Studying the dynamics of weather regimes and their causal influences using a model-based clustering approach.

To study continental-scale weather, the concept of weather regimes has been commonly used for decades. Weather regimes correspond to recurrent and quasi-persistent large-scale atmospheric patterns that typically have a strong impact on surface weather. Once the recurrent large-scale states of the atmosphere are identified, one can study their characteristics, their occurrences, their impacts on local meteorology, or their persistence to gain information on atmospheric dynamics. The goal of this project is to use a data-driven approach to identify dynamical drivers of systematic changes between weather regimes, using statistical causality estimation combined with clustering.

The identification of weather regimes typically relies on statistical methods such as k-means clustering or hierarchical clustering of geopotential or zonal wind fields.
Such approaches allow to identify regimes in different zones of the globe, such as the North Atlantic Oscillation (NAO) in North Atlantic, the Madden-Julian Oscillation in the tropics or the El-Nino Southern Oscillation in the Pacific Ocean.

Such geometrical clustering methods are powerful to identify patterns in the data, but do not directly give dynamical information such as the persistence of the weather regimes or about the transitions between them. Instead, one has to resort to postprocessing to extract dynamical knowledge.

In this work the clustering approach is supplemented by a dynamical model structure, leading to dynamical, model-based clustering. The chosen approach uses the finite element clustering with bounded variation (FEM-BV) framework supplemented by a statistical vector autoregressive model with exogenous variables (VARX) that describes the dynamics within each cluster state. The resulting FEM-BV-VARX method, on top of identifying persistent states, allows 2 further aspects:

- With Linear Vector AutoRegressive (VAR) processes, the model can provide information on the persistence of the states and on the dynamics within and between those states. This was studied in depth applied to the North Atlantic Oscillation by Quinn et al., 2021
- The model can take into account, on top of input fields, exogenous (X) variables. These are variables that we suspect to have an impact on the dynamics of weather regimes, and the method enables the quantification of statistical causality (in the sense of Granger causality).

The FEM-BV-VARX method will be applied on ERA5 reanalysis data. Exogenous variables available in the ERA5 reanalysis will be included to quantify causal relationships. Tested variables will include sea surface temperature and stratospheric ozone, but also variables that incorporate spatial information such as the Rossby waves pattern, the NAO index and the stratospheric polar vortex strength.

Computing power and storage capacity will hence be used to test as many exogenous variables as possible on different regions of the globe. Those results will then be processed to extract meteorological interpretations on the regimes, on their precursors, dynamics or persistence.