Learning evapotranspiration dataset corrections from a water cycle closure supervision

NeurIPS 2022 Workshop Tackling Climate Change with Machine Learning

Tristan Hascoet[1], Victor Pellet[2], Filipe Aires[2]

[1] Kobe University

[2] LERMA, Observatoire de Paris, PSL

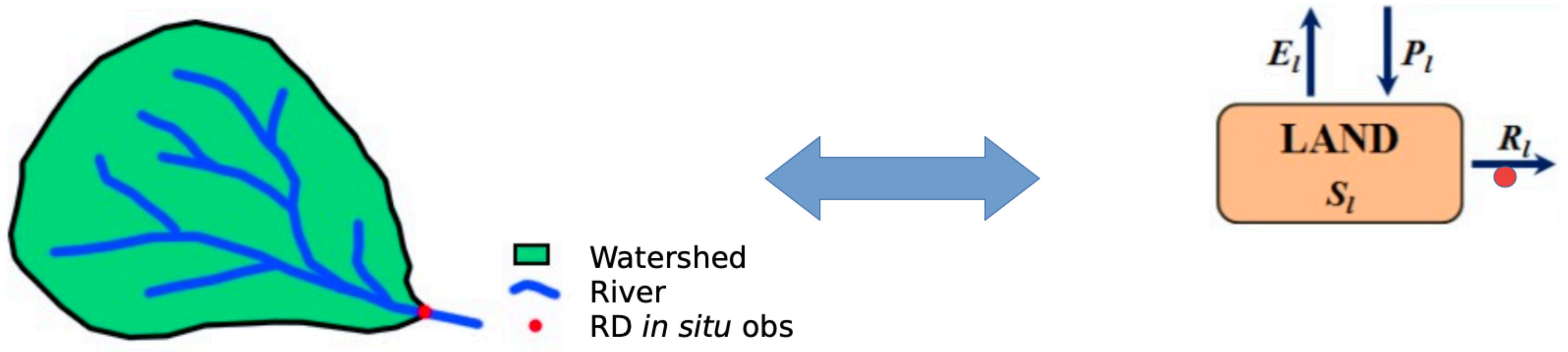
Background: The Water Cycle

Water Resources in a Changing Climate

- In the context of Climate Change (CC), the demand on water resources is increasing as both flood and drought related damages increase
- There is a need to better understand the dynamics of water surface distribution to improve both climate modeling and water resource management operations.
- Evapotranspiration (ET) is one of the most uncertain components of the global Water Cycle (WC).
- Improving global ET estimates is thus needed to forecast the consequences of climate change on the future of global water resource distribution.

The Water Cycle

- The Water Cycle is modeled at the catchment scale as River Discharge (RD) measurements are only available from point-wise in-situ observations
- At the catchment-scale, the water cycle is modeled through the four components of Precipitation (P), ET, R and water Storage Differential dS. These components should sum to 0 in the absence of observational uncertainties.



$$P - ET - R + dS = 0$$

The Water Cycle

- The Optimal Interpolation (OI) framework has been shown to efficiently leverage the WC closure constraint to correct each WC component at the catchment scale.

- To improve global ET datasets, we aim to generalize the corrections on ET estimates brought at the catchment-scale by the OI framework to the pixel scale.

