A blended TROPOMI+GOSAT satellite data product for atmospheric methane using machine learning to correct retrieval biases

Nicholas Balasus[1], Daniel J. Jacob[1,2], Alba Lorente[3], Joannes D. Maasakkers[3], Robert J. Parker[4,5],
Hartmut Boesch[4,5,a], Zichong Chen[1], Makoto M. Kelp[2], Hannah Nesser[1], Daniel J. Varon[1]

[1]School of Engineering and Applied Sciences, Harvard University, Cambridge, USA
[2]Department of Earth and Planetary Sciences, Harvard University, Cambridge, USA
[3]SRON Netherlands Institute for Space Research, Leiden, The Netherlands
[4]National Centre for Earth Observation, University of Leicester, Leicester, UK
[5]Earth Observation Science, School of Physics and Astronomy, University of Leicester, Leicester, UK
[a]now at: Institute of Environmental Physics (IUP), University of Bremen FB1, Bremen, Germany

Correspondence to: Nicholas Balasus (nicholasbalasus@g.harvard.edu)

Abstract. Satellite observations of dry column methane mixing ratios (XCH_4) from shortwave infrared (SWIR) solar backscatter radiation provide a powerful resource to quantify methane emissions in service of climate action. The TROPOMI instrument launched in October 2017 provides global daily coverage at 5.5×7 km^2 nadir pixel resolution but its methane retrievals can suffer from biases associated with SWIR surface albedo, scattering from aerosols and cirrus clouds, and across-track variability (striping). The GOSAT instrument launched in 2009 has better spectral characteristics and its methane retrieval is much less subject to biases, but its data density is 250 times sparser than TROPOMI. Here we present a blended TROPOMI+GOSAT methane product obtained by training a machine learning (ML) model to predict the difference between TROPOMI and GOSAT co-located measurements, using only predictor variables included in the TROPOMI retrieval, and then applying the correction to the complete TROPOMI record from April 2018 to present. We find that the largest corrections are associated with coarse aerosol particles, high SWIR surface albedo, and across-track pixel index. Our blended product corrects a systematic difference between TROPOMI and GOSAT over water, and it features corrections exceeding 10 ppb over arid land, persistently cloudy regions, and high northern latitudes. It reduces the TROPOMI spatially variable bias over land (referenced to GOSAT data) from 14.3 to 10.4 ppb at $0.25° \times 0.3125°$ resolution. Validation with TCCON ground-based column measurements shows reductions in variable bias compared to the original TROPOMI data from 4.7 to 4.4 ppb and in single-retrieval precision from 14.5 to 11.9 ppb. TCCON data are all in locations of SWIR surface albedo below 0.4 (where TROPOMI biases tend to be relatively low), but they confirm the dependence of TROPOMI biases on SWIR surface albedo and coarse aerosol particles, as well as the reduction of these biases in the blended product. Fine-scale inspection of the Arabian Peninsula shows that a number of hotspots in the original TROPOMI data are removed as artifacts in the blended product. The blended product also corrects striping and aerosol/cloud biases in single-orbit TROPOMI data, enabling better detection and quantification of ultra-emitters. Residual coastal biases can be removed by applying additional filters. The ML method

presented here can be applied more generally to validate and correct data from any new satellite instrument by reference to a more established instrument.

1 Introduction

Methane is a strong greenhouse gas, responsible for a third of the increase in global mean surface air temperature from 1750 to 2019 (Szopa et al., 2021). Its high global warming potential and short atmospheric lifetime of only 9 years (Prather et al., 2012) make it an attractive mitigation target to address near-term climate change (Nisbet et al., 2020). Monitoring progress in methane mitigation requires knowledge of worldwide emissions, but these are still highly uncertain (Saunois et al., 2020). Global satellite observations of atmospheric methane provide important top-down information to improve emission inventories by inversion of chemical transport models (CTMs) to relate concentrations to emissions (Palmer et al., 2021). GOSAT has been in space since 2009 and provides mature and accurate retrievals, but they are relatively sparse (Parker et al., 2020). The TROPOMI instrument was launched in 2017 and provides global daily coverage, but it is more subject to biases than GOSAT because it uses a different spectral viewing window, has coarser spectral resolution, and relies on an array of detectors (Jacob et al., 2022). Here we apply machine learning (ML) to produce a blended TROPOMI+GOSAT product that uses GOSAT to correct biases in the TROPOMI data and enables more reliable application of these TROPOMI data for global inference of methane emissions.

Methane can be observed from space by nadir measurement of the spectrum of backscattered sunlight in the shortwave infrared (SWIR) spectral range. There are strong methane absorption features at 1.65 μm and 2.3 μm, enabling retrieval of the atmospheric methane column with near vertically-uniform sensitivity from the top of the atmosphere down to the surface under clear-sky conditions (Frankenberg et al., 2005). Normalization of this methane column to the dry air mass yields a dry total column averaged mixing ratio of methane (XCH_4) as the standard retrieved quantity (Jacob et al., 2016). Retrievals can be biased when spectral structure in the surface albedo is misinterpreted as methane absorption (Jongaramrungruang et al., 2021). Poorly resolved optically thin scatterers, including aerosols and cirrus clouds, as well as stray light from adjacent reflective surfaces can also bias methane retrievals (Aben et al., 2007; Butz et al., 2010; Schepers et al., 2012).

The susceptibility of methane retrievals to surface and atmospheric scattering effects depend on several factors including the spectral resolution of the instrument and the choice of SWIR band. GOSAT measures in the 1.65 μm band with 0.06 nm spectral resolution, enabling accurate retrieval of methane using the proxy approach that takes advantage of CO_2 absorption in that same band (Parker et al., 2011). The CO_2 proxy approach multiplies the XCH_4/XCO_2 ratio retrieved without consideration of atmospheric scattering by a local XCO_2 value from a CTM calibrated with observations. This takes advantage of the much smaller variability of XCO_2 than XCH_4 and largely cancels surface and atmospheric artifacts. A limitation of the proxy approach is the assumption of accurate prior XCO_2, which can introduce biases when CO_2 and methane are co-emitted from a

65   flare, for example. The proxy approach has demonstrated accuracy (Buchwitz et al., 2015) and the GOSAT retrievals are mature. GOSAT utilizes a Fourier transform spectrometer with mechanical cross-track pointing, providing a uniform spectral response for its observations and consistent high-quality data from 2009 to present (Kuze et al., 2016). The main limitation of GOSAT is that its observations are sparse, taken in 10.5 km diameter pixels spaced about 270 km apart with a return time of 3 days. The GOSAT data have been used extensively for inversions of methane emissions on global and continental scales

70   with 100-500 km resolution (Turner et al., 2015; Maasakkers et al., 2019; Janardanan et al., 2020; Western et al., 2021; Maasakkers et al., 2021; Qu et al., 2021; Worden et al., 2022; Feng et al., 2022), but they cannot effectively access finer scales.

    TROPOMI provides global daily coverage in continuous 5.5×7 km^2 (nadir) pixels, increasing the data density relative to GOSAT by more than two orders of magnitude through the use of an imaging grating spectrometer. It measures in the 2.3 μm

75   band, where the CO_2 proxy approach is not possible, with a spectral resolution of 0.25 nm. Retrieval of XCH_4 by TROPOMI employs a full-physics approach in which surface albedo and atmospheric scattering properties are retrieved together with XCH_4, utilizing additional information from the near-infrared (NIR) band of TROPOMI (Butz et al., 2012). Aliasing between these parameters in the retrieval can produce artifacts that bias the inference of methane emissions (Barré et al., 2021; Qu et al., 2021; Jacob et al., 2022). Recent improvements to the operational retrieval produced by the Netherlands Institute for Space

80   Research (SRON) have reduced some of these biases (Lorente et al., 2021; 2023). An independent TROPOMI retrieval by the University of Bremen (Schneising et al., 2019; 2023) applied a ML correction to a methane climatology to remove retrieval biases, but this may bias the product if the correction to climatology is not appropriate.

    Our blended TROPOMI+GOSAT methane product aims to eliminate biases from the TROPOMI data by using co-located

85   GOSAT methane retrievals from four years of observations (2018-2021) to train a ML model for predicting the TROPOMI-GOSAT XCH_4 difference, relying only on predictor variables included in the TROPOMI methane product. This allows us to apply the TROPOMI-GOSAT correction to the complete TROPOMI dataset (2018-present) to form the blended product. The ML model also identifies the main sources of bias in the TROPOMI data to guide further improvements in the retrieval. The methods presented here are not specific to TROPOMI and GOSAT and could be applied to any other satellite instrument pairs.

90 **2 Construction of the blended TROPOMI+GOSAT product**

    Table 1 summarizes the GOSAT and TROPOMI data used in the construction of our blended TROPOMI+GOSAT product including the GOSAT v9.0 proxy retrieval from Parker et al. (2020) (quality flag = 0) and the TROPOMI v02.04.00 operational product based on Lorente et al. (2023) (quality assurance value = 1, albedo bias-correction applied). It is standard practice to evaluate satellite methane products with ground-based XCH_4 observations from the Total Carbon Column Observing Network

95   (TCCON) (Wunch et al. 2011). We do so here for the GOSAT and TROPOMI retrievals using the 24 TCCON sites available in the GGG2020 version of the data during 2018-2021, adjusting all retrievals to common vertical profiles and averaging kernel

sensitivities as described in Appendix A. Details of the evaluation with TCCON data are given in Appendix B and results are given in Table 1. All of the TCCON sites are over land and most are at northern mid-latitudes. We calibrate GOSAT to have a global mean bias of 0 ppb relative to GGG2020 TCCON data, subtracting 9.2 ppb from all retrievals. This follows Parker et al. (2020) but updates the TCCON data version that is calibrated against from GGG2014 to GGG2020. TROPOMI has a global mean bias of 6.0 ppb relative to TCCON. The standard deviation of the satellite-TCCON difference for individual retrievals gives a measure of retrieval precision and is 14.9 ppb for GOSAT and 14.5 ppb for TROPOMI. Most critical for inversions is the spatially variable bias, which reflects artifact data features that inversions could interpret as emissions. Variable bias is commonly diagnosed with TCCON data as the standard deviation of the temporally averaged satellite-TCCON differences for individual stations. Table 1 gives variable biases relative to TCCON of 5.2 ppb for GOSAT and 4.7 ppb for TROPOMI, which are lower than the 10 ppb threshold defined by Buchwitz et al. (2015) for successful regional inversions. However, the spatial coverage of TCCON stations for estimating this variable bias is very limited as the stations are mainly located in regions of moderate SWIR surface albedo where retrieval biases tend to be low (Lorente et al., 2021). A full global diagnostic of variable bias based on TROPOMI-GOSAT differences shows much larger values depending on region (Qu et al., 2021; Jacob et al., 2022).

We compute TROPOMI-GOSAT differences Δ(TROPOMI-GOSAT) for all co-located individual retrievals from 30 April 2018 to 31 December 2021, adjusting all retrievals to common prior vertical profiles and averaging kernel sensitivities as described in Appendix A. Co-location is defined by pixel centers ≤ 5 km apart and retrieval times ≤ 1 hour apart, resulting in 170576 pairs for the four years including 156939 pairs over land and 13637 over water. Data are much sparser over the oceans and limited to lower latitudes because of requirement for specular reflectance in the glint retrieval. The standard deviation of the difference for individual data pairs is 17.4 ppb.

Table 1. TROPOMI and GOSAT data used for the blended TROPOMI+GOSAT product.

	GOSAT	TROPOMI
Retrieval Version	UoL v9.0 [a]	Operational v02.04.00 [b]
Local Overpass Time	13:00	13:30
Pixel Size	10.5 km diameter	5.5×7 km^2 [c]
Pixel Separation	260-280 km	none
Coverage	global	global
Return Time	3 days	1 day
Retrieval Type	CO_2 proxy at 1.65 μm	full physics at 2.3 μm
Number of Retrievals per Day [d]	1443	372167
Mean Bias [e]	0.0 ppb [f]	6.0 ppb

Variable Bias [e]	5.2 ppb	4.7 ppb
Single-Retrieval Precision [e]	14.9 ppb	14.5 ppb

120

[a] Parker et al. (2020). Only observations with quality flag = 0 are used.

[b] Lorente et al. (2023). Only observations with quality assurance value = 1 are used. Albedo bias-corrected data ("methane-mixing_ratio_bias_corrected") is used.

[c] At nadir; 7×7 km^2 before 6 August 2019.

125   [d] Average for 30 April 2018-31 December 2021.

[e] Based on differences with TCCON data (version GGG2020) as derived in this work. See Appendix B for details. Variable bias is the spatial standard deviation of the temporally averaged differences for individual TCCON stations. Retrieval precision is the standard deviation of the differences for individual retrievals.

[f] In this work, we calibrate the UoL v9.0 product to zero global mean bias relative to TCCON (version GGG2020),

130   subtracting 9.2 ppb from all retrievals.

Figure 1 shows the average difference Δ(TROPOMI-GOSAT) for 2018-2021, plotted on a $2° \times 2.5°$ grid for visualization purposes. The global mean bias of TROPOMI relative to GOSAT taken as reference is 3.6 ppb over land and 12.0 ppb over water. Despite the low global mean bias over land, there are large areas with over 20 ppb differences including over bright

135   surfaces (North Africa), persistently cloudy areas (Amazon, Congo, Southeast Asia), and snow-covered surfaces (high northern latitudes).

We quantify a spatially variable bias for TROPOMI relative to GOSAT using the same definition as used for TCCON (spatial standard deviation of the temporally averaged differences). We do this for spatial resolutions of $0.25° \times 0.3125°$ and

140   $2° \times 2.5°$, typical of regional and global inversions respectively. We separate land and water because users conducting inversions may choose not to use the glint data over water. At $0.25° \times 0.3125°$ resolution we find variable biases of 13.5 ppb over land and 14.7 ppb over water, while at $2° \times 2.5°$ resolution we find variable biases of 12.1 ppb over land and 11.6 ppb over water. These variable biases imply that inversions using GOSAT or TROPOMI would produce significantly different results. Variable biases do not decrease much in going from $0.25° \times 0.3125°$ to $2° \times 2.5°$, suggesting that most of the biases

145   are large-scale regional features.

