https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions

Functional ANOVA for Carbon Flux Estimates from Remote Sensing Data

Jonathan Hobbs[1], Matthias Katzfuss[2], Hai Nguyen[1], Vineet Yadav[1], and Junjie Liu[1]

[1]Jet Propulsion Laboratory, California Institute of Technology
[2]Texas A&M University

Correspondence: Jonathan Hobbs (Jonathan.M.Hobbs@jpl.nasa.gov)

Abstract. The constellation of Earth-observing satellites now produces atmospheric greenhouse gas concentration estimates across multiple years. Their global coverage is providing additional information on the global carbon cycle. These products are combined with complex inversion systems to infer the magnitude of carbon sources and sinks around the globe. Multiple factors, including the atmospheric transport model and satellite product aggregation method, can impact flux estimates. Functional analysis of variance (ANOVA) invokes a spatio-temporal statistical model to efficiently estimate common flux signals across multiple inversions, and partitions variability across the discrete factors considered. The approach is illustrated on inversion experiments with different satellite retrieval aggregation methods and identifies significant flux anomalies in the presence of mode differences across aggregation methods. Functional ANOVA is also applied to a recent flux model intercomparison project (MIP), and the relative magnitudes of transport model effects and data source (satellite versus in situ) are similar but exhibit slightly different importance for inversions over different continents.

1 Introduction

Many of the key processes in the global carbon cycle have undergone substantial change in recent decades, yet their impacts remain challenging to estimate. This is due in large part to the sparsity of direct observations of carbon fluxes. In particular, a lack of global coverage requires alternative approaches for understanding the global carbon cycle. Fluxes can be inferred indirectly with atmospheric transport models in combination with information on atmospheric carbon dioxide concentration. Frequent global CO_2 estimates from satellites Greenhouse Gases Observing Satellite (GOSAT; Kuze et al., 2009) and the Orbiting Carbon Observatory-2 (OCO-2; Eldering et al., 2017) have enhanced the tractability and precision of flux inversion. Since the satellite estimates arise from a retrieval (O'Dell et al., 2018), the end-to-end inference from satellite radiance spectra to flux estimates involves two complex inverse problems subject to multiple sources of uncertainty (Cressie, 2018), including observational errors, spatio-temporal representation uncertainty, and model transport error (Engelen et al., 2002). Some flux solutions attempt to account for these sources in their representation of the posterior uncertainty, but these are not always

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022


Geoscientific Model Development Discussions
Open Access
EGU

available and a coherent probabilistic assessment becomes challenging in the presence of multiple flux estimates with varying assumptions. Further, the spatio-temporal structure of flux estimates is of particular interest, and dependence can be exploited in quantifying uncertainty. The statistical methodology in this work provides a framework for this common situation.

Different flux estimates can arise from combinations of multiple categorical factors, and partitioning their relative contributions to the range of solutions can guide priorities for future research in carbon cycle science. In this work we are particularly interested in flux estimates derived from different inversion systems, such as those investigated in model intercomparison projects (MIPs; Thompson et al., 2016; Gaubert et al., 2019; Crowell et al., 2019). A second factor of interest is the makeup of the CO_2 concentration data used in the inversions. Our effort contrasts inversions that use Level 2 satellite retrievals directly versus inversions that use Level 3 products produced through data fusion (Nguyen et al., 2017).

Given a set of flux maps obtained under different scenarios, or combinations of these factors of interest, our goal is to find common features among the scenarios, and to identify systematic ways or regions in which fluxes from different scenarios differ. Analysis of variance (ANOVA) is a statistical modeling framework that facilitates the estimation of the common and factor-specific effects. It further characterizes the magnitude of the differences within factors relative to the inherent variability within a scenario. Model assumptions dictate the estimation of this within-scenario variability and will be an additional focus of our investigation. ANOVA methodology has been extended to functional data, such as time series and spatial fields, where it can provide a coherent depiction of space/time patterns and anomalies due to various factors (Kaufman and Sain, 2010).

The ANOVA approach can be particularly useful for analyzing output from a collection Earth system models or assimilation systems with a common quantity of interest and similar experimental setup. This setup is often formalized as a MIP, an enterprise becoming commonplace among multiple component process models in Earth system modeling (see Eyring et al., 2016, and associated special issue). Several MIPs have been conducted for carbon flux inversion systems, both for in situ observations (Gaubert et al., 2019) and for the growing satellite record (Crowell et al., 2019; Peiro et al., 2022). These recent flux MIPs report on experiments involving multiple inversions from several modeling groups using different combinations of in situ and spatially aggregated OCO-2 products in multiple observing modes. In addition to diagnosing differences among combinations of data sources and inversion systems, these efforts seek consensus flux estimates that suitably combine the results. Cressie et al. (2022) discuss an ANOVA-based approach with associated statistical model to develop weights for individual results in estimating a consensus flux, and the method is demonstrated for regionally-aggregated fluxes.

Functional ANOVA extends the classical approach to settings with quantities of interest that are functions of known inputs such as space and/or time. The statistical model is typically extended with a specification for the relationships among the ANOVA components across space/time. For applications involving spatial fields, individual ANOVA effects are typically assumed to be spatially correlated, and this structure can be estimated from the data available. This strategy has been applied to output from regional climate models (RCMs; Kaufman and Sain, 2010; Sain et al., 2011; Kang and Cressie, 2013). Kaufman and Sain capture spatial dependence in ANOVA effects with Gaussian process (GP) models. Estimation and inference for GP models can be computationally demanding due to operations such as matrix inversion and Cholesky factorization. Sain et al. develop a Markov random field (MRF) model with a sparse precision matrix, but in two spatial dimensions, the cost for the necessary Cholesky decomposition is $\mathcal{O}(N^{3/2})$, where N is the number of locations. Another option used in the current work is

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022


Geoscientific Model Development Discussions

to model the Cholesky factor as sparse with the Vecchia approximation (Vecchia, 1988; Katzfuss and Guinness, 2021; Schäfer et al., 2021).

In this paper, we implement the functional ANOVA methodology for multiple flux inversion solutions in order to identify meaningful, spatially-coherent, carbon cycle signals and partition variability among various solutions in the multi-model ensemble. We illustrate the approach for a recent model intercomparison (MIP) effort (Peiro et al., 2022) and for flux estimates produced with multiple spatial aggregation approaches (Nguyen et al., 2020). Specifically, in section 2, we formulate the statistical model for functional ANOVA. In section 3, we describe the OCO-2 flux MIP and datasets. Results are presented in section 4, and concluding remarks are provided in section 5.

2 Datasets

In subsequent sections, we employ the functional ANOVA methodology for multiple collections of flux inversions using in situ data and products from OCO-2. For the satellite data, the inversion systems use retrievals of XCO_2, the column-average dry air mole fraction of CO_2. These retrievals, termed Level 2 data products, use the Atmospheric Carbon Observations from Space (ACOS) retrieval algorithm (O'Dell et al., 2018) to infer atmospheric CO_2 from the Level 1 satellite spectra. Additional diagnostic data from the retrievals, including XCO_2 averaging kernels, are used in the inversions to map model states to the retrieval space when the data products are assimilated. For OCO-2 in particular, the large data volume and small spatial footprint are impractical for global flux inversions, so the retrievals are often aggregated spatially within a single polar orbit. These aggregated retrievals are examples of Level 3 products that aim to provide additional utility to a broader user community by providing manageable data volume and a regular spatio-temporal structure. In addition, aggregation can provide a precise estimate of a quantity of interest (CO_2 concentration) at a coarser resolution that is still meaningful for applications. As an illustration of functional ANOVA, we investigate the impact of spatial aggregation on flux estimates in the presence of additional sources of variability.

2.1 Records of Fused CO_2

An ongoing NASA MEaSUREs effort aims to provide inversion-ready data products that use OCO-2 and GOSAT retrievals. The effort produces spatially-aggregated and gap-filled estimates of XCO_2 at daily intervals that span the period of overlap for these satellite records, from 2014 to present. The gridded OCO-2 product and multi-instrument fused product are available from the GES DISC (https://doi.org/10.5067/A613YBQVPHCD). The spatially aggregated XCO_2, along with additional quantities used in flux inversion, is estimated using a local kriging approach (Nguyen et al., 2020). The methodology accounts for and exploits the short-range spatial correlation present in the Level 2 retrievals (Torres et al., 2019; Worden et al., 2017). The spatial dependence and uncertainty associated with the aggregated XCO_2 are estimated from the available Level 2 retrievals and vary in space and time. The spatial aggregation also reduces the data volume, making the OCO-2 record manageable for ingestion into global flux inversion systems.

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022


Geoscientific Model Development Discussions

Other spatial aggregation approaches have been devised for OCO-2 inversions. The averaging outlined in Baker et al. (2022) has been implemented for the OCO-2 flux MIP (Peiro et al., 2022) and will be discussed in the next subsection. In addition, the NASA Carbon Monitoring System Flux (CMS-Flux) four-dimensional variational (4D-Var) inversion framework has used an aggregation approach termed "super-obs" (Liu et al., 2017; Byrne et al., 2020). The model is driven by the Goddard Earth Observing System version 5 of the NASA Global Modeling Assimilation Office (GEOS-FP) meteorology and runs at a $4^\circ \times 5^\circ$ resolution. The CMS-Flux team has performed an experiment with two separate inversions: a run that ingested the traditional OCO-2 super-obs and another that ingested the MEaSUREs gridded product.

The ANOVA methodology is particularly convenient for analyzing the quantitative outcomes of experiments or trials under various discrete combinations of one or more *factors* of interest. The approach formulates a statistical model that is outlined in the next section. The parameters of the model include an overall mean response and additive effects for each combination of factors. *Replication* within factor combinations allows the decomposition of variability between (mean differences) and within combinations (noise). Our demonstration of the functional ANOVA for the records of fused CO_2 incorporates the data aggregation method as one experimental factor, with the two levels being the super-obs approach (Control) and the SSDF (Fused). The second factor in this investigation will be an interannual effect, contrasting June-July-August (JJA) of 2016 with the same time period in 2015. The variable of interest are the CMS-Flux estimates available at monthly increments, so individual months represent replicates in this example. For spatio-temporal data applications, time is often used as a pseudo-replicate (Cressie et al., 2022; Sain et al., 2011). Figure 1 shows the collection of CMS-Flux estimates that are used in the functional ANOVA demonstration.

