Spatio-temporal approach to moving window block kriging of satellite data v1.0

Jovan M. Tadić[1], Xuemei Qiu[1], Scot Miller[1] and Anna M. Michalak[1]

[1]{Department of Global Ecology, Carnegie Institution for Science, Stanford, CA 94305, USA}

Abstract. Numerous existing satellites observe physical or environmental properties of the Earth system. Many of these satellites provide global-scale observations, but these observations are often sparse and noisy. By contrast, contiguous, global maps are often most useful to the scientific community (i.e., level 3 products). We develop a spatiotemporal moving window block kriging method to create contiguous maps from sparse and/or noisy satellite observations. This approach exhibits several advantages over existing methods: 1) it allows for flexibility in setting the spatial resolution of the level 3 map, 2) it is applicable to observations with variable density, 3) it produces a rigorous uncertainty estimate, 4) it exploits both spatial and temporal correlations in the data, and 5) it facilitates estimation in real time. Moreover, this approach only requires ~~a limited number of assumptions~~ the assumption that the observable quantity exhibits spatial and temporal correlations that are inferable from the data. We test this method by creating Level 3 products from satellite observations of CO_2 (XCO_2) from the Greenhouse Gases Observing Satellite (GOSAT), CH_4 (XCH_4) from the Infrared Atmospheric Sounding Interferometer (IASI) and solar-induced chlorophyll fluorescence (SIF) from the Global Ozone Monitoring Experiment–2 (GOME-2). We evaluate and analyze the difference in performance of spatio-temporal vs. recently developed spatial kriging methods.

1. Introduction

Satellite observations of the Earth's surface and atmosphere provide a valuable window into the functioning of the Earth system. Satellites often provide global observations, but these observations are rarely uniform or contiguous in space/time. The observations can be non-contiguous due to satellite orbit geometries and periods, geophysical limitations (e.g. cloud cover), and temporary instrument malfunctions. Furthermore, satellites may provide a large quantity of data, but individual observations can have a large noise-to-signal ratio. It is often necessary to spatially interpolate the data in order to organize the data onto a regular grid, query the data at a particular location of interest, estimate data at unsampled times and/or locations, and/or map the underlying signal in a noisy dataset. These gridded, interpolated maps are commonly named "Level 3" data (e.g. NASA, 2014) and are often part of the standard suite of satellite data products.

CO_2 column observations (XCO_2) from the Greenhouse Gases Observing Satellite (GOSAT), CH_4 column observations (XCH_4) from the Infrared Atmospheric Sounding Interferometer (IASI) and solar-induced chlorophyll fluorescence (SIF) observations from The Global Ozone Monitoring Experiment–2 (GOME-2) provide prototypical examples of these challenges, and these three satellites are the primary application used throughout this work (see Section 3).

The most commonly-used method for creating Level 3 maps from satellite data is binning. This approach involves taking the mean of all observations within a given grid cell or "bin" (see Kulawik et al., 2010, and Crévoisier et al., 2009 for examples). The binning method, however, has a number of shortfalls that can lead to inconsistent or inaccurate results. First, different bins contain variable numbers of observations. As a result, some bins will be well-constrained by the data while others may be based upon sparse, noisy

41 observations. Second, binning does not produce uncertainty estimates. Third, this method cannot
42 extrapolate the unknown quantity to bins without any observations.

43 A broad class of geostatistical methods known as kriging provides an alternative approach to mapping
44 satellite observations. Kriging is a best linear unbiased estimator (for kriging see Chiles and Delfiner, 2012),
45 where covariance functions are used to represent correlations among data. As a result, kriging can account
46 for a variable density of observations and can estimate uncertainties in the resulting maps. Various forms
47 of kriging have recently been used to map satellite Earth observations, particularly for XCO_2 (e.g.,
48 Hammerling et al. 2012a,b; Tadić et al., 2015; Zeng et al., 2013; Guo et al., 2013, Zeng et al., 2016).
49 Hammerling et al. (2012a,b) presented an approach to map Orbiting Carbon Observatory-2 (OCO-2) and
50 GOSAT XCO_2 observations, respectively, with non-stationary properties. In our previous study (Tadić et
51 al., 2015) we extended that approach to create XCO_2 maps that can have a different spatial resolution from
52 the resolution or footprint of the original satellite observations. Our previous study and those of
53 Hammerling et al. (2012a,b) accounted for spatial covariances among observations but did not include a
54 temporal component. The present study extends this geostatistical framework from a purely spatial to a
55 spatiotemporal domain.

56 Spatiotemporal approaches to interpolation can provide a number of advantages relative to purely spatial
57 methods (e.g. Zeng et al., 2016; Guo et al., 2013). A purely spatial approach will usually aggregate
58 observations into temporal blocks; observations within the same block effectively have the same time stamp
59 whether or not those observations are actually synchronous (e.g., Tadić et al., 2015; Hammerling et al.,
60 2012a,b). Any real temporal variability within a block becomes noise. A spatiotemporal approach, by
61 contrast, treats time as an explicit dimension and models covariances among data as a function of time. ~~As~~
62 ~~a result, the spatiotemporal approach can (1) fill in temporal gaps in the observations, (2) create maps at~~
63 ~~higher temporal resolutions than purely spatial approach, (3) produce more accurate estimates when~~
64 ~~observations have variable spatio-temporal coverage, (4) predict future values (i.e. extrapolate temporally).~~

65 A handful of recent studies have considered temporal relationships when mapping satellite observations of
66 XCO_2. These studies have either used various forms of Kalman smoothing (e.g., Katzfuss and Cressie 2011,
67 Katzfuss and Cressie 2012, Nguyen et al. 2014) or geostatistics (e.g., Guo et al. 2013; Zeng et al. 2013;
68 Zeng et al. 2016). The former group of studies leverages Kalman smoothing to improve the computational
69 tractability of mapping dense or abundant datasets, like OCO-2 and the Atmospheric Infrared Sounder
70 (AIRS). The latter group of studies, by contrast, has applied geostatistics to sparse datasets like those from
71 the GOSAT satellite. A detailed review of spatial and spatio-temporal mapping methods has been published
72 recently (Li and Heap, 2014).

73 The ~~model developed in~~ goal of this ~~paper also uses geostatistics~~study is to ~~map~~develop a geostatistical
74 spatio-temporal mapping and upscaling method (applicable, but not limited to,) satellite observations of
75 XCO_2~~, but we present~~) that exhibits several advances relative to previous ~~efforts.~~methods. It can: (1) fill in
76 temporal gaps in the observations, (2) create maps at higher temporal resolutions than purely spatial
77 approach, (3) produce more accurate estimates when observations have variable spatio-temporal coverage,
78 (4) predict future values (i.e. extrapolate temporally). Among other improvements, we develop an efficient
79 method ~~to subsample~~for subsampling satellite observations and utilize the product-sum covariance model
80 (e.g., De Iaco et al., 2001) that is easy to parameterize, which makes it applicable to both ~~abundant~~dense
81 and sparse datasets. The entire work has been conducted in Matlab 2012a.

82 Section 2 of this study describes the presented model in detail; it describes an efficient subsampling
83 procedure that can handle very large datasets and a covariance model that can estimate both spatial and
84 temporal relationships in the data. We then incorporate these components into a spatiotemporal version of
85 moving window block kriging. In sections 3 and 4, we subsequently apply this model to map GOSAT
86 XCO_2, IASI XCH_4 and GOME-2 SIF at multiple time resolutions (including daily).

2

2. Methods

The spatio-temporal block kriging approach presented in this study proceeds in three steps for each model grid cell and estimation time. First, we subsample the observations within a predetermined spatio-temporal domain (section 2.1). Next, we characterize the local spatio-temporal covariance structure (section 2.2). Finally, we interpolate the satellite observations at the desired spatial resolution (section 2.3).

