Weather and Climate Extremes 49 (2025) 100792

Contents lists available at ScienceDirect

Weather and Climate Extremes

journal homepage: www.elsevier.com/locate/wace

Identifying key drivers of heatwaves: A novel spatio-temporal framework for extreme event detection

J. Pérez-Aracil [a,i,*], C. Peláez-Rodríguez [a], Ronan McAdam [b], Antonello Squintu [b], Cosmin M. Marina [a], Eugenio Lorente-Ramos [a], Niklas Luther [e], Verónica Torralba [j], Enrico Scoccimarro [b], Leone Cavicchia [b], Matteo Giuliani [c], Eduardo Zorita [d], Felicitas Hansen [d], David Barriopedro [f], Ricardo García-Herrera [g], Pedro A. Gutiérrez [h], Jürg Luterbacher [e], Elena Xoplaki [b,e], Andrea Castelletti [c,k], S. Salcedo-Sanz [a]

[a] Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, 28805, Spain
[b] CMCC Foundation - Euro-Mediterranean Center on Climate Change, Italy
[c] Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milano, Italy
[d] Helmholtz-Zentrum Hereon, Hamburg, Germany
[e] Department of Geography, Climatology, Climate Dynamics and Climate Change, Justus Liebig University Giessen, Giessen, Germany
[f] Instituto de Geociencias (IGEO), Consejo Superior de Investigaciones Científicas–Universidad Complutense de Madrid, Madrid, Spain
[g] Departamento de Física de la Tierra y Astrofísica, Facultad de Ciencias Físicas, UCM, Madrid, Spain
[h] Departamento de Ciencia de la Computación e Inteligencia Artificial, Universidad de Córdoba, Córdoba, Spain
[i] Programa de doctorado en Computación Avanzada, Energía y Plasmas, Universidad de Córdoba, Córdoba, Spain
[j] Barcelona Supercomputing Center (BSC), Barcelona, Spain
[k] RFF-CMCC European Institute on Economics and the Environment, Euro-Mediterranean Center on Climate Change, Milano, Italy

ARTICLE INFO

Dataset link: GitHub repository

Keywords:
Heatwaves
Spatio-temporal optimization
Large-scale drivers
Cluster-based feature selection
Multi-method ensembles
Optimization

ABSTRACT

Heatwaves (HWs) are extreme atmospheric events that produce significant societal and environmental impacts. Predicting these extreme events remains challenging, as their complex interactions with large-scale atmospheric and climatic variables are difficult to capture with traditional statistical and dynamical models. This work presents a general method for driver identification in extreme climate events. A novel framework named Spatio-Temporal Cluster-Optimized Feature Selection (STCO-FS) is proposed to identify key immediate (short-term) HW drivers by combining clustering algorithms with an ensemble evolutionary algorithm. The framework analyzes spatio-temporal data, reduces dimensionality by grouping similar geographical grid cells for each variable, and develops driver selection in spatial and temporal domains, identifying the best time lags between predictive variables and HW occurrences. The proposed method has been applied to analyze HWs in the Adda river basin in Italy. The approach effectively identifies significant variables influencing HWs in this region. This research can potentially enhance our understanding of HW drivers and predictability.

1. Introduction

The occurrence of heatwaves (HWs), characterized by prolonged periods of abnormally high temperatures exceeding typical local conditions, has become a pressing concern in recent years due to their severe societal and environmental impacts (Perkins, 2015; Easterling et al., 2000). Since 1950, extensive regions worldwide have witnessed numerous prolonged and intense HWs, resulting in significant consequences for human mortality, regional economies, and natural ecosystems (Meehl and Tebaldi, 2004; Zittis et al., 2022; Kuglitsch et al., 2010; Della-Marta et al., 2007; García-Herrera et al., 2010). In agriculture, heat stress on crops can significantly reduce yields, leading to food insecurity. In addition, increased demand for electricity for cooling during HWs substantially strains power grids. The escalation in the frequency of HWs has been documented in various parts of the globe in recent years and is at least partly attributed to the temperature increases driven by anthropogenic warming (Russo et al., 2014, 2015).

Numerous studies (Barriopedro et al., 2011; Sillmann et al., 2013; Stott et al., 2011) have consistently highlighted that the ongoing increase in global surface temperatures will lead to significant alterations in the frequency and intensity of HWs across Europe by the end of this century. This trend is not confined to Europe; globally, there is also a

* Corresponding author.
E-mail address: jorge.perezaracil@uah.es (J. Pérez-Aracil).

https://doi.org/10.1016/j.wace.2025.100792
Received 13 November 2024; Received in revised form 18 May 2025; Accepted 3 July 2025
Available online 26 July 2025

growing prevalence of heat extremes, with projections indicating that these events will continue to increase in the coming decades (Battisti and Naylor, 2009; Coumou and Robinson, 2013; Fischer et al., 2013). Regional differences can be encountered in HW projections. Hence, this leads to diverse drivers and climate forcings on regional scales. The identification of these drivers plays a key role in understanding regional variations and in developing effective mitigation and adaptation strategies, as different regions may experience distinct climate impacts due to a combination of local factors and global climate forces. Moreover, understanding these drivers is crucial for improving forecasts on subseasonal scales, allowing for more accurate predictions of HWs and other extreme events.

When tackling the challenge of HW detection or prediction, it is necessary to understand the mechanisms responsible for these extreme events. Although the underlying processes remain not entirely understood (Perkins, 2015), an increasing number of studies have delved into these mechanisms and physical drivers that contribute to the formation and prediction of HWs (Domeisen et al., 2023; Barriopedro et al., 2023). HWs are the product of intricate interactions between large- and small-scale processes that operate across diverse temporal scales. These events are highly influenced by atmospheric circulation, often regarded as a fast-acting driver, as well as anomalous conditions in slowly changing climate components, which can serve as proximate factors (e.g., land surface) or remote factors (e.g., upper ocean temperature, or sea ice) affecting HWs occurrence (Sillmann et al., 2017; Hoskins and Woollings, 2015; Miralles et al., 2019). In the extratropics, atmospheric circulation patterns that influence HWs include quasi-stationary synoptic-scale high-pressure systems (anticyclones) (Brunner et al., 2018; Sousa et al., 2018), whose predictability at a seasonal scale is low due to the influence of the chaotic variability of the atmosphere (Prodhomme et al., 2021). Finally, long-term trends in frequency, duration, and intensity of HWs are primarily driven by anthropogenic forcings, including global factors such as greenhouse gas concentrations and regional factors like land-use/land-cover changes and aerosol emissions (Seneviratne et al., 2021). However, these are out of the scope of this paper.

In close relation to the previous discussion, and considering the vast volume of available spatial and temporal data, employing data-driven methodologies becomes indispensable for uncovering potential HW drivers. A limited body of literature addresses this subject using Machine Learning (ML) and feature selection and dimensionality-reduction approaches. Some works (Asadollah et al., 2021; Buschow et al., 2023; Loughran et al., 2017) employed Principal Components Analysis (PCA) to reduce and optimize the number of highly correlated variables, using them as inputs in some ML algorithms. In Wehrli et al. (2019), authors aimed to identify the role of the individual drivers for five HWs in the recent decade through factorial experiments, which force the model towards observations for one or several key components at a time, allowing to identify how much of the observed temperature anomaly of each event can be attributed to each driver. Other feature selection approaches have been used for different weather problems in searching for optimal input variables. In Tao et al. (2022), an extreme gradient boosting feature selection algorithm was applied with ML models in a problem of short-term relative humidity prediction. In Hagen et al. (2021), a nested loop of roughly pruned random forests was used for identifying significant drivers of daily streamflow from large-scale atmospheric circulation in Norway. In Chaqdid et al. (2023), a clustering method was applied to divide Morocco into regions that are spatially consistent in terms of extreme precipitation and to identify its drivers by analyzing atmospheric circulation anomalies during the occurrence of regional events. In Orimoloye et al. (2022), ML regression-based algorithms were used to identify the drivers of drought dynamics in the Free State Province. Dalal et al. (2024) shows the influence of different drivers to understand the causal mechanism of HWs over South-West India. For that purpose, climate model simulations and long-term observational data were proposed.

This study proposes a general framework for HW driver identification, which can be applied to other extreme events in the context of detection and event short-term prediction. The framework is illustrated here to detect HWs in a European location. Specifically, the framework proposed in this work follows a two-phase methodology to obtain robust HW driver identification. In the first phase, a clustering algorithm is applied to variables identified as potential drivers, extracted from the ERA5 reanalysis dataset (Hersbach et al., 2018), and presented as time series. This clustering step reduces the dimensionality of the spatial domain by grouping grid cells with similar time series patterns. In the second phase, a wrapper feature selection approach based on a multi-method ensemble evolutionary algorithm (PCRO-SL) (Pérez-Aracil et al., 2023) is employed to identify the most skillful drivers and periods for HW forecast over short-term (days to weeks) and seasonal horizons. The optimization algorithm's fitness function performs a driver selection by evaluating the performance of an ML model for HW classification based on a subset of clustered drivers.

The proposed framework is applied to the agricultural districts in the Adda river basin, located downstream of Lake Como, in the Lombardy region, Northern Italy. These districts are part of the Po Valley, one of the most productive European agricultural areas, which provides one-third of the national agricultural production (Eurostat, 2024). Understanding the crop risks associated with extreme temperatures is becoming increasingly crucial to planning effective climate change adaptation strategies.

The manuscript is organized as follows. First, a description of the data, including potential drivers and target variables used for developing the experiments, is provided in Section 2. Then, the spatio-temporal feature selection methodology is presented and detailed in Section 3. Subsequently, the experimental work and the results obtained are further described in Section 4. Finally, in Section 5, there is a discussion on the potential uses of the framework in wider-scale driver detection and on implications for forecasting.

2. Data description

This section provides a detailed description of the data used for the experiments and the construction of the target. First, regarding HW definition, this issue has been widely discussed in the literature (Barriopedro et al., 2023). This work follows the widely-used HW definition given in Russo et al. (2015) based on cumulative normalized daily maximum temperature (TX) exceedances. A heat wave occurs when T_{max} exceeds a given threshold for three or more consecutive days. The threshold is the daily 90th percentile of T_{max} in the reference period (1981–2010), computed over a 31-day window. Therefore, the threshold for a given calendar day d is $d_{th} = P_{90}(A_d)$, with:

$$A_d = \bigcup_{y=1981}^{2010} \bigcup_{i=d-15}^{d+15} T_{y,i}, \tag{1}$$

where $\bigcup$ denotes the union of sets and $T_{y,i}$ stands for the daily T_{max} of the day i in the year y.

