Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions

Anna M. Michalak,[1,2] Adam Hirsch,[3,4] Lori Bruhwiler,[3] Kevin R. Gurney,[5] Wouter Peters,[3,4] and Pieter P. Tans[3]

Received 15 March 2005; revised 26 July 2005; accepted 31 August 2005; published 23 December 2005.

[1] This paper introduces a Maximum Likelihood (ML) approach for estimating the statistical parameters required for the covariance matrices used in the solution of Bayesian inverse problems aimed at estimating surface fluxes of atmospheric trace gases. The method offers an objective methodology for populating the covariance matrices required in Bayesian inversions, thereby resulting in better estimates of the uncertainty associated with derived fluxes and minimizing the risk of inversions being biased by unrealistic covariance parameters. In addition, a method is presented for estimating the uncertainty associated with these covariance parameters. The ML method is demonstrated using a typical inversion setup with 22 flux regions and 75 observation stations from the National Oceanic and Atmospheric Administration-Climate Monitoring and Diagnostics Laboratory (NOAA-CMDL) Cooperative Air Sampling Network with available monthly averaged carbon dioxide data. Flux regions and observation locations are binned according to various characteristics, and the variances of the model-data mismatch and of the errors associated with the a priori flux distribution are estimated from the available data.

Citation: Michalak, A. M., A. Hirsch, L. Bruhwiler, K. R. Gurney, W. Peters, and P. P. Tans (2005), Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, *J. Geophys. Res.*, *110*, D24107, doi:10.1029/2005JD005970.

1. Introduction

[2] The use of inverse modeling methods as tools for estimating fluxes of atmospheric trace gases has become increasingly common as the need to constrain their global and regional budgets has been recognized [*Intergovernmental Panel on Climate Change* (*IPCC*), 2001; *Committee on the Science of Climate Change, Division on Earth and Life Studies, National Research Council*, 2001; *Wofsy and Harriss*, 2002]. Inverse methods attempt to deconvolute the effects of atmospheric transport and recover source fluxes (typically surface fluxes) based on atmospheric measurements. Information about regions that are not being directly sampled can potentially be inferred from downwind atmospheric measurements. Inverse modeling methods have been used to estimate regional contributions to global budgets of trace gases such as CFCs, CH_4, and CO_2, and a review of recent applications is presented by *Enting* [2002, chap. 14–17].

[3] A vast majority of recent inverse modeling studies have relied on a classical Bayesian approach, where the solution to the inverse problem is defined as the set of flux values that represent an optimal balance between two requirements. The first criterion is that the optimized, or a posteriori, fluxes should be as close as possible to the first-guess (a priori) fluxes. The second is that the measurement values that would result from the inversion-derived (a posteriori) fluxes should agree as closely as possible with the actual measured concentrations. For the case where the advection scheme is linear, this solution corresponds to the minimum of a cost function $L_{\mathbf{s}}$ defined as:

$$L_{\mathbf{s}} = \frac{1}{2}(\mathbf{z} - \mathbf{H}\mathbf{s})^T \mathbf{R}^{-1}(\mathbf{z} - \mathbf{H}\mathbf{s}) + \frac{1}{2}(\mathbf{s} - \mathbf{s}_p)^T \mathbf{Q}^{-1}(\mathbf{s} - \mathbf{s}_p), \quad (1)$$

where $\mathbf{z}$ is an $n \times 1$ vector of observations, $\mathbf{H}$ is an $n \times m$ Jacobian matrix representing the sensitivity of the observations $\mathbf{z}$ to the function $\mathbf{s}$ (i.e., $H_{i,j} = \partial z_i / \partial s_j$), $\mathbf{s}$ is an $m \times 1$ vector of the discretized unknown surface flux distribution, $\mathbf{R}$ is the $n \times n$ model-data mismatch covariance, $\mathbf{s}_p$ is the $m \times 1$ prior estimate of the flux distribution $\mathbf{s}$, $\mathbf{Q}$ is the covariance of the errors associated with the prior estimate $\mathbf{s}_p$, and the superscript T denotes the matrix transpose operation. A solution in the form of a superposition of all statistical distributions involved can be computed, from which a posteriori means and covariances can be derived

[1]Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, Michigan, USA.
[2]Also at Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, Michigan, USA.
[3]Climate Monitoring and Diagnostics Laboratory, NOAA, Boulder, Colorado, USA.
[4]Also at Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA.
[5]Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, USA.

Copyright 2005 by the American Geophysical Union.
0148-0227/05/2005JD005970

[e.g., *Enting et al.*, 1995]. The solution is [*Tarantola*, 1987; *Enting*, 2002],

$$\hat{\mathbf{s}} = \mathbf{s}_p + \mathbf{QH}^T\left(\mathbf{HQH}^T + \mathbf{R}\right)^{-1}\left(\mathbf{z} - \mathbf{Hs}_p\right) \quad (2)$$

$$\mathbf{V}_{\hat{\mathbf{s}}} = \mathbf{Q} - \mathbf{QH}^T\left(\mathbf{HQH}^T + \mathbf{R}\right)^{-1}\mathbf{HQ}, \quad (3)$$

where $\hat{\mathbf{s}}$ is the posterior best estimate of $\mathbf{s}$, and $\mathbf{V}_{\hat{\mathbf{s}}}$ is its posterior covariance.

[4] In most Bayesian studies, both **R** and **Q** have been modeled as diagonal matrices. In this case, the diagonal elements of **R** represent the model-data mismatch variance of each observation, which is the sum of the variances associated with all error components such as, for instance, the observation error, the transport modeling error, and the representation error. The diagonal elements of **Q**, on the other hand, represent the error variance of the prior flux estimates and specify the extent to which the real fluxes are expected to deviate from prior flux estimates. It is important to note that although the vast majority of Bayesian inversion studies have used diagonal covariance matrices, some errors are known to be both temporally and spatially correlated.

[5] One of the challenges of the Bayesian approach is the need to estimate the parameters defining the model-data mismatch covariance matrix **R** and the prior error covariance matrix **Q**. These covariances determine the relative weight of prior information versus available data in estimating individual fluxes, and are therefore key components in estimating the posterior covariance (and thereby the uncertainty) of these fluxes. As a result, identifying appropriate covariance parameters is essential to accurate flux estimation. The importance and challenge of accurately estimating these parameters [*Kaminski et al.*, 1999; *Rayner et al.*, 1999; *Law et al.*, 2002; *Peylin et al.*, 2002; *Engelen et al.*, 2002] and the lack of objective methods for doing so [*Rayner et al.*, 1999] have been increasingly recognized in the literature.

[6] Past inversion studies have relied on a variety of methods for estimating these parameters. In the original synthesis inversion work of *Enting et al.* [1995], the data uncertainty was based on a statistical characteristics of the NOAA flask sampling procedures [*Tans et al.*, 1990], but the uncertainty associated with errors in the atmospheric transport model could not be quantified. Many more recent studies have derived model-data mismatch from the residual standard deviation of flask samples around a smooth curve fit [e.g., *Hein et al.*, 1997; *Bousquet et al.*, 1999; *Gurney et al.*, 2002; *Peylin et al.*, 2002]. Others have relied on values independently obtained from the literature [e.g., *Kandlikar*, 1997]. The a priori flux errors have been even more difficult to quantify [*Kaminski et al.*, 1999], and the choice of prior errors has even been described as "mostly arbitrary" in some studies [*Bousquet et al.*, 1999], even though it is recognized that these parameters are crucial to the inversion. Often, researchers have applied what are considered to be "loose" priors [e.g., *Peylin et al.*, 2002; *Law et al.*, 2002] in order to yield conservative estimates of the flux uncertainties. On the basis of assessments of available data, oceanic fluxes have usually been considered more certain than terrestrial fluxes [e.g., *Kaminski et al.*, 1999]. Although considerable effort has been put into estimating covariance parameters, the specification of the prior uncertainties has been described as the "greatest single weakness" in some studies [*Rayner et al.*, 1999].

[7] Recently, several researchers have scaled the model-data mismatch and/or prior error covariance parameters to obtain a data misfit function that follows a χ^2 distribution with a given number of degrees of freedom [e.g., *Rayner et al.*, 1999; *Gurney et al.*, 2002; *Peylin et al.*, 2002; *Rödenbeck et al.*, 2003]. The correct number of degrees of freedom to be used in such an analysis is equal to the total number of independent pieces of information introduced into the system (equal to the number of observational data plus the number of prior flux estimates in the case of diagonal **Q** and **R** matrices), minus the number of variables estimated in the inversion (typically equal to the number of estimated fluxes). If the covariance parameters are reasonable, the sum of the squared residuals, scaled by their uncertainties and normalized by the number of degrees of freedom, should be close to 1 (reduced chi-squared $\chi_r^2 = 1$). Some researchers have applied this test to the residuals between the available observations and those that would result from the a posteriori fluxes obtained from the inversion, while others have applied it both to these observation residuals and to residuals of the a posteriori fluxes from their prior estimates. Such tuning, however, does not yield a unique solution because more than one combination of covariance parameters can lead to the residuals having an acceptable χ_r^2. In addition, looking at the χ_r^2 cannot guide the relative allocation of error between the model-data mismatch and prior flux estimates [*Rayner et al.*, 1999]. Therefore examining the variance of the residuals is a necessary, but not a sufficient, condition for evaluating the appropriateness of covariance parameters.

[8] A few recent studies have attempted to systematically quantify one or more components of the model-data mismatch. *Engelen et al.* [2002] divided the model-data mismatch covariance into four additive covariances (which were modeled as diagonal matrices) representing: (1) the observation error, (2) the error in mapping concentrations at a specific location into the measured quantity, (3) the model transport error, and (4) the representation error describing the effects of model resolution. Each of these error sources was then quantified based either on available additional information (e.g., known precision of analytical methods) or numerical experiments (e.g., comparing inversion results obtained using various transport models in order to estimate transport error). *Kaminski et al.* [2001] focused on the errors introduced as a result of imposing potentially erroneous fixed flux patterns within regions, and estimated the additional variance that should be added to the model-data mismatch covariance to account for this effect. *Krakauer et al.* [2004] estimated single scaling factors for both the model-data mismatch and prior error covariances using a generalized cross-validation approach, to show that the Transcom inversion results for the global land/ocean partitioning, particularly in the Southern Hemisphere and tropical regions, were influenced strongly by the Transcom choice of parameters. The fluxes resulting from this analysis and their associated residuals were not analyzed to determine whether these parameters result in fluxes that are consistent with the underlying statistical assumptions, and the uncertainty associated with the estimated covariance parameters was not quantified.

[9] The method that will be developed in this paper allows the available data to shed light on the covariance parameters to be used in the inversion, both in defining the model-data mismatch and the error associated with the prior flux distribution. This can be done in a consistent manner by identifying the Maximum Likelihood (ML) estimates of these parameters, given the prior flux estimates $\mathbf{s}_p$, the available measurements $\mathbf{z}$, and the sensitivity $\mathbf{H}$ of these measurements to the fluxes to be estimated. The method is applicable whenever estimates of the covariance parameters are to be obtained from the atmospheric data themselves. A ML approach has previously been used in estimating spatial drift and covariance parameters of hydrologic variables, based on limited measurements of these parameters [*Kitanidis and Lane*, 1985]. Also, a related restricted maximum likelihood (RML) approach has been used for estimating covariance parameters in geostatistical inverse modeling [e.g., *Kitanidis*, 1995; *Michalak and Kitanidis*, 2004]. Recently, this technique was applied to the estimation of covariance parameters in a geostatistical implementation of the atmospheric trace-gas inversion problem [*Michalak et al.*, 2004]. In this paper, we develop and demonstrate a maximum likelihood methodology for estimating the model-data mismatch and prior-error covariance matrix parameters needed in the application of the classical Bayesian inverse modeling approach. This method can be directly applied to all studies where the objective function is of the form presented in equation (1), which is the most common setup currently being used in atmospheric inversion studies. This method allows, for the first time, for an objective and data-driven estimation of both the model-data mismatch and prior error covariance parameters required for the solution of the flux estimation inverse problem. Several covariance parameters can be estimated simultaneously, the resulting fluxes yield residuals that follow the assumed distribution (e.g., $\chi^2_r = 1$), and the uncertainty associated with the covariance parameters can also be estimated. In the sample application presented in this work, both $\mathbf{R}$ and $\mathbf{Q}$ are modeled as diagonal matrices, but the method is directly applicable to cases where these matrices include spatial and temporal correlation (i.e., off-diagonal terms). In such cases, parameters such as the correlation length of flux deviations from their prior estimates can also be estimated.