Global ET Errors

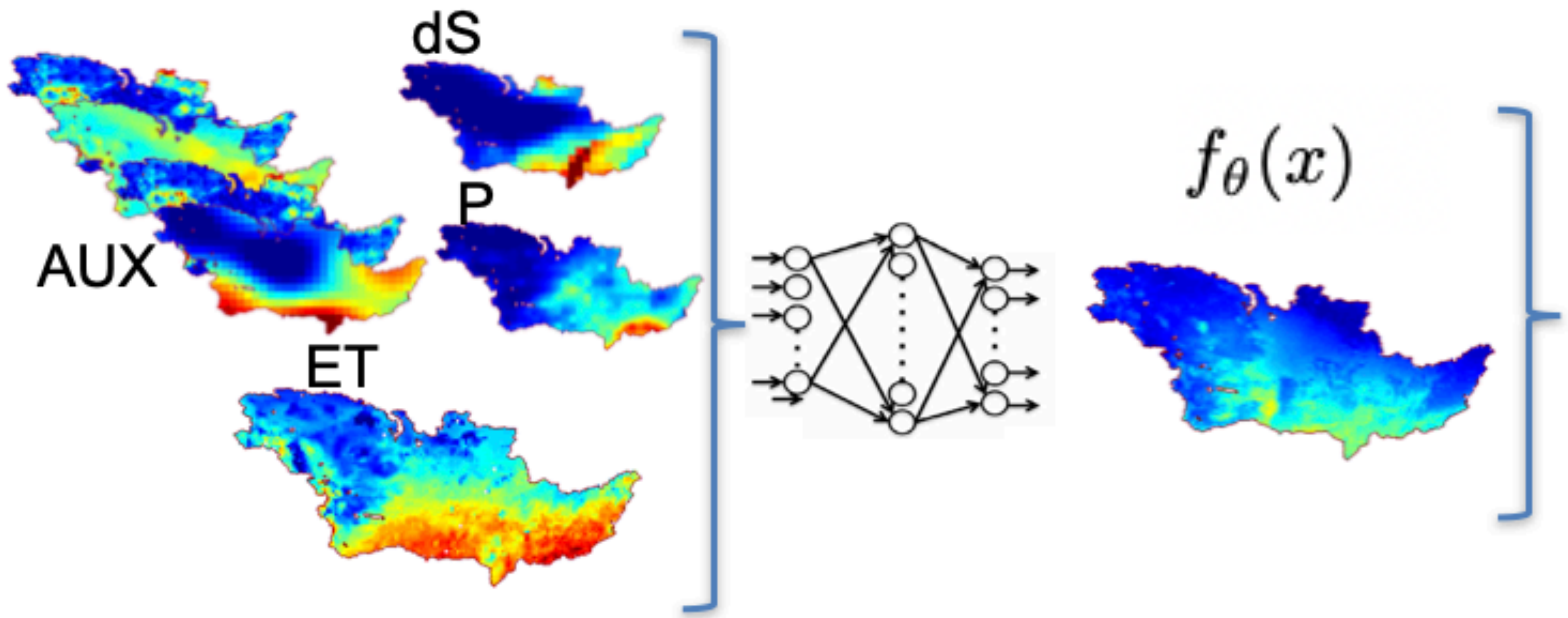
- Global ET datasets do not account for secondary ET: evaporations from wetlands, waterbodies and irrigated areas
- Global ET datasets have been shown to suffer from systematic errors in semiarid regimes and tropical forests, as well as imperfect representations of water stress and canopy interception.

- We thus propose to regress ET corrections (provided by the OI) from both WC components and climatic indices reflecting the above expected sources of errors

Proposed Method

A problem of two scales

- We aim to regress catchment-scale corrections GT from pixel-level inputs.
- We thus define a pixel-wise correction model f, and aggregate its outputs over a catchment c, for each time stamp t:



$$F_\theta(c,t) = \sum_{x \in X(c,t)} f_\theta(x)$$

X

A problem of two scales

- We then regress the aggregated output F to the ground-truth correction y over our dataset

$$e_\theta(c,t) = F_\theta(c,t) - y(c,t)$$
$$\mathcal{L}(\theta) = \frac{1}{T}\sum_{c\in C}\sum_{t\in T} e_\theta(c,t)$$
$$\theta* = min_{\theta\in\Theta}\mathcal{L}(\theta)$$

- We parameterized the pixel-wise ET correction function f with a 4 layers MLP of width 512, which we trained using the Adam optimizer