Next, we will provide details on the potential drivers considered and the target variables considered.

2.1. Potential drivers and target

The variables considered as potential drivers to perform the HW prediction may be categorized into three groups: (1) meteorological variables, local or remote, (2) climate indices, (3) other variables.

The first group consists of atmospheric, ocean and other variables which influence climate on various timescales: mean sea level pressure (MSLP), outgoing longwave radiation (OLR), total precipitation (TP), height of the 500 hPa geopotential (Z500), 2 m temperature (T2M), as well as sea surface temperature (SST), sea ice concentration (SIC) and volumetric soil moisture in the upper 7 cm (SM).

The second group, climate indices, are included because long-term indices are linked to large-scale atmospheric patterns that influence temperature over extended periods (Kenyon and Hegerl, 2008). In addition, long-term indices help distinguish between climate change variability and natural variability. However, the role of large-scale drivers and teleconnections in the Adda river basin, as for much of Europe, is still unclear (Giuliani et al., 2019). First, the NINO3.4 index (area-averaged SST anomaly in the region 5 ° S-5 ° N, 120 ° W-170 ° W) is used to represent the El Nino Southern Oscillation (ENSO), which is strongly linked with the occurrence of extreme heat in northern continents (Luo and Lau, 2020, 2019; Martija-Díez et al., 2021). The Indian Ocean Dipole (IOD), whose association with HWs has been investigated in recent years (Reddy et al., 2021; Dalal et al., 2024), is calculated as the difference of area-averaged SST anomaly between the western tropical Indian Ocean (50 ° E–70 ° E, 10 ° N–10 ° S) and the southeastern tropical Indian Ocean (90 ° E–110 ° E, O°N-10 ° S). Lastly, the North Atlantic Oscillation (NAO), whose impact on European heatwaves has been previously studied in Li et al. (2020), Mukherjee et al. (2020), Kueh and Lin (2020), is derived from the first principal component of Z500 in the North Atlantic domain.

The eight predictor fields, such as soil moisture, sea surface temperature, sea ice coverage, mean sea level pressure, geopotential height at 500 hPa, total precipitation, outgoing longwave radiation, and temperature at 2 m along with climate variability modes (e.g., NAO, ENSO), were selected based solely on their well-established physical relevance as drivers of heatwaves (see Barriopedro et al., 2023 and references therein). In other words, expert knowledge guided the identification of this initial set of potential heatwave drivers.

The third group covers miscellaneous variables such as mean atmospheric CO_2 levels (Lemordant et al., 2016) and the specific calendar day of the year (DOY) (Li et al., 2023).

Table 1 describes the geographical domain considered for each meteorological variable. Some have been studied in two domains to account for their varying influence in various geographical scales, according to previous studies. SM and precipitation are considered local-scale influences on European HWs (Stefanon et al., 2012). European temperatures are indirectly linked to Arctic SIC (Zhang et al., 2018; Sun et al., 2022) through impacts on atmospheric circulation, but we assume no link to Antarctic SIC. North Atlantic SSTs affect the European climate, particularly in winter (Gastineau and Frankignoul, 2015), but there is some evidence to suggest a lagged influence on summer heatwaves (Duchez et al., 2016; Bischof et al., 2023). Meanwhile, global modes of climate variability, which influence distant continents via teleconnections, are represented by global SST patterns. For an explicit consideration of atmospheric dynamics and teleconnections, z500 and MSLP are used locally and globally. Extreme heat is often linked to regional scale circulation anomalies (Sousa et al., 2018), which can be excited and/or amplified remotely via atmospheric teleconnections (Barriopedro et al., 2023). Also, land variables available over the local region under study are considered as independent potential drivers (MSLP, OLR, SM, T2M, TP and Z500).

Regarding the target variable, we have selected the agricultural districts in the Adda river basin (including Lake Como) in the north of Italy (centered around 46° N, 9° E). For this location, a daily time series of binary HW occurrence index over 1950–2022, for the warmest months of the year (May, June, July, and August) using the HW definition given in Russo et al. (2015).

Table 1
Predictive variables considered from the ERA-5 reanalysis dataset. The specific coordinates corresponding to the geographical limits: Europe: [30-70N, 16W-44E], Arctic: [48-90N, 180W-180E], North Atlantic: [0-66N, 90W-40E] and local: [46° N, 9° E].

textbfVariable	Domain
Mean Sea Level Pressure (MLSP)	Global, Europe, local
Soil Moisture (SM)	Europe, local
Sea Ice Cover (SIC)	Arctic
Sea Surface Temperature (SST)	Global, North Atlantic
Height of the 500 hPa Geopotential (Z500)	Global, Europe, local
Total Precipitation (TP)	Europe, local
Outgoing Longwave Radiation (OLR)	Global, North Atlantic, local
2 m Temperature (T2M)	Europe, local

2.2. Data extraction and preparation

The detection of HWs presented in this paper is performed based on physical variables data extracted from ERA5 reanalysis (Hersbach et al., 2018). ERA5 provides hourly information on a broad set of variables, such as temperature, pressure, precipitation, and snowfall, with a resolution of 0.25 degrees in both longitude and latitude. Daily average values were considered in this case, with a horizontal resolution of 0.5 degrees. The 72-year ERA5 database, based on data from 1950 to 2022, has been considered for both target and predictive variables. These data have been divided into a training period from 1950 to 2010 and a test period from 2011 to 2022. For the training split, the positive cases (HW occurrence) represent 5.1% of the cases, while for the test split, the positive cases represent 15.1% of the cases.

The maximum temperature of each calendar day is considered for the 1981–2010 period, and a rolling average of 30 days is applied to smooth this annual cycle. Regarding this point, note that the choice of rolling average varies greatly in the literature (Perkins et al., 2015; Barriopedro et al., 2023). Existing studies employ different methods depending on data availability: for example, Stefanon et al. (2012) used a longer 60-year climatology and opted for a 20-day window (1200 days), while studies on seasonal forecast output are restricted to use a polynomial regressor (Torralba et al., 2024). A 30-day rolling average is now a common choice for heatwave studies (e.g., Russo et al., 2015), although 15-day has been used too (Perkins et al., 2015). Other studies (Sánchez-Benítez et al., 2020) have also employed a 31-day window to detect, predict, and attribute Mediterranean heatwaves as a compromise to obtain a smoothed evolution of that threshold without underestimating its seasonal cycle's amplitude. A recent study has even pointed out possible issues with the percentile-based approach, which could lead to underestimating the expected extreme frequency of heat wave events (Brunner and Voigt, 2024). In this case, we have considered a rolling average of 30 days because, as previously stated, it is a common choice in the literature and fosters comparability with other studies. However, we recognize that this question deserves attention for future works.

3. Spatio-temporal cluster-optimized feature selection (STCO-FS)

This section presents the proposed framework to identify optimal HW drivers in spatial and temporal domains. Fig. 1 illustrates the methodology flow, where it is worthwhile to distinguish between the data treatment of potential predictor variables and the target variable. The possible drivers have been defined in Section 2.1. The proposed framework has two steps (plus a preprocessing stage), as shown in Fig. 1. The first step consists of clustering the drivers to reduce the spatial dimension. The area-weighted spatial average time series of clusters are then merged with the local variables and climate indices. In the second step, a wrapper feature selection method is applied using an evolutionary optimization algorithm and a ML method for selecting the optimal time frames of each potential driver.

Once the input and target variables have been processed, they are used to feed supervised ML classification algorithms, which conduct the detection of the HW occurrence over the Adda river basin in the considered period. The framework steps are described in the subsequent sections.

3.1. Dimensionality reduction through clustering

The classic K-means clustering algorithm has been employed to reduce the spatial dimensionality of the predictive data. Initially introduced by MacQueen (MacQueen et al., 1967), K-means has become one of the most widely used and extensively studied clustering algorithms. The key input parameter in the K-means approach is the number of clusters, denoted by k. The algorithm then partitions the data into k clusters by following defined steps. It is important to note that the proposed methodology allows other possible clustering algorithms (different from K-means) to cluster the input data.

The proposed method is designed to be flexible and adaptable, allowing for the inclusion of any other clustering technique that may better suit the specific problem under study. It is important to emphasize that, at this stage of the process, the primary objective is to reduce the dimensionality of the problem, particularly in the spatial domain. The choice of clustering method in this study is somewhat arbitrary, as it was primarily selected for its ease of implementation and computational efficiency. However, the modular nature of the framework ensures that alternative clustering methods can be easily integrated without compromising the overall structure. Potential alternatives include hierarchical clustering, spectral clustering, or physically guided clustering based on domain-specific knowledge (e.g., IPCC regions). Future research could explore the impact of using different clustering approaches to enhance the robustness and interpretability of the results.

This study implements the clustering algorithm for the 12 predictors outlined in Table 1. Next, the area-weighted spatial average for each cluster is computed. Consequently, a time series is generated for each cluster under consideration. A value of $k = 5$ clusters has been considered for each predictor field, giving a total of 60 clusters involved in the prediction ($5 \text{ clusters} \times 12 \text{ variables} = 60$ potential drivers). The local seasonal cycle is removed from each candidate driver to generate a time series of anomalies, which represent deviations from the expected seasonal pattern. This preprocessing step ensures that the analysis focuses on anomalous variations rather than regular seasonal fluctuations, thereby highlighting potential predictors related to the heatwave occurrence. Clustering was applied considering the training data (from 1950 to 2010) of daily time series anomalies relative to 1981–2010. For each variable, five clusters are obtained, displayed in Fig. 3. For T2M, values over land are only considered. For SIC, areas historically free of sea ice have been masked out.

Although the number of clusters is arbitrary, it is essential to note that this work aims to present the methodology but not its best configuration, which may depend on the location of the study and the different variables considered.

The proposed method allows the introduction of unclustered variables as potential drivers of the problem. Thus, those variables that do not need to be spatially grouped can be selected, and their time series are included in the analysis. This study considers 11 additional features included as potential drivers, including the climate indexes, the local meteorological variables (MSLP, OLR, SM, T2M, TP and Z500 taken over the Adda river basin) and the other variables defined in Section 2.1. Thus, these results in a database of predictor variables comprising 71 variables arranged as a time series.

