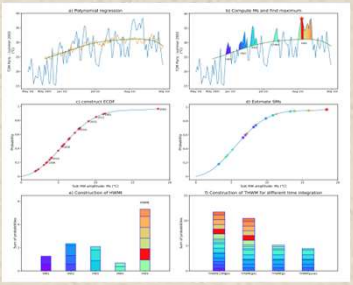
Hybrid modeling for seasonal forecast of European heat wave propensity

S. Materia[1], M. Donat[1], M. Jung[2], C. Gomez-Gonzalez[1]

- Dynamical seasonal prediction skill of **heat waves** is still low in mid-latitudes, and with particular regard in Europe where a large part of the predictability is associated with the global warming trend.
- Statistical techniques, also based on **machine learning** algorithms, have recently demonstrated that **improving dynamical climate prediction** is possible.
- Land drivers of seasonal predictability, such as **soil moisture, snow cover**, etc. can enhance seasonal prediction, but have often been neglected and their interaction with the atmosphere is poorly modeled.

Heat wave propensity

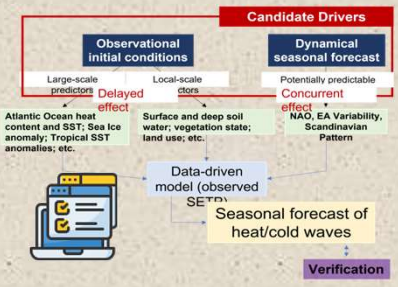


Exact **location, duration and intensity** unpredictable at the **seasonal** scale. Forecasts may have skill for seasonal ***heat wave propensity***

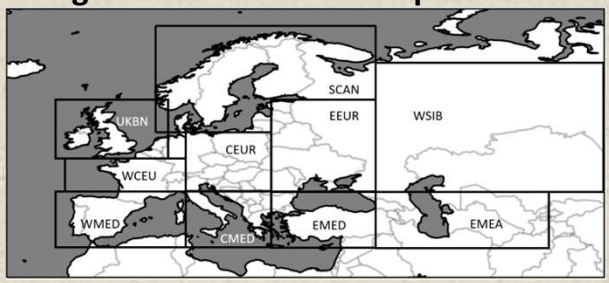
Feature selection approach and workflow

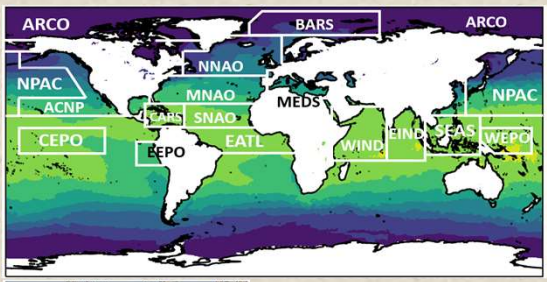
From a list of candidate precursors, a feature selection approach (Guided Hybrid Genetic Algorithm, GHGA, Jung and Zschleischer, 2013) will be used to identify the best variable subset for the prediction of heat wave propensity.

GHGA is wrapped around a Random Forest, that repeatedly works with a different variable subset to minimize a cost function.



Target area and candidate predictors





soil moisture (3 levels)

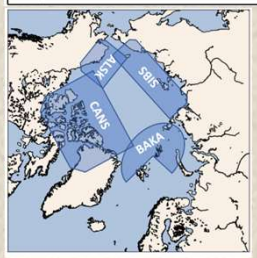
soil temperature (3 levels)

solar radiation

snow water equivalent

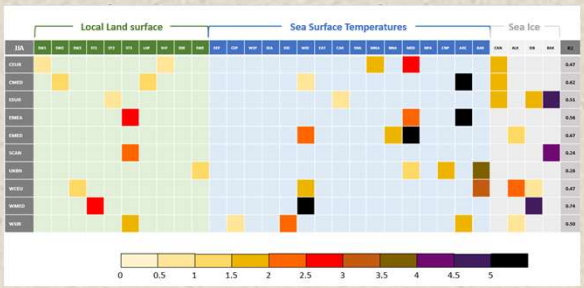
surface heat fluxes

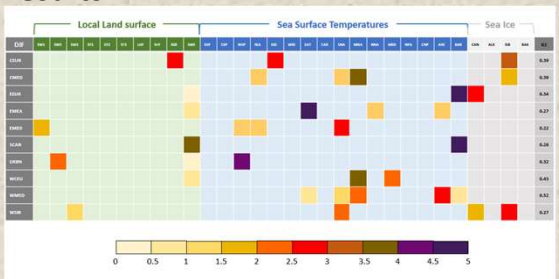
Land surface candidate predictors are local, meaning we assume their effect on heat waves does not propagate in space, but only in time.
Lag: 2,5 days

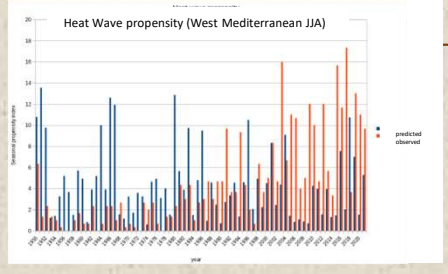


The remote candidate predictors are sea surface temperatures from 20 ocean regions and sea ice concentration from four Arctic sectors.
Lag: 15 days

Preliminary results







↑ Feature importance for each of the regions in JJA/DJF seasons
← Comparison between predicted and observed heat propensity

Discussion

This is a first very preliminary attempt. Many improvements will be done to increase the skill of the statistical model.

- Assumptions on delay time will be tested.
- New local (vegetation) and remote (stratosphere) predictors.
- Dynamical forecasts will be also used as predictors.
- No detrend has been performed so far.

Barcelona Supercomputing Center
Centro Nacional de Supercomputación
Max Planck Institute for Biogeochemistry
MARIE CURIE ACTIONS
This project has received funding from the EU's H2020 research and innovation program under the MSCA grant agreement No 101033654.