Learning from data to model the Earth system and predict climate impacts on the biosphere

Prof. Dr. Benjamin Stocker

INPUTBern Research Meeting, 23. Nov. 2022

FNSNF

SCHMIDT FUTURES

GECO

UNIVERSITÄT BERN

OESCHGER CENTRE
CLIMATE CHANGE RESEARCH

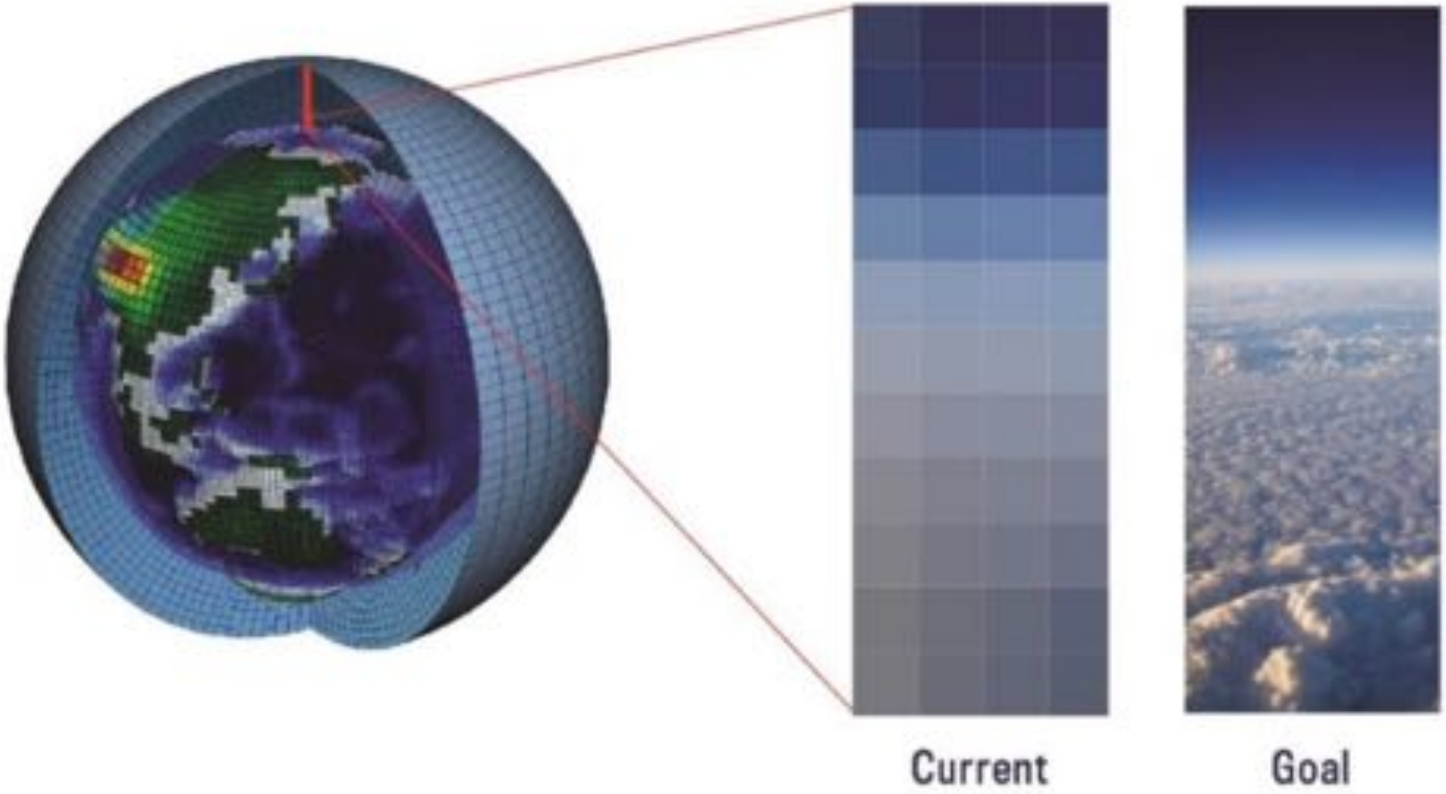
https://pixy.org/372871/

Earth System Model
- Description of physics and mechanisms
- Systems dynamics
- Multiple state variables
- Feedbacks

Data
- Data assimilation
- “Nudging” the state of the system towards observations

- Numerical weather prediction models

https://pixy.org/372871/


Current
Goal

Data revolution in Earth and Environmental Sciences

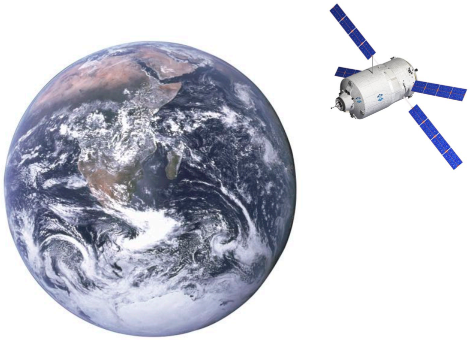


Image Markus Staudinger,

Kattge et al., 2019

Remote sensing

Example: NASA missions

- 12 TB / day (in 2017)

Continuous monitoring

Example: FLUXNET

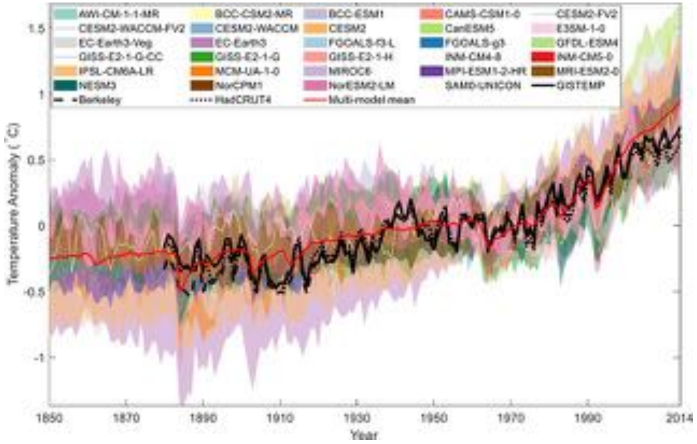
- 166 sites with open access data
- <25 years

Data compilations

Example: TRY plant traits database

- 12 mio. records
- 280’000 species
- 21’000 sites

Data revolution in Earth and Environmental Sciences



Climate model outputs

Example: CMIP6

- 30 PB

Data + Machine Learning = ✔

Promise

- Learn more than possible with traditional data assimilation methods

Challenge

- How to extract interpretable information and knowledge from data?
- How to keep track with increasing data availability?