2.1 Subsampling of observations

The ultimate goal of the proposed subsampling strategy is to reduce the number of observations in the spatio-temporal vicinity of an estimation location to a representative, computationally feasible subset of data. We use a subset of observations (M) to estimate a local set of covariance parameters and use another subset (N) to estimate the desired quantity and associated uncertainty. Note that, for the method presented here, M and N can refer to either the same subset of data or different subsets.

The total number of observations used for covariance parameter estimation (M), is selected to be small enough to make this estimation computationally feasible but large enough to yield a sample representative of both local and regional variability. The optimal subset of N observations used for mapping depends on the actually observed covariance structure which is not known prior to covariance parametrization step. In the example presented in Sect. 3, the optimal observational subset used in a mapping step for each grid cell comprised N points having the highest covariance with the estimation location. In the example below, we set both M and N at 500; larger values of M and N did not have a substantial impact on the estimated parameters and mapped quantity, respectively. ~~Furthermore, *M* should represent local variability, and larger values of *M* would encompass more distant, non-local regions.~~

We select subset of observations M for each estimation grid cell by assigning a relative selection probability to each observation based on that observation's spatial and temporal 'separation distances' from the centroid of the grid cell. In the absence of a proper metric for distance in space-time, we model the spatial and temporal components of the overall selection probability separately.

The selection probability (and its components) is described by the following equation:

$$P = P_s \times P_t \propto 1/(A_s h_s)^2 \times e^{-(A_t h_t)^2} \tag{1}$$

where P_s is the spatial component of the relative probability of a given observation being selected, P_t is temporal component, h_s and h_t are distances between estimation location and observations, in space and time, respectively, and A_s and A_t are unit dependent, user defined weighting factors between separation distance in space vs. in time (how deep in space vs. time the sampling should occur). The unit dependent choice of A_s and A_t can be initially based on user expectations of the decorrelation distances in space vs. time and, if necessary, subsequently corrected accounting for actually computed decorrelation lengths in space and time in an iterative fashion. In this way temporal and spatial sampling depths could even be locally optimized and become location-specific. In the examples below (Section 3), A_s and A_t were set to 1 km^{-1}, and 0.5 day^{-1}, respectively, based on the observed average decorrelation distances in space and time (see Fig. 1 and Section 4.1).

[Figure 1]

h_s is calculated as the great circle distance between the centroid x_j of the estimation grid cell and the location x_i of an observation:

3

127 $$h_s(x_i, x_j) = r\cos^{-1}(\sin\varphi_i \sin\varphi_j + \cos\varphi_i \cos\varphi_j \cos(\lambda_i - \lambda_j)) \qquad (2)$$

128 where φ_i and λ_i are the latitude and longitude of location x_i and r is the radius of the Earth.

129 The temporal and spatial components of the probability function have different functional forms out of
130 necessity. The measurements often come pre-aggregated in time slices corresponding to hours, days, or
131 longer aggregation time periods, which multiplies the number of observations with the same time stamp.
132 As a result, it is not possible to assign sampling probability along a temporal axis in a manner equivalent to
133 the spatial approach; doing so would result in infinite probabilities assigned to all observations within the
134 time slice of the actual estimation location ($P_t \sim 1/0^2 = \infty$). The same holds for spatially co-located
135 observations. However, since each observation comes with unique spatial coordinates (not pre-binned like
136 in temporal case), we select a simpler form of the spatial component of the sampling function. The defined
137 form of P (Eq. 1) ensures that pairs of observations close to estimation location define the shape of the
138 variogram at short separation distances (the variogram should reflect variability in the spatio-temporal
139 vicinity of the estimation grid cell. See Section 2.2). Different forms of P can be used if directional
140 anisotropy is expected or if more/fewer observations along a given direction are desired to better represent
141 expected correlations.

142 Previous approaches required the user to choose spatial and temporal windows that determine which
143 neighboring observations to use (see, for comparison, Alkhaled et al. 2008; Hammerling et al. 2012a,b).
144 The approach proposed in this paper, by contrast, requires fewer subjective choices – only the form of
145 sampling function and unit dependent choice of normalizing coefficients A_s and A_t. In addition, our
146 approach is computationally feasible even for very large data sets.

2.2 Characterization of Spatio-temporal Covariance

148 Existing studies have used a number of models to estimate spatio-temporal covariances for a variety of
149 applications. Models used include the metric model (Dimitrakopoulos and Luo, 1994), linear model
150 (Rouhani and Hall, 1989), product model (De Cesare et al., 1996), non-separable model (Cressie and
151 Huang, 1999), and generalized product-sum covariance model (De Iaco et al., 2001). The approach
152 developed in this paper uses a generalized product-sum covariance model (De Iaco et al., 2001). This model
153 affords a number of advantages relative to other covariance models: (1) a product sum covariance model
154 outperformed other models in terms of prediction accuracy in a recent study using GOSAT satellite data
155 (Guo et al., 2013), (2) it is relatively easy to implement (De Iaco et al., 2001), and (3) it is more flexible
156 than a non-separable covariance model (De Cesare, 2001a).

157 The product-sum model, as it has been applied in the past, has one important area for improvement. The
158 original procedure (De Iaco et al., 2001) assumed separate modeling of the spatial and temporal covariance
159 (variograms) and their later unification into a spatio-temporal model in the final step. The procedure
160 requires observations approximately in the same location at multiple different times. However, satellite
161 observations are often not perfectly collocated in consequent measurement cycles over the same region. As
162 a result, we would need to assume that each measurement cycle is perfectly co-located with previous/future
163 cycles, or define an arbitrary tolerance, in order to apply the original approach. This assumption becomes
164 more prone to error if the observations are very sparse, as is often the case with satellites.

165 Thus, in this study, we cater to specific properties of satellite data and alter the original procedure by
166 estimating all covariance parameters simultaneously, thereby avoiding the aforementioned problem.

167 We broadly define the covariance as follows:

4

168
$$C_{s,t}(h_s,h_t) = \text{Cov}(Z(s_+h_s,t_+h_t), Z(s,t)) \tag{3}$$

169 The equation shows that covariance between two points (Z) separated in space-time (s,t) depends on their
170 distance in space (h_s) and distance in time (h_t). The following class of valid product–sum covariance models
171 was introduced in De Cesare et al. (2001b) and further developed in De Iaco et al. (2001):

172
$$C_{s,t}(h_s,h_t)= k_1 C_s(h_s)C_t(h_t)+ k_2 C_s(h_s)+ k_3 C_t(h_t) \tag{4}$$

173 where C_t and C_s are valid temporal and spatial covariance models, respectively. De Iaco et al. (2001) proved
174 that for positive definiteness it is sufficient that $k_1 > 0$, $k_2 \geq 0$ and $k_3 \geq 0$. It is interesting to note that from
175 Eq. 4 follows that spatio-temporal covariance models collapses down to purely spatial model in cases where
176 temporal covariance does not exist. Thus, the spatial approach could be viewed as a special case of spatio-
177 temporal modeling.

178 The model in Eq. 4 corresponds to the spatio-temporal variogram shown in Equation 5. In the original
179 procedure, De Iaco et al., 2001 estimated separate spatial (h_t=0) and temporal (h_s=0) variograms using the
180 data. De Iaco et al., 2001 ~~then~~ then combined these models to obtain the final spatio-temporal variogram
181 model:

182
$$\gamma_{s,t}(h_s,h_t) = \gamma_{s,t}(h_s,0) + \gamma_{s,t}(0,h_t) - k\gamma_{s,t}(h_s,0)\gamma_{s,t}(0,h_t) \tag{5}$$

183 where $\gamma_{s,t}(h_s,0)$ and $\gamma_{s,t}(0,h_t)$ are spatio-temporal variograms for h_t=0 and h_s=0, respectively (Figure 2).
184 Parameter k is estimated from the data which makes the model easily applicable:

185
$$k = \frac{k_s C_s(0)+ k_t C_t(0)- C_{s,t}(0,0)}{k_s C_s(0)k_t C_t(0)} \tag{6}$$

186 where $k_s C_s(0)$ and $k_t C_t(0)$ are spatial and temporal sills (variances) obtained in modeling of separate
187 spatial and temporal variograms. The only condition k has to fulfill in order to create an admissible
188 covariance model is

189
$$0 < k \leq \frac{1}{max\{\sigma_s^2(\gamma_{s,t}(h_s,0)); \sigma_t^2(\gamma_{s,t}(0,h_t))\}} \tag{7}$$

190 Due to the specifics of satellite data, we estimate both the covariance parameters and parameter k
191 simultaneously. This approach accounts for constraints that assure a positive definiteness of the model (De
192 Iaco et al., 2001). This simultaneous approach makes the model more applicable to sparse data and data
193 with variable spatial coverage, as is often the case with satellite observations.

