Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

Global Vegetation Modeling with Pre-trained Weather Transformers

Pascal Janetzky Florian Gallusser Simon Hentschel
Andreas Hotho Anna Krause

data science

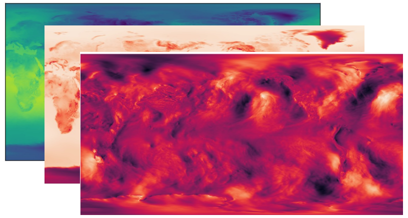
CAIDAS

Co-funded by
the European Union

bdg 2.0

(How) Can We Use Weather Models for Ecosystems?

Weather variables x_t



Deep Learning Weather Model

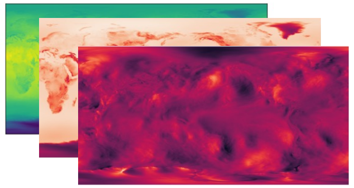
Deep Learning Weather Forecasting

- Deep learning weather models accurately predict weather at high spatial and temporal resolution
- Atmospheric representation learned during pre-training usable for novel tasks

Vegetation modeling

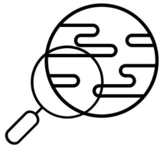
- Vegetation models provide insights into the interplay of vegetation state and environmental conditions
- Short-term and small-scale weather phenomena affect vegetation state

Pre-training



Weather variables x_{t+1}

Finetuning weather applications



Downscaling

...

Regional Modeling

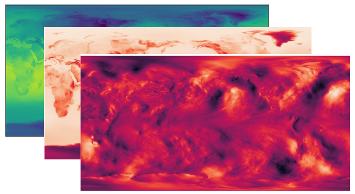
Finetuning for vegetation modeling

Contribution

Data-driven vegetation model at …

- global scale
- high spatial resolution
- high temporal resolution

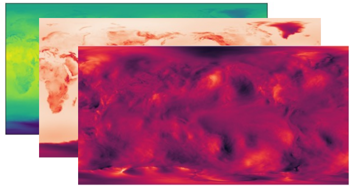
Weather variables x_t



Deep Learning Weather Model

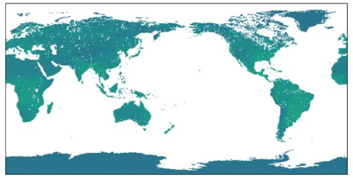
… exploiting the learnt atmospheric representations of a pretrained weather model

Pre-training



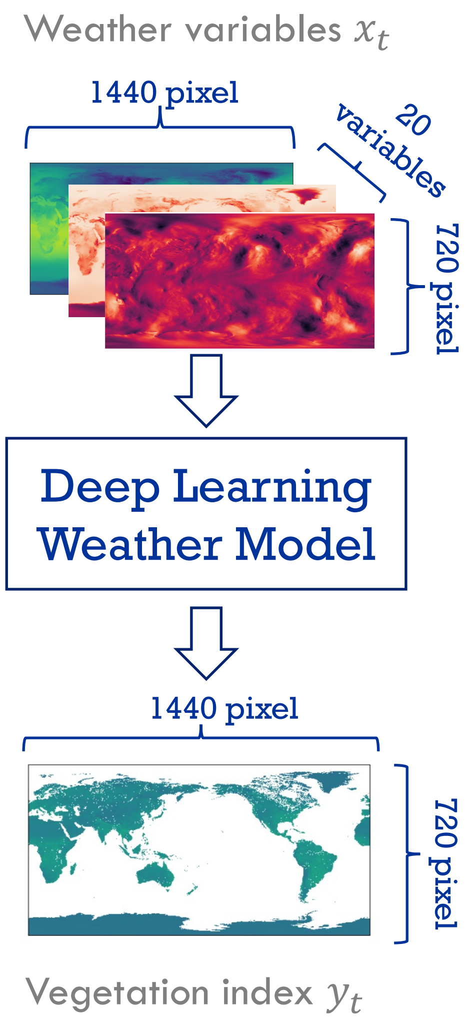
Weather variables x_{t+1}

Our work: Finetuning



Vegetation index y_t

Dataset



Input Dataset

- ERA5 [1]: Global reanalysis weather data
- Multiple atmospherics variables at different vertical levels
  - Temperature
  - Geopotential
  - Wind speeds
  - Relative humidity

