

CarbonSense: A Multimodal Dataset and Baseline for Carbon Flux Modelling

Matthew Fortier

Mats L. Richter

Oliver Sonnentag

Chris Pal



POLYTECHNIQUE
MONTREAL

UNIVERSITÉ
D'INGÉNIERIE

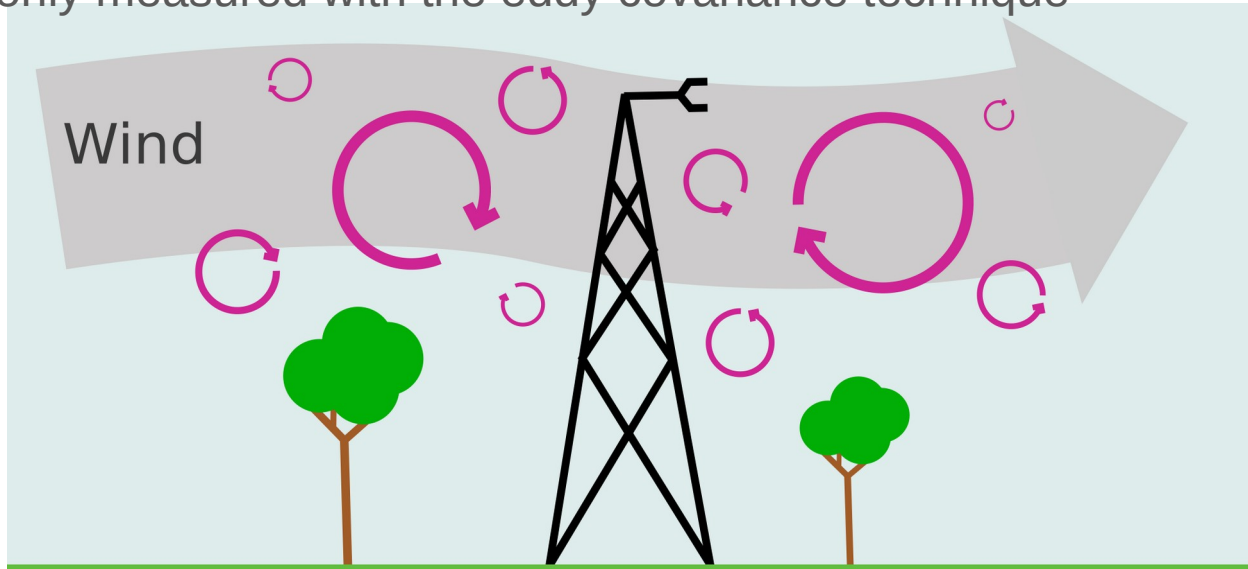


Université 
de Montréal

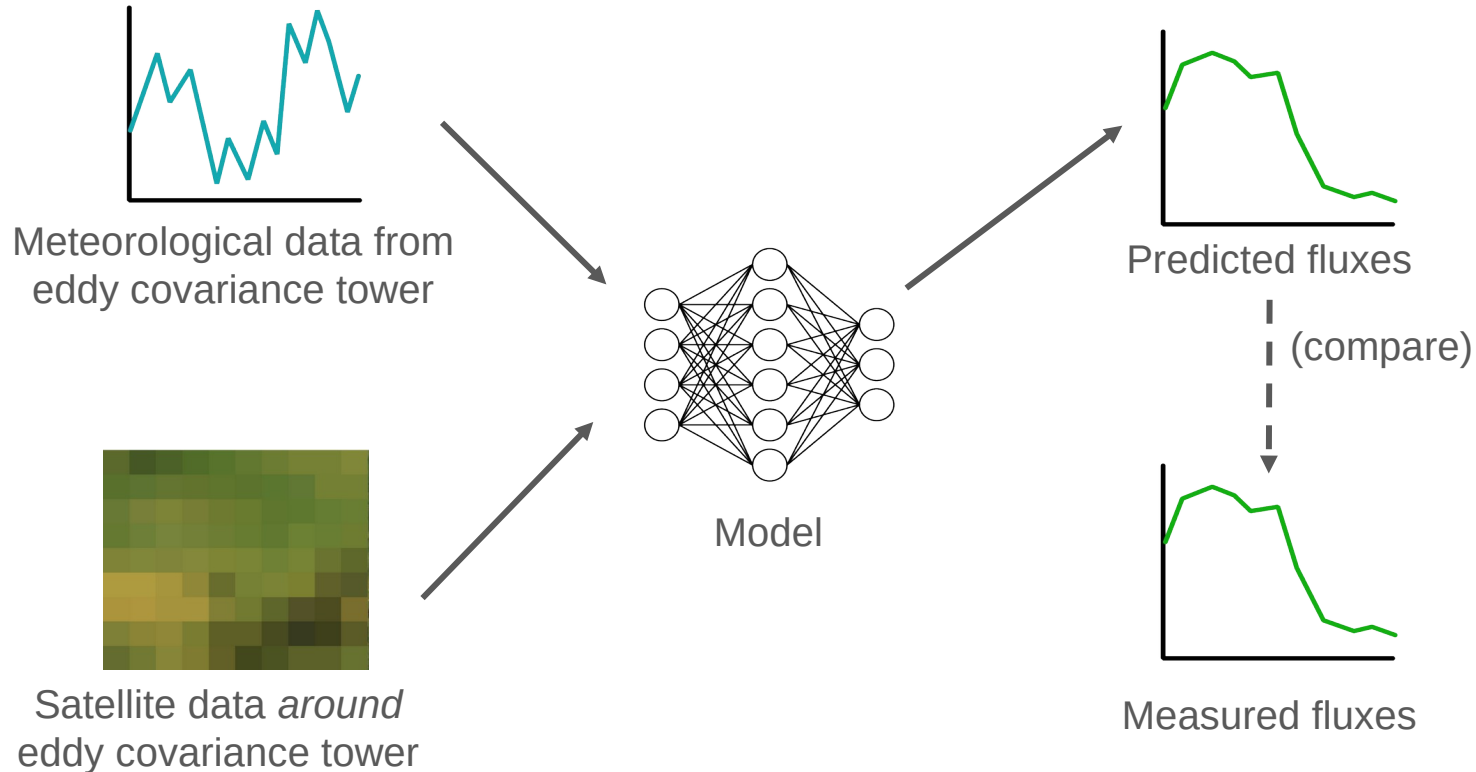


Motivation

- Carbon fluxes are vital to understanding our changing biosphere
- Commonly measured with the eddy covariance technique

