ISC – workshop
Deep Learning for Science
Frankfurt, June 2019

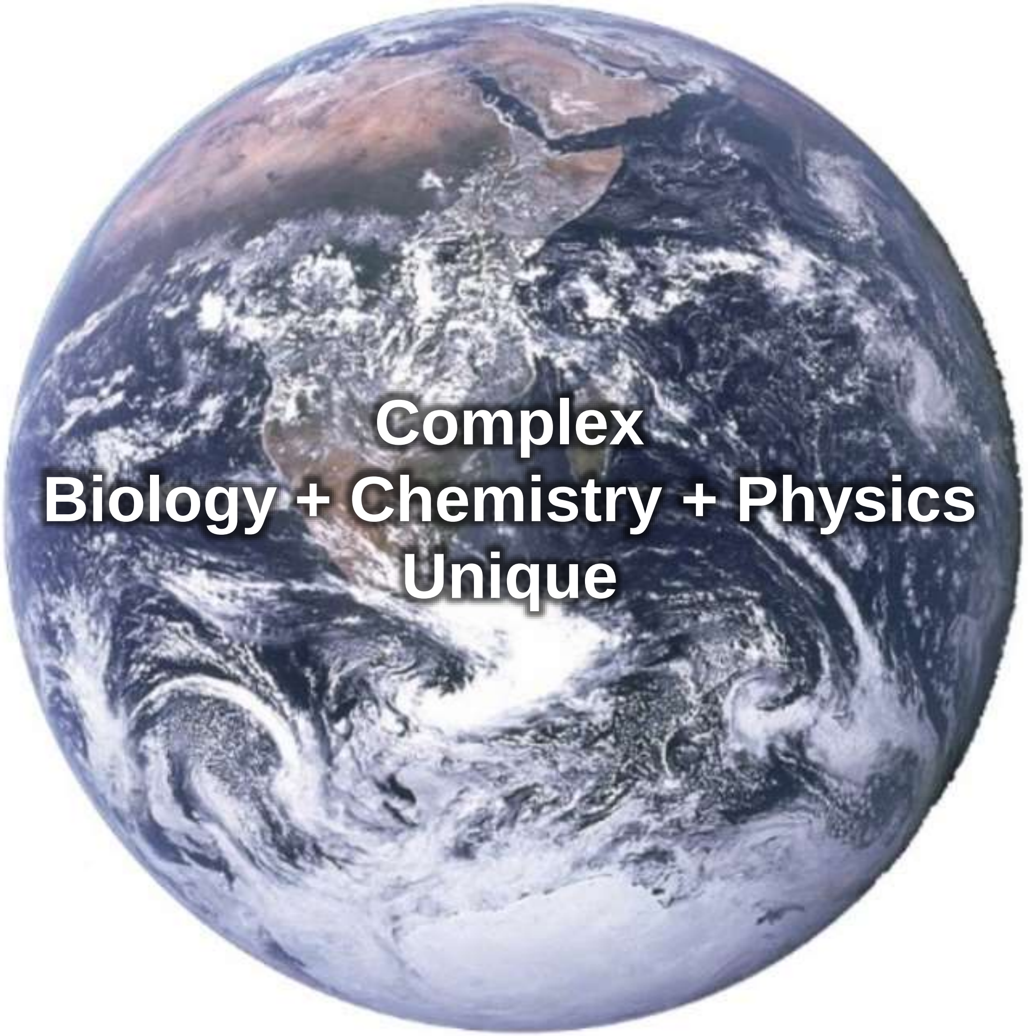
Understanding the Earth system with machine learning

