



RESEARCH ARTICLE

10.1029/2025JD043489

Special Collection:

The Arctic: An AGU Joint
Special Collection

Key Points:

- Arctic atmospheric carbon monoxide (CO) and methane (CH₄) simulations by multiple models are assessed by comparing to column averages retrieved at three Total Carbon Column Observing Network stations
- Model results for CO varied by location and by model, with four of the eight models agreeing with measurements within the uncertainty range
- Results for CH₄ were consistent by location, but varied by model, with biases beyond the relatively small measurement uncertainty range

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

E. McGee,
e.mcgee@mail.utoronto.ca

Citation:

McGee, E., Strong, K., Walker, K. A., Whaley, C., Kivi, R., Notholt, J., et al. (2025). Evaluation of modeled carbon monoxide and methane columns in the high Arctic using TCCON measurements. *Journal of Geophysical Research: Atmospheres*, 130, e2025JD043489. <https://doi.org/10.1029/2025JD043489>

Received 23 JAN 2025

Accepted 4 OCT 2025

Evaluation of Modeled Carbon Monoxide and Methane Columns in the High Arctic Using TCCON Measurements

Erin McGee¹, Kimberly Strong¹, Kaley A. Walker¹, Cynthia Whaley², Rigel Kivi³, Justus Notholt⁴, Gabriel Cassidy¹, Stephen Beagley⁵, Rong-You Chien⁶, Srdan Dobricic⁷, Xinyi Dong⁶, Joshua S. Fu^{6,8}, Michael Gauss⁹, Wanmin Gong⁵, Joakim Langner¹⁰, Kathy S. Law¹¹, Louis Marelle¹¹, Tatsuo Onishi¹¹, Naga Oshima¹², David A. Plummer², Luca Pozzoli^{7,13}, Jean-Christophe Raut¹¹, and Svetlana Tsyro⁹
¹Department of Physics, University of Toronto, Toronto, ON, Canada, ²Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Victoria, BC, Canada, ³Space and Earth Observation Centre, Finnish Meteorological Institute, Sodankylä, Finland, ⁴Institute of Environmental Physics, University of Bremen, Bremen, Germany, ⁵Air Quality Research Division, Environment and Climate Change Canada, Toronto, ON, Canada, ⁶Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, USA, ⁷Joint Research Centre, European Commission, Ispra, Italy, ⁸Computational Earth Science Group, Oak Ridge National Laboratory, Oak Ridge, TN, USA, ⁹Division for Climate Modelling and Air Pollution, Norwegian Meteorological Institute, Oslo, Norway, ¹⁰Swedish Meteorological and Hydrological Institute, Norrköping, Sweden, ¹¹UVSQ, CNRS, LATMOS, Sorbonne Université, Paris, France, ¹²Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan, ¹³FINCONS SPA, Vimerate, Italy

Abstract Methane (CH₄) and carbon monoxide (CO) are gases with important climate impacts as direct and indirect greenhouse gases, respectively. Methane has a warming potential 28 times that of carbon dioxide on a 100-year timescale, and carbon monoxide is a precursor to ozone in the troposphere. Modeling trace gas concentrations in the Arctic atmosphere can be challenging due to Arctic conditions and sensitivity to long-range transport, and comparing model outputs to remote sensing measurements is essential for ensuring that models are performing well. Ground-based Arctic measurements are spatially sparse, so it is important to make use of all such available data sets. In this study, we assess eight atmospheric models, comparing their simulations of atmospheric CO and CH₄ column-averaged dry-air mole fractions for 2014 and 2015 with ground-based retrievals of these species at three Arctic stations in the Total Carbon Column Observing Network (TCCON). The multi-model mean had mean biases ($\pm$ one standard deviation of the mean) of $-5.4\% \pm 8\%$ at Eureka, Canada, $-6.5\% \pm 8\%$ at Ny-Ålesund, Norway, and $-11\% \pm 7\%$ at Sodankylä, Finland for CO, and mean biases of $-0.25\% \pm 0.5\%$ at Eureka, $-0.90\% \pm 0.5\%$ at Ny-Ålesund, and $-1.0\% \pm 0.5\%$ at Sodankylä for CH₄. Individual model mean biases range from -33% to $+35\%$ for CO and -2.5% to $+1.9\%$ for CH₄. These results indicate that models could benefit from improvements targeting simulations of Arctic CO.

Plain Language Summary Carbon monoxide and methane are important trace gases present in our atmosphere. Both have impacts on the climate and carbon monoxide also affects air quality. Atmospheric models that are applied to climate and air quality problems provide simulations of the concentrations of these gases, but to ensure their accuracy, they must be evaluated against measurements. In this paper, we compare model simulations of carbon monoxide and methane to measurements from three Arctic stations in the Total Carbon Column Observing Network. We find that the average of our models underestimates the amount of carbon monoxide at one of the three stations, but not at the other two. We also find that the average of the models underestimates methane at two of our three sites, but by a smaller amount. The results of this study show that carbon monoxide and methane distributions across the sensitive Arctic region are not simple and that models could be further improved.

1. Introduction

The Arctic is a sensitive ecosystem that is warming at three to four times the global average rate (IPCC, 2021; Rantanen et al., 2022). Significant intervention is necessary to mitigate the drivers of anthropogenic climate change. Carbon dioxide (CO₂) is the most important anthropogenic greenhouse gas but presents unique challenges for policymakers as it has a very long atmospheric lifetime, meaning that the climate impacts of emission mitigation are most evident on long timescales. Short-Lived Climate Forcers, or SLCFs, are atmospheric species

© 2025 The Author(s).

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

with lifetimes shorter than that of CO₂, and thus the climate impacts of policy changes with regard to these species can be seen on much shorter timescales (AMAP, 2021). Carbon monoxide (CO) is chiefly of concern as an SLCF due to its role as a precursor to ozone in the troposphere (AMAP, 2021). Modeling of CO requires accurate emissions inventories, as wildfire smoke and other pollution plumes can have major impacts on total column CO in the Arctic (e.g., Fisher et al., 2010; Lutsch et al., 2020; Viatte et al., 2015). The models used in this analysis include Earth system models, climate chemistry models, and chemical transport models. Previous validation studies of these models with in situ surface measurements, ground-based measurements, and satellite data found that the models tended to overestimate summertime Arctic free tropospheric CO, but underestimate it in the spring (Flood et al., 2024; Whaley et al., 2022).

Methane (CH₄) is also a precursor to tropospheric ozone and is a greenhouse gas with a global warming potential about 28 times that of CO₂ on a 100-year time scale (Forster et al., 2021). CH₄ has a longer atmospheric lifetime than CO, and its atmospheric burden is increasing, particularly since 2006 (Nisbet et al., 2023). Previous studies have found that AMAP model simulations of CH₄ have smaller biases than those of CO (Flood et al., 2024; Whaley et al., 2022).

Models can cover spatial and temporal ranges for which it is difficult or impossible to obtain measurements. The Arctic is particularly sensitive to this issue as most long-running remote sensing measurement records of atmospheric trace gas concentrations rely on sunlight, and the high Arctic receives no sunlight for several months of the year, during polar night. In order to be confident in model predictions of the Arctic atmosphere, we must validate model simulations against measurements. The Total Carbon Column Observing Network (TCCON) is a network of high-resolution Fourier transform infrared spectrometers that was established to validate satellite measurements of greenhouse gases, but has also been used to validate model simulations (e.g., Cogan et al., 2012; Messerschmidt et al., 2013; Ostler et al., 2016).

Previous studies have compared the models evaluated in this study to satellite data (Whaley et al., 2022, 2023), in situ surface data (Whaley et al., 2022, 2023) and partial columns from ground-based FTIR stations in the Network for the Detection of Atmospheric Composition Change (NDACC) (Flood et al., 2024). TCCON data, as used in this study, are tightly quality controlled across a global network, and their accuracy and precision are high as a result. The data product, a column-averaged dry-air mole fraction (X_{gas}), is different from the data products used in the other papers in several notable and important ways. Firstly, the TCCON X_{gas} is a column average from 0 to 70 km, which is much higher than the 0–7 km partial columns used in Flood et al. (2024), and the surface measurements used in Whaley et al. (2022). The CH₄ column in particular has a significant contribution above 7 km, meaning that the inclusion of parts of the atmosphere above 7 km are important for evaluating model performance. None of the comparisons to satellite data for CO and CH₄ have included a profile that covers the entire troposphere—TES and MOPITT comparisons used 600 hPa levels, and ACE-FTS profiles do not include information below 5 km of altitude. Therefore, this is the first model-measurement comparison for this specific suite of models that includes both the troposphere and the whole stratosphere.

In this study, we aim to quantify the biases of eight atmospheric models that recently conducted simulations for the Arctic Monitoring and Assessment Programme (AMAP) SLCF Assessment Report (AMAP, 2021). The model outputs were compared to retrieved quantities of CO and CH₄ at three Arctic TCCON stations. This paper is organized as follows: in Section 2, we describe the models and ground-based data used and define the products used for comparison. In Section 3, we present and discuss the results of our analysis, including monthly means, correlations, and biases. Conclusions are summarized in Section 4.

2. Methods

2.1. Models

AMAP is a working group of the Arctic Council that is using a suite of models to inform policy recommendations for SLCFs. The eight atmospheric models used for this study are CESM, CMAM, EMEP-MSC-W, GEM-MACH, GEOS-Chem, MATCH, MRI-ESM2, and WRF-Chem, all of which provided CO (and, for CMAM, GEOS-Chem, and MRI-ESM2, CH₄) output for the 2021 AMAP SLCF Assessment Report (AMAP, 2021). These models were chosen from the set of models that participated in the AMAP SLCF assessment for their simulation of species that TCCON retrieves. For each model, outputs were given as vertical volume mixing ratio profiles at 3-hourly intervals on various model-specific grids for the years 2014 and 2015, with the exception of GEM-

MACH, for which profiles were provided only for 2015. These model outputs were provided by the AMAP SLCP modeling team (CCCma, 2023), and were used in AMAP (2021) and by Whaley et al. (2022, 2023). Table 1 provides a summary of the models used in this study, including their output species and spatial grids. These models include two Earth System Models (ESMs), one Chemistry Climate Model (CCM), four Chemical Transport Models (CTMs), and one combined CCM + CTM. All models used the same anthropogenic emissions inventory, ECLIPSE v6b (Evaluating the Climate and Air Quality Impacts of Short-Lived Pollutants, Stohl et al., 2015) from the IIASA GAINS (International Institute for Applied Systems Analysis—Greenhouse gas—Air pollution Interactions and Synergies) model (Höglund-Isaksson et al., 2020). Table 2 provides the inventories for biogenic, volcanic, forest fire, and agricultural emissions used by each model, as well as the meteorology for each. All eight models used prescribed CH_4 concentrations based on box model output from Meinshausen et al. (2017) for 2014 and from Olivie et al. (2021) for 2015. The prescribed concentrations are input at the bottom model layer and the global distributions of CH_4 then evolve based on the processes in each model.