Target Variable

- Normalized difference vegetation index (NDVI) as proxy for vegetation state from GIMMS [2]
- Remotely sensed data
- Index range: -1 (water) to 1 (dense vegetation); around 0 means only sparse vegetation
- Globally available for terrestrial ecosystems
- Sparingly available in high-latitude regions

Global, high-resolution dataset to model short-term and small-scale meteorological effects on vegetation

- Spatial resolution: 0.25°
- Temporal resolution: daily
- Train period: 1982 – 2010
- Val period: 2011 - 2012
- Test period: 2013

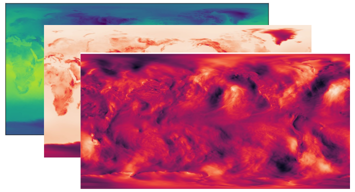
→ Efficient model required to process large-scale dataset

[1] Hersbach et al. "The ERA5 global reanalysis." (2020)
[2] Vermote et al. „Noaa climate data record (cdr) of avhrr normalized difference vegetation index (ndvi), version 5“ (2019)

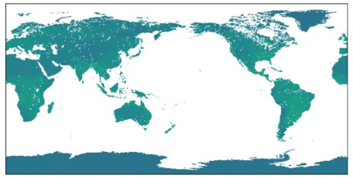
Model

Efficient global vegetation model by replacing the head and finetuning the entire pre-trained FourCastNet [3]

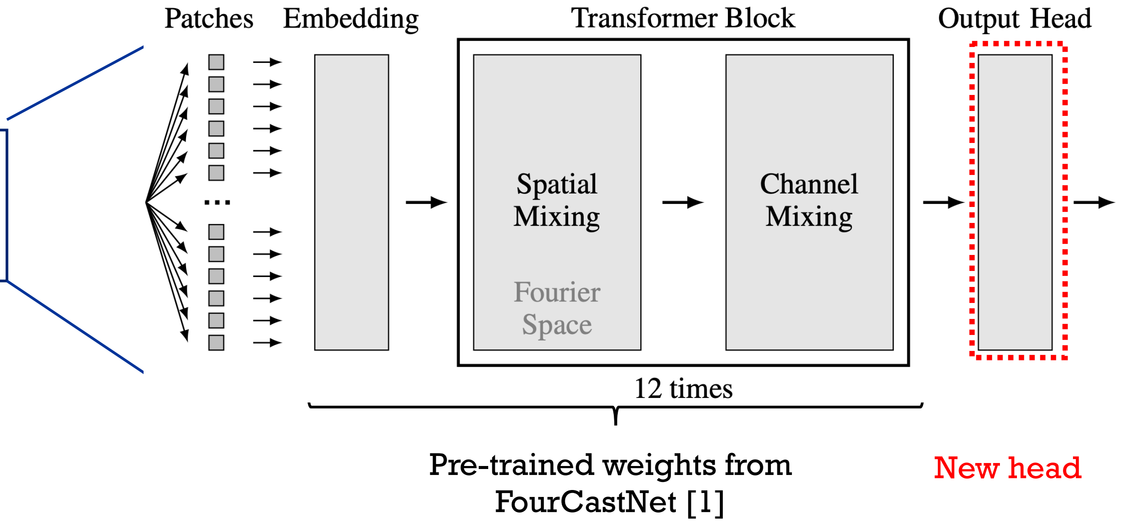
Weather variables x_t



Adapted FourCastNet



Vegetation index y_t



FourCastNet (FCN) info:

- Weather model based on VisionTransformer
- Adaptive Fourier Neural Operator used for efficient global spatial mixing
- Pretrained to predict $x_{t+\Delta 6h}$ from x_t on ERA5
- 73 million parameters

[3]: Pathak et al. "Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators." (2022).

Evaluation

Comparison against results from the literature by adopting their evaluation scheme

LSTM [4]

- Models local NDVI y_t^i from pixel time-series $x_{t-n}^i, x_{t-n+1}^i, \dots, x_t^i$
- Global weight sharing
- 15-daily resolution at 0.5°
- Also uses soil variables

State Space Model [5]

- Guided by plant-growth equations
- Separate, local models for 100 locations
- 7-daily resolution

Our Models

Global input x_t and output y_t with daily resolution at 0.25°

- FourCastNet finetuned
- FourCastNet from scratch (all weights randomly intialized)
- CNN (hyperparameter optimized)

