1 **Estimation of CH$_4$ emission based on advanced 4D-LETKF assimilation system**

2 Jagat S. H. Bisht[1]*, Prabir K. Patra[1,2], Masayuki Takigawa[1], Takashi Sekiya[1], Yugo Kanaya[1], Naoko

3 Saitoh[2], and Kazuyuki Miyazaki[3]

4 [1]Research Institute for Global Change, JAMSTEC, Yokohama, 235-0019, Japan

5 [2]Center for Environmental Remote Sensing, Chiba University, Chiba, 263-8522, Japan

6 [3]Jet Propulsion Laboratory/California Institute for Technology, Pasadena, CA, USA,

7 *corresponding author's e-mail: jagatbisht@jamstec.go.jp

8

9 **Abstract**

10 Methane (CH$_4$) is the second major greenhouse gas after carbon dioxide (CO$_2$) which has

11 substantially increased during last decades in the atmosphere, raising serious sustainability and

12 climate change issues. Here, we develop a data assimilation system for in situ and column averaged

13 concentrations using Local ensemble transform Kalman filter (LETKF) to estimate surface emissions

14 of CH$_4$. The data assimilation performance is tested and optimized based on idealized settings using

15 Observation System Simulation Experiments (OSSEs) where a known surface emission distribution

16 (the truth) is retrieved from synthetic observations. We tested three covariance inflation methods to

17 avoid covariance underestimation in the emission estimates, namely; fixed multiplicative (FM),

18 relaxation to prior spread (RTPS) and adaptive multiplicative. First, we assimilate the synthetic

19 observations at every grid point at the surface level. In such a case of dense observational data, the

20 normalized Root Mean Square Error (RMSE) in the analyses over global land regions are smaller by

21 10-15% in case of RTPS covariance inflation method compared to FM. We have shown that

22 integrated estimated flux seasonal cycles over 15 regions using RTPS inflation are in reasonable

23 agreement between true and estimated flux with 0.04 global normalized annual mean bias. We have

24 then assimilated the column averaged CH$_4$ concentration by sampling the model simulations at

25 GOSAT observation locations and time for another OSSE experiment. Similar to the case of dense

26 observational data, RTPS covariance inflation method performs better than FM for GOSAT synthetic

27 observation in terms of normalized RMSE (2-3%) and integrated flux estimation comparison with the

28 true flux. The annual mean averaged normalized RMSE (normalized mean bias) in LETKF CH$_4$ flux

29 estimation in case of RTPS and FM covariance inflation is found to be 0.59 (0.18) and 0.61 (0.23)

30 respectively. The chi-square test performed for GOSAT synthetic observations assimilation suggests

31 high underestimation of background error covariance in both RTPS and FM covariance inflation

32 methods, however, the underestimation is much higher (>100% always) for FM compared to RTPS
33 covariance inflation method.

1. Introduction

35 Methane (CH_4) is the second major greenhouse gas, after carbon dioxide (CO_2), that has
36 anthropogenic sources. According to the contemporary record of the global CH_4 budget, the total of
37 all CH_4 sources ranges 538–593 Tg yr^{-1} during 2008–2017 (Saunois et al., 2020). The primary natural
38 sources are from wetlands (~40%). The main anthropogenic CH_4 emissions are from microbial
39 emissions associated with ruminant (livestock and waste), rice cultivation, fugitive emissions (oil and
40 gas production and use), and incomplete combustion of bio and fossil fuels. The major fraction of
41 atmospheric CH_4 sinks (range: 474 - 532 Tg yr^{-1}) occurs in the troposphere by oxidation via reaction
42 with hydroxyl (OH) radicals (Patra, et al., 2011; Saunois et al., 2020); other loss processes include
43 oxidation by soil, and reactions with O^1D and Cl. The lifetime of CH_4 in the atmosphere is estimated
44 to be 9.1 ± 0.9 years (Szopa et al. 2021).

45 Regional CH_4 emissions can be estimated from CH_4 concentration fields and chemistry transport
46 models using Bayesian synthesis approaches based on inverse modeling techniques (e.g., Enting,
47 2002). In such approach, emissions are optimized on a coarse resolution (e.g., for a limited number of
48 pre-defined regions) mostly using surface-based observations. CH_4 concentrations are provided by the
49 NOAA cooperative air sampling network sites (Dlugokencky et al., 2020) and other networks by the
50 World Data Centre for Greenhouse Gases (WDCGG) website, hosted by the Japan Meteorological
51 Agency. In the recent years, satellite measurements are made from the Greenhouse Gases Observing
52 Satellite (GOSAT) or the TROPOspheric Monitoring Instrument (TROPOMI) (Lorente et al., 2021),
53 covering the globe with fine spatio-temporal scales. GOSAT provide an extensive global observations
54 of column CH_4 concentrations since 2009 (Yoshida et al., 2013). Some of the inverse modeling
55 studies utilize the satellite observations for CH_4 flux estimation (Zhang et al., 2021; Maasakkers et al.,
56 2016), but, it requires enormous computational resources while dealing with more flux regions and
57 more observations.

58 Grid-based CH_4 flux optimization is also performed using adjoint technique (4-D Var data
59 assimilation) and Ensemble Kalman Filter (EnKF), but was limited to small sets of observations
60 (Houweling et al., 1999; Meirink et al., 2008; Bruhwiler et al., 2014). Bruhwiler et al. (2014) followed
61 the EnKF method of Peters et al. (2005) to estimate the CH_4 surface fluxes that utilizes an off-line
62 ACTM framework. Techniques such as 4-D Var and EnKF are important to estimate CH_4 fluxes since
63 they can assimilate a large number of observations, manage high-resolution fluxes. In the EnKF
64 system, a flow-dependent forecast error covariance structure is provided by ensemble model forecasts,
65 while it does not need an adjoint model that makes it simple but powerful tool for flux estimation.

66   One of the limitations in EnKF method is the dependence of the resolution of state vector on

67   ensemble size, which can give spurious results if the number of ensemble members is much smaller

68   than the rank of the error covariance matrix (Houtekamer and Zhang, 2016).

69   LETKF is a type of square-root EnKF that performs analysis locally in space without perturbing the

70   observations (Ott et al., 2002, 2004; Hunt et al., 2007). LETKF is computationally efficient since the

71   observations are assimilated simultaneously not serially, it is simple to account for observation error

72   correlation. Miyazaki et al. (2011) and Kang et al. (2012) demonstrated the implementation of

73   LETKF data assimilation system by coupling an ACTM for carbon-cycle research using atmospheric

74   CO_2 observations. It is also extensively applied for the emission estimation of short-lived species

75   using satellite data (Skachko et al., 2016; Miyazaki et al., 2019; Sekiya et al., 2021). In this work, we

76   will estimate the CH_4 fluxes using a LETKF data assimilation system. Assimilation windows ranging

77   from 6 hours (Kang et al., 2012) to several months (Bruhwiler et al., 2014) have been used, depending

78   on the desired time resolution of the estimated emissions, which is often limited by the observational

79   data density. The time frame over which the system behaves linearly, and in what time frame the

80   observations respond to the control variables such as, atmospheric transport, as well as observation

81   abundance, must also be taken into consideration. Within an assimilation window, where and when

82   the fluxes would be constrained by specific observations is to be ascertained by the correlation

83   between ensemble prior fluxes and the ensemble CH_4 concentrations simulation from a forward

84   model (Liu et al., 2016).

85   Main objective of this work is to develop an advanced 4-D data assimilation system based on LETKF

86   that simultaneously estimates atmospheric distributions and surface fluxes of CH_4. OSSEs are

87   conducted to assess the performance of LETKF since it is important to test the system against the

88   known emissions or the truth. The OSSE LETKF set-up of top-down CH_4 flux estimation using online

89   ACTM is an essential step before implementing on real in situ and satellite observation.

90   **2. Formulation of LETKF system**

91   We briefly describe the LETKF in the application of CH_4 flux estimation, while detailed derivation of

92   equations and code implementation are given elsewhere (Hunt et al., 2007; Miyazaki et al., 2011;

93   Miyoshi et al., 2010). The notation used here for LETKF formulation is adopted from Kotsuki et al.

94   (2017). In the LETKF, the background ensemble (columns of matrix x^b) in a local region evolved

95   from a set of perturbed initial conditions. The background ensemble mean, $\bar{x}^b$, and its perturbation,

96   X^b, are estimated from the ensemble forecast such as:

$$\bar{x}^b = \frac{1}{m}\sum_{i=1}^{m} x_i^b \; ; \quad X_i^b = x_i^b - \bar{x}^b \tag{1}$$

97 Where 'm' indicates the ensemble size. The background error covariance matrix P^b in the m-
98 dimensional ensemble is defined as:

$$P^b = \frac{1}{m-1} X^b [X^b]^T \tag{2}$$

99 The analysis ensemble mean $\bar{x}^a$ is derived using background ensemble mean $\bar{x}^b$ and ensemble
100 perturbations X^b such as:

$$\bar{x}^a = \bar{x}^b + X^b \widetilde{P}^a (Y^b)^T R^{-1} (y^o - H\bar{x}^b) = \bar{x}^b + X^b w^a \tag{3}$$

101 where H, Y, R, and $\widetilde{P}^a$ denote the linear observation operator, ensemble perturbation matrix in the
102 observation space ($Y \equiv Hx$), observation error covariance matrix, and analysis error covariance matrix
103 in the ensemble space, respectively. The superscripts 'o', 'b' and 'a' denote the observations,
104 background (prior), and analysis (posterior), respectively. w^a defines the analysis increment (or
105 analysis weight) in observation space and is derived using the information about observational
106 increment $y^o - H\bar{x}^b$. The analysis error covariance matrix ($\widetilde{P}^a$) in the m-dimensional ensemble space
107 is spanned by ensemble perturbation (Hunt et al., 2007) and defined as:

$$\widetilde{P}^a = \{(m-1)I + (HX^b)^T R^{-1} HX^b\}^{-1} \tag{4}$$

108 Finally, the analysis ensemble perturbations X^a at the central grid point are derived such as:

$$X^a = X^b \{(m-1)\widetilde{P}^a\}^{1/2} \tag{5}$$

109 Where, $\{(m-1)\widetilde{P}^a\}^{1/2}$ is a multiple of the symmetric square root of the local analysis error
110 covariance matrix in ensemble space and could be computed by singular vector decomposition
111 method. The LETKF solves the analysis update equations 3 and 5 at every model grid point
112 independently by assimilating local observations within the localization cut-off radius.

113 We have applied a gross error check as a quality control to exclude observations that are far from the
114 first guess, the appropriate degrees of the gross error check are also examined. Figure 1 shows the
115 schematic diagram of our LETKF set-up with two ensemble members for 3 consecutive assimilation
116 cycles with 8 days assimilation window. The analysis is obtained at mid-point time of the assimilation
117 window (Figure 1). The analyzed (updated) surface flux is used for next data assimilation cycle
118 starting from the mid-point time of the previous data assimilation window. The state vector
119 augmentation approach is used to estimate the atmospheric CH_4 surface flux (Kang et al., 2012;
120 Miyazaki et al., 2011).