0.25° × 0.3125°	Land	Water		2° × 2.5°	Land	Water
Mean Bias	3.7 ppb	14.1 ppb		Mean Bias	3.6 ppb	12.0 ppb
Variable Bias	13.5 ppb	14.7 ppb		Variable Bias	12.1 ppb	11.6 ppb

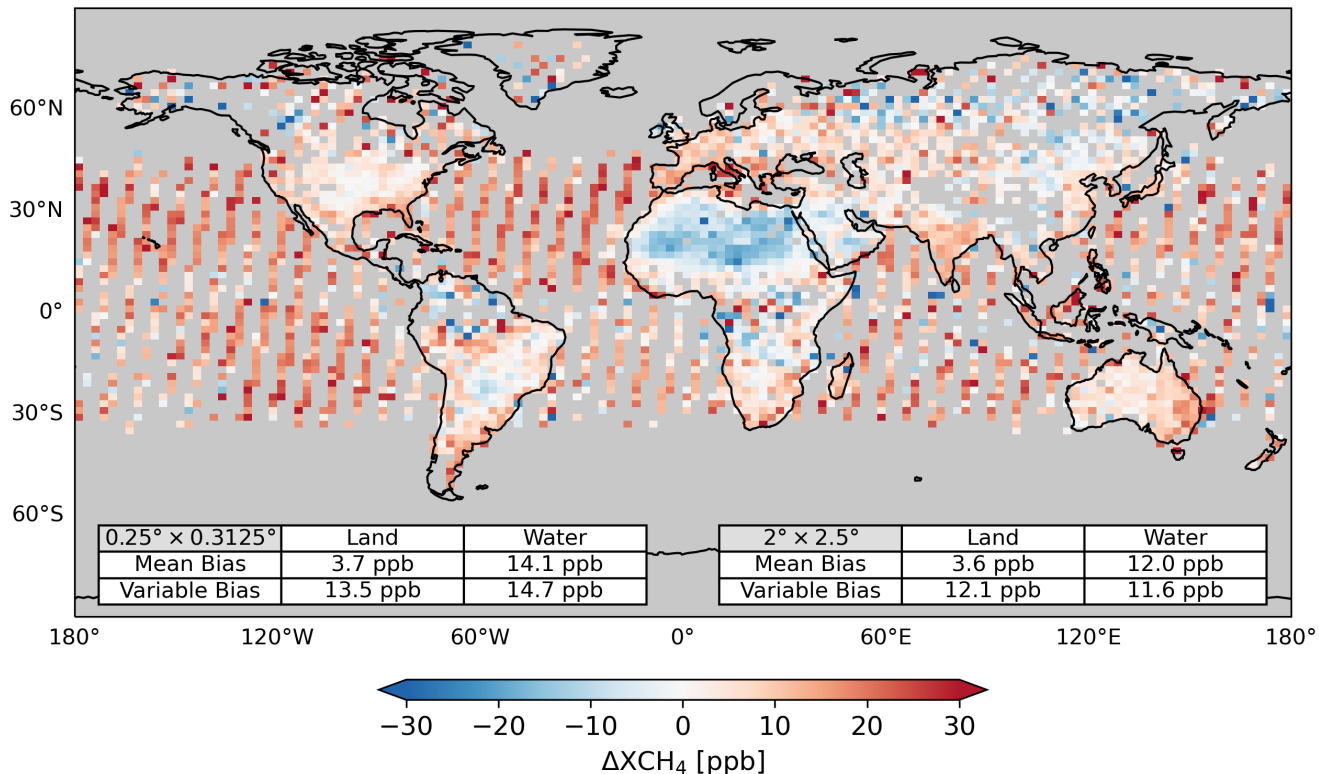



Figure 1: Average difference Δ(TROPOMI-GOSAT) between co-located GOSAT and TROPOMI observations for 2018-2021, plotted on a 2° × 2.5° grid for visibility. Co-location criteria are observation times within 1 hour and pixel centers within 5 km. The GOSAT and TROPOMI observations have been adjusted to common prior estimates and averaging kernel sensitivities to enable meaningful computation of differences (Appendix A). Mean bias and variable bias of TROPOMI relative to GOSAT are shown inset separately over land and water (data over water are from the glint product). Mean bias and variable bias are calculated respectively as the spatial average and standard deviation of the temporally averaged Δ(TROPOMI-GOSAT) data on the specified grid (0.25° × 0.3125° or 2° × 2.5°).

We use the co-located GOSAT and TROPOMI data for 2018-2021 to develop a predictive ML model for Δ(TROPOMI-GOSAT) that can be applied to correct the TROPOMI data with reference to the GOSAT data. The model uses the 30 predictor variables listed in Table 2, which are all TROPOMI retrieval parameters included with the individual XCH$_4$ observations, so that the correction can then be applied to the full TROPOMI dataset as a function of those parameters. We split the 170576 co-located data pairs into two sets. The pairs for 2018-2020 are used to train the predictive model (training dataset). The

training minimizes a loss function of mean squared error that describes the difference between predicted and true Δ(TROPOMI-GOSAT) values. The pairs for 2021 are used for independent evaluation of the predictive model (test dataset).

165   We considered three candidate ML methods (Random Forest, LightGBM, and XGBoost) that rely on ensembles of decision trees (Kingsford and Salzberg, 2008). Random Forest grows an ensemble of decision trees using a bootstrapped sample of training data and subset of features for each decision tree. The averaged predictions from the forest of trees form the model prediction (Breiman, 2001). LightGBM and XGBoost are different implementations of gradient-boosted decision tree algorithms in which decision trees are grown sequentially with each iteration predicting the residual between the observation

170 and the sum of all previous decision trees (Ke et al., 2017; Chen and Guestrin, 2016). It is not necessary to normalize the predictor variables from Table 2 for any of these methods. To choose the best predictive model, we used their implementation in Microsoft's Fast and Lightweight AutoML Library (FLAML) (Wang et al., 2021). FLAML is designed to select the ML model (both method and hyperparameters) that would perform best on the test data. To keep the test data independent, FLAML evaluates models with 10-fold cross validation on the training data (with the 10 folds determined by dividing the data

175 sequentially). LightGBM performed the best, XGBoost was a close second, and Random Forest did the worst by a significant margin. We perform a Z-score transform on model predictions to account for systematic bias in ensemble-tree based approaches at extreme values as was done by Ouyang et al. (2023) and is explained by Belitz and Stackelberg (2021).

Table 2. TROPOMI retrieval parameters used to predict Δ(TROPOMI-GOSAT). [a]

Predictor Variable	Units
1. Solar Zenith Angle	degree
2. Relative Azimuth Angle	degree
3. Across-Track Pixel Index [b]	--
4. Surface Classification [c]	--
5. Surface Altitude	m
6. Surface Roughness	m
7. U10 Wind Speed [d]	$m\ s^{-1}$
8. V10 Wind Speed [d]	$m\ s^{-1}$
9. XCH_4 *a priori*	ppb
10. Cirrus Reflectance [e]	--
11. XCH_4 Precision [f]	ppb
12. Fluorescence [g]	$photons\ s^{-1}\ cm^{-2}\ nm^{-1}\ sr^{-1}$
13/14. CO Column and Precision	$molecules\ cm^{-2}$
15/16. H_2O Column and Precision	$molecules\ cm^{-2}$

17/18. Aerosol Size Distribution Parameter and Precision [h]	--
19/20. Aerosol Height and Precision [i]	m
21/22. Aerosol Column and Precision	particles cm^{-2}
23/24. SWIR Surface Albedo and Precision [j]	--
25/26. NIR Surface Albedo and Precision [k]	--
27. SWIR Aerosol Optical Thickness	--
28. NIR Aerosol Optical Thickness	--
29. SWIR Chi-Squared [l]	--
30. NIR Chi-Squared [l]	--

180

[a] All 30 parameters in this Table are provided together with XCH$_4$ as part of the individual operational v02.04.00 TROPOMI methane retrievals. They are used in the LightGBM machine learning (ML) algorithm to predict Δ(TROPOMI-GOSAT) for individual TROPOMI retrievals.

[b] The retrieval also provides satellite viewing angle but this is redundant with the across-track pixel index.

185   [c] Surface classification for land is from the 1-km resolution Global Land Cover Characteristics Data Base Version 2.0 (USGS, 2018) and for water from the 250-m resolution data from Carroll et al. (2009) as explained by Apituley et al. (2022). We use the information from bit 0 and 1, giving this parameter four possible values: 0 = land, 1 = water, 2 = mostly land (with some water), 3 = mostly water (with some land).

[d] Zonal and meridional wind speeds at 10-meter altitude.

190   [e] From the Visible Infrared Imaging Radiometer Suite (VIIRS).

[f] Precision as given in the TROPOMI retrieval product only includes the effect of noise in the measured radiance and is much smaller than the retrieval precision given in Table 1 from validation with TCCON data (Lorente et al., 2021).

[g] Fluorescence emission at 755 nm.

[h] Negative power law exponent (α) for the aerosol size distribution represented as $n(r) \sim r^{-\alpha}$ where n is the number size

195   distribution function and r is particle radius (Hasekamp et al., 2022). Larger values of α correspond to a finer aerosol.

[i] Central height of Gaussian aerosol altitude distribution (Hasekamp et al., 2022).

[j] Shortwave infrared (SWIR) at 2305-2385 nm.

[k] Near-infrared (NIR) at 757-774 nm.

[l] Quantifies goodness of fit for the retrieval in the SWIR or NIR spectral band.

200

     We applied the SHapley Additive exPlanations (SHAP) approach to determine the contributions of the individual variables in Table 2 to the prediction of Δ(TROPOMI-GOSAT). SHAP analysis partitions individual model predictions to the different predictor variables, giving each a SHAP value (in unit of ppb) that add up to the deviation of the model prediction from the average prediction across a given dataset. We use the TreeExplainer method for our SHAP analysis (Lundberg et al., 2020).

205   The SHAP values for the predictor variables can be used to understand individual predictions, or they can be aggregated across a larger set of data. They do not fully resolve correlation across predictor variables, which can complicate interpretability (Aas et al., 2021; Silva et al., 2022).

      Figure 2 ranks the predictor variables of Table 2 by their average absolute SHAP values across the training dataset. The
210 most important predictors of Δ(TROPOMI-GOSAT) are the aerosol size distribution parameter (given by the negative power law exponent), the across-track pixel index, the CO column precision, and the SWIR surface albedo. SWIR aerosol optical thickness is strongly correlated with the aerosol size distribution parameter ($R^2 = 0.72$), but the SHAP analysis does not resolve this correlation. Similarly, CO column precision is driven by the underlying SWIR surface albedo ($R^2 = 0.68$) which we view as the actual cause of Δ(TROPOMI-GOSAT). Although one might expect arid surfaces to generate large dust particles, we
215 find that SWIR surface albedo is not correlated with the aerosol size distribution parameter ($R^2 = 0.04$). The importance of the across-track pixel index reflects the striping patterns present in TROPOMI retrievals, which change from orbit to orbit (Borsdorff et al., 2018; 2019). Additionally, the across-track pixel index provides information about the viewing geometry of TROPOMI. Cirrus reflectance does not rank among the top ten predictors of Δ(TROPOMI-GOSAT) in Figure 2 but can be important for individual scenes as will be shown later.

220

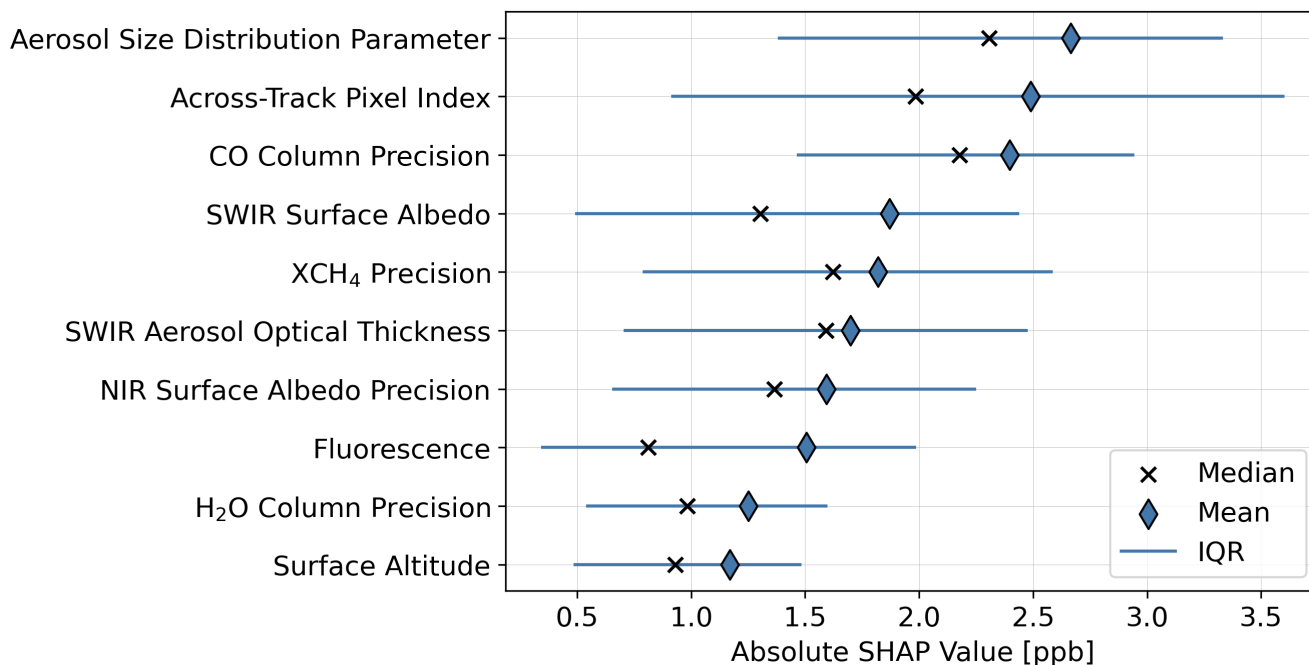



Figure 2: Predictors of Δ(TROPOMI-GOSAT) ranked in order of importance. The Figure shows the top ten predictor variables for the ML model of Δ(TROPOMI-GOSAT) among all predictor variables included in the

225   **TROPOMI retrieval dataset (Table 2). The contributions of individual variables are defined by their mean absolute**
      SHAP values in unit of ppb and are shown here as global mean absolute values along with medians and the
      interquartile range (IQR) for the training data of 2018-2020.

      Figure 3 further examines the SHAP values for the aerosol size distribution parameter (with smaller values indicating
230   larger particles) and SWIR surface albedo. TROPOMI data appear to be biased low with respect to GOSAT when particles are
      large, which is a recognized source of error for full-physics retrievals (Butz et al., 2010; Schepers et al., 2012). Despite already
      undergoing a bias correction with respect to albedo (Lorente et al., 2021), TROPOMI data are biased low relative to GOSAT
      at high SWIR surface albedo.

235

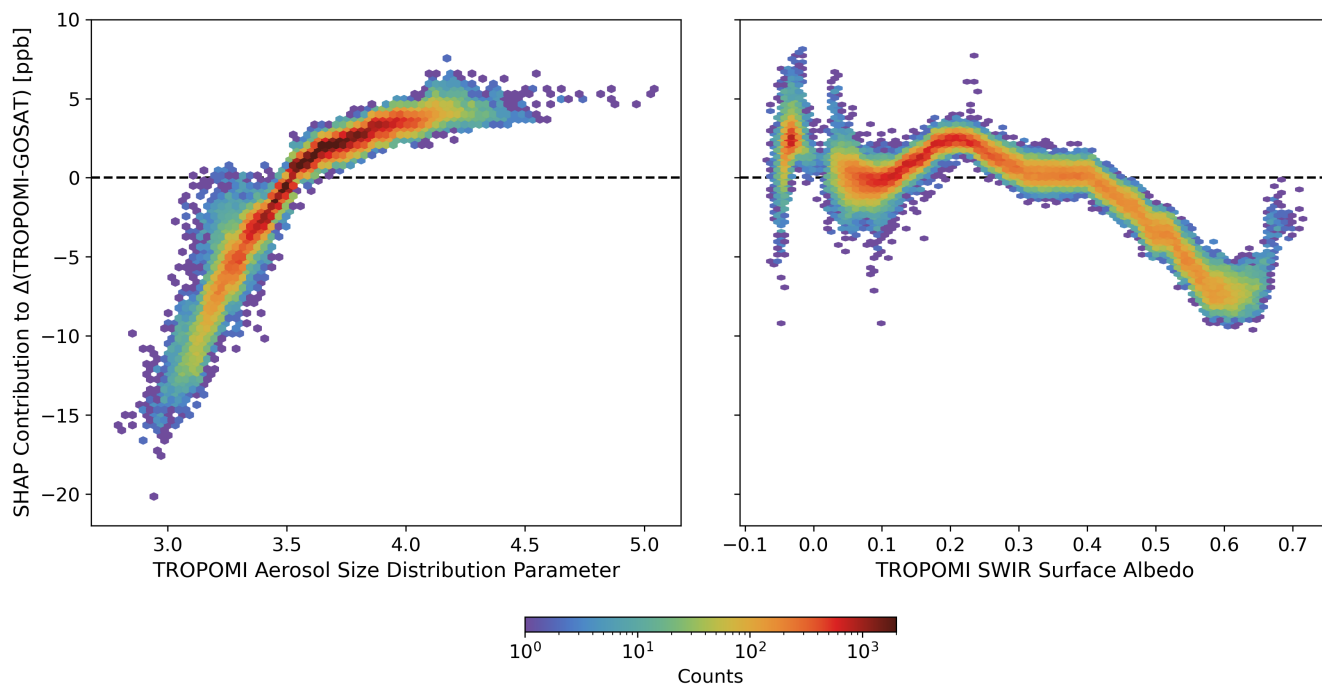



Figure 3: Contributions to Δ(TROPOMI-GOSAT) from the aerosol size distribution parameter and the SWIR surface
albedo in the TROPOMI retrievals. The aerosol size distribution parameter is the negative exponent of the assumed
power law aerosol size distribution in the TROPOMI retrieval (Table 2) and decreases as the contribution from large
240   **particles increases. The SWIR surface albedo is for the 2305-2385 nm wavelength range. Negative values are for water**
      scenes where the SWIR surface albedo is calculated differently in the retrieval (Lorente et al., 2022). The figure shows
      the SHAP-inferred contributions of the two parameters to the predicted Δ(TROPOMI-GOSAT) values for individual
      data pairs (counts) in the 2018-2020 training dataset of the ML model.

