Understanding catchment influences on flood generating processes

Accounting for correlated attributes

Lina Stein

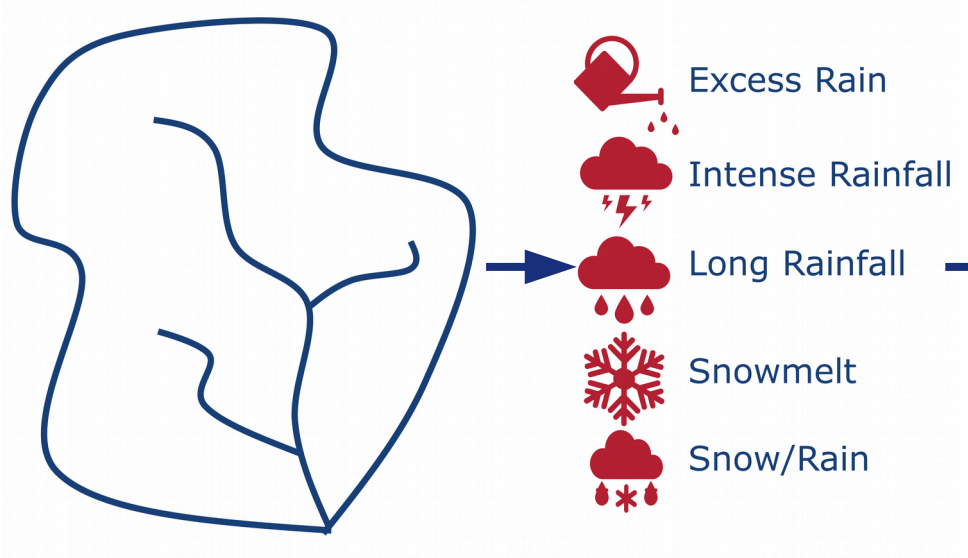
lina.stein@bristol.ac.uk

Martyn Clark

Francesca Pianosi

Wouter Knoben

Ross Woods



University of Bristol

WATER INFORMATICS SCIENCE & ENGINEERING
EPSRC CENTRE FOR DOCTORAL TRAINING

UKRI Engineering and Physical Sciences Research Council

GWF GLOBAL WATER FUTURES
SOLUTIONS TO WATER THREATS IN AN ERA OF GLOBAL CHANGE

University of Saskatchewan

Classified flood generating process for 61,764 flood events using a decision tree.

Read more

Stein, L., Pianosi, F. and Woods, R., 2020. Event-based classification for global study of river flood generating processes. Hydrological Processes, 34(7), pp.1514-1529.

