Improving forecast of precipitation extremes with machine learning

F. Grazzini (LMU/ARPAE), J. Dorrington (KIT), G. C. Craig(LMU), C. Grams (KIT),

L. Magnusson (ECMWF), F. Vitart(ECMWF)

Project T2: Towards seamless prediction of EXtremes (TEX)

5–8 June 2023

UEF 2023

Ensemble Forecasting

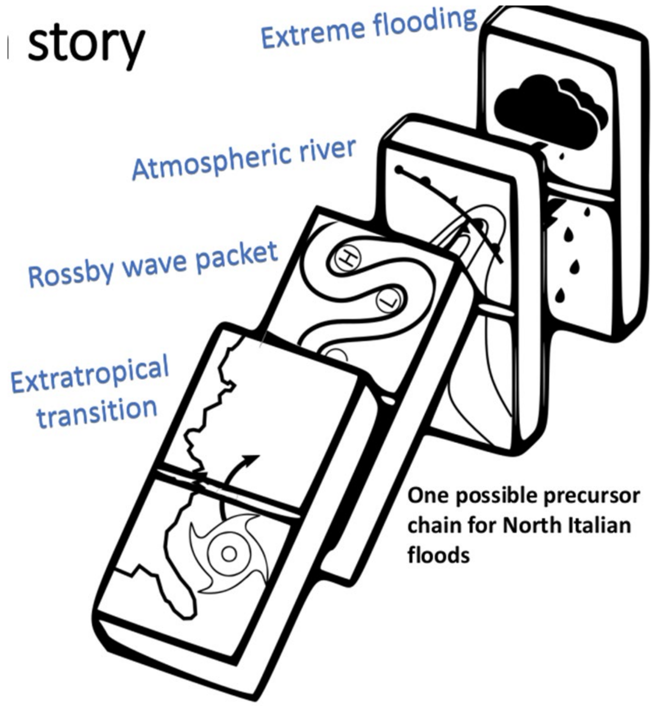
#UEF2023

Concept: every EPE has its own dynamical history but also common drivers.
We use early precursors to improve EPE prediction in the medium-range

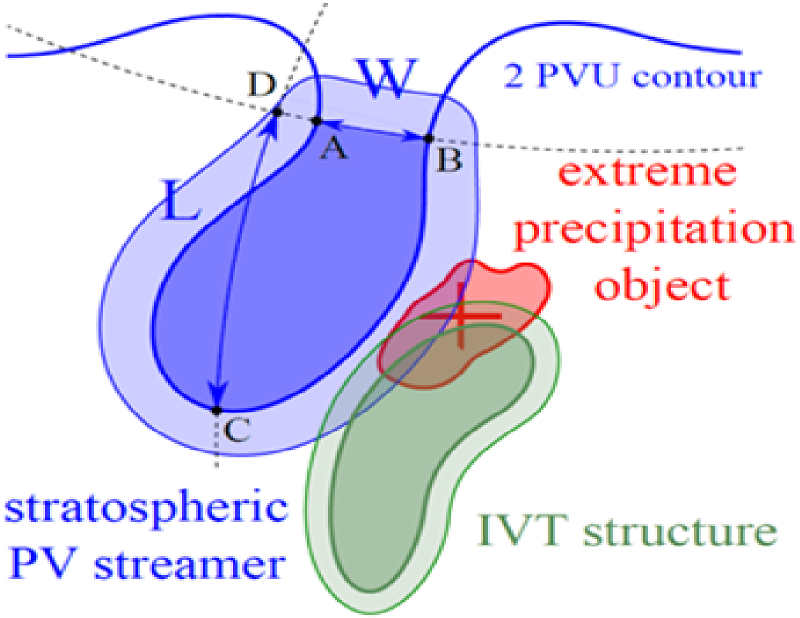
<u>Non-local</u>
large-scale dynamics

<u>Local</u>
dynamical setting

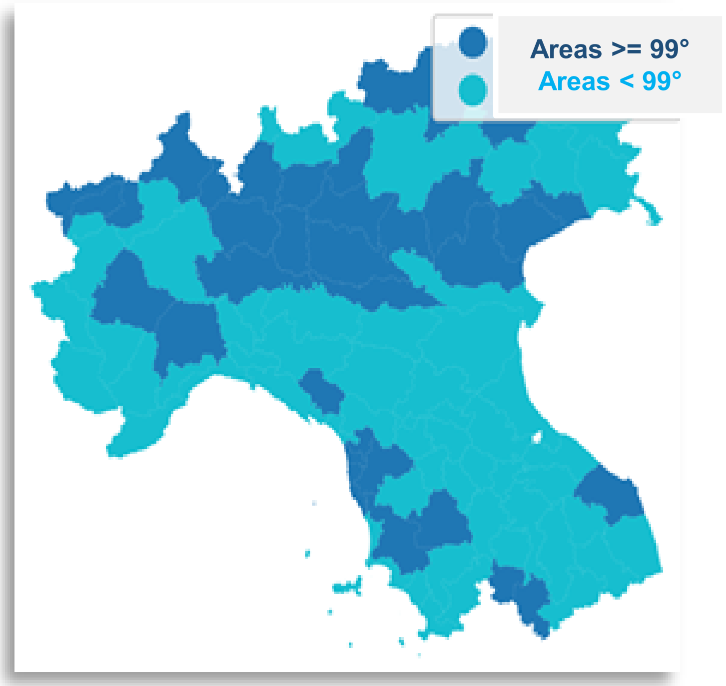
Extreme precipitation event (EPE)



