Atmos. Meas. Tech., 14, 7511–7524, 2021
https://doi.org/10.5194/amt-14-7511-2021

Atmospheric Measurement Techniques
Open Access
EGU

Assessing the feasibility of using a neural network to filter Orbiting Carbon Observatory 2 (OCO-2) retrievals at northern high latitudes

Joseph Mendonca[1], Ray Nassar[1], Christopher W. O'Dell[2], Rigel Kivi[3], Isamu Morino[4], Justus Notholt[5], Christof Petri[5], Kimberly Strong[6], and Debra Wunch[6]

[1]Environment and Climate Change Canada, Toronto, ON, Canada
[2]Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA
[3]Finnish Meteorological Institute, Sodankylä, Finland
[4]National Institute for Environmental Studies (NIES), Tsukuba, Japan
[5]Institute of Environmental Physics, University of Bremen, Bremen, Germany
[6]Department of Physics, University of Toronto, Toronto, ON, Canada

Correspondence: Joseph Mendonca (joe.mendonca@canada.ca)

Received: 19 May 2021 – Discussion started: 19 July 2021
Revised: 6 October 2021 – Accepted: 22 October 2021 – Published: 3 December 2021

Abstract. Satellite retrievals of XCO_2 at northern high latitudes currently have sparser coverage and lower data quality than most other regions of the world. We use a neural network (NN) to filter Orbiting Carbon Observatory 2 (OCO-2) B10 bias-corrected XCO_2 retrievals and compare the quality of the filtered data to the quality of the data filtered with the standard B10 quality control filter. To assess the performance of the NN filter, we use Total Carbon Column Observing Network (TCCON) data at selected northern high latitude sites as a truth proxy. We found that the NN filter decreases the overall bias by 0.25 ppm (∼ 50 %), improves the precision by 0.18 ppm (∼ 12 %), and increases the throughput by 16 % at these sites when compared to the standard B10 quality control filter. Most of the increased throughput was due to an increase in throughput during the spring, fall, and winter seasons. There was a decrease in throughput during the summer, but as a result the bias and precision were improved during the summer months. The main drawback of using the NN filter is that it lets through fewer retrievals at the highest-latitude Arctic TCCON sites compared to the B10 quality control filter, but the lower throughput improves the bias and precision.

1 Introduction

Northern high-latitude regions are undergoing considerable changes related to climate change. The Arctic has seen the annual average temperature increase 3 times more than the global annual average (Stocker et al., 2013). The boreal forest (an important driver of the CO_2 seasonal cycle) has seen its growing season lengthen due to climate change (Pulliainen et al., 2017), with an increase in the frequency and severity of forest fires (Seidl et al., 2017). Permafrost soils of the northern high latitudes are a large carbon reservoir and some fraction of this carbon is vulnerable to release as CO_2 and CH_4 as the climate warms (Schuur et al., 2015). Changes in the carbon cycle will impact the climate, which in turn will impact the carbon cycle. Understanding how the carbon cycle is changing at boreal and Arctic latitudes, including this feedback loop, will be key to predicting future climate change.

In situ atmospheric measurements of CO_2 can be used to study how the carbon cycle is changing. However, cost and logistical challenges present barriers to establishing measurement sites at high northern latitudes, limiting the amount of information available about the carbon cycle in the Arctic and boreal regions. Remote sensing measurements from space can be used to complement coverage to the current in situ networks (Olsen and Randerson, 2004). Current satellite missions such as the Greenhouse Gases Observing Satellite (GOSAT) (Yokota et al., 2009) and the Orbiting Carbon Ob-

servatory 2 (OCO-2) (Crisp et al., 2004) record solar absorption spectra reflected off the Earth's surface, which are used to retrieve column-averaged dry-air mole fractions of CO_2 (XCO_2), giving regional information on atmospheric CO_2. These data can be used to learn about the carbon cycle but require low bias and high precision to be useful (Rayner and O'Brien, 2001).

The density of satellite retrievals of XCO_2 from current missions is limited by the amount of available sunlight and the inability to measure through clouds. At high latitudes there is less sunlight available during the colder seasons, decreasing the number of spectra obtained when compared to the midlatitudes. Furthermore, filtering and bias-correction schemes are optimized for midlatitudes where more validation data sets are available. This has led to a filter that removes a larger fraction of the high-latitude data than data at midlatitudes. Scenes with snow are also filtered out, because they are thought to be problematic for the retrievals, which decreases the throughput during the colder seasons. In order to improve the quality and throughput of retrievals at high latitudes, in this study we focus on using high-latitude validation XCO_2 retrievals to improve the filtering of northern high-latitude OCO-2 bias-corrected XCO_2 retrievals.

The study by Jacobs et al. (2020) showed that when making modifications to the quality control filtering scheme and bias correction used by OCO-2, one can increase the throughput of OCO-2 retrievals (data version B9) (Kiel et al., 2019; O'Dell et al., 2018) in the boreal region. This was done by changing limits on the features used in the quality control scheme created in O'Dell et al. (2018). These changes were validated by comparing OCO-2 XCO_2 retrievals (Kiel et al., 2019; O'Dell et al., 2018) coincident to XCO_2 retrievals from ground-based solar absorption spectra made by remote sensing instruments used by the Total Carbon Column Observing Network (TCCON) (Wunch et al., 2011a).

Machine learning algorithms are useful for pattern recognition in complex data sets. Mandrake et al. (2013) was the first study to demonstrate the use of machine learning (using a genetic algorithm) to filter ACOS-GOSAT retrievals and multiple versions of the OCO-2 retrievals using warn levels. There is potential to apply different machine learning algorithms to the northern high-latitude OCO-2 data set in order to improve the bias, precision, and throughput.

In this study, we investigate the feasibility of using a simple neural network to filter the current OCO-2 data version (B10) (Osterman et al., 2020) XCO_2 retrievals at northern high latitudes. Section 2 outlines the coincidence criteria between OCO-2 and TCCON retrievals and an explanation of how the retrieved XCO_2 is adjusted for different averaging kernels and a priori information when comparing OCO-2 to TCCON. Section 3 describes the architecture of the neural network and how it is trained to filter the OCO-2 bias-corrected XCO_2 retrievals. In Sect. 4, the neural network (NN)-filtered OCO-2 retrievals are compared to the B10 quality control (qc_flag) retrievals to assess the performance of the NN filter. Finally, we discuss results of the study and future work to improve the NN filtering.

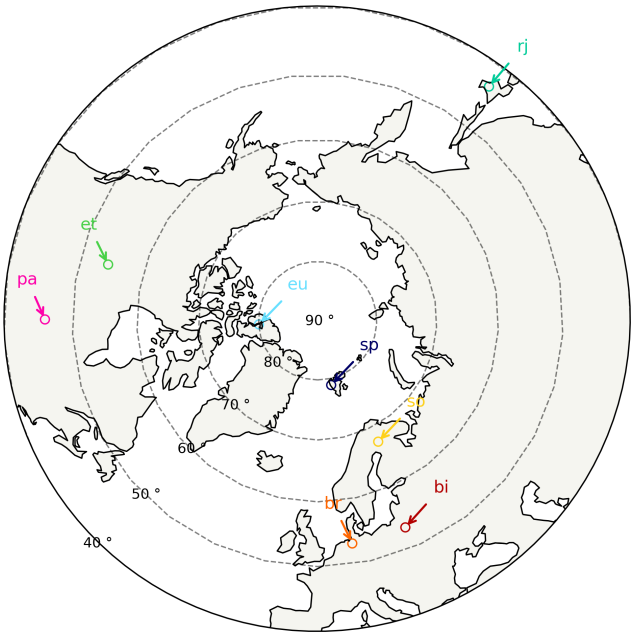

Figure 1. Map of the locations of all TCCON sites used in this study.

2 Coincidence criteria

The OCO-2 satellite was launched on 2 July 2014 and has been making measurements since mid-September 2014. The instrument on board the satellite is a three-channel, imaging, grating spectrometer that records spectra of reflected sunlight in three spectral bands centered at 0.765, 1.62, and 2.04 µm. These spectra are processed using a "full-physics" retrieval algorithm that retrieves a profile of CO_2 (which is used to calculate XCO_2) and other geophysical information. In this study we use OCO-2 data that have been processed using the B10 version of the full-physics retrieval algorithm, with the retrieval output and sounding information contained in the B10 lite files (Osterman et al., 2020). All soundings used in the study were recorded from September 2014 to July 2020.

