Order of magnitude wall time improvement of variational methane inversions by physical parallelization: a demonstration using TM5-4DVAR

5   Sudhanshu Pandey[1], Sander Houweling[2], Arjo Segers[3]

[1]SRON Netherlands Institute for Space Research, Utrecht, the Netherlands
[2]Department of Earth Sciences, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[3]TNO Department of Climate, Air and Sustainability, Utrecht, The Netherlands

Correspondence to: Sudhanshu Pandey (s.pandey@sron.nl)

10   **Abstract.** Atmospheric inversions are used to constrain the emissions of trace gases from atmospheric mole fraction measurements. The variational (4DVAR) inversion approach allows optimization of the emissions at a much higher temporal and spatial resolution than the ensemble or analytical approaches but provides limited opportunities for scalable parallelization as the optimization is performed iteratively. Multidecadal variational inversions are used to optimally extract information from the long measurement records of long-lived atmospheric trace gases like carbon dioxide and methane.

15   However, the wall clock time needed—up to months— complicates these multidecadal inversions. The physical parallelization method introduced by Chevallier (2013) addresses this problem for CO_2 inversions by splitting the time period of the chemical transport model into blocks that are run in parallel. Here we present a new implementation of the physical parallelization for variational inversion (PPVI) approach that is suitable for methane inversions as it accounts for methane's atmospheric lifetime. The performance of PPVI is tested in an 11-year inversion using a TM5-4DVAR inversion setup that

20   assimilates surface observations to optimize methane emissions at grid-scale. We find that the PPVI inversion approach improves the wall clock time performance by a factor of 5 and shows excellent agreement with the posterior emissions of a full serial inversion with identical configuration (global mean emissions difference = 0.06 % with an interannual variation correlation R = 99 %; regional mean emission difference < 5 % and interannual variation R > 0.95). The wall clock time improvement using the PPVI method increases with the size of the inversion period. The PPVI approach is planned to be

25   used in future releases of the CAMS (Copernicus Atmosphere Monitoring Service) multidecadal methane reanalysis.

1 Introduction

Methane (CH_4) is the second most important greenhouse gas after carbon dioxide (CO_2), and its atmospheric abundance has increased more than 250 % since preindustrial times. Due to its strong global warming potential, it is responsible for 25 % of anthropogenic radiative forcing in spite of its 200 times lower abundance than CO_2 (Myhre et al., 2013). Unlike the

30   relatively steady increase in CO_2, mainly due to fossil fuel emissions, the methane observational record shows remarkable

variability in growth rate. The causes of these variations are still debated (Rigby et al., 2017; Schaefer et al., 2016, Worden et al., 2017; Pandey et al., 2017).

Reducing anthropogenic methane emissions has been recognized as an important requirement for achieving the 2015 Paris
35   Agreement target of limiting global temperature rise to below 2° C relative to pre-industrial times (Ganesan et al., 2019; Nisbet et al., 2018). Climate change mitigation and adaptation strategies require reliable knowledge of the methane budget. The methane emissions have been estimated using multidecadal inversions, which optimally combine the information in atmospheric observations and bottom-up emission estimates (process-based models and inventories) along with corresponding error characteristics. Inversions using CTM (chemical transport model) are used to disentangling the
40   influences of atmospheric transport, and sources and sinks on the observed mole fractions (Naus et al., 2019, Pandey et al., 2019). Such multidecadal inversions have been performed, for example, for the "The Global Methane Budget 2000–2017" that was published recently by the Saunois et al. (2020). This study made methane emissions available from nine different inversion setups. The methane emissions reanalysis project under the Copernicus Atmosphere Monitoring Service (CAMS) performs multidecadal inversions using the TM5-4DVAR variational approach to provide regularly updated gridded
45   methane emissions (Segers and Houweling, 2020).

Atmospheric inversions that estimate CH_4 emissions adjust a state vector (consisting of emissions or emission correction factors) to improve the agreement between model simulations and observations. These inversions use a CTM to simulate the spatiotemporal distribution of the tracer in the atmosphere for a given set of emissions while also accounting for its
50   atmospheric sink. A cost function is defined based on the difference between the modelled and observed mole fractions as well as the magnitude of the emission adjustments. The solution of the inverse problem is posterior emission vector, which minimizes the cost function. There are three main minimization approaches used in atmospheric inverse modelling: analytical, ensemble and variational. The analytical approach is based on a closed-form solution of Bayes' theorem (Gurney et al., 2002). It requires the calculation of observation-sensitivies of each of the state vector elements separately. This leads
55   to a large computational cost and restricts the approach's usage to inversion problems with small sized state vectors. The ensemble approach improves the computational performance by representing the sensitivities by a statistical ensemble (Peters et al., 2005). Still only a relatively small sized state vector can currently be afforded using this approach. The variational approach was introduced to lift the state vector size restriction, using the adjoint of the CTM (Chevallier et al., 2005).

60

In the variational approach, the minimum of the cost function is computed using an iterative procedure that comprises of a forward and an adjoint CTM run. As each iteration uses the output of the previous iteration, there are limited opportunities for scalable parallelization in variational inversions, and these calculations can take months depending on the spatial and

temporal resolution of the inversion. The long wall-clock time limits the resolution, the maximum time range, or the number of iterations used in multidecadal inversions.

To improve the computational efficiency of multidecadal variational CO_2 inversions, Chevallier (2013) introduced the physical parallelization method in which the forward and adjoint CTM runs within each iteration are divided into blocks and run in parallel. Correction factors are applied after the CTM block runs are finished to account for changes in the background mole fractions due to net emission changes in earlier blocks. This method reduces the wall-clock time by an order of magnitude (seven-fold improvement for a 32-year inversion) while keeping the inversion-derived emission adjustments statistically consistent with a serial inversion. However, their method cannot be applied directly to methane as it does not account for the limited chemical lifetime of methane of about 10 years, due to oxidation by the OH radicals in the atmosphere. Here we report an extension of the method that accounts for the atmospheric lifetime. The intention is to use this new implementation for the CAMS methane flux reanalysis, which aims to provide every year an updated multi-decadal inversion within a production window of only a few months. The method is referred to as PPVI (physical parallelization for variational inversion) from here on.

In the next section, we present the PPVI method. In Section 3, we test the performance of the PPVI method using 11-year test inversions. We compare the wall clock time and optimized emissions of a serial inversion with a PPVI inversion. We discuss the current CAMS methane inversion setup, and possible improvements and applications of the PPVI method in Section 4. Our conclusions are summarized in Section 5.