2. Maximum Likelihood Estimation

[10] We are interested in finding the maximum likelihood estimate of the model-data mismatch and prior error covariance parameters in Bayesian inverse modeling. Note that throughout this discussion, the term prior errors refers to the errors between the actual fluxes and the prior flux estimates. In the case where both the model-data mismatch matrix and the prior error covariance matrix are diagonal matrices of variances, this can be thought of as optimizing the variances in

$$\mathbf{R} = \begin{bmatrix} \sigma^2_{R,1} & 0 & \cdots & 0 \\ 0 & \sigma^2_{R,2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma^2_{R,n} \end{bmatrix}, \mathbf{Q} = \begin{bmatrix} \sigma^2_{Q,1} & 0 & \cdots & 0 \\ 0 & \sigma^2_{Q,2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma^2_{Q,m} \end{bmatrix}, \tag{4}$$

where $\sigma^2_{R,i}$ are the model-data mismatch variances of individual observations and $\sigma^2_{Q,i}$ are the prior-error variances of individual prior flux estimates. In the simplest case, we could assume that a single variance describes the model-data mismatch at all sites, and another variance describes the prior error for all source regions, yielding a total of two parameters to be estimated. On the other end of the spectrum, a different model-data mismatch could be estimated for each measurement location and each source region. Other options would be to populate the covariance matrices with scaling factors that we believe to be representative of the relative errors of certain measurements or priors and solve for overall proportionality constants that would adjust the magnitude of these variances, or to estimate the covariance (in space and/or in time) of the errors. Various options will be demonstrated in section 3 of this paper. It is important to note that one would not go as far as estimating a different variance for each individual flask measurement, because one cannot estimate a variance using a single measurement.

2.1. Probability Density Function

[11] We are interested in the probability density function of the covariance parameters of $\mathbf{R}$ and $\mathbf{Q}$, which we will jointly call $\boldsymbol{\theta}$, given the available observations $\mathbf{z}$, the prior estimates $\mathbf{s}_p$, and the transport matrix $\mathbf{H}$. In the example given in equation (4) $\boldsymbol{\theta} = \{\sigma^2_{R,1},\ldots, \sigma^2_{R,n}, \sigma^2_{Q,1},\ldots, \sigma^2_{Q,m}\}$, which would be binned into subgroups based on flux regions and measurement stations, and potentially further grouped by regions or stations that are expected to have similar properties (e.g., marine boundary layer stations). According to Bayes' rule, the pdf of the parameters $\boldsymbol{\theta}$ can be defined as

$$p''(\boldsymbol{\theta}|\mathbf{H}, \mathbf{s}_p, \mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{H}, \mathbf{s}_p, \boldsymbol{\theta})p'(\boldsymbol{\theta}|\mathbf{H}, \mathbf{s}_p)}{\int p(\mathbf{z}|\mathbf{H}, \mathbf{s}_p, \boldsymbol{\theta})p'(\boldsymbol{\theta}|\mathbf{H}, \mathbf{s}_p)\, d\boldsymbol{\theta}}, \tag{5}$$

where the denominator is simply a normalizing constant equal to the probability of the data $p(\mathbf{z}|\mathbf{H}, \mathbf{s}_p)$, and where a prime denotes a prior distribution while a double prime denotes a posterior distribution. We assume that, a priori, the probability of the parameters $\boldsymbol{\theta}$ is uniform over all values, $p'(\boldsymbol{\theta}|\mathbf{H},\mathbf{s}_p) \propto 1$, which corresponds to the assumption that we do not have any prior knowledge about these parameters. Then

$$p''(\boldsymbol{\theta}|\mathbf{H}, \mathbf{s}_p, \mathbf{z}) \propto p(\mathbf{z}|\mathbf{H}, \mathbf{s}_p, \boldsymbol{\theta}), \tag{6}$$

yielding the likelihood function of $\boldsymbol{\theta}$. In other words, if we assume that we have no prior information on $\boldsymbol{\theta}$, its pdf is defined by the likelihood of the observations, and the best estimate corresponds to the set of values that maximize this likelihood.

[12] The observations $\mathbf{z}$ are modeled as

$$\mathbf{z} = \mathbf{H}\mathbf{s} + \varepsilon_z, \tag{7}$$

where ε_z is a vector representing the model-data mismatch, typically modeled as a vector of $n \times 1$ normally distributed

random numbers with variance(s) defined by the matrix **R**. A priori, the expected value of **z** is

$$E[\mathbf{z}] = E[\mathbf{Hs} + \boldsymbol{\varepsilon}_z] = \mathbf{Hs}_p. \tag{8}$$

The covariance matrix of **z** is

$$\begin{aligned} &E\left[(\mathbf{z} - E[\mathbf{z}])(\mathbf{z} - E[\mathbf{z}])^T\right] \\ &= E\left[(\mathbf{Hs} + \boldsymbol{\varepsilon}_z - E[\mathbf{Hs} + \boldsymbol{\varepsilon}_z])(\mathbf{Hs} + \boldsymbol{\varepsilon}_z - E[\mathbf{Hs} + \boldsymbol{\varepsilon}_z])^T\right] \\ &= \mathbf{HQH}^T + \mathbf{R}. \end{aligned} \tag{9}$$

This expected value and covariance can be used to define the Gaussian probability density function $p(\mathbf{z}|\mathbf{H}, \mathbf{s}_p, \boldsymbol{\theta})$, which, from (6) is proportional to $p''(\boldsymbol{\theta}|\mathbf{H}, \mathbf{s}_p, \mathbf{z})$:

$$\begin{aligned} p''(\boldsymbol{\theta}|\mathbf{H}, \mathbf{s}_p, \mathbf{z}) &\propto p(\mathbf{z}|\mathbf{H}, \mathbf{s}_p, \boldsymbol{\theta}) \\ &\propto \left|\mathbf{HQH}^T + \mathbf{R}\right|^{-1/2} \\ &\cdot \exp\left(-\frac{1}{2}(\mathbf{z} - \mathbf{Hs}_p)^T(\mathbf{HQH}^T + \mathbf{R})^{-1}(\mathbf{z} - \mathbf{Hs}_p)\right), \end{aligned} \tag{10}$$

where $|\ |$ denotes a matrix determinant, and **Q** and **R** are functions of $\boldsymbol{\theta}$. Note that the above equation is normally distributed with respect to **z** but not $\boldsymbol{\theta}$. Also, this pdf is independent of the actual flux distribution **s**.

2.2. Best Estimate

[13] In order to obtain the maximum likelihood estimate of the covariance parameters, we want to maximize equation (10) with respect to $\boldsymbol{\theta}$, or alternately, minimize its negative logarithm,

$$L_\theta = \frac{1}{2}\ln\left|\mathbf{HQH}^T + \mathbf{R}\right| + \frac{1}{2}(\mathbf{z} - \mathbf{Hs}_p)^T(\mathbf{HQH}^T + \mathbf{R})^{-1}(\mathbf{z} - \mathbf{Hs}_p). \tag{11}$$

Note that higher variance values in **Q** and **R** will tend to increase the value of the first term and decrease the value of the second term in this equation. The $\boldsymbol{\theta}$ values that minimize this overall objective function are the ML estimates of these parameters.

[14] Conceptually, differences between the available observations and observations predicted by the prior flux estimates result from two types of error. First, the model-data mismatch error, parameterized by the covariance matrix **R**, results from random measurement and transport errors. Second, the prior error that is parameterized by the covariance matrix **Q** accounts for more systematic errors in the predicted observations because, as a result of the mixing that takes place as gases are transported downwind, error in a single flux region will be sampled at multiple observation locations. Although these two sources of error cannot be separated perfectly, the maximum likelihood approach offers a rigorous method for estimating the contributions of these two forms of error. Looking at the objective function in equation (11), the model-data mismatch error in **R** is included directly in the optimization, whereas the prior error in **Q** is 'filtered' using the transport matrix **H**. Therefore, the effect of the model-data mismatch is expected to behave as an independent, identically distributed (i.i.d.) additive error for the case of a diagonal **R**, whereas the effect of the prior error will have a more systematic effect on a group of observations. The intuitive separation between the model-data mismatch and prior error is illustrated in Figure 1.

2.3. Parameter Uncertainty

[15] An estimate of the uncertainty of parameters $\boldsymbol{\theta}$ can be obtained from the inverse of the Hessian of equation (11),

$$\mathbf{V}_\theta = (\mathcal{H}_{ij})^{-1}, \tag{12}$$

where the Hessian is defined as

$$\mathcal{H}_{ij} = \frac{\partial^2 L_\theta}{\partial\theta_i\partial\theta_j} \tag{13}$$

and has dimensions $p \times p$, where p is the total number of parameters to be estimated. In order to simplify the notation, we define the additional variables,

$$\boldsymbol{\Psi} = \mathbf{HQH}^T + \mathbf{R} \tag{14}$$

$$\boldsymbol{\Psi}_j = \mathbf{H}\frac{\partial\mathbf{Q}}{\partial\theta_j}\mathbf{H}^T + \frac{\partial\mathbf{R}}{\partial\theta_j} \tag{15}$$

$$\boldsymbol{\Psi}_{ij} = \mathbf{H}\frac{\partial^2\mathbf{Q}}{\partial\theta_i\partial\theta_j}\mathbf{H}^T + \frac{\partial^2\mathbf{R}}{\partial\theta_i\partial\theta_j}, \tag{16}$$

which all have dimensions $n \times n$. Also, we will use the following properties [*Schweppe*, 1973]:

$$\frac{\partial}{\partial\theta_j}\boldsymbol{\Psi}^{-1} = -\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_j\boldsymbol{\Psi}^{-1} \tag{17}$$

$$\frac{\partial}{\partial\theta_j}(\ln|\boldsymbol{\Psi}|) = Tr\left[\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_j\right], \tag{18}$$

where Tr is the "trace" operator, which is the sum of the diagonal elements of a matrix. Given these additional variables, the Hessian becomes

$$\begin{aligned} \mathcal{H}_{ij} = &-\frac{1}{2}Tr\left[\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_i\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_j\right] + \frac{1}{2}Tr\left[\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_{ij}\right] \\ &+ (\mathbf{z} - \mathbf{Hs}_p)^T\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_i\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_j\boldsymbol{\Psi}^{-1}(\mathbf{z} - \mathbf{Hs}_p) \\ &- \frac{1}{2}(\mathbf{z} - \mathbf{Hs}_p)^T\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_{ij}\boldsymbol{\Psi}^{-1}(\mathbf{z} - \mathbf{Hs}_p). \end{aligned} \tag{19}$$

Note that in the case where both **R** and **Q** are diagonal matrices of variances, $\boldsymbol{\Psi}_{ij} = [0]$, and $\mathcal{H}_{ij}$ simplifies to

$$\begin{aligned} \mathcal{H}_{ij} = &-\frac{1}{2}Tr\left[\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_i\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_j\right] \\ &+ (\mathbf{z} - \mathbf{Hs}_p)^T\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_i\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_j\boldsymbol{\Psi}^{-1}(\mathbf{z} - \mathbf{Hs}_p). \end{aligned} \tag{20}$$

[16] In some other cases, the second derivative $\boldsymbol{\Psi}_{ij}$ may be difficult or expensive to compute, in which case the