3.2. Candidate selection: the optimization problem

After reducing the dimensionality of the problem in the spatial domain, our focus shifts to the temporal dimension. Here, the objective is to identify periods and lags exhibiting the highest predictive skill for each potential driver. First, the prediction time horizon is defined, determining how far in advance predictor data should be considered. This work sets the time horizon to zero since it is configured as a detection problem. Based on the prediction horizon, the time lag and sequence length values are searched for each potential driver under study using evolutionary computation. They represent the lead time and the window length considered for each variable, which enables us to distinguish between short-term (low time lag) and long-term (high time lag) predictors. An illustration is shown in Fig. 2. In this work, the maximum time lag is set as 180 days, and the maximum window length is set as 60 days. Therefore, the evolutionary search could account for a lead time of up to 8 months for each of the variables.

The process of determining the time lag and the sequence length has been conducted using a robust and well-established optimization algorithm: the Probabilistic Coral Reef Optimization Algorithm with Substrate Layers (PCRO-SL) (Salcedo-Sanz, 2017; Pérez-Aracil et al., 2023). It is a low-level ensemble for optimization (Wu et al., 2019), based on evolutionary computation. It was first proposed as an advanced version of the original CRO algorithm (Salcedo-Sanz et al., 2014), which was an evolutionary-type meta-heuristic, proposed as a class of hybrid between Evolutionary Algorithms (Del Ser et al., 2019) and Simulated Annealing (Kirkpatrick et al., 1983). Specifically, the CRO algorithm for optimization can be described as follows: The CRO algorithm is an evolutionary-type algorithm that mimics the biological processes of coral reproduction. Due to its features, it is considered a meta-heuristic in between Simulated Annealing and Genetic Algorithms. The main data structure of CRO is the reef (population of potential solutions to the problem): it is represented as a square grid of dimensions (M, N), such that each of its cells contains a coral (a solution to the optimization problem). The reef starts with a fraction ρ_0 of its cells occupied by randomly generated solutions. Next, coral reproduction is simulated according to the following procedures:

1. **Larvae formation**: new individuals (called *larvae*) are produced by one of two processes:
   (a) **Broadcast spawning**: Mimics the external reproduction of real-life corals. A fraction F_b of the reef's corals are selected, and separated into pairs, and then each pair produces a new larva through crossover (or other strategy).
   (b) **Brooding**: Mimics the internal reproduction of real-life corals. The remaining $(1 - F_b)$ fraction of the reef's corals employ a mutation mechanism to produce a new larva each.
2. **Larvae settling**: given the group of larvae produced in the previous step, each tries to settle down by moving to a position in the reef. CRO simulates this by iterating over the larvae, selecting a random cell for each, and checking whether (1) the spot is empty, or (2) the larva's fitness is higher than the fitness of the coral which currently occupies the spot. If either condition is met, the larva settles in the cell and becomes a new coral, otherwise, the larva selects another random cell and tries again (for a maximum of k times).
3. **Fragmentation**: this step simulates the asexual process through which the corals also reproduce. Every coral on the reef is sorted according to its fitness, and the F_a fittest are selected for this type of reproduction. Then, each of the selected individuals is duplicated to produce a new larva, which will try to settle down in the reef in the same way as described in the previous step.

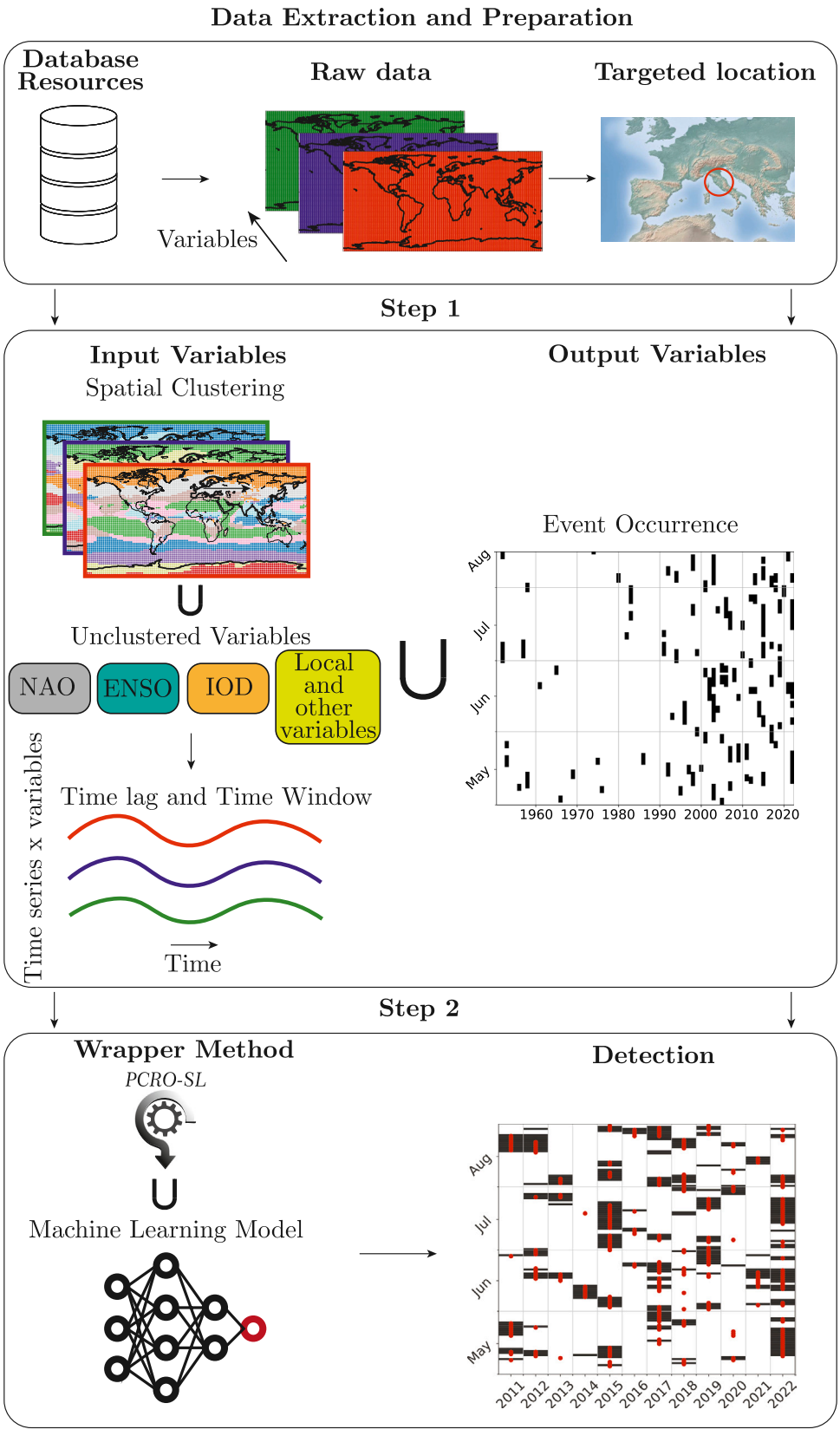

Fig. 1. Scheme of the proposed feature selection framework. The framework is structured in two main steps to identify key heatwaves (HWs) drivers. Step 1: Data extraction and preparation involve collecting climate variables from multiple sources, focusing on the targeted location. Input variables are processed through spatial clustering to group regions with similar temporal behaviour, while unclustered variables (e.g., NAO, ENSO, IOD, local variables) are directly included. The framework considers time lags and windows to capture short-term and long-term effects on HW events. The output variable is the binary HW occurrence index derived from temperature data. Step 2: A wrapper method (PCRO-SL) selects the most relevant variables fed into a machine learning model to detect HW occurrences. The framework efficiently combines clustering, feature selection, and ML methods to identify key spatial-temporal drivers of heatwaves.

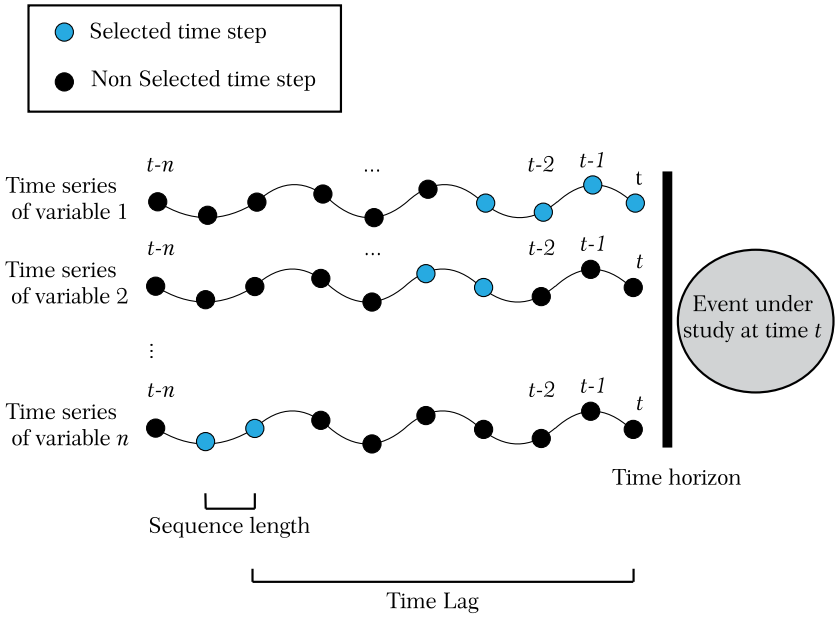

Fig. 2. Illustration of time series: time lag, forecast horizon and sequence length parameters.

