

Supplement of Biogeosciences Discuss., 12, 9341–9368, 2015
<http://www.biogeosciences-discuss.net/12/9341/2015/>
doi:10.5194/bgd-12-9341-2015-supplement
© Author(s) 2015. CC Attribution 3.0 License.



Supplement of

The ability of atmospheric data to resolve discrepancies in wetland methane estimates over North America

S. M. Miller et al.

Correspondence to: S. M. Miller (scot.m.miller@gmail.com)

The copyright of individual parts of the supplement might differ from the CC-BY 3.0 licence.

This supplement provides more detail on the atmospheric observations, the wetland methane flux estimates, and the statistical methods used throughout the paper.

S1 Atmospheric observation sites

Here we describe, in greater depth, the atmospheric methane observations collected across the US and Canada in 2007–2008. The observations used here are identical to those in Miller et al. (2013) and Miller et al. (2014b), and the discussion below summarizes the data descriptions in those papers.

The methane analysis in the main article uses either real or synthetic data at US and Canadian observation sites – a total of 14,703 observations. Of those measurements, 2,009 are from observation towers in Canada. These towers (from east to west) include Chibougamau, Quebec (CHM, 50°N, 74°W, 30m above ground level); Fraserdale, Ontario (FSD, 50°N, 83°W, 40m agl); East Trout Lake, Saskatchewan (ETL, 54°N, 104°W, 105m agl); and Candle Lake, Saskatchewan (CDL, 54°N, 105°W, 30m agl, 2007 only). These sites, operated by Environment Canada, measure methane continuously. In this study, as in Miller et al. (2014b), we use only afternoon averages of the methane data and WRF-STILT model output (1pm - 7pm local time); small scale heterogeneities in the continuous data caused by turbulent eddies and incomplete mixing make it difficult to model finer scale temporal patterns in the data. The 2,009 observations at these Canadian sites represent the total after averaging.

An additional 4,984 methane observations were collected from US towers operated by the National Oceanic and Atmospheric Administration (NOAA) and its partners. These observations include daily flask samples from a number of tower sites (weekly at Argyle and Ponca City): Argyle, Maine (AMT, 45 °N, 69 °W, 107m above ground level (agl)); Erie, Colorado (BAO, 40 °N, 105°W, 300m agl); Park Falls, Wisconsin (LEF, 46°N, 90°W, 244m agl), Martha’s Vineyard, Massachusetts (MVY, 41°N, 71°W, 12m agl); Niwot Ridge and Niwot Forest, Colorado (NWF, NWR, 40°N, 105°W, 2,3,23m agl); Ponca City, Oklahoma (SGP, 37°N, 97°W, 60m agl); West Branch, Iowa (WBI, 42°N, 93°W, 379m agl); Walnut Grove, California (WGC, 38°N, 121°W, 91m agl), and Moody, Texas (WKT, 31°N, 97°W, 122, 457m agl).

A further 7710 methane measurements were obtained from flask samples on regular NOAA aircraft flights and from the START08 (Stratosphere-Troposphere Analyses of Regional Transport 2008) measurement campaign (Pan et al., 2010). As in Miller et al. (2013), we only use aircraft observations up to 2500m above ground level. Observations at higher altitudes are less sensitive to surface emissions and were instead used by Miller et al. (2013) to optimize the empirical methane boundary condition. In this study, we only use aircraft and tower-based observations collected during daytime hours.

We further screen the data for biomass burning influence at the Canadian sites and at Park Falls, Wisconsin. At the these sites, we remove all days with CO that peaks above 200 ppb, as was done in Miller et al. (2014b). When these sites see influence from distant anthropogenic emissions, CO is often elevated, but it rarely exceeds 200 ppb except during time periods with known fires (Miller et al., 2008).

S2 WETCHIMP methane flux models

This section of the supplement details the WETCHIMP methane estimates from Melton et al. (2013) and Wania et al. (2013). The seven methane estimates used in this study are shown in Fig. S1. The wetland methane fluxes estimated by these models varies widely – both in magnitude and in spatial distribution. For example, the SDGVM model places large fluxes over the US Corn Belt relative to other regions while other models like Orchidee place large

fluxes in Northern Canada that extent far into the Northwest Territories. For a more in-depth inter-comparison of these flux estimates, refer to Melton et al. (2013) and Wania et al. (2013).

