, s)\delta e(s)ds, \tag{3}$$

where  $M(\cdot, \cdot)$  is the resolvent obtained from the spatial discretization of  $\mathbf{A}$ . Without loss of generality, we assume  $\delta c(t) \in \mathbf{R}^n$ ,  $\delta e(t) \in \mathbf{R}^n$ , where  $n$  is the dimension of the partial phase space of concentrations and emission rates. Obviously,  $M(\cdot, \cdot) \in \mathbf{R}^{n \times n}$ .

In addition, let  $y(t)$  be the observation vector of  $c(t)$  and define

$$\delta y(t) = y(t) - \mathcal{H}(c_b)(t), \tag{4}$$

where  $\delta y(t) \in \mathbf{R}^{m(t)}$ ,  $m(t)$  is the dimension of the phase space of observation configurations at time  $t$ .  $\mathcal{H}(t)$  is a nonlinear forward observation operator mapping the model space to the observation space. Linearizing the nonlinear operator  $\mathcal{H}$  as  $H$ , we present the observation system as

$$\delta y(t) = H(t)\delta c(t) + v(t), \tag{5}$$

where the observation error  $v(t)$  of the Gaussian distribution has zero mean and variance  $R(t) \in \mathbf{R}^{m(t) \times m(t)}$ .

The Kalman smoother is a recursive estimator to provide the best linear unbiased estimates (BLUE) of the unknown variables with error estimates, using a sequence of observations [e.g. Gelb (1974)]. In addition to 4D-Var approaches, Kalman smoothers not only can provide the best linear unbiased estimate by a series of observations over time for the state vector, but also update the forecasting error covariances of that estimate.

It is clear to see that if the initial state of concentrations is the only parameter to be optimized, we can only consider the concentrations as the model states and apply the Kalman filter and smoother into the tangent linear model (2)

with observations (5) directly. However, in most cases the exact values of emission rates are poorly known and also considered as parameters which need to be optimized. It has been shown by Elbern et al. (2007) that the diurnal profiles are better known than the exact amplitude of emission rates. Hence, we can only consider the amplitude of the diurnal emission cycle as optimization parameters. Thus, we first reformulate the background evolution of emission rates from time  $s$  to  $t$  in a dynamic form as an emission model

$$e_b(t) = M_e(t, s)e_b(s), \quad s \leq t, \tag{6}$$

where  $e_b(\cdot)$  is a  $n$ -dimensional vector, the  $i^{th}$  element of  $e_b(\cdot)$  is denoted by  $e_b^i(\cdot)$  and  $M_e(t, s)$  is the scaling diagonal matrix defined as

$$M_e(t, s) = \begin{pmatrix} \frac{e_b^1(t)}{e_b^1(s)} & & & & \\ & \frac{e_b^2(t)}{e_b^2(s)} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \frac{e_b^n(t)}{e_b^n(s)} \end{pmatrix}. \tag{7}$$

We assume that the amplitude of emission rates is the only parameter to be optimized and then establish the dynamic model of emission rates subject to the above background evolution

$$\delta e(t) = M_e(t, s)\delta e(s), \quad s \leq t. \tag{8}$$

Several studies [e.g. Gelb (1974)] stated that the estimation of the variable  $x$  by the fix-interval Kalman smoother generally equals to the conditional expectation based on the observations within the entire time interval, denoted by  $\mathbf{E}[x|\{y(t_{obs}), t_{obs} \in [t_0, t_N]\}]$ . With the emission model (8), the estimate of  $e(t)$  by Kalman smoother on  $[t_0, t_N]$  follows the linear property of the conditional expectation,

$$\begin{aligned} & \mathbf{E}[e(t)|\{y(t_{obs}), t_{obs} \in [t_0, t_N]\}] \\ &= \mathbf{E}[M_e(t, s)e(s)|\{y(t_{obs}), t_{obs} \in [t_0, t_N]\}] \\ &= M_e(t, s)\mathbf{E}[e(s)|\{y(t_{obs}), t_{obs} \in [t_0, t_N]\}]. \end{aligned} \tag{9}$$

It implies that BLUEs of emission rates with the dynamic model (8) preserve the same diurnal profiles of the background of emission rates.

By rewriting (3) as

$$\delta c(t) = M(t, t_0)\delta c(t_0) + \int_{t_0}^t M(t, s)M_e(s, t_0)\delta e(t_0)ds,$$

we obtain the transport model with the state vector extended by emission rates

$$\begin{pmatrix} \delta c(t) \\ \delta e(t) \end{pmatrix} = \begin{pmatrix} M(t, t_0) & \int_{t_0}^t M(t, s) M_e(s, t_0) ds \\ 0 & M_e(t, t_0) \end{pmatrix} \begin{pmatrix} \delta c(t_0) \\ \delta e(t_0) \end{pmatrix}. \tag{10}$$

Typically, there is no direct observation for emissions, apart from the flux tower observations used for carbon dioxide, which are not considered here. Therefore, we can reformulate the observation mapping as

$$\delta y(t) = (H(t), 0_{n \times n}) \begin{pmatrix} \delta c(t) \\ \delta e(t) \end{pmatrix} + v(t), \tag{11}$$

where  $0_{n \times n}$  is a  $n \times n$  matrix with zero elements.

It is now clear that both concentrations and emission rates are included in the state vector of the homogeneous model (10). It allows us to apply the Kalman smoother in a fixed time interval  $[t_0, t_N]$  in order to optimize both parameters. Besides, a more general case of the transport model extended by emission is shown in Appendix A.

### 3 The degree of freedom for signal of concentrations and emissions

In this section we will introduce the theoretical approach to determine the DFS of concentrations and emissions, resting on the extended model in Sect. 2. This approach gives us access to determine the potential ability of observations to optimize each variable of the above extended model, based on the Kalman smoother within a finite-time interval.

For convenience, we generalize the atmospheric transport model (10) by the following discrete-time linear system on the time interval  $[t_0, t_1, \dots, t_N]$  as

$$x(t_{k+1}) = M(t_{k+1}, t_k)x(t_k) + \varepsilon(t_k), \tag{12}$$

$$y(t_k) = H(t_k)x(t_k) + v(t_k), \tag{13}$$

where  $x(\cdot) \in \mathbf{R}^n$  is the state variable and  $y(t_k) \in \mathbf{R}^{m(t_k)}$  is the observation vector at time  $t_k$ . The model error  $\varepsilon(t_k)$  and the observation error  $v(t_k)$ ,  $k = 1, \dots, N$  of Gaussian distributions have zero means. The model error covariance matrix is denoted by  $Q(t_k)$  and the observation error covariance matrix is denoted by  $R(t_k)$ .

According to Appendix B, applying the singular value decomposition into  $P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^T\mathcal{R}^{-\frac{1}{2}} = VSU^T$ , we obtain

$$\begin{aligned} \tilde{P} &= P^{-\frac{1}{2}}(t_0|t_{-1})(P(t_0|t_{-1}) - P(t_0|t_N))P^{-\frac{1}{2}}(t_0|t_{-1}) \\ &= \sum_{i=1}^r \frac{s_i^2}{1 + s_i^2} v_i v_i^T, \end{aligned} \tag{14}$$

where  $v_i$  is the  $i^{th}$  left singular vector in  $V$  related to the singular value  $s_i$ , which is the  $i^{th}$  element on the diagonal of  $S$ .

It is clear that the trace of  $\tilde{P}$  can be used to evaluate the total improvements of model states. Thus, the nuclear norm is appropriately taken as the metric, which is defined as

$$\|A\|_1 = \text{tr}(\sqrt{A^T A}), \tag{15}$$

where  $A$  is any matrix and  $\text{tr}(\cdot)$  denotes the trace of the matrix.

From (14), we obtain

$$\|\tilde{P}\|_1 = \text{tr}(\tilde{P}) = \sum_{i=1}^r \frac{s_i^2}{1 + s_i^2}. \tag{16}$$

This is well-known as the degree of freedom for signal (DFS) of the model (e.g. Rodgers (2000)).

It is obvious that  $\|\tilde{P}\|_1 < \|I\|_1 = n$ . Here  $n$  can be considered as the total relative improvement if the system is definitely observed. Thus, if we consider the ratio

$$\tilde{p} = \frac{\|\tilde{P}\|_1}{\|I\|_1} = \frac{\|\tilde{P}\|_1}{n} \in [0, 1), \tag{17}$$

the percentage of the total improvement of the model is obtained, which is henceforth called the *relative degree of freedom for signal*.

In order to get a deeper insight into the potential capacity of the observation network to improve the estimation of all model states, we consider the corresponding value in the diagonal of  $\tilde{P}$  as *the contribution of the degree of freedom for signal*. We denote the  $j^{th}$  element on the diagonal of  $\tilde{P}$  by  $\tilde{P}_j$ . From (47), the contribution of the  $j^{th}$  element of  $x(t_0)$  to the degree of freedom for signal can be expressed as

$$\tilde{P}_j = \sum_{i=1}^r \frac{s_i^2}{1 + s_i^2} (v_{ij})^2, \tag{18}$$

where  $v_{ij}$  is the  $j^{th}$  element of  $v_i$ .

Besides, we can see that (14) enables us to discriminate the DFS contributed to different optimization parameters, which are here emission rates and the initial value. Without loss of generality, we divide (14) into the following block matrix according to the dimension of  $c$  and  $e$

$$\tilde{P} = \begin{pmatrix} \tilde{P}^c & \tilde{P}^{ce} \\ \tilde{P}^{ec} & \tilde{P}^e \end{pmatrix} = \sum_{i=1}^{2n} \frac{s_i^2}{1 + s_i^2} \begin{pmatrix} v_i^c \\ v_i^e \end{pmatrix} (v_i^c{}^T, v_i^e{}^T) \in \mathbf{R}^{2n \times 2n}, \tag{19}$$

where  $(v_i^c{}^T, v_i^e{}^T)^T = v_i$ .

It is easy to see that

$$\tilde{P}^c = \sum_{i=1}^{2n} \frac{s_i^2}{1 + s_i^2} v_i^c v_i^c{}^T, \quad \tilde{P}^e = \sum_{i=1}^{2n} \frac{s_i^2}{1 + s_i^2} v_i^e v_i^e{}^T. \tag{20}$$

Further, the degree of freedom for signal of  $j^{th}$  element in  $c(t_0)$  and  $e(t_0)$  are given by

$$\tilde{P}_j^c = \sum_{i=1}^{2n} \frac{s_i^2}{1 + s_i^2} (v_{ij}^c)^2, \quad \tilde{P}_j^e = \sum_{i=1}^{2n} \frac{s_i^2}{1 + s_i^2} (v_{ij}^e)^2, \quad (21)$$

where  $v_{ij}^c$  and  $v_{ij}^e$  are the  $j^{th}$  elements of  $v_i^c$  and  $v_i^e$ , respectively.