Systems modelling paradigm *vs* data-driven

Theory *vs* data

Attempt at classification of ML in Earth system modelling

A: Making sense of information

B: (Black box) prediction

C: Hybrid modelling

This talk

1. Some success stories

    A: Making sense of information

    B: (Black box) prediction

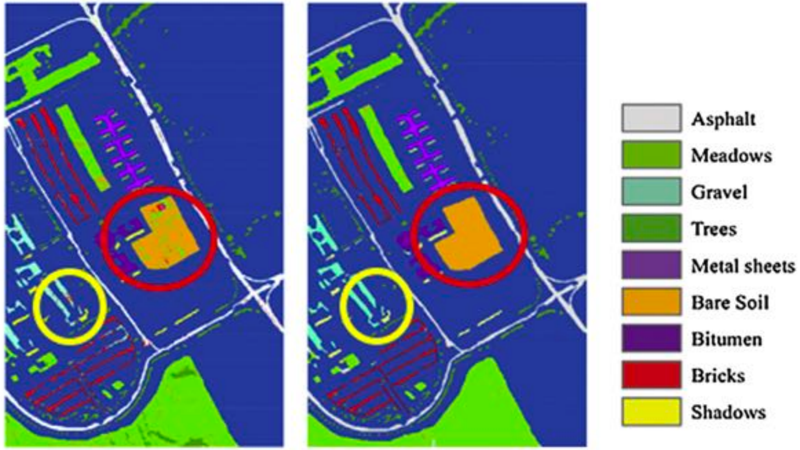
    C: Hybrid modelling

2. Our research in Geocomputation and Earth Observation

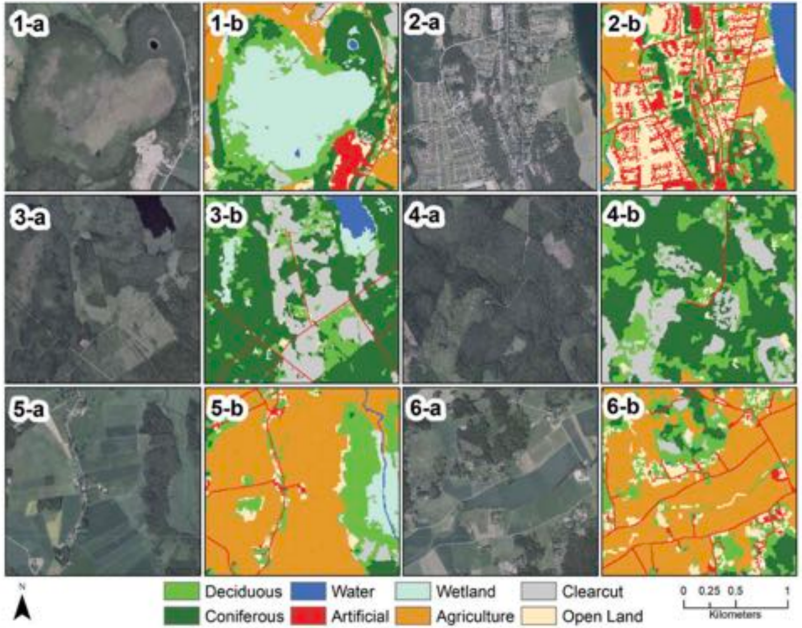
3. Outlook

A: Making sense of information

- **Land cover classification**
  - **Pixel-wise:** Y = *f*(**x**)
  - **Accounting for spatial dependencies**



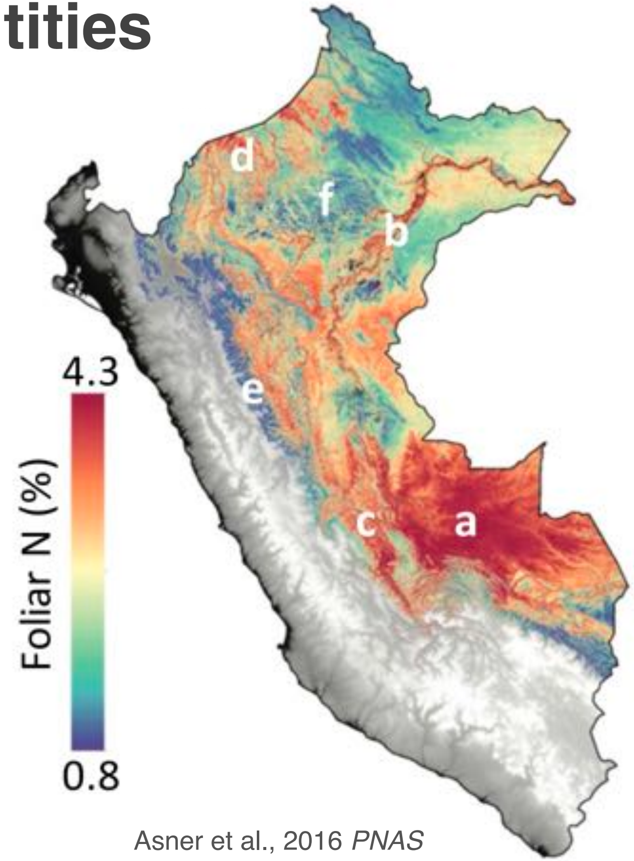
Zhao & Du, 2016



Abdulhakim et al., 2020 *GIScience*

A: Making sense of information

- **Regression of biogeophysical quantities**
  - **Pixel-wise:** Y = $f(\mathbf{x})$

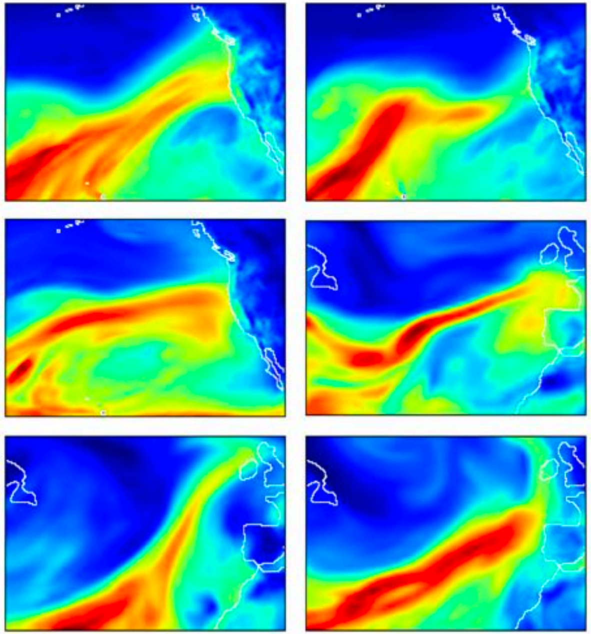


Asner et al., 2016 *PNAS*

A: Making sense of information

- **Detecting and classifying (extreme) weather patterns**
  - **Accounting for spatial dependencies**

