



**Balancing aggregation and smoothing errors**

A. J. Turner and  
D. J. Jacob

This discussion paper is/has been under review for the journal Atmospheric Chemistry and Physics (ACP). Please refer to the corresponding final paper in ACP if available.

# Balancing aggregation and smoothing errors in inverse models

A. J. Turner<sup>1</sup> and D. J. Jacob<sup>1,2</sup>

<sup>1</sup>School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA

<sup>2</sup>Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA

Received: 4 December 2014 – Accepted: 12 December 2014 – Published: 13 January 2015

Correspondence to: A. J. Turner (aturner@fas.harvard.edu)

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

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Abstract

Inverse models use observations of a system (observation vector) to quantify the variables driving that system (state vector) by statistical optimization. When the observation vector is large, such as with satellite data, selecting a suitable dimension for the state vector is a challenge. A state vector that is too large cannot be effectively constrained by the observations, leading to smoothing error. However, reducing the dimension of the state vector leads to aggregation error as prior relationships between state vector elements are imposed rather than optimized. Here we present a method for quantifying aggregation and smoothing errors as a function of state vector dimension, so that a suitable dimension can be selected by minimizing the combined error. Reducing the state vector within the aggregation error constraints can have the added advantage of enabling analytical solution to the inverse problem with full error characterization. We compare three methods for reducing the dimension of the state vector from its native resolution: (1) merging adjacent elements (grid coarsening), (2) clustering with principal component analysis (PCA), and (3) applying a Gaussian mixture model (GMM) with Gaussian pdfs as state vector elements on which the native-resolution state vector elements are projected using radial basis functions (RBFs). The GMM method leads to somewhat lower aggregation error than the other methods, but more importantly it retains resolution of major local features in the state vector while smoothing weak and broad features.

1 Introduction

Inverse models quantify the state variables driving the evolution of a physical system by using observations of that system. This requires a physical model  $\mathbf{F}$ , known as the forward model, that relates a set of input variables  $\mathbf{x}$  (state vector) to a set of output variables  $\mathbf{y}$  (observation vector),

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) + \epsilon \tag{1}$$

Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

AbstractIntroduction

ConclusionsReferences

TablesFigures

⏮⏭

⏪⏩

BackClose

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The observational error  $\epsilon$  includes contributions from both the forward model and the measurements. Solution to the inverse problem involves statistical optimization to achieve a best error-weighted estimate of  $\mathbf{x}$  given  $\mathbf{y}$ .

A critical step in solving the inverse problem is determining the amount of information contained in the observations and choosing the state vector accordingly. This is a non-trivial problem when using large observational datasets with large errors. An example that will guide our discussion is the inversion of methane emissions on the basis of satellite observations of atmospheric methane concentrations (Turner et al., 2014). Methane concentrations can be predicted on the basis of emissions by using a chemical transport model (CTM) that solves the 3-D continuity equation for methane concentrations. Here the CTM is the forward model  $\mathbf{F}$ , the satellite provides a large observation vector  $\mathbf{y}$ , and we need to choose the resolution at which to optimize the methane emission vector  $\mathbf{x}$ .

The simplest approach would be to use the native resolution of the CTM in order to extract the maximum information from the observations. However, the observations may not be sufficiently dense or precise to optimize emissions at that level of detail, resulting in an underdetermined problem. Bocquet et al. (2011) refer to this as the “resolution problem”. The inverse solution must then rely on some prior estimate for the state vector and may not be able to depart sufficiently from that knowledge. This is known as the smoothing error (Rodgers, 2000) and increases with size of the state vector (Bousquet et al., 2000; Kaminski and Heimann, 2001; Kaminski et al., 2001; Bocquet et al., 2011; von Clarmann, 2014). Wecht et al. (2014) illustrate the severity of this problem in their inversion of methane emissions using satellite data.

An additional drawback of using a large state vector is the computational cost of the inversion. Analytical solution to the inverse problem requires calculation of the Jacobian matrix,  $\nabla_{\mathbf{x}}\mathbf{F}$ , and inversion and multiplication of the error covariance matrices (Rodgers, 2000). It has the advantage of providing complete error statistics on the solution but it becomes impractical as the state vector becomes too large. Numerical solutions using

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

variational methods circumvent this problem but do not inherently provide error characterization as part of the solution.

Reducing the dimensionality of the state vector in the inverse problem thus has two advantages. It improves the observational constraints on individual state vector elements and it facilitates analytical solution. Reduction can be achieved by aggregating state vector elements. For a state vector of gridded time-dependent emissions, the state vector can be reduced by aggregating grid cells and time periods. However, this introduces error in the inversion as the underlying spatial and temporal patterns of the aggregated emissions are now imposed from prior knowledge and not allowed to be optimized as part of the inversion. The resulting error is called the aggregation error (Kaminski and Heimann, 2001; Kaminski et al., 2001; Schuh et al., 2009).

Here we present a method for optimizing the selection of the state vector in the solution of the inverse problem for a given ensemble of observations. As the dimension of the state vector decreases, the smoothing error decreases while the aggregation error increases. We show how to derive an optimum where the overall error is minimized. We derive an analytical expression for the aggregation error covariance matrix and show how this can guide selection of a reduced-dimension state vector where the aggregation error remains below an acceptable threshold. We also show how intelligent selection of the state vector can extract more information from the observations for a given state vector dimension.

