Supplement of Weather Clim. Dynam., 7, 223–245, 2026
https://doi.org/10.5194/wcd-7-223-2026-supplement

Supplement of

Linking European droughts to year-round weather regimes

Onaïa Savary et al.

Correspondence to: Onaïa Savary (onaia.savary@meteo.fr)

S1 : Choosing the optimal number of weather regimes

Building on the methodology proposed by [Lee et al., 2023], we use four metrics to determine the optimal number of weather regimes. These four metrics are chosen according to the following criteria:

- Limiting the number of weather regimes for the sake of simplicity and robustness
- Minimizing the distance of each point of a given cluster to the centroid of this cluster
- Maximizing the distance between each centroid

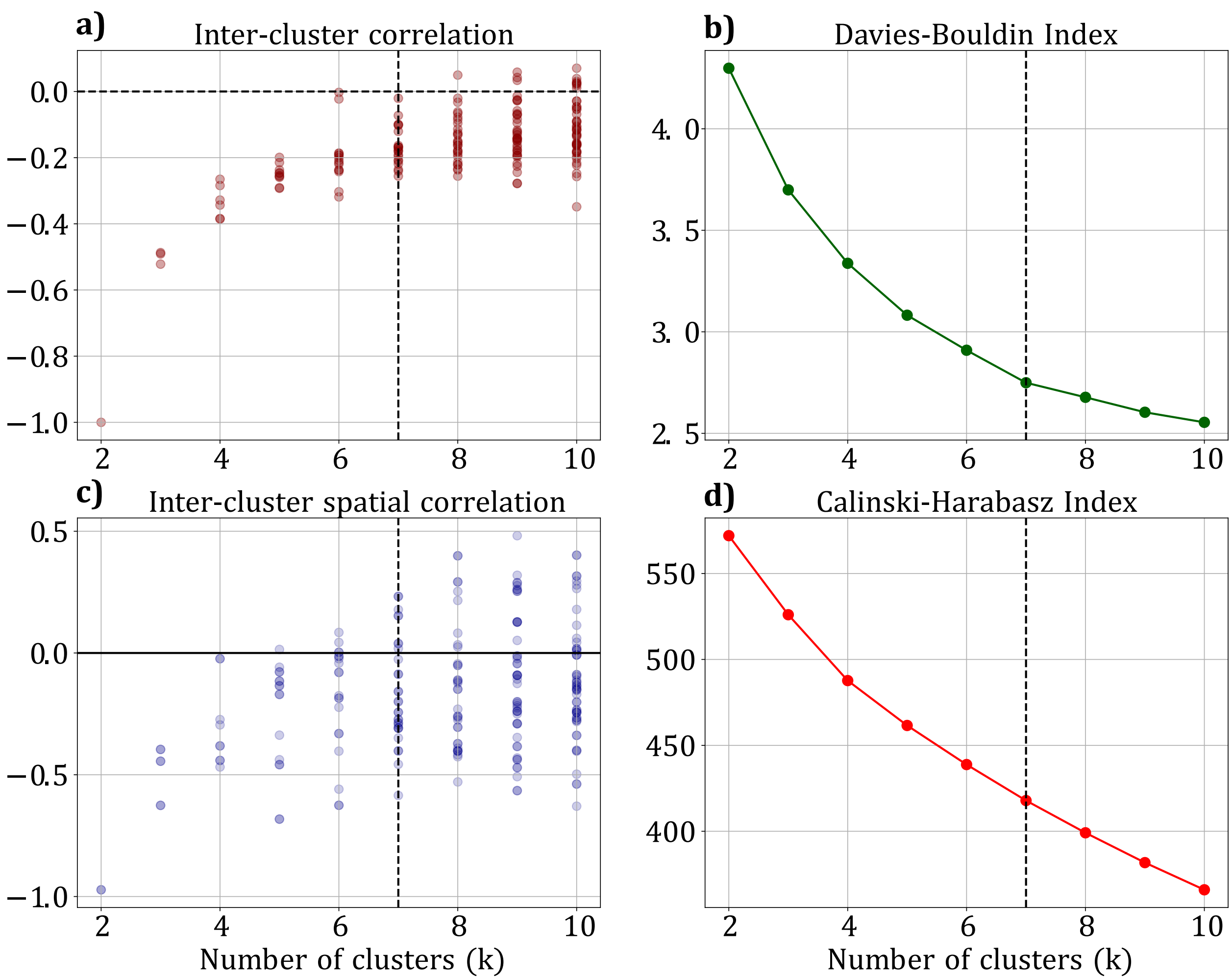

Fig. S1: (a) Intercluster correlation: arithmetic distance between the centroids of each cluster, calculated in the PCs space. (b) Davies–Bouldin index, expressing the ratio of a point from a cluster to the centroid of this cluster over the distance of this centroid to other centroids. This metric is averaged for all clusters. (c) Intercluster spatial correlation: anomaly correlation coefficient between all centroids, calculated in the EOF space. (d) Calinski–Harabasz index: ratio between intercluster variance and intracluster variance.

For each number of clusters, we compute the following indices:

- **Intercluster correlation**: we compute the Pearson correlation coefficient between the centroid coordinates in PC space for each pair of centroids. We aim to get the maximal number of clusters such that the Pearson correlation coefficient is negative, hence all clusters are anticorrelated. This method leads to $k = 7$.

- **Davies–Bouldin index**: proposed by [Davies and Bouldin, 1979], the Davies–Bouldin index is computed as

$$DB(k) = \frac{1}{k}\sum_{i=1}^{k}\max_{i\neq j}\left(\frac{s_i + s_j}{d_{ij}}\right) \tag{S1}$$

  where s_i is the average distance between members of cluster i and the centroid of that cluster, and d_{ij} is the distance between the centroids of clusters i and j. This metric is to be minimized. The computation of this metric does not lead to a clear choice of the number of clusters, but using the elbow method, the optimal number appears to be around 7.

- **Intercluster spatial correlation**: the Pearson correlation coefficient computed above does not take into account the spatial patterns of the weather regimes. We compute the anomaly correlation coefficient (ACC) between each pair of weather regimes for each number of clusters. We require negative ACC values to ensure anticorrelated patterns. This metric suggests an optimal number closer to $k = 4$.

- **Calinski–Harabasz index**: expressed as the ratio of intercluster variance over intracluster variance, this index should be maximized. No clear optimum appears here, as the index decreases monotonically with increasing k.

In conclusion, the optimal number of weather regimes is less clear than for seasonal regimes ($k = 4$). The metrics used here do not rule out the $k = 7$ option, which represents the best compromise between the number of clusters and the criteria defined above.

