Upscaling Global Hourly CO_2 Capture with Temporal Fusion Transformer

Team Carbon Chaser

Our Team

Rumi Nakagawa

Tokyo, Japan

John Calzaretta

Denver, CO

Mary Chau

Seattle, WA

Our Advisors

Berkeley iSchool (MIDS)

Alberto Todeschini

Faculty

Puya H. Vahabi

Faculty

The Quantitative Ecosystem Dynamics Lab

Yanghui Kang

Postdoctoral Researcher

Maoya Bassiouni

Postdoctoral Researcher

Problem Background

Quantifying Carbon Uptake / Gross Primary Productivity (GPP) Matters

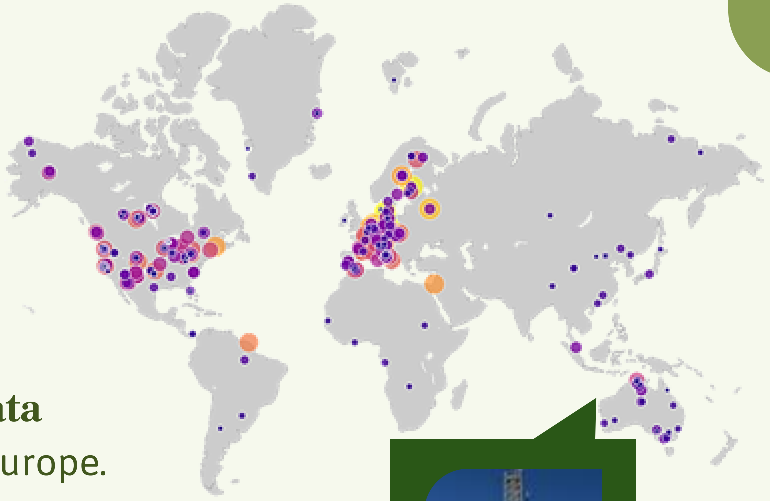
Rapidly growing concern on CO2 reduction in governments and industries

Regional Discrepancy Around Real GPP Data

Most flux towers are located in North America and Europe.

High burden to add new sites to measure actual GPP

High equipment and operational cost to sustainably measure actual GPP



FLUX TOWER

Limitation of Existing Solutions
to Predict Hourly GPP

Limited Data in High Temporal Resolutions

- Hourly data are limited.
- Features are often in larger granularity and high variance.

Lack of Long-Term Memory

Past memory is lost quickly.
Chronological order not captured

Computationally Expensive Resources

Complex models requires
significant compute power

Temporal Fusion Transformer (TFT)

How TFT Can Overcome the Challenges

Short- & Long-Term Memory

Combination of LSTMs and self-attention allows for models to recognize short- and long- term patterns

Capability of Upscaling

Global inference is possible as model can work on new sites that unseen in training data.

Variety of Input Features

TFT accepts time-varying known and unknown features, as well as static features for model input

Novel Components: Efficiency & Interpretability

Compute efficient and model interpretability via feature importance and interpretable self attention layers

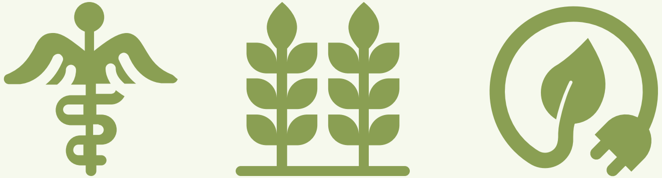
Our Objectives

- Improve model performance of hourly GPP models through Temporal Fusion Transformer (TFT) across global regions with no flux towers.
- Expand to scientific understanding of GPP trends and their contributing factors through TFT's interpretability.

Our Accomplishments

- Outperformed past studies by 2% NSE (Efficiency) and **10% RMSE**
- Successfully implemented TFT with **unconventional** application to solve time-series forecasting problem **without past target values.**
- **Novel approach of analytics** based on TFT's Interpretability outputs to provide insight on **temporal dynamics of feature importance** before predictions

These findings are applicable to other fields to studies!



