Supplemental information for:

Deep Transfer Learning Method for Seasonal TROPOMI XCH4 Albedo Correction

Alexander C. Bradley[1, 2]; Barbara Dix[2]; Fergus Mackenzie[3]; J. Pepijn Veefkind[4, 5]; and Joost A. de Gouw[1,2]

[1]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, 80309
[2]Chemistry Department, University of Colorado, Boulder, CO, 80309
[3]BlueSky Resources, Boulder CO, 80302, USA
[4]Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands
[5]Delft University of Technology, Delft, The Netherlands

Correspondence to: Joost de Gouw (Joost.deGouw@colorado.edu)

Contents:

Table S1. Variables with units used in the machine learning models.

Variable	Units
Solar Zenith Angle	°
Relative Azimuth Angle	°
XCH4	ppb
XCH4 Corrected	ppb
Ground Pixel	-
Surface Pressure	Pa
Pressure Interval	Pa
Surface Altitude	m
Surface Altitude Precision	m
XCH4 Apriori	ppb
Reflectance Cirrus VIIRS SWIR	-
XCH4 Precision	ppb
Fluorescence	mol s^{-1} m^{-2} nm^{-1} sr^{-1}
CO Column	Molecules cm^{-2}
CO Column Precision	Molecules cm^{-2}
H2O Column	Molecules cm^{-2}
H2O Column Precision	Molecules cm^{-2}
Aerosol Size	-
Aerosol Size Precision	-
Aerosol Height	m
Aerosol Height Precision	m
Aerosol Column	m^{-2}
Aerosol Column Precision	m^{-2}
Surface Albedo SWIR	-
Surface Albedo SWIR Precision	-
Surface Albedo NIR	-
Surface Albedo NIR Precision	-
Aerosol Optical Thickness SWIR	-
Aerosol Optical Thickness NIR	-
Chi Square SWIR	-
Chi Square NIR	-