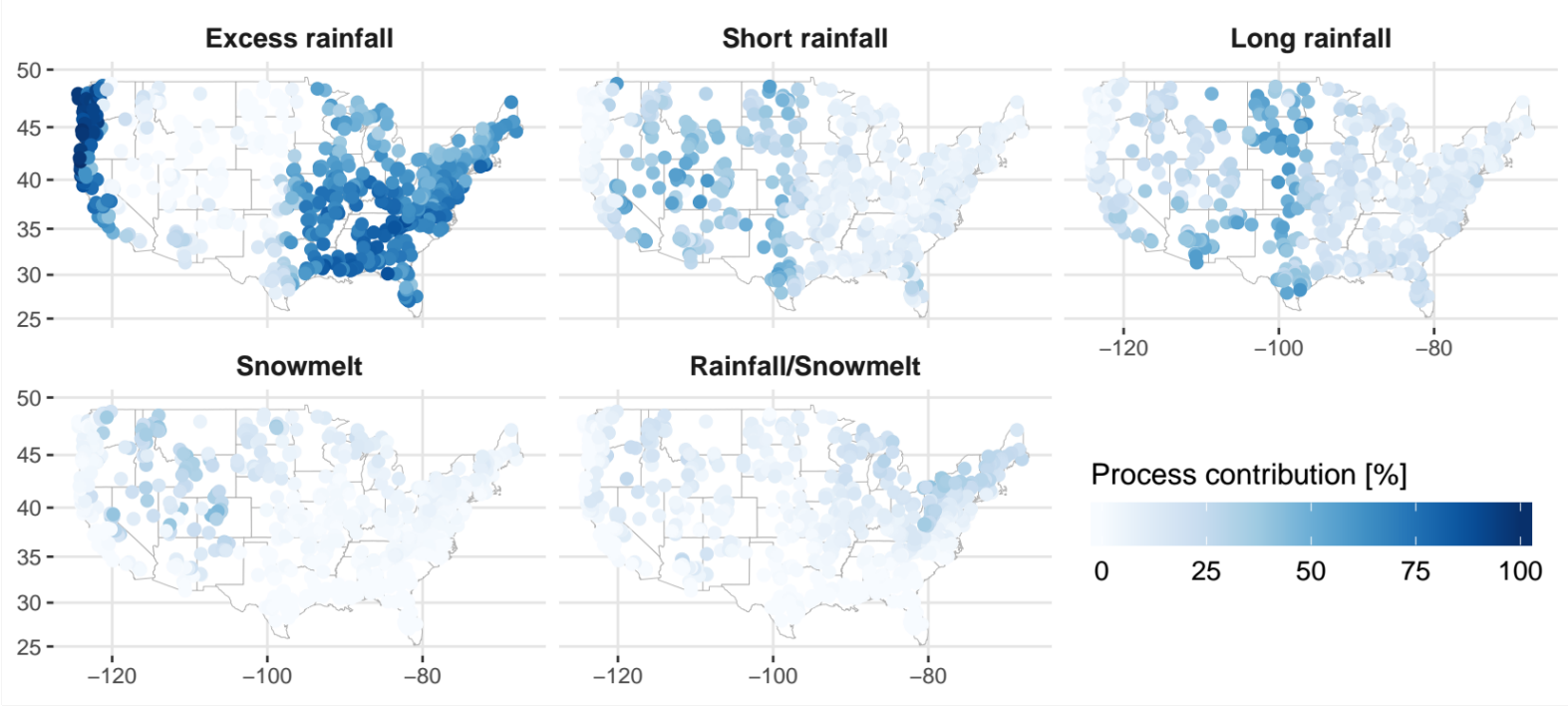


Figure 1: Contribution in percent for each flood generating process. Hydro-climatological input data for the classification taken from the CAMELS catchment dataset (Addor et al, 2017). Flood events are defined as peak-over-threshold with an average of 3 events per year.

Flood events driven by a mix of different processes

➔ Which catchment and climate attributes influence this mix?

Several catchment attributes in the data are highly correlated (Figure 2).

Accumulated local effects (Apley, 2016):

- Interpretable machine learning method
- Not biased toward correlated attributes
- Applied to random forest model
- Determines attribute influence on model prediction

Read more

Molnar, C., 2019. Interpretable machine learning - Accumulated Local Effects (ALE) Plot. Lulu.com.

