Temporal Prediction of Net Ecosystem Exchange (NEE) by Transformer Model

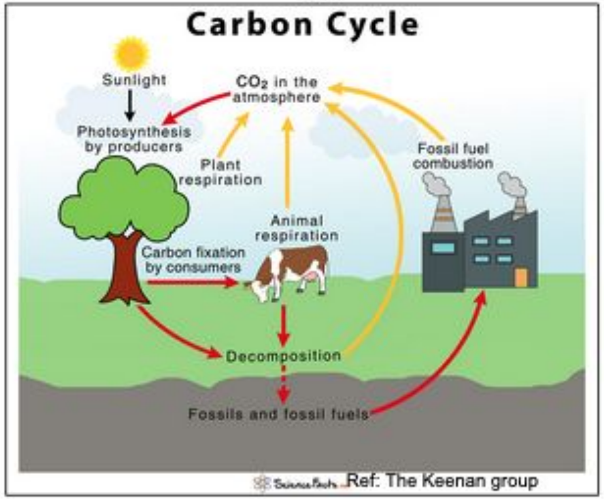
MIDS 2023 Spring - DATASCI 210
Sweta Bhattacharya
Joshua Dunn
Marcia (Yiying) Liu

Berkeley
UNIVERSITY OF CALIFORNIA

Problem Statement

- **Net ecosystem carbon exchange (*NEE*)** measures the carbon interchanges between the Earth's biosphere and the atmosphere.The best result would be a negative value.

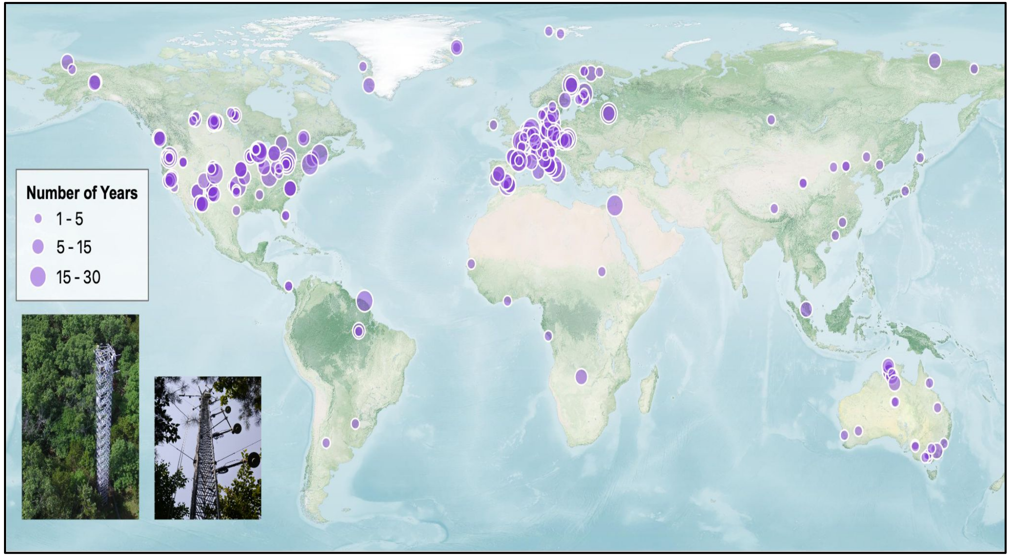
$$GPP = R_{eco} - NEE$$

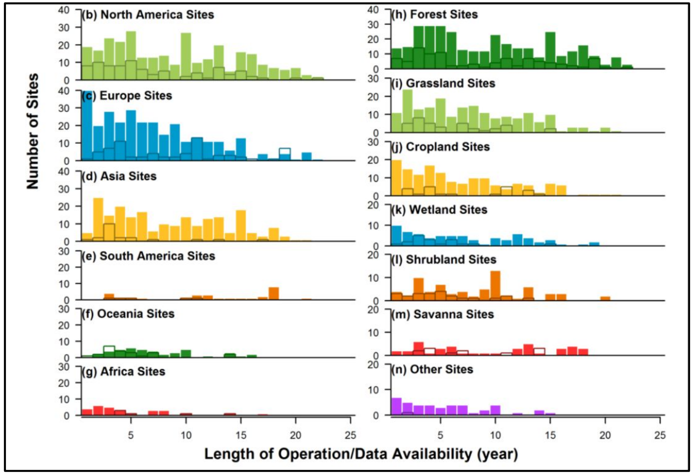


- ❑ Factors affecting carbon flux: Climate, vegetation, soil, etc.
- ❑ NEE derived from Eddy Covariance Measurements
- ❑ Sparse local measurements.
- ❑ Need improved models with better accuracy