Project Workflow

EDA
- Understand the dataset
- Observe distribution and correlation
- Initial feature selection

Preprocessing
- Build pipeline
- Gap-filling/Imputation
- Train/Test split strategy
- Feature engineering

Modeling
- Baseline with Random Forest
- Develop mvp with past TFT model
- Assess cross validation
- 5-Step Experimental Design

Advanced study
- Hybrid models
- Hyperparameter tunings
- Custom TFT models

Analysis
- Study model interpretability by land cover type/ Feature

Agile Operation

Major Features and Sources

Category	Features	Resolution	Source(s)	Availability at Upscaling
Target Variable	**Gross Primary Productivity (GPP)**	Half-Hourly, Sensor-Level	FLUXNET	NO
Radiation	Surface reflectance Band, SWIR, PAR, Longwave/shortwave radiation, SIF etc.	Daily, 7km	MODIS/ BESS_Rad	YES
Temperature	Land surface temperature(Daytime, Nighttime), Air/ Skin temperature	Hourly, 9km	ERA5-Land	YES
Water	Precipitation, Soil Moisture, Snow, Water index(NDWI) etc	Hourly, 9 km/ Daily/7km	ERA5-Land/ MODIS	YES
Pressure	Atmospheric pressure, Vapor pressure	Hourly, 9km	ERA5-Land	YES
Time	Year, Month, Day, Hour	-	-	YES
Land/Climate Type	IGBP, MODIS IGBP, koppen, koppen sub	Static, 1km	MODIS	YES

Data Pipeline

Retrieving Dataset → Aggregation → EDA Feature Selection → Gap-filling → Imputation → Stratified Split

Scope of the Capstone

- FlUXNET
- MODIS
- BESS_Rad
- ERA5-Land
- CSIF
- Köppen-Geiger

→ Hourly / Monthly

Problem: Flux sites have frequent outages due to weather, maintenance, and budget issues, leaving gaps in data sequences

Site A:

Solution: Use KNN-Imputer to fill in missing timesteps and other missing input features

Site A:

= **Gap-Filled w/ KNN Imputer**

- Train Set
- Validation Set
- Test Set

Categorize Earth by Land Cover Types


Deciduous Needleleaf Forest


Evergreen Broadleaf Forest


Evergreen Needleleaf Forest


Deciduous Broadleaf Forest


Mixed Forests


Croplands


Savannas


Grasslands


Snow and Ice


Shrublands


Barren or Sparsely Vegetated


Permanent Wetlands

https://eoimages.gsfc.nasa.gov/images/news/NasaNews/ReleaseImages/LCC/Images/lcc_key.jpg

Stratified Split by Land Cover Type

Data Distribution

129* Available Flux Tower Sites

- **ENF** Evergreen Needleleaf Forests
- **MF** Mixed Forests
- **DBF** Deciduous Broadleaf Forests
- **CRO** Croplands
- **WET** Permanent Wetlands
- **GRA** Grasslands
- **EBF** Evergreen Broadleaf Forests
- **OSH** Open Shrublands
- **WSA** Woody Savannas
- **SAV** Savannas
- **CSH** Closed Shrublands

Stratified Split

Train Set

ENF, CRO, EBF, SAV, MF, WET, OSH, CSH, DBF, GRA, WSA

Validation Set

ENF, MF, DBF, CRO, WET, GRA, EBF, OSH, WSA, SAV, CSH

Test Set

ENF, MF, DBF, CRO, WET, GRA, EBF, OSH, WSA, SAV, CSH

Study Overview & Performance Metrics

Supervised Learning

- Train with flux tower sites
- Validate/test with **"new, unseen"** sites
- Mimic upscaling scenario: areas with no flux tower

3 Key Metrics of the Domain

Nash Sutcliffe Equilibrium (NSE)