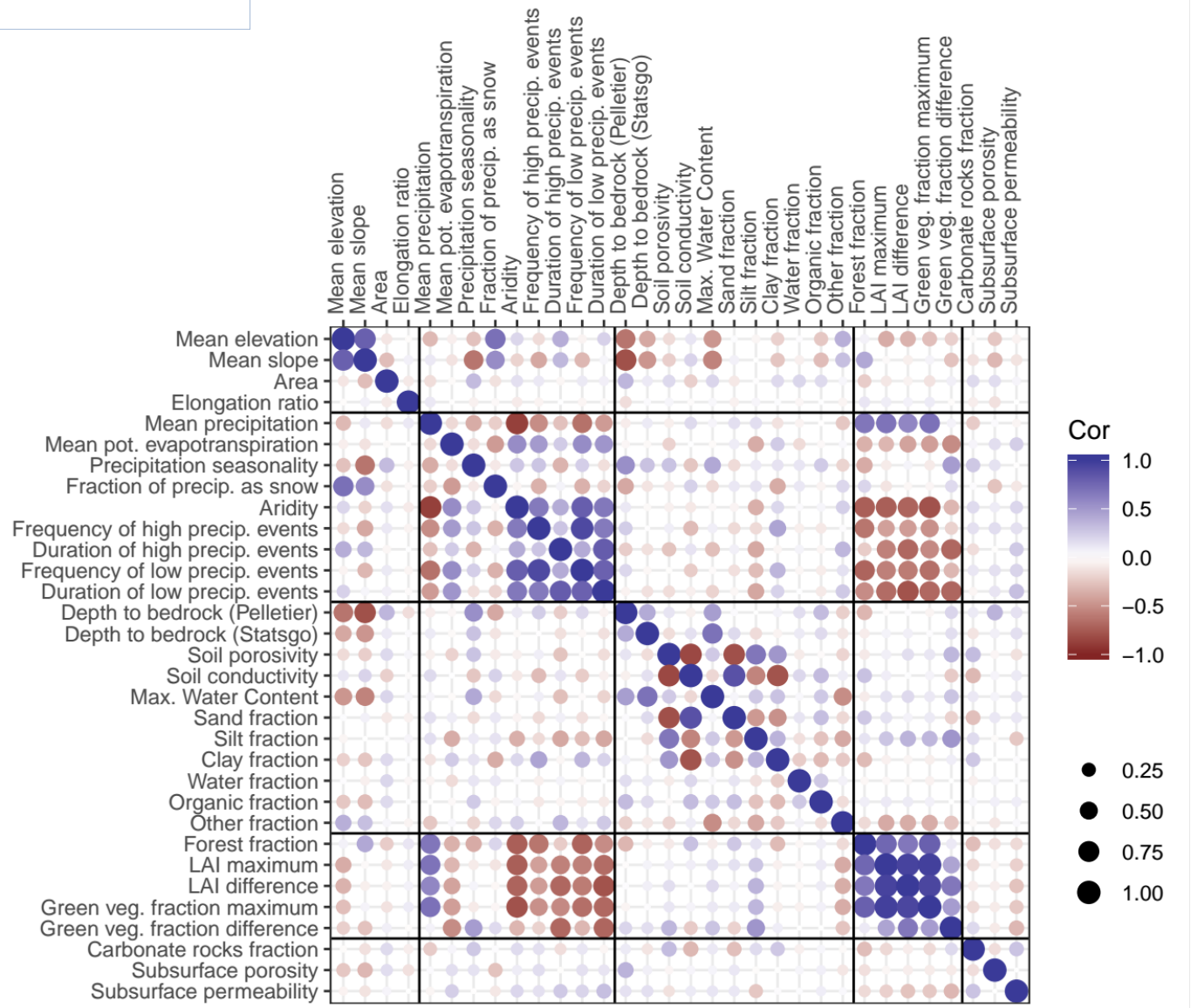


Figure 2: Spearman correlation coefficient for 671 catchments in the CAMELS dataset. Black lines indicate groups (Topology, Climate, Soils, Vegetation, Geology).