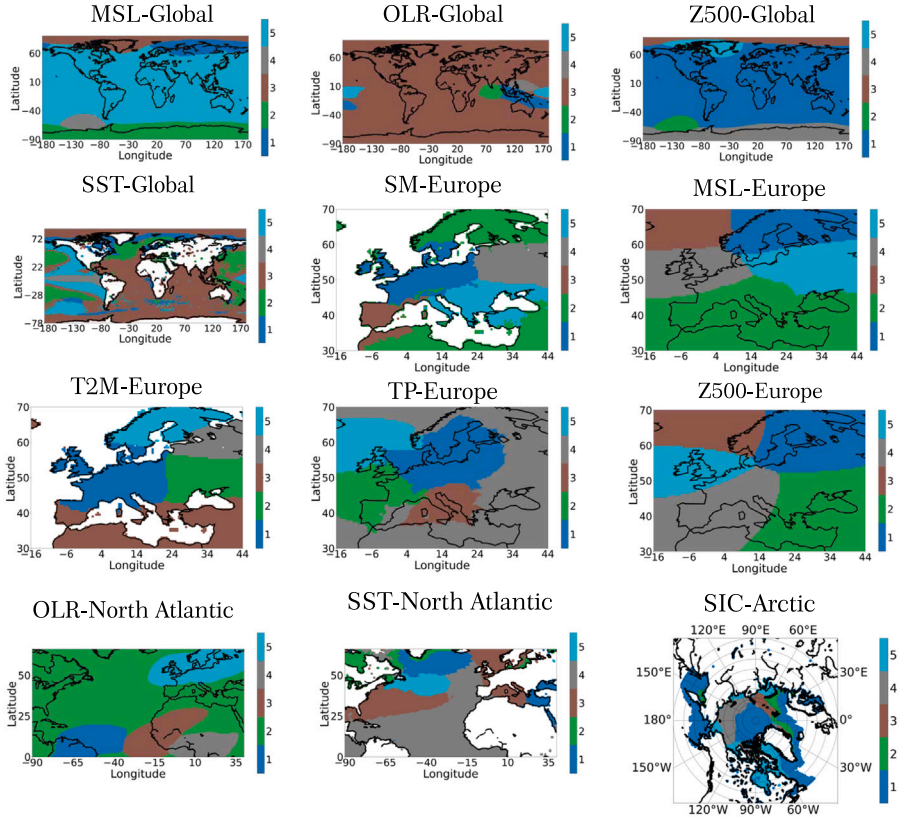

Fig. 3. Clusters provided by the K-Means+ algorithm ($K = 5$) for the first group of variables: meteorological variables.

4. **Depredation**: in this final step, the worst corals in the reef become candidates for elimination. Namely, a fraction F_d of the lowest-fitness individuals is selected, and each one has a probability P_d of being eliminated and freeing up their reef cell.

The process described above is repeated until a stopping criterion is reached, which is typically a maximum number of generations or function evaluations. In turn, the PCRO-SL algorithm further evolves the CRO approach towards a multi-method ensemble. It generally proceeds as the original CRO described above, but with a significant difference: Instead of having a single way of evolving, it considers several *substrate layers* of the approximately same size in the reef. Each substrate, in turn, represents a particular evolution strategy or searching procedure to be applied. Thus, the PCRO-SL is a multi-method ensemble algorithm (Wu et al., 2019), where several searching strategies are carried out within a single population.

Different combinations of well-known meta-heuristics may be implemented. In this case, we considered regular combinations of previously defined meta-heuristics. Specifically, we have defined and applied the following substrates in the PCRO-SL: Harmony Search (HS), Multipoint crossover (MPx), XOR operation (XOR) and BLX-α crossover (BLX).

The optimization problem formulation for selecting the optimal time-domain features for each driver is structured as follows. Each possible solution generated by the PCRO-SL consists of an array containing three key variables for each candidate driver: time lag, sequence length, and a binary indicator determining whether the driver is included or discarded. Including this binary variable encourages solutions that prioritize minimal information, helping reduce the impact of noise. The time lag is constrained within $[h, 180]$ days, where h is the time horizon, while the sequence length is limited to $[1, 60]$ days. In this initial study, we focus on a prediction time horizon of $h = 0$ days, aiming to identify drivers informing on HW detection (nowcasting) up to time scales longer than six months. The variation in sequence length allows us to investigate drivers that maintain significant influence over different periods within the time domain.

The achieved F1-score dictates the optimal lag selection for each variable. The F1-score is the harmonic mean of the measures: precision (ratio of correctly predicted positives to all predicted positives) and recall (ratio between the correctly predicted positives to all observed positives) (Sasaki et al., 2007). The F1-score is widely used to assess the quality of binary predictions (Goutte and Gaussier, 2005). In our case, the F1-score achieved by each candidate combination of drivers served as the fitness function for selecting optimal clustered and unclustered variables, with their corresponding time lags and sequence length, and other variables. To accomplish this, a deterministic and fast training classifier, namely the popular Logistic Regressor (LR) (Kleinbaum et al., 2002), was used (other ML algorithms were tested). When the optimization algorithm provides a potential solution (comprising three values per driver), the chosen time lags are concatenated into a tabular format, and LR training is performed. This process was conducted using a Cross-Validation (CV) approach: the entire training data is divided into five validation folds, and the average error encountered for these folds has been used as the fitness function of the optimization algorithm.

3.3. Machine learning classifiers

Although a fast-training ML algorithm such as LR is used during the optimization process, a variety of more sophisticated models are subsequently implemented to evaluate the optimal solution the PCRO-SL algorithm provides. These models include: Light Gradient Boosting Machine (LGBM) (Ke et al., 2017), Support Vector Classifier (SVC) (Schölkopf et al., 2000), Decision Trees (DTs) (Loh, 2011), Random Forest (RF) (Breiman, 2001), Gaussian Naive Bayes (GNB) (Zhang, 2004), K-Nearest Neighbors (KNN) (Mucherino et al., 2009), Adaptive Boosting (AB) (Freund and Schapire, 1996), Multi-Layer Perceptron (MLP) (Gardner and Dorling, 1998), Gradient Boost (GB) (Friedman, 2001) and Extreme Learning Machine (ELM) (Huang et al., 2006).

Table 2
Parameters of the experimental setup. The hyperparameters of each model are described for the models used in this work: Light Gradient Boosting Machine (LGBM), Support Vector Classifier (SVC), Decision Trees (DTs), Random Forest (RF), Gaussian Naive Bayes (GNB), K-Nearest Neighbors (KNN), Adaptive Boosting (AB), Multi-Layer Perceptron (MLP), Gradient Boost (GB), Extreme Learning Machine (ELM).

ML Methods	**LGBM**		**SVC**	
	num leaves	20–200	C	0.1–1000
	n estimators	50–500	Gamma	0.001–1
			Kernel	*rbf*
	DT		**RF**	
	max depth	1–50	n estimators	100–600
	min samples leaf	1–50	bootstraps	True/False
	GNB		**KNN**	
	var smoothing	−9–0	n neighbors	3–30
	AB		**ELM**	
	n estimators	50–200	n neurons	10–500
	learning rate	0.001–10		
	GB		**MLP**	
	n estimators	50–300	n layers	1–4
	learning rate	0.01–0.2	n neurons	32-512
	max depth	1–9	activation	*relu*
			solver	*adam*
			alpha	0.0001–0.01
			batch size	16–64
			learning rate	0.0001–0.01
			max iters	200–600

These methods are implemented in Python using the libraries `sklearn`, `skelm`, and `lightgbm`. The hyperparameters of the classifiers are optimized through a random hyperparameter search, with the parameter ranges detailed in Table 2. A five-fold cross-validation (CV) is performed to assess model performance, and standardization of the input variables is applied as part of the preprocessing pipeline to ensure that features are on a comparable scale.

4. Experimental work and results

This section describes the experimental work, the results obtained, and the corresponding discussion. First, Section 4.1 details the drivers the optimization algorithm selects. Second, Section 4.2 shows the results provided by the ML models in the detection task. In all experiments, given that the evolutionary search could account for a lead time of up to 8 months, the first year of training data for the target variable is excluded, resulting in a training period spanning from 1951 to 2010, while the test period covers 2011 to 2022, comprising 18% of the total data. The training dataset identifies optimal drivers and model training, whereas the test dataset is reserved exclusively for evaluating model performance.

4.1. The selected variables

The optimization algorithm is initialized once the potential predictors are made up, including clusters, unclustered variables, and the local and climate indices. The clustered variables are shown in Fig. 3.

The solution provided by the proposed method is a vector of length 3×number of variables. It represents the time lag, sequence length and a binary variable with the time steps for which the variable is selected. In this case, the forecast horizon has been set to 0 since the goal is to validate the methodology for unveiling potential drivers before attempting a more complex forecasting problem.

The PCRO-SL optimization algorithm is executed in ten independent runs to prevent false positives caused by the inherent randomness in this optimization problem. Fig. 4 shows a potential solution the algorithm provides, corresponding to the best solution of all the runs,

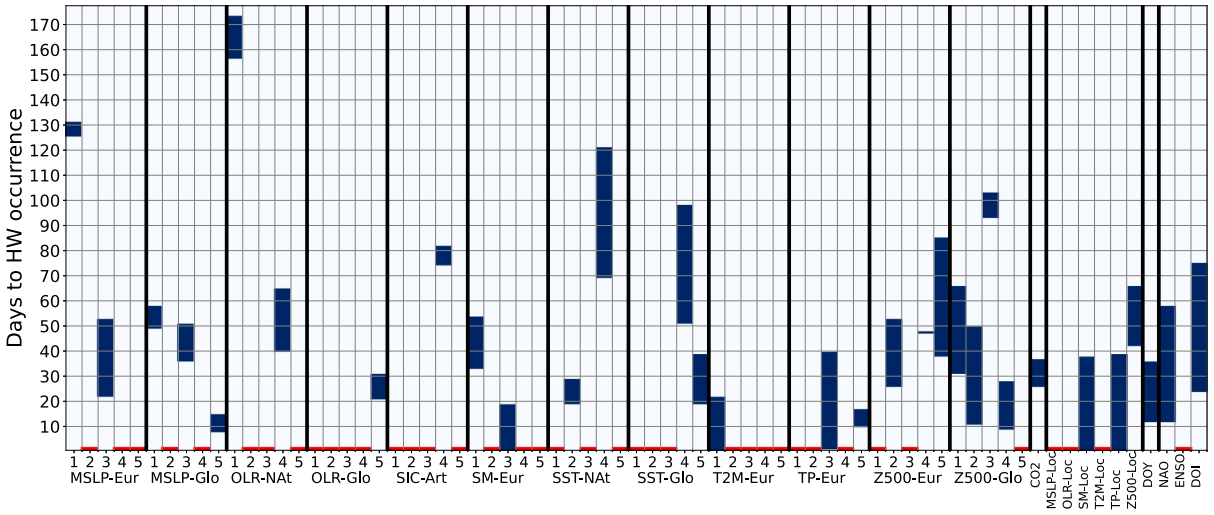

Fig. 4. Example of the individual solution provided by the optimization algorithm. Red boxes represent variables that have not been selected in this particular solution.