## 2 Formulating the inverse problem

Inverse problems are commonly solved using Bayes' theorem,

$$P(\mathbf{x}|\mathbf{y}) \propto P(\mathbf{y}|\mathbf{x})P(\mathbf{x}) \quad (2)$$

where  $P(\mathbf{x}|\mathbf{y})$  is the posterior probability density function (pdf) of the state vector  $\mathbf{x}$  ( $n \times 1$ ) given a vector of observations  $\mathbf{y}$  ( $m \times 1$ ),  $P(\mathbf{x})$  is the prior pdf of  $\mathbf{x}$ , and  $P(\mathbf{y}|\mathbf{x})$  is

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



the conditional pdf of  $\mathbf{y}$  given the true value of  $\mathbf{x}$ . Assuming Gaussian distributions for  $P(\mathbf{y}|\mathbf{x})$  and  $P(\mathbf{x})$  allows us to write the posterior pdf as

$$P(\mathbf{x}|\mathbf{y}) \propto \exp \left\{ -\frac{1}{2}(\mathbf{y} - \mathbf{F}(\mathbf{x}))^T \mathbf{S}_O^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x})) - \frac{1}{2}(\mathbf{x}_a - \mathbf{x})^T \mathbf{S}_a^{-1} (\mathbf{x}_a - \mathbf{x}) \right\} \quad (3)$$

where  $\mathbf{x}_a$  is the  $n \times 1$  prior state vector,  $\mathbf{S}_O$  is the  $m \times m$  observational error covariance matrix, and  $\mathbf{S}_a$  is the  $n \times n$  prior error covariance matrix. The most probable solution  $\hat{\mathbf{x}}$  (called the maximum a posteriori or MAP) is defined by the maximum of  $P(\mathbf{x}|\mathbf{y})$ , i.e., the minimum of the cost function  $\mathcal{J}(\mathbf{x})$ :

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{y} - \mathbf{F}(\mathbf{x}))^T \mathbf{S}_O^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x})) + \frac{1}{2}(\mathbf{x}_a - \mathbf{x})^T \mathbf{S}_a^{-1} (\mathbf{x}_a - \mathbf{x}) \quad (4)$$

This involves solving

$$\nabla_{\mathbf{x}} \mathcal{J} = \nabla_{\mathbf{x}} \mathbf{F}(\mathbf{x})^T \mathbf{S}_O^{-1} (\mathbf{F}(\mathbf{x}) - \mathbf{y}) + \mathbf{S}_a^{-1} (\mathbf{x}_a - \mathbf{x}) = \mathbf{0} \quad (5)$$

Solution to Eq. (5) can be done analytically if  $\mathbf{F}$  is linear, i.e.,  $\mathbf{F}(\mathbf{x}) = \mathbf{K}\mathbf{x} + \mathbf{c}$  where  $\mathbf{K} \equiv \nabla_{\mathbf{x}} \mathbf{F} = \partial \mathbf{y} / \partial \mathbf{x}$  is the Jacobian of  $\mathbf{F}$  and  $\mathbf{c}$  is a constant that can be set to zero in the general case by subtracting  $\mathbf{c}$  from the observations. This yields

$$\hat{\mathbf{x}} = \mathbf{x}_a + \mathbf{G}(\mathbf{y} - \mathbf{K}\mathbf{x}_a) \quad (6)$$

where  $\mathbf{G} = \hat{\mathbf{S}}\mathbf{K}^T \mathbf{S}_O^{-1}$  is the gain matrix and  $\hat{\mathbf{S}}$  is the posterior covariance matrix,

$$\hat{\mathbf{S}} = \left( \mathbf{K}^T \mathbf{S}_O^{-1} \mathbf{K} + \mathbf{S}_a^{-1} \right)^{-1} \quad (7)$$

The MAP solution can also be expressed in terms of the true value  $\mathbf{x}$  as

$$\hat{\mathbf{x}} = \mathbf{x}_a + \mathbf{A}(\mathbf{x} - \mathbf{x}_a) + \mathbf{G}\epsilon \quad (8)$$

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



where  $\mathbf{A}$  is the averaging kernel matrix that measures the error reduction resulting from the observations:

$$\mathbf{A} = \mathbf{GK} = \mathbf{I} - \hat{\mathbf{S}}\mathbf{S}_a^{-1} \quad (9)$$

and  $\mathbf{G}\boldsymbol{\epsilon}$  is the observation error in state space with error covariance matrix  $\mathbf{G}\mathbf{S}_o\mathbf{G}^T$ .

The analytical solution to the inverse problem thus provides full error characterization as part of the solution. It does require that the forward model be linear. The Jacobian matrix must be generally constructed numerically, requiring  $n$  sensitivity simulations with the forward model, and subsequent matrix operations are also of dimension  $n$ . This limits the practical size of the state vector. The matrix operations also depend on the dimension  $m$  of the observation vector but this can be easily addressed by splitting that vector into uncorrelated packets, a method known as sequential updating (Rodgers, 2000).

The limitation on the state vector size can be lifted by finding the solution to  $\nabla_x \mathcal{J} = \mathbf{0}$  numerically, rather than analytically, for example by using the adjoint of the forward model to calculate  $\nabla_x \mathcal{J}$  iteratively at successive approaches to the solution (e.g., Henze et al., 2007). This variational method allows for optimization of state vectors of any size because the Jacobian is not explicitly constructed. But it only yields the MAP solution,  $\hat{\mathbf{x}}$ , with no error statistics. Several approaches have been presented to obtain approximate error characterization (e.g., Desroziers et al., 2005; Chevallier et al., 2007) but they are computationally expensive. An excessively large state vector relative to the strength of the observational constraints also incurs smoothing error, as discussed above.

### 3 Quantifying aggregation and smoothing errors

The resolution of the forward model (e.g., grid resolution of the CTM) places an upper limit on the dimension for the state vector, which we call the native dimension. As we reduce the dimension of the state vector from this native resolution, the smoothing

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



error decreases while the aggregation error increases. Here we present analytical expressions for the aggregation and smoothing error covariance matrices and show how they can be used to select an optimal state vector dimension.

### 3.1 Aggregation error

- 5 As in Bocquet et al. (2011), we define a restriction (aggregation) operator that maps the native-resolution state vector  $\mathbf{x}$  of dimension  $n$  to a reduced-resolution vector  $\mathbf{x}_\omega$  of dimension  $p$ . We assume a linear restriction operator  $\mathbf{\Gamma}_\omega$  as a  $p \times n$  matrix relating  $\mathbf{x}_\omega$  to  $\mathbf{x}$ :

$$\mathbf{x}_\omega = \mathbf{\Gamma}_\omega \mathbf{x} \quad (10)$$

- 10 Bocquet et al. (2011) provide a detailed analysis of aggregation error for reduced-resolution state vectors. Their analysis relies heavily on the probabilistic construction of a prolongation operator ( $\mathbf{\Gamma}^*$ ) mapping  $\mathbf{x}_\omega$  back to  $\mathbf{x}$ :  $\mathbf{x} = \mathbf{\Gamma}^* \mathbf{x}_\omega$ . However, construction of this prolongation operator is not a well-posed problem because the operator is not unique. We present here a simpler and more robust method.