Joshua Dorrington, 2022



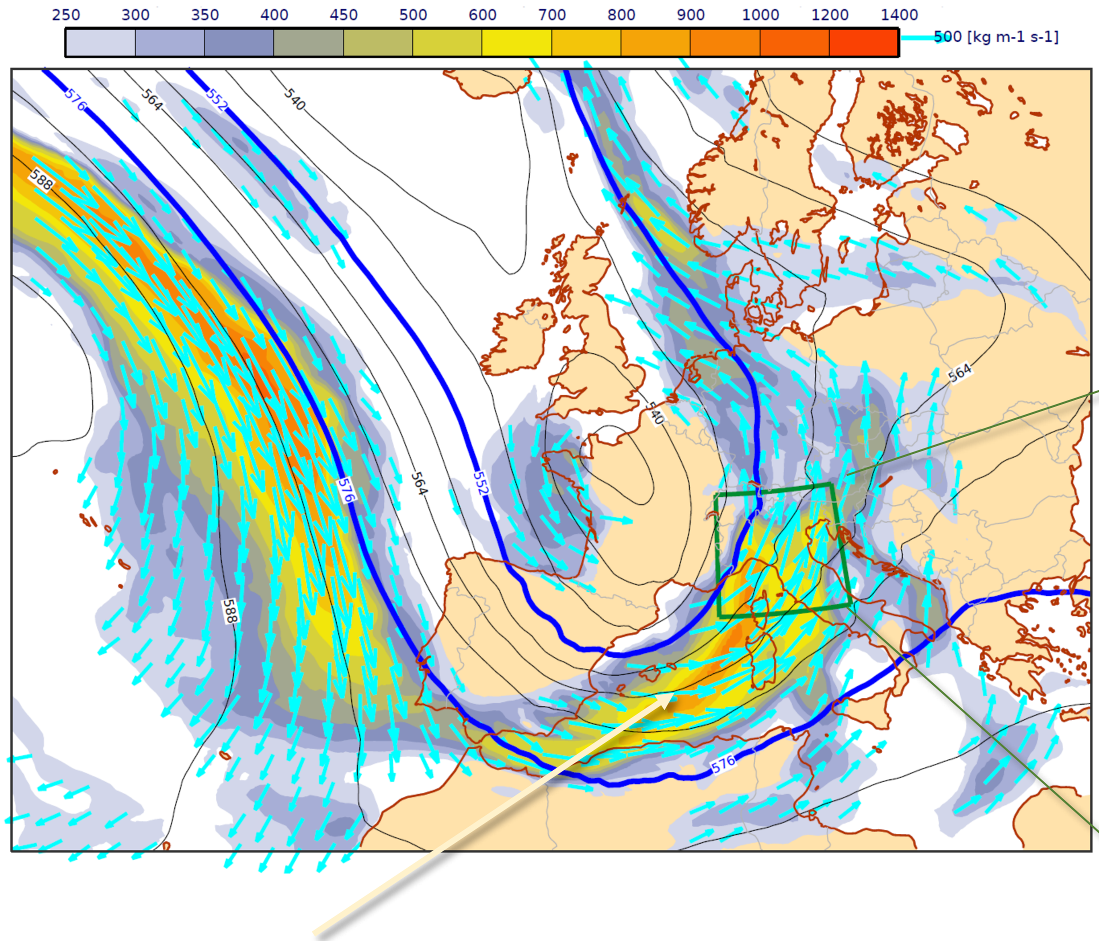
Adapted from De Vries 2020

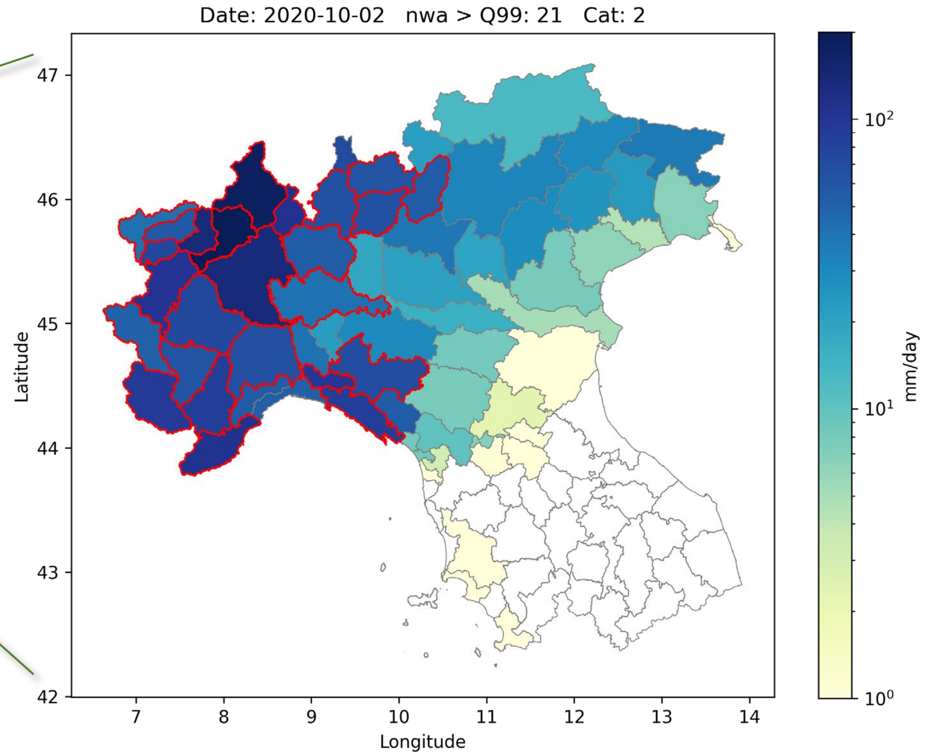


EPE is defined as a day with at least a Warning area of N-IT above the 99° of daily precip