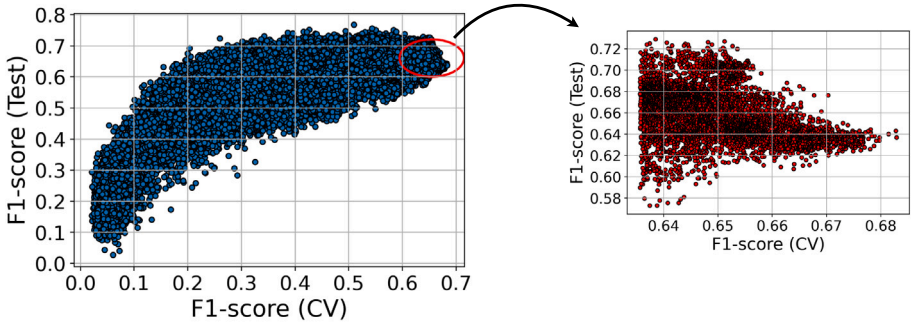

Fig. 5. Test vs train (CV) performance of all the potential solutions tried by the optimization algorithm. The solutions represented in red are the ones selected for an in-depth study.

in terms of CV error. Here, in the x-axis, the 71 potential drivers are listed. The y-axis shows the temporal scale (in days relative to the HW occurrence). This plot highlights in blue the time steps the optimization algorithm selects for each predictor in the case of this specific example. The red square means that the algorithm has discarded the specific potential driver.

Fig. 4 helps the user to interpret, visually and intuitively, the different variables that are being chosen, as well as distinguish between variables that have a sort term influence (in this case, T2M in the European domain over the cluster 1, SM in the European domain, cluster 3, and TP in the European domain over the cluster 3 (see Fig. 3), and the T2M and Z500 at the local node of Lake Como); and others variable with a delayed impact that influence with a delay of up to 170 days (e.g. OLR over the North Atlantic domain in cluster 1).

The question that can arise when examining individual solutions is whether or not all the selected time steps provide predictability to the problem or whether they are noise that has been added to the truly significant variables. To analyze this aspect, we evaluated all potential solutions generated by each independent run of the optimization algorithm. A total of 150,000 potential solutions are generated (15,000 evaluations of the fitness function × 10 independent runs). Each solution represents a combination of predictor variables, tested on the train (by 5-fold CV) and test data. The metrics of each combination of drivers are plotted in Fig. 5. Out of these 150,000 possible solutions, the best 10% (in terms of CV error) have been selected to analyze further the predictor variables that provide significance for the prediction problem under evaluation. These selected combinations are represented with red points in the scatter plot of Fig. 5.

The best solutions are then analyzed in the frequency map reported in Fig. 6. The darker the color, the more frequently that variable has been selected in that time lag. Low intensity means that the corresponding time step has been barely chosen among the best solutions and is therefore considered noise.

It becomes evident that certain variables selected by some solutions merely introduce noise, obscuring the truly significant variables. Generally, the most frequently selected features are found on the short-term and sub-seasonal time scales (i.e. <45 days). Three of the most frequently chosen drivers all have short lag times (<20 days): T2M-Eur-1 (the regional cluster in which the Adda basin is found), and local values of SM and TP. These predictors are considered important over 20 days before the event rather than the day, indicating the importance of persisting conditions. The selection of nearby short-term drivers is not surprising, particularly given the previously identified roles of soil moisture and precipitation in summer temperatures (Stefanon et al., 2012; Ardilouze et al., 2019). On the sub-seasonal scale, OLR-Glo-5 (western-central Pacific; 20–30 days), Z500-Eur-2 (eastern Mediterranean; 30–50 days), NAO and IOD (both over 20–55 days) are the most frequently selected predictors. On the seasonal scale, Z500-Eur-5 (North Atlantic; 70–85 days) and SST-Glo-4 (Tropical Pacific; 90–100 days) are the most selected. While the scope of this study is not to perform a process-based study on how detected features physically impact HW occurrence, it is shown here that the selection of some features is at least partially supported by evidence. For other variables, this framework may act as a first step in understanding which processes need to be studied further.

Regional-scale atmospheric circulation is considered key, judging by the frequent selection of Z500-Eur 2 (eastern Mediterranean; 30–50

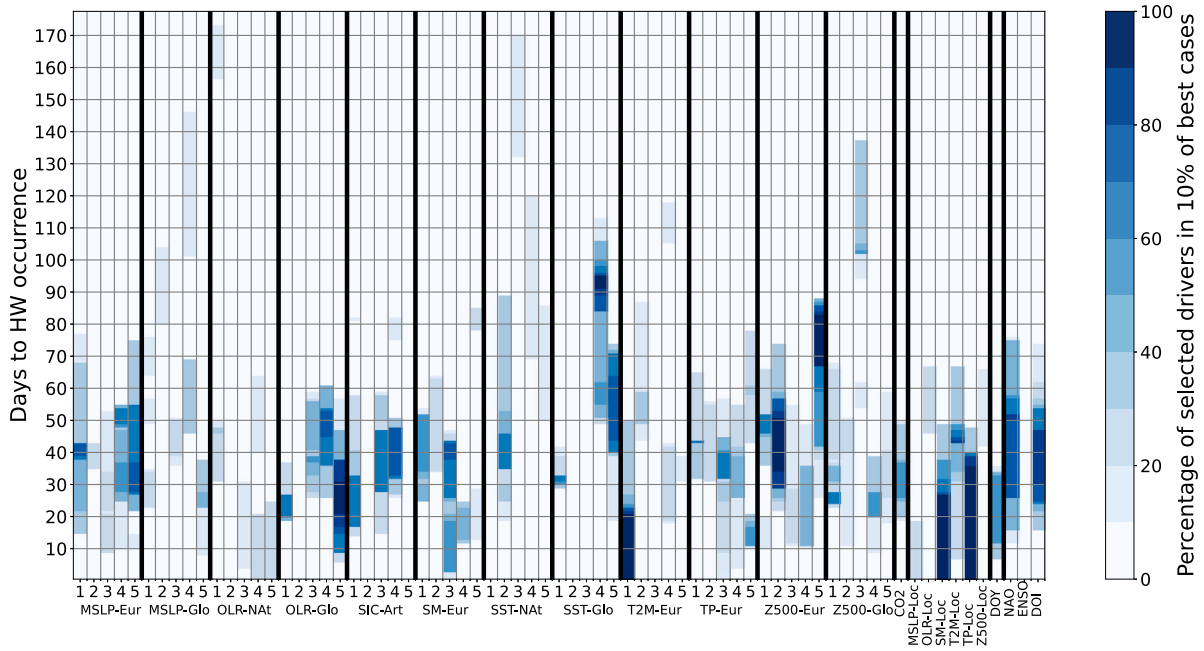

Fig. 6. Heatmap representing the best 10% of solutions the optimization algorithm provides. For a specific time step, darker colors mean that most proposed solutions select the variable, while light colors denote that it is barely selected.

days) and Z500-Eur-5 (British Isles; 70–85 days). However, features such as blocks and ridges are known to determine the occurrence and intensity of summer HWs in the days before an event (Sousa et al., 2018), instead of the subseasonal-to-seasonal (S2S) timescales detected here. Meanwhile, the NAO index, representing circulation over the North Atlantic with impacts on weather across the continent, is also detected as an important predictor on the S2S timescale (20–55 days). Persistence of NAO (specifically, of the positive phase), and of the blocking with which it interacts, has been found before severe heatwave events in the region (Drouard et al., 2019; Kueh and Lin, 2020). It is unclear why NAO was also not selected in the 10 days before HWs. Overall, the selected circulation-based features (in z500 and NAO) likely represent precursor wave trains and potentially very persistent blocking (Domeisen et al., 2023), but the exact mechanisms require further inspection. This shows how the framework can provide motivation and direction for further analysis of extreme event drivers.

Other selected features include the IOD, 20–55 days prior, another interconnection with influences on European climate and extremes in particular (Behera et al., 2013). Lastly, the DOY and, to a lesser extent, the global mean atmospheric CO_2 concentration are widely selected and therefore considered important for the algorithm to identify HWs, but the selection of a specific lag time for each is considered arbitrary.

The next step involves determining the variables that contribute robustly as potential drivers. For this purpose, a threshold has been established in the frequency map of the selected variable, such that only those time steps of each variable selected more frequently than the threshold are considered. As the threshold value increases, the number of predictor variables considered decreases, ultimately isolating those consistently present variables in nearly all of the best solutions. The impact of varying the input variables on classifier performance can be observed in Table 3, which presents the evolution of the F1-score metric on the test dataset as the threshold increases. The threshold is expressed in percentiles: for instance, a threshold of 0.5 indicates selecting features that appear in at least 50% of the solutions depicted in the heat map shown in Fig. 6. Fig. 7 depicts how the performance of the LR classifier improves with the threshold until an upper limit (0.85, i.e. 85% of the time is selected), beyond which predictions worsen.

The group of input variables that are involved in the optimum threshold (0.85, Fig. 8(c)) is the same as previously indicated when analyzing Fig. 6, with some variables concerning the local conditions at short term, and other variables involving broader geographical scales and remote regions with predictive skill on the medium and long ranges (e.g. SSTs).

Table 3
Evolution of test F1-scores when increasing the agreement threshold in the selected drivers across the experiments.

Threshold	Test F1-score	CV F1-score	N° of features
0.50	0.6785	0.6374	585
0.75	0.7389	0.6377	214
0.85	0.7615	0.6475	146
0.95	0.7462	0.6147	105

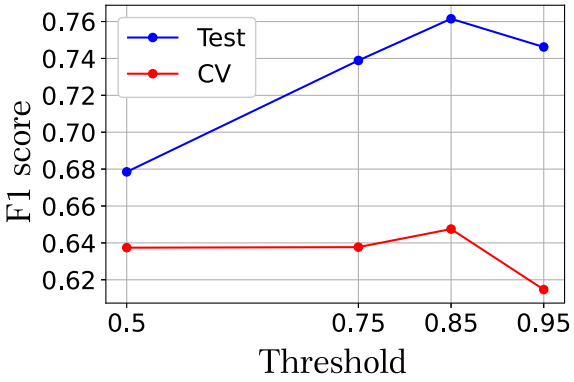

Fig. 7. Evolution of test F1-scores in HW detection when increasing the agreement threshold (reducing the number of selected drivers).

