Irving Fisher Committee on
Central Bank Statistics

IFC Satellite Seminar on "Sustainability data issues and central banks' experience"

4 October 2025

Beat the heat: the role of heat waves and droughts on regional EU economies

Sarah Spiteri,
Central Bank of Malta

Léonore Lebouteiller,
Central Bank of Luxembourg

Nicole Vorderobermeier,
Deutsche Bundesbank

Mar Delgado-Téllez,
Bank of Spain

Beat the Heat

The Role of Heat Waves and Droughts on Regional EU Economies

Sarah Spiteri, Léonore Lebouteiller, Nicole Vorderobermeier, Mar Delgado-Téllez

Abstract

Heat waves and droughts pose growing risks to regional economic stability in the EU, requiring climate-aware nowcasting tools for climate policy-relevant sectors. Applying machine learning (ML) models (Random Forest and XGBoost) to mixed-frequency climate and economic indicators from 1,117 NUTS-3 regions between 2002 and 2022, this study demonstrates that predictions of real gross value added (GVA) growth is significantly improved with climate-augmented ML models, particularly in agriculture. A complementary simulation exercise quantifies the predictive contributions of concurrent extreme weather events to regional economies, revealing that agriculture could occasionally face yearly GVA growth reductions of up to −10 percentage points in Northern and Central Europe. Conversely, the broader industrial sector shows moderate simulated losses and manufacturing remains largely resilient. These findings underscore asymmetric vulnerabilities between sectors and regions and the value of predictive modelling for climate adaptation policy.

Keywords: regional productivity, weather extremes, heat waves, droughts, machine learning

JEL classification: C6, C8, E37, Q5

Contents

1. Introduction

Climate change presents asymmetric risks to economic and political stability, exacerbating regional disparities. In the European Union (EU), hot extremes are intensifying, particularly in Mediterranean countries (Seneviratne et al., 2021), underscoring the need for region-specific adaptation policies. Extreme heat already reduced European[1] Gross Domestic Product (GDP) by 0.3-0.5% and could reduce it by 1.5–2.5% till 2060, with respect to historical average GDP levels between 1981 to 2010 (García-León et al., 2021). Under a 3°C warming scenario, drought-related damages are projected to increase fivefold by 2100, relative to the baseline average of drought-related damages during 1981 to 2010, Mediterranean countries remaining the most affected (European Commission et al., 2020). Heat exposure could also reduce working hours by 15–60% in the RCP8.5 scenario[2] (Casanueva et al., 2020).

The literature studies how climate change affects the economy through multiple channels, including productivity losses, capital damage, and increased energy demand. Early research focused on long-term impacts, while recent studies indicate growing short- and medium-term risks from extreme weather. Macroeconomic models show persistent GDP losses, especially in vulnerable regions (Motta et al., 2025; Kahn et al., 2021; Dellink et al., 2019). Much of this work focused on agriculture, where rising temperatures and shifting rainfall have already slowed productivity (Wang et al., 2018; Praveen and Sharma, 2019). However, many studies rely on low-resolution data and linear methods, with machine learning (ML) tools still mostly limited to carbon tracking and pricing contexts (Li et al., 2022).

Few studies use ML to integrate climate and economic data for macroeconomic forecasting. One exception is Rossi et al. (2023), who applied ML to assess drought risks across interconnected EU systems[3], estimating sector- and region-specific losses under current and projected climate conditions. Their approach used Decision Tree Classifiers trained on detrended impact data, aggregated into vulnerability groups, and applied to historical (1979-2021) and future risk scenarios. Building on this, we employ Random Forest and XGBoost, ensemble tree-based methods. Random Forest mitigates overfitting via bootstrapped sampling and random feature selection, while XGBoost sequentially corrects residuals and applies regularisation.

Usman et al. (2024) analysed the impact of heat waves, droughts, and floods on regional production in the EU, following the nomenclature of territorial units for statistics (NUTS), over five years using a local projections difference-in-differences framework. They found that summer heat waves and droughts reduced annual production by 1.4 and 2.4 percentage points, mainly through declines in employment, hours worked, and productivity. Droughts caused persistent losses in agriculture, lower investment, and reduced productivity. Conversely, heat waves boosted capital investment but lowered productivity. The impact of floods varied: high-income regions saw reconstruction-driven growth, while low-income areas struggled to recover. Usman et al. (2025) extended this work, estimating that extreme weather in the summer of 2025 reduced gross value added (GVA) by €43 billion (0.26% of EU output), with cumulative losses projected to reach €126 billion by 2029, suggesting both immediate and lasting regional effects of extreme weather.

Based on Usman et al. (2024, 2025), we adopt a predictive framework using similar climate indicators to model heat waves and droughts. We compare linear regression with two ML models, Random Forest and XGBoost, to predict per capita real GVA growth in all NUTS-3 EU regions, focussing on climate-

1 Including EU-27 countries, the United Kingdom and EFTA countries.

2 RCP8.5 is one of the Representative Concentration Pathways (RCPs) used in climate science to model possible future greenhouse gas concentration trajectories and their impacts. It is often recognised as a worst-case scenario.

3 Including agriculture, water supply, energy, river transport, and terrestrial and freshwater ecosystems.

relevant sectoral NACE aggregations A, B-E, and C.[4] These sectors, identified as possible targets of climate policies (Battiston et al., 2022), provide a rich dataset encompassing cross-country, within-country, and cross-sectoral variation. This heterogeneity enables a rigorous evaluation of how economic activity responds to climate conditions in linear and non-linear predictive frameworks. We also assess the compounded predicted consequences of overlapping climate events and regional spillovers. Model performance is evaluated through a nowcasting exercise that leverages mixed-frequency data and integrates feature importance analysis with climate event simulations.

Our results show that ML models can significantly improve the accuracy of regional economic nowcasts under climate stress, especially in agriculture. We show that ML models significantly outperform traditional linear regressions in nowcasting real growth in per capita GVA, particularly for the agriculture (A) and industrial (B–E) sectors when climate indicators for heat waves and droughts are included. Sectoral analysis highlights agriculture's pronounced sensitivity to climate extremes. In addition, compound and spatially correlated climate shocks amplify predicted economic losses, especially in low-GVA regions, although their associations remain heterogeneous across Europe. These findings support the use of predictive modelling for timely policy responses and stress the importance of sectoral and regional differentiation in climate adaptation strategies.

To complement the nowcasting analysis, we use a predictive model to simulate the potential contributions of compound scenarios that combine extreme heat waves and droughts to the predicted output, using a nowcasting model. Comparing a no-anomaly climate baseline with a compounded scenario, we find that agriculture suffers the largest simulated losses, with occasional reductions in yearly growth of up to ten percentage points in Northern and Central Europe, while Southern Europe exhibits less pronounced negative consequences, which contrasts with the findings of (European Commission et al., 2020). The industrial sector shows moderate reductions, while manufacturing remains largely resilient, highlighting sectoral differences in climate vulnerability and adaptive capacity.

The paper is structured as follows: Section 2 introduces the data; Section 3 develops the construction of compound climate indicators; Section 4 details the nowcasting methodologies and results; Section 5 presents the design of a simulation and its outcomes; and Section 6 concludes.

2. Data Description

This study builds predictive models for sectoral GVA using economic and climate data from 1,117 NUTS-3 regions between 2002 and 2022, including regional GVA by sector and employment-to-population ratios, and country-level GDP from Eurostat.

A key contribution of this paper is the integration of climate indicators from the Copernicus European Drought Observatory. Table 1 summarises the selected variables: Heat and Cold Waves Index (HCWI) for heat waves, Standardised Precipitation Index (SPI) for meteorological droughts (short- (SPI01), medium- (SPI06), and long-term (SPI12)), Low-Flow Index (LFI) for hydrological droughts, and Soil Moisture Anomaly (SMA) and Fraction of absorbed photosynthetically active radiation (fAPAR) for agricultural droughts.[5]

[4] Sector A covers agriculture; B–E includes mining and quarrying (B), manufacturing (C), electricity, gas, steam and air conditioning supply (D), and water supply, sewerage, waste management and remediation activities (E). Sector C is also modelled separately.

[5] The indices' methodologies are available at https://drought.emergency.copernicus.eu/data/factsheets.

Main Climate Variables

Table 1

Variable	Definition	Methodological Reference
HCWI	Heat and Cold Waves Index: Monitors daily minimum and maximum temperatures, identifying events lasting at least three consecutive days based on percentile thresholds from a 30-year climatology.	Lavaysse et al. (2018)
SPI	Standardised Precipitation Index: Detects meteorological droughts and floods by measuring precipitation anomalies over accumulation periods of 1, 6, and 12 months, comparing observed rainfall to long-term averages. Thresholds below −1 indicate drought conditions, while values above +1 signal excess rainfall or potential flooding.	McKee et al. (1993)
LFI	Low Flow Index: Detects periods of unusually low river discharge compared to long- term thresholds, updated every 10 days. Severity is classified from low to very high.	Cammalleri et al. (2017)
SMA	Soil Moisture Anomaly: Measures deviations in soil moisture from long-term averages, using daily data and 10-day updates. Values below −1 indicate drought stress, while values above +1 suggest unusually wet conditions.	de Roo et al. (2000)
fAPAR	Fraction of Absorbed Photosynthetically Active Radiation: Measures deviations in photosynthetic activity from long-term averages, updated every 10 days. Negative anomalies, below −1, indicate drought-related stress, while positive anomalies, above +1 reflect enhanced vegetation growth under wetter or favourable conditions.	Bojinski et al. (2014)

Source: Copernicus European Drought Observatory

The climate variables used in the empirical analysis capture periods of abnormal conditions relative to long-term averages. Heat wave and drought severity are measured using standardised indices (Table 1), which define threshold values for anomalous events. Accordingly, the variables were coded to indicate only the occurrence of heat waves or droughts, while other conditions were set to zero.[6]

Our data show that most European regions have experienced an increase in the average annual heat wave days (Figure A.1). Summer heat waves were already pronounced in the Mediterranean and Nordic regions at the beginning of the century (Russo et al., 2015). Since then, the largest increases occurred in the Spanish Mediterranean, southern Italy, and northern and eastern Europe. Furthermore, comparing the average HCWI scores between the first and last five years (Table A.1) of the sample confirms that all countries now face more frequent and/or intense heat waves.

