

INTERPRETABLE ML FOR EE

DETECTION: AN APPLICATION TO DROUGHTS IN THE PO RIVER BASIN

PAOLO BONETTI, MATTEO GIULIANI, VERONICA CARDIGLIANO, ALBERTO M. METELLI, MARCELLO RESTELLI AND ANDREA CASTELLETTI

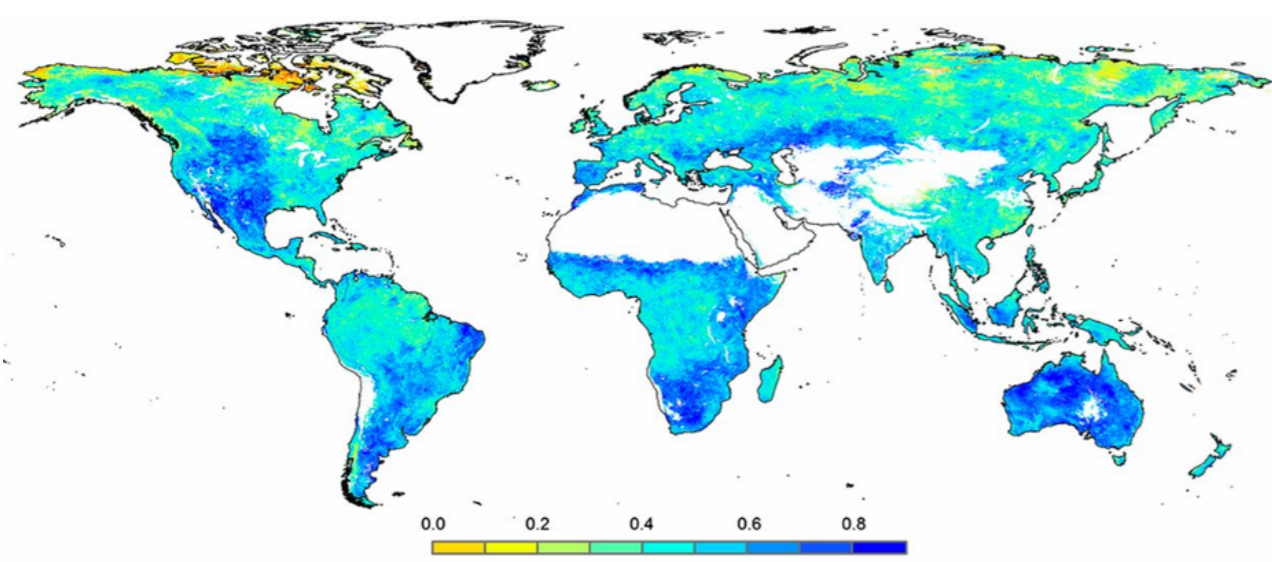


POLITECNICO MILANO 1863



MOTIVATION

- Climate Change: increasing drought frequency and intensity
- Traditional indicators: not impact based (Vicente-Serrano et al., 2010)
- Satellite indices: only real-time
- Existing ML data-driven indices: region-specific and/or black box (Zaniolo et al., 2018)