Total column water vapour

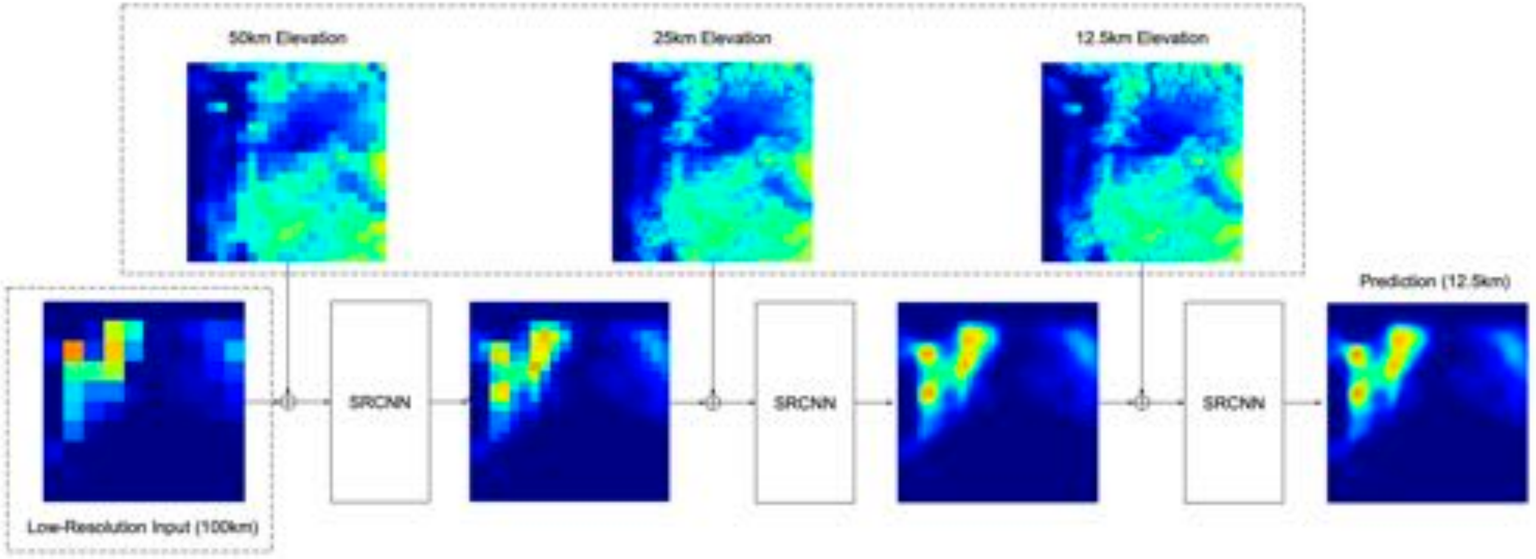


- **Anomaly detection**

Liu et al., 2016 *arXiv*

A: Making sense of information

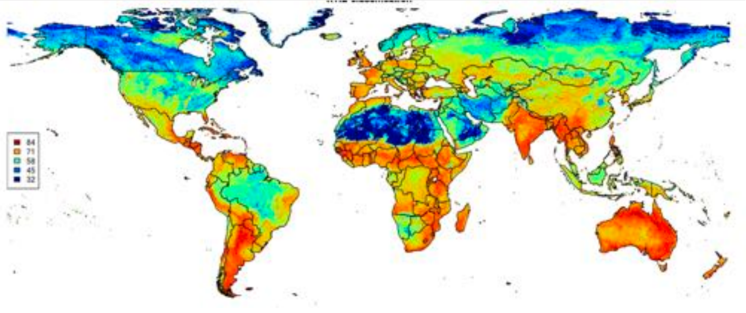
- **Downscaling, gap-filling, making data analysis-ready**



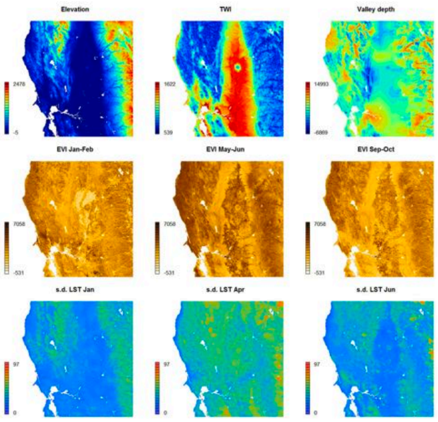
Vandal et al., 2018

A: Making sense of information

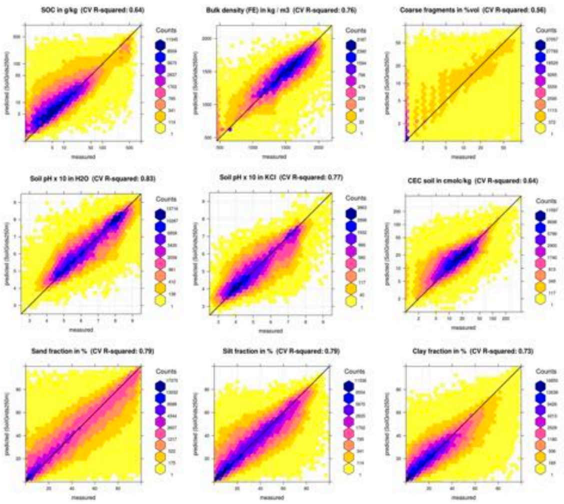
- **Spatial upscaling**
  - **Digital soil mapping**
  - **Pixel-wise:** Y = *f*(**x**)




Training data:
150,000 soil profiles



Environmental covariates
with global coverage

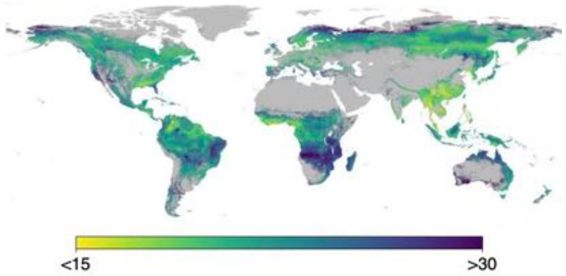

Out-of-sample prediction
and globally upscaled
product

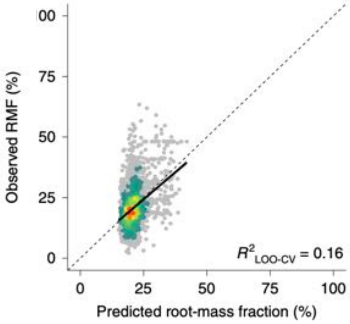
Hengl et al., 2017

A: Making sense of information

- **Spatial upscaling everything**

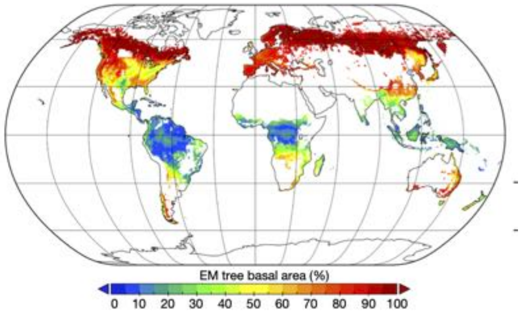
Root mass fraction (%)

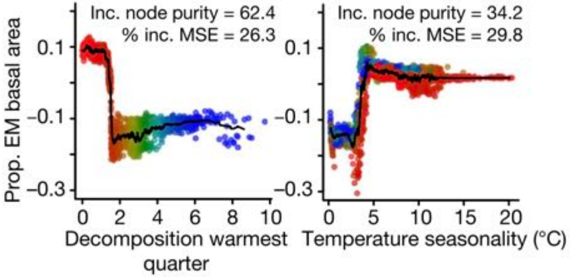




Ma et al., 2020 *Nature Eco. Evo.*

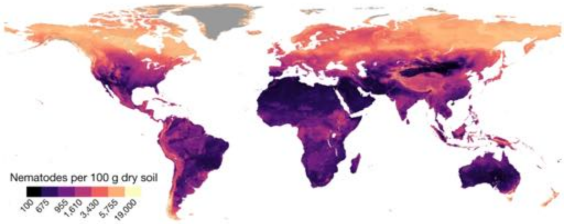
Fraction of trees associated with ectomycorrhizal fungi (%)