2 Physical parallelization for variational inversions

The solution of a methane inverse problems is calculated by minimizing the cost function of the state vector x:

$$J(x) = \frac{1}{2}(x - x^a)B^{-1}(x - x^a) + \frac{1}{2}(H(x) - y)R^{-1}(H(x) - y) \quad \ldots\ldots\ldots(1)$$

In here, y is the observation vector, x^a is the a priori vector. The observation operator H consists of the a CTM run that simulates methane mole fractions at time and location of y. B and R are the error covariance matrices of the a priori emissions and the observations, respectively. In a variational inversion setup, the posterior solution x^p of Eq. (1) is found by minimizing J using an iterative procedure that estimates a new x^i in each iteration i using the gradient of J:

$$\nabla J(x^i) = B^{-1}(x^i - x^a) + H^T R^{-1}(H(x^i) - y) \quad \ldots\ldots\ldots(2)$$

95

where, H^T represents the adjoint operator, which is implemented using the adjoint code of the CTM. The inversion finishes when a predefined convergence criterion is met, such as a desired gradient norm reduction or simply a maximum number of iterations.

100 In a serial variational inversion iteration i, the CTM H simulates mole fraction vector m^i for observation vector y using the initial mole fraction field c_0 and emissions x^i:

$$m^i = H\ (c_0\ , x^i)\ \ldots\ldots\ldots(3)$$

105 In the PPVI method, the full period of the CTM is broken into r overlapping time blocks, which are run in parallel. Figure 1 shows a schematic diagram of the main steps of the PPVI method used in the forward mode to calculate m^i. At the start of the inversion, a CTM run is performed serially (without blocks) to calculate initial mole fraction fields c_k^a for each block k using the a priori emissions x^a. This CTM run can be performed at coarser resolution than the main inversion to save time. The overlaps between consecutive blocks are needed for methane perturbations to uniformly distribute over the spatial
110 domain of the CTM, such that each perturbation could be diagnosed by at least some observation sites. The mole fraction vector m_k^i for the observations y_k could now be calculated by using the small CTM block H_k with emissions x_k^i and a correction factor vector n_k^i which accounts for the state vector innovation in the preceding block.

$$m_k^i = H_k\left(c_k^a, x_k^i\right) + n_k^i\ \ldots\ldots\ldots(4)$$

115

The correction n_k^i is calculated using an emission to mole fraction conversion factor f = 0.361 ppb/Tg and a methane sink operator S:

$$n_k^i = \mathbf{H}_k\ \sum_{l=1}^{k-1} S_{l,k}\ \text{f}\left(x_l^i - x_l^a\right),\ \ldots\ldots\ldots(5)$$

120

In here, $\mathbf{H}_k$ is the CTM block sensitivity to a uniform initial mole fraction field perturbation, which is calculated at the start of the inversion by running each block with an initial unit mole fraction field and a zero emissions field. $S_{l,k}$ account for the impact of atmospheric sinks on the emission perturbations during block l till the start of the block k. Thereon the impact of atmospheric sink, as well as atmospheric transport, is accounted for by $\mathbf{H}_k$. We parameterize $S_{l,k}$ with an e-folding decay
125 function with atmospheric lifetime of methane of 9 years, which was found to be sufficient for our test inversion (Section 3).

In each iteration of a variational inversion, the modelled mole fractions from the forward CTM run are used to calculate departures, i.e., the difference between observations and the modelled mole fractions scaled with the respective errors. Thereafter, the adjoint CTM uses these departures to calculate the local gradient of the cost function (Equation 2). In the

130 PPVI method, the adjoint CTM blocks are kept the same as the forward CTM blocks and Equation 2 is applied in adjoint mode. First, each adjoint block is run with the respective departures. Then, the modelled adjoint sensitivities δx_l^i of block k are adjusted for the effects of departures of successive blocks by adding adjoint corrections δn_l^i calculated using as follows:

$$\delta n_k^i = \ \text{f} \ \sum_{l=k+1}^{r} S_{k,l}^* \ \ \mathbf{H}_l^* \ \ (\delta m_l^i) \ \ ..... (6)$$

135

In here '*' represents the adjoint of an operator. $\mathbf{H}_k^*$ is applied to the departures of each block δm_l^i to specify adjoint mole fraction increments. $S_{k,l}^*$ are applied to the adjoint mole fraction increments, and the result is added to each element δx_l^i. The correct implementation of the adjoint part of the PPVI method was verified using the adjoint test.

140 In a PPVI inversion, the initial mole fraction field c_0 needs to be consistent with the observations as a discrepancy between the two leads to large emission adjustments in the early months of the inversion period. This issue can easily be dealt with in a serial inversion using a spin-up period and rejecting this period from the posterior solution. However, in a PPVI inversion, the large emission adjustments may result in large correction factors, which increases the error in the PPVI approximation (Equation 4 & 5). This can be avoided by taking a realistic c_0 from another inversion covering the period before the PPVI

145 inversion.

In summary, the main steps of a PPVI inversion are as follows:

1. Construct an initial mole fraction field c_0 consistent with observations at the start of the inversion.
2. Divide the full period of the inversion into r over-lapping time blocks.
150 3. Calculate the initial mole fraction fields (c_k^0) for each block by running the forward CTM serially with the a priori emissions x^a and saving the simulated mole fraction fields at the start time of each block.
4. Calculate the CTM block sensitivities ($\mathbf{H}_k$) by running the CTM blocks with a unity initial mole fraction field and zero emissions, and sampling the model output at the observations.
5. Perform the inversion by iteratively minimizing the cost function until the convergence criteria is met using a
155 forward and an adjoint run in each iteration:
    a. Forward run:
       i. Run the forward CTM for each block in parallel with the initial mole fraction fields from the step 3.
       ii. Account for the emission differences from the a priori in preceding blocks, by applying the mole fraction corrections n_k^i (Equation 4).

b. Adjoint run:

    i. Run the adjoint CTM for each block in parallel to calculate the adjoint emissions sensitivities δx_i^i.

    ii. Apply the adjoint corrections δn_i^i to account for the effect of departures in successive blocks (Equation 6).