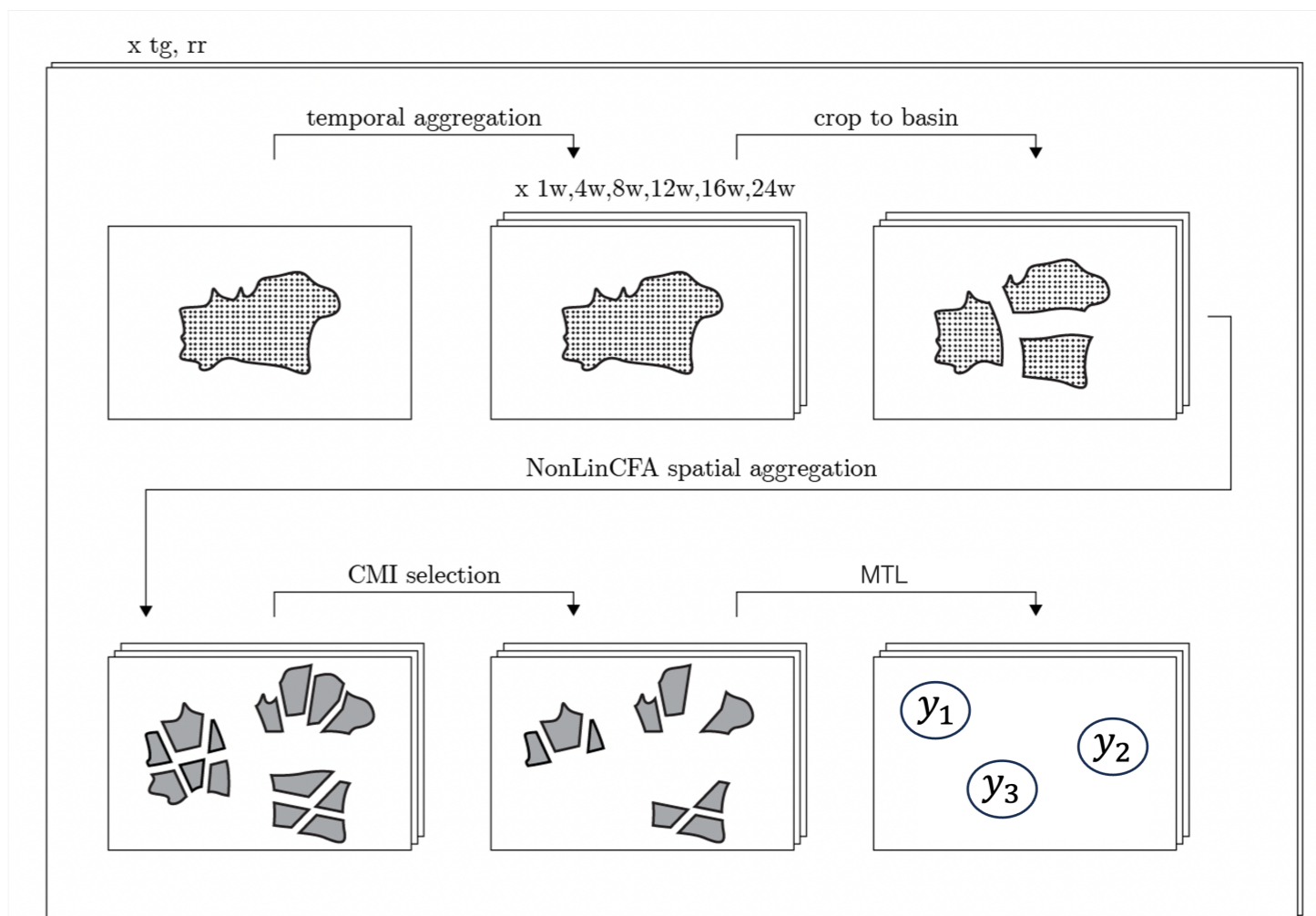
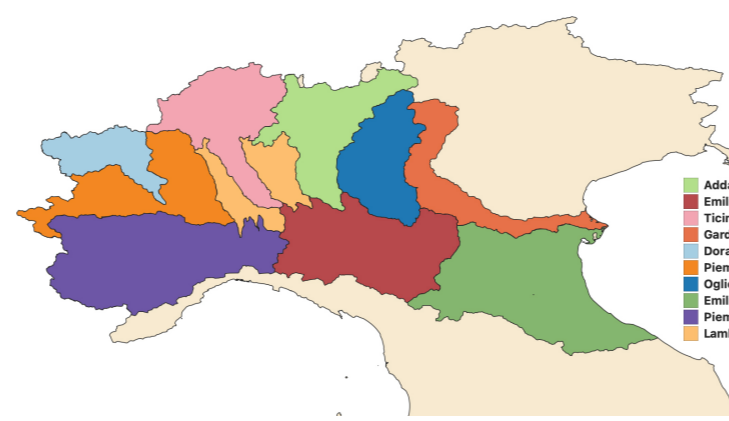


