Changes in European windstorm characteristics

Christian Passow[(1,*)], Uwe Ulbrich[(1)], Henning Rust[(1)]

(1) Freie Universität Berlin

Research questions (RQ)

1. Can windstorm characteristics be used to meaningfully classify windstorm events?
2. Can we quantify the influence of large- and small-scale atmospheric features on the classified windstorms using modern machine learning approaches?
3. Are there significant trends in windstorm characteristics within the observed period? (not shown)

Data

Climate reanalysis - ERA5 and ERAINT:

- Horizontal resolution: 30km (ERA5), 80km (ERAINT)
- Period: 1981-2017, extended winter ONDJFM
- Temporal resolution: 6 hours
- ERAINT is only used for cyclone tracking!
- Matching following Nissen et al. (2010)

Identification & tracking:

- Windstorm tracking using the storm severity index (Leckebusch et al., 2008)
- Cyclone tracking following Murray and Simmonds (1991)
- Filter: EURO-CORDEX region (Fig. 1)

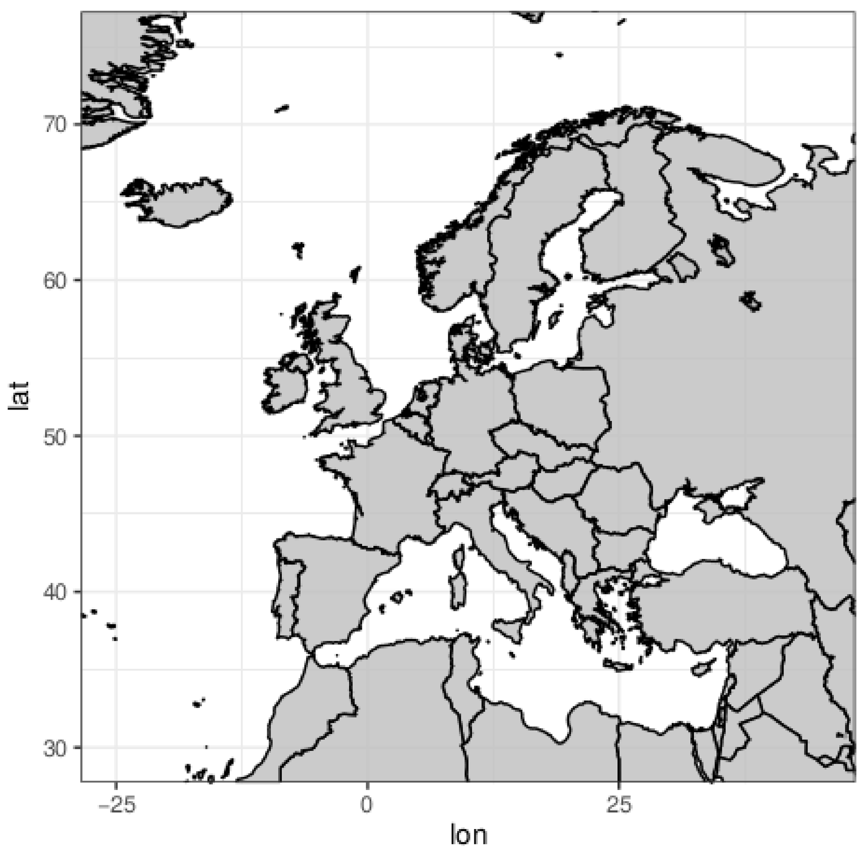

Figure 1: Study region, EURO-CORDEX (40.25°W–75.25°E, 25.25°N–75.75°N)

Summary & conclusion

- Novel semi-supervised clustering method QSKM
- Method produces three windstorm classes of which one shows similar characteristics than the events from the XWS catalog
- Size and duration appear to play a greater role than wind speed in severe events
- Resulting classes can be described through a statistical model out of large scale predictors
- Although large-scale patterns can influence the occurrence of storms, they only have a small impact in our model
- The jet stream and the location and depth of the parent cyclone are determined for the type of event

Outlook

- Dependences of individual windstorm characteristics
- Trends for clusters and characteristics
- Temporal structure and sequence of events