$$NSE = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

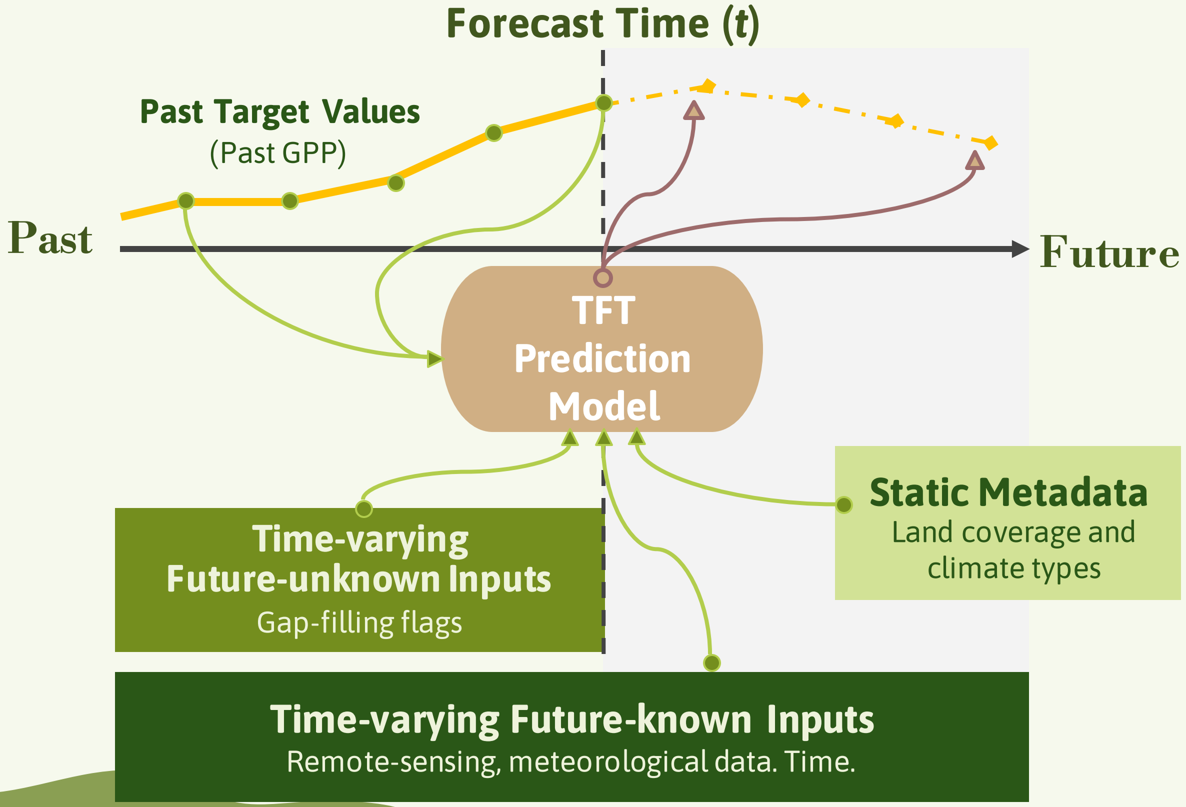
Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Experimental Design

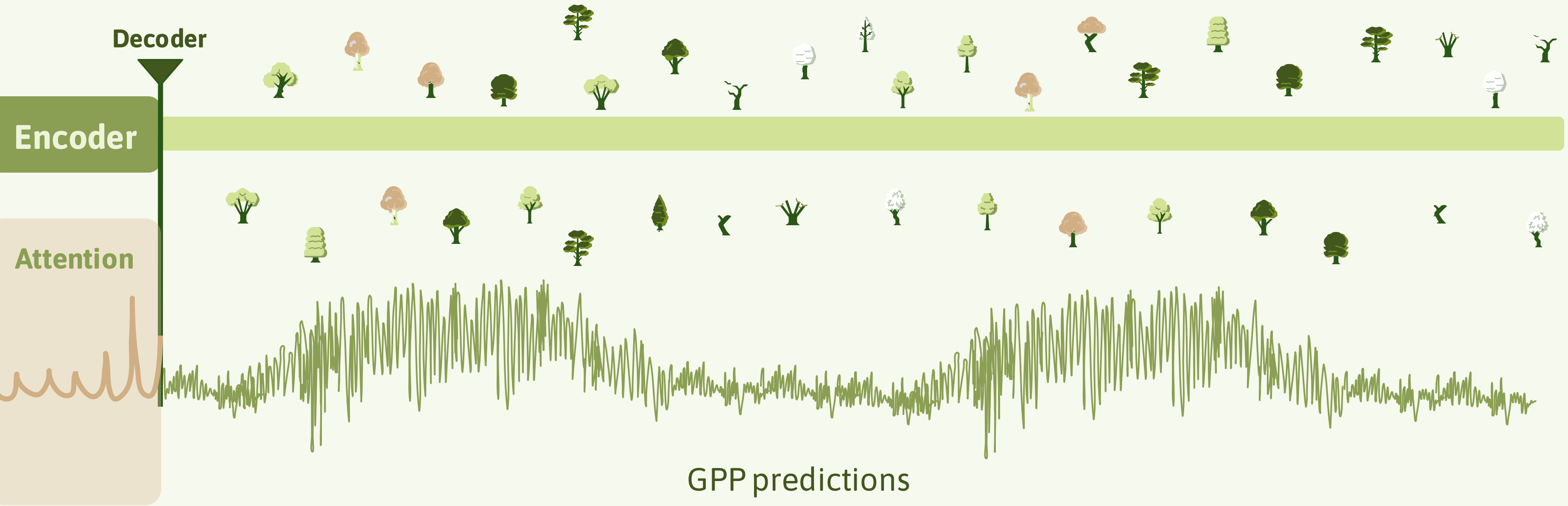
Algorithms	Model
RFR	1. Random Forest Regressor (Baseline)
Tree 🌳	2. Tuned RFR / XGBoost
TFT	3. GPP-TFT
	4. No-GPP-TFT
Hybrid	5. Tree-FT 🍄