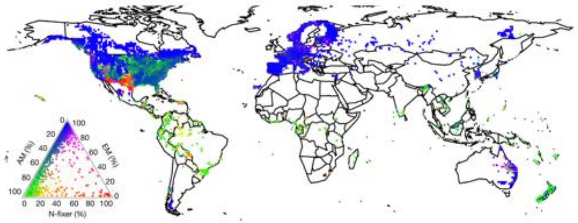


Steidinger et al., 2019 *Nature*

Soil nematode abundance



Van der Hoogen et al., 2019 *Nature*



A: Making sense of information

State estimation

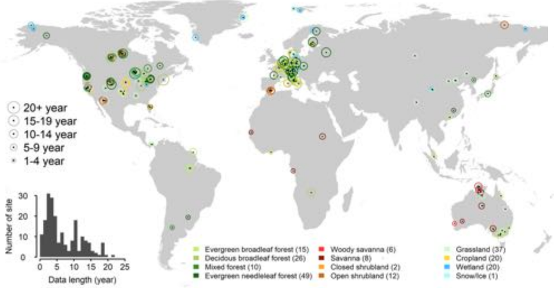
B: (Black box) prediction

Dynamics prediction

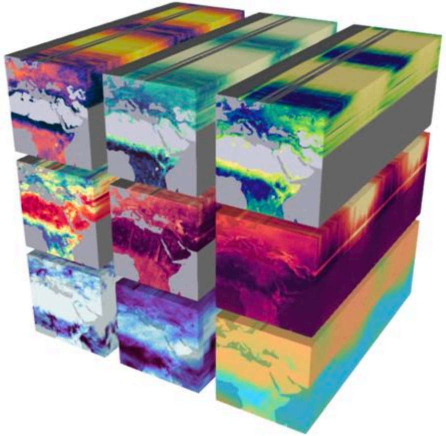
B: Prediction

- **Temporal-spatial upscaling**
  - **Pixel-wise** $Y_t = f(\mathbf{x}_t)$

Image Markus Staudinger,

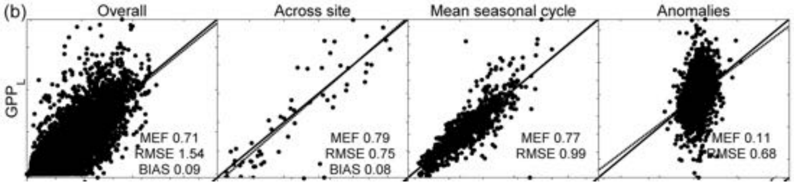


Pastorello et al., 2020 *Sci. Dat.*



Mahecha et al., 2020 *ESD*

Jung et al., 2020 *BG*



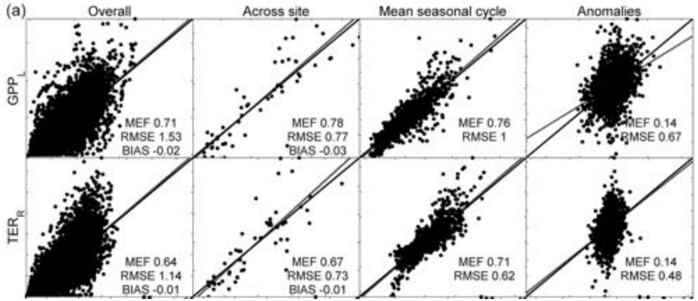
Tramontana et al., 2016

B: Prediction

- **Adding physics into the black box**

NEE = TER - GPP

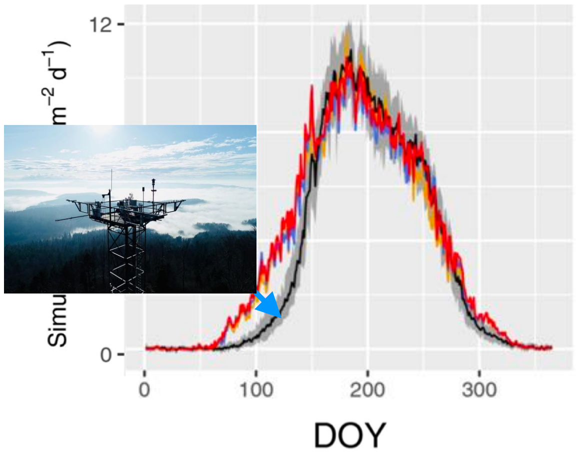
But: TER and GPP are predicted independently

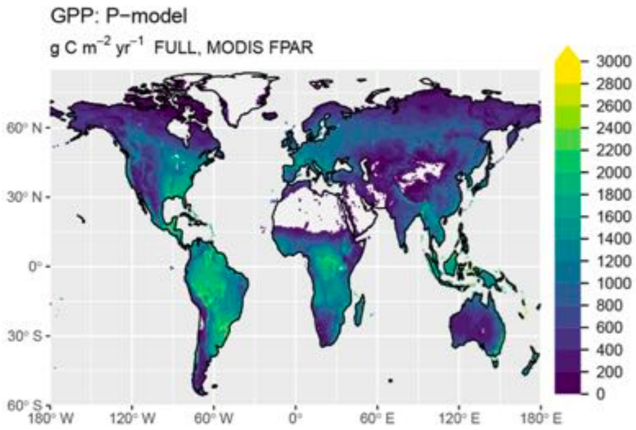


Tramontana et al., 2016

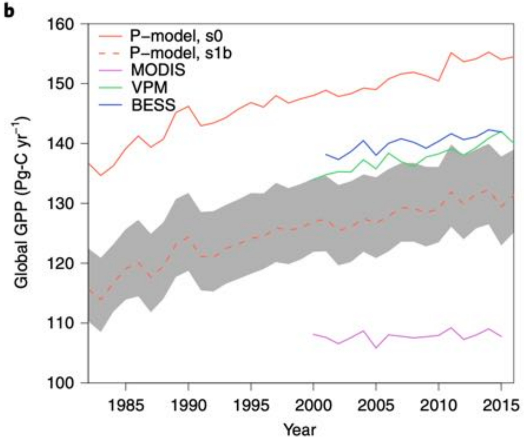
- Penalise “unphysical” predictions in the cost function
- Account for conservation laws in the network

- **ML for error correction**
  - E.g., numerical weather prediction to remove known biases
  - Combines advantages of theoretical understanding and data