Markus Reichstein

Max-Planck-Institute for Biogeochemistry, Jena

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

The Earth System



GLOBE

Biome

Region

Landscape

ECOSYSTEM

Organism

Organ

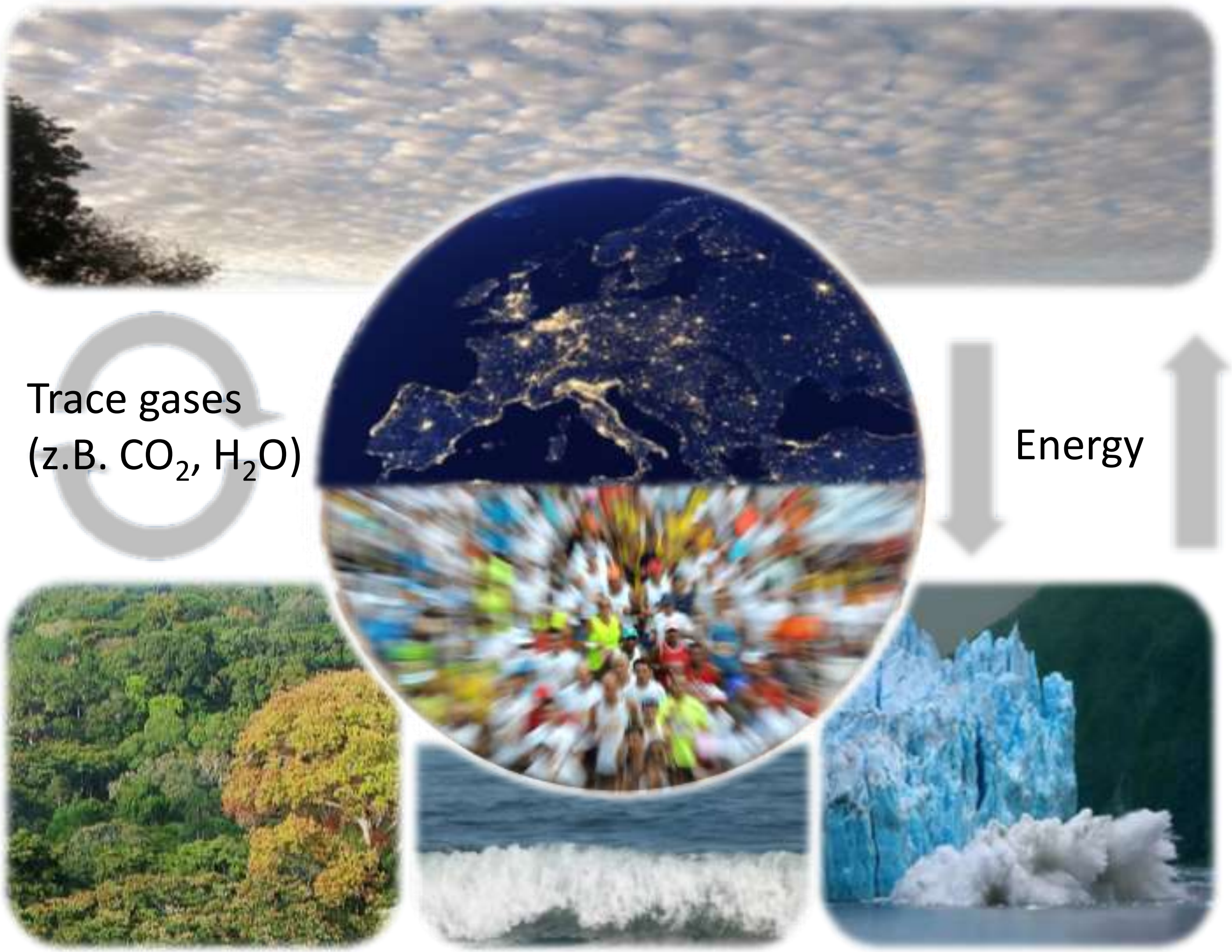
Cell

Molecule

$10^{8}\,m$

$10^{-9}\,m$

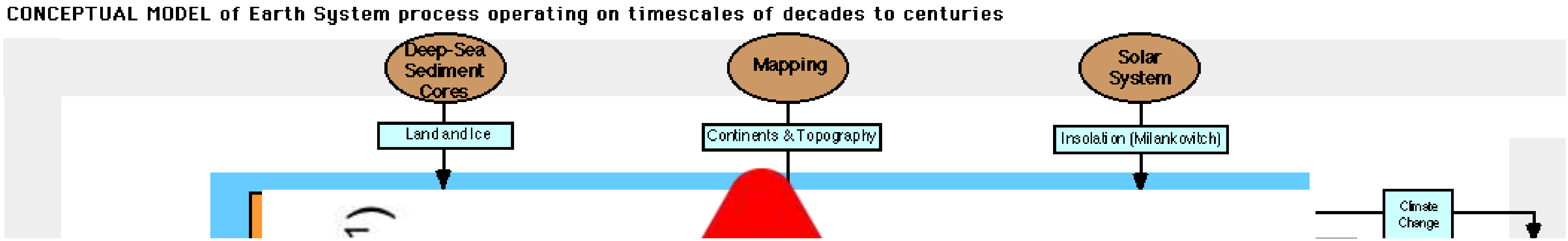
“Spheres” in the Earth System



Earth System Science established in the Max-Planck-Society

"Reductionistic dream": wiring all together

Max Planck Institute for Biogeochemistry

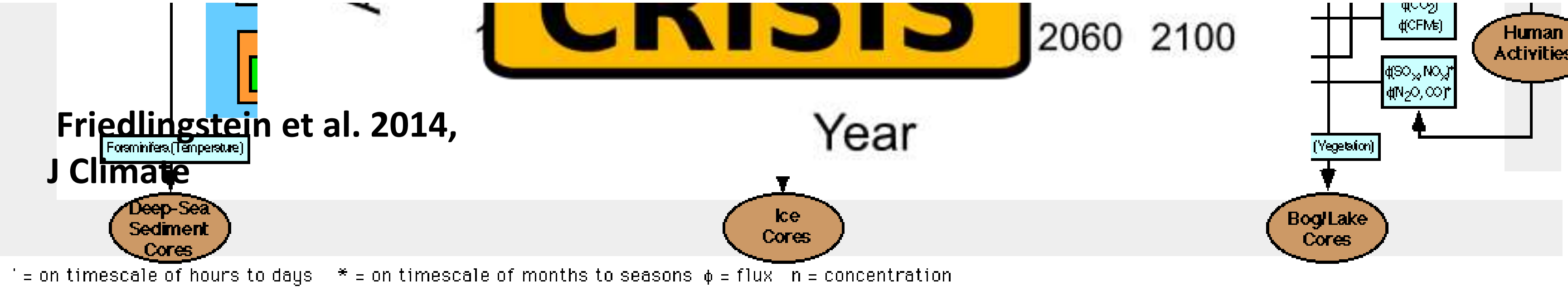


PERSPECTIVE

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

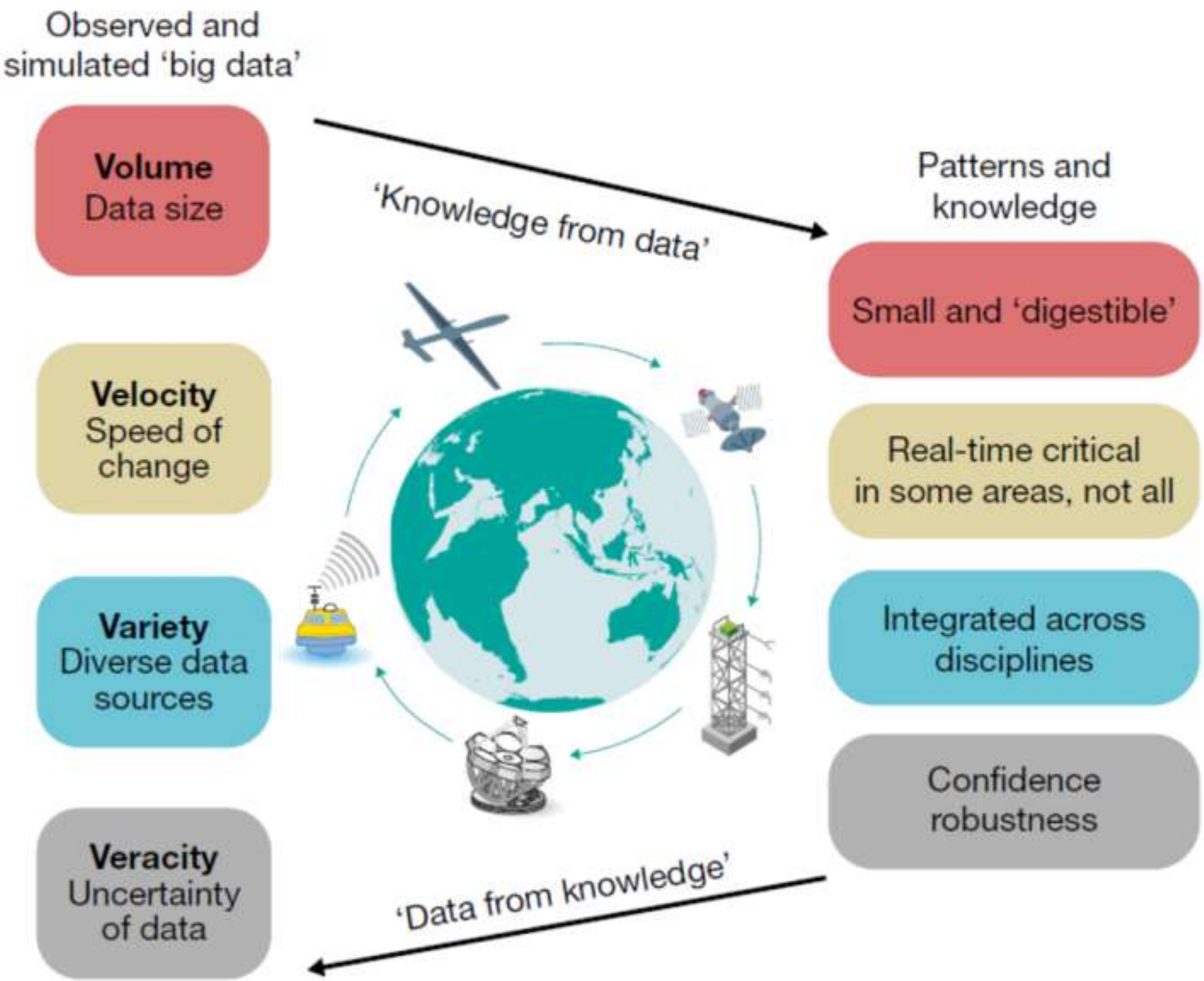
Deep learning and process understanding for data-driven Earth system science