121 Assimilation window size and ensemble members are chosen based on computational efficiency and
122 estimation accuracy. A larger assimilation window means fluxes are constrained by more
123 observations, however, it requires handling of large matrix optimization which is difficult in cases of
124 dense observation and introduces sampling errors related to transport errors. In this study, few
125 sensitivity experiments performed to demonstrate the choice of assimilation window length and
126 ensemble size when GOSAT synthetic observation are assimilated in Section 4.2.
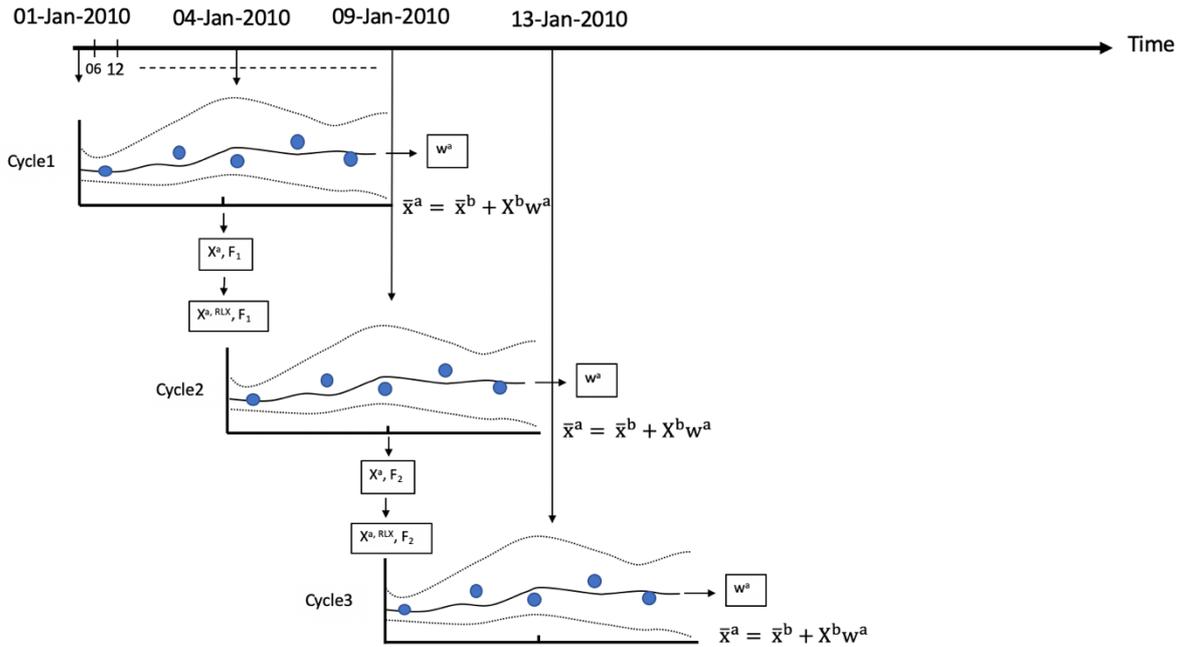
127

128 **Figure 1:** Schematic represents the temporal evolution of LETKF cycle. In the first assimilation
129 window (Cycle1), the dotted lines show the ensemble forecast of CH_4 concentrations (with 2
130 ensemble members), the solid line shows the linear combination of the forecasts, the filled circles
131 show the observations of CH_4 concentration. The data assimilation finds the linear combination of the
132 ensemble forecast by estimating the weight (w^a) that best fits the observations throughout the
133 assimilation window. The analysis weight is applied to obtain optimal surface fluxes (F) and the
134 concentration of CH_4 at the intermediate time of the data assimilation window. The updated analyzed
135 concentration ensembles are used as initial condition after relaxation ($X^{a, RLX}$) (Eq. 8) for the next
136 ensemble forecast. The spread of the ensemble members represents the forecast error. The schematic
137 is adapted from Kalnay and Yang (2010) and Miyazaki et al. (2011).

138 **2.1 Covariance inflation**

139 The LETKF data assimilation needs variance inflation to mitigate the under dispersive ensemble. We
140 tested three methods; fixed multiplicative (FM), relaxation-to-prior spread (RTPS), and adaptive
141 multiplicative covariance inflation.

142   The fixed multiplicative (FM) inflation method (Anderson and Anderson, 1999) inflates the prior

143   ensemble by inflating the background error covariance matrix P^b defined in equation (Eq. 2) such as:

$$P_{inf}^b = \gamma P_{tmp}^b \tag{6}$$

144   where P_{tmp}^b represents the temporary background error covariance matrix which is inflated by a factor

145   γ.

146   The other inflation methods used to prevent the reduction of ensemble spread are relaxation-to-prior

147   perturbation (RTPP) (Zhang et al., 2004) and relaxation-to-prior spread (RTPS) (Whitaker and

148   Hamill, 2012). The RTPP methods relax the reduction of the ensemble spread after updating the

149   ensemble perturbations which blends the background and analysis ensemble perturbations as:

$$X_{inf}^a = \alpha_{RTPP} X^b + (1 - \alpha_{RTPP}) X_{tmp}^a \tag{7}$$

150   where α_{RTPP} denotes the relaxation parameter of the RTPP.

151   The RTPS inflation method relaxes the reduction of ensemble spread by relaxing the analysis spread

152   to prior spread such as:

$$X_{RLX}^a = \left(\frac{\alpha_{RTPS} \sigma^b + (1 - \alpha_{RTPS}) \sigma^a}{\sigma^a} \right) X_{tmp}^a \tag{8}$$

153   where σ and α_{RTPS} denote the ensemble spread, and relaxation parameter of the RTPS, respectively.

154   The range of parameter α_{RTPS} is bounded by [0, 1]. This study focuses mainly on the FM and RTPS

155   covariance inflation methods.

156   In addition, Miyoshi (2011) applied adaptive inflation by determining the multiplicative inflation

157   factors at every grid point at every analysis step using the observation-space statistics derived by

158   Daley (1992) and Desroziers et al. (2005).

$$< dd^T > = H P_{inf}^b H^T + R \tag{9}$$

159   Where the operator '$<\bullet>$' denotes the statistical expectation and $d = y^o - H\bar{x}^b$ (observation-minus-

160   first-guess), and R is the error observation covariance matrix.

161   The impact of using the adaptive multiplication inflation method is discussed in the GOSAT synthetic

162   observation assimilation experiments in Section 4.2.

163   **2.2 MIROC4-ACTM**

164   Model for Interdisciplinary Research on Climate, version 4.0 (MIROC4) based ACTM (hereafter

165   referred to as MIROC4-ACTM) (Patra et al., 2018; Bisht et al., 2021) is used here for CH_4

166 concentration simulations. The model simulations have been performed at horizontal grid resolution
167 of approximately 2.8×2.8º latitude-longitude grid (T42 spectral truncations) and hybrid vertical
168 coordinate of 67 levels (Earth's surface to 0.0128 hPa, Watanabe et al., 2008). Bisht et al., 2021
169 performed the multi-tracer analysis and demonstrated the importance of very well-resolved
170 stratosphere in the MIROC4-ACTM that illustrates better extratropical stratospheric variabilities, and
171 simulated tropospheric dynamical fields. The meteorological fields in MIROC4-ACTM are nudged to
172 the JMA Re-analysis (JRA-55) data (Kobayashi et al., 2015).

173 **3. Experimental set-up**

174 **3.1 Construction of known surface emissions (truth)**

175 Present OSSEs intend to develop basic tuning strategies before the actual data to be assimilated which
176 is useful to accelerate the operational use of real observations. The OSSE has been discussed here by
177 exploiting the known "truth". The synthetic observations to be assimilated in the OSSE are generated
178 from nature runs which uses bottom-up surface emission (true) data to simulate global 3-D CH_4
179 concentrations. The true surface CH_4 emissions are prepared on the monthly scale using
180 anthropogenic and natural sectors, minus the surface sinks due to bacterial consumption in the soil
181 (Chandra et al., 2021). The anthropogenic emissions were obtained from the Emission Database for
182 Global Atmospheric Research, version 4.3.2 inventory (EDGARv4.3.2) (Maenhout et al., 2019) that
183 includes the emissions from the major sectors such as; fugitive, enteric fermentation and manure
184 management, solid waste and wastewater handling. The biomass burning emissions are taken from the
185 Global Fire Database (GFEDv4s) (van der Werf et al., 2017) and Goddard Institute for Space Studies
186 emissions (Fung et al., 1991). The wetland and rice emissions are taken from the process-based model
187 of the terrestrial biogeochemical cycle, Vegetation Integrated Simulator of Trace gases (VISIT) (Ito,
188 2019) that is based on Cao et al. (1996). The other natural emission such as, ocean, termites, mud
189 volcano are taken from TransCom-CH_4 inter-comparison experiment (Patra et al., 2011). The total
190 emissions are taken as the truth for the OSSEs and the concentration simulated by MIROC4-ACTM
191 will be referred to as synthetic observations.

192 **3.2 Prior flux preparation and LETKF setting**

193 Based on our understanding of CH_4 inverse modelling, the uncertainty in regional flux estimation is
194 found to be 30% or lower (Chandra et al., 2021). Therefore, we attempted to reproduce the true flux
195 by starting with a prior flux that is lower by 30% of the true flux (prior flux has same seasonal cycles
196 as true flux). The MIROC4-ACTM is initialized with the spin-up of 3 years (2007 – 2009) with prior
197 flux distribution. The initial CH_4 distribution on 01 January 2007 was taken from an earlier simulation
198 of 27 years. An initial perturbation with standard deviation of approximately 6-8% spread is applied

199   to the a priori flux as the initial ensemble spread, whereas no ensemble perturbation was applied to the

200   initial CH_4 concentration. The sensitivity of the initial ensemble spread to CH_4 flux estimation is

201   discussed in Section 4.2. The uncertainty to perturb prior fluxes is generated based on random

202   positive values with normal distribution. The monthly scale prior emission is linearly interpolated at 6

203   hourly intervals to be used in the MIROC4-ACTM simulation for data assimilation. This study

204   performs two LETKF data assimilation experiments. In these experiments, we provided initial

205   perturbation on regional basis over land (53 different land regions; Chandra et al., 2021) and at every

206   grid over ocean, no spatial error correlation between grid points is considered among ensemble

207   members. However, in Section 4.2.5, we also discussed the sensitivity of CH_4 data assimilation by

208   providing initial ensemble spread at every grid by considering horizontal spatial error correlation

209   between grid points among ensemble members, with a global mean correlation of 20%.

210   **3.3 Experiment 1: Synthetic dense observation formulation**

211   The OSSE setting with very accurate and dense observation surface data is an attempt to demonstrate

212   that data assimilation system works reasonably in the estimation of the true surface flux. Errors in the

213   estimated flux could arise due to the insufficient ensemble size and also the implemented inflation

214   methods to overcome the under-sampling, along with simplified forecast process of emissions. In real

215   data assimilation, there are additional sources of potential errors, such as, atmospheric transports, and

216   inappropriate prior or observation uncertainties. In our OSSEs, CH_4 fluxes as mentioned in Section

217   3.2 are used as "true" fluxes in generating synthetic observations (CH_4 concentrations). In the

218   experiment 1, the simulated surface layer CH_4 concentrations at each grid for the entire globe were

219   used as synthetic assimilated observations. We added a constant measurement uncertainty of 5 ppb,

220   which is typically achieved by the present-day measurement systems (e.g., Dlugokencky et. al, 2020).

