Earth vision 2020 workshop
CVPR conference, Seattle, June 14-19, 2000
Devis Tuia, Jan Wegner, Ronny Hänsch

Linking machine learning and physical-biological dynamic system modelling

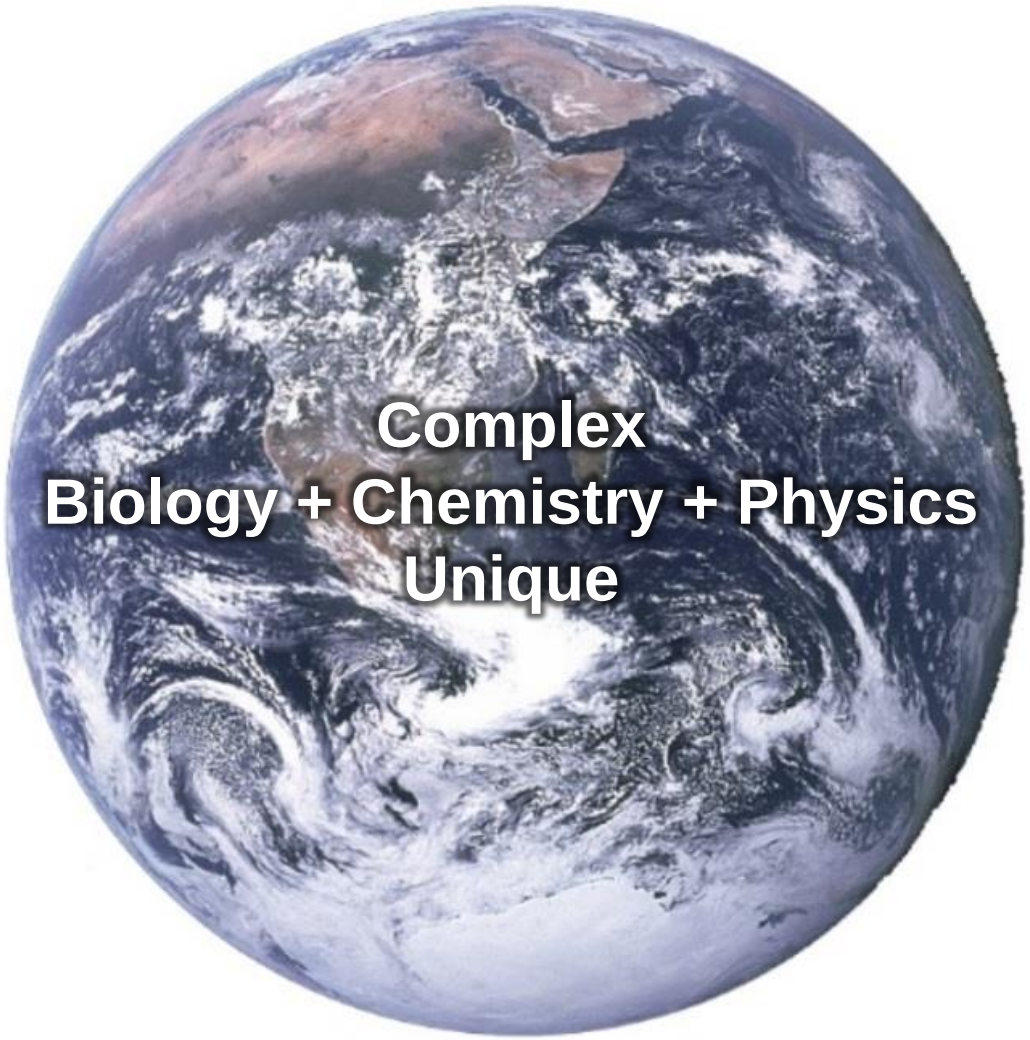
Markus Reichstein*, Fabian Gans, Basil Kraft, Nuno Carvalhais, Martin Jung

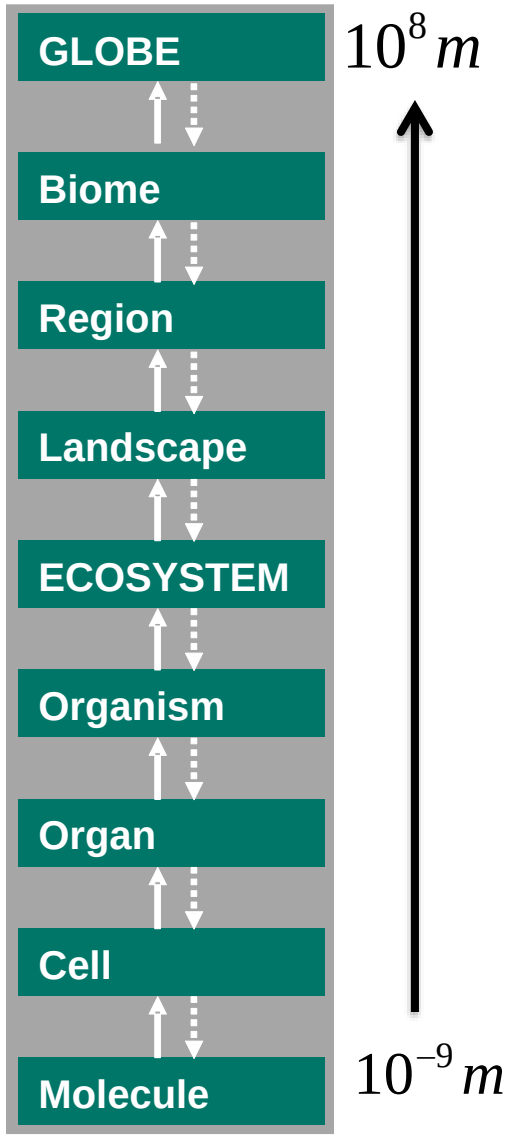
Max-Planck-Institute for Biogeochemistry, Jena

Michael-Stifel-Center Jena for Data-driven and Simulation Science

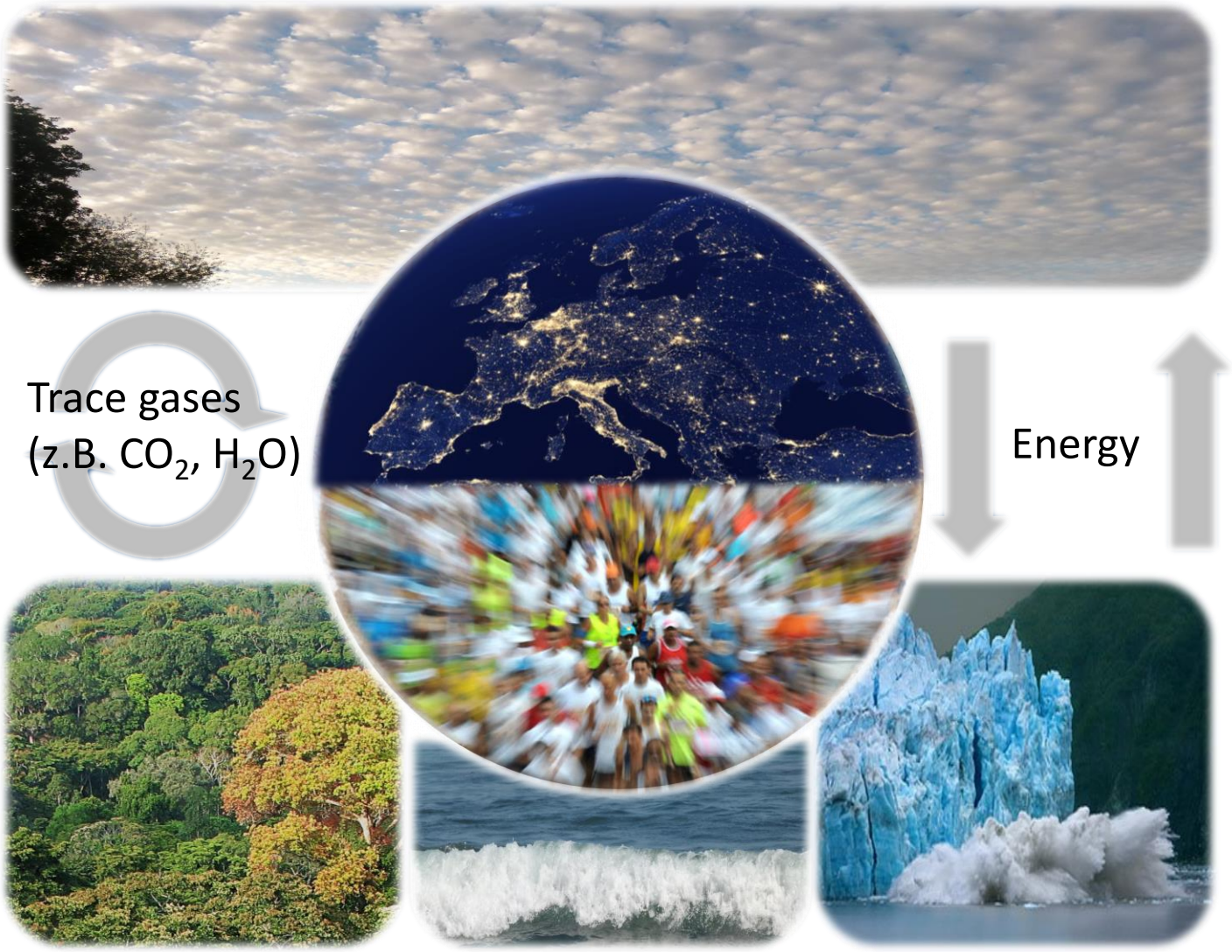
**mreichstein@bgc-jena.mpg.de, @Reichstein_BGC*

The Earth System



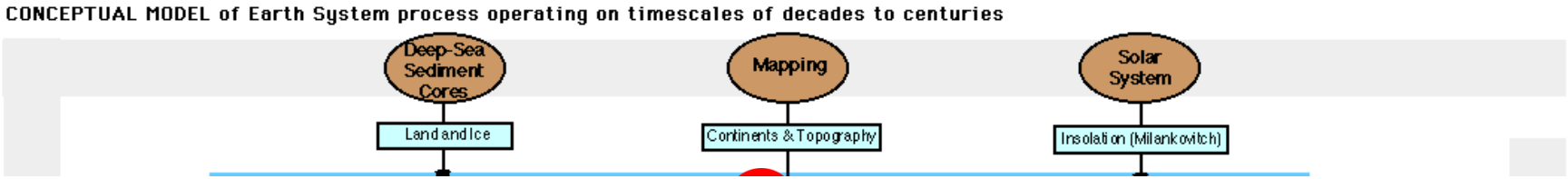


“Spheres” in the Earth System



Earth System Science established in the Max-Planck-Society

Reductionistic approach: ”wiring all together”