RQ1: Classifying windstorms

Quasi-supervised k-means (QSKM) clustering

- Novel clustering technique that partitions data w.r.t a reference (see Alg. 1)
  - + Allows for feature selection
  - + Produces relevant clusters
  - + Includes automatized processes for determining the optimal No. of cluster
- Reference is the XWS open access catalogue of extreme European windstorms from 1979 to 2012 (Roberts et al., 2014)
- Cluster No. 1 has multiple features close to the reference (Fig. 2)
- Most severe events are found over central Europe (Fig. 3b)
- Key characteristics are size, duration and travel speed/distance (Fig. 3a)

Algorithm 1 Quasi-Supervised k-Means

1. Draw J subsets (with replacement) $S_1^{n\times q}, \ldots, S_J^{n\times q}$ from $X^{n\times p}$ such that $q < p$ and $S_1 \neq \ldots \neq S_J$.
2. For $j = 1 \rightarrow J$ do:
   - Enhanced k-means on $S_j \Rightarrow C_{j,1}, \ldots, C_{j,k}$
   - Calculate $\gamma_j = \max\left[\mathrm{count}(C_{j,1} \in R), \ldots, \mathrm{count}(C_{j,k} \in R)\right]$
   - Calculate $\delta_j = \mathrm{dist}(C_{j,p}, R)$ with $C_{j,p}$ being the cluster where $\mathrm{count}(C_{j,p} \in R) = \gamma_j$
3. Select clustering result with the smallest δ_j from the S_j with the largest γ_j