S2

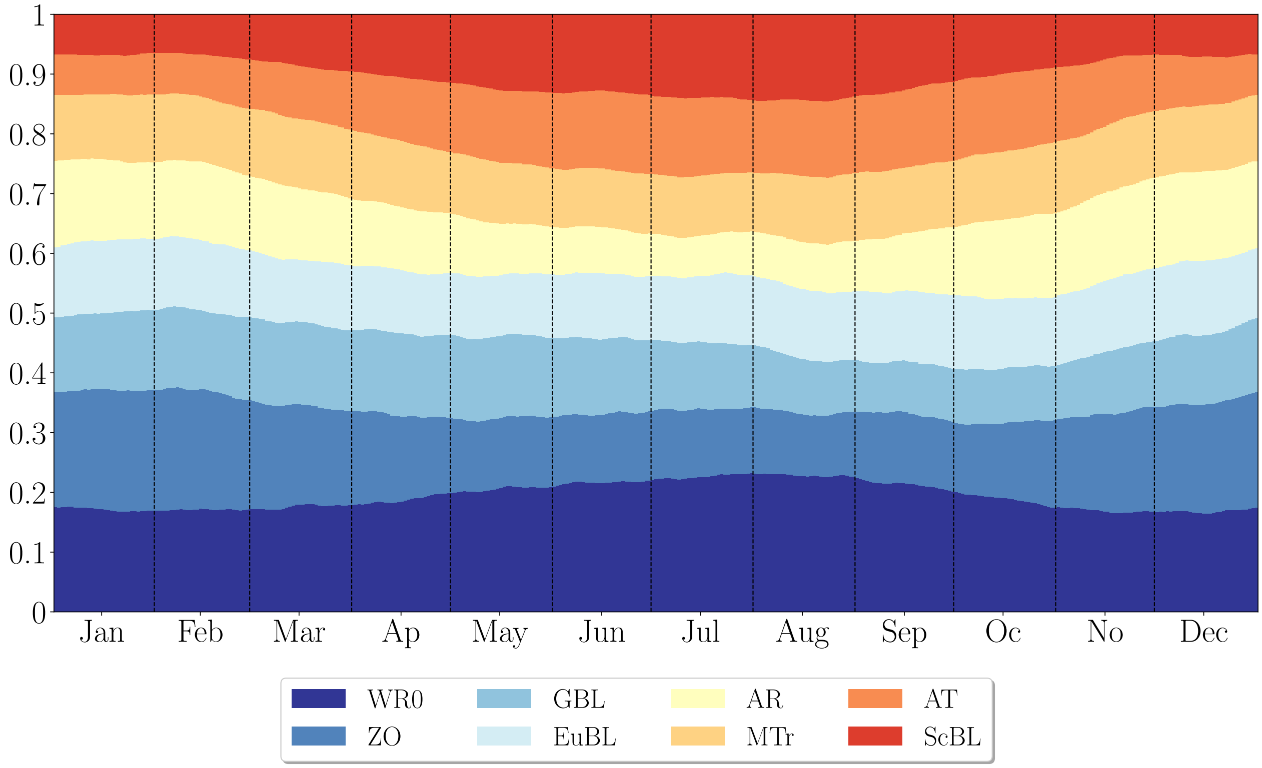

Fig. S2: For each day of the year, average frequency of weather regimes computed over the period 1960–2022.