IDEA AND ML WORKFLOW

- State of vegetation proxy: remotely sensed Vegetation Health Index (VHI)
- General ML pipeline to reconstruct the satellite index with forecastable inputs:

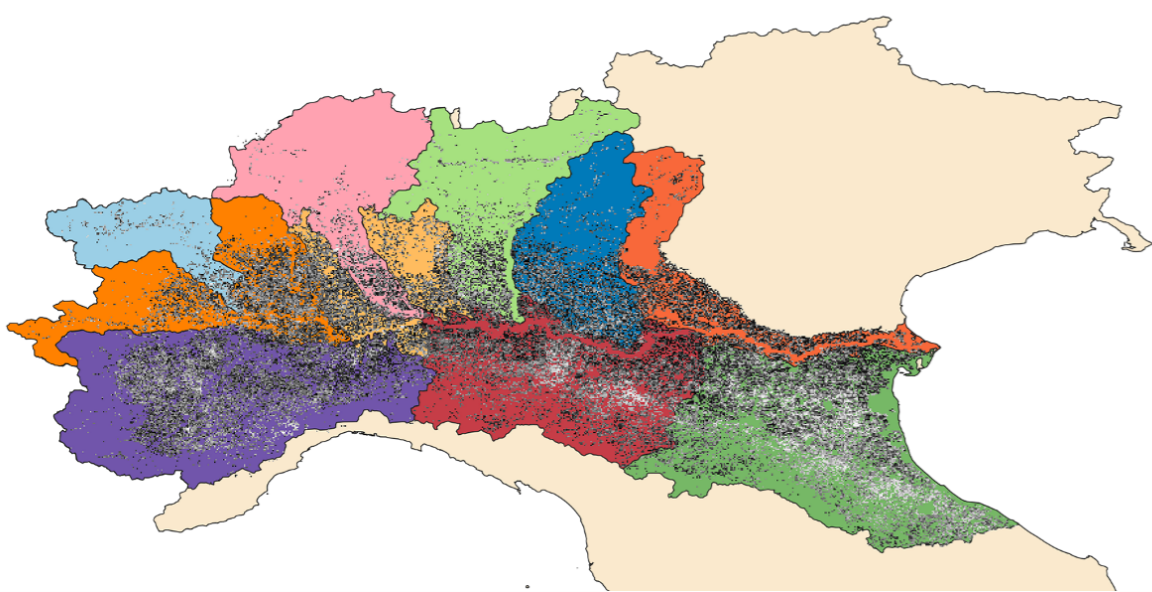


- Interpretable pipeline for trustable index reconstruction with ML:
- Focus: Po River Basin, divided into ten sub-basins