Moreover, the degree of freedom for signal of the concentration  $\|\tilde{P}^c\|_1$  and emission rates  $\|\tilde{P}^e\|_1$  are calculated by

$$\|\tilde{P}^c\|_1 = \sum_{i=1}^{2n} \frac{s_i^2}{1 + s_i^2} \text{tr}(v_i^c v_i^{c\top}), \quad \|\tilde{P}^e\|_1 = \sum_{i=1}^{2n} \frac{s_i^2}{1 + s_i^2} \text{tr}(v_i^e v_i^{e\top}). \quad (22)$$

It is worth noticing that

$$\tilde{P}^c = (P^c(t_0|t_{-1}))^{-\frac{1}{2}} (P^c(t_0|t_{-1}) - P^c(t_0|t_N)) (P^c(t_0|t_{-1}))^{-\frac{1}{2}} \quad (23)$$

$$\tilde{P}^e = (P^e(t_0|t_{-1}))^{-\frac{1}{2}} (P^e(t_0|t_{-1}) - P^e(t_0|t_N)) (P^e(t_0|t_{-1}))^{-\frac{1}{2}} \quad (24)$$

if and only if there is no prior correlation between the initial concentration and emission rates. In this case  $P^{ce}(t_0|t_{-1}) = 0_{n \times n}$ , the corresponding relative degrees of freedom for signal of the concentration and emission rates are defined as

$$\tilde{p}^c = \frac{\|\tilde{P}^c\|_1}{n}, \quad \tilde{p}^e = \frac{\|\tilde{P}^e\|_1}{n}. \quad (25)$$

From (17),  $\tilde{p}^c \in [0, 1]$  and  $\tilde{p}^e \in [0, 1]$  seem like percentages of the relative improvements of concentration and emission rates, respectively. However, efficient observation networks ideally lead to values are close to 1 for both of them, such that

$$\frac{\|\tilde{P}^c\|_1}{n} + \frac{\|\tilde{P}^e\|_1}{n} > 1. \quad (26)$$

It results from the reason that the normalization of  $\tilde{P}$  is only with respect to the extended covariance matrix  $P(t_0|t_{-1})$  rather than specified to the covariance matrices  $\tilde{P}^c(t_0|t_{-1})$  and  $\tilde{P}^e(t_0|t_{-1})$  individually. The relative degree of freedom for signal cannot serve our objective to distinguish the observability of the concentration and emission rates. By observing the block form of  $\tilde{P}$ , we have

$$\|\tilde{P}^c\|_1 + \|\tilde{P}^e\|_1 = \|\tilde{P}\|_1. \quad (27)$$

Thus, in order to compare the potential improvements of the concentration and emission rates separately, we define a relative ratio of the degree of freedom for signal for the concentration or emission rates as

$$\tilde{p}^c = \frac{\|\tilde{P}^c\|_1}{\|\tilde{P}\|_1}, \quad \tilde{p}^e = \frac{\|\tilde{P}^e\|_1}{\|\tilde{P}\|_1}, \quad \tilde{p}^c + \tilde{p}^e \equiv 1. \quad (28)$$

If the degree or relative degree of freedom for signal of the observation network within the assimilation window is almost zero, an improvement cannot be expected. In contrast,  $\{\tilde{P}_j^c\}_{j=1}^n$  and  $\{\tilde{P}_j^e\}_{j=1}^n$ , which show the improvement of each parameter  $j$  of concentrations and emission rates respectively, can help us determining which parameters can be expected to be optimized by the existing observation configurations. Furthermore, comparing  $\tilde{p}^c$  with  $\tilde{p}^e$ , we can conclude that the estimate of the one with the larger relative ratio of freedom for signal can be improved more efficiently by the existing observation configurations than the other. In other words, if  $\tilde{p}^c > \tilde{p}^e$ , the existing observation configuration is more sensitive to the initial values of concentrations. Conversely, if  $\tilde{p}^c < \tilde{p}^e$ , the observation configurations can improve the estimate of emission rates better. According to  $\tilde{p}^c$  and  $\tilde{p}^e$ , the relative weights between the concentration and emission rates can be identified quantitatively. In a data assimilation context, where observations are in a weighted relation to the background, the BLUE favors those parameters with higher observation efficiency.

The special case that  $\tilde{p}^e$  is very close to zero implies that observation network is not detectable for the emission-rate optimization.

### 4 The ensemble approach to determine the DFS

The ensemble Kalman smoother (EnKS) is a frequently applied tool for problems with a large number of control variables in the field of data assimilation [e.g. Anderson (2001), Evensen (2009)]. In this section, in order to identify the potential capacities of observation networks to optimize the concentration and especially the poorly known emission rates for high-dimensional problems, we will introduce the ensemble-based version of the approach in Sect. 3.

According to Appendix C, analog to Sect. 3, we have

$$\begin{aligned} \bar{P} &= \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})(\bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_N))\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) \\ &= \sum_{i=1}^r \frac{s_i^2}{1 + s_i^2} \bar{v}_i \bar{v}_i^\top. \end{aligned} \quad (29)$$

Similar to Sect. 3, we can also divide  $\bar{P}$  into the block form according to the dimensions of the concentration and emission rates. Correspondingly, we obtain the ensemble degree of freedom for signal of  $j^{th}$  element in  $c(t_0)$  and  $e(t_0)$

$$\bar{P}_j^c = \sum_{i=1}^r \frac{s_i^2}{1 + s_i^2} (\bar{v}_{ij}^c)^2, \quad \bar{P}_j^e = \sum_{i=1}^r \frac{s_i^2}{1 + s_i^2} (\bar{v}_{ij}^e)^2, \quad (30)$$

where  $\bar{v}_{ij}^c$  and  $\bar{v}_{ij}^e$  are the  $j^{th}$  elements of  $\bar{v}_i^c$  and  $\bar{v}_i^e$ , respectively and

$$(\bar{v}_i^{c\top}, \bar{v}_i^{e\top})^\top = \bar{v}_i.$$

We observe that (29) and (14) have a similar form. By virtue of

$$\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}_{xy}^f\mathcal{R}^{\leftarrow\frac{1}{2}} = \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^\top\mathcal{R}^{\leftarrow\frac{1}{2}}, \quad (31)$$

we can find that the final results of (14) and (63) are equivalent. However, compared with  $\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^\top\mathcal{R}^{\leftarrow\frac{1}{2}}$ , the ensemble expression  $\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}_{xy}^f\mathcal{R}^{\leftarrow\frac{1}{2}}$  processes the absolute advantage that in the calculation of  $\bar{P}_{xy}^f$  since we do not need the explicit form of  $\mathcal{G}$ . It allows us to code it line by line such that our approach is computationally more efficient.

Analog to Sect. 3, we can similarly define the ensemble degree of freedom for signal (EnDFS) as  $\|\bar{P}\|_1$  and consider each element on the diagonal of  $\bar{P}$  as the contribution to EnDFS of the corresponding model state.

In practice,  $\bar{P}(t_0|t_{-1})$  is typically singular, thus we have

$$\begin{aligned} &\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})(\bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_N))\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) \\ &= V_0\hat{S}_0^\dagger V_0^\top (V_0\hat{S}_0^2 V_0^\top) V_0\hat{S}_0^\dagger V_0^\top - \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}(t_0|t_N)\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) \\ &= V_0 I_{r_0} V_0^\top - \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}(t_0|t_N)\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}), \end{aligned} \quad (32)$$

where  $I_{r_0}$  is the diagonal matrix with the diagonal  $(\mathbf{1}_{1 \times r_0}, \mathbf{0}_{1 \times (n-r_0)})$ . It is clear from (60) that  $\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}(t_0|t_N)\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})$  is still a nonnegative definite matrix.

Thus, the ensemble relative degree of freedom for signal (EnRDFS) is defined by

$$\bar{p} = \frac{\|\bar{P}\|_1}{\|I_{r_0}\|_1} = \frac{\|\bar{P}\|_1}{r_0} \in [0, 1]. \quad (33)$$

In order to distinguish the potential observabilities for the concentration and emission rates, the ensemble relative ratios of DFS remain

$$\bar{p}^c = \frac{\|\bar{P}^c\|_1}{\|\bar{P}\|_1}, \quad \bar{p}^e = \frac{\|\bar{P}^e\|_1}{\|\bar{P}\|_1}. \quad (34)$$

### 5 The sensitivity of observation networks

The above discussion about DFS aims to evaluate the capacity of a predefined measurement network to optimize the initial value and emission rates simultaneously. In Appendix D, independent of any concrete data assimilation method, we use the singular vector approach [see Buizza and Montani (1999), Buizza and Palmer (1995), Liao et al. (2006) etc.] to identify sensitive directions of observation networks to the initial value and emission rates separately and show the association with Sect. 3.

From Appendix D, we can see that the singular value  $s_k$  shows the amplification of the impact of the initial state to observation configurations in the entire time interval. The associated singular vector in the state space  $v_k$  is the direction of the  $k^{th}$  growth of the perturbation of observations evolving from the initial perturbation. With the special choice  $W_0 = P^{-1}(t_0|t_{-1})$  and  $\mathcal{W} = \mathcal{R}^{-1}$ , we compare the sensitivity analysis with the discussion in Sect. 3. It is clear that the vector  $v_k$  also points to the  $k^{th}$  direction which maximizes the relative improvement of estimates based on the Kalman smoother. It indicates that the states with higher contributions to DFS are the same with the states, which are more sensitive to the observation networks. Besides, the leading singular value  $s_1$  is related to the operator norm of  $\tilde{P}$  as

$$\|\tilde{P}\| = \max_{\|x\|=1} \|\tilde{P}x\| = \frac{s_1^2}{1 + s_1^2}, \quad (35)$$

which implies the upper boundedness of  $\tilde{P}$ . It gives us an access to approximate and target sensitive parameters or areas with the metric of the leading singular vectors weighted by the corresponding singular values.

Moreover, due to the homogeneity of the atmospheric transport model state vector extended with emissions, the above sensitivity analysis can be easily applied by dividing singular vectors into the block form according to the dimensions of the initial state and emissions. The corresponding blocks of different singular vectors indicate the different sensitive directions of the initial state and emissions and allow for this relative quantification. Correspondingly, we can approximate and target parameters sensitive to the existing observation networks for both the initial value and emission rates, respectively.

### 6 Experiment

In this section, we apply the approaches in Sects. 4 and 5 into an elementary advection-diffusion model to show how to assess the potential observability of concentrations and emission rates through EnDFS. We can see how it helps to

identify the sensitive parameters of both concentrations and emission rates to the given observations. We consider a linear advection-diffusion model with Dirichlet horizontal (lateral) boundary condition and Neumann lower (surface) boundary condition in the vertical direction on the domain  $[0, 14] \times [0, 14] \times [0, 4]$  as follows,

$$\frac{\partial \delta c}{\partial t} = -v_x \frac{\partial \delta c}{\partial x} - v_y \frac{\partial \delta c}{\partial y} + \frac{\partial}{\partial z} (K(z) \frac{\partial \delta c}{\partial z}) + \delta e, \tag{36}$$

where  $\delta c, \delta e$  are the perturbations of the concentration and the emission rates respectively.  $K(z)$  is a differentiable function of height  $z$ .