The CESM configuration used for this study is a CESM2.0 version using the Community Atmosphere Model version 6 chemistry (CAMv6-chem) coupled with the Community Land Model version 5 (CLMv5) to represent chemical reactions and biogenic emissions respectively. It uses the Modal Aerosol Model (MAM4) to track aerosol species including sulfate, black carbon, organic and secondary organic aerosols, dust, and sea salt. The runs used in this study are forced with prescribed sea surface temperatures and sea-ice concentrations from merged Reynolds/HADISST products (AMAP, 2021).

CMAM runs for this study use the specified dynamics setup, where model horizontal winds are nudged toward reanalysis data. CH_4 - NO_x chemistry is used for the troposphere, where it does not include larger VOCs but does include wet and dry deposition, interactive NO_x emissions from lightning, corrections of clear-sky photolysis rates for clouds, and N_2O_5 hydrolysis on sulfate aerosols. For the stratosphere, the chemistry includes the effects of HO_x , NO_x , Cl_x , and Br_x on O_3 , CH_4 , N_2O , and CFCs. The model was also nudged to ECMWF ERA-Interim reanalysis 6-hourly fields at all model levels below 1 hPa (AMAP, 2021). For CH_4 , the tracer is set to a single global-averaged value in the bottom two model levels but is prognostic elsewhere, being advected around and destroyed in the chemistry of the model. The prescribed value in the bottom two layers evolves over time, linearly interpolated between values given at the midpoint of each year.

EMEP-MSC-W rv4.33 is used for this study. The top of the model is set to be at 100 hPa, and 20 vertical levels are used, on eta-hybrid coordinates. EMEP-MSC-W chemistry includes coarse nitrate formation, N_2O_5 hydrolysis on aerosols, as well as other gas-aerosol loss processes, but this is confined to the troposphere; this run did not have stratospheric chemistry (Whaley et al., 2022). ECLIPSEv6b is used for $\text{PM}_{2.5}$ and coarse PM emissions as well as agricultural waste burning, but not black carbon and organic carbon—these are derived from the $\text{PM}_{2.5}$ and coarse PM emissions (AMAP, 2021).

GEM-MACH has 84 vertical levels in a hybrid coordinate between the surface and 0.1 hPa; the model output submitted to AMAP and used for this study consists of the lowest 35 of these levels, which extend roughly between the surface and 200 hPa. GEM-MACH includes a chemistry module that includes gas-phase, aqueous-phase, and heterogeneous chemistry and aerosol processes. Aerosols included SO_4 , NO_3 , NH_4 , primary and secondary organic aerosols, organic carbon, crustal material, and sea salt. It used a limited-area domain to cover the Arctic at high spatial resolution. The emissions inventories for GEM-MACH are unique among the models used for this paper—they are a combination of the 2015 Canadian national Air Pollution Emission Inventories and global ECLIPSE v6b 2015 baseline emissions. The North American wildfire inventories used the Canadian Forest Fire Emission Prediction System (CFFEPS) to process fire hotspot data from the MODIS, AVHRR, and VIIRS satellite instruments. Biogenic emissions for GEM-MACH were calculated online (AMAP, 2021).

GEOS-Chem includes wet and dry deposition schemes, with the wet deposition scheme including both water-soluble aerosols and gases. It relies on local meteorology for dust aerosol, lightning NO_x , biogenic volatile organic carbons (VOCs), soil NO_x , and sea salt aerosols. In the case of this study, it is GEOS-FP meteorological data that is assimilated, though GEOS-Chem generally has the ability to assimilate other meteorologies. Tropospheric gas chemistry for GEOS-Chem includes HO_x - NO_x -VOC-ozone-halogen-aerosol mechanisms, and the stratospheric chemistry includes ozone (AMAP, 2021). For methane, GEOS-Chem includes a north-south gradient in the near-surface prescribed concentrations. The monthly production and loss rates are based on the NASA Global Modeling Initiative, but CH_4 seasonal cycles and chemistry are not included in this run.

Table 1 Summary of AMAP Model Outputs Used in This Study								
Model	Type of model	Primary reference	Output species	Latitude range provided	Latitude grid spacing	Longitude grid spacing	Pressure altitude range (to center of highest grid cell)	Altitude levels per vertical profile
CESM	Earth System Model	Danabasoglu et al. (2020), Liu et al. (2016)	CO	0° to 90°N	~1.9°	2.5°	Surface to 3.64 hPa	32
CMAM	Chemistry Climate Model	Jonsson et al. (2004), Scinocca et al. (2008)	CO, CH ₄	87.159°S to 87.159°N	~3.7°	3.75°	Surface to 0.0008 hPa	71
EMEP MSC-W	Chemical Transport Model	Simpson et al. (2012, 2019)	CO	0° to 90°N	0.5°	0.5°	Surface to 21 hPa	20
GEM-MACH	Online Chemical Transport Model	Gong et al. (2015), Makar et al. (2015a, b), Moran et al. (2018)	CO	32° to 89.96°N	0.14°	0.14°	Surface to 200 hPa	35
GEOS-Chem	Chemical Transport Model	Bey et al. (2001)	CO, CH ₄	90°S to 90°N	2°	2.5°	Surface to 0.01 hPa	72
MATCH	Chemical Transport Model	Robertson et al. (1999)	CO	20°N to 90°N	0.5°	1°	Surface to 66.6 hPa	38
MRI-ESM2	Earth System Model	Kawai et al. (2019), Yukimoto et al. (2019), Oshima et al. (2020)	CO, CH ₄	0° to 90°N	2.8125°	2.8125°	Surface to 0.01 hPa	80
WRF-Chem	Chemistry Climate Model + Chemical Transport Model	Marelle et al. (2017, 2018)	CO	10°N to 90°N	100 km	100 km	Surface to 52 hPa	49

Table 2
Summary of Emissions Inputs and Meteorology Used by Models in This Study

Model	Biogenic emissions	Volcanic emissions	Forest fire emissions	Agricultural waste burning	Meteorology
CESM2.0	MEGANv2.1	CMIP6	CMIP6	ECLIPSEv6b	Fully interactive with prescribed sea surface temperature and sea ice concentration
CMAM	None	None	CMIP6	ECLIPSEv6b	Nudged to ERA-Interim reanalysis
EMEP MSC-W	EMEP scheme (Simpson et al., 2012)	Degassing from Etna, Stromboli (as reported to EMEP), Eyjafjallajökull (2010), Grimsvotn (2011), Holuhraun (2014, 2015)	FINN (Wiedinmyer et al., 2011)	ECLIPSEv6b	Driven by 3-hourly data from the Integrated Forecast System (IFS) at ECMWF
GEM-MACH	BEIS v3.09	None	CFEPPS (North America only)	ECLIPSEv6b outside North America	Driven by GEM Model numerical forecast, initialized every 24 hr using the Canadian Meteorological Centre's global objective analysis
GEOS-Chem	MEGANv2.1 with update from Guenther et al. (2012)	NASA/GMAO	GFEDv4.1	ECLIPSEv6b	Driven by the GEOS meteorology from the NASA data assimilation office
MATCH	MEGANv2	Climatological + Holuhraun eruption	CMIP6	ECLIPSEv6b	ERA-Interim 6-hourly reanalysis
MRI-ESM2	VOC emissions taken from Horowitz et al. (2003)	CMIP6 stratospheric aerosol data set	CMIP6	ECLIPSEv6b	Nudged to Japanese 55-year Reanalysis (JRA55)
WRF-Chem	MEGANv2.1	None	GFED	ECLIPSEv6b	Nudged to NCEP Final Analysis (FNL)

Note. Adapted from Table A7.2 in AMAP (2021).

MATCH used ECMWF ERA-Interim reanalysis data (3-hourly for precipitation, 6-hourly otherwise) for meteorological outputs. Only the lowest 38 of 60 levels are used, reaching up to about 16 km altitude in the Arctic (AMAP, 2021). Its chemistry is described by Andersson et al. (2007).

The MRI-ESM2 simulation horizontal wind fields are nudged toward the 6-hourly Japanese 55-year Reanalysis (JRA-55). CMIP6 provided both monthly biomass burning emissions and major volcanic aerosols (the latter from the stratospheric aerosol data set). The aerosol component of this model includes black carbon, organic carbon, sea salt, dust, and aerosol precursor gases (SO_2 and dimethyl sulfide) (AMAP, 2021). For CH_4 , MRI-ESM2 provides prescribed global annual averages for the surface to 900 hPa levels. It includes a relatively complete description of stratospheric ozone chemistry including HO_x , NO_x , Cl_x , and Br_x ; its stratospheric chemistry also includes N_2O and CFCs (Whaley et al., 2022). The model then includes CH_4 in advection and chemical reactions such that CH_4 model-specific outputs are given above these levels.

WRF-Chem version 3.8.1 is used for this study. Pressure levels follow the terrain, the model top is set to be 50 hPa, and stratospheric concentrations are taken from climatologies. This model run uses Morrison double-moment scheme microphysics, and short- and long-wave radiative effects from the Rapid Radiative Transfer Model for General Circulation Models. Temperature, relative humidity, and winds are all nudged toward global NCEP Final Analysis meteorology, which is also used to set boundary and initial conditions (AMAP, 2021). More detail about the models can be found in the references listed in Table 1, in the AMAP 2021 Short-Lived Climate Forcer Report (AMAP, 2021), and in Whaley et al. (2022).