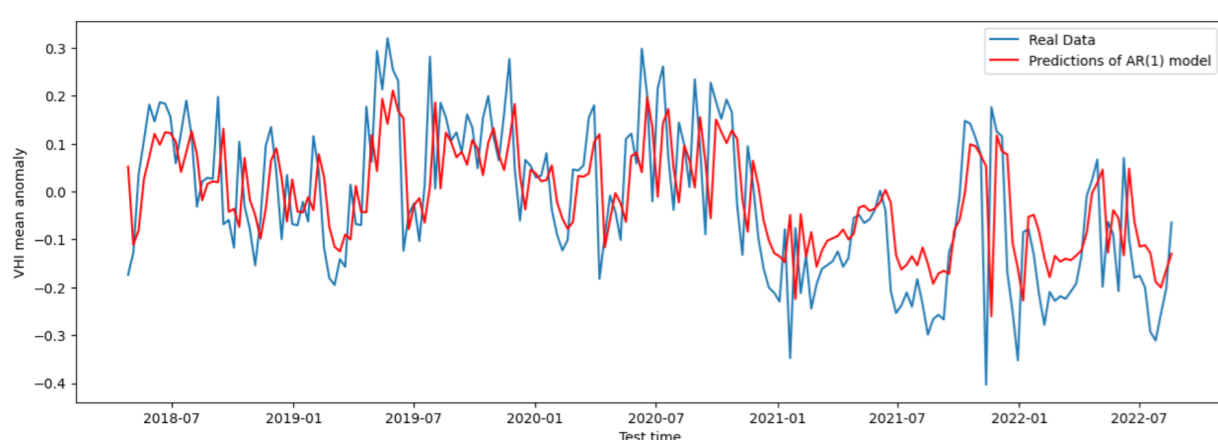


TARGET

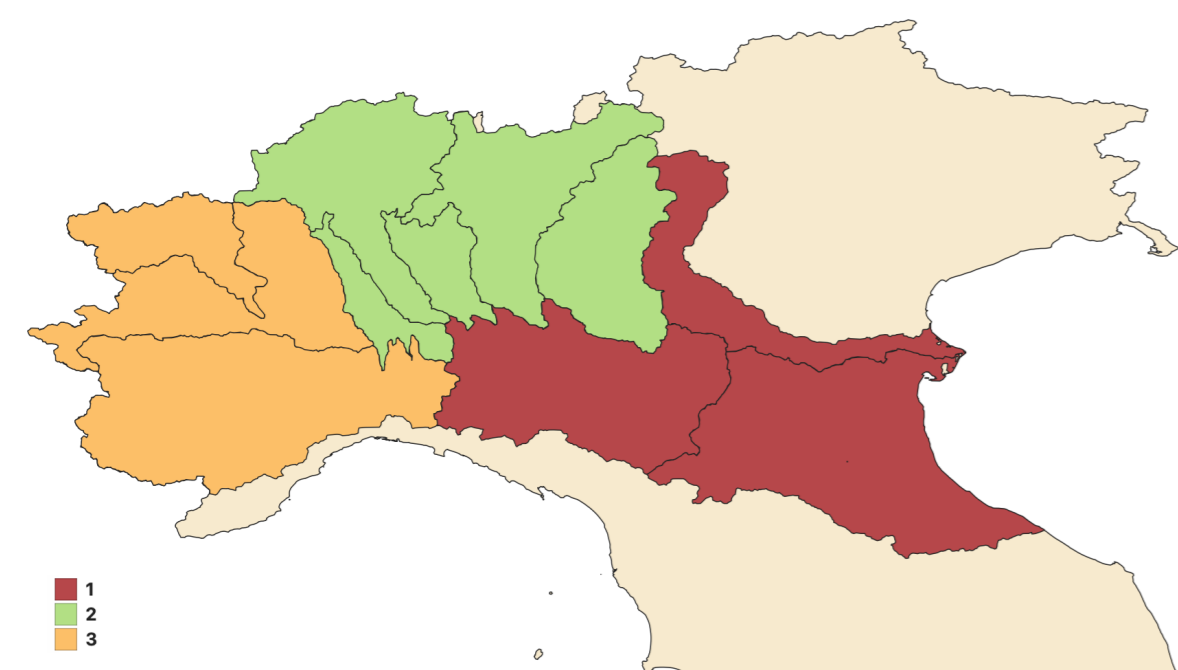
- Average VHI of cultivable pixels



- 8 days temporal aggregation
- Samples: 411 train (2001-2009), 228 val (2010-2014), 228 test (2015-2019)