2.2 OCO-2 Flux MIP

Our second demonstration of the functional ANOVA approach comes from a multi-institution flux MIP using data products from OCO-2, which provides global estimates of XCO_2 suitable for assimilation into a inversion systems that estimate carbon fluxes at regional to global scales. The maturity and diversity of inversion systems continue to grow with scientific interest in the carbon cycle, and a sizeable collection of research groups have the capability to provide global flux estimates based on satellite data products. Like many inferences for the Earth system, flux estimation is an ill-posed inverse problem. An inversion system combines available atmospheric CO_2 concentration data with a transport model, prior assumptions on fluxes, and a statistical/computational inverse method. Various combinations of these system components are employed in satellite CO_2 inversion frameworks.

Multiple flux inversion teams applied a common inversion protocol to their individual inversion systems as part of the OCO-2 Version 9 Model Intercomparison Project (V9 MIP; Peiro et al., 2022). The MIP was designed in part to quantify the impacts of the above inversion system elements on flux estimates. In addition, each team conducted multiple inversion experiments using the same collections of atmospheric CO_2 data. The data collections represent combinations of in situ (IS) surface-based CO_2 observations and aggregated OCO-2 retrievals from the Version 9 products (Kiel et al., 2019). The OCO-2 collections use combinations of its primary observing modes and surface types: land nadir (LN), land glint (LG), and ocean glint (OG). As noted previously, the individual retrievals are both variable and moderately correlated in space and time (Worden et al.,

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

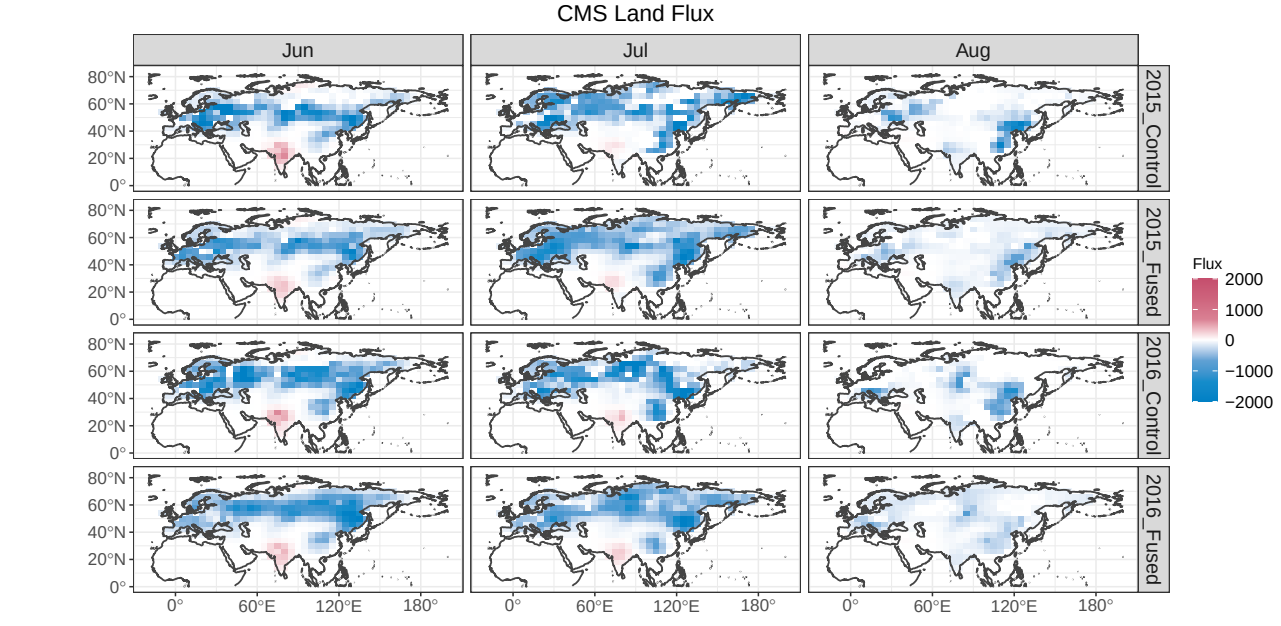

Figure 1. Monthly flux estimates from CMS-Flux for combinations of year and aggregation method. The top two rows depict fluxes for JJA of 2015 from the two aggregation methods, and the bottom two rows depict fluxes for JJA 2016. Fluxes are in units of gC m^{-2} yr^{-1}.

2017). The V9 MIP used spatially aggregated OCO-2 retrievals following the methodology outlined in Baker et al. (2022). The aggregated retrievals include uncertainty estimates that incorporate assumed spatial correlation in retrieval errors and transport uncertainty.

The V9 MIP flux experiment suite includes estimates from ten inversion systems (Peiro et al., 2022, Tables 1-2) and four combinations of data collections. Our investigation focuses on flux estimates using IS and LNLG data collections, which were also the focus in the MIP. Further, we illustrate the functional ANOVA for a subset of four inversion systems as outlined in Table 1. This subset was selected to provide a contrast among different atmospheric transport models that have similar resolution for the flux solution. The functional ANOVA is implemented for the spatially referenced monthly flux estimates for JJA 2016 over North America and separately for the same time period over Africa. Figures 2 and 3 provide maps of these collections of flux estimates. In the subsequent implementation of functional ANOVA, the first factor is the flux model and the second factor is the data source (IS and LNLG).

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions
Open Access
EGU

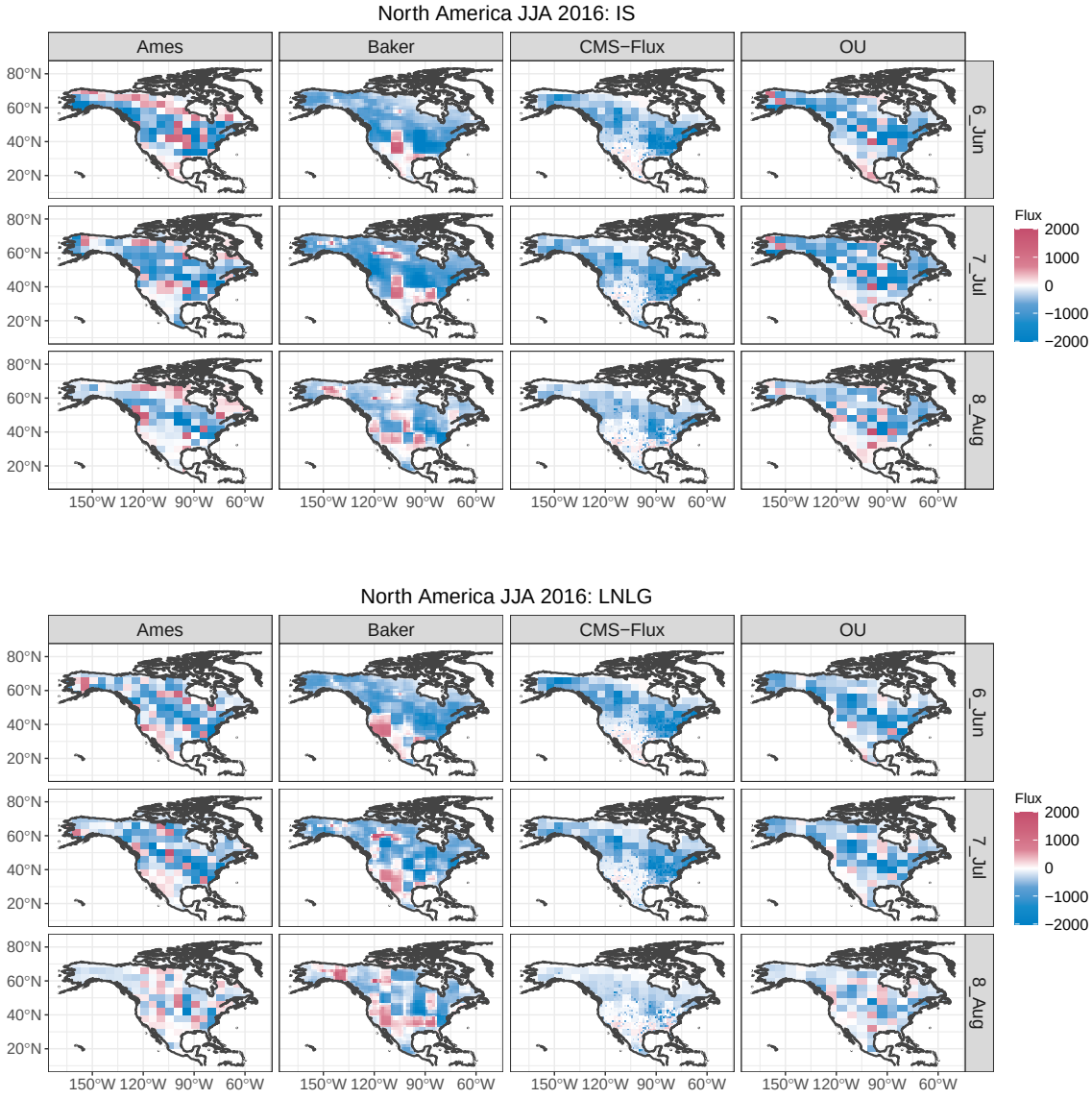

Figure 2. Monthly flux estimates from four inversion systems for JJA 2016 based over North America. The top three rows show flux estimates using in situ (IS) data, and the bottom three rows show flux estimates from OCO-2 land nadir and glint (LNLG) retrievals.

