https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

Enhancing GOSAT methane observations over the high latitudes using a multi-objective genetic algorithm optimisation approach

Lakshmi N Bharathan[1], Robert J Parker[1], Michael Cartwright[1], Dan Orr[1], Antonio Di Noia[2], Peter Somkuti[3,4], Alex Webb[3,4], and Hartmut Bösch[5]

[1]National Centre for Earth Observation, University of Leicester, Leicester, UK
[2]Department of Civil and Computer Engineering, Tor Vergata University, Rome, Italy
[3]Earth System Science Interdisciplinary Center, University of Maryland, Maryland, USA
[4]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Maryland, USA
[5]Institute of Environmental Physics, University of Bremen, Bremen, Germany

Correspondence: Robert J Parker (rjp23@leicester.ac.uk)

Abstract. The Greenhouse Observing Satellite (GOSAT) is the world's first satellite mission dedicated for greenhouse gas monitoring and the University of Leicester has generated a well-validated, global data set of methane column dry-air mole fraction (XCH_4) which has been extensively used for regional and global methane emission attribution and trend analyses. However, satellite remote sensing for greenhouse gases is inherently challenging over high latitudes, due to severe cloud cover, high solar zenith angles and unfavourable surface conditions resulting in a major deficit of high-latitude winter data. A significant portion of otherwise successful retrievals are lost during the post-retrieval quality filtering process because the standard quality filtering, optimised for global data throughput, disproportionately affects, is overly restrictive for the high latitudes. Relaxing quality filters naturally leads to degradation in data quality due to increased contamination from clouds, aerosols and dark surfaces, and optimising quality filters is essential to obtain the best balance between data quality and data quantity. This study successfully improves the high-latitude throughput of the University of Leicester GOSAT Proxy XCH_4 dataset using a multi-objective genetic algorithm (GA) approach by optimising the post-retrieval filtering process, with the least impact on the data quality in comparison with the ground-based observations from Total Carbon Column Observing Network (TCCON) stations. We have found that the GA-optimised quality filtering can significantly increase the number of valid GOSAT methane observations over high latitudes by up to 20% with a compromise of less than 1 ppb in single measurement precision. The optimisation enhances data throughput across the high-latitudes and preserves the statistical distribution and climatology of the original dataset. The optimisation process more than doubled the data in December, significantly contributing to mitigating the winter data-deficit in the high latitudes. This genetic algorithm optimisation approach holds potential for wider applicability including optimising the observation throughput for future satellite missions like CO2M, MicroCarb and GOSAT-GW.

1 Introduction

Monitoring greenhouse gases (GHG) in northern high-latitude regions is important as these regions respond rapidly to global warming at a rate twice the global average with the highest warming in the cold season (Bekryaev et al., 2010; IPCC, 2023).

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

Northern permafrost contains twice the amount of carbon that is present in the atmosphere in the form of organic matter stored in the frozen soil over thousands of years (Tamocai et al., 2009). Climate change-induced warming over the northern high-latitudes leads to extensive thawing of the permafrost, reduced surface albedo, and exposure of organic matter that can release large amounts of methane (CH_4) and carbon dioxide (CO_2), potentially leading to positive feedback loops and associated impacts on climate (Miner et al., 2022; Schuur et al., 2015). In the northern hemisphere, permafrost ground temperatures have increased by approximately $1°C$ since 2000, followed by a decreasing trend in surface water extent, continuous land-subsidence and accelerating glacier melting (Bartsch et al., 2023; Smith et al., 2022). A record high temperature was observed in station-based measurements during 2018-2019 in the Arctic and sub-Arctic regions with warming rates climbing to $1°C$ per decade (Smith et al., 2022). Rising temperatures can lead to the transitioning of many high-latitude ecosystems from being carbon sinks to sources (Vrese et al., 2021) which can trigger climate tipping points producing cascading and irreversible changes in the Earth-atmosphere system (Lenton, 2012; McKay et al., 2022). Studies have reported methane emissions are progressing steadily over the Arctic especially in the cold season due to the presence of a microbially active "zero-curtain" unfrozen layer below the frozen ground (McGuire et al., 2012; Ward et al., 2024; Zona et al., 2016). Station based atmospheric methane measurements are sparse over the Arctic due to challenging adverse weather conditions especially during winter limiting accurate information on seasonal trends and spatial variability (Pallandt et al., 2022; Wittig et al., 2023). Thus, its crucial to have regular, robust satellite GHG measurements of the pan-Arctic atmosphere to keep track of baselines, growth rates, regional hotspots, and seasonal variations, critical for both top-down inversion of emission fluxes and validating process-based modelling.

Satellites like the Greenhouse Gases Observing Satellite (GOSAT) (Kuze et al., 2009) and the TROPOspheric Monitoring Instrument (TROPOMI) onboard Sentinel-5P (Veefkind et al., 2012; Hu et al., 2018) have been regularly tracking global methane levels for long-periods providing long-term climate data records(Buchwitz et al., 2017, 2018). However, satellite monitoring of GHGs over high-latitudes is limited by diverse environmental factors, despite these regions being highly vulnerable to global warming. High-latitude regions, especially the Arctic, experience persistent clouds, which are expected to increase under warming climate (Vavrus et al., 2009) with a low solar angle much of the year that limits the amount of radiation available for passive optical and short-wave infrared satellite sensors (Marshall et al., 1993). This results in a low signal to noise ratio and decreased satellite data throughput over these regions. Optical and short-wave infrared measurements during winter are especially difficult due to decreased daytime and extended periods of a total absence of sunlight. High absorption of incoming short-wave infrared radiation over ice also contributes to biases in the retrieval of greenhouse gases over high-latitudes. Since these natural limitations cannot be changed, we need to focus on optimising the available satellite measurements to get maximum data throughput. However, increasing the throughput by relaxing filtering conditions can result in degrading the data quality. This trade-off between coverage and quality is common in satellite retrieval applications. Thus, the challenge lies in optimising retrieval systems to maximise data coverage while minimising its impact on data quality while comparing with the ground truth.

Genetic algorithms (GA) are being increasingly used for optimising various aspects of satellite remote sensing of the Earth-atmosphere system, from satellite constellation design (Qin et al., 2025), tuning of sensors and antenna (Sánchez-Sevilleja

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

et al., 2025), geo-referencing of satellite images (Jannati and Zoej, 2015) and optimising remote sensing models (Zhou et al., 2014). Many studies have demonstrated the effectiveness of genetic algorithms specifically in the area of satellite imaging which include land cover change detection (Celik, 2010; Pati et al., 2020; Ranjani, 2017; Slimani and Hedjam, 2022), texture feature detection in satellite radar images (Akbarizadeh, 2012), detecting sea clutter and small objects in satellite images (Leung et al., 2002), and assessing building detection and damage during natural disasters (Khodaverdizahraee et al., 2020; Sumer and Turker, 2013). There are different variations of the GA such as a weight-based GA, a vector-evaluated genetic algorithm (VEGA), Multi-objective Genetic Algorithm (MOGA), Niched Pareto Genetic Algorithm (NPGA), Strength Pareto Evolutionary Algorithm (SPEA) and Non-dominated Sorting Genetic Algorithm (NSGA) (Deb et al., 2002; Srinivas and Deb, 1994). NSGA is considered the most efficient amongst them when it comes to solving problems with competing objectives (Deb et al., 2002; Circiu and Leon, 2010). In this study, our objective is to maximise the data density of the University of Leicester GOSAT Proxy methane data retrieval specifically over high latitudes while minimising the degradation of data quality that arises when retrieval filters are relaxed. A non-dominating sorting approach must be well-suited for this application as our objective involve exploring the right trade-off between data coverage and data quality. Thus, our study explores the application of, the latest improved version of the non-dominated sorting genetic algorithm II (NSGA II), to optimise the GOSAT methane total column dry-air mole fraction (XCH4) data retrieved by the University of Leicester GOSAT Proxy methane algorithm (Parker et al., 2011, 2020) with a focus on high-latitude regions north of 50°N.

There have been only a few attempts to increase satellite greenhouse gas data coverage over high-latitudes (Jacobs et al., 2020; Mendonca et al., 2021; Noël et al., 2022). Notably, Mendonca et al., 2021 utilised a neural network approach to filter the OCO-2 XCO_2 data based on binary classification criteria that require training the algorithm with multiple, labelled proxy parameters measured at a reference ground site whereas a genetic algorithm directly optimises the threshold filters with no training with large labelled classification datasets.

2 Data and methodology

2.1 GOSAT - The University of Leicester GOSAT Proxy XCH_4 Retrieval data product

Launched in January 2009 jointly by The Japan Aerospace Exploration Agency (JAXA), The Japanese Ministry of the Environment (MOE) and the National Institute for Environmental Studies (NIES), the Greenhouse Gases Observing Satellite "IBUKI" (GOSAT) has played a significant role in advancing our knowledge of regional and global GHG emissions (Kuze et al., 2009; Kuz, 2014; Kuze et al., 2016). GOSAT employs the Thermal and Near-infrared Sensor for carbon Observation (TANSO) instrument, comprising the Fourier Transform Spectrometer (FTS) and the Cloud and Aerosol Imager, to measure infrared radiation from the Earth's surface and the atmosphere. TANSO-FTS observes sunlight in four spectral bands: three Shortwave Infrared (SWIR) bands (0.758–0.775 μm, 1.56–1.72 μm, 1.92–2.08 μm) with near surface-sensitivity and one Thermal Infrared (TIR) band (5.5 - 14.3 μm) with mid-tropospheric sensitivity, each band targeting specific atmospheric components. The first SWIR band (0.758–0.775 μm) is sensitive to oxygen (O_2), enabling aerosols, cloud and surface pressure measurements. The second SWIR band (1.56–1.72 μm) is sensitive to both carbon dioxide (CO_2) and methane (CH_4) concentrations, while the third

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

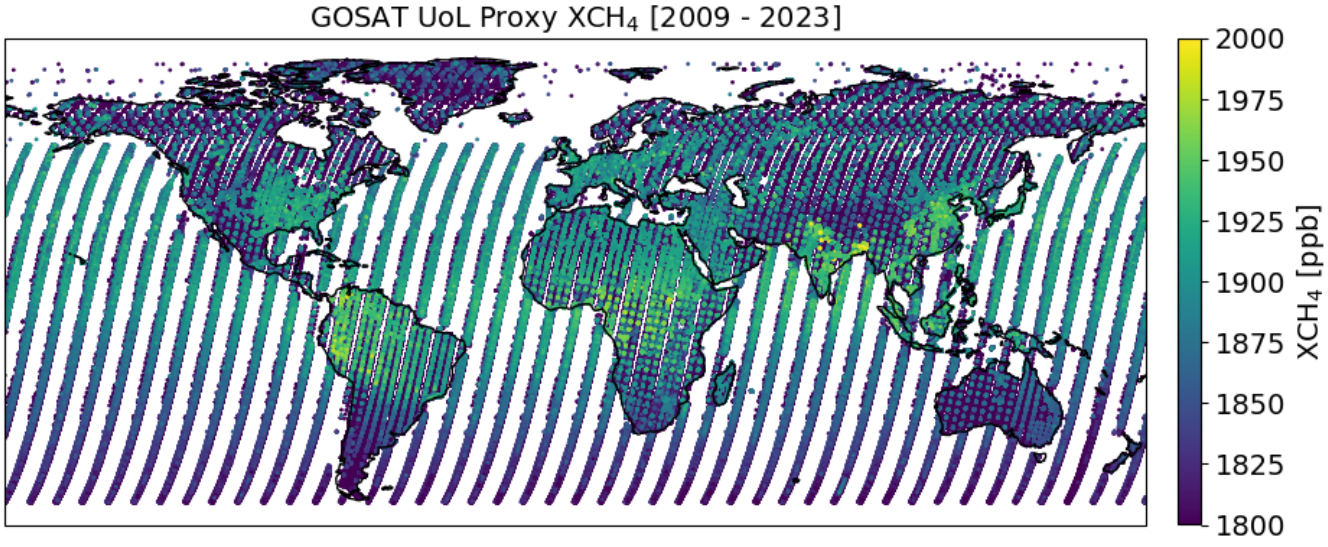

Figure 1. Global distribution of total column-averaged methane dry-air mole fraction [XCH4] from GOSAT at 2° x 2° grid resolution, retrieved using the University of Leicester GOSAT Proxy XCH4 retrieval algorithm during the period 2009 to 2023.