The average number of dry months per year, defined as SPI01 values below −1, suggests a trend toward increased dryness in Europe over the past two decades (Figure A.2). Southern Spain and parts of eastern Europe were already prone to frequent dry months in 2002–2006, but the rise between 2018–2022 and the earlier period highlights growing dryness in central Europe. Whereas regions in Finland saw early anomalies that later subsided, areas such as Germany experienced a persistent increase. Country-level SPI data across accumulation periods (Table A.1) reinforce this spatial pattern, indicating a shift toward drier conditions consistent with intensifying meteorological drought in central Europe.

An assessment of hydrological and agricultural drought, indicated by mean LFI values greater than 0 and mean SMA figures below −1, reveals patterns similar to meteorological drought (Figures A.3 and A.4). Between 2002 and 2006, Spain, eastern Europe, and Scandinavia on average recorded up to eight hydrological and five agricultural dry months annually. From 2018 to 2022, these regions, along with parts of France and central Europe, experienced further increases, indicating increased intensity and spread of hydrological and agricultural drought conditions.

6 Heat waves and droughts are characterised by HCWI and LFI > 0 and SPI, SMA, and fAPAR ≤ −1.

3. Calculating Compounding Indicators

We account for the compounding nature of climate events shaped by temporal and geographic factors. Temporal compounding is addressed by incorporating event duration into climate variables, so that longer events yield higher severity indices, while geographic compounding is addressed by incorporating a variable accounting for the climate conditions across neighbouring regions. Lastly, we introduce a feature to account for the coincidence of heat waves and droughts.

Localised climate shocks can trigger spillover price effects, such as the 50% increase in EU olive oil prices after droughts in Spain and Italy (Kotz et al., 2025), while heat-induced productivity losses cascade through supply chains (Sun et al., 2024). To analyse how spatially compounding heat waves and droughts are associated with GVA, we compute regional averages of climate features across neighbouring areas. For each region r, geographical compounding of a climate indicator $GC_{t,r}^{X}$ is defined as the mean value of that indicator in neighbouring regions N_{r^*}:

$$GC_{t,r}^{X} = \frac{1}{|N_{r^*}|} \sum_{i \in N_{r^*}} X_{t,i}\, , \text{ where } X_{t,r} \in \{HCWI_{t,r}, SPI01_{t,r}, SPI06_{t,r}, SPI12_{t,r}, LFI_{t,r}\} \tag{1}$$

where variables are indexed by time (t) and region (r), $X_{t,i}$ represents the climate indicator in region i from the subset of neighbouring regions (r^*) of r. We exclude agricultural drought indicators (SMA and fAPAR) from this aggregation since, within the short- to medium-term nowcasting window, they exhibit minimal cross-regional compounding effects (Sepulcre-Canto et al., 2012). Geographic compounding captures shared climate systems and cross-border influences, refining regional exposure measures. Correlation analysis shows weak intercorrelations among individual indicators but strong associations (0.6–0.9) between geographic compounding variables and their respective features (X). Such multicollinearity challenges linear models but is mitigated in ML through regularisation techniques that shrink or penalise correlated coefficients.

Usman et al. (2024) show that concurrent heat waves and droughts significantly reduce medium-term regional output, especially in less wealthy areas. Studying these compound events is essential to understand their predictive contributions. In our models, compounding is captured by variables interacting HCWI and SPI. To ensure comparability, each indicator (V) is scaled using Min-Max normalisation with a zero lower bound as follows:

$$\tilde{V}_{t,r} = \frac{V_{t,r} - \min(V)}{\max(V) - \min(V)}, \tag{2}$$

We model heat wave–drought compounding by interacting the scaled HCWI and SPI across accumulation periods $p \in \{1,6,12\}$. The compounding indicator $EC_{t,r}$ for each p is defined as:

$$EC_{t,r}^{p} = \begin{cases} \widetilde{HCWI}_{t,r} \times \left|\widetilde{SPI}_{t,r}^{p}\right|, & \text{if } HCWI_{t,r} > 0 \text{ and } SPI_{t,r}^{p} \leq -1 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

with the indicators for each p aggregated as:

$$EC_{t,r} = f_{p \in P}\left(EC_{t,r}^{p}\right) \tag{4}$$

This formulation captures concurrent heat wave–drought events by multiplying heat stress (HCWI) and drought severity (SPI) across accumulation periods p. The indicator activates only when a heat wave occurs ($HCWI > 0$) along with a meteorological drought ($SPI \leq -1$). Multiple SPI periods reflect differing drought-related ecological outcomes. However, due to sparsity, indicators from all periods are combined into a single measure as shown in Equation 4[7]. Coinciding events were most frequent in Mediterranean

[7] The functional form of $f_{p \in P}(\cdot)$ in Equation 4 depends on the mixed-frequency approach in Section 4.1, whereby a singular factor is extracted from the monthly aggregates of the different $EC_{t,r}^{p}$ for every p.

countries, primarily Italy, Spain, Cyprus and Malta, and parts of Central and Eastern Europe, notably Germany and Hungary, predominantly in later years.

4. Nowcasting Exercise

Predictive Approaches

For sector s, with variables indexed by time t, region r and country c, Equation 5 shows our baseline economic model and Equation 6 presents the climate-augmented version. Model 5 predicts real GVA per capita growth ($Y_{t,r,s}$) using its lag ($Y_{t-1,r,s}$), sectoral employment share ($E_{t,r,s}$), lagged country GDP growth ($G_{t-1,c}$), geographic controls (M_r, R_r, C_r), and a NUTS-3 indicator (N_r). Model 6 extends f_{ECON} with climate variables capturing heat waves, droughts, and their geographical[8] and event compounding. Parameters and hyperparameters are denoted by $\theta_{ECON,s}$ and $\theta_{CLIM,s}$.

$$Y_{t,r,s} = f_{ECON}\left(Y_{t-1,r,s}, E_{t,r,s}, G_{t-1,c}, M_r, R_r, C_r, N_r;\ \theta_{ECON,s}\right) \tag{5}$$

$$Y_{t,r,s} = f_{CLIM}\begin{pmatrix} Y_{t-1,r,s}, E_{t,r,s}, G_{t-1,c}, HCWI_{t,r}, SPI01_{t,r}, SPI06_{t,r}, SPI12_{t,r}, \\ LFI_{t,r}, SMA_{t,r}, fAPAR_{t,r}, GC_{t,r}, EC_{t,r}, M_r, R_r, C_r, N_r;\ \theta_{CLIM,s} \end{pmatrix} \tag{6}$$

Geographic controls identify mountainous (M_r), rural (R_r), and coastal (C_r) regions. Overall, 44.0% of regions are mixed urban–rural, 35.3% predominantly rural, 20.7% urban; 26.7% include mountain areas, and 29.1% are coastal. These fixed factors influence production, connectivity, and regional heterogeneity. To capture additional heterogeneity, we encode NUTS-3 labels (N_r)[9]. This approach avoids transforming the target and offers a low-dimensional alternative to fixed effects, following Yang et al. (2024) on incorporating cross-sectional attributes in ML panel models.

Our approach addresses the frequency mismatch between yearly economic and daily climate data by applying three mixed-frequency methods, grouped into aggregation and feature extraction techniques (Li et al., 2022) as follows:

1. **Yearly Aggregation:** Climate variables are aggregated to yearly medians and standard deviations, capturing levels and volatility. Deviations from historical averages and within-month variability are measured using the median of monthly standard deviations, following Ciccarelli et al. (2023).[10]

2. **Feature Extraction:** Annual climate indicators are derived from monthly medians data using principal component analysis (PCA) and mixed data sampling (MIDAS), which condense high-frequency variables for each indicator into a single interpretable feature.[11]

We employ two complementary feature-extraction methods: an unsupervised approach (PCA) that summarises climate variation independently of the target variable, and a supervised approach (MIDAS) that aggregates high-frequency indicators using beta-distributed lag weights. The beta weighting

8 Geographical compounders, $GC_{t,r}^{X}$ for $X_{t,r} \in \{HCWI_{t,r}, SPI01_{t,r}, SPI06_{t,r}, SPI12_{t,r}, LFI_{t,r}\}$, are represented by $GC_{t,r}$.

9 CatBoost encoding replaces categories with smoothed conditional means of the target, while applying Bayesian smoothing to reduce overfitting and carefully incorporating the global mean to prevent data leakage.

10 Geographic compounding variables remain separate, while event-compounding indicators across drought durations are averaged via $f_{p\in P}(\cdot)$ in Equation 4 to reduce sparsity.

11 Geographic compounding variables ($GC_{t,r}^{X}$) are included with their base indicator X. This setup reduces multicollinearity by extracting a singular feature for each climate indicator. Event compounding indicators ($EC_{t,r}^{p}$) are combined using PCA or MIDAS, so $f_{p\in P}(\cdot)$ in Equation 4 applies the chosen extraction method, reducing sparsity.

scheme parsimoniously captures the delayed, non-linear influence of past climate conditions, i.e. the gradual build-up, peak, and decline of droughts and heat waves (Duo, 2025).

To estimate and compare nowcasting performance for $Y_{t,r,s}$, we employ a linear regression benchmark (Martín Cervantes et al., 2020; Kahn et al., 2021; Mohaddes et al., 2022) and two tree-based ML models, Random Forest and XGBoost. The models are trained on data up to 2017 and evaluated out-of-sample. These ensemble methods are well suited to capture non-linearities and interactions between climate and economic variables (Rossi et al., 2023).