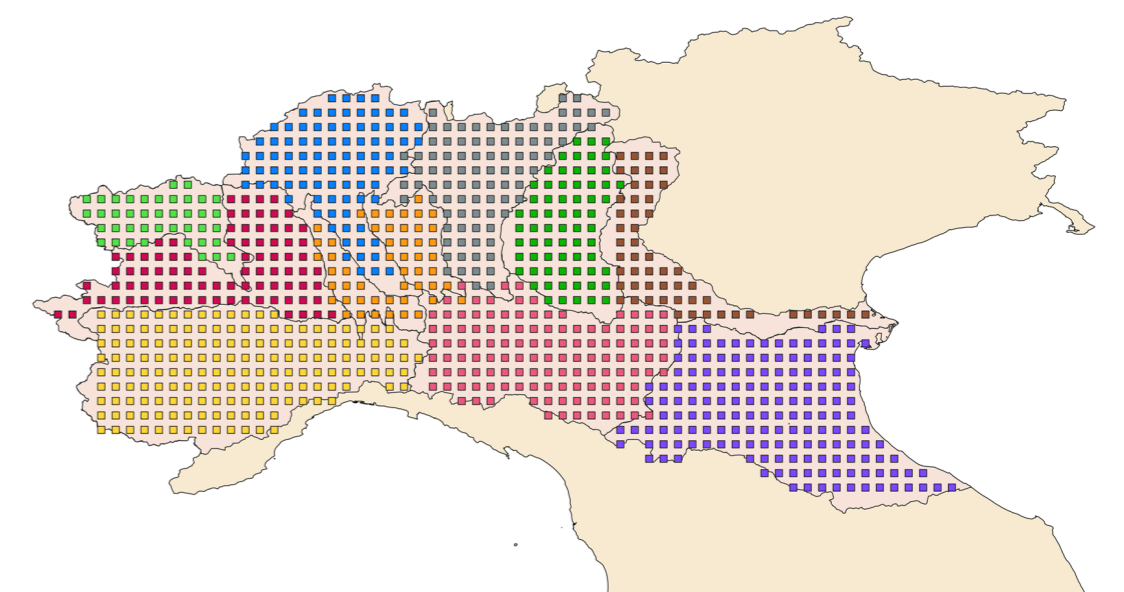


- Clusters: DBSCAN and correlation threshold



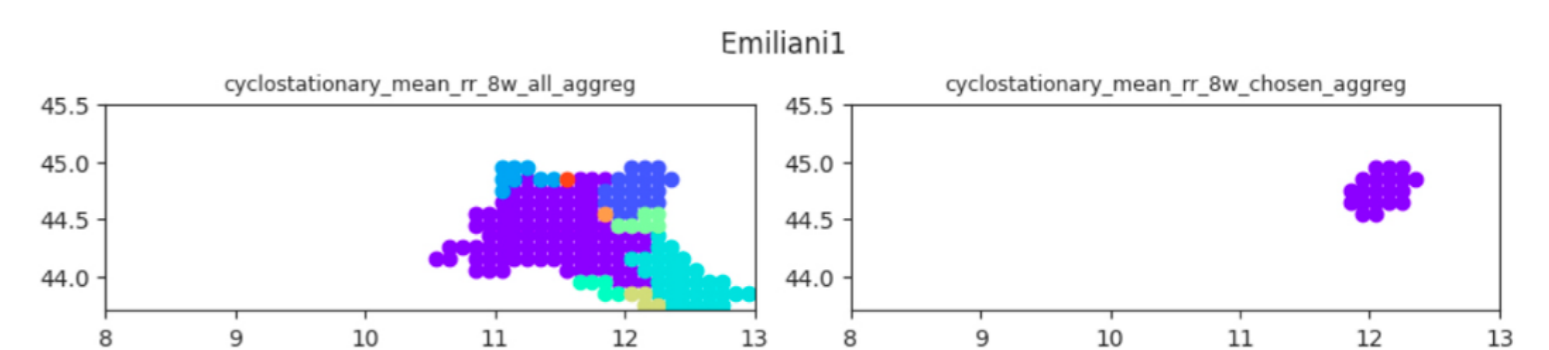
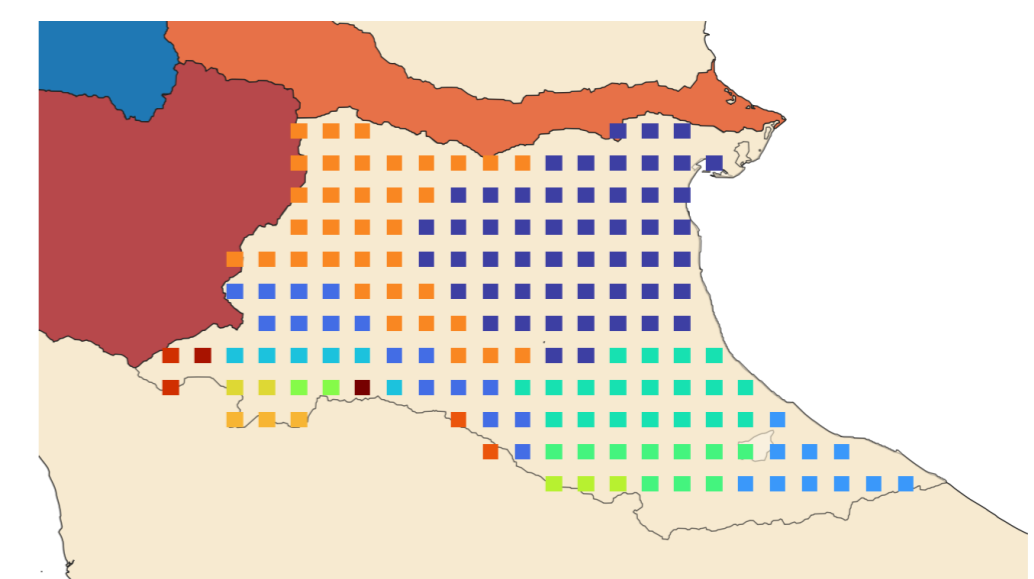
FEATURES