Storm Alex, 2-3 Oct 2020, 500mm in 12h

Local precursors

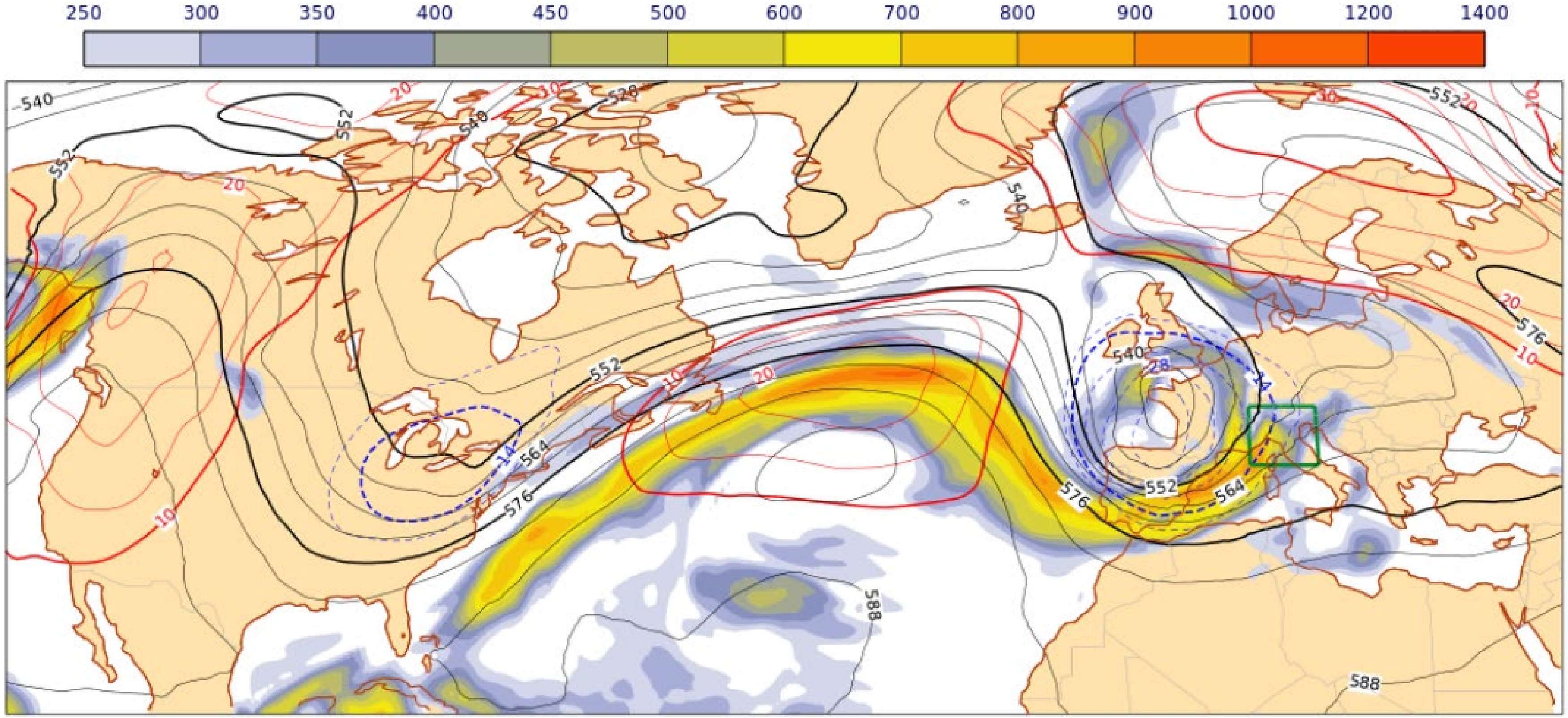




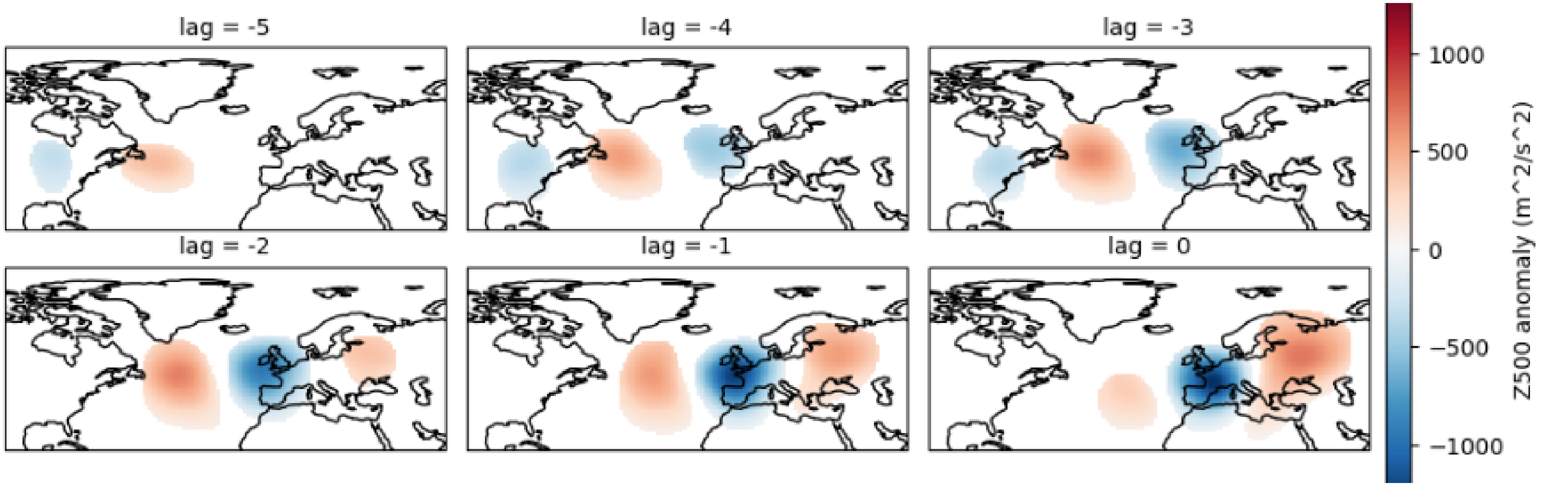
Integrated Water Vapour Transport

Non-local predictors for storm Alex

Non - local precursors

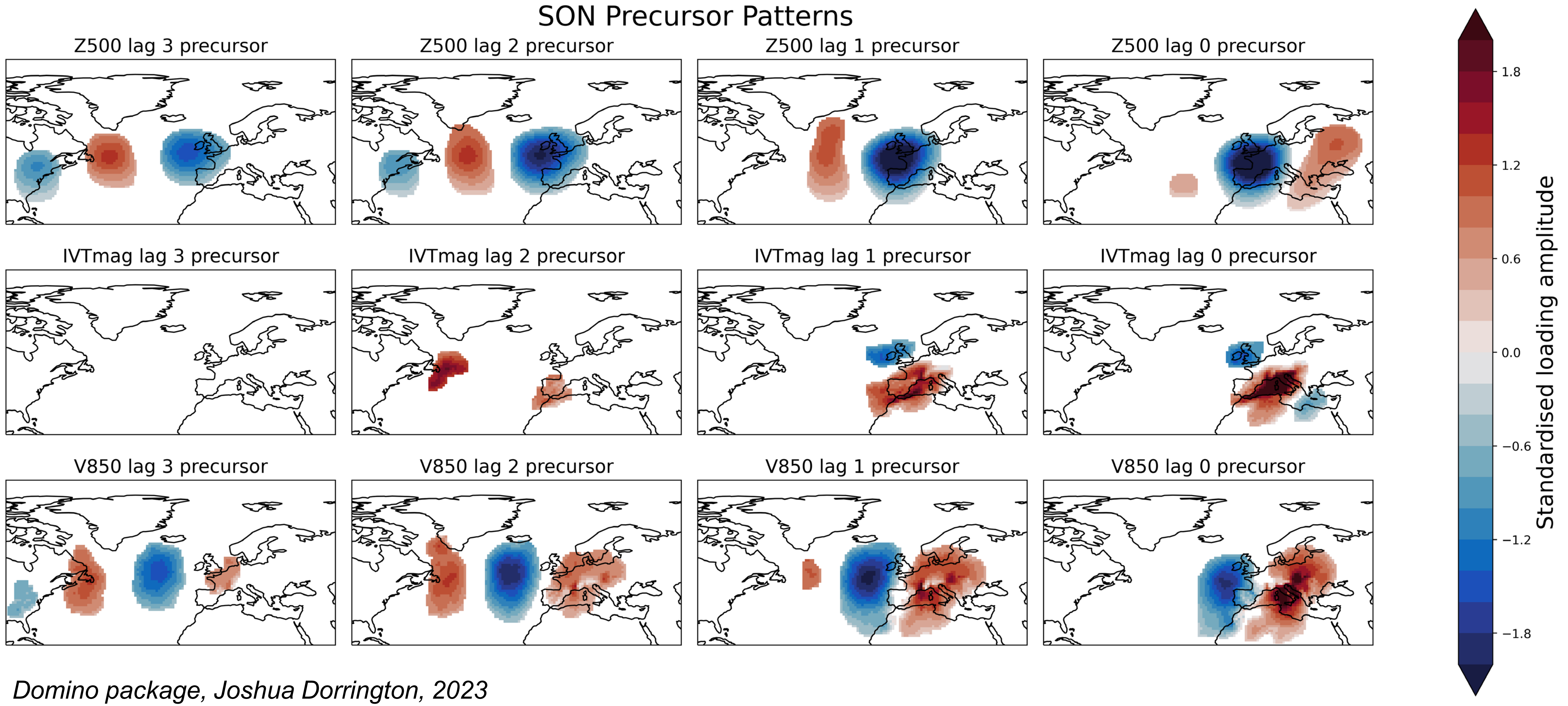


Synoptic situation 2020-10-02 12:00 UTC associated with storm Alex. Geopot height at 500 hPa and IVT. Blue dashed isolines and red solid lines are showing the corresponding 500 hPa anomaly with respect to the seasonal climatological average which is then projected onto the non-dimensional Z500 non-local index



SON **EPE composites of Z500 anomalies** at an increasing number of days of lag respect of EPEs days

Non-local EPE precursors for N-Italy



Domino package, Joshua Dorrington, 2023

Set up of the hybrid model (MaLCox)

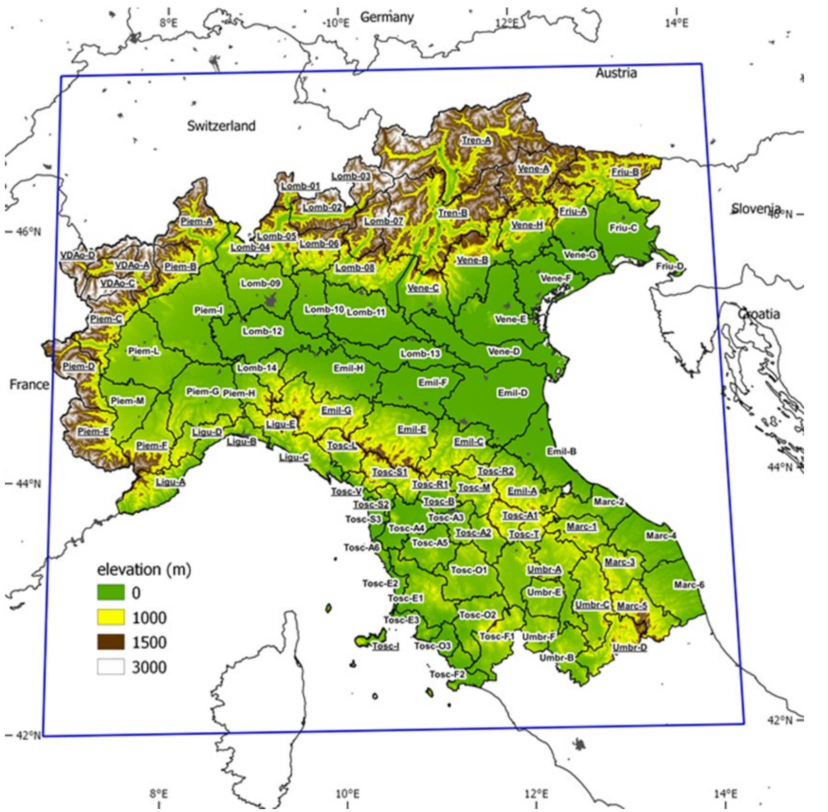
A Machine Learning model predicting Conditions for eXtreme precipitation **(MaLCox)** has been trained from IFS CTRL hindcast 24-360h (2000-2020) predictors, on each forecast lead time.
For each lead time ~ 5000 days of which ~350 EPEs days (7% occurrence)

Non-Local Predictors; (Z500, V850, IVT) EU-ATL EPE composites anomalies
Local Predictors; based on domain averages over N-IT
Direct Predictors; volume of rain predicted over N-IT
Climatological ; day of the year

Target : EPE (yes/no) defined as a day in which one or more warning areas aare exceeding the 99° of daily precipitation during rainy days in 1991-2020. Based ARCIS dataset (high-resolution gridded precipitation of N-IT).