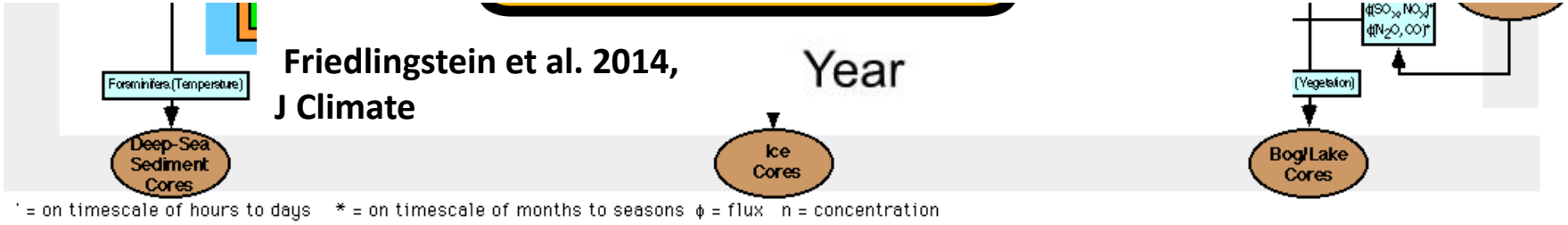


PERSPECTIVE

https://doi.org/10.1038/s41586-019-0912-1

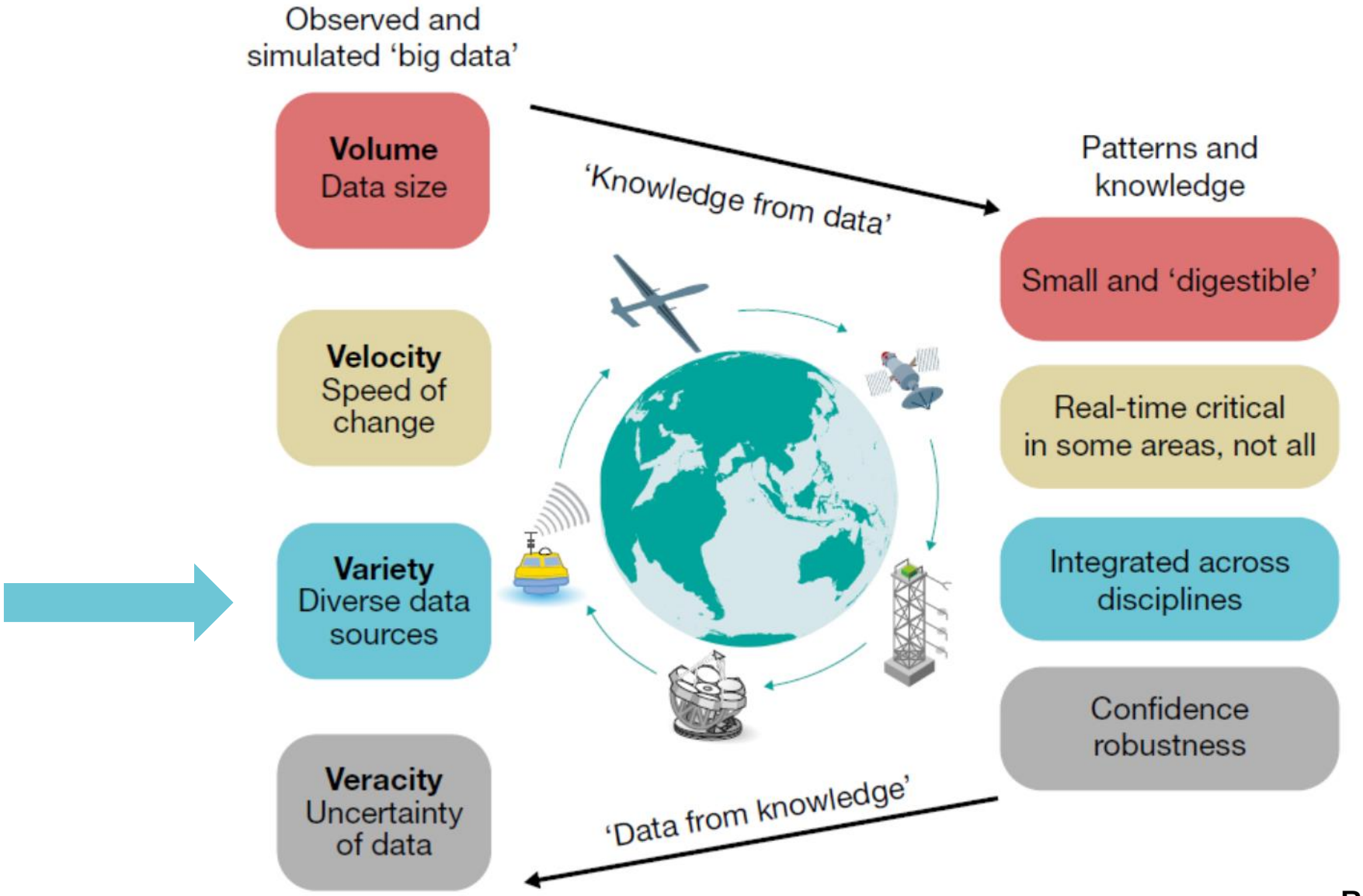
Deep learning and process understanding for data-driven Earth system science

Markus Reichstein[1,2*], Gustau Camps-Valls[3], Bjorn Stevens[4], Martin Jung[1], Joachim Denzler[2,5], Nuno Carvalhais[1,6] & Prabhat[7]



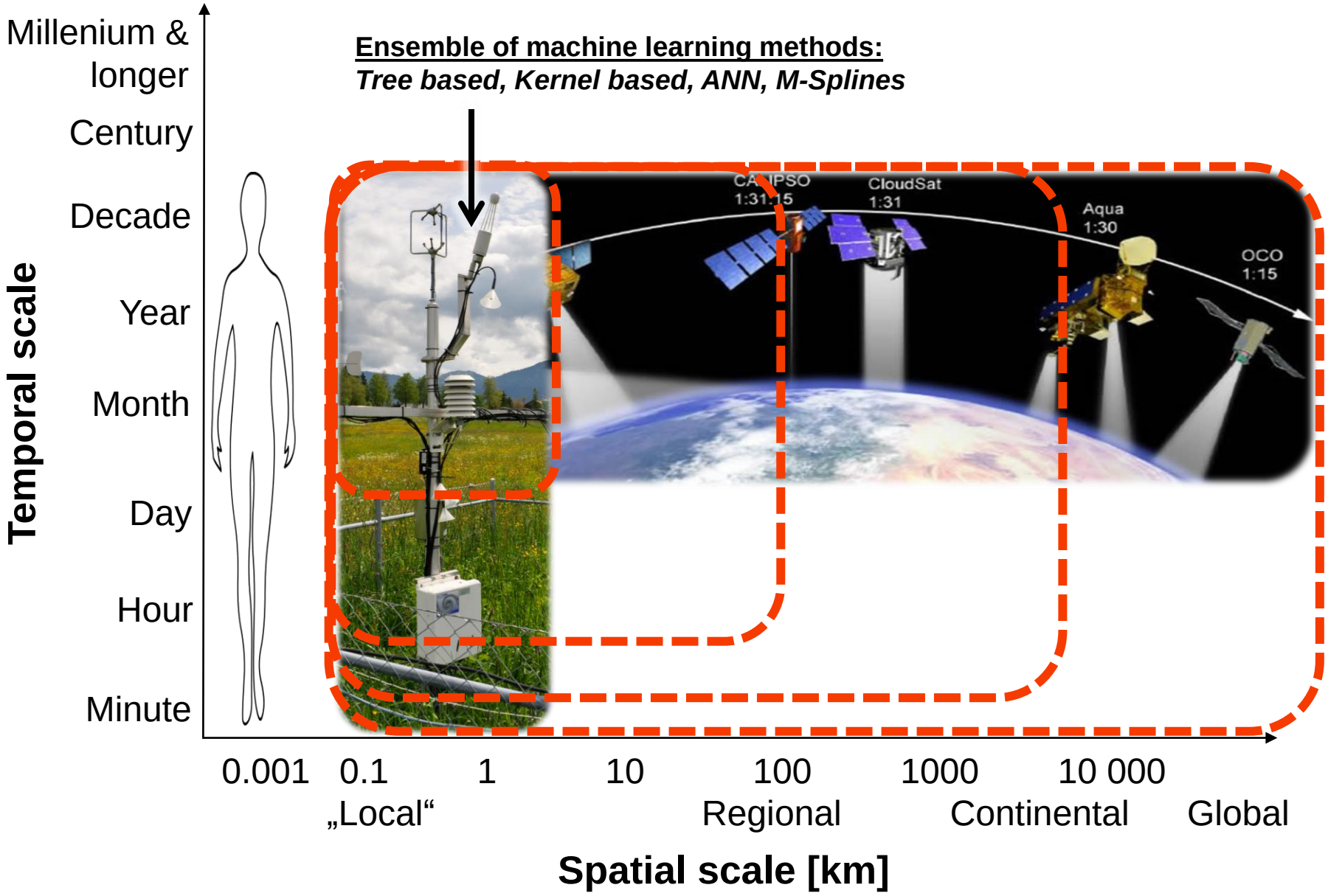
Friedlingstein et al. 2014, J Climate

Data-driven Earth System Science: prototypical for „big data“



Reichstein et al. 2019, Nature

Integrating the variety: “up-scaling” example



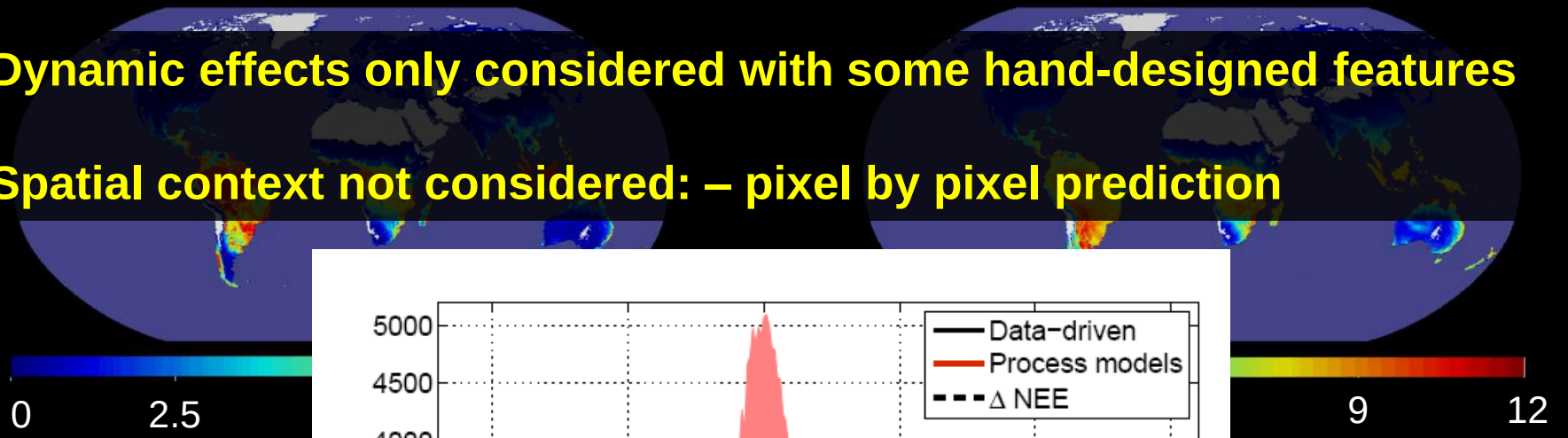
Data-driven view on dynamic Biosphere-Atmosphere Exchange