S3

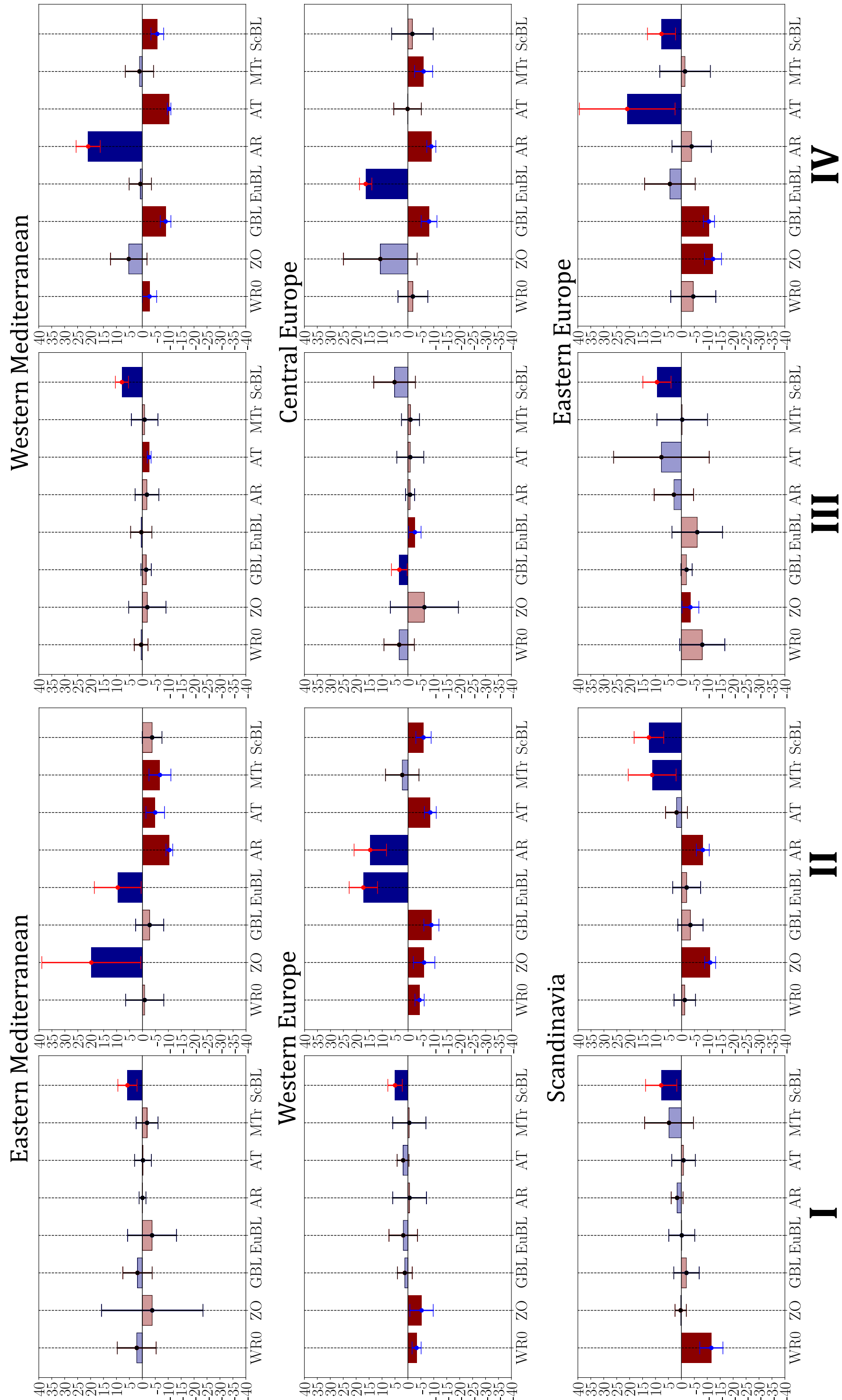

Fig. S3: Frequency anomalies of weather regimes for droughts that cannot be explained by the WR approach (columns I and III) and those that can be explained (columns II and IV).