TCCON is a global network of ground-based Fourier transform infrared (FTIR) spectrometers that record direct solar absorption spectra. The high-resolution spectra are processed using the GGG2014 retrieval algorithm which scales the a priori profile of the gas of interest until the spectrum calculated by forward model best matches the spectrum recorded by the FTIR (Wunch, et al., 2015). GGG2014 retrieves XCO_2, XCH_4, XCO, XN_2O, XHF, and XH_2O from a single spectrum. Selected XCO_2 TCCON retrievals made in the boreal and Arctic regions were used as a truth proxy to compare to OCO-2 retrievals. The TCCON sites used in this study and the date range of the data are East Trout

Lake, Canada (et) (Wunch et al., 2018), from October 2016 to June 2020; Eureka, Canada (eu) (Strong et al., 2019), from September 2014 to July 2020; Park Falls, USA (pa) (Wennberg et al., 2017), from September 2014 to June 2020; Bremen, Germany (br) (Notholt et al., 2019a), from September 2014 to August 2018; Białystok, Poland (bi) (Deutscher et al., 2019), from September 2014 to August 2018; Sodankylä, Finland (so) (Kivi et al., 2014, and Kivi and Heikkinen, 2016), from September 2014 to November 2019; Ny Ålesund, Spitzbergen, Norway (sp) (Notholt et al., 2019b), from September 2014 to August 2018; and Rikubestu, Japan (rj) (Morino et al., 2018), from September 2014 to September 2019. Figure 1 shows the locations of all the TCCON sites used in this study. All TCCON spectra were processed using the GGG2014 algorithm (Wunch et al., 2015) to retrieve XCO_2 and other gases of interest. Data were filtered for standard FLAG = 0 and additionally XHF $\leq$ 150 ppt and XCO $\leq$ 125 ppb.

Filtering for XHF $\leq$ 150 ppt was done to avoid the impact of the polar vortex on the TCCON retrievals. Arctic sites such as Eureka and Ny Ålesund routinely record solar absorption spectra while under polar vortex conditions during the spring months. In some years the polar vortex can reach as far south as 40° N (Whaley et al., 2013). Boreal sites such as East Trout Lake have recorded solar spectra under polar vortex conditions but on fewer days than at the Arctic sites. Since the GGG2014 retrieval algorithm does a profile scaling retrieval (Wunch et al., 2015), it relies on good knowledge of the shape of the profile of the gases of interest. The GGG2014 profiles are built without knowledge of the impact of the polar vortex on the shape of the profiles. When XCO_2 is retrieved from spectra measured through polar vortex conditions, the shape of the a priori profile generated by the GGG2014 retrieval algorithm will likely be incorrect. This is less of an issue for OCO-2 retrievals because OCO-2 performs a profile retrieval (O'Dell et al., 2018).

The TCCON sites used in this study have no direct influence due to anthropogenic pollution but are still influenced by biomass burning plumes. At sites like East Trout Lake, major enhancements in XCO over background levels are measured, typically in late summer when measurements are made through forest fire plumes. Even a remote Arctic site like Eureka sees forest fire plumes during the summer months (Viatte et al., 2013). In an attempt to avoid a situation where a coincident TCCON measurement is influenced by a plume and the OCO-2 measurement is not, we filter any TCCON measurement where XCO is elevated above the value of $\sim$ 150 ppb or more.

We use the B10 lite OCO-2 data product (Osterman et al., 2020), where the XCO_2 values have been corrected for various biases, such as footprint-to-footprint biases and biases that are dependent on features of the atmosphere, surface, or retrieval algorithm. OCO-2 XCO_2 data are also scaled by a global offset term that was derived using the OCO-2 target mode retrievals coincident with TCCON retrievals (Osterman et al., 2020). In our study, we use all OCO-2 spectra that are coincident with the TCCON spectra acquired in nadir and glint modes over land. The coincidence criteria are the distance of an OCO-2 measurement must be $\leq$ 150 km of a TCCON station, the temperature difference between the TCCON and OCO-2 temperature profiles at 700 hPa must be $\leq$ absolute value of 2 K (Wunch et al., 2011b), and the time difference between the TCCON and OCO-2 measurements must be $\leq$ 2 h to avoid the impact of the XCO_2 diurnal cycle.

To compare TCCON and OCO-2 retrievals, one has to take into account that GGG2014 and the OCO-2 retrievals obtain information about atmospheric CO_2 from different spectral regions (which have peak sensitivity at different altitudes) and use different a priori information. To adjust the OCO-2 bias-corrected XCO_2 retrievals to take into account the a priori profile used in the GGG2014 retrieval, the following formula is used:

$$\mathrm{XCO_2}_{\mathrm{adj}}^{\mathrm{OCO\text{-}2}} = \mathrm{XCO_2}_{\mathrm{bc}}^{\mathrm{OCO\text{-}2}} + \sum_j \boldsymbol{h}_j^{\mathrm{OCO\text{-}2}} \left(1 - \boldsymbol{a}_j^{\mathrm{OCO\text{-}2}}\right) \cdot \left(\boldsymbol{x}^{\mathrm{TCCON}} - \boldsymbol{x}^{\mathrm{OCO\text{-}2}}\right)_j, \quad (1)$$

where $\mathrm{XCO_2}_{\mathrm{bc}}^{\mathrm{OCO\text{-}2}}$ is the original bias-corrected XCO_2 value found in the lite files, $\boldsymbol{h}_j^{\mathrm{OCO\text{-}2}}$ is the OCO-2 pressure weighting vector, $\boldsymbol{a}_j^{\mathrm{OCO\text{-}2}}$ is the OCO-2 total column averaging kernel vector, $\boldsymbol{x}^{\mathrm{OCO\text{-}2}}$ is the XCO_2 a priori profile used in the OCO-2 retrieval, and $\boldsymbol{x}^{\mathrm{TCCON}}$ is the XCO_2 a priori profile used in the GGG2014 retrieval but interpolated onto the OCO-2 retrieval pressure grid.

OCO-2 retrieves information about CO_2 from the strong CO_2 band (centered at 2.04 µm) and the weak CO_2 band (centered at 1.62 µm) (O'Dell et al., 2018). TCCON retrieves information from two weak CO_2 bands, centered at 1.62 and 1.57 µm (Wunch et al., 2011b), but not in the strong CO_2 band. This results in the OCO-2 retrievals, having different vertical sensitivities compared to the TCCON retrievals. To take this into account when comparing OCO-2 and TCCON retrievals, the following formula is used to adjust the TCCON retrieved XCO_2:

$$\mathrm{XCO_2}_{\mathrm{adj}}^{\mathrm{TCCON}} = \mathrm{XCO_2}_{\mathrm{a\ priori}}^{\mathrm{TCCON}} + \sum_j \boldsymbol{h}_j^{\mathrm{OCO\text{-}2}} \boldsymbol{a}_j^{\mathrm{OCO\text{-}2}} \cdot \left(\gamma \boldsymbol{x}^{\mathrm{TCCON}} - \boldsymbol{x}^{\mathrm{TCCON}}\right)_j, \quad (2)$$

where $\mathrm{XCO_2}_{\mathrm{a\ priori}}^{\mathrm{TCCON}}$ is the integrated a priori profile used in the GGG2014 retrieval, $\boldsymbol{h}_j^{\mathrm{OCO\text{-}2}}$ is the OCO-2 pressure weighting vector, $\boldsymbol{a}_j^{\mathrm{OCO\text{-}2}}$ is the OCO-2 total column averaging kernel vector, $\boldsymbol{x}^{\mathrm{TCCON}}$ is the XCO_2 a priori profile used in the GGG2014 retrieval, and γ is the TCCON XCO_2 value divided by $\mathrm{XCO_2}_{\mathrm{a\ priori}}^{\mathrm{TCCON}}$. Ideally γ should be the scaling factor determined by the GGG2014 retrieval, but this value does not take into account the air mass dependence correction and aircraft calibration factor applied in post processing to the retrieved XCO_2. The vectors $\boldsymbol{a}_j^{\mathrm{OCO\text{-}2}}$ and $\boldsymbol{x}^{\mathrm{TCCON}}$ have been

interpolated onto a 20-layer pressure grid using the surface pressure measured at the TCCON site.

The bias between coincident TCCON and OCO-2 retrievals is calculated by taking the difference between $XCO_2{}_{adj}^{OCO\text{-}2}$ (Eq. 1) and $XCO_2{}_{adj}^{TCCON}$ (Eq. 2) and resulting in:

$$XCO_2{}^{Diff} = XCO_2{}_{adj}^{OCO\text{-}2} - XCO_2{}_{adj}^{TCCON}. \quad (3)$$

3 Neural network architecture and training

To filter the OCO-2 data, we use a three-layer neural network (NN) that consists of an input layer, a hidden layer, and an output layer. The design of the NN is based on the book by Nielsen (2015). The input layer is the value of the features of the OCO-2 retrievals that are given in the B10 lite files. Table 1 lists all the features used by the NN. An initial feature list was built by combing all features in the OCO-2 qc_flag filter (Osterman et al., 2020), with the features contained in the retrieval state vector. Features that provide information about the quality of the spectral fit, the quality of the recorded spectrum, and air mass were also included in the initial features list. To reduce the total number of features used, each feature that was thought to provide redundant information to others was removed by testing how the NN performed with and without the feature. The bias, precision, number of outliers (absolute value of $XCO_2{}^{Diff} > 2.5$ ppm) getting through, and throughput of the training data set were used as the metrics to judge the NN performance with and without the feature. The hidden layer contains the "neurons" where the calculations are done. Each input is connected to a neuron by a weight. The calculation for single neuron (N) in a NN with k neurons is given by

$$N^k = \sum_{i=1}^{n} w_i^k I_i + b^k, \quad (4)$$

where I_i is the value of feature i, w_i^k is the weight on feature i for neuron k, and b^k is the bias associated with neuron k. There is a total of 37 neurons, which is the total number of features plus one. An activation function is commonly applied to the neuron in order to introduce some nonlinearity into the neuron calculation and to make sure that small changes in the values of w_i^k and b^k result in small changes in the final output values when training the NN (Nielsen, 2015). The sigmoid function

$$f(N^k) = \frac{1}{1 + e^{-N^k}} \quad (5)$$

is used as the activation function. Each neuron is linked to the final output value by a weight (w_k). The output value is given by

$$\hat{Y} = f\left(\sum_{i=1}^{k} w_k f(N^k) + b\right), \quad (6)$$

where b is the offset, and everything else is as described as before.