194 We use a Gaussian variogram function with a nugget effect to model temporal covariance in the example
195 presented here (for an overview of variogram models see Chiles and Delfiner, 2012). We use an exponential
196 model for the spatial variogram. In both cases, we make this choice based upon visual inspection of local
197 variograms at multiple estimation locations:

198
$$\gamma_t(h_t)(Gaussian) = \begin{cases} 0, \ for \ h_t = 0 \\ \sigma_t^2(1 - \exp\left(-\frac{h_t^2}{l_t^2}\right) + \sigma_{nug}^2, \ for \ h_t > 0 \end{cases} \tag{8}$$

199
$$\gamma_s(h_s)(exponential) = \begin{cases} 0, \ for \ h_s = 0 \\ \sigma_s^2(1 - \exp\left(-\frac{h_s}{l_s}\right) + \sigma_{nug}^2, \ for \ h_s > 0 \end{cases} \tag{9}$$

5

200 where σ^2 and l are the variance and correlation length of the quantity being mapped, and σ^2_{nug} is the nugget
201 variance, typically representative of measurement and retrieval errors in the case of satellite observations.

202 [Figure 2]

203 Unlike the original procedure in De Iaco et al. (2001), we model the variogram using only two steps. First,
204 we calculate a raw spatio-temporal variogram based on the subsampled observations for each estimation
205 grid cell:

$$\gamma(h_s, h_t) = \frac{1}{2}[y(x_i) - y(x_j)]^2 \tag{10}$$

207 where γ is the raw spatio-temporal variogram value for a given pair of observations $y(x_i)$ and $y(\underline{x}_j)$, and h_s
208 and h_t are, respectively, the great circle distance and temporal distance between the spatio-temporal
209 locations (x_i and x_j) of these observations.

210 Second, we fit the theoretical variogram defined in Eq. 5 to the raw variogram using non-linear least
211 squares. We subsequently calculate the spatiotemporal covariance using the following equation:

$$C_{s,t}(h_s, h_t) = C_{s,t}(0,0) - \gamma_{s,t}(h_s, h_t)) \tag{11}$$

213 **Validity on the sphere.** Most covariance models were originally designed for Euclidean space, and their
214 validity in other coordinate systems cannot be assumed *per se*. Huang et al. (2011) examined the validity
215 of several theoretical covariance models in spherical coordinate systems. However, this evaluation has not
216 been done for the spatio-temporal product-sum covariance model. Other studies that use a product-sum
217 covariance model typically assume the validity of this covariance model on a sphere (e.g., Zeng et al., 2013;
218 Zeng et al., 2016). Results from Huang et al. (2011) explicitly validate the exponential covariance model
219 on a sphere, as well as sums of the products of exponential covariance models and constants (provided that
220 the constants are positive). The first term of the product-sum covariance model used in this study (Eq. 4)
221 represents a Hadamard product (Million, 2007) of two positive definite matrices. According to Schur
222 product theorem, a Hadamard product of two positive definite matrices necessarily gives a positive definite
223 matrix (Mathias, 1993). It therefore follows that a generalized product-sum model (Equation 4) is valid on
224 a sphere if its spatial component is valid on a sphere.

225 **2.3 Mapping using spatio-temporal moving window block kriging**

226 This section leverages the sampling function (Sect. 2.1) and the product-sum covariance model (Sect. 2.2)
227 to implement a spatio-temporal version of moving window block kriging. A primary advantage of block
228 kriging is its ability to estimate contiguous maps at any spatial resolution equal to or coarser than the spatial
229 support (i.e. footprint size) of observations (refer to Sect. 1 and Tadić et. al. 2015). Unlike ordinary kriging
230 method, the spatial support in block kriging corresponds to the average value within each chosen grid cell.

231 Moving window block kriging requires solving a set of linear equations to obtain a set of weights (λ). These
232 weights must be estimated for each prediction location using N associated observations:

$$\begin{bmatrix} \mathbf{Q} + \mathbf{R} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ -\nu \end{bmatrix} = \begin{bmatrix} \mathbf{q_A} \\ 1 \end{bmatrix} \tag{12}$$

234 In this equation, $\mathbf{R}$ is a diagonal $N \times N$ nugget covariance matrix that describes measurement and retrieval
235 errors, $\mathbf{Q}$ is a $N \times N$ covariance matrix among the N observations with individual entries as defined in Eqn.
236 11, $\mathbf{1}$ is an $N \times 1$ unity vector, ν is a Lagrange multiplier, and $\mathbf{q}_A$ is an $N \times 1$ vector of the spatio-temporal
237 covariances between the N observation locations and the estimation grid cell, defined as:

6

$$238 \qquad q_{A,i} = \frac{1}{n}\sum_{j=1}^{n} q\left(h_{s_{i,j}}, h_{t_{i,j}}\right) \qquad (13)$$

239 where $q_{A,i}$ is the covariance between the grid cell and observation i. $q(h_{i,j})$ is defined as $C_{s,t}$ in Eqn. 11
240 based on the distances $h_{s_{i,j}}$ and $h_{t_{i,j}}$ between observation i and n regularly-spaced locations within the grid
241 cell. In the context of satellite measurements, n is a highest number of non-overlapping footprints contained
242 within a grid cell and was calculated based on the relative size of the satellite footprint compared to the
243 size of the estimation grid cells. n varies with latitude, as the size of grid cells decreases with the distance
244 from the equator. The system in Eqn. 12 is solved for the weights (λ) and the Lagrange multiplier (v). We
245 subsequently use these parameters to define the estimate ($\hat{z}$) and estimation uncertainty ($\sigma^2_{\hat{z}}$) for the grid
246 cell:

$$247 \qquad \hat{z} = \lambda^T \mathbf{y} \qquad (14)$$

$$248 \qquad \sigma^2_{\hat{z}} = \sigma_{AA} - \lambda^T \mathbf{q}_A + v \qquad (15)$$

249 where $\mathbf{y}$ is the $N\times 1$ vector of subsampled observations, and σ_{AA} is the variance of the observations at the
250 resolution of the estimation grid cell, defined as:

$$251 \qquad \sigma_{AA} = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} q\left(h_{j,k}\right) \qquad (16)$$

252 In that equation, $q\left(h_{s_{i,j}}, h_{t_{i,j}}\right)$ is defined as $C_{s,t}$ in Eqn. 11 based on the distances $h_{s_{i,j}}$ and $h_{t_{i,j}}$ between
253 any combination of the n regularly spaced locations within the grid cell defined previously.

254   3. Example applications

255 We select three case studies of satellite Level 2 data to demonstrate the properties of the method developed
256 in this paper: column-integrated dry air model fraction of CO_2 (XCO_2) from the Japanese Greenhouse Gas
257 Observing SATellite (GOSAT), CH_4 (XCH_4) from the Infrared Atmospheric Sounding Interferometer
258 (IASI), and solar-induced fluorescence (SIF) the Global Ozone Monitoring Experiment–2 (GOME-2).
259 Level 2 datasets from GOSAT, IASI and GOME-2 have relatively different characteristics. For example,
260 GOSAT observations are sparse while IASI and GOME-2 are abundant. These diverse datasets are therefore
261 ideal for testing the method developed here.