Global evaluation

- Aggregate model output and target to 15-daily resolution
- Exclude pixels with >50% missing NDVI values
- Compute pixel area-weighted average R^2 and RMSE

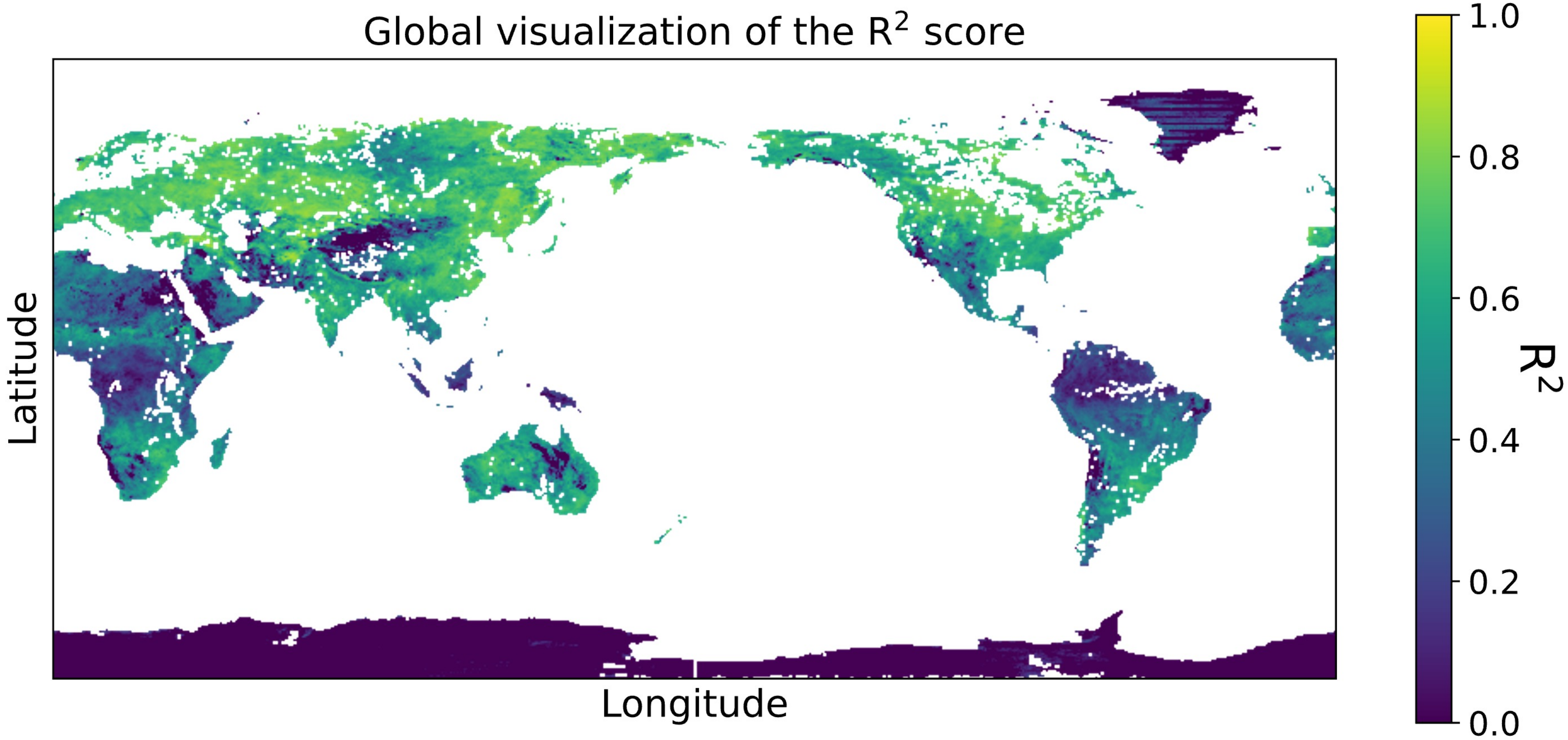
Local evaluation

- Aggregate model output to weekly resolution at the same 100 locations
- Compute unweighted average R^2 and RMSE per location

[4]: Kraft et al. "Identifying dynamic memory effects on vegetation state using recurrent neural networks." (2019)
[5]: Higgins et al. "Shifts in vegetation activity of terrestrial ecosystems attributable to climate trends." (2023)