Datasets - FluxNet

- Global distribution of FluxNet Eddy Covariance data
- Data source for target variables: GPP and NEE
- 276 sites worldwide from 2001-2020





Vegetation Groups by IGBP

Forest: ENF, DBF, EBF, MF
Grassland: GRA
Cropland: CRO, CVM
Wetland: WET
Shrubland: OSH, CSH
Savanna: SAV, WSA
Other: BSV, URB, WAT, SNO

Correlation and Feature Selection

<u>FluxNet data:</u>

** NEE_VUT_REF ⟵ Target Variable

GPP_NT_VUT_REF

RECO_NT_VUT_REF

RECO_DT_VUT_REF

<u>Climate data:</u>

TA_ERA (Temperature)

P_ERA (Precipitation)

VPD_ERA (Vapor Pressure)

SW_IN_ERA (solar radiation)

<u>Remote sensing data:</u>

b1,b2,b3,b4,b5,b6,b7

NDVI

EVI

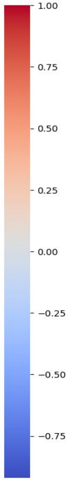
NIRv

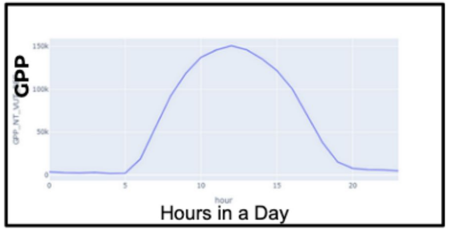
IGBP – Categorical

Koppen – Categorical

Site_ID – Categorical

	NEE_VUT_REF	NEE_VUT_REF_QC	NEE_CUT_REF	NEE_CUT_REF_QC	GPP_NT_VUT_REF	GPP_DT_VUT_REF	GPP_NT_CUT_REF	GPP_DT_CUT_REF	RECO_NT_VUT_REF	RECO_DT_VUT_REF	RECO_NT_CUT_REF	RECO_DT_CUT_REF	NEE_VUT_REF_qa	b1	b2	b3	b4	b5	b6	b7
NEE_VUT_REF	1.00	0.19	0.99	0.19	-0.92	-0.84	-0.91	-0.84	-0.16	-0.18	-0.16	-0.18	1.00	0.08	-0.09	0.06	0.06	-0.07	0.02	0.05
NEE_VUT_REF_QC	0.19	1.00	0.18	0.90	-0.18	-0.18	-0.18	-0.18	-0.05	-0.04	-0.06	-0.05	0.23	0.03	0.00	0.04	0.04	-0.04	-0.06	-0.05
NEE_CUT_REF	0.99	0.18	1.00	0.19	-0.91	-0.84	-0.92	-0.84	-0.16	-0.18	-0.16	-0.18	0.99	0.07	-0.09	0.06	0.06	-0.07	0.02	0.05
NEE_CUT_REF_QC	0.19	0.90	0.19	1.00	-0.18	-0.19	-0.18	-0.19	-0.05	-0.05	-0.04	-0.05	0.22	0.02	0.00	0.03	0.03	-0.05	-0.07	-0.07
GPP_NT_VUT_REF	-0.92	-0.18	-0.91	-0.18	1.00	0.91	0.99	0.91	0.52	0.47	0.52	0.47	-0.92	-0.19	0.16	-0.16	-0.15	0.15	-0.05	-0.13
GPP_DT_VUT_REF	-0.84	-0.18	-0.84	-0.19	0.91	1.00	0.91	0.99	0.49	0.53	0.49	0.52	-0.84	-0.20	0.17	-0.16	-0.15	0.16	-0.05	-0.13