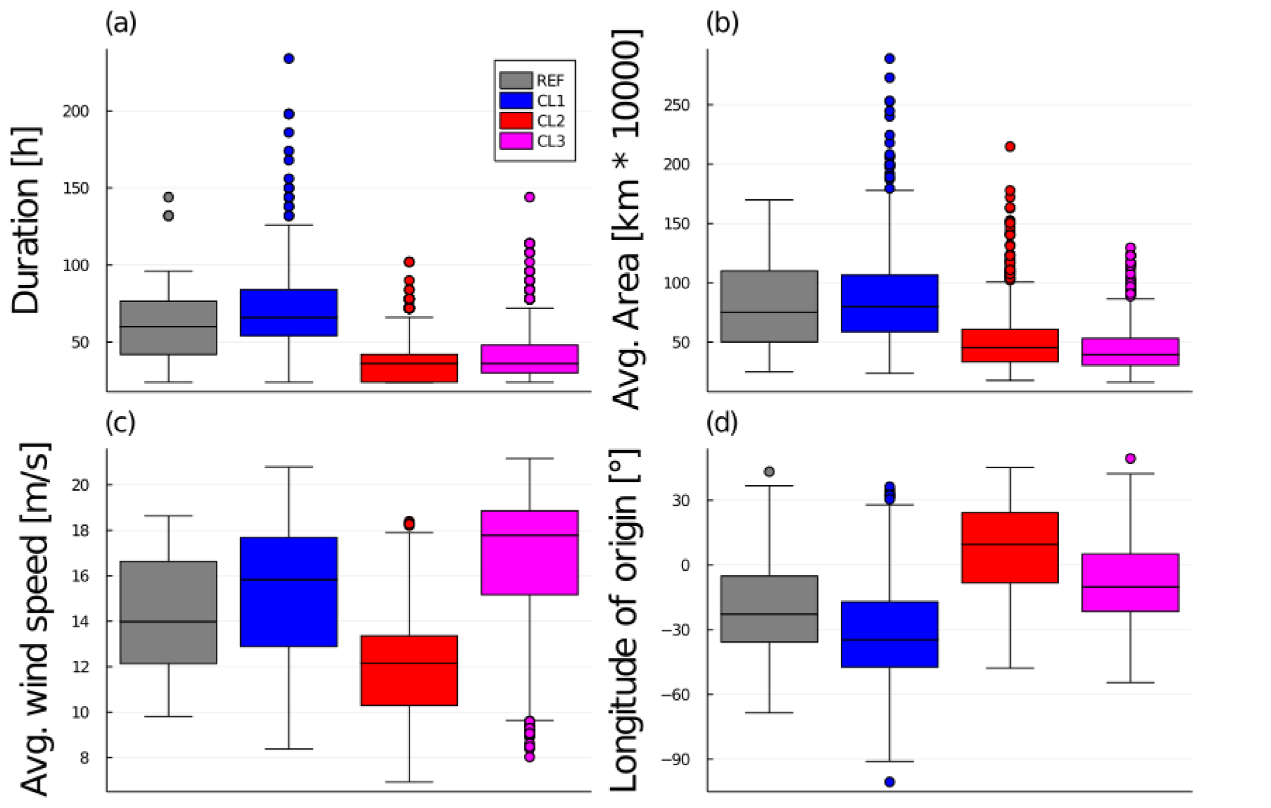

Figure 2: Comparison of the cluster against the reference for four different characteristics.

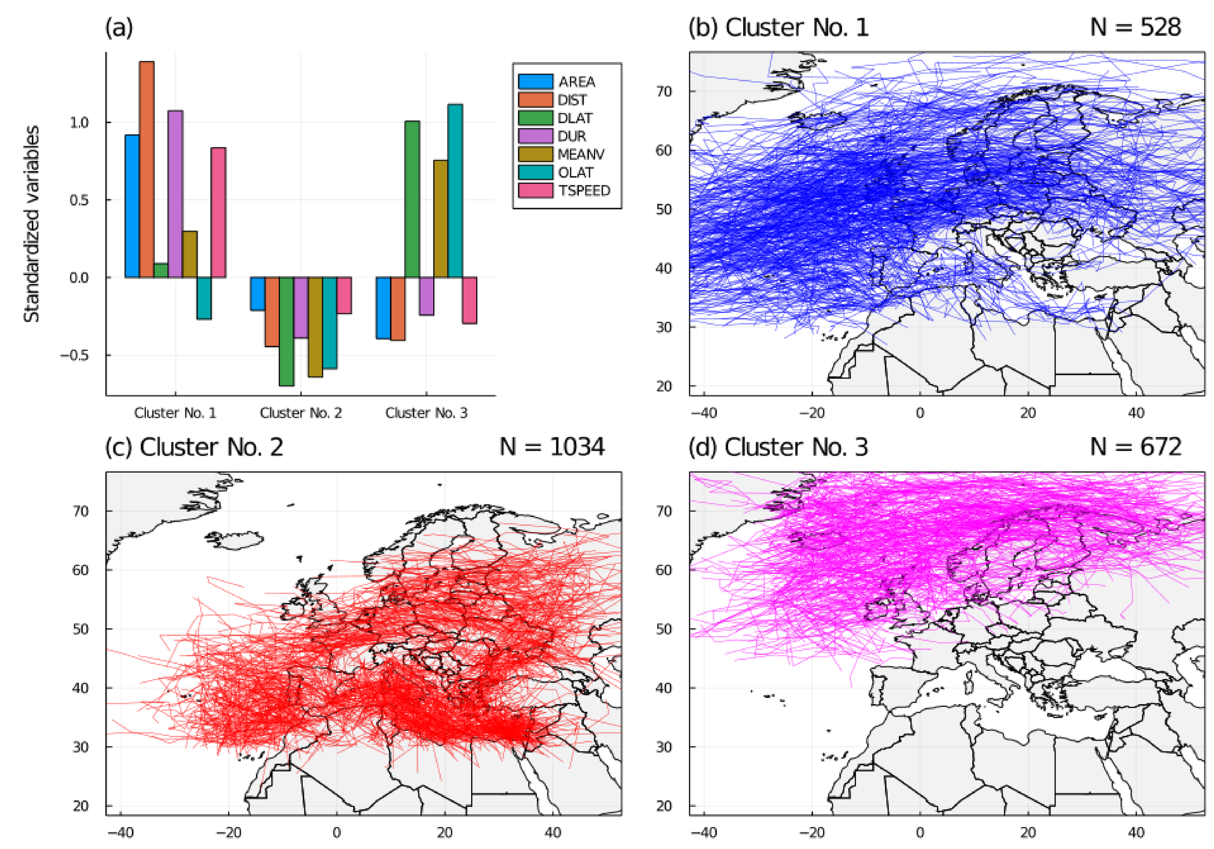

Figure 3: Clustering result after using QSKM where (a) shows the standardized characteristics and (b-d) the tracks for each cluster. Cluster No. 1 is closest to the reference.