S3 The synthetic data

In the main article, we use synthetic methane data to explore the sensitivity of atmospheric observations to wetland fluxes (sections 2.3 and 3). This section describes in greater detail how we construct this synthetic data. The methods described here are adapted from Fang et al. (2014) and Shiga et al. (2014), and the discussion below parallels the descriptions in those studies.

The synthetic observations include contributions from anthropogenic sources, from wetlands, and from simulated model and measurement errors:

$$\mathbf{z}_{\text{synthetic}} = \mathbf{H}(\mathbf{s}_{\text{anthro}} + \mathbf{s}_{\text{wetland}}) + \boldsymbol{\epsilon} \quad (\text{S1})$$

In this equation, $\mathbf{z}_{\text{synthetic}}$ ($n \times 1$) represents the synthetic observations generated for an observation site. The vector $\mathbf{s}_{\text{anthro}}$ ($m \times 1$) represents emissions from anthropogenic sources and $\mathbf{s}_{\text{wetland}}$ ($m \times 1$) represents wetland fluxes. The footprint or sensitivity matrix \mathbf{H} ($n \times m$), generated from WRF-STILT, models the impact of these emissions at the observation sites.

In this study, we use the a priori anthropogenic emissions estimates from Miller et al. (2013) and Miller et al. (2014b) for $\mathbf{s}_{\text{anthro}}$. Those studies used activity data from the EDGAR inventory and a model selection framework to construct a prior anthropogenic emissions estimate. These EDGAR activity datasets include economic or demographic data that may predict the spatial distribution of methane emissions (e.g., human or ruminant population maps).

The wetland fluxes ($\mathbf{s}_{\text{wetland}}$) in Eq. S1 are taken from the WETCHIMP methane flux models (experiment two in Melton et al. (2013)). We use only four of the seven WETCHIMP models to generate synthetic data: CLM4Me, DLEM, LPJ-WSL, and SDGVM. These models have an overall magnitude that most closely matches the methane budgets estimated by three recent top-down studies over Canada’s Hudson Bay Lowlands (HBL) (Pickett-Heaps et al., 2011; Miller et al., 2014b; Wecht et al., 2014). The magnitude of these four models is likely the most realistic among the WETCHIMP flux estimates. The other WETCHIMP models, in contrast, predict much higher fluxes (Fig. 4).

As in Miller et al. (2013) and Miller et al. (2014b), the emissions ($\mathbf{s}_{\text{anthro}}$ and $\mathbf{s}_{\text{wetland}}$) are regridded to a spatial resolution of 1° latitude by 1° longitude. The EDGAR activity data do not have any seasonality, so the anthropogenic emissions ($\mathbf{s}_{\text{anthro}}$) are seasonally invariant. The WETCHIMP models have a monthly temporal resolution, as in Melton et al. (2013). That study provides flux estimates for the years 1993-2004; we use the mean of these ten years for all analysis in this study.

The final term in equation S1, $\boldsymbol{\epsilon}$ ($n \times 1$), represents simulated errors in the measurements, in WRF-STILT, and in the fluxes ($\mathbf{s}_{\text{anthro}}$ and $\mathbf{s}_{\text{wetland}}$). The magnitude and spatial/temporal structure of these errors were estimated in Miller et al. (2013) for the US and Miller et al. (2014b) for Canada. The remainder of this section details the specific calculations for simulating $\boldsymbol{\epsilon}$.

The errors in $\boldsymbol{\epsilon}$ are distributed according to the covariance matrix $\boldsymbol{\Psi}$ ($n \times n$) (Eq. 1):

$$\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}) \quad (\text{S2})$$

$$\boldsymbol{\Psi} = \mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R} \quad (\text{S3})$$

The variances and covariances within $\boldsymbol{\Psi}$ fall into two different categories. The first category are errors due to imperfect emissions, described by covariance matrix \mathbf{Q} ($m \times m$). In atmospheric inversion studies, this matrix is typically termed the a priori covariance matrix. The diagonal elements of \mathbf{Q} describe a set of variances – differences between the prior and the unknown true