GPP_NT_CUT_REF	-0.91	-0.18	-0.92	-0.18	0.99	0.91	1.00	0.91	0.52	0.47	0.52	0.47	-0.91	-0.19	0.16	-0.16	-0.15	0.15	-0.05	-0.13
GPP_DT_CUT_REF	-0.84	-0.18	-0.84	-0.19	0.91	0.99	0.91	1.00	0.49	0.52	0.49	0.53	-0.84	-0.20	0.16	-0.17	-0.15	0.16	-0.05	-0.13
RECO_NT_VUT_REF	-0.16	-0.05	-0.16	-0.05	0.52	0.49	0.52	0.49	1.00	0.80	0.98	0.79	-0.17	-0.33	0.22	-0.28	-0.26	0.22	-0.08	-0.21
RECO_DT_VUT_REF	-0.18	-0.04	-0.18	-0.05	0.47	0.53	0.47	0.52	0.80	1.00	0.79	0.95	-0.18	-0.31	0.21	-0.26	-0.25	0.23	-0.07	-0.20
RECO_NT_CUT_REF	-0.16	-0.06	-0.16	-0.04	0.52	0.49	0.52	0.49	0.98	0.79	1.00	0.80	-0.17	-0.33	0.21	-0.28	-0.26	0.22	-0.08	-0.22
RECO_DT_CUT_REF	-0.18	-0.05	-0.18	-0.05	0.47	0.52	0.47	0.53	0.79	0.95	0.80	1.00	-0.18	-0.32	0.21	-0.27	-0.25	0.23	-0.07	-0.20
NEE_VUT_REF_qa	1.00	0.23	0.99	0.22	-0.92	-0.84	-0.91	-0.84	-0.17	-0.18	-0.17	-0.18	1.00	0.08	-0.09	0.06	0.06	-0.07	0.02	0.06
b1	0.08	0.03	0.07	0.02	-0.19	-0.20	-0.19	-0.20	-0.33	-0.31	-0.33	-0.32	0.08	1.00	0.60	0.96	0.98	0.19	0.09	0.17
b2	-0.09	0.00	-0.09	0.00	0.16	0.17	0.16	0.16	0.22	0.21	0.21	0.21	-0.09	0.60	1.00	0.59	0.65	0.66	0.16	0.03
b3	0.06	0.04	0.06	0.03	-0.16	-0.16	-0.16	-0.17	-0.28	-0.26	-0.28	-0.27	0.06	0.96	0.59	1.00	0.99	0.05	-0.14	-0.06
b4	0.06	0.04	0.06	0.03	-0.15	-0.15	-0.15	-0.15	-0.26	-0.25	-0.26	-0.25	0.06	0.98	0.65	0.99	1.00	0.15	-0.04	0.03
b5	-0.07	-0.04	-0.07	-0.05	0.15	0.16	0.15	0.16	0.22	0.23	0.22	0.23	-0.07	0.19	0.66	0.05	0.15	1.00	0.76	0.58
b6	0.02	-0.06	0.02	-0.07	-0.05	-0.05	-0.05	-0.05	-0.08	-0.07	-0.08	-0.07	0.02	0.09	0.16	-0.14	-0.04	0.76	1.00	0.93
b7	0.05	-0.05	0.05	-0.07	-0.13	-0.13	-0.13	-0.13	-0.21	-0.20	-0.22	-0.20	0.06	0.17	0.03	-0.06	0.03	0.58	0.93	1.00