Benchmark to beat: Forecast of EPE obtained with IFS HRES direct model output

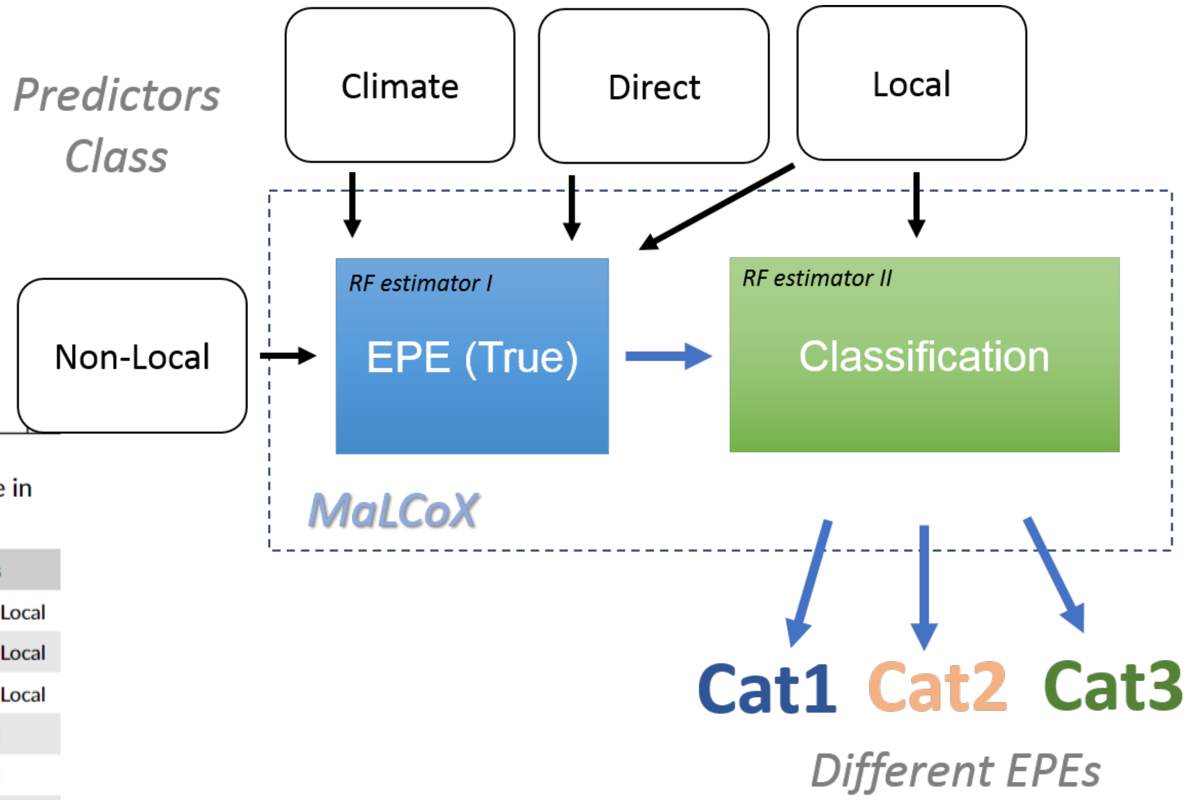
Test period: 2018-2022 with HRES predictors (~70 EPEs days)



PS
The tricky problem of reducing complexity (and overfitting), and accounting for classes imbalance, was addressed setting class_weight to "balanced_subsample" and ccp_alpha to 0.001

Hybrid model architecture and predictors

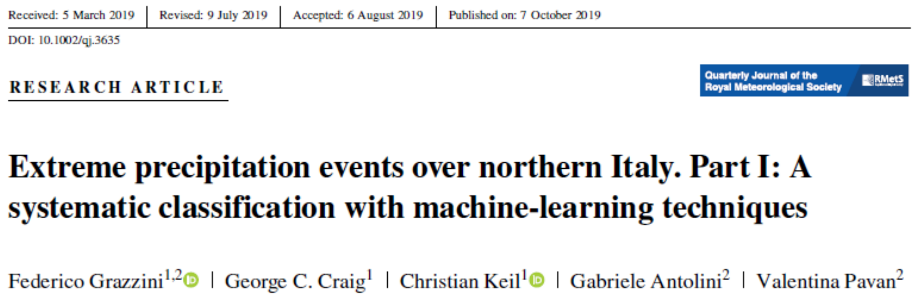
Grazzini et al. 2023, in preparation



Grazzini et al. 2023

TABLE 1 Table showing the list of predictors. Predictors are subdivided according with their type and usage in the two step of MalCoX model

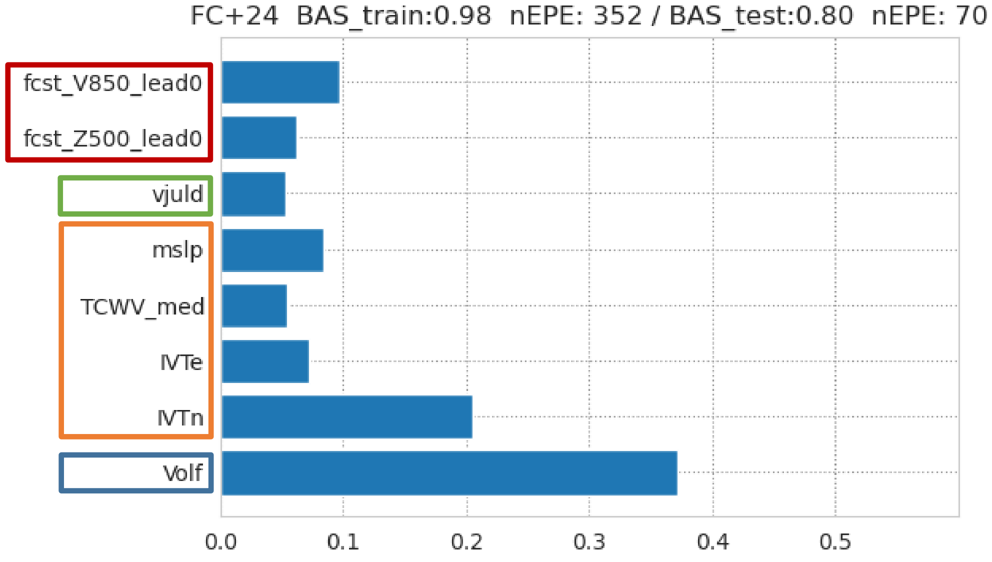
Variable	Description	Units	RF model	Class
fcst_IVTmag	IVT (module) normalized anomaly index lead times from 0 to -2 days		EPE(y/n)	Non-Local
fcst_Z500	Gepot. normalized anomaly index at 500 hPa lead times from 0 to -2 days		EPE(y/n)	Non-Local
fcst_V850	Meridional wind normalized anomaly index at 850 hPa lead times from 0 to -2 days		EPE(y/n)	Non-Local
IVTe	Daily mean of zonal component of integrated water vapour transport	$kgs^{-1}m^{-1}$	EPE(y/n)	Local
IVTn	Daily mean of meridional component of integrated water vapour transport	$kgs^{-1}m^{-1}$	EPE(y/n)	Local
TCWV	Daily mean of total column water vapour	kgm^{-2}	EPE(y/n)	Local
Mslp	Daily mean of mean sea level pressure	hPa	EPE(y/n)	Local
Volf	Daily volume of rain over the target domain	m^3	EPE(y/n)	Direct
Juld	Day of the year (Julian date)		EPE(y/n)	Climate
Θ_{e850}	Daily mean of equivalent potential temperature at 850hPa	K	Classification	Local
$\Delta\Theta_{e500-850}$	Daily minimum of delta $\Theta e(500-850)$hPa	K	Classification	Local
$\Theta pv2$	Daily mean of Θ on dynamical tropopause (pv2)	K	Classification	Local
Taudmax	Daily maximum of convective adjustment time scale	h	Classification	Local
CAPEdmax	Daily maximum of CAPE	Jkg^{-1}	Classification	Local

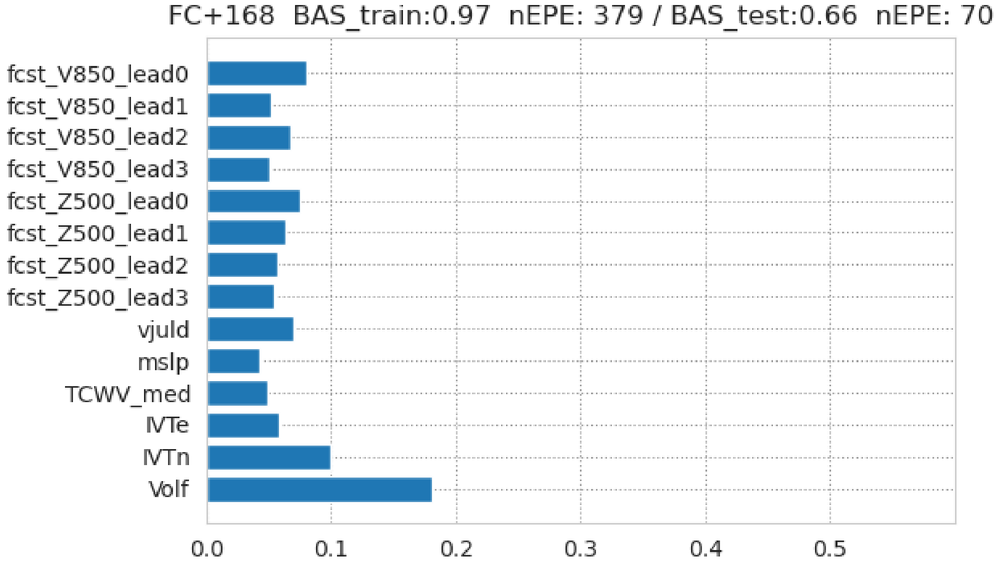

Received: 5 March 2019 | Revised: 9 July 2019 | Accepted: 6 August 2019 | Published on: 7 October 2019
DOI: 10.1002/qj.3635
RESEARCH ARTICLE
Quarterly Journal of the Royal Meteorological Society RMetS