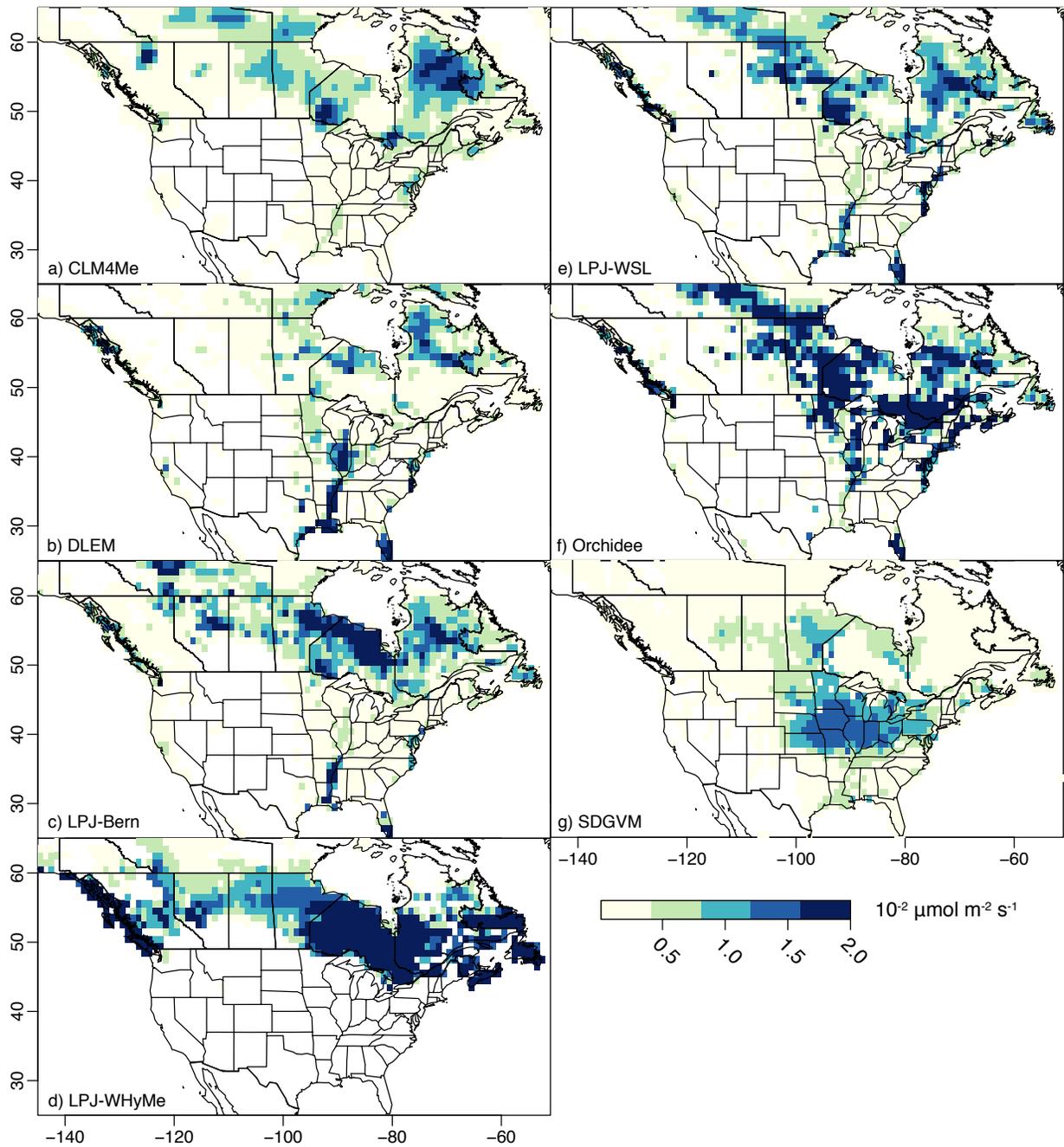


Figure S1: Annual mean wetland methane fluxes from seven different WETCHIMP estimates (Melton et al., 2013; Wania et al., 2013). The fluxes shown here are averaged over the 1993-2004 study period. Note that the fluxes shown above are averaged over the entire grid cell, not per m^2 of wetlands.