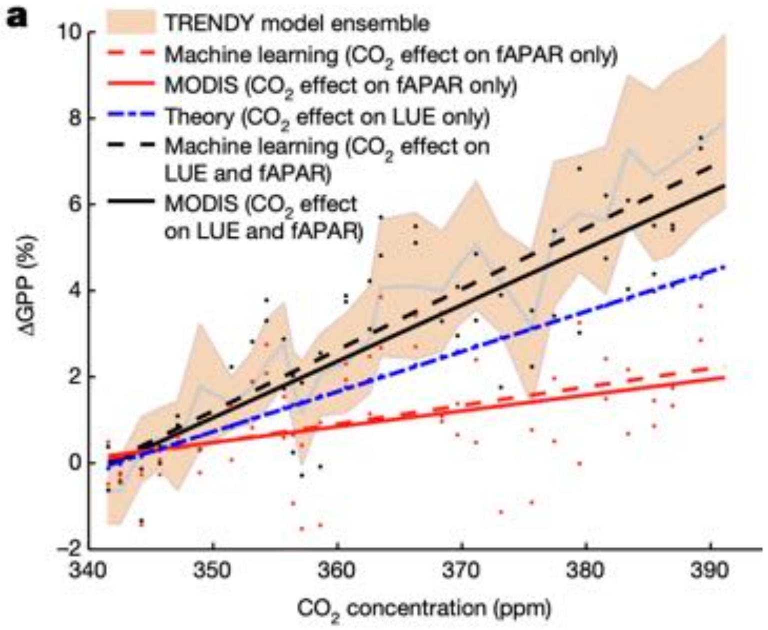
Stocker et al., 2020 *GMD*



Stocker et al., 2019 *Nature Geosci.*

B: Prediction

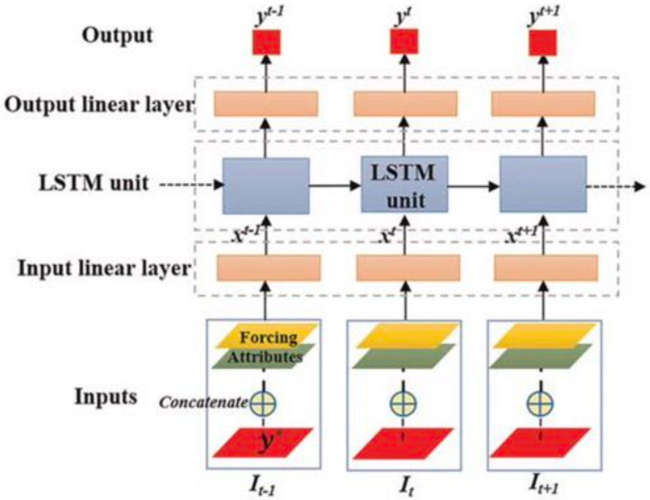
- **Adding theory to ML**

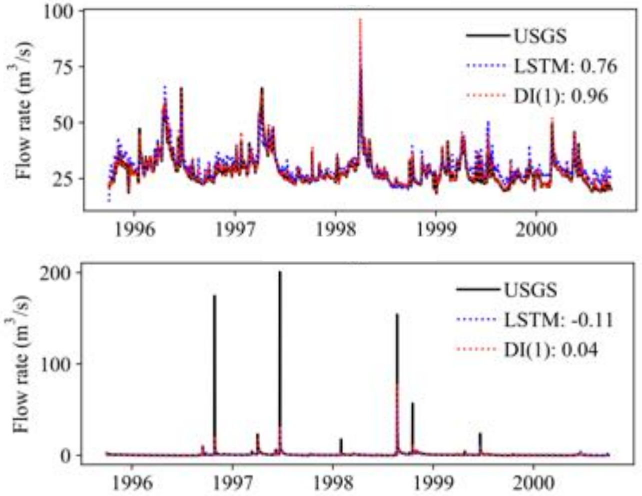


Keenan et al., *in prep.*

B: Prediction

- **Streamflow modelling**
  - **Catchment scale** $Y_t = f(\mathbf{x}_t, \mathbf{x}_{t-1}, \ldots$
    $\mathbf{x}_N$