Primary production (GPP) [g m^{-2} day^{-1}]

Evapotranspiration [MJ m^{-2} day^{-1}]

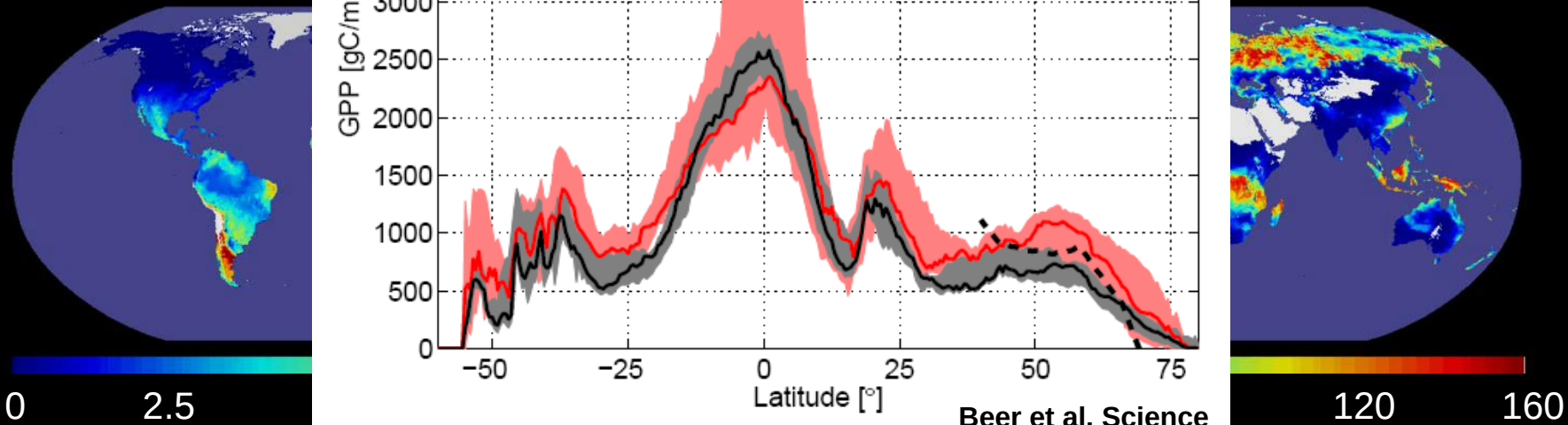
Dynamic effects only considered with some hand-designed features

Spatial context not considered: – pixel by pixel prediction



Sensible heat

bility [index]



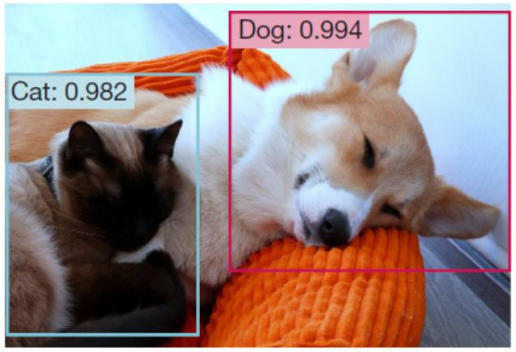
Data: Jung et al. (2010, 2017), Nature. Animations: F. Gans, MPI-BGC

Deep learning for Earth System Science: Examples

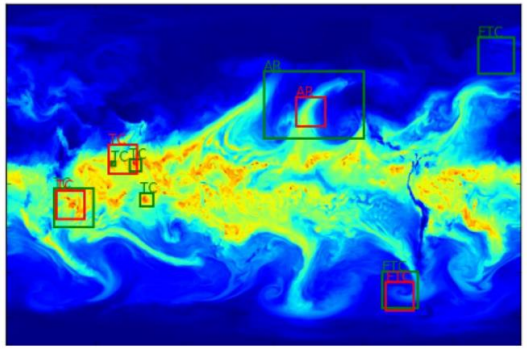
Max Planck Institute for Biogeochemistry

Machine learning tasks	Earth science tasks	Machine learning tasks	Earth science tasks

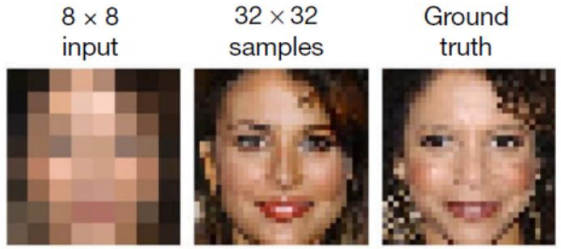
a Object classification and localization



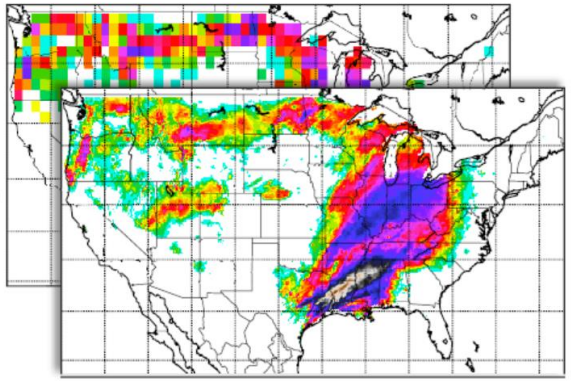
Pattern classification



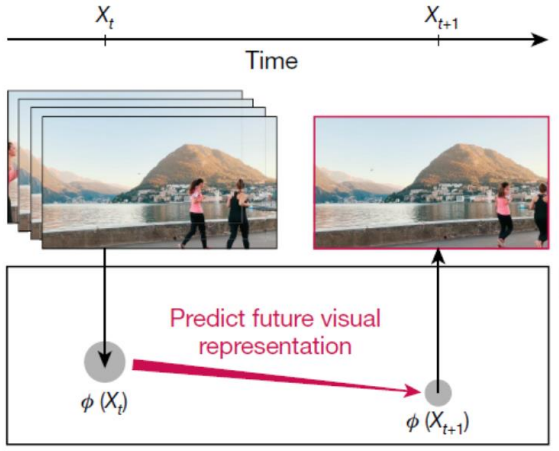
b Super-resolution and fusion



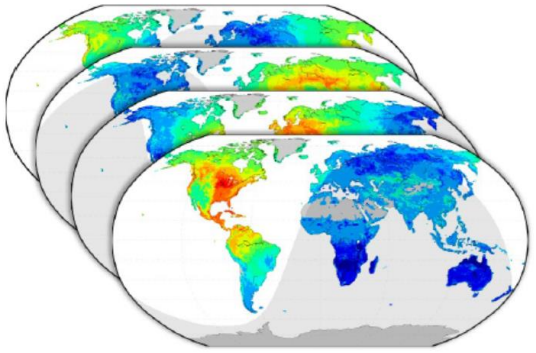
Statistical downscaling and blending



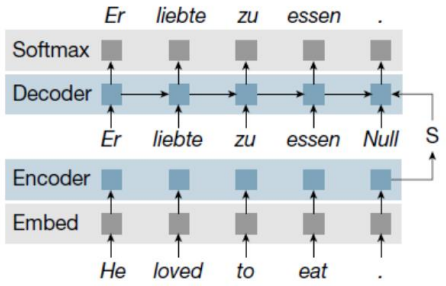
c Video prediction



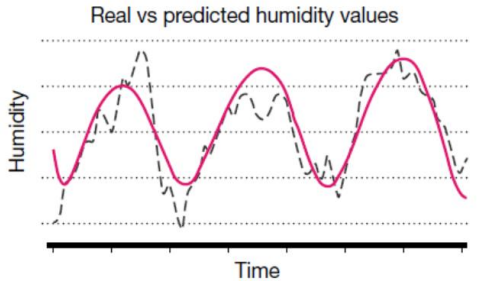
Short-term forecasting



d Language translation



Dynamic time series modelling

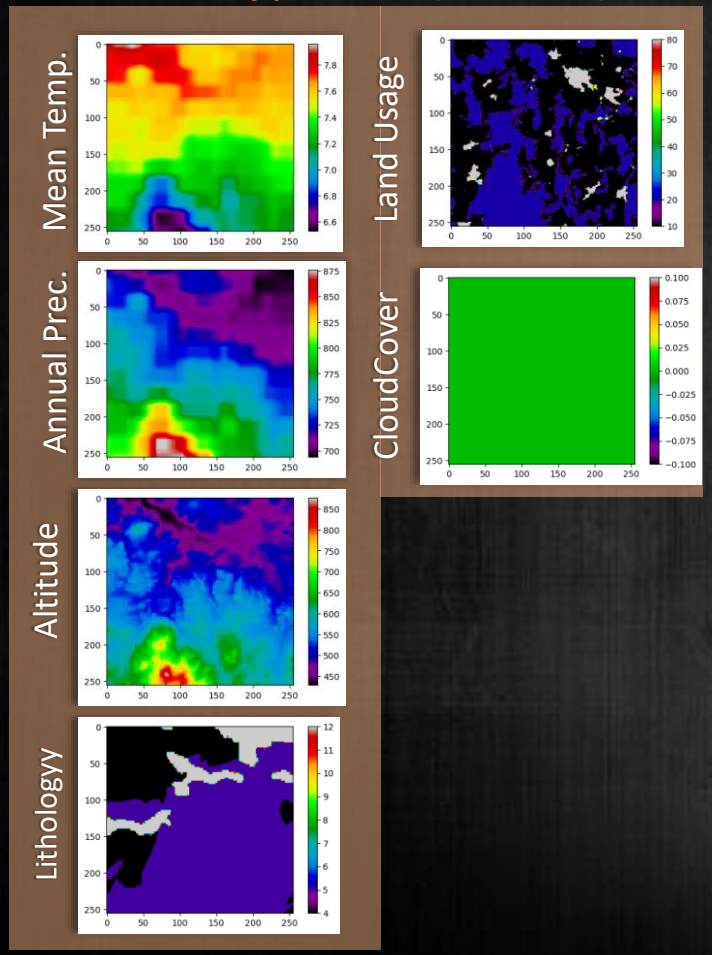


Reichstein et al. (2019)