- 15 Aggregation error is the error introduced by aggregating state vector elements in the inversion. The relationship between the aggregated elements is not optimized as part of the inversion anymore and instead becomes an unoptimized parameter in the forward model, effectively increasing the forward model error and inhibiting the ability of the model to fit the observations. The aggregation error is thus a component of the observational error.

- 20 The aggregation error can be quantified by comparing the observational error incurred by using the native-resolution state vector,

$$\boldsymbol{\epsilon} = \mathbf{y} - \mathbf{K}\mathbf{x} \quad (11)$$

to that using the aggregated state vector,

$$\boldsymbol{\epsilon}_\omega = \mathbf{y} - \mathbf{K}_\omega \mathbf{x}_\omega \quad (12)$$

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Here  $\mathbf{y}$  is the observation vector (common in both cases), and  $\mathbf{x}$  and  $\mathbf{x}_\omega$  are the true values of the native-resolution and aggregated state vectors. The only difference between  $\epsilon$  and  $\epsilon_\omega$  is the aggregation of state vector elements. As such,

$$\epsilon_\omega = \epsilon + \epsilon_A \quad (13)$$

5 where  $\epsilon_A$  is the aggregation error. Rearranging,

$$\epsilon_A = (\mathbf{K} - \mathbf{K}_\omega \mathbf{\Gamma}_\omega) \mathbf{x} \quad (14)$$

Obtaining the error statistics for  $\epsilon_A$  requires knowledge of the pdf of  $\mathbf{x}$  for the ensemble of possible true states. Let  $\bar{\mathbf{x}}$  represent the mean value of this ensemble and  $\mathbf{S}_e$  the corresponding covariance matrix. The aggregation error covariance matrix is:

$$10 \quad \mathbf{S}_A = E \left[ (\epsilon_A - E[\epsilon_A]) (\epsilon_A - E[\epsilon_A])^T \right] \quad (15)$$

where  $E[\ ]$  is the expected value operator.  $E[\epsilon_A] = (\mathbf{K} - \mathbf{K}_\omega \mathbf{\Gamma}_\omega) \bar{\mathbf{x}}$  is the bias introduced by the aggregation. Replacing into Eq. (15):

$$\begin{aligned} \mathbf{S}_A &= (\mathbf{K} - \mathbf{K}_\omega \mathbf{\Gamma}_\omega) E \left[ (\mathbf{x} - \bar{\mathbf{x}}) (\mathbf{x} - \bar{\mathbf{x}})^T \right] (\mathbf{K} - \mathbf{K}_\omega \mathbf{\Gamma}_\omega)^T \\ &= (\mathbf{K} - \mathbf{K}_\omega \mathbf{\Gamma}_\omega) \mathbf{S}_e (\mathbf{K} - \mathbf{K}_\omega \mathbf{\Gamma}_\omega)^T \end{aligned} \quad (16)$$

15 In designing our inversion system we use  $\mathbf{x}_a$  as our best estimate of  $\bar{\mathbf{x}}$  and  $\mathbf{S}_a$  as our best estimate of  $\mathbf{S}_e$ . If  $\mathbf{x}_a = \mathbf{x}$  there is no aggregation error since the prior relationship assumed between state vector elements is correct, thus  $\mathbf{K} = \mathbf{K}_\omega \mathbf{\Gamma}_\omega$  and the aggregation bias is zero. Furthermore, assuming  $\mathbf{S}_a = \mathbf{S}_e$  allows us to calculate the aggregation error covariance matrix as

$$20 \quad \mathbf{S}_A = (\mathbf{K} - \mathbf{K}_\omega \mathbf{\Gamma}_\omega) \mathbf{S}_a (\mathbf{K} - \mathbf{K}_\omega \mathbf{\Gamma}_\omega)^T \quad (17)$$

and we will use this expression in the analysis that follows.

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

ACPD

## Balancing aggregation and smoothing errors

5 where  $\mathbf{e}_S = (\mathbf{I} - \mathbf{A})(\mathbf{x}_a - \mathbf{x})$  is the smoothing error. As pointed out by Rodgers (2000), the smoothing error statistics must be derived from the pdf of possible true states, in the same way as for the aggregation error and characterized by the error covariance matrix  $\mathbf{S}_e$ . For purposes of designing the inverse system we assume that  $\mathbf{S}_e = \mathbf{S}_a$ . Thus we have

We can also express the smoothing error in observation space,  $\mathbf{e}_S^*$ , (i.e., as a difference between  $\mathbf{y}$  and  $\mathbf{K}\hat{\mathbf{x}}$ ) by multiplying both sides of Eq. (18) by the Jacobian matrix:

so that

The corresponding smoothing error covariance matrix in observation space is

This expression can be generalized to compute the smoothing error covariance matrix in observation space for any reduced-dimension state vector  $\mathbf{x}_\omega$  with Jacobian  $\mathbf{K}_\omega$ , prior error covariance matrix  $\mathbf{S}_{\hat{\mathbf{x}}_\omega}$ , and averaging kernel matrix  $\mathbf{A}_\omega$ :

1009

### 3.3 Total error budget

From Eq. (18) we can see that the total error on  $\hat{x}$  without aggregation is  $\epsilon_T = \epsilon_S + \mathbf{G}\epsilon$  in the state space, or  $\epsilon_T^* = \epsilon_S^* + \mathbf{K}\mathbf{G}\epsilon$  in the observation space. The  $\mathbf{K}\mathbf{G}$  term in the observation space appears because we are interested in the error on  $\hat{x}$ . If  $\hat{x} = x$  then  $\mathbf{K}\mathbf{G} = \mathbf{I}$  and  $\mathbf{A} = \mathbf{I}$ , thus  $\epsilon_S = \mathbf{0}$  and our total error reverts to  $\epsilon$ ,

$$\epsilon_{\hat{x}=x}^* = \mathbf{K}(\mathbf{I} - \mathbf{A})(x_a - x) + \mathbf{K}\mathbf{G}\epsilon = \epsilon \quad (24)$$