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development
Discussions
Open Access
EGU

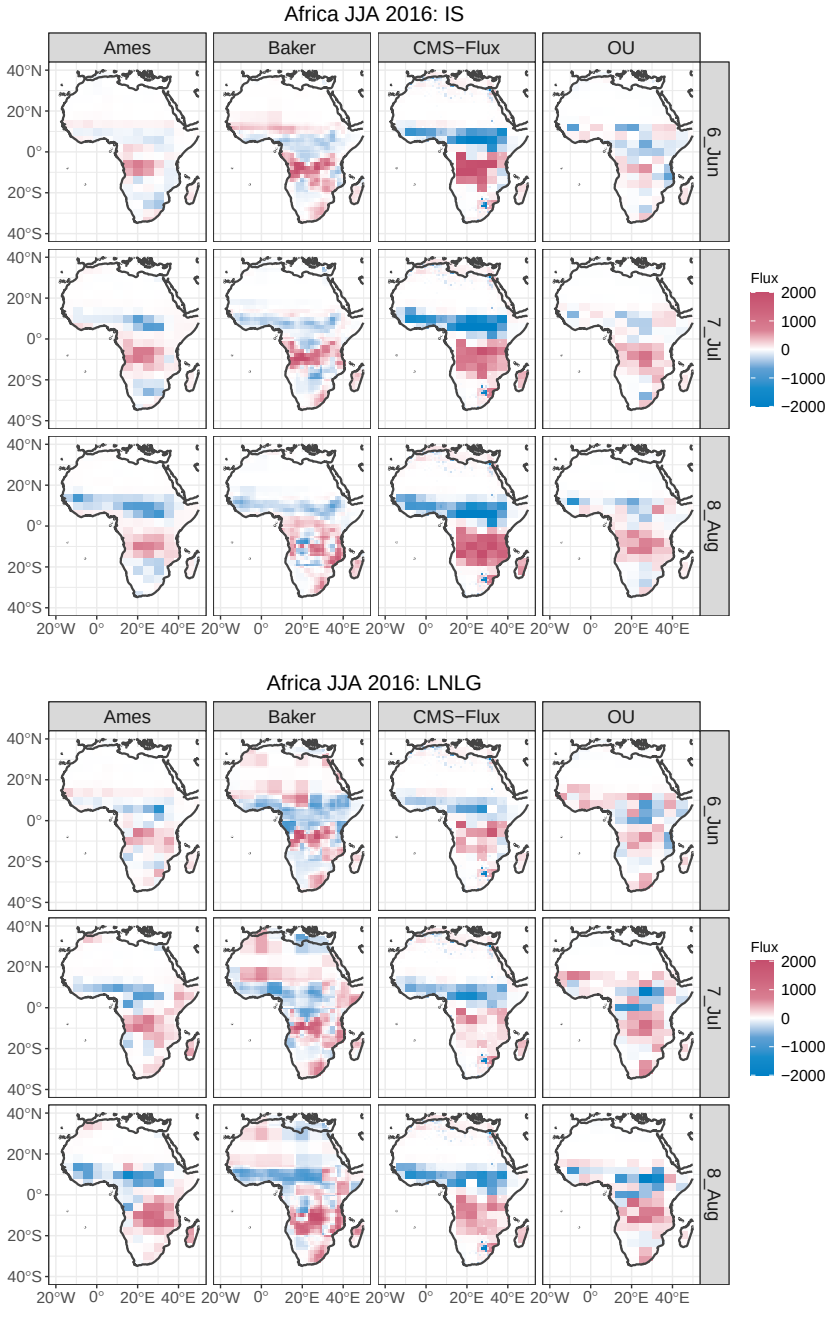

Figure 3. Monthly flux estimates from four inversion systems for JJA 2016 over Africa. The top three rows show flux estimates using in situ (IS) data, and the bottom three rows show flux estimates from OCO-2 land nadir and glint (LNLG) retrievals.

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions
Open Access
EGU

Table 1. Flux inversion systems in this study; see Tables 1-2 of Peiro et al. (2022) for further details.

Model name	Institution	Transport model	Inverse method
Ames	NASA Ames Research Center	GEOS-Chem	4D-Var
Baker	Colorado State University	PCTM	4D-Var
CMS-Flux	NASA Jet Propulsion Laboratory	GEOS-Chem	4D-Var
OU	University of Oklahoma	TM5	4D-Var

3 Functional ANOVA

Analysis of variance (ANOVA) is a statistical method with a long history connected to designed experiments. In such experiments, one or more factors can be controlled at levels selected by the the experimenters, and ANOVA provides a framework for estimating the factors' impact on response variables of interest. The method relies on replication within combinations of factors, or treatments, in order to estimate a mean response for each combination of factors, along with a partitioning of variability between and within treatments. The treatment means are typically re-parameterized into an overall mean and individual effects for each level of the factors, as well as interaction effects.

The classic implementation of ANOVA considers a univariate response, such as an integrated or average carbon flux over a region of interest. This is frequently extended to a multivariate response with MANOVA, and the decomposition of variance is accompanied by estimation of the correlation structure among the multivariate responses (Johnson and Wichern, 2002). As the dimension of the multivariate response grows, the number of parameters to be estimated from the available data grows as well. The ANOVA approach can be extended to spatial fields using tools from functional data analysis and spatial statistical modeling. In this setting, the dimension can be large but the parameter space can be managed through a hierarchical approach and by exploiting the spatial dependence present in the data. This functional ANOVA approach has been implemented for spatial fields of output from climate model experiments (Kaufman and Sain, 2010; Kang and Cressie, 2013). Our implementation and notation for the carbon flux inversion results generally follows that from Kaufman and Sain.

3.1 Statistical Model

In the current work, we invoke a two-way functional ANOVA in the context of carbon flux fields over land. In the two-way model, there are two experimental factors examined, generically termed Factor A and Factor B. In the records of fused CO_2 example, year is factor A, and aggregation method is factor B. In the OCO-2 flux MIP example, the modeling group is factor A, and data source is factor B. Then $y_{ijk}(\mathbf{s})$ represents the flux field at location $\mathbf{s}$ for level (setting) i of factor A, level j of factor B, and replicate k. In addition, flux inversions incorporate a space-time varying prior flux field, which we denote $y^{(0)}_{ijk}(\mathbf{s})$. The prior fluxes incorporate biospheric contributions, fossil fuel emissions, and fires. The CMS-Flux prior methodology is summarized in Liu et al. (2017) and Byrne et al. (2020), and the OCO-2 MIP prior specifications are summarized in Table 1 of

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions

Table 2. Summary of flux inversion results used in functional ANOVA examples. Region numbers indicate TRANSCOM regions used in each example (e.g. Crowell et al., 2019).

Experiment	Domain	Time Period	Factor A	Factor B
Records of Fused CO_2	Eurasia Regions 7, 8, 11	JJA 2015, 2016	Year $n_\alpha = 2$	Aggregation Method $n_\beta = 2$
OCO-2 V9 MIP	North America Regions 1, 2	JJA 2016	Inversion System $n_\alpha = 4$	Data Source $n_\beta = 2$
OCO-2 V9 MIP	Africa Regions 5, 6	JJA 2016	Inversion System $n_\alpha = 4$	Data Source $n_\beta = 2$

Peiro et al. (2022). The functional ANOVA statistical model can be written in two equivalent forms,

$$y_{ijk}(\mathbf{s}) - y^{(0)}_{ijk}(\mathbf{s}) = \mu_{ij}(\mathbf{s}) + \epsilon_{ijk}(\mathbf{s}) \tag{1}$$

$$= \mu(\mathbf{s}) + \alpha_i(\mathbf{s}) + \beta_j(\mathbf{s}) + (\alpha\beta)_{ij}(\mathbf{s}) + \epsilon_{ijk}(\mathbf{s}). \tag{2}$$

The *cell means* formulation in (1) clearly defines a unique mean field for each treatment, μ_{ij}, at level i of factor A and level j of factor B. With n_α levels of factor A and n_β levels of factor B, there are $n_\alpha \times n_\beta$ cell mean fields. The *effects* model in (2) separates the mean field into the additive effects of the experimental factors. In the effects model, μ is the mean field representing spatial features in the common response, α_i quantifies the variation around μ due to level i of factor A, β_j quantifies the variation around μ due to level j of factor B, and $(\alpha\beta)_{ij}$ is an interaction effect. In both forms, ϵ_{ijk} quantifies the internal variability within each scenario. In all examples, the replication within each treatment, indexed by k, is across months within a season. Table 2 defines the factors for the carbon flux functional ANOVA examples.

The effects model (2) is commonly used in ANOVA because it provides a convenient setup for partitioning variability among the factors and their interaction. Further, inference for the overall mean effect μ and linear contrasts among the factors is straightforward. However, the model is over-parameterized, so constraints among the effects are often enforced to ensure identifiability. Following Kaufman and Sain (2010), the flux inversion examples invoke sum-to-zero constraints, meaning that the factor and interaction effects add to zero, e.g.

$$\sum_{i=1}^{n_\alpha} \alpha_i(\mathbf{s}) = 0$$

$$\sum_{j=1}^{n_\beta} \beta_j(\mathbf{s}) = 0$$

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions
Open Access
EGU

for all locations $\mathbf{s}$. For interaction effects, the constraints apply across all levels of each factor,

$$\sum_{i=1}^{n_\alpha}(\alpha\beta)_{ij}(\mathbf{s}) = 0; \quad j = 1,\ldots,n_\beta,$$
$$\sum_{j=1}^{n_\beta}(\alpha\beta)_{ij}(\mathbf{s}) = 0; \quad i = 1,\ldots,n_\alpha.$$

In classic univariate ANOVA, the effects model parameters are estimated by assembling a series of contrast effects of reduced dimension to ensure identifiability. For factor A, there are $n_\alpha - 1$ of these contrast effects, denoted α^*. In the case where $n_\alpha = 2$, there is a single contrast effect and can be interpreted as the difference in mean response for the two levels of Factor A; in the records of fused CO_2 example, this contrast effect is the difference between the two years studied, 2016 and 2015. Similarly, factor B has $n_\beta - 1$ contrast effects β^*, and there are $(n_\alpha - 1) \times (n_\beta - 1)$ contrast effects for interaction, $(\alpha\beta)^*$. The contrast effects are related to the original effects model (2) parameters through linear transformations. This is discussed further in the supplement and in Kaufman and Sain (2010).