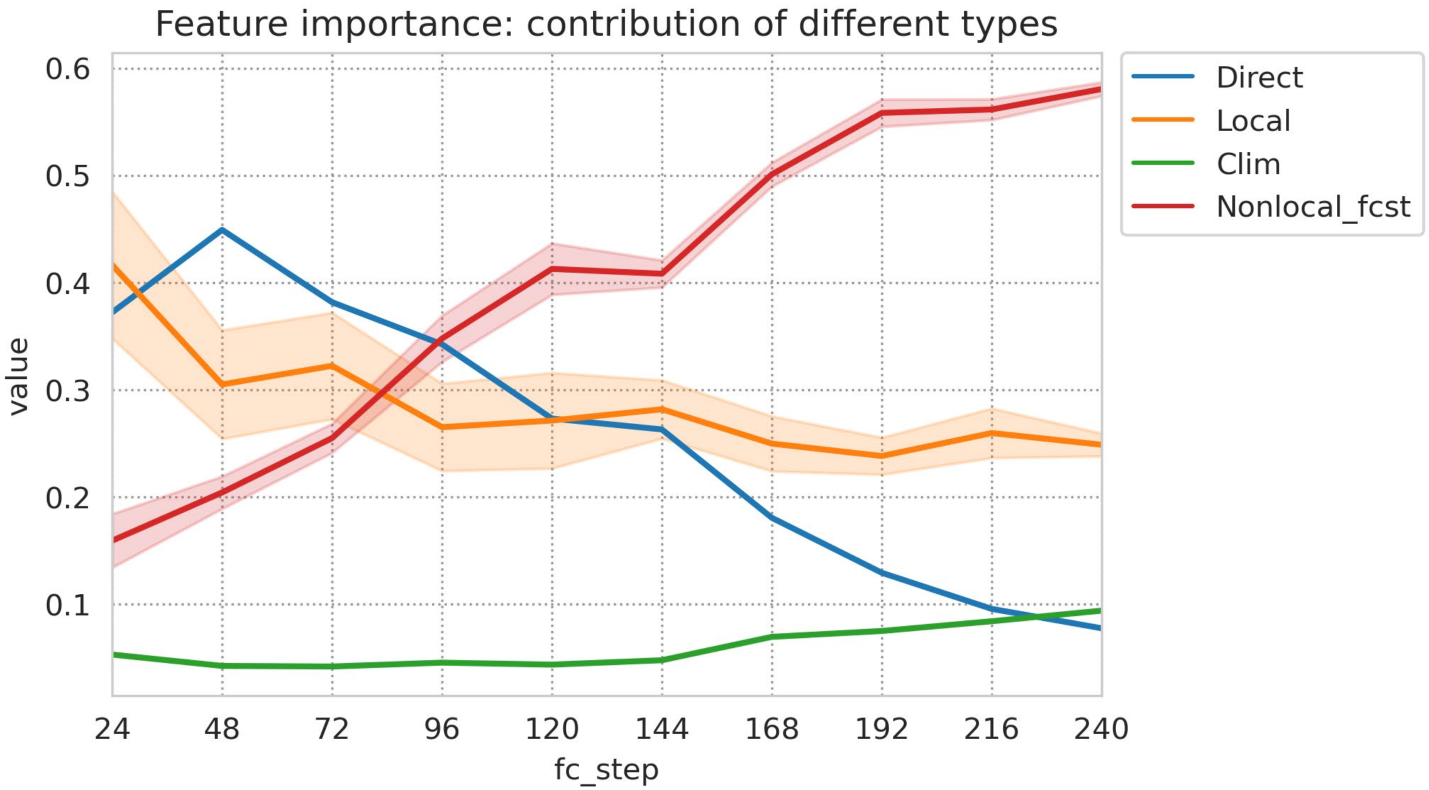
Extreme precipitation events over northern Italy. Part I: A systematic classification with machine-learning techniques

Federico Grazzini[1,2] | George C. Craig[1] | Christian Keil[1] | Gabriele Antolini[2] | Valentina Pavan[2]

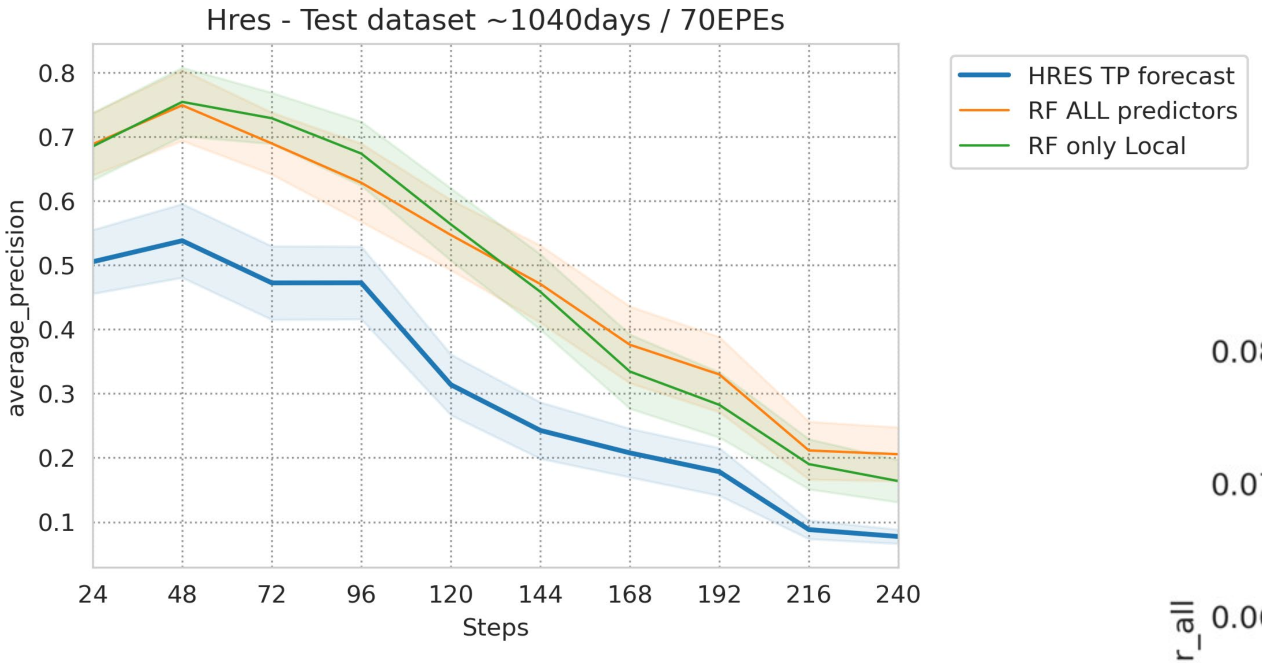
Predictors importance vs lead time: predictability implications