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



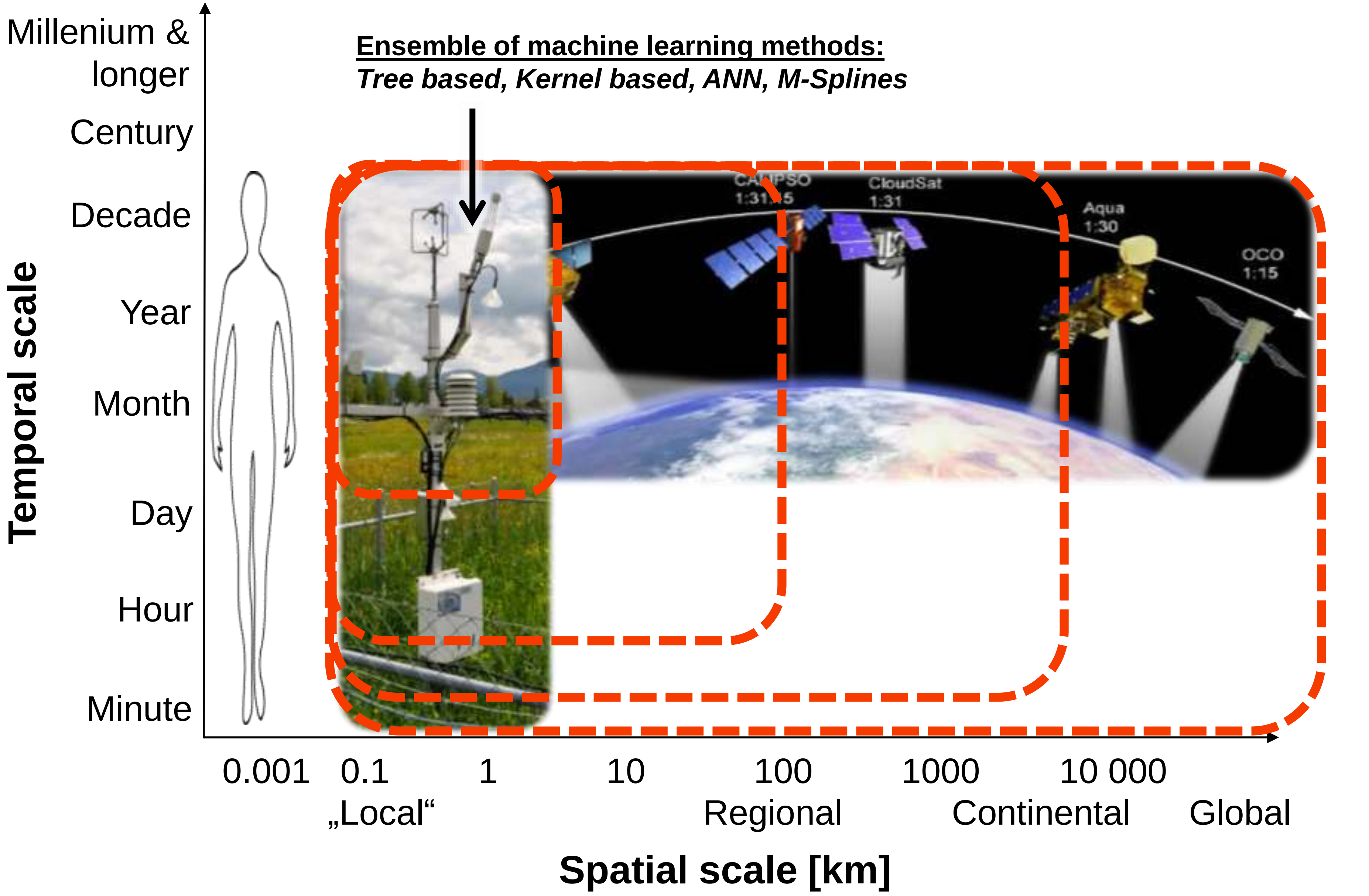
Friedlingstein et al. 2014, J Climate

Data-driven Earth System Science: prototypical



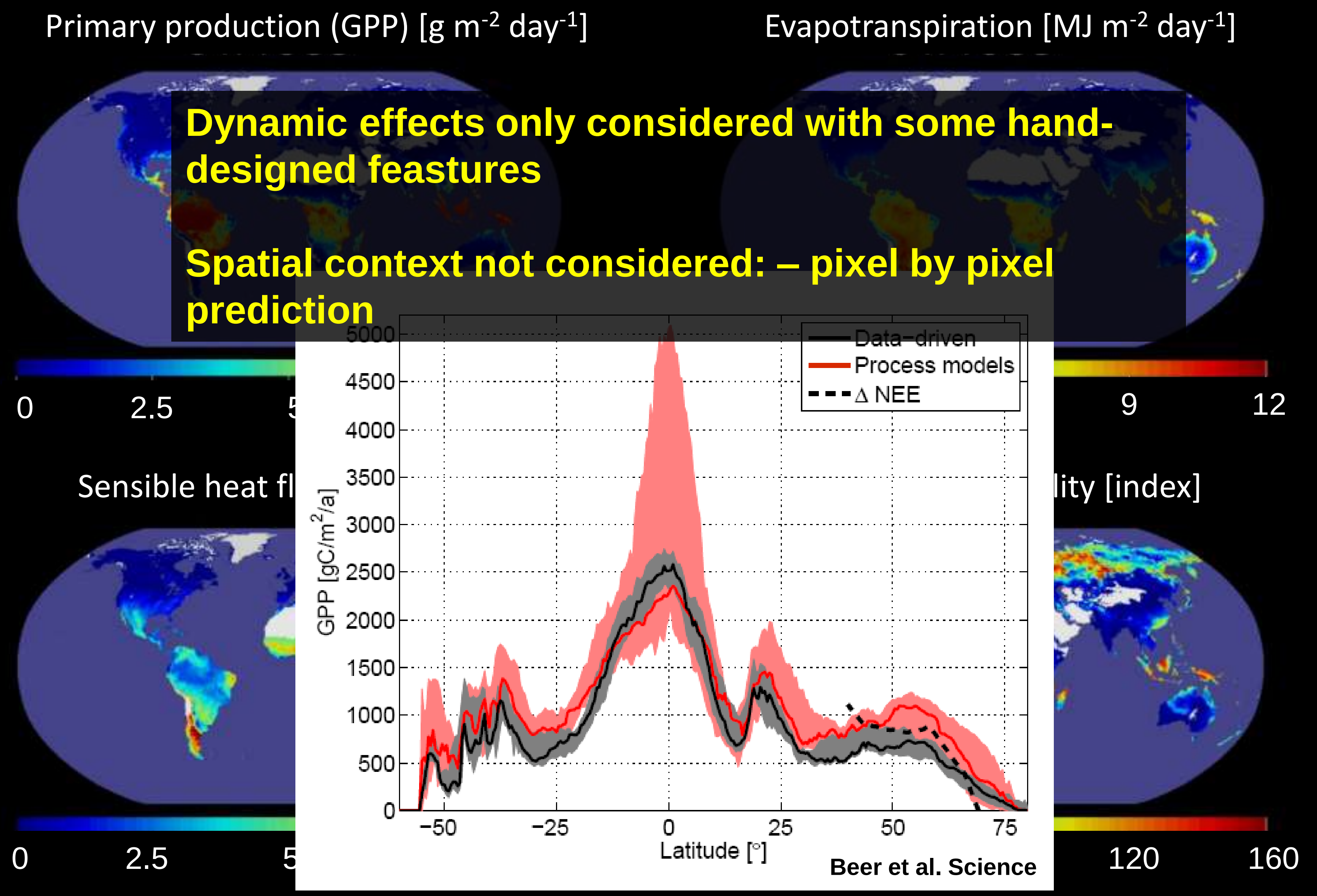
Reichstein et al. (2019)

Scaling from flux-towers to globe



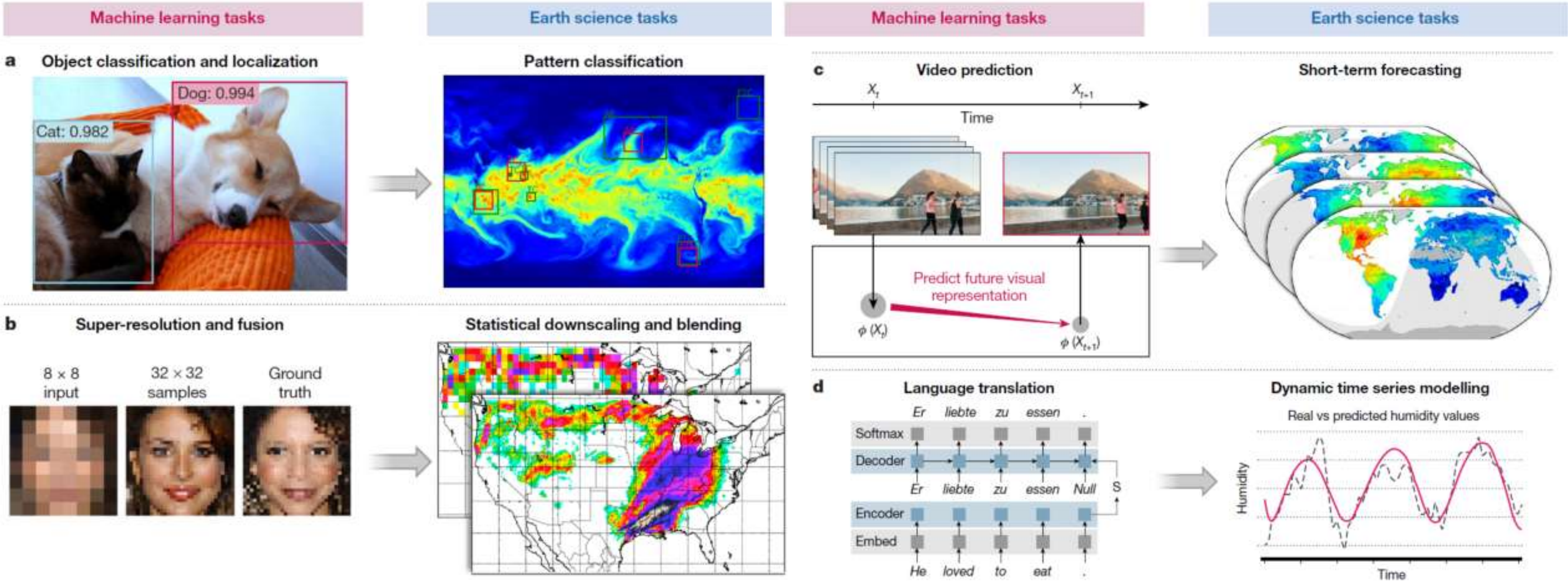
Max Planck Institute for Biogeochemistry