221   In this study, the CH_4 observations are assimilated by applying the observation error covariance

222   localization (Kotsuki et al., 2020) to reduce the spurious spatial correlation due to smaller ensemble

223   size than the degrees of freedom of the system ($R \leftarrow R \times \exp\left(-\frac{1}{2}\{(d_h/\sigma_h)^2 + (d_v/\sigma_v)^2\}\right)$). Where

224   d_h and d_v denote the horizontal distance (km) and vertical difference (log[Pa]) between the analysis

225   model grid point and observation location. The tunable parameters σ_h and σ_v are the horizontal

226   localization scale (km) and vertical localization scale (log[Pa]), respectively. Using the spatial

227   localization technique, we have estimated the CH_4 flux for each grid by choosing the CH_4

228   observations that influence the grid point using optimal cutoff radius ($\simeq 3.65\sigma_{h,v}$; Miyoshi et al.,

229   2007) with horizontal covariance localization (σ_h) of 2200 km and vertical covariance localization

230   (σ_v) of 0.3 in the natural logarithmic pressure (log[Pa]) coordinate. The localization is performed to

231   improve the signal to noise ratio of ensemble-based covariance. Numerous sensitivity experiments

232   have been performed by varying the horizontal and vertical localization length in order to obtain the

233   optimized CH_4 flux that best compare with the truth. The LETKF assimilates the observations within

8

234 the specified radius to solve the analysis state at each grid point independently (Liu et al., 2016;

235 Kotsuki et al., 2020). State vector of the analysis includes the atmospheric CH_4 concentration, which

236 is the prognostic variable of forecast model and the state vector is further augmented by surface CH_4

237 flux, which is not a model prognostic variable. This augmentation enables the LETKF to directly

238 estimate the parameter through the background error covariance with observed variables (Baek et al.,

239 2006). The state vector augmentation is implemented similar to that used by Miyazaki et al. (2011).

240 This approach analyses CH_4 flux during the analysis step. The purpose of the simultaneous CH_4

241 emission and concentration optimization is to reduce the uncertainty of the initial CH_4 concentrations

242 on the CH_4 evolution during the assimilation window and to maximize the observations potential

243 (Tian et al., 2014).

244 The atmospheric CH_4 concentration is changed during both the analysis and forecast steps. A

245 challenge of this scheme is that, the analysis increment is added to the model state at each analysis

246 step, without considering the global total CH_4 mass conservation in the model, but consistent with the

247 observed local CH_4 abundance.

248 In this case, surface flux at every model grid point is analyzed with 8-days assimilation window

249 during the year 2010 with the 100 ensemble members. The ensemble size and assimilation window

250 are chosen based on the CH_4 flux estimation accuracy calculated by performing sensitivity experiment

251 for ensemble size (60, 80, and 100) and assimilation window (3-days and 8-days), respectively (not

252 shown).

253 **3.4 Experiment2: synthetic satellite observation formulation**

254 One way to address the real-world CH_4 flux estimation problem is to first make the OSSE dataset like

255 real observations. In this OSSE experiment, we have assimilated synthetic column average CH_4

256 concentrations with a coverage mimicking GOSAT satellite observations. We prepared a model

257 simulated column averaged CH_4 concentrations (XCH_4) dataset that is spatiotemporally sampled with

258 GOSAT-observations as follows:

$$XCH_4 = XCH_{4(a\ priori)} + \sum_j h_j\, a_j (CH_{4(ACTM)} - CH_{4(a\ priori)})_j \qquad (10)$$

259 Where, XCH_4 is the column-averaged model simulated CH_4 concentration. $XCH_{4(a\ priori)}$ is a priori

260 column-averaged concentration. $CH_{4(ACTM)}$ and $CH_{4(a\ priori)}$ are the CH_4 profile from ACTM and a

261 priori, respectively. h_j is the pressure weighting function (j is the vertical layer index), and a_j

262 represents averaging kernel matrix for the column retrieval which is the sensitivity of the retrieved

263 total column at the various ('j') atmospheric levels. In the next step, we added the same retrieval

264   (XCH₄) error as GOSAT to the XCH4 (ACTM simulated) to make the OSSE more realistic and then

265   attempt to estimate the true fluxes.

266   In this case, the CH₄ flux has been estimated for each grid by choosing the CH₄ observation with

267   cutoff radius ($\simeq 3.65 \, \sigma_{h,v}$) with horizontal covariance localization (σ_h) of 5000 km and vertical

268   covariance localization (σ_v) of 0.35 in the natural logarithmic pressure (log[Pa]) coordinate. The

269   optimal horizontal and vertical covariance localization values are chosen based on trial and error

270   method (those best fits to estimate CH₄ flux when compared with truth). A long cutoff radius has been

271   chosen due to sparse observational coverage of GOSAT. Covariance localization is necessary to

272   remove long-range erroneous correlations and for mitigating sampling errors in the ensemble-based

273   error covariance with a limited ensemble size (Miyoshi et al., 2007; Greybush et al., 2011; Kotsuki et

274   al., 2020). The surface flux is analyzed at every model grid point with 8-days assimilation window

275   and 100 ensemble members those are chosen based on sensitivity experiments discussed in Section

276   4.2.

277   **4. Results and Discussion**

278   **4.1 Experiment with dense OSSE**

279   The time series of normalized RMSE ($\sqrt{\sum_{i=1}^{n}(x_i^a - x_i^t)^2/n}/\tilde{x}^t$; x_i^a and x_i^t is the analysis and true

280   state at ith model grid point, n is the total number of grid points, and $\tilde{x}^t$ represents the mean of true

281   flux) in the analyses over global landmass region is shown in Figure 2. The normalized global RMSE

282   is calculated using FM and RTPS inflation methods (Fig. 2) after assimilating synthetic observation at

283   every grid (Section 3.4). Noteworthy is that the experiment with FM inflation method shows 10-15%

284   larger error in estimating the atmospheric surface CH₄ flux compared to RTPS inflation method. One

285   of the reasons of better RMSE using RTPS inflation method is due to the more degrees of freedom

286   provided by relaxation (α_{RTPS}) in ensemble spread (Eq. 8) that could nudge the ensemble of CH₄

287   concentrations towards observations. The initial flux analysis spread using RTPS and FM is shown in

288   supporting information (Fig. S1) which shows larger initial analysis flux spread over Brazil, tropical

289   America, and Asia in RTPS inflation compared to FM inflation method. We performed numerous

290   sensitivity test with RTPS inflation method and found that uniform relaxation is not substantial, for

291   some of the regions. Figure 2 shows the RMSE for FM, fixed RTPS ($\alpha_{RTPS} = 0.4$, applied globally,

292   the optimized value is obtained by manual fine tuning) and conditional RTPS ($\alpha_{RTPS} = 0.3\text{-}0.7$

293   applied different α_{RTPS} regionally by manual fine tuning). In case of conditional RTPS, the optimal

294   values of α_{RTPS}, i.e., 0.6, 0.3, and 0.7 for the regions south of 20ºS, 20ºS-20ºN, and north of 20ºN,

295   respectively, were obtained from data assimilation sensitivity calculations with varying α_{RTPS} for the

296   three regions separately to best match the true states. We find that the conditional RTPS method

297   improves the accuracy by ~5% compared to fixed RTPS and 10-15% compared to FM. In the

298   following, we discuss the results obtained using the conditional RTPS and FM inflation methods.

299   We have also shown RMSE (not normalized) of surface flux in supplementary information (Fig. S2).

300   Flux RMSE has been estimated globally for both the inflation methods, and also for south of 20°N (by

301   considering only those land grids which fall into south of 20°N; Fig. S2) for comparative purposes. It

302   could be noticed that (supporting information Fig. S2), above north of 20°N, the flux estimation error

303   is higher, specifically during spring-summer when CH_4 emissions peak over most of the northern

304   hemispheric regions (Fig. 3). The high uncertainty during spring-summer (Fig. S2) in the flux

305   estimation over these regions could appear due to the attenuation of surface observations as a result of

306   active vertical mixing. The RMSE during autumn (Fig. S2) is comparable in case of global and south

307   of 20°N, which indicates RMSE arising from southern hemispheric regions, likely over Brazil as it

308   peaks during autumn (Fig. 3).
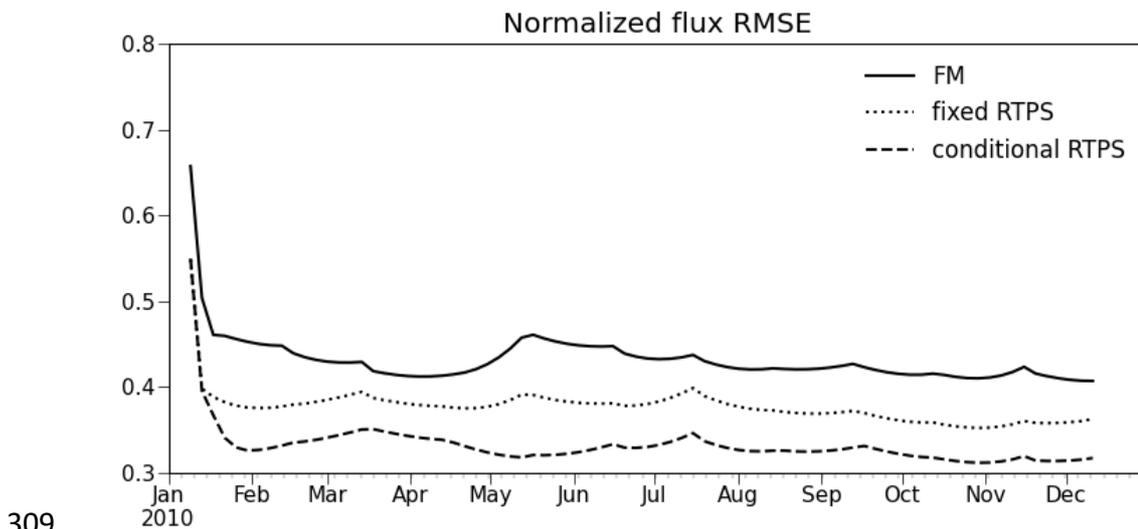
309

310   **Figure 2.** Time series of normalized RMSE of surface CH_4 flux analysis, for 1 year of data

311   assimilation using FM, fixed RTPS, and conditional RTPS inflation methods over global landmass

312   region.