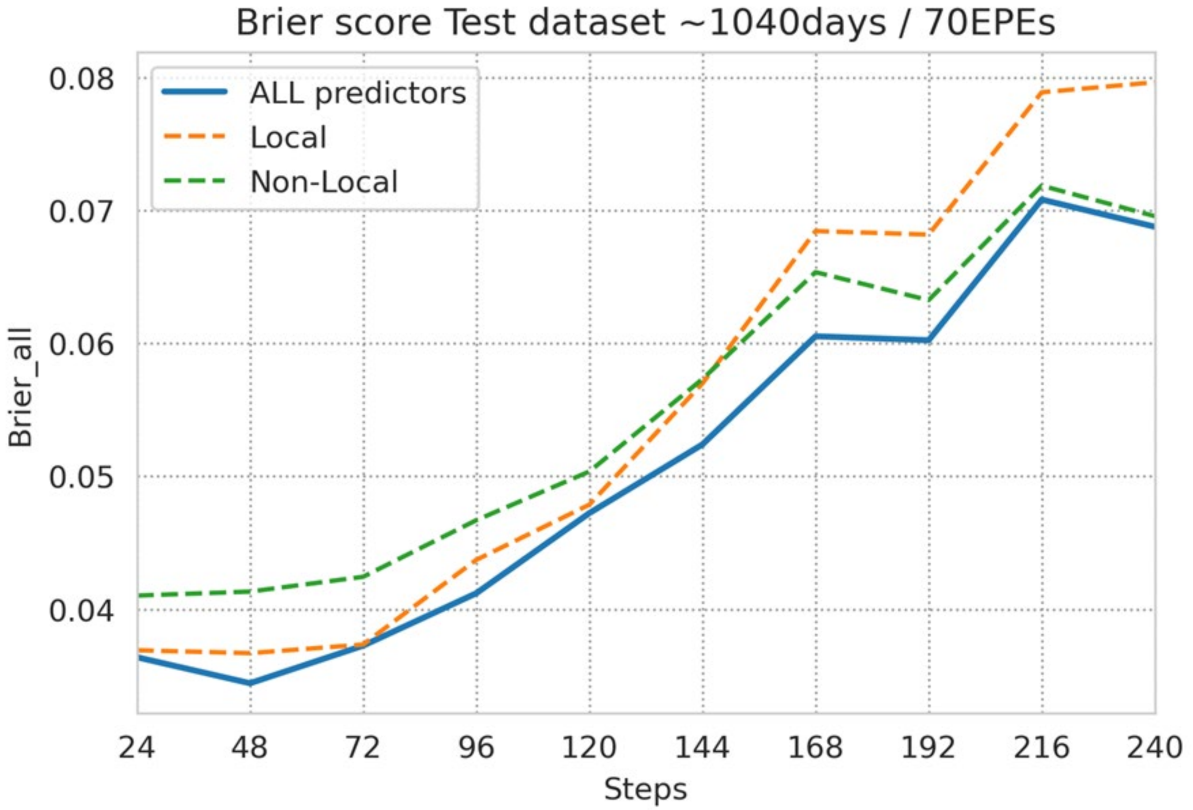


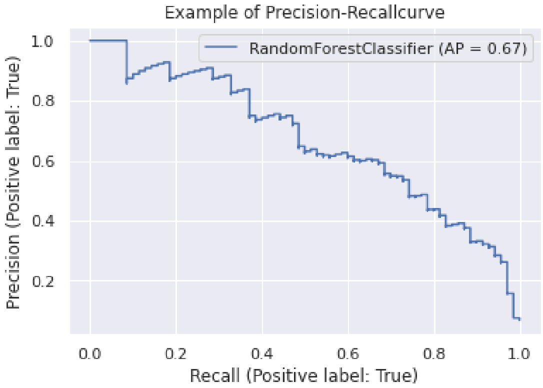




Verification : test dataset, average precision score







positive predictive value

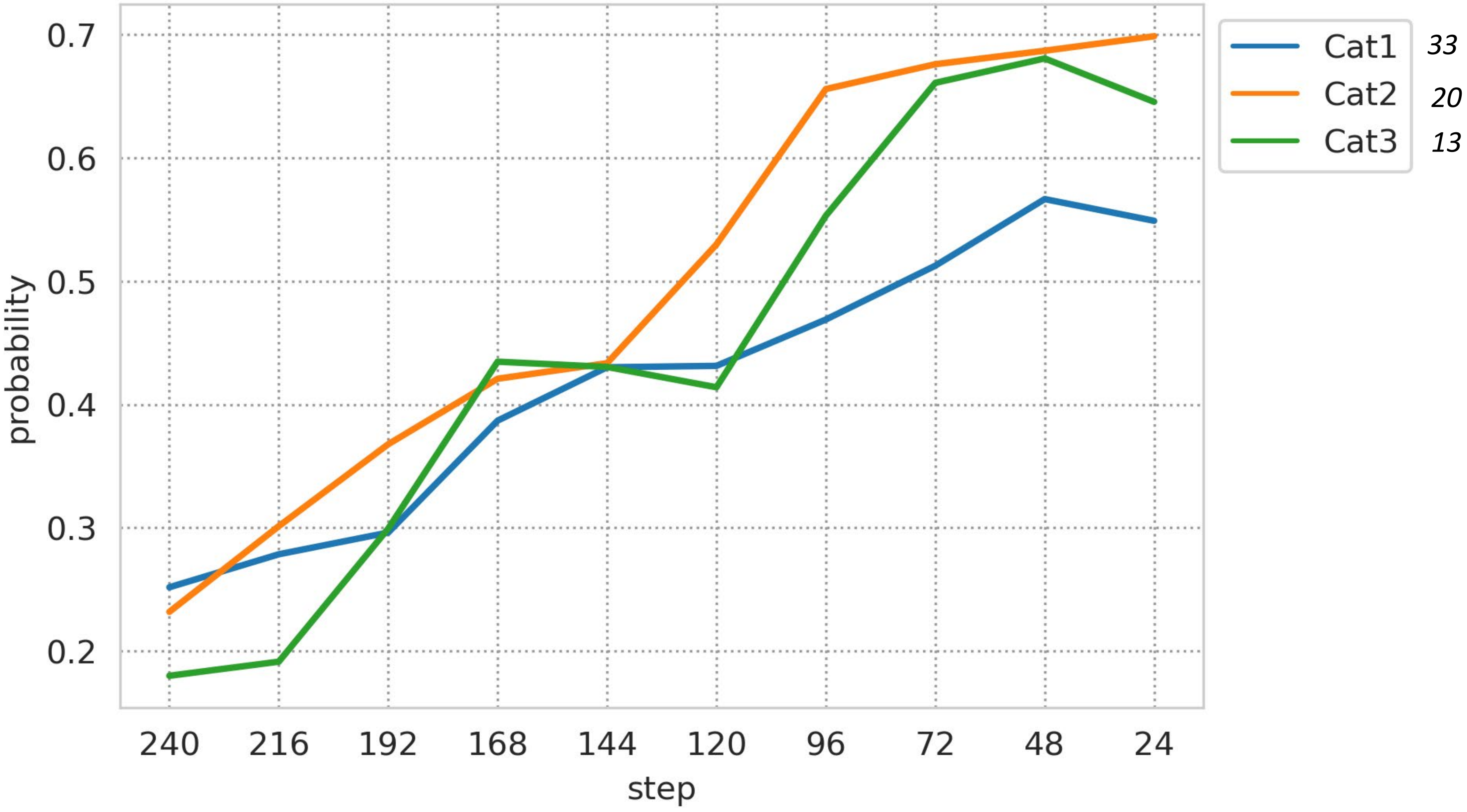
$$\text{precision} = \frac{tp}{tp + fp},$$

