

LETTER • OPEN ACCESS

Tracking shifts in European drought hotspots

To cite this article: Martina Merlo *et al* 2026 *Environ. Res. Lett.* **21** 064006

View the [article online](#) for updates and enhancements.

You may also like

- [JWST Observations Rule Out 2032 Lunar Impact for Asteroid 2024 YR4](#)
A. S. Rivkin, J. de Wit, M. Micheli et al.
- [A Strong Anticorrelation between Sunspot Maximum and the Lead Time from Hemispheric-average First-spot Emergence to Maximum Epoch: Significant Only in Even Cycles?](#)
Jean-Guillaume Richard
- [Photometric Study of Type Ibn SN 2022ablg Based on g, r, i-band Observations](#)
M. Soliman, F. Kerrache, K. A. K. Gadallah et al.

ENVIRONMENTAL RESEARCH
LETTERS

LETTER

Tracking shifts in European drought hotspots

Martina Merlo¹ , Matteo Giuliani^{1,2,*} , Yiheng Du³ , Ilias Pechlivanidis³  and Andrea Castelletti^{1,2} ¹ Department of Electronics, Information, and Bioengineering Politecnico di Milano, Milano, Italy² CMCC Foundation—Euro-Mediterranean Center on Climate Change, Milano, Italy³ Hydrology Research Unit, Swedish Meteorological and Hydrological Institute (SMHI), Norrköping, Sweden

* Author to whom any correspondence should be addressed.

E-mail: matteo.giuliani@polimi.it**Keywords:** drought detection, climate change, machine learning, impact-based combined drought indicesSupplementary material for this article is available [online](#)

OPEN ACCESS

RECEIVED

13 November 2025

REVISED

2 March 2026

ACCEPTED FOR PUBLICATION

3 March 2026

PUBLISHED

17 March 2026

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Abstract

Drought is a complex hazard with far-reaching socio-economic and ecological impacts that are often inadequately captured by conventional indices based solely on hydroclimatic anomalies. To bridge this gap, this study develops novel impact-based combined drought indices (iCDIs) that directly relate hydroclimatic drivers to remotely sensed vegetation stress, here used as a proxy for observed agricultural impacts. The proposed framework employs a machine learning framework to optimize the selection of hydroclimatic predictors including precipitation, temperature, and soil moisture across distinct hydrological clusters of the European domain. This bottom-up machine learning-based approach outperforms traditional indices in reproducing observed vegetation responses across European river basins. Crucially, future projections of iCDIs under warming scenarios unveil a distinct northward shift in drought impacts, with Central Europe emerging as an increasingly vulnerable hotspot. Notably, this contrasts with patterns suggested by standardized indices, which typically emphasize Mediterranean regions as the primary areas of concern. This spatial reconfiguration of risk underscores the urgency of adopting impact-oriented drought diagnostics to support climate-resilient planning and adaptive risk governance in a rapidly changing climate.

1. Introduction

Drought is a slow-onset natural hazard affecting all climatic regions, with wide-ranging socio-economic and ecological consequences. It propagates through the hydrological cycle, from precipitation deficits to soil moisture depletion, reduced streamflow, and groundwater decline, often intensified by rising temperatures [1, 2]. Its gradual onset, long duration, and spatially heterogeneous impacts have hindered the formulation of a unique and operational definition [3], leading to multiple drought characterizations depending on the variable and time scale considered [4–7].

The impacts of drought events extend across many water-dependent sectors, including agriculture, industry, energy, and public supply [7, 8], with severe associated economic losses estimated at €9 billion per year in the EU and UK (1981–2010

baseline) [9]. Among the affected sectors, agriculture and vegetation-dependent ecosystems are particularly vulnerable, as they respond rapidly to water stress through reduced photosynthetic activity, yield losses, and increased vegetation mortality [10–12]. Recent European droughts, such as those in 2018 and 2022, have shown how vegetation stress cascades into major socio-economic impacts, disrupting food production, forestry, and rural economies. With global warming projected to intensify drought severity [13–15], the ability to detect, quantify, and predict vegetation stress becomes central to drought risk assessment and adaptation planning. Annual drought-related damages in the EU and UK are projected to reach €24.7 billion and €45.4 billion by 2100 under 1.5 °C and 3 °C warming levels, respectively, an increase of 170% and 400% relative to the historical baseline, a large share of which is expected to stem from agricultural and ecosystem losses [9].

To reduce these costs, the first crucial step is to characterize the drought itself. Effective drought characterization enables the development of early warning systems and risk analysis, thereby improving preparedness and adaptation strategies [16, 17]. Various methodologies exist for this purpose, with drought indices being among the most widely used. Model-based indices (e.g. the palmer drought severity index, [18]) and statistical drought indices (e.g. standardized precipitation index, SPI [19]); standardized precipitation and evapotranspiration index, SPEI [20]; standardized streamflow index, SSI [21]; soil moisture anomaly, SMA [22]) are designed to describe hydroclimatic anomalies by measuring the departure of a physical variable from its historical norm. However, these anomalies are not necessarily representative of actual drought impacts. In many cases, the drought signal is decoupled from meteorological drivers, particularly in highly regulated systems where water storage, allocation rules, and human interventions can delay, attenuate, or redistribute effects across time and space [23, 24]. To overcome the limitations of univariate indices, multivariate drought indices have been proposed to integrate multiple hydroclimatic variables into a single metric, thereby providing a more comprehensive representation of drought evolution. Examples include the multivariate standardized drought index ([25]), the hybrid drought index ([26]), and combined drought index (CDI [27]). Although these approaches better capture compound drought conditions, they generally remain impact-agnostic, as they rely on predefined combinations of physical variables and rarely incorporate direct observations of ecosystem or socio-economic responses.

Remote sensing-based indices provide a complementary approach, offering a powerful tool for large-scale, (near) real-time monitoring of drought evolution and, potentially, associated impacts. Metrics like the fraction of absorbed photosynthetically active radiation anomaly (FAPAN) and the normalized difference vegetation index offer global coverage and long-term consistency, enabling retrospective analyses of drought-induced vegetation stress [28]. However, while these indices effectively track environmental responses to drought, translating them into actual socio-economic or ecological impacts remains a challenge due to the scarcity and inconsistency of impact data. Databases such as (emergency events database [29]), GDIS (geocoded disasters dataset [30]) and (European drought impact database [31]), as well as more recent text mining approaches [32], aim at documenting drought-related damages, but they are still limited in terms of temporal resolution and sectoral detail needed for developing a solid drought risk assessment. Moreover, both remote sensing indices and impact data remain constrained to real-time monitoring and retrospective analysis,

without the capability to project future drought conditions and inform adaptation design.