Global Results: R^2 score

Finetuning FourCastNet captures a substantial amount of the NDVI's variability



Global Results

Our finetuned model outperforms training from scratch

Model	FCN finetune	FCN scratch
RMSE	0.0403	0.0512
R^2	0.6331	0.4977

Recap: evaluation setting

- Aggregate model output and target to 15-daily resolution

→ Structural atmospheric knowledge gained during pre-training is beneficial for NDVI modeling

→ Structural knowledge might not be attained in training from scratch

Global Results

Our finetuned model outperforms a hyperparameter-optimized CNN

Model	FCN finetune	CNN	LSTM
RMSE	0.0403	0.0431	0.017
R^2	0.6331	0.6061	0.904

Evaluation setting and recap:

- Aggregate model output and target to 15-daily resolution

- LSTM performance is a strong baseline and currently not reached
    - LSTM uses data from multiple past timesteps (FCN uses only current *t*)
    - LSTM uses soil variables

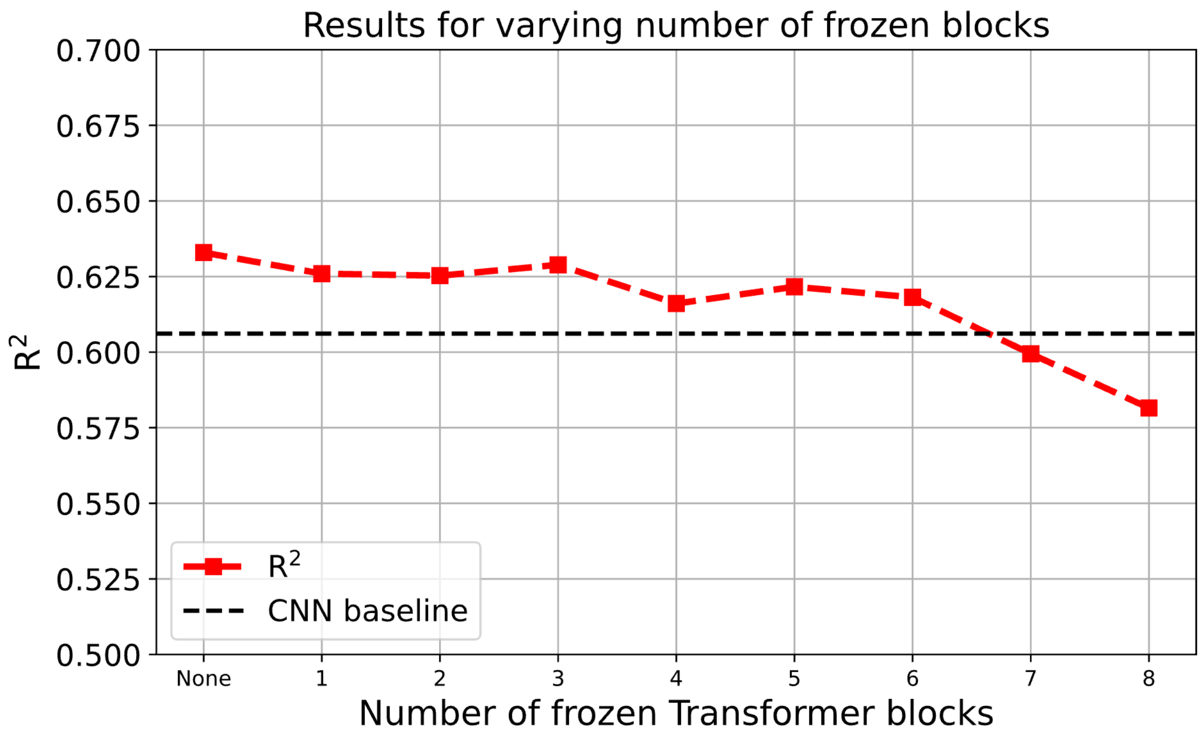
→ FCN performance might improve by incorporating past and soil-related variables

Ablation Study: Freezing Blocks

Finetuning only 50% of FourCastNet's parameters creates strong global ecosystem model

- Vary the number of trainable parameters by freezing Transformer blocks
- Until 50% frozen parameters, FCN outperforms CNN baseline

→ Features extracted from (frozen) pre-trained weather models beneficial for ecosystem modeling