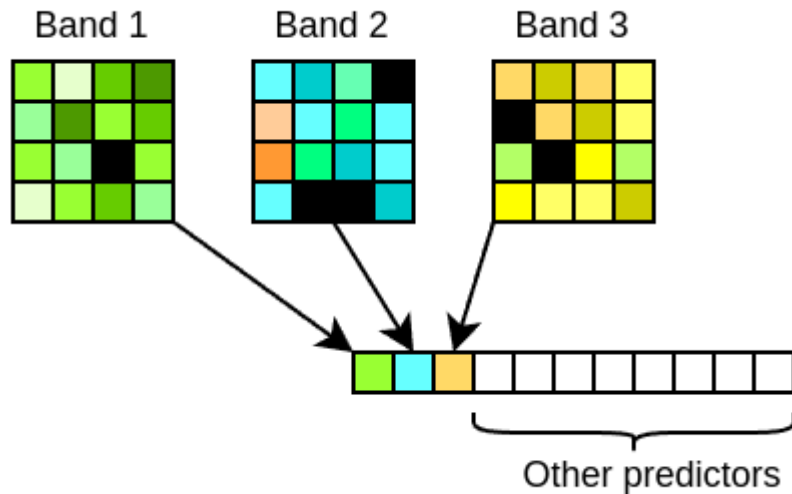


Background - Carbon Flux Modelling



Current State-of-the-Art

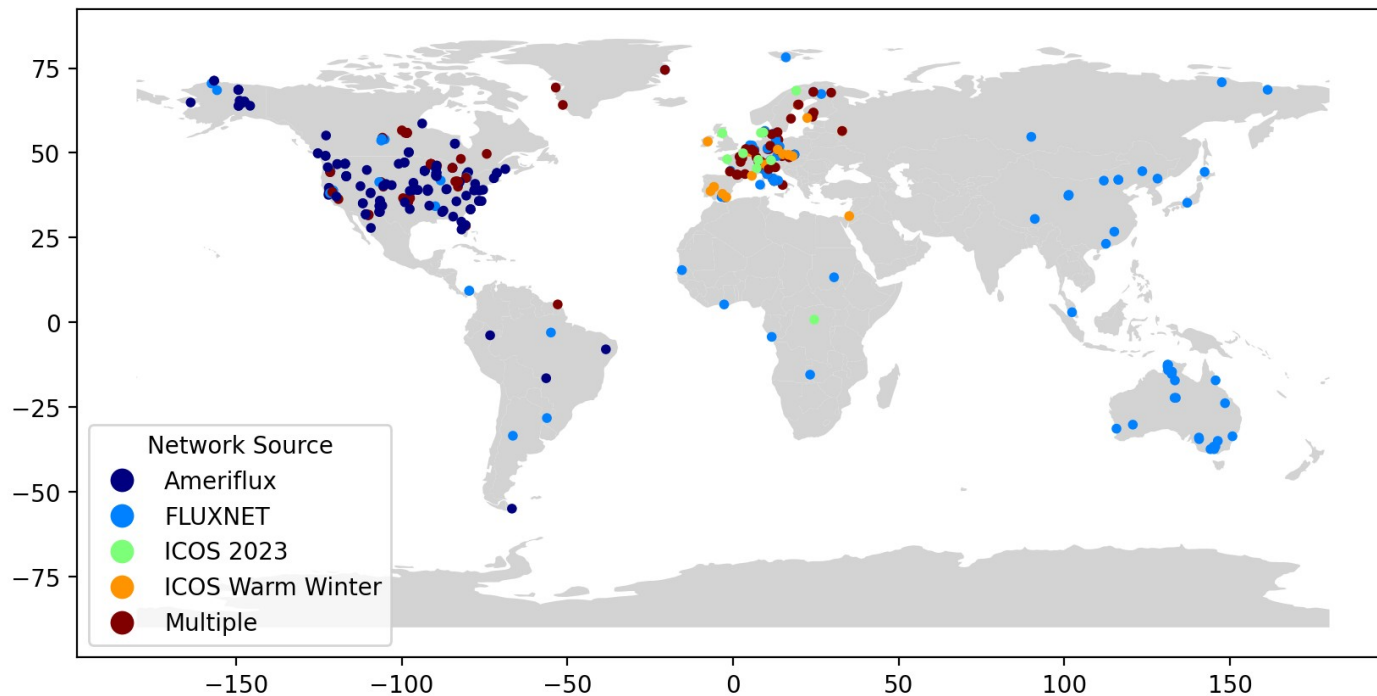
- Decision tree models like XGBoost
- Take strictly tabular input
- Spectral bands are compressed to single weighted-average values