2.2. The Total Carbon Column Observing Network

TCCON is a global network of high-resolution Bruker IFS 125HR Fourier-transform infrared (FTIR) spectrometers. TCCON has very good site-to-site consistency due to the use of nearly identical instruments at all sites and the use of the same retrieval algorithm (Wunch et al., 2011). All column-averaged dry-air mole fractions (X_{gas}) are retrieved from near-infrared solar absorption spectra using the GGG2020 version of the GGG/GFIR algorithm (publicly available at <https://tccodata.org/>) (Laughner et al., 2024). This algorithm uses a scaling profile retrieval that assumes that the shape of the profile of a given species is the same as that of the prior profile and then scales those profiles (Wunch et al., 2011). While not all species retrieved by TCCON are SLCFs, TCCON retrieves both CO and CH_4 , the two species evaluated in this work.

The TCCON data used in this study have a lot in common with the retrieved partial profiles in Flood et al. (2024) but there are several notable differences. TCCON data are retrieved using a profile scaling retrieval that includes profile information up to 70 km in altitude, much higher than the 0–7 km partial columns compared in Flood et al. (2024) and so contain more information about the stratosphere and above. TCCON retrieved X_{gas} are also more precise and accurate than the retrieved quantities of other ground-based networks (Wunch et al., 2011). Scaling retrievals are less sensitive than a full profile retrieval to certain errors that can arise due to the shapes of calculated spectral lines. However, if the profile shape is not accurately predicted a priori, there is no opportunity to correct this. In order to minimize this error, the prior profiles used in the GGG2020 retrieval are generated every 3 hours. One particularity of the GGG2020 retrieval algorithm is that in order to most accurately capture the shape of the CO profile in the stratosphere, outputs from CMAM30 (a version of CMAM with specified dynamics) are used for the prior for altitudes with pressures <10 hPa, and blended with the GEOS FP-IT prior used for the rest of the priors between 30 and 10 hPa (Kolonjari et al., 2018; Laughner et al., 2023). Notably, this specified-dynamics setup of CMAM is also used for the CMAM model outputs analyzed in this study.

Data from the three Arctic TCCON sites, summarized in Table 3 with locations shown in Figure 1, were used in this study. The Eureka TCCON instrument is housed in the Polar Environment Atmospheric Research Laboratory (PEARL) Ridge Lab, which is located in a semi-arid environment on a ridge 610 m above sea level on the shores of Slide Fiord (Fogal et al., 2013). The FTIR instrument at Ny-Ålesund is placed inside the dedicated NDACC building at 23 m above sea level. Ny-Ålesund is on the west coast of Spitsbergen at the Kongsfjord; the local topography is flat, the topographic context is within a valley, and the surface cover is barren or sparsely vegetated. The TCCON station at Sodankylä is located at an altitude of 188 m above sea level, surrounded by Boreal Forest (Kivi & Heikkinen, 2016).

Table 3
Arctic TCCON Sites Included in This Study

TCCON site	Latitude	Longitude	Altitude (m.a.s.l.)	Data collection period of year (approximate)	Color code for plots	Reference
Eureka (Nunavut, Canada)	80.05°N	86.42°W	610	March–September		Strong et al. (2022)
Ny-Ålesund (Svalbard, Norway)	78.92°N	11.92°E	27	March–September		Buschmann et al. (2022)
Sodankylä (Finland)	67.37°N	26.63°E	188	February–October		Kivi et al. (2022)



Figure 1. A map showing the Arctic TCCON stations used in this analysis.

2.3. Comparisons

The 3-hourly model vertical profiles in the grid cells that included the three TCCON stations were extracted from the model files. These vertical profiles were not sampled along the line of sight of the instruments but selected from the grid cells that contained the TCCON stations. Given the model grid cell size and altitude range, we do not expect there to be significant errors from not accounting for the line of sight in the comparisons. TCCON priors and averaging kernels are produced on a 3-hourly basis that fortunately match with the model timestamps, except for those of GEM-MACH. GEM-MACH outputs are at 02:00, 05:00, 08:00..., etc., instead of 00:00, 03:00, 06:00, etc., the latter of which was the case for every other model. For GEM-MACH, smoothing was done with the TCCON averaging kernels and priors that were closest in time (i.e., 1 hour later). Integration of the model data to create an Xgas data product was done in accordance with the recommended procedure for model comparison with TCCON data (Laughner, 2023) such that:

$$X_{\text{gas}} = \mathbf{h}^T \mathbf{x}_{\text{prior}} + \mathbf{d}^T (\mathbf{x}_{\text{model}} - \mathbf{x}_{\text{prior}}). \quad (1)$$

here, $\mathbf{x}_{\text{prior}}$ is the TCCON prior at the time of the model vertical profile $\mathbf{x}_{\text{model}}$, $\mathbf{d}$ is the element-wise product of the TCCON averaging kernel $\mathbf{a}$, and pressure weighting function $\mathbf{h}$.

$\mathbf{h}$ is defined as follows:

$$h_j = \frac{\Delta p_j}{g_j \cdot m_{\text{air}}^{\text{dry}} \cdot \left(1 + f_{\text{H}_2\text{O},j}^{\text{dry}} \cdot \frac{m_{\text{H}_2\text{O}}^{\text{dry}}}{m_{\text{air}}^{\text{dry}}}\right)} \quad (2)$$

such that Δp_j is the pressure thickness for level j , calculated as:

$$\Delta p_j = \frac{p_j + (p_{j+1} - p_j)}{\log \frac{p_{j+1}}{p_j}} + \frac{p_j - (p_j - p_{j-1})}{\log \frac{p_j}{p_{j-1}}} \quad (3)$$

setting the first term to 0 where no $j + 1$ term exists and setting the second term to 0 for the first pressure level ($j = 1$). In Equation 2, g is the TCCON gravity profile, $m_{\text{air}}^{\text{dry}}$ is the molar mass of dry air (0.0289644 kg/mol), $f_{\text{H}_2\text{O}}^{\text{dry}}$ is the dry water profile, defined to be:

$$\left(\frac{1}{f_{\text{H}_2\text{O}}^{\text{wet}}} - 1\right)^{-1} \quad (4)$$

where $f_{\text{H}_2\text{O}}^{\text{wet}}$ is the wet water profile (a TCCON product), and $m_{\text{H}_2\text{O}}^{\text{dry}}$ is the molar mass of water (0.01801528 kg/mol). The TCCON prior data product is a wet mole fraction, which was dried using:

$$f_{\text{gas}}^{\text{dry}} = \frac{f_{\text{gas}}^{\text{wet}}}{1 - f_{\text{H}_2\text{O}}^{\text{dry}}} \quad (5)$$

Model profiles are dry mole fractions, and therefore did not need to be dried.

The vertical coordinate for TCCON averaging kernels and priors is altitude, with pressures provided at each level. For the models, no altitudes were provided—pressures were given as the vertical coordinate and the pressure could vary slightly from profile to profile, depending on how levels were defined in each model. All calculations were performed using the TCCON averaging kernel pressures as the index (which extend from the surface to 0.046 hPa), interpolating model profiles and TCCON priors onto that pressure grid in order to ensure all vectors used a common vertical coordinate. Where the vertical range of a model profile did not cover that of the TCCON averaging kernel (i.e., the TCCON averaging kernel either started lower in the atmosphere, reached higher, or both), TCCON a priori values were used to “fill in the blanks.” For cases where differences and percent differences between models and TCCON data were calculated, a 3-hourly mean of the TCCON data, centered on the model time, was calculated at the time of the model output. Percent differences were calculated as:

$$\frac{X_{\text{gas, model}} - X_{\text{gas, TCCON}}}{X_{\text{gas, TCCON}}} \times 100\%, \quad (6)$$

such that a positive percent difference indicates that a model is biased high relative to TCCON, and vice versa. Mean percent differences were calculated for each model at each site, and the standard deviations of these means are used as error bars.

For correlations, we plot TCCON on the x -axis and models on the y -axis. Correlation coefficients (R^2 values) are calculated to assess the consistency of the model compared to TCCON measurements, that is, the likelihood of the model predicting a proportionally higher X_{gas} value when TCCON measures a higher X_{gas} value. Monthly means for each model at each site were also calculated, such that the monthly mean for a model is the average of all predictions that have a corresponding TCCON prior and averaging kernel for that month and year. Monthly means were split by year (e.g., a mean is calculated for both June 2014 and June 2015, not June in general) as this better accounts for conditions that may change from year to year, such as wildfires, and because GEM-MACH outputs were only provided for 2015.

Root mean squared error (RMSE) was calculated as follows, where N is the number of model predictions used:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_{\text{gas, model}_i} - X_{\text{gas, TCCON}_i})^2}. \quad (7)$$

A multi-model mean (MMM) was calculated by taking the mean of all model outputs at a specific timestamp. In order to account for the time offset of GEM-MACH, for this calculation only, GEM-MACH outputs were interpolated onto the timestamps that all other models used for 2015. This time series was then compared to TCCON as above. In the case of the monthly mean of the multi-model mean, the average of all model monthly means was used, as in Flood et al. (2024). This was done so that the standard deviation of this mean would better reflect the variation in the models.

3. Results and Discussion

3.1. Carbon Monoxide

Figure 2 shows the monthly mean column-averaged dry-air mole fraction of carbon monoxide (X_{CO}) for each model simulation and the TCCON FTIR measurements at each station. The complete 3-hourly data sets are plotted in Figure S1 in Supporting Information S1. Figure 2 shows that most models replicate the seasonal cycle that is observed in the FTIR X_{CO} product, but some models exhibit varying bias over the course of the year. MRI-

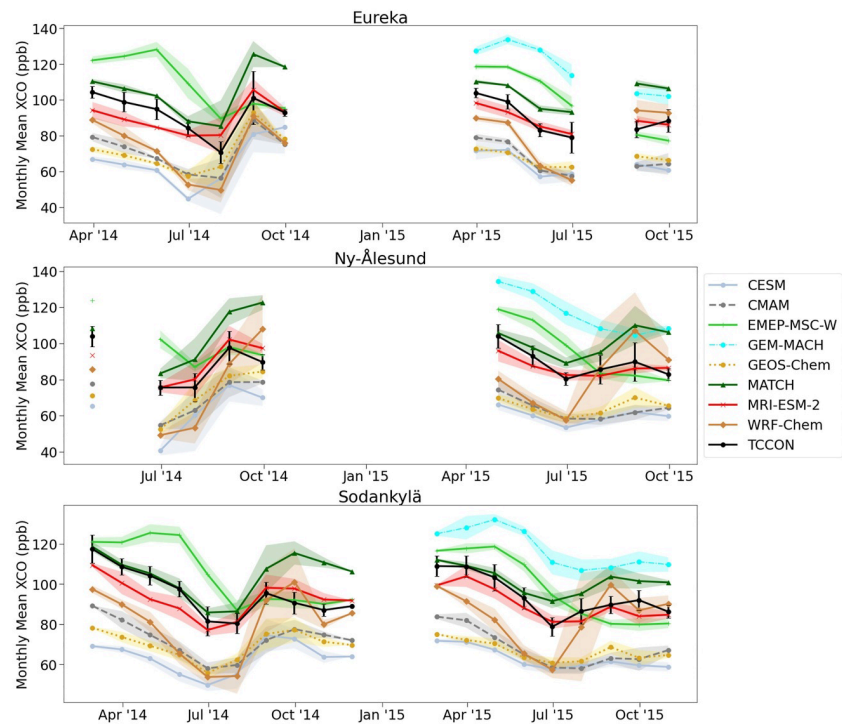


Figure 2. Monthly mean XCO from the model simulations and the TCCON measurements for the three Arctic sites. Error bars (for TCCON) and shaded areas (for models) represent the standard deviations for each monthly mean.

