train 12 separate monthly models, starting from the annual base model; a schematic representing this training process is presented in Fig. 1. The remaining 20 % of the data were used to calculate final fit statistics for each monthly model. There is an unequal distribution of points across the months, which introduces some seasonal bias to the annual model. This bias is subsequently removed when the monthly models are trained. DTL is especially suited to this type of learning because (1) the initial learning phase trained on all training data helps the lower levels of the model learn to generalize the task and (2) the subsequent training occurs on much smaller monthly training datasets that help train higher, more specific levels of the model. Various hyperparameters were tuned using Optuna, a hyperparameter tuning package for Python (Takuya et al., 2023). To monitor against overfitting, training and validation loss for the training period of each model were calculated and are presented in Fig. S3.

2.3 Python

All of the computations were completed using Python, a general-use, interpreted, object-oriented programming language ideal for building and implementing machine learning models and algorithms. A number of third-party packages were useful in the computations completed in this work: Matplotlib for figure generation (Hunter, 2007); the machine learning packages TensorFlow (Abadi et al., 2015), Keras (Chollet et al., 2015), XGBoost (Chen and Guestrin, 2016), LightGBM (Ke et al., 2017), and CatBoost (Dorogush et al., 2018); Pandas and GeoPandas for tabular and geospatial data organization (Jordahl et al., 2020; The pandas development team, 2024); Optuna and Fast and Lightweight AutoML Library (FLAML) for tuning ML models hyperparameters (Takuya et al., 2023); SciPy for scientific and statistical functions (Virtanen et al., 2020); Shapely for manipulation of geometric objects (Gillies et al., 2007); NumPy for array manipulation (Harris et al., 2020); netcdf4 for opening and reading satellite data (Whitaker, 2008); Rasterio for raster manipulation (Gillies et al., 2013); and tqdm to visualize data processing progress (da Costa-Luis, 2019). Figure 1 was created using the Google Drawings suite, Fig. 2 was created using Python Matplotlib, and Figs. 3–7 were created in Igor Pro 8.04.

2.4 Other geospatial data

River paths and extent data for the South Platte River and North Platte River were downloaded from NOAA (National Weather Service, 2024). Crop data were downloaded from CropScape, a geospatial thematic agricultural mapping software (Han et al., 2014). Cartographic shapefiles containing state, county, and urbanized area boundary lines were downloaded from the US Census Bureau (2024). Finally, data visualizations were made to be color-accessible by Fabio Crameri's scientific color maps (Crameri et al., 2020).

3 Results and discussion

The seasonal biases of the current TROPOMI operational product, which includes the albedo correction from Lorente et al., are studied in Fig. 2 for the area of interest. Figure 2 shows the ratio between co-located GOSAT and TROPOMI methane retrievals as a function of surface albedo in the shortwave infrared. In the ideal case, these ratios are equal to 1 and there is no correlation between this ratio and surface albedo ($R = 0$). When all data are used (Fig. 2a) the Pearson correlation is indeed calculated to be low, i.e., below a threshold of 0.1, which we chose here as a target value for minimal correlation between SWIR surface albedo and the albedo corrected methane retrieval. Though the significance of Pearson coefficients is up to interpretation, most would agree that a value of < 0.1 signifies negligible correlation (Akoglu, 2018; Schober et al., 2018). When the data are shown by season, this is no longer true – Pearson correlations with an absolute value greater than 0.1 indicate that there exists some correlation between the SWIR surface albedo and the albedo corrected methane retrieval. The Lorente et al. correction algorithm does account for some seasonality because the TROPOMI-retrieved variables include the surface albedo SWIR which is used to calculate a correction value. However, the seasonal correlation reappears after the Lorente et al. correction because this correction assumes that the relationship between surface albedo SWIR and correction value is static over time. Figure 2b and c demonstrate the change in surface albedo as a function of season, with the density of counts shifting from the left side of the plot, indicating smaller albedos, to the center of the plot, indicating higher albedos on average from summer to winter. The different seasons also have different directions of change, with summers having an inverse correlation and winters having a positive correlation. The QA value used in processing the TROPOMI retrieval data retained high-quality snow-covered scenes, so some of this shift could be attributed to the SWIR reflectance of snow over bare soil. Regardless of the reason, the shifting albedo and seasonally variable albedo effect biases methane retrieval data from TROPOMI at finer timescales. In order to correct for this bias we employed a DTL neural network machine learning algorithm.

3.1 Model evaluation

The DTL neural network models were trained and evaluated as described in Sect. 2.2 and compared against the uncorrected methane retrieval, the Lorente et al. corrected methane retrieval, and the blended TROPOMI–GOSAT product produced by Balasus et al., commonly referred to as the "Harvard dataset", for their effectiveness in methane correction. To evaluate against the other models, Pearson correlations were calculated and presented in Fig. 3a, where different constructions of Pearson values have been unified according to Table S2. Pearson correlations have been calculated the