SWIR band (1.92–2.08 μm) primarily focuses on CO_2 measurements. Including the two orthogonal polarisation channels in the SWIR, it observes incoming light in 7 different channels. GOSAT's large footprint of 10.5 km, combined with its three-day repeat cycle, enables frequent sampling of the atmosphere. The University of Leicester has developed a Proxy algorithm for the column-averaged dry-air mole fraction of methane (XCH_4) retrievals from GOSAT observations and has generated long term total column methane data from April 2009 to January 2025. (Parker et al., 2011, 2015, 2020). Level 1B (version 220.221) calibrated radiance data is processed with the University of Leicester Full-Physics retrieval scheme (UoL-FP) using the Proxy retrieval method to generate XCH_4 dataset (Parker et al., 2011). Atmospheric light scattering is one of the major sources of error in space-based gas retrievals, and as CO_2 is a relatively uniformly distributed gas (compared to CH_4) in the atmosphere, it can act as a suitable proxy for atmospheric light scattering for the methane retrieval. Both CH_4 and CO_2 absorb light at 1.6 μm, allowing for distinct retrievals while being subject to similar scattering effects. The ratio of measured XCH_4 and XCO_2 cancels out the effects of atmospheric light scattering around the common absorption band at 1.6 μm, and the ratio is subsequently multiplied by a model-based estimate of XCO_2 to calculate the corrected proxy XCH_4. The UoL proxy XCH_4 retrieval algorithm begins with pre-processing of L1B data files using the Leicester Retrieval Preparation Toolset, followed by cloud screening, forward modelling, spectral fitting, correcting XCH_4 using proxy XCO_2 and the quality control filtering. Retrieved XCH_4 data has been validated against collocated ground-based observations from the Total Carbon Column Observing Network (TCCON) network with 22 stations over the globe with an overall correlation coefficient of 0.92 and a single-sounding precision of 13.72 ppb (Parker et al., 2020). A bias correction of 9.062 ppb is applied to the data based on comparison with TCCON observations(Parker et al., 2020). Figure 1 shows the spatial distribution of mean GOSAT UoL proxy XCH_4 over the

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

period 2009 to 2023 at a 2˚ x 2˚ grid resolution. It should be noted that the GOSAT sampling is not necessarily uniform in space or time. For example, GOSAT was not operational for certain periods (e.g., January 2015), and the sampling pattern has been adjusted/extended at various stages of the mission.

2.2 Data coverage over high latitudes

Figure 2 shows the a) spatial, b) latitudinal and c) monthly distribution of the observation count (N) of the final quality-checked GOSAT XCH_4. The observation count peaks around the mid-latitudes while the tropics show a decrease which further diminishes significantly at high latitudes with almost no observations during winter months (November - January). Despite frequent orbit passes over the high latitudes, persistent cloud cover, low solar zenith angles and long polar nights limit available observations over this region. In addition, surface albedo effects and extreme atmospheric conditions affect the retrieval quality over high latitudes. About a third of cloud-cleared successful retrievals are lost during the quality filtering process (Parker et al., 2020). Although that includes complete omission of data over Antarctica, data loss over the Arctic is critical for methane monitoring and optimising post-retrieval quality filters is an essential first step towards improving the data throughput over the northern high latitudes.

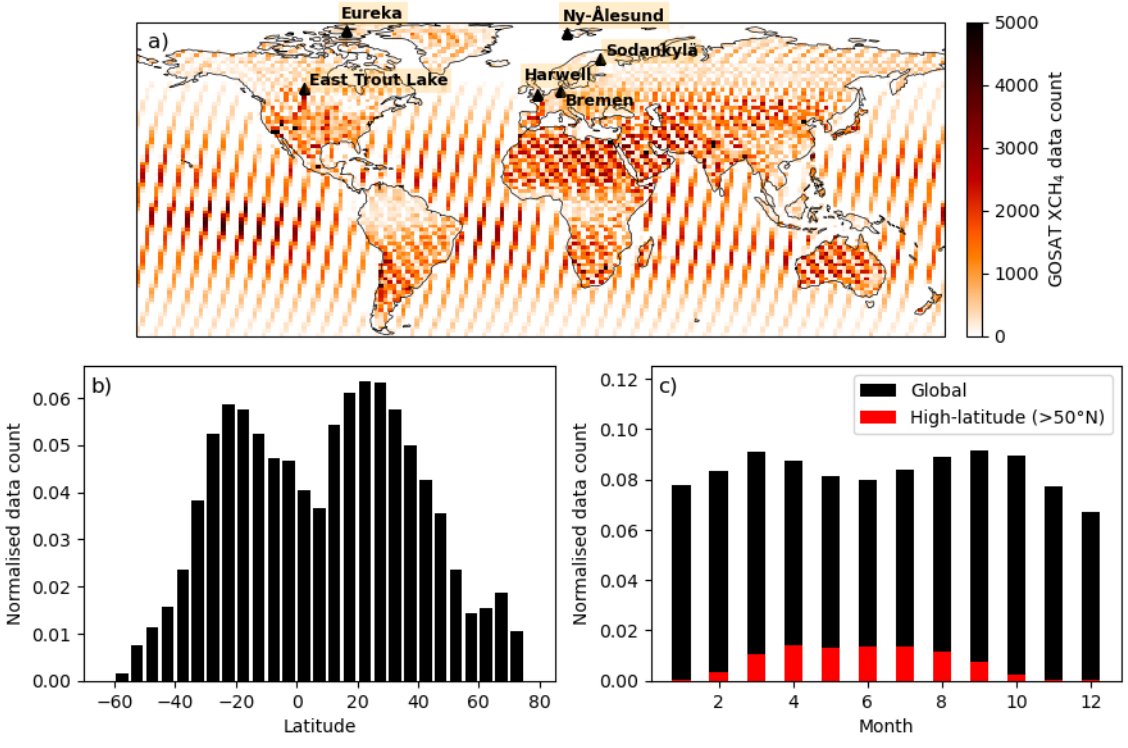

Figure 2. a) Global distribution of GOSAT XCH_4 data count from 2009 to 2023, at 2˚ x 2˚ grid resolution, with high-latitude TCCON stations used in the present study b) latitudinal variation of normalised data count, and c) monthly normalised observation counts during the period 2009 and 2023.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

2.3 TCCON

The Total Carbon Column Observing Network (TCCON) employs ground-based Fourier Transform Spectrometers (FTS) over the globe that measure column concentration of atmospheric trace gases (CO_2, CH_4, N_2O, HF, CO, H_2O, and HDO) from direct solar spectra in the near-IR region (Laughner et al., 2024). We use column-average dry-air mole fraction of methane (GGG2020 data release), primarily at six high-latitude ($> 50°N$) stations, namely Bremen ($53.10°N$) (Notholt et al., 2022), East Trout Lake ($54.35°N$) (Wunch et al., 2022), Harwell ($51.57°N$) (Weidmann et al., 2023), Ny-Ålesund ($78.9°N$) (Buschmann et al., 2022), Sodankylä ($67.36°N$) (Kivi et al., 2022) and Eureka ($80.05°N$) (Strong et al., 2022), to train the genetic algorithm to maximise GOSAT XCH_4 observation density over the high latitudes. The locations of these six stations are shown in Fig. 2a. Measurement precision for TCCON XCH_4 is generally < 0.3% for single measurements meaning they are highly consistent, though it can vary from site to site. Potential sources of errors include spectroscopic uncertainties at high solar zenith angles and instrumental changes (such as instrument line shape) over time. TCCON removes spectroscopic uncertainties through airmass correction and calibration with respect to aircraft measurements, and instrument changes are identified by flagging long time-periods over which these changes occur. TCCON carry out cloud filtering based on the fractional variation in solar intensity (fvsi) expressed as the percentage change in solar intensity measured every 90 seconds, and data points with fvsi > 5% are ignored as a sudden change in solar intensity indicates clouds in the column measurement.

3 Multi-objective genetic algorithm

Optimisation algorithms aim at finding the parameters of a system that are "best" with respect to a given criterion. An optimisation problem is formulated by: 1. Defining the parameters to be optimised (in our case, the quality filtering parameters); 2. Defining a function that measures the "quality" of a set of variables with respect to the given criterion (in our case, a combination between throughput and retrieval precision); 3. Choosing an algorithm for the maximisation or the minimisation of the function defined at point 2. A satellite retrieval algorithm itself can be seen as an optimisation algorithm, as it typically looks for the inputs of a forward model (e.g., the CH_4 profile) that minimise the discrepancy between the modelled signal and the signal that is actually measured by a satellite instrument (e.g., top of atmosphere radiance). An optimisation algorithm typically begins with an initial guess of the parameters to be optimised, generates an iterative sequence of estimates, and evaluates the objective function at each iteration. The process is stopped when convergence is achieved.

Traditional optimisation algorithms, such as Gauss-Newton or Levenberg-Marquardt, compute the derivatives of a cost function and update the solution at each step after identifying a direction of descent (Bekryaev et al., 2010; Nocedal and Wright, 2006). When derivatives are costly to evaluate, or when the problem is hard to formulate using a differentiable objective function, derivative-free algorithms (Rios and Sahinidis, 2013) may be a useful choice. The idea behind derivative-free optimisation algorithms is to randomly generate a population of "candidate" estimates, evaluate the objective function of these estimates, discard those that score worse, and use the remaining ones as a basis for the generation of a new population. Once again, the process is repeated until convergence is achieved. Genetic algorithms (GA) belong to this class of optimisation methods. GA optimisation is based on Darwin's Theory of natural selection (Darwin, 1872), and treats potential solutions to the problem

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

as individuals/chromosomes within a population, which are mixed together to generate better and better solutions (Goldberg, 1989; Holland, 1992; Katoch et al., 2021).

A GA starts with an initial population of random solutions, ranks each solution based on a fitness function (which measures their quality with respect to the chosen objective), selects the fitter individuals as parents, and creates the next generation by producing offspring solutions through the process of crossover and mutation. This process is iterated over different generational cycles until the solution converges. When multiple objectives are involved, the fitness of each solution/individual is typically determined by combining the objectives into a single fitness function (e.g., a weighted sum) that assigns a fitness score to each individual solution which is used to select the best fit individuals for the next generation (iteration). Since the overall fitness is determined by the weights assigned, a user defined weight vector is crucial for an accurate result (Karami and Dariane, 2022). In addition, for problems with conflicting objectives (improvement in one objective results in worsening of the other), a weighted sum fitness approach can be biased towards a certain objective masking the effect of the other objective even when weights are assigned (Das and Dennis, 1997). Thus, this method is not ideal for trade-off problems where improvement in one objective can be achieved only at the expense of the other. The solution is to use a Pareto analysis where multiple objectives are kept separately in an objective space (multi-dimensional space where objective variables are represented along different axes) and each solution is compared with other solutions in the population to determine if they are non-dominated in the population (Deb, 2011; Knowles and Corne, 1999; Srinivas and Deb, 1994). A solution is non-dominated when it is better at least in one objective and is no worse in all other objectives compared to another solution. A set of non-dominated solutions in a population provides the optimal trade-off between the objectives and are considered as fit individuals and will be passed on to the next generation. In addition to non-dominated sorting, selection of fit individuals is also based on crowding distance, which is a measure of how close the solution is close to its neighbours in the objective space. A higher crowding distance is generally preferred, as it avoids solutions from crowded in the objective space, and thereby preserves the diversity of the solutions and avoids premature convergence. Our study uses the Non-dominated Sorting Genetic Algorithm (NSGA-II), a multi-objective evolutionary algorithm widely used in various fields best suited for handling problems with competing objectives (Deb et al., 2002). We have used the Python framework named Pymoo to carry out NSGA II (Blank and Deb, 2020).