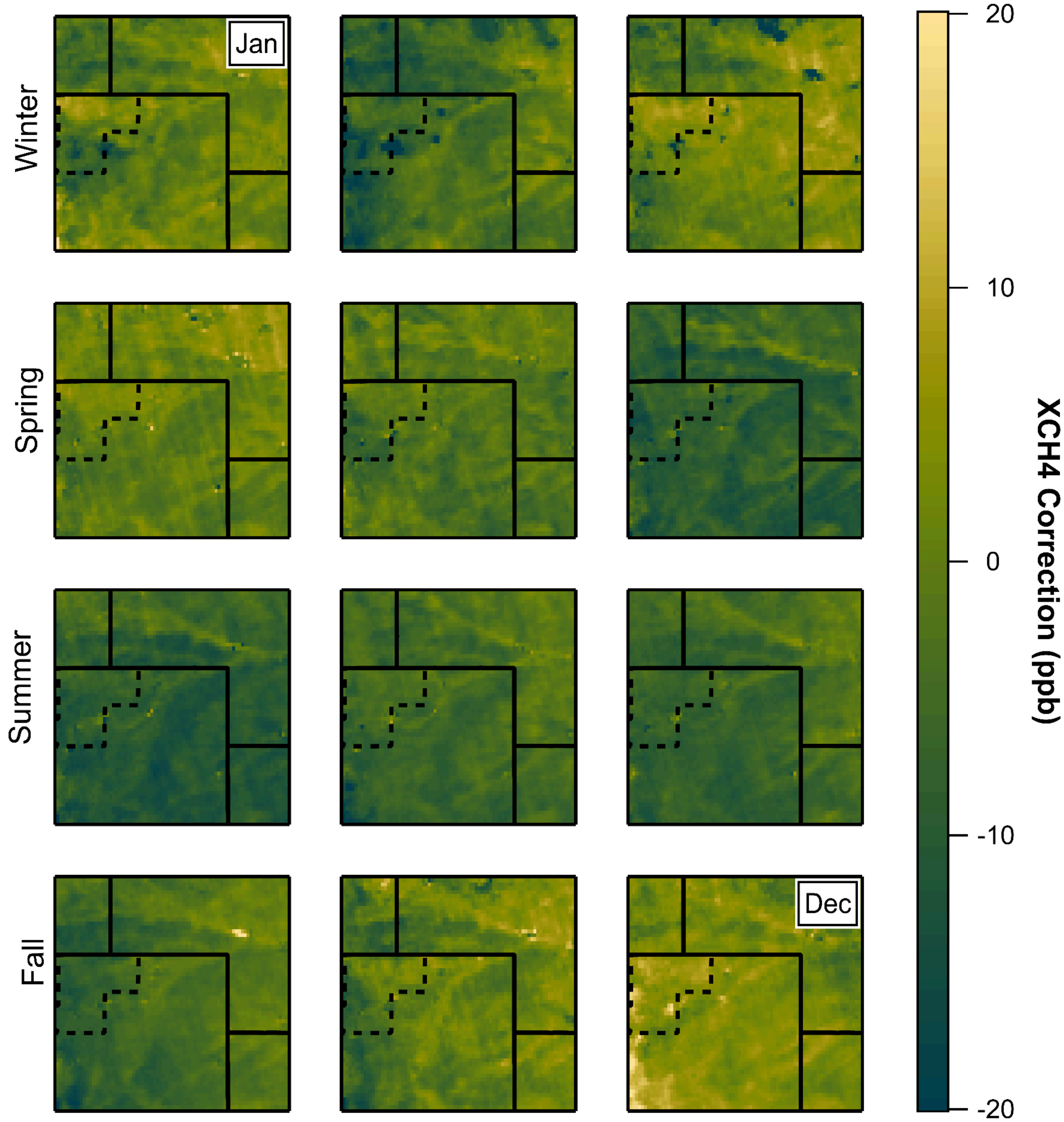

Figure S1. Monthly corrections for each month of the year 2019. Solid lines depict states in the region around the DJ basin (clockwise from top left: Wyoming, Nebraska, Kansas, Colorado). Lighter colors in colder months indicate more positive corrections while darker colors in warmer months indicate more negative corrections.

Some of the plots in **Fig S1** show structures that are being corrected for. Prominent around June, influences from the North and South Platte rivers can be seen as brighter lines cutting through darker regions. April has visually detectable data striping that is being corrected for

.