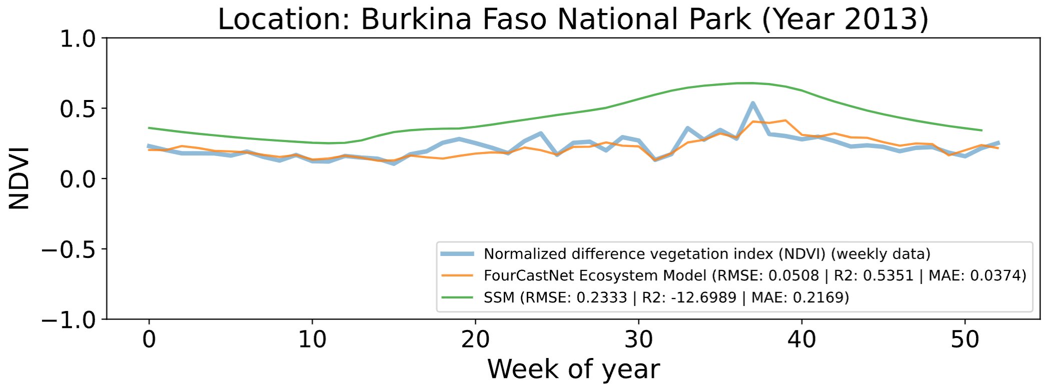


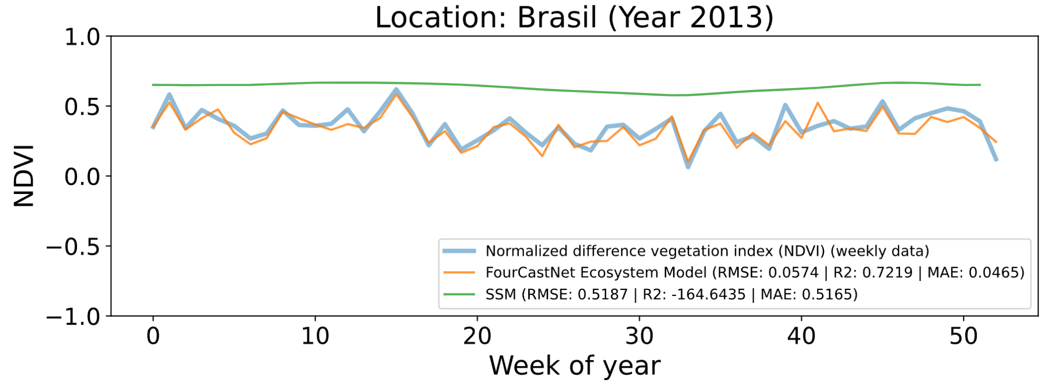
Local Results

Our global model outperforms location-specific state space models

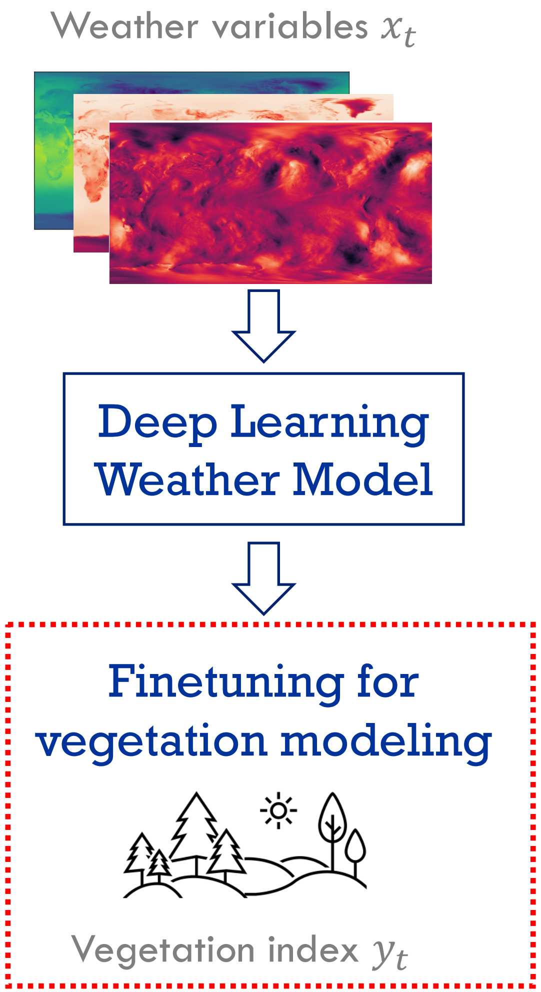
Model	FCN finetune	SSM
RMSE	0.0547	0.0548
R^2	0.5151	0.4038