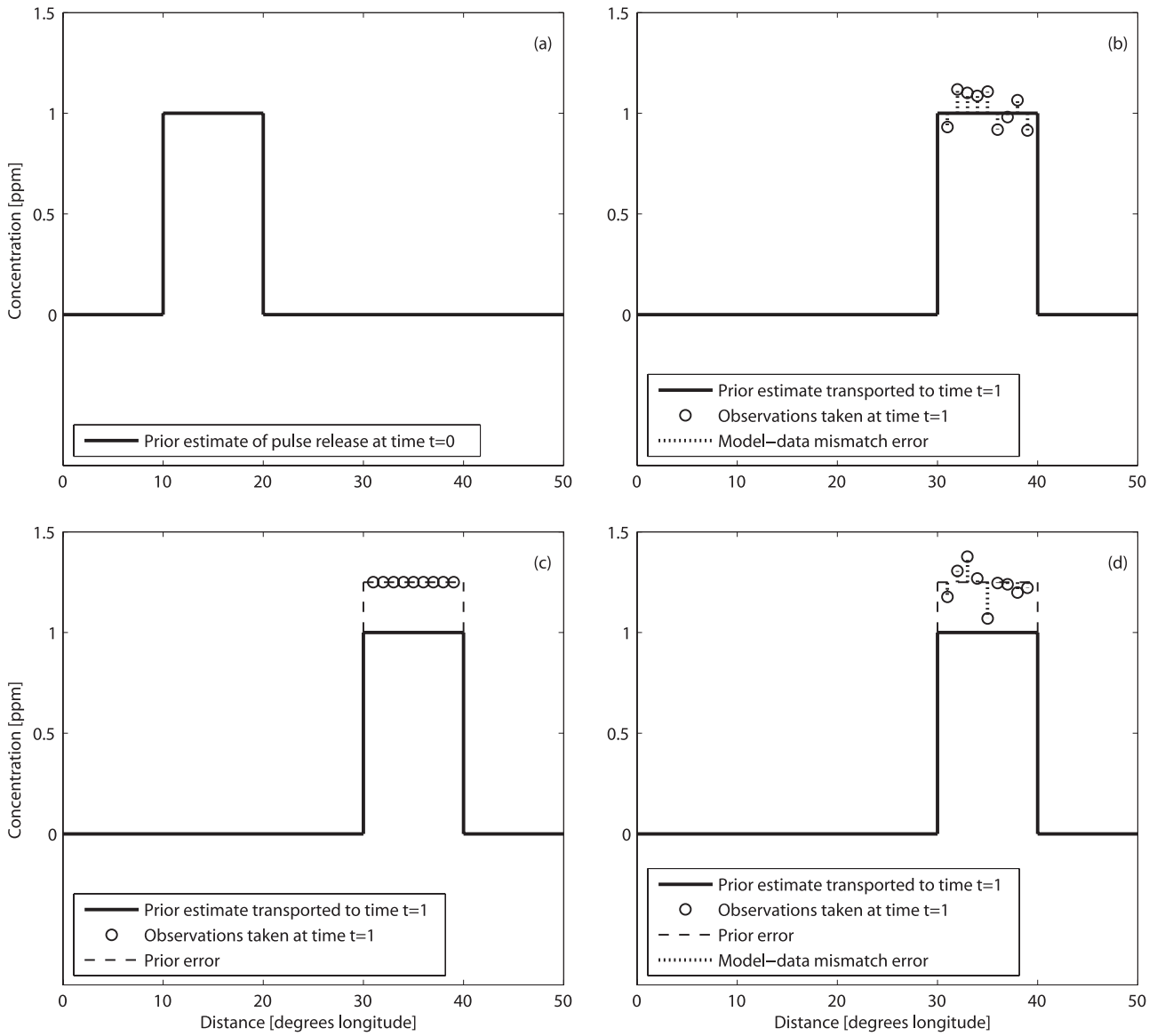

Figure 1. This figure presents a conceptualization of (a) a prior flux estimate and (b, c, d) three possible scenarios of how actual observations relate to this prior flux (transported to the time when observations are made). It illustrates the difference between the effects of model-data mismatch error and prior error. The prior estimate of a flux is depicted in Figure 1a. After a given time, the compound is advected (assuming only advection and no mixing) to a downwind location. On the basis of the prior estimate in Figure 1a, the a priori predicted distribution at the new time is presented as the thick line in Figures 1b, 1c, and 1d. The effect of model-data mismatch is expected to be relatively independent for each observation and is presented in Figure 1b. The effect of prior error in the magnitude of the flux is correlated among several observations and is presented in Figure 1c. Their combined effect is presented in Figure 1d.

covariance of the parameters can be approximated using the Cramer-Rao inequality [*Schweppe*, 1973] by

$$\mathbf{V}_\theta \geq \left(\mathcal{F}_{ij}\right)^{-1}, \tag{21}$$

where $\mathcal{F}_{ij}$ is the Fisher information matrix and is the expected value, with respect to $\mathbf{z}$, of the Hessian $\mathcal{H}_{ij}$:

$$\begin{aligned}\mathcal{F}_{ij} = E\left[\mathcal{H}_{ij}\right] &= -\frac{1}{2}Tr\left[\mathbf{\Psi}^{-1}\mathbf{\Psi}_i\mathbf{\Psi}^{-1}\mathbf{\Psi}_j\right] + \frac{1}{2}Tr\left[\mathbf{\Psi}^{-1}\mathbf{\Psi}_{ij}\right] \\ &+ Tr\left[\mathbf{\Psi}^{-1}\mathbf{\Psi}_i\mathbf{\Psi}^{-1}\mathbf{\Psi}_j\mathbf{\Psi}^{-1}\right]E\left[\left(\mathbf{z}-\mathbf{H}\mathbf{s}_p\right)^T\left(\mathbf{z}-\mathbf{H}\mathbf{s}_p\right)\right] \\ &- \frac{1}{2}Tr\left[\mathbf{\Psi}^{-1}\mathbf{\Psi}_{ij}\mathbf{\Psi}^{-1}\right]E\left[\left(\mathbf{z}-\mathbf{H}\mathbf{s}_p\right)^T\left(\mathbf{z}-\mathbf{H}\mathbf{s}_p\right)\right] \\ &= \frac{1}{2}Tr\left[\mathbf{\Psi}^{-1}\mathbf{\Psi}_i\mathbf{\Psi}^{-1}\mathbf{\Psi}_j\right].\end{aligned} \tag{22}$$

The diagonal elements of $\mathbf{V}_\theta$ represent the uncertainty of the estimates of the covariance parameters, as defined by their estimation variance. The uncertainty of the covariance parameters could be incorporated into the subsequent inversion by drawing an ensemble of samples of the covariance parameters based on their best estimates and uncertainty, and solving the inversion for each set of parameters [e.g., *Kitanidis*, 1986].

2.4. Gauss-Newton Method

[17] In general, an iterative method is needed to find the minimum of equation (11) with respect to the vector of values $\boldsymbol{\theta}$. One efficient option is the Gauss-Newton algorithm [e.g., *Gill et al.*, 1986]. Starting from an initial estimate of $\boldsymbol{\theta}$ denoted $\hat{\boldsymbol{\theta}}_k$ the algorithm proceeds as

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k - \mathcal{F}_k^{-1}\mathbf{g}_k, \tag{23}$$

where $\mathcal{F}$ is the expected value, with respect to $\mathbf{z}$, of the Hessian $\mathcal{H}$ of the likelihood function, $\mathcal{H}_{ij}$ and $\mathcal{F}_{ij}$ are as defined in equations (19) and (22), $\mathbf{g}$ is a vector of the first derivatives of the likelihood function L_θ with respect to $\boldsymbol{\theta}$,

$$g_j = \frac{\partial L_\theta}{\partial \theta_j} = \frac{1}{2} Tr\left[\boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_j\right] - \frac{1}{2}\left(\mathbf{z} - \mathbf{H}\mathbf{s}_p\right)^T \boldsymbol{\Psi}^{-1}\boldsymbol{\Psi}_j\boldsymbol{\Psi}^{-1}\left(\mathbf{z} - \mathbf{H}\mathbf{s}_p\right) \tag{24}$$

and the subscript k in $\mathbf{g}$ and $\mathcal{F}$ means that they are calculated using $\hat{\boldsymbol{\theta}}_k$. The indices take on the values $i = 1,\ldots, p$ and $j = 1,\ldots, p$, where p is the total number of parameters $\boldsymbol{\theta}$ to be estimated. Note that $\mathbf{g}_k$ is calculated using the latest estimate $\hat{\boldsymbol{\theta}}_k$. In the special case where both $\mathbf{R}$ and $\mathbf{Q}$ are linear in $\boldsymbol{\theta}$, $\mathbf{g}_k$ is a constant vector.

2.5. Method Validation

[18] In order to demonstrate and validate the proposed method, sample applications are presented in the section 3. For each case, once the covariance parameters and their uncertainties have been estimated, the results are examined by performing a standard Bayesian inversion (equations (2) and (3)) using the estimated parameters. The residuals from these inversions are analyzed to ensure that the method has successfully identified covariance parameters consistent with the statistical setup of the inversion.

[19] As discussed by *Tarantola* [1987], the squared residuals from the inversions should follow a χ^2 distribution. If both $\mathbf{R}$ and $\mathbf{Q}$ are diagonal and the residuals are calculated using the posterior best estimate of the flux distribution, the sum of the squared data and flux residuals, normalized by the variances in $\mathbf{R}$ and $\mathbf{Q}$, should follow a χ^2 distribution with n degrees of freedom [*Tarantola*, 1987]. This results from the fact that the residuals are not independent, and there are only n degrees of freedom among the $n + m$ residuals for the case where the $\mathbf{R}$ and $\mathbf{Q}$ matrices are diagonal. This criterion is applicable to the full set of residuals, and cannot be applied directly to only flux or observation residuals, or to residuals from individual stations or regions. Note that the covariance parameters are treated as deterministic parameters in the inversion step, and the number of covariance parameters therefore does not affect the number of degrees of freedom of the residual χ^2 distribution.

[20] If, instead of using the best estimates of the fluxes, the residuals from prior fluxes and observations are calculated using conditional realizations of the a posteriori fluxes (see Appendix A), the residuals are expected to follow the statistical distributions specified in the covariance matrices $\mathbf{R}$ and $\mathbf{Q}$. If these matrices are diagonal, this implies that the residuals from the actual fluxes (and from the conditional realizations) are expected to be independent. The conditional realizations represent the range of possible flux distributions, given the assumptions and data incorporated into the inversion. For the case where residuals are expected to be independent, the number of degrees of freedom is equal to the number of residuals. Therefore the sum of the squared residuals from individual conditional realizations, normalized by the variances in $\mathbf{R}$ and $\mathbf{Q}$, should follow a χ^2 distribution with $n + m$ degrees of freedom. More importantly, because the residuals should follow the distributions specified in $\mathbf{R}$ and $\mathbf{Q}$, which do not include a cross-covariance between data and flux residuals, observation residuals can be analyzed separately from flux residuals to determine whether both the observations and the prior fluxes are reproduced to the extent assumed by the parameters in the model-data mismatch and prior error covariance matrices. In addition, residuals from individual stations and/or regions can also be analyzed. Whether residuals are calculated from the best estimate or conditional realizations, the χ_r^2 statistic (with the appropriate number of degrees of freedom) should be equal to 1 in the case of a sufficiently large number of observations.

[21] Once a conditional realization $\mathbf{s}_{ci}$ has been generated, the χ_r^2 statistic can be calculated both for the observations and the prior flux estimates,

$$\chi_{r,\mathbf{z}}^2 = \frac{1}{n}(\mathbf{z} - \mathbf{H}\mathbf{s}_{ci})^T\mathbf{R}^{-1}(\mathbf{z} - \mathbf{H}\mathbf{s}_{ci}) \tag{25}$$

$$\chi_{r,\mathbf{s}}^2 = \frac{1}{m}\left(\mathbf{s}_{ci} - \mathbf{s}_p\right)^T\mathbf{Q}^{-1}\left(\mathbf{s}_{ci} - \mathbf{s}_p\right), \tag{26}$$

where $\chi_{r,\mathbf{z}}^2$ is the statistic for the full set of observations used in the inversion, and $\chi_{r,\mathbf{s}}^2$ is the statistic for the full set of fluxes estimated in the inversion. As mentioned earlier, a $\chi_r^2 = 1$ statistic is not a sufficient condition for identifying the optimal set of covariance parameters, but it is a necessary condition. Therefore, although this statistic is not used in estimating the covariance parameters, inversions using the optimal parameter values should yield residuals with $\chi_r^2 = 1$.