3 PPVI Performance test

In this section, we evaluate the performance of the PPVI method against a serial inversion. We perform the inversions for a 11-year period (1999-2010) using the TM5-4DVAR inversion system (Bergamaschi et al., 2010; Meirink et al., 2008), which consists of the TM5 (Transport Model version 5; Krol et al., 2005) with the settings used in Pandey et al. (2016). TM5 is run at $6° \times 4°$ horizontal resolution and 25 vertical hybrid sigma-pressure levels from the surface to the top of the atmosphere. We optimize a single category ('total') of methane emissions at $6° \times 4°$ spatial resolution and monthly temporal resolution. The posterior emissions of the two inversions are compared after integrating over the TRANSCOM regions shown in Figure 2a. The inversion assimilates surface observations from the NOAA Earth System Research Laboratory (ESRL) global cooperative air sampling network at on- and off-shore sites (Dlugokencky et al., 2011; Dlugokencky et al., 2020). The locations of the observations are shown in Figure 2b. The prior covariance matrix is constructed assuming relative emission uncertainties of 50% per grid box per month. The emissions are assumed to be correlated temporally using an exponential correlation function with e-folding time scale of 3 months, and spatially with a Gaussian correlation function using a length scale of 500 km (Houweling et al., 2014). Uncertainties of 1.4 ppb are assigned to the CH_4 observations. Our system also assigns a modeling representation error based on simulated local mole fractions gradients (Basu et al., 2013). The prior emissions, same as in Pandey et al. (2016), of 2008 are applied to every year in the inversion time period, hence there is no interannual variability in the prior emissions.

In the PPVI inversion, we divide the inversion period of 1999-2010, into 11 blocks of 21 months with 9 months overlap between successive blocks. Effectively, each block provides modelled mole fractions for one year. The successful implementations of the PPVI method on this TM5-4DVAR setup was verified using the adjoint and gradient tests. The next section (3.1) compares the observation-model mismatches and posterior emissions differences between the two inversions, and the section thereafter (3.2) presents the wall clock time improvement by the PPVI method.

3.1 Emission estimation errors

190 Here we compare the observation fit and the posterior emissions of the PPVI and serial inversions. Figure 3 shows the timeseries of the posterior simulations and the observations for two background sites, one for each hemisphere: Barrow (Alaska) and South Pole. The observation-model RMSE (root mean square error) for Barrow (78 ppb) is 3 time higher than for South Pole (28 pbb). The Barrow observations shows a larger high-frequency variation than the South Pole as the Northern Hemisphere station is closer to methane sources. The PPVI results are good agreement with the results from serial:

195 the RMSE between the two is 2 ppb and 1 pbb respectively for Barrow and South pole, which are only 2.5 % and 3.2 % of the initial observation-prior mismatch. This shows that, starting from an identical prior, the PPVI inversion is able to match the observations about as good as the serial inversion. Figure 4 shows the probability density functions of the observation-model mismatch weighted with the uncertainties used in the inversion. The observation-prior mismatch is -6.7 ± 6, with negative mean because the 2008 bottom-up emissions used as prior are larger than the mean posterior emission over 1999-

200 2010. In the posterior solution of the serial inversion, the mismatch is reduced to –0.06 ± 1.24. PPVI mismatch (–0.06 ± 1.26) is also very small and similar to the serial inversion. The mismatch between the PPVI and serial inversions (0.005 ± 0.23) is an order of magnitude smaller than the observation-prior mismatch, which shows that the implementation of PPVI method does not have a significant impact on the inversion system's ability to fit the observations. For both inversions, the good fit to the observations also confirms that a gradient norm reduction of 1000 is sufficient.

205

A good agreement between observations and posterior models do not necessary mean that the inversions have produced similar posterior emissions. The parallelized transport model of the PPVI is a simplification of the "perfect" transport model of used in the serial inversion. If the impact of this simplification is small, the posterior emissions of the two inversions should be in good agreement. Figure 5 shows the mean emission estimates of the inversions integrated over globe and the

210 TRANSCOM regions. The mean global emissions of the PPVI and serial inversions are in excellent agreement with < 0.3 Tg yr^{-1} (0.05 %) difference. The global methane emissions are generally well constrained by the NOAA observations in the serial and PPVI inversions meaning that the error in the PPVI approach does not impact the constraint on emissions at the global scale. The performance of PPVI in other TRANSCOM regions is also good for mean emissions with < 5 % deviation from the serial inversion. On average, the deviations are within 30% of the posterior uncertainty. Figure 6 shows the

215 interannual variability of the emission estimates. Due to the large observational constraint, the global emissions of the two inversions show the best agreement with a correlation coefficient $R = 0.99$. Over the TRANSCOM regions, the agreements are also good with $R > 0.95$.

3.2 Computational cost

220   Table 1 compares the wall clock times needed for the PPVI and serial inversions. Our TM5 model runs use OpenMP parallelization and gave best wall clock performance on 4 CPUs on a single node. Using more CPUs reduces the performance as the communication overhead within the CPUs becomes the bottleneck. (Note that the TM5-MP version described in Williams et al., 2017, with improved parallel scaling, was not used in this study). In this configuration, a forward or adjoint TM5 run of one year takes about 15 minutes. Hence an iteration of the serial inversion, consisting of 11

225   years forward and adjoint runs, requires 5 hours. PPVI inversion iterations are performed in 11 parallel blocks of 21 months each on 4 CPUs. A single PPVI iteration takes 55 mins, which is > 5 time faster than the serial inversion. Both inversions achieved a gradient norm reduction of 1000 in 19 iterations. The PPVI inversion runtime is given in Table 1, including the time needed for (1) a serial TM5 forward run for the initial mole fraction fields (2) a block run for the initial mole fraction sensitivities. Overall, the PPVI inversion takes 20 hours, or 5 time less than the serial inversion (101 hours). Table 1 also

230   provides an estimate of the wall clock time of a hypothetical 35-year inversions using the TM5-4DVAR setup. For such a long period, a PPVI inversion would be 15 times faster.

   Overall, we find that the PPVI method, which account for the atmospheric lifetime of methane, is able to reproduce the posterior emissions of a traditional 11 years serial inversion well within its uncertainties in 5 times less wall clock time.