RQ2: Quantifying influences

- Random forest model to fit cluster against various predictors (see variable names in Fig. 4)
- The model shows high skill for Cluster No. 2 and 3, but struggles to predict Cluster No. 1 (Tab. 1)
- Evaluation of the feature importance and dependence using the Shapley value proposed by Shapley (1953)
- Latitude of minimum core pressure and strength of the jet stream are the two most important features (Fig. 4)
- The stronger the jet, the more likely a event is from Cluster No. 1 (Fig. 5, top left)
- Although core pressure is important, the relationship to the cluster is confuse (Fig. 5, bottom left)

Table 1: Confusion matrix for the trained random forests model.

	Ground Truth		
Predicted	1	2	3
1	64	15	15
2	23	254	12
3	30	18	135

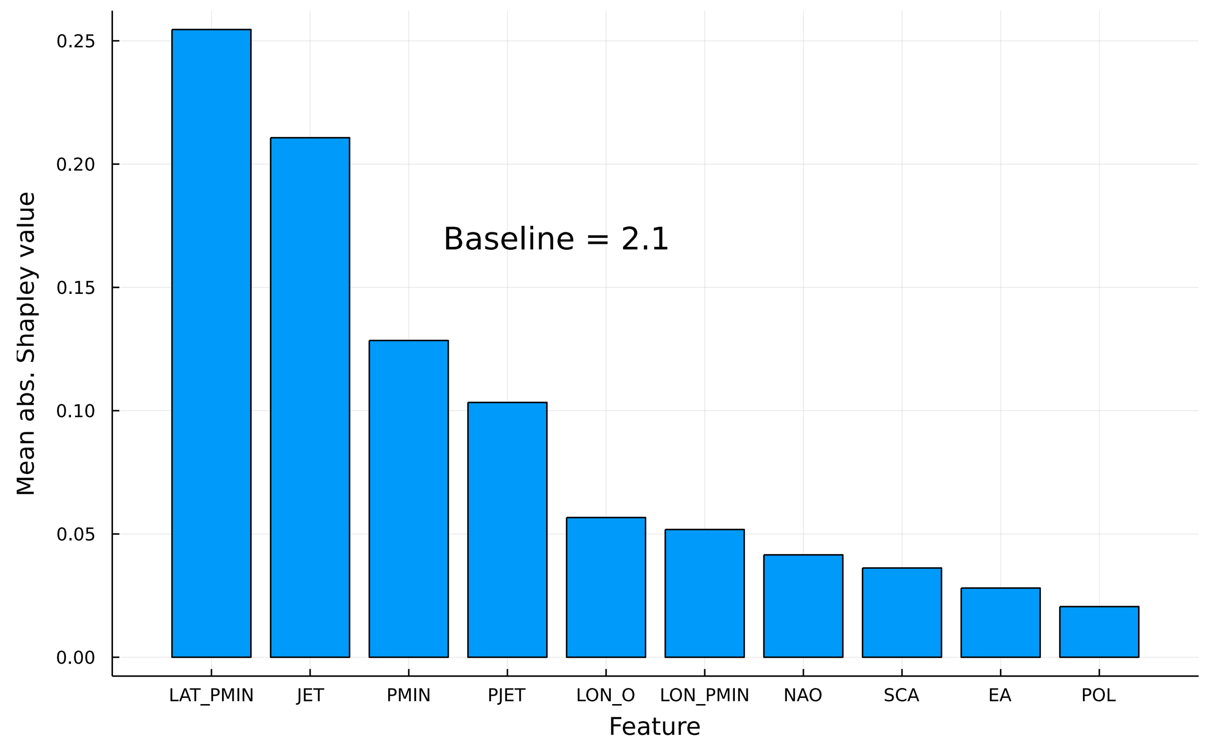

Figure 4: Feature importance derived from the random forests model using Shapley values. The base line is the average prediction of the model (here Cluster No. 2).