3 Evaluation of the blended TROPOMI+GOSAT product

245   Figure 4 shows the ability of the ML model to predict Δ(TROPOMI-GOSAT) for the 2021 testing data that the model was not trained on. The correction is overall successful, with a coefficient of determination (R^2) of 0.53 and a root-mean-square-error (RMSE) of 12.4 ppb. Random noise necessarily limits the quality of the fit for individual pairs. The RMSE is smaller than would be expected from the precision of Δ(TROPOMI-GOSAT) derived by adding the precisions of the individual TROPOMI (14.5 ppb) and GOSAT (14.9 ppb) retrievals relative to TCCON in quadrature (20.8 ppb; Table 1). This implies

250   that the TROPOMI and GOSAT retrieval precisions derived from TCCON are not fully random but are partly predictable on the basis of the TROPOMI retrieval parameters.

       Despite the bias correction applied by the Z-score transform from Belitz and Stackelberg (2021), we see from Figure 4 a tendency for the ML model to underestimate the high tail of the observed distribution and overestimate the low tail. This is a

255   recognized problem in ML algorithms that aim to provide a good model of the mean (Zhang and Lu, 2012).

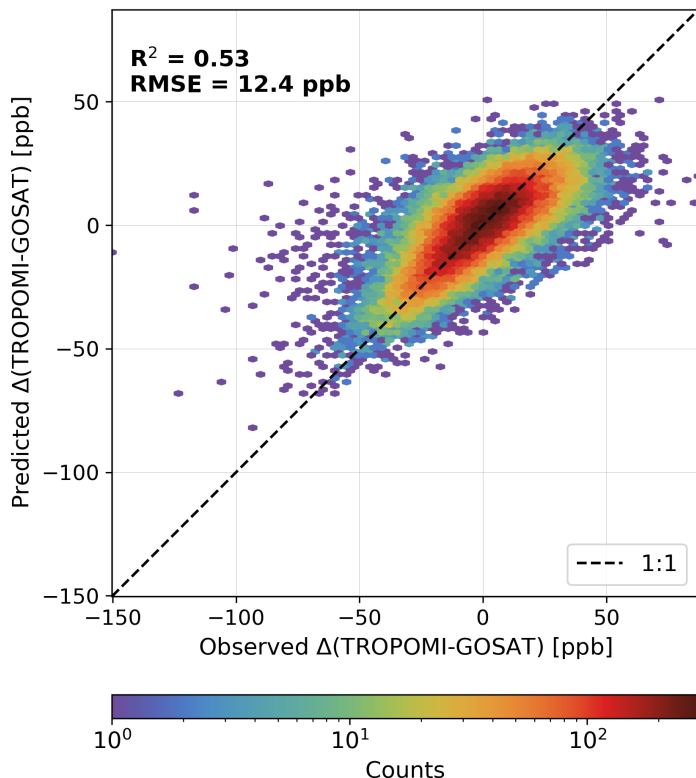
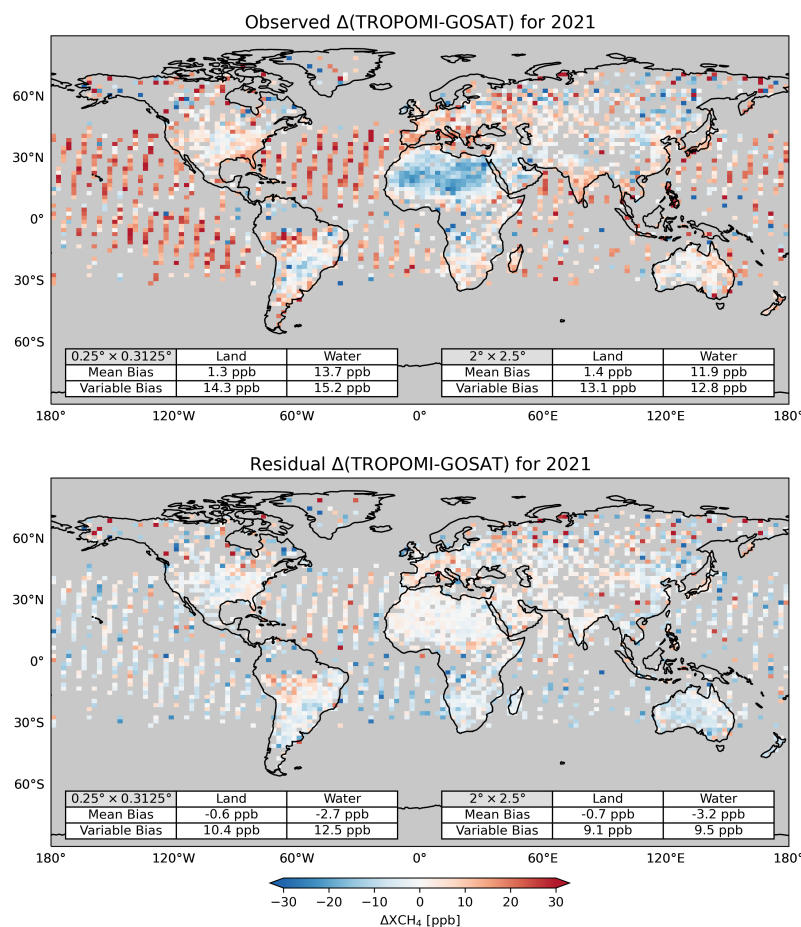



Figure 4: Ability of the ML model to predict Δ(TROPOMI-GOSAT) on the test data from 2021. Coefficient of determination (R^2) and root-mean-square error (RMSE) are shown inset.

260

Figure 5 displays the observed global distribution of Δ(TROPOMI-GOSAT) for the 2021 test data and the residual distribution (observed-predicted) after correction with the ML model. As the model was not trained on these data, the data can be used for an independent global evaluation of the reduction in the mean and variable bias of TROPOMI relative to GOSAT. Much of the original regional structure in the TROPOMI bias has disappeared or is greatly reduced. The variable bias over

265   land decreases from 14.3 to 10.4 ppb at 0.25° × 0.3125° resolution, and from 13.1 to 9.1 ppb at 2° × 2.5° resolution. The mean bias is reduced over water, going from 13.7 to -2.7 ppb at 0.25° × 0.3125° resolution and from 11.9 to -3.2 ppb at 2° × 2.5° resolution.

270   **Figure 5: TROPOMI-GOSAT XCH₄ differences (Δ(TROPOMI-GOSAT)) for co-located data in 2021, plotted on a 2° × 2.5° grid for visibility. Values are annual means. The 2021 observations (top panel) are used as test data for the ML model trained to predict Δ(TROPOMI-GOSAT) from 2018-2020 data. The bottom panel shows the residual Δ(TROPOMI-GOSAT) after removing the predicted values from the observations. The bottom panel is equivalent to Δ(Blended-GOSAT). Mean bias and variable bias are calculated as described in Figure 1.**

After forming the full blended TROPOMI+GOSAT product for the 2018-2021 period (described below), we perform an independent evaluation with the TCCON data for that period covering 24 sites (Figure 6). The evaluation procedure is described in Appendix B. This allows us to compare to the evaluations of the original TROPOMI and GOSAT retrievals with the same TCCON data. We find that the retrieval precision is improved from 14.5 ppb in the TROPOMI data to 11.9 ppb in

280   the blended TROPOMI+GOSAT product, both surpassing GOSAT's precision of 14.9 ppb. The variable bias is slightly reduced from 4.7 ppb in the TROPOMI data to 4.4 ppb for the blended TROPOMI+GOSAT product. The mean bias decreases from 6.0 ppb to -2.9 ppb. Individual station comparisons are in Table B2. All stations except for two see a reduction in the standard deviation of Δ(satellite-TCCON).

285

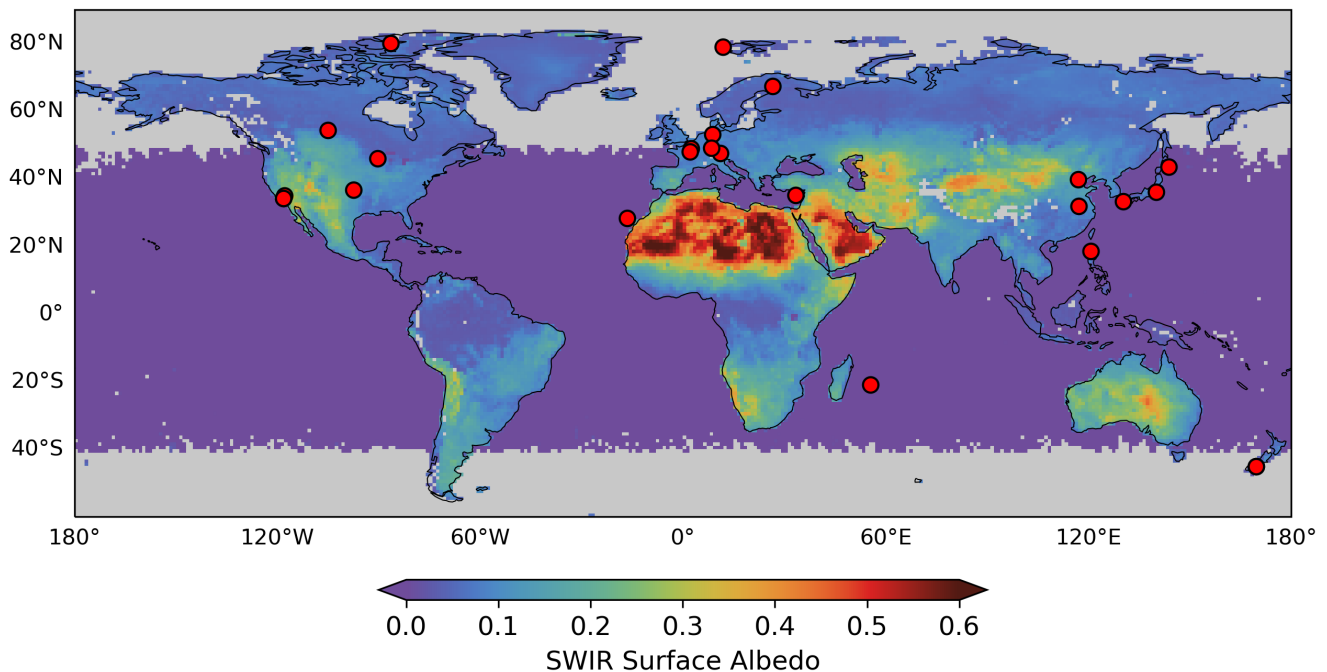



Figure 6: TCCON stations with 2018-2021 data mapped on the mean TROPOMI SWIR surface albedo for 2021. The TROPOMI data are annual means and plotted on a 1° × 1° grid with data south of 60°S cropped for visualization purposes. Grey areas have no TROPOMI data. Site locations are listed in Table B1.

290

The blended TROPOMI+GOSAT product shows only a modest improvement in error statistics at TCCON sites, but this is because these sites are all in locations of SWIR surface albedo lower than 0.4. As shown in Figure 3, the largest TROPOMI biases are for SWIR surface albedos higher than 0.4 (15% of all TROPOMI data). Beyond the simple evaluation, the TCCON

data allow us to test our previously derived relationships of TROPOMI retrieval biases to retrieval parameters, including SWIR
surface albedo and aerosol size parameter found to be important (Figures 2 and 3). We show in Figure 7 the standard deviation
of satellite-TCCON differences as a function of these two parameters. For both parameters, the spread of the satellite-TCCON
difference is lower at all parameter values. Consistent with Figure 3, the largest improvements are for large aerosols (small
values of the aerosol size parameters) and away from moderate SWIR albedo.

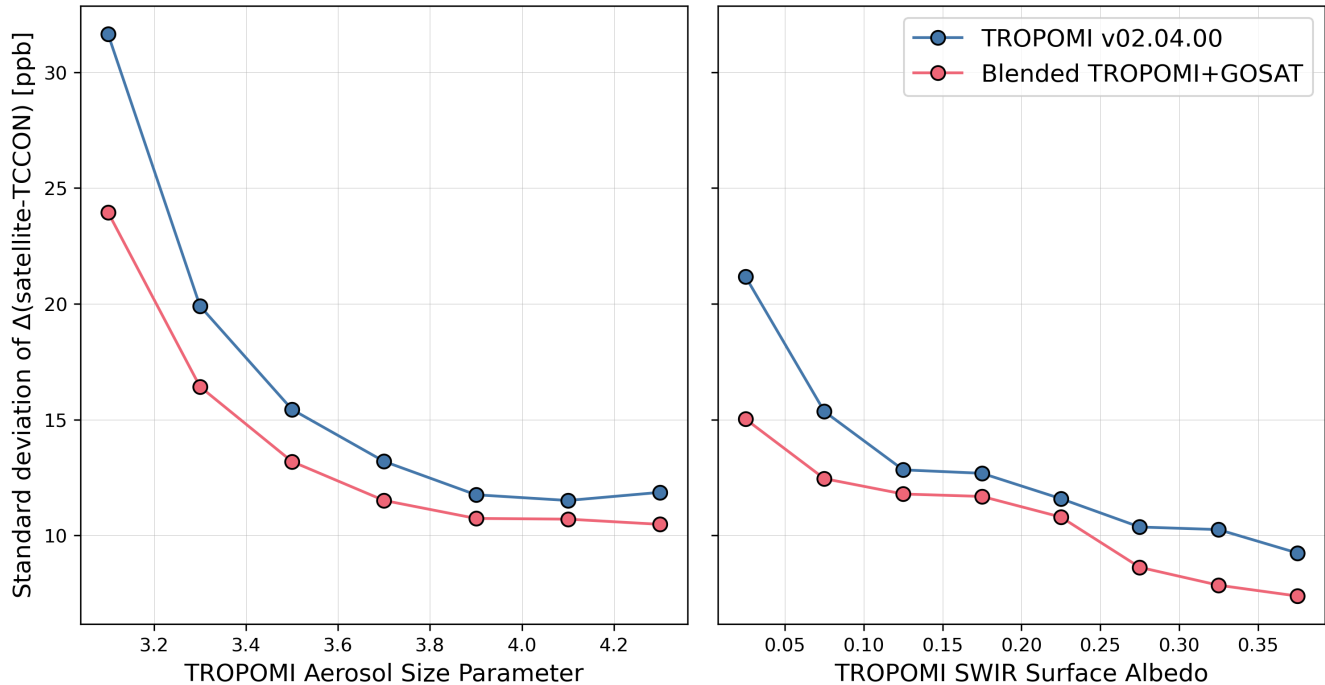



Figure 7: Application of TCCON data to evaluate satellite XCH$_4$ retrieval biases in relation to retrieval parameters. The figure shows the standard deviations of satellite-TCCON differences for different bins of the TROPOMI retrieval parameters found to be important causes of retrieval bias (Figure 2): aerosol size parameter (left, 0.2 bins) and SWIR surface albedo (right, 0.05 bins). Results are for the 2018-2021 period at the 24 TCCON sites of Figure 6 (excluded sites are explained in Appendix B) and compare the TROPOMI operational v02.04.00 retrieval and the blended TROPOMI+GOSAT product.