Applying the sigmoid activation function in Eq. (6) ensures that $\hat{Y}$ will have a value between 0 and 1. This is useful for binary classification, which in this case we would use the NN to classify the OCO-2 retrieval as either "good" or "bad" by equating a calculated value that is close to 0 as good and a calculated value that is close to 1 as bad.

For the NN to work, the values of w_i^k, b^k, w_k, and b need to be determined. This was done by using a subset of the OCO-2 coincident retrievals to train the NN. The coincident data set consists of co-located OCO-2 soundings at the following TCCON sites: East Trout Lake (et), Eureka (eu), Bremen (br), Białystok (bi), Sodankylä (so), Ny Ålesund (sp), and Rikubestu (rj). We withhold the Park Falls (pa) data set so that it can be a completely independent source of validation. The coincident data were split into three data sets: training, testing, and validation. For the training and testing data, 20 % of the data were randomly selected to go into each data set, with the remaining 60 % used for validating the results. In order to train the NN, one needs to know the input values of the training data set (which are the values of the features in the B10 lite files) and the expected output value (Y). The expected output value was set to $Y = 0$ if the difference between a coincident OCO-2 and TCCON retrieval is $\leq \pm 2.5$ ppm and set to $Y = 1$ if the difference between the retrievals is > 2.5 ppm. Figure 2a shows the histogram of the difference between coincident OCO-2 and TCCON retrievals as well as the boundaries separating data into expected values of 0 and 1. All data between the dashed red lines were set to $Y = 0$ (or good) and set to $Y = 1$ (or bad) if outside of the boundary.

To achieve the best results when training the NN, we standardize the values of the input features so that each feature has a similar range of values. This is helpful because the features have different units and orders of magnitude, and if left as is the NN, they will place much more importance on features that have large absolute values than other features with smaller values. To standardize the input features, the following formula is used:

$$z_i = \frac{I_i - \mu_i}{\sigma_i}, \quad (7)$$

where I_i is as before, μ_i is the mean of I_i values from the training data set, and σ_i is the standard deviation of the I_i values from the training data set. This means that z_i is used in Eq. (4) instead of I_i. The Excel file in the Supplement (sheet "Standardize values") contains the μ_i and σ_i for each of the features to be used to standardize the data before inputted into the NN.

Table 1. List of all the features available to the input layer of the neural network and a brief description of the features based on the descriptions found in Osterman et al. (2020).

Feature name	Description of feature
Retrieval co2_grad_del	Measure of how much the retrieved profile shape is different compared to the a priori profile
Retrieval dpfrac	Correction of XCO_2 due to satellite pointing error
Retrieval eof3_3_rel	Scale factor for of the third empirical orthogonal function (eof) in the CO_2 strong band
Retrieval deltaT	Retrieved offset of a priori temperature profile
Retrieval h2o_scale	Scale factor of retrieved H_2O column
Retrieval aod_oc	Retrieved aerosol optical depth of organic carbon
Retrieval aod_water	Retrieved aerosol optical depth of water
Retrieval aod_dust	Retrieved aerosol optical depth of dust
Retrieval aod_bc	Retrieved aerosol optical depth of black carbon
Retrieval aod_strataer	Retrieved aerosol optical depth of stratospheric aerosol
Retrieval aod_seasalt	Retrieved aerosol optical depth of sea salt
Retrieval aod_sulfate	Retrieved aerosol optical depth of sulfate
Retrieval aod_ice	Retrieved aerosol optical depth of ice
Retrieval water_height	Retrieved central pressure of cloud water layer relative to surface pressure
Retrieval ice_height	Retrieved central pressure of cloud ice layer relative to surface pressure
Retrieval dust_height	Retrieved central pressure of dust aerosol layer relative to surface pressure
Retrieval albedo_wco2	Retrieved albedo of weak CO_2 band
Retrieval albedo_slope_wco2	Retrieved albedo slope of weak CO_2 band
Retrieval albedo_quad_wco2	Retrieved albedo quadratic coefficient of weak CO_2 band
Retrieval albedo_sco2	Retrieved albedo of strong CO_2 band
Retrieval albedo_slope_sco2	Retrieved albedo slope of strong CO_2 band
Retrieval albedo_quad_sco2	Retrieved albedo quadratic coefficient of strong CO_2 band
Retrieval albedo_o2a	Retrieved albedo of O_2 A-band
Retrieval albedo_slope_o2a	Retrieved albedo slope of O_2 A-band
Retrieval albedo_quad_o2a	Retrieved albedo quadratic coefficient of O_2 A-band
Retrieval rms_rel_o2a	Root mean square residual of O_2 A-band relative to continuum signal
Retrieval rms_rel_wco2	Root mean square residual of weak CO_2 band relative to continuum signal
Retrieval rms_rel_sco2	Root mean square residual of strong CO_2 band relative to continuum signal
Sounding altitude_stddev	How much the surface elevation changes within the soundings field of view
Preprocessors max_declocking_sco2	Estimate of the clocking error in the strong CO_2 band
Preprocessors max_declocking_o2a	Estimate of the clocking error in the O_2 A-band
Preprocessors max_declocking_wco2	Estimate of the clocking error in the weak CO_2 band
Preprocessors co2_ratio	Ratio of retrieved CO_2 column from the strong and weak CO_2 bands
Preprocessors h2o_ratio	Ratio of retrieved H_2O column from the strong and weak CO_2 bands
solar_zenith_angle	Solar zenith angle of sounding
sensor_zenith_angle	Sensor zenith angle of sounding

To determine the values of w_i^k, b^k, w_k, and b, they are initially set randomly to be between a value of ± 1. Using the training data set, $\hat{Y}$ is calculated for all the data using the initial values of w_i^k, b^k, w_k, and b. The performance of the NN is then determined by comparing the calculated value ($\hat{Y}$) to the expected output value (Y) using the log loss entropy cost function:

$$C = \frac{-1}{n} \sum_{i}^{n} Y^i \log(\hat{Y}^i) + (1 - Y^i) \log(1 - \hat{Y}^i), \tag{8}$$

where n is the total number of OCO-2 retrievals in the training data set. If $\hat{Y} = Y$, then C will equal zero, meaning the values of w_i^k, b^k, w_k, and b are set to the best values that perfectly determine whether an OCO-2 retrieval is good or bad. This is unlikely to happen for the initial values of those variables since they are set randomly, so C will be > 0. To minimize the value of C, the values of w_i^k, b^k, w_k, and b are adjusted. The adjustments are done by taking the partial derivative with respect to the cost function (i.e., $\frac{\partial C}{\partial w_i^k}$, $\frac{\partial C}{\partial b^k}$, $\frac{\partial C}{\partial w_k}$, and $\frac{\partial C}{\partial b}$). In principle, this should be iterated until $C = 0$, but in practice, the classification of the training data setup is not perfect. The assumption made when setting up the classification of the training data is that if $-2.5\,\text{ppm} < \text{XCO}_2^{\text{Diff}}$ (Eq. 3) $< 2.5\,\text{ppm}$, then it is a good OCO-2 retrieval, but this might not be true. It could be the case that the OCO-2 retrieval has adjusted parameters as much as possible to achieve the best possible fit to the measured spectra but that the retrieved parameters deviate from the true values while still providing an integrated profile that is close to the TC-