Since the quantities of interest are spatial fields, the ANOVA effects are functions of location. The estimation can account for this structure and exploit potential spatial correlation if a suitable spatial statistical model is incorporated in a hierarchical fashion. To that end, a Gaussian process (GP) is assumed for each spatial field. For the spatial field of mean effects,

$$\boldsymbol{\mu} \sim GP(\mathbf{0}, \mathbf{C}_{\boldsymbol{\theta}_\mu}).$$

The covariance function $\mathbf{C}$ describes the covariance between any realizations as a function of their locations. For the carbon flux inversion examples, the Matérn covariance model is used and is parameterized as $\boldsymbol{\theta}_\mu \equiv \{\sigma_\mu, \lambda_\mu, \nu_\mu\}$, where σ_μ is a standard deviation, λ_μ is a range parameter describing the rate of decay of spatial correlation with distance, and ν_μ is a smoothness parameter (Stein, 1999). The GP mean here is taken to be zero because the flux prior has been subtracted in (1)-(2).

Analogous GP assumptions are made for the remaining ANOVA model components. Some of these components, particularly the error process $\boldsymbol{\epsilon}_{ijk}$ have multiple realizations, which are assumed to be independent and identically distributed (iid) GP realizations,

$$\boldsymbol{\epsilon}_{ijk} \sim GP(\mathbf{0}, \mathbf{C}_{\boldsymbol{\theta}_\epsilon}); \quad i = 1,\ldots,n_\alpha;\ j = 1,\ldots,n_\beta;\ k = 1,\ldots,n_\epsilon$$

The GP assumptions are applied to the contrast effect specification of the main effects and interactions,

$$\boldsymbol{\alpha}^*_{i'} \sim GP(\mathbf{0}, \mathbf{C}_{\boldsymbol{\theta}_\alpha}); \quad i' = 1,\ldots,n_\alpha - 1$$
$$\boldsymbol{\beta}^*_{j'} \sim GP(\mathbf{0}, \mathbf{C}_{\boldsymbol{\theta}_\beta}); \quad j' = 1,\ldots,n_\beta - 1$$
$$(\boldsymbol{\alpha\beta})^*_{i'j'} \sim GP(\mathbf{0}, \mathbf{C}_{\boldsymbol{\theta}_{\alpha\beta}}); \quad i' = 1,\ldots,n_\alpha - 1; \quad j' = 1,\ldots,n_\beta - 1$$

3.2 Estimation

Bayesian inference is commonly used for hierarchical spatio-temporal statistical models and has been implemented in previous work on functional ANOVA incorporating spatial dependence (Kaufman and Sain, 2010; Kang and Cressie, 2013). The com-

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022


Geoscientific Model Development Discussions

putational overhead for Bayesian inference can be substantial, particularly when working with GP models for spatial processes, since operations must be performed on large, dense covariance matrices numerous times. Therefore, the GP models in the functional ANOVA are formulated using the Vecchia approximation of Katzfuss and Guinness (2021). The representation yields sparse matrices that allow for more efficient computations in Bayesian inference. The GP approximations are still functions of the same parameters as the original Matérn models, and each ANOVA component has a unique set of parameters. Bayesian inference interrogates the joint posterior distribution of the ANOVA components and parameters, $p()$, given the available flux fields. This is written as

$$p\left(\boldsymbol{\mu}, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, (\boldsymbol{\alpha}\boldsymbol{\beta})^*, \boldsymbol{\theta} | \boldsymbol{y} - \boldsymbol{y}^{(0)}\right) \propto f\left(\boldsymbol{y} - \boldsymbol{y}^{(0)} | \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, (\boldsymbol{\alpha}\boldsymbol{\beta})^*, \boldsymbol{\theta}\right) \pi\left(\boldsymbol{\theta}\right), \tag{3}$$

where $f(\boldsymbol{y} - \boldsymbol{y}^{(0)}|\cdot)$ is a joint Gaussian likelihood arising from the GP models for the ANOVA components, including the noise, and $\pi\left(\boldsymbol{\theta}\right)$ is a prior distribution for the collection of GP parameters. Prior distributions are independent for all elements of $\boldsymbol{\theta}$. The distributional forms and parameters are selected to maintain proper, yet diffuse, prior distributions guided by previous work on Bayesian analysis of hierarchical models (Gelman, 2006; Kaufman and Sain, 2010). Further details on the prior distribution assumptions are provided in the supplement.

The posterior distribution (3) is complex and high-dimensional, but it can be sampled using Markov chain Monte Carlo methods (MCMC). In particular, a Metropolis-within-Gibbs MCMC algorithm is used (Gelman et al., 2013). This approach uses the general Gibbs sampler to sample sequentially at each iteration from individual-component conditional posterior distributions, $p(\boldsymbol{\mu}|\boldsymbol{\theta}_\mu, \boldsymbol{\theta}_\epsilon, \boldsymbol{y} - \boldsymbol{y}^{(0)})$, $p(\boldsymbol{\alpha}^*|\boldsymbol{\theta}_\alpha, \boldsymbol{\theta}_\epsilon, \boldsymbol{y} - \boldsymbol{y}^{(0)})$, $p(\boldsymbol{\beta}^*|\boldsymbol{\theta}_\beta, \boldsymbol{\theta}_\epsilon, \boldsymbol{y} - \boldsymbol{y}^{(0)})$, and $p((\boldsymbol{\alpha}\boldsymbol{\beta})^*|\boldsymbol{\theta}_{\alpha\beta}, \boldsymbol{\theta}_\epsilon, \boldsymbol{y} - \boldsymbol{y}^{(0)})$. As outlined in the supplement, the individual component distributions are multivariate Gaussian and depend on summary statistics of the data $\boldsymbol{y} - \boldsymbol{y}^{(0)}$, the GP parameters for the component and the noise, but not on other ANOVA components. In addition to these distributions, the Gibbs sampler cycles through draws from the GP parameters conditional distributions: $p(\boldsymbol{\theta}_\mu|\boldsymbol{\mu})$, $p(\boldsymbol{\theta}_\alpha|\boldsymbol{\alpha}^*)$, $p(\boldsymbol{\theta}_\beta|\boldsymbol{\beta}^*)$, $p(\boldsymbol{\theta}_{\alpha\beta}|(\boldsymbol{\alpha}\boldsymbol{\beta})^*)$, and $p(\boldsymbol{\theta}_\epsilon|(\boldsymbol{y} - \boldsymbol{y}^{(0)})$. Each of these distributions are sampled with a Metropolis-Hastings (MH) proposal. Further details on the MCMC procedure can be found in the supplement.

The functional ANOVA model and MCMC algorithm for the carbon flux examples is broadly similar to previous demonstrations with climate model output (Kaufman and Sain, 2010; Sain et al., 2011; Kang and Cressie, 2013), but there are a few notable extensions in the current work. The OCO-2 MIP examples use more than two levels per factor, which is addressed through the contrast effects. All examples include Bayesian inference for all Matérn parameters; in previous work the smoothness parameter ν is often fixed. Finally, the current work invokes the Vecchia approximation (Katzfuss and Guinness, 2021) for the Gaussian processes, which provides necessary computational efficiency for the MCMC algorithm.

4 Results

The MCMC algorithm outlined in the previous section yields a large collection of random draws from the posterior distribution of the GP parameters and ANOVA components. The posterior samples can be summarized for individual parameters, as well as for arbitrary functions of them. For example, while the MCMC samples the contrast effects α^*, the draws can be transformed

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions

Table 3. Functional ANOVA Gaussian process parameter estimates for the CMS-Flux inversions in the records of fused CO_2 experiment. The posterior median is the value listed first in each cell, and the values in parentheses are the lower and upper endpoints of 95% posterior credible intervals. Data and standard deviations have units of gC m^{-2} yr^{-1}. Range parameters have units of km and smoothness parameters are unitless.

CMS-Flux Fusion Experiment, Eurasia JJA 2015-16					
Parameter	Mean μ	Year α 2016 - 2015	Aggregation β Fused - SuperObs	Interaction $(\alpha\beta)$	Error ϵ
Standard Deviation σ	55 (34, 87)	20 (1, 58)	2.5 (0.1, 8.6)	1.9 (0.1, 7.0)	233 (225, 242)
Range λ [km]	1100 (520, 3000)	1550 (170, 6800)	150 (50, 410)	160 (60, 440)	420 (370, 490)
Smoothness ν	0.88 (0.53, 1.51)	1.08 (0.60, 1.67)	0.94 (0.50, 1.54)	0.95 (0.50, 1.58)	0.78 (0.69, 0.88)

to summarize the main effects α. The functional ANOVA results for the experiments outlined in Table 2 are summarized in various ways in this section.

4.1 Records of Fused CO_2

The CMS-Flux inversion results over Eurasia for JJA 2015 and 2016 using the super-obs and data fusion aggregation methods were incorporated into the first functional ANOVA implementation. Table 3 summarizes the posterior distributions for the GP parameters for each of the ANOVA components. The estimated GP standard deviation is largest for the error fields ϵ, indicating that the month-to-month variability within each treatment combination is relatively large. However, the estimated range parameter λ_ϵ is relatively small. The estimated range for the overall mean μ and year effect α exceed 1000 km and are an order of magnitude larger.

The location-specific posterior means of the ANOVA components are summarized in Figure 4. The upper left panel provides the estimates of $\mu(\mathbf{s})$, the overall mean deviation from the prior flux for JJA across the two years, 2015-16. There are broad swaths of the domain with sizeable negative mean effects, particularly over central and eastern Asia. The estimated difference between 2016 and 2015, $\alpha^*(\mathbf{s})$, shown in the upper right panel also exhibits large-scale coherence but of a modest magnitude. The contrast effects for the aggregation method, $\beta^*(\mathbf{s})$, and for interaction, $(\alpha\beta)^*(\mathbf{s})$, are shown in the bottom left and right panels, respectively. In both cases, the magnitudes are small with limited spatial coherence.