Random Forest builds an ensemble of decision trees on bootstrapped samples, introducing randomness by selecting predictor subsets at each split. For each input, the prediction is the average across trees, reducing variance and making Random Forests robust to overfitting in the presence of collinear climate indicators, while providing interpretable variable importance (Breiman, 2001; Pedregosa et al., 2011). XGBoost, in contrast, sequentially adds trees to correct residuals, based on the learning rate. It improves gradient boosting with second-order gradients and regularisation, making it effective for modelling non-linear climate–economic dynamics (Chen and Guestrin, 2016).

Model performance, based on Equations 5 and 6, is evaluated through a nowcasting exercise using an expanding rolling window and out-of-sample root mean squared errors (RMSEs). To interpret predictions, we apply SHAP (SHapley Additive exPlanations) values to identify key predictive drivers, particularly climate variables influencing $Y_{t,r,s}$. SHAP values fairly attribute feature contributions across all combinations, offering transparent insights (Shapley, 1953; Lundberg et al., 2018).

Nowcasting Results

Table 2 reports RMSEs from the nowcasting exercise, revealing clear contrasts between linear and ML models across sectors. Predictive ability testing indicates that climate-augmented linear models were unable to significantly outperform f_{ECON} for all sectors, whereas ML models show sector-specific gains. For agriculture, climate-augmented ML models significantly improved accuracy over f_{ECON}, driven by heat wave and drought indicators. In contrast, f_{CLIM} did not surpass f_{ECON} for sectors B–E and C within the same model class and sector. However, ML predictions for B–E significantly outperformed all linear counterparts, while most ML predictions for sector C did not.

Nowcast RMSEs

Table 2

Model	Linear Regression			Random Forest			XGBoost		
	A	B-E	C	A	B-E	C	A	B-E	C
Economic (f_{ECON})	21.64	14.11	77.9	23.05	13.83†	72.81	28.15	13.89†	76.73†
Climate (f_{CLIM})									
Medians & St. Deviations	22.12	14.20	77.95	20.93*†§	14.01†	73.64	20.54*†§	13.95†	77.69
PCA Features	22.23	14.17	77.91	21.05*†	13.94†	74.03	21.52*†	14.03†	77.80
MIDAS Features	22.98	14.36	78.03	21.18*†	14.06†	75.19	21.72*†	13.97†	85.34

Note: RMSEs are calculated over annual nowcasts of $Y_{t,r,s}$ from 2018 to 2022. Sector A covers agriculture; B–E includes mining (B), manufacturing (C), utilities (D), and waste management (E), with C also modelled separately to obtain direct predictions for manufacturing. The symbols to denote statistically significant improvements in predictive performance are: $* \, p < 0.05$ indicates predictive improvements relative to the respective economic counterpart of the same model type; † $p < 0.05$ indicates predictive improvements relative to the respective linear regression counterpart; and § $p < 0.05$ indicates predictive improvements relative to the respective linear economic model for the same sector.

Among the climate-augmented models, the simpler aggregation of climate variables into yearly medians and standard deviations performs at least as well as, and often better than, the more complex PCA and MIDAS feature-extraction techniques. As shown in Table 2, this specification generally yields the most accurate climate-augmented nowcasts of $Y_{t,r,s}$, particularly in agriculture, where it significantly outperforms both the corresponding economic model and the linear regression benchmark.

SHAP values combined with factor loadings reveal key patterns in model predictions (Figures A.5 and A.6). Economic variables remain most predictive, but climate and geographic factors also matter, varying by sector and specification. HCWI is consistently influential in climate-augmented models, while drought indicators show sector-specific relevance. In agriculture, heat wave indicators[12] dominate, alongside short- to medium-term meteorological drought (SPI01, SPI06). In sectors B–E and C, heat waves retain influence, but drought importance shifts to hydrological measures and long-term droughts (LFI, SPI12). These differences highlight sectoral dynamics and ML's ability to adjust weights via regularisation, enhancing climate-augmented models over linear ones. The prominence of heat waves aligns with Usman et al. (2024), showing strong short-term contributions, while droughts indicate medium- to long-term contributions.

Strong regional heterogeneity in how heat waves, droughts, and their interactions affect outcomes is revealed by SHAP values. Heat waves influence both high- and low-GVA regions[13], sometimes benefiting higher-GVA areas, while droughts disproportionately harm lower-GVA regions. Climate-related predictive contributions are more pronounced in below-median GVA regions, underscoring their vulnerability. Consistent with this, Usman et al. (2025) find that droughts cause persistent losses in low-GVA regions, while some high-GVA regions may gain short-term benefits from heat waves, highlighting asymmetric climate vulnerability.

Geographic compounding indicators have a pronounced influence on predicted outcomes, reflecting both competitive and complementary dynamics[14]. Regions often exhibit beneficial associations when neighbours face heat waves, consistent with competitive dynamics, while the contributions of hydrological drought are mixed, suggesting complementarity shaped by geography and trade routes and competitive behaviours being evident when drought conditions force deviations from established routes. Dellink et al. (2019) demonstrate that climate-induced economic effects vary significantly between regions, with some benefiting from neighbouring disruptions due to trade and sectoral shifts, while others suffer compounded losses, underscoring these dynamics and the need to study them within integrated ecological–economic systems.

The interplay between heat waves and droughts can amplify, offset, or mirror their individual contributions, highlighting the complexity of compound climate shocks[14]. Event-based indicators are especially relevant in models using feature extraction, where factor loadings vary non-linearly across drought horizons. Under the Medians & St. Deviations definition of f_{CLIM}, which weights drought horizons equally, SHAP values suggest that the joint contribution is generally smaller than the sum of individual values. In contrast, feature extraction yields more diverse outcomes: agricultural predictions are driven by short-term droughts, while sectors B–E and C respond more to long-term droughts. This reflects sectoral sensitivities, where agriculture reacts to immediate vegetation stress, whereas industry is more affected by hydrological drought. These findings align with Zhu and Li (2024), who show climate risks have non-linear,

12 Includes yearly median, standard deviation, and geographic compounding for the Medians & St. Deviations version of f_{CLIM} , and extracted features for PCA- and MIDAS-based implementations.

13 High- and low-GVA regions are defined by a time-invariant distribution of real GVA per capita growth, with low (high) regions below (above) the midpoint.

14 Supporting data is reported in Figures A.5 and A.6, and Tables A.2, A.3 and A.4.

sector-specific economic effects, and Yin and Slater (2023), who find compound events often diverge from additive impacts, underscoring the need for models that capture sectoral and temporal nuances.

5. Application: A Simulation of Extreme Compounded Events

Simulation Design

To further examine the influence of heat waves and long-term meteorological droughts on GVA growth, we design a comparative analysis of two distinct scenarios: the *non-event* and *event* scenarios. These simulations are estimated using XGBoost with the yearly aggregation approach.[15]

The *non-event scenario* is a counterfactual where sectoral real GVA growth evolves without contributions from heat waves or droughts. All climate variables are set to baseline values, where the heat wave and drought indicators, including their standard deviations, are set to zero, while economic variables remain as observed. In contrast, the *event scenario* keeps economic variables as observed but simulates climate conditions for heat waves and droughts. Specifically, we set HCWI and LFI to their 99th percentile values, whilst all other climate indices were set to their 1st percentile values, of each variable's pooled series across regions and years. This simulation design ensures the reflection of the compounding effect of heat waves and hydrological drought, whilst mitigating the influence of other forms of droughts. Additionally, given that our climate variables are expressed as indices rather than absolute measures of temperature or precipitation, applying the same percentile threshold of heat waves and droughts represents comparable extremity levels across all NUTS-3 regions.

The predicted GVA growth rates under the event and non-event scenarios are denoted by $\hat{y}^*_{t,r,s}$ and $\hat{y}_{t,r,s}$, respectively. The simulation of the compounded heat wave and long-term drought event on $Y_{t,r,s}$, $\Delta\hat{y}_{t,r,s}$, is calculated as the difference between these predictions, such that:

$$\Delta\hat{y}_{t,r,s} = \hat{y}^*_{t,r,s} - \hat{y}_{t,r,s} \, .$$

Simulated Outcomes

The simulated values of the compounded heat wave and hydrological drought event on $Y_{t,r,s}$, described in Section 5, is averaged over the timespan of 2002 to 2022 at the NUTS-3 level for each sector.

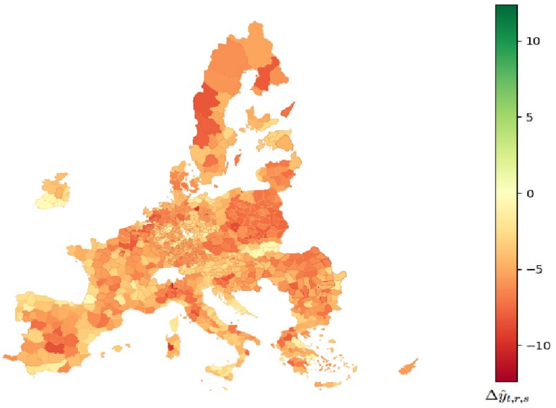
The outcomes of the simulation, depicted in Figure 1, show that EU agriculture would experience widespread negative contributions from compounded heat wave and drought events, with simulated values ranging from approximately -10 to 0 percentage points,[16] indicating substantial regional heterogeneity in the magnitude of $\Delta\hat{y}_{t,r,s}$. Northern Europe, especially Scandinavia and the Baltics, faces the most severe simulated losses (around -8 percentage points), while Central regions, such as Germany, Poland and France, experience moderate to severe simulated consequences (-4 to -8 percentage points). Southern Europe would experience less severe negative outcomes. The spatial distribution suggests that agricultural systems in cooler, traditionally less drought-adapted regions may be disproportionately sensitive to extreme compound climate events, possibly due to lower baseline resilience to simultaneous heat and water stress conditions, as indicated by Lesk et al. (2022) and Tabari and Willems (2023).