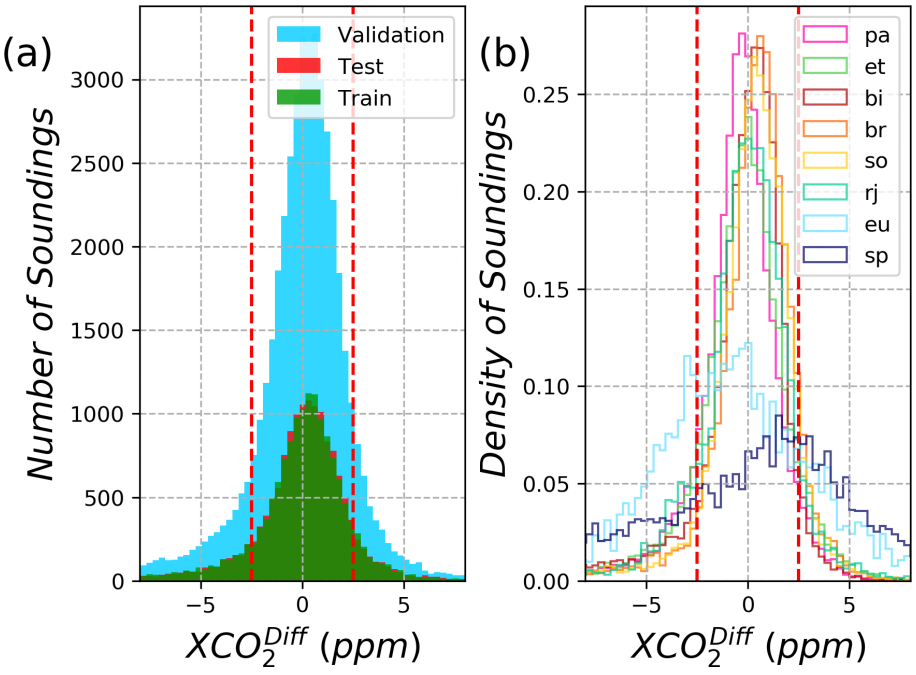

Figure 2. (a) Histogram of the bias between coincident TCCON and OCO-2 retrievals $XCO_2{}^{Diff}$ (Eq. 3) for the three data sets. The dashed red lines represent the boundary between setting the classification of the data. **(b)** Same as plot **(a)** but shows the density of soundings for each of the sites.

CON XCO_2. This retrieval would be misclassified as good, so the cost function will never reach 0.

To stop training the NN, a few cutoffs were placed: the maximum number of iterations is 5000 or the accuracy between the training and testing data $< 3\,\%$. When training the NN, the accuracy of the training data and the testing data is calculated on each iteration and compared. Since the data were set up in a binary classification (i.e., 0 or 1), on each iteration the classification was set to 0 if a calculated value was ≤ 0.1 (unitless) and the classification was set to 1.0 if a calculated value was > 0.1. This threshold of 0.1 was determined by trying to balance throughput with degradation of precision as well as limiting the amount of individual retrievals with high absolute $XCO_2{}^{Diff} > 2.5$ ppm passing the NN filter. These classification values were compared to the expected classification value on each iteration to get the accuracy of the training and testing data sets. The testing data set is not used when determining the values of w_i^k, b^k, w_k, and b, but rather it is used as an independent data source to make sure that the NN is not overfitting the training data. The derived values of w_i^k, b^k, w_k, and b can be found in the Excel file in the Supplement with values of w_i^k in sheet "w1", b^k in sheet "b1", w_k in sheet "w2", and b in sheet "b2".

Figure 3 shows the $XCO_2{}^{Diff}$ as a function of the value calculated by the NN for all three data sets. Figure 3a shows that the OCO-2 retrievals with calculated values close to 0 have the smallest spread in $XCO_2{}^{Diff}$, while calculated values close to 1 have the largest spread. This pattern is seen in all three of the data sets. The density plot shown in Fig. 3b confirms that for most of the data the calculated values are ≤ 0.1. There is no clear separation of data (i.e., good retrievals ≤ 0.1 and bad retrievals ≤ 0.9) as one would expect from a binary classifier. Clearly there are retrievals with an ambiguous classification ($\hat{Y} > 0$ and < 1.0) even though all retrievals in the training data set were assigned a value of 0 or 1. For the NN to achieve an ambiguous value when training, there would have to be retrievals with similar feature values, with no clear majority between good and bad examples in the training data. This could happen because $XCO_2{}^{Diff}$ is not a perfect classification metric, which could lead to some portion of retrievals being incorrectly classified. Another possibility is that a lack of data combined with the real variance in the scene could result in the no clear majority case leading to an ambiguous classification. Since the NN is not showing any confidence in the classification of these retrievals, they are manually classified as bad to err on the side of caution. There is also the possibility that actual bad retrievals can get through the NN filter due to insufficient training examples as well as incorrect classification of training data.

4 Validation

To validate the NN filtering, the NN filter was applied to the validation data set and compared to the same validation data set but with the B10 qc_flag = 0 applied to the soundings. Since the validation data set was not used in the training of

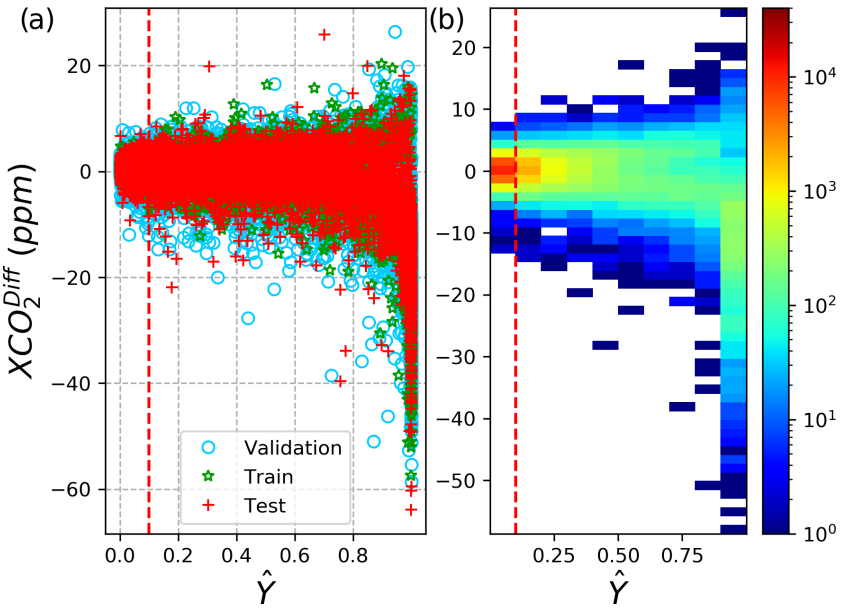

Figure 3. **(a)** The difference between the coincident TCCON and OCO-2 retrievals ($XCO_2{}^{Diff}$) as a function of $\hat{Y}$ calculated by the NN (after training) for the training, testing, and validation data sets. **(b)** Same as plot **(a)** but shows the density of all three data sets combined, with the color bar given on a log scale. The dashed red lines represent the boundary between setting the binary classification of the data.

Table 2. The XCO_2 bias and scatter (ppm) and number of OCO-2 retrievals at each TCCON site as well as overall for the OCO-2 bias-corrected XCO_2 after applying either the NN or qc_flag filter to the validation data set. The row labeled "All" excludes data from Park Falls.

	Neural network		B10 qc_flag	
Site	Bias ± precision	Number of retrievals	Bias ± precision	Number of retrievals
All	0.25 ± 1.27	23 429	0.52 ± 1.45	20 198
Eureka (eu)	−0.53 ± 2.35	59	0.34 ± 2.94	634
Ny Ålesund (sp)	0.87 ± 2.30	91	2.09 ± 2.65	92
Sodankylä (so)	0.34 ± 1.23	5118	0.64 ± 1.30	4736
East Trout Lake (et)	0.01 ± 1.34	5261	0.44 ± 1.48	3186
Białystok (bi)	0.27 ± 1.16	6237	0.40 ± 1.18	5609
Bremen (br)	0.42 ± 1.19	4066	0.85 ± 1.28	3672
Rikubetsu (rj)	0.23 ± 1.39	2597	0.13 ± 1.70	2269
Park Falls (pa)	−0.12 ± 1.27	14 859	−0.12 ± 1.18	12 406

the NN, it is an independent data set kept aside to assess the performance of the NN filter. Table 2 shows the bias, scatter, and number of retrievals for the entire validation data set (All) and at each site when applying either the NN or qc_flag filter. The overall XCO_2 bias using the NN filter is half of the qc_flag filter, the scatter has been decreased by 0.18 ppm, and the throughput has been increased by 16 %. The NN filter reduces the bias at every site except at Eureka and Rikubetsu. The precision is better at every site when the NN filter is applied to the validation data. The throughput has increased at every site, when the NN filter is used, except for the Arctic sites (Eureka and Ny Ålesund). Park Falls data were not used to train the NN filter as Park Falls is slightly outside of the boreal domain, but the data set is used as a completely independent data set to validate the NN filter. When the NN filter is applied to Park Falls data, the bias remains the same, the precision decreases by 0.09 ppm, and the throughput increases by ∼ 20 %.