262 The method was demonstrated by producing two different sets of maps. First, it was applied at resolutions
263 coarser than native (1 × 1°, 2.5 × 2°, and 1 × 1° for GOSAT, IASI and GOME-2, respectively) to
264 demonstrate block kriging capabilities of the method (Section 3). Second, it was applied at the native
265 resolution of the satellites for cross-validation (method evaluation) purposes (Section 4).

266   3.1 Total column CO_2 (XCO_2) observed by GOSAT

267 The Japanese Greenhouse Gas Observing SATellite (GOSAT) (e.g., Kuze et al., 2009), the first satellite
268 dedicated to global greenhouse gas monitoring, was launched in 2009. Basic information about the satellite,
269 its orbit configuration, and the CO_2 column observations are given in our previous study (Tadić et al., 2014).
270 It flies in a polar, sun-synchronous orbit with a 3-day repeat cycle and an approximate 13:00 LT overpass
271 time. GOSAT has a nadir footprint of about 10.5 km diameter at sea level (Kuze et al., 2009) and 2×10^3

272   observations per week. The XCO_2 observations from GOSAT have large retrieval uncertainties (e.g.,
273   O'Dell et al. 2012) and exhibit large spatial and temporal gaps (e.g., Fig. 3a). Although these XCO_2
274   observations are sparse and noisy, contiguous Level 3 maps are often desirable for environmental and
275   ecological applications. (Maksyutov et al., 2013; Liu et al., 2012). To this end, we generate global daily
276   estimates for XCO_2 (August 2-7, 2009) to match the timeframe used in Tadić et al., 2014.

277   [Figure 3]

278   We obtain bias-corrected and filtered GOSAT Level 2 observations using NASA's Atmospheric CO_2
279   Observations from Space (ACOS) algorithm v3.4 release 3 (e.g., O'Dell et al., 2012; Crisp et al., 2012). In
280   this study, we use spatio-temporal moving window block kriging to create a series of contiguous, in-filled
281   global daily maps and associated uncertainties for 2-7 August 2009 (two repeat cycles) (Fig. 3a-c) at 1×1°
282   resolution. We select the time period to match the time period from our previous study (Tadić et al., 2014).
283   Unlike results from our previous study and other similar studies, which created estimates at 6-day or longer
284   time periods (Hammerling et al., 2012a), we leverage the method developed here to produce maps at the
285   daily scale.

3.2 Total column CH_4 (XCH_4) observed by IASI

287   The Infrared Atmospheric Sounding Interferometer (IASI) developed by the Centre National d'Etudes
288   Spatiales (CNES) in collaboration with the European Organisation for the Exploitation of Meteorological
289   Satellites (EUMETSAT) is a Fourier Transform Spectrometer based on a Michelson Interferometer coupled
290   to an integrated imaging system that measures infrared radiation emitted from the Earth. It is carried by
291   MetOp-A, a sun-synchronous polar orbit satellite which flows at an altitude of 817 km. Detailed information
292   about the IASI instrument could be found elsewhere (Crévoisier et al., 2009a,b; Massart et al., 2014). IASI
293   has an instantaneous field of view of 50×50 km, composed of four pixels each 12 km in radius, delivering
294   $\sim56\times10^3$ XCH_4 observations per week.

295   [Figure 4]

296   Methane Level 2 IASI (0-4 km) data were retrieved at the NOAA/NESDIS using the NUCAPS (NOAA
297   Unique CrIS/ATMS Processing System) algorithm (Gambacorta, 2013; Xiong et al., 2013). For the ice-
298   covered ocean the data for the lower troposphere (0-4 km) are unreliable due to insufficient thermal contrast
299   between the surface and the atmosphere. Filtering parameters have been provided by Xiong (2014, private
300   communication). The data are available at http://www.nsof.class.noaa.gov/. Using the new method, we
301   created a series of contiguous global daily maps and associated uncertainties for the Northern Hemisphere,
302   for February 26-March 4, 2013 (i.e. Figure 4a-c) at $1^o\times1^o$ resolution. We chose this time period to match
303   the occurrence of the methane "anomaly" North of the coast of Scandinavia.

3.3 Global land solar-induced fluorescence fields observed by GOME-2

305   The GOME-2 (The Global Ozone Monitoring Experiment–2) instrument on board METOP-A (e.g., Joiner
306   et al., 2013) observes solar-induced fluorescence (SIF). The GOME-2 spatial footprint (i.e. support) of the
307   observations is 40 km $\times$ 80 km (Joiner et al, 2013), and the volume of available data is approximately 2×10^5
308   SIF observations per week.

8

309   [Figure 5]

310   Multiple recent studies have demonstrated the potential use of satellite observations of solar-induced
311   fluorescence (SIF) for understanding the photosynthetic CO_2 uptake at large scales (Joiner et al., 2011;
312   Joiner et al., 2012; Joiner et al., 2013; Frankenberg et al., 2011; Frankenberg et al., 2012; Guanter et al.,
313   2012, Lee et al., 2013; Frankenberg et al., 2014). Satellite SIF measurements can be used with land surface
314   models to understand GPP response to environmental stress (e.g., Lee et al., 2013) and to improve the
315   representation of GPP. GOME-2 provides the highest spatial and temporal density of data, among all
316   available datasets.

317   In the example presented here we use SIF GOME-2 v.14 data (Joiner et al., 2013) with the approach
318   described in Section 2 to create contiguous maps of SIF at a single spatial resolution ($1° \times 1°$) and daily
319   temporal resolutions. Maps of SIF and associated uncertainties are created at daily temporal resolutions
320   covering 5-14 May, 2012, some of which are shown on Figures 5a-c.

4. Method evaluation: accuracy, precision and bias

4.1 Accuracy, precision and bias

323   We use a leave-one-out cross validation technique to assess the performance of spatio-temporal (ST) versus
324   spatial moving window block kriging. We produce these estimates at the native resolution of GOSAT, IASI
325   and GOME-2 satellites/instruments, which allowed a direct comparison to measured values. For IASI and
326   GOME-2, for each day in February 26-March 4, 2013, and May 5-14, 2012, respectively, 10% of available
327   observational data were randomly selected for use in leave-one-out cross-validation and their coordinates
328   extracted. For XCO_2, all GOSAT XCO_2 observations for each day in August 2-7, 2009, were used. We
329   assess the accuracy (the difference between estimates and withheld observations) of both methods using
330   two common measures: (1) Mean Absolute Error (MAE), and (2) Root Mean Squared Error (RMSE). ~~We~~
331   ~~also~~We also use two more recently proposed measures (Li and Heap, 2011; Li, 2016) that remove the effect
332   of unit/scal. The first is relative mean absolute error (RMAE) that is given as:

$$RMAE = \frac{1}{n}\sum_{i=1}^{n}|(\hat{z}_i - y_i)/o_i| \times 100 \quad\quad\quad (17)$$

336   and the second is relative root mean square error (RRMSE), as follows:

$$RRMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(|y_i - \hat{z}_i|/y_i)^2\right]^{1/2} \times 100 \quad\quad\quad (18)$$

339   where n is the number of observations or samples, o is observed value, and p is predicted or estimated value.

340   We assess the performance of each method using two additional measures: (3) the accuracy of the
341   uncertainty bounds (the degree to which the reported uncertainties capture the difference between estimates
342   and withheld observations) and (4) bias (the mean difference between estimates and withheld observations).