313   Figure 3 shows regional total flux seasonal cycles comparison of the estimated fluxes for 15 terrestrial

314   regions with those of the prior and true fluxes. The estimated flux retrieved using RTPS inflation

315   method over different regions agrees well with that of true flux. We intend to show the capability of

316   LETKF estimated fluxes over these regions using surface observations to mimic the true fluxes in our

317   understanding of terrestrial biosphere CH_4 cycle. These results are consistent with Figure 2 with

318   annual global normalized mean bias ($\sum_{i=1}^{n}(x_i^a - x_i^t)/\sum_{i=1}^{n}(x_i^t)$) of -0.04. It could also be noticed

319   from Figure 3 that estimated fluxes converge to true fluxes over most of the regions after about 2-3

320   months.
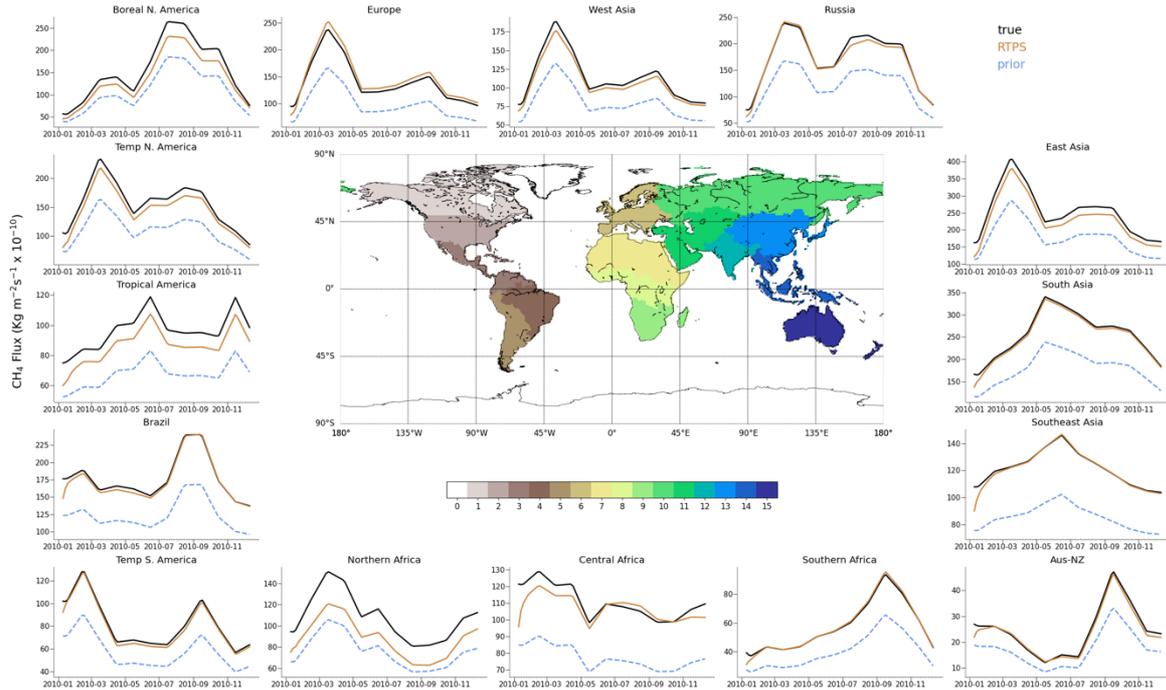
Figure 3. The 1-year CH$_4$ total flux seasonal cycles of true (black), prior (blue), and estimated from the LETKF (orange) conditional RTPS inflation method in 15 regions after assimilating dense synthetic surface CH$_4$ observations.

To see the degree of similarity in the flux distribution between the estimated and true fluxes, we show monthly mean spatial flux distribution for June, and November in Figure 4 and 5, respectively, along with the bias in prior and estimated flux. As shown in Figures 4 and 5, the general spatial patterns of the true flux are estimated well. These results suggest that, our LETKF system is capable of reproducing continental spatial flux patterns by using such an idealized dense surface observational data. However, some clear differences in flux estimation could be noticed from FM and RTPS inflation method (Figs. 4 and 5), for e.g., over Eurasian and American continent, analysis with RTPS shows clear improvement compared to FM covariance inflation method. We calculated the global mean normalized bias with RTPS and FM covariance inflation method which is found to be -0.04 and -0.11, respectively over land regions that shows RTPS significantly improved the flux estimation compared to FM covariance inflation method.

12
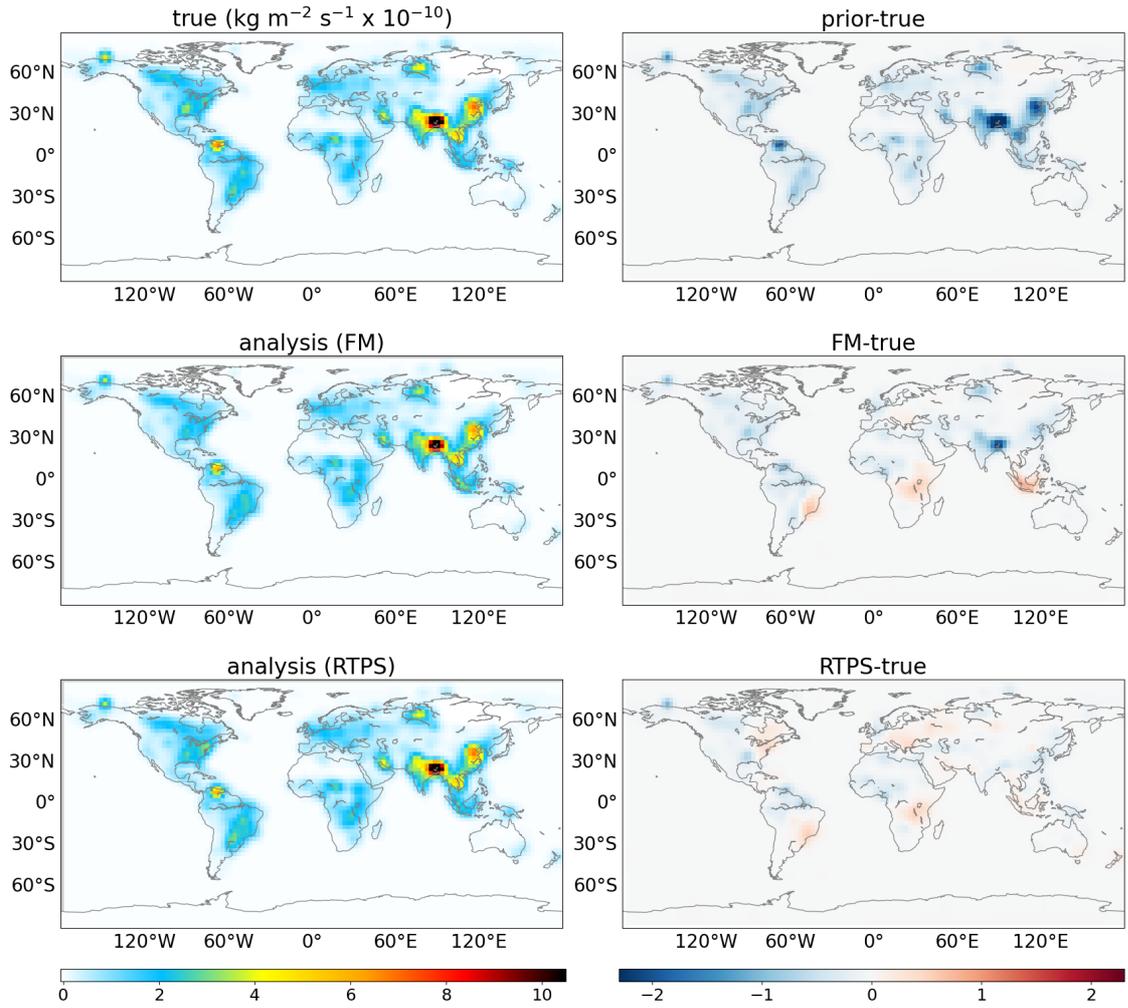
Figure 4. Spatial distribution of surface CH_4 fluxes (true; top left panel, FM analysis; middle left panel, RTPS analysis; bottom left panel) and the associated bias in prior (prior-true; top right panel) and estimated (FM-true; middle right panel, RTPS-true; bottom right panel) fluxes during June, 2010.
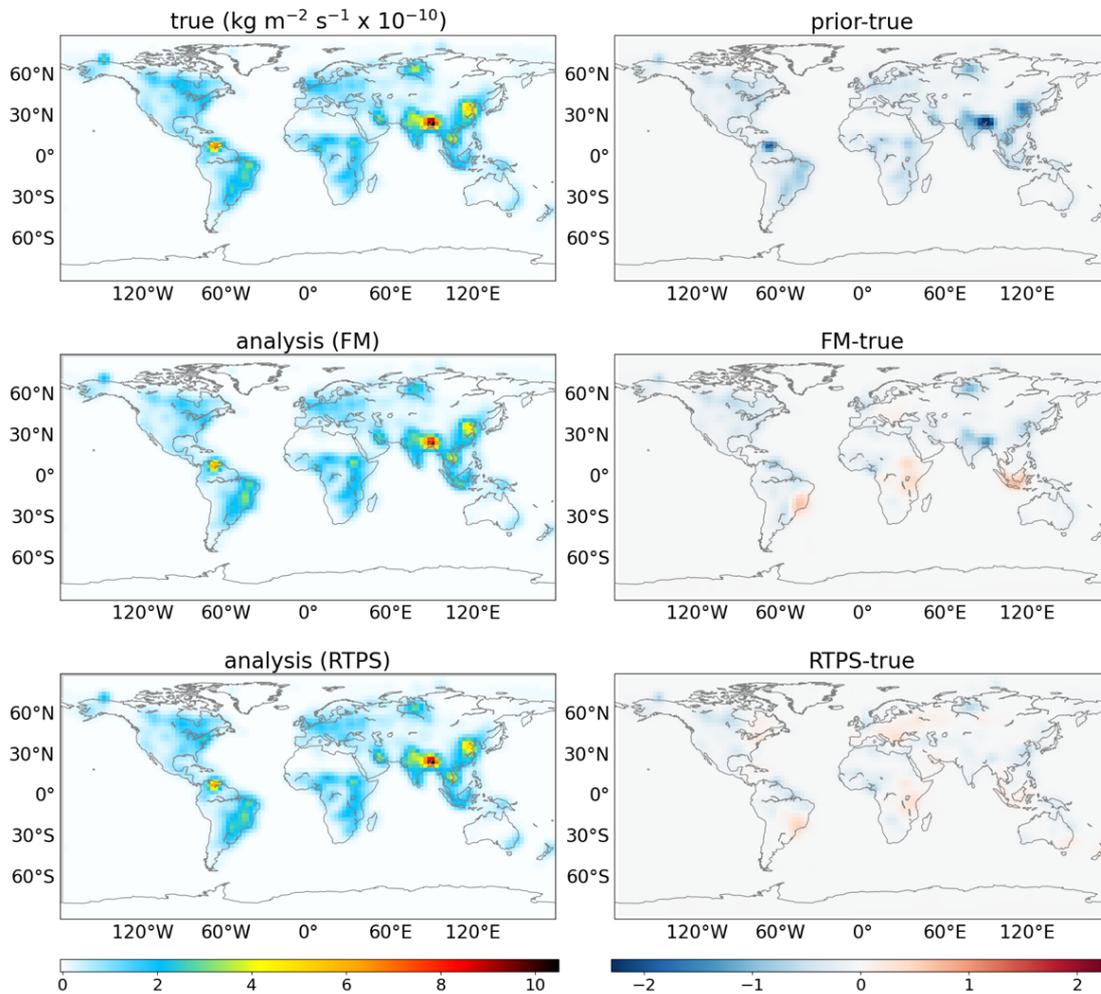
Figure 5. Same as Figure 4 but for November, 2010.

4.2 Experiment by mimicking the real satellite observational data set

In this section we discuss the LETKF flux estimation by assimilation of GOSAT synthetic CH_4 concentration observations. Figure 6 shows the model simulated mean XCH_4 concentration sampled spatiotemporally with GOSAT observations during January and July for the year 2010 (sampling method discussed in Section 3.4). In this case we have shown different LETKF sensitivity experiments such as; LETKF sensitivity to (1) FM, RTPS, adaptive multiplicative inflation (2) assimilation window (3) ensemble size, (4) chi-square test, (5) prior ensemble spread. In the LETKF sensitivity experiments from 1-4, the initial ensemble spread provided similar way as Experiment 1 and conditional RTPS inflation method is used. Conditional RTPS method is also used in Section 4.2.6 for CH_4 flux estimation.