TFT Model Overview



How Does TFT Work?

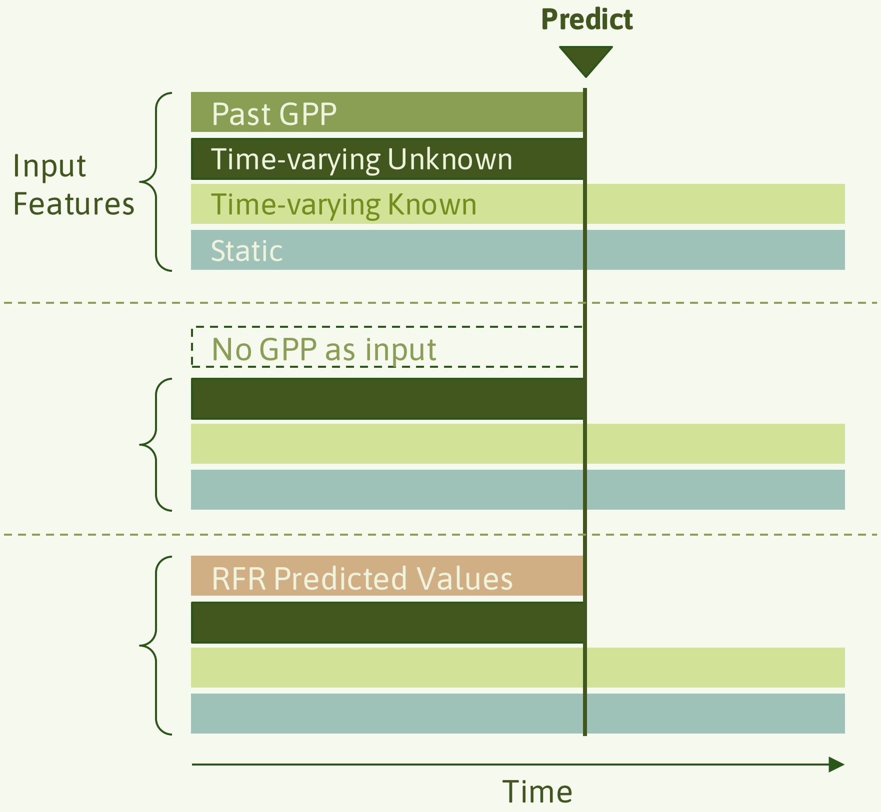
Train data(Ex. 2 years)



*Shape of the attention changes by each snapshot of time

TFT Experiments Overview

Applicable Sites	Model
Flux Tower Sites	3. GPP-TFT
Upscaling (Global Sites)	4. NO-GPP-TFT
	5. Tree-FT



Model Results

Algorithms	Model	RMSE	MAE	NSE/R^2
Past Literature	Bodesheim(2018) - RFR	3.94	-	0.67
Tree	1. Baseline RFR	3.66	2.01	0.68
	2-1. Tuned RFR	3.54	1.83	0.70
	2-2. Tuned XGBoost	3.53	1.85	0.69
TFT	3. GPP-TFT*	2.23*	1.05*	0.88*
	4. No-GPP-TFT	3.60	1.90	0.68
Hybrid	5-1. Tree-FT (RFR)	3.61	1.88	0.67
	5-2. Tree-FT (XGBOOST)	3.81	2.00	0.64

* Trained with 1 year of data.