343   We parameterize the temporal component of the spatio-temporal sampling function in such a way
344   that observations located +/- 3 days from the actual date had 10% probability of being sampled

9

compared to observations from the actual day (see Fig 1a). We compare the results to spatial kriging estimates obtained in two different ways, based on observations only from the actual day (1d) and based on observations from +/-3 days from the actual day (7d). This latter case is analogous to the +/- 3 -day window that we use for the ST approach. In this 7d case, we obtain these spatial kriging results by assuming the entire observational dataset collected within the selected time period (actual day +/- 3 days) is perfectly temporally correlated. In other words, we use all observations as though they were collected at the same time. We then produce estimates at locations of observations collected within the selected timeframe and compare the performance of the two methods. We repeat procedure described in Section 2 for every observation selected for cross-validation, and we average the statistics, displayed in Table 1.

[Table 1]

According to the results, the spatio-temporal approach performs better than the spatial (7d) approach in all three cases and in all performance measures (for example, spatial (7d) MAE was 6-10% larger). The comparison clearly shows that proper characterization of the temporal covariance between two points residing in different time periods (days), embedded into spatio-temporal approach, improves kriging performance. In IASI ~~case~~and GOME-2 cases, the spatio-temporal method also performed better than spatial (1d). However, in case of GOSAT ~~and GOME 2~~ data, spatio-temporal approach slightly underperformed the spatial (1d) approach having 12% higher MAE (please see Section 4.2 for discussion).

We observed that RMAE and RRMSE error measures should be used with caution in cases when observations can take real zero values, like in the GOME-2 case. In such cases the division by close-to-zero values result in extremely high RMAE and RRMSE values, which overall limits the applicability of these error measures.

We evaluate the accuracy of the uncertainty bounds by examining how often those bounds encapsulate withheld observations. The percentage of observations that fall outside the uncertainty bounds in spatio-temporal approach is comparable to that of the spatial method, confirming the accuracy of the estimated uncertainty bounds (for normally-distributed data the percentage of observations that fall outside of the one, two, and three ~~estimation~~estimations standard deviation ($\sigma_{\hat{z}}$) uncertainty bounds should be 32%, 5% and 0.3%, respectively). The fraction of observations that fall outside the uncertainty bound is generally lower than would be expected for normally-distributed data, and our results may indicate non-normal features in the data.

4.2 When is spatio-temporal modeling recommended?

A ST approach can afford advantages over purely spatial methods when temporal data correlations and data coverage are strong. Indeed, in many cases, the ST approach is more accurate than a purely spatial method (Table 1). This result is consistent with existing literature which uniformly reports that ST approaches are more accurate than spatial approaches (Zeng et al., 2013; Guo et al., 2013; Zeng et al., 2016).

However, although considering information from days preceding and following the target estimation day should in principle always provide a further constraint on the estimate, this does not guarantee that an ST method will always outperform a spatial-only method in practice. The prime reasons for this are two-fold. First, because computational limitations cap the number of observations that can be considered, considering observations across multiple days necessarily leads to a reduction in the spatial density of observations

10

385 being considered. This first factor can be partially alleviated by carefully designing the selection probability
386 function (Eqn. 1). The second reason is that implementing a ST approach involves the estimation of a larger
387 number of covariance parameters (Eqn. 4-9) relative to a spatial-only approach, which can introduce
388 additional uncertainty. Indeed, we observe that the purely spatial approach performs better than the ST
389 method in some cases (e.g., the GOSAT ~~and GOME 2 1d cases).~~case).

390 Overall, a ST approach is likely to outperform a spatial-only approach when the data exhibit one (or more)
391 of three characteristics. First, a ST approach is likely better when the data are sparse or unequally
392 distributed. In these cases, a ST approach can intelligently leverage data in adjacent time periods to
393 compensate for the sparsity of data in the time period of interest. Second, an ST approach works well for
394 datasets with temporal gaps (e.g., due to cloud cover or instrument malfunction). An ST approach can fill
395 these gaps while a spatial-only approach cannot be used for temporal gap-filling. Third, an ST-approach is
396 well-suited to datasets with regional biases that manifest in one time slice but that do not repeat in adjacent
397 time slices. ~~Phrased differently, an ST-approach is well-suited to datasets with~~The difference between the
398 performance of ST and S-approaches obtained through cross-validation becomes most pronounced in
399 processing datasets with measurement errors that are spatially but not temporally correlated. In these cases,
400 an ST approach can use data from adjacent time periods to ~~create the~~obtain an estimate, data that do not
401 have the same regional, spatially-correlated biases. Although the resulting estimate may appear inferior
402 during cross-validation, this is because that estimate will not reproduce regional biases in data from the
403 time slice of interest. A spatial-only approach, by contrast, will reproduce these regional biases because it
404 does not use data from adjacent times when creating the estimate. As a result, a spatial-only approach will
405 appear to perform better in cross validation, but the ST approach will more accurately reflect the true,
406 underlying process.

5. Conclusions

408 In this study, we develop a method to create high spatio-temporal resolution maps from satellite data using
409 spatio-temporal moving window block kriging based on product-sum covariance model. The method relies
410 on a limited number of assumptions: that the observed physical quantity is spatio-temporally auto-
411 correlated, and that its nature can be inferred from the observations.

412 The method has several advantages over previously applied methods~~, as alluded to in Sect. 1: 1) it allows~~
413 ~~for the creation of contiguous maps at varying spatio-temporal resolution, 2) it can create maps at temporal~~
414 ~~resolutions shorter than achievable by other binning or kriging methods, 3) it can be applied for creating~~
415 ~~contiguous maps for physical quantities with varying spatio-temporal coverage (i.e., density of~~
416 ~~measurements), 4) it provides assessments of the uncertainty of interpolated values, 5) it utilizes all spatio-~~
417 ~~temporally available information to generate estimates, 6~~. Apart from the advances alluded to in Sect. 1: 1)
418 it improves covariance parameters estimation procedure because it does not model spatial and temporal
419 covariance separately, ~~7~~2) it allows for great flexibility in the choice of sampling function and ~~8~~3) it
420 provides estimates even for the time periods when measurements are not available. It can exploit
421 correlations with both past and future periods of the observed time spot to provide the most accurate
422 estimates.

423 We demonstrate the applicability of this method by creating Level 3 products from the GOSAT XCO_2, IASI
424 CH_4 and GOME-2 SIF data. Sparse XCO_2 observations from GOSAT and dense XCH_4 and SIF
425 observations from IASI and GOME-2 make a perfect test ground for the method. We show that the proposed
426 method can even map XCO_2 on daily time scales. The method generally yields more precise and accurate
427 (and unbiased) estimates compared to spatial method which used the same observations but assumed perfect

11

428 temporal correlation between data. The factors which could affect the performance of the ST method are
429 discussed in Section 4.2.

430 This approach could be used in the future to produce real-time estimates not only of XCO_2, XCH_4 or SIF,
431 but of other environmental data observed by satellites which exhibit spatio-temporal autocorrelations.
432 Especially important could be satellite datasets that have spatially, but not temporally, correlated errors. In
433 such cases, sampling across several time periods could perhaps help isolate and remove them, which should
434 be a subject of further studies.

435 The method could be applied in a standalone mode or as part of a broader satellite data processing package.
436 Maps produced by the spatio-temporal approach could then be incorporated into physical and
437 biogeochemical models of the Earth system.

438 6. Code availability

449 References

450 Chiles, J.-P. and Delfiner, P.: Geostatistics, second edition, Wiley, 2012.

451 Cressie, N. and Huang, H. C.: Classes of nonseperable, spatio-temporal stationary covariance functions,
452 Journal of the American Statistical Association, 94, 1 – 53, 1999.

453 Crévoisier, C., Nobileau, D., Fiore, A. M., Armante, R., Chédin, A., and Scott, N. A.: Tropospheric
454 methane in the tropics – first year from IASI hyperspectral infrared observations, Atmos. Chem. Phys., 9,
455 6337–6350, doi:10.5194/acp-9-6337-2009, 2009a.