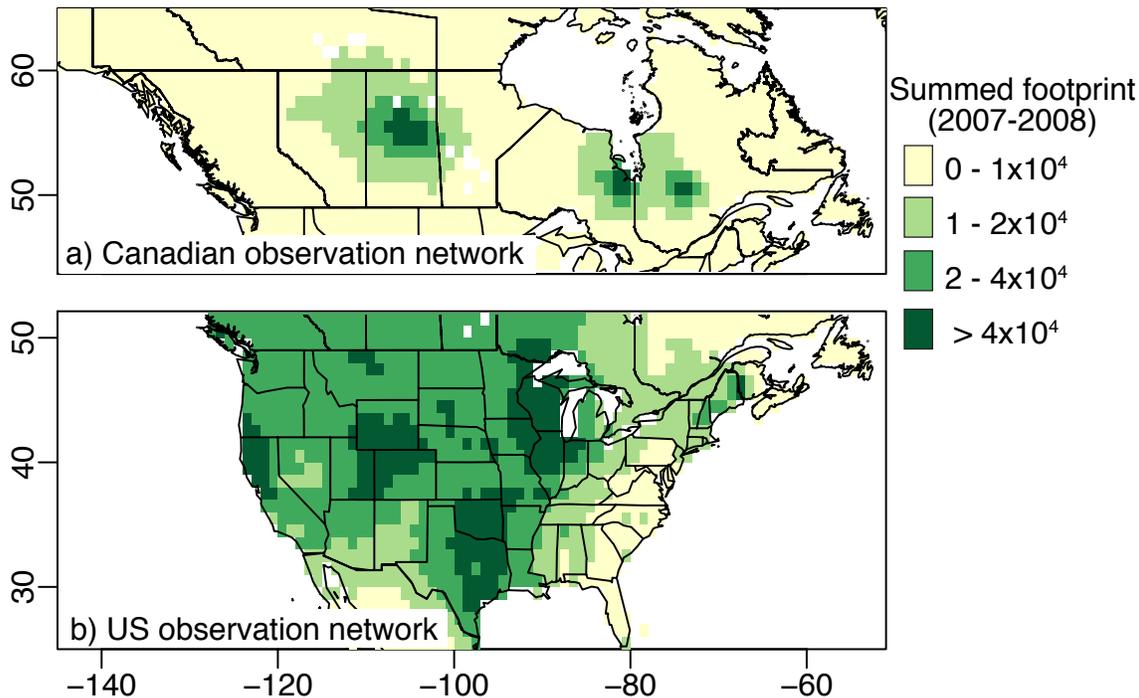


Figure S2: Total, summed footprint from the (a) Canadian and (b) US observation networks. The observation sites incorporated into this figure are shown in Fig. 2. Each individual footprint (associated with an individual atmospheric observation) has units of concentration per unit flux (ppb per $\mu\text{mol m}^{-2} \text{s}^{-1}$). In this figure, we sum all footprints for 2007–2008.

emissions over long spatial or temporal scales. The off-diagonal elements of \mathbf{Q} describe any spatial and/or temporal covariances in these differences. In Eq. S3, the footprint or sensitivity matrix (\mathbf{H}) projects \mathbf{Q} from units of $(\text{flux})^2$ into units of parts per billion squared, $(\text{ppb})^2$.

We refer to the second type of errors as model-data mismatch errors, denoted by covariance matrix \mathbf{R} ($n \times n$). These include all errors in the WRF-STILT model or the measurements that are unrelated to an imperfect flux estimate. Examples of model-data mismatch errors include measurement error, atmospheric transport error, or errors due to the spatial or temporal resolution of WRF-STILT. Over the United States, we use values for \mathbf{R} and \mathbf{Q} that were estimated by Miller et al. (2013) using WRF-STILT and the same atmospheric methane observations used in this study. Similarly, we use values for \mathbf{R} and \mathbf{Q} over Canada that were estimated in Miller et al. (2014b), a parallel inverse modeling study over that country.

In order to simulate ϵ , we next compute the Cholesky decomposition of Ψ :

$$\Psi = \mathbf{C}\mathbf{C}^T \quad (\text{S4})$$

The covariance matrix Ψ has units of $(\text{ppb})^2$, but its Cholesky decomposition (\mathbf{C}) has units of ppb, a fact that will become useful in the next step. With this decomposition in hand, we simulate a set of errors, ϵ (e.g., Fang et al., 2014; Shiga et al., 2014):

$$\epsilon = \mathbf{C}\mathbf{u} \quad (\text{S5})$$

$$\mathbf{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{1}) \quad (\text{S6})$$

where \mathbf{u} represents a set of randomly-generated numbers with a mean of zero and variance of one.