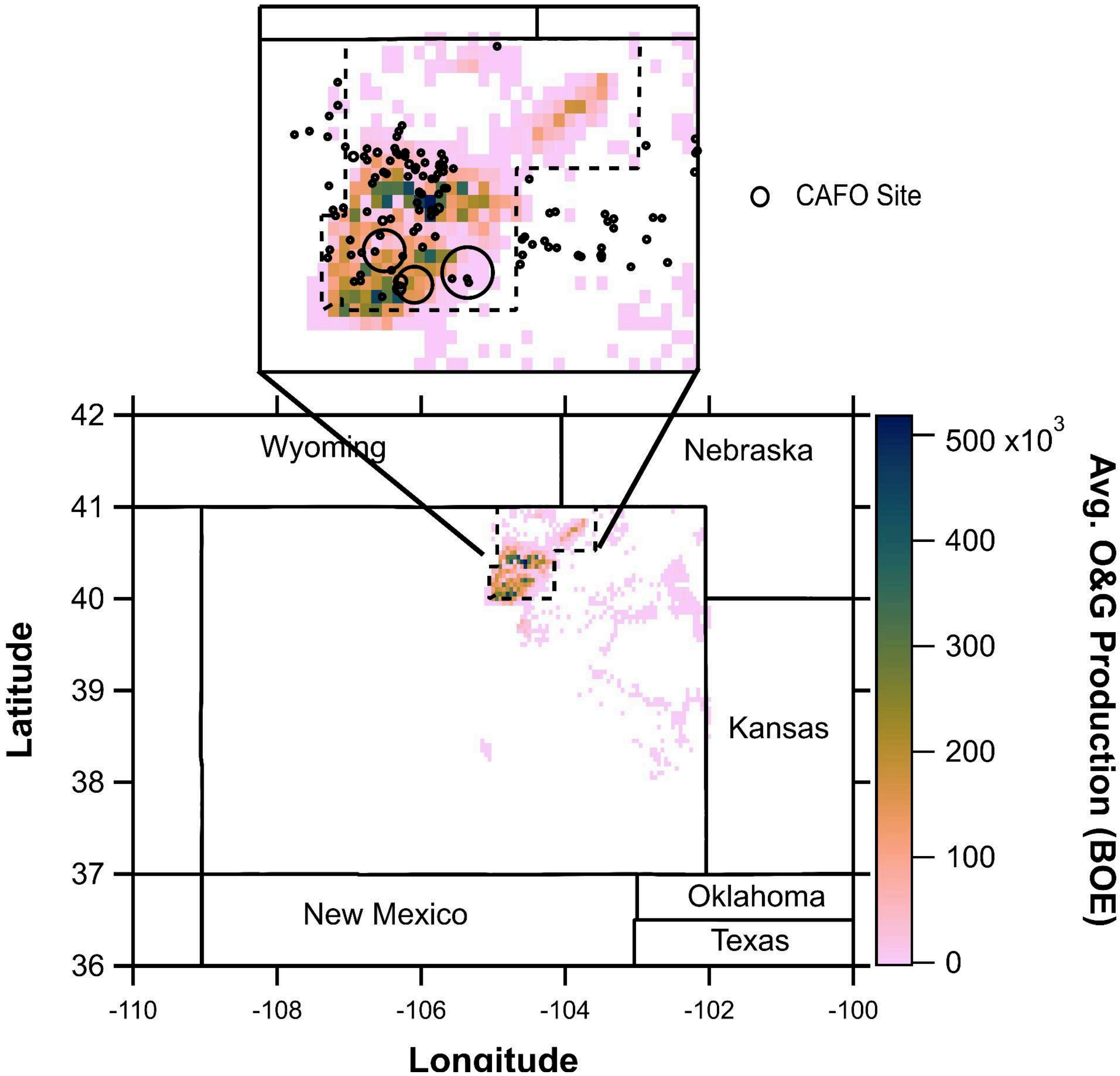

Figure S2. Drillinginfo production averages in northeast colorado. Data were downloaded on a statewide basis, so only production within Colorado is shown. Solid black lines represent state boundaries, dashed black lines represent Weld county Colorado. Zoomed in plot (top) depicts Weld county and adjacent Confined Animal Feeding Operations (CAFOs) with the size of the marker representing the relative maximum animal capacity of the operation, scaled by species.

Figure S2. shows where the Denver Julesburg basin sees the most production in Colorado. As stated in the main text, the highest production region of the DJ basin lies within Weld county, co-located with the agricultural powerhouse of Greely. Many Confined Animal Feeding Operations (CAFOs) are active in this area as well.