The product development team for the records of fused CO_2 is particularly interested in the impact of the aggregation method on the estimated fluxes from an inversion system. Figure 5 illustrates a metric from the posterior distribution for detecting meaningful flux anomalies in the presence of different aggregation methods. The left panel shows $Pr(|\mu(\mathbf{s})| > |\beta^*(\mathbf{s})|)$, the location-specific posterior probability that the overall mean flux anomaly has a magnitude greater than the magnitude of the effect due to aggregation method. The probabilities are nearly 1 for the entire domain, with some slight regional differences. This suggests that the overall mean flux anomalies are meaningfully distinguishable from differences due to use of the fused product versus the super-obs approach. Similarly, the right panel of Figure 5 shows $Pr(|\alpha^*(\mathbf{s})| > |\beta^*(\mathbf{s})|)$, the posterior

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions
Open Access
EGU

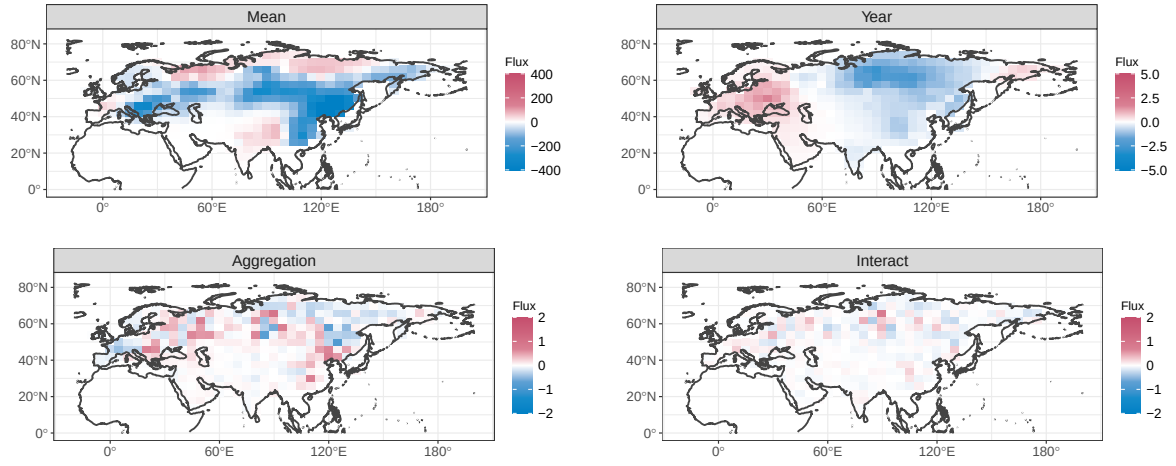

Figure 4. Posterior means for functional ANOVA model components for the records of fused CO_2 example. Note the different color scales for the panels.

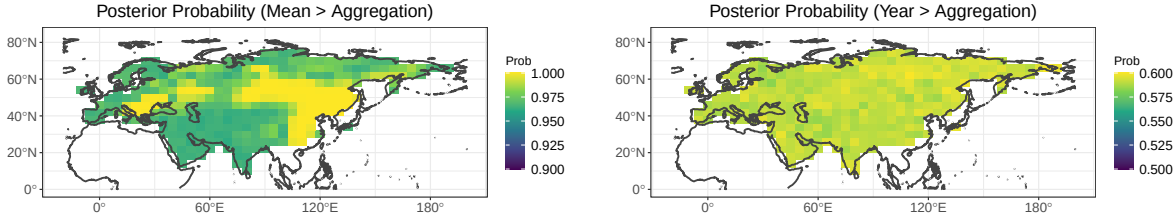

Figure 5. Location specific posterior probabilities for the records of fused CO_2 example. Left panel shows the probability that the magnitude of the overall mean exceeds the aggregation method effect, $Pr(|\mu(\mathbf{s})| > |\beta^*(\mathbf{s})|)$. Right panel shows the probability that the magnitude of the year effect exceeds the aggregation method effect, $Pr(|\alpha^*(\mathbf{s})| > |\beta^*(\mathbf{s})|)$

probability that the year-to-year difference has a magnitude greater than the aggregation effect. These probabilities are fairly uniform spatially at around 0.6.

4.2 OCO-2 Flux MIP

The functional ANOVA inference was carried out separately for flux fields from the OCO-2 V9 flux MIP over North America and Africa for JJA 2016. In both cases the two ANOVA factors are the flux inversion system with four levels (modeling groups) and the data source with two levels (IS and LNLG). These two regions represent distinct scenarios for the methodology for a number of reasons. The carbon cycle of the temperate and boreal land regions, and transitions therein, of North America differ from the tropical and subtropical areas of Africa. In addition, data availability for the two regions is markedly different. As shown in Figure 1 of Peiro et al. (2022), the density of in situ CO_2 observations is substantially higher over North America than

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions
Open Access
EGU

Table 4. Functional ANOVA Gaussian process parameter estimates for North America and Africa for JJA 2016. The posterior median is the value listed first in each cell, and the values in parentheses are the lower and upper endpoints of 95% posterior credible intervals. Data and standard deviations have units of gC m^{-2} yr^{-1}. Range parameters have units of km and smoothness parameters are unitless.

OCO-2 MIP, North America JJA 2016

Parameter	Mean μ	Flux Model α	Data Source β	Interaction $(\alpha\beta)$	Error ϵ
Standard Deviation σ	153 (137, 173)	22 (14, 35)	19 (2, 39)	0.6 (0.2, 2.1)	409 (403, 414)
Range λ [km]	196 (159, 248)	2300 (1000, 6900)	340 (80, 740)	160 (70, 360)	104 (101, 108)
Smoothness ν	1.10 (0.99, 1.22)	1.06 (0.67, 1.48)	1.30 (0.64, 1.82)	0.89 (0.53, 1.33)	1.28 (1.25, 1.30)

OCO-2 MIP, Africa JJA 2016

Parameter	Mean μ	Flux Model α	Data Source β	Interaction $(\alpha\beta)$	Error ϵ
Standard Deviation σ	68 (60, 77)	6.6 (1.1, 13.5)	27 (20, 39)	11.2 (6.6, 16.1)	199 (197, 202)
Range λ [km]	203 (160, 264)	700 (260, 1770)	500 (330, 830)	580 (420, 810)	70.5 (67.8, 73.6)
Smoothness ν	1.48 (1.27, 1.72)	1.64 (0.93, 2.12)	1.68 (1.29, 2.18)	1.80 (1.44, 2.29)	2.96 (2.81, 3.10)

over Africa. OCO-2 has dense coverage over both continents with some regional disparities (O'Dell et al., 2018). For example, OCO-2 has substantially more successful retrievals over northern and southern Africa than over the tropics. Despite some of these differences in behavior, both continents should exhibit some spatial heterogeneity in the overall flux signal.

Table 4 summarizes the posterior distributions for the GP parameters for the OCO-2 MIP functional ANOVA. Once again, the estimated GP standard deviation σ_ϵ is largest for the error fields ϵ for both regions, indicating sizable month-to-month variability after accounting for the overall mean anomaly along with the model, data source, and interaction effects. In addition, the spatial range λ_ϵ is relatively short for the error component. The estimated spatial range is largest among all components for the flux model effect (λ_α) for both regions. This result is consistent with large-scale regional flux differences among inversion systems with different driving atmospheric transport (Peiro et al., 2022; Basu et al., 2018). The two regions differ slightly in the relative variability of the model, data source, and interaction effects ($\sigma_\alpha, \sigma_\beta, \sigma_{(\alpha\beta)}$). Over North America, the model and data source effects have similar variability, and the interaction effect is an order of magnitude smaller. On the other hand, the variability in the model and interaction effects are similar over Africa, but the data source effect standard deviation, σ_β, is larger than these. The difference in coverage between in situ and OCO-2 likely contributes to this relatively large data source effect.

Figure 6 shows the estimated spatial pattern for the overall mean flux anomaly, $\mu(\mathbf{s})$, for the OCO-2 MIP North America functional ANOVA. Noting again that the analysis is carried out the inversions' deviations from their respective priors, the map depicts a consensus flux anomaly from the collection of inversions analyzed. The patterns indicate generally negative anomalies (increased uptake) over the eastern United States and positive anomalies over the southwestern U.S. While some spatial coherence is present, the estimated range is around 200 km and less than that for the model effect in particular. This difference is evident in Figure 7, which provides maps of the estimated (posterior mean) individual model effects, $\alpha(\mathbf{s})$, as

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions
Open Access
EGU

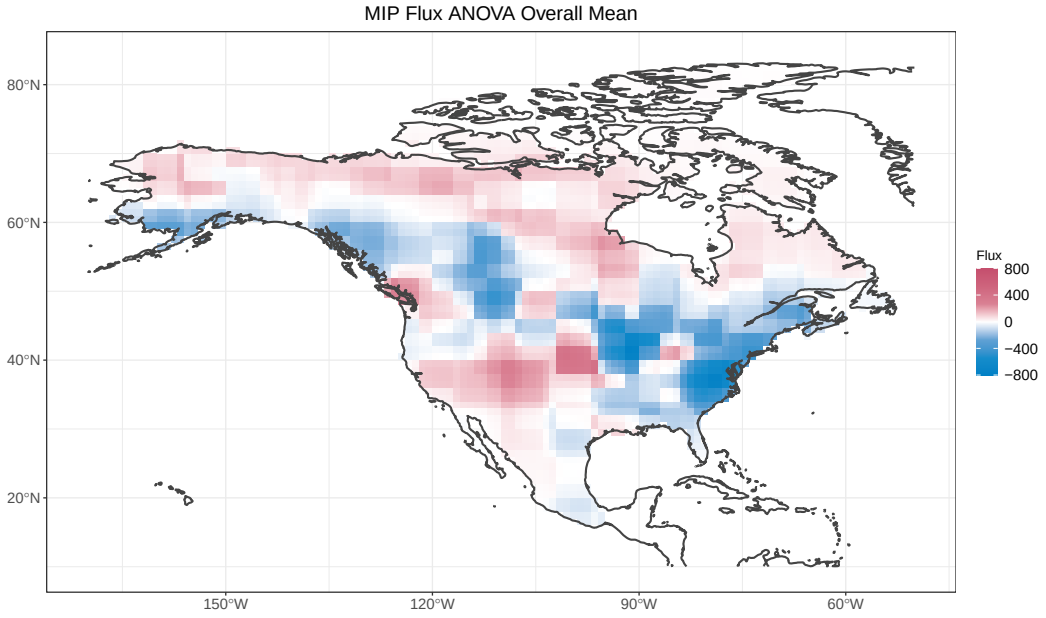

Figure 6. Posterior mean for the functional ANOVA overall mean μ for JJA 2016 over North America.

well as the data source effects, $\beta(\mathbf{s})$. While generally smaller in magnitude than some of the localized mean anomalies, the estimated model effects in the top panels of Figure 7 are quite coherent across the continent for each model. The data source effect exhibits spatial dependence that decays at shorter distances.