3.1 Problem formulation and objectives

In our problem, we are trying to optimize the post-retrieval quality filtering of the GOSAT $X\mathrm{CH_4}$ data; hence, the potential solutions (chromosomes) are sets of data quality filter parameters. Our chromosome consists of eight filter values as shown in Table 1. Four of these comprise the low and high thresholds for the goodness of fit (χ^2) between measured and modelled spectra for $\mathrm{CH_4}$ ($\chi^2_{\mathrm{low}}[\mathrm{CH_4}]$ and $\chi^2_{\mathrm{high}}[\mathrm{CH_4}]$) and for $\mathrm{CO_2}$ ($\chi^2_{\mathrm{low}}[\mathrm{CO_2}]$ and $\chi^2_{\mathrm{high}}[\mathrm{CO_2}]$). The remaining four parameters are the lower threshold values of the raw retrieved $X\mathrm{CH_4}$ ($X\mathrm{CH_{4,low}}$) and $X\mathrm{CO_2}$ ($X\mathrm{CO_{2,low}}$), and the uncertainty values of the retrieved $X\mathrm{CH_4}$ ($X\mathrm{CH_{4,error}}$) and $X\mathrm{CO_2}$ ($X\mathrm{CO_{2,error}}$).

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

Table 1. Quality filters to be optimised to improve data throughput. This forms the chromosome/individual for GA.

$\chi^2_{low}[CH_4]$	$\chi^2_{high}[CH_4]$	$\chi^2_{low}[CO_2]$	$\chi^2_{high}[CO_2]$	$XCH_{4,low}$	$XCO_{2,low}$	$XCH_{4,error}$	$XCO_{2,error}$

Figure 3 shows the global maps of average filter parameters we use in the quality filtering of UoL GOSAT XCH_4 data. Relaxing filter parameters to increase the XCH_4 data count (N), impacts the quality of the data as measured by the root mean square error (RMSE) between GOSAT and TCCON XCH_4 observations ($RMSE_{GOSAT\text{-}TCCON}$). So, our objective is to find an optimised solution that corresponds to a higher number of observations over high latitudes and with minimal increase in $RMSE_{GOSAT\text{-}TCCON}$. So essentially, we have two objectives here for the optimisation problem, 1) to maximise the XCH_4 data count (N) and 2) to minimise $RMSE_{GOSAT\text{-}TCCON}$. These objectives will drive the GA towards the solution that provides the highest N with smallest increase in $RMSE_{GOSAT\text{-}TCCON}$.

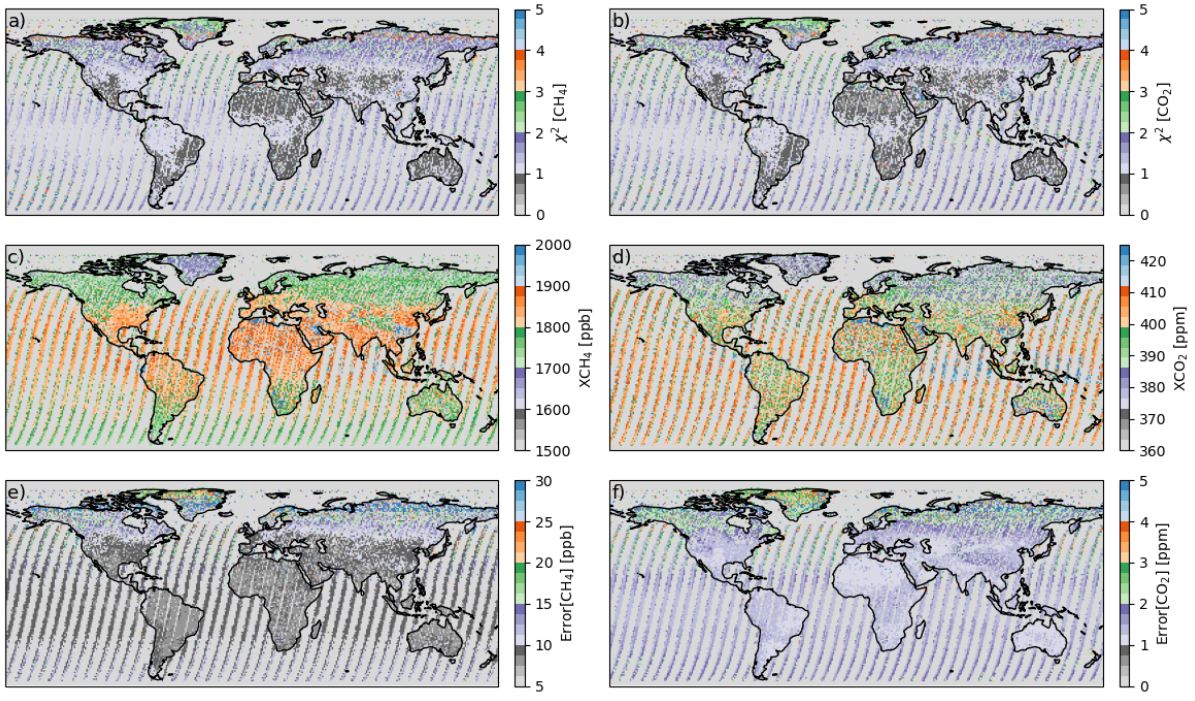

Figure 3. Global distribution of mean filter parameters - a) and b) indicate goodness-of-fit χ^2 parameter between measured and simulated spectra for CH_4 and CO_2 respectively, c), d), e) and f) indicate raw retrieved XCH_4 and XCO_2 values and their associated error estimates respectively.

The GA starts with a population of random solutions/individuals (filter combinations) and iterates through generations by selecting the fitter individuals which are chosen as "parents" and carries out crossover (combining genetic material from the parent chromosomes/solutions) and mutation (random modifications promoting genetic diversity) to produce new solutions/offspring to evolve towards the best solution by evaluating the fitness of each individual. Fitter individuals are selected based

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

on non-dominated sorting and crowding distance in the objective space (multi-dimensional space where each axis represents each objective). It is to be noted that Greenland is omitted from the analysis as anomalous filter values in that region were biasing the GA towards an enhanced number of observations over Greenland.

3.2 Selection of mating pool: non-dominated sorting

We start with an initial population of 100 random individuals/chromosomes (combinations of the 8 filters) and apply this filter solution set to the XCH_4 data to calculate the observation count over high latitudes and compare with collocated TCCON observations over high latitudes to estimate the root mean square error ($RMSE_{GOSAT\text{-}TCCON}$). Even though the individuals of the population are selected randomly, it is made sure that the values are within a predefined lower and upper limit so that GA excludes impossible or irrelevant solutions thereby converging faster. The upper and lower bounds for the filters are given in the parentheses for the following filters, $\chi^2_{low}[CH_4]$ (0 - 1), $\chi^2_{high}[CH_4]$ (1 - 3), $\chi^2_{low}[CO_2]$ (0 -1), $\chi^2_{high}[CO_2]$ (1 - 3), $XCH_{4,low}$ (1400 – 2100 ppb), $XCO_{2,low}$ (250 - 450 ppm), $XCH_{4,error}$ (0 – 30 ppb), and $XCO_{2,error}$ (0 – 4 ppm). While filter values are the decision variables, available observation count (N) and $RMSE_{GOSAT\text{-}TCCON}$ forms the objective variables of our problem. Subsequently, all members of this population undergo crossover and mutation operations to produce 100 new offspring members to form a population of 200-members. We use a simulated binary crossover operator combining 2 parent chromosomes, to create two new offspring chromosomes based on the parents' gene values. Similarly, mutation is carried out using a polynomial mutation operator that introduces small, random changes to the gene values. The genetic parameters that have been used to drive the NSGA-II have been listed in Table 2.

Table 2. Input genetic parameters used in NSGA-II.

Parameters	Value	Description
Population size	100	Number of individuals in each population
Crossover	SBX	Simulated Binary Crossover
Crossover Probability	90 %	90 % of the population will undergo crossover
Mutation	PM	Polynomial Mutation (10 % of the population will undergo mutation)

As mentioned, above, the selection of fit individuals/parents are carried out using non-dominated sorting combined with crowding distance comparison. Each individual is compared with every other individual to rank them based on their domination count. Rank 1: solutions with a domination count of 0 (i.e., not dominated by any other solution). Rank 2: solutions with a domination count of 1 (i.e., dominated by exactly one set of solutions, which are in rank 1). Rank 3: solutions with a domination count of 2, and so on, which are termed as pareto fronts. Out of 200 members of the population, next generation parents are selected by adding individuals sequentially from these fronts until the population size reaches just 100, that contain non-dominated solutions based on their rank. If the size of a front (for e.g., 120) is greater than population size, in this case 100,

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

220 solutions with the highest crowding distance in the objective space will be used to select the fit individuals. The schematic of the optimisation methodology is represented in Fig. 4.

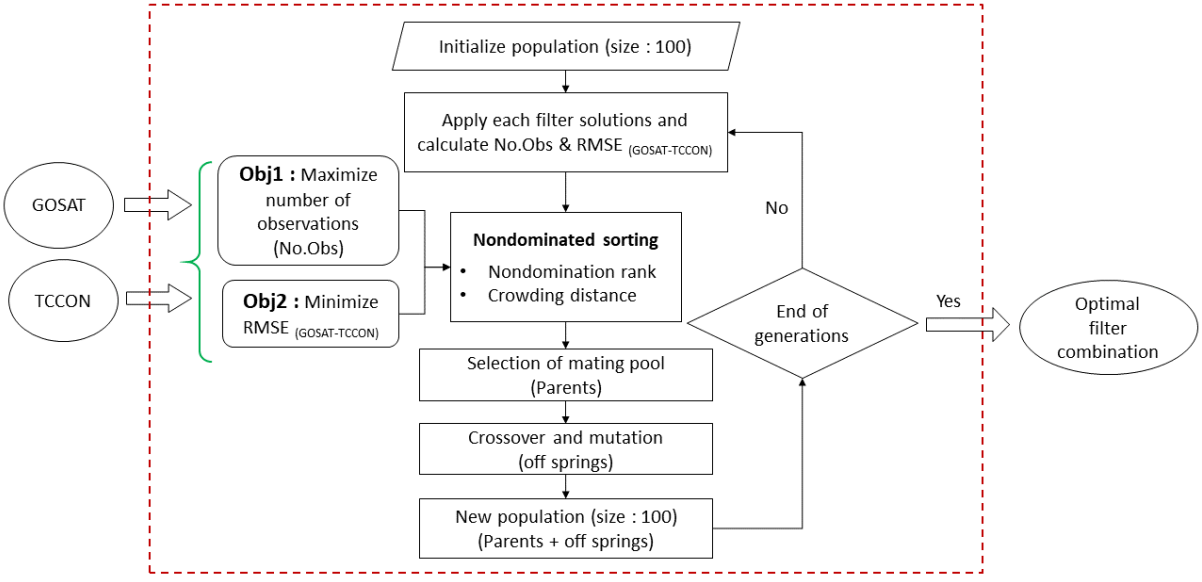

Figure 4. Schematic representation of the post-retrieval filter optimisation methodology.