This paper addresses the limitations of existing drought indices by leveraging machine learning (ML) algorithms to design a set of impact-based combined drought indices (iCDIs) and to project their spatiotemporal evolution under future warming scenarios, thereby tracking shifts in European drought hotspots. In this study, hotspots are defined as regions experiencing the most severe and recurrent drought impacts, as represented by the iCDIs. Instead of categorizing drought into specific types, these indices focus on the cumulative consequences of combined hydroclimatic stressors. To do so, they integrate ML with remote sensing and hydroclimatic data, describing drought impacts in a river basin as a function of a subset of informative hydroclimatic drivers, selected using feature-extraction techniques. Recent advances in ML significantly improved drought analysis by enabling the processing of large, multi-variable datasets (including satellite imagery, weather stations, and historical climate records), uncovering patterns and relationships that traditional methods might overlook, and extracting key features that enhance the accuracy of drought assessments [33, 34]. Although ML-based methods have demonstrated their potential in drought forecasting [35–37], a major challenge remains in connecting these predictions to real-world impacts. Here, we use the framework for index-based drought analysis (FRIDA [38]), a methodology that relies on the wrapper for quasi-equally informative subset selection [39, 40] to fully automate and generalize the variable selection process for constructing basin-specific indices that effectively support the characterization of drought conditions and their associated impacts. In contrast to typical top-down approaches, which assume that the drought processes and drivers are known, FRIDA contributes a bottom-up perspective starting from observed impacts to identify relevant drought drivers and formulate indices [41]. Specifically, we use the remotely sensed FAPAN as the target variable, indicating vegetation stress. Future drought impacts are further explored under warming scenarios using two regional climate projections from the EURO-CORDEX ensemble addressing wet and dry conditions.

2. Methods

2.1. Standardized drought indices

Drought detection is traditionally based on the analysis of multiple standardized drought indices, each representing the conditions of distinct components of the hydrological cycle (e.g. precipitation, soil moisture, streamflow) and associated with specific drought types [16]. These indices quantify statistical anomalies relative to long-term climatology at a given

location and time period, providing a probabilistic measure of drought severity.

In this study, we calculate eight drought indices at a monthly resolution over the period 1993–2018 across 35 408 basins spanning the pan-European domain:

- SPI at 1 month (SPI-1) and 3 month (SPI-3) timescales.
- SPEI at 1 month (SPEI-1) and 3 month (SPEI-3) timescales.
- SMA at 1 month (SMA-1), 3 month (SMA-3), and 6 month (SMA-6) timescales.
- SSI at a 6 month timescale (SSI-6).

All indices are computed following a standardized three-step procedure involving aggregation, distribution fitting, and normalization to a standard normal variable. Detailed information is provided in supplementary text S1.

Each index reflects different drought characteristics and potential impacts: SPI and SPEI, which characterize meteorological drought, are typically computed at short accumulation periods, capturing immediate impacts. SSI, linked to streamflow, characterizes hydrological drought, which evolves over longer timescales. SMA represents agricultural drought, which develops on intermediate time periods. Following McKee *et al* [19], a drought event is defined as a period during which a standardized index remains below -1 for at least two consecutive months ending when it rises above 0.

Computation of these indices requires hydroclimatic time series. Here, we used precipitation and temperature data from HydroGFD (Hydrological Global Forcing Data) version 2.0 reanalysis [42] combined with simulations of evapotranspiration, streamflow, and soil moisture from the European Hydrological Predictions for the Environment model (E-HYPE [43]) forced with HydroGFD2.0. Detailed information is available in the supplementary text S2. Building on the work by Pechlivanidis *et al* [44], the 35 408 E-HYPE basins were clustered based on their hydrological behavior. This process involved categorizing 15 model-based hydrological signatures, resulting in 11 clusters of varying sizes and distributions of these signatures (figure S1 in supplementary material).

2.2. FRIDA

The (FRIDA [38]) was developed to identify and model the links between hydroclimatic variability and vegetation response. The framework integrates multiple standardized drought indices and vegetation anomaly data to quantify drought impacts and to derive new iCDIs that can be projected under future climate scenarios. By explicitly learning the statistical relationships between climatic drivers and observed

vegetation stress, FRIDA aims to enhance drought-impact detection beyond what is achievable with traditional standardized indices alone.

FRIDA is structured into three main steps (a simplified diagram illustrating the logical flow of these steps is provided in figure S2 of the supplementary material):

1. Identification of basin characteristics

The first step of FRIDA involves selecting a target variable to represent drought impacts and assembling a dataset of candidate predictors. In this study, we use the FAPAN index, a standardized anomaly of the fraction of absorbed photosynthetically active radiation (FAPAR) monitored by the European drought observatory, as a proxy for agricultural drought impacts. Crucially, FAPAN is computed by rescaling FAPAR values against their local, time-varying climatology. This pixel-based standardization effectively homogenizes the drought stress signal across diverse biomes and growth stages, allowing for a consistent comparison of relative stress regardless of the specific vegetation type or absolute productivity. Based on an analysis of the time-series distributions, we adopted a threshold of -0.5 to identify impactful events across both historical and future time horizons. Detailed information regarding the FAPAN index is provided in the supplementary text S3.

The dataset of candidate predictors includes both hydro-meteorological variables and standardized indices. In particular, we consider 4 hydro-meteorological variables: precipitation, precipitation minus potential evapotranspiration (PET), soil moisture and streamflow. These variables represent the input for the calculation of the drought indices under analysis. Each of these has been aggregated over 1 and 3 months, except for soil moisture and streamflow which have been also aggregated over 6 months. This temporal aggregation of the variables is implemented as a cumulated value for precipitation and precipitation minus evapotranspiration, while the average is considered for soil moisture and streamflow. In addition, we also include in this dataset 8 different statistical indices, namely: SPI-1, SPI-3, SPEI-1, SPEI-3, SMA-1, SMA-3, SMA-6 and SSI-6. Longer SSI accumulation periods were initially tested but eventually excluded as preliminary feature selection showed they provided no additional explanatory power (see figure S3 in supplementary material for an example).

The resulting dataset is therefore composed of 18 features (table S1 in supplementary material). To align the different spatial scales, all hydro-meteorological variables, standardized indices, and FAPAN values were spatially aggregated for each of the 11 clusters. Moreover, since the FAPAN index is available only from the middle of 2001, the length of the time series is equal to 210 samples (July 2001–December 2018). This window represents the

maximum available record for consistent satellite-derived FAPAR observations, yet it encompasses several major extreme events necessary for robust model training.