4.2. Results for different machine learning models

Finally, the optimum combination of drivers, corresponding to a threshold equal to 0.85, is used to train a pool of ML classifiers (Section 3.3). The test error metrics for these methods are shown in Table 4. The predictive results provided by the best-performing model, GB (an F1-score of 0.7906), are further detailed. Firstly, Fig. 9 illustrates the

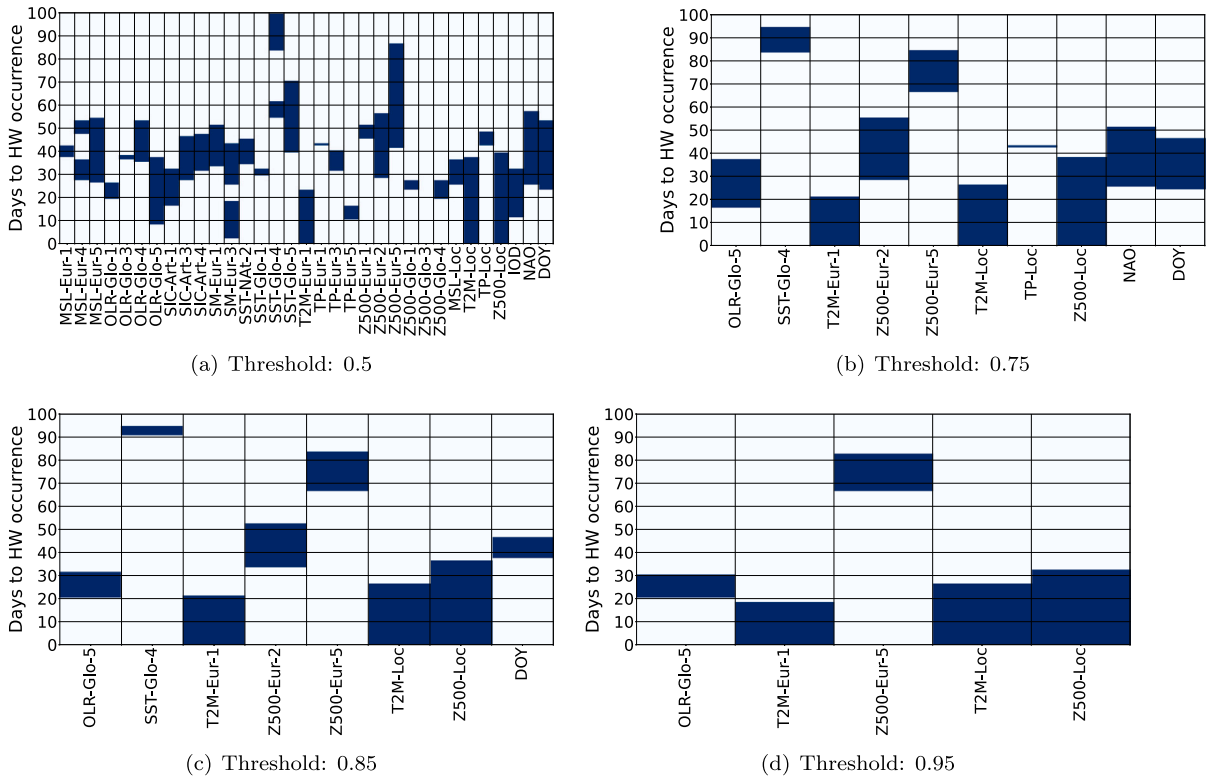

Fig. 8. Features that are most often chosen by the solutions proposed by the optimization algorithm, in the case of HW index. A threshold of 0.5 represents that these variables appear in at least 50% of the 15000 best solutions.

Table 4
Test error metrics for the different ML classification methods assessed, considering the HW index as the target variable. Models are trained with the selected features corresponding to threshold 0.85.

	Recall	Precision	F1-score
LR	0.7436	0.7803	0.7615
LGBM	0.7094	0.8177	0.7597
SVC	0.7051	0.7933	0.7466
DT	0.7350	0.7818	0.7577
RF	0.6068	**0.8658**	0.7136
GNB	**0.9658**	0.3435	0.5067
KNN	0.0983	0.7931	0.1749
AB	0.7009	0.7923	0.7438
MLP	0.6966	0.7376	0.7165
GB	0.7906	0.7906	**0.7906**
ELM	0.1752	0.7736	0.2857

strong performance of the classification task, highlighting the model's ability to accurately detect HW days across the test domain. Although some false alarms are present (instances where the model predicts an HW event that does not occur) the majority of actual HW events are successfully captured. Furthermore, in order to assess the model's temporal consistency, Fig. 10 presents the aggregated HW events at both monthly and yearly resolutions, comparing the actual number of HW occurrences with the corresponding predictions from the GB model. A high correlation is observed between the actual and predicted time series at both time scales, supporting the model's reliability over time.

5. Conclusions and further research

Understanding the drivers behind the formation of Heatwaves (HWs) is vital to enhance our ability to anticipate, forecast and mitigate the impacts of these extreme events, ultimately reducing the risks to human health, economies, and ecosystems. Researchers can develop more sophisticated models that improve predictive capabilities by recognizing the complex interactions between atmospheric conditions, oceanic patterns, and terrestrial processes. This enhanced understanding also informs the development of effective adaptation and mitigation strategies, ensuring societies are better equipped to handle the increasing frequency and severity of HWs driven by climate change.

This study proposes a comprehensive framework to investigate the interactions between HWs and potential physical drivers across multiple spatio-temporal scales (STCO-FS). The proposed methodology follows a two-phase approach: initially, a clustering algorithm is applied to reduce spatial dimensionality by grouping similar time series data from the ERA5 reanalysis database. Additional variables, such as climate indices and local meteorological factors, are then incorporated into the database. In the second phase, a multi-method ensemble evolutionary algorithm (PCRO-SL) identifies significant periods and clusters relevant to HW occurrences. This approach determines which variables are crucial for HW prediction and the specific time frames in which they are most influential, distinguishing between short-term and long-term drivers.

The framework has been successfully applied to an agriculturally intensive region in North Italy, demonstrating its ability to detect key HW drivers effectively. A standard definition of HW has been considered. Regarding the potential drivers considered, 8 variables covering different geographic domains, together with three climate indexes (ENSO, NAO and IOD) and local meteorological conditions, have been incorporated into the study. The results indicate strong HW detection capabilities, with nowcasting error metrics of 0.8363.

For the specific drivers identified in each case, relationships have been established between the occurrence of heat waves and various variables across different spatio-temporal scales. The key selected predictors for the Adda River basin are regional-scale temperature and atmospheric circulation in the 20 days before the event and ENSO, IOD and NAO on sub-seasonal to seasonal timescales. While some of these connections have been suggested in previous work, this study indicates

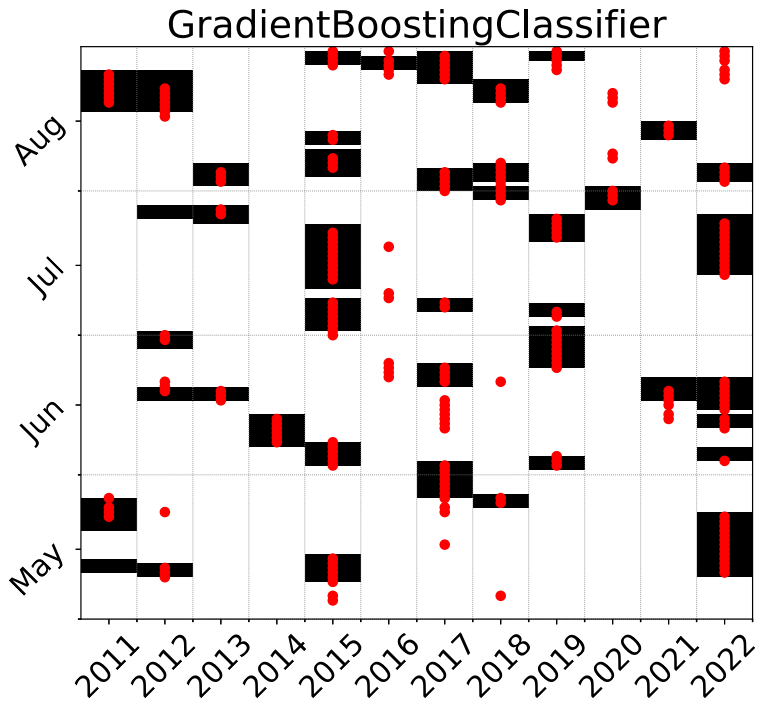

Fig. 9. Results from the top-performing ML classifier (GB) over the test period, considering the HW index as the target variable. Black boxes correspond to days with extreme temperature, while red circles denote days predicted by the model as extreme.

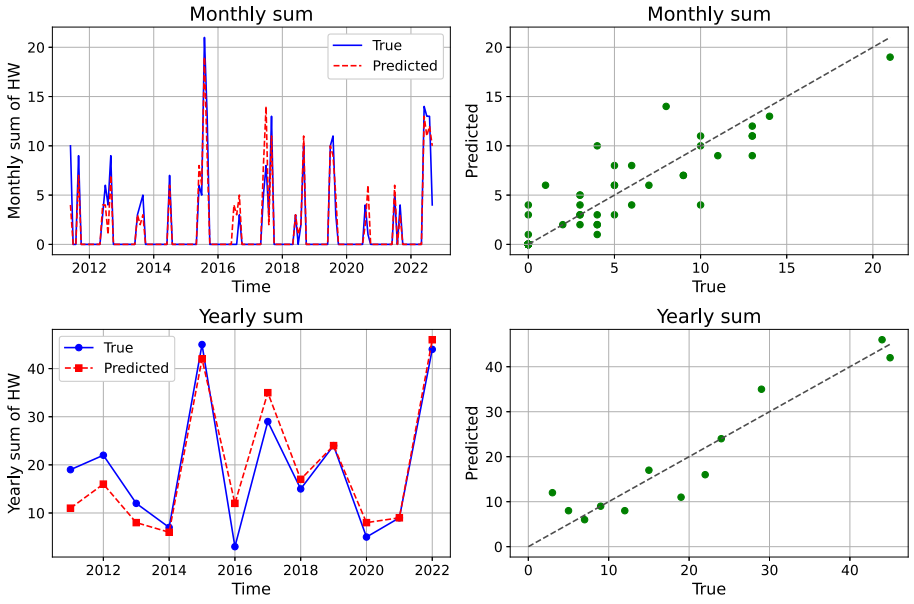

Fig. 10. Results provided by the GB model over the test period, considering the HW index as the target variable. Each year, the frequency of actual and predicted HW days is accumulated monthly (top figures) and seasonal (bottom figures). The figures on the left represent the real and predicted time series. The right figures show the scatter plot between the actual vs predicted HW days, featuring a high correlation R2 value of 0.80 for the monthly scale and of 0.74 for the anual.

the locations and time frames in which specific mechanisms can be studied and is thus a powerful tool and first step in understanding processes and predictability.