[22] Similarly, a χ_r^2 statistic can be calculated for individual stations and/or flux regions. Because we are classifying observation locations and flux regions into a small number of groups, we do not expect the χ_r^2 for individual stations or regions to be exactly equal to one. This is because this setup yields a single variance for a subset of stations/regions, whereas in reality these stations/regions may exhibit slightly different error characteristics. Instead, if we have grouped the observation locations and flux regions in a manner that yields groups with similar characteristics, we expect the χ_r^2 statistic for individual locations and regions to be distributed around 1. As we allow for more groups (i.e., more estimated covariance parameters) the χ_r^2 statistic of individual members of these groups will cluster more closely around 1. The χ_r^2 statistics for individ-

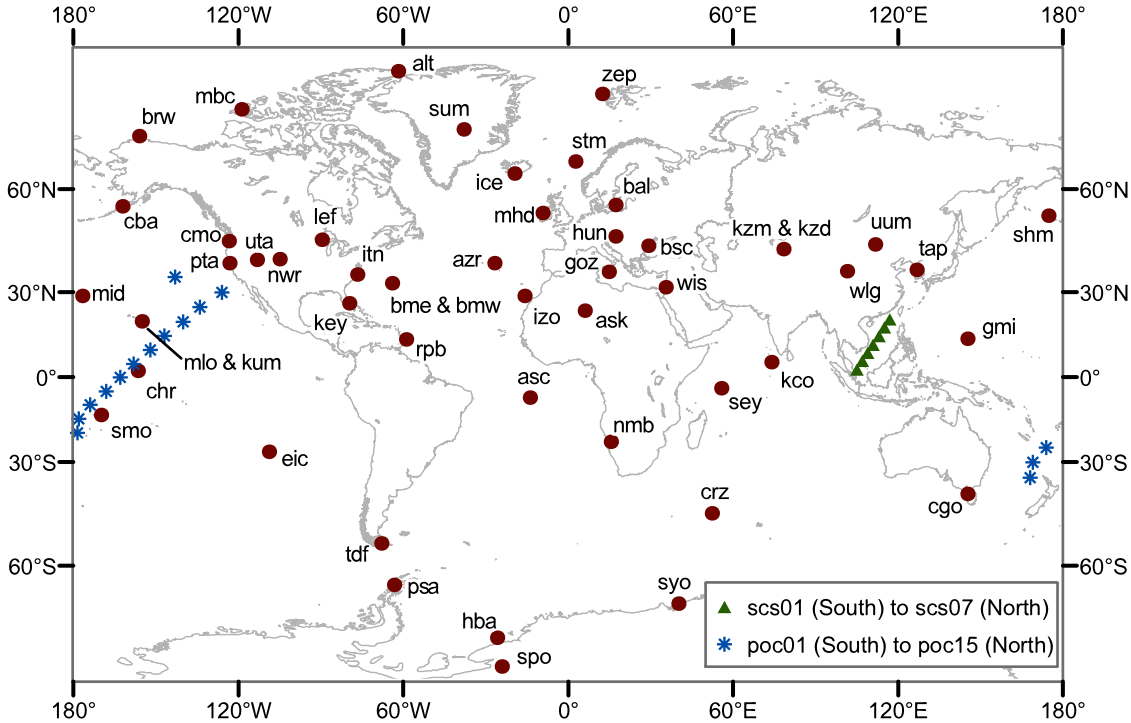

Figure 2. NOAA CMDL cooperative air sampling network sites used in the analysis. Sites indicated in blue and in green represent ship tracks.

ual locations ($\chi^2_{r,\mathbf{z}_j}$) and individual flux regions ($\chi^2_{r,\mathbf{s}_k}$) are defined as

$$\chi^2_{r,\mathbf{z}_j} = \frac{1}{n_j}\left(\mathbf{z}_j - \mathbf{H}_j\mathbf{s}_{ci}\right)^T \mathbf{R}_j^{-1}\left(\mathbf{z}_j - \mathbf{H}_j\mathbf{s}_{ci}\right) \tag{27}$$

$$\chi^2_{r,\mathbf{s}_k} = \frac{1}{m_k}\left(\mathbf{s}_{ci,k} - \mathbf{s}_{p,k}\right)^T \mathbf{Q}_k^{-1}\left(\mathbf{s}_{ci,k} - \mathbf{s}_{p,k}\right), \tag{28}$$

where $\mathbf{z}_j$ are the observations taken at a given location j, $\mathbf{H}_j$ is the transport matrix relating these observations to all fluxes, $\mathbf{R}_j$ is the portion of the model-data mismatch covariance matrix relating to observations $\mathbf{z}_j$, $\mathbf{s}_k$ are the flux components for a given region k, $\mathbf{s}_{ci,k}$ and $\mathbf{s}_{p,k}$ are the flux conditional realization values and the prior flux estimates for the same region, and $\mathbf{Q}_k$ is the portion of the prior error covariance matrix relating to fluxes $\mathbf{s}_k$.

3. Application to CO_2 Data

[23] The method is illustrated using monthly averaged carbon dioxide data for 1996 through 2000, from 75 Climate Monitoring and Diagnostics Laboratory (CMDL) cooperative air sampling network sites (Figure 2). These data represent a total of 2698 monthly averaged observations. The earth is broken up into 22 flux regions (Figure 3), according to the divisions used in the TransCom3 experiment, an atmospheric transport inversion intercomparison study [*Gurney et al.*, 2002, 2003, 2004]. The transport matrix **H** was obtained by running month-long unit pulses originating from each region and in each month of the study period, and recording the monthly averaged response at each sampling time and location. The spatial flux patterns within regions were based on work by *Randerson et al.* [1997] for net ecosystem production (NEP), and *Takahashi et al.* [2002] for net oceanic carbon exchange. The influence of fossil fuel emissions, based on work by *Brenkert* [1998] and *Andres et al.* [1996], was presubtracted from the measurements. This setup is similar to that used in TransCom3 [*Gurney et al.*, 2002]. The transport model used was TM3 [*Heimann*, 1996] run at 7.5° latitude by 10.0° longitude resolution, with nine vertical levels spanning the surface to 10 hPa and a 3-hour integration time step. The model was implemented with National Center for Environmental Prediction (NCEP) windfields corresponding to the modeled years. It is important to note that the obtained covariance parameters are a function of the particular setup, and are not necessarily representative of optimal values to be used in significantly different inversion setups. We recommend that researchers implement the ML method with their particular data and setup in order to obtain optimal values to be used in their own work.

3.1. Setup and Results

[24] In the presented applications, observation locations and flux regions are grouped according to various characteristics commonly thought to be indicative of their relative uncertainty. This is done partly to limit the number of parameters to be estimated, but, more importantly, to demonstrate that the method is able to identify differences between various observation locations and regions that we expect (from our conceptual understanding of the problem) to behave differently. The number of covariance parameters estimated in the presented examples ranges from 2 to 7, but a larger number of parameters could be estimated if deemed necessary in future studies.

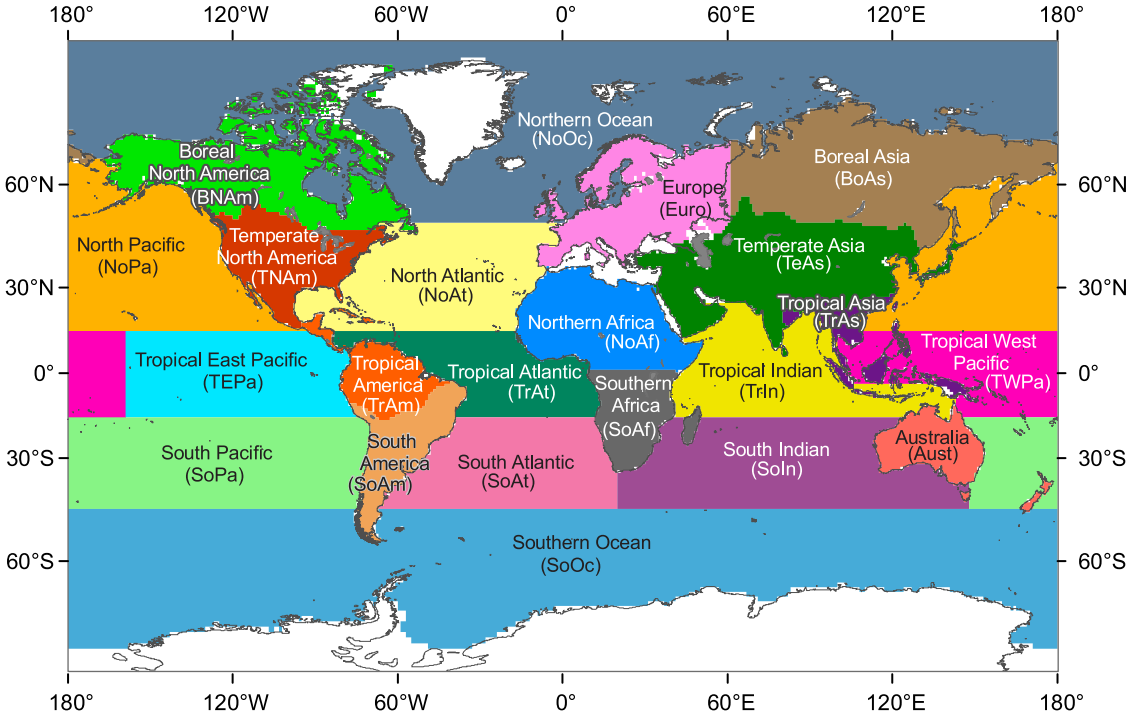

Figure 3. Definition of 22 Transcom3 regions on a 1° latitude by 1° longitude scale.

[25] Four different cases were considered. These cases are designed to (1) demonstrate the applicability of the derived method given different assumptions regarding the uncertainty associated with measurements and flux regions, (2) verify that the method can be used to distinguish the covariance characteristics of different groups of observations and/or flux regions, (3) verify that the fluxes resulting from inversions incorporating ML estimates of covariance parameters yield residuals with the specified covariance structure, and (4) demonstrate the versatility of the proposed method.

[26] In the first case, the uncertainty associated with all measurements is assumed to be the same for all measurement locations, and the uncertainty of the prior flux estimates is assumed not to vary with flux regions. In the second case, observations are grouped into marine boundary layer sites, high elevation or desert sites, and all other remaining (i.e., continental) sites, because such a distinction is commonly thought to strongly contribute to the degree to which measurements can be reproduced. Measurements from marine boundary layer sites can typically be reproduced most closely, and high elevation and desert sites may also exhibit lower model-data mismatch relative to other continental sites. Regions in this and subsequent cases are separated into land and ocean regions, because prior fluxes for ocean regions are considered to be less uncertain than those in land regions, and separating the regions in this way allows the ML algorithm to potentially identify this feature. In the third case, observation locations are separated into five groups. Contrary to case 2, however, the distinction is not based on specific physical characteristics of the locations. Instead the observation locations are separated in an ad-hoc fashion based on the observed model-data mismatch at those locations. In the fourth case, finally, the residual standard deviation from a smooth curve of measurements taken at individual locations is assumed to be indicative of the expected model-data mismatch. Therefore, instead of separating observation locations into a set number of groups with a constant model-data mismatch variance for each group, the relative variance for each observation location is assumed to be proportional to the residual variance from a smooth curve. The ML method is then used to solve for a single scaling factor which determines the model-data mismatch variance at each site.