→ Our global model captures biome-related patterns solely from global data





Summary



- Pre-trained weather model for globally modeling NDVI at high spatial and temporal resolution
- Global model also shows strong local performance
- Structural atmospheric weather knowledge of weather models beneficial for ecosystem modeling
- FCN performance might be improved by incorporating past and soil-related variables

Thank You For Your Attention

Code

→ github.com/LSX-UniWue/Global-Ecosystem-Modeling

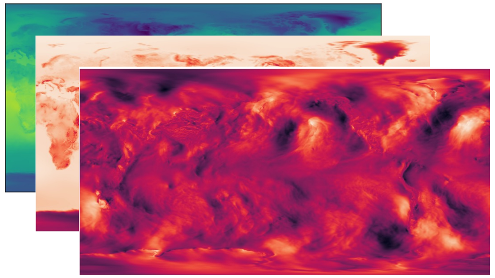
Deep Learning for Dynamical Systems Group

→ anna.krause@informatik.uni-wuerzburg.de (Head)

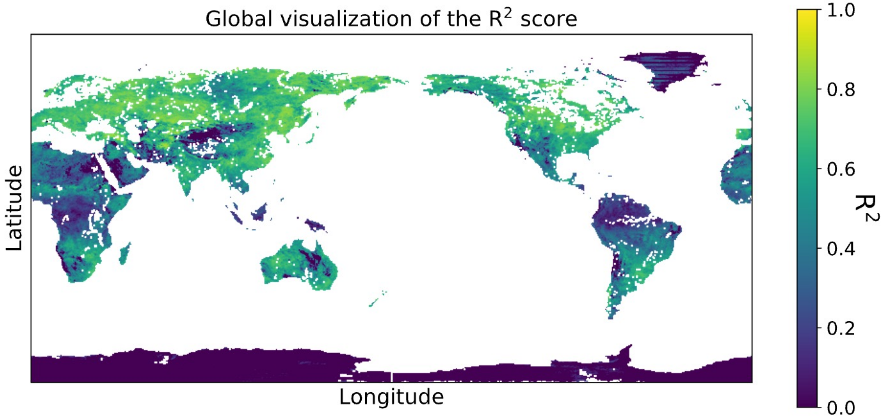
→ janetzky@informatik.uni-wuerzburg.de

→ gallusser@informatik.uni-wuerzburg.de

Co-funded by the European Union



Pre-trained Weather Model



References

[1] Hersbach, Hans, et al. "The ERA5 global reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146.730 (2020): 1999-2049.

[2] Vermote, Eric „Noaa climate data record (cdr) of avhrr normalized difference vegetation index (ndvi), version 5.“ (2019). Downloaded from https://www.ncei.noaa.gov/data/land- normalized-difference-vegetation-index/access/, last accessed February 2024.

[3] Pathak, Jaideep, et al. "Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators." *arXiv preprint arXiv:2202.11214* (2022).

[4] Kraft, Basil, et al. "Identifying dynamic memory effects on vegetation state using recurrent neural networks." *Frontiers in Big Data* 2 (2019): 31.

[5] Higgins, Steven I., Timo Conradi, and Edward Muhoko. "Shifts in vegetation activity of terrestrial ecosystems attributable to climate trends." *Nature Geoscience* 16.2 (2023): 147-153.

Supplementary Slides

Local Results

Modeling vegetation activity in tropical regions is challenging

Biome	RMSE FCN	RMSE SSM	R2 FCN	R2 SSM	Samples
Boreal forest	0.0656	0.0834	0.7244	0.8321	16
Grassland	0.0416	0.0449	0.4716	0.4484	14
Mediterranean-type	0.0364	0.0393	0.1921	–0.6008	5
Tropical forest	0.0789	0.0508	0.1405	–0.0515	16
Savanna	0.0512	0.0516	0.7151	0.6628	18
Shrubland	0.0478	0.0445	0.3021	0.1966	16
Temperate forest	0.0598	0.0693	0.7726	0.5247	12
Tundra	0.0423	0.0600	0.8312	0.9149	9

NDVI modeling performance of our global FCN model (evaluated locally) and the SSM approach [5] at 100 selected locations, covering 8 different terrestrial ecosystems. Both approaches perform notably worse in tropical and Mediterranean-type ecosystems.

Global Results

Number of valid samples decreases towards the poles