ESM2 generally performs quite well, while MATCH performs very well in the spring, tending toward a slight high bias in the fall. Three models—CESM, CMAM, and GEOS-Chem—are consistently biased low but follow the seasonal pattern observed in the TCCON data well. GEM-MACH is consistently biased high, with this bias being slightly stronger in the spring compared to the fall at Eureka and Ny-Ålesund. EMEP-MSC-W performs similarly to GEM-MACH, with a high positive bias in the spring and a lower bias in the fall. WRF-Chem, by contrast, is biased low in the spring but high in the fall. The multi-model mean, shown in Figure 3, performs well when compared to TCCON values, especially at the two most northerly sites, Eureka and Ny-Ålesund. Other analyses have found that model means tend to be biased low in the spring and high in late summer for CO₂ (AMAP, 2021; Flood et al., 2024; Whaley et al., 2022); we do not see this effect to a statistically significant degree, although the multi-model mean is lower in the spring compared to the late summer. At Ny-Ålesund, the multi-model mean is biased low in May 2014, but we lack data for April and June 2014, so it is difficult to determine whether this is representative of the rest of the early season. However, the bias compared to TCCON is consistent over the spring and fall of 2015, where it is insignificant.

The model XCO is plotted against TCCON XCO in Figure 4 for all eight models and three sites. Linear fits and correlation coefficients (R^2 values) show a large degree of variation both between models and between the same model at different sites. R^2 values tend to be highest at Sodankylä for most models, with the exception of CMAM and MATCH, and WRF-Chem, for which the largest R^2 values are at Eureka. Due to the shorter polar night at Sodankylä, there are more observational data points at this station compared to Eureka and Ny-Ålesund, which may have an impact on the quality of the fit. The models with the strongest correlations (largest R^2 values) are CMAM at Eureka and Sodankylä, with R^2 values of 0.61 for Eureka and 0.56 for Sodankylä, and MRI-ESM2 at Ny-Ålesund, with an R^2 value of 0.53. EMEP-MSC-W and WRF-Chem have slopes closest to one, with values of 0.72 and 0.82, respectively, for Eureka, 0.70 and 0.86 for Ny-Ålesund, and 1.00 and 0.80 for Sodankylä. A summary of the slope, y-intercept, R^2 values, as well as mean percent difference for each model, and RMSE (both discussed below), can be found Table S1 in Supporting Information S1.

Figure 5 presents the mean percent differences in XCO between each model and TCCON for the three stations. CESM, CMAM, and GEOS-Chem are significantly biased low, while EMEP-MSC-W, MRI-ESM2, and WRF-

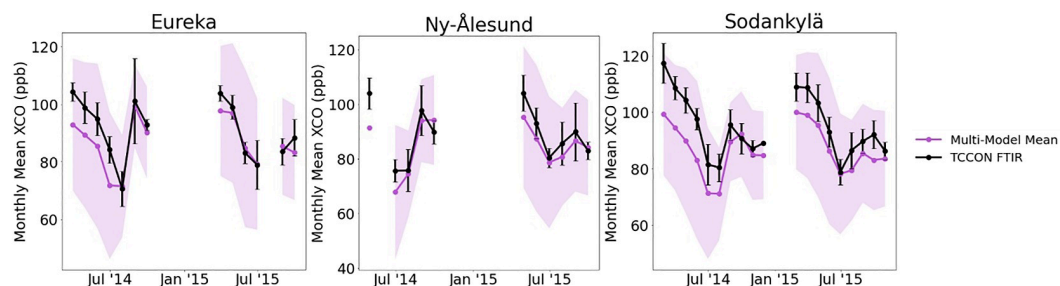


Figure 3. The multi-model mean of the monthly mean XCO from the model simulations and the TCCON measurements for the three Arctic sites. The TCCON error bars represent the standard deviation for each monthly mean (same as in Figure 2), while the multi-model mean shaded areas are the standard deviations of the set of model monthly means.

Chem agree with TCCON data within their standard deviations (although this can obscure seasonal biases that vary in sign, such as with WRF-Chem). MATCH is biased high at Eureka but agrees within error bars for Ny-Ålesund and Sodankylä, and GEM-MACH is biased high for all three sites. The error bars here also offer an indication of the consistency of the model, for example, while WRF-Chem agrees with TCCON measurements on average within uncertainty, it has large standard deviations and has a larger variation in bias by season. Models with a low mean bias but a high RMSE indicate that the model was biased high for some of the year and low for another portion of it. Overall, the mean percent differences are relatively large, with biases around 20% not being uncommon. Other analyses have found that surface and mid-tropospheric CO are underestimated at high latitudes by models including the ones used in our analysis (AMAP, 2021; Flood et al., 2024; Whaley et al., 2022, 2023). The low bias of most models for CO may be due to underestimating wintertime CO emissions (Jiang et al., 2015; Kasibhatla et al., 2002; Pétron et al., 2002), or an overestimation of OH, leading to underestimating CO lifetime (Miyazaki et al., 2012; Monks et al., 2015; Quennehen et al., 2016).

No one type of model outperforms the others—CMAM and MRI-ESM2 are the two models with the highest R^2 values, and CMAM is a Chemistry Climate Model while MRI-ESM2 is an Earth System Model. MRI-ESM2 also has the lowest absolute mean bias, but EMEP-MSC-W and MATCH are the next lowest, and EMEP-MSC-W and MATCH are Chemical Transport Models. CESM, EMEP-MSC-W, GEM-MACH, GEOS-Chem, MATCH, MRI-ESM2, and WRF-Chem all have detailed tropospheric HO_x - NO_x -hydrocarbon O_3 chemistry with VOCs and secondary aerosols (Whaley et al., 2022); CMAM has a more limited treatment of tropospheric chemistry (CMAM chemistry is limited to CH_4 - NO_x - O_3 without VOCs). CMAM and MRI-ESM2 both have a full simulated stratosphere and are the only models to include a relatively complete description of stratospheric chemistry (Whaley et al., 2022). Since these models have the largest R^2 values, this appears to be important for accurately reproducing CO variations. As mentioned in Section 2.2, outputs from CMAM are used for the TCCON CO prior at altitudes with pressures <10 hPa, and blended with the GEOS FP-IT prior between 30 hPa and 10 hPa. This is fairly high in the atmosphere, especially for CO, which is primarily tropospheric; the amount of CO in the column at altitudes above 10 hPa is very small compared to the total column of CO. While it is possible that this commonality between the TCCON priors and CMAM contributes to CMAM's higher correlation coefficients, it seems unlikely to us that this is a major factor in the agreement.

GEM-MACH may be biased high in part due to its use of MOZART-4 for the lateral boundary condition (Gong et al., 2018); MOZART-4 CO is high compared to MOPITT in the southern hemisphere and some more southern parts of the northern hemisphere (Emmons et al., 2010). It is difficult to determine why MATCH may be biased high at Eureka—it is the only model to use MEGANv2 (although many other models use MEGANv2.1) for biogenic emissions (Table 2), and its chemistry is unique as well (Andersson et al., 2007)—either of these could be contributing to the high bias. All models have a more positive bias at Eureka and a more negative bias at Sodankylä regardless of whether the model is overall biased high or low, though these inter-site differences are not statistically significant. Short timescale effects such as wildfire plumes can be seen in Figure S1 in Supporting Information S1; each model handles fire plume transport differently and many have varying atmospheric chemistry schemes. It is therefore difficult to determine which aspect of model treatment of fire plumes is driving the inter-model differences.