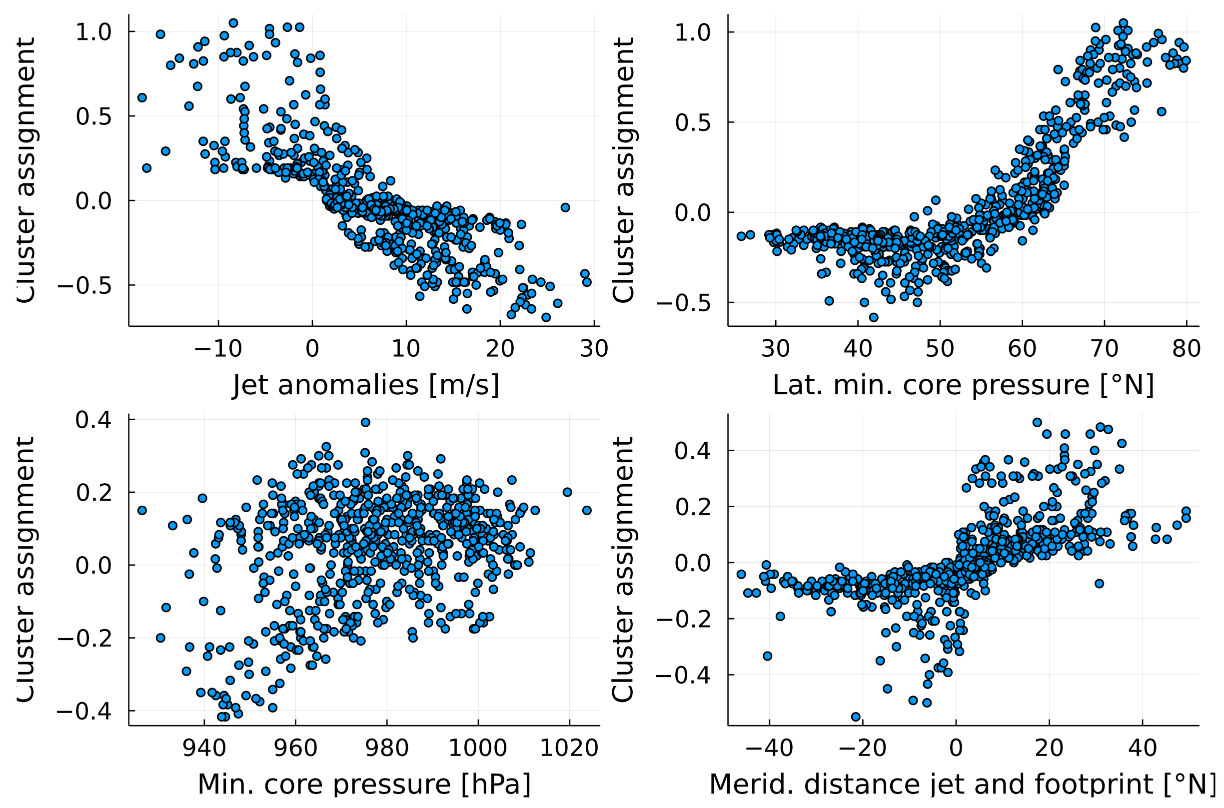

Figure 5: Feature dependence derived from the random forests model using Shapley values. Cluster assignments are added to the baseline of 2.1.

References

Leckebusch, Gregor C. et al. (2008). "Development and application of an objective storm severity measure for the Northeast Atlantic region". In: *Meteorol. Zeitschrift* 17.5, pp. 575–587.
Murray, Ross J. and Ian Simmonds (1991). "A numerical scheme for tracking cyclone centres from digital data. Part I: development and operation of the scheme". In: *Aust. Meteorol. Mag.* 39.3, pp. 155–166.
Nissen, K. M. et al. (2010). "Cyclones causing wind storms in the Mediterranean: characteristics, trends and links to large-scale patterns". In: *Nat. Hazards Earth Syst. Sci.* 10.7, pp. 1379–1391.
Roberts, J. F. et al. (2014). "The XWS open access catalogue of extreme European windstorms from 1979 to 2012". In: *Nat. Hazards Earth Syst. Sci.* 14.9, pp. 2487–2501.
Shapley, L. S. (1953). "17. A Value for n-Person Games". In: *Contrib. to Theory Games (AM-28), Vol. II*. Princeton University Press, pp. 307–318.

*christian.passow@fu-berlin.de