Data-driven view on dynamic Biosphere-Atmosphere Exchange



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

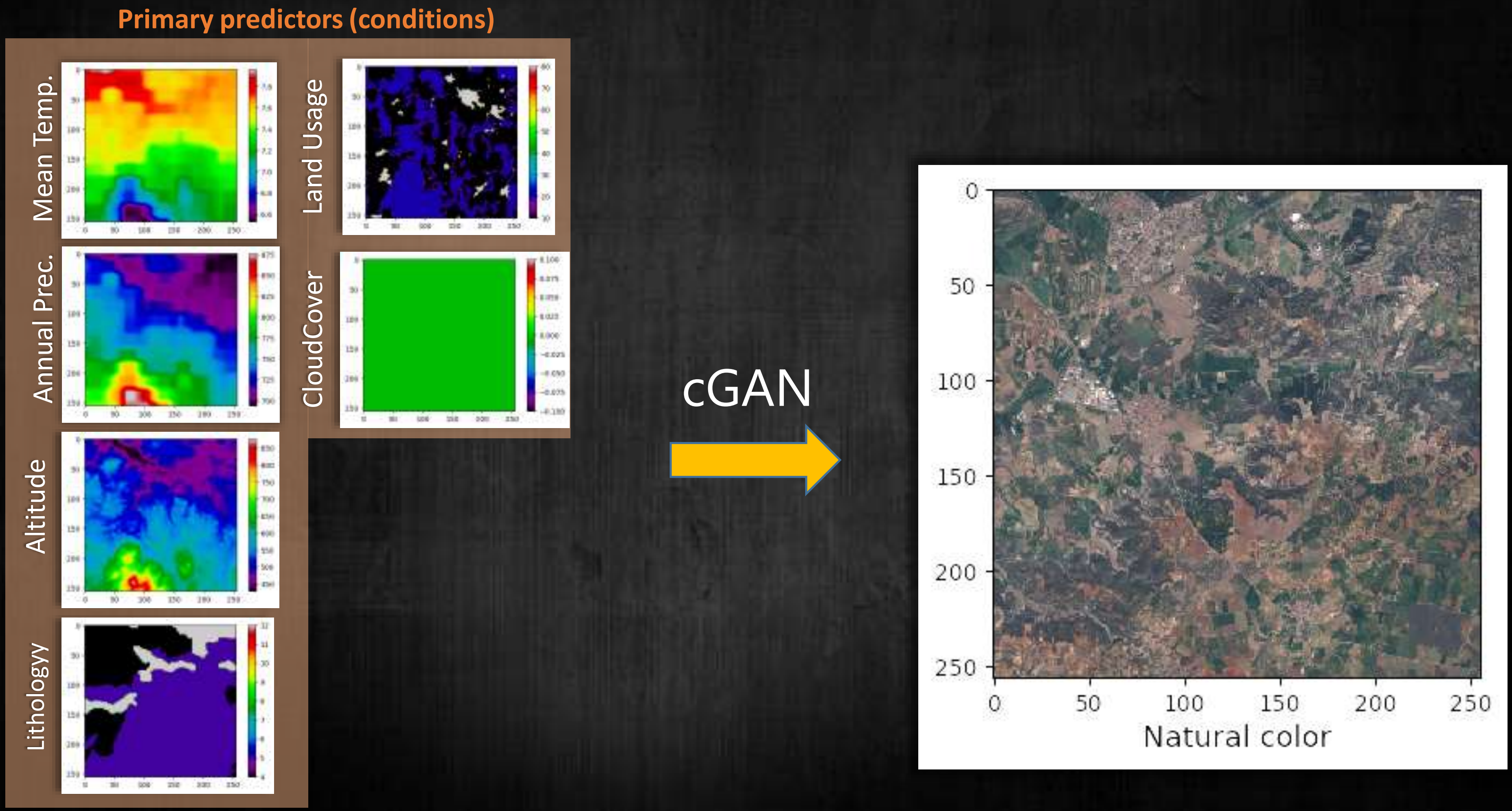
Deep learning for Earth System Science…



Reichstein et al. (2019)

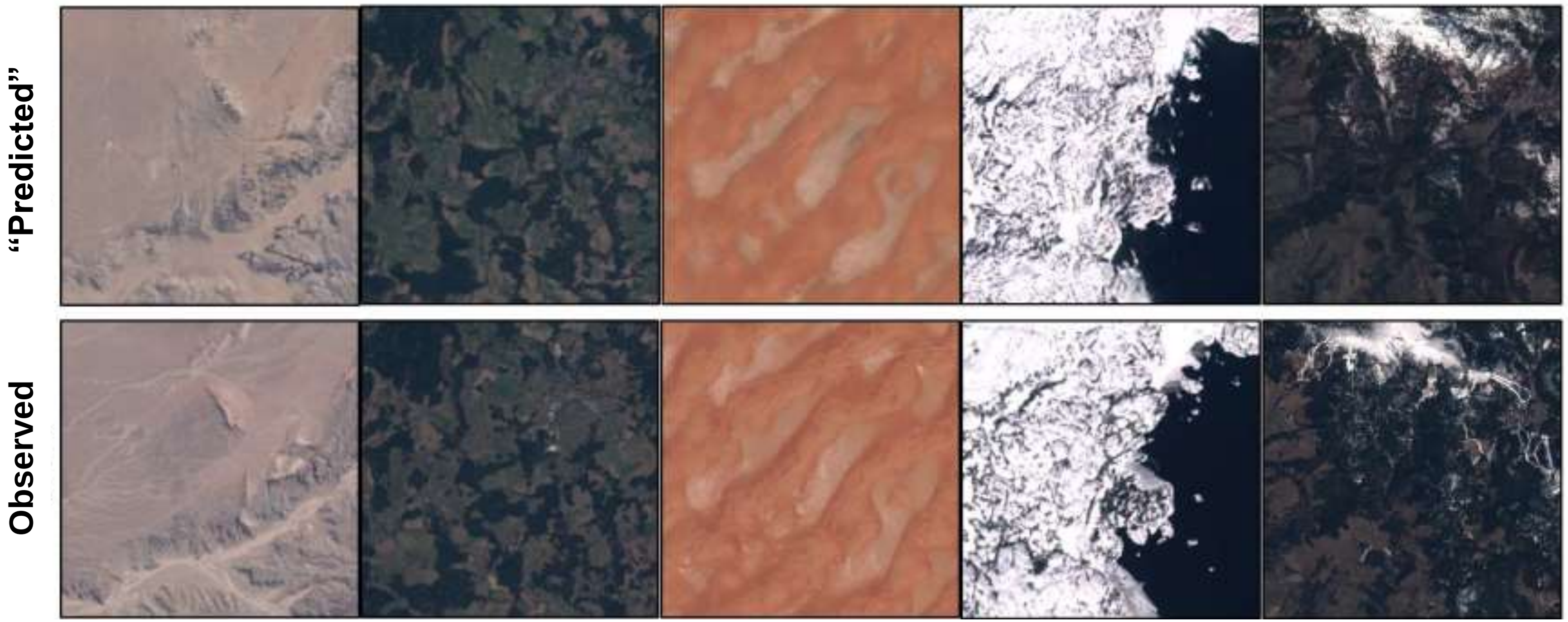
"Predicting" whole landscapes as seen from space

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



Requena et al. 2018

“Predicting” whole landscapes as seen from space



<u>Major limitation:</u> Understanding and physical consistency

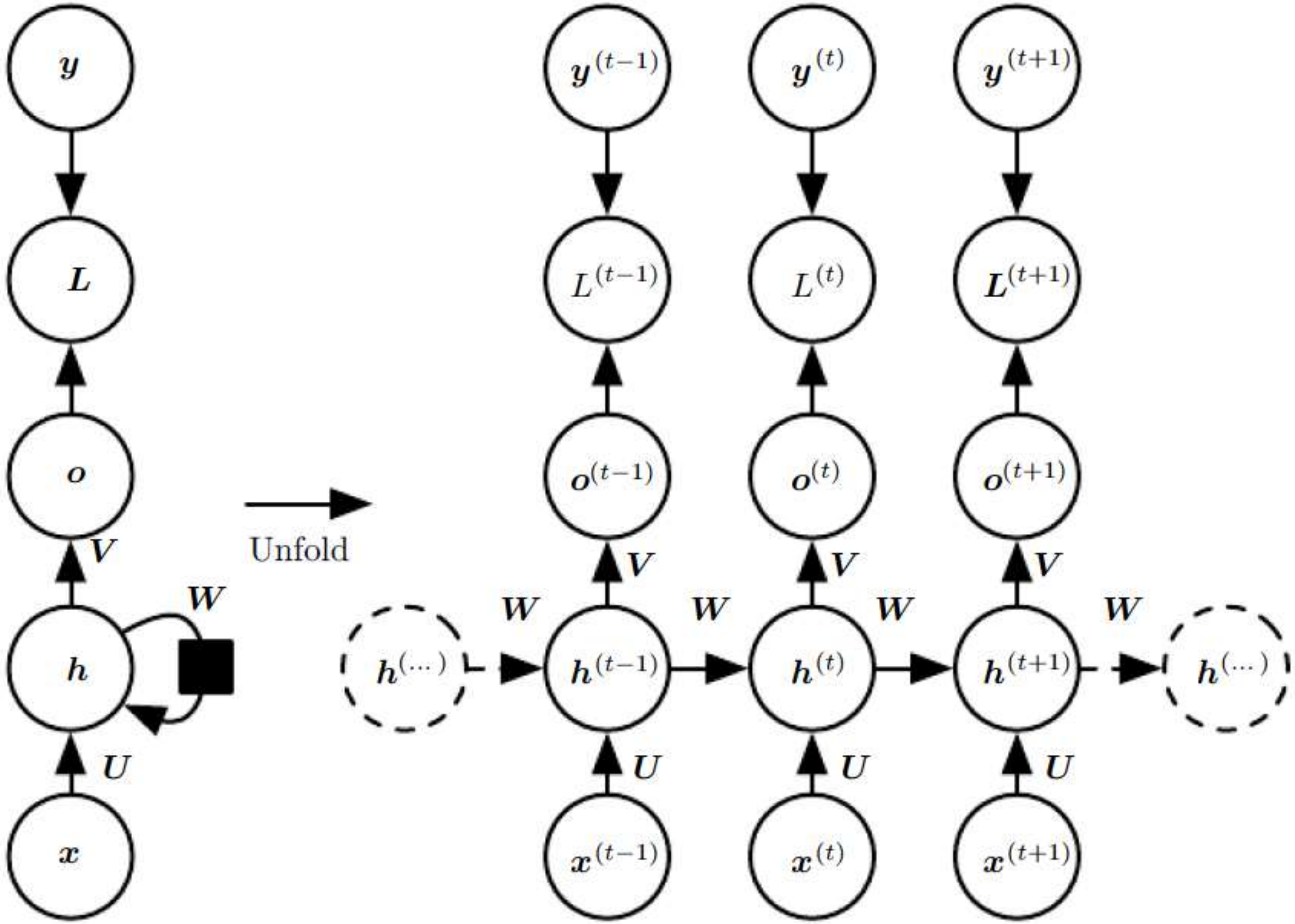
Requena et al. (2018) [conditional GAN]

Throwing away all knowledge?