4.2.1 LETKF sensitivity to FM, RTPS, and adaptive multiplicative inflation

353   This study mainly emphasizes on FM and RTPS inflation methods used in CH_4 LETKF data

354   assimilation. The annual average normalized RMSE (absolute bias) with RTPS and FM covariance

355   inflation is found to be 0.59 (0.18) and 0.64 (0.22), respectively. The RTPS inflation method performs

356   better than the FM inflation method overall. In addition to RTPS inflation, sensitivity test is also

357   performed using adaptive multiplicative inflation methods.

358   In the adaptive inflation, we need to provide an initial multiplicative inflation factor at the beginning

359   of data assimilation cycle (Cycle 1 in Fig. 1). Following the method of Miyoshi (2011), the

360   multiplication inflation factor information calculated in previous cycle (i.e. Cycle1 in Fig. 1) is used

361   for next data assimilation cycle at every grid point (Cycle 2 in Fig. 1). We perform two sensitivity

362   experiments. In the first (second) case we provided 50% (40%) initial inflation in the beginning of

363   Cycle 1 (Fig. 1). The normalized RMSE in the both the adaptive inflation sensitivity experiments are

364   comparable (0.65, Supporting information Fig. S3a) till July, but from the beginning of August,

365   RMSE increases exponentially in the first experiment. However, in terms of chi-square distribution

366   CH_4 flux estimation with first sensitivity adaptive multiplicative inflation experiment (50% initial

367   inflation case) is better than second sensitivity experiment (Supporting information Fig. S3b; chi-

368   square test described in Section 4.2.4). To identify the regions of high estimated CH_4 flux error, we

369   have shown the background error spread in CH_4 flux estimation over 15 regions (Supporting

370   information Fig. S3c) and found that spread over west and south east Asia rises exponentially post

371   July that indicates the rise of estimated CH_4 flux error over these regions in the first sensitivity

372   adaptive multiplicative inflation experiment. Our analysis suggests that CH_4 flux estimation is

373   depending on the initial inflation factor provided in the beginning of data assimilation cycle (Cycle 1,

374   Fig. 1) in adaptive multiplication method. Also, we need to be very careful to monitor the background

375   error spread evolution with time to estimate the CH_4 flux with adaptive inflation, chi-square

376   distribution analysis is not sufficient.

377   In case of RTPP inflation, we found the parameter α_{RTPP} is very difficult to fine-tuned due to its very

378   high sensitivity to estimate the CH_4 flux. We fail to obtain an optimized α_{RTPP} value to estimate the

379   CH_4 flux. Whitaker and Hamill (2012), also demonstrated the better accuracy in LETKF

380   meteorological data assimilation with RTPS compared to RTPP covariance inflation method. They

381   found RTPP method produces very large errors if the inflation parameter exceeds the optimal value.

382   **4.2.2 Assimilation window**

383   The LETKF data assimilation window length determines the time span of the observations assimilated

384   in each assimilation cycle. We have shown the sensitivity of two assimilation window size

385   configurations; 3 days and 8 days in supporting information Figure S4. Our sensitivity experiments

386   with window size configurations show that 8 days long assimilation window estimates the CH_4 flux

with better accuracy (~10%) compared to 3 days assimilation window, because more observational
information is incorporated into the system with 8 days long assimilation window. This study uses 8
days assimilation window for CH_4 LETKF data assimilation.
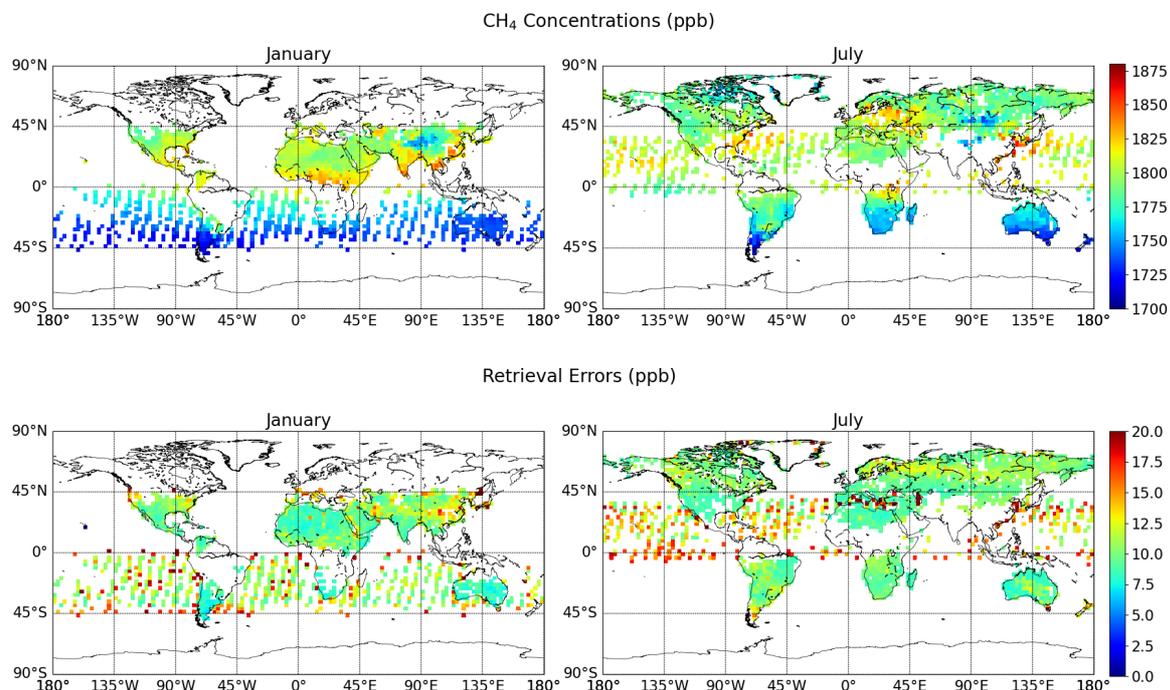
Figure 6. Monthly mean ACTM simulated XCH_4 (ppb) sampled with GOSAT observations to be
assimilated (valid during the year 2010). The actual retrieval errors are added in the synthetic GOSAT
observations. Data are shown for two representative months, depicting the southern and northern
hemisphere differences in data coverage.

4.2.3 Ensemble size

Figure 7a shows the RMSE using different ensemble members. The RMSE stabilizes gradually as the
ensemble size increases from 60 to 80 to 100 ensemble members. The ensemble size dependency of
flux estimation suggests the further scope of the improvement in flux estimation by increasing the
ensemble members. In this study we stick to 100 ensemble members due to high computational cost
while solving large covariance matrices. The larger error in flux estimation in case of column
averaged synthetic GOSAT CH_4 observations assimilation compared to dense observations (Fig. 2) is
likely due to the weaker constraint on surface fluxes provided by satellite observations and sparse
observations.

404
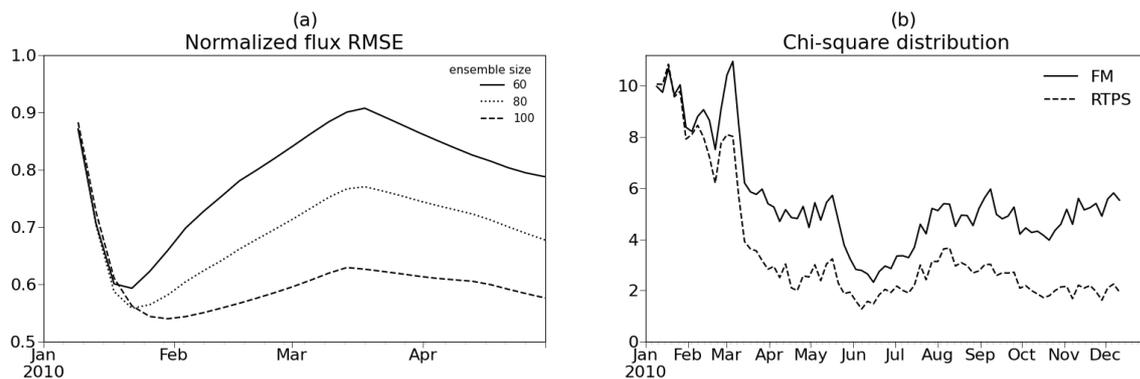
Figure 7: (a) Flux estimation RMSE using different ensemble size with RTPS covariance inflation. (b) Chi-square distribution using FM and RTPS covariance inflation methods with the ensemble size of 100.

4.2.4 Chi-square test

We have carried out chi-square test for the evaluation of background error covariance matrix (Miyazaki et al., 2012). For the χ^2 test, the innovation statistics are diagnosed from the observation minus forecast $(y^o - Hx^b)$, the estimated error covariance in the observation space $(HP^bH^T + R)$, and the number of observations k, such as:

$$Y = \frac{1}{\sqrt{k}}\left(HP^bH^T + R\right)^{-1/2}\left(y^o - Hx^b\right) \tag{11}$$

Using this statistic, the χ^2 is defined as follow:

$$\chi^2 = \text{trace}YY^T \tag{12}$$

The performance of background error covariance matrix determined based on the high and lower value of chi-square. Chi-square value should converge to 1, a value higher (lower) than 1 indicates underestimation (overestimation) of the background error covariance matrices. Our results suggest that, background error covariance matrix is highly underestimated in both RTPS and FM covariance inflation methods (Fig. 7b). However, the chi-square values convergence towards 1 is better in the case of RTPS compared to FM covariance inflation method which indicates the improved representation of background errors and then more appropriate data assimilation corrections in the case of the RTPS inflation method. The chi-square distribution starts saturating after the month of March. Post March analysis shows the background error covariance matrix underestimation is much higher (>100%) in case of FM compared to RTPS covariance inflation method.

4.2.5 CH$_4$ LETKF sensitivity to initial ensemble spread

17

425   A test case for CH_4 LETKF data assimilation has been performed where the initial spread is provided

426   by considering the initial perturbation on each model grid with spatial error correlation between grid

427   points among ensemble members, with global mean correlation of 20%. In this case, we found that

428   the analysis fluxes are extremely sensitive to the initial ensemble spread if prior fluxes perturbed with

429   more than 5% prior uncertainty. Therefore, we used initial ensemble perturbation with only 2% prior

430   uncertainty. Reducing the initial ensemble spread reduces the CH_4 flux estimation sensitivity (>60%).