4 Results and discussion

4.1 Evolution of objective variables

NSGA II carry out a non-dominated sorting procedure for each population/generation to maximise the GOSAT XCH_4 data
225 count (N) and minimise $RMSE_{GOSAT\text{-}TCCON}$ (which represents the quality of the filtered data and is labelled RMSE in figures), which forms the objective variables of our problem. For each generation, the two objective variables are plotted along two dimensions of the objective space (Fig. 5c) and the point of maximum curvature (knee point) in the objective space (RMSE Vs N) is calculated. This point represents the best trade-off between RMSE and N in a solution population. Figure 5a shows the evolution knee point objective variables of each generation over 100 generations. The initial 10 generations show a rapid
230 increase in both knee RMSE and N which later settles to a trade-off state of high N and low RMSE for the rest of the generations. By the 20th generation, the evolution stabilises with the entire population filled with 100 non-dominated, highly fit solutions and the ratio of RMSE to N minimises considerably favouring the objectives of the problem (Fig. 5b). This indicates the successful convergence of GA to achieve filter combinations that would provide the highest possible gain in N for a minimal increase in RMSE as generations evolve further.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

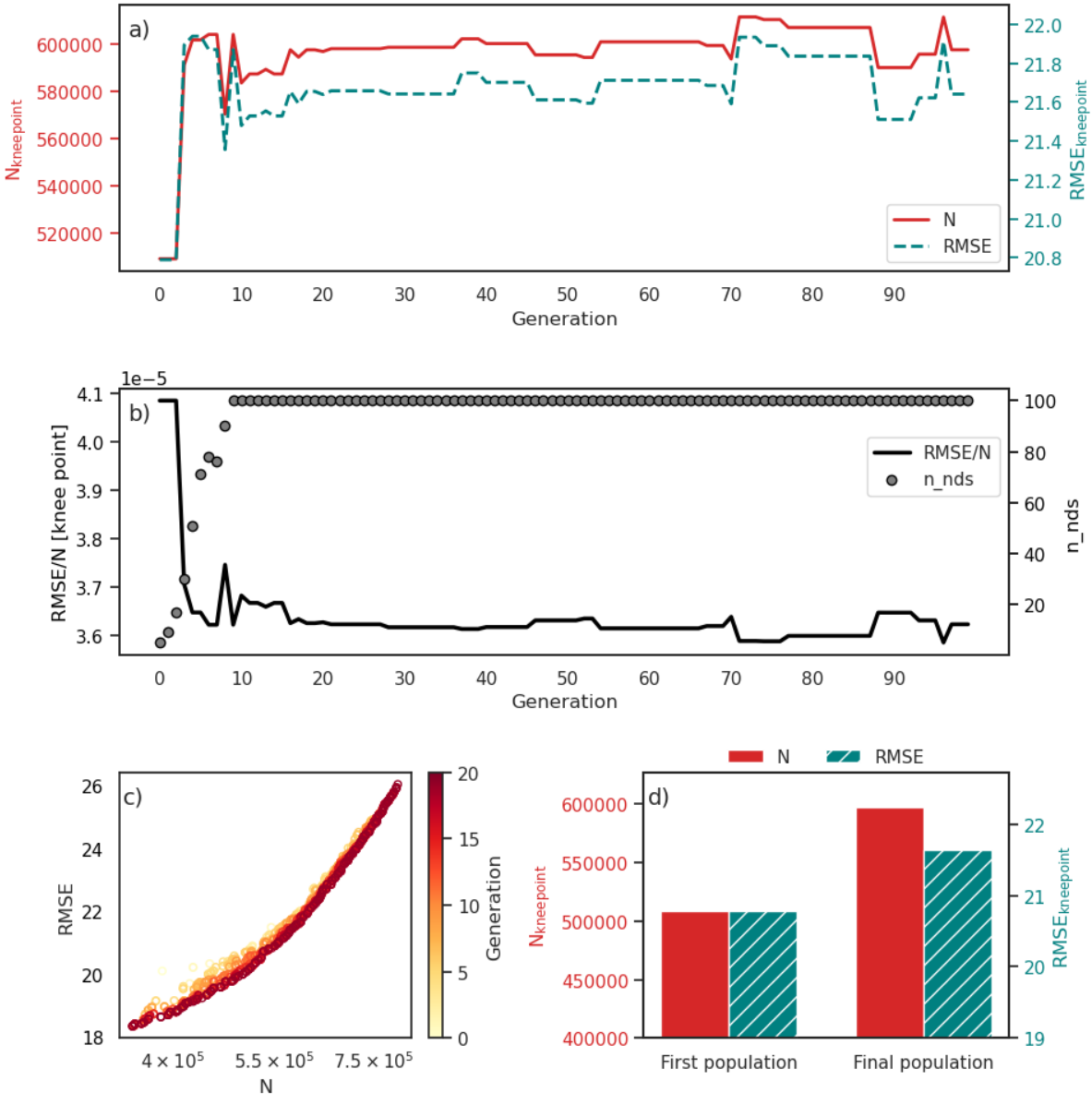

Figure 5. Generational evolution of a) objective variables N (data count over high latitudes (>50˚N)) and RMSE (GOSAT vs TCCON) b) ratio of RMSE and N, and number of nondominated solutions in each population (n_nds) c) Pareto front evolution in first 15 generations, and d) comparison of N and RMSE with standard filters and GA optimised filters. RMSE refers to Root Mean Square Error between collocated GOSAT and TCCON station measurements.

235 The evolution of the pareto front during the first 20 generations is shown in Fig. 5c. Scattered solutions in the initial populations are aligned with a curvature, towards increasing N and decreasing $RMSE_{GOSAT\text{-}TCCON}$ showing the evolution of solutions

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

in the required direction. From the initial population to the final population, the knee N shows a greater enhancement than that in RMSE indicating that GA is capable of maximising N for a minimal increase in RMSE (Fig. 5d).

4.2 Evolution of decision variables - selection of optimised filter combination

240 Figure 6. shows the evolution of the filter values, which are the decision variables, over 100 generations.

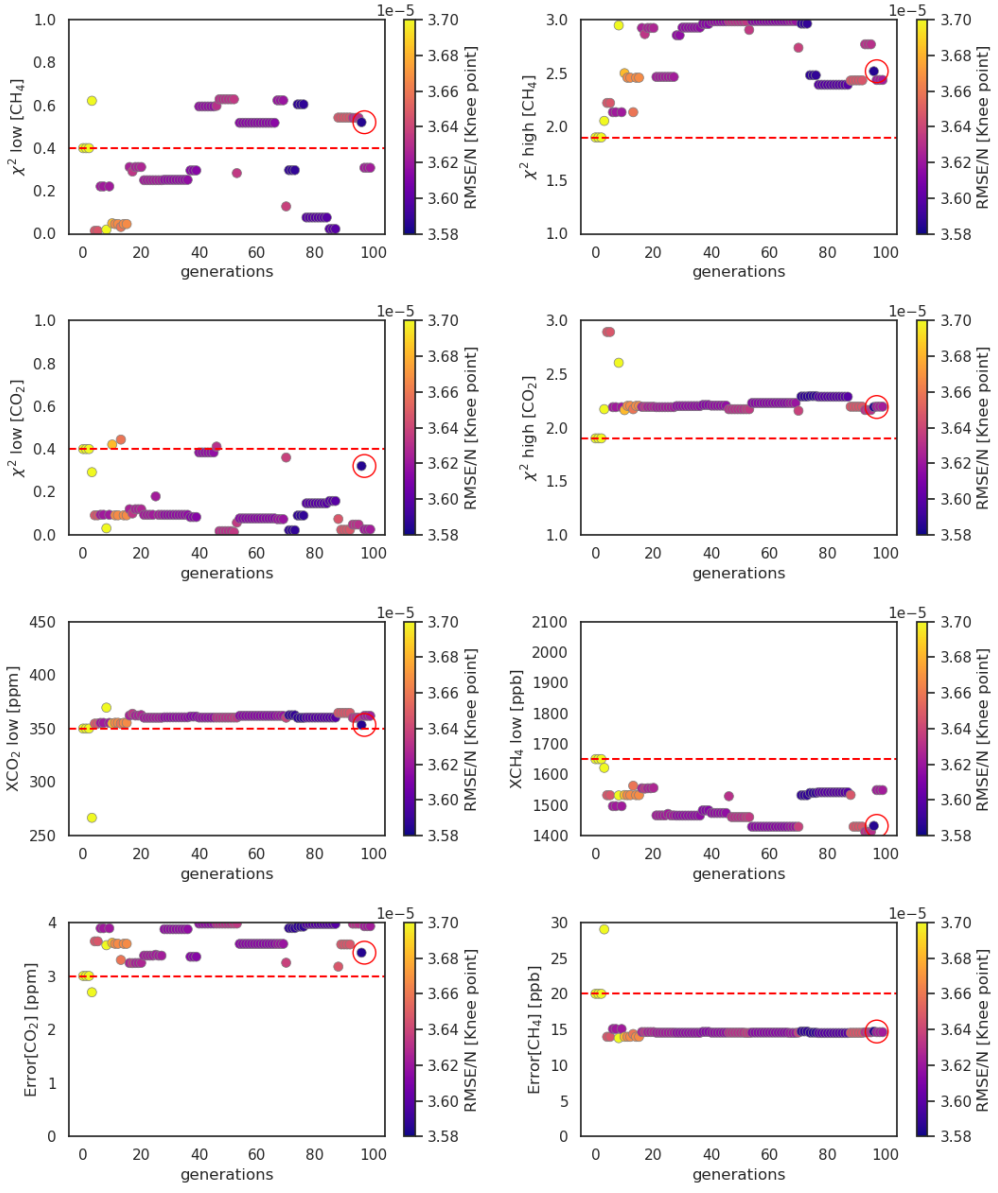

Figure 6. Evaluation of decision variables (filter values as given in Table 1) with generations. Red dotted line shows the initial standard filters. The red circle shows the selected optimised filter values correspond to the generation of minimum of the ratio of RMSE to N.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

Even though the objective variables converge systematically (Fig. 5a), the decision variables show instability over generations as NSGA II focuses on improving the objective variables and not on the decision variables. Since our problem involves multiple (8) filter values, multiple combinations can lead to the same outcome. However, all of the filter combinations corresponding to the convergence zone of the objective variables (Fig. 5), are equally good and can provide the same objective outcome. This is one of the challenges with multi-objective optimisations since they provide multiple optimal solutions. In the present analysis optimisation converges quite well from around the 15th generation. Hence, to select the best suited filter combination for our problem, we have selected the minimum of the ratio of RMSE to N in the convergence zone (from 20 to 100th generation). We assume that this filter combination can provide the best outcome of high number observations and low RMSE. The filter values corresponding to the minimum RMSE to N in the convergence zone are given in Table 3.

Table 3. Comparison between standard filter values and GA-optimised filter values.

Filters	$\chi^2_{\text{low}}[CH_4]$	$\chi^2_{\text{high}}[CH_4]$	$\chi^2_{\text{low}}[CO_2]$	$\chi^2_{\text{high}}[CO_2]$	$XCH_{4,\text{low}}$	$XCO_{2,\text{low}}$	$XCH_{4,\text{error}}$	$XCO_{2,\text{error}}$
Standard	0.40	1.90	0.40	1.90	350.00	1650.00	20.00	3.00
GA	0.52	2.52	0.32	2.19	353.38	1431.65	14.71	3.43

4.3 Enhancement in data density

Figures 7a and 7b show the histograms of GOSAT XCH4 [2009 - 2023] applied with GA-modified quality filters (given in Table 2) along with standard/initial filter values; both globally and for high latitudes ($> 50^\circ N$). On applying the GA optimised filters, the total observation count (N_{GA}) increases by 20 % over the high latitudes and by 5 % globally. The histograms of GOSAT XCH_4 using the initial filters and GA-optimised filters are nearly identical in shape, spread and central tendency, indicating the data density has enhanced without altering the statistics and climatology of the dataset. GA optimisation has enhanced data density globally except in the equatorial tropics where data density remained the same, with no regional/localised data loss (Fig. 7c). Enhancement is greater over the northern hemisphere land-masses especially over Eurasia and North America, with no spurious spikes. The optimisation yielded more data ($\Delta N = N_{\text{GA}} - N_{\text{initial}}$) over high latitudes and mid-latitudes, compared to the tropics. This regional difference might be due to the optimisation process being centred on ground-based TCCON observations from high latitudes.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