[27] All cases presented here involve diagonal model-data mismatch (**R**) and prior error (**Q**) covariance matrices. Estimated covariance parameters are summarized in Table 1. For all the presented cases, implementing an inversion with the optimized parameters resulted in an overall $\chi^2_{r,\mathbf{z}}$ and $\chi^2_{r,\mathbf{s}}$ of 1.0. This indicates that the ML method yields covariance parameters that lead to fluxes and modeled observations with overall variances that are exactly consistent with the specified model-data mismatch and prior error covariance parameters. As already outlined, this is not the criterion used by the ML method to select parameters, but simply a side-effect of identifying the correct parameters.

3.1.1. Case 1 (Figure 4)

[28] The first case considered a single variance for the model-data mismatch for all stations, and a second single variance for the prior error for all regions. The model-data mismatch standard deviation was determined to be $\sigma_R = 1.63 \pm 0.03$ ppm, where the intervals represent 1 standard deviation. This error range is consistent with typical values used in inversions. The overall prior error standard deviation was $\sigma_Q = 2.17 \pm 0.08$ GtC/yr.

[29] The $\chi^2_{r,\mathbf{z}_j}$ statistic, broken down by individual stations is presented in Figure 4 (top) and the $\chi^2_{r,\mathbf{s}_k}$ statistic broken

Table 1. Estimated Covariance Parameters for Examined Cases[a]

	Model-Data Mismatch by Station		Prior Error by Region	
Case	Station	Mismatch	Region	Prior Error
1	all stations	σ_R = 1.63 ± 0.03 ppm	All regions	σ_Q = 2.17 ± 0.08 GtC/yr
2	marine boundary layer	σ_R = 0.71 ± 0.02 ppm	ocean	σ_Q = 1.07 ± 0.07 GtC/yr
	high elevation and desert	σ_R = 1.49 ± 0.06 ppm	land	σ_Q = 2.02 ± 0.12 GtC/yr
	other	σ_R = 3.16 ± 0.10 ppm		
3	subset 1	σ_R = 0.58 ± 0.01 ppm	ocean	σ_Q = 1.21 ± 0.06 GtC/yr
	subset 2	σ_R = 1.04 ± 0.04 ppm	land	σ_Q = 1.76 ± 0.10 GtC/yr
	subset 3	σ_R = 1.91 ± 0.07 ppm		
	subset 4	σ_R = 4.36 ± 0.23 ppm		
	subset 5	σ_R = 7.64 ± 0.72 ppm		
4	all stations, weighted	σ_R = 0.11 ± 0.002 ppm	ocean	σ_Q = 0.89 ± 0.08 GtC/yr
		to 5.96 ± 0.13 ppm	land	σ_Q = 2.08 ± 0.15 GtC/yr

[a]Intervals represent 1 standard deviation.

down by individual regions is presented in Figure 4 (bottom). As can be seen in the figure, the fact that measurements taken at some stations cannot be reproduced as precisely as those from certain other stations cannot be captured by this simple setup. As a result, a few observation stations, and especially Hungary (HUN), have a large overall impact on the model-data mismatch variance, because they are much more difficult to reproduce relative to other stations. Also, the Europe flux region deviates from its prior flux estimate much more strongly than all other regions. As will be seen in the other cases, however, this effect disappears once certain stations such as HUN are

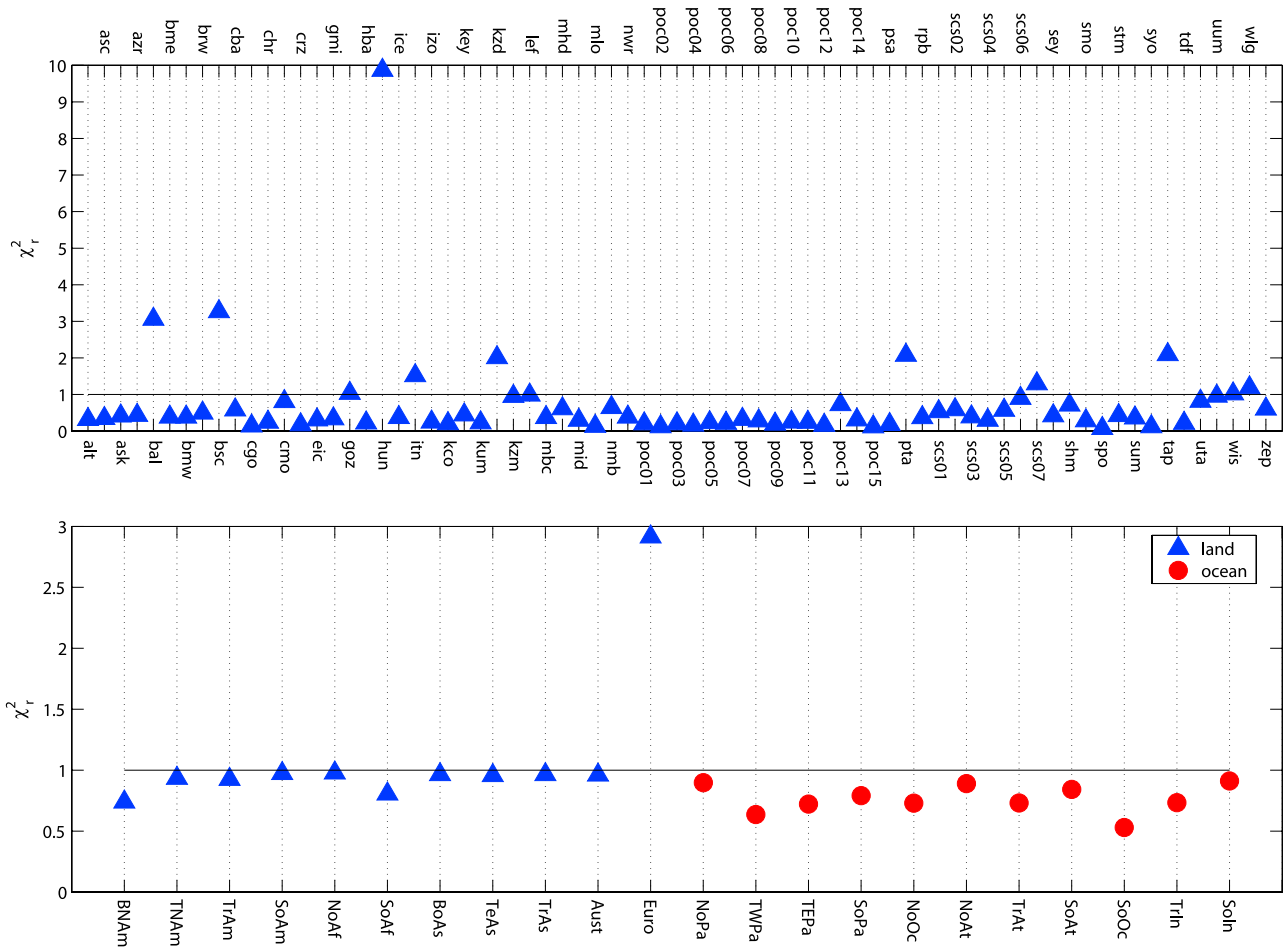

Figure 4. The χ^2_r statistics calculated from conditional realizations of the a posteriori fluxes resulting from an inversion with covariance parameters optimized according to case 1. (top) The $\chi^2_{r,\mathbf{z}_j}$ for individual observation stations. (bottom) The $\chi^2_{r,\mathbf{s}_k}$ for individual flux regions. The solid lines represent χ^2_r = 1.0, which is the mean χ^2_r in both panels (across all stations or regions).

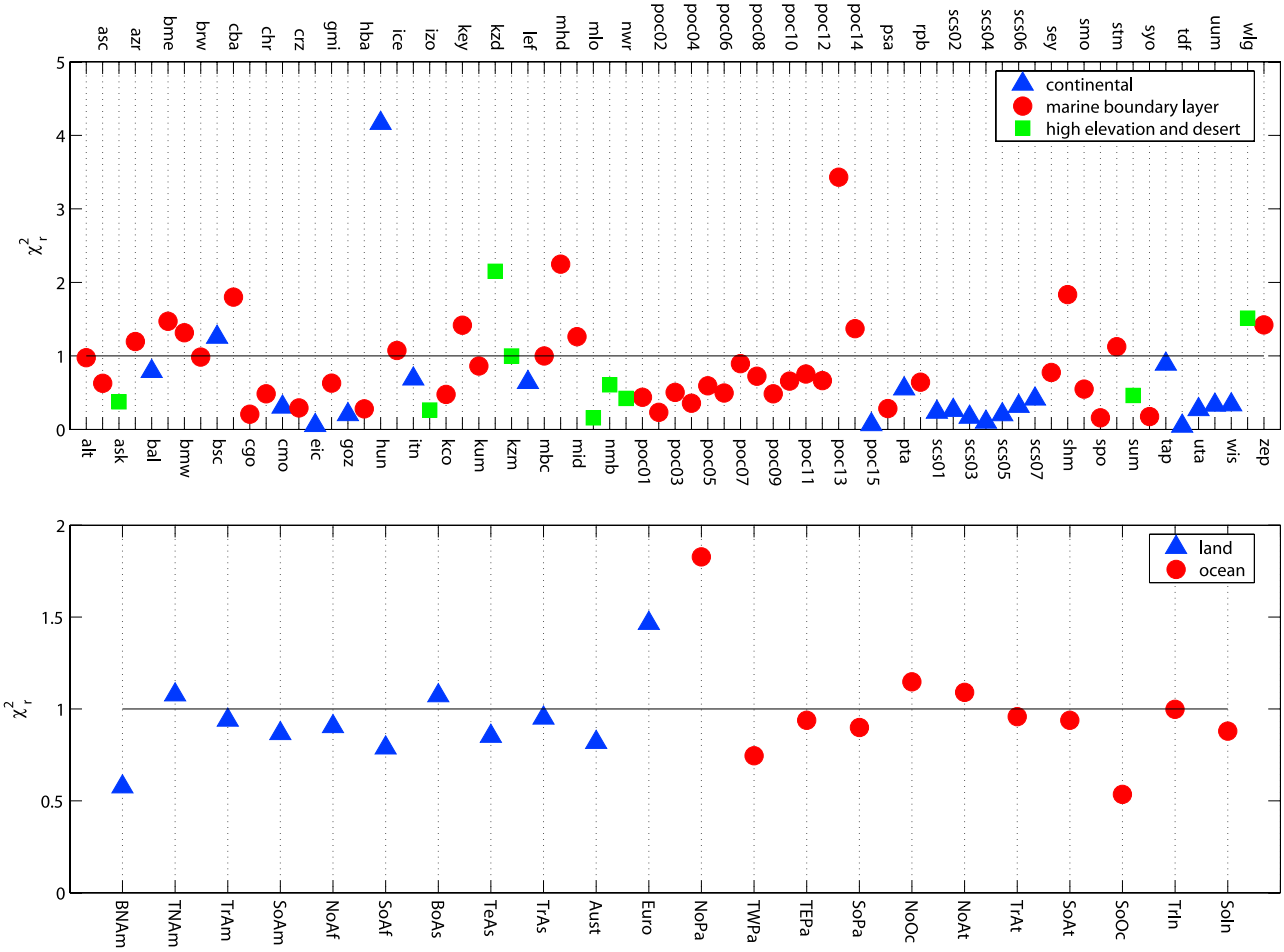

Figure 5. The χ^2_r statistics calculated from conditional realizations of the a posteriori fluxes resulting from an inversion with covariance parameters optimized according to case 2. (top) The $\chi^2_{r,\mathbf{z}_j}$ for individual observation stations. (bottom) The $\chi^2_{r,\mathbf{s}_k}$ for individual flux regions. Color coding designates grouping of observation locations and flux regions according to the definitions presented in case 2. The solid lines represent $\chi^2_r = 1.0$, which is the mean χ^2_r in both panels (across all stations or regions).

allowed to have higher model-data mismatch variances relative to other stations. The covariance parameters estimated using this simplest setup still guarantee that, on average, measurements and prior flux estimates are reproduced to the degree prescribed by the covariance parameter. This is evidenced by the fact that, in this and all subsequent cases, implementing an inversion with the optimized parameters results in an overall $\chi^2_{r,\mathbf{z}}$ and $\chi^2_{r,\mathbf{s}}$ of 1.0.