NEE = RECO - GPP

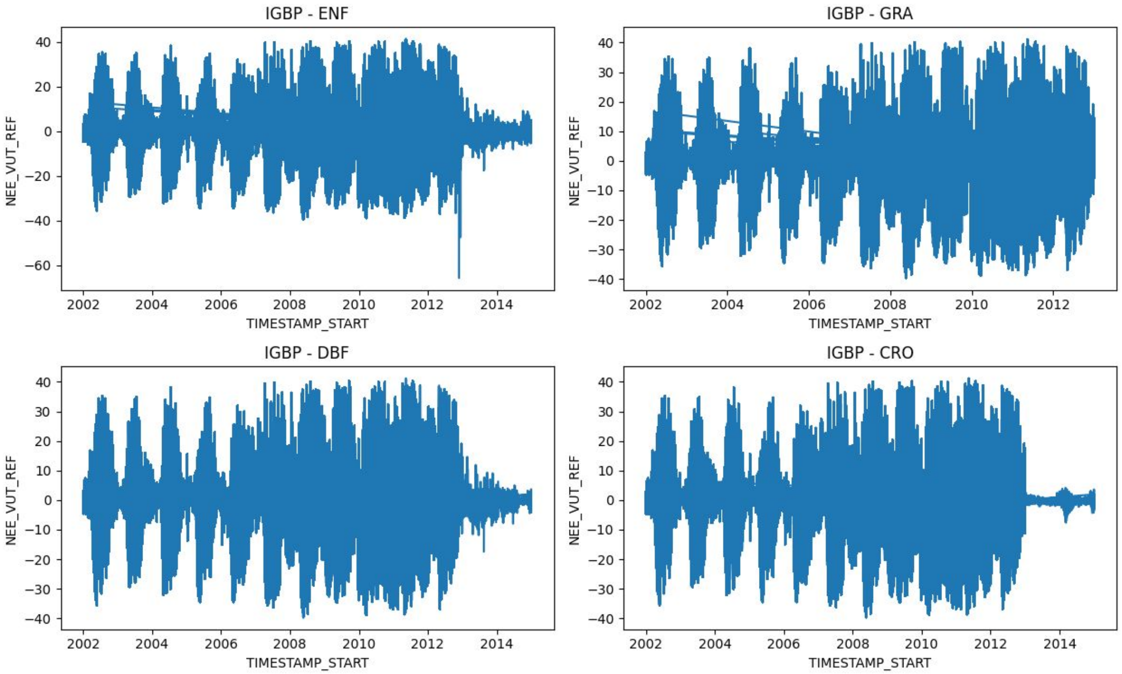
IGBP Schema (The International Geosphere–Biosphere Programme)