431   However, it also poses a challenge to mitigate the under-dispersive background error covariance

432   matrix. We performed LETKF data assimilations in this case with RTPS covariance inflation method

433   (α_{RTPS} = 0.9 optimized value is used here uniformly) with 8-days long assimilation window and 100

434   ensemble members and calculated the normalized RMSE between analysis and true fluxes

435   (Supporting information Fig. S5). Noteworthy that, the estimated error between analysis and true

436   fluxes (Fig. S5) with this setting (grid-wise initial ensemble spread) is still larger (25%) than the case

437   when region-wise initial ensemble spread provided (Fig. 7a; 100 ensemble size). It suggests that,

438   initial ensemble spreads among ensemble members needs to be carefully provided that best represents

439   CH_4 variability among ensembles to estimate the CH_4 flux.

440   Note that, the OSSEs used in this study did not consider the effects of model errors other than CH_4

441   fluxes, such as model transport errors. In real situations, model errors can have a substantial impact on

442   flux estimates (Locatelli et al., 2013), which needs to be taken into account in background

443   covariances. Therefore, the optimal data assimilation setting can differ between the OSSEs presented

444   in this study and real observation cases. Further efforts, e.g., by conducting a more comprehensive

445   OSSE that accounts for various model errors and by performing various sensitivity calculations in real

446   cases, would provide an improved understanding of the optimal inflation settings to improve CH_4 flux

447   estimates in following study.

448   **4.2.6 Estimated CH_4 flux analysis**

449   Figure 8 shows the regional fluxes seasonal cycle comparison for the estimated fluxes over 15

450   terrestrial regions with those of the prior and true fluxes. We have also shown assimilation results in

451   case of FM inflation method in supporting information (Fig. S6), which shows the flux estimation

452   disagreement over more regions compared to RTPS inflation method; e.g., for Tropical and North

453   America, whole African continent, Australia-New Zealand.

454   We have shown the GOSAT observations in Figure 6 and supporting information Figure S7. We

455   found very marginal flux estimation improvement over Central Africa after May (Fig. 8), that could

456   be associated with the less GOSAT coverage over this region (Fig. 6). On the other hand, over

457   Northern Africa, no improvement in flux estimation is found. In case of dense OSSE too (Fig. 3), we

458   didn't find satisfactory flux estimation over Northern Africa which is most probably related to the

insufficient initial spread among ensemble members over this region (we have used same initial ensemble spread in both OSSE cases). Over Europe, GOSAT observations are remarkably less, specifically for first few months (January-April; supporting information Fig. S7). Therefore, the flux update over Europe would be influenced by the observations from neighboring regions falling under the chosen cutoff radius that are mainly in Northern Africa where the flux estimation itself not satisfactory. It could also be noticed that the retrieval error added in this OSSE case are high over Europe (September-October; supporting information Fig. S7),) and its adjacent Sea (Mediterranean Sea; June-August) which could also affect the surface CH_4 flux estimation.
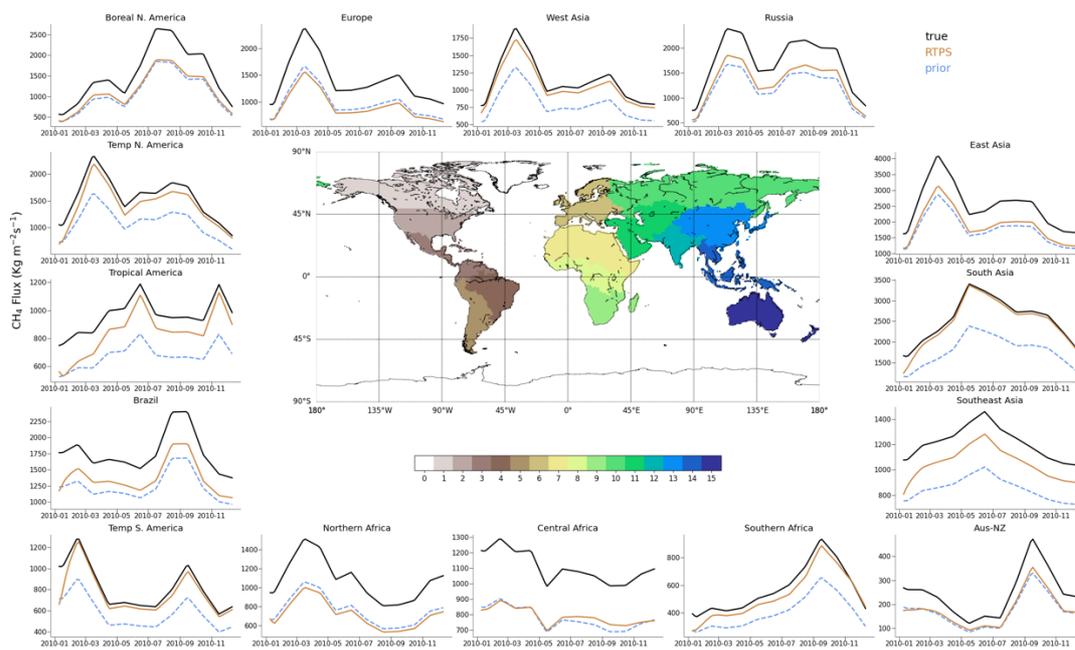
Figure 8. Same as Figure3 but after assimilating synthetic GOSAT observations.

Figure 9 and 10 show spatial patterns of the true and estimated fluxes by assimilating the column averaged CH_4 concentrations during June and November (Fig. 6). It may be noticed that RTPS covariance inflation method better able to estimate the true flux pattern compared to FM covariance inflation method. The spatial pattern shown using RTPS inflation method emphasizes the positive and negative bias in the estimated flux (Figs. 9 and 10), but generally agrees with the flux seasonal cycle plots shown in Figure 8.

Our LETKF CH_4 data assimilation experiment by assimilating GOSAT synthetic observation with the implementation of the advanced RTPS covariance inflation method better estimate the time-evolving surface CH_4 fluxes compared to FM covariance inflation method. The difficulty to estimate the surface CH_4 flux over a few regions may be overcome by applying additional methodologies, such as the assimilation of surface observations simultaneously, and the use of information about the CH_4 fluxes climatology. A correction factor derived based on empirical formulation that could use CH_4

481 flux climatology information is needed to apply to maintain the CH_4 mass conservation. This could be

482 implemented by the checking the simulated CH_4 burden gain between years in comparison with the

483 observed CH_4 growth rates.

484
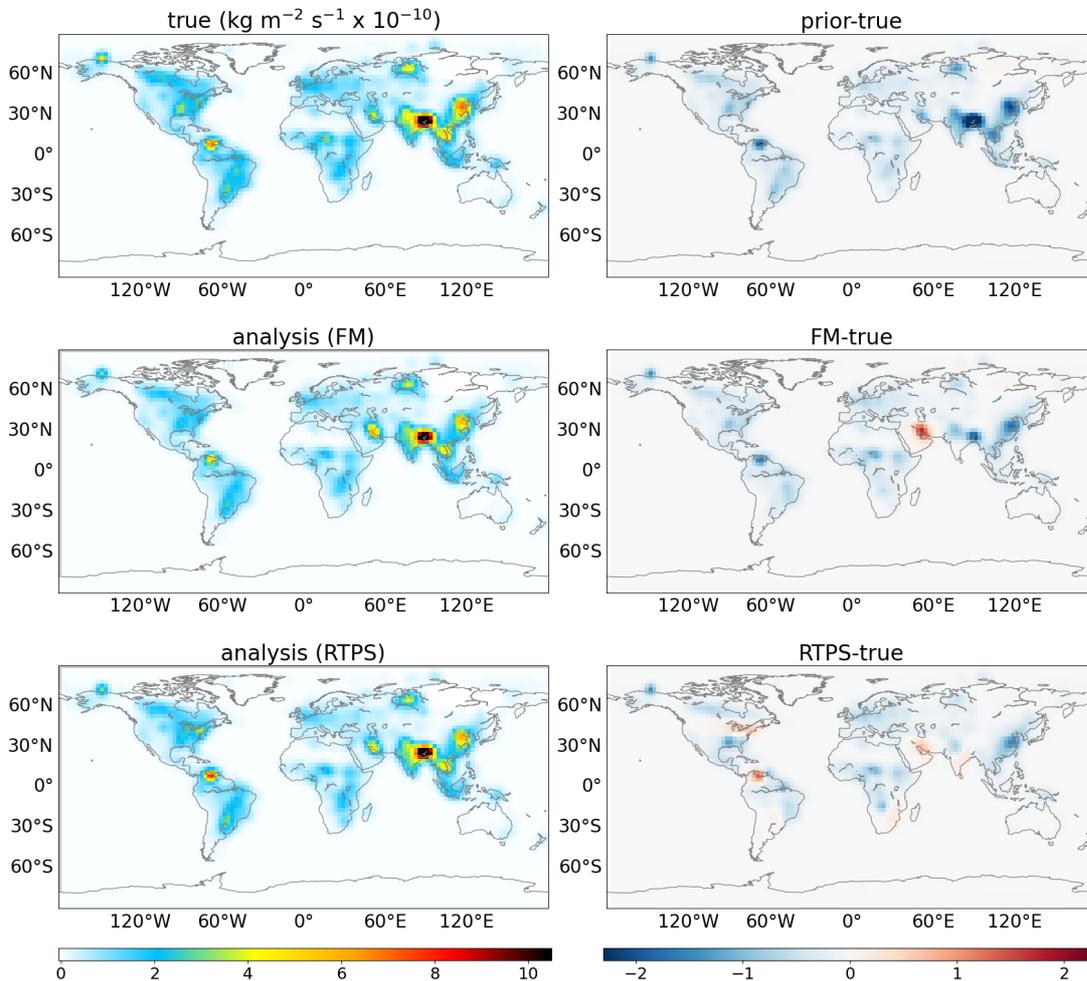
Figure 9. Monthly mean true (true; top left panel) and estimated (FM analysis; middle left panel,

RTPS analysis; bottom left panel) CH_4 flux after assimilating column averaged synthetic CH_4

concentrations (Fig. 6) during June using FM and RTPS inflation methods. The associated bias with

prior and estimated fluxes is also shown (prior-true; top right panel; FM-true; middle right panel,

RTPS-true; bottom right panel).