University of BRISTOL

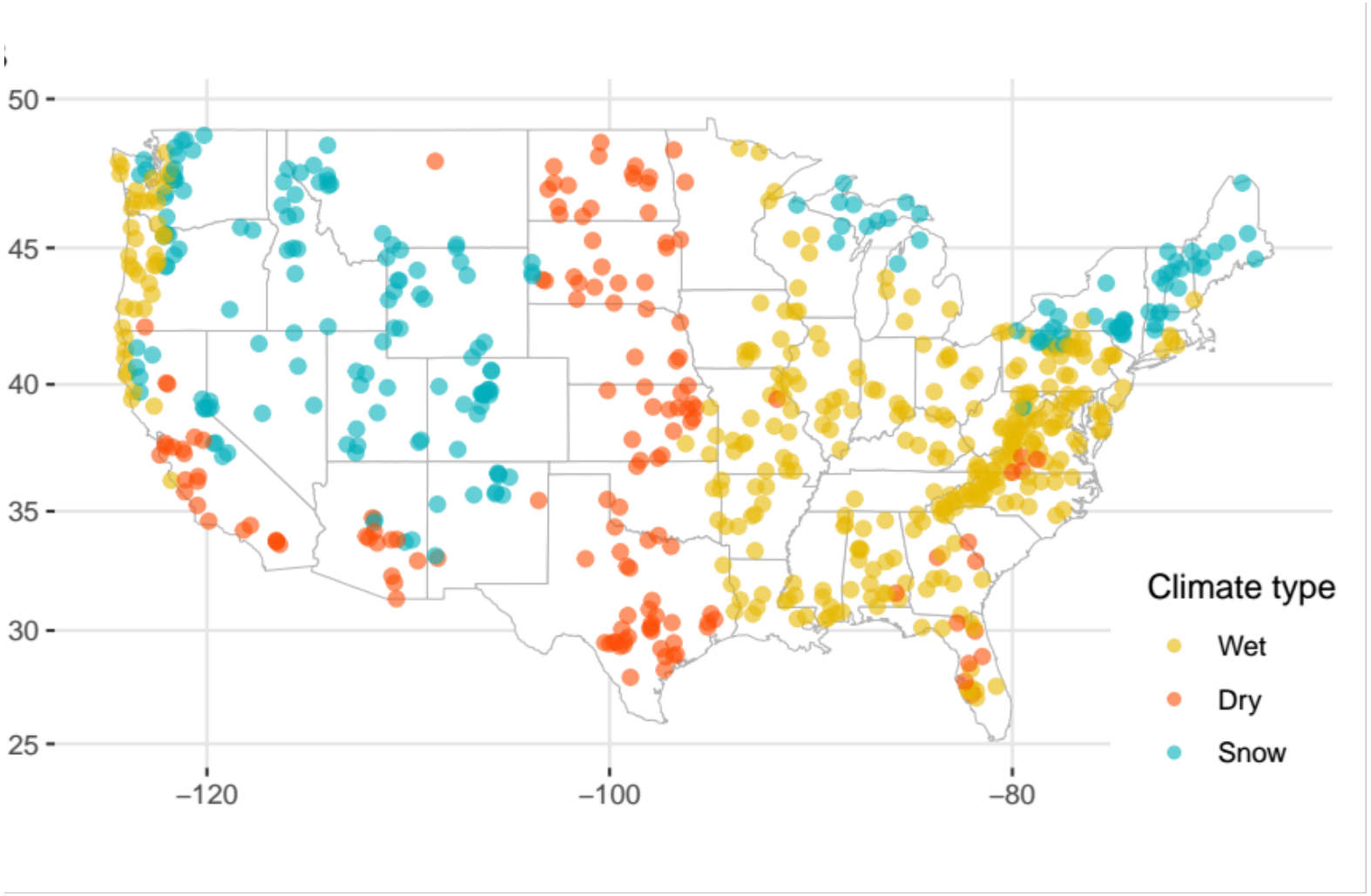


Figure 3: 671 CAMELS catchments split into three different climate types.

We check if attribute importance varies between climates, by dividing the catchments into three climate types based on

- Aridity
- Fraction of precipitation as snow.

- Climate attributes influence flood process distribution most
- For wet and dry catchments:
  - Fraction of precipitation as snow
  - Aridity
  - Mean precipitation
  - Precipitation seasonality
- In snow influenced catchments also:
  - Elevation