2. Feature extraction

The second step of FRIDA focuses on selecting the most relevant subsets of predictors that best explain the target variable. This is achieved using the wrapper for quasi-equally informative subset selection (W-QEISS [39, 40]; see supplementary text S4 for a more detailed description). The wrapper selects Pareto-efficient subsets of predictors based on four metrics—accuracy, cardinality, relevance and redundancy—using the Borg multi-objective evolutionary algorithm [45]. The iterative search concludes upon meeting the termination criterion, i.e. the number of iterations. Low-accuracy solutions are excluded from the pool of Pareto-efficient subsets, resulting in a set of quasi-equally accurate subsets.

3. Drought index modeling

In the final step, a regression model is fitted to the optimal solution identified within the quasi-equally accurate subsets. The choice of the model class is guided by the specific application requirements, balancing accuracy, interpretability and flexibility. To this end, different model structures were considered, including artificial neural networks (ANNs), extreme learning machines (ELMs), and linear models. ANNs offer a strong balance between accuracy and flexibility [46] and remain a cornerstone for identifying complex non-linearities between hydroclimatic drivers and vegetation response [47]. However they are often criticized for their black-box nature, which limits interpretability [48]. ELMs, as a variant of ANNs, provide fast training times while maintaining good predictive performance [49], demonstrating high efficiency in the characterization of drought events [50]. In contrast, linear models are more interpretable and have immediate physical meaning, though they generally have lower approximation capabilities [51]. The final iCDI for each cluster was derived from the model structure that best reproduced the observed FAPAN trajectory using the predictor variables selected by W-QEISS. The performance of the new index was evaluated based on the correlation between predicted and observed FAPAN values, ensuring a reliable representation of drought impacts on vegetation.

2.3. Climate projections

Future drought impacts were assessed using two end-member GCM–RCM combinations from the EURO-CORDEX ensemble based on CMIP5, selected to bracket the range of plausible hydroclimatic conditions: ICHEC-EC-EARTH—RACMO22E (wet scenario) and MPI-M-MPI-ESM-LR—REMO2009 (dry scenario). The rationale for using the CMIP5 ensemble, along with detailed information on climate

model selection and bias-adjustment procedures, is provided in the supplementary text S5.

Projections were limited to 2050 to maintain consistency with current vegetation–climate relationships. This choice avoids uncertainties introduced by long-term adaptation processes (e.g. land-use change or shifts in crop types), which are expected to become more relevant toward the end of the century. Moreover, to mitigate the risk of out-of-distribution extrapolation, the iCDI was recalibrated for future applications using a joint normalization of combined historical (2002–2018) and projected (2019–2050) trajectories.

3. Results

3.1. Limited detection of drought impacts using standardized indices

To assess the ability of traditional indices to detect drought events and their impacts, we first computed a suite of standardized indices (section 2.1) to cover the diverse dynamics of meteorological, agricultural and hydrological droughts. A comparison of drought occurrence (figure S15 in supplementary material) and average duration (figure S16 in supplementary material) across the pan-European domain shows that all indices consistently identify the Baltic Sea region and as experiencing the longest drought between 1993 and 2018. However, the indices show less agreement in capturing drought frequency across Europe. While there is broad coherence in identifying Southern England, Northern France, and Northern Italy as drought-prone regions, other areas, such as Southern Spain, exhibit substantial variability among indices. This variability reflects the multi-faceted nature of droughts, as each standardized index emphasizes different variables and accumulation periods, leading to distinct and often non-overlapping representations of drought occurrence and duration when considered individually.

To further assess the ability of traditional drought indices to capture impacts, we used FAPAN as a proxy for agricultural stress. The reliability of this proxy was validated by comparing FAPAN data with drought events (crop failures and food shortages) reported in the GDIS database (2001–2018), showing a strong correspondence between negative anomalies and documented losses. This is consistent with previous evidence on the sensitivity of FAPAR-based metrics to drought-induced reductions in vegetation productivity and yields [52]. Detailed validation against specific GDIS events, including the 2007 pan-European drought and events in East Africa, is provided in supplementary figure S17.

Using FAPAN as a benchmark, our results provide quantitative evidence of the limited ability of individual standardized indices to capture drought impacts, with average maximum correlations remaining below 0.25 and never exceeding 0.40 (figure 1).

Central Europe exhibits the highest correlation levels, while the Alpine arc, the Mediterranean region, and the far eastern part of Europe show values not exceeding 0.12.

Notably, this analysis uses the best-performing index for each cluster; any other choice of standardized index would yield even poorer results. For instance, in countries with warm and temperate Mediterranean climate (i.e. Spain, Italy, Albania and Greece), the Pearson correlation coefficient drops to just 0.004 when the least effective standardized index is considered (see the example in figure S18 in supplementary material). Overall, these results demonstrate the very limited ability of traditional standardized indices to detect drought related impacts on vegetation.

3.2. iCDIs

Given the low correlation between standardized indices and observed impacts (figure 1), we applied FRIDA (section 2.2) to derive iCDIs. FRIDA employs automated feature-extraction algorithms to identify the most informative combinations of predictors for modeling drought impacts. Predictor selection was performed independently for each cluster of sub-basins, yielding multiple quasi-equally informative subsets of predictors (figures S4–S14 in supplementary material). For example, 13 alternative predictor subsets were identified in the case of Cluster 7 (figure S4). Among these, subset 9 was selected as the preferred solution, as it combines high predictive accuracy with high predictor relevance and low cardinality. This subset includes the SPI and SMA aggregated over one month (SPI-1 and SMA-1 respectively) as the most relevant predictors, along with the mean soil moisture aggregated over 6 months (meanSRFF6) and the SPI-3. The prominence of short-term predictors reflects the strong influence of rapid precipitation and soil moisture dynamics in this cluster, which is characterized by high runoff coefficients and low actual evapotranspiration, resulting in fast hydrological responses [44]. Importantly, this selection does not reflect the physiology of a specific vegetation type, but rather indicates that in these basins any short-term water deficit rapidly translates into departures from normal photosynthetic activity across vegetation types. Moreover, it is important to note that while the standardized indices show limited skill when used individually (figure 1), their relevance emerges when these selected indices are jointly combined with other hydroclimatic variables, allowing the iCDI to capture the compound nature of impact-driven drought dynamics. Similar interpretations apply to the remaining clusters (figures S5–S14 in supplementary material). Using the selected predictors for each cluster, we formulated the iCDI as a