Additional consideration of aggregation error for a reduced-dimension state vector  $x_\omega$  yields a total error in the state space

$$\epsilon_T = \epsilon_S + \mathbf{G}_\omega \epsilon + \mathbf{G}_\omega \epsilon_A \quad (25)$$

where  $\mathbf{G}_\omega$  is the gain matrix for the reduced-dimension state vector. In the observation space we get

$$\epsilon_T^* = \epsilon_S^* + \mathbf{K}_\omega \mathbf{G}_\omega \epsilon + \mathbf{K}_\omega \mathbf{G}_\omega \epsilon_A \quad (26)$$

From these relationships we derive the total error covariance matrix as

$$\begin{aligned} \mathbf{S}_{T,\omega} = & \underbrace{(\mathbf{I} - \mathbf{A}_\omega) \mathbf{S}_{a,\omega} (\mathbf{I} - \mathbf{A}_\omega)^T}_{\text{Smoothing Error}} \\ & + \underbrace{\mathbf{G}_\omega (\mathbf{K} - \mathbf{K}_\omega \mathbf{\Gamma}_\omega) \mathbf{S}_a (\mathbf{K} - \mathbf{K}_\omega \mathbf{\Gamma}_\omega)^T \mathbf{G}_\omega^T}_{\text{Aggregation Error}} \\ & + \underbrace{\mathbf{G}_\omega \mathbf{S}_O \mathbf{G}_\omega^T}_{\text{Observation Error}} \end{aligned} \quad (27)$$

in the state space and

$$\begin{aligned}
 \mathbf{S}_{T,\omega}^* = & \underbrace{\mathbf{K}_{\omega}(\mathbf{I} - \mathbf{A}_{\omega})\mathbf{S}_{a,\omega}(\mathbf{I} - \mathbf{A}_{\omega})^T\mathbf{K}_{\omega}^T}_{\text{Smoothing Error}} \\
 & + \underbrace{\mathbf{K}_{\omega}\mathbf{G}_{\omega}(\mathbf{K} - \mathbf{K}_{\omega}\mathbf{\Gamma}_{\omega})\mathbf{S}_a(\mathbf{K} - \mathbf{K}_{\omega}\mathbf{\Gamma}_{\omega})^T\mathbf{G}_{\omega}^T\mathbf{K}_{\omega}^T}_{\text{Aggregation Error}} \\
 & + \underbrace{\mathbf{K}_{\omega}\mathbf{G}_{\omega}\mathbf{S}_O\mathbf{G}_{\omega}^T\mathbf{K}_{\omega}^T}_{\text{Observation Error}}
 \end{aligned} \tag{28}$$

in the observation space.

Each of these three error terms depends on state vector dimension. Because the smoothing error increases with state vector dimension while the aggregation error decreases, we expect to find an optimal dimension where the total error is minimum. To enable an analytical inversion we may wish to decrease the state vector dimension further within a tolerance on aggregation error, such as requiring that the aggregation error remain smaller than the observation error. We give an example in Sect. 5.

A caveat in the above expressions for the aggregation and smoothing error covariance matrices is that they are valid only if the prior  $\mathbf{x}_a$  is the mean value  $\bar{\mathbf{x}}$  for the pdf of true states and if the error covariance matrix  $\mathbf{S}_a$  is the covariance matrix for that pdf. These conditions define the assumption for the prior, so the expressions can be taken as valid for the purpose of selecting an appropriate state vector dimension in an inverse problem. However, they should not be used to diagnose errors on the inversion results.

## 4 Aggregation methods

Aggregation of state vector elements to reduce the state vector dimension introduces aggregation errors, as described in Sect. 3.1. The aggregation error can be reduced by

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



grouping elements with correlated errors. Analyzing the off-diagonal structure of a precisely constructed prior error correlation matrix would provide the best objective way to carry out the aggregation. We generally lack such information but do have some qualitative knowledge of prior error correlation that can be used to optimize the aggregation. Bocquet et al. (2011) used tiling and tree-based aggregation methods, while Wecht et al. (2014) used a hierarchical clustering method based on prior error patterns. Here we compare three aggregation methods: (1) simple grid coarsening, (2) principal component analysis (PCA) clustering, and (3) a Gaussian mixture model (GMM) with radial basis functions (RBFs) to project native-resolution state vector elements to Gaussian pdfs. A qualitative illustration of these methods is shown in Fig. 1 for the aggregation of a native-resolution state vector of methane emissions with  $\frac{1}{2}^\circ \times \frac{2}{3}^\circ$  native grid resolution over North America (Turner et al., 2014). We focus here on spatial aggregation and assume that the state vector has no temporal dimension. However, the same methods can be used for temporal aggregation.

The simplest method for reducing the dimension of the state vector is to merge adjacent elements, i.e., neighboring grid cells. This method considers only spatial proximity as source of error correlation. It may induce large aggregation errors if proximal but otherwise dissimilar regions are aggregated together. In the case of methane emissions, aggregating neighboring wetlands and farmland would induce large errors because different processes drive methane emissions from these two source types.

The other two methods enable consideration of additional similarity factors besides spatial proximity when aggregating state vector elements. These similarity factors are expressed by vectors of dimension  $n$  describing correlative properties of the original native-resolution state vector elements. In the case of a methane source inversion, for example, we can choose as similarity vectors latitude and longitude to account for spatial proximity, but also wetland fraction to account for error correlations in the bottom-up wetland emission estimate used as prior.