We simulate 1000 synthetic datasets for each experiment to adequately sample the random errors in ϵ . We then use the model selection framework to find the optimal candidate model for

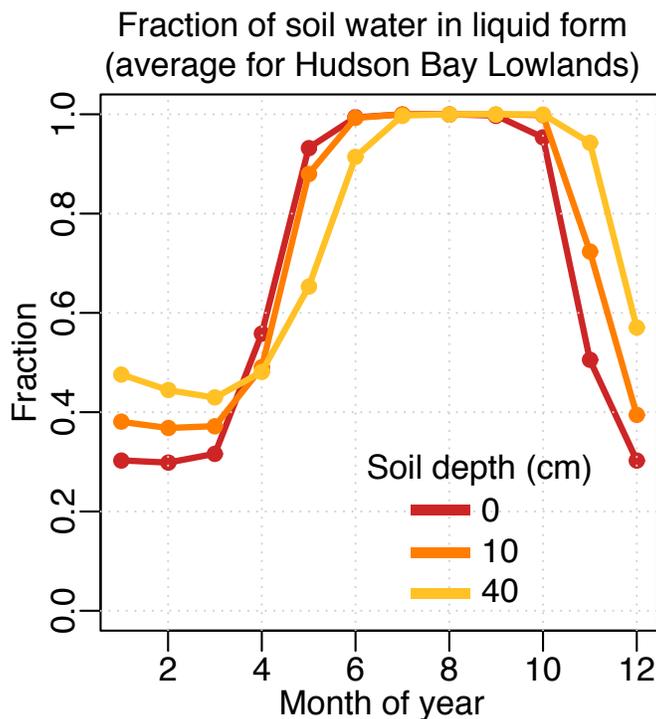


Figure S3: This figure displays the fraction of soil water that is unfrozen for the HBL in different seasons and at different soil depths. Estimates are taken from NARR (Mesinger et al., 2006).

each of these datasets. The results presented in Fig. 3 are therefore the composite of thousands of model selection runs: one model selection run for each synthetic dataset. We use a branch and bound algorithm from Yadav et al. (2013) to improve the computational efficiency of these model selection runs. Furthermore, we estimate the coefficients (β) in Eq. 1 using Lagrange multipliers to ensure that none of the estimated coefficients have unrealistic negative values (e.g., Miller et al., 2014a).

S4 Sensitivity of the observation network to surface fluxes

In this section, we describe the overall sensitivity of the observation network to methane fluxes. This sensitivity will play at least some role in network’s ability to identify a signal from wetlands. The WRF-STILT model quantifies this sensitivity in terms of a footprint. Each row the matrix \mathbf{H} is the footprint associated with a different atmospheric methane observation. In Fig. S2, we plot these footprints, summed over all of 2007–2008.

This figure show several distinctive patterns. First, the US network has a higher sensitivity than the Canadian network. This pattern is due to the larger number of observation sites over the US. Second, the highest sensitivities are clustered in distinctive regions with multiple observation sites – Wisconsin, Texas/Oklahoma, and California, among other regions.

S5 Soil freeze/thaw estimates from NARR

Figure S3 shows the soil freeze/thaw cycle at different depths averaged across the HBL. These estimates are taken from North American Regional Reanalysis (NARR) (Mesinger et al., 2006), and the values shown in Fig. S3 are average values for each month. The main article references this figure in a discussion of the methane flux seasonal cycle (Sect. 4.3).