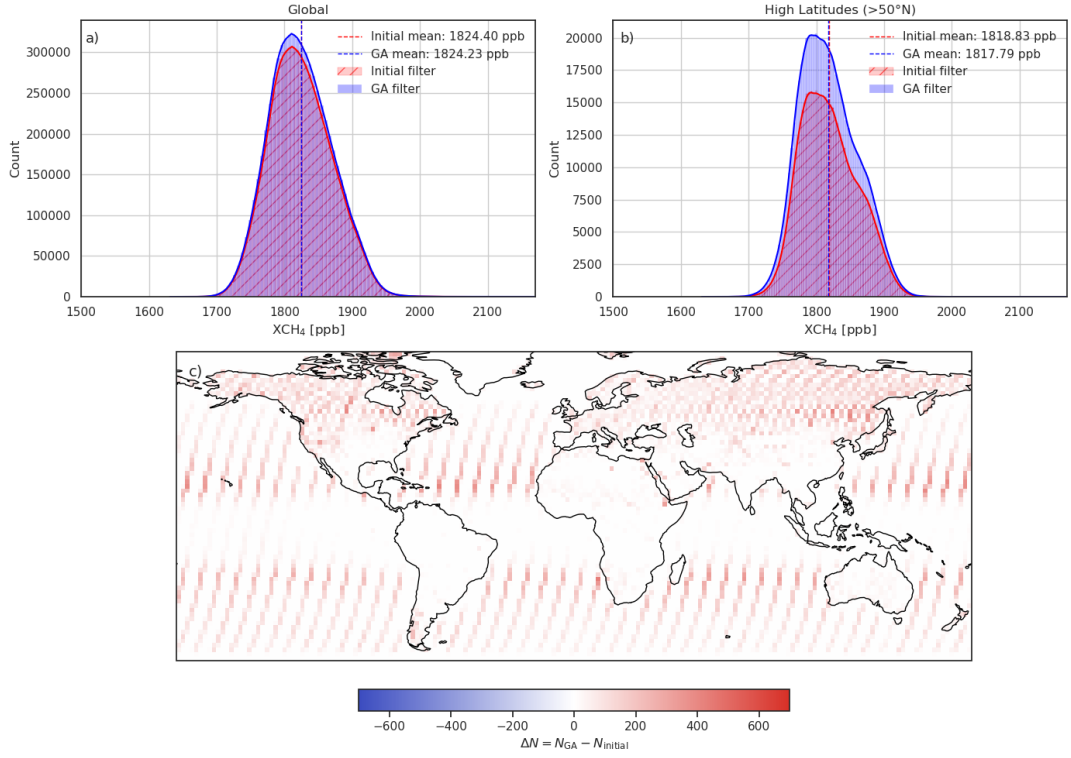

Figure 7. Frequency distribution of GOSAT Proxy XCH_4 [2009 - 2023] on applying standared filters and GA optimised filters a) globally, b) over high latitudes (> 50˚N) and, c) global distribution of change in observation count ($DeltaN$) upon optimisation at 2° x 2° grid resolution.

Data acquisition over high latitudes during the winter season is challenging due to extended periods of low sun over the horizon or the total absence of solar illumination (polar night). Also the presence of frequent and widespread cloud cover, that covers the Arctic more than 50 % throughout the year causes large gaps in atmospheric and surface data sets (Eastman and Warren, 2010; Marshall et al., 1993). Over snow, even though incoming SWIR signals are strongly absorbed, it creates
265 powerful scattering at specific angles due to its non-Lambertian nature leading to errors in the retrieval (Mikkonen et al., 2024). Altogether these factors contribute to severe reduction in atmospheric methane data density during high-latitude winter and make it much harder to quantify the methane emission flux from the seasonal freeze-thaw processes of highly heterogeneous permafrost landscapes (Bartsch et al., 2023). Even though the thawing emissions are mainly confined to the summer season, large bursts of methane emissions are observed in early winter due to methane production in the unfrozen near $0°C$ sub-
270 surface soil layer (zero-curtain layer) beneath the frozen top active layer which can persists for weeks to months (Mastepanov et al., 2008). Wintertime methane emissions over the Arctic tundra are found to contribute to more than 50 % of the annual flux, controlling the Arctic tundra methane budget (Zona et al., 2016), yet in-situ measurements are sparse during this season leaving a large observational data gap (Miner et al., 2022). To assess the performance of the GA optimisation in mitigating the winter deficit of satellite data, we show the enhancement in data density in different seasons in Fig. 8.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

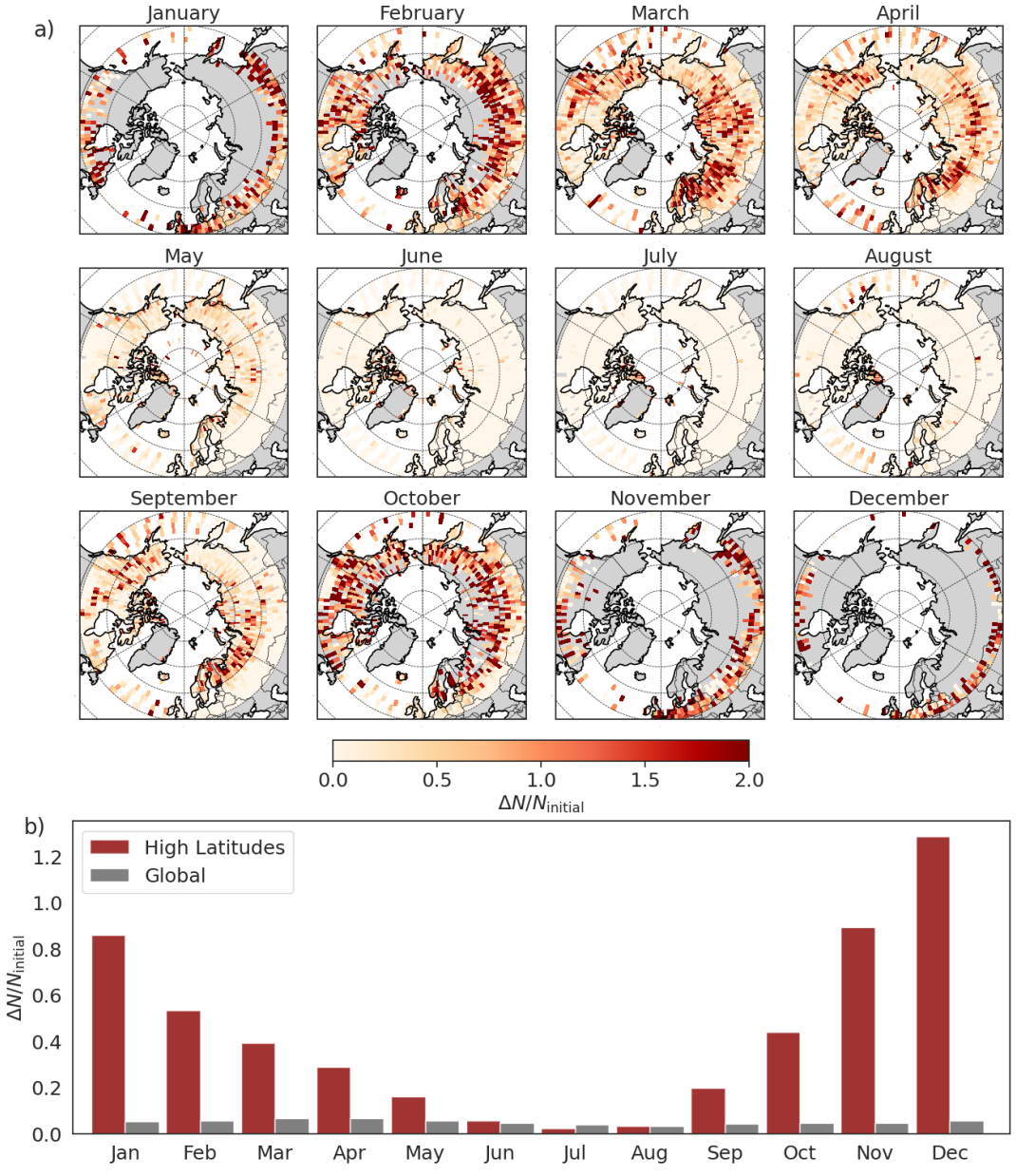

Figure 8. Fractional increase in GOSAT XCH_4 [2009 - 2023] observation count on optimisation - a) Spatial distribution in enhancement at 2° x 2° grid resolution over the high latitudes, grid cells with $N_{\text{initial}} = 0$, are excluded from the estimation of fractional increase calculations and, b) total increase above 50°N during different months.

Figure 8a shows the spatial distribution of fractional enhancement in the optimised GOSAT XCH_4 [2009 - 2023] data over the high latitude regions at $2° \times 2°$ grid resolution during different months. A fractional enhancement of more than 1 shows that the data count has more than doubled across many pixels during most months except in the summer period. Low summertime

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

enhancement indicates that initial filters likely provide optimal performance during summer when data acquisition is most efficient due to favourable sky conditions. Optimised filtering provides the highest enhancement during the winter months, especially in December when the total data count above $50°N$ has increased to more than double (Fig. 8b). Globally, the total optimised data count shows only a minimal fractional increase of less than 0.1 (10 %) almost all months while over high latitudes it goes above 1 (100 %) during winter months, showing that the optimisation has maximum impact over the winter high latitudes and minimum impact over the global data. Thus, the GA optimisation algorithm is capable of selectively targeting the gap regions and periods of low data acquisition. It is to be noted that fractional enhancement is calculated relative to the initial data density (baseline data). Thus, grid cells with zero initial observations ($N_{\text{initial}} = 0$) were excluded from Fig. 8a even though they have new data points upon optimisation. They are considered as newly sampled regions and have been shown in Fig. A1 in appendix A.

4.4 Validation with ground-based TCCON measurements

To validate the optimised GOSAT XCH_4 data with ground-based TCCON measurements, GOSAT soundings collocated within a spatial window of $\pm$ 5° and within 2 hours around 6 TCCON sites over high-latitude region are used including Bremen (Notholt et al., 2022), East Trout Lake (Wunch et al., 2022), Harwell (Weidmann et al., 2023), Ny-Ålesund (Buschmann et al., 2022), Sodankylä (Kivi et al., 2022) and Eureka (Strong et al., 2022). Locations of the high-latitude TCCON stations used in this study are shown in Fig. 2a. and 15 more stations are used for global validation, including: Burgos, Garmisch, Karlsruhe, Lamont, Lauder, Orleans, Paris, Park Falls, Réunion, Rikubetsu, Saga, Tsukuba, Białystok, Darwin and Wollongong (Deutscher et al., 2023b, a; Hase et al., 2024; MaziÃ¨re et al., 2022; Morino et al., 2022a, c, b; Pollard et al., 2022; Sherlock et al., 2022a, b; Shiomi et al., 2022; Sussmann et al., 2025; TÃ© et al., 2022; Warneke et al., 2024; Wennberg et al., 2022, 2025). Figure 9 shows scatter plots of collocated GA optimised XCH_4 data and ground-based TCCON data with mean GOSAT-TCCON difference (Δ), the standard deviation of the GOSAT-TCCON difference (σ) or single-measurement precision, the correlation coefficient (R) and the total number of GOSAT-TCCON pairs (N).

The offset is calculated as the GOSAT-TCCON average difference across the observed pairs (Δ) and remains almost the same after optimisation, with a small increase of 0.44 ppb over high latitudes and 0.1 ppb globally. The correlation coefficient, R has slightly decreased from 0.89 to 0.88 over high latitudes and from 0.94 to 0.93 globally showing minor increase in scatter and more variable data points. The use of optimised filters has increased the number of GOSAT-TCCON data pairs over high latitudes by 19 % and by 5.8 % globally, consistent with the overall increase in data count. The single measurement precision, σ has increased by 1.06 ppb for high latitudes and by 0.58 ppb globally. Thus, the optimisation has achieved an enhancement in high-latitude winter data coverage in GOSAT XCH_4 without notable bias and without compromising precision relative to TCCON. It is to be noted that the dataset still preserves the 'breakthrough' (B) requirement of remaining better than 17 ppb for the single-measurement precision as per ESA GHG-CCI User Requirements Document (Buchwitz et al., 2017).

