# Supplement Material

1. Sampling

Figure 1a demonstrates that the area ratio of sampled to ice-free grids of 1×1 degree is less than 1% in most months in 1980s. The largest ratio is less than 7% in 2010s. The sample distribution map of 1981 shows that most samples in the year were taken in the north Atlantic in warm seasons (Fig. 1b). Even when every month had samples as in 2017, much more samples were taken in the northern hemisphere and in warm seasons (Fig. 1c).
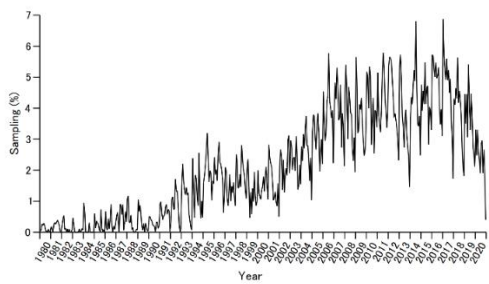

(a)
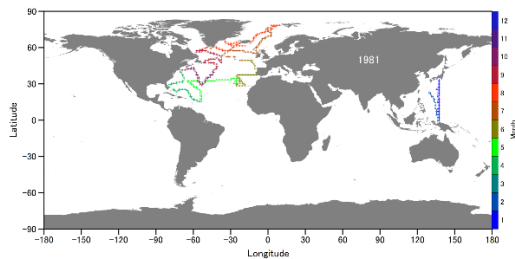
(b)
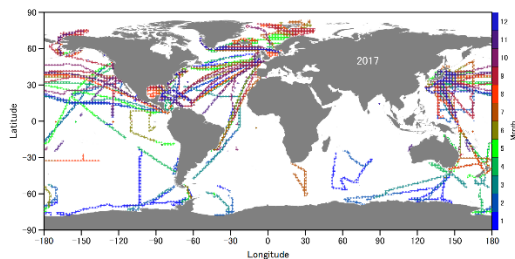
(c)

Figure 1. Sampling statistics and distribution examples. (a) area ratio of monthly sampled grids to all ice-free grids in the SOCAT 2021 product with a spatial resolution of 1x1 degree. (b) sample distribution in 1981. (c) sample distribution in 2017.

2. LOYO Test

We evaluated the leave-one-year-out (LOYO) validation method by using it with the raw ocean CO2 data. LOYO means to set aside one year's data for validation and use the rest to train the models. The residues of prediction minus observation show a significant ($R^2=0.971$) negative correlation with time (Figure 2a). As we expected, it can detect the increase of CO_2 with time because the CO_2 in early years tends to be overestimated by the models trained with increased CO_2 in later years and vis versa. The negative slope (-1.352 µatm yr-1) can be considered as an approximation of the increase trend. We normalized the data using the trend and repeated the process. The second around yielded a regression slope of -0.208 µatm yr-1 with $R^2=0.495$ (Figure 2b). The second around repeated the process with data normalized by the sum of previous trends. After three rounds, the slope became negligible (Figure 2c).
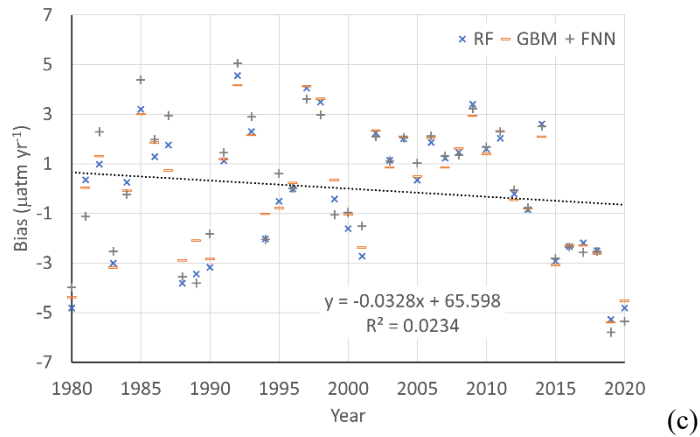

(a)

(b)

(c)
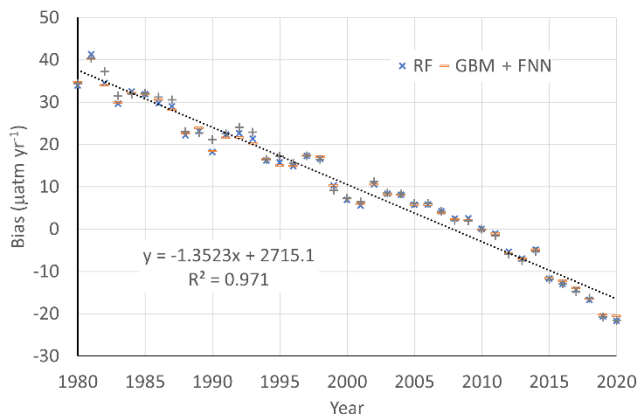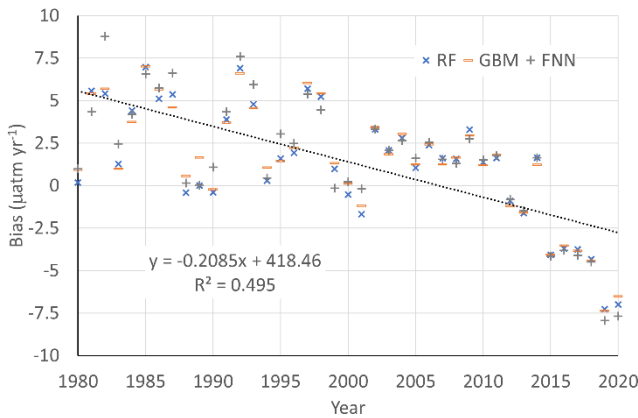
Figure 2. Biases resulted from LOYO validations. (a) using LOYO with raw CO2 data; (b) with data normalized by the trend in (a); (c) with data normalized by the trend in (a) and (b).