The proposed method offers a series of advantages that are outlined as follows: (1) It enables the discovery of new potential drivers in both temporal and spatial domains simultaneously; (2) By grouping meteorological variables into clusters, the dimensionality of the problem is significantly reduced, and the level of granularity can be adjusted as a parameter of the algorithm; (3) Encoding the problem by establishing lag time and the length of the selected window for each predictor variable reduces the dimensionality in the time domain, thereby simplifying the evolutionary process; (4) The use of a robust evolutionary algorithm (PCRO-SL), along with a fast, efficient, and deterministic classifier, allows the resolution of a complex optimization problem within a short period.

We also highlight some potential weak aspects of the proposed approach, which could be further studied and improved. First, note that the proposed method does not present a physical relationship between the occurrence of heatwave and heatwave drivers, i.e., it does not explain the link between predictor and target but only identifies it (which is also an important step for future studies of mechanisms and processes). The clustering task could also be improved with other

methods or including some other metrics in the measurement of the distance between samples since the regions are defined based on the climatological behavior of the predictor fields, with little information about the target (local heatwaves) and the relationships between the predictors (e.g., non-stationary interactions, the multiplicity of optimal solutions depending on the state of specific drivers, etc.)

Future lines of research will focus on several key areas. First, the framework will be extended to predict HWs with a longer prediction horizon (i.e., S2S) and to tackle predicting the number of HWs on a seasonal scale. Additionally, future research will examine how the selected drivers vary depending on the region under investigation and the data they are applied to (e.g., historical or future climate simulations). This continued research will further enhance the predictive power and applicability of the framework, contributing to more effective HW management and mitigation strategies. Moreover, given the flexibility and modularity of the framework, both predictors and target data can be changed, meaning it can be applied to other extremes with ease. Finally, following a previous note on potential weaknesses on the method, how the clusters are created can also be analyzed and improved. Unconnected clusters often arise from variability unrelated to the seasonal cycle. (i.e. not removed when calculating the anomaly). A new way to avoid this is being under study.

CRediT authorship contribution statement

J. Pérez-Aracil: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **C. Peláez-Rodríguez:** Writing – original draft, Visualization, Software, Methodology, Investigation, Data curation. **Ronan McAdam:** Supervision, Software, Resources, Data curation. **Antonello Squintu:** Supervision, Software, Formal analysis, Data curation. **Cosmin M. Marina:** Software, Methodology, Data curation. **Eugenio Lorente-Ramos:** Software, Methodology, Data curation. **Niklas Luther:** Resources, Methodology, Data curation. **Verónica Torralba:** Resources, Methodology, Investigation, Data curation. **Enrico Scoccimarro:** Writing – review & editing, Supervision, Funding acquisition, Formal analysis. **Leone Cavicchia:** Supervision, Methodology, Investigation. **Matteo Giuliani:** Resources, Methodology, Investigation. **Eduardo Zorita:** Supervision, Resources, Methodology, Investigation. **Felicitas Hansen:** Supervision, Software, Investigation. **David Barriopedro:** Software, Methodology, Investigation. **Ricardo García-Herrera:** Supervision, Investigation. **Pedro A. Gutiérrez:** Supervision, Software, Methodology. **Jürg Luterbacher:** Writing – review & editing, Supervision, Investigation. **Elena Xoplaki:** Writing – review & editing, Supervision. **Andrea Castelletti:** Writing – review & editing, Supervision, Funding acquisition, Data curation. **S. Salcedo-Sanz:** Writing – review & editing, Validation, Supervision, Funding acquisition, Formal analysis.

Code and data availability

https://github.com/GheodeAI/STCO-FS.git

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work has been partially supported by the European Commission, project "CLImate INTelligence: Extreme events detection, attribution and adaptation design using machine learning, CLINT" (grant ref.: H2020-LC-CLA-2020-2, 101003876), by the "Agencia Estatal de Investigación (España)", Spanish Ministry of Research and Innovation through NEXO project (grant ref.: PID2023-150663NB-C21), and the European Union through the European Regional Development Fund; "project IA-CONV" with ref. CM/DEMG/2024-039 Funded by the Community of Madrid through the grant agreement for the promotion and advancement of research and technology transfer at the University of Alcalá. Jürg Luterbacher and Elena Xoplaki acknowledge supported by the Horizon Europe project MedEWSa (grant no. 101121192). Verónica Torralba acknowledges the Beatriu de Pinós program (2022 BP 00227) and the Ministry of Research and Universities of the Government of Catalonia. All authors have read and approved the final manuscript.

Data availability

The datasets and code used in this study are available at GitHub repository.

References

Ardilouze, C., Batté, L., Déqué, M., van Meijgaard, E., van den Hurk, B., 2019. Investigating the impact of soil moisture on European summer climate in ensemble numerical experiments. Clim. Dyn. 52, 4011–4026.

Asadollah, S.B.H.S., Khan, N., Sharafati, A., Shahid, S., Chung, E.-S., Wang, X.-J., 2021. Prediction of heat waves using meteorological variables in diverse regions of Iran with advanced machine learning models. Stoch. Environ. Res. Risk Assess. 1–16.

Barriopedro, D., Fischer, E.M., Luterbacher, J., Trigo, R.M., García-Herrera, R., 2011. The hot summer of 2010: redrawing the temperature record map of Europe. Science 332 (6026), 220–224.

Barriopedro, D., García-Herrera, R., Ordóñez, C., Miralles, D., Salcedo-Sanz, S., 2023. Heat waves: Physical understanding and scientific challenges. Rev. Geophys. e2022RG000780.

Battisti, D.S., Naylor, R.L., 2009. Historical warnings of future food insecurity with unprecedented seasonal heat. Science 323 (5911), 240–244.

Behera, S., Ratnam, J.V., Masumoto, Y., Yamagata, T., 2013. Origin of extreme summers in Europe: the indo-Pacific connection. Clim. Dyn. 41, 663–676.

Bischof, S., Pilch Kedzierski, R., Hänsch, M., Wahl, S., Matthes, K., 2023. The role of the North Atlantic for heat wave characteristics in Europe, an ECHAM6 study. Geophys. Res. Lett. 50 (23), e2023GL105280.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Brunner, L., Schaller, N., Anstey, J., Sillmann, J., Steiner, A.K., 2018. Dependence of present and future European temperature extremes on the location of atmospheric blocking. Geophys. Res. Lett. 45 (12), 6311–6320.

Brunner, L., Voigt, A., 2024. Pitfalls in diagnosing temperature extremes. Nat commun15: 2087. Nat. Commun. 15, 1–9.

Buschow, S., Keller, J., Wahl, S., 2023. Explaining heatwaves with machine learning. arXiv preprint arXiv:2305.15170.

Chaqdid, A., Tuel, A., El Fatimy, A., El Moçayd, N., 2023. Extreme rainfall events in Morocco: Spatial dependence and climate drivers. Weather. Clim. Extrem. 40, 100556.

Coumou, D., Robinson, A., 2013. Historic and future increase in the global land area affected by monthly heat extremes. Environ. Res. Lett. 8 (3), 034018.

Dalal, G., Pathania, T., Koppa, A., Hari, V., 2024. Drivers and mechanisms of heatwaves in south west India. Clim. Dyn. 1–15.

Del Ser, J., Osaba, E., Molina, D., Yang, X.-S., Salcedo-Sanz, S., Camacho, D., Das, S., Suganthan, P.N., Coello, C.A.C., Herrera, F., 2019. Bio-inspired computation: Where we stand and what's next. Swarm Evol. Comput. 48, 220–250.

Della-Marta, P.M., Haylock, M.R., Luterbacher, J., Wanner, H., 2007. Doubled length of western European summer heat waves since 1880. J. Geophys. Res.: Atmospheres 112 (D15).

Domeisen, D.I., Eltahir, E.A., Fischer, E.M., Knutti, R., Perkins-Kirkpatrick, S.E., Schär, C., Seneviratne, S.I., Weisheimer, A., Wernli, H., 2023. Prediction and projection of heatwaves. Nat. Rev. Earth Environ. 4 (1), 36–50.

Drouard, M., Kornhuber, K., Woollings, T., 2019. Disentangling dynamic contributions to summer 2018 anomalous weather over Europe. Geophys. Res. Lett. 46 (21), 12537–12546.

Duchez, A., Frajka-Williams, E., Josey, S.A., Evans, D.G., Grist, J.P., Marsh, R., McCarthy, G.D., Sinha, B., Berry, D.I., Hirschi, J.J., 2016. Drivers of exceptionally cold North Atlantic ocean temperatures and their link to the 2015 European heat wave. Environ. Res. Lett. 11 (7), 074004.

Easterling, D.R., Evans, J., Groisman, P.Y., Karl, T.R., Kunkel, K.E., Ambenje, P., 2000. Observed variability and trends in extreme climate events: a brief review. Bull. Am. Meteorol. Soc. 81 (3), 417–426.

Eurostat, 2024. Agri-environmental indicator – irrigation. https://ec.europa.eu/eurostat/statistics-explained/index.php/Agri-environmental_indicator_-_irrigation#Analysis_at_regional_level. (Accessed 7 November 2024).

Fischer, E.M., Beyerle, U., Knutti, R., 2013. Robust spatially aggregated projections of climate extremes. Nat. Clim. Chang. 3 (12), 1033–1038.

Freund, Y., Schapire, R., 1996. Experiments with a new boosting algorithm. In: Machine Learning: Proceedings of the Thirteenth International Conference. pp. 148–156.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232.

García-Herrera, R., Díaz, J., Trigo, R.M., Luterbacher, J., Fischer, E.M., 2010. A review of the European summer heat wave of 2003. Crit. Rev. Environ. Sci. Technol. 40 (4), 267–306.

Gardner, M.W., Dorling, S.R., 1998. Artificial neural networks (multilayer perceptron)– A review of applications in the atmospheric sciences. Atmos. Environ. 32, 2627–2636.

Gastineau, G., Frankignoul, C., 2015. Influence of the North Atlantic SST variability on the atmospheric circulation during the twentieth century. J. Clim. 28 (4), 1396–1416.

Giuliani, M., Zaniolo, M., Castelletti, A., Davoli, G., Block, P., 2019. Detecting the state of the climate system via artificial intelligence to improve seasonal forecasts and inform reservoir operations. Water Resour. Res. 55 (11), 9133–9147.