NO-GPP-TFT: Performance By Land Cover Type

Land Cover Type (IGBP)	RMSE	MAE	NSE/R^2
Deciduous Broadleaf Forests (DBF)	3.315	1.742	**0.859**
Grasslands (GRA)	2.939	1.525	**0.733**
Mixed Forest (MF)	3.441	2.154	0.650
Evergreen Needleleaf Forests (ENF)	4.025	2.211	0.634
Savanas - Woody savannas (WSA)	2.933	1.632	0.595
Wetlands (WET)	3.900	2.126	0.571
Cropland (CRO)	5.246	2.941	0.536
Evergreen Broadleaf Forests (EBF)	1.252	0.698	0.193
Shrublands - Open Shrublands (OSH)	4.391	2.607	0.050

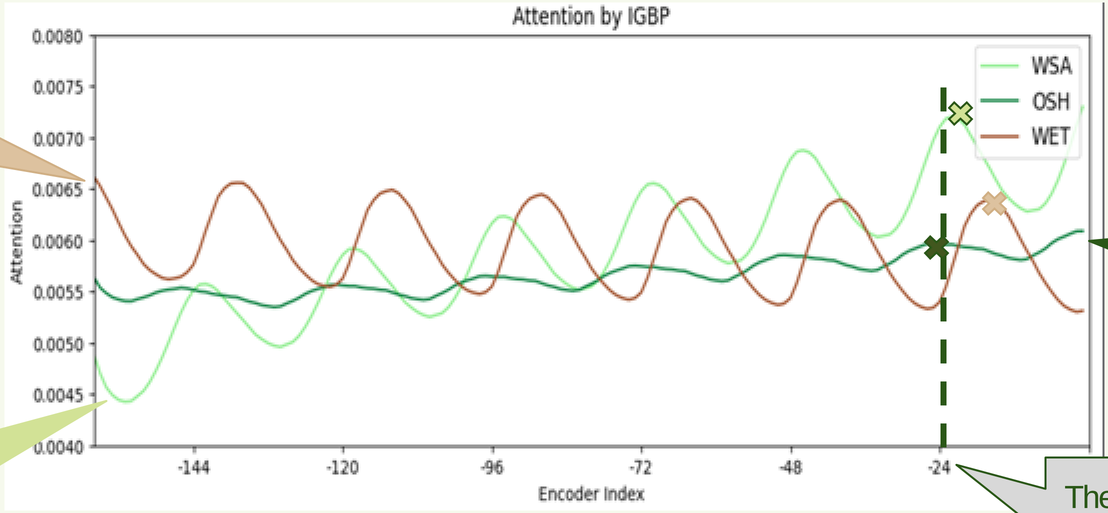
Demo: TFT Analysis

On Carbon Chaser Website

NO-GPP-TFT Analysis: IGBP Attention

Finding: The TFT model applies attention differently across IGBP groups. Future work should explore training models for each distinct IGBP group