In this example, the vertical coupling of horizontal grid layers is accomplished only by a diffusion operator to avoid signal imprints following some arbitrarily designed small scale vertical advection patterns. This is considered as valid, since the information loss by diffusion induced reduction of the noise ratio due to the signal diffusion analogue is significantly stronger than in case of advection

We assume the velocities  $v_x = v_y = 0.5$  and the time step  $\Delta t = 0.5$  and the numerical solution is based on the symmetric operator splitting technique [see Yanenko (1971)] with the following operator sequence

$$\delta c(t + \Delta t) = T_x T_y D_z A D_z T_y T_x \delta c(t), \tag{37}$$

where  $T_x$  and  $T_y$  are transport operators in horizontal directions  $x$  and  $y$ ,  $D_z$  is the diffusion operator in the vertical direction  $z$ . The parameters of emission and deposition rates are included in  $A$ . The Lax-Wendroff algorithm is chosen as the discretization method for horizontal advection with  $\Delta x = \Delta y = 1$ . The vertical diffusion is discretized with  $\Delta z = 1$  by Crank-Nicolson scheme with the Thomas algorithm [see Higham (2002)] as solver. The number of the grid points  $N_g = 1125$ .

With the same temporal and spatial discretization of the concentration, the background knowledge of the emission rates is given by  $e_b(t_n, i, j, l)$ ,  $n = 1, \dots, N$ . We establish

the discrete dynamic model of the emission rates according to (8)

$$\delta e(t_{n+1}) = M_e(t_{n+1}, t_n) \delta e(t_n), \quad n = 1, \dots, N, \tag{38}$$

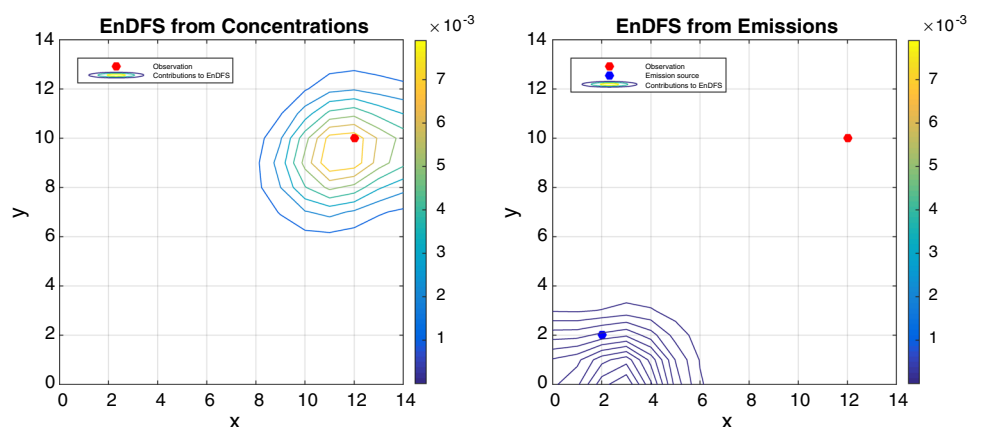
where  $M_e(t_{n+1}, t_n) = e_b(t_{n+1})/e_b(t_n)$ .

In this section, we assume  $\delta d$  is a constant over time and the observation operator  $H(t)$  mapping the state space to the observation space is a  $1 \times 2N_g$  time-invariant matrix. In Wu et al. (2016), the convergence of the numerical solution based on the above splitting and discretization method to the original solution of the partial differential equation (36) has been proved.

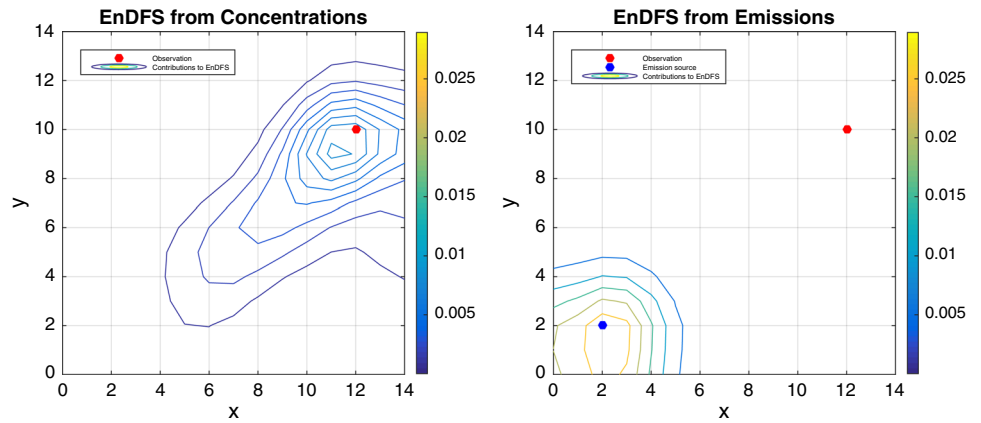
In our simulations, we produce  $q = 500$  (the ensemble number) samples for the initial concentration and emission rates respectively by pseudo independent random numbers and make the states correlated by the moving average technique. It has been tested that the computation cost of our approach is linearly increasing with the number of ensembles. In the following, we present three different tests, aiming to demonstrate roles of variable winds, emissions, and vertical diffusion.

*Advection tests:* The following part demonstrates the potential capacity and limits of the DFS analysis tool. The prototypical examples are designed to show the expected elementary outcomes of the following situations. They exhibit the effects of assimilation window length in relation to emission location. These include (i) an assimilation window too short to capture emission impacts at the observation site, (ii) an extended assimilation window with balanced signal of impacts of concentrations and emissions at the observation site, (iii) a further increased assimilation window featuring a declining impact of initial values and growing emission impact. The first elementary advection test (Figs. 1, 2, 3, 4, 5, 6, and 7) identifies the sensitivities of parameters subject to different wind direction and data assimilation window through the EnDFS of each element of the concentrations and emission rates. Focusing on the

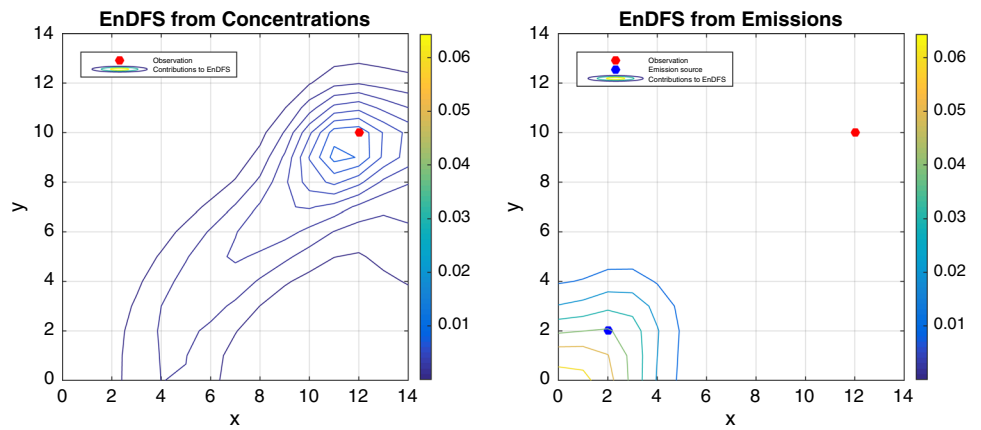
**Fig. 1** Advection test with 10Δt data assimilation window (DAW) and southwesterly wind. The contributions from the concentration and emission to EnDFS are shown in the left and right figure panels respectively. The blue point located at (12, 10, 0) shows the time-invariant observation configuration. The blue point located at (2, 2, 0) is the source of the emission rate



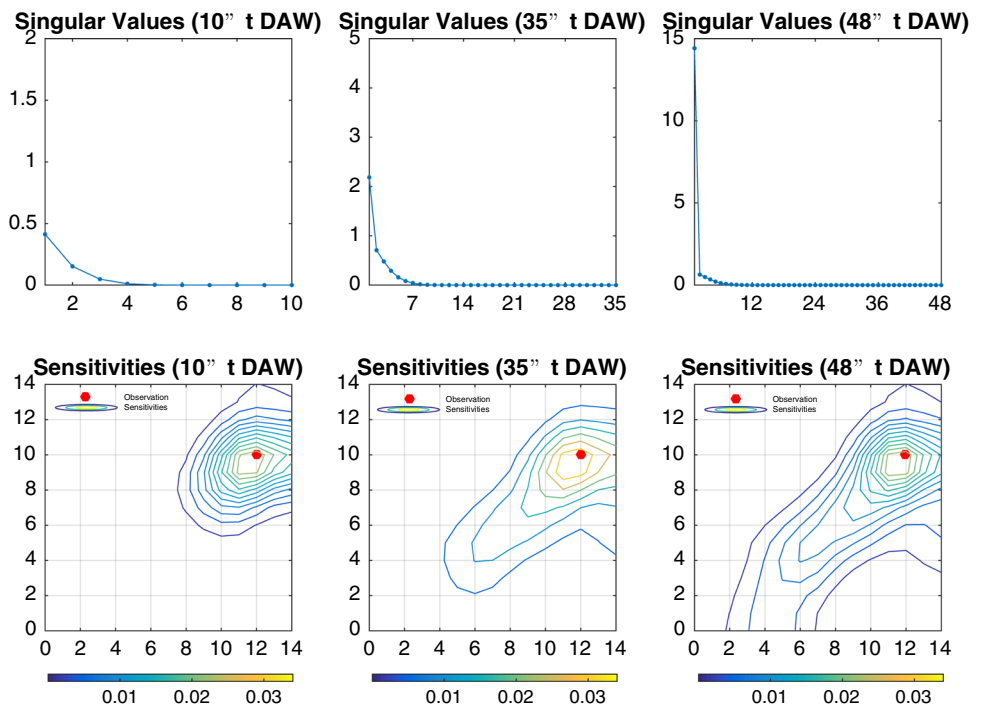
**Fig. 2** Advection test with  $35\Delta t$  DAW and southwesterly wind. Plotting conventions are as in Fig. 1



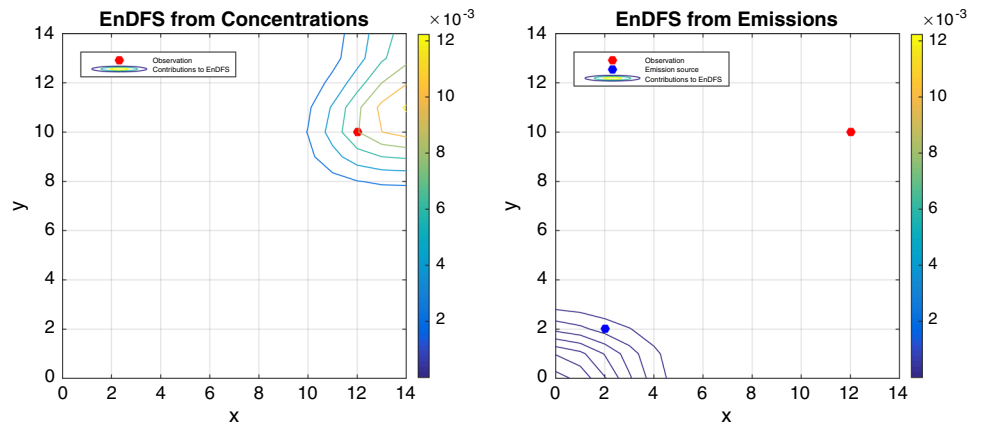
**Fig. 3** Advection test with  $48\Delta t$  DAW and southwesterly wind. Plotting conventions are as in Fig. 1