Evergreen Needleleaf Forests
(ENF): 9,933,360 rows

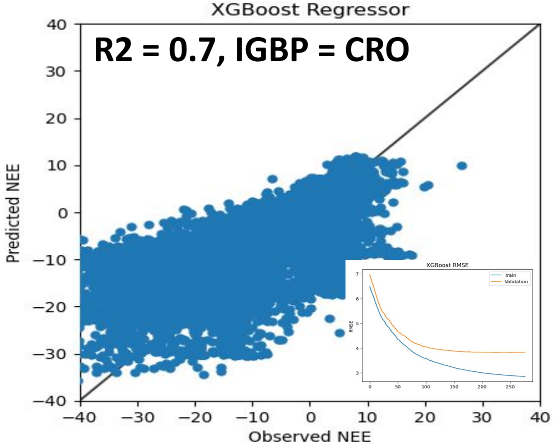
Grasslands
(GRA): 4,771,104 rows

Deciduous Broadleaf Forests
(DBF): 4,312,512 rows

Croplands
(CRO): 3,607,104 rows



XGBoost Model

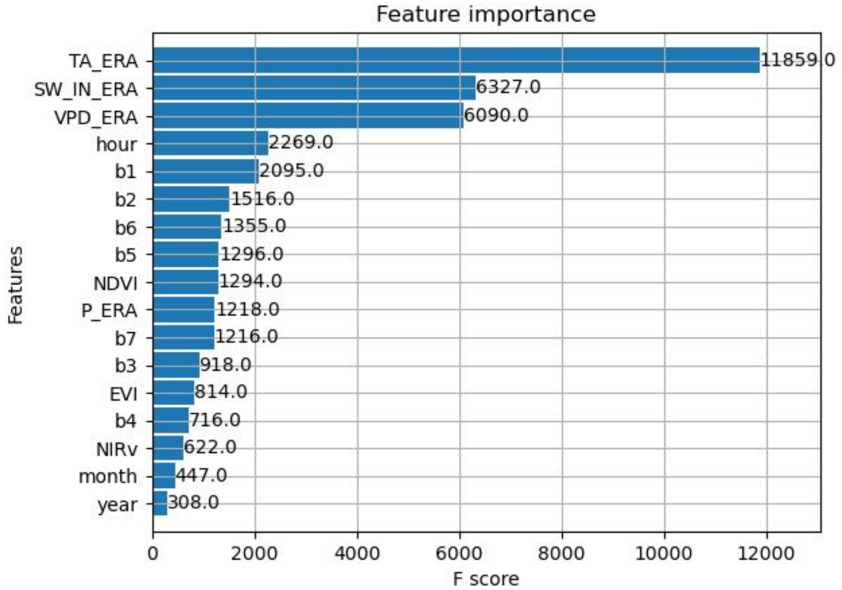


<u>Baseline Model:</u>

Training and validation Data: 80-20 split on **Site Ids**

TimeSeries Data Frequency: half-hourly

Evaluation Metrics: **MAE, RMSE, and R^2**

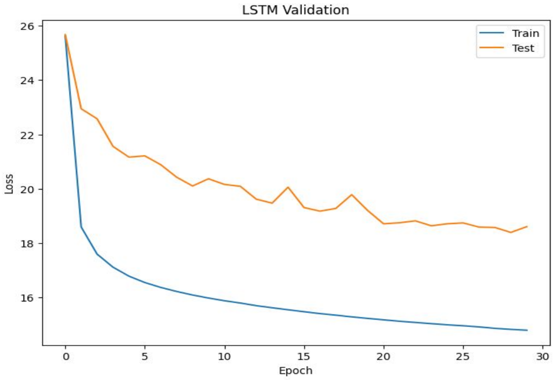


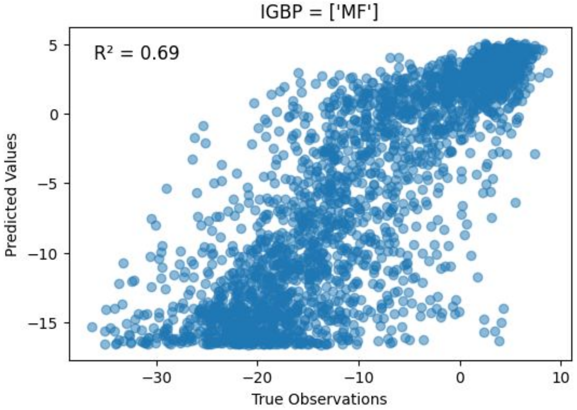
Current State-of-the-art

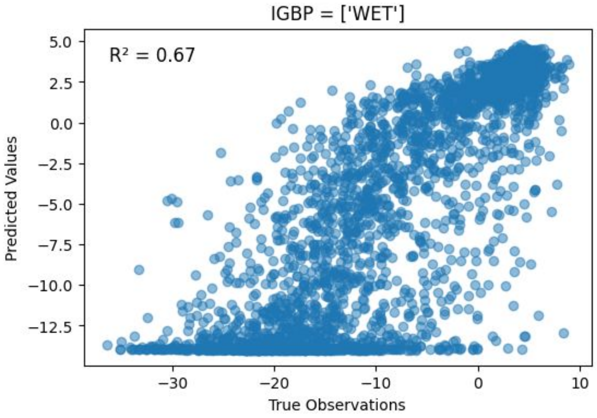
		XGB
12	DBF	0.80
5	EBF	0.58
17	ENF	0.79
4	MF	0.67
15	GRA	0.57
5	CRO	0.43
2	OSH	0.27
1	CSH	0.73
4	SAV	0.61
3	WSA	0.50

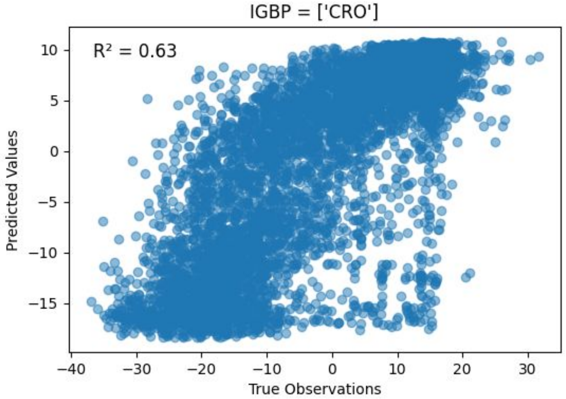
Liu, J.; Zuo, Y.; Wang, N.; Yuan, F.; Zhu, X.; Zhang, L.; Zhang, J.; Sun, Y.; Guo, Z.; Guo, Y.; Song, X.; Song, C.; Xu, X. Comparative Analysis of Two Machine Learning Algorithms in Predicting Site-Level Net Ecosystem Exchange in Major Biomes. *Remote Sens.* **2021**, *13*, 2242. https://doi.org/10.3390/rs13122242