S4 : Evolution of the number of days between two successive droughts

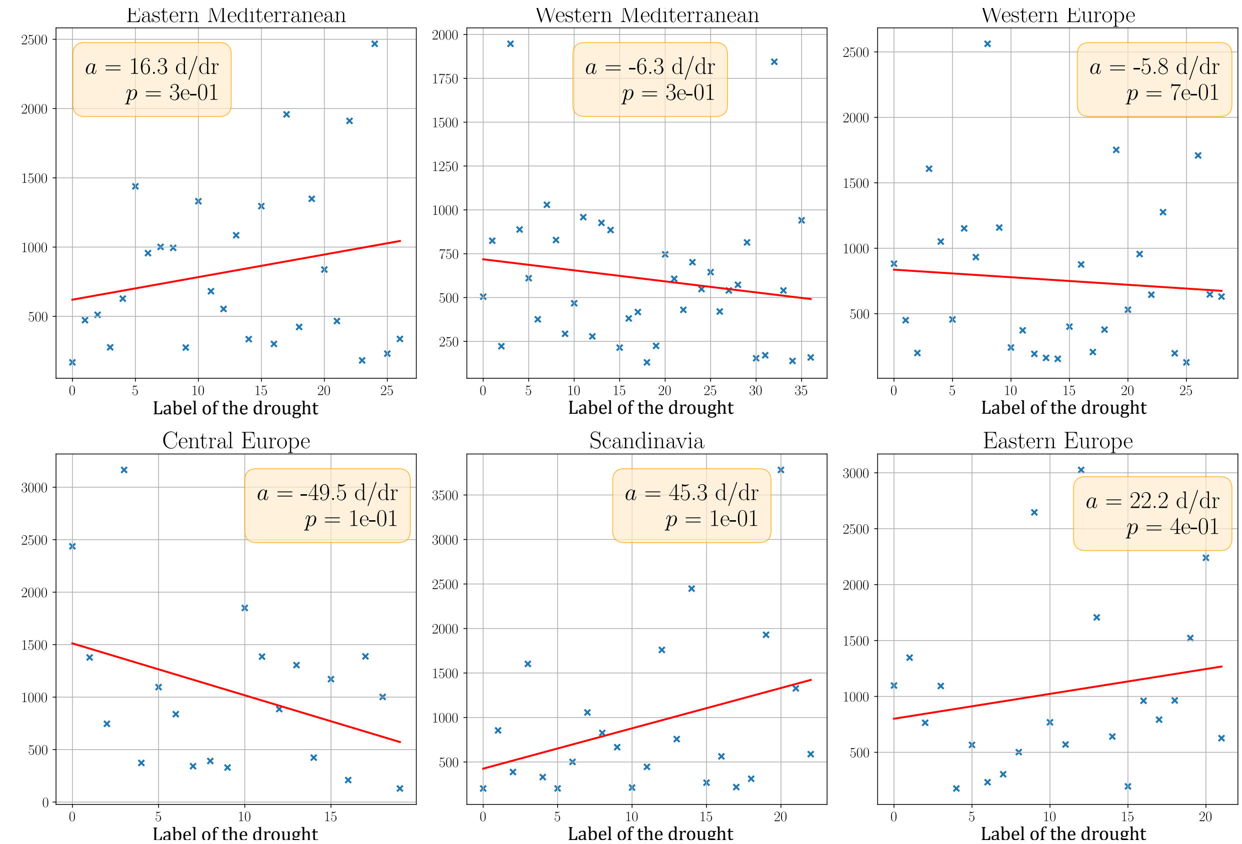

Fig. S4: Evolution of the number of days between two successive droughts for each region. Linear trends obtained by regression are shown in red.

S5 : Representativeness of precipitation and Z500 anomalies by WRs

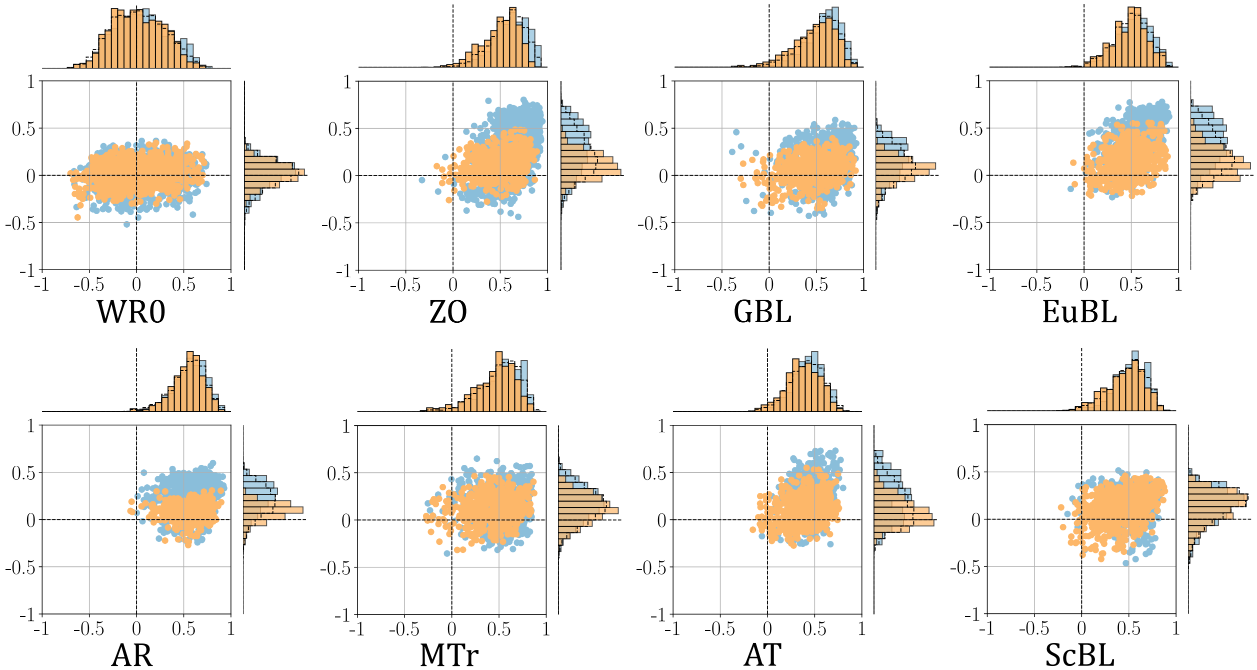

Fig. S5: For each WR: scatterplots (dots) and distributions (bars) of daily anomaly correlation coefficients (ACCs) between observed fields and the centroids of their assigned weather regimes, for precipitation (y-axis) and Z500 (x-axis), in DJF (light blue) and JJA (orange). Days assigned to WR0 are excluded. Dotted lines show the distributions for all days.