Describing and detecting dynamic memory effects

- Time-varying properties which depend on past (possibly latent) variables
- Typically described with differential equations or time-discrete analogue equations
- Examples:
  - Vegetation development depends on cumulative temperature over winter-spring
  - Simple water balance: $SM(t) = \int [P(t) - E(t) - D(t)]dt$

Describing & detecting dynamic memory effects

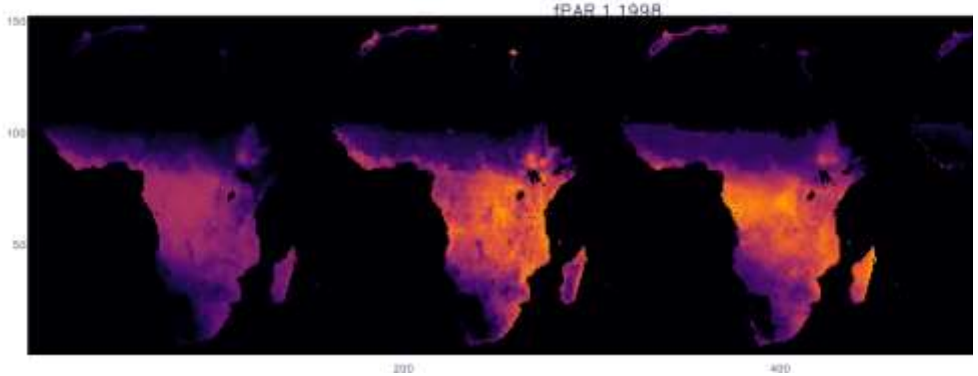


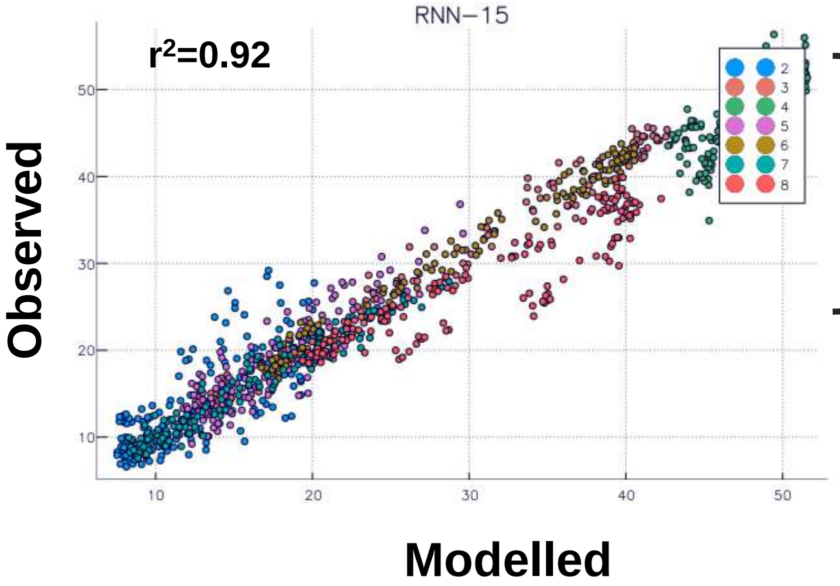
Vegetation state (“leaves” or “greenness”)

Max Planck Institute
for Biogeochemistry

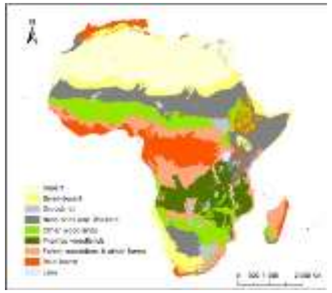
- Target: GIMMS fPAR variability over Africa, 0.5° lat/lon, monthly, 1982-2012
- Two approaches:
  1. Random Forest with standard meteorological predictors **plus lagged and cumulative water variables** (e.g. relative humidity, soil moisture previous months) 🡐 from Feature selection algorithm [Jung et al.]
  2. Recurrent neural network with **only** standard meteorological drivers and vegetation type etc. (trained on 4% of the pixels only)