456 Crévoisier, C., Chédin, A., Matsueda, H., Machida, T., Armante, R.,and Scott, N. A.: First year of upper
457 tropospheric integrated content of CO_2 from IASI hyperspectral infrared observations, Atmos. Chem. Phys.,
458 9, 4797–4810, doi:10.5194/acp-9-4797-2009, 2009b.

459 Crisp, D., Fisher, B. M., O'Dell, C., Frankenberg, C., Basilio, R., Bösch, H., Brown, L. R., Castano, R.,
460 Connor, B., Deutscher, N. M., Eldering, A., Griffith, D., Gunson, M., Kuze, A., Man10 drake, L., McDuffie,
461 J., Messerschmidt, J., Miller, C. E., Morino, I., Natraj, V., Notholt, J., O'Brien, D. M., Oyafuso, F.,
462 Polonsky, I., Robinson, J., Salawitch, R., Sherlock, V., Smyth, M., Suto, H., Taylor, T. E., Thompson, D.

12

463 R., Wennberg, P. O., Wunch, D., and Yung, Y. L.: The ACOS CO_2 retrieval algorithm – Part II: Global
464 XCO_2 data characterization, Atmos. Meas. Tech., 5, 687–707, doi:10.5194/amt-5-687-2012, 2012.

465 De Cesare, L., Myers, D. E., and Posa, D.: Spatio-temporal modelling of SO_2 in Milan district. In: Baafi,
466 E. Y., Schofield, N. A. (Eds.), Geostatistics Wollongong. Kluwer Academic Publishing, Dordrecht, pp.
467 1031 – 1042, 1996.

468 De Cesare, L., Myers, D. E., and Posa, D.: Product–sum covariance for space–time modeling: an
469 environmental application. Environmetrics 12, 11 – 23, 2001a.

470 De Cesare, L., Myers, D., and Posa, D.: Estimating and modeling space–time correlation structures: Stat.
471 Prob. Lett., v. 51, no. 1, 9–14, 2001b.

472 Dimitrakopoulos, R., and Luo, X.: Spatiotemporal Modeling: Covariances and Ordinary Kriging Systems,
473 Geostatistics for the Next Century, Kluwer Academic Publishers, Dordrecht, 88-93, 1994.

474 Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J.-E., Toon, G. C., Butz, A.,
475 Jung, M., Kuze, A., and Yokota, T.: New global observations of the terrestrial carbon cycle from GOSAT:
476 Patterns of plant fluorescence with gross primary productivity, Geophys. Res. Lett., 38, L17706,
477 doi:10.1029/2011GL048738, 2011.

478 Frankenberg C, O'Dell C, Guanter L, McDue J.: Remote sensing of near-infrared chlorophyll fluorescence
479 from space in scattering atmospheres: implications for its retrieval and interferences with atmospheric CO2
480 retrievals. Meas. Tech. 5, 2081–2094., doi:10.5194/amt-5-2081-2012, 2012.

481 Frankenberg, C., O'Dell, C., Berry, J., Guanter, L., Joiner, J., Köhler, P., et al.: Prospects for chlorophyll
482 fluorescence remote sensing from the Orbiting Carbon Observatory-2, Remote Sensing of Environment,
483 147, 1-12, doi: 10.1016/j.rse.2014.02.007, 2014.

484 Gambacorta A.: The NOAA Unique CrIS/ATMS Processing System (NUCAPS): Algorithm Theoretical
485 Basis Documentation, NOAA Center for Weather and Climate Prediction (NCWCP), Version 1.0, August
486 21, 2013, http://www.star.nesdis.noaa.gov/jpss/documents/ATBD/NUCAPS_ATBD_20130821.pdf

487 Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P. E., Gomez-Dans, J., Kuze, A., Suto, H., and Grainger,
488 R. G.: Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space
489 measurements, Remote Sens. Environ., 121, 236– 251, doi:10.1016/j.rse.2012.02.006, 2012.

490 Guo, L., Lei, L. and Zeng, Z.: Spatiotemporal correlation analysis of satellite-observed CO_2: Case studies
491 in China and USA. Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International, 21-
492 26 July, Melbourne, VIC, 2013.

493 Hammerling, D. M., Michalak, A. M., and Kawa, S. R.: Mapping of CO_2 at high spatiotemporal resolution
494 using satellite observations: Global distributions from OCO-2, J. Geophys. Res., 117, D06306,
495 doi:10.1029/2011JD017015, 2012a.

496 Hammerling, D. M., Michalak, A. M. O'Dell, C., and Kawa, S. R.: Global CO_2 distributions over land from
497 the Greenhouse Gases Observing Satellite (GOSAT), Geophysical Research Letters, 39, L08804,
498 doi:10.1029/2012GL051203, 2012b.

499 Huang, C., Zhang, H., and Robeson, S. M.: On the validity of commonly used covariance and variogram
500 functions on the sphere, Math. Geosci., 43, MR2824128, 721–733, doi: 10.1007/s11004-011-9344-7, 2011.

13

501 De Iaco, S., Myers, D., and Posa, D.: Space-time analysis using a general product–sum model: Stat. Probab.
502 Lett., v. 52, no. 1, 21–28, 2001.

503 Joiner, J., Yoshida, Y., Vasilkov, A. P., Yoshida, Y., Corp, L. A., and Middleton, E. M.: First observations
504 of global and seasonal terrestrial chlorophyll fluorescence from space, Biogeosciences, 8, 637-651,
505 doi:10.5194/bg-8-637-2011, 2011.

506 Joiner, J., Yoshida, Y., Vasilkov, A. P., Middleton, E. M., Campbell, P. K. E., Yoshida, Y., Kuze, A., and
507 Corp, L. A.: Filling-in of near-infrared solar lines by terrestrial fluorescence and other geophysical effects:
508 simulations and space-based observations from SCIAMACHY and GOSAT, Atmos. Meas. Tech., 5, 809–
509 829, doi:10.5194/amt-5-809-2012, 2012.

510 Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., Huemmrich, K. F.,
511 Yoshida, Y., and Frankenberg, C.: Global monitoring of terrestrial chlorophyll fluorescence from moderate
512 spectral resolution near-infrared satellite measurements: methodology, simulations, and application to
513 GOME-2, Atmos. Meas. Tech., 6, 2803-2823, doi:10.5194/amt-6-2803-2013, 2013.

514 Liu, J., Fung, I., Kalnay, E., Kang, J.-S., Olsen, E. T., and Chen, L.: Simultaneous assimilation of AIRS
515 Xco2 and meteorological observations in a carbon climate model with an ensemble Kalman filter, J.
516 Geophys. Res., 117, D05309,doi:10.1029/2011JD016642, 2012.

517 Matthias, K., and Cressie, N.: Spatio-temporal smoothing and EM estimation for massive remote-sensing
518 data sets, Journal of Time Series Analysis 32.4, 430-446, doi: 10.1111/j.1467-9892.2011.00732.x, 2011.

519 Katzfuss, M. and Cressie, N.: Bayesian hierarchical spatio-temporal smoothing for very large datasets.
520 Environmetrics, 23: 94–107. doi: 10.1002/env.1147, 2012.

521 Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near infrared sensor for carbon
522 observation Fourier-transform spectrometer on the Greenhouse Gases 5 Observing Satellite for greenhouse
523 gases monitoring, Appl. Optics, 48, 6716–6733, doi:10.1364/AO.48.006716, 2009.

524 Lee, J.-E., Frankenberg, C., van der Tol, C., Berry, J. A., Guanter, L., Boyce, C. K., Fisher, J. B., Morrow,
525 E., Worden, J. R., Asefi, S., Badgley, G., and Saatchi S.: Forest productivity and water stress in Amazonia:
526 observations from GOSAT chlorophyll fluorescence Proc. R. Soc. B, 280(1762), doi:
527 10.1098/rspb.2013.0171, 2013.