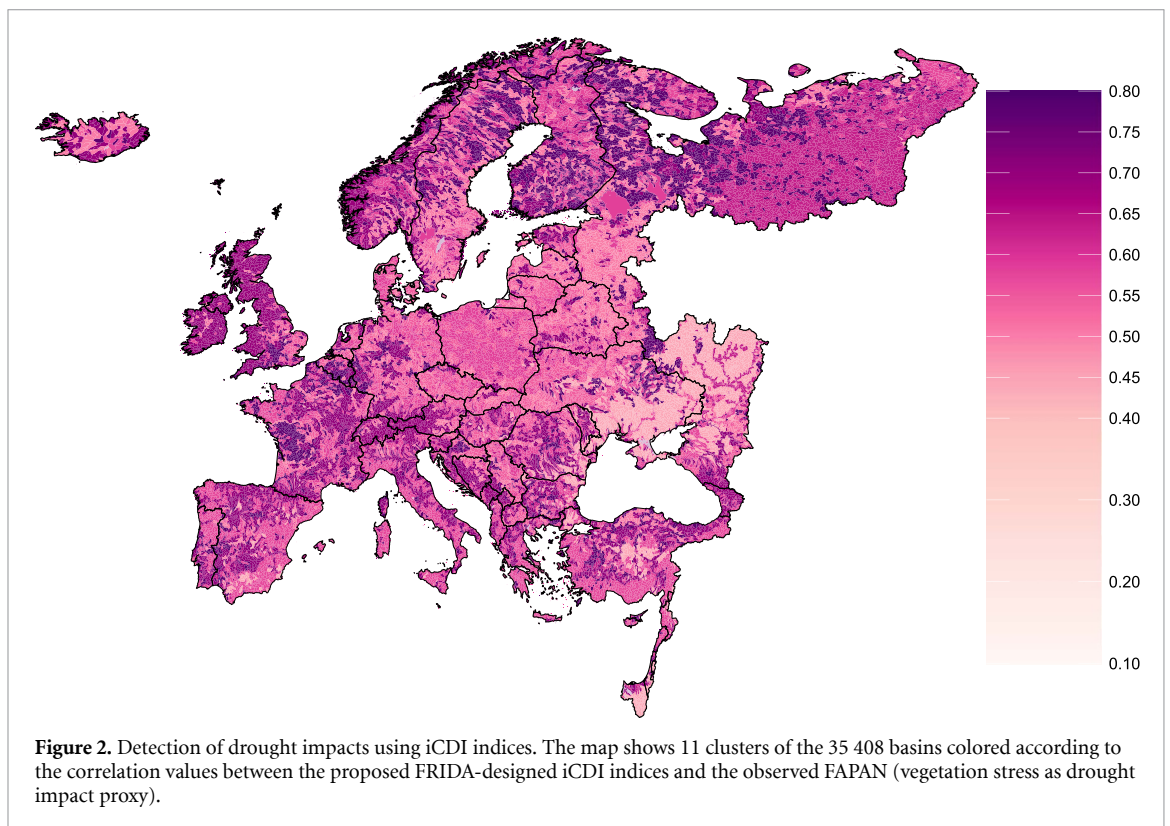
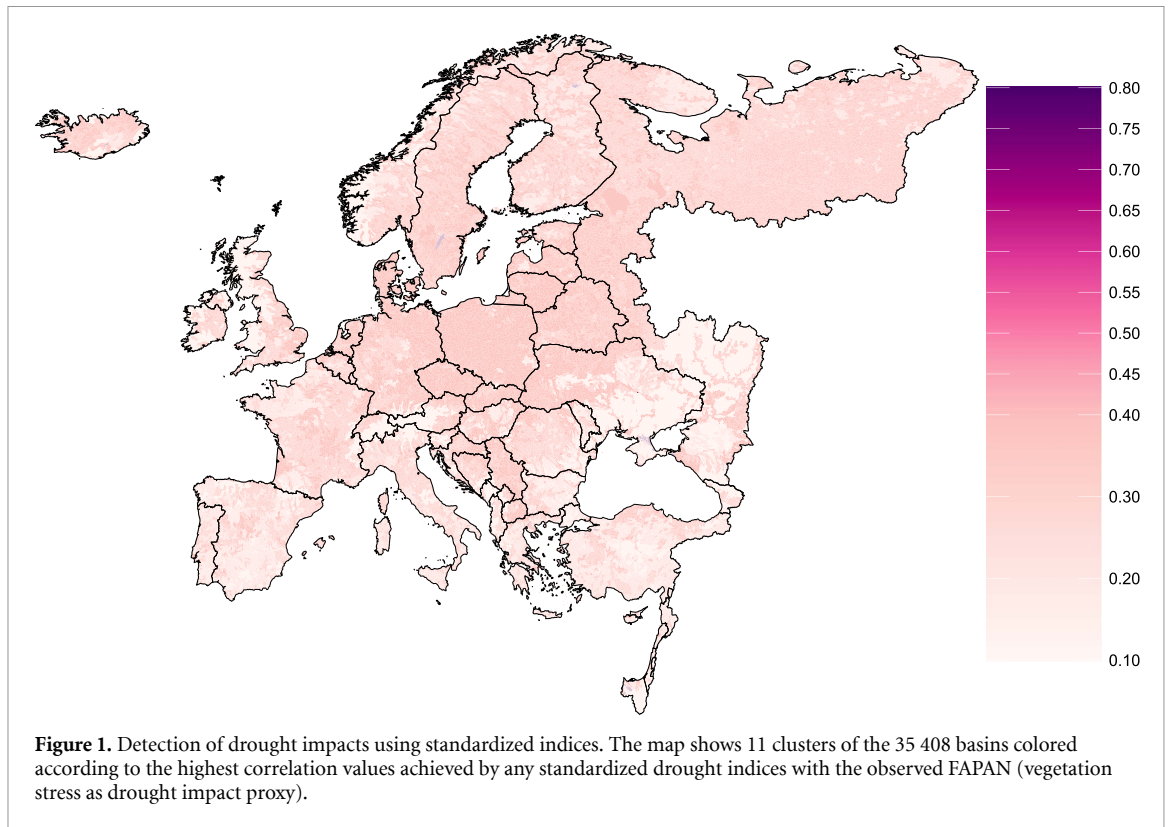
feed-forward neural network with the model's architecture tuned to each cluster of basins (table S2 in supplementary material). The resulting iCDIs attain an average correlation of 0.61 over the pan-European domain (figure 2), corresponding to an improvement of 0.35 over the use of standardized indices (see figure S19 in supplementary material for the Cluster 7 temporal trajectories).

3.3. Evolution of drought impacts under projected climate

Building on the high skill of iCDI in modeling the observed FAPAN dynamics, we extend its application to project future drought impacts under different climate change scenarios (section 2.2). Under the RCP4.5 scenario, Central Europe emerges as a clear hotspot of increasing vulnerability, with a marked escalation in drought impact frequency (figure 3).