The reduction in throughput at the Arctic sites is because the distribution of data at the Arctic sites is different compared to all other sites as shown in Fig. 2b. The peaks of the histograms for the Arctic sites are closer to the boundaries used to classify the training data as good or bad, so almost half of the data are set to bad when training the NN. Figure 4 shows the pass rate for the NN filter given the value of the solar zenith angle (Fig. 4a), sensor zenith angle (Fig. 4b), and altitude standard deviation (Fig. 4c). In all three plots the data are binned, with the blue dots showing the number of

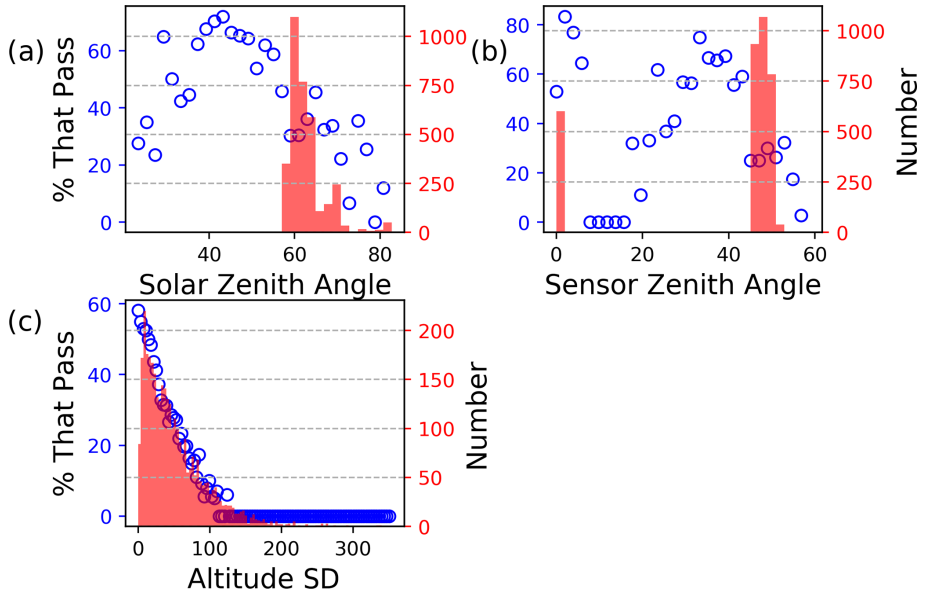

Figure 4. The percentage of data that passes the NN filter (blue dots) for a given solar zenith angle **(a)**, sensor zenith angle **(b)** and altitude standard deviation (stddev) **(c)**. The bars are the histograms of the OCO-2 soundings coincident with the Eureka TCCON data.

OCO-2 soundings that pass the NN filter divided by the total amount of data multiplied by 100 in each bin. The pink bars plot the histogram of OCO-2 soundings coincident with Eureka TCCON data. Figure 4a shows that the coincident OCO-2 soundings are made at solar zenith angles between 58 and 85°, with the blue dots showing that 30 % to 0 % of the soundings that have these values pass the NN filter. Similarly, Fig. 4b shows that the coincident OCO-2 soundings are made at high sensor zenith angles, which are less likely to pass the NN filter. Most of the coincident OCO-2 soundings at Eureka are made over land that contains significant topographic variability. Figure 4c shows the altitude standard deviation, which is the standard deviation of the elevation (in meters) of the field of view of the sounding. The plot shows that at an altitude standard deviation of $\sim$50 m only 30 % of the soundings pass the NN filter. The combination of high air mass and variable topography decreases the throughput at Eureka.

For further validation, the seasonal bias, scatter, and number of retrievals that pass the filters at each site are compared. Figure 5 shows the bias at each site for spring, summer, fall, and winter when the NN filter is applied to the validation data (solid bars) and also when the qc_flag filter (dashed bars) is used on the same validation data set. For most sites and seasons, the magnitude of the biases for the two different filtering schemes are similar, although in most cases the NN filter has a lower absolute bias compared to the qc_flag filter. The NN filter significantly improves the bias at Sodankylä and Rikubetsu during spring, Ny Ålesund during summer, and Rikubetsu and Bremen during winter. Both the NN filter and the qc_flag show that there is a positive bias between OCO-2 and TCCON in summer. The NN filter is able to reduce this

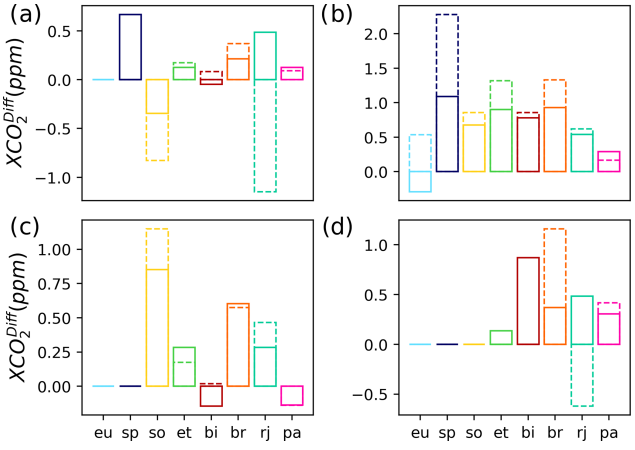

Figure 5. The bias at each site for the different seasons when the NN filter (bars with solid lines) and qc_flag (bars with dashed lines) are used to filter the OCO-2 retrievals in the validation data set. **(a)** Spring (March, April, May); **(b)** summer (June, July, August); **(c)** fall (September, October, November); **(d)** winter (December, January, February). Note the different y-axis ranges for each plot. Note that bars that show a bias of zero are due to no data rather than a bias of zero.

summer bias, but it still remains. At Park Falls the bias between the two filters is similar for the different scenes, with the qc_flag showing a lower bias in summer, and the NN filter decreasing the bias in winter.

Figure 6 shows the precision at each site for spring, summer, fall, and winter when the different filters are applied to the validation data. The precision is very similar (i.e., within 0.2 ppm) for most sites during the different seasons. The NN

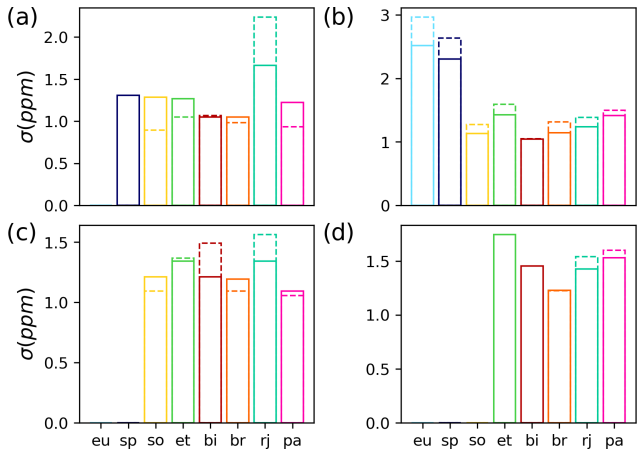

Figure 6. Same as Fig. 5 but shows the precision at each site for each season. **(a)** Spring (March, April, May); **(b)** summer (June, July, August); **(c)** fall (September, October, November); **(d)** winter (December, January, February). Solid bars indicate the NN filter, and dashed bars indicate the original B10 qc_filter.

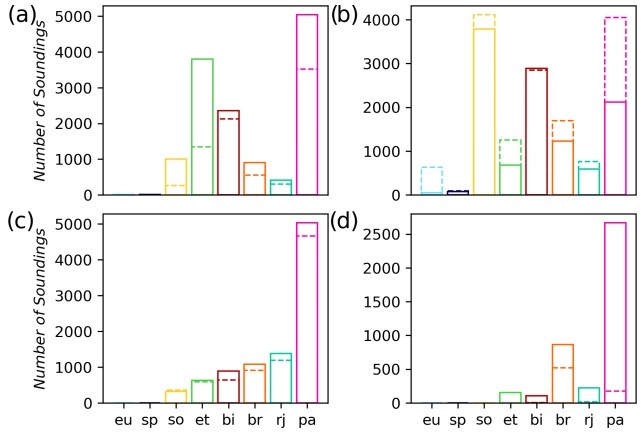

Figure 7. Same as Fig. 5 but shows the number of soundings that pass each filter at each site for the different seasons. **(a)** Spring (March, April, May); **(b)** summer (June, July, August); **(c)** fall (September, October, and November); **(d)** winter (December, January, February). Solid bars indicate the NN filter, and dashed bars indicate the original B10 qc_filter.

filter improves the precision (by more than 0.2 ppm) at Rikubetsu during spring, Eureka and Ny Ålesund during summer, and Białystok and Rikubetsu during fall. However, the qc_flag filter has a much better precision at Sodankylä during spring when compared to the NN filter.

Figure 7 shows the number of retrievals that pass each filter for the different sites during spring, summer, fall and winter. At most sites, the NN filter lets through more retrievals compared to the qc_flag filter during spring, fall, and winter. In summer the qc_flag filter has a slightly higher throughput compared to the NN filter at most sites. This decrease in throughput during summer helps improve the bias and precision as seen in Figs. 5b and 6b. There is a significant increase in throughput at East Trout Lake during spring with the NN filter, and it even produces some retrievals in winter. At Park Falls the throughput has increased in spring, fall, and winter but is significantly decreased during summer. The decrease in summer is because the NN filter is trained on data that show a bias during summer, which it decreases by filtering out more data compared to the qc_flag filter. Even though Park Falls is not in the boreal domain, its scene type (forest) is similar to East Trout Lake. The NN has no information on time of year, but it does have information on the surface type through the albedo values, which change due to the time of year. It is most likely that what the NN learned from the East Trout Lake data is influencing how the NN filters the data at Park Falls.