235

4 Discussion

4.1 Current CAMS inversion setup

   In the future, the PPVI method will be implemented in the CAMS multidecadal methane emissions reanalysis setup. The European Commission has anticipated the need for reliable information about atmospheric composition of greenhouse gases

240   through development of numerical systems that combine sophisticated physical models with measurements from a wide range of observing systems for an operational service, which is being implemented. The current CAMS methane flux reanalysis product (Segers and Houweling, 2020) uses the TM5-4DVAR inverse modelling system and provides measurement-informed monthly methane emission estimates. The product has two sets of methane emissions: (1) release v19r1 for 1990-2019 using surface observation; (2) release v19r1s for 2010-2019 using surface and GOSAT satellite

245   observations. The surface observations are mainly from the NOAA network (Dlugokencky et al., 2011). Methane emissions are optimized at $3° \times 2°$ spatial resolution and monthly temporal resolution using TM5 with 34 vertical layers. If performed in serial mode each iteration of the 1990-2019 inversion would take about 5-10 days, and the full inversion will require multiple months to finish. Segers and Houweling (2020) circumvent this issue by breaking the full inversion into smaller inversions of 3-year time windows that are performed in parallel. The target inversion on high resolution ($3° \times 2°$ degrees, 34

250   layers) is preceded by a coarse resolution inversion (6° × 4°, 25 layers) that provides the initial mole fraction fields and is processed serially. The high-resolution inversion optimizes only the emissions and uses initial mole fractions for each 3-year block obtained from mole fraction fields of a coarse resolution inversion, which optimizes both emission and initial mole fractions. The 1990-2019 inversion using this approach still takes 3-4 months to finish, and requires about 40 smaller inversions to provide the end result. These numbers depend of course on the parallel efficiency of the model and the

255   computing server, but even if these are improved, the need for a serial sequence of inversions to provide a timeseries of initial mole fractions imposes a limitation to model resolution that can be used. With the implementation of the PPVI method presented in this study, the computational performance of the CAMS reanalysis inversions will improve if future.

4.2 Possible further improvements

260   In the PPVI method, the wall clock time of a CTM run in an inversion iteration is reduced by physical parallelization of CTM into blocks. To account for changes in the background mole fractions due to emission changes in pervious blocks, the sink operator S, CTM block sensitivity $\mathbf{H}$ and the overlaps between the consecutive blocks are used. In our test experiment, S is assumed to be a e-folding decay function with atmospheric lifetime of methane of 9 years, which we find is sufficient for the annually-repeating OH field used in our 11-year CTM runs. This might not be the case for multi-decadal inversions,

265   in which the methane lifetime will vary due to climatological influences, as well as possible trends and interannual variations in the hydroxyl radical abundance. In such cases, S can be defined as a function of an annual lifetime vector for the specific CTM run. The lifetime vector can be calculated as the ratio of the annual sink and mean methane burden simulated in a serial CTM run.

270   The overlap period between the blocks in the PPVI method allows a uniform mole fraction perturbations to mix within the CTM domain according the atmospheric transport. We use a 9-month overlap in our test experiment and found it to be sufficient for emissions from large regions at annual scale (Figure 5 & 6) that are optimized using the surface observations. A shorter overlap, which improves the computational efficiently but reduces the accuracy of the physical CTM parallelization, could be used depending on the scales that are addressed by the inversion.

275 ## 4.3 Methane sink optimization

The hydroxyl radical OH is the main sink of methane in atmosphere. Zhang et al. (2018) showed that the satellite-observed atmospheric signature of the methane sink from oxidation by OH is sufficiently distinct from that of methane emissions, hence OH mole fractions can be optimized using synthetic SWIR and TIR satellite observations. Following up on this, Maasakkers et al. (2019) and Zhang et al. (2021) used methane observations from the GOSAT satellite to optimize

280   atmospheric OH fields along with methane emissions. These studies assume a quasi-linearity as the changes to the methane mole fractions and OH are small in an inversion. Under such conditions, the PPVI method can also be implemented in

inversions optimizing OH. In such a PPVI implementation, the methane lifetimes in the *S* operator would be scaled in each iteration to reflect the adjustments to the OH mole fractions. Such an implementation can also be used in inversions optimizing OH using methyl chloroform (CH_3CCl_3) and a CTM, for example, Naus et al. (2021).

5 Conclusions

Regular surface observations of methane mole fractions started in the early 1984, and by now the measurement record spans more than 35 years (Dlugokencky et al., 2011). An atmospheric inversion with a very large state vector is needed to properly utilize the information in such long measurement records. The variational inversion approach allows for optimization of state vector of a larger size than the ensemble or analytical approaches. However, each iteration step of a variational inversion uses the CTM output of the previous iteration, limiting the opportunity for scalable parallelization. At the same time, an increase in the spatio-temporal resolution of CTMs, which is needed to take full advantage of the rapidly improving precision and coverage of surface and satellite measurements, results in an exponential increase in wall clock time.

We have developed the PPVI method which improves the wall clock time of variational methane inversions by the application of physical parallelization while accounting for the atmospheric lifetime in forward and adjoint variational modes. We have tested the performance of this method using a TM5-4DVAR inverse modeling setup that consists of a traditional serial inversion and a PPVI inversion of identical configuration performed for a period of 11 years. The PPVI method reduced the wall clock time by 5 times and showed excellent agreement with the posterior emissions of the serial inversion. The wall clock time improvement of using PPVI will be even larger for longer inversions, for example, 15 times for a 35 years inversion. The PPVI method makes multi-decadal inversions more feasible. It will be implemented in the CAMS reanalysis setup which provides multidecadal emission estimates by assimilating surface and satellite observations.

Data Availability. NOAA GMD ESRL methane mole fraction observations are available at https://www.esrl.noaa.gov/gmd/dv/data/. last access: 27-03-2021.

Code availability. The TM5 model is described in detail on http://tm5.sourceforge.net/ and the source codes of the TM5-4DVAR inversion system are available at https://sourceforge.net/projects/tm5/.

Author contributions. The study was designed by SH and FC. SP and AJ developed the PPVI. SP implemented the PPVI method on TM5-4DVAR and did the performance test simulations. SP wrote the manuscript using contributions from all the co-authors.

315 *Acknowledgements*. We thank the efforts of NOAA and other surface observations networks for producing and maintaining the vital long record of global methane surface observations. The computations for this study were carried out on the Dutch national supercomputer Cartesius (https://userinfo.surfsara.nl/systems/cartesius;_last_access:_27-03-2020) maintained by SURFSara. SP and AJ were funded for this study by the Copernicus Atmosphere Monitoring Service, implemented by the European Centre for Medium-Range Weather Forecasts on behalf of the European Commission (grant no. CAMS73).
320

Competing interests. The authors declare that they have no competing interests.

References

Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M.,
325 Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global CO2 fluxes estimated from GOSAT retrievals of total column CO2, Atmos. Chem. Phys., 13, 8695–8717, doi:10.5194/acp-13-8695- 2013, 2013.