Table 3 summarizes the error statistics of the blended TROPOMI+GOSAT product referenced to the GOSAT dataset for
2021 (separately for land and water, 0.25° × 0.3125° and 2° × 2.5°) and the TCCON dataset for 2018-2021. Mean biases are
low. Variable bias relative to TCCON is low (4.4 ppb), but this reflects the favorable locations of the TCCON stations as
discussed above. Variable bias relative to GOSAT is about 10 ppb. This is sufficiently low that inversions of the blended

TROPOMI+GOSAT product to infer methane emissions should be consistent with inversions of GOSAT data (Buchwitz et al., 2015). The blended TROPOMI+GOSAT product benefits from the TROPOMI coverage to produce a data density 250× higher than GOSAT.

Table 3. Summary of error statistics for the blended TROPOMI+GOSAT product. [a]

Reference Dataset	Mean Bias (ppb)	Variable Bias (ppb)
GOSAT (2021, 0.25° × 0.3125°, Land)	-0.6	10.4
GOSAT (2021, 0.25° × 0.3125°, Water)	-2.7	12.5
GOSAT (2021, 2° × 2.5°, Land)	-0.7	9.1
GOSAT (2021, 2° × 2.5°, Water)	-3.2	9.5
TCCON (2018-2021, GGG2020)	-2.9	4.4

[a] Error statistics are for the differences between the blended TROPOMI+GOSAT product and the reference dataset.

4 Overview of the blended TROPOMI+GOSAT product

We produced a blended TROPOMI+GOSAT product by applying the predictive model for Δ(TROPOMI-GOSAT) to the complete operational v02.04.00 TROPOMI data product. The correction is implemented as Δ(TROPOMI-GOSAT) subtracted from the TROPOMI data. The blended product contains all successful TROPOMI retrievals from April 2018 to present. Figure 8 shows the global distribution of the blended product for 2021 and the corrections to the TROPOMI retrieval. We see a systematic downward correction over the oceans (-12.9 ± 8.3 ppb) except in persistently cloudy regions near the equator. Over land, the correction averages -5.5 ± 8.0 ppb. It is highest over bright arid surfaces, which are known to be difficult for TROPOMI retrievals (Lorente et al., 2021; Schneising et al., 2019). We also see large corrections at high northern latitudes that are seasonally driven (see below), and over tropical wetlands (Amazon, central Africa) where TROPOMI data are particularly sparse (Qu et al., 2021). Data south of 60°S (where the correction averages -16.2 ± 9.6 ppb) are excluded from these statistics and visualizations because of a lack of GOSAT data for evaluation. However, they are included in the blended TROPOMI+GOSAT data available for download.

Our correction is built on top of the TROPOMI operational v02.04.00 data that has already been bias-corrected with respect to SWIR surface albedo (Lorente et al., 2021). We compare these corrections in Appendix C.

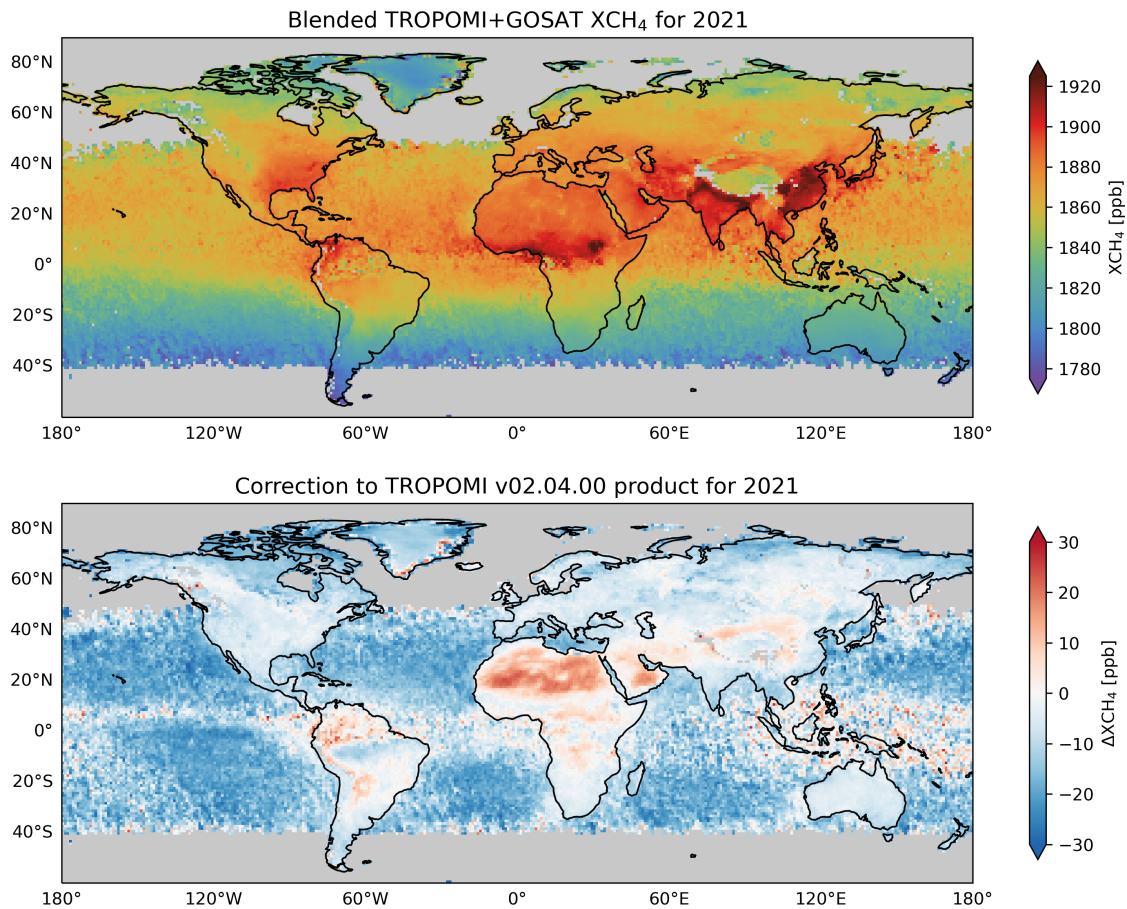



Figure 8: Blended TROPOMI+GOSAT product for 2021 (top) and correction to the TROPOMI operational v02.04.00 product (bottom). The data are annual means and plotted on a 1° × 1° grid with data south of 60°S cropped for visualization purposes. Grey areas have no TROPOMI data. The color bar in the bottom panel saturates at ± 30 ppb but there are outliers ranging from -48.1 ppb (East Siberian Sea) to +44.5 ppb (over the equatorial western Pacific).

Figure 9 shows the seasonal variation of the correction for 2021. Upward correction over arid surfaces due to SWIR surface albedo is consistent across seasons, but there is still seasonality in the correction over these regions driven by dust emission (and thus the aerosol size parameter). As a result, the correction over North Africa is largest in the boreal spring-summer, and the correction over East Asian deserts is largest in late boreal winter, reflecting the seasonality of dust emission (Shao and Dong, 2006; Senghor et al., 2017). There is large seasonal variation in the correction at high northern latitudes because of the low SWIR albedo of snow and ice-covered surfaces.

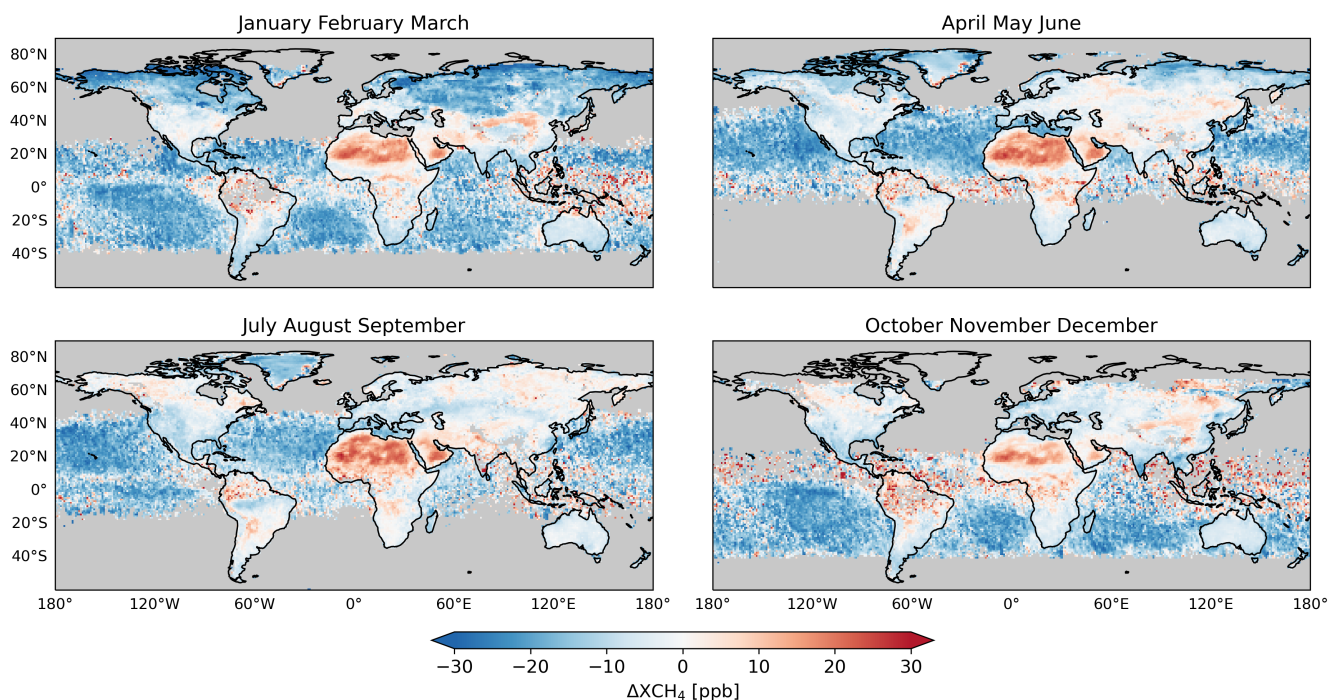



Figure 9: Seasonal correction to the TROPOMI data in the blended TROPOMI+GOSAT product. The Figure shows the differences with the TROPOMI operational v02.04.00 product averaged for each season in 2021. Data are plotted on a 1° × 1° grid and data south of 60°S are cropped for visualization purposes.

Figure 10 illustrates the correction over the Arabian Peninsula with annual mean oversampled data on a 0.01° × 0.01° grid. The TROPOMI operational v02.04.00 XCH_4 data show patterns that correlate with SWIR surface albedo, such as the XCH_4 gradients across Sudan and Saudi Arabia. These are removed in the blended product. The original data also show a number of hotspots over Iraq and Saudi Arabia that are removed in the blended product. These artifact enhancements tend to be related to the TROPOMI aerosol size distribution parameter, but they are more persistent than would be expected from aerosol plumes, suggesting that surface features might be aliasing into the aerosol retrieval. Other hotspots, such as those over Iran, are intensified in the blended product due to the high albedo in the region.

The TROPOMI operational v02.04.00 data show some coastal artifacts that are apparent in Figure 10 and that are not always fully corrected in the blended TROPOMI+GOSAT product. Coastal scenes are difficult to retrieve in full-physics algorithms because of the subpixel albedo contrast between dark water and bright land. Despite our correction, Figure 10 shows that some coastal areas have persistent biases, most evidently along the southern coast of the Gulf of Aden. The ability of the ML model to correct coastal biases may be limited by the diversity of coastal conditions and the small number of

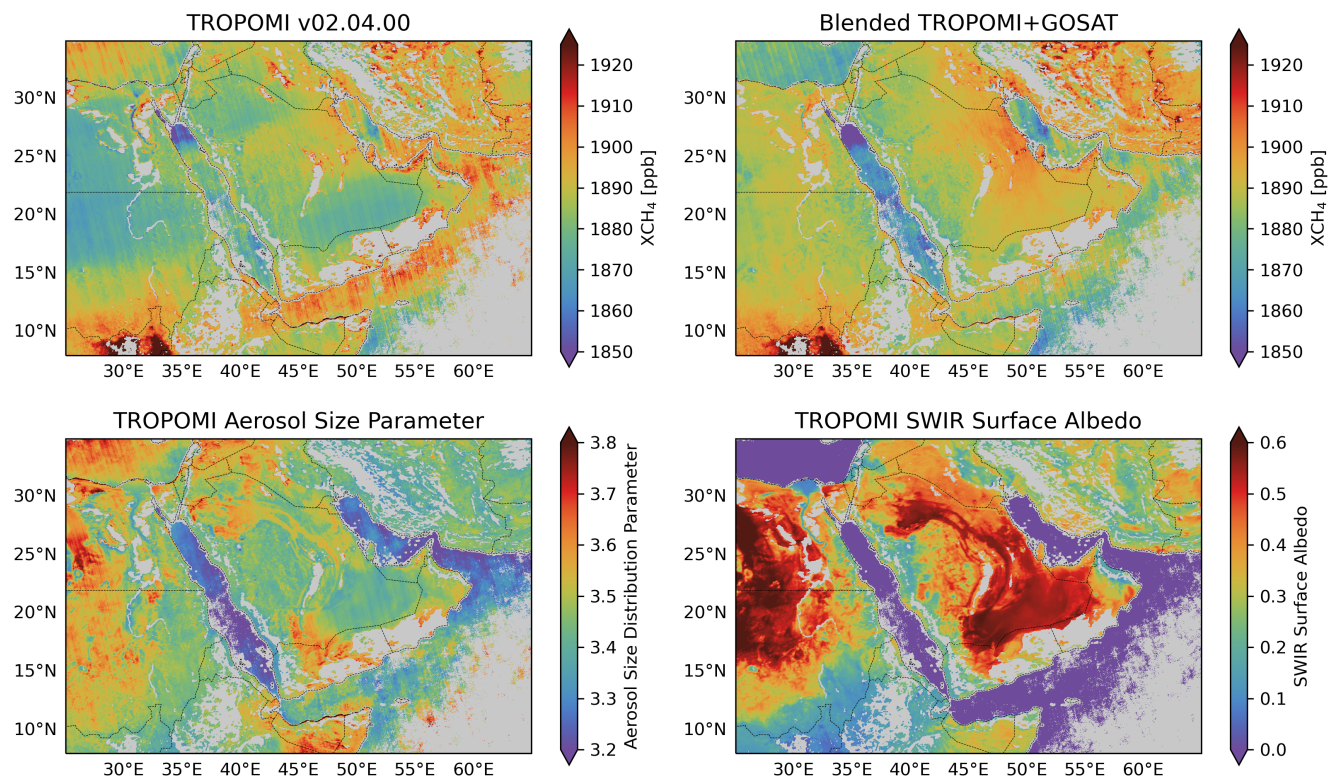
TROPOMI and GOSAT coastal data pairs available for training. Data users can choose to mitigate coastal bias by filtering out a subset of TROPOMI scenes that contain both land and water (Appendix D).