Some of the increase in throughput with the NN filter during spring, fall, and winter can be explained by the fact that the qc_flag filter tries to filter out spectra that have been recorded over snow scenes (Osterman et al., 2020). The snow_flag, found in the B10 lite files, is used to indicate the presence of snow in the scene. We applied this snow_flag to the NN-filtered data to see if the NN filter removes all soundings over snow. From the validation data set, 3219 retrievals have snow_flag = 1, with 785 of those retrievals passing the NN filter. This means that the NN filter passes about 24 % of the OCO-2 soundings made over snow. This is much lower compared to the general case (all scenes) where greater than 40 % of the data pass both the NN and qc_flag filters. The bias over snow scenes compared to TCCON from all the retrievals that pass the NN filter in the validation data set is 0.13 ± 1.44. Since the precision is lower over snow, it makes sense that the throughput over snow is lower compared to the general case. At Park Falls 1032 soundings that pass the NN filter, with 727 in winter and 302 in spring, are snow scenes. So a significant amount of throughput during winter at Park Falls are made over snow scenes. The bias of snow scenes at Park Falls was found to be 0.12 ± 1.41.

The NN filter was applied to all OCO-2 B10 data at latitudes greater than 45° N to determine the throughput in the boreal and Arctic regions. Figure 8 shows the percent difference (NN minus qc_flag, divided by qc_flag, and multiplied by 100) between the number of soundings that pass the filters. The maximum value for the percent difference was capped at 100 %. The throughput with the NN filter is greater than the qc_flag filter in spring and winter, while the throughput with the qc_flag filter is greater than the NN filter in summer and fall. This is consistent with what is seen at the individual TCCON sites. Over Greenland the throughput has increased with the NN filter regardless of season, because the qc_flag filter removes all data over Greenland with the snow_flag filter. During fall, the throughput has increased at greater than 70° N, because the NN filter is letting through soundings that were recorded over snow scenes.

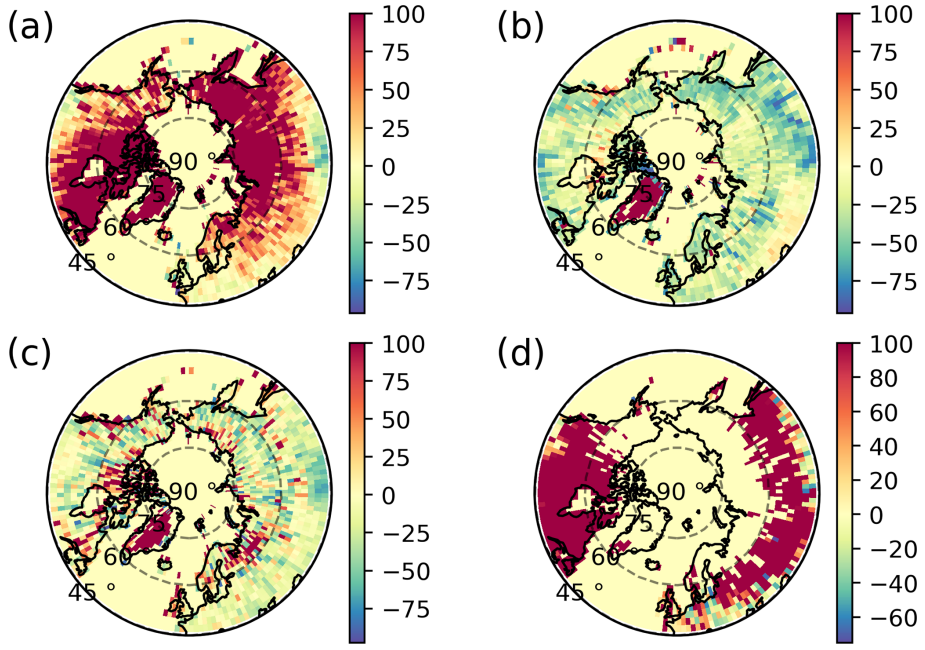

Figure 8. Polar plots of the percent difference in the number of soundings that pass the NN filter compared to the qc_flag filter for spring **(a)**, summer **(b)**, fall **(c)**, and winter **(d)**. The data have been binned by 2° longitude by 2° latitude.

5 Discussion and conclusions

In this study, a neural network was used to filter the OCO-2 bias-corrected XCO_2 data collected near northern high-latitude TCCON stations as described in Sect. 3. The performance of the NN filter was assessed by comparing the bias, precision, and throughput to the quality control filtered data. There was an improvement in the bias, precision, and throughput both overall and at most sites, as well as improvements in the bias in different seasons. However, the NN filter decreases the throughput at Eureka, because it finds that OCO-2 soundings made at high solar zenith angles, high sensor zenith angles, and over variable topography are problematic.

The main downside to using a neural network to filter OCO-2 retrievals is that it does not readily provide information on why a retrieval was classified as good or bad, which would be useful for improving the retrieval algorithm. Decision tree algorithms are binary classifiers which do provide information on the classification of data. However, a drawback to decision trees is that they overfit the training data. In Sect. 3, it was explained why some of the training data might be incorrectly classified, leading to the NN filter determining an ambiguous classification for some retrievals. The assumption with decision trees is that all the training data are correctly classified, which is not entirely true in this case. The NN calculating values that make the classification of some retrievals ambiguous can be interpreted as a confidence in the classification of retrievals. With a decision tree, there is no metric to measure the confidence in the classification of retrievals, so you could get retrievals passing the decision tree filter that are actually closer to bad retrievals than good.

There are also ways that the implementation of a NN to filter OCO-2 retrievals can be improved upon. Increasing the amount of retrievals used in the training data set would improve the performance of the NN filter. This can be done by incorporating new coincident TCCON measurements from the sites in this study, retrievals from new TCCON sites coming into operation, and the potential use of other similar truth proxies such as the COllaborative Carbon Column Observing Network (COCCON) (Frey et al., 2019). To improve the training classification, other helpful truth proxies such as cloud and aerosol information can be combined with $XCO_2{}^{Diff}$ when classifying the retrievals for training. The current implementation of the NN as a binary classifier was done to make the problem as simple as possible in order to filter out retrievals where the forward model of the retrieval algorithm is suboptimal rather than scenes of high variance. A possible alteration to the algorithm would be to do the classification on a continuum where the output of the NN ($\hat{Y}$) would be related to the expected precision of the data. The downside to this configuration would be that the NN filter would be filtering out not only bad retrievals but also scenes with high variance. For example, if you wanted a precision of better than 1 ppm, chances are you would not have any retrievals over snow getting through the NN, greatly reducing throughput in the winter and shoulder season at high latitudes.

This study shows the potential of using a neural network to filter OCO-2 retrievals that could be useful in future filter-

ing schemes for OCO-2 or other satellite missions. However, there are potential drawbacks to the methodology presented in this study. In this study, we focus on data near northern high-latitude TCCON stations and so do not sample globally representative ranges of surface properties or air masses. Figure 1 shows the limited coverage that the TCCON sites provide, with no coverage over Greenland and most of the Eurasian boreal region. The effectiveness of the NN filter is dependent on how well the NN is trained. We train the NN using OCO-2 data coincident with TCCON data, so the NN filter is trained only under atmospheric conditions observed at the northern high-latitude TCCON sites. We have shown that this way of training the NN is effective when validated against northern high-latitude TCCON data. When the NN filter was applied to Park Falls data, which were not used in the training of the NN, we found that the bias was similar to the qc_flag filter, with a decrease of 0.09 ppm in precision but a 20 % increase in throughput. Although the throughput increased in spring, fall, and winter seasons, it decreased a lot during summer. The decrease in throughput in summer led to improved bias and precision values at all the TCCON sites used in the training of the NN but not at Park Falls. This is because the NN has found a pattern that improves the training data set which is not as applicable to Park Falls data. The qc_flag filter lets through almost twice as much data compared to the NN filter during summer with a decrease in precision of only 0.09 ppm compared to the NN filter. The NN filter is suboptimal at Park Falls (during summer) and if one applied the NN filter to data that are not similar to northern high-latitude data used to train the NN, it will not be as effective.

The effectiveness of the NN filter is dependent on the data set used to train the NN. In this study we assume that the TCCON data represent the truth and that any bias that we see is in the OCO-2 retrievals. The NN is trying to decrease the bias it sees as much as possible, and if there is a bias in the TCCON data, it will attribute this to a bias in the OCO-2 data and treat that data as bad. One way to be less influenced by TCCON data is to use a small area approximation, where XCO_2 is assumed constant within a small region (O'Dell et al., 2018). While the absolute value of the retrieval cannot be evaluated using a small area analysis, variability within the small area can, and this would vastly increase the data set size used in the NN and improve the range of surface properties, atmospheric conditions, and air masses represented by the training data set. This small area approach will be investigated in a future study.