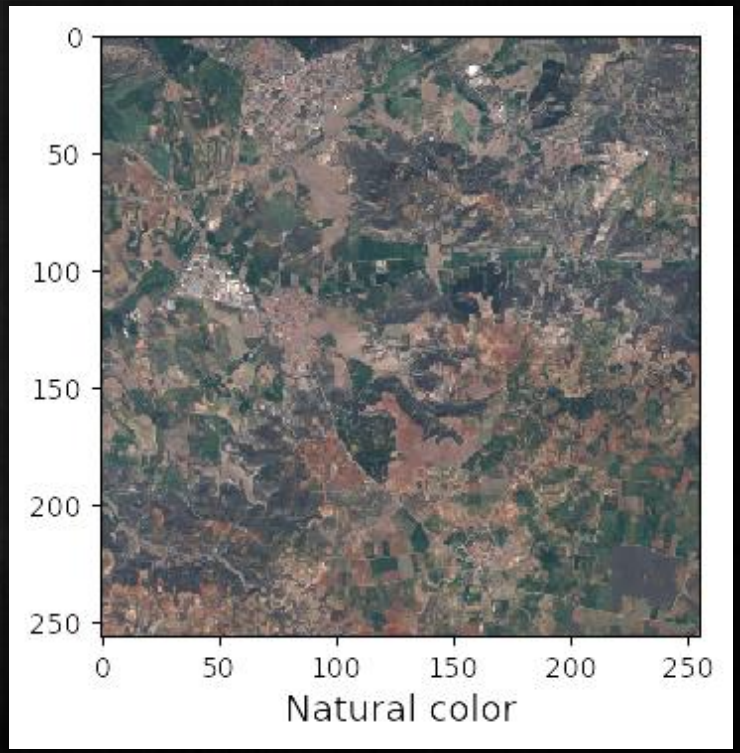
“Predicting” whole landscapes as seen from space

Example data sample : Tile 33UVQ86 (1st April 2017)

Primary predictors (conditions)

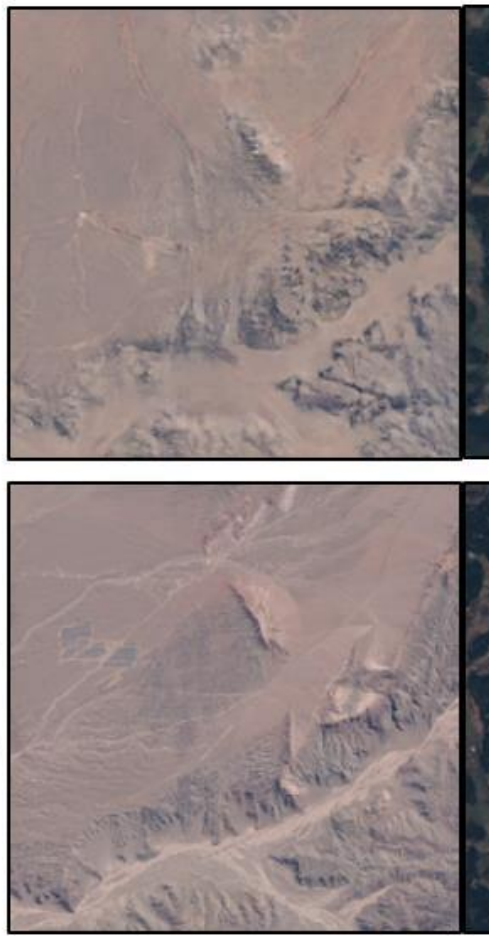


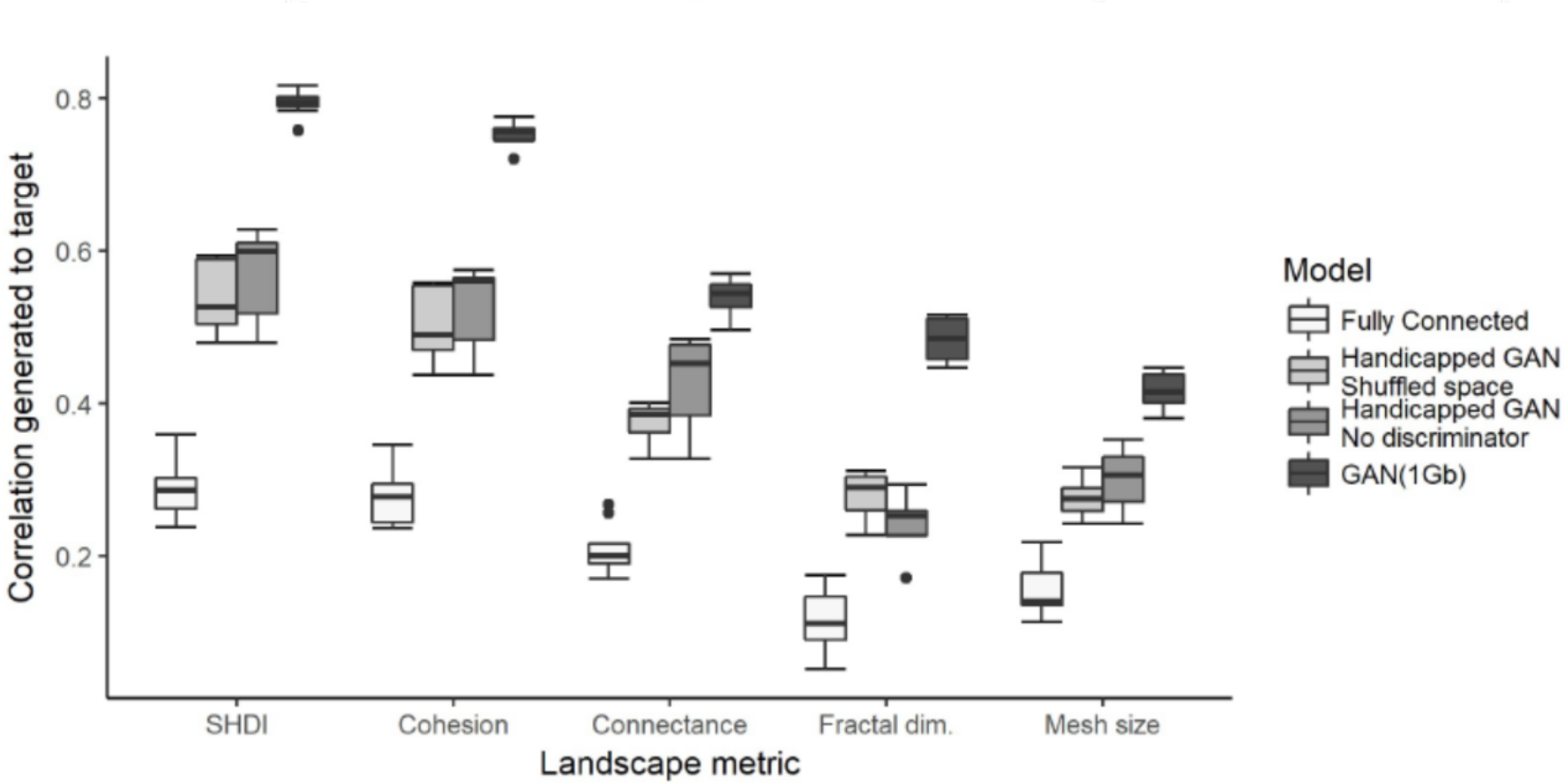
cGAN



Requena et al. 2018, IEEE

“Predicting” whole landscapes as seen from space

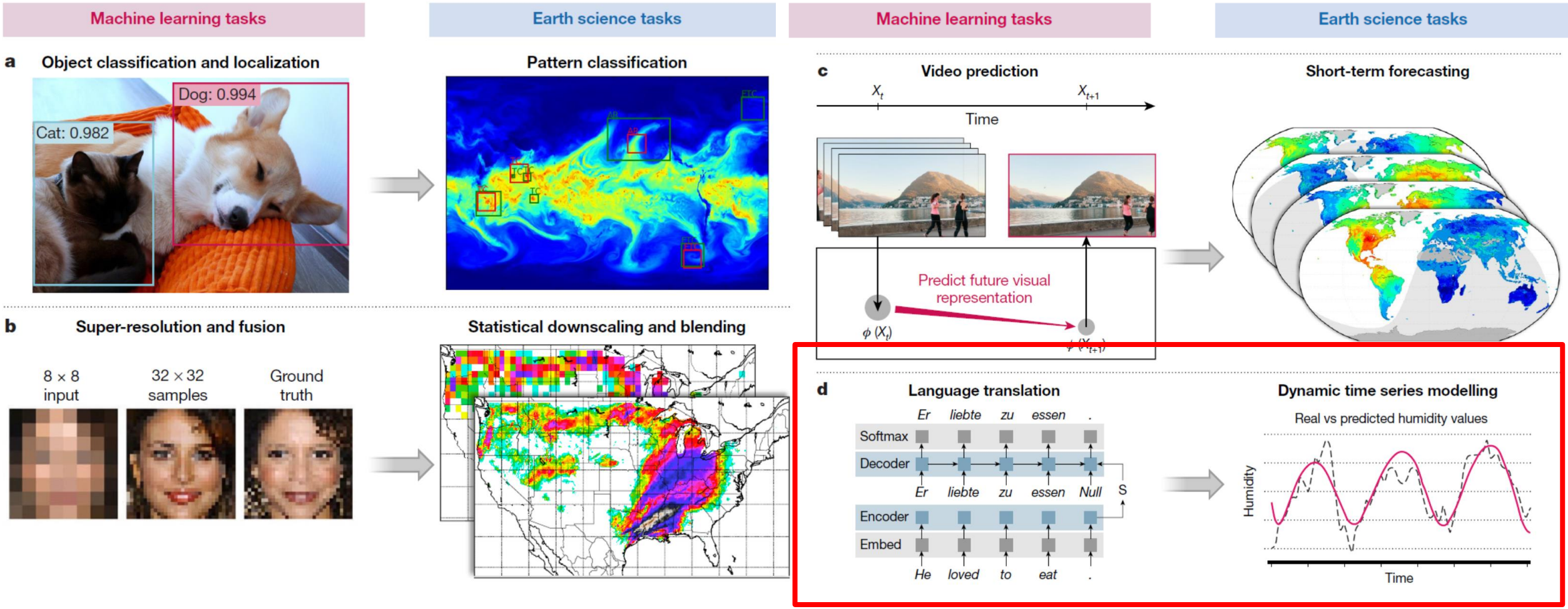




<u>But:</u> Understanding and physical consistency???