370

Figure 10: Blended TROPOMI+GOSAT and TROPOMI operational v02.04.00 data oversampled to a 0.01° × 0.01° grid over the Arabian Peninsula. Values are 2021 annual means. Also shown are the aerosol size distribution parameter and the SWIR surface albedo from the TROPOMI operational v02.04.00 retrieval. Oversampling to increase spatial resolution was done with the tessellation method following Zhu et al. (2017). Grey areas have less than 10 individual
375 **satellite observations contributing to the average.**

Our blended product also corrects transient biases from striping and atmospheric scattering that may affect single-scene observations of hotspots and plumes. This is illustrated in Figure 11 with a single-orbit scene (no temporal or spatial averaging) over Algeria on 15 December 2019 featuring a plume from an oil/gas ultra-emitting facility previously identified in the
380 TROPOMI data by Lauvaux et al. (2022). There is strong striping along the orbit track in the original single-orbit data from TROPOMI (Liu et al., 2021; Schneising et al., 2023) and this is substantially reduced in our product. The plume, shown in the center of Figure 11, was partially overlain by cirrus clouds (observed by VIIRS) that were not filtered out in the TROPOMI

retrieval and caused a low bias in the retrieved XCH₄. Our blended product corrects this cloud bias, enabling a better characterization of the plume to infer the source rate.

385

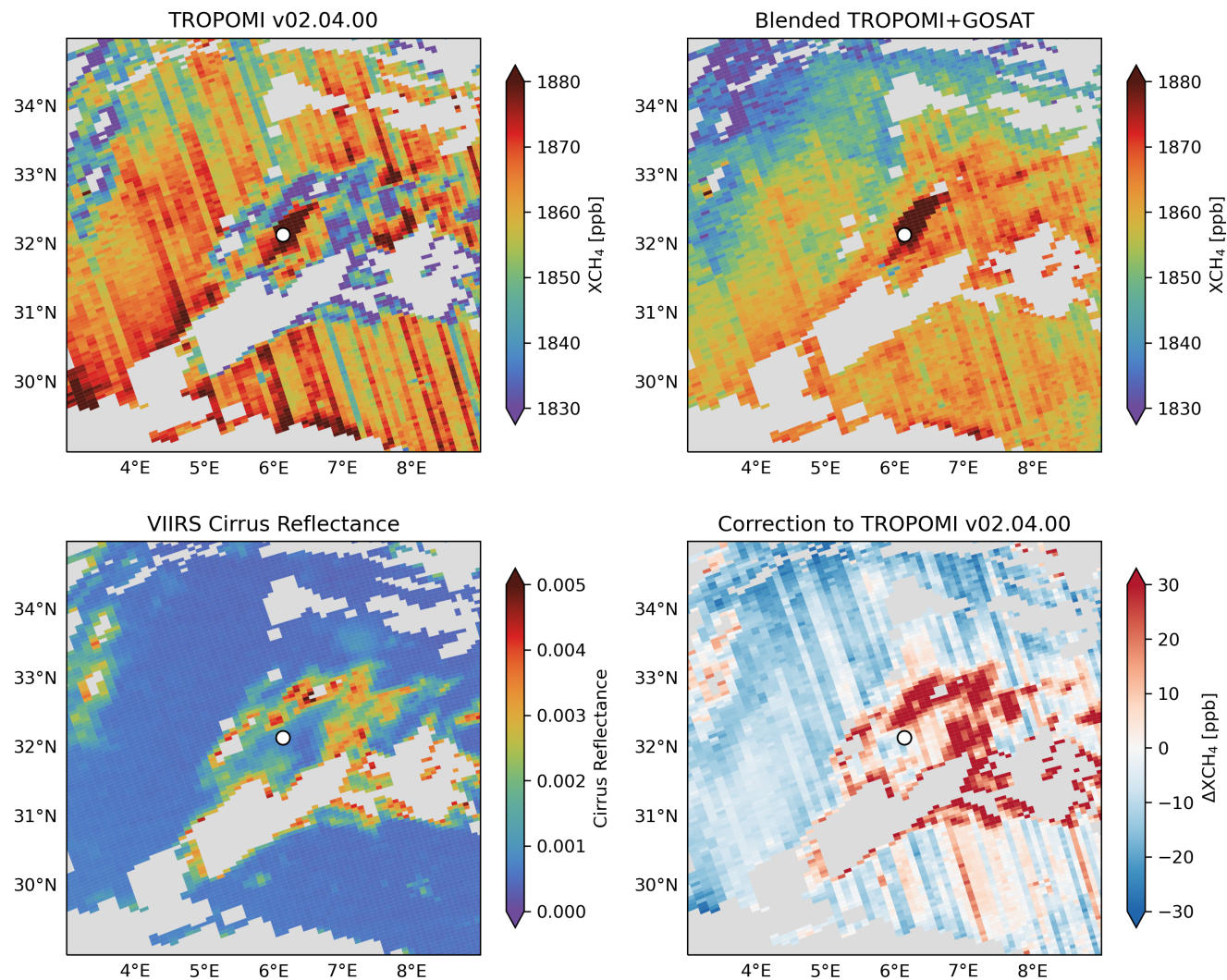



Figure 11: Methane ultra-emitter plume detection in single-orbit TROPOMI data. The Figure shows a scene over Algeria sampled by orbit number 11252 on 15 December 2019. Missing data are shown in grey. The enhancement in the center of the image was identified by Lauvaux et al. (2022) from TROPOMI data as a plume from an ultra-emitting oil/gas facility with an approximate location at the white circle marker. The TROPOMI operational v02.04.00 retrieval (top left panel) shows extensive striping and low values surrounding the plume that are biases from cirrus cloud reflectance (bottom left panel). This is effectively corrected in the blended TROPOMI+GOSAT product (right panels).

5 Conclusions

395 We presented a new blended TROPOMI+GOSAT methane product that corrects spatially variable biases and artifacts in the TROPOMI satellite instrument observations of atmospheric methane (XCH_4) by referencing to high-quality GOSAT observations. Our blended product improves the reliability of inversions of TROPOMI data to infer methane emissions and identify methane super-emitters in single-orbit and time-averaged observations. It includes the full dataset of TROPOMI retrievals from April 2018 to present and is available for download (see Data Availability section).

400

The blended product was generated by training a machine learning (ML) model (LightGBM) to predict the difference Δ(TROPOMI-GOSAT) between co-located GOSAT and TROPOMI methane retrievals for 2018-2020, using TROPOMI retrieval parameters as the sole predictor variables. This enabled subsequent application of the ML model to compute Δ(TROPOMI-GOSAT) for the full ensemble of TROPOMI data. The most important predictors of Δ(TROPOMI-GOSAT)

405 were found to be aerosol size, SWIR surface albedo, and across-track pixel index. The corrections are largest for observations with high albedos and coarse particles. Systematic downward correction averaging 12.9 ppb was found over water where the GOSAT XCH_4 glint retrievals are lower than TROPOMI.

Evaluation with independent ground-based TCCON XCH_4 data shows that our blended TROPOMI+GOSAT product

410 reduces the global mean bias in the TROPOMI data from 6.0 to -2.9 ppb, the variable bias from 4.7 to 4.4 ppb, and the single-retrieval precision from 14.5 to 11.9 ppb. However, the TCCON data are spatially limited and in particular do not sample regions with SWIR surface albedos larger than 0.4 where the largest TROPOMI biases relative to GOSAT are found. Nevertheless, evaluation with the TCCON data confirms that TROPOMI retrieval biases are related to SWIR surface albedo and aerosol size parameter, and that the blended TROPOMI+GOSAT product successfully reduces these biases. Global

415 evaluation of the blended TROPOMI+GOSAT product relative to GOSAT for 2021 as an independent test dataset shows a reduction in variable bias over land from 14.3 to 10.4 ppb on a 0.25° × 0.3125° grid (as might be used for regional inversions) and from 13.1 to 9.1 ppb on a 2° × 2.5° grid (as might be used for global inversions).

Annual mean corrections in the blended product relative to the TROPOMI operational v02.04.00 data exceed 10 ppb over

420 the oceans, desert regions (notably North Africa), persistently cloudy regions (notably tropical wetlands), and seasonally snow-covered regions (notably high northern latitudes). Large-scale corrections are mostly driven by SWIR surface albedo. Fine-scale inspection of the Arabian Peninsula reveals a number of annual mean hotspots in the original TROPOMI data that are removed in the blended product as artifacts. Some coastal artifacts remain in the blended product that can be filtered out at the discretion of the user.

425

The blended product also increases the quality of the single-orbit TROPOMI data by reducing striping and removing transient biases from aerosol plumes and cirrus clouds. This can increase confidence in the identification of ultra-emitters from TROPOMI hotspots and the quantitative interpretation of plume observations to infer point source rates.

430   Our correction will be applicable to all past and future operational TROPOMI v02.04.00 data (as well as v02.05.00, as it features no changes relative to v02.04.00). Users can download our blended TROPOMI+GOSAT product for 2018-present (see Data Availability) or apply the correction themselves to the operational product. A new version of the TROPOMI retrieval would require retraining the ML algorithm.

435   The ML framework presented here can be extended to any pair of satellite instruments in which one instrument provides a dense dataset while the other provides a more accurate but sparser dataset for the same variable. This situation often arises with a new satellite launch, as retrievals take time to mature, and an older, more established instrument may have been previously validated. Application of our approach to identify biases with the new instrument provides a far more spatially extensive evaluation than the traditional approach using surface sites or aircraft profiles. It also enables the identification of
440   the critical retrieval parameters that should be improved in the new instrument. Finally, it generates a blended product that corrects data from the new instrument.

Appendix A: Adjustment of TROPOMI, GOSAT, and TCCON data to common averaging kernel sensitivities and prior vertical profiles

Unbiased intercomparison of XCH$_4$ values retrieved from TROPOMI, GOSAT, and TCCON requires adjustments for the
445   different averaging kernel vertical sensitivities and prior vertical CH$_4$ concentration profiles used in the retrievals and reported as part of the retrieval products. We follow Schneising et al. (2019) and Buchwitz et al. (2022) to make these adjustments.

The following notation will be used. Column-averaged dry air mixing ratios of methane (XCH$_4$) are denoted as c with units of ppb. Vertical profiles of CH$_4$ sub column mixing ratios for pressure levels indexed by l are denoted as x^l and are either
450   retrieved (x_r^l) or prior estimates (x_a^l) with units of ppb. Vertical profiles of averaging kernels describing sensitivity are denoted by A^l and are dimensionless. Pressure weights that map vertical profiles x^l to XCH$_4$ are denoted as h and are dimensionless. Subscripts G, T, and F denote GOSAT, TROPOMI, and TCCON, respectively. Δ denotes the XCH$_4$ differences between pairs of instruments after adjustment to the same vertical sensitivities and prior estimates.

455   The GOSAT retrieval has more vertical pressure levels (19 or 20 depending on the retrieval) than TROPOMI (12, denoted pressure layers). It is therefore better to interpolate x^l from GOSAT to TROPOMI, following the principle of using the coarser vertical grid when comparing two different satellite retrievals (Rodgers and Connor, 2003). To calculate Δ(TROPOMI-

GOSAT), we first calculate what value c_T^* TROPOMI would have retrieved with GOSAT's prior profile.

$$c_T^* = c_{T,r} + \sum_l h_T^l (1 - A_T^l)(x_{G,a}^l - x_{T,a}^l) \quad \text{(A1)}$$

460

Next, we calculate what value c_G^* GOSAT would have retrieved with TROPOMI's vertical sensitivity.

465

$$c_G^* = \sum_l h_T^l \left(x_{G,a}^l + \left(x_{G,r}^l - x_{G,a}^l \right) A_T^l \right) \quad \text{(A2)}$$

Because the retrieved vertical profile of CH$_4$ is not reported for GOSAT, we estimate it here by scaling the prior profile by the ratio of retrieved to prior XCH$_4$ values.

$$x_{G,r} = x_{G,a} \frac{c_{G,r}}{c_{G,a}} \quad \text{(A3)}$$

470

Equations (A1) and (A2) require GOSAT's prior profile to be on the same pressure grid as TROPOMI. Interpolation is conducted from the 19 or 20 GOSAT pressure levels to the 12 TROPOMI pressure layers for this purpose. Equations (A1), (A2), and (A3) are then used to calculate Δ(TROPOMI-GOSAT).

475

$$\Delta(\text{TROPOMI} - \text{GOSAT}) = c_T^* - c_G^* \quad \text{(A4)}$$

The same procedure is used to calculate Δ(GOSAT-TCCON) and Δ(TROPOMI-TCCON). TCCON uses 51 pressure levels for its retrieval. For Δ(GOSAT-TCCON), we use TCCON's prior profile and GOSAT's averaging kernel sensitivities. For Δ(TROPOMI-TCCON), we use TCCON's prior profile and TROPOMI's averaging kernel sensitivities. An equation

480 analogous to equation (A3) is used to estimate the retrieved TCCON CH$_4$ profile. Equations (A5) and (A6) thus calculate Δ(GOSAT-TCCON) and Δ(TROPOMI-TCCON).

$$\Delta(\text{GOSAT} - \text{TCCON}) = \left[c_{G,r} + \sum_l h_G^l (1 - A_G^l)\left(x_{F,a}^l - x_{G,a}^l \right) \right] - \left[\sum_l h_G^l \left(x_{F,a}^l + \left(x_{F,r}^l - x_{F,a}^l \right) A_G^l \right) \right] \quad \text{(A5)}$$

485

$$\Delta(\text{TROPOMI} - \text{TCCON}) = \left[c_{T,r} + \sum_l h_T^l (1 - A_T^l)\left(x_{F,a}^l - x_{T,a}^l \right) \right] - \left[\sum_l h_T^l \left(x_{F,a}^l + \left(x_{F,r}^l - x_{F,a}^l \right) A_T^l \right) \right] \quad \text{(A6)}$$

Appendix B: Evaluation with TCCON data

We evaluated the GOSAT, TROPOMI, and blended TROPOMI+GOSAT products with the independent TCCON data for 30 April 2018-31 December 2021, correcting for retrieval differences in prior information and vertical sensitivities (Appendix A). We use the TCCON GGG2020 data version (https://tccondata.org, last accessed 18 March 2023) and consider all 24 stations that have reported measurements covering our study period (Table B1 and Figure 6).

Table B1. TCCON stations used for evaluation of the satellite data.