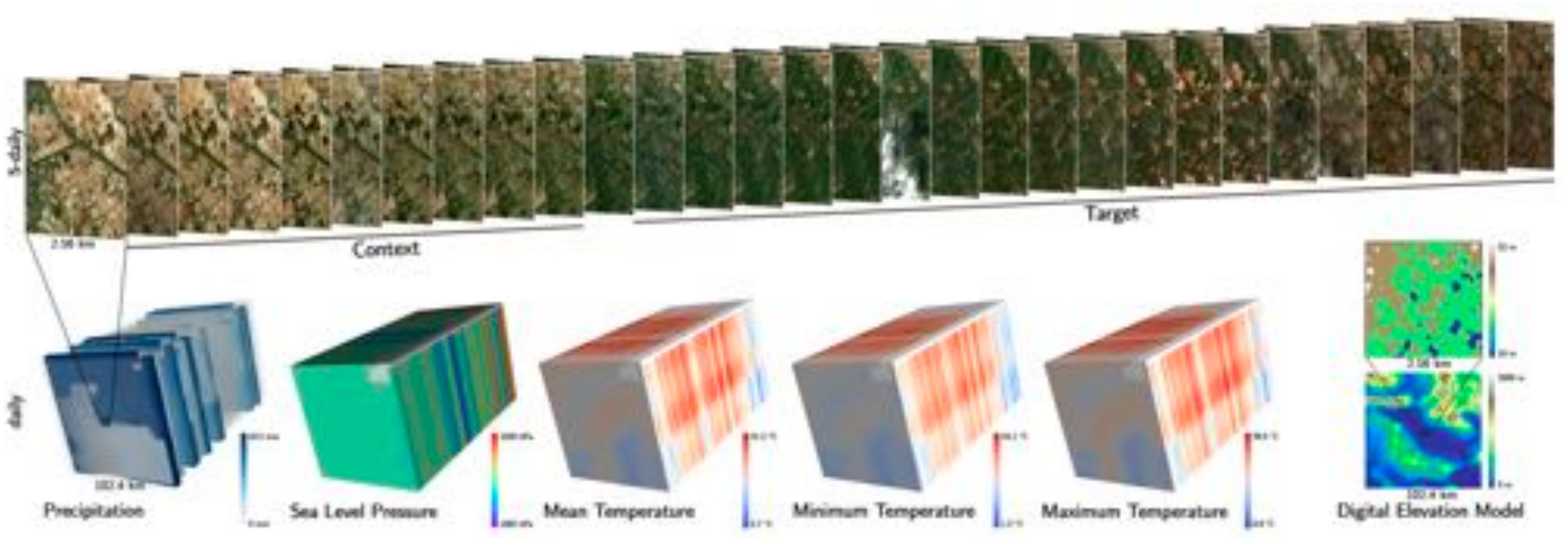




Feng et al., 2020 *WRR*

B: Prediction

- **Learning spatial-temporal dependencies**



EarthNet2021 Challenge
Requena-Mesa et al., 2021 *arXiv*

Attempt at classification of ML in Earth system modelling

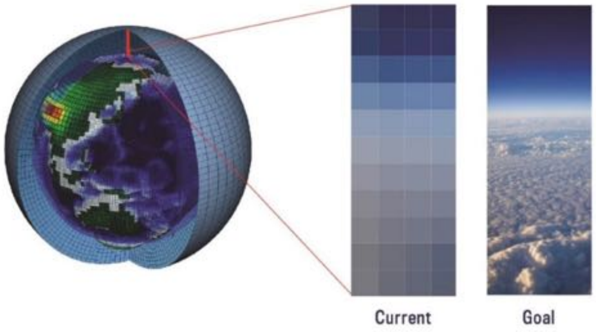
A: Making sense of information

B: (Black box) prediction

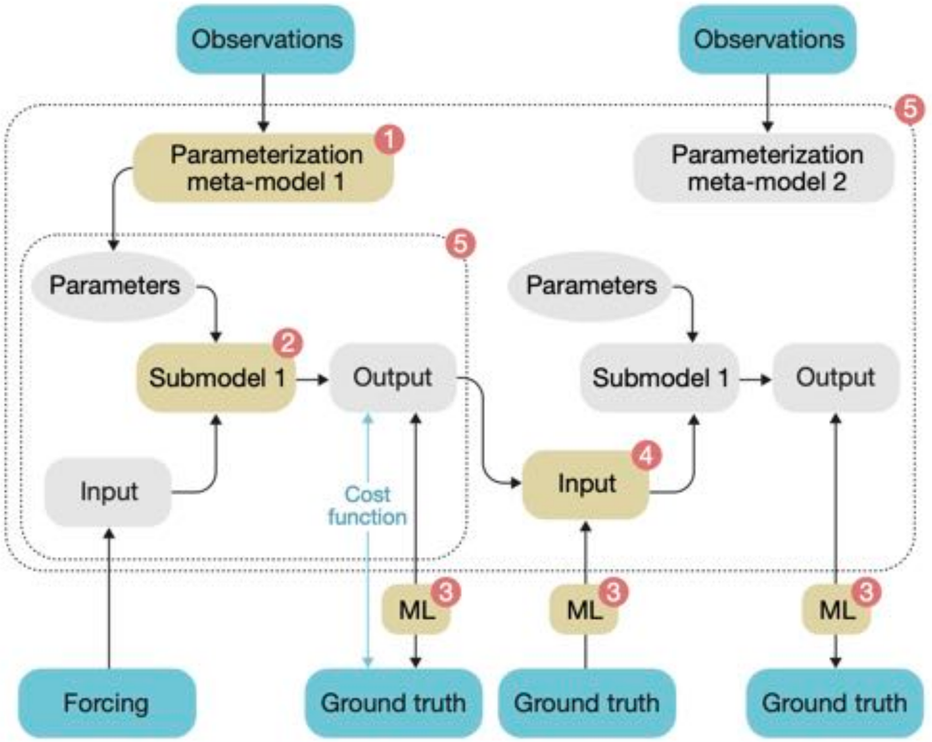
C: Hybrid modelling

C: Hybrid modelling

- ML informed by costly high-resolution model



C: Hybrid modelling



1. ML to learn the parameters of a physical model
2. ML for parametrisation of expensive process representations
3. Error interpretation (and correction)
4. Avoiding error propagation between submodules
5. Emulating the full model

Reichstein et al., 2019 *Nature*

This talk

1. Some success stories

2. **Our research in Geocomputation and Earth Observation**

3. Outlook

The drivers of ecosystem CO_2 exchange and gross primary production

Temperature

Radiation

Vapour pressure deficit

Soil moisture

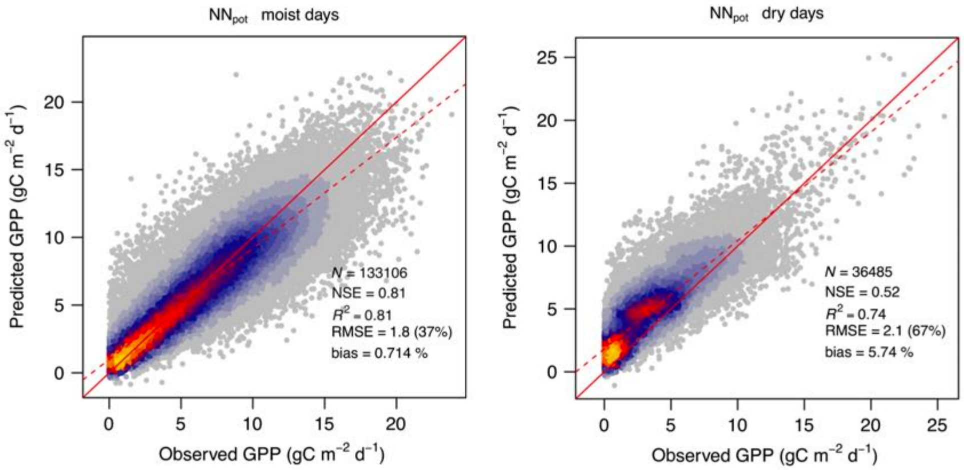
Net CO_2 exchange

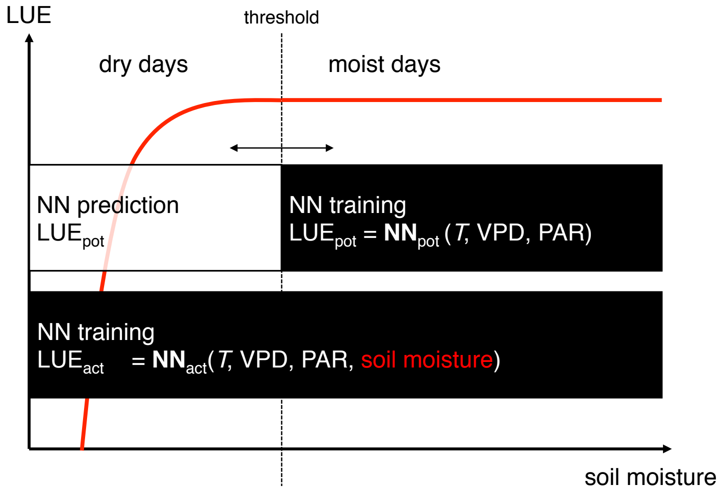
GPP

Theory-informed NN modelling

$$\text{GPP} = \text{PAR} \times \text{fAPAR} \times \text{LUE}$$

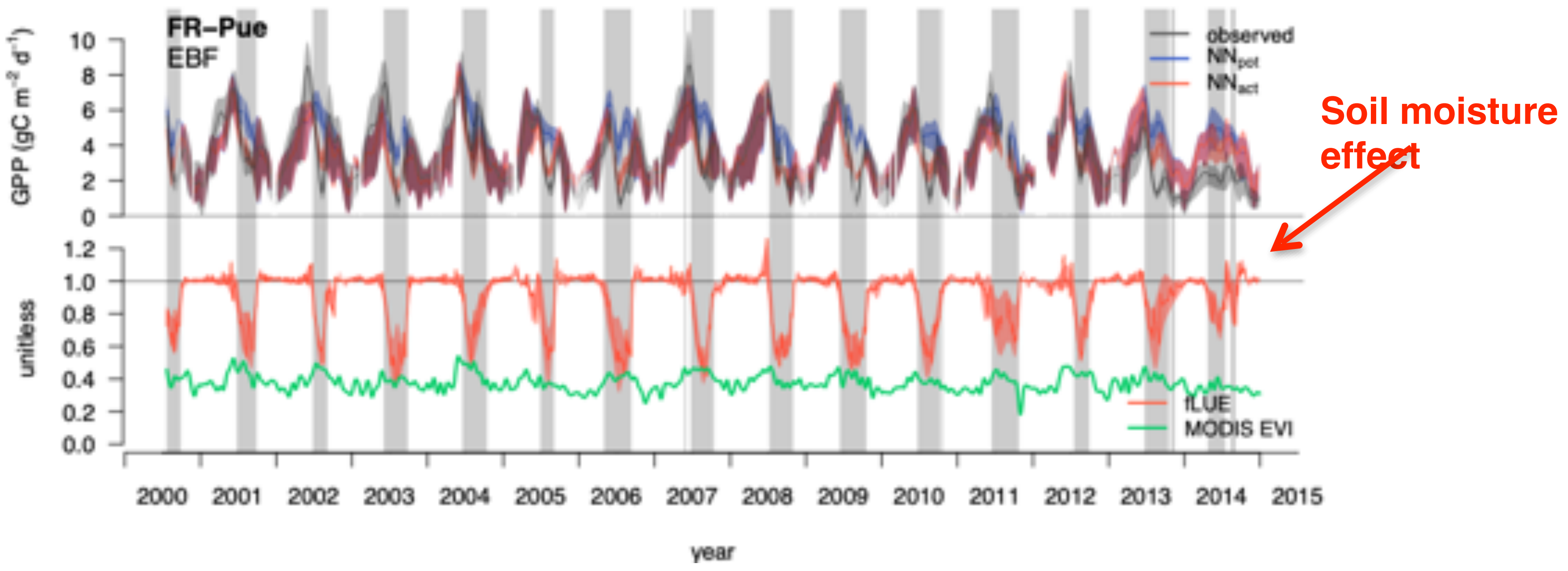
$$\text{fLUE} = \text{LUE}_{\text{act}} / \text{LUE}_{\text{pot}}$$