528 Li, J., and Heap, A.D.: Spatial interpolation methods applied in the environmental sciences: a review.
529 Environ. Model. Softw. 53, 173–189, doi: http://dx.doi.org/10.1016/j.envsoft.2013.12.008, 2014.

530 Li, J. and Heap, A.: A review of comparative studies of spatial interpolation methods in environmental
531 sciences: performance and impact factors. Ecol. Inf. 6, 228e241, 2011.

532 Li, J.: Assessing spatial predictive models in the environmental sciences: Accuracy measures, data variation
533 and variance explained, Environmental Modelling & Software, Volume 80, Pages 1-8, ISSN 1364-8152,
534 http://dx.doi.org/10.1016/j.envsoft.2016.02.004, 2016.

535 Maksyutov, S., Takagi, H., Valsala, V. K., Saito, M., Oda, T., Saeki, T., Belikov, D. A., Saito, R.,
536 Ito, A., Yoshida, Y., Morino, I., Uchino, O., Andres, R. J., and Yokota, T.: Regional CO_2 flux
537 estimates for 2009–2010 based on GOSAT and ground-based CO2 observations, Atmos. Chem. Phys.,
538 13, 9351–9373, doi:10.5194/acp-13-9351-2013, 2013.

Mathias, R.: Matrix completions, norms and Hadamard products, Proc. Amer. Math. Soc. 117 (4), 905–918, 1993.

Million, E.: The Hadamard Product. http://buzzard.ups.edu/courses/2007spring/projects/million-paper.pdf, 2007. Accessed on 15 January 2015.

NASA Earth Science: http://science.nasa.gov/earth-science/earth-science-data/data-processing-levels-for-eosdis-data-products/, last access: 23 July, 2014.

Nguyen, H., Katzfuss, M., Cressie, N. and Braverman, A.: Spatio-Temporal Data Fusion for Very Large Remote Sensing Datasets, Technometrics, Vol. 56, Iss. 2, doi: 10.1080/00401706.2013.831774, 2014.

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polon15 sky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO_2 retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmos. Meas. Tech., 5, 99–121, doi:10.5194/amt-5-99-2012, 2012.

Rouhani, S., and Hall, T.J.: Space-Time Kriging of Groundwater Data. In: Armstrong, M. (Ed.), Geostatistics. Kluwer Academic Publishers, Dordrecht, Vol. 2, 639-651, 1989.

Tadić, J., Qiu, X., Yadav, V. and Michalak, A.: Mapping of satellite Earth observations using moving window block kriging, Geosci. Model Dev., 8, 1–9, doi:10.5194/gmd-8-1-2015, 2015.

Xiong X., Barnet C., Maddy E., Gambacorta A., King T., and Wofsy S.: Mid-upper tropospheric methane retrieval from IASI and its validation. Atmos Meas Tech. 6: 2255-2265, 2013.

Zeng, Z., LiPing, L., L. LiJie, G., Li, Z., Bing, Z.,: Incorporating temporal variability to improve geostatistical analysis of satellite-observed CO_2 in China, Chinese Science Bulletin, 58(16), 1948-1954, 2013.

Zeng, Z., Lei, L., Strong, K., Jones, D. B. A., Guo, L., Liu, ., Deng, F., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Henderson, B., Kivi, R., Lindenmaier, R., Morino, I., Notholt, J., Ohyama,H., Petri, C., Sussmann, R., Velazco, V., A., Wennberg, P., O., and Lin, H.: Global land mapping of satellite-observed CO_2 total columns using spatio-temporal geostatistics, International Journal of Digital Earth, DOI: 10.1080/17538947.2016.1156777, 2016.

15

574 **Table 1**. Cross-validation results of GOSAT XCO₂, IASI XCH₄ and GOME-2 SIF datasets using spatio-
575 temporal and spatial methods, including mean absolute error (MAE), root mean squared error (RMSE),
576 relative mean absolute error (RMAE), relative root mean square error (RRMSE), percent of observations
577 lying outside of one, two, and three standard deviations ($\sigma_{\hat{z}}$) of the mapping uncertainty, and mean
578 difference. MAE, RMSE and bias units for GOSAT, IASI and GOME-2 are ppm, ppb and mW/m²/sr/nm,
579 respectively. RMAE and RRMSE are unitless, and due to the reasons explained in Section 4.1 given only
580 for GOSAT and IASI. Shaded fields represent best estimate in each category for every satellite.

|		GOSAT XCO₂			IASI XCH₄			GOME-2 SIF		
		ST	1d	7d	ST	1d	7d	ST	1d	7d
Estimates	Mean absolute error (MAE)	0.83	0.74	0.89	19.19	20.23	21.04	0.52	0.54	0.54
	Root mean squared error (RMSE)	1.12	0.98	1.21	25.25	27.10	27.77	0.68	0.69	0.69
	Relative mean absolute error (RMAE)	0.22	0.19	0.23	1.04	1.09	1.14	-	-	-
	Relative root mean square error (RRMSE)	0.29	0.25	0.31	1.37	1.46	1.50	-	-	-
Uncertainties	% observations falling outside 1$\sigma_{\hat{z}}$ uncertainty	9.13	15.03	10.70	11.02	9.06	13.84	14.60	12.14	24.80
	% observations falling outside 2$\sigma_{\hat{z}}$ uncertainty	1.12	3.01	1.39	0.48	0.51	0.86	1.20	0.64	4.33
	% observations falling outside 3$\sigma_{\hat{z}}$ uncertainty	0.067	0.52	0.13	0.04	0.046	0.022	0.11	0.05	0.83
Bias	Mean difference	-0.012	0.0066	-0.034	0.28	-0.14	0.58	0.016	0.0013	0.032

581

582

583

16

584
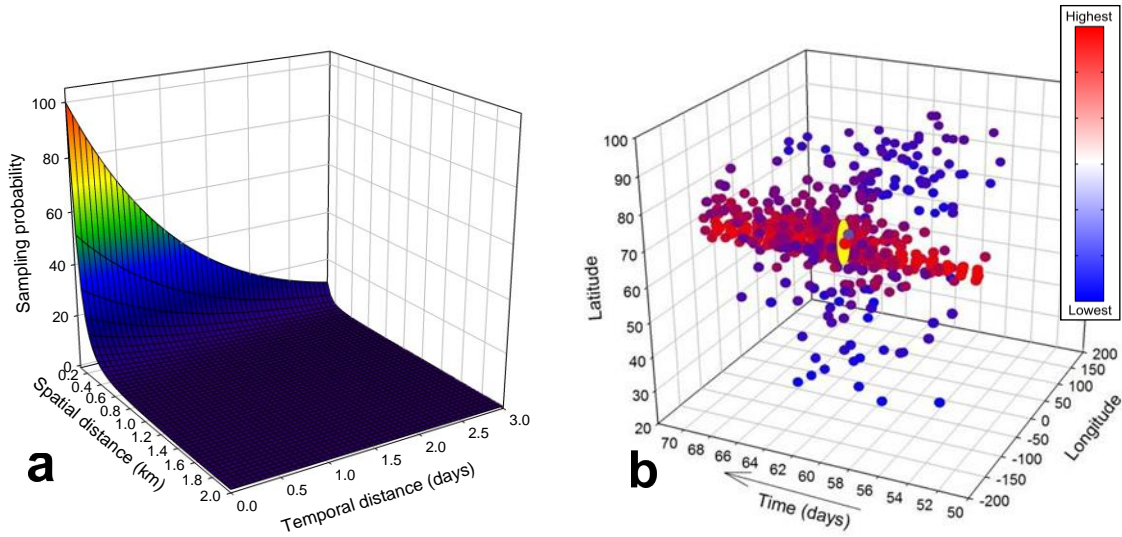
Figure 1. (a) Sampling probability as a decreasing function of spatial and temporal distance as used in this study, (b) The typical example of subsampled IASI Level 2 XCH_4 (altitude below 4 km) data for a selected estimation location (yellow circle). Color of observations shows semivariance between observation and estimation location (blue-lowest, red-highest). Due to stronger temporal covariance, the relative decrease of the sampling probability along temporal axis is smaller than with spatial distance.