The NN filter passes some retrievals of soundings made over snow albeit at a lower throughput compared to non-snow scenes. This is expected because retrievals over snow are difficult due to the low albedo in the spectral regions of the CO_2 bands, often providing insufficient signal for a good retrieval. However, the albedo of snow is dependent on the age of the snow with fresh snow having higher albedo compared to old snow, so there is a possibility that some of the soundings recorded over snow have enough signal to produce a good retrieval. Nevertheless, these retrievals are further complicated by the fact that the spectra are usually recorded through large solar zenith angles (SZAs) due to the soundings being made at either high latitudes or at times of the year when the SZA is large, which is challenging for the radiative transfer model of the retrieval algorithm to deal with. The results of this study show the potential of a machine learning algorithm to tease apart these factors and recover some of the retrievals over snow, although in this study there was not enough coincident data over snow to get meaningful site statistics (bias and precision). A future study will investigate the potential of a machine learning algorithm to filter the retrievals over snow by folding in more training and validation data.

The accuracy of XCO_2 observations over the northern high latitudes and the loss of data there due to filtering has been a long-standing issue with OCO-2 and GOSAT data versions to date, which has limited the scientific community's ability to apply their data to investigate important northern high-latitude carbon cycle science questions. This paper demonstrates that a neural network approach can be used to increase the number of soundings at northern high latitudes, while also improving the bias, precision, and throughput depending on the site. One possible future application of the NN (or other machine learning algorithms) could be to improve the bias correction of OCO-2 retrievals. Le et al. (2020) used a convolution NN for spatiotemporal bias correction of satellite precipitation data, and air quality forecasts have moved towards bias correction using a decision tree algorithm (Ivatt and Evans, 2020). Continual efforts at improving northern high-latitude retrievals and filtering will be beneficial not only for current missions but also for future XCO_2 missions like MicroCarb (Pasternak et al., 2016), GOSAT-GW (Kasahara et al., 2020), and CO_2M (Sierk et al., 2018), which will make global observations that include northern high latitudes, and even more so for missions under consideration like AIM-North (Nassar et al., 2019), which is dedicated to observing the Arctic and boreal atmosphere.

Data availability. The OCO-2 B10 lite files were obtained from the NASA Goddard Earth Science Data and Information Services Center (GES DISC; https://oco2.gesdisc.eosdis.nasa.gov/data/OCO2_DATA/OCO2_L2_Lite_FP.10r/, NASA, 2020). TCCON data are available from the TCCON Data Archive, hosted by CaltechDATA (https://tccondata.org/, last access: 1 August 2020). DOIs: https://doi.org/10.14291/TCCON.GGG2014.BREMEN01.R1 (Notholt et al., 2019a); https://doi.org/10.14291/TCCON.GGG2014.BIALYSTOK01.R2 (Deutscher et al., 2019); https://doi.org/10.14291/TCCON.GGG2014.SODANKYLA01.R0/1149280 (Kivi et al., 2014); https://doi.org/10.14291/TCCON.GGG2014.NYALESUND01.R1 (Notholt et al., 2019b); https://doi.org/10.14291/TCCON.GGG2014.EASTTROUTLAKE01.R1 (Wunch et al., 2018); https://doi.org/10.14291/TCCON.GGG2014.EUREKA01.R3

(Strong et al., 2019); https://doi.org/10.14291/TCCON.GGG2014.PARKFALLS01.R1 (Wennberg et al., 2017); https://doi.org/10.14291/TCCON.GGG2014.RIKUBETSU01.R2 (Morino et al., 2018).

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/amt-14-7511-2021-supplement.

Author contributions. JM designed the study, developed the neural network code, analyzed the results, and wrote the paper. RN provided input into the analysis. DW provided insight into the use of TCCON data, comparisons between TCCON and OCO-2, and the overall analysis. CWO'D provided critical analysis of choice of coincident criteria, filtering of OCO-2 data, and insight into OCO-2 bias correction. RK, IM, JN, CP, KS, and DW are involved with the operation of the TCCON sites, data processing, and use of TCCON data in this study. All authors read and provided feedback to JM on the paper.

Competing interests. Some authors are members of the editorial board of *Atmospheric Measurement Techniques*. The peer-review process was guided by an independent editor, and the authors have also no other competing interests to declare.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. The OCO-2 data were produced by the OCO-2 project at the Jet Propulsion Laboratory, California Institute of Technology, and obtained from the OCO-2 data archive maintained at the NASA Goddard Earth Science Data and Information Services Center. TCCON data were obtained from the TCCON Data Archive, hosted by CaltechDATA, California Institute of Technology. We would like to thank Paul O. Wennberg and his team for processing TCCON spectra at Park Falls and providing that data to the TCCON Data Archive.

Financial support. The Rikubetsu TCCON site is supported in part by the GOSAT series project. East Trout Lake support is provided by CFI/ORF and NSERC. The Eureka TCCON measurements were made at the Polar Environment Atmospheric Research Laboratory (PEARL) by the Canadian Network for the Detection of Atmospheric Change (CANDAC), primarily supported by the Natural Sciences and Engineering Research Council of Canada, Environment and Climate Change Canada, and the Canadian Space Agency. The University of Bremen acknowledges support by the ESA project IDEAS and by DLR within the Sentinal-S5P validation projects.

Review statement. This paper was edited by Joanna Joiner and reviewed by François-Marie Bréon and two anonymous referees.

References

Crisp, D., Atlas, R. M., Breon, F.-M., Brown, L. R., Burrows, J. P., Ciais, P., Connor, B. J., Doney, S. C., Fung, I. Y., Jacob, D. J., Miller, C. E., O'Brien, D., Pawson, S., Randerson, J. T., Rayner, P., Salawitch, R. J., Sander, S. P., Sen, B., Stephens, G. L., Tans, P. P., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Yung, Y. L., Kuang, Z., Chudasama, B., Sprague, G., Weiss, B., Pollock, R., Kenyon, D., and Schroll, S.: The Orbiting Carbon Observatory (OCO) mission, Adv. Space Res., 34, 700–709, https://doi.org/10.1016/j.asr.2003.08.062, Special Issue: Trace Constituents in the Troposphere and Lower Stratosphere, 2004.

Deutscher, N. M., Notholt, J., Messerschmidt, J., Weinzierl, C., Warneke, T., Petri, C., and Grupe, P.: TCCON data from Bialystok (PL), Release GGG2014.R2 (Version R2), TCCON Data Archive [data set], hosted by CaltechDATA, https://doi.org/10.14291/TCCON.GGG2014.BIALYSTOK01.R2, 2019.

Frey, M., Sha, M. K., Hase, F., Kiel, M., Blumenstock, T., Harig, R., Surawicz, G., Deutscher, N. M., Shiomi, K., Franklin, J. E., Bösch, H., Chen, J., Grutter, M., Ohyama, H., Sun, Y., Butz, A., Mengistu Tsidu, G., Ene, D., Wunch, D., Cao, Z., Garcia, O., Ramonet, M., Vogel, F., and Orphal, J.: Building the COllaborative Carbon Column Observing Network (COCCON): long-term stability and ensemble performance of the EM27/SUN Fourier transform spectrometer, Atmos. Meas. Tech., 12, 1513–1530, https://doi.org/10.5194/amt-12-1513-2019, 2019.

Ivatt, P. D. and Evans, M. J.: Improving the prediction of an atmospheric chemistry transport model using gradient-boosted regression trees, Atmos. Chem. Phys., 20, 8063–8082, https://doi.org/10.5194/acp-20-8063-2020, 2020.

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 CO_2 columns in the boreal forest with Orbiting Carbon Observatory-2, Total Carbon Column Observing Network, and EM27/SUN measurements, Atmos. Meas. Tech., 13, 5033–5063, https://doi.org/10.5194/amt-13-5033-2020, 2020.

Kasahara, M., Kachi, M., Inaoka, K., Fujii, H., Kubota, T., Shimada, R., and Kojima, Y.: Overview and current status of GOSAT-GW mission and AMSR3 instrument, Sensors, Systems, and Next-Generation Satellites XXIV, Proc. SPIE 11530, 1153007, https://doi.org/10.1117/12.2573914, 2020.

Kiel, M., O'Dell, C. W., Fisher, B., Eldering, A., Nassar, R., MacDonald, C. G., and Wennberg, P. O.: How bias correction goes wrong: measurement of X_{CO_2} affected by erroneous surface pressure estimates, Atmos. Meas. Tech., 12, 2241–2259, https://doi.org/10.5194/amt-12-2241-2019, 2019.

Kivi, R. and Heikkinen, P.: Fourier transform spectrometer measurements of column CO_2 at Sodankylä, Finland, Geosci. Instrum. Method. Data Syst., 5, 271–279, https://doi.org/10.5194/gi-5-271-2016, 2016.

Kivi, R., Heikkinen, P., and Kyrö, E.: TCCON data from Sodankylä (FI), Release GGG2014.R0 (Version GGG2014.R0), TCCON

Data Archive [data set], hosted by CaltechDATA, https://doi.org/10.14291/TCCON.GGG2014.SODANKYLA01.R0/1149280, 2014.