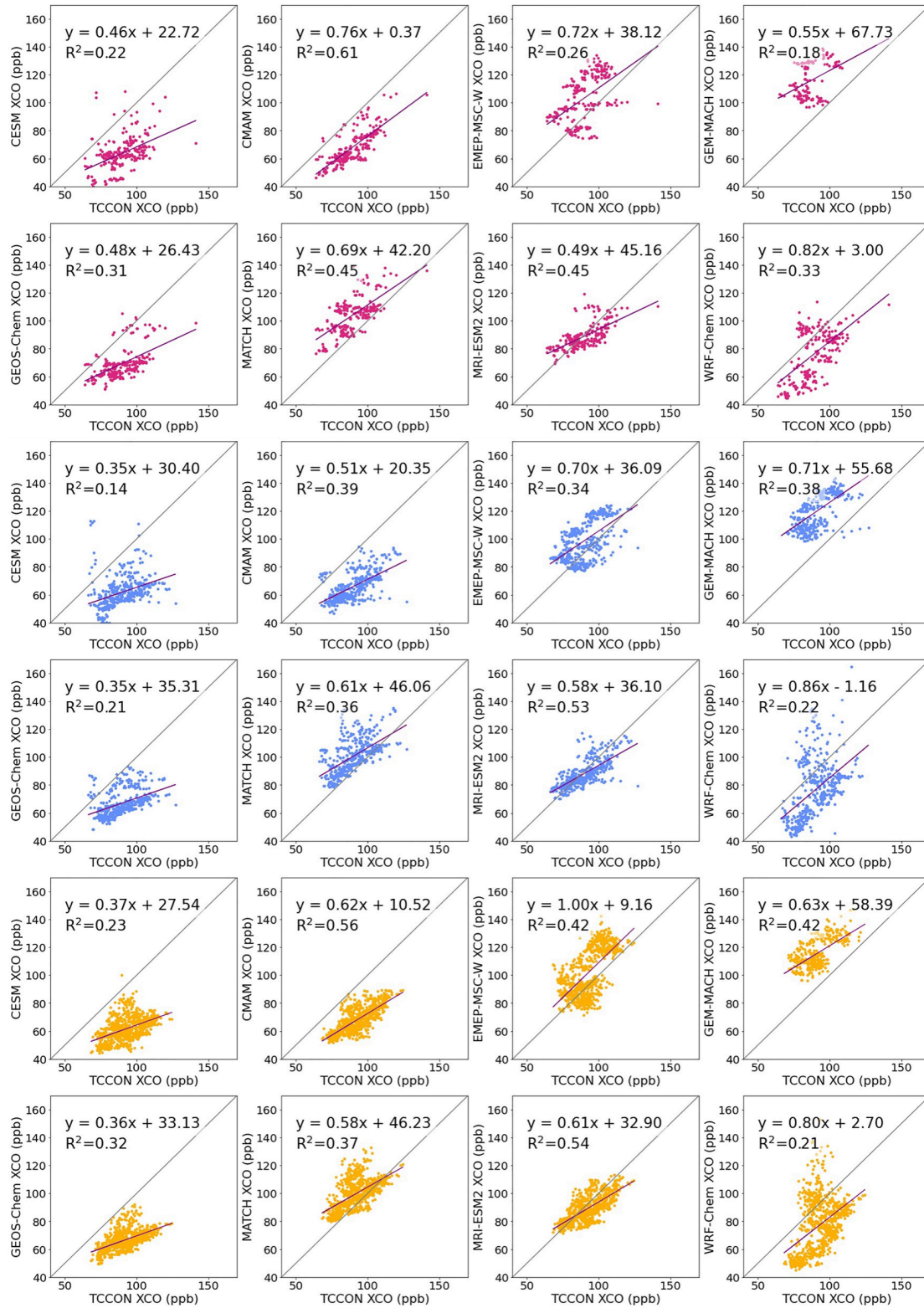


Figure 4. Correlation plots of model versus TCCON XCO for all 3-hourly points. The 1:1 correlation line (model = TCCON) is shown in light gray, and lines of best fit obtained by linear regression are shown in magenta. The equations of the lines of best fit and the associated R^2 values are included in the panels. Pink dots indicate Eureka, blue dots indicate Ny-Ålesund, and yellow dots indicate Sodankylä.

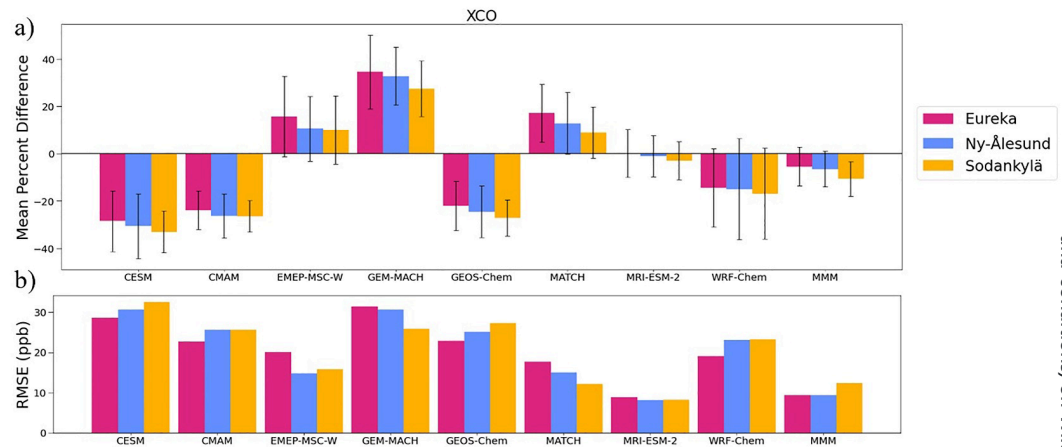


Figure 5. (a) Mean percent differences in XCO between each model and TCCON data set (calculated using Equation 6) for all 3-hourly points at each site. Error bars represent the standard deviation of the set of percent differences used for the mean. MMM indicates the multi-model mean. (b) Root Mean Squared Error (RMSE) (ppb) by model and site.

3.2. Methane

Figure 6 shows the monthly mean column-averaged dry-air mole fractions of methane (XCH_4) for three AMIP models and the TCCON FTIR measurements at each station, with the 3-hourly data sets plotted in Figure S2 Supporting Information S1. Combined, the three panels in Figure 6 show that one of the models, GEOS-Chem, overestimates CH_4 at all three sites, and the other two models, CMAM and MRI-ESM2, both underestimate CH_4 at all three sites. CMAM and MRI-ESM2 perform similarly, and both reproduce the variation seen in the TCCON FTIR data. None of the three models show the same degree of increase in CH_4 from spring to fall as seen in the TCCON data. This can be seen more clearly in the multi-model mean (Figure 7), where we see that while the multi-model monthly mean agrees with TCCON data within uncertainty for all months, the absolute difference between TCCON monthly means and the multi-model monthly mean is larger in the fall than in the spring at all three sites. From the CH_4 correlation plots, shown in Figure 8, we find that CMAM and MRI-ESM2 have R^2 values that are closer to 1 than for GEOS-Chem. GEOS-Chem also shows a very limited range of predicted CH_4 (much smaller slope) compared to the other two models considered.

Previous studies validating a suite of models, including the ones used for this study, using satellite and in situ measurements found that CH_4 from the models were biased slightly low at high latitudes, but that models failed to reproduce the variability seen in surface-level measurements near CH_4 sources (Whaley et al., 2022). Our results generally agree with Whaley et al. (2022), with CMAM and MRI-ESM2 consistently biased low compared to TCCON observations and GEOS-Chem being biased high. The TCCON column-averaged measurements may be more representative of the finite-resolution models than in situ surface point measurements, but the lack of vertical profile information from the measurements limits the type of comparisons that can be made. The low bias of CMAM and MRI-ESM2 CH_4 may be due to a missing or inaccurate north-south gradient (which tends to cause a low bias in the Arctic); GEOS-Chem does use a north-south gradient, but this may also increase the variability of the model prediction (Flood et al., 2024; Whaley et al., 2022). Flood et al. (2024) found that all three models were biased low, with CMAM and MRI-ESM2 being biased low by about 10%, and GEOS-Chem being biased low by about 7%. We find that CMAM and MRI-ESM2 are biased low by about 2%, and GEOS-Chem is biased high by about 1%, meaning that the model results compared to each other are overall consistent, but show better agreement with TCCON measurements. Given that TCCON Xgas measurements are not the same as the NDACC 0–7 km partial columns used in Flood et al. (2024), it is difficult to determine what is driving the difference in agreement. GEOS-Chem is the only model for which CH_4 does not participate in stratospheric chemistry, but instead follows a prescribed monthly loss rate. This may be limiting the dynamic range of XCH_4 for this model. The multi-model mean reproduces the seasonal variation in CH_4 well, and, when averaged monthly, agrees with TCCON monthly means within uncertainty for all sites (Figure 7).

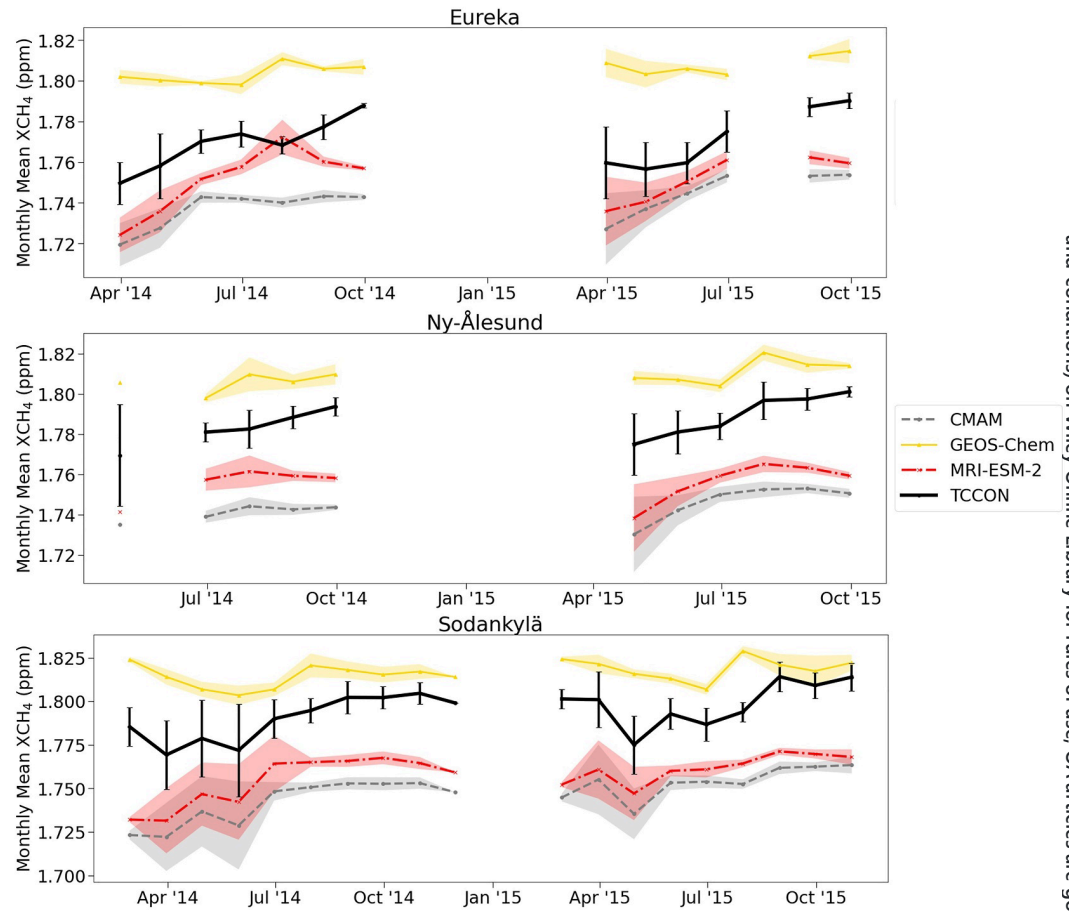


Figure 6. Monthly mean XCH_4 from the model simulations and TCCON measurements at the three Arctic sites. Error bars (for TCCON) and shaded areas (for models) represent standard deviations for each monthly mean.