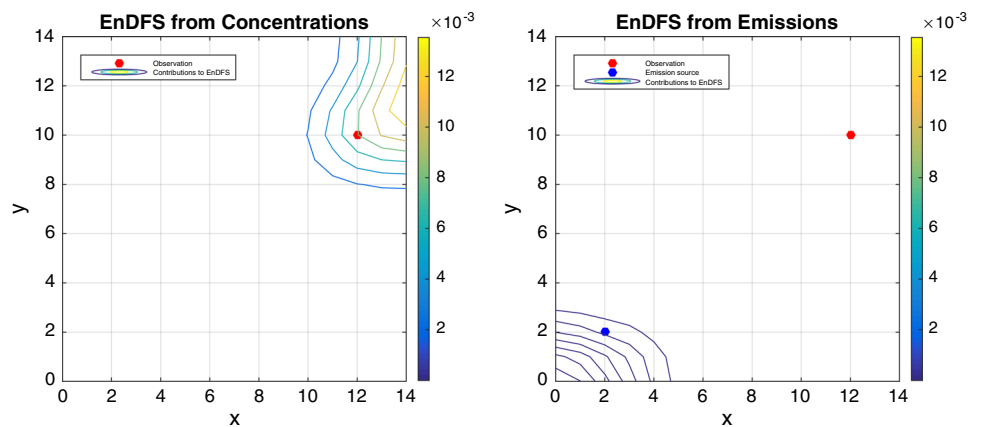
**Fig. 4** The panels in the first row are singular values of the advection tests shown in the left panels of Fig. 1 ~ 3. The panels in the second row show the corresponding sensitivities of concentrations at the initial time approximated by 5 leading singular vectors



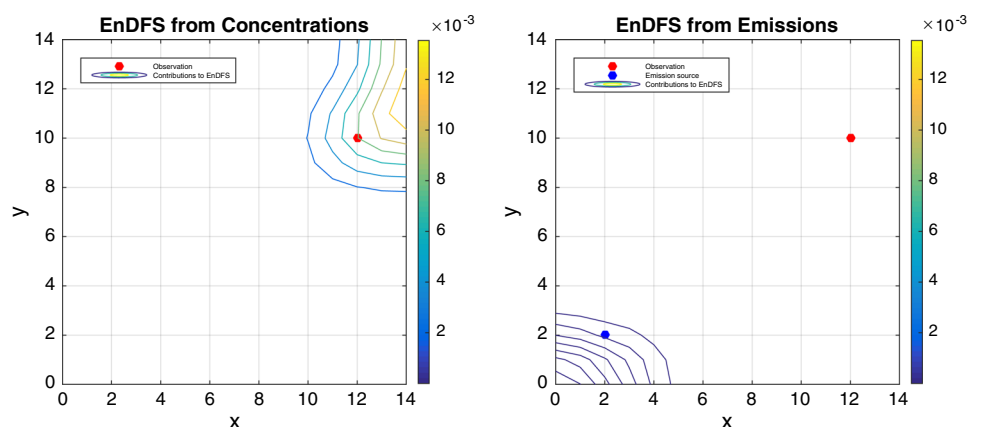
**Fig. 5** Advection test with  $10\Delta t$  DAW and northeasterly wind. Plotting conventions are as in Fig. 1



**Fig. 6** Advection test with  $35\Delta t$  DAW and northeasterly wind. Plotting conventions are as in Fig. 1



**Fig. 7** Advection test with  $48\Delta t$  DAW and northeasterly wind. Plotting conventions are as in Fig. 1



advection effects, we apply the model with a weak diffusion process ( $K(z) = 0.5e^{-z^2}$ ).

In Figs. 1, 2, and 3 we assume southwesterly winds and data assimilation windows are  $10\Delta t$ ,  $35\Delta t$  and  $48\Delta t$ , respectively. The computation times are approximately 8.1s, 28.5s and 39.4s in our tests with the above three different assimilation windows, from which we can verify that the computation cost is nearly linearly increasing with the length of data assimilation window. The contributions

to EnDFS of the initial states are shown in the left panels of Figs. 1, 2, and 3. We can find that in the horizontal field at lowest layer ( $z = 0$ ), the optimized field of the concentration is enlarged with the extension of data assimilation windows. This is because an increased domain of the concentration are controlled with longer data assimilation windows.

The right panels of Figs. 1, 2, and 3 show EnDFS of the emission rate at each grid point with  $z = 0$ . From the right

**Table 1** Ensemble relative ratios of the initial value and emission rate at the lowest layer

	Fig. 1	Fig. 2	Fig. 3	Fig. 5	Fig. 6	Fig. 7	Fig. 8	Fig. 9	Fig. 10	Fig. 11
$\bar{p}^c$	0.992	0.455	0.285	0.994	0.994	0.994	0.225	0.181	0.993	0.191
$\bar{p}^e$	0.008	0.545	0.715	0.006	0.006	0.006	0.775	0.8192	0.007	0.890

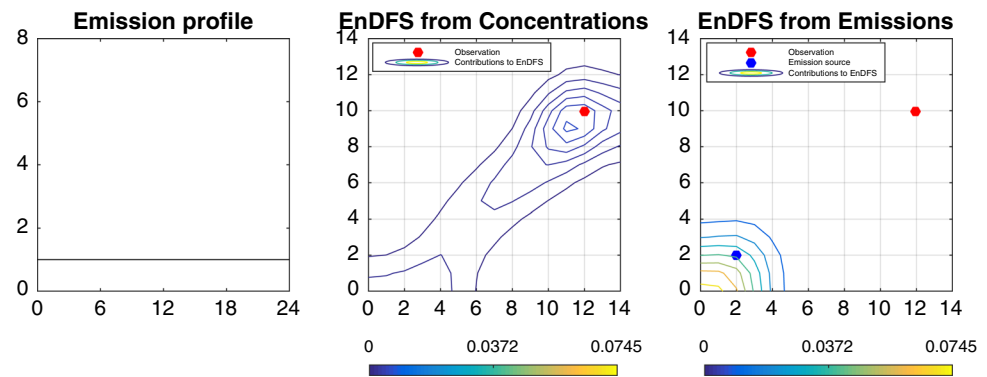
panel of Fig. 1, we can observe that contributions to EnDFS from emissions are less than  $2 \times 10^{-3}$ . Compared with the right panel of Fig. 1, the EnDFS of the emissions are obviously smaller than the EnDFS of the concentration in the influenced area. It indicates that the observations cannot detect the emission rates within  $10\Delta t$  data assimilation window. Thus, in this case initial values of the area adjacent to the observation site are alone optimized. It is shown in the right panels of Figs. 2 and 3 that emission rates play a more and more important role on the impact of observations. In this two cases, we consider both the concentration and emission rate as optimizable parameters. The quantitative balance between the concentration and emission rates is provided in Table 1.

The upper row panels of Fig. 4 exhibit singular values corresponding to results shown in Figs. 1, 2 and 3. We approximate sensitivities of the initial concentration by the first five leading singular vectors weighted by associated singular values in the nuclear norm and show results in the

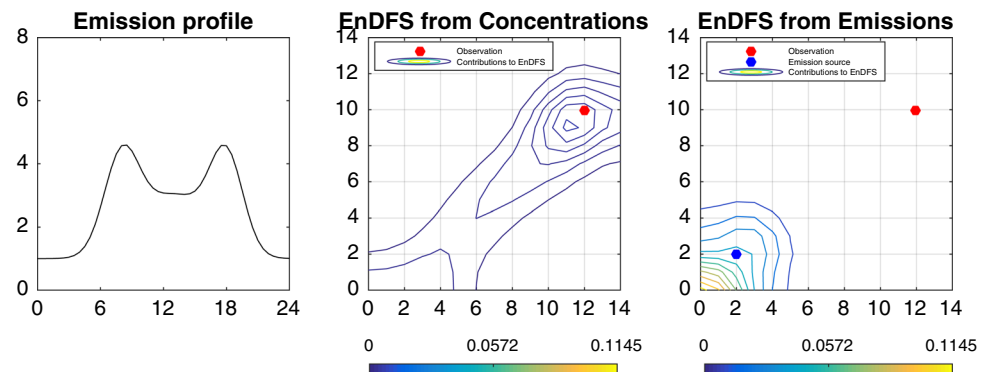
lower row panels of Fig. 4. It is verified that the sensitive area can be well targeted by only few singular vectors, although the sensitivity analysis cannot provide the quantitative solutions with a clear statistical significance as the DFS of the model. Besides, in line with expectations, the area influenced by the observation configuration depends on wind direction and assimilation window lengths.

As the counter examples, Figs. 5, 6, and 7 also show the EnDFS of the concentration and emission rates under the same assumptions as Figs. 1, 2 and 3 respectively, except that northeasterly wind is assumed. As expected, our approach can demonstrate that with the adverse wind direction, emission rates are not detectable and improvable by the given observation configuration whatever the duration of the assimilation window is. The quantitative balances of related figures are exposed in Table 1. It can be seen that the insensitivity to emission rate optimization remains equally low and affected by numerical noises.

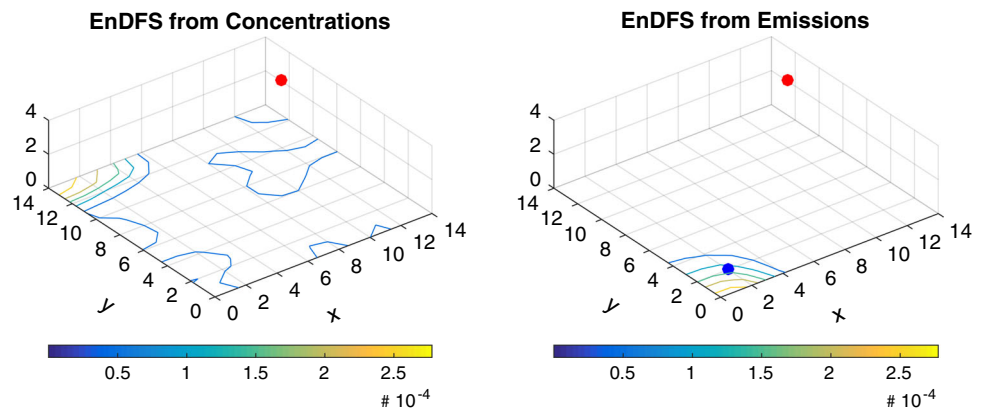
**Fig. 8** Emission signal test (weak) with  $48\Delta t$  DAW and southwesterly wind ( $v_x = 1$  and  $v_y = 1$ ). Plotting conventions are as in Fig. 1