The MCMC procedure provides samples from the full joint posterior distribution, and the samples can be summarized in various ways to describe the uncertainty for quantities of interest. Figure 8 summarizes the posterior distribution for the overall mean $\mu(\mathbf{s})$ through spatially-referenced credible intervals. Some of the inferred local anomalies are evident here, including negative anomalies over the U.S. Midwest and Atlantic coast, and modest positive anomalies over the Southwest.

Figure 9 shows the estimated spatial pattern for the overall mean flux anomaly, $\mu(\mathbf{s})$, for the OCO-2 MIP Africa functional ANOVA. This posterior mean map suggests broad negative anomalies over western tropical Africa. The prior fluxes over the continent (not shown) contrast uptake north of the Equator and a net source to the south. This contrast manifests in some of the remaining ANOVA effects, as shown in Figure 10, where the estimated model effects (top panels) change sign across the Equator. It should also be noted that the magnitude of these model effects is generally smaller over Africa than over North America. A north-south contrast is also evident in the data source effect estimates in the lower panel of Figure 10. The contrast between the two data sources (IS - LNLG) is captured in the contrast effect $\beta^*(\mathbf{s})$. Posterior credible intervals for this contrast are mapped in Figure 11. For most of the continent the intervals cover zero, but the in situ inversions appear to have consistently higher fluxes over southern tropical Africa.

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions
Open Access
EGU

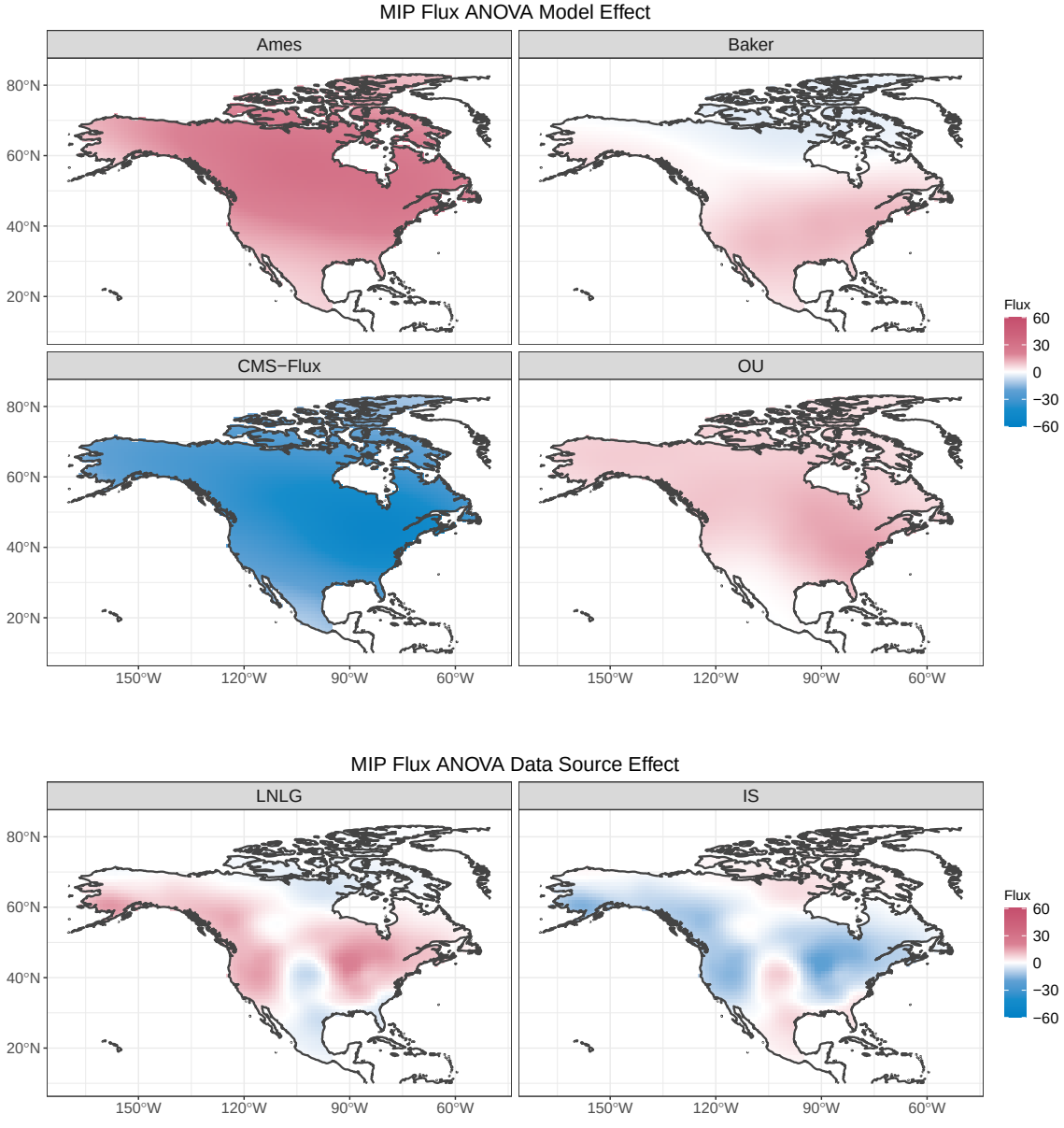

Figure 7. Posterior mean for the functional ANOVA main effect for flux model (α, top) and for data source (β, bottom) for JJA 2016 over North America.

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development
Discussions
Open Access
EGU

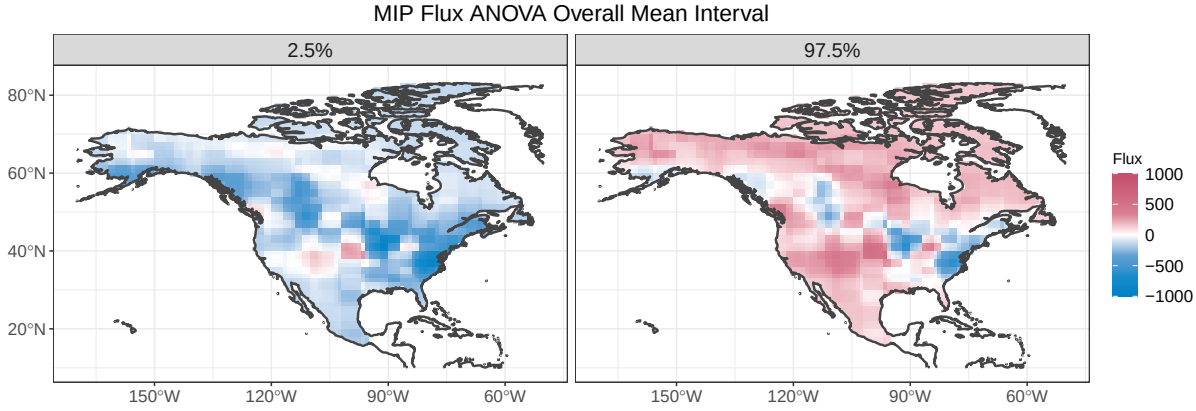

Figure 8. Posterior credible intervals for the functional ANOVA overall mean μ for JJA 2016 over North America. Left (right) panel depicts the 2.5th (97.5th) percentile of the posterior distribution for each location.

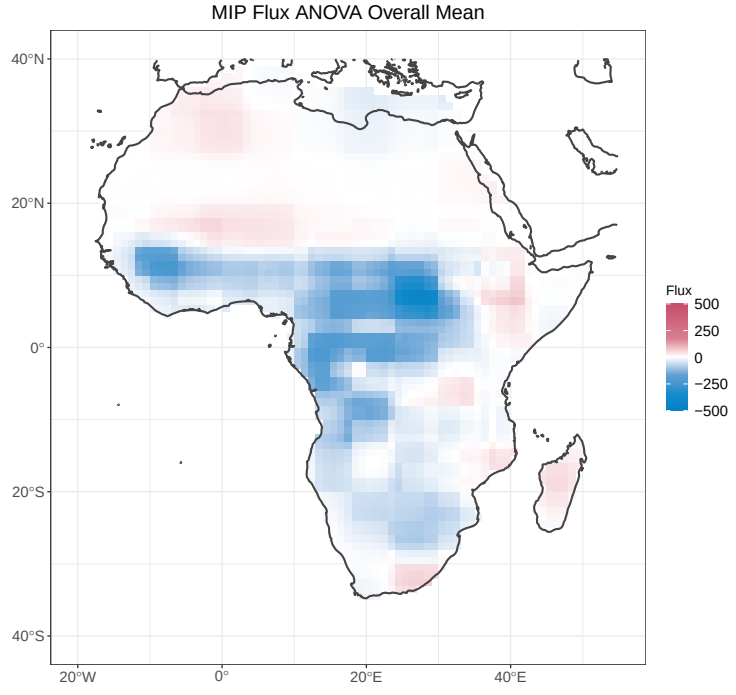

Figure 9. Posterior mean for the functional ANOVA overall mean μ for JJA 2016 over Africa.