Rand Forest **Observation** **RNN**

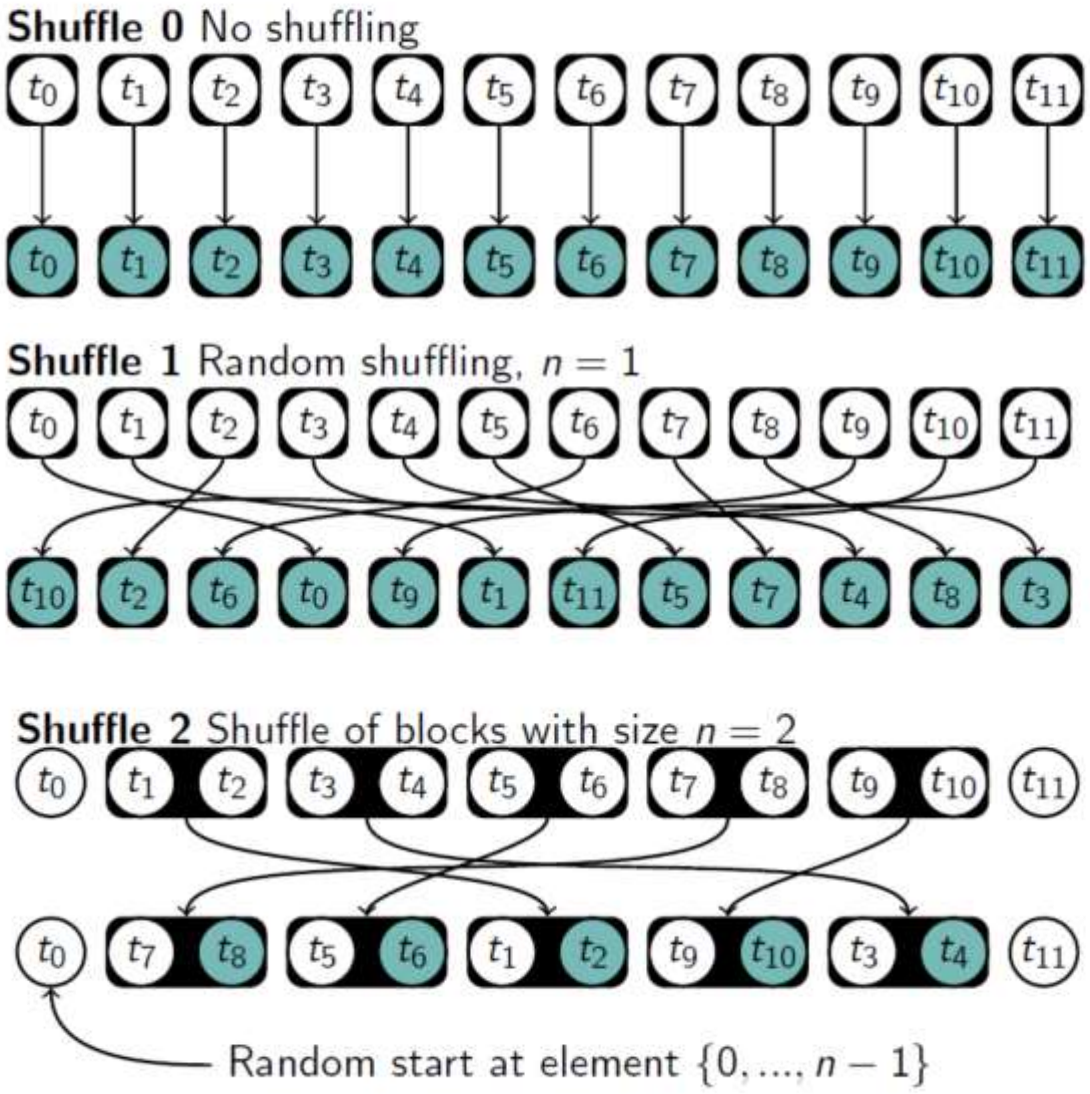




Ecoregions



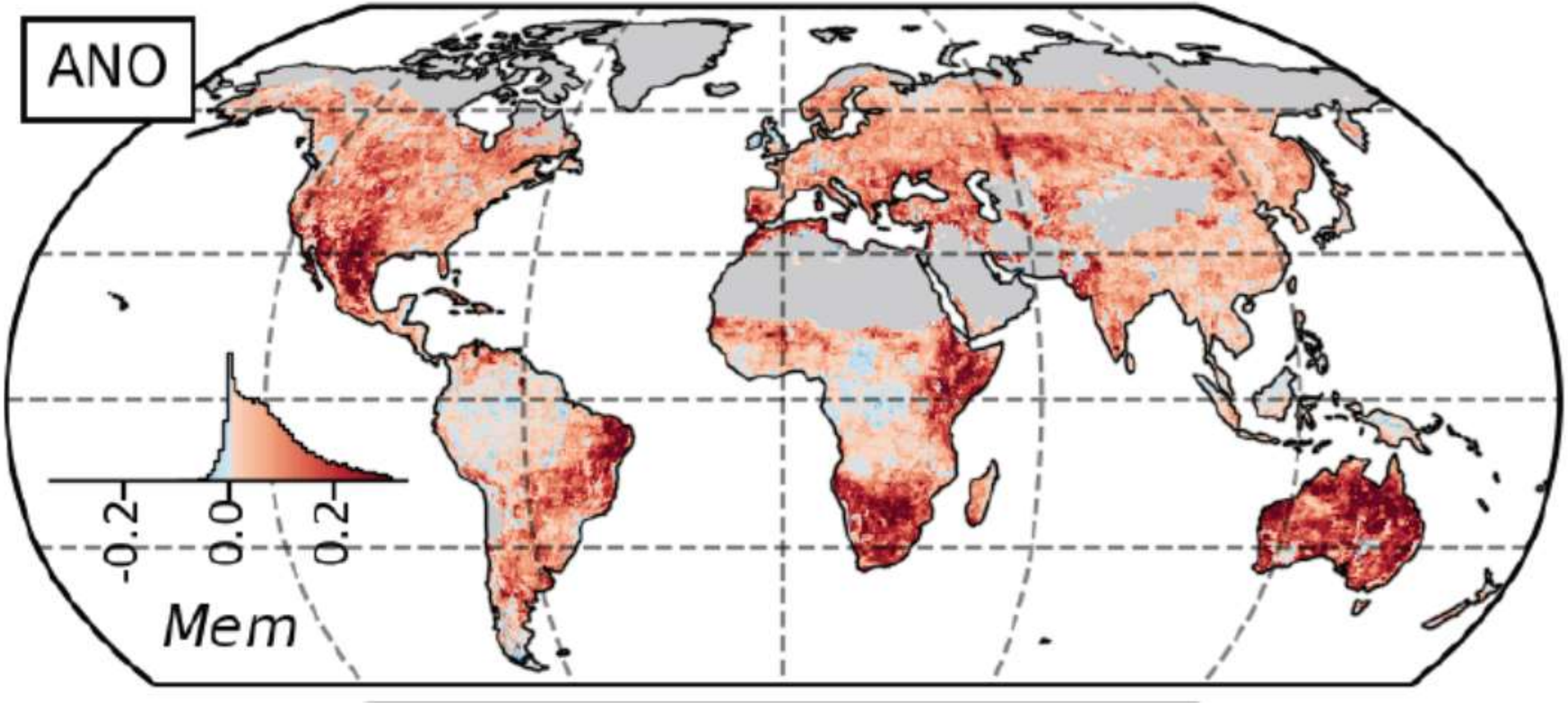
Detecting memory effects by permutation experiments



Memory effect
=
Error with permutation
–
Error without permutation

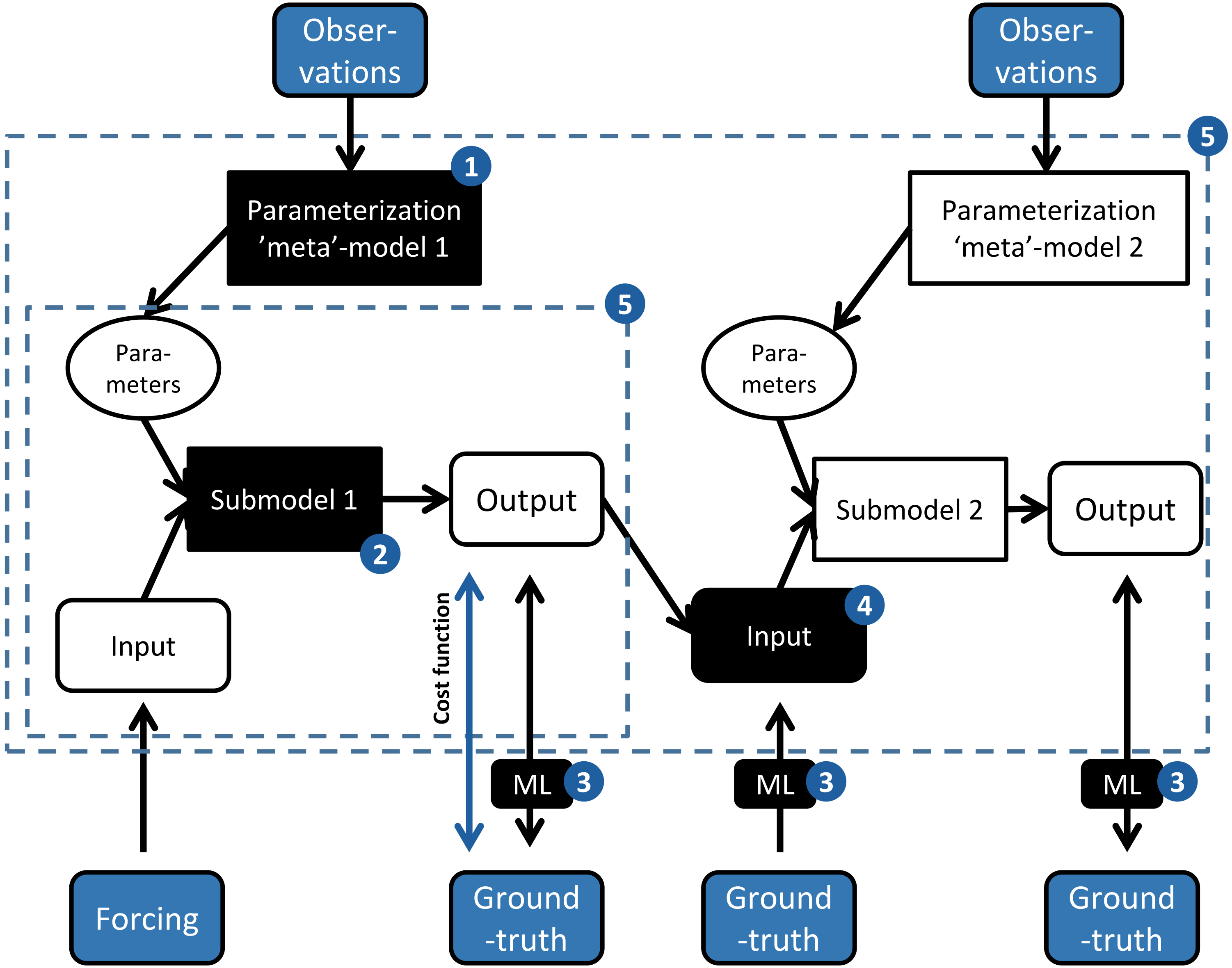
Kraft et al. (2019) in review

Map of memory effects on vegetation



Kraft et al. (2019) in review

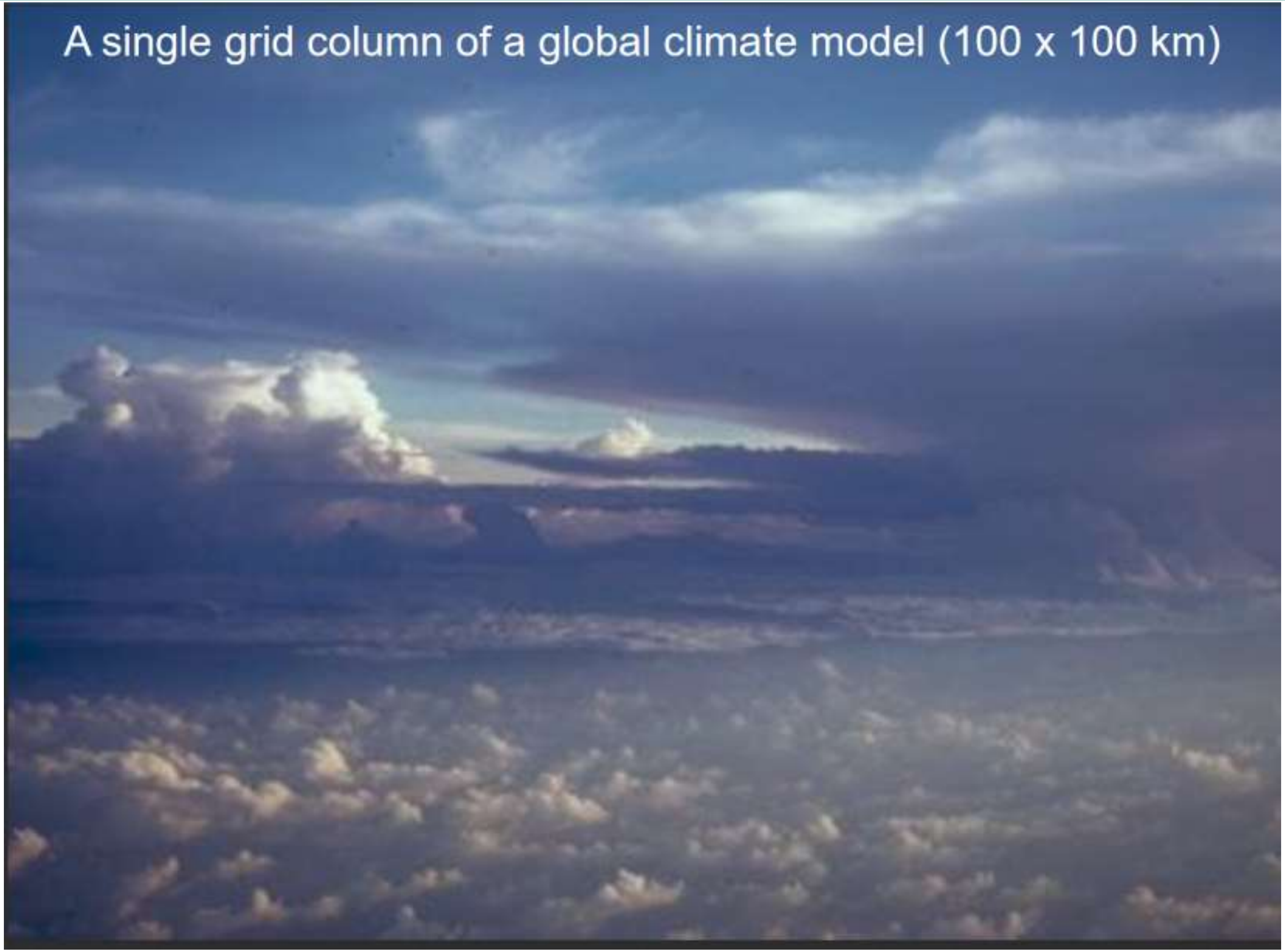
Model-data-machine-learning integration…