Bergamaschi, P., Krol, M., Meirink, J. F., Dentener, F., Segers, A., van Aardenne, J., Monni, S., Vermeulen, A. T., Schmidt, M., Ramonet, M., Yver, C., Meinhardt, F., Nisbet, E. G., Fisher, R. E., O'Doherty, S., and Dlugokencky, E. J.: Inverse
330 modeling of European CH4 emissions 2001–2006, J. Geophys. Res., 115, D22309, doi:10.1029/2010JD014180, 2010.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F. M., Chédin, A. and Ciais, P.: Inferring CO$_2$ sources and sinks from satellite observations: Method and application to TOVS data, J. Geophys. Res. Atmos., 110(24), 1–13, doi:10.1029/2005JD006390, 2005.
335
Chevallier, F.: On the parallelization of atmospheric inversions of CO_2 surface fluxes within a variational framework, Geosci. Model Dev., 6, 783–790, https://doi.org/10.5194/gmd-6-783-2013, 2013.

Dlugokencky, E. J., Nisbet, E. G., Fisher, R. and Lowry, D.: Global atmospheric methane: Budget, changes and dangers,
340 Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., 369(1943), 2058–2072, doi:10.1098/rsta.2010.0341, 2011.

Dlugokencky, E. J., Crotwell, A. M., Mund, J. W., Crotwell, M. J., and Thoning, K. W.: Atmospheric Methane Dry Air Mole Fractions from the NOAA GML Carbon Cycle Cooperative Global Air Sampling Network, Version: 2020-07, https://doi.org/10.15138/VNCZ-M766, 2020.
345

Ganesan, A. L., Schwietzke, S., Poulter, B., Arnold, T., Lan, X., Rigby, M., Vogel, F. R., van der Werf, G. R., Janssens-Maenhout, G., Boesch, H., Pandey, S., Manning, A. J., Jackson, R. B., Nisbet, E. G. and Manning, M. R.: Advancing Scientific Understanding of the Global Methane Budget in Support of the Paris Agreement, Global Biogeochem. Cycles, 33(12), 1475–1512, doi:10.1029/2018GB006065, 2019.

350

Gurney, K., Law, R., Denning, A., Rayner, P., Baker, D., Bousquet, P., Bruhwiler, L., Chen, H., Ciais, P., Fan, S., Fung, I., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyu- tov, S., Ken, M., Peylin, P., Prather, M., Pak, B. C., Rander- son, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.-W.: Towards robust regional estimates of CO2 sources and sinks using atmospheric transport models, Nature, 415, 626–630, doi:10.1038/415626a, 2002.

355

Houweling, S., Krol, M., Bergamaschi, P., Frankenberg, C., Dlugokencky, E. J., Morino, I., Notholt, J., Sherlock, V., Wunch, D., Beck, V., Gerbig, C., Chen, H., Kort, E. A., Röckmann, T. and Aben, I.: A multi-year methane inversion using SCIAMACHY, accounting for systematic errors using TCCON measurements, Atmos. Chem. Phys., 14(8), 3991–4012, doi:10.5194/acp-14-3991-2014, 2014.

360

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., and Bergamaschi, P.: The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, Atmos. Chem. Phys., 5, 417– 432, doi:10.5194/acp-5-417-2005, 2005.

365 Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J. X., Zhang, Y., Hersher, M., Anthony Bloom, A., Bowman, K. W., Worden, J. R., Janssens-Maenhout, G. and Parker, R. J.: Global distribution of methane emissions, emission trends, and OH concentrations and trends inferred from an inversion of GOSAT satellite data for 2010-2015, Atmos. Chem. Phys., 19(11), 7859–7881, doi:10.5194/acp-19-7859-2019, 2019.

370 Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and Natural Radiative Forcing, in: Climate Change 2013: The Physical Science Basis. Contribution ofWorking Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2013.

375 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, doi:10.5194/acp-8-6341-2008, 2008.

Naus, S., Montzka, S. A., Pandey, S., Basu, S., Dlugokencky, E. J. and Krol, M.: Constraints and biases in a tropospheric
380 two-box model of OH, Atmos. Chem. Phys., 407–424, 2019.

Naus, S., Montzka, S. A., Patra, P. K. and Krol, M. C.: A three-dimensional-model inversion of methyl chloroform to
constrain the atmospheric oxidative capacity, Atmos. Chem. Phys., 21(6), 4809–4824, doi:10.5194/acp-21-4809-2021, 2021.

385 Nisbet, E. G., Manning, M. R., Dlugokencky, E. J., Fisher, R. E., Lowry, D., Michel, S. E., … and White, J. W. C.: Very
Strong Atmospheric Methane Growth in the 4 Years 2014–2017: Implications for the Paris Agreement, Global Biogeochem.
Cycles, 33(3), 318–342, doi:10.1029/2018GB006009, 2019.

Pandey, S., Houweling, S., Krol, M., Aben, I., Monteil, G., Nechita-Banda, N., Dlugokencky, E. J., Detmers, R., Hasekamp,
390 O., Xu, X., Riley, W. J., Poulter, B., Zhang, Z., McDonald, K. C., White, J. W. C., Bousquet, P. and Röckmann, T.:
Enhanced methane emissions from tropical wetlands during the 2011 la Niña, Sci. Rep., 7 (October 2016), 1–8,
doi:10.1038/srep45759, 2017.

Pandey, S., Houweling, S., Krol, M., Aben, I., Chevallier, F., Dlugokencky, E. J., Gatti, L. V., Gloor, E., Miller, J. B.,
395 Detmers, R., Machida, T. and Röckmann, T.: Inverse modeling of GOSAT-retrieved ratios of total column CH4 and CO2 for
2009 and 2010, Atmos. Chem. Phys., 16(8), 5043–5062, doi:10.5194/acp-16-5043-2016, 2016.

Pandey, S., Houweling, S., Krol, M. and Aben, I.: Influence of Atmospheric Transport on Estimates of Variability in the
Global Methane Burden Geophysical Research Letters, Geophys. Res. Lett., doi:10.1029/2018GL081092, 2019.
400
Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D., Bruhwiler, L., and Tans, P. P.:
An ensemble data assimilation system to estimate CO2 surface fluxes from atmospheric trace gas observations, J. Geophys.
Res., 110, D24304, doi:10.1029/2005JD006157, 2005.