3.1.2. Case 2 (Figure 5)

[30] The second case broke observation locations into three groups: Marine boundary layer (MBL) sites, high elevation or desert sites, and other sites. The flux regions were separated into land regions and ocean regions. Therefore setup 2 required the estimation of a total of five variances. The $\chi^2_{r,\mathbf{z}_j}$ and $\chi^2_{r,\mathbf{s}_k}$ statistics for individual measurement locations and regions, respectively, are presented in Figure 5.

[31] One interesting thing to notice is that the ML algorithm clearly recognizes that observations at marine boundary layer stations can be reproduced with a higher precision relative to non-MBL sites. In addition, the model-data mismatch for high elevation and desert sites is lower than that for other continental sites. The model-data mismatch standard deviation for MBL sites estimated by the ML routine and then incorporated into the **R** matrix is $\sigma_R = 0.71 \pm 0.02$ ppm, whereas it was $\sigma_R = 1.49 \pm 0.06$ ppm for high elevation and desert sites, and $\sigma_R = 3.16 \pm 0.10$ ppm for other non-MBL sites. This is consistent with both our intuitive understanding of the problem and with past inversion studies that have found that a higher variance needed to be used for continental sites.

[32] Similarly, the ML algorithm assigned a higher prior error to land flux regions relative to ocean regions. The prior error standard deviation was $\sigma_Q = 2.02 \pm 0.12$ GtC/yr for land regions, and $\sigma_Q = 1.07 \pm 0.07$ GtC/yr for ocean regions. Again, past inversion studies have also assigned greater uncertainties to the land priors relative to ocean priors. In this study, however, this conclusion is derived directly from the available data. This confirms that the ML method is able to identify and corroborate the common assumption that land flux prior estimates are more uncertain that ocean flux prior estimates.

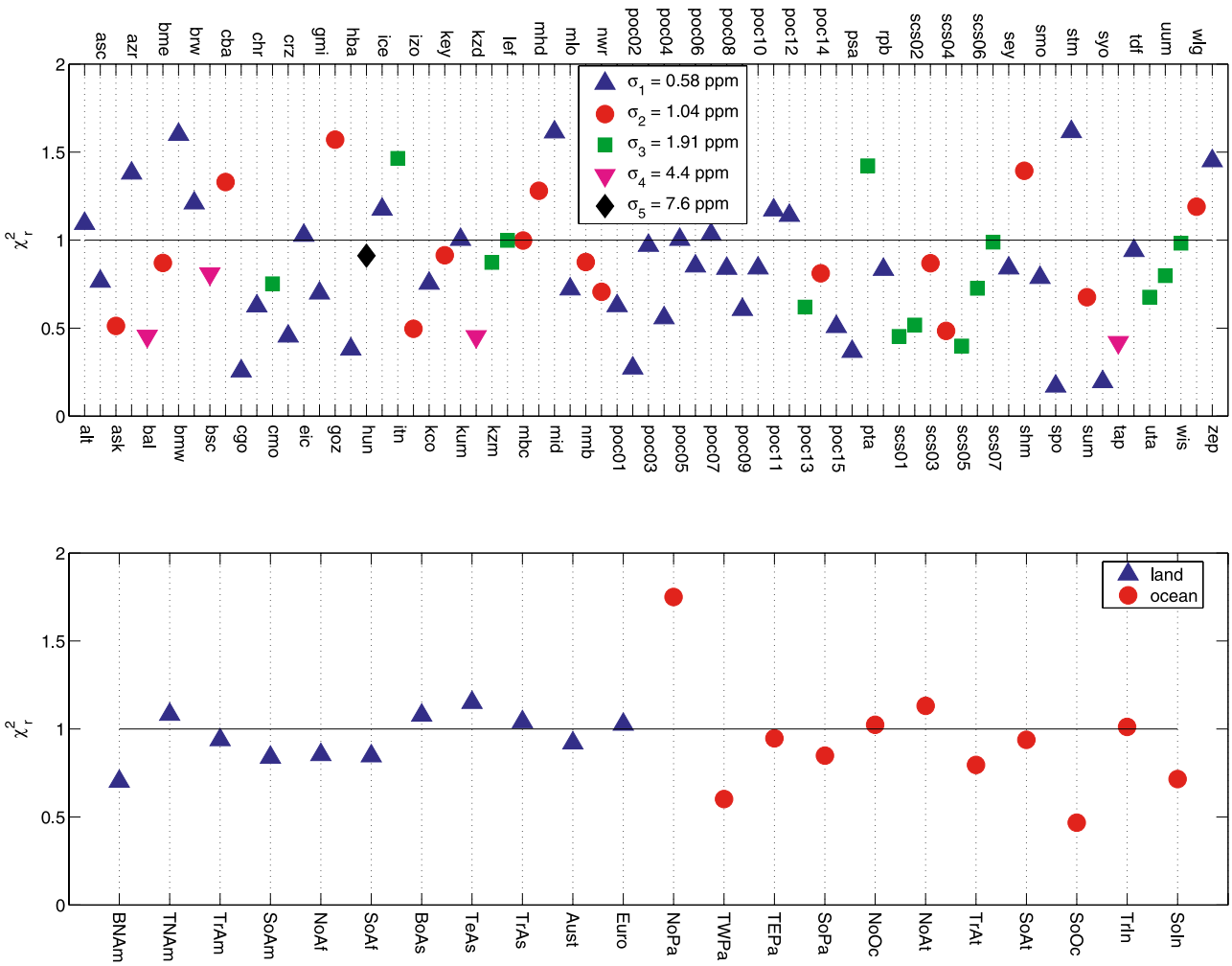

Figure 6. The χ^2_r statistics calculated from conditional realizations of the a posteriori fluxes resulting from an inversion with covariance parameters optimized according to case 3. (top) The $\chi^2_{r,\mathbf{z}_j}$ for individual observation stations. (bottom) The $\chi^2_{r,\mathbf{s}_k}$ for individual flux regions. Color coding designates grouping of observation locations and flux regions according to the definitions presented in case 3. The solid lines represent $\chi^2_r = 1.0$, which is the mean χ^2_r in both panels (across all stations or regions).

3.1.3. Case 3 (Figure 6)

[33] In case 3, observation locations were separated into groups purely based on the model's ability to reproduce the observations. Five different groups were formed. The resulting $\chi^2_{r,\mathbf{z}_j}$ and $\chi^2_{r,\mathbf{s}_k}$ statistics for individual measurement locations and regions, respectively, are presented in Figure 6, with the colors indicating the various groups. The resulting model-data mismatch and prior error standard deviations are presented in Table 1.

[34] This fourth case results in the $\chi^2_{r,\mathbf{z}_j}$ and $\chi^2_{r,\mathbf{s}_k}$ statistics being most closely clustered around 1.0. Part of this is due to the fact that we are estimating the largest number of parameters, getting closer to a case where each stations and each regions would have its own variance, and where these statistics would be expected to be very close to 1.0 for all stations and regions. The second reason for this clustering is that we are using the performance of the model itself to help group observation locations, instead of relying exclusively on our physical understanding of the problem. In such a setup, sites for which we may expect lower errors due to their physical locations may be assigned to groups with higher errors, recognizing potential transport errors due to, for instance, biased wind fields, incorrect assigned regional flux patterns, and/or sampling error at coastal sites. In practice, when applying this method, researchers can choose how they want to group sites, whether they want to rely on physical differences for classification, or allow the numerical results to direct their choices.

3.1.4. Case 4 (Figure 7)

[35] In this last case (Figure 7), instead of breaking the observation locations into groups and estimating a single model-data mismatch variance for each group, each station was assigned a scaling factor based on the average variance between observations taken at that station and a smooth curve fit. These constants were taken from results obtained during the TransCom3 experiment, and were scaled by the proportion of real data in the *GLOBALVIEW-CO_2* [2000] record [*Gurney et al.*, 2003]. The ML routines were then

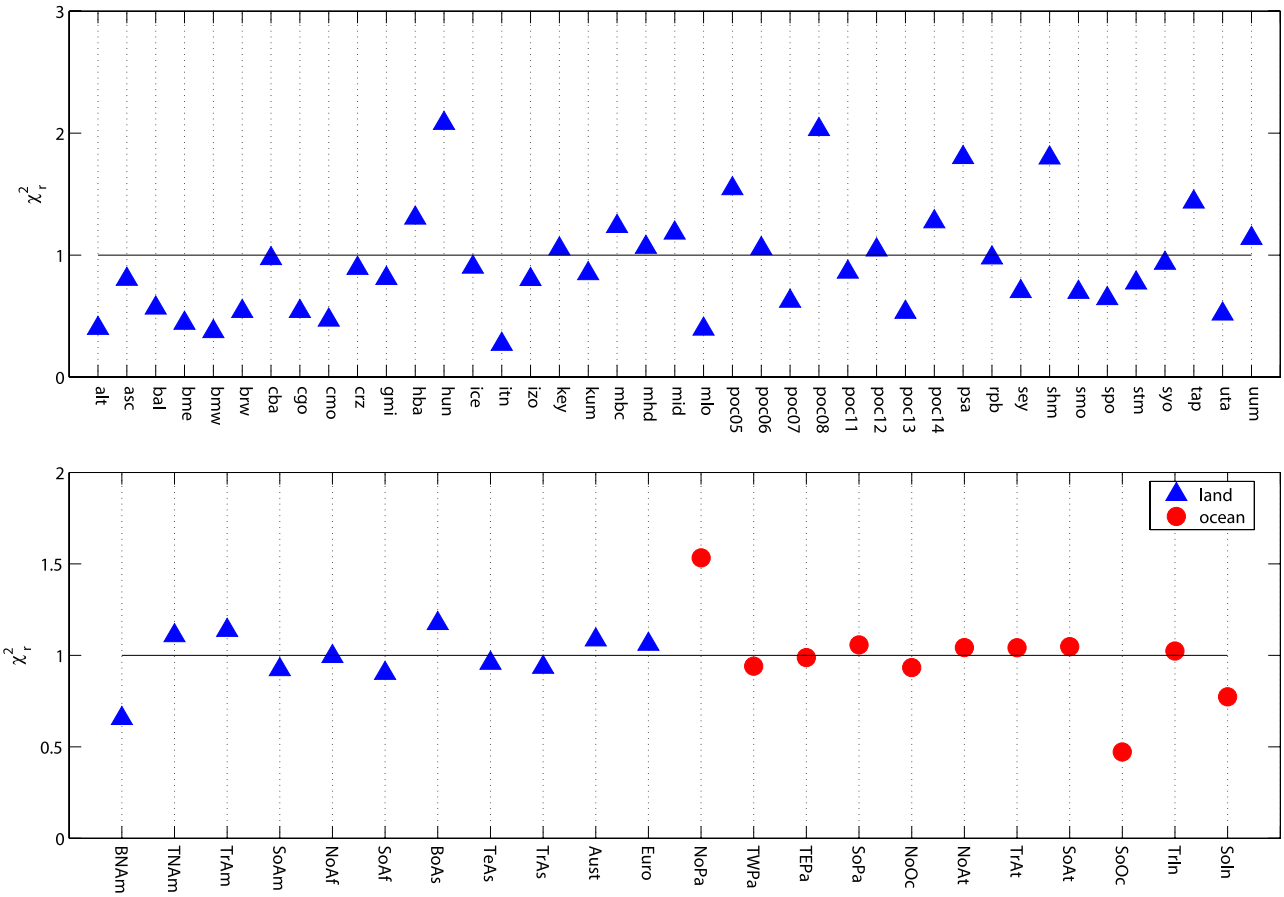

Figure 7. The χ^2_r statistics calculated from conditional realizations of the a posteriori fluxes resulting from an inversion with covariance parameters optimized according to case 4. (top) The $\chi^2_{r,\mathbf{z}_j}$ for individual observation stations. (bottom) The $\chi^2_{r,\mathbf{s}_k}$ for individual flux regions. The solid lines represent $\chi^2_r = 1.0$, which is the mean χ^2_r in both panels (across all stations or regions).

used to derive a single proportionality constant that would scale these variances to obtain an optimal overall model-data mismatch for each site. The resulting model-data mismatch covariance function was

$$\mathbf{R} = \begin{bmatrix} C\sigma^2_{S,1} & 0 & \cdots & 0 \\ 0 & C\sigma^2_{S,2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & C\sigma^2_{S,n} \end{bmatrix}, \tag{29}$$

where $\sigma^2_{S,i}$ are the variances of the deviations of samples taken at individual sites from a smooth curve fit, and C is the single multiplicative factor that we will estimate using the ML method. The $\mathbf{R}$ matrix still has dimensions $n \times n$, where n is the total number of samples considered, but the values on the diagonal correspond to the residual variances observed at the site where each observation was taken. The prior error was broken into land and ocean groups, as in cases 2 and 3. Note that this final case uses only 41 stations, which are the subset of the examined 75 stations used here which were also used in the TransCom3 annual-average inversion study [*Gurney et al.*, 2003]. This also means that there are a total of 41 different variances that populate the matrix in equation (29).