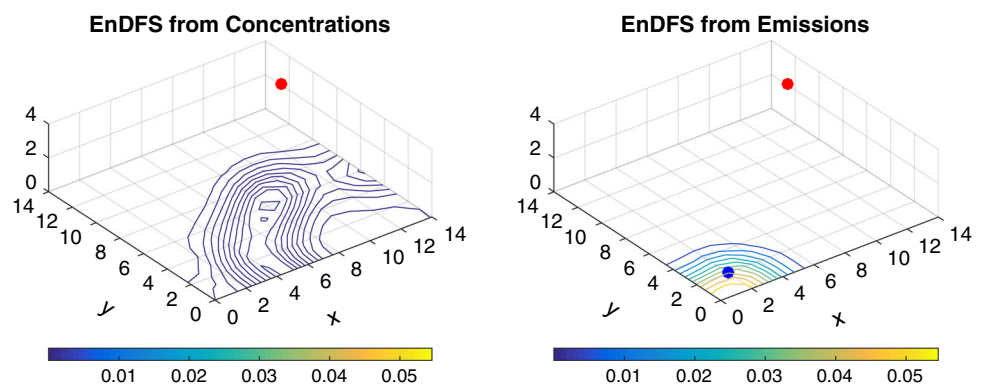
**Fig. 9** Emission signal test (strong) with  $48\Delta t$  DAW and southwest wind ( $v_x = 1$  and  $v_y = 1$ ). Plotting conventions are as in Fig. 1



**Fig. 10** Diffusion test (weak) with  $48\Delta t$  DAW and southwesterly wind. Plotting conventions are as in Fig. 1



**Fig. 11** Diffusion test (strong) with  $48\Delta t$  DAW and southwesterly wind. Plotting conventions are as in Fig. 1



*Emission signal tests:* The purpose of emission signal tests (Figs. 8 and 9) is to assess the impact of observation configurations to the emission rates evolved with different diurnal profiles. We make the same assumptions as for Fig. 3, except that the wind speed in Figs. 8 and 9 is accelerated such that the profiles of the emission rate is better detectable in relation to the observation within the assimilation window  $48\Delta t$ . The only distinction between situations in Figs. 8 and 9 is the pronounced diurnal cycle background profile of the emission rate during the assimilation window  $48\Delta t$ . The different profiles of emission rates are correlated with the different emitted amount of that species during the data assimilation window. It is clearly shown in Table 1 that the distinct variation of the emission rates during the data assimilation window acts to level  $\bar{p}^c$  and  $\bar{p}^e$ , and thus helps to improve the optimization results.

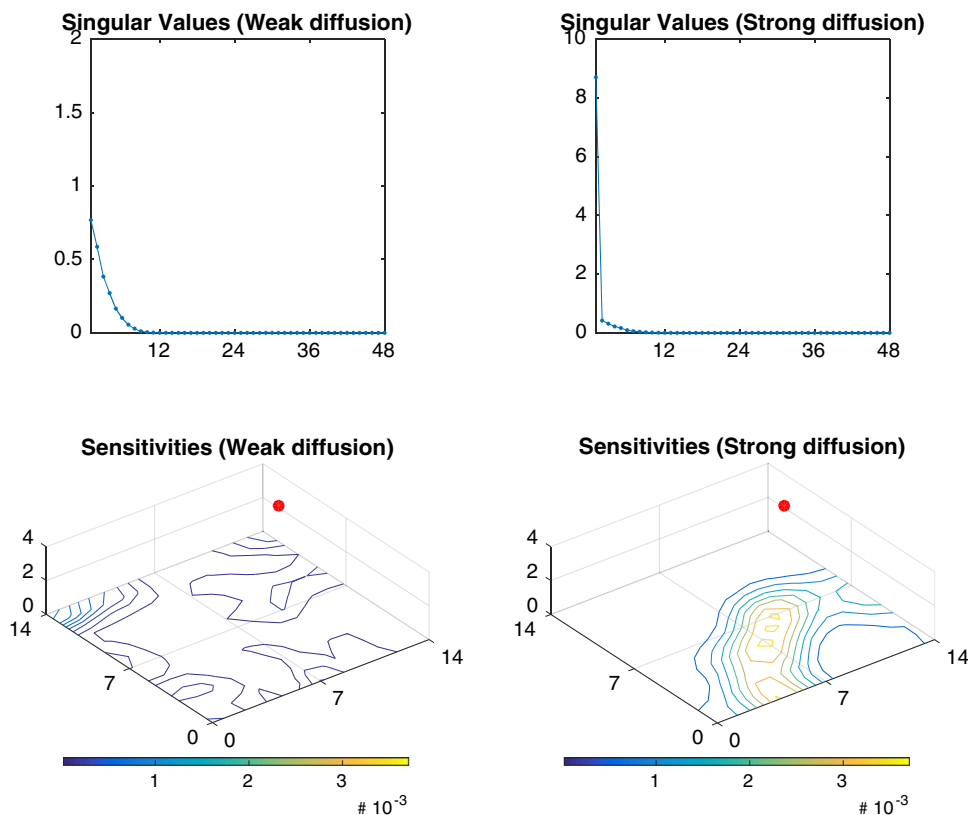
*Diffusion tests:* The vertical exchange of trace gases can be described by advection and diffusion, dependent on the nature of the process and the model grid resolution. In this study we confine our simulation to include the vertical diffusion only for the vertical coupling. The diffusion tests (Figs. 10, 11, and 12) aims to test our approach by comparing the EnDFS of the concentration and the emission

rate at the layer  $z = 0$  with a weak diffusion process and a strong diffusion process. We assume that the observation configuration at each time step is located at (12, 10, 4) in the diffusion test, with  $K(z) = 0.5e^{-z^2}$  in Fig. 10 and  $K(z) = 0.5e^{-z^2} + 1$  in Fig. 11. Besides, Figs. 10 and 12 preserve the same assumptions with Fig. 3.

It is obvious from Figs. 3 and 10 that the different observation locations strongly influence the distribution of the concentration. Table 2 shows that with the same diffusion coefficient the EnDFS of the concentration in the lowest layer in Fig. 3 is definitely larger than the one in Fig. 10. Moreover, it can be seen from Table 1 that the observation configuration at the top layer is not efficient to emission rates with such weak diffusion within  $48\Delta t$  data assimilation window.

Comparing Fig. 10 with Fig. 11, we can see how the EnDFS of concentration and emission rates increase with the stronger diffusion process. The increasing impact of the observation configuration with the stronger diffusion is also verified by the EnDFS and ensemble relative ratios of DFS of the concentration and emission rate for Figs. 10 and 11 in Table 2. The balances between the concentration and emission rate for Figs. 10 and 11 are shown in Table 1. The significant difference of weights of emission rates in

**Fig. 12** The panels in the first row show singular values of the diffusion tests shown in the left panels of Figs. 10 and 11. The panels in the second row show sensitivities of concentrations at the initial time approximated by 5 leading singular vectors



**Table 2** The ensemble degrees of freedom for signal of the initial concentration and emission at the lowest layer

	$\bar{P}_{low}^c$	$\bar{P}_{low}^e$
Fig. 3	0.277	0.885
Fig. 10	0.010	0.003
Fig. 11	0.050	0.789

Table 1 implies that the observation configuration cannot detect emission rates at the lowest layer with such a weak diffusion in Fig. 10, while with the stronger in Fig. 11 both the concentration and emission rates should be considered as optimized parameters with the corresponding weights.

Finally, similar to Fig. 4, the singular values of Figs. 10, 11 and the approximating targeting results of sensitive parameters are shown in Fig. 12. It shows that the sensitive parameters can be also caught by few leading singular vectors in the diffusion tests.

### 7 Conclusions and outlooks

In this study we extended the transport-diffusion models forced by emission rates in a novel way. Based on the Kalman smoother, we developed an approach to quantitatively identify the impact of a given observation network on the optimization of the initial trace gas state and emission rates. The contribution to the degree of freedom

for signal is adopted as a criterion to evaluate the potential assessment of observability to each element in the extended state vector. The degree of freedom for signal and a number of metrics was taken as a quantitative solutions to measure to what extent the parameters can be optimized in advance of the data assimilation procedure. It provides the opportunity to select the suitable and sensitive parameters to fulfill the optimization more efficiently. The ensemble case of the approach gave us the feasibility to determine the assessment of the potential observability jointly for initial values and emission rates for high-dimensional models in practical applications. Besides, we formulated sensitivities of observational networks by seeking the fastest directions of the perturbation ratio between initial states and observation configurations during the entire time interval. It facilitates to target the sensitive parameters to the observation networks by few leading singular values and vectors so that the computation costs can be further reduced. A series of experiments based on an elementary advection-diffusion model illustrated the significance of our approach in different situations.

In the future, we plan to apply this approach into the real atmospheric transport model to solve practical network validation problems prior to the solution of the inversion task, as far as the validity of the tangent linear assumption holds.

### Appendix A

In this appendix, we show a more general case to extend the control vector by emissions. As we know, the initial state and emission rates do not have the same dimension in some practical cases. Compared with (2), the general situation leads us to consider the following model

$$\frac{d\delta c}{dt} = \mathbf{A}\delta c + B(t)\delta e(t), \tag{39}$$

where  $B(t)$  is an operator transforming the emission state vector into the concentration-state space. Combining with (8), we obtain the extended model

$$\begin{pmatrix} \delta c(t) \\ \delta e(t) \end{pmatrix} = \begin{pmatrix} M(t, t_0) & \int_{t_0}^t M(t, s)B(s)M_e(s, t_0)ds \\ 0 & M_e(t, t_0) \end{pmatrix} \begin{pmatrix} \delta c(t_0) \\ \delta e(t_0) \end{pmatrix}. \tag{40}$$

### Appendix B

In this Appendix, we derive the foundation matrix of DFS based on the Kalman smoother and evaluate the potential improvement of the estimate via SVD.

For the discrete-time linear system (12) with the observation mapping (13) in Sect. 3, we denote the BLUE of  $x(t_i)$  based on  $\{y(t_0), \dots, y(t_k)\}$  by  $\hat{x}(t_i|t_k)$ ,  $t_i, t_k \in [t_0, \dots, t_N]$ . Especially, the prior estimation, or background state, of  $x(t_0)$  is denoted by  $\hat{x}(t_0|t_{-1})$ . Correspondingly,  $P(t_i|t_k)$  is defined as the error covariance of  $\hat{x}(t_i|t_k)$ ,  $t_i \in [t_0, \dots, t_N]$  and  $t_k \in [t_{-1}, \dots, t_N]$ . It is known that the inverse of the analysis error covariance matrix at initial time,  $P^{-1}(t_0|t_N)$  of a fixed-interval Kalman smoother is the optimal Hessian of the underlying cost function of 4D-Var (see Li and Navon (2001)). Thus, we have

$$\begin{aligned} P^{-1}(t_0|t_N) &= P^{-1}(t_0|t_{-1}) \\ &+ \sum_{i=0}^N M^T(t_i, t_0)H^T(t_i)R^{-1}(t_i)H(t_i)M(t_i, t_0) \\ &= P^{-1}(t_0|t_{-1}) + \mathcal{G}^T \mathcal{R}^{-1} \mathcal{G}, \end{aligned} \tag{41}$$

where

$$\mathcal{G} = \begin{pmatrix} H(t_0)M(t_0, t_0) \\ H(t_1)M(t_1, t_0) \\ \vdots \\ H(t_N)M(t_N, t_0) \end{pmatrix}, \mathcal{R}^{-1} = \begin{pmatrix} R^{-1}(t_0) & & & \\ & R^{-1}(t_1) & & \\ & & \ddots & \\ & & & R^{-1}(t_N) \end{pmatrix}. \tag{42}$$

It is clear that (41) comprises the information of the initial condition, model evolution, observation configurations and errors over the entire time interval  $[t_0, \dots, t_N]$ . At the same time, it is independent of any specific data and state vector, apart from the reference model evolution  $M(\cdot, \cdot)$  as well as

the observation operator  $H(\cdot)$ . Besides,  $\mathcal{G}^T \mathcal{R}^{-1} \mathcal{G}$  is the observability Gramian with respect to  $\mathcal{R}^{-1}$  in control theory [see Brockett (1970)]. It represents the observation capacity of the observation networks with respect to the model.