15 This model and approach were chosen since it generally resulted in the lowest RMSEs for the climate-augmented models as described in Section 4.2.

16 The 5th and 95th percentiles of $\Delta\hat{y}_{t,r,s}$ for agriculture equals -9.78 and 0, respectively.

Simulation of Heat Waves & long-term Drought on Sector A

Figure 1

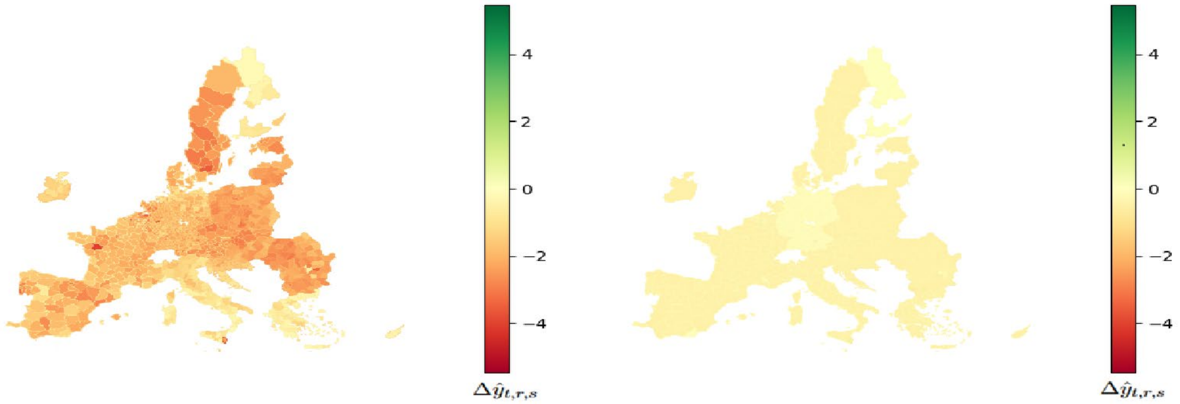


Note: Figure 1 illustrates $\Delta\hat{y}_{t,r,s}$ for agricultural (A) GVA growth from a compounded heat wave and long-term drought across the EU. The figure displays the differences between the non-event and event scenarios defined in section 5, using a green-to-red gradient (positive to negative differences of ten percentage points).

Source: Authors' calculations

Simulation of Heat Waves and long-term Drought on Sectoral Aggregates B-E (left) & C (right)

Figure 2



Note: Figure 2 illustrates $\Delta\hat{y}_{t,r,s}$ on industrial (B-E, left) and manufacturing (C, right) GVA growth from a compounded heat wave and long-term drought across the EU. The figure displays the differences between the non-event and event scenarios defined in section 5, using a green-to-red gradient (positive to negative differences of four percentage points).

Source: Authors' calculations

The simulation results for the broader industrial and energy sector, shown on the left of Figure 2, reveal moderate but spatially heterogeneous negative outcomes across the EU, with most NUTS-3 regions experiencing GVA growth reductions of 0-2 percentage points under our compound heat wave and drought scenario.[17] Northern and Central Europe, especially Germany, northern France, Poland, and Scandinavia, are most affected, reflecting the vulnerability of energy-intensive industries and thermal power generation to heat stress. Southern regions show milder simulated values, likely due to existing adaptation measures and infrastructure suited to higher baseline temperatures.

17 The 5th and 95th percentiles of $\Delta\hat{y}_{t,r,s}$ for sectors B-E equal -2.3 and 0, respectively.

The right panel of Figure 2 indicates that the manufacturing sector remains largely unaffected by compound heat wave and drought events, with real GVA per capita simulated losses up to 0.5 percentage point across the EU.[18] This resilience from extreme weather events likely stems from controlled indoor environments, lower water dependency, and advanced climate control systems. The uniform pale-yellow shading reflects stable productivity, supported by automation and reduced exposure to ambient conditions, unlike agriculture or energy sectors.

6. Conclusion

This paper examines how extreme heat waves and droughts affect regional economic performance of three NACE sectoral aggregations across EU NUTS-3 regions. Using mixed-frequency data and ML models (Random Forest and XGBoost), we find statistically significant improvements in nowcasting accuracy over linear regression, especially in agriculture. Climate variables, such as heat wave intensity and short- to medium-term drought indicators, play a key role in predicting real GVA per capita growth in this sector, compared to other sectors.

Our nowcasting results suggest that climate-augmented ML models outperform both economic and linear regression benchmarks in out-of-sample RMSEs. In agriculture, using climate indicators within ML models, especially yearly medians and standard deviations, significantly improves accuracy. For industrial sectors (B–E and C), climate data adds less value, though ML models still outperform regressions for sectoral aggregation B-E. Complex feature extraction methods like PCA and MIDAS offer no clear advantage over simpler aggregations, suggesting that interpretability and parsimony are preferable.

The findings reveal substantial heterogeneity in climate vulnerability between sectors and regions. Agriculture is highly sensitive to climate stress, while industrial sectors show more complex responses, particularly to hydrological droughts. Geographic and event-based compounding tend to further amplify negative contributions to predicted economic outcomes, especially in low-GVA regions, underscoring the asymmetric nature of climate contributions across Europe. These results support the use of predictive modelling to guide timely and targeted policy responses and highlight the importance of sectoral and regional differentiation in climate adaptation strategies.

Simulation results confirm sectoral differences in climate vulnerability. Under compound heat wave and hydrological drought scenarios, agriculture faces widespread simulated real GVA growth losses, with Northern and Central Europe seeing reductions up to -10 percentage points. The broader industrial sector (B-E) shows moderately negative $\Delta\hat{y}_{t,r,s}$ values (up to 2 percentage points), especially in energy-intensive regions. In contrast, manufacturing remains resilient, with minimal reductions (up to 0.5 percentage points), reflecting its controlled environments and lower exposure to weather stress.

Nonetheless, several limitations constrain the interpretability and generalisability of our results. First, the use of yearly aggregations and feature extraction techniques may obscure the temporal dynamics and lagged contributions of climate variables. Second, the study employs a limited set of models; as results may vary with alternative ML architectures or more advanced econometric approaches. Third, economic benchmark models rely on a narrow set of variables, and more sophisticated specifications could present more challenging benchmarks for ML models. Finally, the nowcasting results and feature importance values are sensitive to the chosen time periods and may vary across forecast horizons and regional contexts.

[18] The 5th and 95th percentiles of $\Delta\hat{y}_{t,r,s}$ for sector C equal -0.5 and 0, respectively.

Future research could address these limitations by expanding the set of models, both ML and econometric, and refining climate indicators to better reflect sector-specific vulnerabilities. This includes tuning the selection of heat wave and drought indicators to the sector being nowcasted, incorporating additional economic variables, and extending the analysis to additional sectors beyond agriculture and industry. Moreover, shifting the focus to price dynamics and regional migration patterns could offer deeper insights into the socio-economic consequences of weather extremes and their transmission across regions.

This study contributes to the growing literature on climate–economic modelling by demonstrating the value of predictive modelling approaches to better understand and respond to climate-induced economic risks. The integration of climate data into regional economic nowcasting may not only enhance accuracy but also support the development of more resilient, timely and adaptive policy frameworks across the EU.

References

Battiston, S., Monasterolo, I., van Ruijven, B., and Krey, V. (2022). The NACE – CPRS – IAM mapping: A tool to support climate risk analysis of financial portfolio using NGFS scenarios. SSRN. Available at SSRN: https://ssrn.com/abstract=4223606.

Bojinski, S., Verstraete, M., Peterson, T. C., Richter, C., Simmons, A., and Zemp, M. (2014). The concept of essential climate variables in support of climate research, applications, and policy. Bulletin of the American Meteorological Society, 95(9):1431–1443.

Breiman, L. (2001). Random forests. Machine Learning, 45(1):5–32.

Cammalleri, C., Vogt, J., and Salamon, P. (2017). Development of an operational low-flow index for hydrological drought monitoring over Europe. Hydrological Sciences Journal, 62(3):346–358.

Casanueva, A., Kotlarski, S., Fischer, A. M., Flouris, A. D., Kjellstrom, T., Lemke, B., Nybo, L., Schwierz, C., and Liniger, M. A. (2020). Escalating environmental summer heat exposure - a future threat for the European workforce. Regional Environmental Change, 20(40).

Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 785–794. ACM.

Ciccarelli, M., Kuik, F., and Martínez Hernández, C. (2023). The asymmetric effects of weather shocks on euro area inflation. ECB Working Paper 2798, European Central Bank. ECB Working Paper Series No. 2798.

de Roo, A., Wesseling, C., and van Deursen, W. (2000). Physically based river basin modelling within a GIS: the LISFLOOD model. Hydrological Processes, 14:1981–1992.

Dellink, R., Lanzi, E., and Chateau, J. (2019). The sectoral and regional economic consequences of climate change to 2060. Environmental and resource economics, 72:309–363.

Duo, J. (2025). Impact of climate policy uncertainty on energy price volatility: Evidence from China. Modern Economy, 16(6):884–903.

European Commission, Joint Research Centre, Cammalleri, C., Naumann, G., Mentaschi, L., Formetta, G., Forzieri, G., Gosling, S., Bisselink, B., De Roo, A., and Feyen, L. (2020). Global warming and drought impacts in the EU – JRC PESETA IV project – Task 7. Publications Office.

García-León, D., Casanueva, A., Standardi, G., Burgstall, A., Flouris, A. D., and Nybo, L. (2021). Current and projected regional economic impacts of heatwaves in Europe. Nature Communications, 12(5807).

Kahn, M. E., Mohaddes, K., Ng, R. N., Pesaran, M. H., Raissi, M., and Yang, J.-C. (2021). Long-term macroeconomic effects of climate change: A cross-country analysis. Energy Economics, 104:105624.