## 4.1 Similarity matrix for aggregation

Table 1 lists the similarity vectors chosen for our example problem of estimating methane emissions (Turner et al., 2014). The first two vectors account for spatial proximity, the third represents the scaling factors from the first iteration of an adjoint-based inversion at native resolution (Wecht et al., 2014), and the others are the source type patterns from the bottom-up inventories used as prior. All similarity vectors are normalized and then weighted by judgment of their importance. Let  $\{\mathbf{c}_1, \dots, \mathbf{c}_K\}$  represent the  $K$  similarity vectors chosen for the problem ( $K = 14$  in our example of Table 1). We assemble them into a  $n \times K$  similarity matrix  $\mathbf{C}$ . We will also make use of the ensemble of similarity vector values for individual state vector elements, which we assemble into vectors  $\{\mathbf{c}'_1, \dots, \mathbf{c}'_n\}$  representing the rows of  $\mathbf{C}$ . Thus:

$$\mathbf{C} = \left[ \begin{pmatrix} \vdots \\ \mathbf{c}_1 \\ \vdots \end{pmatrix} \begin{pmatrix} \vdots \\ \mathbf{c}_2 \\ \vdots \end{pmatrix} \cdots \begin{pmatrix} \vdots \\ \mathbf{c}_K \\ \vdots \end{pmatrix} \right] = \begin{bmatrix} (\cdots \mathbf{c}'_1 \cdots) \\ (\cdots \mathbf{c}'_2 \cdots) \\ \vdots \\ (\cdots \mathbf{c}'_n \cdots) \end{bmatrix} \quad (29)$$

In this work all of the aggregation methods except for grid coarsening will use the same similarity matrix to construct the restriction operator.

This approach of using a similarity matrix  $\mathbf{C}$  to account for prior error covariances bears some resemblance to the geostatistical approach for inverse modeling (e.g., Michalak et al., 2004, 2005; Gourdji et al., 2008; Miller et al., 2012). The geostatistical approach specifies the prior estimate as  $\mathbf{x}_a = \mathbf{C}\boldsymbol{\beta}$  where  $\boldsymbol{\beta}$  is a vector of unknown drift coefficients to be optimized as part of the inversion. Here we use the similarity matrix to reduce the dimension of the state vector, rather than just as a choice of prior constraints.

## 4.2 Clustering with principal component analysis

In this method we cluster state vector elements following the principal components of the similarity matrix. It is generally not practical to derive the principal components in state vector space because the  $n$ -dimension is large. Instead we derive them in similarity space (dimension  $K$ ) as the eigenvectors of  $\mathbf{C}^T \mathbf{C}$  sorted in order of importance by their eigenvalues. The leading  $j$  principal components are kept for clustering. The reduced state vector is then constructed by grouping state vector elements that have the same sign patterns for all  $j$  principal components. Each unique  $j$ -dimensional sign pattern constitutes a cluster. The number of clusters defined in that way ranges between  $j$  and  $2^j$ . Figure 1b shows an example of applying this method to methane emissions in North America with reduction of the state vector to  $n = 8$ . The separation into four quadrants reflects the importance of latitude and longitude as error correlation factors. The additional separation within each quadrant isolates large from weak sources as defined by the prior.

## 4.3 Gaussian mixture model (GMM)

Here we use a Gaussian mixture model (GMM; Bishop, 2007) to project the native-resolution state vector onto  $p$  Gaussian pdfs using radial basis functions (RBFs). Mixture models are probabilistic models for representing a population comprised of  $p$  subpopulations. Each subpopulation is assumed to follow a pdf, in this case Gaussian. The Gaussians are  $K$ -dimensional where  $K$  is the number of similarity criteria. Each native-resolution state vector element is fit to this ensemble of Gaussians using RBFs as weighting factors.

The first step in constructing the GMM is to define a  $p \times n$  weighting matrix  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_p]^T$ . Each element  $w_{i,j}$  of this weighting matrix is the relative probability for native-resolution state vector element  $j$  to be described by Gaussian subpopulation

$i$ , i.e., “how much does element  $j$  look like Gaussian  $i$ ?”. It is given by

$$w_{i,j} = \frac{\pi_i \mathcal{N}(\mathbf{c}'_j | \boldsymbol{\mu}_i, \boldsymbol{\Lambda}_i)}{\sum_{k=1}^p \pi_k \mathcal{N}(\mathbf{c}'_j | \boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k)} \quad (30)$$

Here  $\mathbf{c}'_j$  is the  $j$ th row of the similarity matrix  $\mathbf{C}$ ,  $\boldsymbol{\mu}_i$  is a  $1 \times K$  row vector of means for the  $i$ th Gaussian,  $\boldsymbol{\Lambda}_i$  is a  $K \times K$  covariance matrix for the  $i$ th Gaussian, and  $\boldsymbol{\pi} = [\pi_1, \dots, \pi_p]^T$

5 is the relative weight of the  $p$  Gaussians in the mixture.  $\mathcal{N}(\mathbf{c}'_j | \boldsymbol{\mu}_i, \boldsymbol{\Lambda}_i)$  denotes the probability density of vector  $\mathbf{c}'_j$  on the normal distribution of Gaussian  $i$ . We define a  $p \times K$  matrix  $\mathbf{M}$  with rows  $\boldsymbol{\mu}_i$  and a  $K \times K \times p$  third-order tensor  $\mathcal{L} = [\boldsymbol{\Lambda}_1, \dots, \boldsymbol{\Lambda}_p]$  as the set of covariance matrices.

10 Projection of the native-resolution state vector onto the GMM involves four unknowns:  $\mathbf{W}$ ,  $\boldsymbol{\pi}$ ,  $\mathbf{M}$ , and  $\mathcal{L}$ . This is solved by constructing a cost function to estimate the parameters of the Gaussians in the mixture model using maximum likelihood:

$$\mathcal{J}_{\text{GMM}}(\mathbf{C} | \boldsymbol{\pi}, \mathbf{M}, \mathcal{L}) = \sum_{j=1}^n \ln \left\{ \sum_{i=1}^p \pi_i \mathcal{N}(\mathbf{c}'_j | \boldsymbol{\mu}_i, \boldsymbol{\Lambda}_i) \right\} \quad (31)$$

15 Starting from an initial guess for  $\boldsymbol{\pi}$ ,  $\mathbf{M}$ , and  $\mathcal{L}$  we compute the weight matrix  $\mathbf{W}$  using Eq. (30). We then differentiate the cost function with respect to  $\boldsymbol{\pi}$ ,  $\mathbf{M}$ , and  $\mathcal{L}$ , and set the derivative to zero to obtain (see Bishop, 2007):

$$\boldsymbol{\mu}_i = \Psi_i \sum_{j=1}^n w_{i,j} \mathbf{c}'_j \quad (32)$$

$$\boldsymbol{\Lambda}_i = \Psi_i \sum_{j=1}^n w_{i,j} (\mathbf{c}'_j - \boldsymbol{\mu}_i)^T (\mathbf{c}'_j - \boldsymbol{\mu}_i) \quad (33)$$

$$\pi_i = \frac{1}{n \Psi_i} \quad (34)$$

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



where:

$$\Psi_i = \sum_{j=1}^n \frac{1}{w_{i,j}} \tag{35}$$

The weights are re-calculated from the updated guesses of  $\mathbf{W}$ ,  $\boldsymbol{\pi}$ ,  $\mathcal{M}$ , and  $\mathcal{L}$  from Eqs. (32)–(35), and so on until convergence. The final weights define the restriction operator as  $\boldsymbol{\Gamma}_{\omega} = \mathbf{W}$ .