- 991 temperature and precipitation pixels (also snow and lake considered)
- 7 aggregations: 8 days (week), 1 – 4 – 8 – 12 – 16 – 24 weeks, total of 991×14
- Benchmark: full averages, 10×14



EXTRACTION OF INTERPRETABLE RELEVANT FEATURES

- **Dimensionality Reduction:**
 - Lin/NonLinCFA, mean (Bonetti et al., 2024)
 - Only Neighbours
 - Only same basins
 - Only same features
 - Only same timescale
 - Internal ordering
- **Feature Selection:** CMI up to 5 (Beraha et al., 2019)
- Benchmark: wrapper



E1cyclostationary_mean_rr_1w_16
E1cyclostationary_mean_rr_24w_2
E1cyclostationary_mean_rr_4w_1
E1cyclostationary_mean_tg_0

DATA-DRIVEN, IMPACT-BASED, DROUGHT INDICES

Test R ²	Adda	Emiliani1	Emiliani2	Garda_Mincio	Lambro_Olona	Oglio_Iseo	Ticino
NonLinCFA	0.10	0.29	0.24	0.17	0.26	0.20	0.16
+ MTL	0.21	0.33	0.29	0.23	0.24	0.22	0.23
Averages	0.16	0.32	0.22	0.19	0.21	0.20	0.17
Wrapper	0.09	0.24	0.27	0.21	0.20	0.18	0.22
SMA	0.01	0.03	0.02	0.05	0.08	0.05	0.06

REFERENCES

- M. Beraha et al. Feature selection via mutual information: New theoretical insights. In *IJCNN*, pages 1–9. IEEE, 2019.
- P. Bonetti, A. M. Metelli, and M. Restelli. Interpretable linear dimensionality reduction based on bias-variance analysis. *Data Mining and Knowledge Discovery*, pages 1–69, 2024.
- S. M. Vicente-Serrano, S. Beguería, and J. I. López-Moreno. A multiscale drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, 23(7):1696–1718, 2010.
- M. Zaniolo et al. Automatic design of basin-specific drought indexes for highly regulated water systems. *Hydrology and Earth System Sciences*, 2018.