Kotz, M., Donat, M. G., Lancaster, T., Parker, M., Smith, P., Taylor, A., and Vetter, S. H. (2025). Climate extremes, food price spikes, and their wider societal risks. Environmental Research Letters, 20(8):081001.

Lavaysse, C., Cammalleri, C., Dosio, A., van der Schrier, G., Toreti, A., and Vogt, J. (2018). Towards a monitoring system of temperature extremes in Europe. Natural Hazards and Earth System Sciences, 18:91–104.

Lesk, C., Anderson, W., Rigden, A., Coast, O., Jägermeyr, J., McDermid, S., Davis, K. F., and Konar, M. (2022). Compound heat and moisture extreme impacts on global crop yields under climate change. Nature Reviews Earth & Environment, 3(12):872–889.

Li, Z., Long, J., and Li, L. (2022). A novel machine learning ensemble forecasting model based on mixed frequency technology and multi-objective optimization for carbon trading price. Applied Energy, 309:118376.

Lundberg, S. M., Erion, G. G., and Lee, S.-I. (2018). Consistent individualized feature attribution for tree ensembles. arXiv preprint arXiv:1802.03888.

Martín Cervantes, P. A., Rueda L´opez, N., and Cruz Rambaud, S. (2020). The effect of globalization on economic development indicators: An inter-regional approach. Sustainability, 12(5):1942.

McKee, T. B., Doesken, N. J., and Kleist, J. (1993). The relationship of drought frequency and duration to time scale. In Proceedings of the Eighth Conference on Applied Climatology, pages 179–184, Anaheim, California. American Meteorological Society.

Mohaddes, K., Ng, R. N., Pesaran, M. H., Raissi, M., and Yang, J.-C. (2022). Climate change and economic activity: evidence from US states. Oxford Open Economics, 2:odac010.

Motta, C., Naumann, G., Gomez, D., Formetta, G., and Feyen, L. (2025). Assessing the economic impact of droughts in Europe in a changing climate: A multi-sectoral analysis at regional scale. Journal of Hydrology: Regional Studies, 59:102296.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in python. Journal of Machine Learning Research, 12:2825–2830.

Praveen, B. and Sharma, P. (2019). A review of literature on climate change and its impacts on agriculture productivity. Journal of Public Affairs, 19(4):e1960.

Rossi, L., Wens, M., De Moel, H., Cotti, D., Sabino Siemons, A., Toreti, A., Maetens, W., Masante, D., Van Loon, A., Hagenlocher, M., Rudari, R., Naumann, G., Meroni, M., Avanzi, F., Isabellon, M., and Barbosa, P. (2023). European drought risk atlas. (KJ-NA-31-682-EN-N (online), KJ-NA-31-682-EN-C (print)).

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

Seneviratne, S., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S., Wehner, M., and Zhou, B. (2021). Weather and Climate Extreme Events in a Changing Climate. In Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J., Maycock, T., Waterfield, T., Yelekçi, O., Yu, R., and Zho, B., editors, Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, chapter 11, pages 1513–1766. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

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. Natural Hazards and Earth System Sciences, 12:3519–3531.

Shapley, L. S. (1953). A value for n-person games. Contributions to the Theory of Games, 2:307–317.

Sun, P.-Y., Otto, F. E. L., Kornhuber, K., and Levermann, A. (2024). Heat exposure, labour productivity losses and global supply chain risks. Nature, 629:63–69.

Tabari, H. and Willems, P. (2023). Global risk assessment of compound hot-dry events in the context of future climate change and socioeconomic factors. NPJ Climate and Atmospheric Science, 6(74):1–10.

Usman, S., González-Torres Fernández, G., and Parker, M. (2024). Going nuts: the regional impact of extreme climate events over the medium term. ECB Working Paper Series, 3002.

Usman, S., Parker, M., and Vallat, M. (2025). Dry-roasted nuts: Early estimates of the regional impact of 2025 extreme weather. Available at SSRN: http://dx.doi.org/10.2139/ssrn.5484206.

Wang, S. L., Ball, E., Nehring, R., Williams, R., and Chau, T. (2018). Impacts of climate change and extreme weather on us agricultural productivity: Evidence and projection. In Agricultural Productivity and Producer Behaviour, pages 41–75. University of Chicago Press.

Yang, B., Long, W., and Cai, Z. (2024). Machine learning based panel data models. Technical report, University of Kansas, Department of Economics.

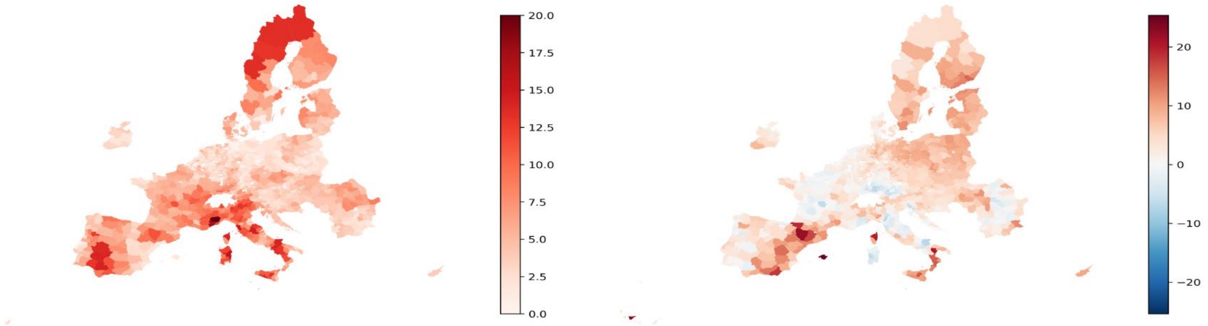
Yin, J. and Slater, L. (2023). Understanding heatwave-drought compound hazards and impacts on socio-ecosystems. The Innovation Geoscience, 1(3):100042.

Zhu, W. and Li, S. (2024). Nonlinear effects of climate risks on climate-sensitive sectors. Economic Change and Restructuring, 57(167).

Appendix

Heat Waves over the Hotter Months (May to September) — Figure A.1

(a) Average Heat Wave Days in 2002-2006 (b) Change in Averages between 2018-2022 & 2002-2006

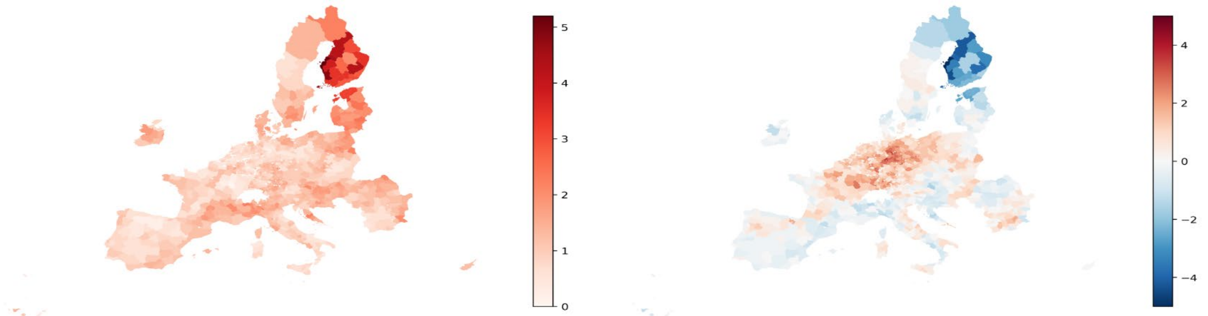


Note: Figure A.1 illustrates heat wave patterns across Europe during the summer months (May to September). Panel (a) presents the average annual number of heat wave days recorded between 2002 and 2006, using a white-to-dark red gradient (0 to 20 average heat wave days). Panel (b) depicts the change in frequency between 2018 and 2022 relative to the earlier period—calculated as the average annual number of heat wave days in 2018–2022 minus that of 2002–2006. The colour gradients convey the intensity of heat wave activity, with red indicating an increase in frequency and blue representing a decline.

Source: Authors' calculations using HCWI from the Copernicus European Drought Observatory.

Average Number of Meteorological Dry Months (Mean SPI01 < -1) per Year — Figure A.2

(a) Average Number of Dry Months in 2002-2006 (b) Change in Averages between 2018-2022 & 2002-2006



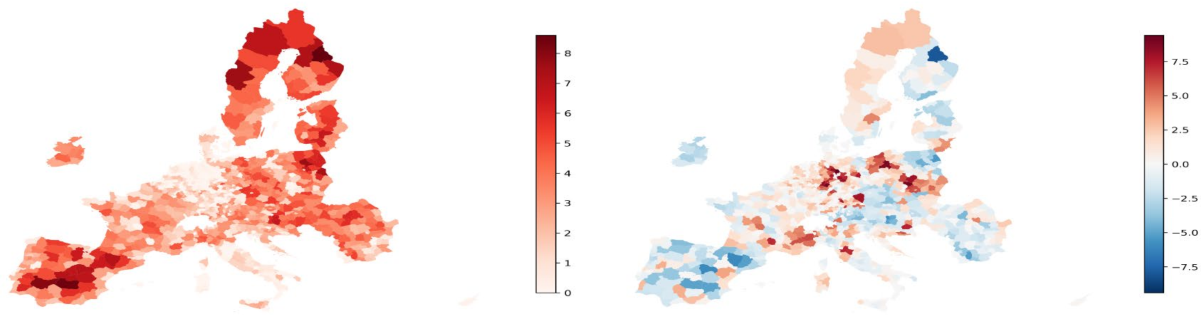
Note: Figure A.2 compares meteorological dry month patterns across Europe. Panel (a) shows the average number of dry months per year from 2002–2006, using a white-to-dark red gradient (0 to 5 months). Panel (b) displays the change in averages between 2018–2022 and 2002–2006, with a blue-to-red gradient indicating decreases and increases, respectively. Dry months are defined as those with mean SPI01 values below −1.