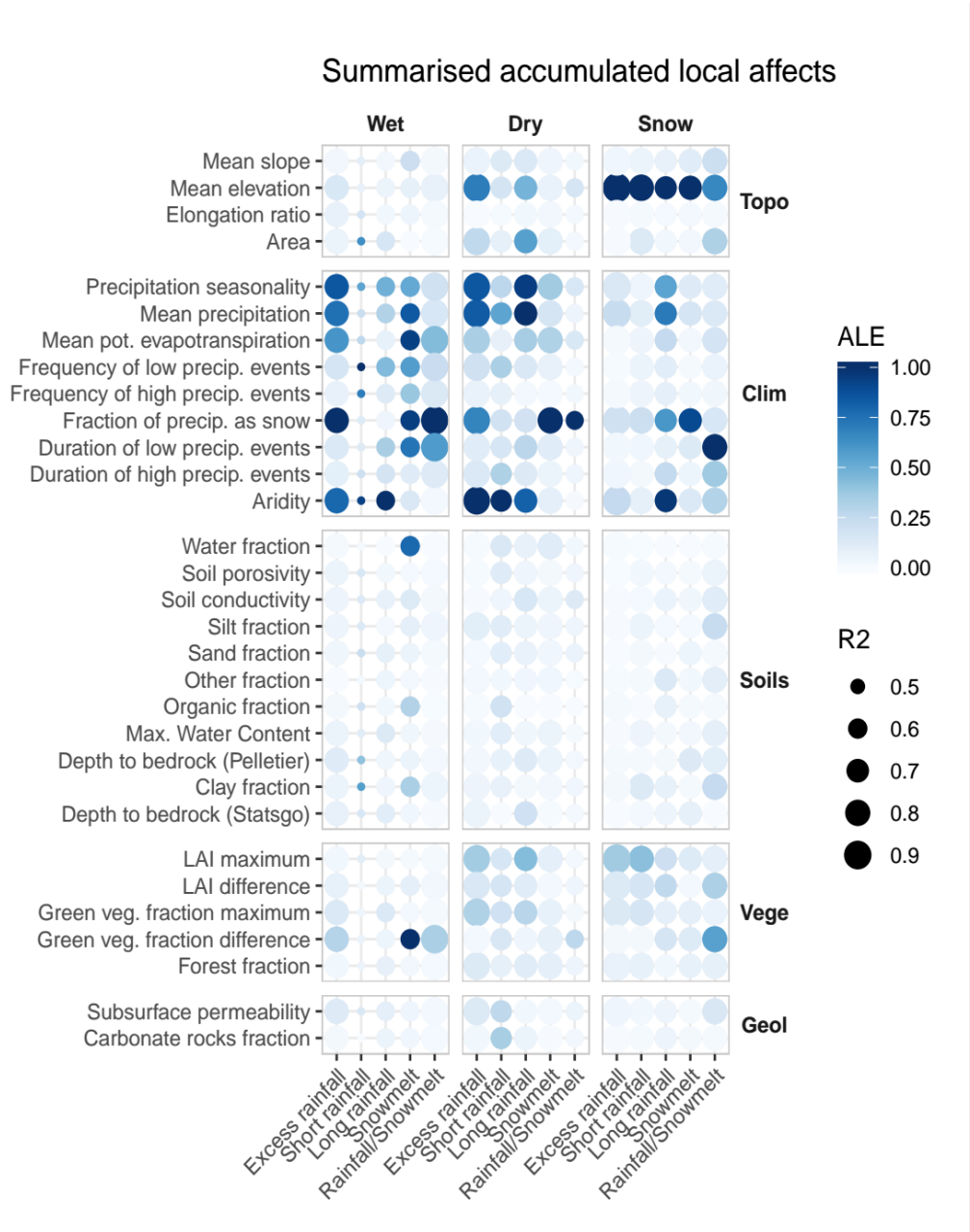


Figure 4: The point colour shows the mean absolute values for each accumulated local effects curve. Higher values indicate increased importance. Point sizes represent cross-validation R2 prediction accuracy for the random forest model.

- Which catchment attributes most influence flood generating processes varies between climate types and processes
- So what? Looking at the impact of climate change on flooding requires us to take the different processes and climate regions into account

Read more soon
Stein, L., Clark, M., Pianosi, F., Knoben, W. and Woods, R., Process oriented insights from interpretable machine learning - what influences flood generating processes?. In preparation.

References
Addor, N., Newman, A.J., Mizukami, N. and Clark, M.P., 2017. The CAMELS data set: catchment attributes and meteorology for large-sample studies. Hydrology and Earth System Sciences (HESS), 21(10), pp.5293-5313.
Apley, D.W., 2016. Visualizing the effects of predictor variables in black box supervised learning models. arXiv preprint arXiv:1612.08468.
Stein, L., Pianosi, F. and Woods, R., 2020. Event-based classification for global study of river flood generating processes. Hydrological Processes, 34(7), pp.1514-1529.

Acknowledgements
WISE CDT is supported by EPSRC grant EP/L016214/1.

University of BRISTOL
WATER INFORMATICS SCIENCE & ENGINEERING
EPSRC CENTRE FOR DOCTORAL TRAINING
UKRI Engineering and Physical Sciences Research Council
GWF GLOBAL WATER FUTURES
SOLUTIONS TO WATER THREATS IN AN ERA OF GLOBAL CHANGE
University of Saskatchewan