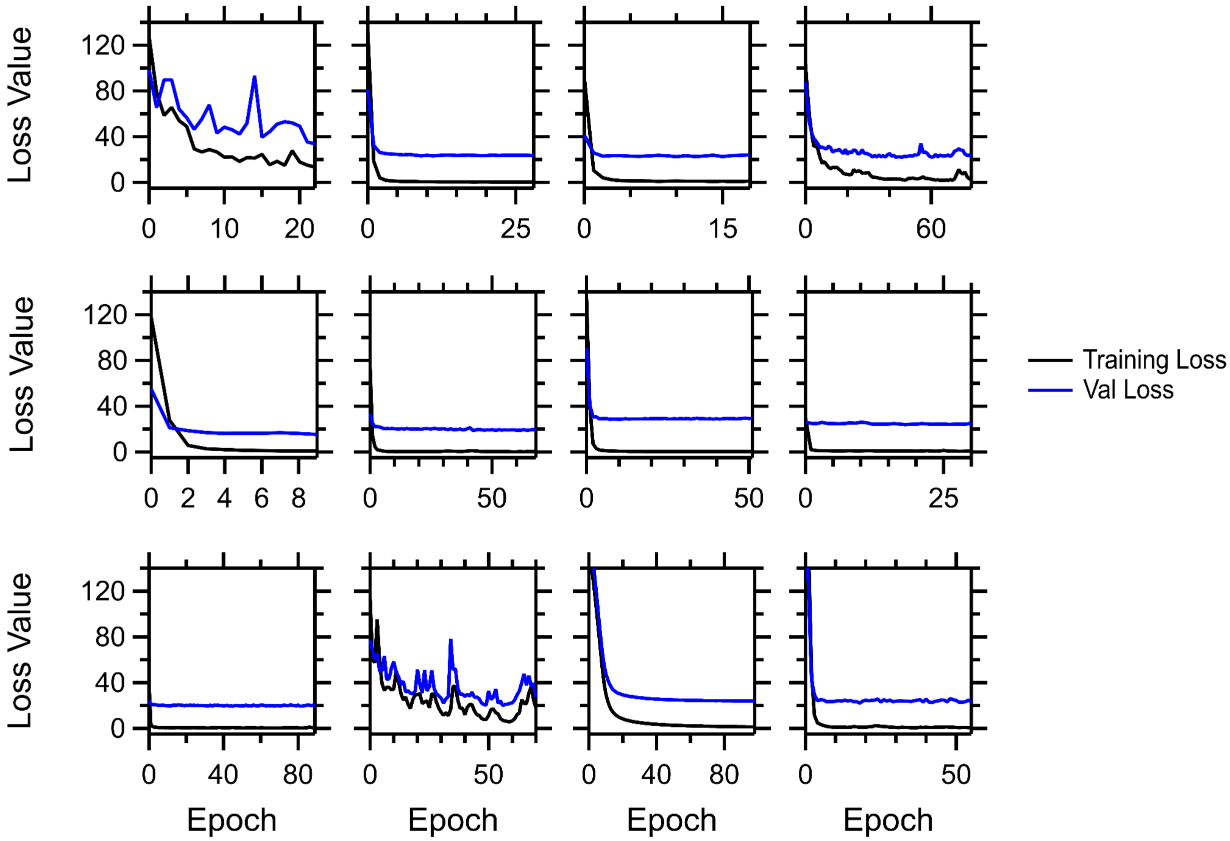

Figure S3. Training loss and validation loss for the monthly models. Training loss (black) shows how the MSE loss function is reduced through the training process, while validation loss (blue) shows how the validation dataset's loss progresses while the model is trained.

To ensure the monthly models are not over- or under-fitting the data, training and validation loss plots were constructed for the training process for each monthly model. The number of epochs for these training processes are generally quite low, but the lack of loss value fluctuation later in the training process demonstrates that more training is unnecessary. Models that are underfit will show significant gaps between the training and validation loss, and one could argue that the models presented here show some signs of underfitting. However, the solution to underfitting is generally to incorporate more data into the model, and the co-located GOSAt and TROPOMI points used for training were already stretched to their limit. An overfit model will appear in these plots as a divergence, generally later in the training process, of the training and validation losses. These models do not demonstrate any patterns of overfitting, suggesting that the models were trained accurately and successfully.

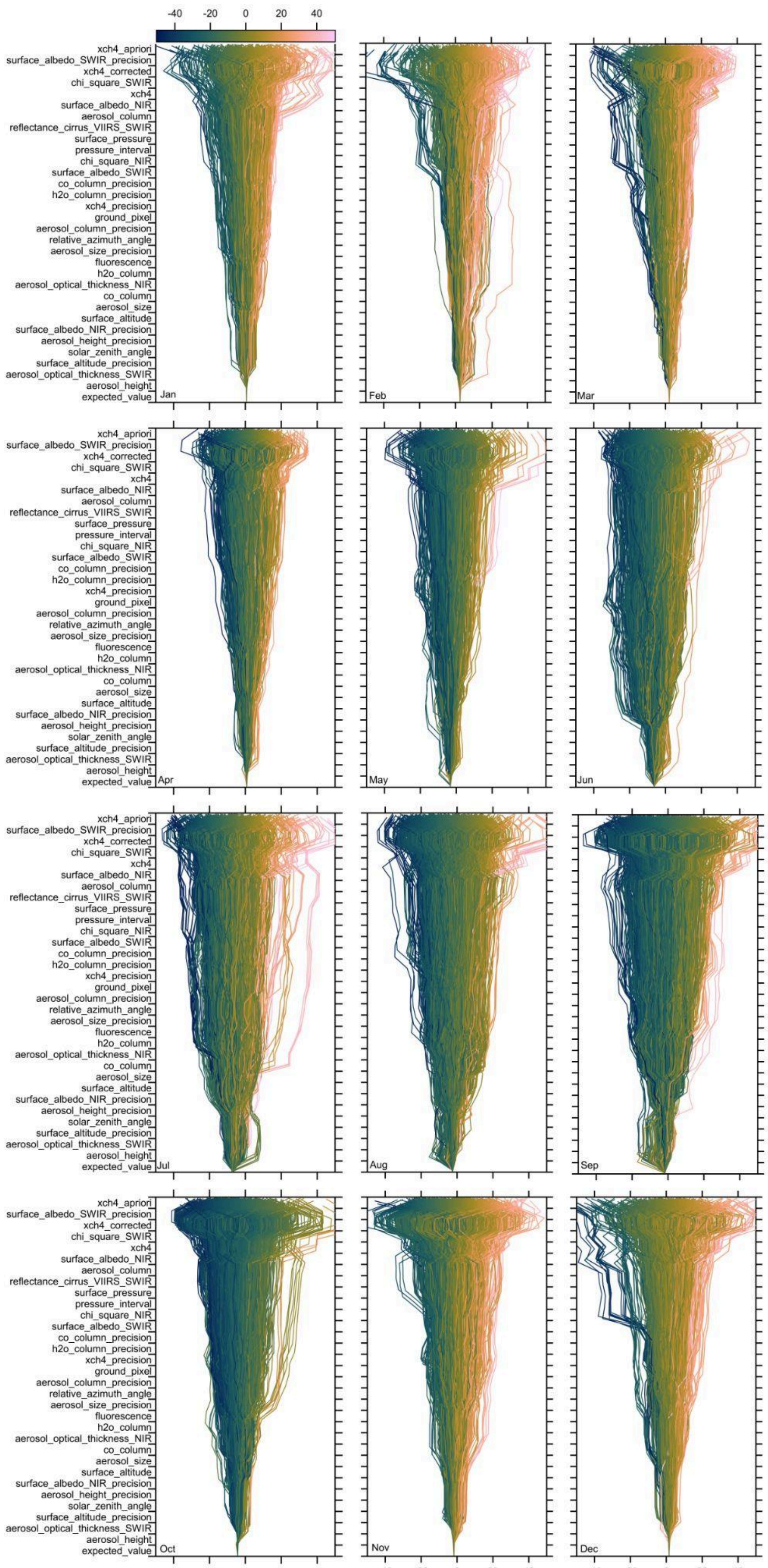
-40
-20
0
20
40
xch4_apriori
surface_albedo_SWIR_precision
xch4_corrected
chi_square_SWIR
xch4
surface_albedo_NIR
aerosol_column
reflectance_cirrus_VIIRS_SWIR
surface_pressure
pressure_interval
chi_square_NIR
surface_albedo_SWIR
co_column_precision
h2o_column_precision
xch4_precision
ground_pixel
aerosol_column_precision
relative_azimuth_angle
aerosol_size_precision
fluorescence
h2o_column
aerosol_optical_thickness_NIR
co_column
aerosol_size
surface_altitude
surface_albedo_NIR_precision
aerosol_height_precision
solar_zenith_angle
surface_altitude_precision
aerosol_optical_thickness_SWIR
aerosol_height
expected_value
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