Source: Authors' calculations using SPI01 from the Copernicus European Drought Observatory.

Average Number of Hydrological Dry Months (Mean LFI > 0) per Year

Figure A.3

(a) Average Number of Dry Months in 2002-2006 (b) Change in Averages between 2018-2022 & 2002-2006



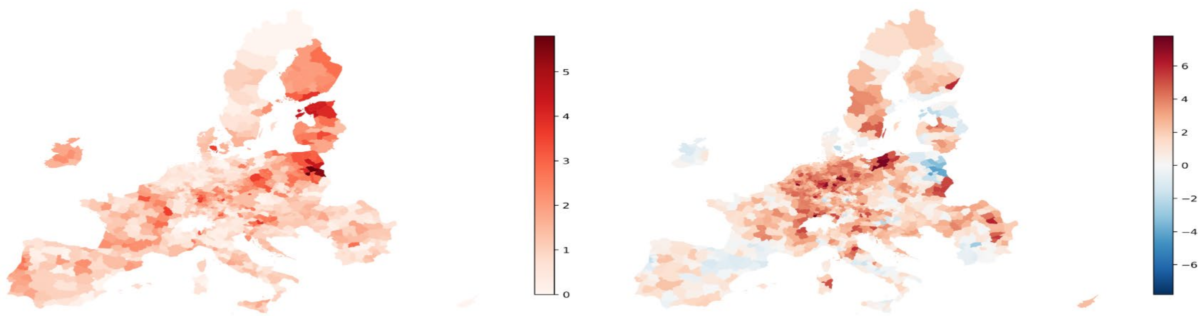
Note: Figure A.3 compares hydrological dry month patterns across Europe at the NUTS-3 regional level. Panel (a) shows the average number of dry months per year from 2002–2006, using a white-to-dark red gradient (0 to 8 months). Panel (b) displays the change in averages between 2018–2022 and 2002–2006, with a blue-to-red gradient indicating decreases and increases, respectively. Dry months are defined as those with mean LFI values above 0.

Source: Authors' calculations using LFI from the Copernicus European Drought Observatory.

Average Number of Agricultural Dry Months (Mean SMA < -1) per Year

Figure A.4

(a) Average Number of Dry Months in 2002-2006 (b) Change in Averages between 2018-2022 & 2002-2006



Note: Figure A.4 compares agricultural dry month patterns across Europe at the NUTS-3 regional level. Panel (a) shows the average number of dry months per year from 2002–2006, using a white-to-dark red gradient (0 to 5 months). Panel (b) displays the change in averages between 2018–2022 and 2002–2006, with a blue-to-red gradient indicating decreases and increases, respectively. Dry months are defined as those with mean SMA values below −1.

Source: Authors' calculations using SMA from the Copernicus European Drought Observatory.

Average SPI and HCWI values by country and period

Table A.1

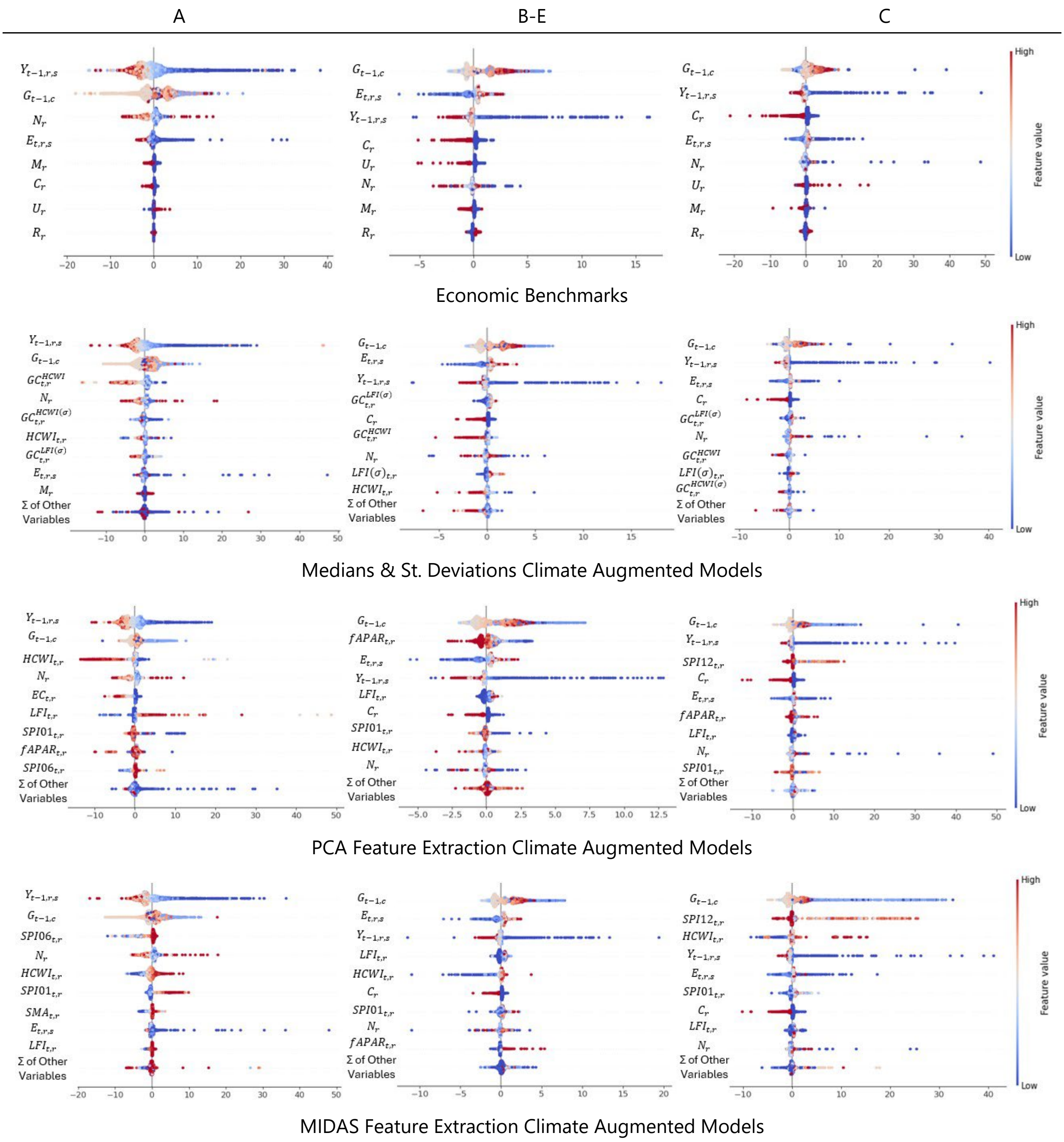
Country	Average Values Over 2002–2006				Average Values Over 2018–2022			
	SPI01	SPI06	SPI12	HCWI	SPI01	SPI06	SPI12	HCWI
Austria	0.0000	-0.0058	-0.0084	0.6131	0.0000	-0.0268	-0.0558	1.2414
Belgium	0.0000	-0.0125	-0.0438	0.7507	-0.0180	-0.2210	-0.2585	0.9366
Bulgaria	0.0000	0.0000	0.0000	0.5963	-0.0146	0.0000	-0.0147	0.8442
Cyprus	0.0000	0.0000	-0.2091	0.6297	0.0000	0.0000	0.0000	1.5645
Czech Republic	0.0000	0.0000	0.0000	0.9195	0.0000	-0.0167	-0.0383	1.5161
Germany	0.0000	-0.0044	-0.0329	0.6947	-0.0136	-0.1796	-0.2810	0.9897
Denmark	0.0000	0.0000	-0.0478	0.7650	0.0000	0.0000	-0.0202	1.2358
Estonia	0.0000	-0.4699	-0.4995	1.0835	0.0000	0.0000	0.0000	1.9151
Greece	0.0000	0.0000	0.0000	0.3528	-0.0069	-0.0418	-0.0488	0.8192
Spain	0.0000	-0.0888	-0.1499	0.6050	0.0000	-0.0045	-0.0153	1.4302
Finland	-0.1511	-0.6627	-0.8369	1.1448	0.0000	0.0000	-0.0119	1.7958
France	0.0000	-0.0257	-0.0638	0.8064	-0.0199	-0.0300	-0.0342	1.3075
Croatia	0.0000	0.0000	0.0000	0.5485	0.0000	0.0000	0.0000	0.8604
Hungary	0.0000	-0.0224	-0.0749	0.6807	0.0000	-0.0552	-0.0913	1.1856
Ireland	0.0000	0.0000	0.0000	0.4099	0.0000	0.0000	0.0000	0.6715
Italy	0.0000	-0.0079	-0.0823	1.0521	-0.0205	-0.0783	-0.0795	1.4173
Lithuania	0.0000	-0.1965	-0.2204	0.8041	-0.0430	0.0000	-0.1123	1.6761
Luxembourg	0.0000	0.0000	-0.4149	0.9569	0.0000	0.0000	0.0000	1.4145
Latvia	0.0000	-0.2935	-0.2907	1.0554	0.0000	-0.0471	-0.2489	1.9054
Malta	0.0000	-0.2156	-0.3345	0.4721	0.0000	0.0000	0.0000	0.8198
Netherlands	0.0000	0.0000	0.0000	0.6678	-0.0069	-0.0326	-0.1974	0.7780
Poland	0.0000	0.0000	-0.0091	0.7130	-0.0948	-0.0103	-0.0170	1.3940
Portugal	0.0000	-0.1631	-0.2619	0.5536	-0.0449	-0.0376	-0.1162	0.7040
Romania	0.0000	0.0000	0.0000	0.6957	-0.0055	0.0000	0.0000	1.3178
Sweden	0.0000	-0.0100	-0.0228	1.0263	-0.0417	0.0000	0.0000	1.6997
Slovenia	0.0000	0.0000	0.0000	0.5543	0.0000	0.0000	0.0000	0.8222
Slovakia	0.0000	0.0000	0.0000	0.7944	0.0000	0.0000	0.0000	1.4916

Note: Table A.1 displays the average values of SPI across the accumulation periods of 1, 6 and 12 months, and HCWI during the five-year periods of 2002 to 2006 and 2018 to 2022. Dryer (Hotter) conditions over time are suggested by smaller (larger) values in the 2018-2022 period, relative to that of 2002-2006.