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

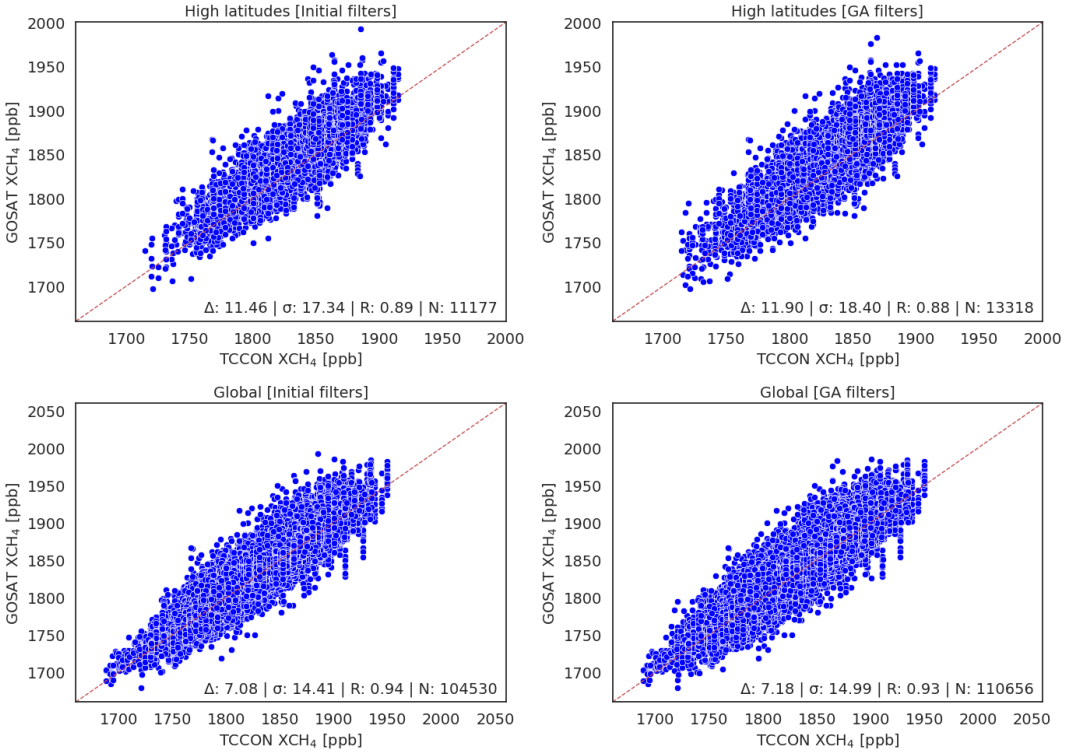

Figure 9. Comparison of GOSAT vs TCCON XCH_4 on applying initial/standard filters and GA optimised filters over High latitudes (upper panels) and globally (lower panels).

Table 4 shows the changes in the optimised GOSAT XCH4 data metrics, such as total data count, RMSE between GOSAT and TCCON collocated observations and single-measurement precision compared to initial values. Thus, the GA optimisation has improved the high-latitude data density and coverage, without altering the overall global data significantly, thus maintaining the robustness of the global dataset.

Table 4. Summary of GOSAT XCH_4 data metrics after applying GA-optimized filters. HL refers to high latitudes ($> 50^\circ$ N).

Parameter	Initial filters	GA filters	Difference	Change (%)
N [Global]	6,259,719	6,571,028	311,309	$+5\%$
N [HL]	509,077	610,723	151,412	$+20\%$
RMSE [Global] (ppb)	16.06	16.62	0.56	$+3.5\%$
RMSE [HL] (ppb)	20.79	21.91	1.12	$+5.4\%$
Precision [Global]	14.41	14.99	0.58	$+4\%$
Precision [HL]	17.34	18.40	1.06	$+6\%$

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

5 Conclusions

Satellite remote sensing of GHGs is challenging over high latitudes (here defined as >50° latitude) due to persistent cloud cover and low solar illumination. This results in poor spatial coverage and lack of sufficient data density especially during the winter season. This sparsity restricts accurate analysis of the spatio-temporal variability and long-term regional trends in atmospheric methane, especially over the Arctic where methane-induced warming and associated climate feedbacks are very relevant. Methane emissions from vastly heterogenous Arctic permafrost and wetlands need to be quantified with continuous data year-round with sufficient spatial coverage. It is critical for top-down emission inversions, model validation and for keeping track of anomalous emissions, growth rates and permafrost feedbacks. Thus, enhancing satellite wintertime methane emissions is crucial which is unfortunately challenging due to natural constrains for satellite measurements especially during winter.

This study successfully improves data throughput of the University of Leicester GOSAT Proxy column methane data in the high-latitude regions using a genetic algorithm approach to optimise the post-retrieval quality-filters, recovering a significant amount of data which is otherwise discarded when using standard restrictive filter values. Thus, the objective of the study is to enhance the observation density and spatial coverage with the least impact on data quality when validated against ground-based observations at TCCON stations. However, relaxing quality filters naturally leads to a worsening of data quality due to the inclusion of marginal quality data. To realise these conflicting objectives, we have adopted a non-dominated sorting genetic algorithm approach, ideal for analysing trade-offs between multiple conflicting objectives to optimise the QA filtering of the UoL GOSAT XCH4 data product. We have used ground-based TCCON methane observations to drive the algorithm towards an optimal solution that ensures minimal impact on data quality.

Upon optimisation, it is observed that GA-optimised filters can provide a 20 % increase in the number of satellite soundings over high latitudes and an increase of 5 % globally, with a minimal compromise to data quality. The optimised GOSAT XCH_4 dataset has been validated against ground-based TCCON observations, showing that use of optimised filters causes negligible changes to the offset bias and correlation whilst increasing the number of observations. Single measurement precision of the global dataset has increased (i.e., worsened) by 1.06 ppb for high-latitude data and 0.58 ppb for the global data; well within the validity of the dataset. The optimised data set retains the statistical distribution of the original data set and improves data density globally, with a specific focus on optimisation over the high latitudes. Ground-based observations over high-latitude regions are sparse and studies like ours would be greatly benefitted from initiatives like GEMINI-UK, which aims to establish a dense network of GHG observatories across UK as a part of Greenhouse Gas Emissions Measurement and Modelling Advancement (GEMMA) framework (Humpage et al., 2025; Kurganskiy et al., 2025). This GA optimisation algorithm benefits regions and seasons with historically poor coverage while minimally impacting data quality without data loss elsewhere, and is suitable for selectively targeting regional and temporal gaps in data for data gain by relaxing the filters. At the same time, it avoids compromising the data quality by not adding data unnecessarily or spuriously, maintaining the quality of the global dataset. The study demonstrates the potential of genetic algorithms to develop improved atmospheric datasets from upcoming missions like GOSAT-GW, MicroCarb and CO2M and would support inverse modelling and better emission flux estimates especially over high latitudes.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

Data availability. The University of Leicester GOSAT Proxy v9.0 XCH_4 data are available from the Centre for Environmental Data Analysis data repository at https://catalogue.ceda.ac.uk/uuid/18ef8247f52a4cb6a14013f8235cc1eb. The TCCON data were obtained from the TCCON Data Archive hosted by CaltechDATA at https://doi.org/10.14291/TCCON.GGG2020

Appendix A: A

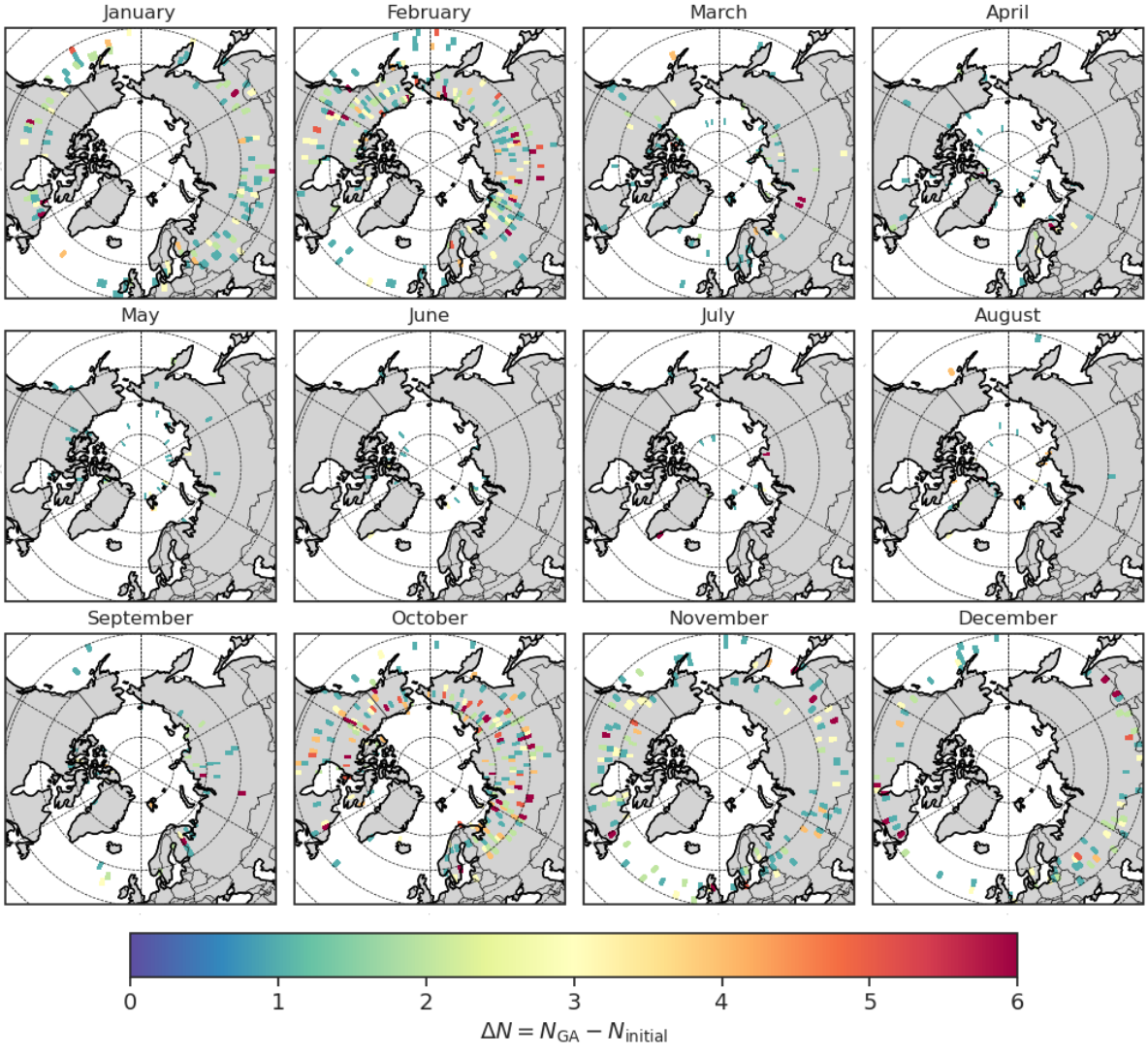

Figure A1. Newly sampled areas upon optimisation at 2˚ x 2˚ grid resolution with colour variation showing the absolute increase in observation count (ΔN). These are the grid cells where $N_{initial} = 0$ and $N_{GA} > 0$

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

Author contributions. LNB, developed the methodology, performed the analysis, and wrote the original manuscript. RJP developed the GOSAT Proxy XCH_4 data, conceived the research idea and provided supervision. MC,DO,AND, PS,AW, and HB all contributed to development and analysis at different stages of the processing chain of GOSAT Proxy XCH_4 data. All authors contributed towards revision of the final manuscript.