Le, X.-H., Lee, G., Jung, K., An, H., Lee, S., and Jung, Y.: Application of Convolutional Neural Network for Spatiotemporal Bias Correction of Daily Satellite-Based Precipitation, Remote Sens.-Basel, 12, 2731, https://doi.org/10.3390/rs12172731, 2020.

Mandrake, L., Frankenberg, C., O'Dell, C. W., Osterman, G., Wennberg, P., and Wunch, D.: Semi-autonomous sounding selection for OCO-2, Atmos. Meas. Tech., 6, 2851–2864, https://doi.org/10.5194/amt-6-2851-2013, 2013.

Morino, I., Yokozeki, N., Matsuzaki, T., and Horikawa, M.: TCCON data from Rikubetsu (JP), Release GGG2014.R2 (Version R2), TCCON Data Archive [data set], hosted by CaltechDATA, https://doi.org/10.14291/TCCON.GGG2014.RIKUBETSU01.R2, 2018.

NASA: GES DISC data, NASA [data set], available at: https://oco2.gesdisc.eosdis.nasa.gov/data/OCO2_DATA/OCO2_L2_Lite_FP.10r/, last access: 1 October 2020.

Nassar, R., McLinden, C., Sioris, C. E., McElroy, C. T., Mendonca, J., Tamminen, J., MacDonald, C. G., Adams, C., Boisvenue, C., Bourassa, A., Cooney, R., Degenstein, D., Drolet, G., Garand, L., Girard, R., Johnson, M., Jones, D. B. A., Kolonjari, F., Kuwahara, B., Martin, R. V., Miller, C. E., O'Neill, N., Riihelä, A., Roche, S., Sander, S. P., Simpson, W. R., Singh, G., Strong, K., Trishchenko, A. P., Mierlo, H. van, Zanjani, Z. V., Walker, K. A., and Wunch, D.: The Atmospheric Imaging Mission for Northern Regions: AIM-North, Can. J. Remote Sens., 45, 423–442, https://doi.org/10.1080/07038992.2019.1643707, 2019.

Nielsen, M. A.: Neural Networks and Deep Learning, Determination Press, San Francisco, CA, USA, 2015.

Notholt, J., Petri, C., Warneke, T., Deutscher, N. M., Palm, M., Buschmann, M., Weinzierl, C., Macatangay, R. C., and Grupe, P.: TCCON data from Bremen (DE), Release GGG2014.R1 (Version R1), TCCON Data Archive [data set], hosted by CaltechDATA, https://doi.org/10.14291/TCCON.GGG2014.BREMEN01.R1, 2019a.

Notholt, J., Warneke, T., Petri, C., Deutscher, N. M., Weinzierl, C., Palm, M., and Buschmann, M.: TCCON data from Ny Ålesund, Spitsbergen (NO), Release GGG2014.R1 (Version R1), TCCON Data Archive [data set], hosted by CaltechDATA, https://doi.org/10.14291/TCCON.GGG2014.NYALESUND01.R1, 2019b.

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O., and Velazco, V. A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11, 6539–6576, https://doi.org/10.5194/amt-11-6539-2018, 2018.

Olsen, S. C. and Randerson, J. T.: Differences between surface and column atmospheric CO_2 and implications for carbon cycle research, J. Geophys. Res.-Atmos., 109, D02301, https://doi.org/10.1029/2003JD003968, 2004.

Osterman, G. B., O'Dell, C. W., Eldering, A., Fisher, B., Crisp, D., Cheng, C., Frankenberg, C., Lambert, A., Gunson, M. R., Mandrake, L., and Wunch, D.: Orbiting Carbon Observatory-2 & 3 (OCO-2 & OCO-3) Data Product User's Guide, Operational Level 2 Data Versions 10 and Lite File Version 10 and VEarly, National Aeronautics and Space Administration Jet Propulsion Laboratory California Institute of Technology Pasadena, California, available at: https://docserver.gesdisc.eosdis.nasa.gov/public/project/OCO/OCO2_OCO3_B10_DUG.pdf (last access: 1 August 2020), 2020.

Pasternak, F., Bernard, P., Georges, L., and Pascal, V.: The MicroCarb instrument, ICSO 2016, Proc. SPIE, 10562, 105621P-13, https://doi.org/10.1117/12.2296225, 2016.

Pulliainen, J., Aurela, M., Laurila, T., Aalto, T., Takala, M., Salminen, M., Kulmala, M., Barr, A., Heimann, M., Lindroth, A., Laaksonen, A., Derksen, C., Mäkelä, A., Markkanen, T., Lemmetyinen, J., Susiluoto, J., Dengel, S., Mammarella, I., Tuovinen, J.-P., and Vesala, T.: Early snowmelt significantly enhances boreal springtime carbon uptake, P. Natl. Acad. Sci. USA, 114, 11081–11086, https://doi.org/10.1073/pnas.1707889114, 2017.

Rayner, P. J. and O'Brien, D. M.: The utility of remotely sensed CO_2 concentration data in surface source inversions, Geophys. Res. Lett., 28, 175–178, https://doi.org/10.1029/2000GL011912, 2001.

Schuur, E. A. G., 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, Nature, 520, 171–179, https://doi.org/10.1038/nature14338, 2015.

Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M., Vacchiano, G., Wild, J., Ascoli, D., Petr, M., Honkaniemi, J., Lexer, M. J., Trotsiuk, V., Mairota, P., Svoboda, M., Fabrika, M., Nagel, T. A., and Reyer, C. P. O.: Forest disturbances under climate change, Nat. Clim. Change, 7, 395–402, https://doi.org/10.1038/nclimate3303, 2017.

Sierk, B., Beźy, J.-L., Löscher, A., and Meijer, Y.: The European CO_2 Monitoring Mission: Observing anthropogenic greenhouse gas emissions from space, International Conference on Space Optics, SPIE, 237–250, https://doi.org/10.1117/12.2535941, 2018.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M.: IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK, N. Y., NY, USA, 1535 pp., https://doi.org/10.1017/CBO9781107415324, 2013.

Strong, K., Roche, S., Franklin, J. E., Mendonca, J., Lutsch, E., Weaver, D., Fogal, P. F., Drummond, J. R., Batchelor, R., and Lindenmaier, R.: TCCON data from Eureka (CA), Release GGG2014.R3 (Version R3), TCCON Data Archive [data set], hosted by CaltechDATA, https://doi.org/10.14291/TCCON.GGG2014.EUREKA01.R3, 2019.

Viatte, C., Strong, K., Paton-Walsh, C., Mendonca, J., O'Neill, N. T., and Drummond, J. R.: Measurements of

CO, HCN, and C_2H_6 Total Columns in Smoke Plumes Transported from the 2010 Russian Boreal Forest Fires to the Canadian High Arctic, Atmos.-Ocean, 51, 522–531, https://doi.org/10.1080/07055900.2013.823373, 2013.

Wennberg, P. O., Roehl, C. M., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R., Keppel-Aleks, G., Allen, N. T., and Ayers, J.: TCCON data from Park Falls (US), Release GGG2014.R1 (Version GGG2014.R1), TCCON Data Archive [data set], hosted by CaltechDATA, https://doi.org/10.14291/TCCON.GGG2014.PARKFALLS01.R1, 2017.

Whaley, C., Strong, K., Adams, C., Bourassa, A. E., Daffer, W. H., Degenstein, D. A., Fast, H., Fogal, P. F., Manney, G. L., Mittermeier, R. L., Pavlovic, B., and Wiacek, A.: Using FTIR measurements of stratospheric composition to identify midlatitude polar vortex intrusions over Toronto, J. Geophys. Res.-Atmos., 118, 12766–12783, https://doi.org/10.1002/2013JD020577, 2013.

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

Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Rettinger, M., Robinson, J., Roehl, C. M., Salawitch, R. J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R., Uchino, O., Warneke, T., and Wofsy, S. C.: A method for evaluating bias in global measurements of CO_2 total columns from space, Atmos. Chem. Phys., 11, 12317–12337, https://doi.org/10.5194/acp-11-12317-2011, 2011b.

Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, C., Feist, D. G., and Wennberg, P. O.: The Total Carbon Column Observing Network's GGG2014 Data Version, Tech. rep., Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, https://doi.org/10.14291/TCCON.GGG2014.DOCUMENTATION.R0/1221662, 2015.

Wunch, D., Mendonca, J., Colebatch, O., Allen, N. T., Blavier, J.-F., Roche, S., Hedelius, J., Neufeld, G., Springett, S., Worthy, D., Kessler, R., and Strong, K.: TCCON data from East Trout Lake, SK (CA), Release GGG2014.R1 (Version R1), TCCON Data Archive [data set], hosted by CaltechDATA, https://doi.org/10.14291/TCCON.GGG2014.EASTTROUTLAKE01.R1, 2018.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global Concentrations of CO_2 and CH_4 Retrieved from GOSAT: First Preliminary Results, Sola, 5, 160–163, https://doi.org/10.2151/sola.2009-041, 2009.