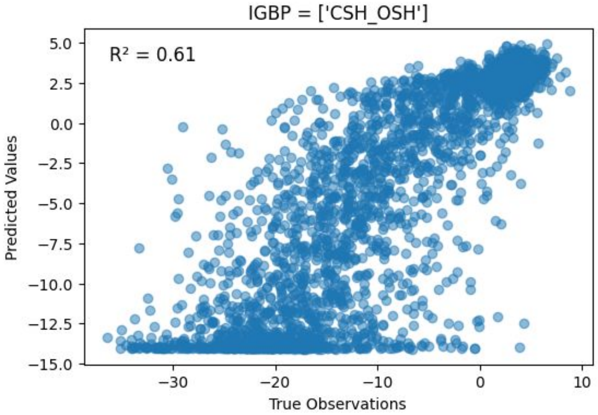
LSTM Model Performance by IGBP

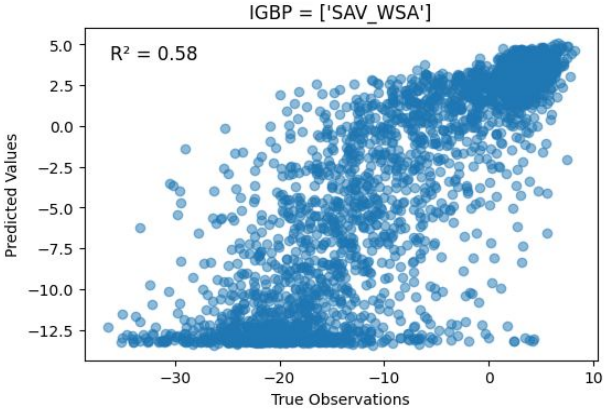












Enter Temporal Fusion Transformer

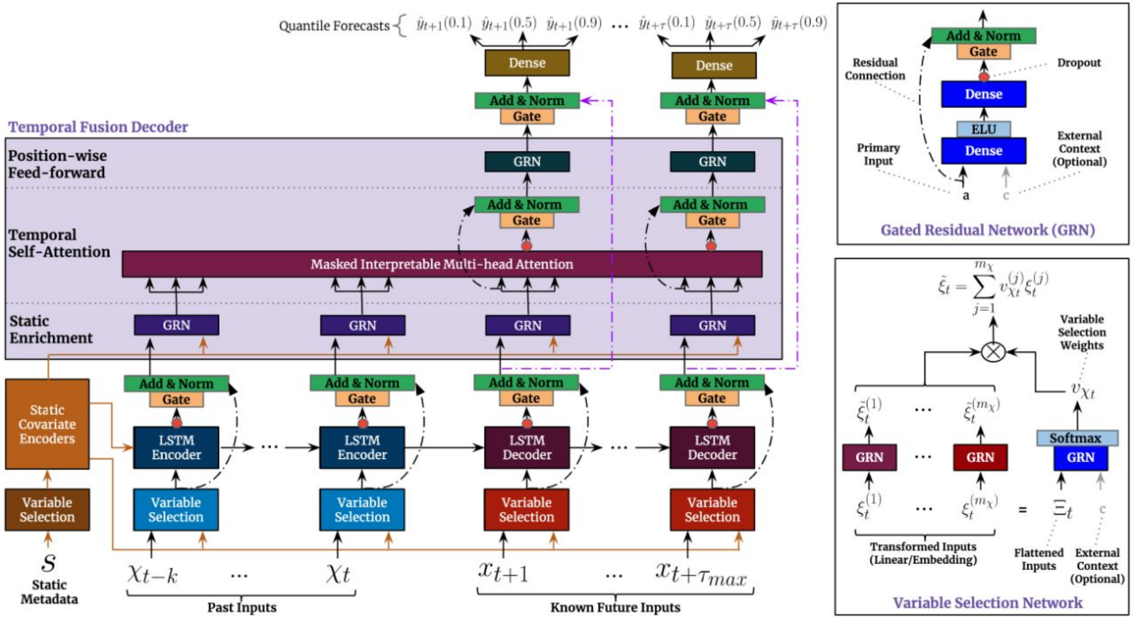
Temporal Fusion Transformer (TFT) is Google’s state-of-the-art transformer model for series prediction.