Requena et al. 2019, GCPR

Deep learning for Earth System Science…

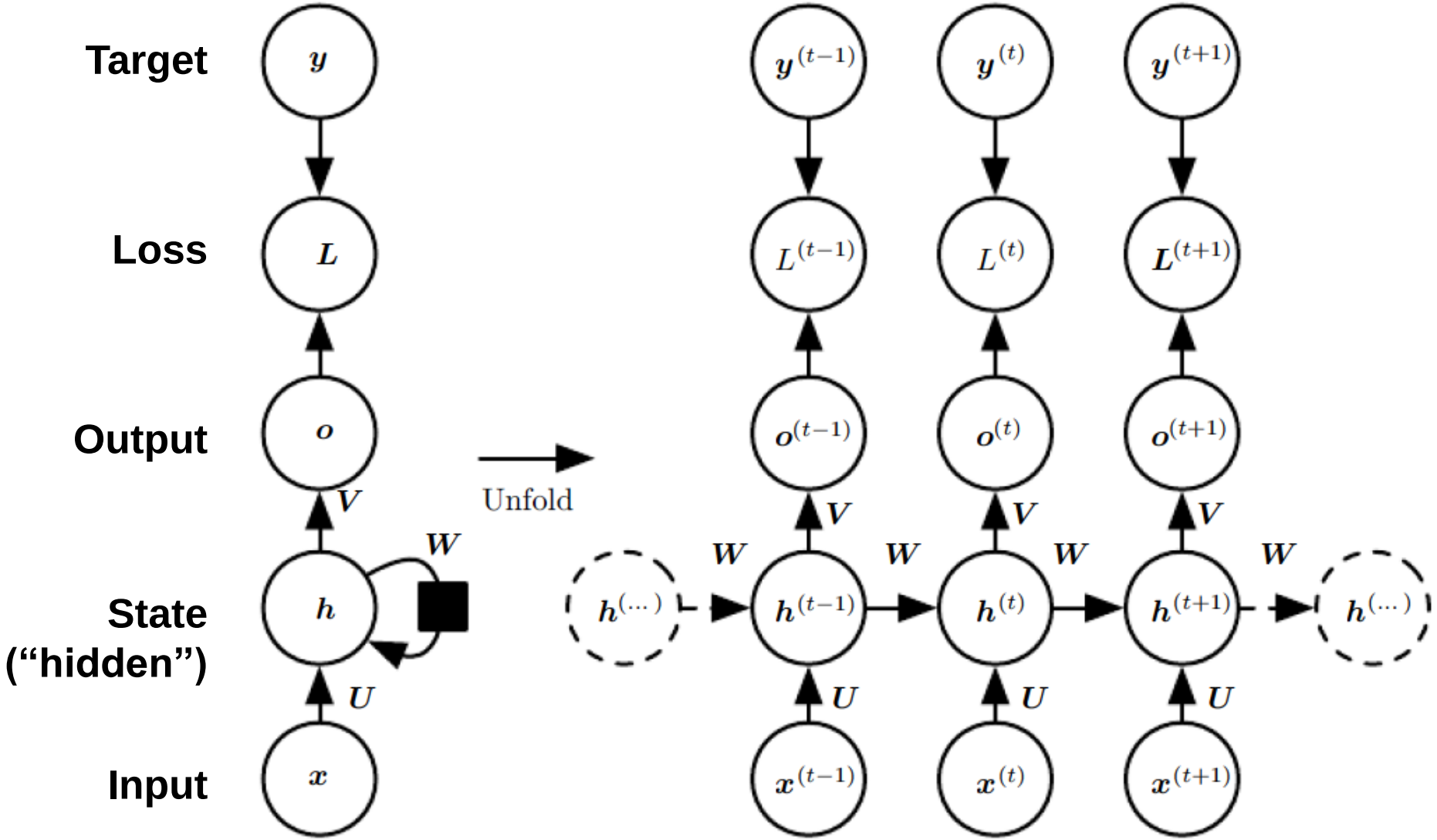


Reichstein et al. (2019)

Describing and detecting dynamic 'memory' effects

- Time-varying properties which depend on past (possibly latent) variables
- Typically described with differential or time-discrete analogue equations
- Examples:
  - Vegetation development depends on cumulative temperature over winter-spring ("temperature sums")
  - Simple water balance: $SM(t) = \int_{t_0}^{t} [Q(t) - ET(t) - D(t)]dt$
  - Depletion of carbohydrate reserves in a drought and carry-over effects

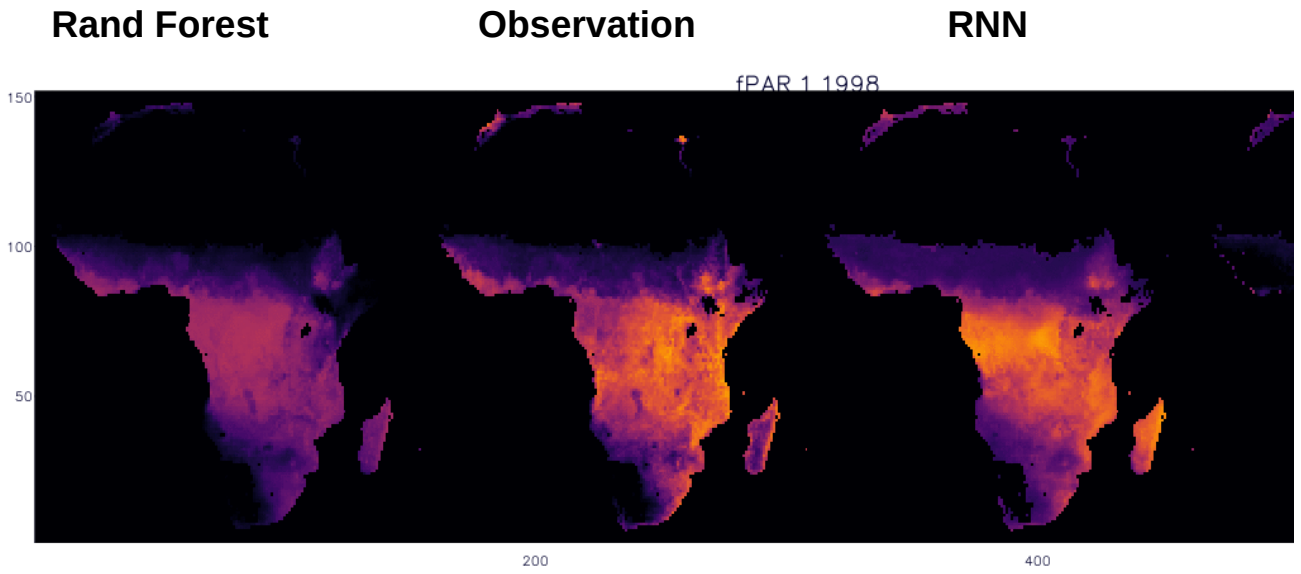
Describing & detecting dynamic memory effects: RNNs

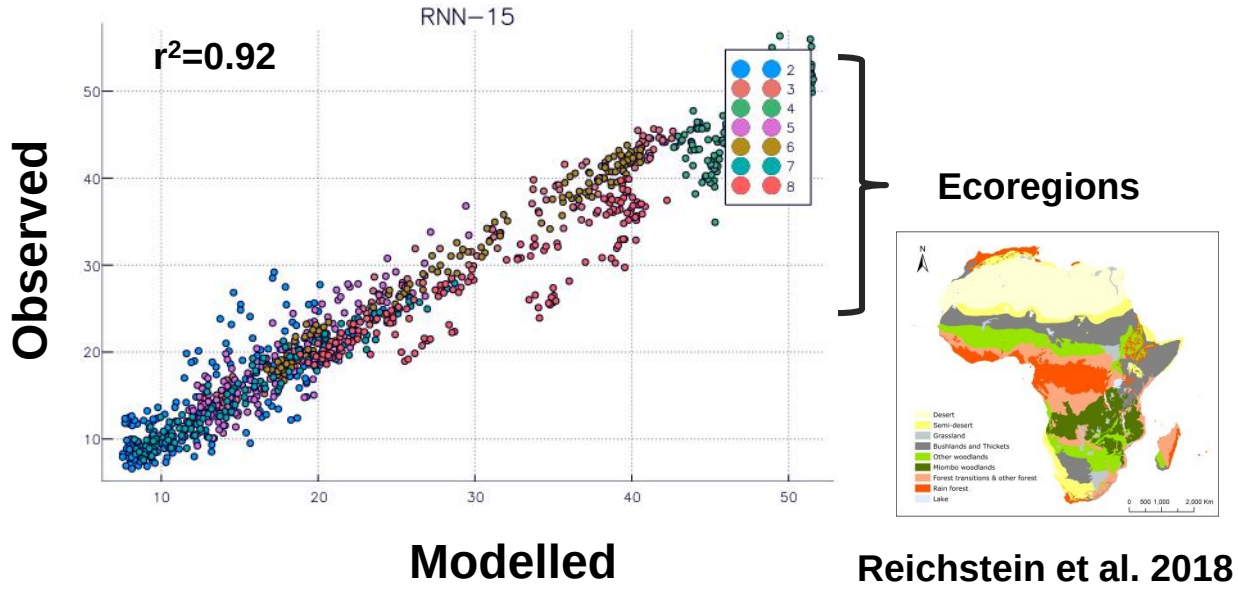


after Karpathy (2

Modelling vegetation state ("leaves" or "greenness")

- Target: GIMMS fPAR variability over Africa, 0.5° lat/lon, monthly, 1982-2014
- Two approaches:
  1. Random Forest with standard meteorological predictors **plus lagged and cumulative water variables** (e.g. relative humidity, soil moisture previous months) ← from Feature selection algorithm [Jung et al.]
  2. Recurrent neural network with **only** standard meteorological drivers and vegetation type etc. (trained on 4% of the pixels only)

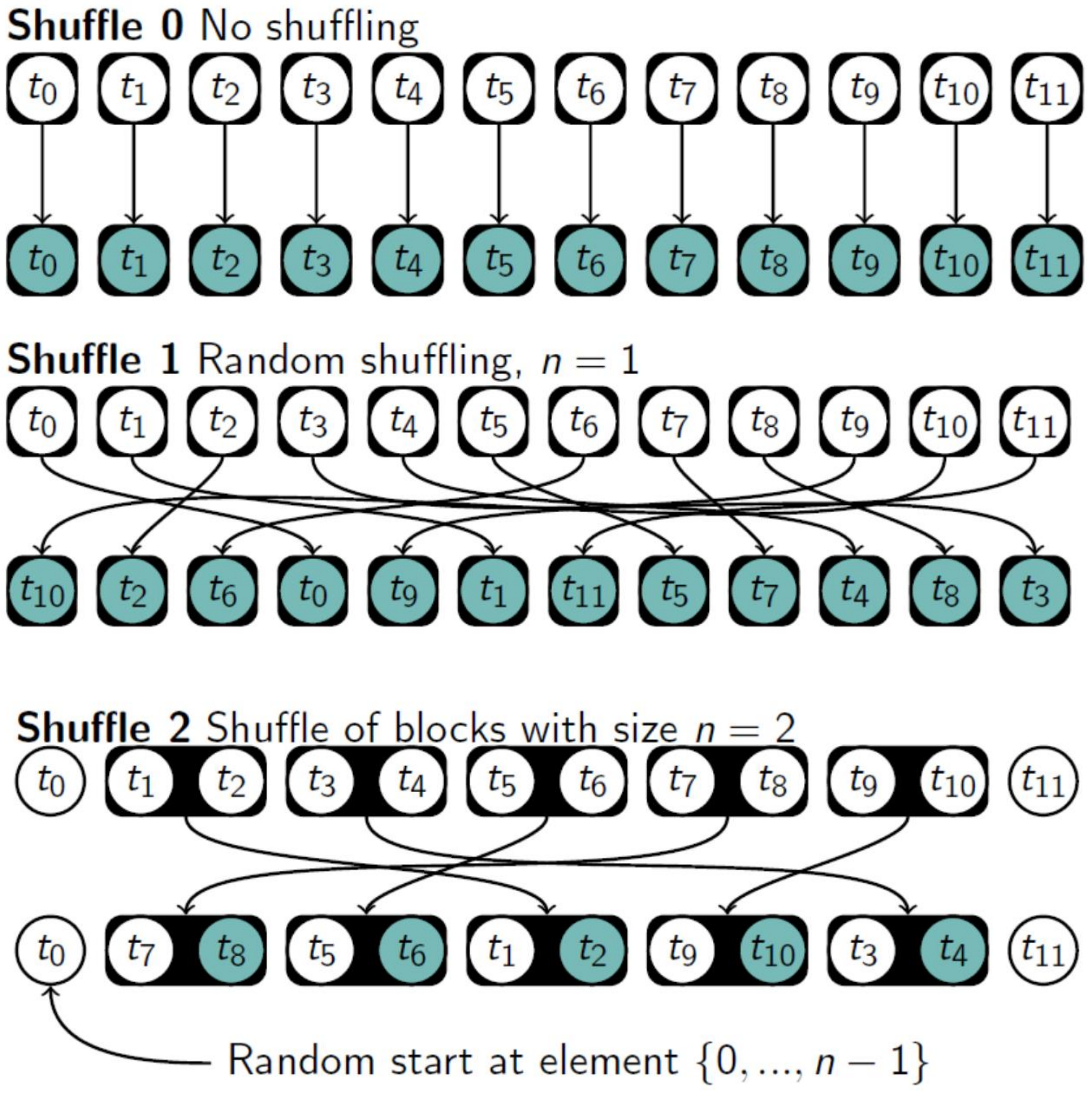




Ecoregions

Reichstein et al. 2018 IEEE)