The GMM allows each native-resolution state vector element to be represented by a unique linear combination of the Gaussians through the RBFs. For a state vector of a given dimension, defined by the number of Gaussian pdfs, we can achieve high resolution for large localized sources by sacrificing resolution for weak or uniform source regions where resolution is not needed. This is illustrated in Fig. 2 with the resolution of southern California in an inversion of methane sources for North America. The figure shows the three dominant Gaussians describing emissions in Southern California and the corresponding RBF weights for each native-resolution grid square. Gaussian 1 is centered over Los Angeles and is highly localized, Gaussian 2 covers the Los Angeles Basin, and Gaussian 3 is a Southern California background. The sum of these three Gaussians accounts for most of the emissions in southern California and Nevada (which is mostly background). Additional Gaussians (not shown) resolve the southern San Joaquin Valley (large livestock and oil/gas emissions) and Las Vegas (large emissions from waste).

## 5 Application

We apply the aggregation methods described above to our example problem of estimating methane emissions from satellite observations of methane concentrations, focusing on selecting a reduced-dimension state vector that minimizes aggregation and smoothing errors. The inversion is described in detail in Turner et al. (2014) and uses GOSAT

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





satellite observations for 2009–2011 over North America. The forward model for the inversion is the GEOS-Chem CTM with  $\frac{1}{2}^{\circ} \times \frac{2}{3}^{\circ}$  grid resolution. The native-resolution state vector of methane emissions as defined on that grid includes 7366 elements.

For purpose of selecting an aggregated state vector for the inversion we consider a subset of observations for May 2010 ( $m = 6070$ ) so that we can afford to construct the corresponding Jacobian matrix  $\mathbf{K}$  at the native resolution; this is necessary to derive the aggregation error covariance matrix following Eq. (17). The prior error covariance matrix is specified as diagonal with 100 % uncertainty at the native resolution, decreasing with aggregation following the central limit theorem (Turner et al., 2014). The observational error covariance matrix is also diagonal and specified as the scene-specific retrieval error from Parker et al. (2011), which dominates the total observational error as shown by Turner et al. (2014). We compare the three methods presented in Sect. 4 for aggregating the state vector in terms of the implications for aggregation and smoothing errors for different state vector dimensions. In addition to the GMM with RBFs, we also consider a “GMM clustering” method where each native resolution state vector element is assigned exclusively to its dominant Gaussian pdf. This yields sharp boundaries between clusters (Fig. 1) as in the grid coarsening and PCA methods.

Figure 3 shows the mean error SD in the aggregation and smoothing error covariance matrices, computed as the square root of the diagonal terms, as a function of state vector dimension. The aggregation error is zero by definition at the native resolution (7366 state vector elements), and increases as the number  $n$  of state vector elements decreases, following a roughly  $n^{-0.7}$  dependence. Conversely, the smoothing error increases as the number of state vector elements increases, following roughly a  $\log(n)$  dependence. The different aggregation methods of Sect. 4 yield very similar smoothing errors but the aggregation error is somewhat improved using the GMM method. RBF weighting performs slightly better than GMM clustering (sharp boundaries). As discussed above, a major advantage of the GMM method is its ability to retain resolution of large localized sources after aggregation.

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Figure 4 shows the sum of contributions from aggregation, smoothing, and observational error SD as a function of state vector aggregation using the GMM with RBF weighting. In this application, aggregation error dominates for small state vectors ( $n < 100$ ), but drops below the observation error for  $n > 100$  and below the smoothing error for  $n > 1000$ . The smoothing error remains smaller than the observational error even at the native resolution ( $n = 7366$ ). The observational error is not independent of aggregation, as shown in Eq. (28), but we find here that the dependence is small.

From Fig. 4 we can identify a state vector dimension for which the total error is minimum ( $n = 2208$ ; circle in Fig. 4). However, error growth is small until  $n \approx 200$ , below which the aggregation error grows rapidly. A state vector of 369 elements, as adopted by Turner et al. (2014), does not incur significant errors associated with aggregation or smoothing, and enables computation of an analytical solution to the inverse problem with full error characterization.

## 6 Conclusions

We presented a method for optimizing the selection of the state vector in the solution of the inverse problem for a given ensemble of observations. The optimization involves minimizing the total error in the inversion by balancing the aggregation error (which increases as the state vector dimension decreases), the smoothing error (which increases as the state vector dimension increases), and the observational error. We further showed how one can reduce the state vector dimension within the constraints from the aggregation error in order to facilitate an analytical solution to the inverse problem with full error characterization.

We explored different methods for aggregating state vector elements as a means of reducing the dimension of the state vector. Aggregation error can be minimized by grouping state vector elements with the strongest correlated prior errors. We showed that a Gaussian mixture model (GMM), where the state vector elements are multi-dimensional Gaussian pdfs constructed from prior error correlation patterns, is a pow-

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



erful aggregation tool. Reduction of the state vector dimension using the GMM retains fine-scale resolution of important features in the native-resolution state vector while merging weak or uniform features.

*Acknowledgements.* For advice and discussions, we thank Kevin Wecht (Harvard University).

Special thanks to Robert Parker and Hartmut Boesch (University of Leicester) for providing the GOSAT observations. This work was supported by the NASA Carbon Monitoring System and by a Department of Energy (DOE) Computational Science Graduate Fellowship (CSGF) to AJT. We also thank the Harvard SEAS Academic Computing center for access to computing resources.