Figure S4. Decision plots for all models. Variables are ordered from top to bottom by importance in January. Color scale indicates the final model output value, which is the Δ(TROPOMI – GOSAT) value. Expected value is the average prediction made by the model across all possible combinations of features, and is thus the same value for all trials using the same model.

Decision plots for models 1 and 7 appear in the main text to contrast the difference between the decision processes. As shown in **Fig. S4** the trend continues throughout the year with each model being slightly different, but the models of warmer months being similar to other warmer months and colder months being more similar to other colder months. The overall trend is both in color and in general shape of the decision plots. Warmer months appear darker, corresponding with the more negative correction values observed during this time period, while colder months appear more neutral or lighter. The shape of the decision plots indicates how the model is making decisions, and the winter months all appear to be shaped more conically, while the summer months are more rounded near the bottom, appearing more cylindrical. The different shapes indicate that the models in winter and in summer are making decisions about the correction variables in different ways.

Section S1 Additional discussion of the ML training process

The hyperparameters included the learning rate, batch size, number of epochs, and a unique hyperparameter: number of frozen layers. The original base model was allowed to be retrained on a limited basis, with the number of frozen layers hyperparameter describing how many of the base models' hidden layers would not be retrainable. All the monthly models had additional hidden layers added that were untrained on top of the pretrained base model layers. The Optuna algorithm determined the values of the variables to use through testing of multiple values across 25 model training runs, identifying the most effective values. The model was validated based both in the traditional loss function reduction and in its performance in reducing the dependence between corrected TROPOMI / GOSAT data and surface albedo SWIR. The process is described as an albedo correction due to this validation method despite taking into account many other factors.

Python Package version numbers:

- matplotlib: 3.7.2
- tensorflow: 2.15.0
- keras: 2.15.0
- xgboost: 1.7.6
- lightgbm: 4.0.0
- catboost: 1.2.2
- pandas: 2.1.0
- geopandas: 0.13.2
- optuna: 3.5.0
- FLAML: 2.1.1
- scipy: 1.11.3
- shapely: 2.0.1
- numpy: 1.26.0
- netcdf4: 1.6.5
- tqdm: 4.66.1
- rasterio: 1.3.6