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development
Discussions
Open Access
EGU

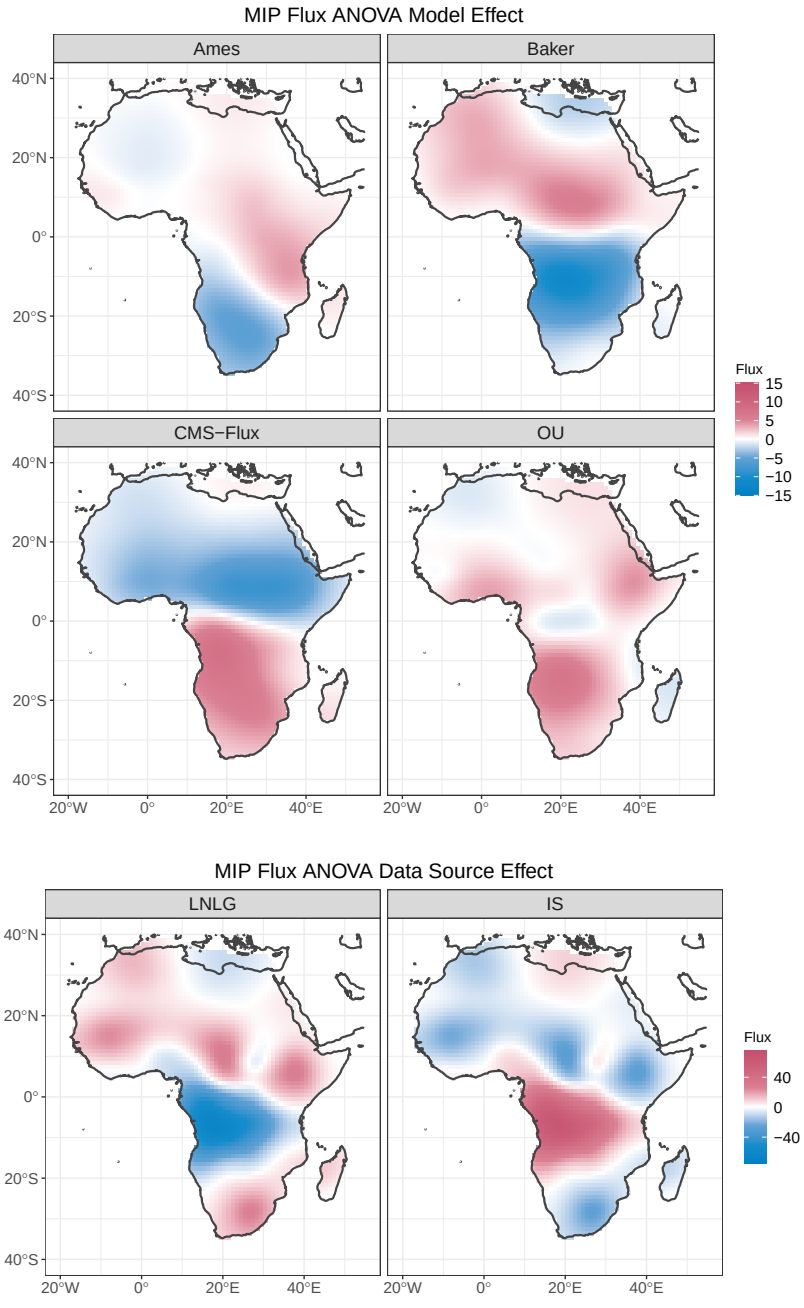

Figure 10. Posterior mean for the functional ANOVA main effect for flux model (α, top) and for data source (β, bottom for JJA 2016 over Africa.)

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions
Open Access
EGU

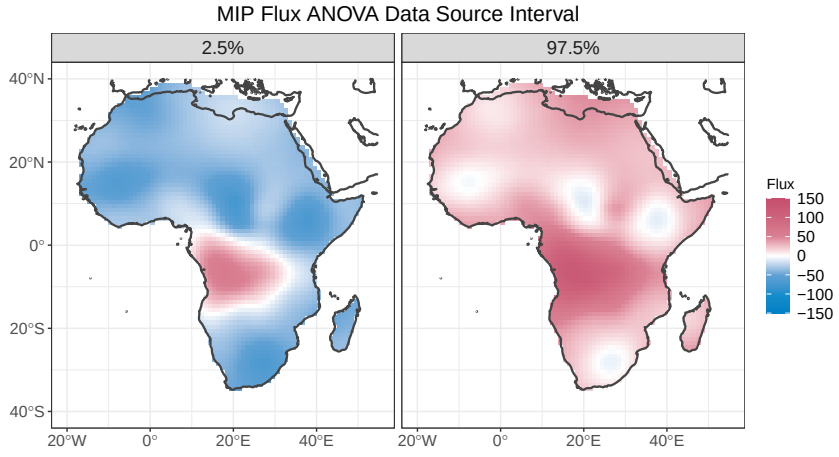

Figure 11. Posterior credible intervals for the functional ANOVA data source effect β^*, the difference between in situ (IS) and OCO-2 land (LNLG) inversions for JJA 2016 over Africa. Left (right) panel depicts the 2.5th (97.5th) percentile of the posterior distribution for each location.

5 Conclusions

Flux inversions produce estimates of the land-ocean-atmosphere exchange of carbon as spatio-temporal fields, providing critical information on the global climate system. These estimates can be variable due to the combinations of factors, such as the atmospheric transport model and the input data source(s), used in the inversions. The functional ANOVA presented and illustrated here provides a statistical model for partitioning variability among these factors while estimating the common signals from the various flux inversions. In addition, the approach accounts for spatial dependence in the flux fields. The extent of spatial dependence is estimated separately for each of the factors considered, along with their interactions. Each of the ANOVA components is represented as a spatial GP using the Vecchia approximation for computational efficiency. This paper has illustrated the functional ANOVA method for flux estimates at continental scale under multiple configurations (Table 2).

The CMS-Flux inversion system was used in a set of inversion experiments to investigate the impact of satellite retrieval aggregation on flux inferences, contrasting a super-obs method with data fusion for aggregating fine-scale OCO-2 retrievals. Overall the aggregation method effect estimated via functional ANOVA was small in magnitude and in its extent of spatial dependence relative to the overall flux anomaly and differences across years. This consistency across aggregation methods bolsters the utility of the fused CO_2 products, which provide an internally consistent summary of the satellite retrievals at a common spatial resolution with a substantial reduction in data volume and complexity.

The functional ANOVA was also implemented for a subset of inversions from the OCO-2 flux MIP over both North America and Africa for JJA 2016. The functional ANOVA identified local consensus flux anomalies for both continents in the presence of variability across inversion systems and atmospheric CO_2 data sources. Over Africa, the data source effect (in situ versus satellite-only) was larger than the flux model effect. This assessment could be a useful diagnostic for understanding the relative

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022


Geoscientific Model Development Discussions

roles of transport model uncertainty and input data challenges such as bias and incomplete sampling. Over both continents, the range of spatial correlation was largest for the model effect, suggesting that model-to-model implementations contribute to differences at large scales, including aggregated regional fluxes (Peiro et al., 2022; Crowell et al., 2019).

The four inversion systems represented in the MIP functional ANOVA use the same inverse method and have similar spatial resolution in their flux solutions. This subset was selected to illustrate the ANOVA, including the Vecchia approximation for GPs, for a factor with more than two levels, where a more complex set of contrasts is employed to preserve the sum-to-zero constraints. This demonstration indicates that the extension to more than two levels per factor is attainable methodologically and computationally. The estimation could be extended to the full collection of inversion systems in the OCO-2 MIP collection. This extension would modestly increase the computational burden of the MCMC, but the intensive operations on the GP precision matrices would still be executed just once per ANOVA component per MCMC iteration, as noted in the supplement. MCMC convergence could be somewhat more challenging with more levels per factor.

ANOVA methods and the resulting estimates can be used to devise potentially unequal weights for combining flux estimates into a consensus flux estimate (Cressie et al., 2022). This weighting is employed with univariate ANOVA when there are different variances, e.g. $\sigma^2_{\alpha,i}$, for all levels of a particular factor. Provided these additional parameters can be estimated reliably from the available data, the weights are taken to be inversely proportional to the level-specific variances. In the functional ANOVA setting in this paper, related extensions for the ANOVA GPs are possible. The GP models used in this paper result in a variance that differs by ANOVA component but is constant across space. Alternative parameterizations with heterogeneous across space could be developed. Kang and Cressie (2013) used a spatial random effects (SRE) model for functional ANOVA components that could result in additional nonstationarity across space.

The current implementation of functional ANOVA for carbon flux estimates has extended related applications to climate models (e.g. Kaufman and Sain, 2010; Sain et al., 2011; Kang and Cressie, 2013) in a number of ways, including the estimation of ANOVA effects for factors with more than two levels and the incorporation of the Vecchia approximation for GPs. The data structure for the examples in this paper differ in some key ways from the previous climate applications as well. These previous studies used multiple years in a climate simulation as replicates to infer a spatially-varying climate signal and model effects in the presence of interannual variability. In these studies, the number of replicates was sizeable, with $n_\epsilon \approx 30$. The carbon flux examples presented here all used $n_\epsilon = 3$, and the ANOVA error term's GP standard deviation, σ_ϵ, is relatively large. This low-replication, high-variance scenario often translates to higher uncertainty in the other ANOVA effects and a tendency for them to shrink to the assumed population mean, zero in this case (Gelman, 2005). Even so, significant anomalies can be inferred in this case. As the satellite CO_2 record, particularly from OCO-2, extends to multiple years, the methodology can be extended to also include replicates across years. The functional ANOVA approach can additionally be modified to analyze groups of time series (Kaufman and Sain, 2010; Cuevas et al., 2004). Finally, extension to spatio-temporal inference is possible with the inclusion of a suitable statistical model and appropriate handling of pseudo-replication. This may require more complex modeling of the error terms in the statistical model. Overall, the functional ANOVA methodology offers suitable flexibility for anomaly detection among discrete collections of Earth system models.