## References

Bishop, C. M.: Pattern Recognition and Machine Learning, Springer, 1st Edn., New York, 2007. 1014, 1015

Bocquet, M., Wu, L., and Chevallier, F.: Bayesian design of control space for optimal assimilation of observations. Part I: Consistent multiscale formalism, Q. J. Roy. Meteor. Soc., 137, 1340–1356, doi:10.1002/qj.837, 2011. 1003, 1007, 1012

Bousquet, P., Peylin, P., Ciais, P., Le Quere, C., Friedlingstein, P., and Tans, P. P.: Regional changes in carbon dioxide fluxes of land and oceans since 1980, Science, 290, 1342–1346, doi:10.1126/Science.290.5495.1342, 2000. 1003

Chevallier, F., Breon, F. M., and Rayner, P. J.: Contribution of the Orbiting Carbon Observatory to the estimation of CO<sub>2</sub> sources and sinks: theoretical study in a variational data assimilation framework, J. Geophys. Res.-Atmos., 112, D09307, doi:10.1029/2006jd007375, 2007. 1006

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Q. J. Roy. Meteor. Soc., 131, 3385–3396, doi:10.1256/qj.05.108, 2005. 1006

Gourdji, S. M., Mueller, K. L., Schaefer, K., and Michalak, A. M.: Global monthly averaged CO<sub>2</sub> fluxes recovered using a geostatistical inverse modeling approach: 2. Results including auxiliary environmental data, J. Geophys. Res., 113, D21115, doi:10.1029/2007jd009733, 2008. 1013

Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, Atmos. Chem. Phys., 7, 2413–2433, doi:10.5194/acp-7-2413-2007, 2007. 1006

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Kaminski, T. and Heimann, M.: Inverse modeling of atmospheric carbon dioxide fluxes, *Science*, 294, 259, doi:10.1126/science.294.5541.259a, 2001. 1003, 1004
- Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G.: On aggregation errors in atmospheric transport inversions, *J. Geophys. Res.*, 106, 4703, doi:10.1029/2000jd900581, 2001. 1003, 1004
- Michalak, A. M., Bruhwiler, L., and Tans, P. P.: A geostatistical approach to surface flux estimation of atmospheric trace gases, *J. Geophys. Res.*, 109, D14109, doi:10.1029/2003jd004422, 2004. 1013
- Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R., Peters, W., and Tans, P. P.: Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, *J. Geophys. Res.*, 110, D24107, doi:10.1029/2005jd005970, 2005. 1013
- Miller, S. M., Kort, E. A., Hirsch, A. I., Dlugokencky, E. J., Andrews, A. E., Xu, X., Tian, H., Nehrkorn, T., Eluszkiewicz, J., Michalak, A. M., and Wofsy, S. C.: Regional sources of nitrous oxide over the United States: seasonal variation and spatial distribution, *J. Geophys. Res.*, 117, D06310, doi:10.1029/2011jd016951, 2012. 1013
- Parker, R., Boesch, H., Cogan, A., Fraser, A., Feng, L., Palmer, P. I., Messerschmidt, J., Deutscher, N., Griffith, D. W. T., Notholt, J., Wennberg, P. O., and Wunch, D.: Methane observations from the Greenhouse Gases Observing SATellite: comparison to ground-based TCCON data and model calculations, *Geophys. Res. Lett.*, 38, L15807, doi:10.1029/2011gl047871, 2011. 1017
- Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding*, World Scientific, Singapore, 2000. 1003, 1006, 1009
- Schuh, A. E., Denning, A. S., Uliasz, M., and Corbin, K. D.: Seeing the forest through the trees: recovering large-scale carbon flux biases in the midst of small-scale variability, *J. Geophys. Res.*, 114, G03007, doi:10.1029/2008jg000842, 2009. 1004
- Turner, A. J., Jacob, D. J., Wecht, K., Maasakkers, J. D., Biraud, S. C., Boesch, H., Bowman, K. W., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Kuze, A., Notholt, J., Ohyama, H., Parker, R., Payne, V. H., Sussmann, R., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, *Atmos. Chem. Phys.*, available at: [http://acmg.seas.harvard.edu/publications/2014/Turner\\_2014\\_Methane.pdf](http://acmg.seas.harvard.edu/publications/2014/Turner_2014_Methane.pdf), submitted, 2014. 1003, 1012, 1013, 1016, 1017, 1018, 1022, 1025, 1026

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



von Clarmann, T.: Smoothing error pitfalls, Atmos. Meas. Tech., 7, 3023–3034, doi:10.5194/amt-7-3023-2014, 2014. 1003

Wecht, K. J., Jacob, D. J., Frankenberg, C., Jiang, Z., and Blake, D. R.: Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data, J. Geophys. Res.-Atmos., 119, 7741–7756, doi:10.1002/2014jd021551, 2014. 1003, 1012, 1013, 1022

ACPD

15, 1001–1026, 2015

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Table 1.** Similarity vectors for inverting methane emissions in North America<sup>a</sup>.

Similarity Vector	Weighting factor <sup>b</sup>
1. Latitude <sup>c</sup>	1.00
2. Longitude <sup>d</sup>	1.00
3. Initial scaling factors <sup>e</sup>	0.15
4. Wetland <sup>f</sup>	0.31
5. Livestock <sup>f</sup>	0.22
6. Oil/gas <sup>f</sup>	0.16
7. Waste <sup>f</sup>	0.15
8. Coal <sup>f</sup>	0.06
9. Soil absorption <sup>f</sup>	0.05
10. Termites <sup>f</sup>	0.02
11. Biomass burning <sup>f</sup>	0.02
12. Biofuel <sup>f</sup>	0.01
13. Rice <sup>f</sup>	0.01
14. Other <sup>f</sup>	0.01

<sup>a</sup> The  $K = 14$  similarity vectors describe prior error correlation criteria for the native-resolution state vector, representing here the methane emission in North America at the  $\frac{1}{2}^\circ \times \frac{2}{3}^\circ$  resolution of the GEOS-Chem chemical transport model. The criteria are normalized and then weighted (weighting factor).

<sup>b</sup> The weighting factors (dimensionless) measure the estimated relative importance of the different similarity criteria in determining prior error correlations in the state vector. For the prior emission patterns these weighting factors are the fractional contributions to total prior emissions in North America.