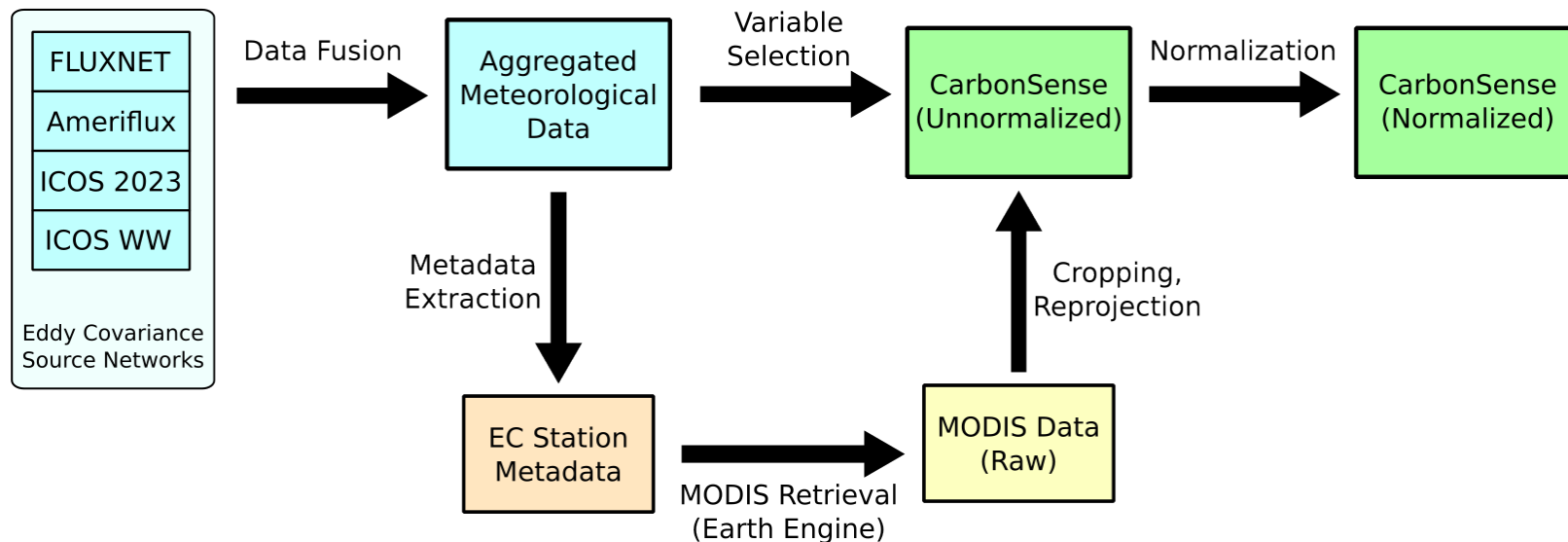
Research Objectives

- Create a large multimodal dataset for carbon flux modelling
- Create a deep learning model tailored to this domain
- Provide benchmarks to encourage future research