It can be seen now that (41) includes all available information before starting the data assimilation procedure. In order to evaluate the potential improvement of the estimate by the Kalman smoother, we aspire a matrix, which allows us for a direct and normalized comparison between sensitivities to initial values and emission rates. To this end, we consider matrix  $\tilde{P}$  with the following form:

$$\begin{aligned} \tilde{P} &= P^{-\frac{1}{2}}(t_0|t_{-1})(P(t_0|t_{-1}) - P(t_0|t_N))P^{-\frac{1}{2}}(t_0|t_{-1}) \\ &= I - P^{-\frac{1}{2}}(t_0|t_{-1})P(t_0|t_N)P^{-\frac{1}{2}}(t_0|t_{-1}), \end{aligned} \tag{43}$$

where  $I$  is the identity matrix and  $P^{\frac{1}{2}}(t_0|t_{-1})$  satisfies  $P^{\frac{1}{2}}(t_0|t_{-1})P^{\frac{1}{2}}(t_0|t_{-1}) = P(t_0|t_{-1})$ .

The matrix  $\tilde{P}$  is a normalized matrix of the difference between the background forecast error covariance matrix  $P(t_0|t_{-1})$  and the analysis error covariance matrix  $P(t_0|t_N)$ . It is the foundation matrix to study the DFS of models (Fisher (2003), Rodgers (2000), Singh et al. (2013)), which shows how much observation networks improve the estimation of model states.

$P(t_0|t_N)$  is unknown prior to the data assimilation procedure, so we use (41) to rewrite  $\tilde{P}$  as

$$\begin{aligned} \tilde{P} &= P^{-\frac{1}{2}}(t_0|t_{-1})(P(t_0|t_{-1}) - P(t_0|t_N))P^{-\frac{1}{2}}(t_0|t_{-1}) \\ &= I - (I + P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^T \mathcal{R}^{-1} \mathcal{G}P^{\frac{1}{2}}(t_0|t_{-1}))^{-1}. \end{aligned} \tag{44}$$

It is worth noting that in (44)

$$I + P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^T \mathcal{R}^{-1} \mathcal{G}P^{\frac{1}{2}}(t_0|t_{-1}) \tag{45}$$

is always invertible even if the observation Gramian  $\mathcal{G}^T \mathcal{G}$  is not full-rank. Thus,  $\tilde{P}$  is well-defined for all models with invertible initial covariance and observation systems with invertible error covariances within the assimilation window  $t_0$  to  $t_N$ . Then, we apply SVD to simplify (44)

$$P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^T \mathcal{R}^{-\frac{1}{2}} = VSU^T, \tag{46}$$

where  $V$  and  $U$  are unitary matrices consisting of the left and right singular vectors respectively, while  $S$  is the rectangular diagonal matrix consisting of the singular values.

Then, (44) can be simplified as

$$\begin{aligned} \tilde{P} &= I - (I + VSS^T V^T)^{-1} = V(I - (I + SS^T)^{-1})V^T \\ &= \sum_{i=1}^r \frac{s_i^2}{1 + s_i^2} v_i v_i^T, \end{aligned} \tag{47}$$

where  $r$  is the rank of (44) and  $v_i$  is the  $i^{th}$  left singular vector in  $V$  related to the singular value  $s_i$ , which is the  $i^{th}$  element on the diagonal of  $S$ .

### Appendix C

In this appendix, we investigate the ensemble version of expressions in Appendix B. For the discrete-time system (12), we denote the ensemble samples of  $\hat{x}(t_i|t_j)$  by  $\hat{x}_k(t_i|t_j)$ ,  $i, j = 1, \dots, N$ ,  $k = 1, \dots, q$ , where  $q$  is the number of ensemble members.

Correspondingly, the ensemble means of  $\hat{x}(t_i|t_j)$  is given by

$$\bar{x}(t_i|t_j) = \frac{1}{q} X(t_i|t_j) \mathbf{1}_{q \times 1}, \tag{48}$$

where  $X(t_i|t_j) = (\hat{x}_1(t_i|t_j), \hat{x}_2(t_i|t_j), \dots, \hat{x}_q(t_i|t_j))$  is the  $n \times q$  ensemble matrix,  $\mathbf{1}_{i \times j}$  is a  $i \times j$  matrix of which each element is equal to 1.

We calculate the ensemble forecast and analysis covariances as

$$\bar{P}(t_i|t_j) = \frac{1}{q-1} \tilde{X}(t_i|t_j) \tilde{X}^\top(t_i|t_j), \tag{49}$$

where  $\tilde{X}(t_i|t_j) = X(t_i|t_j) - \frac{1}{q} X(t_i|t_j) \mathbf{1}_{q \times q}$  is the related perturbation matrix. We define the ensemble observation configurations in the entire assimilation window as

$$y_k^f = \mathcal{G} \hat{x}_k(t_0|t_{-1}), \quad k = 1, \dots, q. \tag{50}$$

Further, the ensemble mean and the forecasting error covariance matrix of the ensemble observation configurations are given by

$$\begin{aligned} \bar{y}^f &= \frac{1}{q} \sum_{k=1}^q y_k^f, \\ \bar{P}_{yy}^f &= \frac{1}{q-1} \sum_{k=1}^q (y_k^f - \bar{y}^f)(y_k^f - \bar{y}^f)^\top = \mathcal{G} \bar{P}(t_0|t_{-1}) \mathcal{G}^\top. \end{aligned} \tag{51}$$

Similarly, we denote the ensemble covariance between the initial states and the forecasting observations by

$$\begin{aligned} \bar{P}_{xy}^f &= \frac{1}{q-1} \sum_{k=1}^q (\hat{x}_k(t_0|t_{-1}) - \bar{x}(t_0|t_{-1}))(y_k^f - \bar{y}^f)^\top \\ &= \bar{P}(t_0|t_{-1}) \mathcal{G}^\top. \end{aligned} \tag{52}$$

Furthermore, we define the ensemble observations as

$$\hat{y}_k(t_i) = y(t_i) + v_k(t_i), \quad k = 1, \dots, q, \quad i = 1, \dots, N, \tag{53}$$

and then assume  $\bar{v}(t_i) = \frac{1}{q} \sum_{k=1}^q v_k(t_i) = 0$ ,  $\bar{R}(t_i) = \frac{1}{q-1} \sum_{k=1}^q v_k(t_i) v_k^\top(t_i)$  and  $\mathcal{R}^{-1}$  is the block diagonal matrix with the diagonal  $(\bar{R}^{-1}(t_0), \dots, \bar{R}^{-1}(t_N))$ .

It is shown by Evensen (2009) that the ensemble forecast and analysis covariances have the same form as the covariances in the standard Kalman filter. However, the ensemble size  $q$  is significantly smaller than the dimension of the model  $n$  in practical applications. As a result, the initial ensemble covariance  $\bar{P}(t_0|t_{-1})$  is not invertible. In this case, the pseudo inverse is a widely used alternative of the inverse of a matrix, due to its optimal uniqueness properties. We denote the pseudo inverse of a matrix  $A$  by  $A^\dagger$ . Then, for the initial ensemble covariance

$$\bar{P}(t_0|t_{-1}) = \frac{1}{q-1} \tilde{X}(t_0|t_{-1}) \tilde{X}^\top(t_0|t_{-1}), \tag{54}$$

we apply the singular value decomposition to

$$\frac{1}{\sqrt{q-1}} \tilde{X}(t_0|t_{-1}) = V_0 S_0 U_0^\top, \tag{55}$$

where  $V_0 \in \mathbf{R}^{n \times n}$  and  $U_0 \in \mathbf{R}^{q \times q}$  consist of the left and right singular vectors respectively,  $S_0 \in \mathbf{R}^{n \times q}$  is a rectangular diagonal matrix with singular values  $\{s_{0i} | s_{0i} \geq 0\}_{i=1}^q$  on its diagonal. Thus,

$$\bar{P}(t_0|t_{-1}) = V_0 S_0 U_0^\top U_0 S_0^\top V_0^\top = V_0 S_0 S_0^\top V_0^\top = V_0 \hat{S}_0^2 V_0^\top, \tag{56}$$

where  $\hat{S}_0^2 = S_0 S_0^\top \in \mathbf{R}^{n \times n}$  is a block diagonal matrix with the diagonal

$$(s_{01}^2, \dots, s_{0r_0}^2, \mathbf{0}_{1 \times (n-r_0)}),$$

with  $r_0$  being the rank of  $S_0$ . Hence, we find a pseudo inverse

$$\bar{P}^{\dagger \frac{1}{2}}(t_0|t_{-1}) = V_0 \hat{S}_0^\dagger V_0^\top, \tag{57}$$

where  $\hat{S}_0^\dagger$  is the pseudo inverse of  $\hat{S}_0$  with the diagonal  $(1/s_{01}, \dots, 1/s_{0r_0}, \mathbf{0}_{1 \times (n-r_0)})$ .