Source: Authors' calculations using SPI and HCWI from the Copernicus European Drought Observatory.

SHAP Values of Random Forest Models

Figure A.5

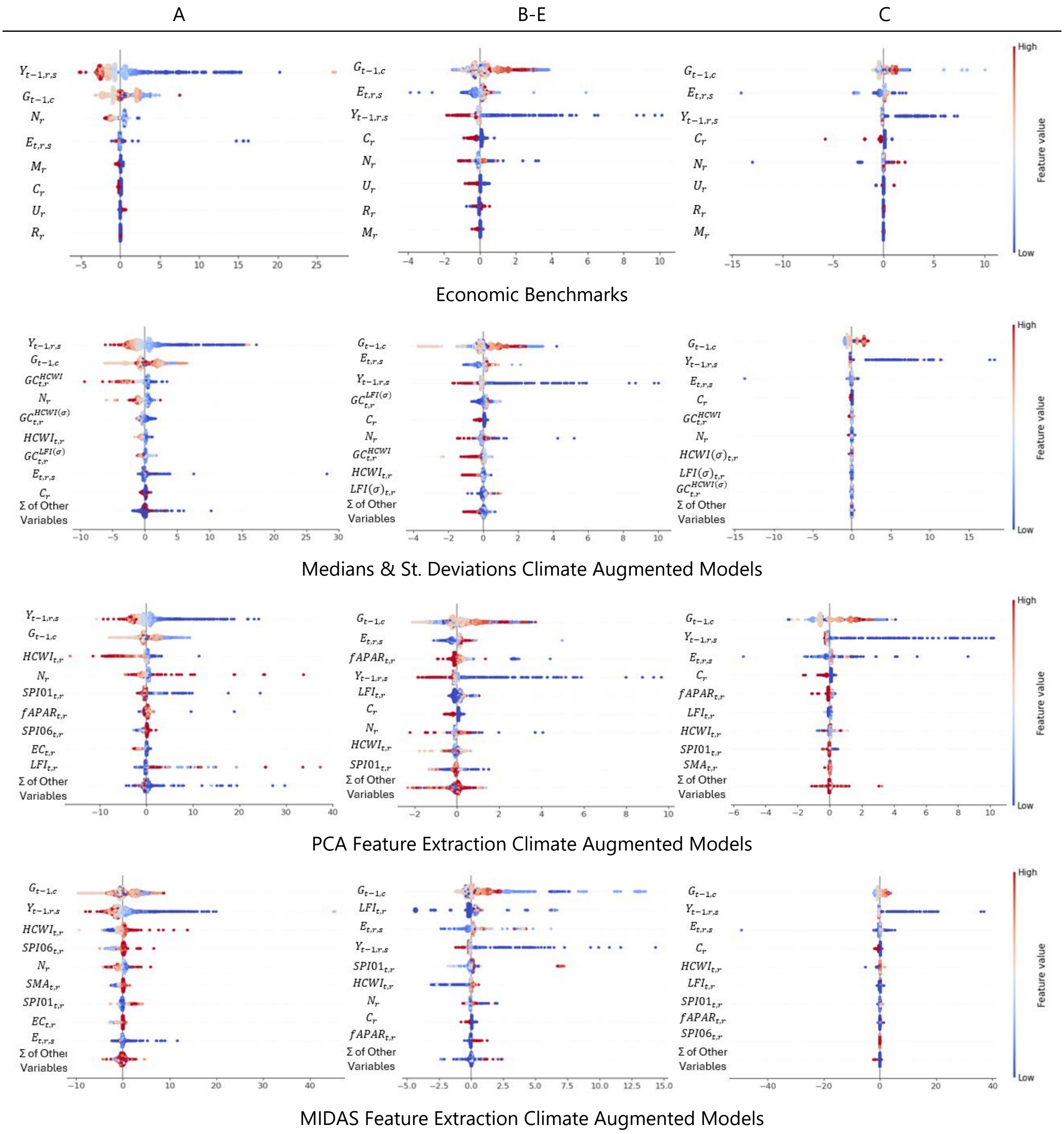


Note: The figure displays the SHAP beeswarm plots of all the estimated Random Forest models. The sub-figures visualise the marginal contribution of a feature to individual predictions.

Source: Authors' calculations.

SHAP Values of XGBoost Models

Figure A.6



Note: The figure displays the SHAP beeswarm plots of all the estimated XGBoost models. The sub-figures visualise the marginal contribution of a feature to individual predictions.

Source: Authors' calculations.

Average Predictive Contributions within the Medians & St. Deviations Models

Table A.2

	Sector A			Sector B-E			Sector C		
	Linear Reg.	Random Forest	XGBoost	Linear Reg.	Random Forest	XGBoost	Linear Reg.	Random Forest	XGBoost
N_r	0.3762	-0.2410	-0.2014	0.1309	-0.0815	-0.0296	0.8213	-0.1562	-0.0076
$E_{t,r,S}$	0.4090	-0.0067	-0.0132	0.5282	0.0084	0.0069	0.0625	0.0293	0.0015
$Y_{t-1,r,S}$	-3.2126	-0.0185	-0.0508	-0.7985	-0.0037	-0.0026	-1.7310	-0.0390	-0.0137
$G_{t-1,c}$	-0.2522	0.0013	-0.0019	0.3931	-0.0358	-0.0194	0.1740	-0.0003	-0.0141
M_r	-0.8854	-0.0272	-0.0117	-0.4552	0.0008	0.0000	-0.6775	0.0016	0.0000
U_r	1.1929	0.0065	0.0003	-1.2317	0.0009	-0.0002	0.0689	0.0014	0.0000
R_r	-0.2779	0.0002	-0.0002	0.2580	-0.0001	0.0005	0.0533	0.0013	0.0000
C_r	-1.4417	-0.0065	0.0004	-0.8270	-0.0071	0.0002	-2.0727	-0.0283	-0.0003
$HCWI_{t,r}$	-4.8562	-0.0958	-0.0065	-1.9701	0.0053	0.0016	0.4587	0.0090	0.0005
$Own\ Region$	1.6550	-0.0063	0.0046	0.2470	0.0018	0.0017	-0.1015	0.0047	0.0001
$GC_{t,r}^{HCWI}$	-6.5112	-0.0895	-0.0111	-2.2171	0.0036	-0.0001	0.5602	0.0043	0.0004
$SPI01_{t,r}$	-1.0812	0.0000	0.0000	2.8646	-0.0002	0.0000	2.3341	-0.0001	0.0000
$Own\ Region$	1.8573	0.0000	0.0000	-4.2821	-0.0002	0.0000	-4.4634	-0.0002	0.0000
$GC_{t,r}^{SPI01}$	-2.9385	0.0000	0.0000	7.1468	0.0000	0.0000	6.7976	0.0001	0.0000
$SPI06_{t,r}$	7.5284	0.0254	0.0000	0.4315	0.0000	0.0001	0.1621	0.0006	0.0000
$Own\ Region$	3.2179	0.0251	0.0010	0.5592	0.0003	0.0000	0.0181	0.0008	0.0000
$GC_{t,r}^{SPI06}$	4.3105	0.0004	-0.0010	-0.1277	-0.0003	0.0001	0.1440	-0.0003	0.0000
$SPI12_{t,r}$	-2.2817	-0.0006	0.0068	0.3016	-0.0007	-0.0001	0.5485	-0.0027	0.0000
$Own\ Region$	-0.9168	0.0013	0.0071	0.2562	-0.0004	0.0002	0.2799	-0.0021	0.0000
$GC_{t,r}^{SPI12}$	-1.3649	-0.0019	-0.0002	0.0455	-0.0003	-0.0004	0.2686	-0.0007	0.0000
$LFI_{t,r}$	-18.3530	0.1007	0.0392	16.2342	0.0169	0.0023	15.1460	0.0194	0.0000
$Own\ Region$	-6.7037	0.0433	0.0111	4.7357	0.0059	0.0014	6.7405	0.0116	0.0001
$GC_{t,r}^{LFI}$	-11.6493	0.0573	0.0281	11.4985	0.0109	0.0010	8.4055	0.0079	0.0000
$SMA_{t,r}$	3.0591	0.0074	0.0061	0.5488	-0.0005	0.0002	0.8182	-0.0007	0.0000
$fAPAR_{t,r}$	-12.1011	-0.0075	-0.0083	0.5397	0.0004	0.0001	1.0788	-0.0002	0.0000
$EC_{t,r}$	-11.7883	0.0000	0.0000	-137.8228	0.0000	0.0001	-114.0757	0.0000	0.0000
$EC_{t,r}^{01}$	-3.9294	0.0000	0.0000	-45.9409	0.0000	0.0000	-38.0252	0.0000	0.0000
$EC_{t,r}^{06}$	-3.9294	0.0000	0.0000	-45.9409	0.0000	0.0000	-38.0252	0.0000	0.0000
$EC_{t,r}^{12}$	-3.9294	0.0000	0.0000	-45.9409	0.0000	0.0000	-38.0252	0.0000	0.0000

Note: The table presents the average contributions to predicted outcomes across all models utilising yearly medians and standard deviations. For linear models, regression coefficients are reported, while for ML models, the average SHAP values are used to quantify feature importance. Heat wave and drought values are represented by the sum of the coefficients, or SHAP values, oof median and standard deviation variables. Climate-related features are decomposed into own region, geographic compounding, and event compounding types (highlighted in grey).

Source: Authors' calculations.