[36] The optimal value of the multiplicative factor was found to be 0.80, with a standard deviation of 0.14, indicating that residual standard deviations should be scaled by a factor of $\sqrt{0.80} = 0.89$ in order to, on average, represent an appropriate model-data mismatch. The $\chi^2_{r,\mathbf{z}_j}$ and $\chi^2_{r,\mathbf{s}_k}$ statistics for individual measurement locations and regions, respectively, are presented in Figure 7. As can be seen from the figure, classification based on the deviation of the observations from a smooth curve is a much better predictor of model-data mismatch relative to using a constant variance for all stations (Figure 4). Better, in this case refers to the amount of scattering of $\chi^2_{r,\mathbf{z}_j}$ and $\chi^2_{r,\mathbf{s}_k}$ around one, because the overall $\chi^2_{r,\mathbf{z}}$ and $\chi^2_{r,\mathbf{s}}$ are 1.0 for both cases. This last case does not behave as well as subdividing the stations into a moderate number of constant-variance groups (Figure 6), however. For a few stations (HUN, POC13, PSA, SHM in Figure 2) the model-data mismatch appears to be significantly higher relative to the variance of deviations from a smooth curve as compared to other stations. This is not to say that the variance of the deviation from a smooth curve could not be used as part of the determinant of the relative variance expected at a site, but it does indicate that this criterion may not be effective if used as the only guiding principle.

[37] Whereas the model-data mismatch standard deviations used in the TransCom3 annual inversion study [*Gurney*

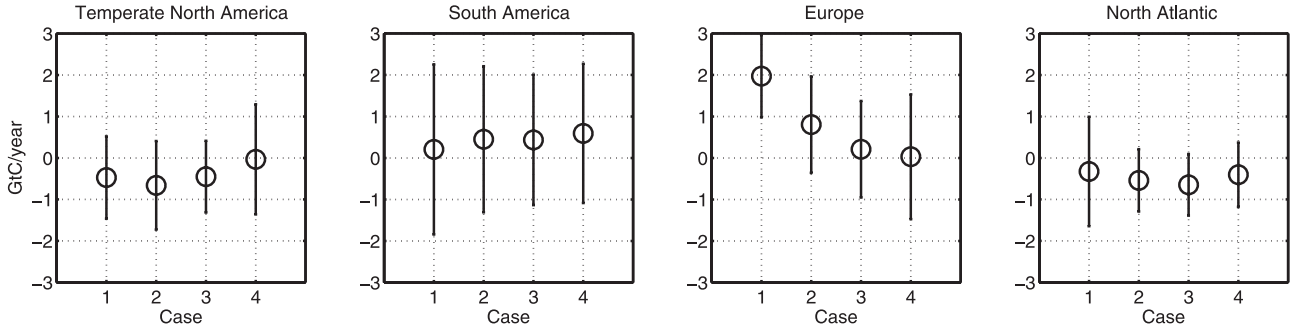

Figure 8. Examples of the effect of the four different covariance parameter cases on estimated fluxes and their uncertainties. The circles represent the estimated average flux for 2000 for four Transcom regions for the four examined cases. The bars represent the estimated standard deviations.

et al., 2003] ranged from 0.04 ppm (SYO) to 2.23 ppm (ITN), depending on the residual standard deviation of samples taken from individual sites from a smooth curve fit, the current analysis suggests that, if all stations are considered to have variances proportional to their residual standard deviations, the model-data mismatch standard deviations range from 0.11 ppm (SYO) to 5.96 ppm (ITN). The stations with the maximum/minimum model-data mismatch are consistent for these two studies because all model-data mismatch standard deviations are determined by multiplying the residual standard deviations of samples taken at a given site by a constant factor. Overall, the model-data mismatch variances inferred using the ML method are higher relative to those used in the TransCom3 annual inversions. This is likely due, at least in part, to the fact that the Transcom study used smoothed Globalview data [*GLOBALVIEW-CO_2*, 2000] in their analysis, whereas the current estimates were obtained using CO_2 flask data directly. In addition, whereas the Transcom study was focusing on annually averaged fluxes, we have instead estimated monthly fluxes. Additional differences such as the prior flux uncertainties, the selected transport models, and the subsets of sampling sites used would also contribute to this difference. Note also that, in the Transcom study, all model-data mismatch standard deviations were ultimately modified to be at least 0.25 ppm, whereas no such adjustments were made here.

3.2. Effect on Flux Estimates

[38] The purpose of this work is to present a method for estimating covariance matrix parameters used in Bayesian inversions, and presenting detailed flux estimates for the examined period falls outside the scope of this work. However, given that each of the sets of covariance parameters is optimal for its respective setup, it is interesting to examine the effect of the choice of covariance structure on estimated fluxes.

[39] The choice of covariance parameter setup has an effect on both the flux magnitudes and their estimated uncertainties. The effect on the posterior covariance of the fluxes is relatively straightforward. From equation (3), it is apparent that once a transport model has been selected (thereby fixing **H**), the covariance matrices fully determine the posterior covariance. What is more difficult to ascertain is the effect on the best estimates of the fluxes. Some of the key differences are illustrated in Figure 8. This figure represents the total non-fossil-fuel flux and uncertainty for the year 2000 for four of the 22 Transcom regions. These estimates were obtained by summing the monthly estimated fluxes for 2000. The uncertainties were obtained by summing the appropriate elements of the posterior covariance matrix, including off-diagonal temporal correlation terms. The results for Temperate North America are typical of relatively well-constrained land regions. The different cases yield different total fluxes, but these fluxes generally fall within the other cases' confidence bounds. The results for South America are typical of poorly constrained land regions. The different cases have little effect on either the fluxes or their uncertainties because the measurements contain little information regarding these regions. The fluxes for Europe show quite dramatic variation, most likely due to how these four cases treat measurement from the Hungary (HUN) observation site. Data from this site are very difficult to reproduce as is also evidenced in Figures 4 to 7. The fluxes for the North Atlantic are typical of results for oceanic regions. The strongest difference is between case 1, where land and ocean prior fluxes are represented by a single error variance, and the other cases, where prior estimates of fluxes over land are assigned a different error variance than the oceanic prior fluxes. In case 1, the prior error on ocean regions is significantly higher relative to the other cases, yielding a higher posterior uncertainty for these regions.

[40] The main conclusions for this analysis are that the choice of covariance matrix setup, and not only the choice of covariance parameter values, can have a strong effect on estimated fluxes. Therefore expert knowledge must be used to guide the selection of an inversion setup.

3.3. Discussion

[41] In all cases, the ML approach is able to estimate covariance parameters that result in residuals that have the variability specified in the covariance matrices (i.e., $\chi_r^2 = 1$). This had been the primary goal of previous attempts to estimate covariance parameters. In addition, the method has been demonstrated as a tool for estimating the overall and relative variance associated with different stations and regions. In all cases, the ML method results in covariance parameters that are qualitatively consistent with our phys-

ical understanding of the relative variance of observations from various stations and fluxes from various regions. The method goes beyond this qualitative assessment, however, and quantifies the differing uncertainties associated with individual stations and/or regions. In addition, the method was demonstrated as being an effective guiding tool for grouping observation locations into groups that behave similarly. The method can also incorporate additional information, such as standard deviations of observations from smooth curve fits, into the estimation of covariance parameters. Last, the method provides uncertainty bounds on the covariance parameters, which could allow for sensitivity runs exploring the effect of covariance parameter uncertainty on the estimated flux distributions.

[42] As expected and as can be seen from Table 1, increasing the number of possible variance parameters allows for individual groups of stations or regions to have different behavior. When the grouping corresponds to real differences in the model-data mismatch, the $\chi^2_{r,\mathbf{z}_j}$ statistics of individual stations are clustered more closely around 1.0, as can be seen then comparing Figures 4 to 7. It appears especially important to assign a different variance to stations that are particularly difficult to match, such as Hungary (HUN). Otherwise, as can be seen in Figure 4, these few stations have too great an influence on the overall covariance parameters, and as a result most stations have $\chi^2_{r,\mathbf{z}_j}$ significantly below 1.

[43] Case 3 also demonstrated that one does not necessarily need to separate stations based on our conceptual understanding of the difference between various sites, but can in fact let the model guide group selections, and thereby identify sites that behave similarly, in terms of the degree to which their observations can be reproduced given the selected inversion setup.

[44] Case 4 demonstrated the possibility of using the ML algorithm to solve for variance proportionality constants, instead of solving for individual group variances. Future studies could explore indices based on other physical attributes, to determine whether they can predict variability at given sites, and whether they can do this better relative to binning sites into fixed-variance groups, as was done in cases 1–3.

[45] Finally, as would be expected, the larger the number of covariance parameters that are to be estimated, the higher their uncertainty. This is a direct result of the fact that, as the number of observation location groups increases, there are fewer measurements to constrain the covariance parameters of each individual group. Overall, however, the uncertainty of the covariance parameters appears quite small. It is important to note that this low uncertainty is only valid if all the assumptions used in the problem setup are valid. As is currently typical in flux inversions we have assumed that model-data mismatch errors are independent, as are the prior flux errors for large regions. In addition, we have binned the covariance parameters into a relatively small number of representative groups. Therefore the uncertainties on the covariance parameters should be interpreted in the context of the specific setup used here. If, for example, a different model-data mismatch uncertainty were to be calculated for each observation site, we would expect a higher uncertainty on these parameters relative to the uncertainty on the covariance parameter averaged over several sites, as was done here.

4. Conclusions

[46] The method presented in this paper uses a maximum likelihood framework with available observations, prior flux estimates and transport information to optimize the covariance parameters needed in the solution of Bayesian inverse problems used to estimate surface fluxes of atmospheric trace gases. The method can also be applied to estimate the uncertainty in these parameters. The strong influence and critical importance of these parameters has been discussed in many studies [*Kaminski et al.*, 1999; *Rayner et al.*, 1999; *Law et al.*, 2002; *Peylin et al.*, 2002; *Engelen et al.*, 2002]. Up to this point, however, no objective method was available to estimate these parameters from the available data and guarantee that the resulting flux and observation residuals would follow the assumed distribution.

[47] In addition to optimizing covariance parameters for a given inversion setup, the method can also be used to evaluate whether grouping certain measurement stations or flux regions is justified given the available information. For example, if stations are grouped according to whether or not they constitute marine boundary layer sites, and the resulting model-data mismatch variances are different for the two groups, this suggests that whether or not a station samples marine boundary layer air is a predictor of the precision with which the data can be reproduced.