Experiments and Results

Dataset

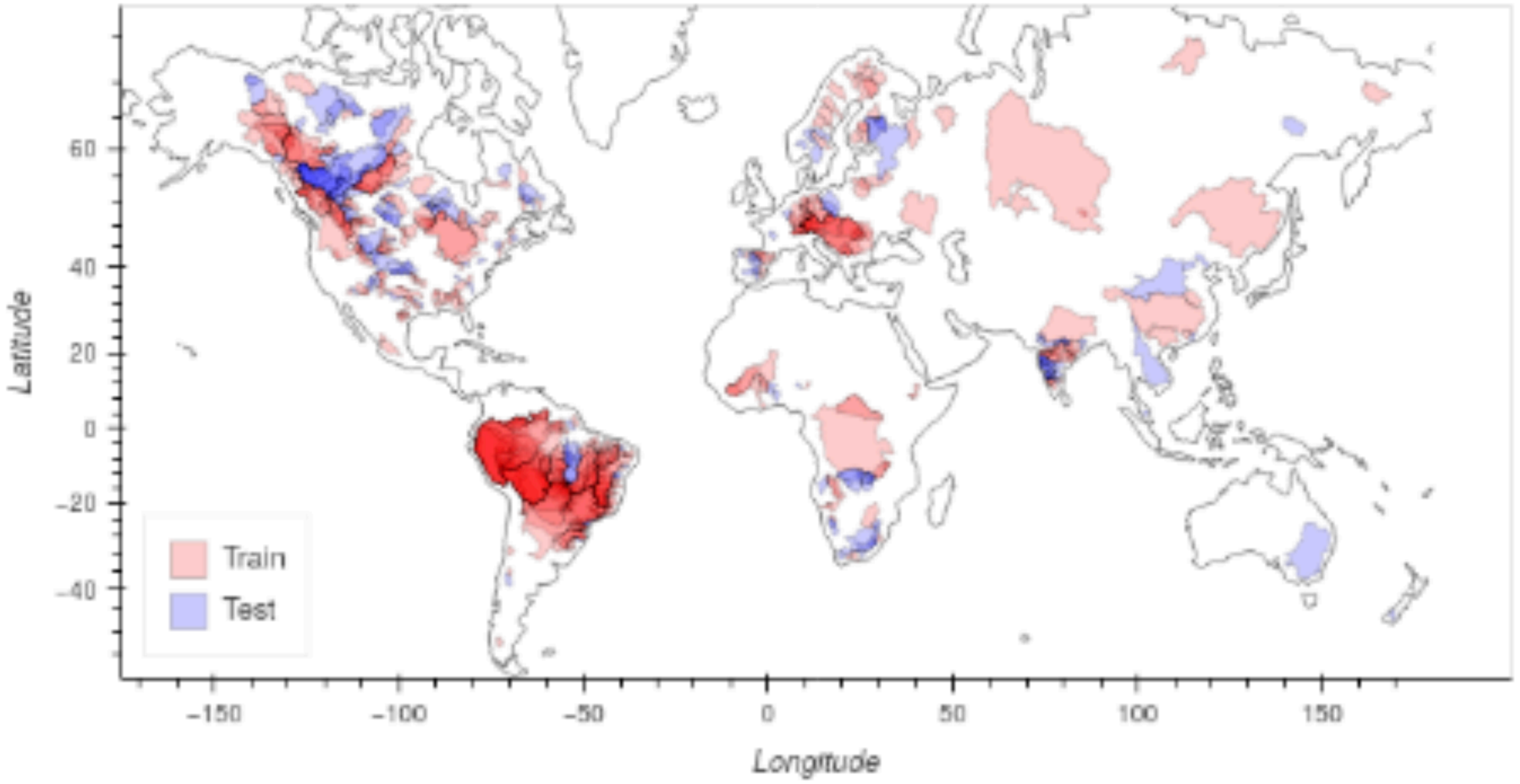


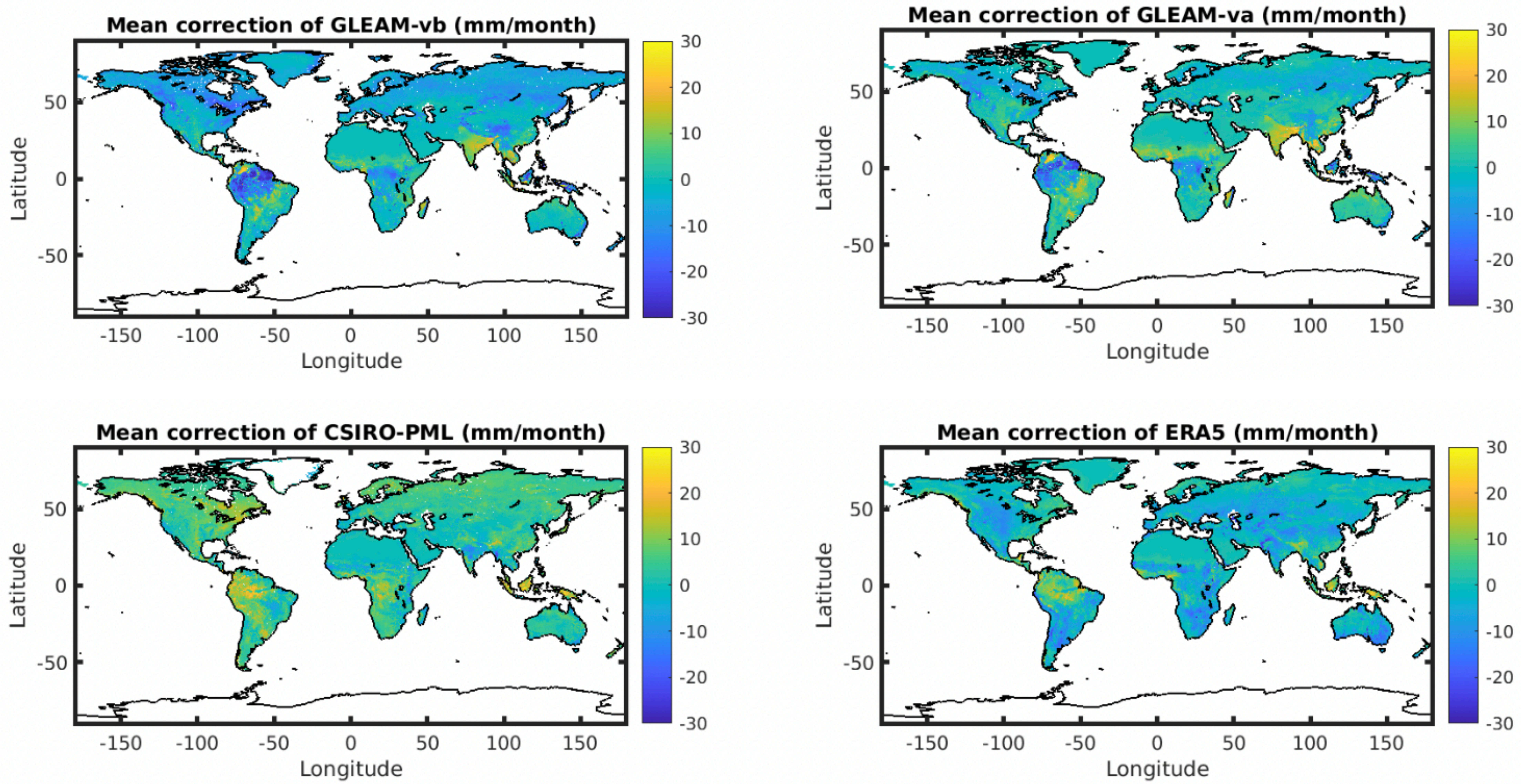
Illustration of our dataset's catchments distribution

- We collected a dataset of 663 catchments globally distributed.
- We collected various climatic indices to use as input to our correction model (detailed table available in slide 17)
- We investigate a period of 192 months spanning from the year 2000 to 2015
- The spatial resolution of the pixel-level data is 0.25 degrees