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022


Geoscientific Model Development Discussions

Code and data availability. The OCO-2 V9 MIP surface gridded fluxes used in the examples are available from https://www.esrl.noaa.gov/gmd/ccgg/OCO2_v9mip. R code for processing the flux fields and implementing the functional ANOVA via MCMC is available at https://github.com/co2anomaly/flux_fanova. The repository release used to produce the examples is available at https://doi.org/10.5281/zenodo.7080750. The supporting datasets for the examples, including the MCMC posterior samples, are available at https://doi.org/10.5281/zenodo.7081161

Author contributions. J.H.: Project administration, Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – Original Draft, Writing - review & editing, Visualization. M.K.: Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing. H.N.: Conceptualization, Methodology, Software, Data curation, Writing - review & editing. V.Y.: Project administration, Conceptualization, Methodology, Supervision, Writing - review & editing. J.L.: Conceptualization, Methodology, Formal analysis, Data curation.

Competing interests. The authors declare no competing interests.

Acknowledgements. The research described in this paper was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. Support was provided by the Making Earth System Data Records for Use in Research Environments (MEaSUREs) program. The authors thank David Baker, Amy Braverman, Noel Cressie, Susan Kulawik, and the OCO-2 flux team for helpful discussions.

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Geoscientific Model Development Discussions

References

Baker, D. F., Bell, E., Davis, K. J., Campbell, J. F., Lin, B., and Dobler, J.: A new exponentially-decaying error correlation model for assimilating OCO-2 column-average CO2 data, using a length scale computed from airborne lidar measurements, Geosci. Model Dev., 15, 649–668, https://doi.org/10.5194/gmd-15-649-2022, 2022.

Basu, S., Baker, D. F., Chevallier, F., Patra, P. K., Liu, J., and Miller, J. B.: The impact of transport model differences on CO_2 surface flux estimates from OCO-2 retrievals of column average CO_2, Atmos. Chem. Phys., 18, 7189—-7215, https://doi.org/10.5194/acp-18-7189-2018, 2018.

Byrne, B., Liu, J., Lee, M., Baker, I., Bowman, K. W., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Iraci, L. T., Kiel, M., Kimball, J. S., Miller, C. E., Morino, I., Parazoo, N. C., Petri, C., Roehl, C. M., Sha, M. K., Storng, K., Velazco, V. A., Wennberg, P. O., and Wunch, D.: Improved Constraints on Northern Extratropical CO_2 Fluxes Obtained by Combining Surface-Based and Spaced-Based Atmospheric CO_2 Measurements, J. Geophys. Res.: Atmos., 125, https://doi.org/10.1029/2019JD032029, 2020.

Cressie, N.: Mission CO_2ntrol: a statistical scientist's role in remote sensing of atmospheric carbon dioxide, J. Am. Stat. Assoc., 113, 152–168, https://doi.org/10.1080/01621459.2017.1419136, 2018.

Cressie, N., Bertolacci, M., and Zammit-Mangion, A.: From Many to One: Consensus Inference in a MIP, Geophys. Res. Lett., 49, e2022GL098 277, https://doi.org/10.1029/2022GL098277, 2022.

Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A. R., Chevallier, F., Liu, J., Deng, F., Feng, L., McKain, K., Chatterjee, A., Miller, J. B., Stephens, B. B., Eldering, A., Crisp, D., Schimel, D., Nassar, R., O'Dell, C. W., Oda, T., Sweeney, C., Palmer, P. I., and Jones, D. B. A.: The 2015-2016 carbon cycle as seen from OCO-2 and the global in situ network, Atmos. Chem. Phys., 19, 9797–9831, https://doi.org/10.5194/acp-19-9797-201, 2019.

Cuevas, A., Febrerob, M., and Fraimanc, R.: An ANOVA test for functional data, Comput. Stat. Data Anal., 47, 111–122, https://doi.org/10.1016/j.csda.2003.10.021, 2004.

Eldering, A., O'Dell, C. W., Wennberg, P. O., Crisp, D., Gunson, M., Viatte, C., Avis, C., Braverman, A., Castano, R., Chang, A., et al.: The Orbiting Carbon Observatory-2: First 18 months of science data products, Atmos. Meas. Tech., 10, 549–563, https://doi.org/10.5194/amt-10-549-2017, 2017.

Engelen, R. J., Denning, A. J., Gurney, K. R., and TransCom3: On error estimation in atmospheric CO_2 invsersions, J. Geophys. Res., 107, https://doi.org/10.1029/2002JD002195, 2002.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016, 2016.

Gaubert, B., Stephens, B. B., Basu, S., Chevallier, F., Deng, F., Kort, E. A., Patra, P. K., Peters, W., Rodenbeck, C., Saeki, T., Schimel, D., van der Laan-Luijkx, I., Wofsy, S., and Yin, Y.: Global atmospheric CO_2 invsere models converging on neutral tropical land exchange, but disagreeing on fossil fuel and atmospheric growth rate, Biogeosciences, 16, 117–134, https://doi.org/10.5194/bg-16-117-2019, 2019.

Gelman, A.: Analysis of Variance — Why it is More Important Than Ever, Ann. Stat., 33, 1–33, https://doi.org/10.1214/009053604000001048, 2005.

Gelman, A.: Prior distributions for variance parameters in hierarchical models, Bayesian Anal., 1, 515–534, https://doi.org/10.1214/06-BA117A, 2006.

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B.: Bayesian Data Analysis, Chapman & Hall/CRC, Boca Raton, FL, third edn., 2013.

Johnson, R. A. and Wichern, D. W.: Applied Multivariate Statistical Analysis, Prentice Hall, Upper Saddle River, NJ, fifth edn., 2002.

Kang, E. L. and Cressie, N.: Bayesian Hierarchical ANOVA of Regional Climate-Change Projections from NARCCAP Phase II, Int. J. Appl. Earth Obs. Geoinf., 22, 3–15, https://doi.org/10.1016/j.jag.2011.12.007, 2013.

Katzfuss, M. and Guinness, J.: A general framework for Vecchia approximations of Gaussian processes, Stat. Sci., 36, 124–141, https://doi.org/10.1214/19-STS755, 2021.

Kaufman, C. G. and Sain, S. R.: Bayesian Functional ANOVA Modeling Using Gaussian Process Prior Distributions, Bayesian Anal., 5, 123–150, https://doi.org/10.1214/10-BA505, 2010.

Kiel, M., O'Dell, C. W., Fisher, B., Eldering, A., Nassar, R., MacDonald, C. G., and Wennberg, P. O.: How bias correction goes wrong: Measurement of XCO2 affected by erroneous surface pressure estimates, Atmos. Meas. Tech., 12, 2241–2259, https://doi.org/10.5194/amt-12-2241-2019, 2019.

Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, Appl. Opt., 48, 6716–6733, https://doi.org/10.1364/AO.48.006716, 2009.

Liu, J., Bowman, K. W., Schimel, D. S., Parazoo, N. C., Jiang, Z., Lee, M., et al.: Contrasting carbon cycle responses of the tropical continents to the 2015-2016 El Niño, Science, 358, 1–7, https://doi.org/10.1126/science.aam5690, 2017.

Nguyen, H., Cressie, N., and Braverman, A.: Multivariate spatial data fusion for very large remote sensing datasets, Remote Sens., 9, https://doi.org/10.3390/rs9020142, 2017.

Nguyen, H., Liu, J., Kulawik, S., Baker, D., Hobbs, J., Braverman, A., Katzfuss, M., and Yadav, V.: MEASURES 2017 Data Fusion (v2) Algorithm Theoretical Basis Document, URL: https://docserver.gesdisc.eosdis.nasa.gov/public/project/MEaSUREs/XCO2_Data_Fusion/MEASURES17_ATBD_dataVersion2.pdf, 2020.

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., et al.: Improved Retrievals of Carbon Dioxide from the Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11, 6539–6576, https://doi.org/10.5194/amt-11-6539-2018, 2018.

Peiro, H., Crowell, S., Schuh, A., Baker, D. F., O'Dell, C., Jacobson, A. R., Chevallier, F., Liu, J., Eldering, A., Crisp, D., Deng, F., Weir, B., Basu, S., Johnson, M. S., Philip, S., and Baker, I.: Four years of global carbon cycle observed from OCO-2 version 9 and in situ data, and comparison to OCO-2 v7, Atmos. Chem. Phys., 22, 1097–1130, https://doi.org/10.5194/acp-22-1097-2022, 2022.

Sain, S. R., Nychka, D., and Mearns, L.: Functional ANOVA and regional climate experiments: a statistical analysis of dynamic downscaling, Environmetrics, 22, 700–711, https://doi.org/10.1002/env.1068, 2011.

Schäfer, F., Katzfuss, M., and Owhadi, H.: Sparse Cholesky factorization by Kullback-Leibler minimization, SIAM J. Sci. Comput., 43, A2019–A2046, https://doi.org/10.1137/20M1336254, 2021.

Stein, M. L.: Interpolation of Spatial Data: Some Theory for Kriging, Springer, New York, NY, 1999.

Thompson, R. L., Patra, P. K., Chevallier, F., Maksyutov, S., Law, R. M., van der Laan-Luijkx, I. T., Peters, W., Ganshin, A., Zhuravlev, R., Maki, T., Nakamura, T., Shirai, T., Ishizawa, M., Saeki, T., Machida, T., Poulter, B., Canadell, J. G., and Ciais, P.: Top-down assessment of the Asian carbon budget since the mid 1990s, Nat. Commun., 7, https://doi.org/10.1038/ncomms10724, 2016.

https://doi.org/10.5194/gmd-2022-230
Preprint. Discussion started: 14 November 2022

Torres, A. D., Keppel-Aleks, G., Doney, S. C., Fendrock, M., Luis, K., Maziére, M. D., et al.: A Geostatistical Framework for Quantifying the Imprint of Mesoscale Atmospheric Transport on Satellite Trace Gas Retrievals, J. Geophys. Res.: Atmos., 124, 1–23, https://doi.org/10.1029/2018JD029933, 2019.

Vecchia, A.: Estimation and model identification for continuous spatial processes, J. R. Stat. Soc. B, 50, 297–312, 1988.

Worden, J. R., Doran, G., Kulawik, S., Eldering, A., Crisp, D., Frankenberg, C., et al.: Evaluation and Attribution of OCO-2 XCO2 Uncertainties, Atmos. Meas. Tech., 10, 2759–2771, https://doi.org/10.5194/amt-10-2759-2017, 2017.