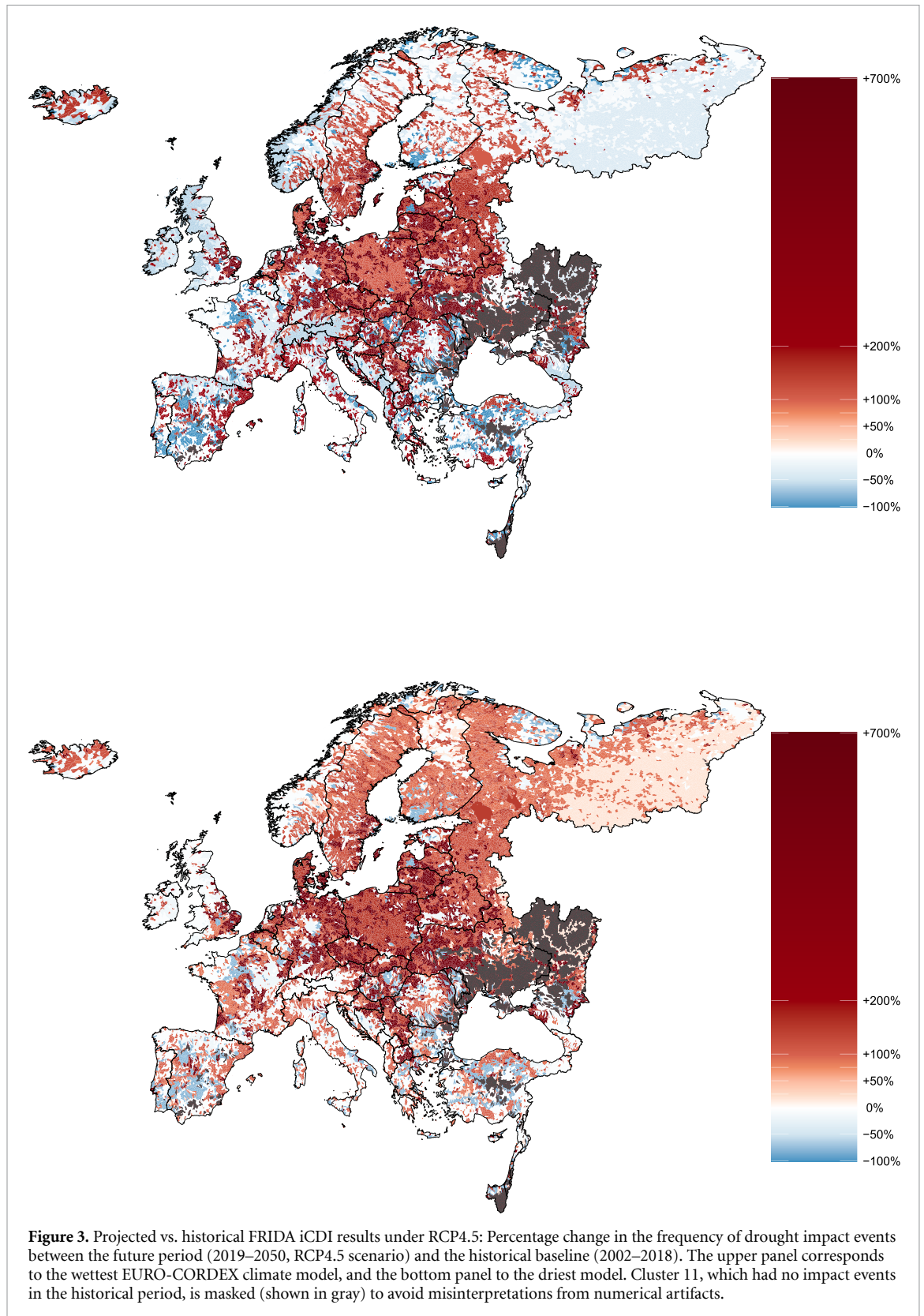
This signal contrasts with the more heterogeneous patterns observed in Northern and Southern Europe. In the north, a distinct east–west dipole is evident; western Norway is projected to experience a reduction in impact frequency, while eastern regions show an opposite trend. In Southern Europe, changes are spatially fragmented, with clusters of both increasing and decreasing impact frequency interspersed across the region. Notably, the Alpine arc stands out as a zone of decreasing drought impact frequency, diverging from the surrounding regions. The most pronounced rise in event frequency relative to historical levels is observed in Cluster 8, spanning temperate, continental Central Europe, with +515% and +542% under the wet and dry models, respectively. In addition, areas that historically exhibited no significant drought impacts, such as Eastern Ukraine and southeastern Russia (Cluster 11; hydrological regime is affected by human impacts, i.e. irrigation), show a clear upward trend in future projections. By contrast, only a few clusters exhibit a reduction in event frequency relative to historical levels, and these reductions are both modest and spatially limited. The most marked decrease occurs in Cluster 10, mainly covering Mediterranean countries such as Portugal, Spain, Greece, and Turkey, with projected declines of –100% and –76% under the wet and dry models, respectively. However, these reductions remain minor when compared to the widespread and substantial increases across most regions. Overall, our findings reveal a clear and spatially consistent intensification of drought impacts under RCP4.5, a pattern that holds across RCP2.6 and RCP8.5 scenarios as well (figures S20 and S21 in supplementary material), with RCP8.5 amplifying the magnitude of change and RCP2.6 showing more moderate but still concerning increases.

A different spatial pattern emerges when considering the SPEI-1 projections, a commonly used