WET has decreasing attention over time

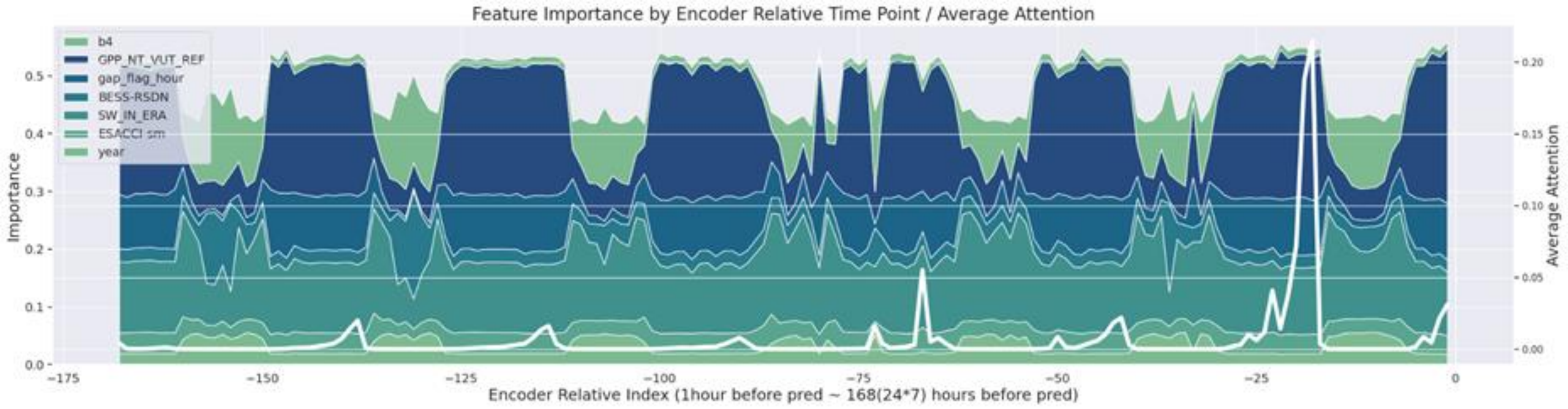


OSH has low variance of attention across each day

WSA has decreasing attention over time

The model pays attention to different timesteps for different groups

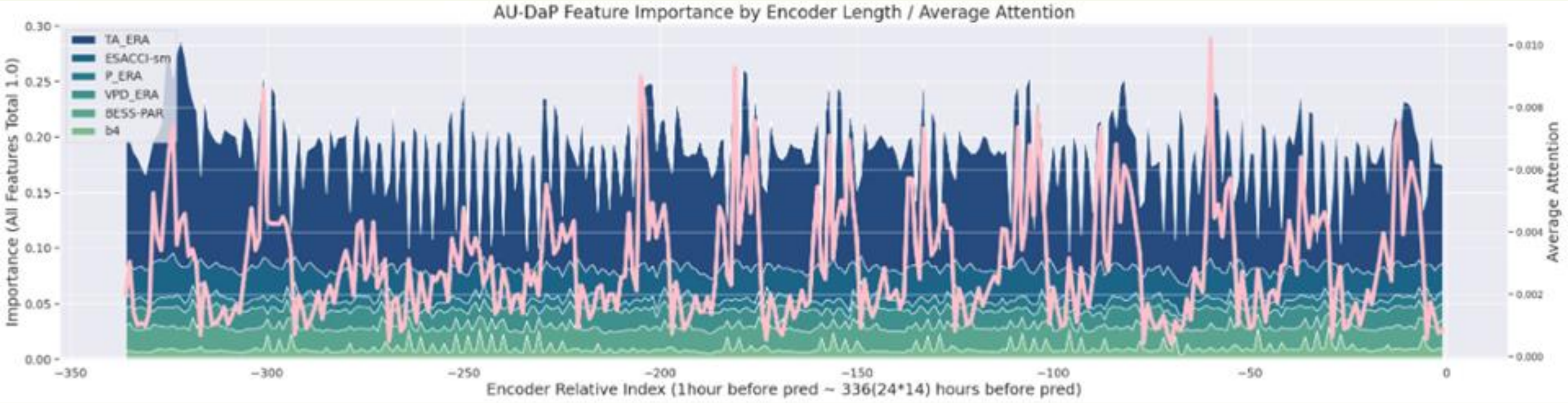
Feature Importance by Encoder Time Step (Including Past GPP)



Influence of b4 is smaller than GPP-TFT, while the influence of TA(Air Temperature) if closer to GPP-TFT

- 'b4','BESS-PAR','VPD_ERA', 'P_ERA', 'ESACCI-sm','TA_ERA'

Model	NO-GPP-TFT	Encoder Length	14 days	Year	2010	Month	1	Day	15	Time	0AM



Conclusion & Next Step

Conclusion

- Improved model performance
- Established groundworks for unconventional TFT application on generic upscaling problems
- Enabled new insights through novel approach of analysis

Future Work

- Multi-model solution by IGBP (land cover) type looks promising
- Apply improved models to upscaling data products
- Upcoming research sharing with:

NASA Carbon Monitoring System

The Quantitative Ecosystem Dynamics Lab
UC Berkeley & Lawrence Berkeley National Lab

Thank You

Question, Feedback, Suggestion?

Carbon Chaser Website

https://carbonchaser.wixsite.com/GlobalUpscale

GitHub & Paper

https://github.com/CarbonChaser/GlobalHourlyGppUpscale