S6 : Role of persistence and intensity of individual WR life cycles

We distinguish between the size and the number of sequences for each regime (Fig. S6). In most cases, frequency anomalies cannot be explained by changes in persistence alone or in the number of sequences alone, but rather by a combination of both.

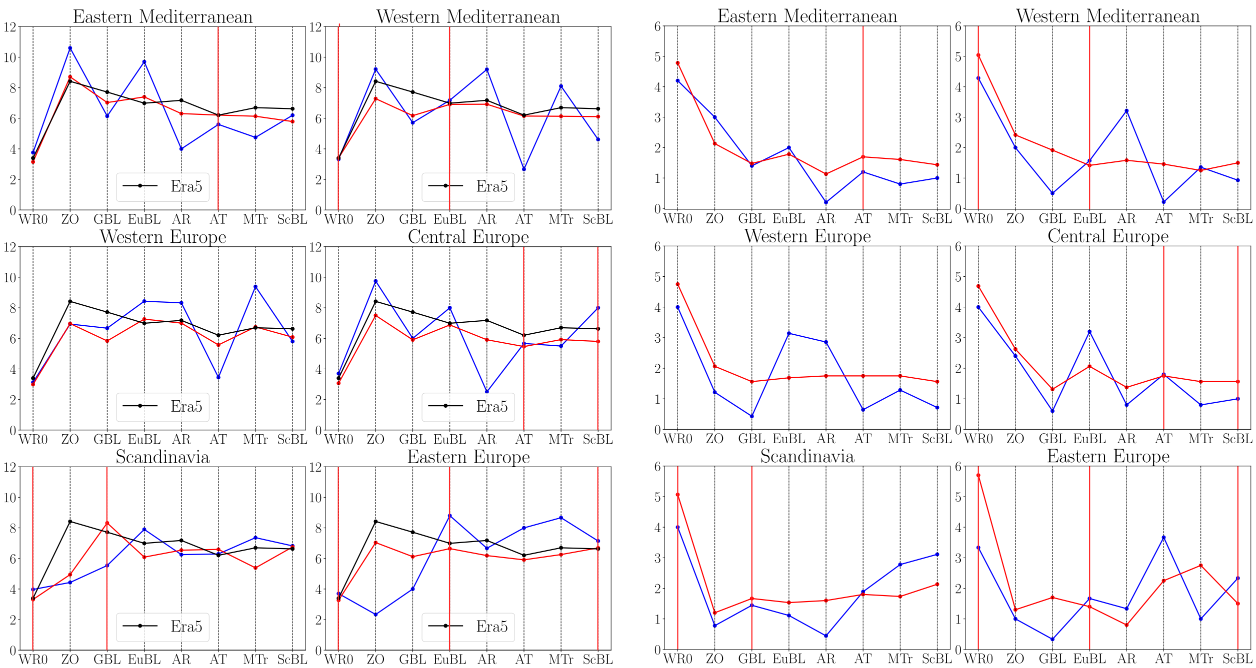

Fig. S6: [LEFT] Mean duration of WR sequences (in days) for droughts explained by WRs (blue), droughts not explained by WRs (red), and the full ERA5 period (black). [RIGHT] Mean number of WR sequences. WRs with non-significant frequency anomalies are marked by a vertical red line.

We also present Z500 composites for each regime under drought conditions in the WMed region, distinguishing between situations well explained by the WR approach and those that are not (Fig. S7).