Figure 9 shows the mean percent differences and RMSE between models and TCCON retrievals for XCH_4 . As with XCO , the models have smaller biases at Eureka and larger biases at Sodankylä when the biases are negative and larger biases at Eureka and smaller biases at Sodankylä when the bias is positive. None of these models alone agree with TCCON observations within uncertainty, but the average percent difference values are small—less than 2% in most cases. The multi-model mean (from the three participating models) performs better than any individual model, agreeing with TCCON within one standard deviation of the mean bias at Eureka, and having the lowest RMSE (see Figure 9).

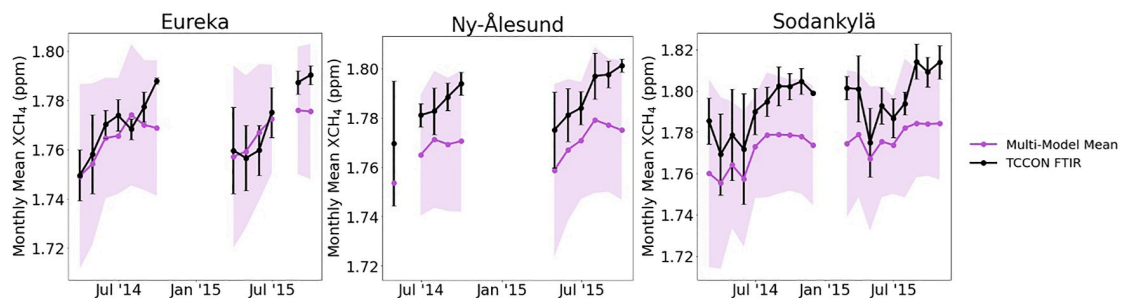


Figure 7. The multi-model mean of the monthly mean XCH_4 from the model simulations and the TCCON measurements for the three Arctic sites. The TCCON error bars represent the standard deviation for each monthly mean (same as in Figure 6), while the multi-model mean shaded areas are the standard deviations of the set of model monthly means.

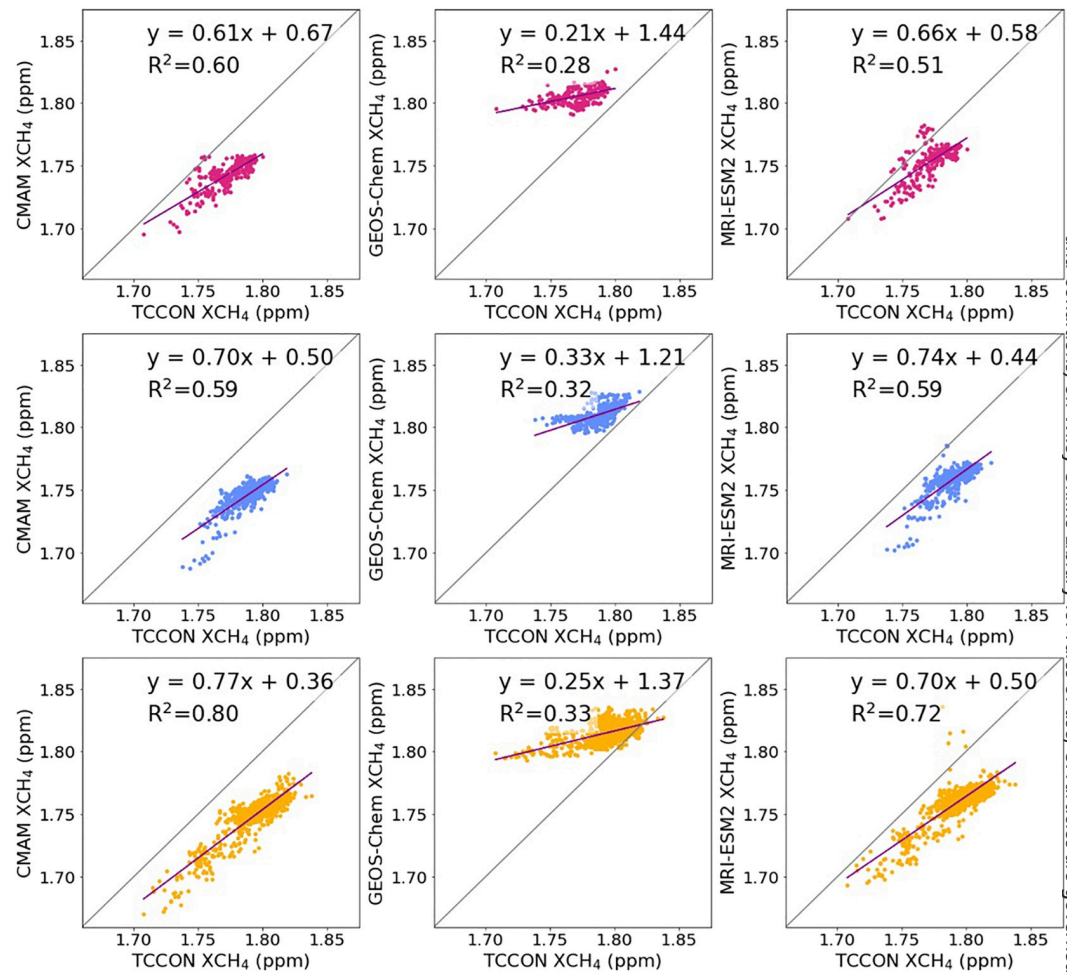


Figure 8. Correlation plots of model versus TCCON XCH₄ for all 3-hourly points. The 1:1 correlation line (model = TCCON) is shown in light gray, and lines of best fit obtained by linear regression are shown in magenta. The equations of the lines of best fit and the associated R^2 values are included in the panels. Pink dots indicate Eureka, blue dots indicate Ny-Ålesund, and yellow dots indicate Sodankylä.

4. Conclusions

This study compared simulations of CO and CH₄ by eight AMAP SLCF models to ground-based TCCON FTIR measurements at three Arctic sites. Column-averaged dry-air mole fractions (X_{gas}) were calculated from the model outputs (3-hourly vertical profiles) and compared to 3-hourly averages of TCCON X_{gas} data products retrieved using the GGG2020 version of the GGG/GFIT retrieval algorithm. TCCON was chosen as a validation measure due to its high site-to-site consistency and history of use for validation of both model and satellite data; it also complements the use of surface measurements and partial columns for validation in previous studies (Flores et al., 2024; Whaley et al., 2022).

Individual model performance varied, with MRI-ESM2 outperforming other models for CO, but no one model clearly outperforming others for CH₄. The multi-model mean agreed well with the TCCON measurements for both species at all sites, having a mean bias of $-5.4\% \pm 8\%$ at Eureka, $-6.5\% \pm 8\%$ at Ny-Ålesund, and $-11\% \pm 7\%$ at Sodankylä for XCO, and a mean bias of -0.25% at Eureka, -0.90% at Ny-Ålesund, and -1.0% at Sodankylä for XCH₄. Individual model mean biases range from -33% to $+35\%$ for XCO and -2.5% to $+1.9\%$ for XCH₄. The mean percent differences show that three out of eight models agree with TCCON within one standard deviation for XCO at all sites, with one more model agreeing within one standard deviation at two out of three sites, while no model agrees with TCCON within one standard deviation for XCH₄. Future studies would benefit from expanding both the latitude range considered (to investigate any possible latitudinal dependence of the

Acknowledgments

Funding for this work was provided by the Canadian Space Agency under the Earth System Science Data Analyses Program (Grant 21SUASABBC). Erin McGee acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC). The authors acknowledge the use of model data sets from the 2021 AMAP SLCF Assessment Report. The authors would like to thank all the people involved with the data collection and maintenance of the high-latitude TCCON FTIR instruments used in this study. The Eureka FTIR measurements were made at the Polar Environment Atmospheric Research Laboratory (PEARL) by the Canadian Network for the Detection of Atmospheric Composition Change (CANDAC), which has been supported by the Atlantic Innovation Fund/Nova Scotia Research Innovation Trust, the Canada Foundation for Innovation, the Canadian Foundation for Climate and Atmospheric Sciences, the Canadian Space Agency, Environment and Climate Change Canada (ECCC), Government of Canada International Polar Year funding, the Natural Sciences and Engineering Research Council, the Northern Scientific Training Program, the Ontario Innovation Trust, the Polar Continental Shelf Program, and the Ontario Research Fund. We thank former CANDAC/PEARL PI James Drummond, PEARL Site Manager Pierre Fogal, CANDAC Data Manager Yan Tsehtik, the CANDAC operators, and the staff at ECCC's Eureka Weather Station for their contributions to data acquisition and for logistical and on-site support. Funding by the DFG within the Transregio TRR 172, Arctic Amplification, (AC)3 is acknowledged by Justus Notholt. Rigel Kivi received financial support by the Academy of Finland through Grant 140408. Kathy S. Law, Jean-Christophe Raut, Louis Marelle, and Tatsuo Onishi (LATMOS) acknowledge support from the EU iCUPE (Integrating and Comprehensive Understanding on Polar Environments) project (grant agreement no. 689443) under the European Network for Observing our Changing Planet (ERA-Planet) and for access to IDRIS HPC resources (GENCI allocations A011017141 and A013017141) as well as the IPSL mesoscale computing center (SPIRIT) for model simulations. Kathy S. Law also acknowledges support from the French Space Agency (CNES) MERLIN (contract no. 9838). Svetlana Tsyro and Michael Gauss have received support from the AMAP Secretariat and the EMEP Trust Fund. Naga Oshima was supported by the Environment Research and Technology Development Fund (JPMEERF20202003 and JPMEERF20232001) of the Environmental Restoration and Conservation Agency provided by Ministry of the Environment of