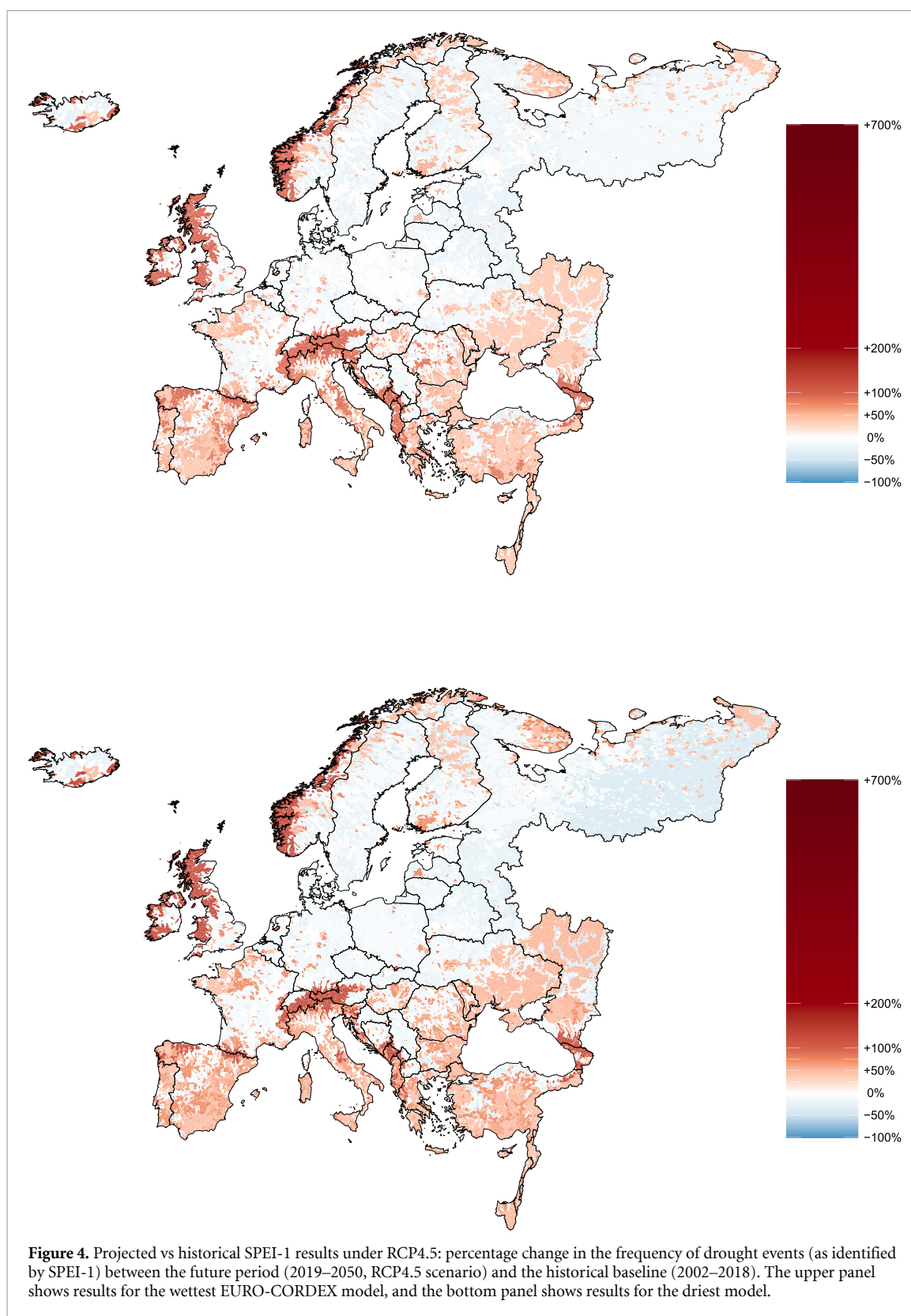
index for drought assessments. While our iCDIs identify Central Europe as the most critical hotspot of increasing drought impacts, SPEI-1 highlights the Mediterranean region as being at highest risk (figure 4). This apparent discrepancy is driven by the temperature-induced increase in PET, which

lowers SPEI-1 values in the Mediterranean without necessarily translating into vegetation stress. This gap between meteorological variables and observed impacts can be partly explained by the presence of water infrastructures. To verify this hypothesis, we quantified the influence of human water management



using the degree of regulation (DoR), defined as the ratio between total reservoir storage capacity and mean annual inflows in each sub-basin [44]. Considering only regulated basins (DoR > 0), the Mediterranean region (i.e. Clusters 7–5–10–11) shows substantially higher regulation intensity than Central Europe (i.e. Clusters 1–8–9), with

a mean DoR of 5.79 compared to 2.93. The spatial distribution of DoR among regulated basins (figure S22) confirms a concentration of strongly regulated catchments across the Mediterranean arc, consistent with independent irrigation data from the European Irrigation Map 2010 (EIM2010 [51]).



These findings suggest that the weaker correlation between SPEI-1 and observed vegetation anomalies in the Mediterranean does not imply lower drought risk, but rather reflects the buffering effect of irrigation and reservoir infrastructure, which can sustain vegetation activity under negative climatic water

balances. In contrast, Central European basins, characterized by more diffuse and less intensive regulation, remain more directly exposed to hydroclimatic stress, explaining why iCDI projections identify them as emerging hotspots of vegetation drought impacts.

4. Discussion and conclusions

This study proposes iCDIs that can characterize the observed vegetation stress across the pan-European domain and project future drought impacts under different climate change scenarios. Numerical results show that our approach successfully constructs a new drought index that aligns much more closely with the observed FAPAN trajectory. Moreover, the identification of the main predictors of drought conditions contributes to obtain a better understanding of the complex dynamics of drought events and resulting impacts. Notably, iCDI projections indicate that Central Europe will suffer the most severe drought impacts, in contrast to traditional indices which identify the Mediterranean as the key hotspot.

Despite promising results, several limitations remain. First, FAPAN was selected as the target variable as it provides an objective and spatially continuous representation of vegetation stress across Europe, addressing the well-known limitations in resolution and coverage of traditional socio-economic drought impact databases. However, this proxy also introduces specific uncertainties. While negative FAPAN anomalies are a robust indicator of reduced photosynthetic activity, they are not exclusively caused by drought. Other disturbances, such as pest infestations, plant diseases, or extreme heatwaves occurring independently of significant water deficits, can also trigger vegetation stress signals. These confounding factors may introduce noise into the training of the iCDI, potentially leading the model to associate certain vegetation responses with hydroclimatic drivers even when other ecological pressures are dominant. By acknowledging these non-drought signals, we emphasize that our impact-based indices might capture a broad spectrum of environmental stress. Future research should focus on integrating multi-source impact data to better disentangle drought-specific responses from other biotic and abiotic stressors, further enhancing the robustness of the iCDI framework. However, this remains a significant challenge, as existing databases (e.g. GDIS) often report drought as a single, large-scale aggregate event, lacking the spatial and temporal granularity required for the calibration of data-driven models.

Second, our analysis was conducted using hydrologically defined clusters [44]. These clusters categorize the river systems according to their dominant hydrological processes, meaning they capture the underlying hydroclimatic predictors of drought conditions. While this ensures physical homogeneity, it may not perfectly align with specific vegetation patterns, meaning a single cluster might host

diverse land-use types [51]. However, this potential mismatch is mitigated by the use of FAPAN anomalies as the target variable: by rescaling the vegetation response against its own local climatology, we account for the intrinsic differences in land-cover types, focusing on their relative stress rather than absolute productivity. Future research could explore multi-criteria clustering, integrating both hydrological signatures and vegetation phenology, to better capture the direct biological and socio-economic consequences of drought.