3. Trend Adjustment

We applied the LOYO validation method to the normalized CO_2 data using the rates obtained by the iteration method (Fig. 1 of the paper). A small trend (0.079 $\mu tm\ yr^{-1}$) exits in the residues of prediction minus normalized CO_2. The trend was added to the rates to form the final rates in Table 2 in the paper.
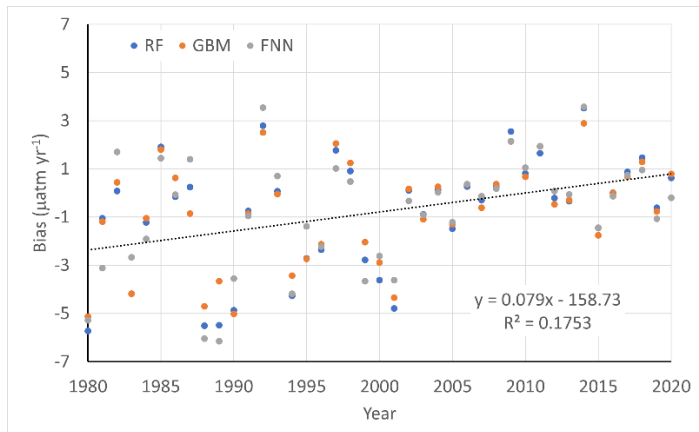

Figure 3. Residues resulted from using LOYO validation with CO2 data normalized by the rates shown in Fig. 1 of the paper.

4. Performance

We used the results of LOYO validation as the performance metric. The correlation coefficient indicates that the RF and GBM models performed equally well and were better than the FNN model (Fig. 4). The overall all biases of predictions minus normalized observations are negligible comparing to the variability ocean CO_2. Note that biases in Fig. 4 equals the weighted mean biases by data points in Table 2 of the paper.
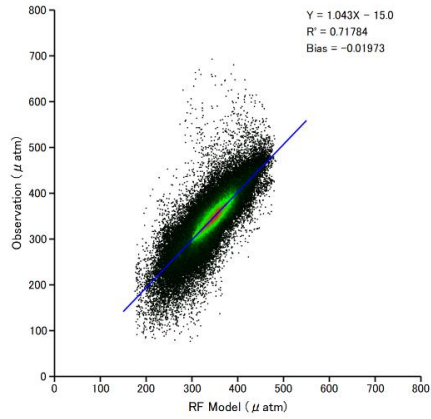
(a)
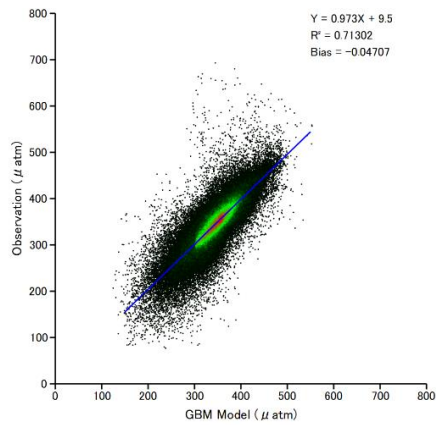
(b)
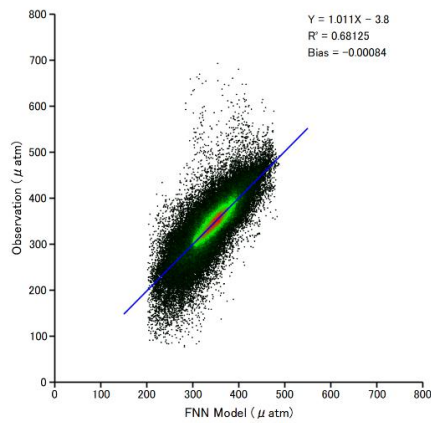
(c)

Figure 4. Predicted vs observed CO_2 resulted from LOYO validations with normalized CO_2 data by the rates in Table 2 of the manuscript.

5. Python Code for Model Initialization

```
#!/usr/local/bin/python3.9


from sklearn.neural_network import MLPRegressor

from skranger.ensemble import RangerForestRegressor

from lightgbm import LGBMRegressor


def regressor(model):

    if (model == 'nn'):

        reg = MLPRegressor(hidden_layer_sizes=(256,),activation='logistic',random_state=1,max_iter=300)

    elif (model == 'rf'):

        reg = RangerForestRegressor(n_estimators=500,min_node_size=100,n_jobs=6)

    elif (model == 'gb'):

        reg = LGBMRegressor(n_estimators=500,num_leaves=100,random_state=1,n_jobs=6)

    return reg
```