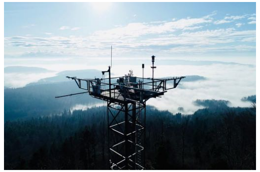
Stocker et al., 2018 *New Phytologist*

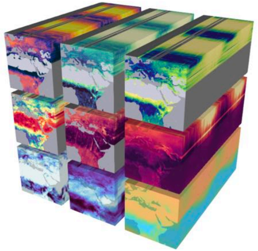
The soil moisture effect, separated



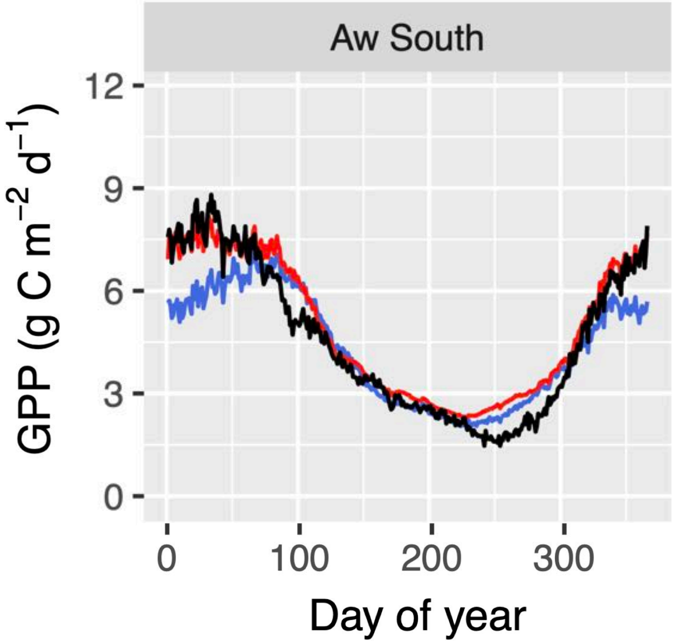
Stocker et al., 2018 *New Phytologist*

LSTM for GPP modelling



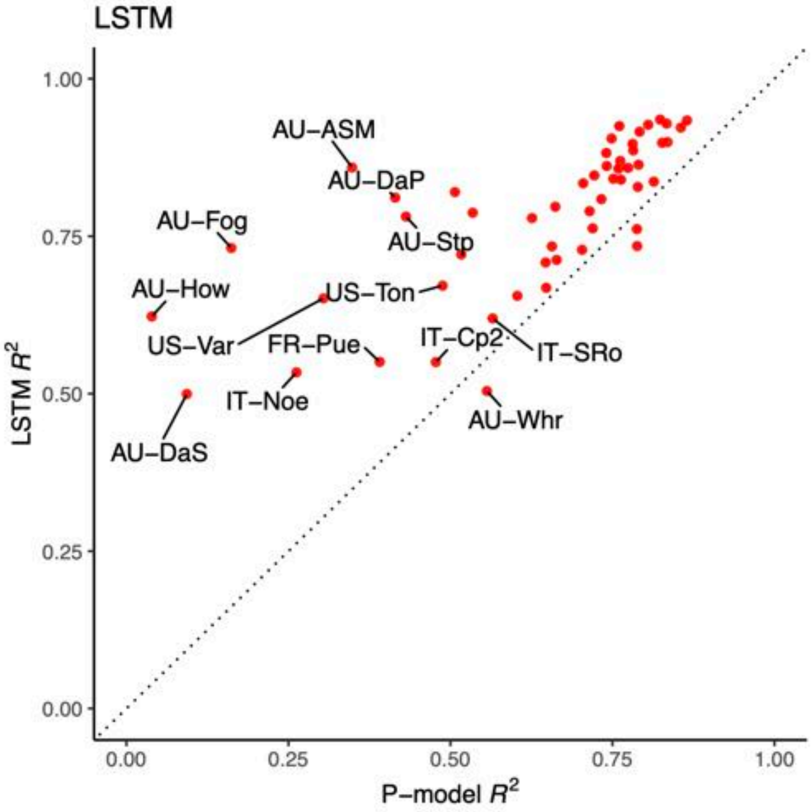


- **Point-scale** $Y_t = f(\mathbf{x}_t, \mathbf{x}_{t-1}, \ldots \mathbf{x}_N)$
- Considering the multivariate temporal context
- Avoiding iid assumption
- Avoiding hand crafted features (e.g, to capture memory effect of precipitation - soil moisture - water stress)



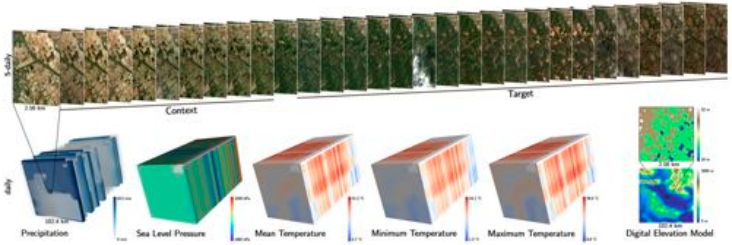
Stocker et al., *in prep.*

LSTM for GPP modelling

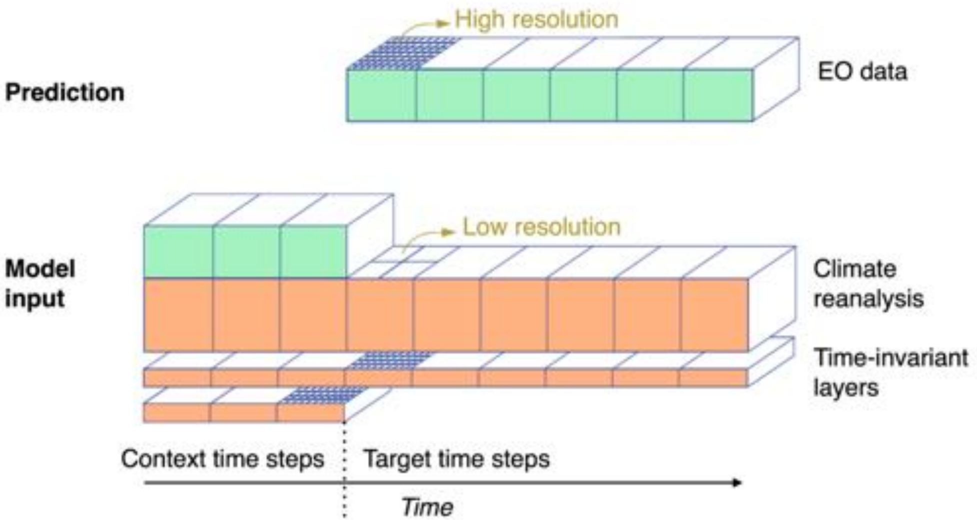


Stocker et al., *in prep.*

Data-driven drought impact prediction



- Considering spatial and temporal context (ConvLSTM)
- Implicit downscaling from coarse climate to local impacts