$$\text{recall} = \frac{tp}{tp + fn},$$

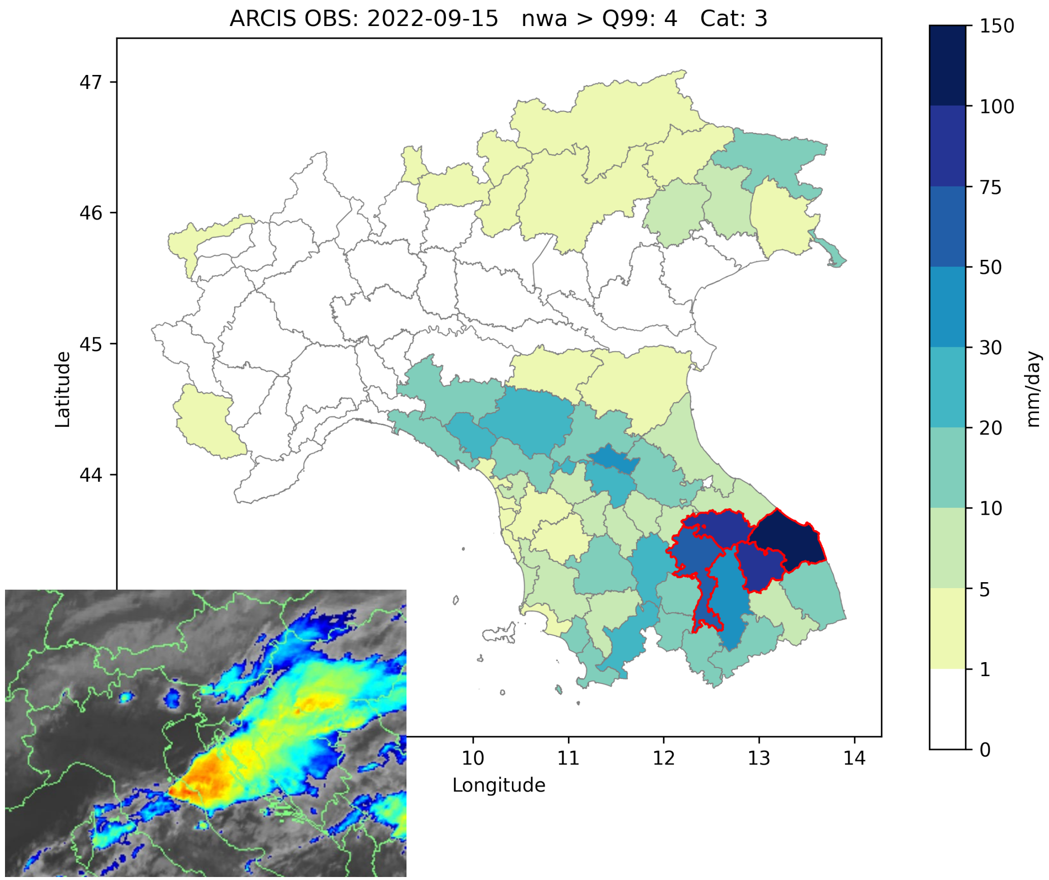
sensitivity

Both measures are based on *relevance*

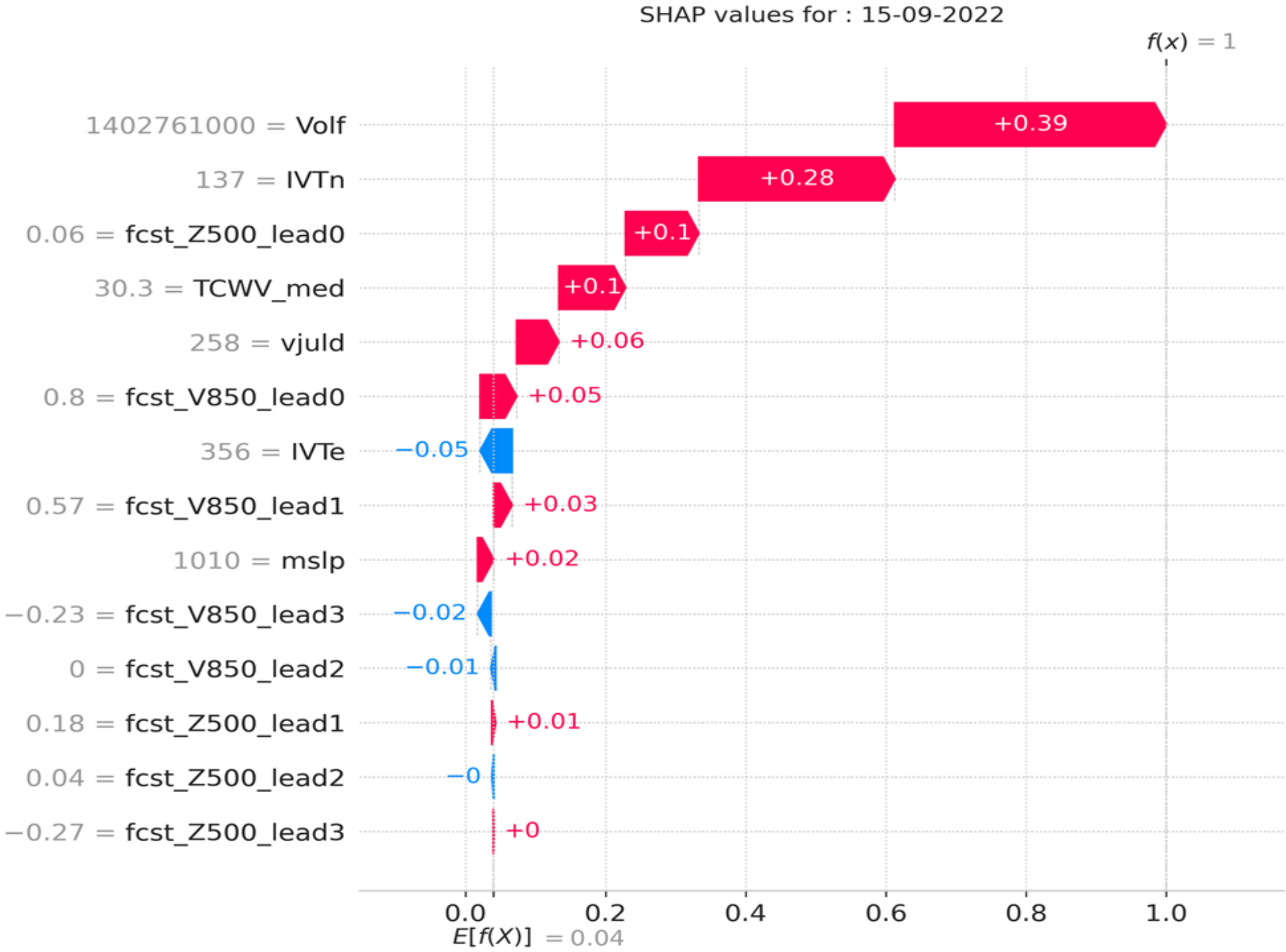
Average probability of MaLCoX prediction when EPE is observed (test period)