Analog to (43), we define  $\bar{P}$  as

$$\bar{P} = \bar{P}^{\dagger \frac{1}{2}}(t_0|t_{-1}) (\bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_N)) \bar{P}^{\dagger \frac{1}{2}}(t_0|t_{-1}). \tag{58}$$

Likewise, corresponding to (13), we present the observation system in the entire time interval as

$$y = \mathcal{G}x(t_0) + v, \tag{59}$$

where  $y = (y^\top(t_0), \dots, y^\top(t_N))^\top$ ,  $v = (v^\top(t_0), \dots, v^\top(t_N))^\top$  and  $\mathcal{G}$  as the observation

configuration for  $x(t_0)$ . Then, for the analysis error covariance matrix, we obtain

$$\begin{aligned} & \bar{P}(t_0|t_N) \\ &= \bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_{-1})\mathcal{G}^\top(\mathcal{G}\bar{P}(t_0|t_{-1})\mathcal{G}^\top + \mathcal{R})^{-1}\mathcal{G}\bar{P}(t_0|t_{-1}) \\ &= \bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_{-1})\mathcal{G}^\top\mathcal{R}^{-\frac{1}{2}}(I + \mathcal{R}^{-\frac{1}{2}}\mathcal{G}\bar{P}(t_0|t_{-1})\mathcal{G}^\top\mathcal{R}^{-\frac{1}{2}})^{-1}\mathcal{R}^{-\frac{1}{2}}\mathcal{G}\bar{P}(t_0|t_{-1}) \\ &= \bar{P}(t_0|t_{-1}) - \bar{P}_{xy}^f\mathcal{R}^{-\frac{1}{2}}(I + \mathcal{R}^{-\frac{1}{2}}\bar{P}_{yy}^f\mathcal{R}^{-\frac{1}{2}})^{-1}\mathcal{R}^{-\frac{1}{2}}(\bar{P}_{xy}^f)^\top. \end{aligned} \tag{60}$$

Further, analog to (47), we obtain

$$\begin{aligned} \bar{P} &= \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})(\bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_N))\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) \\ &= \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}_{xy}^f\mathcal{R}^{-\frac{1}{2}}(I + \mathcal{R}^{-\frac{1}{2}}\bar{P}_{yy}^f\mathcal{R}^{-\frac{1}{2}})^{-1}\mathcal{R}^{-\frac{1}{2}}(\bar{P}_{xy}^f)^\top\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}). \end{aligned} \tag{61}$$

Let  $\sum_{i=1}^N m(t_i) = m$  be the number of observations available within the assimilation window. To proceed with (61), we apply again the singular value decomposition for the following matrix

$$\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}_{xy}^f\mathcal{R}^{-\frac{1}{2}} = \bar{V}\bar{S}\bar{U}^\top \in \mathbf{R}^{n \times m}, \tag{62}$$

where  $\bar{U} \in \mathbf{R}^{m \times m}$  and  $\bar{V} \in \mathbf{R}^{n \times n}$  consists of the right and left singular vectors of  $\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}_{xy}^f\mathcal{R}^{-\frac{1}{2}}$ , respectively.  $S \in \mathbf{R}^{n \times m}$  consists of the singular values on its diagonal.

We denote the rank of (62) by  $r$ . Then, we rewrite  $\bar{P}$  as

$$\begin{aligned} \bar{P} &= \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})(\bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_N))\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) \\ &= \bar{V}\bar{S}^\top\bar{U}^\top(\bar{U}\bar{U}^\top + \bar{U}(\bar{S}\bar{S}^\top)\bar{U}^\top)^{-1}\bar{U}\bar{S}\bar{V}^\top \\ &= \bar{V}\bar{S}^\top(I + \bar{S}^\top\bar{S})^{-1}\bar{S}\bar{V}^\top \\ &= \sum_{i=1}^r \frac{\bar{s}_i^2}{1 + \bar{s}_i^2} \bar{v}_i\bar{v}_i^\top. \end{aligned} \tag{63}$$

where  $r$  is the rank of  $\bar{P}$  as well, and  $\bar{v}_i$  is the  $i^{th}$  left singular vector in  $V$  related to the singular value  $\bar{s}_i$ , which is the  $i^{th}$  element on the diagonal of  $S$ .

### Appendix D

In this appendix, we will introduce the singular vector approach to identify sensitive directions of initial values and emission rates to observation networks.

Similar to Liao et al. (2006), we define the magnitude of the perturbation of the initial state by the norm in the state space with respect to a positive definite matrix  $W_0$ , typically the forecast error covariance matrix.

$$\|\delta x(t_0)\|_{W_0}^2 = \langle \delta x(t_0), W_0\delta x(t_0) \rangle. \tag{64}$$

Likewise, we define the magnitude of the related observations perturbation in the time interval  $[t_0, \dots, t_N]$  by the

norm with respect to a sequence of positive definite matrices  $\{W(t_k)\}_{k=1}^N$

$$\|\delta y_c\|_{\{W(t_k)\}}^2 = \sum_{k=0}^N \langle \delta y_c(t_k), W(t_k)\delta y_c(t_k) \rangle, \tag{65}$$

where  $\delta y_c = (\delta y_c^\top(t_0), \dots, \delta y_c^\top(t_N))^\top$ .

In order to find the direction of observation configuration which minimizes the perturbation of the initial states, the ratio

$$\frac{\|\delta x(t_0)\|_{W_0}^2}{\|\delta y_c\|_{\{W(t_k)\}}^2}, \quad \delta y \neq 0_{m \times 1}. \tag{66}$$

must be minimized. It is equivalent to maximize the ratio of the magnitude of observation perturbation and the initial perturbation

$$\frac{\|\delta y_c\|_{\{W(t_k)\}}^2}{\|\delta x(t_0)\|_{W_0}^2}, \quad \delta x(t_0) \neq 0_{n \times 1}. \tag{67}$$

Thus, we define the measure the perturbation growth as

$$g^2 = \frac{\|\delta y_c\|_{\{W(t_k)\}}^2}{\|\delta x(t_0)\|_{W_0}^2} \tag{68}$$

$$\begin{aligned} &= \sum_{k=0}^N \frac{\langle \delta y_c(t_k), W(t_k)\delta y_c(t_k) \rangle}{\langle \delta x(t_0), W_0\delta x(t_0) \rangle} \\ &= \sum_{k=0}^N \frac{\langle H(t_k)\delta x(t_k), W(t_k)H(t_k)\delta x(t_k) \rangle}{\langle \delta x(t_0), W_0\delta x(t_0) \rangle} \\ &= \sum_{k=0}^N \frac{\langle \delta x(t_k), H(t_k)^\top W(t_k)H(t_k)\delta x(t_k) \rangle}{\langle \delta x(t_0), W_0\delta x(t_0) \rangle} \\ &= \sum_{k=0}^N \frac{\langle \delta x(t_0), M(t_k, t_0)^\top H(t_k)W(t_k)H(t_k)M(t_k, t_0)\delta x(t_0) \rangle}{\langle \delta x(t_0), W_0\delta x(t_0) \rangle} \\ &= \frac{\langle \delta x(t_0), \sum_{k=0}^N M(t_k, t_0)^\top H(t_k)^\top W(t_k)H(t_k)M(t_k, t_0)\delta x(t_0) \rangle}{\langle \delta x(t_0), W_0\delta x(t_0) \rangle} \\ &= \frac{\langle \delta x(t_0), \mathcal{G}^\top \mathcal{W} \mathcal{G} \delta x(t_0) \rangle}{\langle \delta x(t_0), W_0\delta x(t_0) \rangle}, \quad \delta x(t_0) \neq 0, \end{aligned} \tag{69}$$

where  $\mathcal{G}$  has the same definition as given in Sect. 3 and

$$\mathcal{W} = \begin{pmatrix} W(t_0) & & \\ & \ddots & \\ & & W(t_N) \end{pmatrix}. \tag{70}$$

We arrive at the eigenvalue problems Liao et al. (2006)

$$W_0^{-\frac{1}{2}}\mathcal{G}^\top\mathcal{W}\mathcal{G}W_0^{-\frac{1}{2}}v_k = s_k^2v_k, \quad W_0^{\frac{1}{2}}\mathcal{G}W_0^{-1}\mathcal{G}^\top W_0^{\frac{1}{2}}u_k = s_k^2u_k, \tag{71}$$

where  $s_1 \geq s_2 \geq \dots \geq s_n \geq 0$  are singular values,  $\{v_k\}_{k=1}^n$  and  $\{u_k\}_{k=1}^n$  are the corresponding orthogonal singular vectors. Then,  $\max_{\delta x(t_0) \neq 0} g^2 = s_1^2$ .

Especially, if the perturbation norms are provided by the choice  $W_0 = P^{-1}(t_0|t_{-1})$  and  $\mathcal{W} = \mathcal{R}^{-1}$ ,

$$g^2 = \frac{\langle \delta x(t_0), \mathcal{G}^\top \mathcal{R}^{-1} \mathcal{G} \delta x(t_0) \rangle}{\langle \delta x(t_0), P^{-1}(t_0|t_{-1}) \delta x(t_0) \rangle}, \quad \delta x(t_0) \neq 0. \quad (72)$$

Correspondingly, we consider the singular value decomposition of the following matrices

$$P^{\frac{1}{2}}(t_0|t_{-1}) \mathcal{G}^\top \mathcal{R}^{-1} \mathcal{G} P^{\frac{1}{2}}(t_0|t_{-1}) v_k = s_k^2 v_k; \quad (73)$$

$$\mathcal{R}^{-\frac{1}{2}} \mathcal{G} P(t_0|t_{-1}) \mathcal{G}^\top \mathcal{R}^{-\frac{1}{2}} u_k = s_k^2 u_k, \quad k = 1, \dots, n, \quad (74)$$

where  $s_k$ ,  $v_k$  and  $u_k$ ,  $k = 1, \dots, n$  denotes singular values and corresponding singular vectors to simplify the notation. At the same time, it can be seen that the above singular value decomposition problem coincides with (46) in Sect. 3.

**Acknowledgements** This study was supported by HITEC (Helmholtz Interdisciplinary Doctoral Training in Energy and Climate Research). HITEC is the Graduate School of Forschungszentrum Jülich and the five partner universities Aachen, Bochum, Cologne, Düsseldorf and Wuppertal focusing on energy and climate research. Technical and computational support was provided by the Jülich Supercomputer Centre (JSC) of the Research Centre Jülich. A large and central part of the case studies has been computed on JSC supercomputer JUR-ECA. Access given to these computational resources is highly appreciated. Besides, We thank two anonymous reviewers whose valuable suggestions and comments improved and clarified this manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Anderson JL (2001) An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Rev* 129(12):2884–2903. [https://doi.org/10.1175/1520-0493\(2001\)129](https://doi.org/10.1175/1520-0493(2001)129)
- Bellsky T, Kostelich EJ, Mahalov A (2014) Kalman filter data assimilation: targeting observations and parameter estimation. *Chaos* 24(2):024406. <https://doi.org/10.1063/1.4871916>