Dataset

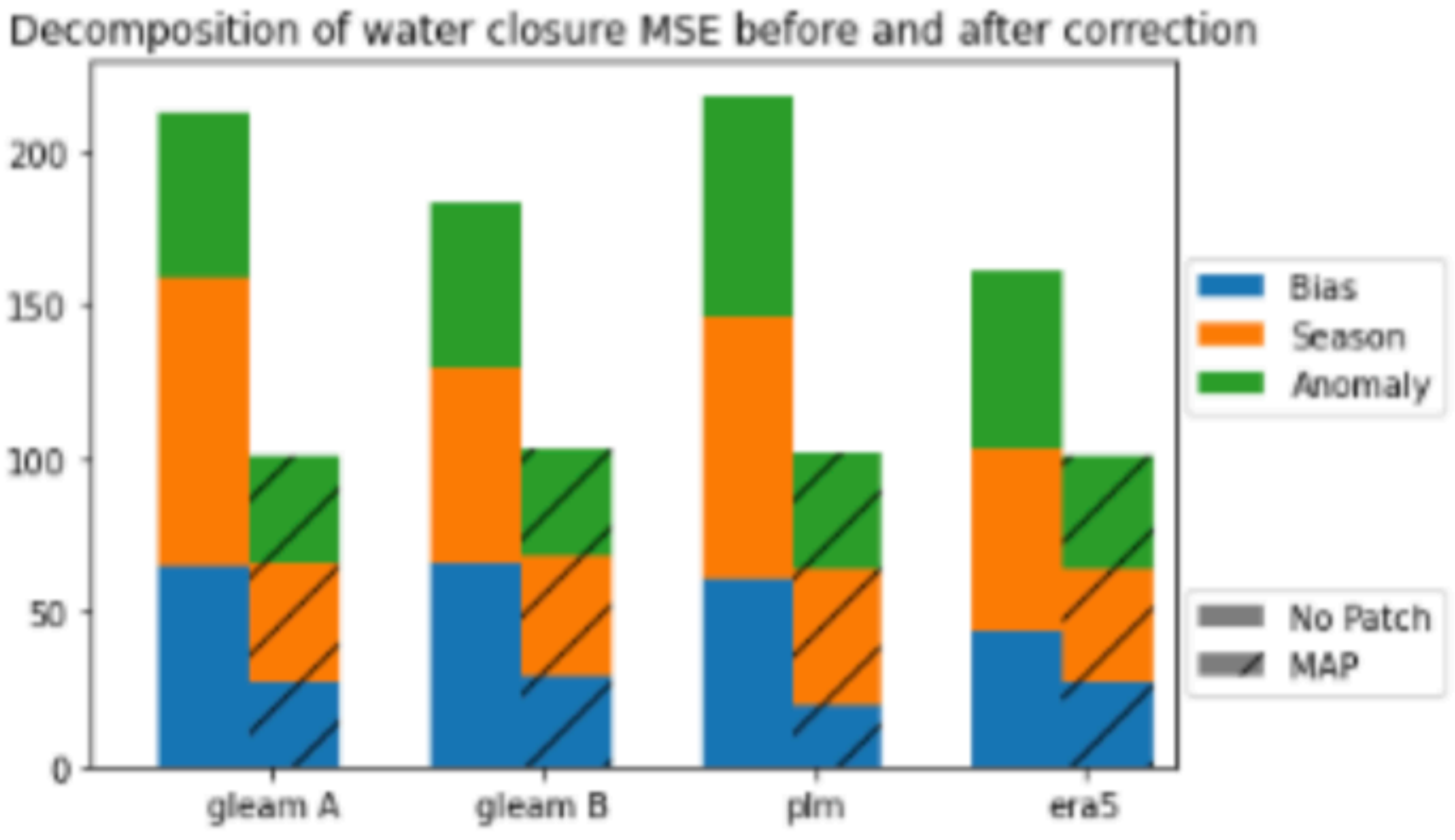
- We applied our methodology to correct 4 different global ET datasets:
  - The Global Land Evaporation Amsterdam Model (GLEAM) version va.3 and vb.3
  - The Penman-Monteith-Leuning (PML) estimate
  - The ET variable of the ERA5 reanalysis.

Correction Visualizations (Bias)