490

true (kg m^{-2} s^{-1} x 10^{-10})

prior-true

analysis (FM)

FM-true

analysis (RTPS)

RTPS-true
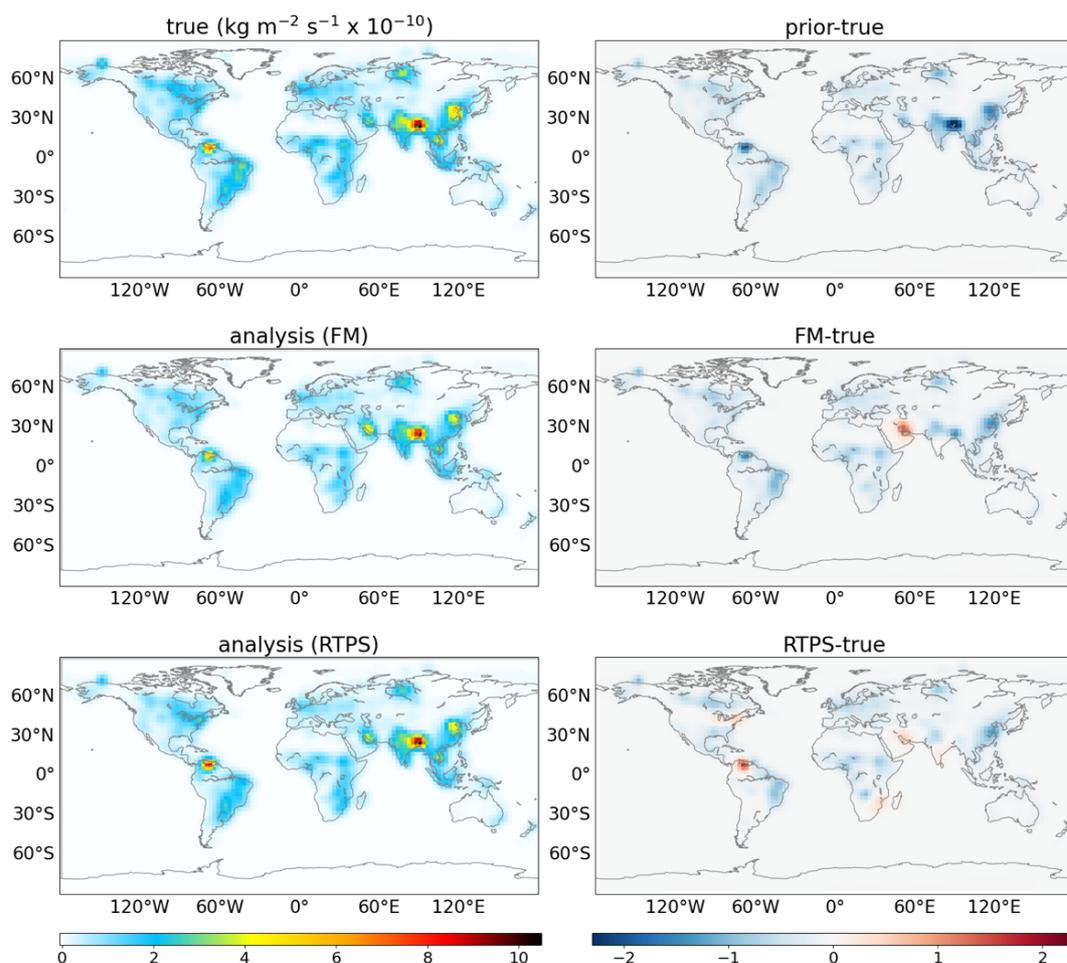
Figure 10. Same as Figure 9 but for November.

5. Summary

In this study, we have introduced 4D-LETKF data assimilation system that utilizes MIROC4-ACTM as a forward model for CH$_4$ flux estimation. This study has extensively tested both FM and RTPS inflation methods for the LETKF CH$_4$ flux estimation. We have conducted two experiments to demonstrate the ability of LETKF system to estimate the CH$_4$ surface flux globally. In Experiment1, we have assimilated the synthetic dense surface CH$_4$ observations. While in Experiment2, synthetic GOSAT CH$_4$ observations are assimilated. Based on the results of the sensitivity tests using FM and RTPS inflation methods in Experiment1, we have found that RTPS inflation produces significantly less normalized RMSE (10-15%) compared to FM inflation method. In Experiment2, we discussed, LETKF parameters such as, different inflation techniques, ensemble size, assimilation window, initial ensemble spread sensitivity, and chi-square test. The ensemble size (this study uses maximum 100 ensemble members) sensitivity test suggests that more ensemble members could help to accurately represent the covariance matrix with more degrees of freedom. The assimilation window sensitivity

test exhibits that 8 days assimilation window reduces the normalized flux RMSE by about 10%
compared to 3 days assimilation window in case of GOSAT synthetic observations assimilation.

Our approach of assimilation with RTPS inflation could provide more degrees of freedom to fit the
ensemble of CH_4 concentrations to the observed ones, resulting the improved analyzed fluxes. The
RTPS inflation method is capable of obtaining reasonable flux estimates with normalized annual
mean bias of 0.04, and 0.61 in case of dense surface synthetic observations and GOSAT synthetic
observations, respectively. We demonstrated in our sensitivity OSSE experiment with synthetic
GOSAT observations that, over American and African continents and also over Australia - New
Zealand, the LETKF data assimilation with FM inflation method does not show much improvement in
the true flux estimation, but RTPS inflation method reasonably estimate the true flux over most of
these regions. One of the reasons for better flux estimates from RTPS inflation method is the
prevention of analysis spread drastically. In the CH_4 LETKF flux estimation, surface CH_4 flux is not a
prognostic state vector in the ACTM, which results in the decay of spread continuously in analysis
steps. RTPS inflation method could mitigate such under disperse spread problem. This study finds
that spatially homogeneous relaxation is not sufficient. It needs to be fine-tuned and applied
conditionally.

The sensitivity of LETKF CH_4 flux estimation to initial ensemble spread needed to be carefully dealt
with when applied to real data assimilation system. A future OSSE with additive covariance inflation
technique could be interesting while applied with RTPS inflation method for CH_4 LETKF data
assimilation since in additive covariance inflation initial estimated flux error cannot propagate. The
state vector augmentation technique used here updates the flux after each data assimilation cycle but it
doesn't conserve the total atmospheric CH_4 amount which is one of the limitations of this work. A
correction factor needs to be implemented to conserve the total atmospheric CH_4 amount after
completion of a few data assimilation cycles. We have not accounted for the transport error due to
meteorological fields in this work (Patra et al., 2011b), in case of real observations data assimilation a
week-long window may introduce transport errors in CH_4 analysis because of nonlinear growth of
ensemble perturbations.

References

Anderson, J. L. and Anderson, S. L.: A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts, Mon. Weather Rev., 127, 2741–2758, https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2, 1999.

Baek, S.-J., Hunt, B. R., Kalnay, E., Ott, E., and Szunyogh, I.: Local ensemble Kalman filtering in the presence of model bias, Tellus A Dyn. Meteorol. Oceanogr., 58, 293–306, https://doi.org/10.1111/j.1600-0870.2006.00178.x, 2006.

Bisht, J. S. H., Machida, T., Chandra, N., Tsuboi, K., Patra, P. K., Umezawa, T., Niwa, Y., Sawa, Y., Morimoto, S., Nakazawa, T., Saitoh, N., and Takigawa, M.: Seasonal Variations of SF_6, CO_2, CH_4, and N_2O in the UT/LS Region due to Emissions, Transport, and Chemistry, J. Geophys. Res. Atmos., 126, https://doi.org/10.1029/2020JD033541, 2021.

Bruhwiler, L., Dlugokencky, E., Masarie, K., Ishizawa, M., Andrews, A., Miller, J., Sweeney, C., Tans, P., and Worthy, D.: CarbonTracker-CH4: an assimilation system for estimating emissions of atmospheric methane, Atmos. Chem. Phys., 14, 8269–8293, https://doi.org/10.5194/acp-14-8269-2014, 2014.

Cao, M., Marshall, S., and Gregson, K.: Global carbon exchange and methane emissions from natural wetlands: Application of a process-based model, J. Geophys. Res. Atmos., 101, 14399–14414, https://doi.org/10.1029/96JD00219, 1996.

Chandra, N., Patra, P. K., Bisht, J. S. H., Ito, A., Umezawa, T., Saigusa, N., Morimoto, S., Aoki, S., Janssens-Maenhout, G., Fujita, R., Takigawa, M., Watanabe, S., Saitoh, N., and Canadell, J. G.: Emissions from the Oil and Gas Sectors, Coal Mining and Ruminant Farming Drive Methane Growth over the Past Three Decades, J. Meteorol. Soc. Japan. Ser. II, 99, 309–337, https://doi.org/10.2151/jmsj.2021-015, 2021.

573   Daley, R.: The lagged-innovation covariance: A performance diagnostic for data assimilation, Mon.

574   Wea. Rev., 120, 178–196, 1992.

575   Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and

576   analysis-error statistics in observation space, Q. J. R. Meteorol. Soc., 131, 3385–3396,

577   https://doi.org/10.1256/qj.05.108, 2005.

578   Dlugokencky, E. J., Crotwell, A. M., Mund, J. W., Crotwell, M. J., & Thoning, K. W.: Atmospheric

579   Methane Dry Air Mole Fractions from the NOAA GML Carbon Cycle Cooperative Global Air

580   Sampling Network, 1983-2019, Version: 2020-07, https://doi.org/https://doi.org/10.15138/VNCZ-

581   M766, 2020.

582   Enting, I. G.: Inverse Problems in Atmospheric Constituent Transport, Cambridge University Press,

583   https://doi.org/10.1017/CBO9780511535741, 2002.

584   Fung, I., John, J., Lerner, J., Matthews, E., Prather, M., Steele, L. P., and Fraser, P. J.: Three-

585   dimensional model synthesis of the global methane cycle, J. Geophys. Res., 96, 13033,

586   https://doi.org/10.1029/91JD01247, 1991.

587   Greybush, S. J., Kalnay, E., Miyoshi, T., Ide, K., and Hunt, B. R.: Balance and Ensemble Kalman

588   Filter Localization Techniques, Mon. Weather Rev., 139, 511–522,

589   https://doi.org/10.1175/2010MWR3328.1, 2011.

590   Houtekamer, P. L. and Zhang, F.: Review of the Ensemble Kalman Filter for Atmospheric Data

591   Assimilation, Mon. Weather Rev., 144, 4489–4532, https://doi.org/10.1175/MWR-D-15-0440.1,

592   2016.

593   Houweling, S., Kaminski, T., Dentener, F., Lelieveld, J., and Heimann, M.: Inverse modeling of

594   methane sources and sinks using the adjoint of a global transport model, J. Geophys. Res. Atmos.,

595   104, 26137–26160, https://doi.org/10.1029/1999JD900428, 1999.

596   Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A

597   local ensemble transform Kalman filter, Phys. D Nonlinear Phenom., 230, 112–126,

598   https://doi.org/10.1016/j.physd.2006.11.008, 2007.

599   Ito, A.: Methane emission from pan-Arctic natural wetlands estimated using a process-based model,

600   1901–2016, Polar Sci., 21, 26–36, https://doi.org/10.1016/j.polar.2018.12.001, 2019.

601   Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F.,

602   Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., van Aardenne, J. A., Monni, S.,

603   Doering, U., Petrescu, A. M. R., Solazzo, E., and Oreggioni, G. D.: EDGAR v4.3.2 Global Atlas of

604   the three major greenhouse gas emissions for the period 1970–2012, Earth Syst. Sci. Data, 11, 959–

605   1002, https://doi.org/10.5194/essd-11-959-2019, 2019.

606   Kalnay, E. and Yang, S.-C.: Accelerating the spin-up of Ensemble Kalman Filtering, Q. J. R.

607   Meteorol. Soc., 136, 1644–1651, https://doi.org/10.1002/qj.652, 2010.

608   Kang, J.-S., Kalnay, E., Miyoshi, T., Liu, J., and Fung, I.: Estimation of surface carbon fluxes with an

609   advanced data assimilation methodology, J. Geophys. Res. Atmos., 117, D24101,

610   https://doi.org/10.1029/2012JD018259, 2012.

611   Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H.,

612   Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: General

613   Specifications and Basic Characteristics, J. Meteorol. Soc. Japan. Ser. II, 93, 5–48,

614   https://doi.org/10.2151/jmsj.2015-001, 2015.

615   Kotsuki, S., Ota, Y., and Miyoshi, T.: Adaptive covariance relaxation methods for ensemble data

616   assimilation: experiments in the real atmosphere, Q. J. R. Meteorol. Soc., 143, 2001–2015,

617   https://doi.org/10.1002/qj.3060, 2017.