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

Acknowledgements. This work was supported by the Natural Environment Research Council through the UK EO Climate Information Service (NE/X019071/1) and the Greenhouse Gas Emissions Measurement and Modelling Advancement (GEMMA) Programme (NE/Y00177X/1). RJP acknowledges funding from the UK National Centre for Earth Observation (Grant: NE/W004895/1) and a UKRI Future Leaders Fellowship (Grant: MR/X033139/1). We acknowledge the NASA-ESA Arctic Methane and Permafrost Challenge (AMPAC) project. 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. We acknowledge the TCCON PIs maintaining the sites and the TCCON data managers for the centralized data archiving. This research used the ALICE High Performance Computing Facility at the University of Leicester for the analysis.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

References

Long-term vicarious calibration of GOSAT short-wave sensors: Techniques for error reduction and new estimates of radiometric degradation factors, IEEE Transactions on Geoscience and Remote Sensing, 52, https://doi.org/10.1109/TGRS.2013.2278696, 2014.

Akbarizadeh, G.: A New Recognition Approach Based on Genetic Algorithm for Classifying Textures in Satellite SAR Images, International Journal of Remote Sensing Applications, 2, 2012.

Bartsch, A., Strozzi, T., and Nitze, I.: Permafrost Monitoring from Space, https://doi.org/10.1007/s10712-023-09770-3, 2023.

Bekryaev, R. V., Polyakov, I. V., and Alexeev, V. A.: Role of polar amplification in long-term surface air temperature variations and modern arctic warming, Journal of Climate, 23, https://doi.org/10.1175/2010JCLI3297.1, 2010.

Blank, J. and Deb, K.: Pymoo: Multi-Objective Optimization in Python, IEEE Access, 8, 89 497–89 509, https://doi.org/10.1109/ACCESS.2020.2990567, 2020.

Buchwitz, M., Reuter, M., Schneising, O., Hewson, W., Detmers, R. G., Boesch, H., Hasekamp, O. P., Aben, I., Bovensmann, H., Burrows, J. P., Butz, A., Chevallier, F., Dils, B., Frankenberg, C., Heymann, J., Lichtenberg, G., Mazière, M. D., Notholt, J., Parker, R., Warneke, T., Zehner, C., Griffith, D. W., Deutscher, N. M., Kuze, A., Suto, H., and Wunch, D.: Global satellite observations of column-averaged carbon dioxide and methane: The GHG-CCI XCO2 and XCH4 CRDP3 data set, Remote Sensing of Environment, 203, https://doi.org/10.1016/j.rse.2016.12.027, 2017.

Buchwitz, M., Reuter, M., Schneising, O., Bovensmann, H., Burrows, J. P., Boesch, H., Anand, J., Parker, R., Detmers, R. G., Aben, I., Hasekamp, O. P., Crevoisier, C., Armante, R., Zehner, C., and Schepers, D.: Copernicus Climate Change Service (C3S) global satellite observations of atmospheric carbon dioxide and methane, Proceedings of the International Astronautical Congress, IAC, 2018-October, https://researchportal.ip-paris.fr/en/publications/copernicus-climate-change-service-c3s-global-satellite-observatio/, 2018.

Buschmann, M., Petri, C., Palm, M., Warneke, T., and Notholt, J.: TCCON data from Ny-Ã. . . lesund, Svalbard (NO), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.NYALESUND01.R0, 2022.

Celik, T.: Change detection in satellite images using a genetic algorithm approach, IEEE Geoscience and Remote Sensing Letters, 7, https://doi.org/10.1109/LGRS.2009.2037024, 2010.

Circiu, M. S. and Leon, F.: Comparative Study of Multiobjective Genetic Algorithms, Bulletin of the Polytechnic Institute of Iasi, tome LVI (LX), section Automatic Control and Computer Science, 2010.

Darwin, C.: The Origin of Species: By Means of Natural Selection Or the Preservation of Favored Races in the Struggle for Life, vol. 2, Modern Library, 1872.

Das, I. and Dennis, J. E.: A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems, Structural Optimization, 14, https://doi.org/10.1007/BF01197559, 1997.

Deb, K.: Multi-objective Optimisation Using Evolutionary Algorithms: An Introduction, https://doi.org/10.1007/978-0-85729-652-8_1, 2011.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Transactions on Evolutionary Computation, 6, 182–197, https://doi.org/10.1109/4235.996017, 2002.

Deutscher, N. M., Griffith, D. W. T., Paton-Walsh, C., Jones, N. B., Velazco, V. A., Wilson, S. R., Macatangay, R. C., Kettlewell, G. C., Buchholz, R. R., Riggenbach, M. O., Bukosa, B., John, S. S., Walker, B. T., and Nawaz, H.: TCCON data from Wollongong (AU), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.WOLLONGONG01.R0, 2023a.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

Deutscher, N. M., Griffith, D. W. T., Paton-Walsh, C., Velazco, V. A., Wennberg, P. O., Blavier, J.-F., Washenfelder, R. A., Yavin, Y., Keppel-Aleks, G., Toon, G. C., Jones, N. B., Kettlewell, G. C., Connor, B. J., Macatangay, R. C., Wunch, D., Roehl, C., and Bryant, G. W.: TCCON data from Darwin (AU), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.DARWIN01.R0, 2023b.

Eastman, R. and Warren, S. G.: Arctic Cloud Changes from Surface and Satellite Observations, Journal of Climate, 23, 4233–4242, https://doi.org/10.1175/2010JCLI3544.1, 2010.

Goldberg, D. E.: Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Reading, MA, 1989, NN Schraudolph and J.., 3, 1989.

Hase, F., Herkommer, B., GroÃŸ, J., Blumenstock, T., Kiel, M., and Dohe, S.: TCCON data from Karlsruhe (DE), Release GGG2020.R2, https://doi.org/10.14291/TCCON.GGG2020.KARLSRUHE01.R2, 2024.

Holland, J. H.: Genetic algorithms, Scientific American, 267, 66–72, https://doi.org/10.1038/SCIENTIFICAMERICAN0792-66, 1992.

Hu, H., Landgraf, J., Detmers, R., Borsdorff, T., de Brugh, J. A., Aben, I., Butz, A., and Hasekamp, O.: Toward Global Mapping of Methane With TROPOMI: First Results and Intersatellite Comparison to GOSAT, Geophysical Research Letters, 45, https://doi.org/10.1002/2018GL077259, 2018.

Humpage, N., Palmer, P., Kurganskiy, A., Feng, L., Woodwark, J., Doniki, S., and Weidmann, D.: First data from GEMINI-UK, the UK national network of ground-based greenhouse gas observing spectrometers, EGU25, https://doi.org/10.5194/EGUSPHERE-EGU25-16209, 2025.

IPCC: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Core writing team, h. lee and j. romero (eds.), Intergovernmental Panel on Climate Change, Geneva, Switzerland, https://doi.org/10.59327/IPCC/AR6-9789291691647, 2023.

Jacobs, N., Simpson, W. R., Wunch, D., O'Dell, C. W., Osterman, G. B., Hase, F., Blumenstock, T., Tu, Q., Frey, M., Dubey, M. K., Parker, H. A., Kivi, R., and Heikkinen, P.: Quality controls, bias, and seasonality of CO2columns in the boreal forest with Orbiting Carbon Observatory-2, Total Carbon Column Observing Network, and EM27/SUN measurements, Atmospheric Measurement Techniques, 13, https://doi.org/10.5194/amt-13-5033-2020, 2020.

Jannati, M. and Zoej, M. J. V.: Introducing genetic modification concept to optimize rational function models (RFMs) for georeferencing of satellite imagery, GIScience and Remote Sensing, 52, https://doi.org/10.1080/15481603.2015.1052634, 2015.

Karami, F. and Dariane, A. B.: A review and evaluation of multi and many-objective optimization: Methods and algorithms, Global Journal of Ecology, 7, 2022.

Katoch, S., Chauhan, S. S., and Kumar, V.: A review on genetic algorithm: past, present, and future, Multimedia Tools and Applications, 80, https://doi.org/10.1007/s11042-020-10139-6, 2021.

Khodaverdizahraee, N., Rastiveis, H., and Jouybari, A.: Segment-by-segment comparison technique for earthquake-induced building damage map generation using satellite imagery, International Journal of Disaster Risk Reduction, 46, https://doi.org/10.1016/j.ijdrr.2020.101505, 2020.

Kivi, R., Heikkinen, P., and KyrÃ¶, E.: TCCON data from SodankylÃ¤ (FI), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.SODANKYLA01.R0, 2022.

Knowles, J. and Corne, D.: The Pareto archived evolution strategy: A new baseline algorithm for Pareto multiobjective optimisation, in: Proceedings of the 1999 Congress on Evolutionary Computation, CEC 1999, vol. 1, https://doi.org/10.1109/CEC.1999.781913, 1999.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

Kurganskiy, A., Feng, L., Humpage, N., Palmer, P. I., Woodwark, A. J. P., Doniki, S., and Weidmann, D.: The Greenhouse gas Emission Monitoring network to Inform Net-zero Initiatives UK (GEMINI-UK): network design, theoretical performance, and initial data, https://doi.org/10.5194/egusphere-2025-94, 2025.

Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, Applied Optics, 48, https://doi.org/10.1364/AO.48.006716, 2009.

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, Atmospheric Measurement Techniques, 9, https://doi.org/10.5194/amt-9-2445-2016, 2016.

Laughner, J. L., Toon, G. C., Mendonca, J., Petri, C., Roche, S., Wunch, D., Blavier, J. F., Griffith, D. W., Heikkinen, P., Keeling, R. F., Kiel, M., Kivi, R., Roehl, C. M., Stephens, B. B., Baier, B. C., Chen, H., Choi, Y., Deutscher, N. M., Digangi, J. P., Gross, J., Herkommer, B., Jeseck, P., Laemmel, T., Lan, X., McGee, E., McKain, K., Miller, J., Morino, I., Notholt, J., Ohyama, H., Pollard, D. F., Rettinger, M., Riris, H., Rousogenous, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Wofsy, S. C., Zhou, M., and Wennberg, P. O.: The Total Carbon Column Observing Network's GGG2020 data version, Earth System Science Data, 16, https://doi.org/10.5194/essd-16-2197-2024, 2024.

Lenton, T. M.: Arctic climate tipping points, https://doi.org/10.1007/s13280-011-0221-x, 2012.

Leung, H., Dubash, N., and Xie, N.: Detection of small objects in clutter using a GA-RBF neural network, IEEE Transactions on Aerospace and Electronic Systems, 38, 98–118, https://doi.org/10.1109/7.993232, 2002.

Marshall, G., Rees, W., and Dowdeswell, J.: Limitations imposed by cloud cover on multitemporal visible band satellite data sets from polar regions, Annals of Glaciology, 17, 113–120, https://doi.org/10.3189/S0260305500012696, 1993.

Mastepanov, M., Sigsgaard, C., Dlugokencky, E. J., Houweling, S., Ström, L., Tamstorf, M. P., and Christensen, T. R.: Large tundra methane burst during onset of freezing, Nature, 456, https://doi.org/10.1038/nature07464, 2008.

MaziÃ¨re, M. D., 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 RÃ©union Island (RE), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.REUNION01.R0, 2022.

McGuire, A. D., Christensen, T. R., Hayes, D., Heroult, A., Euskirchen, E., Kimball, J. S., Koven, C., Lafleur, P., Miller, P. A., Oechel, W., Peylin, P., Williams, M., and Yi, Y.: An assessment of the carbon balance of Arctic tundra: Comparisons among observations, process models, and atmospheric inversions, Biogeosciences, 9, https://doi.org/10.5194/bg-9-3185-2012, 2012.

McKay, D. I., Staal, A., Abrams, J. F., Winkelmann, R., Sakschewski, B., Loriani, S., Fetzer, I., Cornell, S. E., Rockström, J., and Lenton, T. M.: Exceeding 1.5°C global warming could trigger multiple climate tipping points, Science, 377, https://doi.org/10.1126/science.abn7950, 2022.