405 Rigby, M., Montzka, S. A., Prinn, R. G., White, J. W. C., Young, D., O'Doherty, S., Lunt, M. F., Ganesan, A. L., Manning,
A. J., Simmonds, P. G., Salameh, P. K., Harth, C. M., Mühle, J., Weiss, R. F., Fraser, P. J., Steele, L. P., Krummel, P. B.,
McCulloch, A. and Park, S.: Role of atmospheric oxidation in recent methane growth, Proc. Natl. Acad. Sci., 114(21), 5373–
5377, doi:10.1073/pnas.1616426114, 2017.

410 Saunois, M., R. Stavert, A., Poulter, B., Bousquet, P., G. Canadell, J., B. Jackson, … and Zhuang, Q.: The global methane
budget 2000-2017, Earth Syst. Sci. Data, 12(3), 1561–1623, doi:10.5194/essd-12-1561-2020, 2020.

Segers, A. and Houweling, S.: Validation of the CH4 surface flux inversion - reanalysis 1990-2017, 1–40, 2020. Available at https://atmosphere.copernicus.eu/sites/default/files/2021-02/CAMS73_2018SC2_D73.2.4.1-

415   2020_202012_validation_CH4_1990-2019_v2.pdf Last access: 27-3-2021

Schaefer, H., Fletcher, S. E. M., Veidt, C., Lassey, K. R., Brailsford, G. W., Bromley, T. M., Dlugokencky, E. J., Michel, S. E., Miller, J. B., Levin, I., Lowe, D. C., Martin, R. J., Vaughn, B. H., and White, J. W. C.: A 21st century shift from fossil-fuel to biogenic methane emissions indicated by 13CH4, Science, 10.1126/science.aad2705, 2016.

420

Worden, J. R., Bloom, A. A., Pandey, S., Jiang, Z., Worden, H. M., Walker, T. W., Houweling, S. and Röckmann, T.: Reduced biomass burning emissions reconcile conflicting estimates of the post-2006 atmospheric methane budget, Nat. Commun., 8(1), 2227, doi:10.1038/s41467-017-02246-0, 2017.

425   Williams, J. E., Boersma, K. F., Le Sager, P., and Verstraeten, W. W.: The high-resolution version of TM5-MP for optimized satellite retrievals: description and validation, Geosci. Model Dev., 10, 721–750, https://doi.org/10.5194/gmd-10-721-2017, 2017.

Zhang, Y., Jacob, D. J., Maasakkers, J. D., Sulprizio, M. P., Sheng, J.-X., Gautam, R., and Worden, J.: Monitoring global

430   tropospheric OH concentrations using satellite observations of atmospheric methane, Atmos. Chem. Phys., 18, 15959–15973, https://doi.org/10.5194/acp-18-15959-2018, 2018.

Zhang, Y., J. Jacob, D., Lu, X., D. Maasakkers, J., R. Scarpelli, T., Sheng, J. X., Shen, L., Qu, Z., P. Sulprizio, M., Chang, J., Anthony Bloom, A., Ma, S., Worden, J., J. Parker, R. and Boesch, H.: Attribution of the accelerating increase in atmospheric

435   methane during 2010-2018 by inverse analysis of GOSAT observations, Atmos. Chem. Phys., 21(5), 3643–3666, doi:10.5194/acp-21-3643-2021, 2021.

440 **Table 1: Wall clock time comparison for inversions performed in this study. Wall clock time projections for a hypothetical 35 years inversion are also given.**

Model runs		Serial	PPVI
One year forward or adjoint run		15 minutes	
1999-2010 inversion	1 iteration (forward + adjoint TM5 run)	5 hours	55 minutes
	Inversion with 19 iterations	101 hours	20 hours
1985-2020 inversion*	1 iteration (forward + adjoint TM5 run)	16 hours	55 minutes
	Inversion with 50 iterations	34 days	56 hours

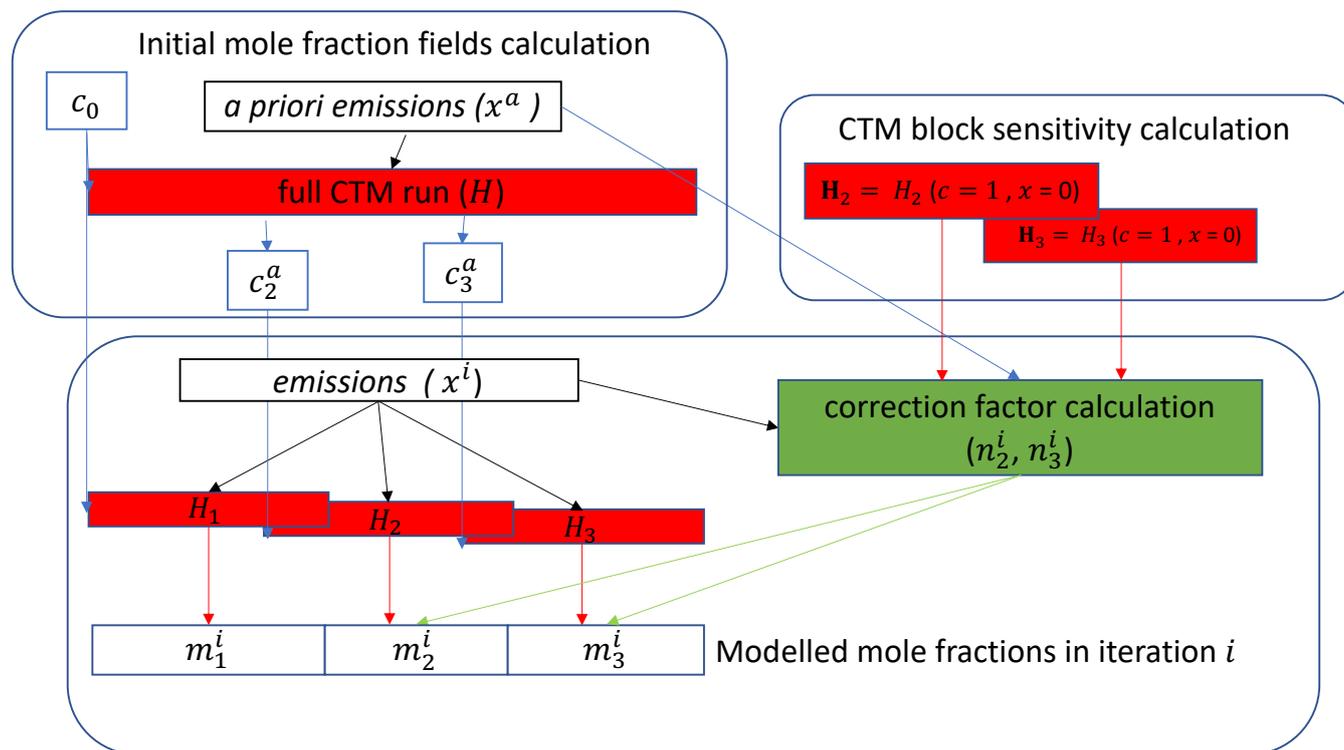
*Projection based on the 1999-2010 inversion