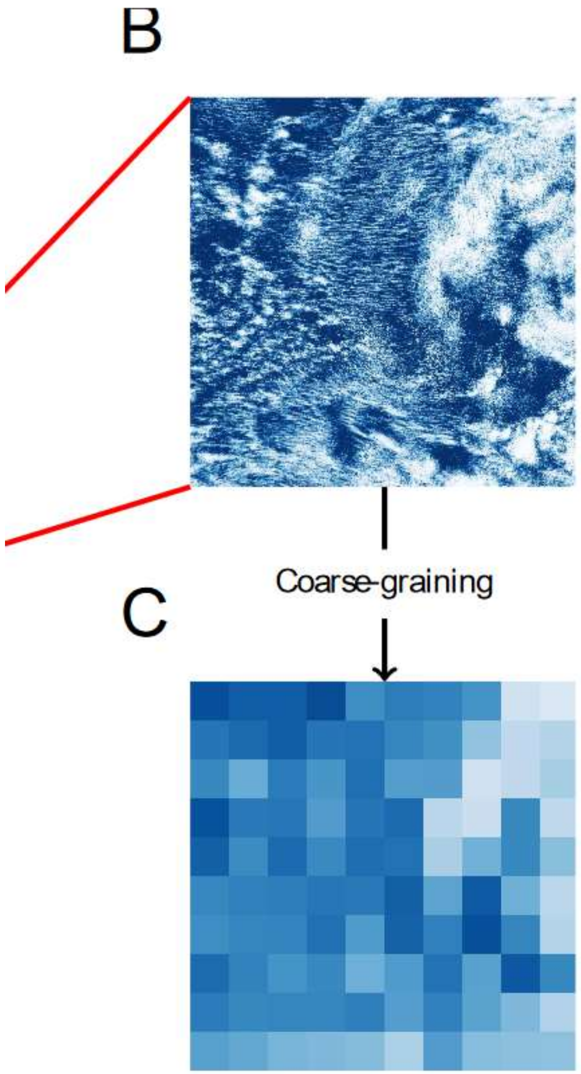


1. **Model parameterization**
2. **Hybrid modelling**
3. **Pattern-oriented model evaluation and calibration**
4. **Driving a model with machine learning output**
5. **Model Emulation**

Reichstein et al. (2019)

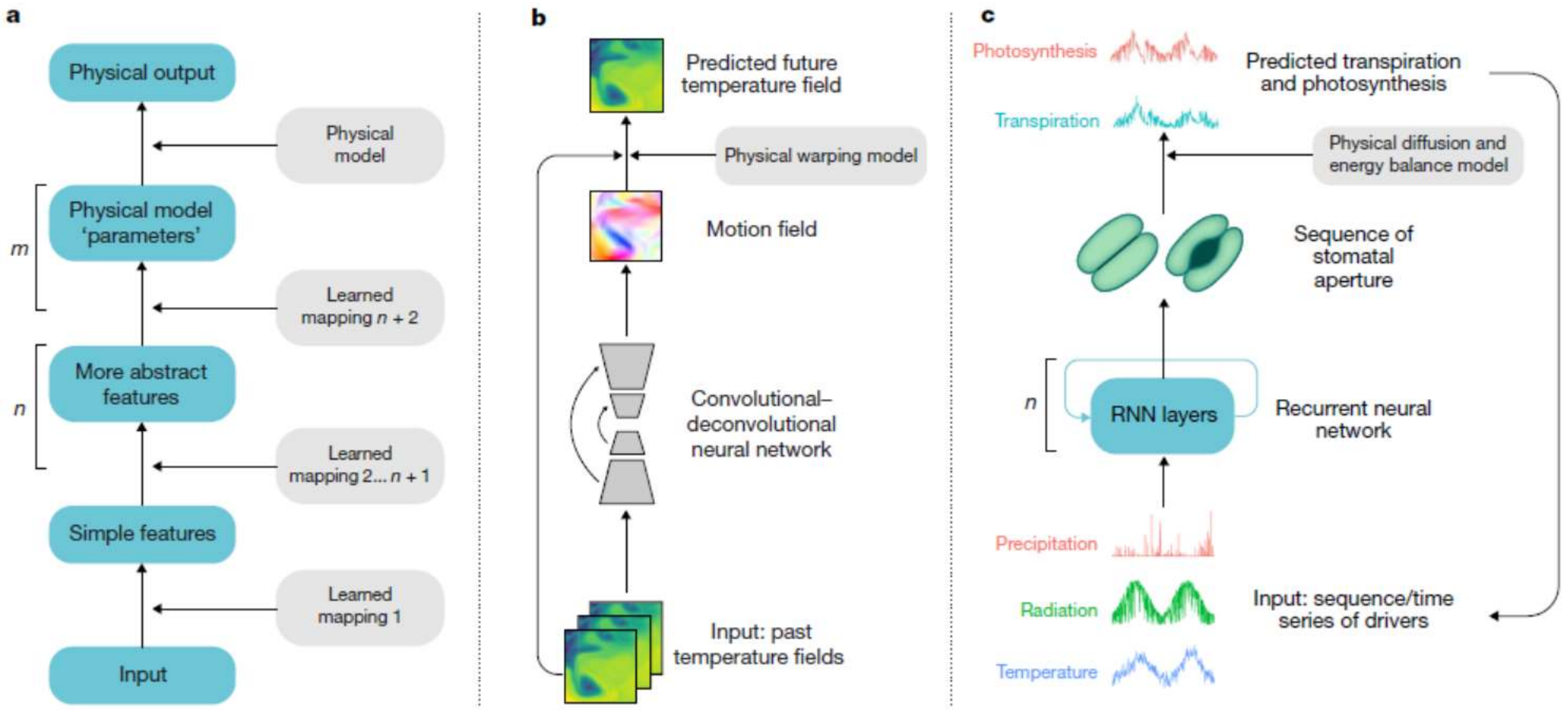
Parameterization, Coarse graining…





Bretherton et al. (2018)