We believed we might be able to achieve better results using a more modern architecture such as TFT.

Temporal Fusion Transformer (TFT)

- Multi-head Attention Layer
- LSTM Encoders/Decoders
- Variable Selection Network
- Gated Residual Network

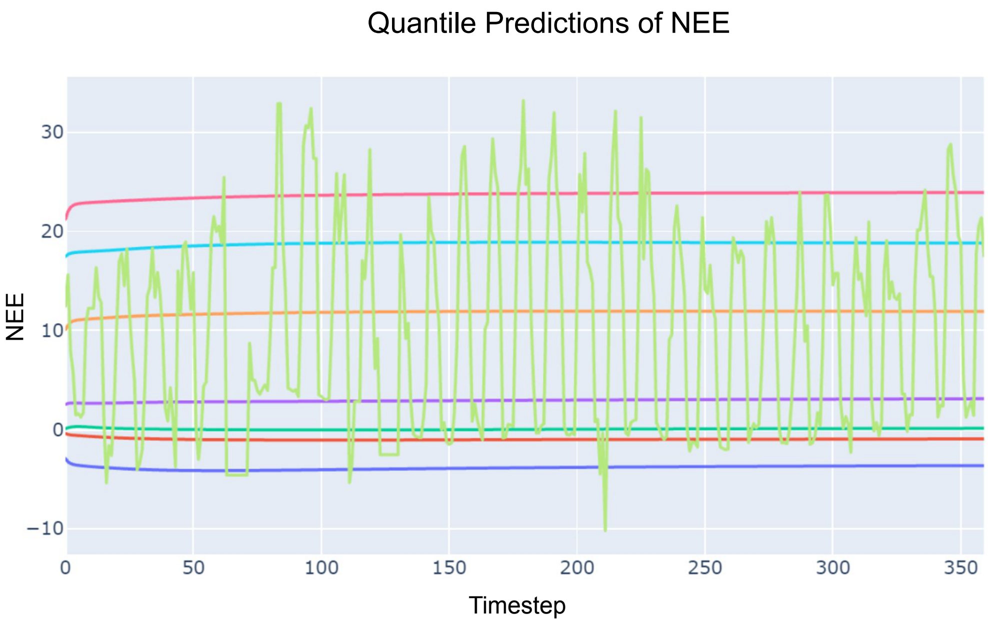
Produces series-to-series predictions.



Initial TFT Model Development

The initial model was designed to provide:

- Long-term predictions
  - 1,440 half-hour predictions
- 6 NEE and GPP target predictions
- 7 quantile values per target
- Removing rows with missing values



But the results were pretty bad!

Final TFT Model Implementation

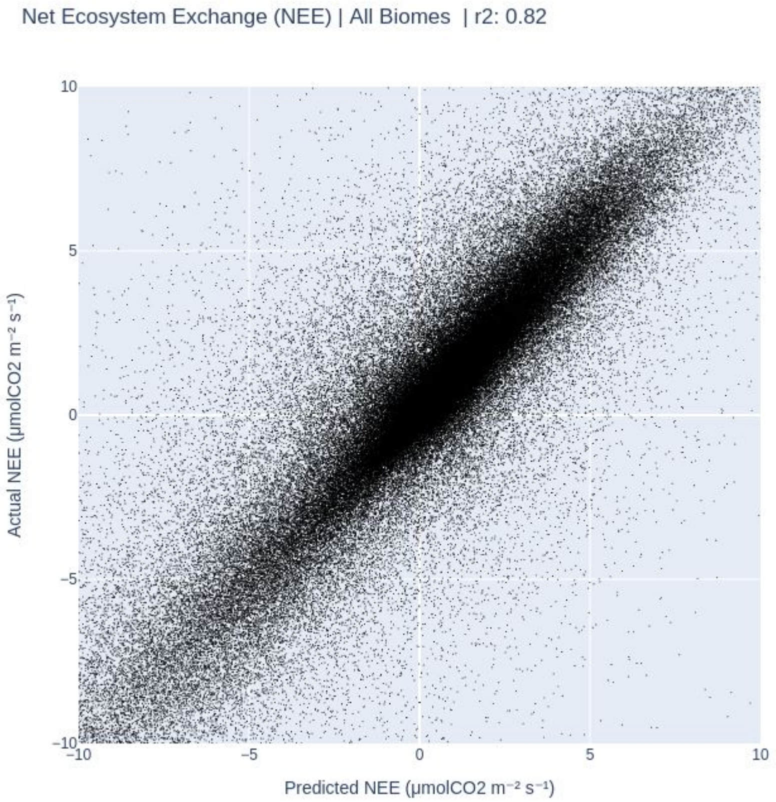
After much collaboration with the subject matter experts and our instructors, we pulled back on the scope of our model significantly.