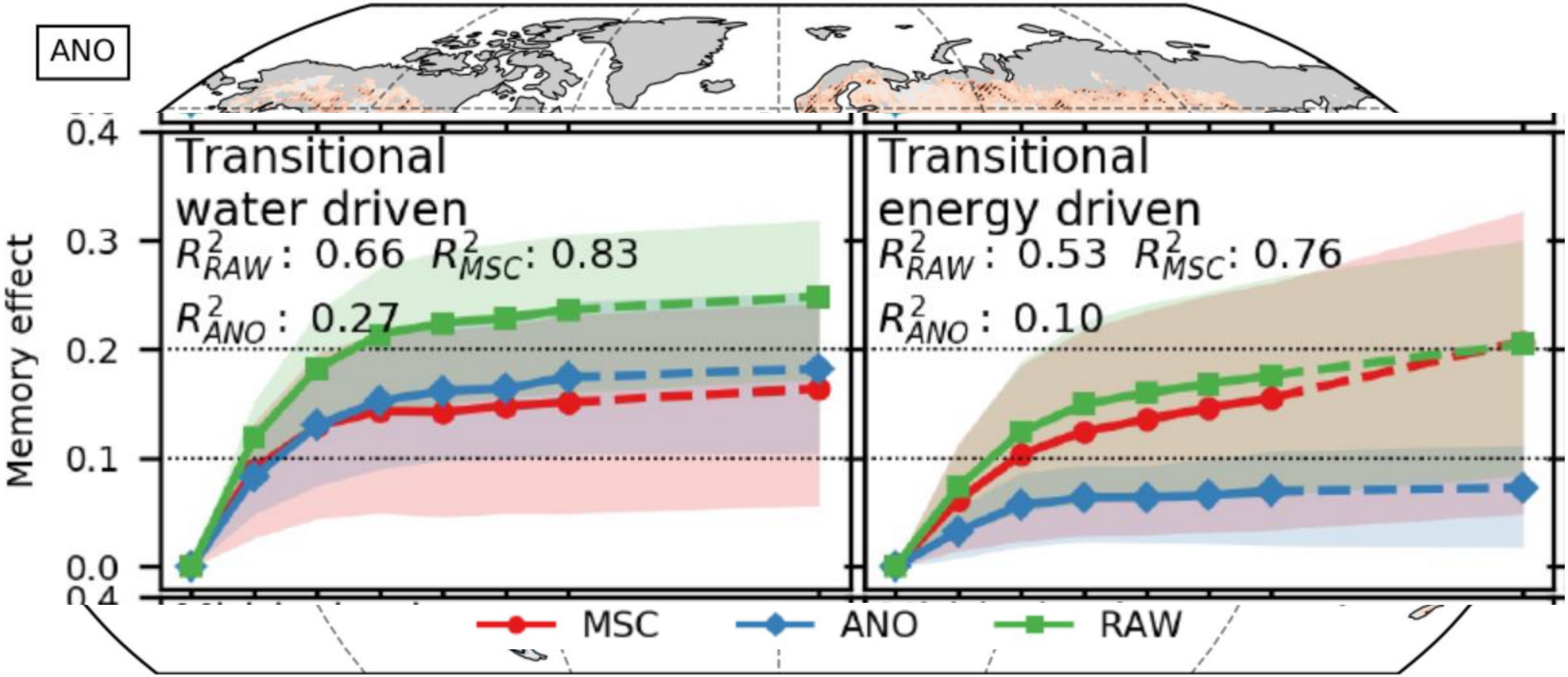
Detecting memory effects by permutation experiments



Memory effect
:=
Error with permutation
minus
Error without permutation

$$Mem := RMSE_{perm} - RMSE_{nonPerm}$$

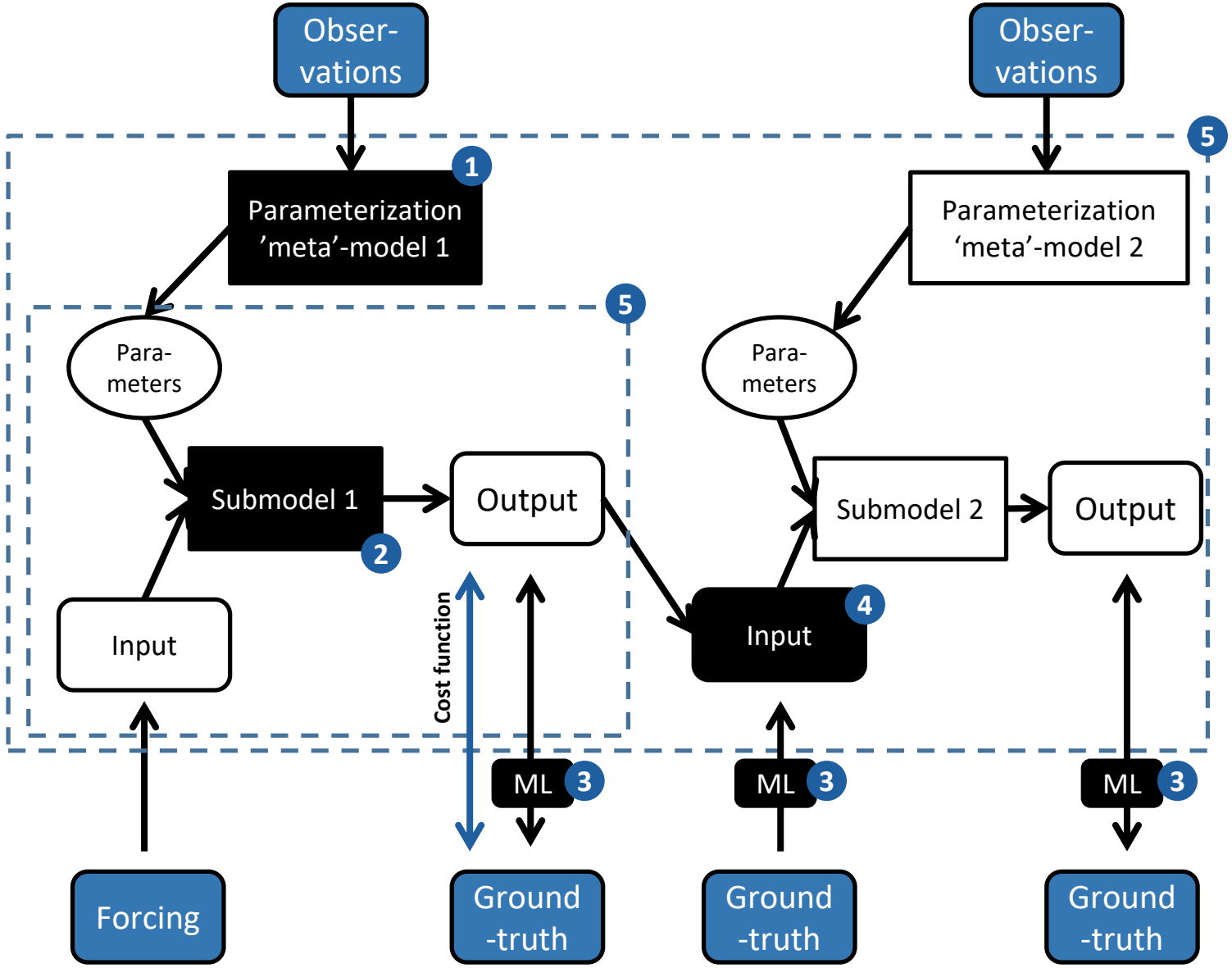
Map of memory effects on vegetation anomalies