Lastly, while selecting the wettest and driest members of the EURO-CORDEX ensemble effectively brackets the potential range of future impacts, it does not fully capture the entire ensemble variability; future work should therefore incorporate the remaining GCM–RCM combinations to provide a comprehensive probability distribution of risks, potentially leveraging the latest CMIP6 projections. In addition, while we employed a joint normalization and recalibration strategy to mitigate the uncertainties related to future climate conditions falling outside historical ranges [53], as climate extremes intensify, all models may still encounter unprecedented conditions for which no analogues exist in the training data. Longer time series of impact observations would therefore be instrumental to improve model robustness and capture rare or emerging drought dynamics.

Despite these limitations, our findings demonstrate the value of integrating ML with remote sensing and hydroclimatic data to build more realistic and actionable drought indices as traditional indices often fail in capturing the actual impacts. Notably, while the proposed iCDIs focus on vegetation stress, our approach is generic and can capture drought impacts across different socio-economic sectors. Shifting to impact-based drought analysis can substantially enhance early warning systems and resilience planning, particularly as Europe faces a shifting and intensifying drought risk landscape.

Acknowledgment

This study was partially supported by the CLINT (Climate Intelligence: Extreme events detection, attribution and adaptation design using machine learning) project funded by the EU Horizon 2020 programme under Grant Agreement 101003876. MM was supported by the RETURN Extended Partnership funded by the European Union Next-GenerationEU (National Recovery and Resilience Plan—NRRP, Mission 4, Component 2, Investment 1.3—D.D. 1243 2/8/2022, PE0000005).

Data availability statement

The hydro-meteorological input data used for the computation of the indices are based on the HydroGFD2.0 reanalysis dataset, available on <https://hypeweb.smhi.se>.

For the projection of the indices under future climate scenarios, regional climate simulations from the EURO-CORDEX project (www.euro-cordex.net) were used and subsequently bias-adjusted as part of this study.

Supplementary data 1 available at <https://doi.org/10.1088/1748-9326/ae4ca7/data1>.

Conflict of interest

The authors declare that they have no competing financial interests.

Code Availability

The wrapper W-QEISS used for the feature selection analysis is available at: <https://github.com/EILab-Polimi/Matlab-Multi-objective-Feature-Selection>.

Author contributions

Martina Merlo  0009-0006-6968-2259
Conceptualization (equal), Investigation (lead),
Methodology (equal), Visualization (lead), Writing –
original draft (lead), Writing – review &
editing (equal)

Matteo Giuliani  0000-0002-4780-9347
Conceptualization (equal), Methodology (equal),
Supervision (equal), Writing – review &
editing (equal)

Yiheng Du  0000-0002-5176-8111
Data curation (equal), Formal analysis (equal),
Writing – review & editing (equal)

Ilias Pechlivanidis  0000-0002-3416-317X
Data curation (equal), Formal analysis (equal),
Writing – review & editing (equal)

Andrea Castelletti  0000-0002-7923-1498
Conceptualization (equal), Methodology (equal),
Supervision (equal), Writing – review &
editing (equal)

References

- [1] Van Loon A F and Van Lanen H A J 2012 A process-based typology of hydrological drought *Hydrol. Earth Syst. Sci.* **16** 1915–46
- [2] Spinoni J, Naumann G and Vogt J V 2017 Pan-European seasonal trends and recent changes of drought frequency and severity *Glob. Planet. Change* **148** 113–30
- [3] Blauhut V, Gudmundsson L and Stahl K 2015 Towards pan-European drought risk maps: quantifying the link between drought indices and reported drought impacts *Environ. Res. Lett.* **10** 014008
- [4] Heim J R R 2002 A review of twentieth-century drought indices used in the United States *Bull. Am. Meteorol. Soc.* **83** 1149–66
- [5] Mishra A K and Singh V P 2010 A review of drought concepts *J. Hydrol.* **391** 202–16
- [6] Pedro-Monzonis M, Solera A, Ferrer J, Estrela T and Paredes-Arquiola J 2015 A review of water scarcity and drought indexes in water resources planning and management *J. Hydrol.* **527** 482–93
- [7] Spinoni J, Naumann G, Vogt J and Barbosa P 2016 *Meteorological Droughts in Europe: Events and Impacts, past Trends and Future Projections* (Publications Office of the European Union)
- [8] Shyrokaya A, Pappenberger F, Messori G, Pechlivanidis I, Cloke H and Di Baldassarre G 2025 How good is my drought index? Evaluating predictability and ability to estimate impacts across Europe *Environ. Res. Lett.* **20** 034051
- [9] Cammalleri C, Naumann G, Mentaschi L, Formetta G, Forzieri G, Gosling S N, Bisselink B, De Roo A and Feyen L, and European Commission. Joint Research Centre 2020 *Global warming and drought impacts in the EU: JRC PESETA IV project : Task 7* (Publications Office) (<https://doi.org/10.1016/j.ejrh.2020.100723>)
- [10] Peña-Gallardo M, Vicente-Serrano S M, Domínguez-Castro F and Beguería S 2019 The impact of drought on the productivity of two rainfed crops in Spain *Nat. Hazards Earth Syst. Sci.* **19** 1215–34
- [11] Sjulgård H, Graf L V, Colombi T, Hirte J, Keller T and Aasen H 2024. Earth observation reveals reduced winter wheat growth and the importance of soil water storing capacity during drought
- [12] Qiao M, Hong C, Jiao Y, Hou S and Gao H 2024 Impacts of drought on photosynthesis in major food crops and the related mechanisms of plant responses to drought *Plants* **13** 1808
- [13] Vicente-Serrano S M, Peña-Angulo D, Beguería S, Domínguez-Castro F, Tomás-Burguera M, Noguera I, Gimeno-Sotelo L and El Kenawy A 2022 Global drought trends and future projections *Phil. Trans. R. Soc. A* **380** 20210285
- [14] Spinoni J et al 2020 Future global meteorological drought hot spots: a study based on CORDEX data *J. Clim.* **33** 3635–61
- [15] Satoh Y et al 2022 The timing of unprecedented hydrological drought under climate change *Nat. Commun.* **13** 3287
- [16] Zargar A, Sadiq R, Naser B and Khan F 2011 A review of drought indices *Environ. Rev.* **19** 333–49
- [17] Van Loon A F et al 2024 Review article: drought as a continuum—memory effects in interlinked hydrological, ecological, and social systems *Nat. Hazards Earth Syst. Sci.* **24** 3173–205
- [18] Palmer W C 1965 *Meteorological Drought* vol 30 (US Department of Commerce, Weather Bureau)
- [19] McKee T B, Doesken N J and Kleist J 1993 The relationship of drought frequency and duration to time scales *Proc. of the 8th Conf. on Applied Climatology* vol 17
- [20] Vicente-Serrano S M, Beguería S and López-Moreno J I 2010 A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration Index *J. Clim.* **23** 1696–718
- [21] Vicente-Serrano S M, López-Moreno J I, Beguería S, Lorenzo-Lacruz J, Azorin-Molina C and Morán-Tejeda E 2012 Accurate computation of a streamflow drought index *J. Hydrol. Eng.* **17** 318–32
- [22] European Commission, Joint Research Centre (JRC) 2019 EDO indicator factsheet: soil moisture anomaly (SMA)
- [23] Kchouk S, Melsen L A, Walker D W and Van Oel P R 2022 A geography of drought indices: mismatch between indicators of drought and its impacts on water and food securities *Nat. Hazards Earth Syst. Sci.* **22** 323–44