592
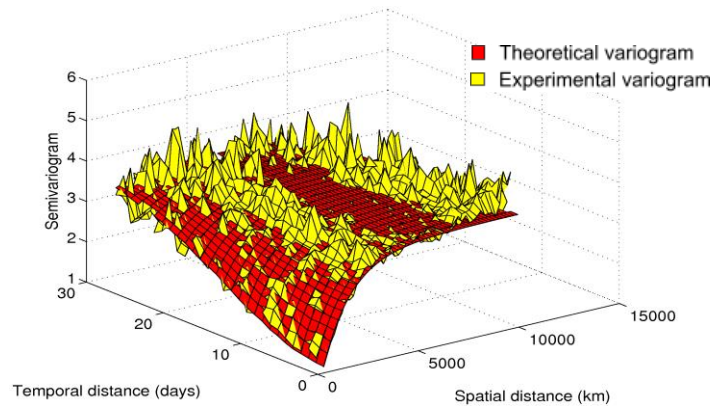
Figure 2. Illustration of experimental and fitted theoretical spatio-temporal variogram for GOSAT XCO_2 data.
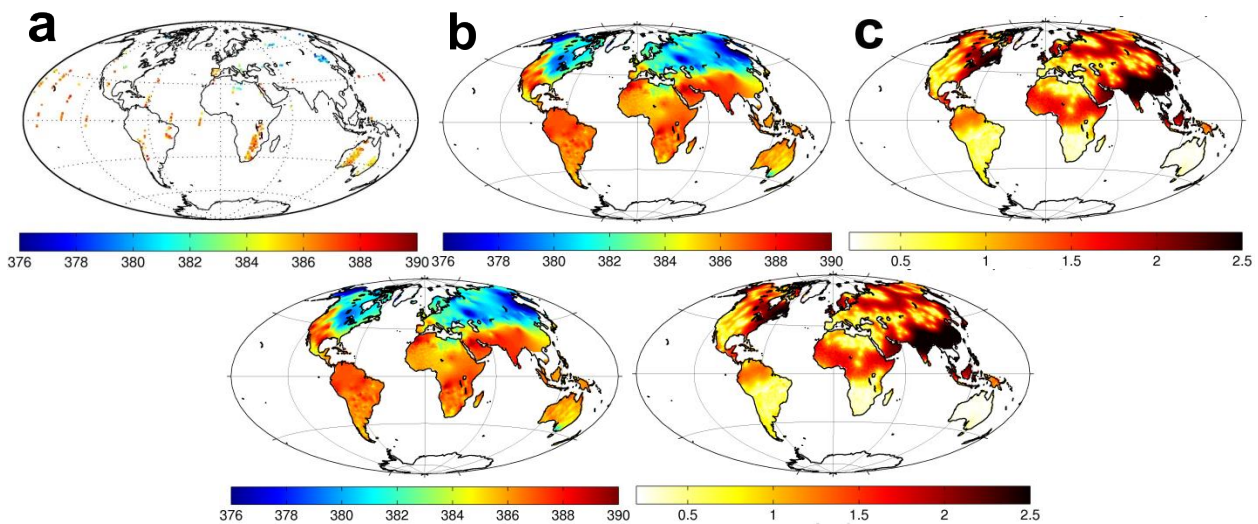
595

17

Figure 3. (a) GOSAT/ACOS v3.4 XCO$_2$ retrievals (Level 2 data) (ppm) for August 3, 2009 (b) Contiguous global GOSAT/ACOS v3.4 maps (Level 3 data) (ppm) for the same day obtained using Spatio-temporal Moving Window Block Kriging at $1 \times 1°$ spatial resolution, (c) associated uncertainties, given as 1-sigma ($\sigma_{\hat{z}}$) (ppm).
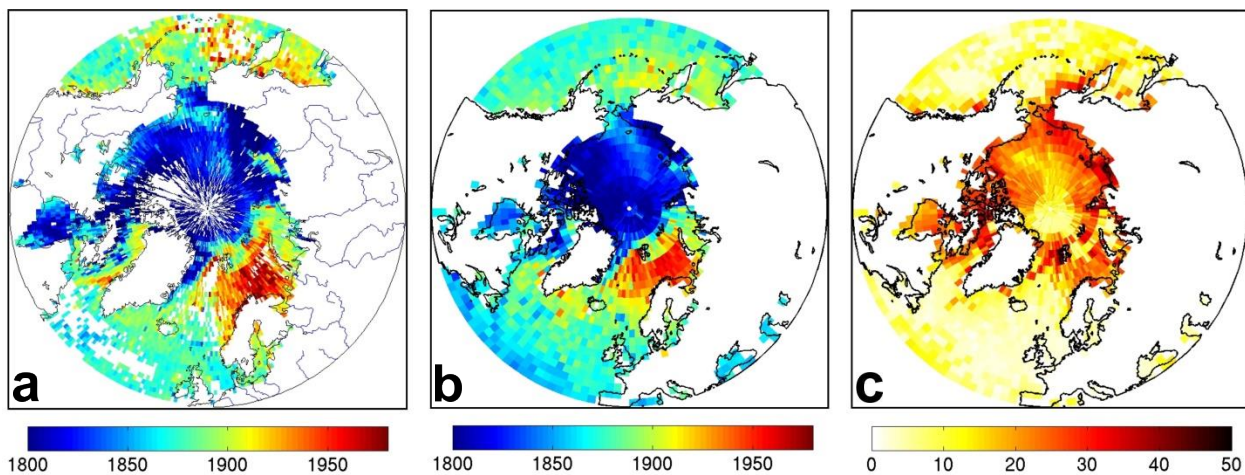

Figure 4. (a) IASI XCH$_4$ (0-4 km) retrievals (ppb) for March 2, 2013 (sea only), (b) Contiguous IASI maps for Northern Hemisphere for the same day obtained using Spatio-temporal Moving Window Block Kriging at $2.5 \times 2°$ spatial resolution and (c) associated uncertainties, given as 1-sigma ($\sigma_{\hat{z}}$) (ppb).
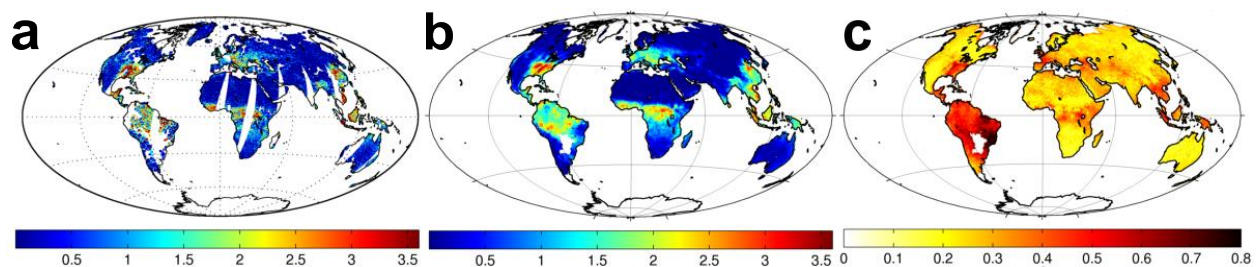
18

Figure 5. (a) GOME-2 SIF v14 retrievals (Level 2 data) (mW/m^2/sr/nm) for May 5, 2012, (b) Contiguous global GOME-2/SIF v14 maps (Level 3 data) (mW/m^2/sr/nm) for the same day obtained using Spatio-temporal Moving Window Block Kriging at $1 \times 1°$ spatial resolution, (c) associated uncertainties, given as 1-sigma ($\sigma_{\hat{z}}$) (mW/m^2/sr/nm).

19