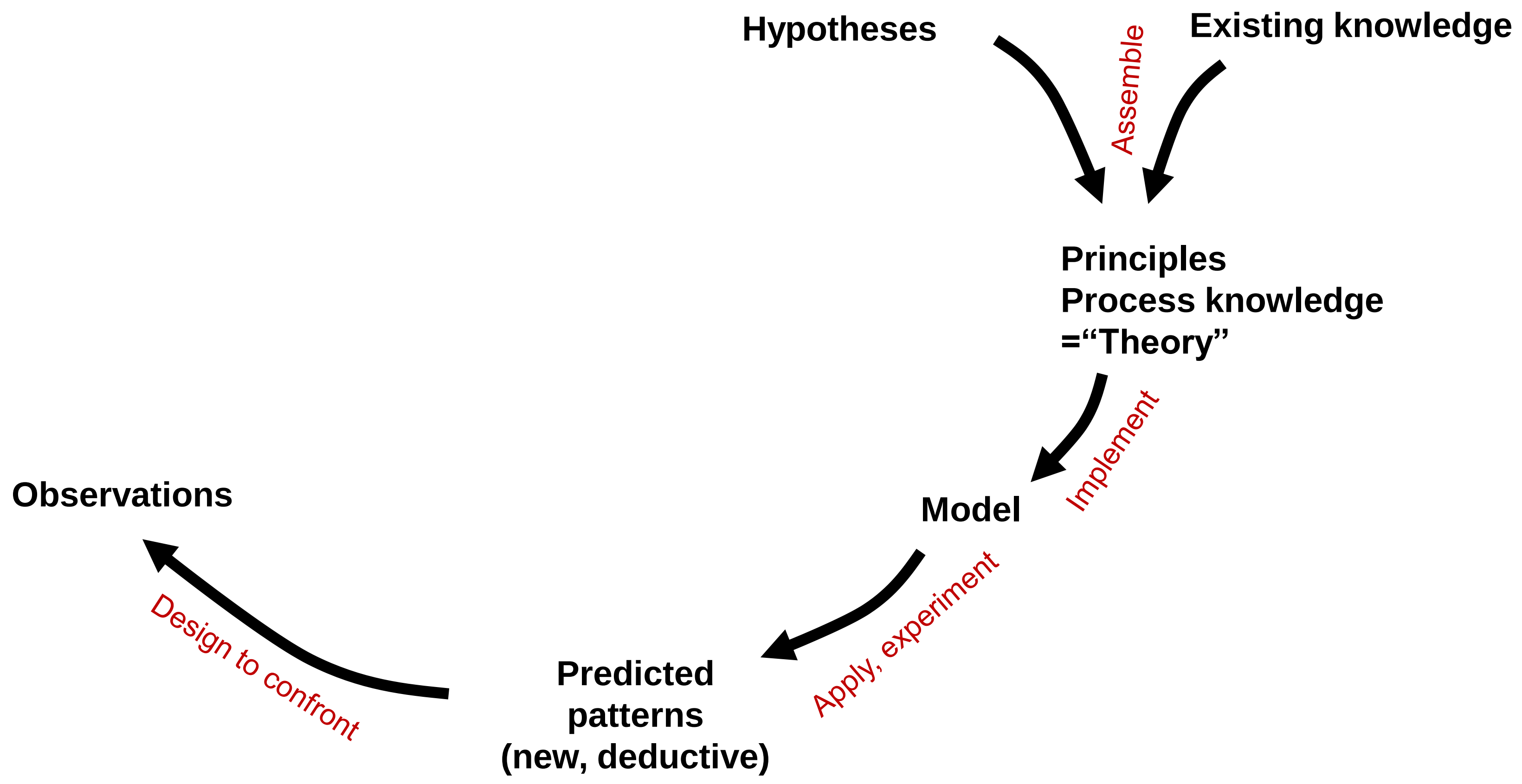
Hybrid modelling – Physicizing Deep learning



Only one perspective: complementary approach

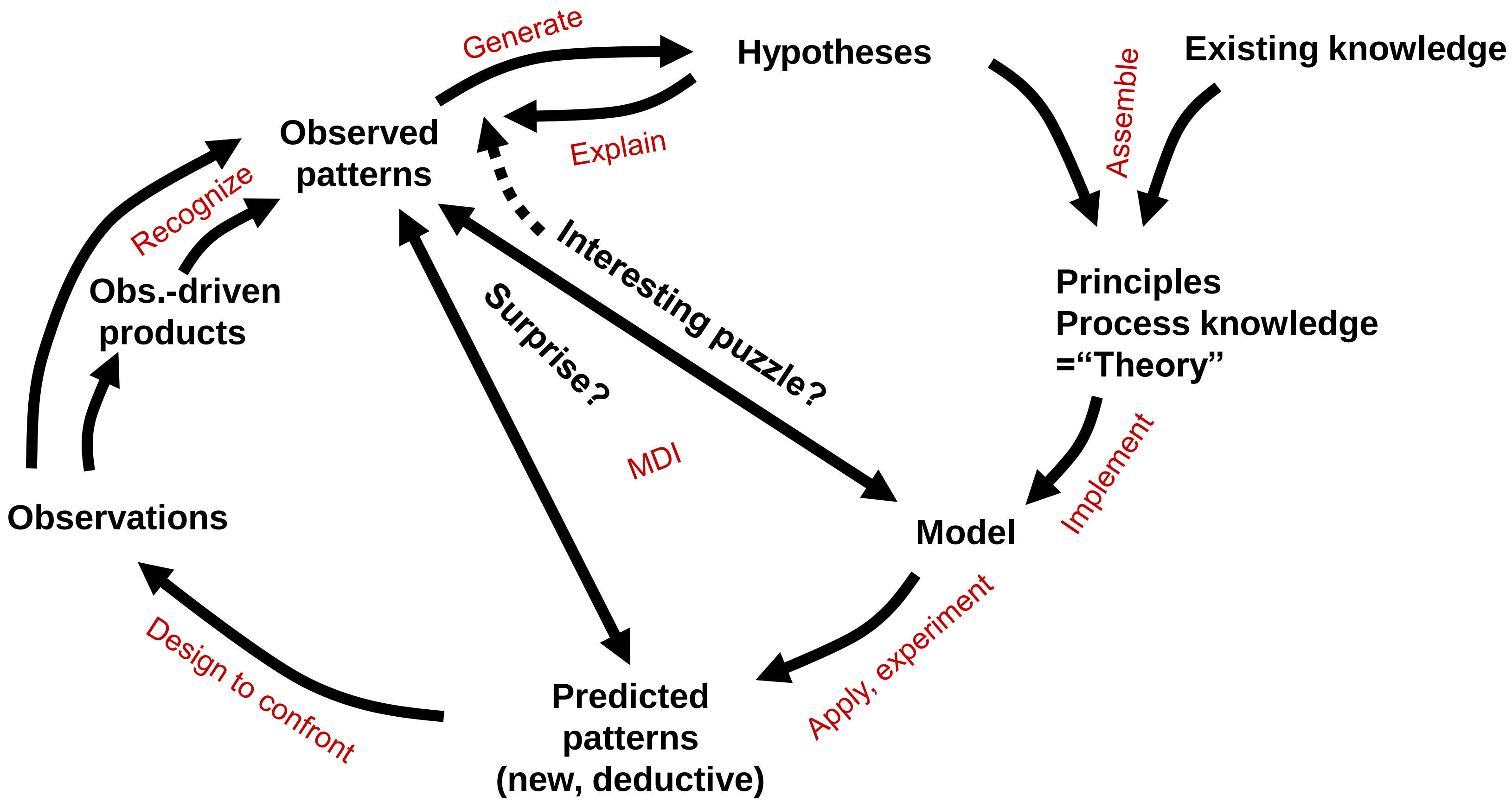
Reichstein et al. (2019)

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



Reichstein et al. (2019)

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



Reichstein et al. (2019)


ELLIS Society
Home News Letter Press Award Members Supporters Board
e l l i s
a
European Laboratory for Learning and Intelligent Systems

www.ellis.eu

Machine Learning for Earth and Climate Sciences

Gustau Camps-Valls (Universitat de València, València, ES)
Markus Reichstein (Max-Planck-Institute for Biogeochemistry, Jena, DE)

- **<u>Goal:</u>** Model and understand the Earth system with Machine Learning and Process Understanding
  - *Spatio-temporal anomaly and extreme events detection, anticipation and attribution*
  - *Data-driven dynamic modelling and forecasting*
  - *Hybrid modeling: linking physics and machine learning models*
  - *Causal inference, Learning and explaining feature representations*
  - *Earth and Climate model emulation, generative modelling and data-model fusion*
  - *Benchmark synthetic and real datasets*

<u>First program workshop:</u> November 2019, tbd

<u>List of Fellows:</u> Joachim Denzler (University of Jena, DE), Veronika Eyring (DLR, DE), Sancho Salcedo (UAH, ES), Kristian Kersting (TU Darmstadt, DE), Miguel Mahecha (MPI-BGC, DE), Jonas Peters (U Aarhus, DK), Rasp (LMU Munich, DE), Jakob Runge (DLR, DE), Dino Sejdinovic (Oxford Univ, UK), Nuno Carvalhais (Univ. Lisboa, PT), Bjorn Stevens (MPI-MET, DE), Devis Tuia (Wageningen Univ, NL), Xiaoxiang Zhu (DLR, DE), Peter van Leeuwen (Reading University, UK)

ellis
European Laboratory for Learning and Intelligent Systems

Literature

- Besnard, S., et al. (2019), Memory effects of climate and vegetation affecting net ecosystem CO2 fluxes in global forests, *Plos One, 14(2), doi:10.1371/journal.pone.0211510.*
- Reichstein, M., G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and Prabhat (2019), Deep learning and process understanding for data-driven Earth System Science, *Nature, 566, 195-204.*
- Requena-Mesa, C., M. Reichstein, M. Mahecha, B. Kraft, and J. Denzler (2018), Predicting Landscapes as Seen from Space from Environmental Conditions, paper presented at IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE.
- Beer, C., et al. (2010), Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate, *Science, 329(5993), 834 - 838, doi:10.1126/science.1184984.*
- Bodesheim, P., M. Jung, F. Gans, M. D. Mahecha, and M. Reichstein (2018), Upscaled diurnal cycles of land-atmosphere fluxes: a new global half-hourly data product, *Earth System Science Data, 10(3), 1327-1365.*
- Jung, M., et al. (2011), Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, *Journal of Geophysical Research - Biogeosciences, 116, G00j07, doi:10.1029/2010jg001566.*