Site (lat., lon.)	Elevation (km a.s.l.)	Data Reference
Bremen (53.10, 8.85)	0.03	Notholt et al. (2022)
Burgos (18.53, 120.65)	0.04	Morino et al. (2022c)
East Trout Lake (54.35, -104.99)	0.50	Wunch et al. (2022)
Edwards (34.96, -117.88)	0.70	Iraci et al. (2022)
Eureka (80.05, -86.42)	0.61	Strong et al. (2022)
Garmisch (47.48, 11.06)	0.74	Sussmann and Rettinger (2023)
Hefei (31.90, 117.17)	0.03	Liu et al. (2022)
Izaña (28.30, -16.50) [a, b]	2.37	García et al. (2022)
JPL (34.20, -118.18) [a, b]	0.39	Wennberg et al. (2022a)
Karlsruhe (49.10, 8.44)	0.12	Hase et al. (2022)
Lamont (36.60, -97.49)	0.32	Wennberg et al. (2022d)
Lauder (-45.04, 169.68)	0.37	Sherlock et al. (2022); Pollard et al. (2022)
Nicosia (35.14, 33.38)	0.19	Petri et al. (2023)
Ny-Ålesund (78.92, 11.92) [a, b]	0.02	Buschmann et al. (2022)
Orléans (47.97, 2.11)	0.13	Warneke et al. (2022)
Paris (48.85, 2.36)	0.06	Te et al. (2022)
Park Falls (45.95, -90.27)	0.44	Wennberg et al. (2022b)
Pasadena (34.14, -118.13)	0.23	Wennberg et al. (2022c)
Réunion Island (-20.90, 55.49) [b]	0.09	De Mazière et al. (2022)
Rikubetsu (43.46, 143.77)	0.38	Morino et al. (2022a)
Saga (33.24, 130.29)	0.01	Shiomi et al. (2022)
Sodankylä (67.37, 26.63)	0.19	Kivi et al. (2022)
Tsukuba (36.05, 140.12)	0.03	Morino et al. (2022b)
Xianghe (39.75, 116.96)	0.04	Yang et al. (2020); Zhou et al. (2022)

495   [a] Excluded from the GOSAT evaluation due to a low number of pairs for comparison.

[b] Excluded from the TROPOMI and blended TROPOMI+GOSAT evaluations due to a low number of pairs for comparison.

The general evaluation framework is to identify co-located satellite and TCCON XCH$_4$ retrievals and compare these pairs. When evaluating TROPOMI or the blended TROPOMI+GOSAT product, satellite and TCCON pairs are defined to be those

500   within 1 hour and 100 km of each other and a surface elevation difference of no more than 250 m (some of the TCCON stations are on mountaintops). When evaluating the GOSAT product, satellite and TCCON pairs are defined to be those within 2 hours and 500 km of each other and a surface elevation difference of no more than 250 m. For all comparisons, a reduced radius of 50 km is used for the Edwards station (Schneising et al., 2019). We find 632683 TROPOMI-TCCON data pairs and 31093 GOSAT-TCCON data pairs. The TROPOMI-TCCON data pairs are also used to evaluate the blended TROPOMI+GOSAT

505   product.

For each station, we take the mean and standard deviation of all values of Δ(satellite-TCCON) to yield a station bias and station precision. The mean bias is the average of the station biases. The variable bias is the standard deviation of the station biases. The retrieval precision is the average of the station precisions. These metrics are calculated for 2018-2021 and for all

510   TCCON stations listed in Table B1 and mapped in Figure 6. Using a threshold of 100 satellite and TCCON pairs for a station to be used, Izaña, JPL, Ny-Ålesund, and Réunion Island are excluded from the TROPOMI and blended TROPOMI+GOSAT analyses, while Izaña, JPL, and Ny-Ålesund are excluded from the GOSAT analysis. The station biases and precisions by station are shown in Table B2.

515   **Table B2.** Comparison of satellite products with XCH$_4$ measured at TCCON stations. [a]

	GOSAT		TROPOMI		Blended	
Site	μ (ppb)	σ (ppb)	μ (ppb)	σ (ppb)	μ (ppb)	σ (ppb)
Bremen	0.3	14.3	7.6	13.4	-2.3	9.7
Burgos	-1.2	11.0	6.6	15.0	-8.1	9.8
East Trout Lake	1.6	18.6	10.4	20.7	4.2	14.0
Edwards	-1.4	10.6	-0.9	10.2	-7.4	8.0
Eureka	14.2	20.2	17.3	15.6	-2.3	16.5
Garmisch	0.8	15.9	13.1	15.1	6.3	13.3
Hefei	3.0	21.3	7.0	13.1	-1.7	12.2
Izaña	--	--	--	--	--	--
JPL	--	--	--	--	--	--

Karlsruhe	-3.9	14.8	4.6	14.2	-4.4	10.6
Lamont	-0.5	14.0	2.6	12.1	-2.9	10.4
Lauder	-1.7	9.7	-0.3	15.1	-8.9	10.9
Nicosia	-0.6	12.6	5.2	13.4	-4.4	11.7
Ny-Ålesund	--	--	--	--	--	--
Orléans	-1.2	12.6	5.3	12.6	-3.6	9.7
Paris	-3.2	13.7	5.0	13.0	-3.8	10.9
Park Falls	3.0	15.3	2.7	17.2	-0.9	13.7
Pasadena	-7.5	15.4	0.4	13.9	-11.7	11.8
Réunion Island	-9.4	10.6	--	--	--	--
Rikubetsu	9.9	17.0	8.3	15.5	3.2	13.2
Saga	-2.1	14.6	12.9	13.1	0.3	12.2
Sodankylä	1.6	16.4	4.9	19.9	0.3	12.9
Tsukuba	-4.9	13.2	3.1	10.8	-4.0	10.1
Xianghe	3.1	20.3	4.1	15.5	-5.4	15.6

[a] Mean (μ) and standard deviation (σ) of the satellite-TCCON difference in XCH_4 for co-located data over the 2018-2021 period. Station locations are listed in Table B1 and shown in Figure 6. Satellite and TCCON data have been corrected to the same prior estimates and averaging kernel sensitivities (Appendix A). Dashes indicate insufficient co-located data (see text).

Appendix C: Comparison to the operational TROPOMI Bias Correction

As described in Lorente et al. (2021), the operational v02.04.00 TROPOMI data include a bias correction for SWIR surface albedo. In our study, we have used these data as our starting point for computation of Δ(TROPOMI-GOSAT). This adds a bias-correction on top of a previous bias-correction. It is of interest to examine to what extent the previous bias correction was insufficient.

The operational SWIR surface albedo correction is derived using the "small area approximation," in which a few regions are selected around the globe where variation in SWIR surface albedo is observed but variation in XCH_4 is not expected. Referencing Aben et al. (2007), an albedo of 0.2 is selected as the best conditions for the retrieval and the correction is derived so that all retrievals in these regions match the retrieval at a SWIR surface albedo of 0.2.

Figure C1 shows the TROPOMI data before the operational albedo bias correction as well as the magnitude of this correction. The sum of the two figures gives the TROPOMI operational v02.04.00 data used in this study. Comparing the operational albedo bias correction to Figure 6 suggests that it is pushing values of XCH_4 to be too low over bright surface (North Africa, Arabian Peninsula) and to be too high over dark surfaces (snow-covered scenes).

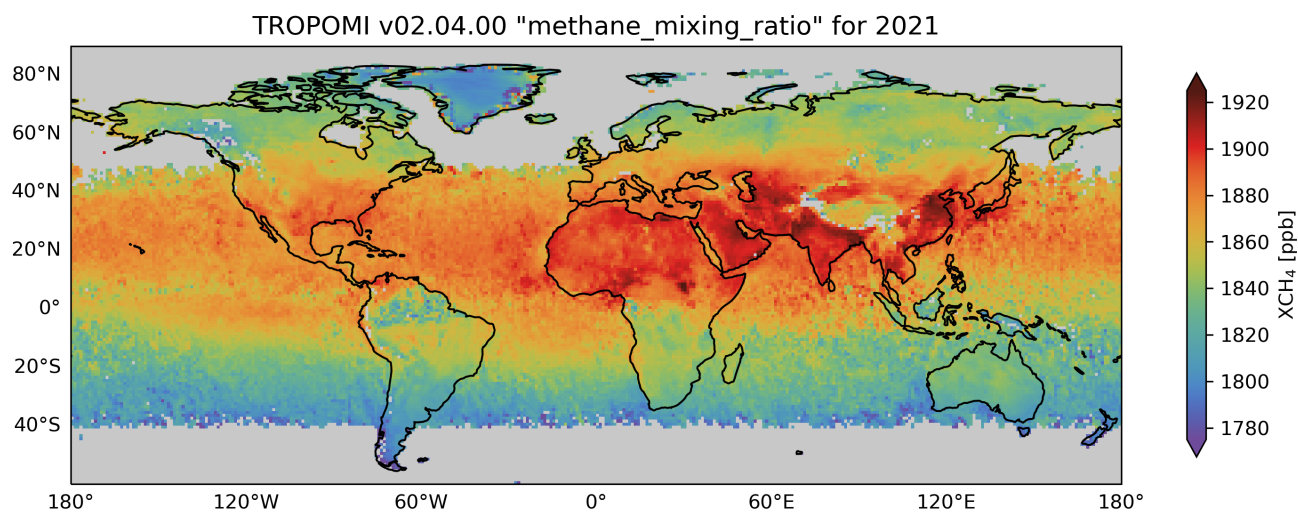
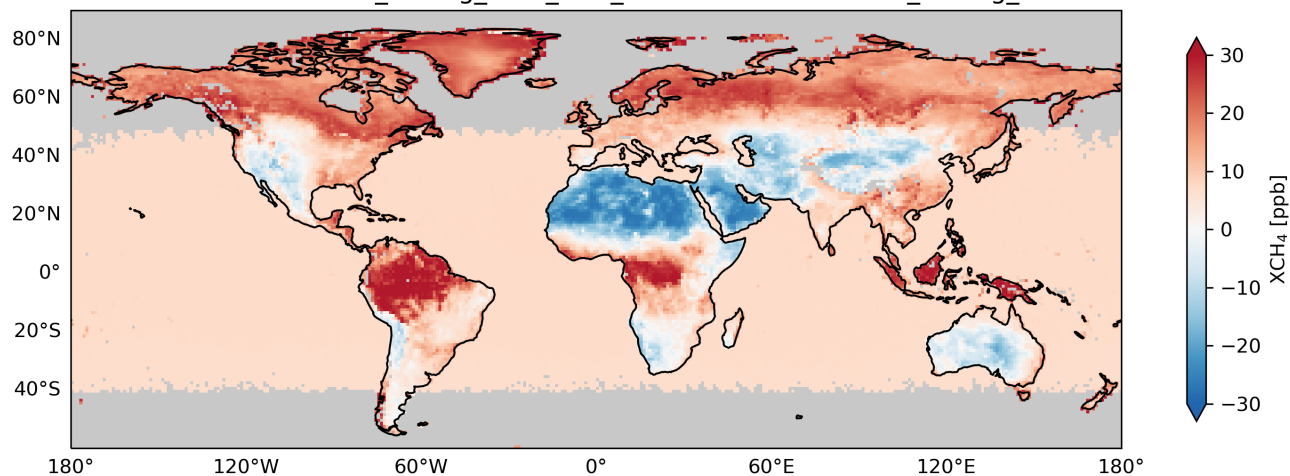



Figure C1. TROPOMI operational v02.04.00 XCH₄ product for 2021 before the operational albedo bias correction (top) and the operational albedo bias correction (bottom). The data are annual means and plotted on a 1° × 1° grid with data south of 60°S cropped for visualization purposes. Grey areas have no TROPOMI data. In the operational files, the top plot corresponds to "methane_mixing_ratio" and the sum of the top and bottom plot corresponds to "methane_mixing_ratio_bias_corrected."

540

Appendix D: Filtering Coastal TROPOMI Scenes

Retrieval pixels that include both land and water are problematic because of the subpixel albedo contrast. This can result in coastal biases, including for lakes and large rivers, that are not always successfully removed in our blended TROPOMI+GOSAT product. For example, in the top panel of Figure D1, there are enhancements of XCH_4 that outline the coast of North Africa. This can be fully avoided by filtering out all retrievals with a "surface_classification" value of 3 (pixel contains mostly water with some land) and retrievals with a "surface_classification" value of 2 (pixel contains mostly land with some water) but this excludes 8% of the global data. Filtering out all pixels with "surface_classification" value of 3 (0.3% of data) and the subset of retrievals with "surface_classification" value of 2 and SWIR Chi-Squared greater than 20000 (see Table 2) largely corrects coastal artifacts while excluding only 0.6% of the data. The middle panel of Figure D1 shows the pixels that are removed by this filter, and the bottom panel shows the blended product after removal of these pixels. We keep these coastal pixels in our blended product and leave it to the user to decide what filters to apply.

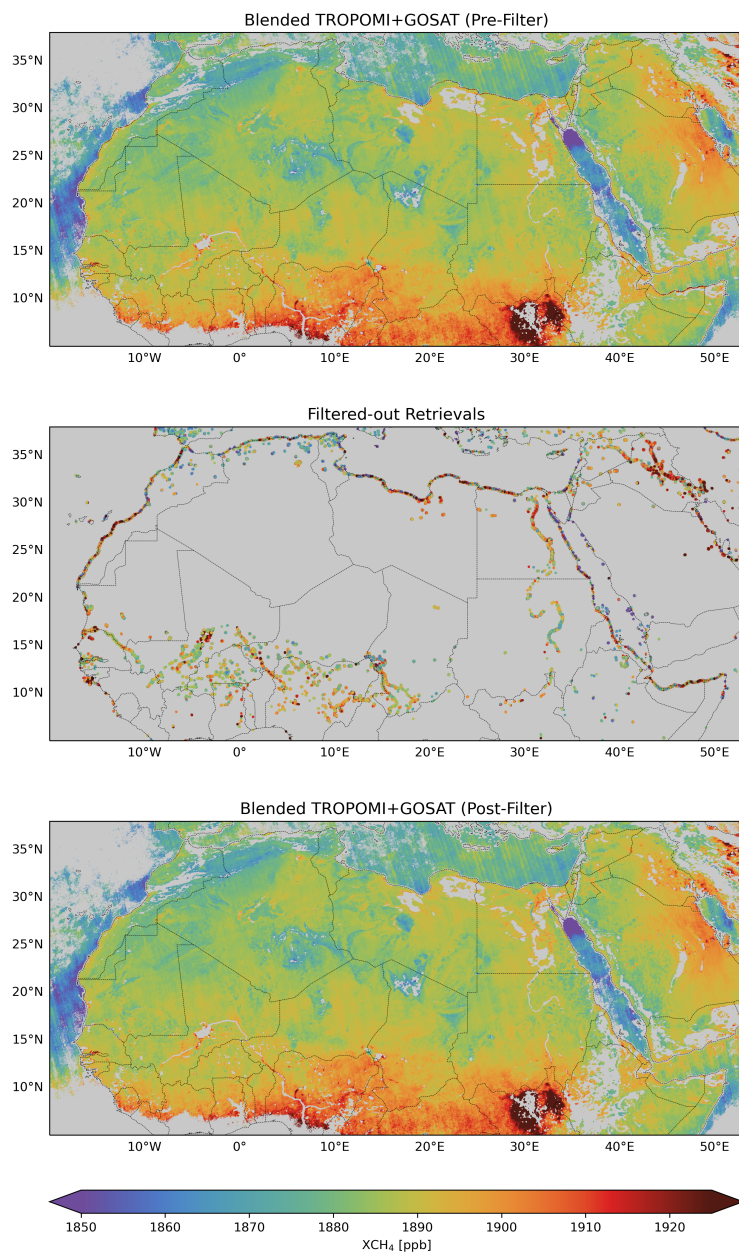



Figure D1: Filtering of coastal pixels in the blended TROPOMI+GOSAT XCH₄ product. The top panel shows the unfiltered data over North Africa, oversampled on a 0.01° × 0.01° grid and averaged for 2021. The middle panel shows our coastal filter excluding all pixels with a "surface_classification" value of 3 and the subset of pixels with a "surface_classification" value of 2 and a SWIR Chi-Squared greater than 20000 (Table 2). The bottom panel shows the filtered data. Grid cells with less than 10 individual observations contributing to the oversampled average are shown in grey in the top and bottom panels.