To do: Attribution to specific drivers and (latent) states

Kraft et al. 2019, Frontiers in Big Data

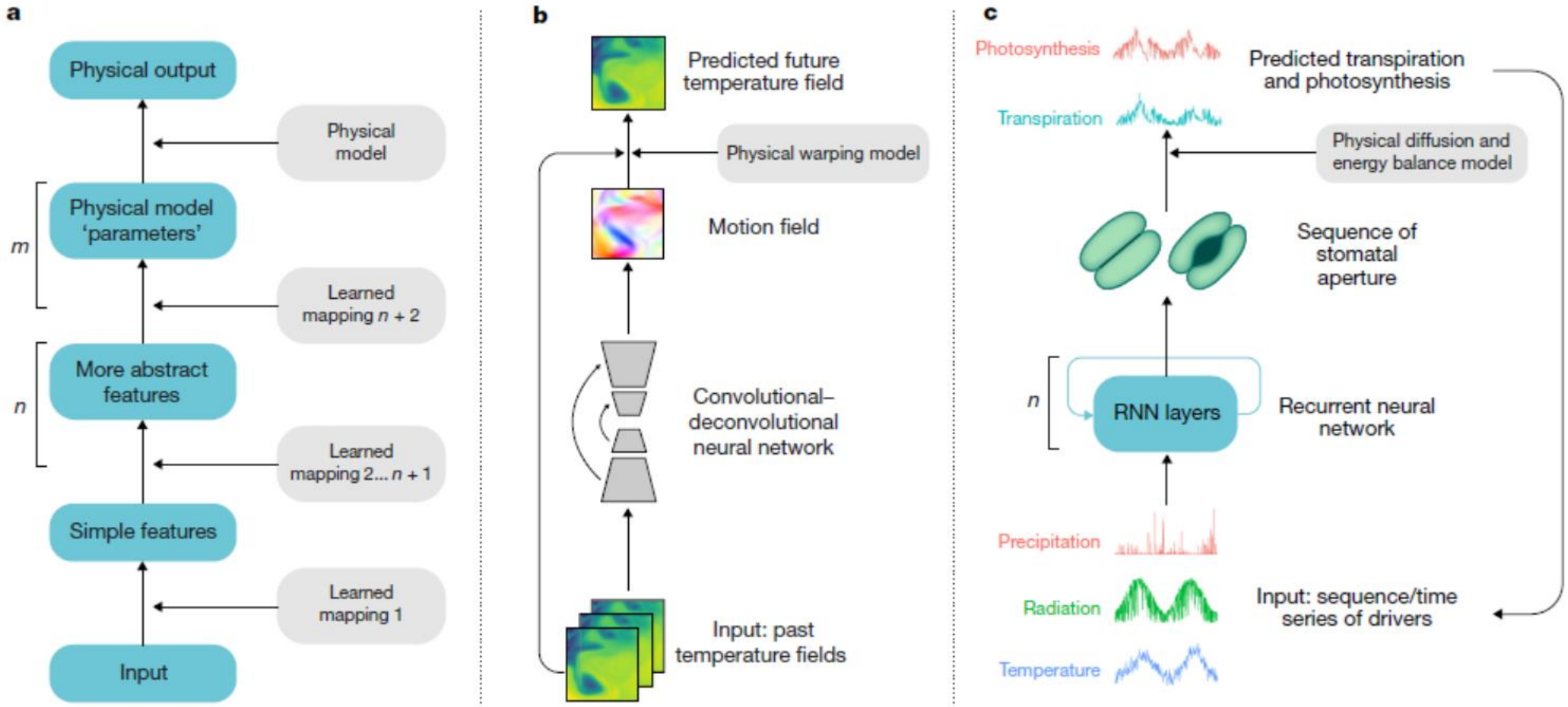
Model-data-machine-learning integration…



1. **Model parameterization**
2. **Hybrid modelling**
3. **Pattern-oriented model evaluation and calibration**
4. **Driving a model with machine learning output**
5. **Model Emulation**

Reichstein et al. (2019)

Hybrid modelling – “Physicizing” Deep learning



Only one perspective: complementary approach is theory-guided ML (e.g. Karpatne et al.)

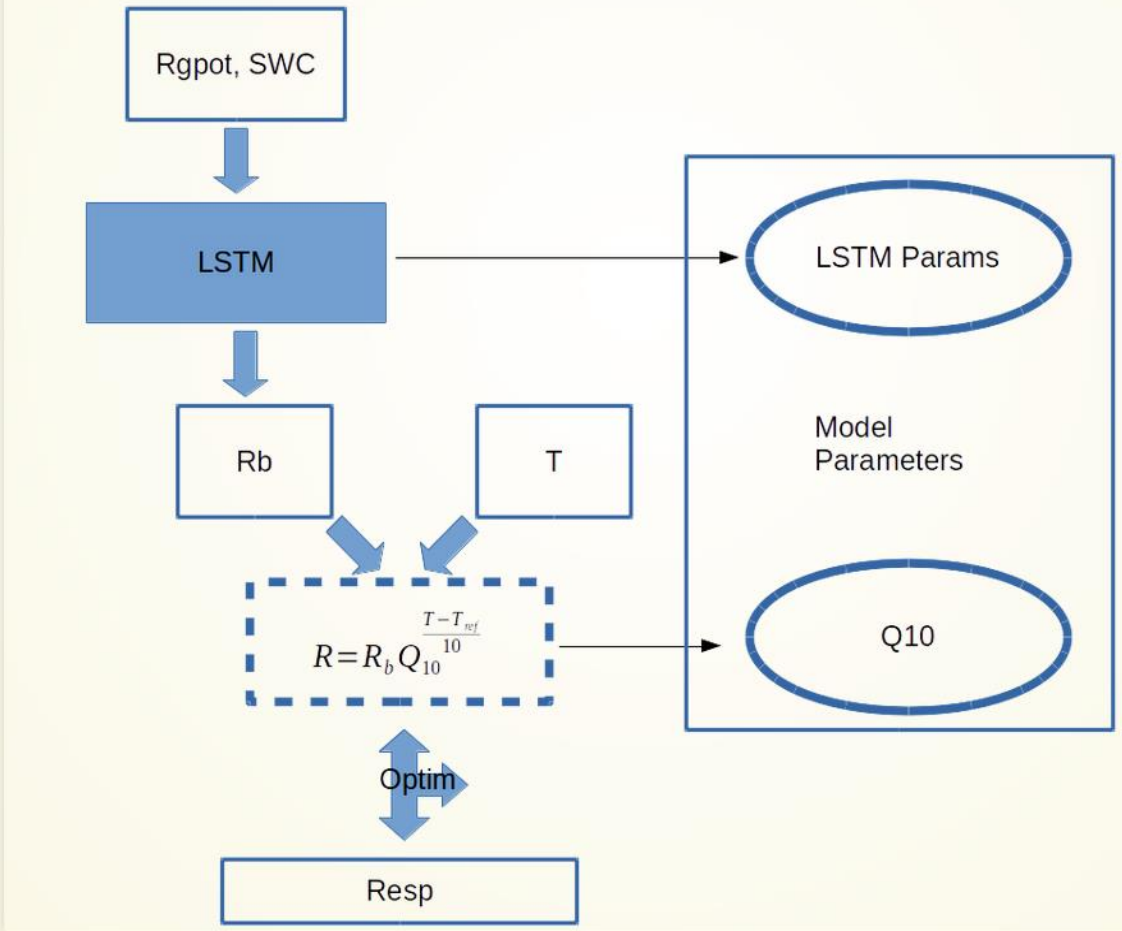
Reichstein et al. (2019)

Hybrid modelling for grasping the carbon cycle: e.g. respiration

$$Resp(t) = R_b(t) * Q_{10}^{\frac{T(t)-T_{ref}}{10}}$$

Biological "base" activity = f(potential factors)

Physico-chemical temperature dependence



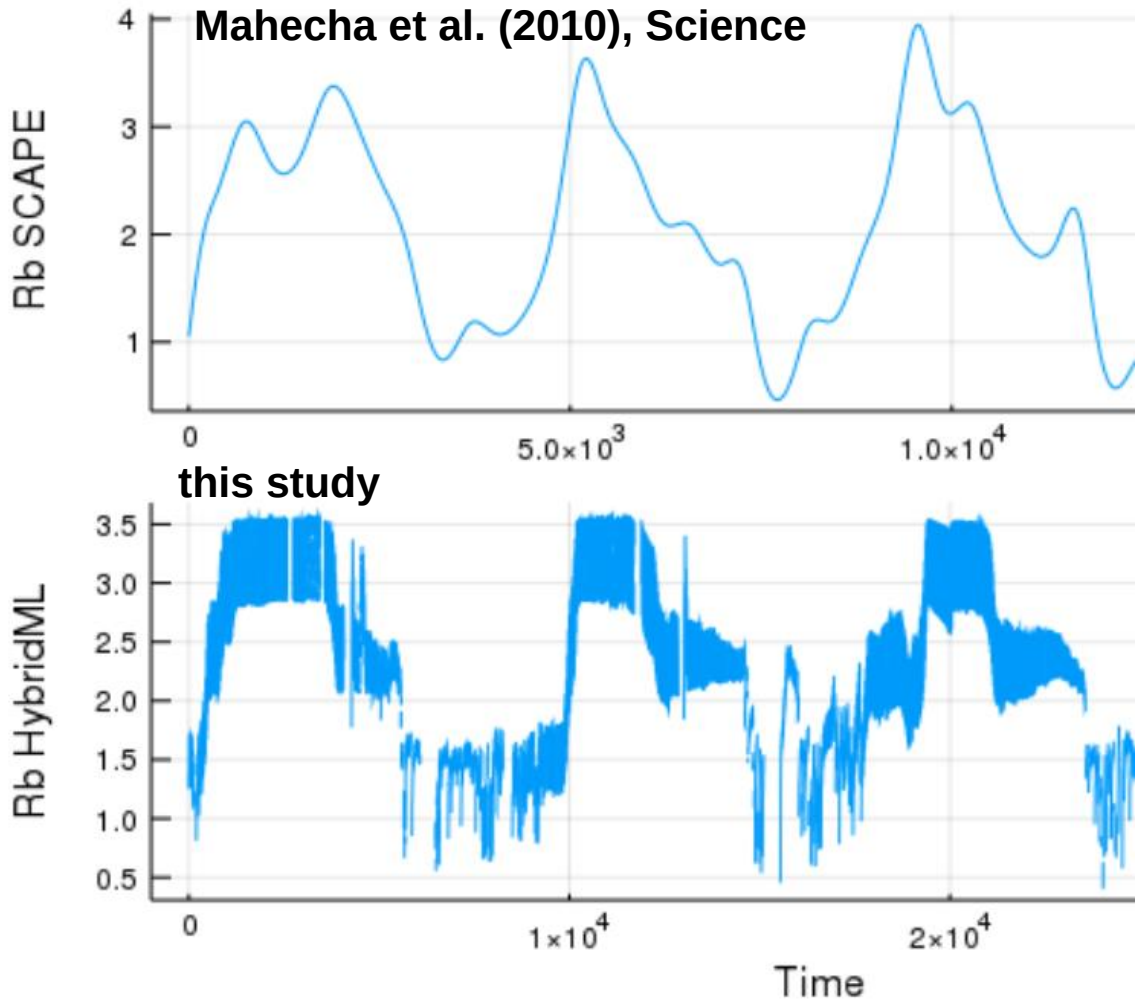
Gans, et al. in prep.