[48] The examples presented in this paper were for diagonal model-data mismatch and prior error covariance matrices. The method can be directly applied, however, to more complex covariance matrices, where correlation lengths or other parameters must be estimated in addition to variances. Such applications have been demonstrated using the restricted maximum likelihood approach used in geostatistical inverse modeling [*Michalak et al.*, 2004].

[49] Note that we do not advocate researchers using the model-data mismatch variances and prior error variances derived in this work directly, because the optimal values will depend not only on the data set used, but also on the transport model, the defined flux regions, flux patterns within regions, etc. Instead, we recommend that researchers implement the maximum likelihood algorithm described in this work in order to gain insight into covariance parameters that are appropriate for their specific applications. The examples presented in this work simply demonstrate the applicability of the presented method to typical inversions.

[50] Most importantly, in all the presented cases, the ML method identified the most likely covariance parameter values, given the selected inversion setup and covariance matrix definitions. These variances are optimized solely using the available data, transport information, and prior flux estimates.

Appendix A: Generation of Conditional Realizations

[51] If the model-data mismatch and prior-flux error covariance parameters are selected appropriately, residuals calculated from conditional realizations of the a posteriori fluxes will follow the distributions specified in **Q** and **R**.

The method presented here generates equally likely conditional realizations of the function **s**. The first step is to generate an unconditional realization $\mathbf{s}_{ui}$ of the unknown function **s** from the prior covariance **Q**. Here the subscript u indicates that the realization has not been conditioned on the data, and the subscript i serves as a counter and a reminder that there is an infinite number of possible realizations. There are several methods for generating $\mathbf{s}_{ui}$, and the method presented here is based on eigenvalue decomposition. In the general case,

$$\mathbf{s}_{ui} = \mathbf{V}\lambda^{1/2}\boldsymbol{\varepsilon}_{\mathbf{s}} \tag{A1}$$

where **V** is an $(m \times m)$ matrix containing the eigenvectors of **Q**, $\lambda^{1/2}$ is an $(m \times m)$ diagonal matrix with the square root of the eigenvalues of **Q** on the diagonal, and $\boldsymbol{\varepsilon}_{\mathbf{s}}$ is an $(m \times 1)$ vector of normally distributed random numbers with mean zero and variance one. For the case where **Q** is a diagonal matrix, as is often the case in classical Bayesian inverse modeling, the above equation simplifies to

$$\mathbf{s}_{ui} = \mathbf{Q}^{1/2}\boldsymbol{\varepsilon}_{\mathbf{s}}, \tag{A2}$$

where $\mathbf{Q}^{1/2}$ is a diagonal matrix with the square root of the prior error variances on the diagonal.

[52] The conditional realization is then obtained through

$$\begin{aligned}\mathbf{s}_{ci} &= \mathbf{s}_p + \mathbf{s}_{ui} + \mathbf{QH}^T\left(\mathbf{HQH}^T + \mathbf{R}\right)^{-1}\left(\mathbf{z} + \boldsymbol{\varepsilon}_z - \mathbf{Hs}_{ui} - \mathbf{Hs}_p\right)\\ &= \hat{\mathbf{s}} + \mathbf{s}_{ui} + \mathbf{QH}^T\left(\mathbf{HQH}^T + \mathbf{R}\right)^{-1}(\boldsymbol{\varepsilon}_z - \mathbf{Hs}_{ui}),\end{aligned} \tag{A3}$$

where the subscript c refers to the fact that the realization has been conditioned on the data **z**, and $\boldsymbol{\varepsilon}_z$ is an $(n \times 1)$ vector sampled from the model-data mismatch covariance matrix **R**, using the same method as was described in equation (A1). The validity of this algorithm can easily be demonstrated numerically by generating a large number of conditional realizations and verifying that their statistics (e.g., median and 95% confidence intervals) are identical to those implied by $\hat{\mathbf{s}}$ and $\mathbf{V}_{\hat{\mathbf{s}}}$.

[53] In addition, substituting the definition of $\hat{\mathbf{s}}$ (equation (2)) and using the definitions

$$\mathbf{z} = \mathbf{Hs} + \boldsymbol{\varepsilon}_z \tag{A4}$$

$$\mathbf{s}_p = \mathbf{s} + \boldsymbol{\varepsilon}_{\mathbf{s}}, \tag{A5}$$

the expected covariance of residuals from conditional realizations can be shown to be

$$cov\left(\begin{bmatrix}\mathbf{s}_{ci} - \mathbf{s}_p\\ \mathbf{Hs}_{ci} - \mathbf{z}\end{bmatrix}\right) = \begin{bmatrix}\mathbf{Q} & 0\\ 0 & \mathbf{R}\end{bmatrix}, \tag{A6}$$

verifying that, given appropriate covariance parameters, we expect these residuals to be sampled from **Q** and **R**.

[54] **Acknowledgments.** Funding for Anna Michalak was partially provided by a NOAA Climate and Global Change postdoctoral fellowship, a program administered by the University Corporation for Atmospheric Research (UCAR). The authors would like to thank all of the collaborators in the NOAA/CMDL Cooperative Air Sampling Network, John B. Miller of NOAA-CMDL for helpful insights during the preparation of this manuscript, and Kimberley L. Mueller for help with figure preparation. This paper was also greatly improved as a result of suggestions offered by Ian G. Enting, Peter Rayner, and Christian Rödenbeck.

References

Andres, R. J., G. Marland, I. Fung, and E. Matthews (1996), Distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950–1990, *Global Biogeochem. Cycles*, *10*(3), 419–430.

Bousquet, P., P. Ciais, P. Peylin, M. Ramonet, and P. Monfray (1999), Inverse modeling of annual atmospheric CO_2 sources and sinks: 1. Method and control inversion, *J. Geophys. Res.*, *104*(D21), 26,161–26,178.

Brenkert, A. L. (1998), Carbon dioxide emission estimates from fossil-fuel burning, hydraulic cement production, and gas flaring for 1995 on a one degree grid cell basis, http://cdiac.esd.ornl.gov/ndps/ndp058a.html, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., Oak Ridge, Tenn.

Committee on the Science of Climate Change, Division on Earth and Life Studies, National Research Council (2001), *Climate Change Science: An Analysis of Some Key Questions*, Natl. Acad. Press, Washington, D. C.

Engelen, R. J., A. S. Denning, and K. R. Gurney (2002), On error estimation in atmospheric CO_2 inversions, *J. Geophys. Res.*, *107*(D22), 4635, doi:10.1029/2002JD002195.

Enting, I. G. (2002), *Inverse Problems in Atmospheric Constituent Transport*, Cambridge Univ. Press, New York.

Enting, I. G., C. M. Trudinger, and R. J. Francey (1995), A synthesis inversion of the concentration and $\delta^{13}C$ of atmospheric CO_2, *Tellus, Ser. B*, *47*, 35–52.

Gill, P. E., W. Murray, and M. H. Wright (1986), *Practical Optimization*, Elsevier, New York.

GLOBALVIEW-CO_2 (2000), *Cooperative Atmospheric Data Integration Project—Carbon Dioxide* [CD-ROM], NOAA Clim. Model. and Diag. Lab., Boulder, Colo.

Gurney, K. R., et al. (2002), Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature*, *415*, 626–630, doi:10.1038/415626a.

Gurney, K. R., et al. (2003), TransCom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, *Tellus, Ser. B*, *55*, 555–579.

Gurney, K. R., et al. (2004), Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks, *Global Biogeochem. Cycles*, *18*, GB1010, doi:10.1029/2003GB002111.

Heimann, M. (1996), The Global Atmospheric Tracer Model TM2, *Tech. Rep. 10*, 52 pp., Dtsch. Klimarechenzent., Hamburg, Germany.

Hein, R., P. J. Crutzen, and M. Heimann (1997), An inverse modeling approach to investigate the global atmospheric methane cycle, *Global Biogeochem. Cycles*, *11*(1), 43–76.

Intergovernmental Panel on Climate Change (2001), *Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, edited by J. T. Houghton et al., Cambridge Univ. Press, New York.

Kaminski, T., M. Heimann, and R. Giering (1999), A coarse grid three-dimensional global inverse model of the atmospheric transport: 1. Adjoint model and Jacobian matrix, *J. Geophys. Res.*, *104*(D15), 18,535–18,554.

Kaminski, T., P. J. Rayner, M. Heimann, and I. G. Enting (2001), On aggregation errors in atmospheric transport inversions, *J. Geophys. Res.*, *106*(D5), 4703–4716.

Kandlikar, M. (1997), Bayesian inversion for reconciling uncertainties in global mass balances, *Tellus, Ser. B*, *59*, 123–135.

Kitanidis, P. K. (1986), Parameter uncertainty in estimation of spatial functions: Bayesian analysis, *Water Resour. Res.*, *22*(4), 499–507.

Kitanidis, P. K. (1995), Quasi-linear geostatistical theory for inversing, *Water Resour. Res.*, *31*(10), 2411–2419.

Kitanidis, P. K., and R. W. Lane (1985), Maximum likelihood parameter estimation of hydrologic spatial processes by the Gauss-Newton method, *J. Hydrol.*, *79*, 53–71.

Krakauer, N. Y., T. Schneider, J. T. Randerson, and S. C. Olsen (2004), Using generalized cross-validation to select parameters in inversions for regional carbon fluxes, *Geophys. Res. Lett.*, *31*, L19108, doi:10.1029/2004GL020323.

Law, R. M., R. J. Rayner, L. P. Steele, and I. G. Enting (2002), Using high temporal frequency data for CO_2 inversions, *Global Biogeochem. Cycles*, *16*(4), 1053, doi:10.1029/2001GB001593.

Michalak, A. M., and P. K. Kitanidis (2004), Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling, *Water Resour. Res.*, *40*, W08302, doi:10.1029/2004WR003214.

Michalak, A. M., L. Bruhwiler, and P. P. Tans (2004), A geostatistical approach to surface flux estimation of atmospheric trace gases, *J. Geophys. Res.*, *109*, D14109, doi:10.1029/2003JD004422.

Peylin, P., D. Baker, J. Sarmiento, P. Ciais, and P. Bousquet (2002), Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric CO_2 data, *J. Geophys. Res.*, *107*(D19), 4385, doi:10.1029/2001JD000857.

Randerson, J., M. V. Thompson, T. J. Conway, I. Y. Fung, and C. B. Field (1997), The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, *Global Biogeochem. Cycles*, *11*(4), 535–560.

Rayner, P. J., I. G. Enting, R. J. Francey, and R. Langenfelds (1999), Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}C$ and O_2/N_2 observations, *Tellus, Ser. B*, *51*, 213–232.

Rödenbeck, C., S. Houweling, M. Gloor, and M. Heimann (2003), CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmos. Chem. Phys.*, *3*, 1919–1964.

Schweppe, F. C. (1973), *Uncertain Dynamic Systems*, Prentice-Hall, Upper Saddle River, N. J.

Takahashi, T., et al. (2002), Global sea-air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects, *Deep Sea Res., Part II*, *49*, 1601–1622.

Tans, P. P., K. W. Thoning, W. P. Elliott, and T. J. Conway (1990), Error estimates of background atmospheric CO_2 patterns from weekly flask samples, *J. Geophys. Res.*, *95*(D9), 14,063–14,070.

Tarantola, A. (1987), *Inverse Problem Theory Methods for Data Fitting and Model Parameter Estimation*, Elsevier, New York.

Wofsy, S. C., and R. C. Harriss (2002), The North American Carbon Program (NACP): A report of the NACP Committee of the U.S. Carbon Cycle Science Steering Group, 56 pp., U.S. Global Change Res. Program, Washington, D. C.

L. Bruhwiler, A. Hirsch, W. Peters, and P. P. Tans, NOAA-CMDL, Mailcode R/CMDL1, 325 Broadway, Boulder, CO 80305-3328, USA. (lori.bruhwiler@noaa.gov; adam.hirsch@noaa.gov; wouter.peters@noaa.gov; pieter.tans@noaa.gov)

K. R. Gurney, Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523-1371, USA. (keving@atmos.colostate.edu)

A. M. Michalak, Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI, USA.