Case study : Marche flash-flood 15/09/2022

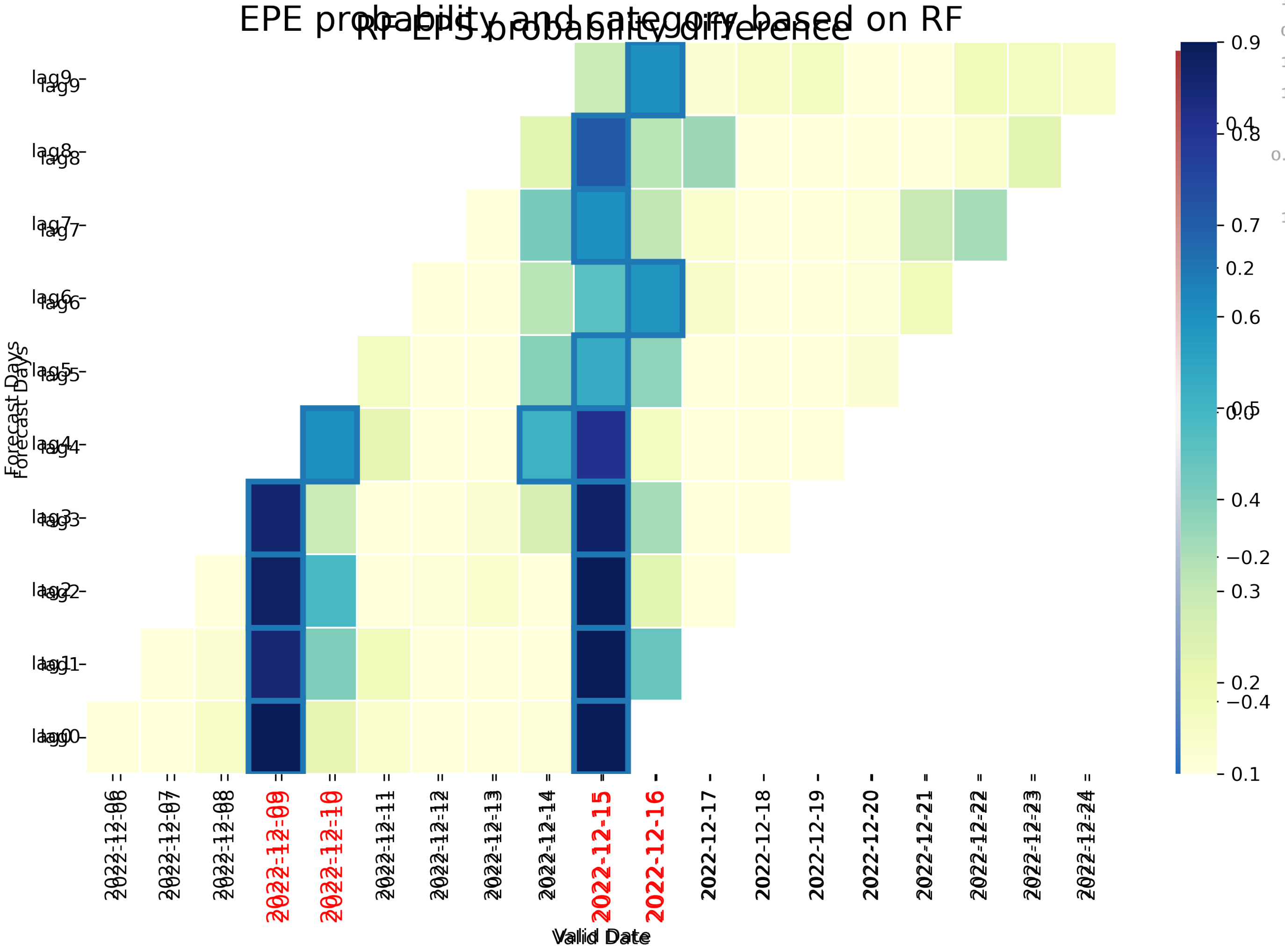


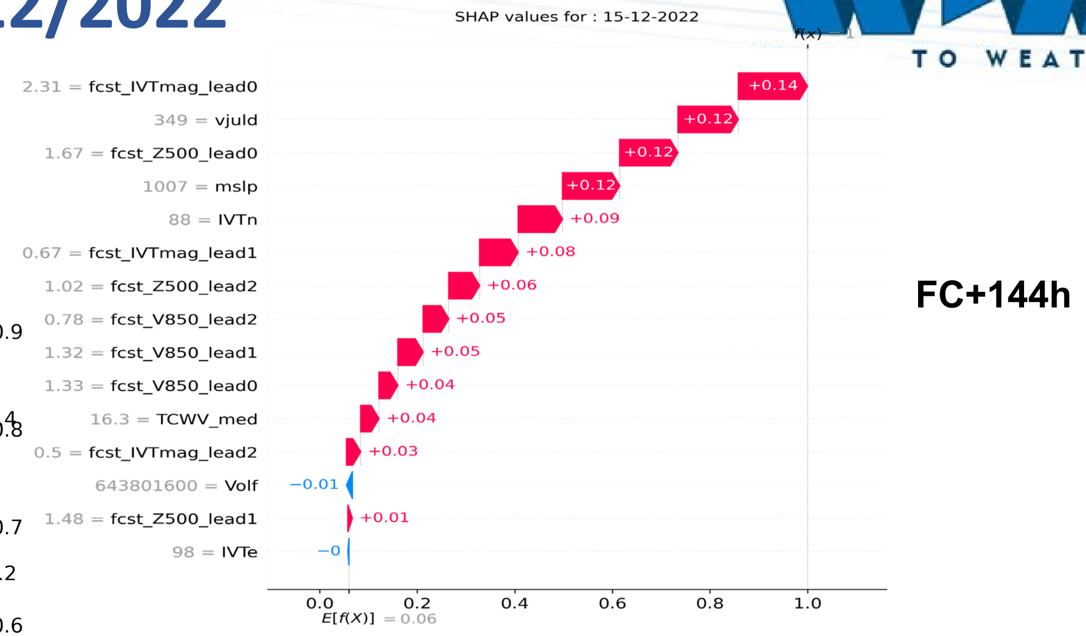
Enhanced IR 15 set 2022, 19:00 UTC

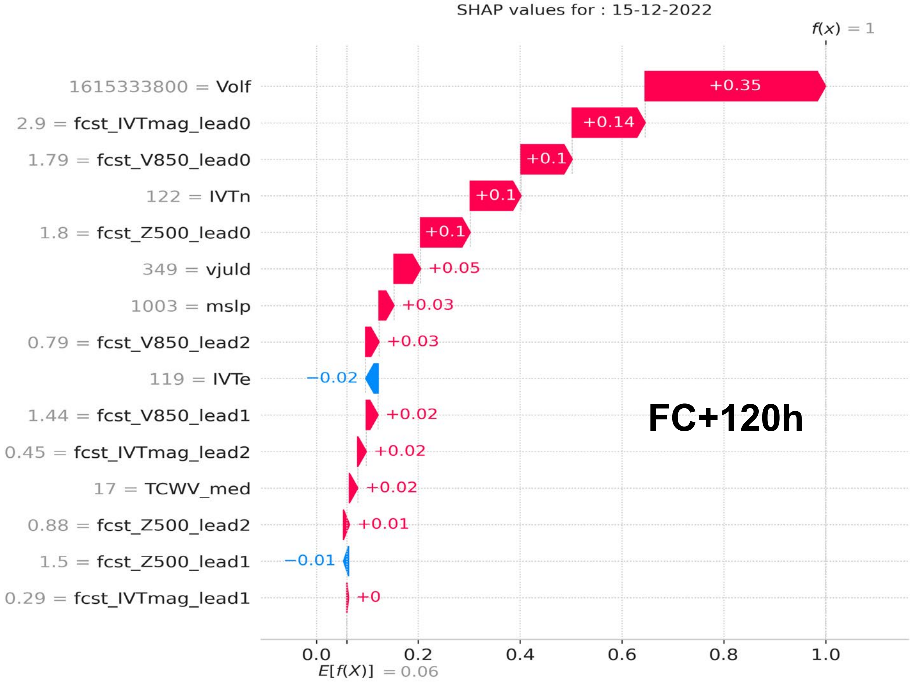


SHapley **A**dditive ex**P**lanations (Lunnberg and Lee, 2017)
how each feature contributed to the RF hybrid forecast

Case studies: Northern Tuscany flood 15/12/2022

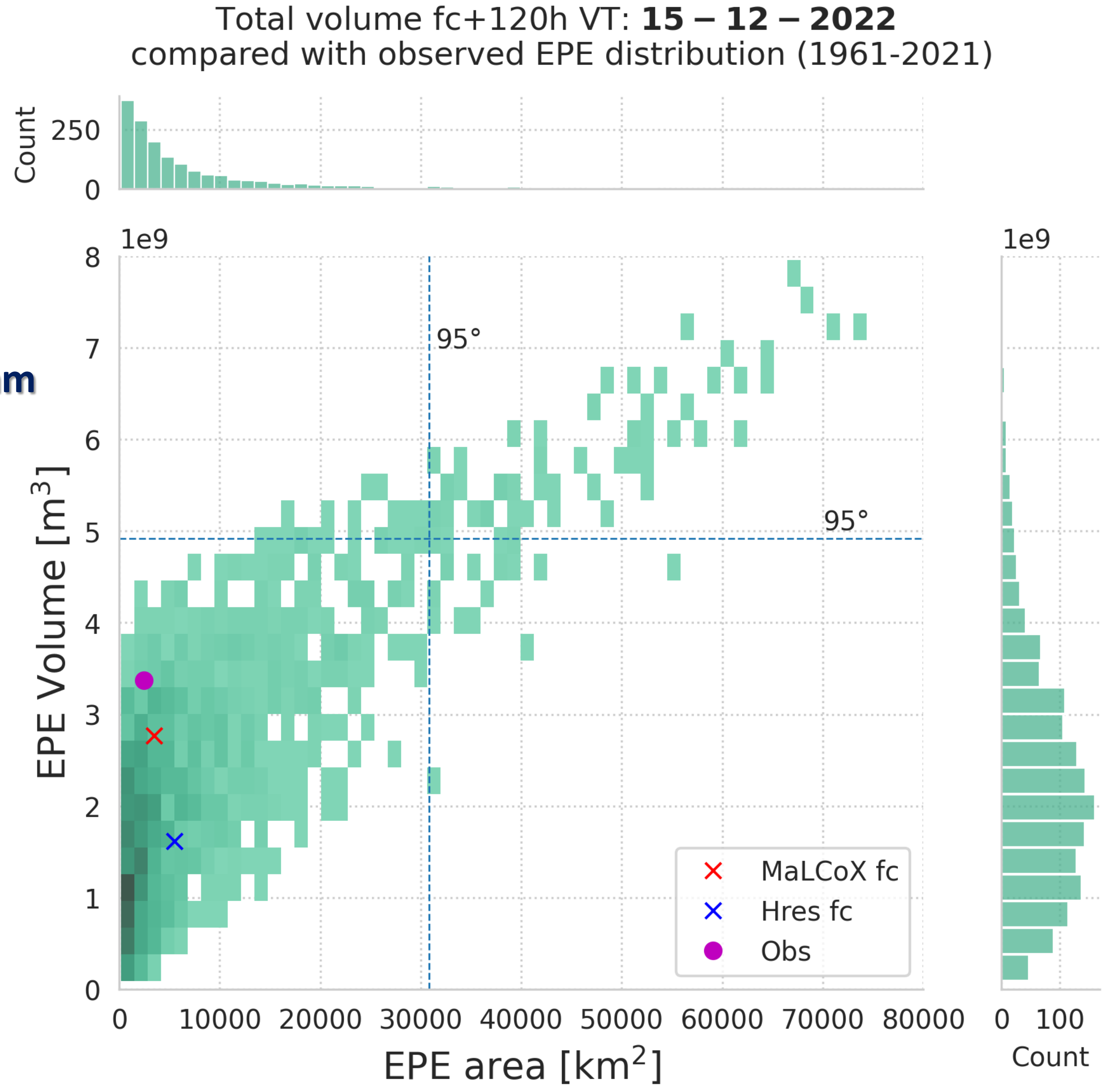






Case studies: Northern Tuscany flood 15/12/2022

Rain volume prediction diagram



Conclusions and outlook

- MaLCoX provides a **complementary way to predict EPE probabilities**. In addition, inform on the atmospheric driver causing the event. ML approach **help to increase trust** in the forecast incorporating past event statistic

- *Expand the training dataset with all (independent) EPS members of ECMWF reforecast*

- *Explore the possibility to build a model for each waning areas or use methods to infer the probability over each WA*

Emilia-Romagna event (II) – 16/17 May 2023

D+4 forecast