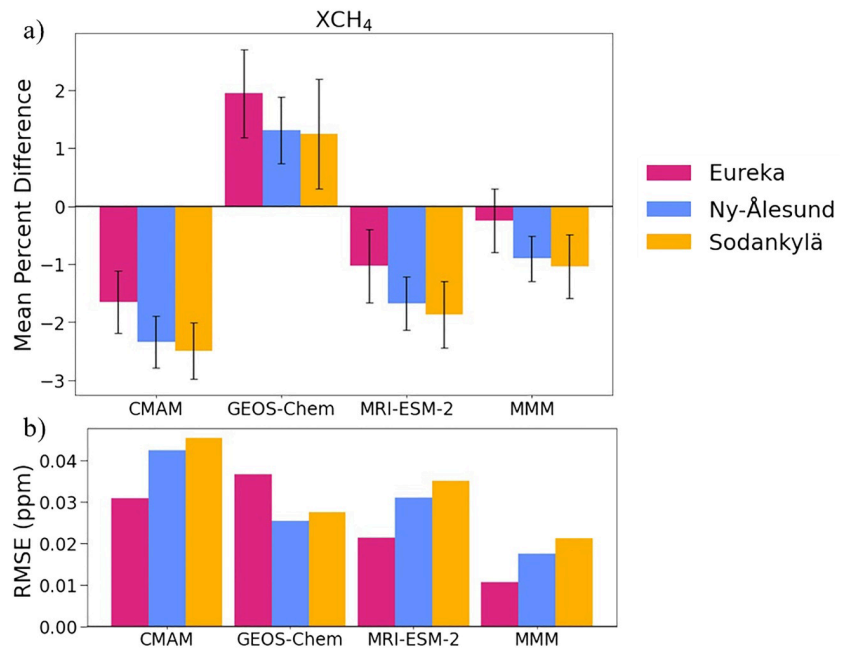


Figure 9. (a) Mean percent differences in XCH₄ between each model and TCCON data set (calculated using Equation 6) for all 3-hourly points at each site. Error bars represent the standard deviation of the set of percent differences used for the mean. MMM indicates the multi-model mean. (b) Root Mean Squared Error (RMSE) (ppm) by model and site.

biases observed) and the time periods modeled, as all three Arctic TCCON sites have data records going back to at least 2010. As more models develop the capacity to simulate emissions-based atmospheric CH₄, we recommend that the TCCON data set continue to be used for these model evaluations.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The TCCON data are publicly available from the TCCON Data Archive, <https://tccondata.org/> (Buschmann et al., 2022; Kivi et al., 2022; Strong et al., 2022), and the model data are available at <https://doi.org/10.18164/e0a0ac5c-d851-45b9-b6d9-4abc29d7d419> (CCCma, 2023). The GGG2020 code is documented in Laughner et al. (2024) and is publicly available at <https://github.com/TCCON/GGG>.

References

- AMAP. (2021). *AMAP Assessment 2021: Impacts of short-lived climate forcers on Arctic climate, air quality, and human health* (p. 375). Arctic Monitoring and Assessment Programme (AMAP), Tromsø, Norway. x +.
- Andersson, C., Langner, J., & Bergström, R. (2007). Interannual variation and trends in air pollution over Europe due to climate variability during 1958–2001 simulated with a regional CTM coupled to the ERA-40 reanalysis. *Tellus B: Chemical and Physical Meteorology*, 59(1), 77–98. <https://doi.org/10.1111/j.1600-0889.2006.00231.x>
- Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., et al. (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research*, 106(D19), 23073–23095. <https://doi.org/10.1029/2001JD000807>
- Buschmann, M., Petri, C., Palm, M., Warneke, T., & Notholt, J. (2022). TCCON data from Ny-Ålesund, Svalbard (NO), release GGG2020.R0 (version R0) [Dataset]. *CaltechDATA*. <https://doi.org/10.14291/tcccon.ggg2020.nyalesund01.R0>
- CCCma. (2023). Canadian Centre for Climate Modelling and Analysis. AMAP SLCF models output in NetCDF format [Dataset]. <https://doi.org/10.18164/e0a0ac5c-d851-45b9-b6d9-4abc29d7d419>
- Cogan, A. J., Boesch, H., Parker, R. J., Feng, L., Palmer, P. I., Blavier, J.-F. L., et al. (2012). Atmospheric carbon dioxide retrieved from the Greenhouse Gases Observing Satellite (GOSAT): Comparison with ground-based TCCON observations and GEOS-Chem model calculations. *Journal of Geophysical Research*, 117(D21), D21301. <https://doi.org/10.1029/2012jd018087>
- Danabasoglu, G., Lamarque, J., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modelling Earth Systems*, 12, e2019MS001916. <https://doi.org/10.1029/2019MS001916>

Japan, the Arctic Challenge for Sustainability 2 and 3 (ArCS-2 and ArCS-3), Program Grant JPMXD1420318865, and a grant for the Global Environmental Research Coordination System from Ministry of the Environment of Japan (MLIT2253).

- Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J. F., Pfister, G. G., Fillmore, D., et al. (2010). Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4). *Geoscientific Model Development*, 3(1), 43–67. <https://doi.org/10.5194/GMD-3-43-2010>
- Fisher, J. A., Jacob, D. J., Purdy, M. T., Kopacz, M., Sager, P. L., Carouge, C., et al. (2010). Source attribution and interannual variability of Arctic pollution in spring constrained by aircraft (ARCTAS, ARCPAC) and satellite (AIRS) observations of carbon monoxide. *Atmospheric Chemistry and Physics*, 10(3), 977–996. <https://doi.org/10.5194/acp-10-977-2010>
- Flood, V. A., Strong, K., Whaley, C. H., Walker, K. A., Blumenstock, T., Hannigan, J. W., et al. (2024). Evaluating modelled tropospheric columns of CH₄, CO, and O₃ in the Arctic using ground-based Fourier transform infrared (FTIR) measurements. *Atmospheric Chemistry and Physics*, 24(2), 1079–1118. <https://doi.org/10.5194/acp-24-1079-2024>
- Fogal, P. F., Leblanc, L. M., & Drummond, J. R. (2013). The Polar Environment Atmospheric Research Laboratory (PEARL): Sounding the atmosphere at 80° north. *Arctic*, 66(3), 377–386. <https://doi.org/10.14430/arctic4321>
- Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., et al. (2021). The Earth's energy budget, climate feedbacks, and climate sensitivity. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Eds.), *Climate Change 2021: The Physical Science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 923–1054). Cambridge University Press. <https://doi.org/10.1017/9781009157896.009>
- Gong, W., Beagley, S., Zhang, J., Staebler, R., Aliabadi, A. A., Sharma, S., et al. (2018). Modelling regional air quality in the Canadian Arctic: Simulation of an Arctic summer field campaign. *Springer Proceedings in Complexity*, 401–406. https://doi.org/10.1007/978-3-319-57645-9_63/FIGURES/3
- Gong, W., Makar, P. A., Zhang, J., Milbrandt, J., Gravel, S., Hayden, K. L., et al. (2015). Modelling aerosol cloud meteorology interaction: A case study with a fully coupled air quality model GEM-MACH. *Atmospheric Environment*, 115, 695–715. <https://doi.org/10.1016/j.atmosenv.2015.05.062>
- Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., & Wang, X. (2012). The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions. *Geoscientific Model Development*, 5(6), 1471–1492. <https://doi.org/10.5194/gmd-5-1471-2012>
- Höglund-Isaksson, L., Gómez-Sanabria, A., Klimont, Z., Rafaj, P., & Schöpp, W. (2020). Technical potentials and costs for reducing global anthropogenic methane emissions in the 2050 timeframe – Results from the GAINS model. *Environmental Research Communications*, 3(1), 025004. <https://doi.org/10.1088/2515-7620/ab7457>
- Horowitz, L. W., Walters, S., Mauzerall, D. L., Emmons, L. K., Rasch, P. J., Granier, C., et al. (2003). A global simulation of tropospheric ozone and related tracers: Description and evaluation of MOZART, version 2. *Journal of Geophysical Research*, 108(D24), 4784. <https://doi.org/10.1029/2002JD002853>
- IPCC. (2021). In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, et al. (Eds.), *Climate Change 2021: The Physical Science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (p. 2391). Cambridge University Press. <https://doi.org/10.1017/9781009157896>
- Jiang, Z., Jones, D. B. A., Worden, J., Worden, H. M., Henze, D. K., & Wang, Y. X. (2015). Regional data assimilation of multi-spectral MOPITT observations of CO over North America. *Atmospheric Chemistry and Physics*, 15(12), 6801–6814. <https://doi.org/10.5194/acp-15-6801-2015>
- Jonsson, A. I., de Grandpré, J., Fomichev, V. I., McConnell, J. C., & Beagley, S. R. (2004). Doubled CO₂-induced cooling in the middle atmosphere photochemical analysis of the ozone radiative feedback. *Journal of Geophysical Research*, 109, D24103. <https://doi.org/10.1029/2004JD005093>
- Kasibhatla, P., Arellano, A., Logan, J. A., Palmer, P. I., & Novelli, P. (2002). Top-down estimate of a large source of atmospheric carbon monoxide associated with fuel combustion in Asia. *Geophysical Research Letters*, 29(19), 1900. <https://doi.org/10.1029/2002GL015581>
- Kawai, H., Yukimoto, S., Koshiro, T., Oshima, N., Tanaka, T., Yoshimura, H., & Nagasawa, R. (2019). Significant improvement of cloud representation in the global climate model MRI-ESM2. *Geoscientific Model Development*, 12(7), 2875–2897. <https://doi.org/10.5194/gmd-12-2875-2019>
- Kivi, R., & Heikkinen, P. (2016). Fourier transform spectrometer measurements of column CO₂ at Sodankylä, Finland. *Geoscientific Instrumentation, Methods and Data Systems*, 5(2), 271–279. <https://doi.org/10.5194/gi-5-271-2016>
- Kivi, R., Heikkinen, P., & Kyrö, E. (2022). TCCON data from Sodankylä (FI), release GGG2020.R0 (version R0) [Dataset]. *CaltechDATA*. <https://doi.org/10.14291/tcon.ggg2020.sodankyla01.R0>
- Kolonjari, F., Plummer, D. A., Walker, K. A., Boone, C. D., Elkins, J. W., Hegglin, M. I., et al. (2018). Assessing stratospheric transport in CMAM30 simulations using ACE-FTS measurements. *Atmospheric Chemistry and Physics*, 18(9), 6801–6828. <https://doi.org/10.5194/acp-18-6801-2018>
- Laughner, J. (2023). GGG2020 auxiliary data. *TCCON Wiki*. <https://tcon-wiki.caltech.edu/Main/AuxiliaryDataGGG2020>
- Laughner, J., Roche, S., Kiel, M., Toon, G. C., Wunch, D., Baier, B. C., et al. (2023). A new algorithm to generate a priori trace gas profiles for the GGG2020 retrieval algorithm. *Atmospheric Measurement Techniques*, 16(5), 1121–1146. <https://doi.org/10.5194/amt-16-1121-2023>
- Laughner, J., Toon, G. C., Mendonça, J., Petri, C., Roche, S., Wunch, D., et al. (2024). The Total Carbon Column Observing Network's GGG2020 data version. *Earth System Science Data*, 16(5), 2197–2260. <https://doi.org/10.5194/essd-16-2197-2024>
- Liu, X., Ma, P.-L., Wang, H., Tilmes, S., Singh, B., Easter, R. C., et al. (2016). Description and evaluation of a new four-mode version of the Modal Aerosol Module (MAM4) within version 5.3 of the Community Atmosphere Model. *Geoscientific Model Development*, 9(2), 505–526. <https://doi.org/10.5194/gmd-9-505-2016>
- Lutsch, E., Strong, K., Jones, D. B. A., Blumenstock, T., Conway, S., Fisher, J. A., et al. (2020). Detection and attribution of wildfire pollution in the Arctic and northern midlatitudes using a network of Fourier-transform infrared spectrometers and GEOS-Chem. *Atmospheric Chemistry and Physics*, 20(21), 12813–12851. <https://doi.org/10.5194/acp-20-12813-2020>
- Makar, P. A., Gong, W., Hogrefe, C., Zhang, Y., Curci, G., Zabkar, R., et al. (2015a). Feedbacks between air pollution and weather. Part 2: Effects on chemistry. *Atmospheric Environment*, 115, 499–526. <https://doi.org/10.1016/j.atmosenv.2014.10.021>
- Makar, P. A., Gong, W., Milbrandt, J., Hogrefe, C., Zhang, Y., Curci, G., et al. (2015b). Feedbacks between air pollution and weather. Part 1: Effects on weather. *Atmospheric Environment*, 115, 442–469. <https://doi.org/10.1016/j.atmosenv.2014.12.003>
- Marelle, L., Raut, J.-C., Law, K. S., Berg, L. K., Fast, J. D., Easter, R. C., et al. (2017). Improvements to the WRF-Chem 3.5.1 model for quasi-hemispheric simulations of aerosols and ozone in the Arctic. *Geoscientific Model Development*, 10, 3661–3677. <https://doi.org/10.5194/gmd-10-3661-2017>
- Marelle, L., Raut, J.-C., Law, K. S., & Duclaux, O. (2018). Current and future Arctic aerosols and ozone from remote emissions and emerging local sources – Modeled source contributions and radiative effects. *Journal of Geophysical Research: Atmospheres*, 123(22), 12942–12963. <https://doi.org/10.1029/2018JD028863>

- Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., et al. (2017). Historical greenhouse gas concentrations for climate modelling (CMIP6). *Geoscientific Model Development*, 10(5), 2057–2116. <https://doi.org/10.5194/gmd-10-2057-2017>
- Messerschmidt, J., Parazoo, N., Wunch, D., Deutscher, N. M., Roehl, C., Warneke, T., & Wennberg, P. O. (2013). Evaluation of seasonal atmosphere–biosphere exchange estimations with TCCON measurements. *Atmospheric Chemistry and Physics*, 13(10), 5103–5115. <https://doi.org/10.5194/acp-13-5103-2013>
- Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., van Weele, M., & Boersma, K. F. (2012). Simultaneous assimilation of satellite NO₂, O₃, CO, and HNO₃ data for the analysis of tropospheric chemical composition and emissions. *Atmospheric Chemistry and Physics*, 12(20), 9545–9579. <https://doi.org/10.5194/acp-12-9545-2012>
- Monks, P. S., Archibald, A. T., Colette, A., Cooper, O., Coyle, M., Derwent, R., et al. (2015). Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer. *Atmospheric Chemistry and Physics*, 15, 8889–8973. <https://doi.org/10.5194/acp-15-8889-2015>
- Moran, M. D., Pavlovic, R., & Anselmo, D. (2018). Regional air quality deterministic prediction system (RAQDPS): Update from version 019 to version 020, Environment and Climate Change Canada, Montreal.
- Nisbet, E. G., Manning, M. R., Dlugokencky, E. J., Michel, S. E., Lan, X., Röckmann, T., et al. (2023). Atmospheric methane: Comparison between methane's record in 2006–2022 and during glacial terminations. *Global Biogeochemical Cycles*, 37(8), e2023GB007875. <https://doi.org/10.1029/2023GB007875>
- Olivié, D., Höglund-Isaksson, L., Klimont, Z., & von Salzen, K. (2021). *Boxmodel for calculation of global atmospheric methane concentration*. Zenodo. <https://doi.org/10.5281/zenodo.5293940>
- Oshima, N., Yukimoto, S., Deushi, M., Koshiro, T., Kawai, H., Tanaka, T. Y., & Yoshida, K. (2020). Global and Arctic effective radiative forcing of anthropogenic gases and aerosols in MRI-ESM2.0. *Progress in Earth and Planetary Science*, 7(1), 38. <https://doi.org/10.1186/s40645-020-00348-w>
- Ostler, A., Sussmann, R., Patra, P. K., Houweling, S., De Bruine, M., Stiller, G. P., et al. (2016). Evaluation of column-averaged methane models and TCCON with a focus on the stratosphere. *Atmospheric Measurement Techniques*, 9, 4843–4859. <https://doi.org/10.5194/amt-9-4843-2016>
- Pétron, G., Granier, C., Khattatov, B., Lamarque, J.-F., Yudin, V., Müller, J.-F., & Gille, J. (2002). Inverse modeling of carbon monoxide surface emissions using Climate Monitoring and Diagnostics Laboratory network observations. *Journal of Geophysical Research*, 107, D24. <https://doi.org/10.1029/2001JD001305>
- Quennehan, B., Raut, J.-C., Law, K. S., Daskalakis, N., Ancellet, G., Clerbaux, C., et al. (2016). Multi-model evaluation of short-lived pollutant distributions over east Asia during summer 2008. *Atmospheric Chemistry and Physics*, 16(17), 10765–10792. <https://doi.org/10.5194/acp-16-10765-2016>
- Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., et al. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth & Environment*, 3(1), 168. <https://doi.org/10.1038/s43247-022-00498-3>
- Robertson, L., Langner, J., & Engardt, M. (1999). An Eulerian limited area atmospheric transport model. *Journal of Applied Meteorology and Climatology*, 38(2), 190–210. [https://doi.org/10.1175/1520-0450\(1999\)038<0190:AELAAT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1999)038<0190:AELAAT>2.0.CO;2)
- Scinocca, J. F., McFarlane, N. A., Lazare, M., Li, J., & Plummer, D. (2008). Technical note: The CCCma third generation AGCM and its extension into the middle atmosphere. *Atmospheric Chemistry and Physics*, 8(23), 7055–7074. <https://doi.org/10.5194/acp-8-7055-2008>
- Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., et al. (2012). The EMEP MSC-W chemical transport model – Technical description. *Atmospheric Chemistry and Physics*, 12(16), 7825–7865. <https://doi.org/10.5194/acp-12-7825-2012>
- Simpson, D., Bergström, R., Tsyro, S., & Wind, P. (2019). Updates to the EMEP MSC-W model, 2018–2019, Chap. 10. In *Transboundary particulate matter, photo-oxidants, acidifying and eutrophying components, the Norwegian Meteorological Institute, Oslo, Norway, EMEP status report 1/2019* (pp. 145–152). Retrieved from https://emep.int/publ/reports/2019/EMEP_Status_Report_1_2019.pdf
- Stohl, A., Aamaas, B., Amann, M., Baker, L. H., Bellouin, N., Bernsten, T. K., et al. (2015). Evaluating the climate and air quality impacts of short-lived pollutants. *Atmospheric Chemistry and Physics*, 15(18), 10529–10566. <https://doi.org/10.5194/acp-15-10529-2015>
- Strong, K., Roche, S., Franklin, J. E., Mendonca, J., Lutsch, E., Weaver, D., et al. (2022). TCCON data from Eureka (CA), release GGG2020.R0 (version R0) [Dataset]. *CaltechDATA*. <https://doi.org/10.14291/tcccon.ggg2020.eureka01.R0>
- Viatte, C., Strong, K., Hannigan, J., Nussbaumer, E., Emmons, L. K., Conway, S., et al. (2015). Identifying fire plumes in the Arctic with tropospheric FTIR measurements and transport models. *Atmospheric Chemistry and Physics*, 15(5), 2227–2246. <https://doi.org/10.5194/acp-15-2227-2015>
- Whaley, C. H., Law, K. S., Hjorth, J. L., Skov, H., Arnold, S. R., Langner, J., et al. (2023). Arctic tropospheric ozone: Assessment of current knowledge and model performance. *Atmospheric Chemistry and Physics*, 23(1), 637–661. <https://doi.org/10.5194/acp-23-637-2023>
- Whaley, C. H., Mahmood, R., von Salzen, K., Winter, B., Eckhardt, S., Arnold, S., et al. (2022). Model evaluation of short-lived climate forcers for the Arctic Monitoring and Assessment Programme: A multi-species, multi-model study. *Atmospheric Chemistry and Physics*, 22(9), 5775–5828. <https://doi.org/10.5194/acp-22-5775-2022>
- Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., & Soja, A. J. (2011). The Fire INventory from NCAR (FINN). A high resolution global model to estimate the emissions from open burning. *Geoscientific Model Development*, 4(3), 625–647. <https://doi.org/10.5194/gmd-4-625-2011>
- Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., et al. (2011). The Total Carbon Column Observing Network. *Philosophical Transactions of the Royal Society A*, 369(1943), 2087–2112. <https://doi.org/10.1098/rsta.2010.0240>
- Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., et al. (2019). The Meteorological Research Institute Earth System Model version 2.0, MRI-ESM2.0: Description and basic evaluation of the physical component. *Journal of the Meteorological Society of Japan*, 97(5), 931–965. <https://doi.org/10.2151/jmsj.2019-051>