Stocker, *unpublished*

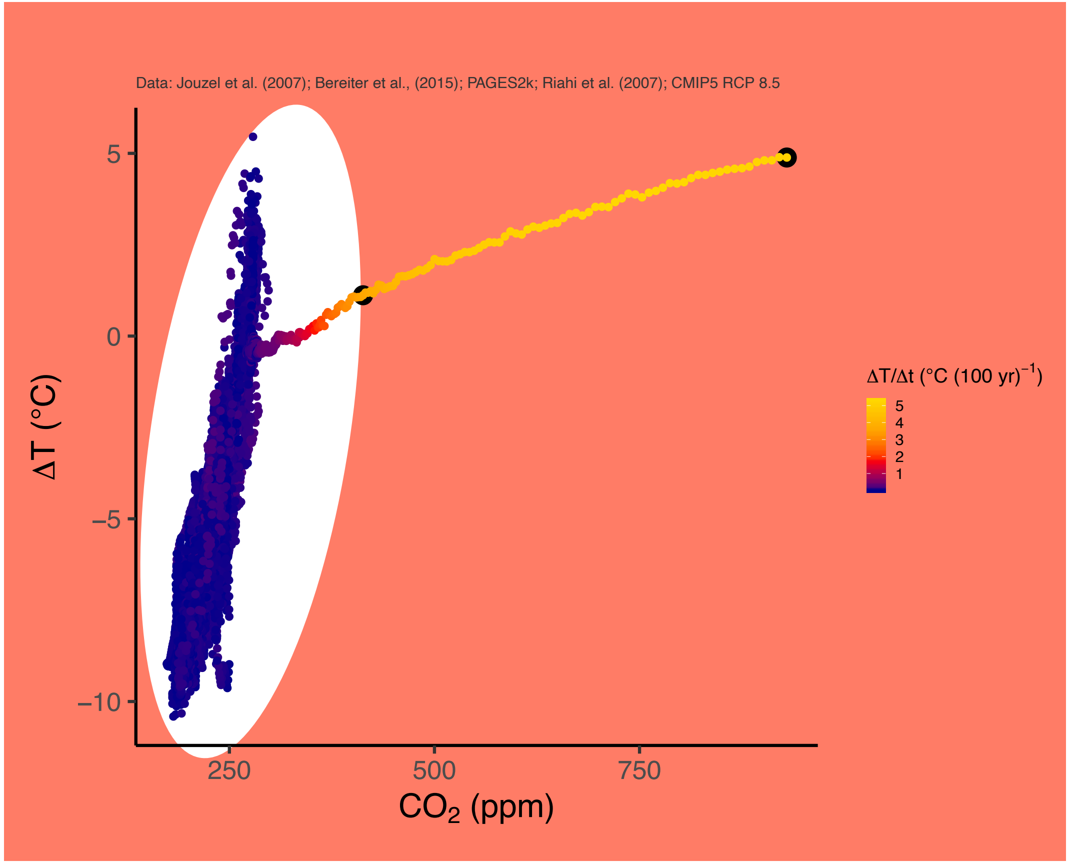
Data-driven drought impact prediction

Kladny et al., 2022 *bioRxiv*

This talk

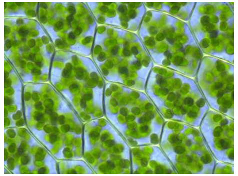
1. Some success stories
2. Our research in Geocomputation and Earth Observation
3. Outlook

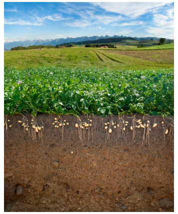
Moving outside the domain of past observations

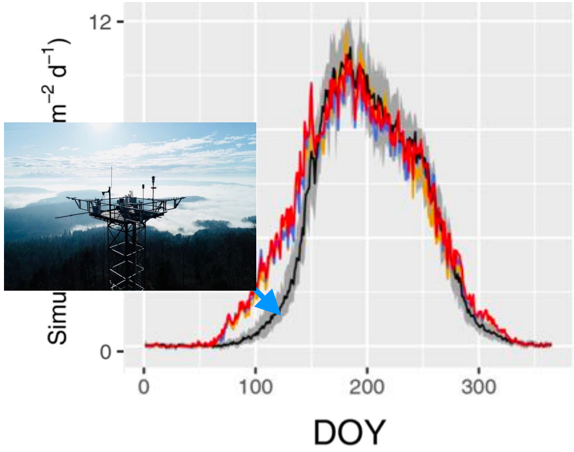


Stocker, *unpublished*

The CO_2 fertilisation cascade



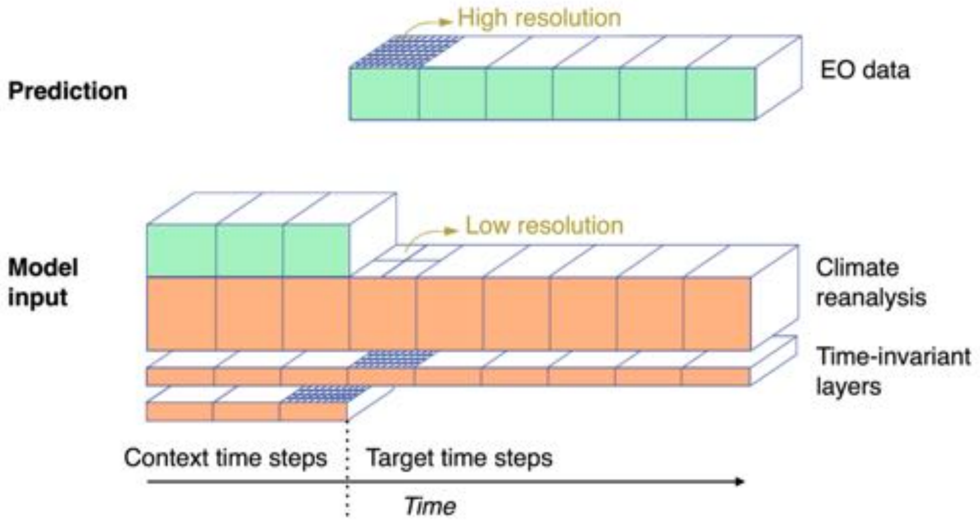




Earth observation-informed near-term forecasting

Promising applications

- Crop and grass yield forecasting
- Wildfire forecasting



Stocker, *unpublished*

“Conclusion”

- Many applications of ML that are complementary and supporting Earth system modelling.
- Interesting: Combining ML with theory and physical laws.
- Interesting: Hybrid modelling where ML can become useful as a part, but not to replace ESM.

GECO Group

GECO

Geocomputation and Earth Observation

https://github.com/computationales

SCHMIDT FUTURES FNSNF