445

Figure 1: Schematic diagram of the PPVI method in forward mode used to calculate modelled mole fractions m^i of iteration i. The subscripts represent the time block number (except for c_0, which is the initial mole fraction field at the start of the inversion). For the block 1, the initial mole fraction field $c_1^a = c_0$, and no correction factor n^i is used.
450   **The overlap between the blocks (H_1, H_2, H_3) represent the block overlap period, where the modelled mole fractions from the preceding block are used in m^i.**

455

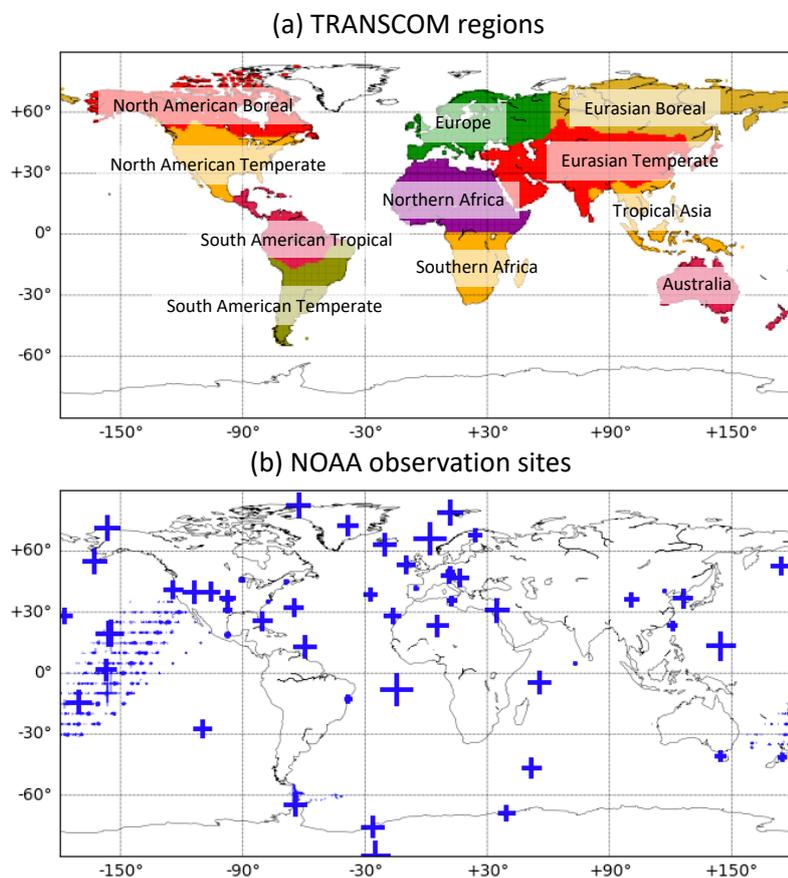



Figure 2: (a) Definition of the TRANSCOM regions (Gurney et al., 2002). **(b)** Locations of NOAA methane observation sites used in this study. The size of the symbol "+" is proportional to the number of observations assimilated from each site.

460

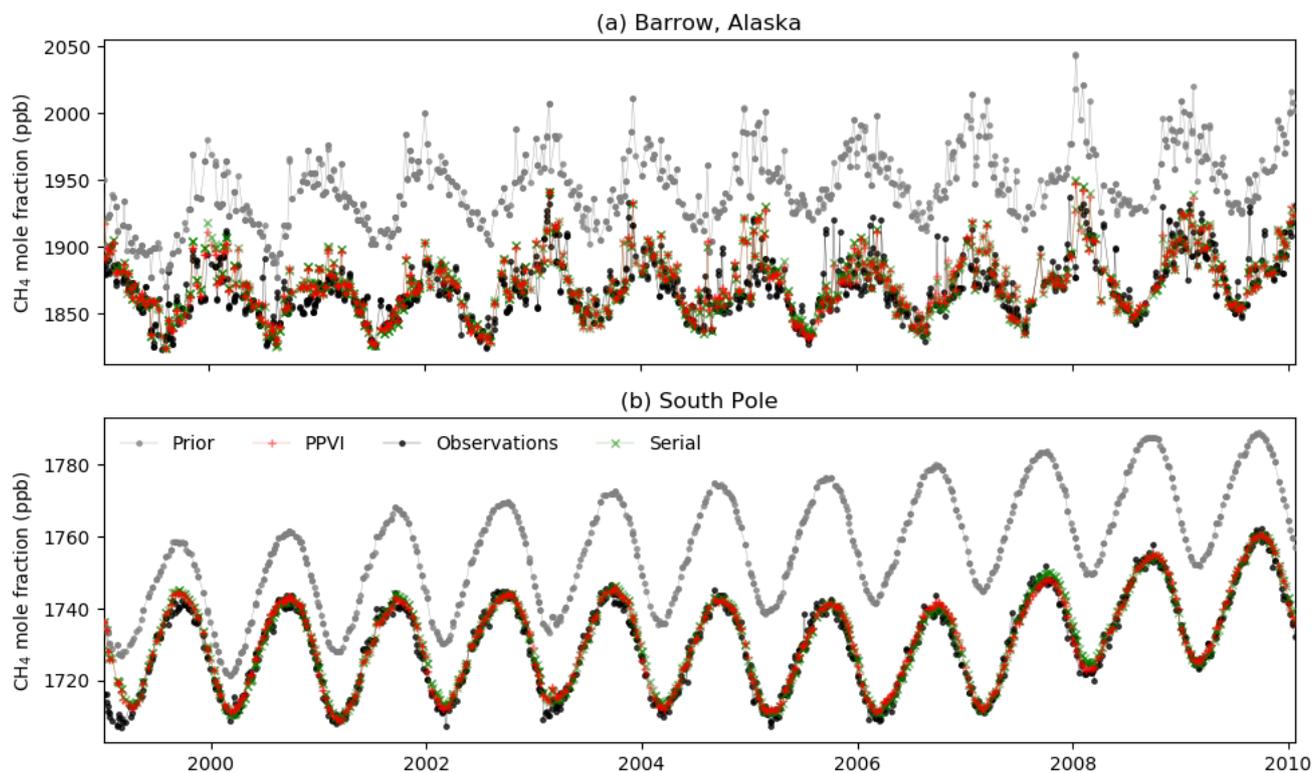



Figure 3: Modelled and observed methane mole fractions at two remote background NOAA stations, in the Northern (a) and Southern (b) hemisphere.