References

- 125 Fang, Y., Michalak, A. M., Shiga, Y. P., and Yadav, V.: Using atmospheric observations to evaluate the spatiotemporal variability of CO₂ fluxes simulated by terrestrial biospheric models, *Biogeosciences*, 11, 6985–6997, doi:10.5194/bg-11-6985-2014, 2014.
- Melton, J. R., Wania, R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C. A., Beerling, D. J., Chen, G., Eliseev, A. V., Denisov, S. N., Hopcroft, P. O., Lettenmaier, 130 D. P., Riley, W. J., Singarayer, J. S., Subin, Z. M., Tian, H., Zürcher, S., Brovkin, V., van Bodegom, P. M., Kleinen, T., Yu, Z. C., and Kaplan, J. O.: Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP), *Biogeosciences*, 10, 753–788, doi:10.5194/bg-10-753-2013, 2013.
- Mesinger, F., Dimego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jovi, D., 135 Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., and Shi, W.: North American Regional Reanalysis., *B. Am. Meteorol. Soc.*, 87, 343–360, doi:10.1175/BAMS-87-3-343, 2006.
- Miller, S. M., Matross, D. M., Andrews, A. E., Millet, D. B., Longo, M., Gottlieb, E. W., Hirsch, A. I., Gerbig, C., Lin, J. C., Daube, B. C., Hudman, R. C., Dias, P. L. S., Chow, 140 V. Y., and Wofsy, S. C.: Sources of carbon monoxide and formaldehyde in North America determined from high-resolution atmospheric data, *Atmos. Chem. Phys.*, 8, 7673–7696, doi:10.5194/acp-8-7673-2008, 2008.
- Miller, S. M., Wofsy, S. C., Michalak, A. M., Kort, E. A., Andrews, A. E., Biraud, S. C., Dlugokencky, E. J., Eluszkiewicz, J., Fischer, M. L., Janssens-Maenhout, G., Miller, B. R., 145 Miller, J. B., Montzka, S. A., Nehrkorn, T., and Sweeney, C.: Anthropogenic emissions of methane in the United States, *P. Natl. Acad. Sci. USA*, 110, 20 018–20 022, doi:10.1073/pnas.1314392110, 2013.
- Miller, S. M., Michalak, A. M., and Levi, P. J.: Atmospheric inverse modeling with known physical bounds: an example from trace gas emissions, *Geoscientific Model Development*, 7, 150 303–315, doi:10.5194/gmd-7-303-2014, 2014a.
- Miller, S. M., Worthy, D. E. J., Michalak, A. M., Wofsy, S. C., Kort, E. A., Havice, T. C., Andrews, A. E., Dlugokencky, E. J., Kaplan, J. O., Levi, P. J., Tian, H., and Zhang, B.: Observational constraints on the distribution, seasonality, and environmental predictors of North American boreal methane emissions, *Global Biogeochem. Cy.*, 28, 146–160, doi:10.1002/2013GB004580, 2014b. 155
- Pan, L. L., Bowman, K. P., Atlas, E. L., Wofsy, S. C., Zhang, F., Bresch, J. F., Ridley, B. A., Pittman, J. V., Homeyer, C. R., Romashkin, P., and Cooper, W. A.: The Stratosphere-Troposphere Analyses of Regional Transport 2008 Experiment, *B. Am. Meteorol. Soc.*, 91, 327–342, doi:10.1175/2009BAMS2865.1, 2010.
- 160 Pickett-Heaps, C. A., Jacob, D. J., Wecht, K. J., Kort, E. A., Wofsy, S. C., Diskin, G. S., Worthy, D. E. J., Kaplan, J. O., Bey, I., and Drevet, J.: Magnitude and seasonality of wetland methane emissions from the Hudson Bay Lowlands (Canada), *Atmos. Chem. Phys.*, 11, 3773–3779, doi:10.5194/acp-11-3773-2011, 2011.
- Shiga, Y. P., Michalak, A. M., Gourdji, S. M., Mueller, K. L., and Yadav, V.: Detecting 165 fossil fuel emissions patterns from subcontinental regions using North American in situ CO₂ measurements, *Geophys. Res. Lett.*, 41, 4381–4388, doi:10.1002/2014GL059684, 2014.

- 170 Wania, R., Melton, J. R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T.,
Avis, C. A., Chen, G., Eliseev, A. V., Hopcroft, P. O., Riley, W. J., Subin, Z. M., Tian,
H., van Bodegom, P. M., Kleinen, T., Yu, Z. C., Singarayer, J. S., Zürcher, S., Lettenmaier,
D. P., Beerling, D. J., Denisov, S. N., Prigent, C., Papa, F., and Kaplan, J. O.: Present
state of global wetland extent and wetland methane modelling: methodology of a model
inter-comparison project (WETCHIMP), *Geoscientific Model Development*, 6, 617–641, doi:
10.5194/gmd-6-617-2013, 2013.
- 175 Wecht, K. J., Jacob, D. J., Frankenberg, C., Jiang, Z., and Blake, D. R.: Mapping of North
American methane emissions with high spatial resolution by inversion of SCIAMACHY satel-
lite data, *J. Geophys. Res.-Atmos.*, 119, 7741–7756, doi:10.1002/2014JD021551, 2014.
- Yadav, V., Mueller, K. L., and Michalak, A. M.: A backward elimination discrete optimization
algorithm for model selection in spatio-temporal regression models, *Environ. Modell. Softw.*,
42, 88 – 98, doi:10.1016/j.envsoft.2012.12.009, 2013.