- Berliner LM, Lu Z, Snyder C (1999) Statistical design for adaptive weather observations. *J Atmosph Sci* 56:2536–2552. [https://doi.org/10.1175/1520-0469\(1999\)056](https://doi.org/10.1175/1520-0469(1999)056)
- Bishop CH, Toth Z (1999) Ensemble transformation and adaptive observations. *J Atmosph Sci* 56:1748–1765. [https://doi.org/10.1175/1520-0469\(1999\)056](https://doi.org/10.1175/1520-0469(1999)056)
- Bocquet M, Sakov P (2013) Joint state and parameter estimation with an iterative ensemble Kalman smoother. *Nonlin Processes Geophys* 20:803–818. <https://doi.org/10.5194/npg-20-803-2013>
- Bousserez N, Henze DK (2018) Optimal and scalable methods to approximate the solutions of large-scale Bayesian problems: theory and application to atmospheric inversion and data assimilation. *Q J R Meteorol Soc* 144:365–390. <https://doi.org/10.1002/qj.3209>
- Brockett RW (1970) Finite dimensional linear systems. John Wiley and Sons Inc
- Buizza R, Montani A (1999) Targeting observations using singular vectors. *J Atmosph Sci* 56:2965–2985. [https://doi.org/10.1175/1520-0469\(1999\)056](https://doi.org/10.1175/1520-0469(1999)056)
- Buizza R, Palmer TN (1995) The singular-vector structure of the atmospheric global circulation. *J Atmosph Sci* 52:1434–1456. [https://doi.org/10.1175/1520-0469\(1995\)052](https://doi.org/10.1175/1520-0469(1995)052)
- Cardinali C, Pezzulli S, Andersson E (2004) Influence-matrix diagnostic of a data assimilation system. *Q J R Meteorol Soc* 130:2767–2786. <https://doi.org/10.1256/qj.03.205>
- Cioaca A, Sandu A (2014) Low-rank approximations for computing observation impact in 4D-Var data assimilation. *Comput Math Appl* 67(12):2112–2126. <https://doi.org/10.1016/j.camwa.2014.01.024>
- Cioaca A, Sandu A (2014) An optimization framework to improve 4D-Var data assimilation system performance. *J Comput Phys* 275:377–389. <https://doi.org/10.1016/j.jcp.2014.07.005>
- Daescu D, Navon IM (2004) Adaptive observations in the context for 4D-Var data assimilation. *Meteorol Atmos Phys* 55:205–236. <https://doi.org/10.1007/s00703-003-0011-5>
- Daley R (1991) Atmospheric data analysis. Cambridge University Press, Cambridge
- Duarte E, Franke P, Lange AC, Friese E, Silva Lopes FJ, Silva JJ, Reis JS, Landulfo E, Silva CMS, Elbern H, Hoelzemann JJ (2021) Evaluation of atmospheric aerosols in the metropolitan area of São Paulo simulated by the regional EURAD-IM model on high-resolution. *Atmosph Pollut Res* 12(2):451–469. <https://doi.org/10.1016/j.apr.2020.12.006>
- Elbern H, Schmidt H (1999) A four-dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling. *J Geophys Res* 104:583–598. <https://doi.org/10.2495/EURO991042>
- Elbern H, Schmidt H, Talagrand O, Ebel A (2000) 4D-variational data assimilation with an adjoint air quality model for emission analysis. *Environ Model Softw* 15:539–548. [https://doi.org/10.1016/S1364-8152\(00\)00049-9](https://doi.org/10.1016/S1364-8152(00)00049-9)
- Elbern H, Strunk A, Schmidt H, Talagrand O (2007) Emission rate and chemical state estimation by 4-dimension variational inversion. *Atmos Chem Phys* 7:3749–3769. <https://doi.org/10.5194/acp-7-3749-2007>
- Evensen G (2009) Data assimilation: the ensemble Kalman filter, 2nd edn. Springer, Berlin
- Eyre JR (1990) The information content of data from satellite sounding systems: a simulation study. *Quart J Roy Meteor Soc* 116:401–434. <https://doi.org/10.1002/qj.49711649209>
- Fisher M (2003) Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems. ECMWF Tech Memo. <https://doi.org/10.21957/2bec9m380>
- Gautam AS, Kumar S, Gautam S, Anand A, Kumar R, Joshi A, Baudh K, Singh K (2021) Pandemic induced lockdown as a boon to the Environment: trends in air pollution concentration

- across India. *Asia Pac J Atmos Sci* 1:1–16. <https://doi.org/10.1007/s13143-021-00232-7>
- Gautam S, Samuel C, Gautam AS, Kumar S (2021) Strong link between coronavirus count and bad air: a case study of India. *Environ Dev Sustain* 3:1–14. <https://doi.org/10.1007/s10668-021-01366-4>
- Gelb A (ed) (1974) *Applied optimal estimation*. The MIT Press, Cambridge
- Goris N, Elbern H (2013) Singular vector decomposition for sensitivity analyses of tropospheric chemical scenarios. *Atmos Chem Phys* 13:5063–5087. <https://doi.org/10.5194/acp-13-5063-2013>
- Goris N, Elbern H (2015) Singular vector based targeted observations of chemical constituents: description and first application of EURAD-IM-SVA. *Geosci Model Dev* 8:3929–3945. <https://doi.org/10.5194/gmd-8-3929-2015>
- Higham NJ (2002) *Accuracy and stability of numerical algorithms*, 2nd edn. SIAM, England
- Kang W, Xu L (2012) Optimal placement of mobile sensors for data assimilations. *Tellus A Dyn Meteorol Oceanogr* 64:1. <https://doi.org/10.3402/tellusa.v64i0.17133>
- Kaskaoutis DG, Kumar S, Sharma D, Singh RP, Kharol SK, Sharma M, Singh AK, Singh S, Singh A, Singh D (2014) Effects of crop residue burning on aerosol properties, plume characteristics, and long-range transport over northern India. *J Geophys Res Atmos* 119:5424–5444. <https://doi.org/10.1002/2013JD021357>
- Khattatov BB, Gille J, Lyjak L, Brasseur G, Dvortsov V, Roche A, Waters J (1999) Assimilation of photochemically active species and a case analysis of UARS data. *J Geophys Res* 22:18715–18738. <https://doi.org/10.1029/1999JD900225>
- Kumar A, Bali K, Singh S, Naja S, Mishra AK (2019) Estimates of reactive trace gases (NMVOCs, CO and NOx) and their ozone forming potentials during forest fire over Southern Himalayan region. *Atmosph Res* 227:41–51. <https://doi.org/10.1016/j.atmosres.2019.04.028>
- Langland RH, Gelaro R, Rohaly GD, Shapiro MA (1999) Targeted observations in FASTEX: adjoint based targeting procedures and data impact experiments in IOP17 and IOP18. *Q J R Meteorol Soc* 125:3241–3270. <https://doi.org/10.1002/qj.49712556107>
- Li H, Kalnay E, Miyoshi T, Danforth CM (2009) Accounting for model errors in ensemble data assimilation. *Mon Weather Rev* 137:3407–3419. <https://doi.org/10.1175/2009MWR2766.1>
- Li Z, Navon IM (2001) Optimality of variational data assimilation and its relationship with the Kalman filter and smoother. *Q J R Meteorol Soc* 127:661–683. <https://doi.org/10.1002/qj.49712757220>
- Liao W, Sandu A, Carmichael GR, Chai T (2006) Singular vector analysis for atmospheric chemical transport models. *Mon Weather Rev* 134:2443–2465. <https://doi.org/10.1175/MWR3158.1>
- Lorenz EN (1965) A study of the predictability of a 28 variable atmospheric model. *Tellus* 17:321–333. <https://doi.org/10.1111/j.2153-3490.1965.tb01424.x>
- Martynenko D, Holzer-Popp T, Elbern H, Schroedter-Homscheidt M (2010) Understanding the aerosol information content in multi-spectral reflectance measurements using a synergistic retrieval algorithm. *Atmos Meas Tech* 3:2579–2602. <https://doi.org/10.5194/amt-3-1589-2010>
- Miyazaki K, Eskes HJ, Sudo K, Takigawa M, van Weele M, Boersma KF (2012) Simultaneous assimilation of satellite NO<sub>2</sub>, O<sub>3</sub>, CO, and HNO<sub>3</sub> data for the analysis of tropospheric chemical composition and emissions. *Atmos Chem Phys* 12:9545–9579. <https://doi.org/10.5194/acp-12-9545-2012>
- Navon IM (1997) Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dyn Atmos Oceans* 27:55–79. [https://doi.org/10.1016/S0377-0265\(97\)00032-8](https://doi.org/10.1016/S0377-0265(97)00032-8)
- Rabier F, Fourrié N, Chafaï D, Prunet P (2002) Channel selection methods for infrared atmospheric sounding interferometer radiances. *Quart J Roy Meteor Soc* 128:1011–1027. <https://doi.org/10.1256/0035900021643638>
- Rodgers CD (2000) *Inverse methods for atmospheric sounding: theory and practice*. World Scientific, Singapore
- Sandu A, Chai T (2011) Chemical data assimilation-an overview. *Atmosphere* 2(3):426–463. <https://doi.org/10.3390/atmos2030426>
- Sandu A, Cioaca A, Rao V (2013) Dynamic sensor network configuration in InfoSymbiotic systems using model singular vectors. *Proc Comput Sci* 18:1909–1918. <https://doi.org/10.1016/j.procs.2013.05.360>
- Singh K, Sandu A, Jurdak M, Bowman KW, Lee M (2013) A practical method to estimate information content in the context of 4D-Var data assimilation. *SIAM/ASA J Uncertain Quantification* 1(1):106–138. <https://doi.org/10.1137/120884523>
- Sharma M, Kaskaoutis DG, Singh RP, Singh S (2014) Seasonal variability of atmospheric aerosol parameters over greater noida using ground sunphotometer observations. *Aerosol Air Qual Res* 14:608–622. <https://doi.org/10.4209/aaqr.2013.06.0219>
- Smith PJ, Thornhill GD, Dance SL, Lawless AS, Mason DC, Nichols NK (2013) Data assimilation for state and parameter estimation: application to morphodynamic modeling. *Q J R Meteorol Soc* 139:314–327. <https://doi.org/10.1002/qj.1944>
- Spantini A, Solonen A, Cui T, Martin J, Tenorio L, Marzouk Y (2015) Optimal low-rank approximations of Bayesian linear inverse problems. *SIAM J Sci Comput* 37(6):A2451–A2487. <https://doi.org/10.1137/140977308>
- Szunyogh I, Toth Z, Emanuel KA, Bishop CH, Woolen J, Marchok T, Morss R, Snyder C (1999) Ensemble based targeting experiments during FASTEX: the impact of dropsonde data from the Lear jet. *Q J R Meteorol Soc* 125:3189–3218. <https://doi.org/10.1002/qj.49712556105>
- Vogel A, Elbern H (2021) Identifying forecast uncertainties for biogenic gases in the Po Valley related to model configuration in EURAD-IM during PEGASOS 2012. *Atmos Chem Phys* 21:4039–4057. <https://doi.org/10.5194/acp-21-4039-2021>
- Vogel A, Elbern H (2021) Efficient ensemble generation for uncertain correlated parameters in atmospheric chemical models: a case study for biogenic emissions from EURAD-IM version 5. *Geosci Model Dev* 14:5583–5605. <https://doi.org/10.5194/gmd-14-5583-2021>
- Winiarek V, Bocquet M, Duhanyan N, Roustan Y, Saunier O, Mathieu A (2014) Estimation of the caesium-137 source term from the Fukushima Daiichi nuclear power plant using a consistent joint assimilation of air concentration and deposition observations. *Atmos Environ* 82:268–279. <https://doi.org/10.1016/j.atmosenv.2013.10.017>
- Wu X, Jacob B, Elbern H (2016) Optimal control and observation locations for time-varying systems on a finite-time horizon. *SIAM J Control Optim* 54(1):291–316. <https://doi.org/10.1137/15M1014759>
- Woodbury MA (1950) *Inverting modified matrices*. Memorandum Rept. 42, Statistical Research Group, Princeton University, Princeton, NJ
- Yanenko NN (1971) *The method of fractional steps: solution of problems of mathematical physics in several variables*. Springer, Berlin

Zupanski D, Hou YA, Zhang QS, Zupanski M, Kummerow DC, Cheung HS (2007) Applications of information theory in ensemble data assimilation. *Q J R Meteorol Soc* 133:1533–1545. <https://doi.org/10.1002/qj.123>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.