465

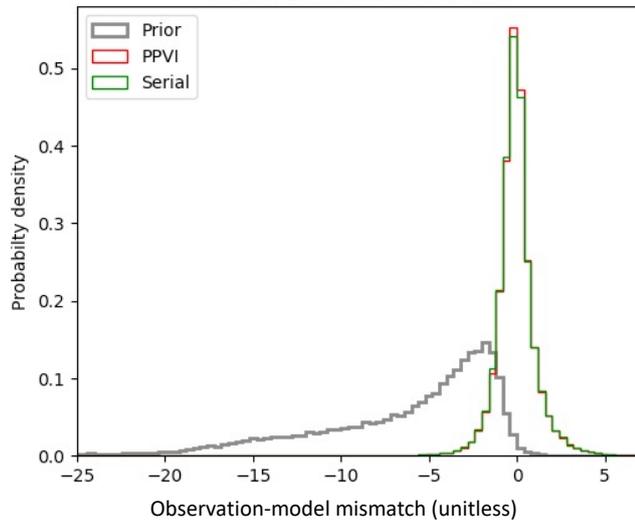



Figure 4: Probability density functions of the observation-model mismatches weighted with their uncertainties used in the inversions.

470

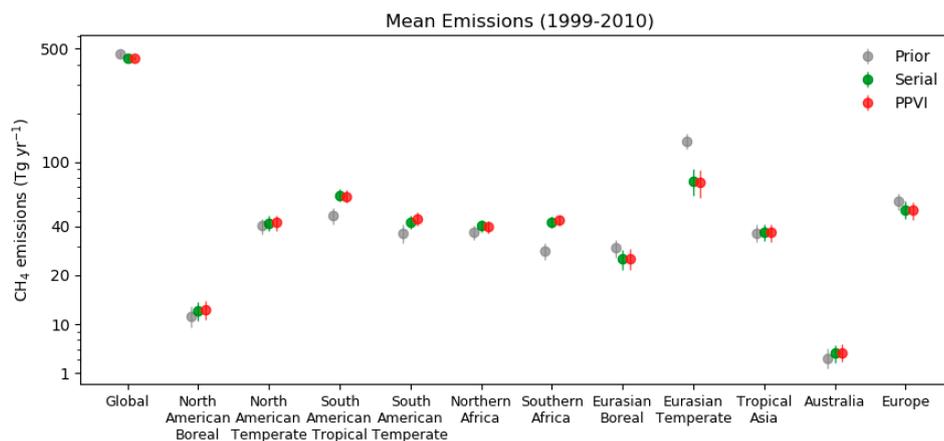



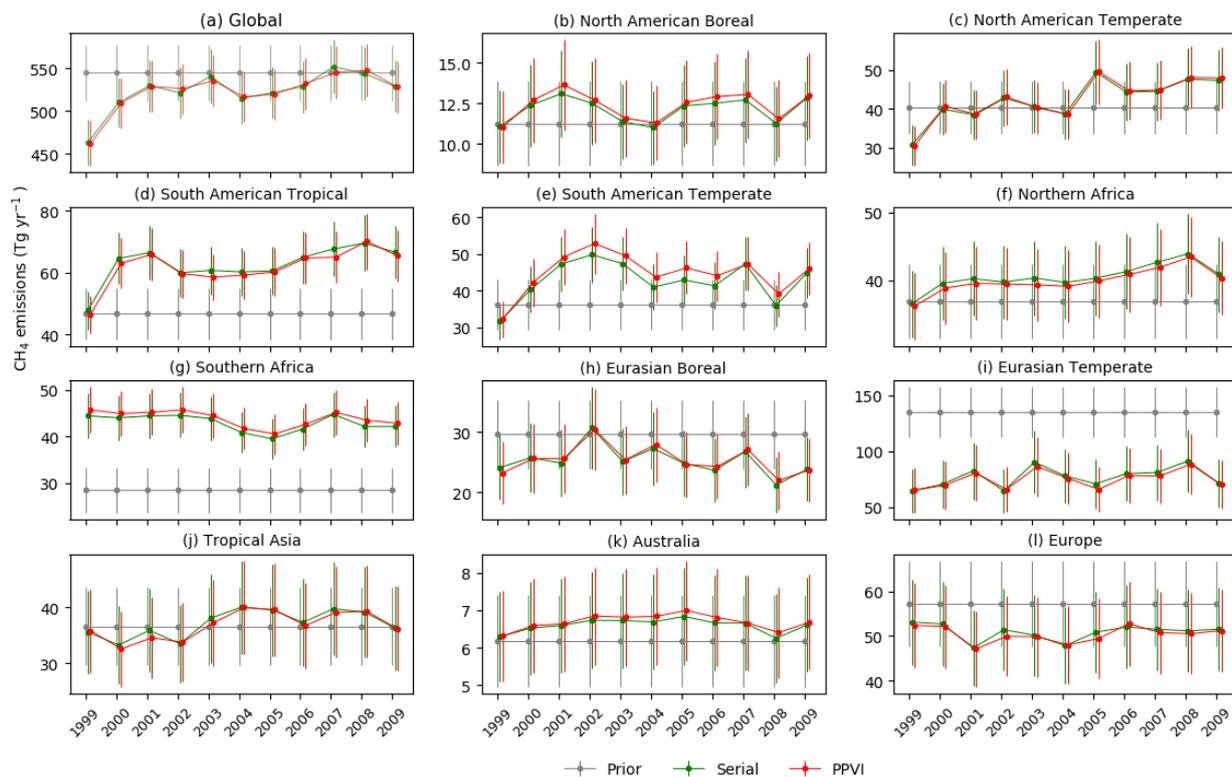
Figure 5: **Mean emission estimates of the inversions for TRANSCOM regions (see Figure 1). The vertical bars show ±2σ uncertainties.**

480

Figure 6: **Emission estimates of the PPVI and serial inversions for TRANSCOM regions. The vertical bars show ±2σ uncertainties.**