<sup>c</sup> Distance in kilometers from the equator.

<sup>d</sup> Distance in kilometers from the prime meridian.

<sup>e</sup> Initial scaling factors from one iteration of an adjoint inversion at the native resolution.

<sup>f</sup> Prior emissions used in the GEOS-Chem model (Wecht et al., 2014; Turner et al., 2014).

## Balancing aggregation and smoothing errors

A. J. Turner and  
D. J. Jacob

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

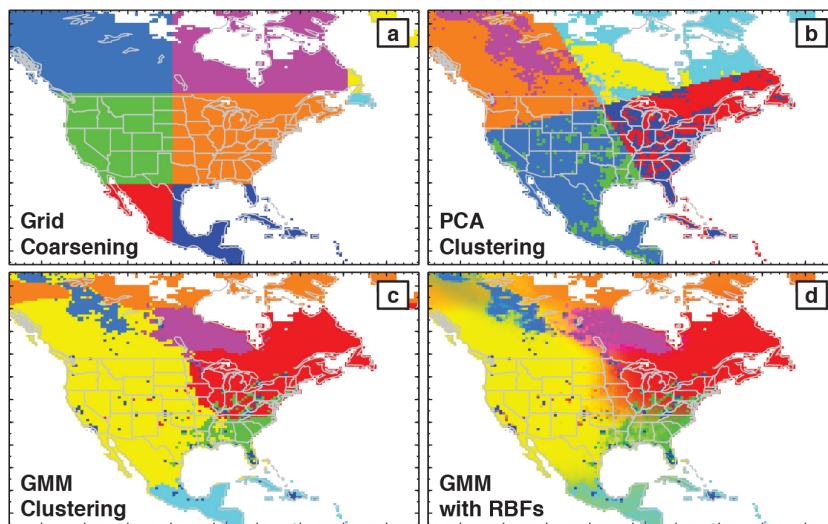
Printer-friendly Version

Interactive Discussion



**Balancing  
aggregation and  
smoothing errors**

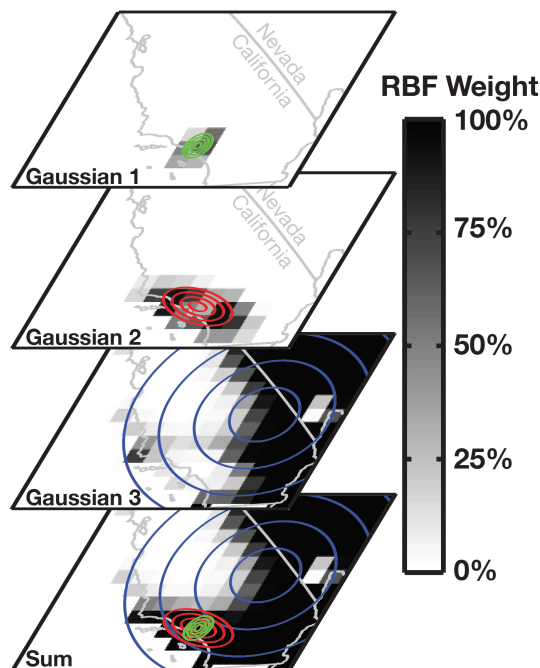
A. J. Turner and  
D. J. Jacob

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

**Figure 1.** Illustration of different approaches for aggregating a state vector. Here the native resolution state vector is a field of gridded methane emissions at  $\frac{1}{2}^\circ \times \frac{2}{3}^\circ$  resolution over North America. Extreme reduction to 8 state vector elements is shown with individual elements distinguished by color.

## Balancing aggregation and smoothing errors

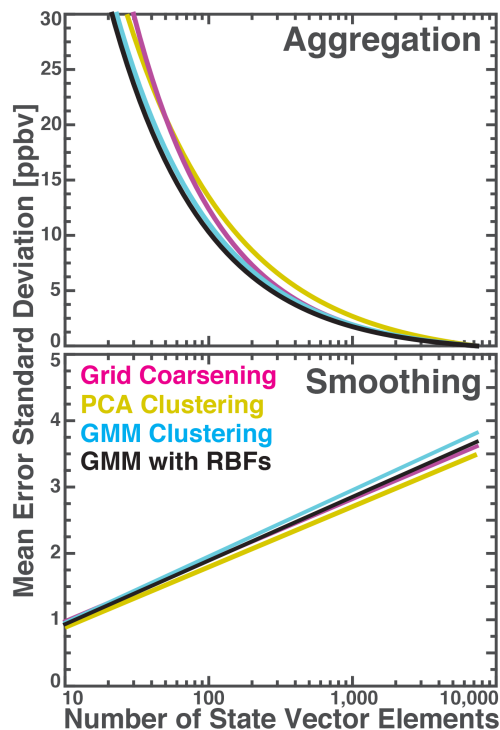
A. J. Turner and  
D. J. Jacob



**Figure 2.** Gaussian Mixture Model (GMM) representation of methane emissions in southern California with Gaussian pdfs as state vector elements. The Gaussians are constructed from a similarity matrix for methane emissions on the  $\frac{1}{2}^\circ \times \frac{2}{3}^\circ$  horizontal resolution of the GEOS-Chem CTM used as forward model for the inversion. The figure shows the dominant three Gaussians for southern California with contours delineating the 0.5, 1.0, 1.5, and 2.0  $\sigma$  spreads for the latitude–longitude dimensions. The RBF weights  $w_1$ ,  $w_2$ , and  $w_3$  of the three Gaussians for each  $\frac{1}{2}^\circ \times \frac{2}{3}^\circ$  grid square are also shown along with their sum.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)



**Figure 3.** Aggregation and smoothing error dependences on the aggregation of state vector elements in an inverse model. The application here is to an inversion of methane emissions over North America using satellite methane data with 7366 native-resolution state vector elements (Sect. 5 and Turner et al., 2014). Results are shown as the square roots of the means of the diagonal terms (mean error SD) in the aggregation and smoothing error covariance matrices. Different methods for aggregating the state vector (Sect. 4) are shown as separate lines. Note the log-scale on the  $x$  axis.