Code availability. The code used for all portions of this project is available at
https://github.com/nicholasbalasus/blended_tropomi_gosat_methane and the code is archived on Zenodo at
565 https://doi.org/10.5281/zenodo.8136738.

Data availability. The blended TROPOMI+GOSAT methane product data are available for April 2018 to present on Harvard
Dataverse at https://dataverse.harvard.edu/dataverse/blended-tropomi-gosat-methane. The TROPOMI data used here are
available at https://s5phub.copernicus.eu/dhus/ for April 2018 to present. The GOSAT data used here are available at
570 http://dx.doi.org/10.5285/18ef8247f52a4cb6a14013f8235cc1eb for 2009 to 2021. The TCCON data were obtained from the
TCCON Data Archive hosted by CaltechDATA at https://tccondata.org.

Author contributions. NB and DJJ designed the study. NB performed the analysis with contributions from AL, JDM, RJP,
HB, ZC, MMK, HN, and DJV. NB and DJJ led the writing of the paper with contributions from all co-authors.
575

Acknowledgements. The authors thank the team that realized the TROPOMI instrument and its data products, consisting of
the partnership between Airbus Defense and Space Netherlands, KNMI, SRON, and TNO, commissioned by NSO and ESA.
Sentinel-5 Precursor is part of the EU Copernicus program, and Coperncius (modified) Sentinel-5P data (2018-2021) have
been used. We gratefully acknowledge TCCON site PIs for the data used in this work and their guidance on its use. This
580 research used the ALICE high-performance computing facility at the University of Leicester for the GOSAT retrievals and
analysis. We thank the Japanese Aerospace Exploration Agency, National Institute for Environmental Studies and the
Ministry of Environment for the GOSAT data and their continuous support as part of the Joint Research Agreement.

Competing interests. The authors declare that they have no conflict of interest.
585

Financial support. This work was supported by the NASA Carbon Monitoring System. Nicholas Balasus was supported by
the Department of Defense (DoD) through the National Defense Science & Engineering Graduate (NDSEG) Fellowship
Program. Robert J. Parker and Hartmut Boesch are funded via the UK National Centre for Earth Observation (grant nos.
NE/R016518/1 and NE/N018079/1), as well as from the ESA GHG-CCI and Copernicus C3S projects (grant no.
590 C3S2_312a_Lot2).

References

Aas, K., Jullum, M., and Løland, A.: Explaining individual predictions when features are dependent: More accurate
approximations to Shapley values, Artif. Intell., 298, 103502, https://doi.org/10.1016/j.artint.2021.103502, 2021.

595   Aben, I., Hasekamp, O., and Hartmann, W.: Uncertainties in the space-based measurements of CO2 columns due to scattering in the Earth's atmosphere, J. Quant. Spectrosc. Radiat. Transf., 104, 450–459, https://doi.org/10.1016/j.jqsrt.2006.09.013, 2007.

Apituley, A., Pedergnana, M., Sneep, M., Veefkind, J. P., Loyola, D., Hasekamp, O., Delgado, A. L., and Borsdorff, T.:
600   Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Methane v2.4.0, https://sentinels.copernicus.eu/documents/247904/2474726/Sentinel-5P-Level-2-Product-User-Manual-Methane.pdf (last access: 3 Jul 2023), 2022.

Barré, J., Aben, I., Agustí-Panareda, A., Balsamo, G., Bousserez, N., Dueben, P., Engelen, R., Inness, A., Lorente, A.,
605   McNorton, J., Peuch, V.-H., Radnoti, G., and Ribas, R.: Systematic detection of local CH4 anomalies by combining satellite measurements with high-resolution forecasts, Atmos. Chem. Phys., 21, 5117–5136, https://doi.org/10.5194/acp-21-5117-2021, 2021.

Belitz, K. and Stackelberg, P. E.: Evaluation of six methods for correcting bias in estimates from ensemble tree machine
610   learning regression models, Environ. Model. Softw., 139, 105006, https://doi.org/10.1016/j.envsoft.2021.105006, 2021.

Borsdorff, T., aan de Brugh, J., Hu, H., Hasekamp, O., Sussmann, R., Rettinger, M., Hase, F., Gross, J., Schneider, M., Garcia, O., Stremme, W., Grutter, M., Feist, D. G., Arnold, S. G., De Mazière, M., Kumar Sha, M., Pollard, D. F., Kiel, M., Roehl, C., Wennberg, P. O., Toon, G. C., and Landgraf, J.: Mapping carbon monoxide pollution from space down to city
615   scales with daily global coverage, Atmos. Meas. Tech., 11, 5507–5518, https://doi.org/10.5194/amt-11-5507-2018, 2018.

Borsdorff, T., aan de Brugh, J., Schneider, A., Lorente, A., Birk, M., Wagner, G., Kivi, R., Hase, F., Feist, D. G., Sussmann, R., Rettinger, M., Wunch, D., Warneke, T., and Landgraf, J.: Improving the TROPOMI CO data product: update of the spectroscopic database and destriping of single orbits, Atmos. Meas. Tech., 10, 5443-5455, https://doi.org/10.5194/amt-12-
620   5443-2019, 2019.

Breiman, L.: Random Forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Guerlet, S., Dils, B., Aben, I., Armante, R., Bergamaschi, P.,
625   Blumenstock, T., Bovensmann, H., Brunner, D., Buchmann, B., Burrows, J. P., Butz, A., Chédin, A., Chevallier, F., Crevoisier, C. D., Deutscher, N. M., Frankenberg, C., Hase, F., Hasekamp, O. P., Heymann, J., Kaminski, T., Laeng, A., Lichtenberg, G., De Mazière, M., Noël, S., Notholt, J., Orphal, J., Popp, C., Parker, R., Scholze, M., Sussmann, R., Stiller,

G. P., Warneke, T., Zehner, C., Bril, A., Crisp, D., Griffith, D. W. T., Kuze, A., O'Dell, C., Oshchepkov, S., Sherlock, V., Suto, H., Wennberg, P., Wunch, D., Yokota, T., and Yoshida, Y.: The Greenhouse Gas Climate Change Initiative (GHG-
630   CCI): Comparison and quality assessment of near-surface-sensitive satellite-derived CO2 and CH4 global data sets, Remote Sens. Environ., 162, 344–362, https://doi.org/10.1016/j.rse.2013.04.024, 2015.

Buchwitz, M., Dils, B., Reuter, M., Schneising, O., Hilker, M., Preval, S., Boesch, H., Borsdorff, T., Landgraf, J., and Krisna, T. C.: ESA Climate Change Initiative "Plus" (CCI+) Product Validation and Intercomparison Report (PVIR) version
635   3 for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), https://climate.esa.int/media/documents/PVIR_GHG-CCIp_v3_reducedsize.pdf (last access: 27 Jan 2023), 2022.

Buschmann, M., Petri, C., Palm, M., Warneke, T., Notholt, J., and Engineers, A. S.: TCCON data from Ny-Ålesund, Svalbard, Norway, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of
640   Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.nyalesund01.R0, 2022.

Butz, A., Hasekamp, O. P., Frankenberg, C., Vidot, J., and Aben, I.: CH4 retrievals from space-based solar backscatter measurements: Performance evaluation against simulated aerosol and cirrus loaded scenes, J. Geophys. Res., 115, D24302, https://doi.org/10.1029/2010JD014514, 2010.
645
Butz, A., Galli, A., Hasekamp, O., Landgraf, J., Tol, P., and Aben, I.: TROPOMI aboard Sentinel-5 Precursor: Prospective performance of CH4 retrievals for aerosol and cirrus loaded atmospheres, Remote Sens. Environ., 120, 267–276, https://doi.org/10.1016/j.rse.2011.05.030, 2012.

650   Carroll, M. L., Townshend, J. R., DiMiceli, C. M., Noojipady, P., and Sohlberg, R. A.: A new global raster water mask at 250 m resolution, Int. J. Digit. Earth, 2, 291–308, https://doi.org/10.1080/17538940902951401, 2009.

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794, https://doi.org/10.1145/2939672.2939785,
655   2016.

De Mazière, M., Sha, M. K., Desmet, F., Hermans, C., Scolas, F., Kumps, N., Zhou, M., Metzger, J.-M., Duflot, V., and Cammas, J.-P.: TCCON data from Réunion Island (La Reunion), France, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set],
660   https://doi.org/10.14291/tccon.ggg2020.reunion01.R0, 2022.

Feng, L., Palmer, P. I., Zhu, S., Parker, R. J., and Liu, Y.: Tropical methane emissions explain large fraction of recent changes in global atmospheric methane growth rate, Nat. Commun., 13, 1378, https://doi.org/10.1038/s41467-022-28989-z, 2022.

665

Frankenberg, C., Meirink, J. F., van Weele, M., Platt, U., and Wagner, T.: Assessing Methane Emissions from Global Space-Borne Observations, Science, 308, 1010–1014, https://doi.org/10.1126/science.1106644, 2005.

García, O. E., Schneider, M., Herkommer, B., Gross, J., Hase, F., Blumenstock, T., and Sepúlveda, E.: TCCON data from

670   Izaña, Tenerife, Spain, Release GGG2020R01. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.izana01.R1, 2022.

Hase, F., Blumenstock, T., Dohe, S., Groß, J., and Kiel, M.: TCCON data from Karlsruhe, Germany, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set],

675   https://doi.org/10.14291/tccon.ggg2020.karlsruhe01.R0, 2022.

Hasekamp, O., Lorente, A., Hu, H., Butz, A., Aan de Brugh, J., and Landgraf, J.: Algorithm Theoretical Baseline Document for Sentinel-5 Precursor Methane Retrieval v2.4.0, https://sentinels.copernicus.eu/documents/247904/2476257/Sentinel-5P-TROPOMI-ATBD-Methane-retrieval.pdf (last access: 3 Jul 2023), 2022.

680

Iraci, L., Podolske, J., Roehl, C., Wennberg, P. O., Blavier, J.-F., Allen, N., Wunch, D., and Osterman, G.: TCCON data from Armstrong Flight Research Center, Edwards, CA, USA, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.edwards01.R0, 2022.

685

Jacob, D. J., Turner, A. J., Maasakkers, J. D., Sheng, J., Sun, K., Liu, X., Chance, K., Aben, I., McKeever, J., and Frankenberg, C.: Satellite observations of atmospheric methane and their value for quantifying methane emissions, Atmos. Chem. Phys., 16, 14371–14396, https://doi.org/10.5194/acp-16-14371-2016, 2016.

690   Jacob, D. J., Varon, D. J., Cusworth, D. H., Dennison, P. E., Frankenberg, C., Gautam, R., Guanter, L., Kelley, J., McKeever, J., Ott, L. E., Poulter, B., Qu, Z., Thorpe, A. K., Worden, J. R., and Duren, R. M.: Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane, Atmos. Chem. Phys., 22, 9617–9646, https://doi.org/10.5194/acp-22-9617-2022, 2022.

695   Janardanan, R., Maksyutov, S., Tsuruta, A., Wang, F., Tiwari, Y. K., Valsala, V., Ito, A., Yoshida, Y., Kaiser, J. W., Janssens-Maenhout, G., Arshinov, M., Sasakawa, M., Tohjima, Y., Worthy, D. E. J., Dlugokencky, E. J., Ramonet, M., Arduini, J., Lavric, J. V., Piacentino, S., Krummel, P. B., Langenfelds, R. L., Mammarella, I., and Matsunaga, T.: Country-Scale Analysis of Methane Emissions with a High-Resolution Inverse Model Using GOSAT and Surface Observations, Remote Sens., 12, https://doi.org/10.3390/rs12030375, 2020.

700

Jongaramrungruang, S., Matheou, G., Thorpe, A. K., Zeng, Z.-C., and Frankenberg, C.: Remote sensing of methane plumes: instrument tradeoff analysis for detecting and quantifying local sources at global scale, Atmos. Meas. Tech., 14, 7999–8017, https://doi.org/10.5194/amt-14-7999-2021, 2021.

705   Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.: LightGBM: A Highly Efficient Gradient Boosting Decision Tree, in: Advances in Neural Information Processing Systems, 2017.

Kingsford, C. and Salzberg, S. L.: What are decision trees?, Nat. Biotechnol., 26, 1011–1013, https://doi.org/10.1038/nbt0908-1011, 2008.

710

Kivi, R., Heikkinen, P., and Kyro, E.: TCCON data from Sodankylä, Finland, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.sodankyla01.R0, 2022.

715   Kuze, A., Suto, H., Shiomi, K., Kawakami, S., Tanaka, M., Ueda, Y., Deguchi, A., Yoshida, J., Yamamoto, Y., Kataoka, F., Taylor, T. E., and Buijs, H. L.: Update on GOSAT TANSO-FTS performance, operations, and data products after more than 6 years in space, Atmos. Meas. Tech., 9, 2445–2461, https://doi.org/10.5194/amt-9-2445-2016, 2016.

Lauvaux, T., Giron, C., Mazzolini, M., d'Aspremont, A., Duren, R., Cusworth, D., Shindell, D., and Ciais, P.: Global
720   assessment of oil and gas methane ultra-emitters, Science, 375, 557–561, https://doi.org/10.1126/science.abj4351, 2022.

Liu, M., van der A, R., van Weele, M., Eskes, H., Lu, X., Veefkind, P., de Laat, J., Kong, H., Wang, J., Sun, J., Ding, J., Zhao, Y., and Weng, H.: A New Divergence Method to Quantify Methane Emissions Using Observations of Sentinel-5P TROPOMI, Geophys. Res. Lett., 48, e2021GL094151, https://doi.org/10.1029/2021GL094151, 2021.

725

Liu, C., Wang, W., and Sun, Y.: TCCON data from Hefei, China, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.hefei01.R0, 2022.

730   Lorente, A., Borsdorff, T., Butz, A., Hasekamp, O., aan de Brugh, J., Schneider, A., Wu, L., Hase, F., Kivi, R., Wunch, D., Pollard, D. F., Shiomi, K., Deutscher, N. M., Velazco, V. A., Roehl, C. M., Wennberg, P. O., Warneke, T., and Landgraf, J.: Methane retrieved from TROPOMI: improvement of the data product and validation of the first 2 years of measurements, Atmos. Meas. Tech., 14, 665–684, https://doi.org/10.5194/amt-14-665-2021, 2021.

735   Lorente, A., Borsdorff, T., Martinez-Velarte, M. C., Butz, A., Hasekamp, O. P., Wu, L., and Landgraf, J.: Evaluation of the methane full-physics retrieval applied to TROPOMI ocean sun glint measurements, Atmos. Meas. Tech., 15, 6585–6603, https://doi.org/10.5194/amt-15-6585-2022, 2022.

       Lorente, A., Borsdorff, T., Martinez-Velarte, M. C., and Landgraf, J.: Accounting for surface reflectance spectral features in
740   TROPOMI methane retrievals, Atmos. Meas. Tech., 16, 1597-1608, https://doi.org/10.5194/amt-2022-255, 2023.

       Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.-I.: From local explanations to global understanding with explainable AI for trees, Nat. Mach. Intell., 2, 56–67, https://doi.org/10.1038/s42256-019-0138-9, 2020.
745

       Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J.-X., Zhang, Y., Hersher, M., Bloom, A. 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, 7859–7881, https://doi.org/10.5194/acp-19-7859-2019, 2019.
750

       Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J., Zhang, Y., Lu, X., Bloom, A. A., Bowman, K. W., Worden, J. R., and Parker, R. J.: 2010–2015 North American methane emissions, sectoral contributions, and trends: a high-resolution inversion of GOSAT observations of atmospheric methane, Atmos. Chem. Phys., 21, 4339–4356, https://doi.org/10.5194/acp-21-4339-2021, 2021.
755

       Morino, I., Ohyama, H., Hori, A., and Ikegami, H.: TCCON data from Rikubetsu, Hokkaido, Japan, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.rikubetsu01.R0, 2022a.

760   Morino, I., Ohyama, H., Hori, A., and Ikegami, H.: TCCON data from Tsukuba, Ibaraki, Japan, 125HR, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.tsukuba02.R0, 2022b.

Morino, I., Velazco, V. A., Hori, A., Uchino, O., and Griffith, D. W. T.: TCCON data from Burgos, Philippines, Release
GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A.
[data set], https://doi.org/10.14291/tccon.ggg2020.burgos01.R0, 2022c.

Nisbet, E. G., Fisher, R. E., Lowry, D., France, J. L., Allen, G., Bakkaloglu, S., Broderick, T. J., Cain, M., Coleman, M.,
Fernandez, J., Forster, G., Griffiths, P. T., Iverach, C. P., Kelly, B. F. J., Manning, M. R., Nisbet-Jones, P. B. R., Pyle, J. A.,
Townsend-Small, A., al-Shalaan, A., Warwick, N., and Zazzeri, G.: Methane Mitigation: Methods to Reduce Emissions, on
the Path to the Paris Agreement, Rev. Geophys., 58, e2019RG000675, https://doi.org/10.1029/2019RG000675, 2020.

Notholt, J., Petri, C., Warneke, T., Deutscher, N., Buschmann, M., Weinzierl, C., Macatangay, R., and Grupe, P.: TCCON
data from Bremen, Germany, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of
Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.bremen01.R0, 2022.

Ouyang, Z., Jackson, R. B., McNicol, G., Fluet-Chouinard, E., Runkle, B. R. K., Papale, D., Knox, S. H., Cooley, S.,
Delwiche, K. B., Feron, S., Irvin, J. A., Malhotra, A., Muddasir, M., Sabbatini, S., Alberto, Ma. C. R., Cescatti, A., Chen, C.-
L., Dong, J., Fong, B. N., Guo, H., Hao, L., Iwata, H., Jia, Q., Ju, W., Kang, M., Li, H., Kim, J., Reba, M. L., Nayak, A. K.,
Roberti, D. R., Ryu, Y., Swain, C. K., Tsuang, B., Xiao, X., Yuan, W., Zhang, G., and Zhang, Y.: Paddy rice methane
emissions across Monsoon Asia, Remote Sens. Environ., 284, 113335, https://doi.org/10.1016/j.rse.2022.113335, 2023.

Palmer, P. I., Feng, L., Lunt, M. F., Parker, R. J., Bösch, H., Lan, X., Lorente, A., and Borsdorff, T.: The added value of
satellite observations of methane for understanding the contemporary methane budget, Phil. Trans. R. Soc. A., 379,
20210106, https://doi.org/10.1098/rsta.2021.0106, 2021.

Parker, R., Boesch, H., Cogan, A., Fraser, A., Feng, L., Palmer, P. I., Messerschmidt, J., Deutscher, N., Griffith, D. W. T.,
Notholt, J., Wennberg, P. O., and Wunch, D.: Methane observations from the Greenhouse Gases Observing SATellite:
Comparison to ground-based TCCON data and model calculations, Geophys. Res. Lett., 38, L15807,
https://doi.org/10.1029/2011GL047871, 2011.

Parker, R. J., Webb, A., Boesch, H., Somkuti, P., Barrio Guillo, R., Di Noia, A., Kalaitzi, N., Anand, J. S., Bergamaschi, P.,
Chevallier, F., Palmer, P. I., Feng, L., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Kivi, R., Morino, I.,
Notholt, J., Oh, Y.-S., Ohyama, H., Petri, C., Pollard, D. F., Roehl, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té,
Y., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: A decade of GOSAT Proxy satellite CH4 observations,
Earth Syst. Sci. Data, 12, 3383–3412, https://doi.org/10.5194/essd-12-3383-2020, 2020.

Petri, C., Vrekoussis, M., Rousogenous, C., Warneke, T., Sciare, J., and Notholt, J.: TCCON data from Nicosia, Cyprus, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.nicosia01.R0, 2023.

800

Pollard, D., Robinson, J., and Shiona, H.: TCCON data from Lauder, New Zealand, 125HR, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.lauder03.R0, 2022.

805

Prather, M. J., Holmes, C. D., and Hsu, J.: Reactive greenhouse gas scenarios: Systematic exploration of uncertainties and the role of atmospheric chemistry, Geophys. Res. Lett., 39, L09803, https://doi.org/10.1029/2012GL051440, 2012.

Qu, Z., Jacob, D. J., Shen, L., Lu, X., Zhang, Y., Scarpelli, T. R., Nesser, H., Sulprizio, M. P., Maasakkers, J. D., Bloom, A.

810 A., Worden, J. R., Parker, R. J., and Delgado, A. L.: Global distribution of methane emissions: a comparative inverse analysis of observations from the TROPOMI and GOSAT satellite instruments, Atmos. Chem. Phys., 21, 14159–14175, https://doi.org/10.5194/acp-21-14159-2021, 2021.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, J. Geophys. Res., 108,

815 https://doi.org/10.1029/2002JD002299, 2003.

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G.,

820 Frankenberg, C., Gedney, N., Hegglin, M. I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L., Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton, J. R., Morino, I., Müller, J., Murguia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R. J., Peng, C., Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J. A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, L. P., Thornton, B. F.,

825 Tian, H., Tohjima, Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der Werf, G. R., Weiss, R. F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: The Global Methane Budget 2000–2017, Earth Syst. Sci. Data, 12, 1561–1623, https://doi.org/10.5194/essd-12-1561-2020, 2020.

830 Schepers, D., Guerlet, S., Butz, A., Landgraf, J., Frankenberg, C., Hasekamp, O., Blavier, J.-F., Deutscher, N. M., Griffith, D. W. T., Hase, F., Kyro, E., Morino, I., Sherlock, V., Sussmann, R., and Aben, I.: Methane retrievals from Greenhouse Gases Observing Satellite (GOSAT) shortwave infrared measurements: Performance comparison of proxy and physics retrieval algorithms, J. Geophys. Res., 117, D10307, https://doi.org/10.1029/2012JD017549, 2012.

835 Schneising, O., Buchwitz, M., Reuter, M., Bovensmann, H., Burrows, J. P., Borsdorff, T., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Hermans, C., Iraci, L. T., Kivi, R., Landgraf, J., Morino, I., Notholt, J., Petri, C., Pollard, D. F., Roche, S., Shiomi, K., Strong, K., Sussmann, R., Velazco, V. A., Warneke, T., and Wunch, D.: A scientific algorithm to simultaneously retrieve carbon monoxide and methane from TROPOMI onboard Sentinel-5 Precursor, Atmos. Meas. Tech., 12, 6771–6802, https://doi.org/10.5194/amt-12-6771-2019, 2019.

840

Schneising, O., Buchwitz, M., Hachmeister, J., Vanselow, S., Reuter, M., Buschmann, M., Bovensmann, H., and Burrows, J. P.: Advances in retrieving XCH4 and XCO from Sentinel-5 Precursor: improvements in the scientific TROPOMI/WFMD algorithm, Atmos. Meas. Tech., 16, 669–694, https://doi.org/10.5194/amt-16-669-2023, 2023.

845 Senghor, H., Machu, É., Hourdin, F., and Gaye, A. T.: Seasonal cycle of desert aerosols in western Africa: analysis of the coastal transition with passive and active sensors, Atmos. Chem. Phys., 17, 8395–8410, https://doi.org/10.5194/acp-17-8395-2017, 2017.

Shao, Y. and Dong, C. H.: A review on East Asian dust storm climate, modelling and monitoring, Glob. Planetary Change,
850 52, 1–22, https://doi.org/10.1016/j.gloplacha.2006.02.011, 2006.

Sherlock, V., Connor, B., Robinson, J., Shiona, H., Smale, D., and Pollard, D. F.: TCCON data from Lauder (NZ), 125HR, Release GGG2020.R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.lauder02.R0, 2022.

855

Shiomi, K., Kawakami, S., Ohyama, H., Arai, K., Okumura, H., Ikegami, H., and Usami, M.: TCCON data from Saga, Japan, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.saga01.R0, 2022.

860 Silva, S. J., Keller, C. A., and Hardin, J.: Using an Explainable Machine Learning Approach to Characterize Earth System Model Errors: Application of SHAP Analysis to Modeling Lightning Flash Occurrence, JAMES, 14, e2021MS002881, https://doi.org/10.1029/2021MS002881, 2022.

Strong, K., Roche, S., Franklin, J. E., Mendonca, J., Lutsch, E., Weaver, D., Fogal, P. F., Drummond, J. R., Batchelor, R.,

865   and Lindenmaier, R.: TCCON data from Eureka, Canada, Release GGG2020R0. TCCON data archive, hosted by
CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set],
https://doi.org/10.14291/tccon.ggg2020.eureka01.R0, 2022.

Sussmann, R. and Rettinger, M.: TCCON data from Garmisch, Germany, Release GGG2020R0. TCCON data archive,

870   hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set],
https://doi.org/10.14291/tccon.ggg2020.garmisch01.R0, 2022.

Szopa, S., Naik, V., Adhikary, B., Artaxo, P., Berntsen, T., Collins, W. D., Fuzzi, S., Gallardo, L., Kiendler-Scharr, A.,
Kilmont, Z., Liao, H., Unger, N., and Zanis, P.: Short-Lived Climate Forcers, in: Climate Change 2021: The Physical

875   Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate
Change, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y.,
Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi,
O., Yu, R., and Zhou, B., Cambridge University Press,
https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Chapter06.pdf (last access: 27 Jan 2023), 2021.

880
Te, Y., Jeseck, P., and Janssen, C.: TCCON data from Paris, France, Release GGG2020R0. TCCON data archive, hosted by
CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set],
https://doi.org/10.14291/tccon.ggg2020.paris01.R0, 2022.

885   Turner, A. J., Jacob, D. J., Wecht, K. J., Maasakkers, J. D., Lundgren, E., Andrews, A. E., Biraud, S. C., Boesch, H.,
Bowman, K. W., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Kuze, A., Notholt, J., Ohyama, H., Parker,
R., Payne, V. H., Sussmann, R., Sweeney, C., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: Estimating
global and North American methane emissions with high spatial resolution using GOSAT satellite data, Atmos. Chem.
Phys., 15, 7049–7069, https://doi.org/10.5194/acp-15-7049-2015, 2015.

890
USGS: Global Land Cover Characteristics Data Base Version 2.0, https://lta.cr.usgs.gov/glcc/globdoc2_0 (last access: 27 Jan
2023), 2012.

Wang, C., Wu, Q., Weimer, M., and Zhu, E.: FLAML: A Fast and Lightweight AutoML Library, in: Proceedings of

895   Machine Learning and Systems, 434–447, 2021.

Warneke, T., Messerschmidt, J., Notholt, J., Weinzierl, C., Deutscher, N., Petri, C., Grupe, P., Vuillemin, C., Truong, F., Schmidt, M., Ramonet, M., and Parmentier, E.: TCCON data from Orléans, France, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set],

900   https://doi.org/10.14291/tccon.ggg2020.orleans01.R0, 2022.

Wennberg, P. O., Roehl, C., Blavier, J.-F., Wunch, D., Landeros, J., and Allen, N.: TCCON data from Jet Propulsion Laboratory, Pasadena, California, USA, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.jpl02.R0, 2022a.

905

Wennberg, P. O., Roehl, C., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R., Keppel-Aleks, G., Allen, N., and Ayers, J.: TCCON data from Park Falls, Wisconsin, USA, Release GGG2020R1. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.parkfalls01.R1, 2022b.

910

Wennberg, P. O., Wunch, D., Roehl, C., Blavier, J.-F., Toon, G. C., and Allen, N.: TCCON data from California Institute of Technology, Pasadena, California, USA, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.pasadena01.R0, 2022c.

915   Wennberg, P. O., Wunch, D., Roehl, C., Blavier, J.-F., Toon, G. C., Allen, N., Dowell, P., Teske, K., Martin, C., and Martin, J.: TCCON data from Lamont, Oklahoma, USA, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.lamont01.R0, 2022d.

920   Western, L. M., Ramsden, A. E., Ganesan, A. L., Boesch, H., Parker, R. J., Scarpelli, T. R., Tunnicliffe, R. L., and Rigby, M.: Estimates of North African Methane Emissions from 2010 to 2017 Using GOSAT Observations, Environ. Sci. Technol. Lett., 8, 626–632, https://doi.org/10.1021/acs.estlett.1c00327, 2021.

Worden, J. R., Cusworth, D. H., Qu, Z., Yin, Y., Zhang, Y., Bloom, A. A., Ma, S., Byrne, B. K., Scarpelli, T., Maasakkers, J.
925   D., Crisp, D., Duren, R., and Jacob, D. J.: The 2019 methane budget and uncertainties at 1° resolution and each country through Bayesian integration Of GOSAT total column methane data and a priori inventory estimates, Atmos. Chem. Phys., 22, 6811–6841, https://doi.org/10.5194/acp-22-6811-2022, 2022.

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network, Phil. Trans. R. Soc. A., 369, 2087–2112, https://doi.org/10.1098/rsta.2010.0240, 2011.

Wunch, D., Mendonca, J., Colebatch, O., Allen, N., Blavier, J.-F. L., Kunz, K., Roche, S., Hedelius, J., Neufeld, G., Springett, S., Worthy, D., Kessler, R., and Strong, K.: TCCON data from East Trout Lake, Canada, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. [data set], https://doi.org/10.14291/tccon.ggg2020.easttroutlake01.R0, 2022.

Yang, Y., Zhou, M., Langerock, B., Sha, M. K., Hermans, C., Wang, T., Ji, D., Vigouroux, C., Kumps, N., Wang, G., De Mazière, M., and Wang, P.: New ground-based Fourier-transform near-infrared solar absorption measurements of XCO2, XCH4 and XCO at Xianghe, China, Earth Syst. Sci. Data, 12, 1679–1696, https://doi.org/10.5194/essd-12-1679-2020, 2020.

Zhang, G. and Lu, Y.: Bias-corrected random forests in regression, J. Appl. Stat., 39, 151–160, https://doi.org/10.1080/02664763.2011.578621, 2012.

Zhou, M., Wang, P., Kumps, N., Hermans, C., and Nan, W.: TCCON data from Xianghe China, Release GGG2020.R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A., https://doi.org/10.14291/tccon.ggg2020.xianghe01.R0, 2022.

Zhu, L., Jacob, D. J., Keutsch, F. N., Mickley, L. J., Scheffe, R., Strum, M., González Abad, G., Chance, K., Yang, K., Rappenglück, B., Millet, D. B., Baasandorj, M., Jaeglé, L., and Shah, V.: Formaldehyde (HCHO) As a Hazardous Air Pollutant: Mapping Surface Air Concentrations from Satellite and Inferring Cancer Risks in the United States, Environ. Sci. Technol., 51, 5650–5657, https://doi.org/10.1021/acs.est.7b01356, 2017.