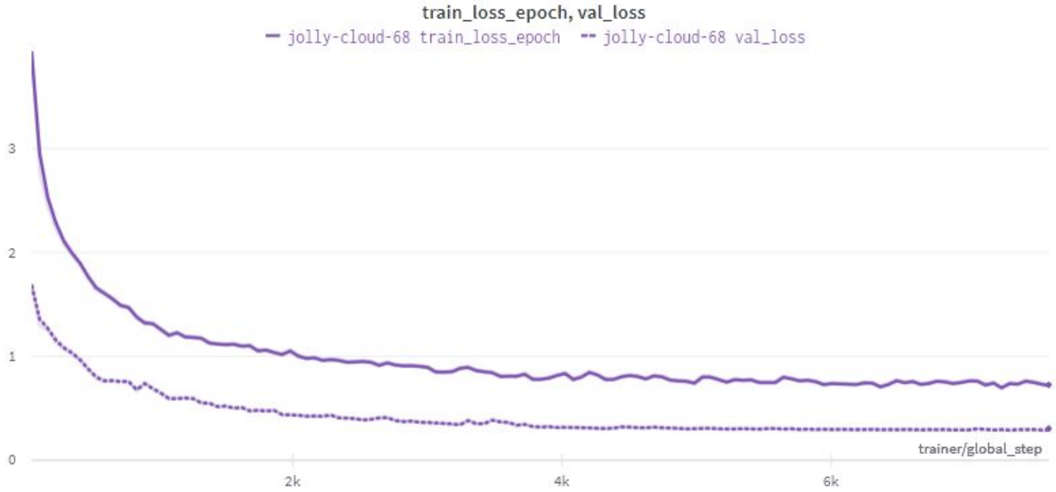
- The model now provides:
  - Only 1 variable prediction (NEE)
  - No quantile predictions
  - No long term prediction
    - Time horizon = 1 half-hour prediction
  - Forward-fill strategy on missing values

TFT Model Training Difference

The final model (purple) performed significantly better than all previous iterations.

TFT Model Performance - Overall





The results were much better!

Model Performance Comparison

Scores for Spatial Prediction (train/test models split based on IGBPs)

IGBP	Number of Sites	Number of training samples	XGB - R^2	LSTM - R^2	TFT - R^2
MF (mixed Forest)	4	116,353	0.75	0.61	0.57
GRA (Grassland)	9	576,534	0.43	0.52	0.87
ENF (Evergreen Needle Leaf)	17	1,199,243	0.66	0.43	0.77
CRO (Croplands)	5	279,361	0.7	0.63	0.89
CSH, OSH (shrublands)	3	379,442	0.26	0.61	0.73
SAV, WSA (Savannah)	7	84,529	0.43	0.58	0.78
WET (Wetlands)	4	253,728	0.72	0.67	0.79
All Biomes			0.7		0.82

Discussion

- **The TFT architecture performed the best and we suggest it as the new state-of-the-art for NEE prediction**

- The multi-head attention layer is interprable and further work can explore the model's feature selection.
- The FluxNet data is geographically sparse in many continents and where present, is restricted to a very small radius of observation.
  - The TFT model can be extended through more work to predict NEE where flux tower data is not available.

Q & A

Thank You

Berkeley
UNIVERSITY OF CALIFORNIA