Mendonca, J., Nassar, R., O'dell, C. W., Kivi, R., Morino, I., Notholt, J., Petri, C., Strong, K., and Wunch, D.: Assessing the feasibility of using a neural network to filter Orbiting Carbon Observatory 2 (OCO-2) retrievals at northern high latitudes, Atmospheric Measurement Techniques, 14, https://doi.org/10.5194/amt-14-7511-2021, 2021.

Mikkonen, A., Lindqvist, H., Peltoniemi, J., and Tamminen, J.: Non-Lambertian snow surface reflection models for simulated top-of-the-atmosphere radiances in the NIR and SWIR wavelengths, Journal of Quantitative Spectroscopy and Radiative Transfer, 315, 108 892, https://doi.org/10.1016/J.JQSRT.2023.108892, 2024.

Miner, K. R., Turetsky, M. R., Malina, E., Bartsch, A., Tamminen, J., McGuire, A. D., Fix, A., Sweeney, C., Elder, C. D., and Miller, C. E.: Permafrost carbon emissions in a changing Arctic, https://doi.org/10.1038/s43017-021-00230-3, 2022.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

Morino, I., Ohyama, H., Hori, A., and Ikegami, H.: TCCON data from Rikubetsu (JP), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.RIKUBETSU01.R0, 2022a.

Morino, I., Ohyama, H., Hori, A., and Ikegami, H.: TCCON data from Tsukuba (JP), 125HR, Release GGG2020.R0, 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, Ilocos Norte (PH), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.BURGOS01.R0, 2022c.

Nocedal, J. and Wright, S. J.: Numerical optimization, Springer, 2006.

Notholt, J., Petri, C., Warneke, T., and Buschmann, M.: TCCON data from Bremen (DE), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.BREMEN01.R0, 2022.

Noël, S., Reuter, M., Buchwitz, M., Borchardt, J., Hilker, M., Schneising, O., Bovensmann, H., Burrows, J. P., Noia, A. D., Parker, R. J., Suto, H., Yoshida, Y., Buschmann, M., Deutscher, N. M., Feist, D. G., Griffith, D. W., Hase, F., Kivi, R., Liu, C., Morino, I., Notholt, J., Oh, Y. S., Ohyama, H., Petri, C., Pollard, D. F., Rettinger, M., Roehl, C., Rousogenous, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Vrekoussis, M., and Warneke, T.: Retrieval of greenhouse gases from GOSAT and GOSAT-2 using the FOCAL algorithm, Atmospheric Measurement Techniques, 15, https://doi.org/10.5194/amt-15-3401-2022, 2022.

Pallandt, M. M., Kumar, J., Mauritz, M., Schuur, E. A., Virkkala, A. M., Celis, G., Hoffman, F. M., and Göckede, M.: Representativeness assessment of the pan-Arctic eddy covariance site network and optimized future enhancements, Biogeosciences, 19, https://doi.org/10.5194/bg-19-559-2022, 2022.

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, Geophysical Research Letters, 38, https://doi.org/10.1029/2011GL047871, 2011.

Parker, R. J., Boesch, H., Byckling, K., Webb, A. J., Palmer, P. I., Feng, L., Bergamaschi, P., Chevallier, F., Notholt, J., Deutscher, N., Warneke, T., Hase, F., Sussmann, R., Kawakami, S., Kivi, R., Griffith, D. W., and Velazco, V.: Assessing 5 years of GOSAT Proxy XCH4 data and associated uncertainties, Atmospheric Measurement Techniques, 8, https://doi.org/10.5194/amt-8-4785-2015, 2015.

Parker, R. J., Webb, A., Boesch, H., Somkuti, P., Guillo, R. B., Noia, A. D., 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 CH_4 observations, Earth System Science Data, 12, 3383–3412, https://doi.org/10.5194/essd-12-3383-2020, 2020.

Pati, C., Panda, A. K., Tripathy, A. K., Pradhan, S. K., and Patnaik, S.: A novel hybrid machine learning approach for change detection in remote sensing images, Engineering Science and Technology, an International Journal, 23, 973–981, https://doi.org/10.1016/j.jestch.2020.01.002, 2020.

Pollard, D. F., Robinson, J., and Shiona, H.: TCCON data from Lauder (NZ), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.LAUDER03.R0, 2022.

Qin, C., Gao, Y., and Wang, Y.: The optimization of low Earth orbit satellite constellation visibility with genetic algorithm for improved navigation potential, Scientific Reports, 15, 1–22, https://doi.org/10.1038/S41598-025-16815-7;SUBJMETA, 2025.

Ranjani, J. J.: A Study on Intelligent Algorithms for Change Detection using Remote Sensing Images, Int. J. Signal Process., 2, http://www.iaras.org/iaras/filedownloads/ijsp/2017/003-0020(2017).pdf, 2017.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

Rios, L. M. and Sahinidis, N. V.: Derivative-free optimization: A review of algorithms and comparison of software implementations, in: Journal of Global Optimization, vol. 56, ISSN 09255001, https://doi.org/10.1007/s10898-012-9951-y, 2013.

Schuur, E. A., McGuire, A. D., Schädel, C., Grosse, G., Harden, J. W., Hayes, D. J., Hugelius, G., Koven, C. D., Kuhry, P., Lawrence, D. M., Natali, S. M., Olefeldt, D., Romanovsky, V. E., Schaefer, K., Turetsky, M. R., Treat, C. C., and Vonk, J. E.: Climate change and the permafrost carbon feedback, https://doi.org/10.1038/nature14338, 2015.

Sherlock, V., Connor, B., Robinson, J., Shiona, H., Smale, D., and Pollard, D. F.: TCCON data from Lauder (NZ), 120HR, Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.LAUDER01.R0, 2022a.

Sherlock, V., Connor, B., Robinson, J., Shiona, H., Smale, D., and Pollard, D. F.: TCCON data from Lauder (NZ), 125HR, Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.LAUDER02.R0, 2022b.

Shiomi, K., Kawakami, S., Ohyama, H., Arai, K., Okumura, H., Ikegami, H., and Usami, M.: TCCON data from Saga (JP), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.SAGA01.R0, 2022.

Slimani, Y. and Hedjam, R.: A Hybrid Metaheuristic and Deep Learning Approach for Change Detection in Remote Sensing Data, Engineering, Technology and Applied Science Research, 12, 9351–9356, https://doi.org/10.48084/etasr.5246, 2022.

Smith, S. L., O'Neill, H. B., Isaksen, K., Noetzli, J., and Romanovsky, V. E.: The changing thermal state of permafrost, https://doi.org/10.1038/s43017-021-00240-1, 2022.

Srinivas, N. and Deb, K.: Muiltiobjective Optimization Using Nondominated Sorting in Genetic Algorithms, Evolutionary Computation, 2, https://doi.org/10.1162/evco.1994.2.3.221, 1994.

Strong, K., Roche, S., Franklin, J. E., Mendonca, J., Lutsch, E., Weaver, D., Fogal, P. F., Drummond, J. R., Batchelor, R., Lindenmaier, R., and McGee, E.: TCCON data from Eureka (CA), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.EUREKA01.R0, 2022.

Sumer, E. and Turker, M.: An adaptive fuzzy-genetic algorithm approach for building detection using high-resolution satellite images, Computers, Environment and Urban Systems, 39, https://doi.org/10.1016/j.compenvurbsys.2013.01.004, 2013.

Sussmann, R., Rettinger, M., and Pak, N. M.: TCCON data from Garmisch (DE), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.GARMISCH01.R1, 2025.

Sánchez-Sevilleja, S., García-Rodríguez, M., Masa-Campos, J. L., and Cuerda-Muñoz, J. M.: Antenna Model with Pattern Optimization Based on Genetic Algorithm for Satellite-Based SAR Mission, Sensors 2025, Vol. 25, Page 4835, 25, 4835, https://doi.org/10.3390/S25154835, 2025.

Tamocai, C., Canadell, J. G., Schuur, E. A., Kuhry, P., Mazhitova, G., and Zimov, S.: Soil organic carbon pools in the northern circumpolar permafrost region, Global Biogeochemical Cycles, 23, https://doi.org/10.1029/2008GB003327, 2009.

TÃ©, Y., Jeseck, P., and Janssen, C.: TCCON data from Paris (FR), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.PARIS01.R0, 2022.

Vavrus, S., Waliser, D., Schweiger, A., and Francis, J.: Simulations of 20th and 21st century Arctic cloud amount in the global climate models assessed in the IPCC AR4, Climate Dynamics, 33, 1099–1115, https://doi.org/10.1007/S00382-008-0475-6/FIGURES/13, 2009.

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sensing of Environment, 120, https://doi.org/10.1016/j.rse.2011.09.027, 2012.

https://doi.org/10.5194/egusphere-2026-1029
Preprint. Discussion started: 16 March 2026

Vrese, P. D., Stacke, T., Kleinen, T., and Brovkin, V.: Diverging responses of high-latitude co2and ch4emissions in idealized climate change scenarios, Cryosphere, 15, 1097–1130, https://doi.org/10.5194/TC-15-1097-2021,, 2021.

Ward, R. H., Sweeney, C., Miller, J. B., Goeckede, M., Laurila, T., Hatakka, J., Ivakov, V., Sasakawa, M., Machida, T., Morimoto, S., Goto, D., and Ganesan, A. L.: Increasing Methane Emissions and Widespread Cold-Season Release From High-Arctic Regions Detected Through Atmospheric Measurements, Journal of Geophysical Research: Atmospheres, 129, https://doi.org/10.1029/2024JD040766, 2024.

Warneke, T., Petri, C., Notholt, J., and Buschmann, M.: TCCON data from OrlÃ©ans (FR), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.ORLEANS01.R1, 2024.

Weidmann, D., Brownsword, R., and Doniki, S.: TCCON data from Harwell, Oxfordshire (UK), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.HARWELL01.R0, 2023.

Wennberg, P. O., Roehl, C. M., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R., Keppel-Aleks, G., and Allen, N. T.: TCCON data from Park Falls (US), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.PARKFALLS01.R1, 2022.

Wennberg, P. O., Wunch, D., Roehl, C. M., Blavier, J.-F., Toon, G. C., and Allen, N. T.: TCCON data from Lamont (US), Release GGG2020.R1, https://doi.org/10.14291/TCCON.GGG2020.LAMONT01.R1, 2025.

Wittig, S., Berchet, A., Pison, I., Saunois, M., Thanwerdas, J., Martinez, A., Paris, J. D., Machida, T., Sasakawa, M., Worthy, D. E., Lan, X., Thompson, R. L., Sollum, E., and Arshinov, M.: Estimating methane emissions in the Arctic nations using surface observations from 2008 to 2019, Atmospheric Chemistry and Physics, 23, https://doi.org/10.5194/acp-23-6457-2023, 2023.

Wunch, D., Mendonca, J., Colebatch, O., Allen, N. T., Blavier, J.-F., Kunz, K., Roche, S., Hedelius, J., Neufeld, G., Springett, S., Worthy, D., Kessler, R., and Strong, K.: TCCON data from East Trout Lake, SK (CA), Release GGG2020.R0, https://doi.org/10.14291/TCCON.GGG2020.EASTTROUTLAKE01.R0, 2022.

Zhou, J., Zhang, X., Zhan, W., and Zhang, H.: Land Surface Temperature Retrieval from MODIS Data by Integrating Regression Models and the Genetic Algorithm in an Arid Region, Remote Sensing 2014, Vol. 6, Pages 5344-5367, 6, 5344–5367, https://doi.org/10.3390/RS6065344, 2014.

Zona, D., Gioli, B., Commane, R., Lindaas, J., Wofsy, S. C., Miller, C. E., Dinardo, S. J., Dengel, S., Sweeney, C., Karion, A., Chang, R. Y., Henderson, J. M., Murphy, P. C., Goodrich, J. P., Moreaux, V., Liljedahl, A., Watts, J. D., Kimball, J. S., Lipson, D. A., and Oechel, W. C.: Cold season emissions dominate the Arctic tundra methane budget, Proceedings of the National Academy of Sciences of the United States of America, 113, https://doi.org/10.1073/pnas.1516017113, 2016.