- [24] Shyrokaya A, Pappenberger F, Pechlivanidis I, Messori G, Khatami S, Mazzoleni M and Di Baldassarre G 2024 Advances and gaps in the science and practice of impact-based forecasting of droughts *WIREs Water* **11** e1698
- [25] Hao Z and AghaKouchak A 2013 Multivariate standardized drought index: a parametric multi-index model *Adv. Water Resour.* **57** 12–18
- [26] Karamouz M, Rasouli K and Nazif S 2009 Development of a hybrid index for drought prediction: case study *J. Hydrol. Eng.* **14** 617–27
- [27] Sepulcre-Canto G, Horion S, Singleton A, Carrao H and Vogt J 2012 Development of a combined drought indicator to detect agricultural drought in Europe *Nat. Hazards Earth Syst. Sci.* **12** 3519–31
- [28] Peng J et al 2019 Can we use satellite-based FAPAR to detect drought? *Sensors* **19** 3662
- [29] Centre for Research on the Epidemiology of Disasters (CRED) 2025 EM-DAT, The international disaster database
- [30] Rosvold E L and Buhaug H 2021 GDIS, a global dataset of geocoded disaster locations *Sci. Data* **8** 61
- [31] Drought Team: European Commission Joint Research Center 2025 European drought impact database (EDID)
- [32] Madruga de Brito M, Kuhlicke C and Marx A 2020 Near-real-time drought impact assessment: a text mining approach on the 2018/19 drought in Germany *Environ. Res. Lett.* **15** 1040a9
- [33] Camps-Valls G et al 2025 Artificial intelligence for modeling and understanding extreme weather and climate events *Nat. Commun.* **16** 1919
- [34] Salcedo-Sanz S et al 2024 Analysis, characterization, prediction, and attribution of extreme atmospheric events with machine learning and deep learning techniques: a review *Theor. Appl. Climatol.* **155** 1–44
- [35] Danandeh Mehr A, Rikhtehgar Ghiasi A, Yaseen Z M, Sorman A U and Abualigah L 2023 A novel intelligent deep learning predictive model for meteorological drought forecasting *J. Ambient Intell. Humaniz. Comput.* **14** 10441–55
- [36] Felsche E and Ludwig R 2021 Applying machine learning for drought prediction in a perfect model framework using data from a large ensemble of climate simulations *Nat. Hazards Earth Syst. Sci.* **21** 3679–91
- [37] Piri J, Abdollahipour M and Keshtegar B 2023 Advanced machine learning model for prediction of drought indices using hybrid SVR-RSM *Water Resour. Manag.* **37** 683–712
- [38] Zaniolo M, Giuliani M, Castelletti A F and Pulido-Velazquez M 2018 Automatic design of basin-specific drought indexes for highly regulated water systems *Hydrol. Earth Syst. Sci.* **22** 2409–24
- [39] Karakaya G, Galelli S, Ahipasaoglu S D and Taormina R 2016 Identifying (Quasi) equally informative subsets in feature selection problems for classification: a max-relevance min-redundancy approach *IEEE Trans. Cybern.* **46** 1424–37
- [40] Taormina R, Galelli S, Karakaya G and Ahipasaoglu S D 2016 An information theoretic approach to select alternate subsets of predictors for data-driven hydrological models *J. Hydrol.* **542** 18–34
- [41] AghaKouchak A et al 2022 Status and prospects for drought forecasting: opportunities in artificial intelligence and hybrid physical–statistical forecasting *Philos. Trans. R. Soc. A* **380** 20210288
- [42] Berg P, Donnelly C and Gustafsson D 2018 Near-real-time adjusted reanalysis forcing data for hydrology *Hydrol. Earth Syst. Sci.* **22** 989–1000
- [43] Hundecha Y, Arheimer B, Donnelly C and Pechlivanidis I 2016 A regional parameter estimation scheme for a pan-European multi-basin model *J. Hydrol.* **6** 90–111
- [44] Pechlivanidis I G, Crochemore L, Rosberg J and Bosshard T 2020 What are the key drivers controlling the quality of seasonal streamflow forecasts? *Water Resour. Res.* **56** e2019WR026987
- [45] Hadka D and Reed P 2013 Borg: an auto-adaptive many-objective evolutionary computing framework *Evol. Comput.* **21** 231–59
- [46] Maier H R and Dandy G C 2000 Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications *Environ. Model. Softw.* **15** 101–24
- [47] Nandgude N, Singh T P, Nandgude S and Tiwari M 2023 Drought prediction: a comprehensive review of different drought prediction models and adopted technologies *Sustainability* **15** 11684
- [48] Olden J D and Jackson D A 2002 Illuminating the “black box”: a randomization approach for understanding variable contributions in artificial neural networks *Ecol. Modell.* **154** 135–50
- [49] Huang G-B, Zhu Q-Y and Siew C-K 2006 Extreme learning machine: theory and applications *Neurocomputing* **70** 489–501
- [50] Wang G C, Zhang Q, Band S S, Dehghani M, Chau K W, Tho Q T, Zhu S, Samadianfard S and Mosavi A 2022 Monthly and seasonal hydrological drought forecasting using multiple extreme learning machine models *Eng. Appl. Comput. Fluid Mech.* **16** 1364–81
- [51] Sohil F, Sohali M U and Shabbir J 2022 An introduction to statistical learning with applications in R *Statistical Theory and Related Fields* vol 6, ed J Gareth, D Witten, T Hastie and R Tibshirani (Springer Science and Business Media) pp 87
- [52] Kuentz A, Arheimer B, Hundecha Y and Wagener T 2017 Understanding hydrologic variability across Europe through catchment classification *Hydrol. Earth Syst. Sci.* **21** 2863–79
- [53] Peres D J, Bonaccorso B, Palazzolo N, Cancelliere A, Mendicino G and Senatore A 2023 A dynamic approach for assessing climate change impacts on drought: an analysis in Southern Italy *Hydrol. Sci. J.* **68** 1213–28