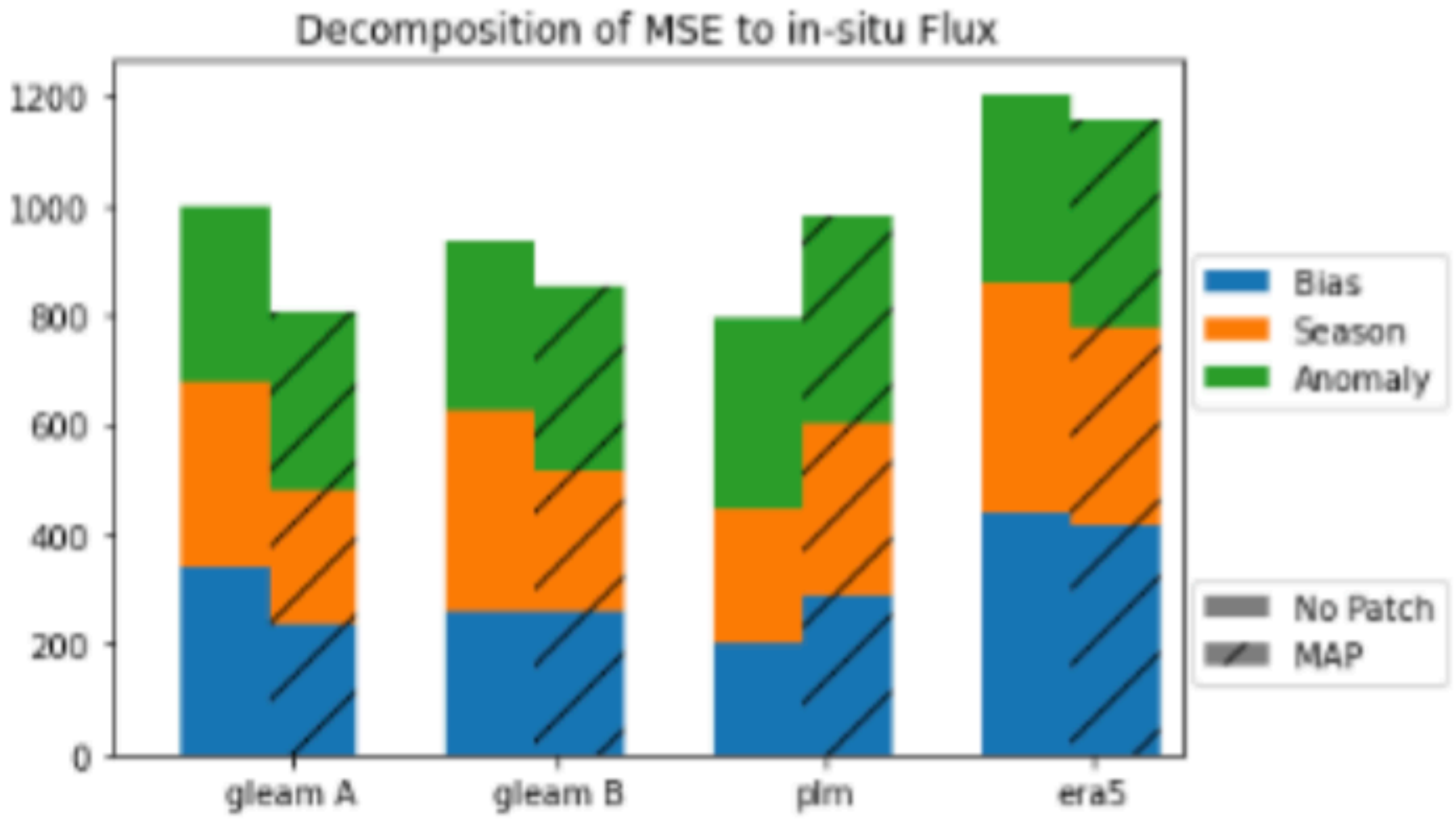


Visualizations

- We validate our results by comparing global ET datasets with and without applying our learned corrections on:
  - The water cycle closure residual on our test catchments (left)
  - Difference between global ET estimates and in-situ measurements (right)



(a) Water Cycle closure MSE



(b) In-situ data MSE

Conclusion

Conclusion

- We proposed a methodology to bring corrections to global pixel-level ET datasets
- Our approach consists in regressing catchment-level corrections computed with OI framework from pixel-level globally available climatic indices.
- Our model was shown to improve the water cycle closure across three datasets and improve on in-situ observations over three of the four datasets

- A long-form version of this paper with additional results and explanations is in preparation for journal publication, please connect with us if you are interested in further details on this work

Data used in this study

Dataset	Coverage	S. res. (°)	T. res.	Reference
Evapotranspiration				
GLEAM vb	2003-2017	0.25	daily	(Martens et al., 2016)
GLEAM va	1980-2017	0.25	daily	(Martens et al., 2016)
CSIRO	1980-2012	0.5	monthly	(Zhang Yongqiang et al., 2016)
ERA-5	1980-2017	0.25	6h	(Hersbach & Dee, 2016)
Precipitation				
GPCP	1979-2015	1	monthly	(Huffman et al., 2001)
TMPA	2002-2015	0.25	daily	(Huffman et al., 2007)
MSWEP	1979-2015	0.5	daily	(Beck et al., 2017)
ERA-5	1980-2015	0.25	6h	(Hersbach & Dee, 2016)
Water storage				
JPL	2002-2017	1	monthly	(Watkins & Yuan, 2014)
CSR	2002-2017	1	monthly	(Bettadpur, 2012)
GFZ	2002-2017	1	monthly	(Dahle et al., 2013)
River network & discharge				
Flow direction	static	0.25	NA	(Yamazaki et al., 2019)
Discharge	1980-2015	NA	monthly	(Do et al., 2018)
Auxiliary information used in the ML-correction model				
Soil moisture	1980-2015	0.25	6h	(Hersbach & Dee, 2016)
Tskin	1980-2015	0.25	6h	(Hersbach & Dee, 2016)
LAI	1980-2015	0.25	6h	(Hersbach & Dee, 2016)
NDVI	1980-2015	0.25	daily	(Mu et al., 2011)