Inferring *multi-scale* variability of processes (seasonal, diurnal)

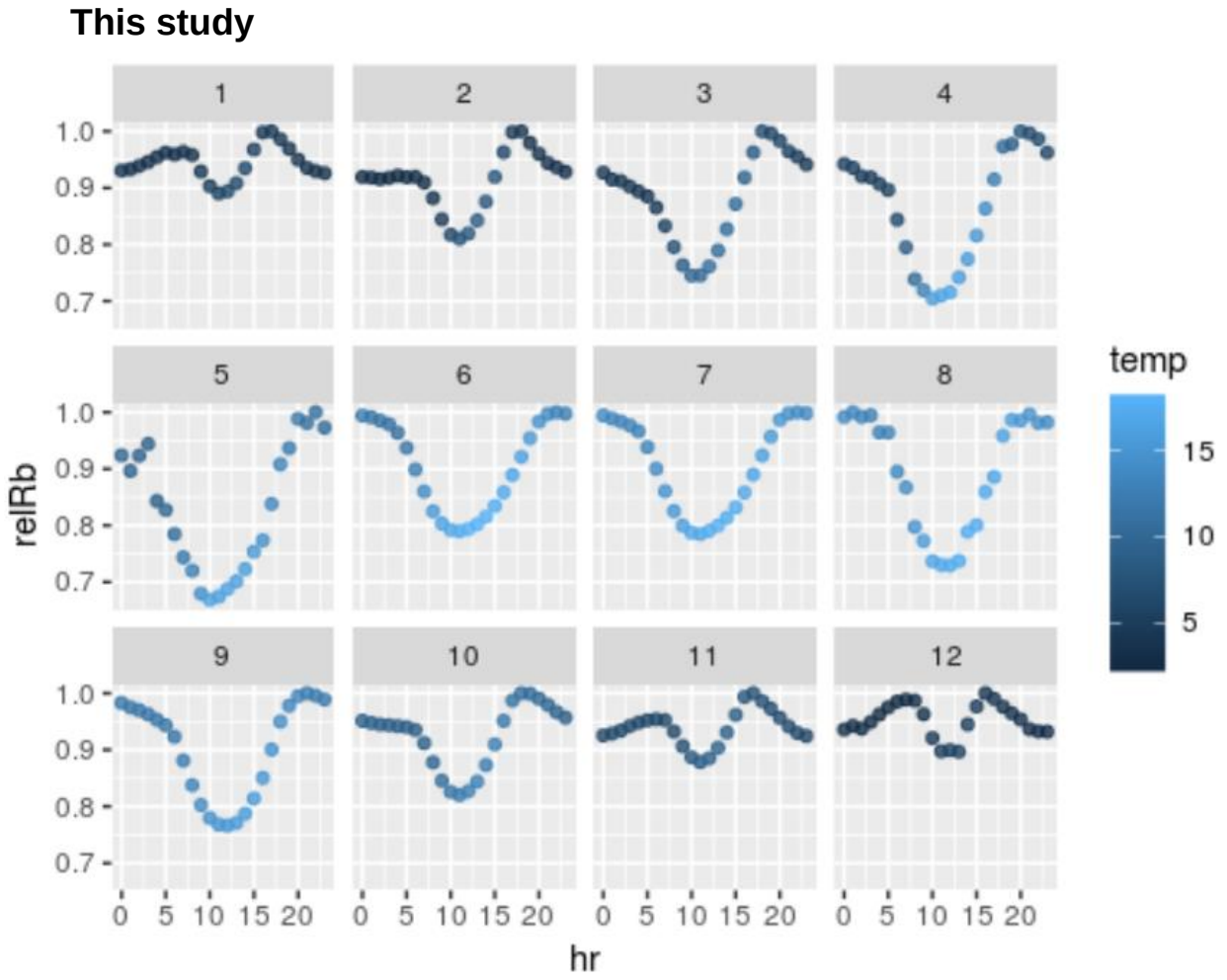
Max Planck Institute for Biogeochemistry

Seasonal scale

[20]:



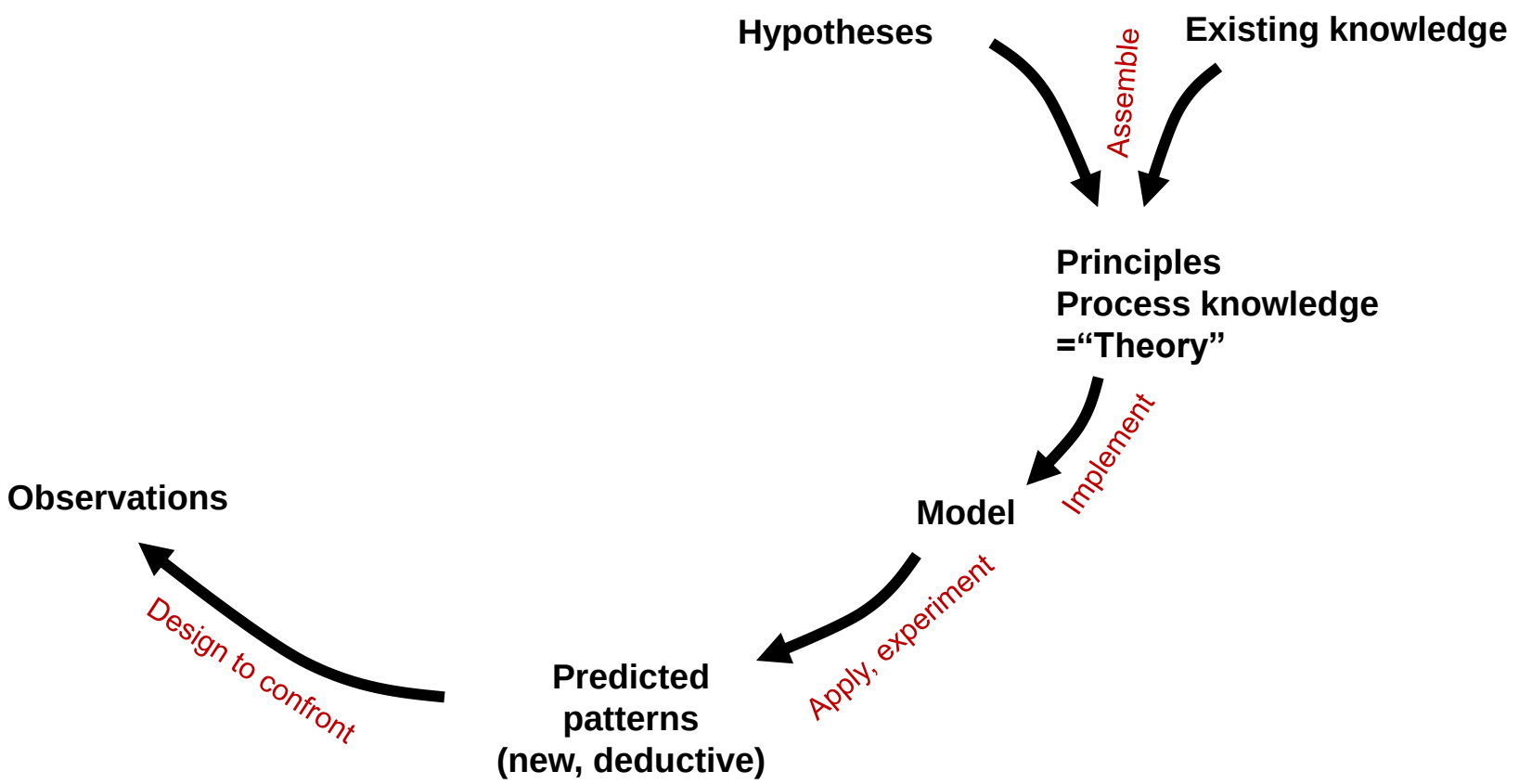
Diurnal scale



Hybrid global hydrological model

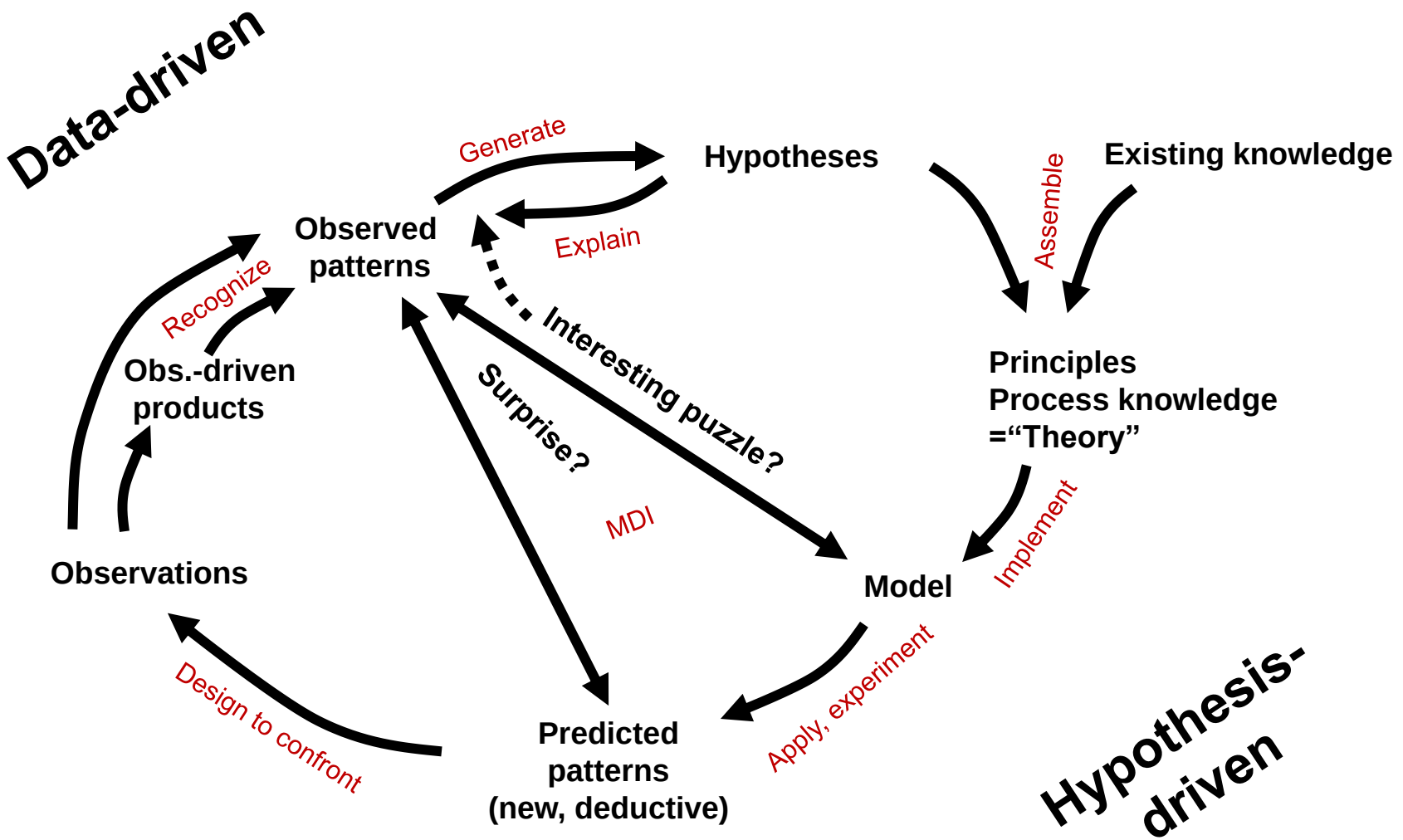
Please contact
Basil Kraft bkraft@bgcbkraft@bgc-jena.mpg.de &
Markus Reichstein mreichstein@bgc-jena.mpg.de
for info on this topic

Wrap-up: Hypothesis-driven / data-driven science



Reichstein et al. (2019)

Wrap-up: Hypothesis-driven / data-driven science



Shameless advertisement…

USMILE ERC Synergy Grant @USMILE_ERC · Oct 14

We are excited to announce the USMILE @ERC_Research #Synergy grant "Understanding and modeling the #Earth System with #MachineLearning". Hybrid #physics-ML modeling and #causality on our Planet! @USMILE_ERC

1 4 19

Through this and other developments:

PhD, PostDoc, group leader opportunities at the interface of machine learning and Earth system science in an int'l environment!

Max Planck Institute for Biogeochemistry

Contact me here or via mreichstein@bgc-jena.mpg.de or

@Reichstein_BGC