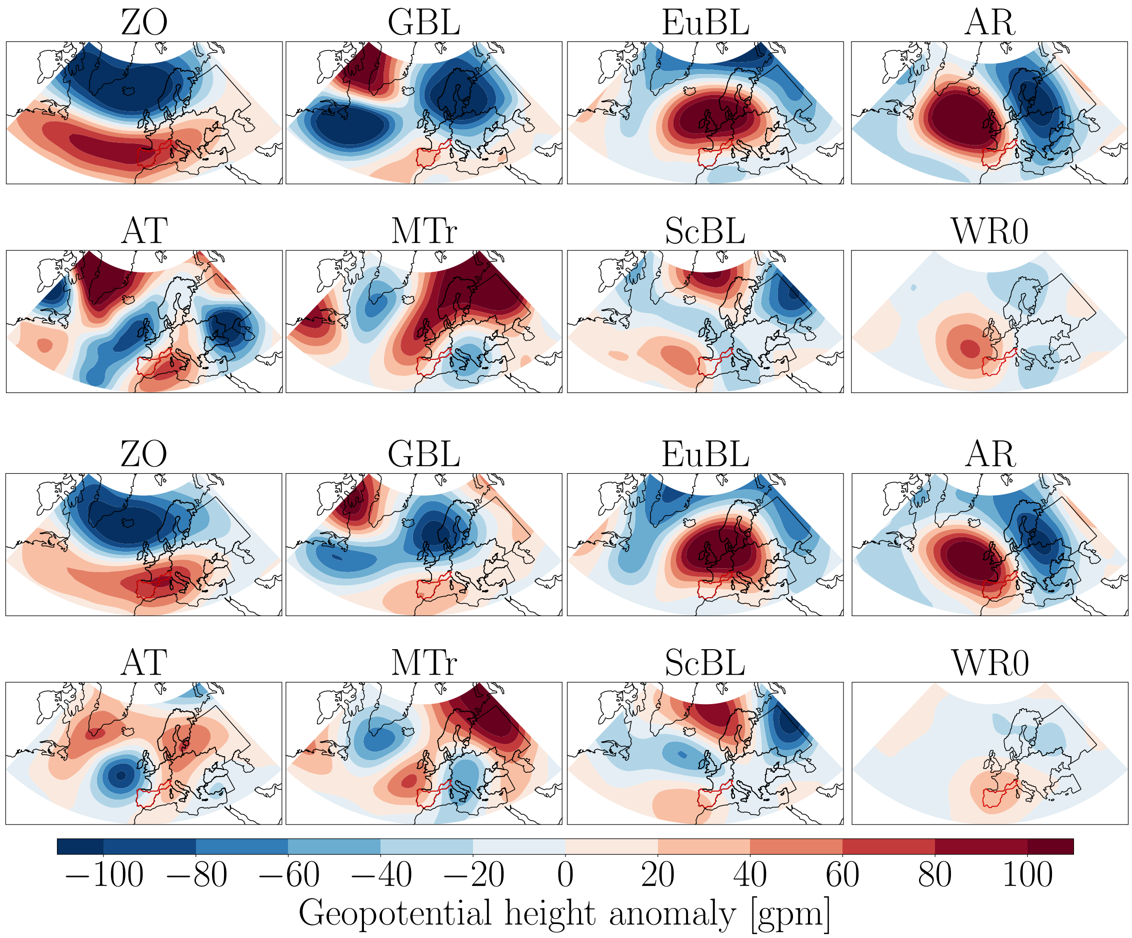

Fig. S7: Z500 patterns for situations well explained by the WR approach (first two rows) and poorly explained (rows three and four).

References

David L. Davies and Donald W. Bouldin. A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2):224–227, April 1979. ISSN 0162-8828, 2160-9292. doi: 10.1109/TPAMI.1979.4766909. URL http://ieeexplore.ieee.org/document/4766909/.

Simon H. Lee, Michael K. Tippett, and Lorenzo M. Polvani. A new year-round weather regime classification for north america. 36(20):7091–7108, 2023. ISSN 0894-8755, 1520-0442. doi: 10.1175/JCLI-D-23-0214.1. URL https://journals.ametsoc.org/view/journals/clim/36/20/JCLI-D-23-0214.1.xml.