618   Kotsuki, S., Pensoneault, A., Okazaki, A., and Miyoshi, T.: Weight structure of the Local Ensemble

619   Transform Kalman Filter: A case with an intermediate atmospheric general circulation model, Q. J. R.

620   Meteorol. Soc., 146, 3399–3415, https://doi.org/10.1002/qj.3852, 2020.

621   Liu, J., Bowman, K. W., and Lee, M.: Comparison between the Local Ensemble Transform Kalman

622   Filter (LETKF) and 4D-Var in atmospheric CO 2 flux inversion with the Goddard Earth Observing

623   System-Chem model and the observation impact diagnostics from the LETKF, J. Geophys. Res.

624   Atmos., 121, 13,066-13,087, https://doi.org/10.1002/2016JD025100, 2016.

625   Locatelli, R., Bousquet, P., Chevallier, F., Fortems-Cheney, A., Szopa, S., Saunois, M., Agusti-

626   Panareda, A., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M. P., Gloor, E., Houweling,

627   S., Kawa, S. R., Krol, M., Patra, P. K., Prinn, R. G., Rigby, M., Saito, R., and Wilson, C.: Impact of

628   transport model errors on the global and regional methane emissions estimated by inverse modelling,

629   Atmos. Chem. Phys., 13, 9917–9937, https://doi.org/10.5194/acp-13-9917-2013, 2013.

630   Lorente, A., Borsdorff, T., Butz, A., Hasekamp, O., aan de Brugh, J., Schneider, A., Wu, L., Hase, F.,

631   Kivi, R., Wunch, D., Pollard, D. F., Shiomi, K., Deutscher, N. M., Velazco, V. A., Roehl, C. M.,

632   Wennberg, P. O., Warneke, T., and Landgraf, J.: Methane retrieved from TROPOMI: improvement of

633   the data product and validation of the first 2 years of measurements, Atmos. Meas. Tech., 14, 665–

634   684, https://doi.org/10.5194/amt-14-665-2021, 2021.

635   Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Turner, A. J., Weitz, M., Wirth, T., Hight, C.,

DeFigueiredo, M., Desai, M., Schmeltz, R., Hockstad, L., Bloom, A. A., Bowman, K. W., Jeong, S., and Fischer, M. L.: Gridded National Inventory of U.S. Methane Emissions, Environ. Sci. Technol., 50, 13123–13133, https://doi.org/10.1021/acs.est.6b02878, 2016.

Meirink, J. F., Bergamaschi, P., and Krol, M. C.: Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion, Atmos. Chem. Phys., 8, 6341–6353, https://doi.org/10.5194/acp-8-6341-2008, 2008.

Miyazaki, K., Maki, T., Patra, P., and Nakazawa, T.: Assessing the impact of satellite, aircraft, and surface observations on CO 2 flux estimation using an ensemble-based 4-D data assimilation system, J. Geophys. Res., 116, D16306, https://doi.org/10.1029/2010JD015366, 2011.

Miyazaki, K., Eskes, H. J., and Sudo, K.: Global NOx emission estimates derived from an assimilation of OMI tropospheric NO2 columns, Atmos. Chem. Phys., 12, 2263–2288, https://doi.org/10.5194/acp-12-2263-2012, 2012.

Miyazaki, K., Sekiya, T., Fu, D., Bowman, K. W., Kulawik, S. S., Sudo, K., Walker, T., Kanaya, Y., Takigawa, M., Ogochi, K., Eskes, H., Boersma, K. F., Thompson, A. M., Gaubert, B., Barre, J., and Emmons, L. K.: Balance of Emission and Dynamical Controls on Ozone During the Korea-United States Air Quality Campaign From Multiconstituent Satellite Data Assimilation, J. Geophys. Res. Atmos., 124, 387–413, https://doi.org/10.1029/2018JD028912, 2019.

Miyoshi, T.: The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation with the Local Ensemble Transform Kalman Filter, Mon. Weather Rev., 139, 1519–1535, https://doi.org/10.1175/2010MWR3570.1, 2011.

Miyoshi, T., Yamane, S., and Enomoto, T.: Localizing the Error Covariance by Physical Distances within a Local Ensemble Transform Kalman Filter (LETKF), SOLA, 3, 89–92, https://doi.org/10.2151/sola.2007-023, 2007.

Miyoshi, T., Sato, Y., and Kadowaki, T.: Ensemble Kalman Filter and 4D-Var Intercomparison with the Japanese Operational Global Analysis and Prediction System, Mon. Weather Rev., 138, 2846–2866, https://doi.org/10.1175/2010MWR3209.1, 2010.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A.: A Local Ensemble Kalman Filter for Atmospheric Data Assimilation, https://doi.org/https://doi.org/10.48550/arXiv.physics/0203058, 2002.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A.: A local ensemble Kalman filter for atmospheric data assimilation, Tellus A Dyn. Meteorol. Oceanogr., 56, 415–428, https://doi.org/10.3402/tellusa.v56i5.14462, 2004.

668   Patra, P. K., Niwa, Y., Schuck, T. J., Brenninkmeijer, C. A. M., Machida, T., Matsueda, H., and

669   Sawa, Y.: Carbon balance of South Asia constrained by passenger aircraft CO2 measurements,

670   Atmos. Chem. Phys., 11, 4163–4175, https://doi.org/10.5194/acp-11-4163-2011, 2011a.

671   Patra, P. K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., Bian, H., Cameron-

672   Smith, P., Chipperfield, M. P., Corbin, K., Fortems-Cheiney, A., Fraser, A., Gloor, E., Hess, P., Ito,

673   A., Kawa, S. R., Law, R. M., Loh, Z., Maksyutov, S., Meng, L., Palmer, P. I., Prinn, R. G., Rigby, M.,

674   Saito, R., and Wilson, C.: TransCom model simulations of CH 4 and related species: Linking

675   transport, surface flux and chemical loss with CH 4 variability in the troposphere and lower

676   stratosphere, Atmos. Chem. Phys., 11, 12813–12837, https://doi.org/10.5194/acp-11-12813-2011,

677   2011b.

678   Patra, P. K., Takigawa, M., Watanabe, S., Chandra, N., Ishijima, K., and Yamashita, Y.: Improved

679   Chemical Tracer Simulation by MIROC4.0-based Atmospheric Chemistry-Transport Model

680   (MIROC4-ACTM), SOLA, 14, 91–96, https://doi.org/10.2151/sola.2018-016, 2018.

681   Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D.,

682   Bruhwiler, L., and Tans, P. P.: An ensemble data assimilation system to estimate CO 2 surface fluxes

683   from atmospheric trace gas observations, J. Geophys. Res., 110, D24304,

684   https://doi.org/10.1029/2005JD006157, 2005.

685   Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A.,

686   Dlugokencky, E. J., Houweling, S., Patra, P. K., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi,

687   P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Castaldi, S., Chandra, N.,

688   Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G., Frankenberg, C., Gedney, N.,

689   Hegglin, M. I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G.,

690   Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L.,

691   Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton, J. R., Morino, I.,

692   Müller, J., Murguia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R. J., Peng, C.,

693   Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J.

694   A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, L. P., Thornton, B. F., Tian, H., Tohjima,

695   Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der

696   Werf, G. R., Weiss, R. F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao,

697   Y., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: The Global Methane Budget 2000–2017, Earth Syst.

698   Sci. Data, 12, 1561–1623, https://doi.org/10.5194/essd-12-1561-2020, 2020.

699   Sekiya, T., Miyazaki, K., Ogochi, K., Sudo, K., Takigawa, M., Eskes, H., and Boersma, K. F.:

700   Impacts of Horizontal Resolution on Global Data Assimilation of Satellite Measurements for

701   Tropospheric Chemistry Analysis, J. Adv. Model. Earth Syst., 13,

702   https://doi.org/10.1029/2020MS002180, 2021.

703   Skachko, S., Ménard, R., Errera, Q., Christophe, Y., and Chabrillat, S.: EnKF and 4D-Var data

704   assimilation with chemical transport model BASCOE (version 05.06), Geosci. Model Dev., 9, 2893–

705   2908, https://doi.org/10.5194/gmd-9-2893-2016, 2016.

706   Szopa, S., V. Naik, B. Adhikary, P. Artaxo, T. Berntsen, W.D. Collins, S. Fuzzi, L. Gallardo, A.

707   Kiendler Scharr, Z. Klimont, H. Liao, N. U. and P. Zanis: Short-Lived Climate Forcers. In Climate

708   Change 2021, Clim. Chang. 2021 Phys. Sci. Basis. Contrib. Work. Gr. I to Sixth Assess. Rep.

709   Intergov. Panel Clim. Chang., 2021.

710   Tian, X., Xie, Z., Liu, Y., Cai, Z., Fu, Y., Zhang, H., and Feng, L.: A joint data assimilation system

711   (Tan-Tracker) to simultaneously estimate surface CO_2 fluxes and 3-D atmospheric CO_2

712   concentrations from observations, Atmos. Chem. Phys., 14, 13281–13293,

713   https://doi.org/10.5194/acp-14-13281-2014, 2014.

714   Watanabe, S., Miura, H., Sekiguchi, M., Nagashima, T., Sudo, K., Emori, S., and Kawamiya, M.:

715   Development of an atmospheric general circulation model for integrated Earth system modeling on

716   the Earth Simulator, J. Earth Simulator, 9, 27–35, 2008.

717   van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu,

718   M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global

719   fire emissions estimates during 1997–2016, Earth Syst. Sci. Data, 9, 697–720,

720   https://doi.org/10.5194/essd-9-697-2017, 2017.

721   Whitaker, J. S. and Hamill, T. M.: Evaluating Methods to Account for System Errors in Ensemble

722   Data Assimilation, Mon. Weather Rev., 140, 3078–3089, https://doi.org/10.1175/MWR-D-11-

723   00276.1, 2012.

724   Yoshida, Y., Kikuchi, N., Morino, I., Uchino, O., Oshchepkov, S., Bril, A., Saeki, T., Schutgens, N.,

725   Toon, G. C., Wunch, D., Roehl, C. M., Wennberg, P. O., Griffith, D. W. T., Deutscher, N. M.,

726   Warneke, T., Notholt, J., Robinson, J., Sherlock, V., Connor, B., Rettinger, M., Sussmann, R.,

727   Ahonen, P., Heikkinen, P., Kyrö, E., Mendonca, J., Strong, K., Hase, F., Dohe, S., and Yokota, T.:

728   Improvement of the retrieval algorithm for GOSAT SWIR XCO 2 and XCH 4 and their validation

729   using TCCON data, Atmos. Meas. Tech., 6, 1533–1547, https://doi.org/10.5194/amt-6-1533-2013,

730   2013.

731   Zhang, F., Snyder, C., and Sun, J.: Impacts of Initial Estimate and Observation Availability on

732   Convective-Scale Data Assimilation with an Ensemble Kalman Filter, Mon. Weather Rev., 132,

733   1238–1253, https://doi.org/10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2, 2004.

734   Zhang, Y., Jacob, D. J., Lu, X., Maasakkers, J. D., Scarpelli, T. R., Sheng, J.-X., Shen, L., Qu, Z.,

735   Sulprizio, M. P., Chang, J., Bloom, A. A., Ma, S., Worden, J., Parker, R. J., and Boesch, H.:

736   Attribution of the accelerating increase in atmospheric methane during 2010–2018 by inverse analysis

737   of GOSAT observations, Atmos. Chem. Phys., 21, 3643–3666, https://doi.org/10.5194/acp-21-3643-

738   2021, 2021.

739