Average Predictive Contributions within PCA-based Features Extraction Models

Table A.3

	Sector A			Sector B-E			Sector C		
	Linear Reg.	Random Forest	XGBoost	Linear Reg.	Random Forest	XGBoost	Linear Reg.	Random Forest	XGBoost
N_r	-4.0549	-0.2976	-0.1316	-0.8570	-0.0653	-0.0269	-1.7629	-0.1504	-0.0141
$E_{t,r,S}$	0.5007	0.0383	0.0024	0.5082	-0.0036	0.0044	0.0446	0.0253	0.0113
$Y_{t-1,r,S}$	-0.8888	0.0491	-0.0010	0.3444	-0.0026	0.0000	0.1142	-0.0398	-0.0115
$G_{t-1,c}$	0.7968	-0.0842	-0.1136	0.1441	-0.0347	0.0088	0.8340	0.0374	-0.0197
M_r	-1.4326	0.0028	-0.0024	-0.6199	0.0005	0.0002	-0.8922	0.0023	0.0000
U_r	1.1591	0.0019	0.0013	-1.1642	0.0008	0.0006	0.1257	0.0053	0.0000
R_r	-0.2013	0.0007	-0.0003	0.2456	-0.0004	0.0003	0.0246	0.0016	0.0001
C_r	-1.7584	-0.0010	0.0039	-0.8752	-0.0035	0.0002	-2.1044	-0.0283	-0.0018
$HCWI_{t,r}$	-0.7032	-0.0584	-0.0074	0.0148	-0.0031	-0.0064	0.0821	-0.0014	-0.0014
$Own\ Region$	-0.3692	-0.0307	-0.0039	0.0078	-0.0016	-0.0034	0.0430	-0.0007	-0.0008
$GC_{t,r}^{HCWI}$	-0.3340	-0.0278	-0.0035	0.0070	-0.0015	-0.0031	0.0392	-0.0007	-0.0007
$SPI01_{t,r}$	2.8046	0.0187	0.0249	0.6473	-0.0174	-0.0101	0.8969	-0.0174	-0.0025
$Own\ Region$	1.3798	0.0092	0.0122	0.3185	-0.0085	-0.0050	0.4413	-0.0085	-0.0012
$GC_{t,r}^{SPI01}$	1.4248	0.0095	0.0126	0.3288	-0.0088	-0.0051	0.4556	-0.0089	-0.0013
$SPI06_{t,r}$	-1.9742	0.0264	0.0055	-0.7002	0.0022	-0.0054	-0.9495	0.0218	0.0039
$Own\ Region$	-1.0172	0.0136	0.0028	-0.3615	0.0011	-0.0028	-0.4857	0.0111	0.0020
$GC_{t,r}^{SPI06}$	-0.9570	0.0129	0.0026	-0.3386	0.0011	-0.0026	-0.4638	0.0106	0.0019
$SPI12_{t,r}$	0.9908	-0.0217	-0.0296	0.2678	0.0171	0.0029	0.4289	-0.0456	0.0098
$Own\ Region$	0.5307	-0.0116	-0.0159	0.1444	0.0092	0.0016	0.2299	-0.0244	0.0053
$GC_{t,r}^{SPI12}$	0.4601	-0.0101	-0.0138	0.1233	0.0080	0.0013	0.1990	-0.0211	0.0045
$LFI_{t,r}$	9.1629	-0.0085	0.0174	1.6230	0.0081	0.0056	1.4679	0.0193	0.0023
$Own\ Region$	5.0917	-0.0047	0.0097	0.9098	0.0045	0.0031	0.8130	0.0107	0.0013
$GC_{t,r}^{LFI}$	4.0712	-0.0038	0.0077	0.7133	0.0036	0.0025	0.6549	0.0086	0.0010
$SMA_{t,r}$	-0.2708	0.0239	0.0287	0.0422	0.0030	-0.0022	-0.0149	0.0069	0.0013
$fAPAR_{t,r}$	-0.9959	0.0864	0.0658	0.3970	0.0352	0.0122	0.1610	0.0445	0.0086
$EC_{t,r}$	-18.9298	-0.0021	0.0012	1.0696	0.0049	0.0027	2.2196	0.0070	0.0005
$EC_{t,r}^{01}$	-9.1976	-0.0010	0.0006	0.5171	0.0024	0.0013	1.0817	0.0034	0.0002
$EC_{t,r}^{06}$	-7.1517	-0.0008	0.0005	0.4060	0.0018	0.0010	0.8375	0.0026	0.0002
$EC_{t,r}^{12}$	-2.5805	-0.0003	0.0002	0.1465	0.0007	0.0004	0.3004	0.0009	0.0001

Note: The table presents the average contributions to predicted outcomes across all models utilising PCA-based feature extraction. For linear models, regression coefficients are reported, while for ML models, the average SHAP values are used to quantify feature importance. Climate-related features are decomposed into own region, geographic compounding, and event compounding types (highlighted in grey), based on factor loadings derived from the PCA kernel.

Source: Authors' calculations.

Average Predictive Contributions within MIDAS-based Features Extraction Models

Table A.4

	Sector A			Sector B-E			Sector C		
	Linear Reg.	Random Forest	XGBoost	Linear Reg.	Random Forest	XGBoost	Linear Reg.	Random Forest	XGBoost
N_r	-3.5201	-0.2815	-0.2119	-0.8314	-0.0726	-0.0270	-1.7328	-0.1662	0.0040
$E_{t,r,S}$	0.2516	0.0480	0.0063	0.5049	0.0038	0.0085	0.0215	0.0284	0.0097
$Y_{t-1,r,S}$	-0.5562	-0.0203	-0.0291	0.3968	-0.0032	-0.0068	0.2238	-0.0350	-0.0181
$G_{t-1,c}$	0.4901	-0.0569	-0.1521	0.1510	-0.0324	0.0025	0.8092	0.0384	-0.0089
M_r	-0.7923	-0.0054	0.0012	-0.2416	0.0011	-0.0001	-0.4998	0.0033	0.0000
U_r	1.1793	0.0030	0.0024	-1.2466	0.0015	0.0008	0.0199	0.0013	-0.0001
R_r	-0.0926	0.0013	0.0012	0.2726	0.0006	0.0000	0.0238	0.0017	0.0000
C_r	-1.8548	-0.0040	0.0038	-0.8945	-0.0051	-0.0008	-2.0580	-0.0256	-0.0227
$HCWI_{t,r}$	0.7173	-0.0470	0.0085	1.1375	0.0064	-0.0006	1.0037	0.0309	0.0074
$Own\ Region$	0.6388	-0.0419	0.0069	0.2571	0.0020	-0.0001	0.1099	0.0155	0.0049
$GC_{t,r}^{HCWI}$	0.0785	-0.0051	0.0016	0.8803	0.0044	-0.0006	0.8939	0.0155	0.0025
$SPI01_{t,r}$	3.1600	-0.0647	-0.0311	0.7348	-0.0015	-0.0053	0.8004	-0.0015	-0.0036
$Own\ Region$	2.8142	-0.0576	-0.0252	0.6543	-0.0013	-0.0043	0.7128	-0.0014	-0.0029
$GC_{t,r}^{SPI01}$	0.3459	-0.0071	-0.0058	0.0804	-0.0002	-0.0010	0.0876	-0.0002	-0.0007
$SPI06_{t,r}$	1.2569	-0.0129	0.0297	-5.7149	0.0037	-0.0015	5.7065	-0.0040	-0.0053
$Own\ Region$	1.1193	-0.0115	0.0241	-1.2918	0.0012	-0.0002	0.6246	-0.0020	-0.0035
$GC_{t,r}^{SPI06}$	0.1376	-0.0014	0.0056	-4.4231	0.0026	-0.0013	5.0819	-0.0020	-0.0018
$SPI12_{t,r}$	-0.3859	0.0043	0.0036	1.2962	0.0003	0.0007	20.8891	-0.0364	-0.0005
$Own\ Region$	-0.3437	0.0038	0.0029	0.2930	0.0001	0.0001	2.2863	-0.0182	-0.0004
$GC_{t,r}^{SPI12}$	-0.0422	0.0005	0.0007	1.0032	0.0002	0.0006	18.6029	-0.0182	-0.0002
$LFI_{t,r}$	0.2515	0.0074	-0.0100	0.8496	0.0136	-0.0033	1.0232	0.0290	0.0104
$Own\ Region$	0.2240	0.0066	-0.0081	0.1921	0.0043	-0.0004	0.1120	0.0145	0.0068
$GC_{t,r}^{LFI}$	0.0275	0.0008	-0.0019	0.6576	0.0093	-0.0029	0.9112	0.0145	0.0036
$SMA_{t,r}$	0.7121	0.0316	0.0250	0.6634	-0.0001	-0.0010	1.3874	-0.0090	0.0000
$fAPAR_{t,r}$	1.9427	0.0126	0.0339	0.7422	0.0066	0.0070	0.6416	0.0020	0.0011
$EC_{t,r}$	0.2448	0.0032	0.0095	12.1794	0.0037	0.0031	-4.7344	0.0039	-0.0001
$EC_{t,r}^{01}$	0.1564	0.0020	0.0050	0.5165	0.0004	0.0001	-0.0788	0.0010	0.0000
$EC_{t,r}^{06}$	0.0843	0.0011	0.0041	6.5674	0.0018	0.0011	-1.6312	0.0019	-0.0001
$EC_{t,r}^{12}$	0.0041	0.0001	0.0004	5.0955	0.0015	0.0020	-3.0245	0.0010	0.0000

Note: The table presents the average contributions to predicted outcomes across all models utilising MIDAS-based feature extraction. For linear models, regression coefficients are reported, while for ML models, the average SHAP values are used to quantify feature importance. Climate-related features are decomposed into own region, geographic compounding, and event compounding types (highlighted in grey), based on factor loadings derived from the beta polynomial function.

Source: Authors' calculations.