Goutte, C., Gaussier, E., 2005. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In: European Conference on Information Retrieval. Springer, pp. 345–359.

Hagen, J.S., Leblois, E., Lawrence, D., Solomatine, D., Sorteberg, A., 2021. Identifying major drivers of daily streamflow from large-scale atmospheric circulation with machine learning. J. Hydrol. 596, 126086.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., et al., 2018. ERA5 hourly data on single levels from 1979 to present. Copernic. Clim. Chang. Serv. (C3S) Clim. Data Store (CDS) 10.

Hoskins, B., Woollings, T., 2015. Persistent extratropical regimes and climate extremes. Curr. Clim. Chang. Rep. 1, 115–124.

Huang, G.B., Zhu, Q.Y., Siew, C.K., 2006. Extreme learning machine: Theory and applications. Neurocomputing 70, 489–501. http://dx.doi.org/10.1016/j.neucom.2005.12.126.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. Lightgbm: A highly efficient gradient boosting decision tree. Adv. Neural Inf. Process. Syst. 30.

Kenyon, J., Hegerl, G.C., 2008. Influence of modes of climate variability on global temperature extremes. J. Clim. 21 (15), 3872–3889.

Kirkpatrick, S., Gelatt, Jr., C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220 (4598), 671–680.

Kleinbaum, D.G., Dietz, K., Gail, M., Klein, M., Klein, M., 2002. Logistic Regression. Springer.

Kueh, M.-T., Lin, C.-Y., 2020. The 2018 summer heatwaves over northwestern Europe and its extended-range prediction. Sci. Rep. 10 (1), 19283.

Kuglitsch, F.G., Toreti, A., Xoplaki, E., Della-Marta, P.M., Zerefos, C.S., Türkeş, M., Luterbacher, J., 2010. Heat wave changes in the eastern Mediterranean since 1960. Geophys. Res. Lett. 37 (4).

Lemordant, L., Gentine, P., Stéfanon, M., Drobinski, P., Fatichi, S., 2016. Modification of land-atmosphere interactions by CO2 effects: Implications for summer dryness and heat wave amplitude. Geophys. Res. Lett. 43 (19), 10–240.

Li, M., Yao, Y., Simmonds, I., Luo, D., Zhong, L., Chen, X., 2020. Collaborative impact of the NAO and atmospheric blocking on European heatwaves, with a focus on the hot summer of 2018. Environ. Res. Lett. 15 (11), 114003.

Li, P., Yu, Y., Huang, D., Wang, Z.-H., Sharma, A., 2023. Regional heatwave prediction using graph neural network and weather station data. Geophys. Res. Lett. 50 (7), e2023GL103405.

Loh, W.-Y., 2011. Classification and regression trees. Wiley Interdiscip. Rev.: Data Min. Knowl. Discov. 1 (1), 14–23.

Loughran, T.F., Perkins-Kirkpatrick, S.E., Alexander, L.V., 2017. Understanding the spatio-temporal influence of climate variability on Australian heatwaves. Int. J. Climatol. 37 (10), 3963–3975.

Luo, M., Lau, N.-C., 2019. Amplifying effect of ENSO on heat waves in China. Clim. Dyn. 52 (5), 3277–3289.

Luo, M., Lau, N.-C., 2020. Summer heat extremes in northern continents linked to developing ENSO events. Environ. Res. Lett. 15 (7), 074042.

MacQueen, J., et al., 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, Oakland, CA, USA, pp. 281–297.

Martija-Díez, M., Rodríguez-Fonseca, B., López-Parages, J., 2021. ENSO influence on western European summer and fall temperatures. J. Clim. 34 (19), 8013–8031.

Meehl, G.A., Tebaldi, C., 2004. More intense, more frequent, and longer lasting heat waves in the 21st century. Science 305 (5686), 994–997.

Miralles, D.G., Gentine, P., Seneviratne, S.I., Teuling, A.J., 2019. Land–atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges. Ann. New York Acad. Sci. 1436 (1), 19–35.

Mucherino, A., Papajorgji, P.J., Pardalos, P.M., 2009. K-nearest neighbor classification. In: Data Mining in Agriculture. Springer, pp. 83–106.

Mukherjee, S., Ashfaq, M., Mishra, A.K., 2020. Compound drought and heatwaves at a global scale: The role of natural climate variability-associated synoptic patterns and land-surface energy budget anomalies. J. Geophys. Res.: Atmospheres 125 (11), e2019JD031943.

Orimoloye, I.R., Olusola, A.O., Belle, J.A., Pande, C.B., Ololade, O.O., 2022. Drought disaster monitoring and land use dynamics: identification of drought drivers using regression-based algorithms. Nat. Hazards 112 (2), 1085–1106.

Pérez-Aracil, J., Camacho-Gómez, C., Lorente-Ramos, E., Marina, C.M., Cornejo-Bueno, L.M., Salcedo-Sanz, S., 2023. New probabilistic, dynamic multi-method ensembles for optimization based on the CRO-SL. Mathematics 11 (7), 1666.

Perkins, S.E., 2015. A review on the scientific understanding of heatwaves-their measurement, driving mechanisms, and changes at the global scale. Atmos. Res. 164, 242–267.

Perkins, S.E., Argüeso, D., White, C.J., 2015. Relationships between climate variability, soil moisture, and Australian heatwaves. J. Geophys. Res.: Atmospheres 120 (16), 8144–8164.

Prodhomme, C., Materia, S., Ardilouze, C., White, R.H., Batté, L., Guemas, V., Fragkoulidis, G., García-Serrano, J., 2021. Seasonal prediction of European summer heatwaves. Clim. Dyn. 1–18.

Reddy, P.J., Perkins-Kirkpatrick, S.E., Sharples, J.J., 2021. Interactive influence of ENSO and IOD on contiguous heatwaves in Australia. Environ. Res. Lett. 17 (1), 014004.

Russo, S., Dosio, A., Graversen, R.G., Sillmann, J., Carrao, H., Dunbar, M.B., Singleton, A., Montagna, P., Barbola, P., Vogt, J.V., 2014. Magnitude of extreme heat waves in present climate and their projection in a warming world. J. Geophys. Res.: Atmospheres 119 (22), 12–500.

Russo, S., Sillmann, J., Fischer, E.M., 2015. Top ten European heatwaves since 1950 and their occurrence in the coming decades. Environ. Res. Lett. 10 (12), 124003.

Salcedo-Sanz, S., 2017. A review on the coral reefs optimization algorithm: new development lines and current applications. Prog. Artif. Intell. 6 (1), 1–15.

Salcedo-Sanz, S., Del Ser, J., Landa-Torres, I., Gil-López, S., Portilla-Figueras, J., 2014. The coral reefs optimization algorithm: a novel metaheuristic for efficiently solving optimization problems. Sci. World J. 2014.

Sánchez-Benítez, A., Barriopedro, D., García-Herrera, R., 2020. Tracking iberian heatwaves from a new perspective. Weather. Clim. Extrem. 28, 100238.

Sasaki, Y., et al., 2007. The truth of the F-measure. Teach Tutor. Mater. 1 (5), 1–5.

Schölkopf, B., Smola, A.J., Williamson, R.C., Bartlett, P.L., 2000. New support vector algorithms. Neural Comput. 12, 1207–1245. http://dx.doi.org/10.1162/089976600300015565.

Seneviratne, S.I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskander, I., Kossin, J., Lewis, S., et al., 2021. Weather and Climate Extreme Events in a Changing Climate (Chapter 11). Cambridge University Press.

Sillmann, J., Kharin, V.V., Zwiers, F.W., Zhang, X., Bronaugh, D., 2013. Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. J. Geophys. Res.: Atmospheres 118 (6), 2473–2493.

Sillmann, J., Thorarinsdottir, T., Keenlyside, N., Schaller, N., Alexander, L.V., Hegerl, G., Seneviratne, S.I., Vautard, R., Zhang, X., Zwiers, F.W., 2017. Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities. Weather. Clim. Extrem. 18, 65–74.

Sousa, P.M., Trigo, R.M., Barriopedro, D., Soares, P.M., Santos, J.A., 2018. European temperature responses to blocking and ridge regional patterns. Clim. Dyn. 50, 457–477.

Stefanon, M., D'Andrea, F., Drobinski, P., 2012. Heatwave classification over Europe and the Mediterranean region. Environ. Res. Lett. 7 (1), 014023.

Stott, P.A., Jones, G.S., Christidis, N., Zwiers, F.W., Hegerl, G., Shiogama, H., 2011. Single-step attribution of increasing frequencies of very warm regional temperatures to human influence. Atmos. Sci. Lett. 12 (2), 220–227.

Sun, J., Liu, S., Cohen, J., Yu, S., 2022. Influence and prediction value of arctic sea ice for spring eurasian extreme heat events. Commun. Earth Environ. 3 (1), 172.

Tao, H., Awadh, S.M., Salih, S.Q., Shafik, S.S., Yaseen, Z.M., 2022. Integration of extreme gradient boosting feature selection approach with machine learning models: application of weather relative humidity prediction. Neural Comput. Appl. 34 (1), 515–533.

Torralba, V., Materia, S., Cavicchia, L., Álvarez-Castro, M.C., Prodhomme, C., McAdam, R., Scoccimarro, E., Gualdi, S., 2024. Nighttime heat waves in the Euro-Mediterranean region: definition, characterisation, and seasonal prediction. Environ. Res. Lett. 19 (3), 034001.

Wehrli, K., Guillod, B.P., Hauser, M., Leclair, M., Seneviratne, S.I., 2019. Identifying key driving processes of major recent heat waves. J. Geophys. Res.: Atmospheres 124 (22), 11746–11765.

Wu, G., Mallipeddi, R., Suganthan, P.N., 2019. Ensemble strategies for population-based optimization algorithms–A survey. Swarm Evol. Comput. 44, 695–711.

Zhang, H., 2004. The optimality of naive Bayes. Aa 1 (2), 3, URL www.aaai.org.

Zhang, J., Yang, Z., Wu, L., 2018. Skillful prediction of hot temperature extremes over the source region of ancient silk road. Sci. Rep. 8, 6677.

Zittis, G., Almazroui, M., Alpert, P., Ciais, P., Cramer, W., Dahdal, Y., Fnais, M., Francis, D., Hadjinicolaou, P., Howari, F., et al., 2022. Climate change and weather extremes in the eastern Mediterranean and middle east. Rev. Geophys. 60 (3), e2021RG000762.