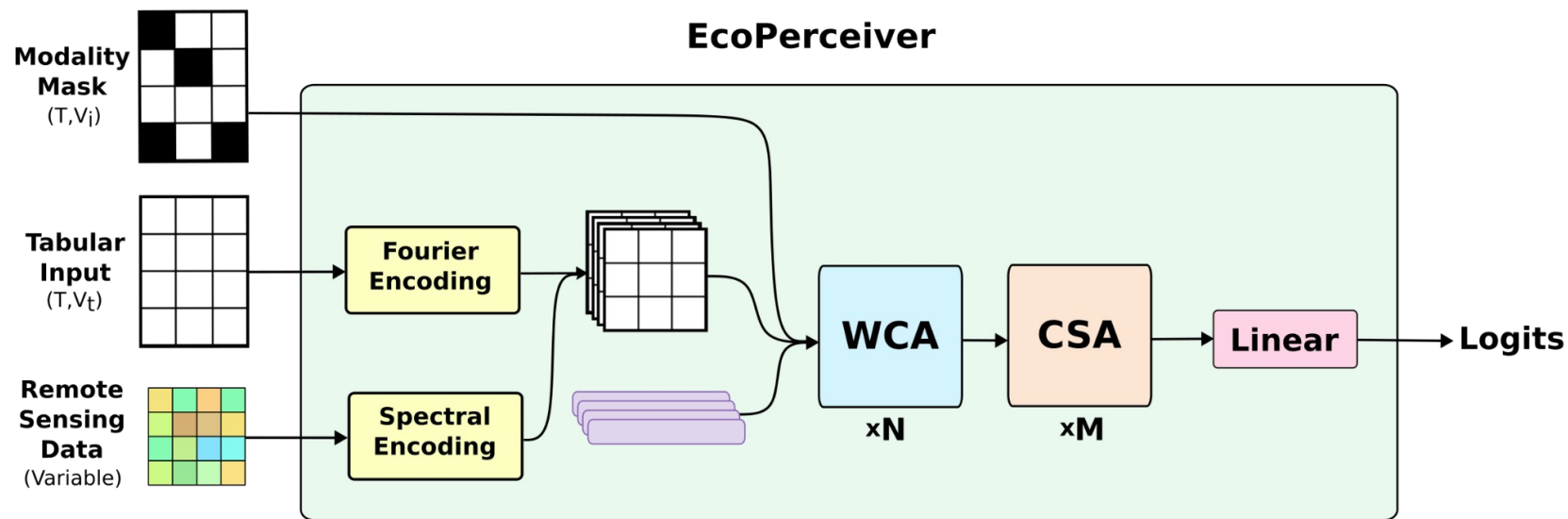
Gathering Eddy Covariance Data



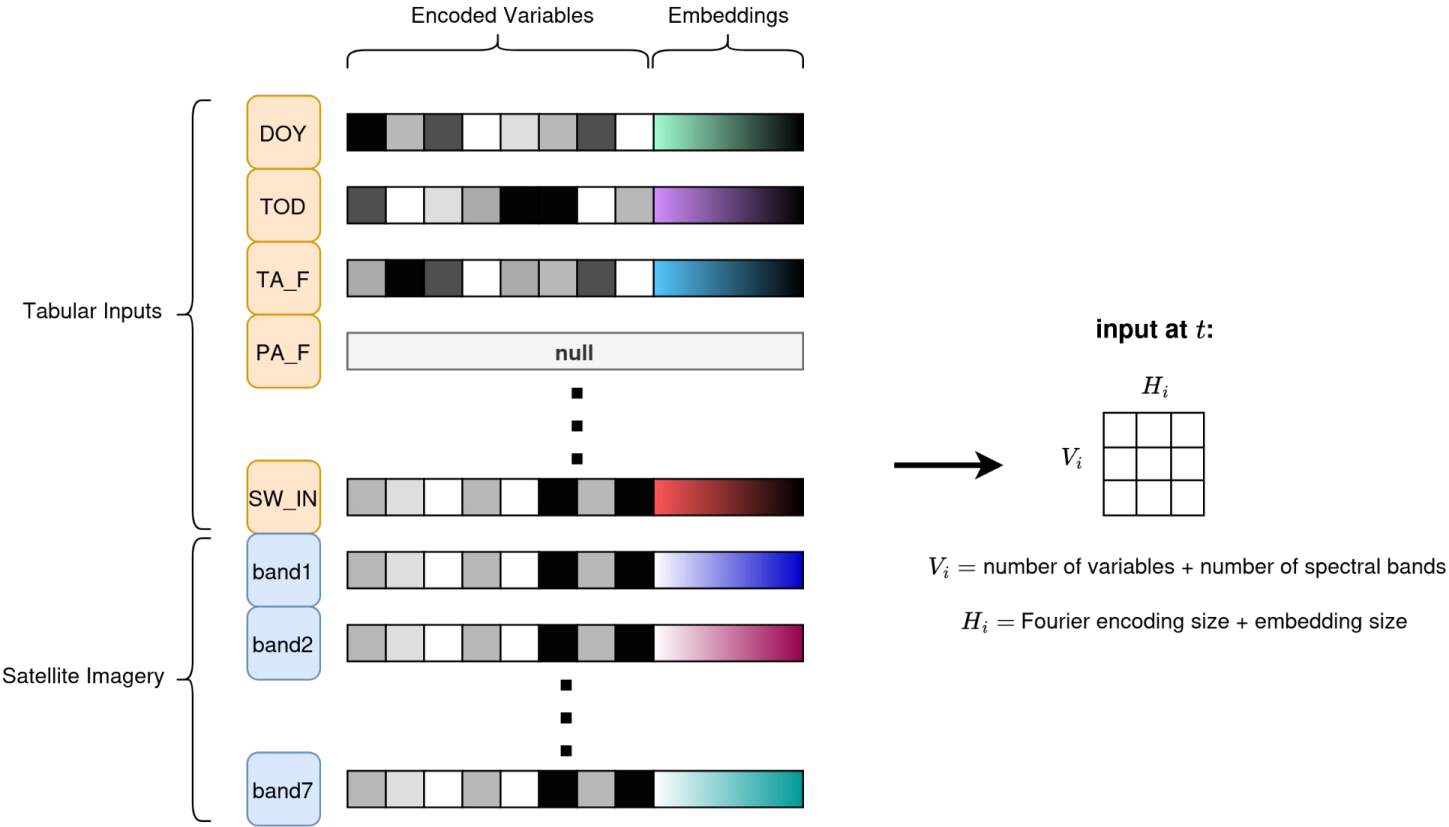
CarbonSense - Data Pipeline



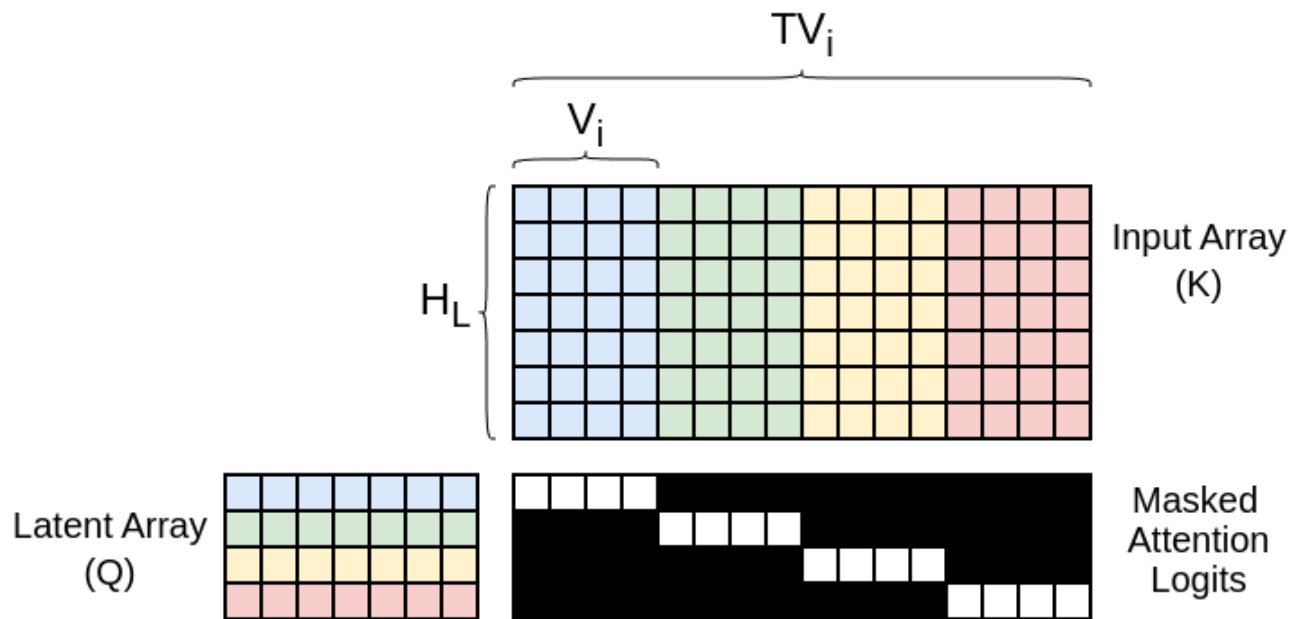
EcoPerceiver - Overview



EcoPerceiver - Input Encoding

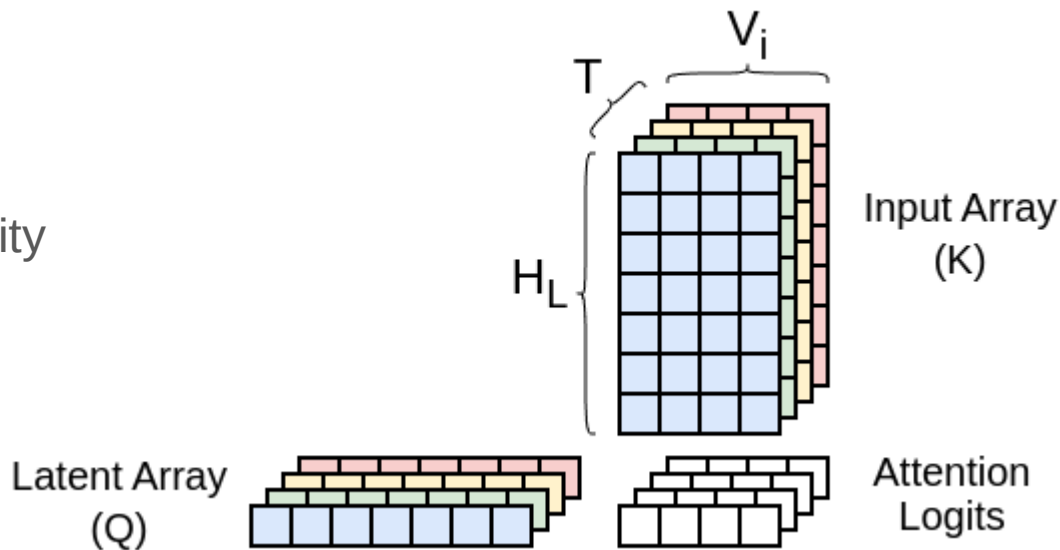


EcoPerceiver - Windowed Cross Attention



EcoPerceiver - Windowed Cross Attention

- Idea: move T into the batch dimension
- Reduces attention complexity by a factor of T



IGBP	XGBoost		EcoPerceiver	
	NSE	RMSE	NSE	RMSE
CRO	0.8066	3.2381	0.8482	2.8677
CSH	0.7510	1.5224	0.7670	1.4709
CVM	0.5277	5.5157	0.5763	5.2236
DBF	0.7250	4.0959	0.7547	3.8678
DNF	0.2803	4.0974	0.4336	3.6322
EBF	0.7966	4.6050	0.8220	4.3070
ENF	0.7765	2.8141	0.7694	2.8579
GRA	0.7461	3.2487	0.7967	2.9059
MF	0.7559	3.8633	0.7717	3.7361
OSH	0.5451	1.8796	0.6060	1.7475
SAV	0.5802	1.6514	0.7368	1.3070
SNO	-0.0370	1.4291	0.2898	1.1816
WAT	-11.0524	3.1838	-14.4010	3.5802
WET	0.4530	2.2073	0.4137	2.2830
WSA	0.6132	2.5153	0.6267	2.4706

Conclusion - What's Next?

- Add more eddy covariance sites from underrepresented areas
- Include new target fluxes like methane (CH_4), nitrous oxide (N_2O)
- Add framework for global inference