npj | climate and atmospheric science

Article

Published in partnership with CECCR at King Abdulaziz University

https://doi.org/10.1038/s41612-024-00722-1

Unveiling teleconnection drivers for heatwave prediction in South Korea using explainable artificial intelligence

Yeonsu Lee[1], **Dongjin Cho**[1], **Jungho Im**[1,2,3] ✉, **Cheolhee Yoo**[4], **Joonlee Lee**[1], **Yoo-Geun Ham**[5] & **Myong-In Lee**[1,3]

Increasing heatwave intensity and mortality demand timely and accurate heatwave prediction. The present study focused on teleconnection, the influence of distant land and ocean variability on local weather events, to drive long-term heatwave predictions. The complexity of teleconnection poses challenges for physical-based prediction models. In this study, we employed a machine learning model and explainable artificial intelligence to identify the teleconnection drivers for heatwaves in South Korea. Drivers were selected based on their statistical significance with annual heatwave frequency ($|R| > 0.3$, $p < 0.05$). Our analysis revealed that two snow depth (SD) variabilities—a decrease in the Gobi Desert and increase in the Tianshan Mountains—are the most important and predictive teleconnection drivers. These drivers exhibit a high correlation with summer climate conditions conducive to heatwaves. Our study lays the groundwork for further research into understanding land–atmosphere interactions over these two SD regions and their significant impact on heatwave patterns in South Korea.

Record-breaking heatwaves are becoming more frequent around the world[1] and have significant detrimental effects on human mortality and economic stability[2–4]. The year 2023 was the hottest year on record and the first year in which the daily global average temperature exceeded 1 °C above pre-industrial levels for every day of the year[5]. South Korea is among the regions where warming is faster than that of the global average[6]. As average summer temperatures rise, the number of people with heat illnesses continues to rise; hence, early and accurate prediction of heat waves is paramount.

Teleconnection, the phenomenon whereby variability in distant land and ocean influences the climate in other regions, is crucial to enhancing the accuracy of long-term climate predictions. Heatwaves in South Korea are significantly influenced by the formation of high-pressure systems, often driven by large-scale atmospheric configurations such as the Scandinavian (SCAND), CircumGlobal Teleconnection (CGT), and Pacific–Japan (PJ) patterns[7–10]. However, there have been changes over time in the impact of these teleconnection patterns on heatwaves in South Korea in recent years, including a weakened impact of the PJ pattern[7] and CGT[9], and a strengthened impact of SCAND[9]. Climate change and its resultant changes in land and ocean conditions are expected to complexify large-scale atmospheric circulations, reducing the influence of traditional teleconnection triggers and allowing new regional variabilities to emerge as potential drivers[11,12].

Various land and ocean variables are known to substantially influence upper atmospheric circulation, including snow depth (SD)[13–15], soil moisture (SM)[16–18], sea surface temperature (SST)[19–21], and sea-ice concentration (SIC)[22,23]. While General Circulation Models (GCMs) attempt to resolve the interactions between these components and the atmosphere, accurately representing these complex interactions remains a challenge[24]. In response to this challenge, data-driven models have recently been adopted as they can more easily capture these complex interactions with higher prediction performance than physics-based approaches. These models utilize a range of algorithms, including multi-linear regression (MLR), support vector regression, and artificial neural networks, and leverage historical data to reveal nonlinear relationships between these land and ocean variables and specific climate conditions such as heatwaves or precipitation[25–27].

In data-driven modeling, the selection of input variables is critical for identifying relationships with output variables, and contributes to increasing forecasting skill[28]. Previous studies conducted correlation analysis on global SST anomalies to identify potential teleconnection drivers for predicting climate variability and identified the regions exhibiting high correlation as

---

[1]Department of Civil, Urban, Earth, and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, South Korea. [2]Artificial Intelligence Graduate School, UNIST, Ulsan, South Korea. [3]Graduate School of Carbon Neutrality, UNIST, Ulsan, South Korea. [4]Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong. [5]Department of Environmental Planning, Graduate School of Environmental Studies, Seoul National University, Seoul, South Korea. ✉e-mail: ersgis@unist.ac.kr

independent variables[25–27]. For example, Lim and Seo[26] selected three regions of spring SST anomalies as input variables, built an MLR model to predict the number of heatwave days, and achieved skillful prediction performance. However, these studies only examined SST and used simple rectangular boundaries, which may have contained a certain amount of noise.

This study aimed to identify predictive teleconnection drivers for heatwaves in South Korea using machine learning (ML) and explainable artificial intelligence (XAI). In constructing the input variables for ML, we proposed a novel approach to finding potential teleconnection drivers from multiple land and ocean variables by applying a clustering method rather than using rectangular boundaries. We introduced SHapley Additive exPlanation (SHAP) analysis[29], an XAI method, to elucidate how selected teleconnection drivers contribute to predictions within an ML model. We also attempted to bridge the gap between theoretically important drivers determined through SHAP analysis and their practical impact on the summer climate. Our study provides a pioneering method for climate science, utilizing the power of ML and XAI to decode the complex teleconnections crucial for predicting heatwaves.

Results

Selected teleconnection drivers for heatwave prediction in South Korea

Using the proposed clustering method for statistically significant pixels with an annual heatwave frequency (HF) in South Korea (see Methods for details), a total of 16 teleconnection drivers were identified (Fig. 1). The variability in the SD was selected for two regions in December to February (DJF) and four regions in March to May (MAM). The two DJF drivers were in the Tianshan Mountains (DJF TM SD) and eastern Tibetan Plateau (DJF ET SD). Four MAM drivers were positioned across the western and eastern Eurasian continents (MAM WE SD and MAM EE SD), the Gobi Desert (MAM GD SD), and northern China (MAM NC SD). Owing to the frozen soil surface in winter, the variability in SM was investigated for MAM, and three regions were selected: northwestern Mongolia and southern Russia (MAM NM SM), northern China (MAM NC China), and central China (MAM CC SM).

Regions associated with well-established teleconnection patterns were selected for SST. The Indo-Pacific warm pool region (DJF IP SST and MAM WP SST), central Pacific (MAM CP SST), and eastern tropical Pacific (MAM EP SST) were related to the El Niño–Southern Oscillation (ENSO). In addition, the North Atlantic (DJF NA SST) was associated with the North Atlantic Oscillation (NAO). The selected regions of SIC variability were across the Barents and Kara seas (DJF BK SIC and MAM BK SIC), which are noted for the greatest SIC decline in the Arctic Sea from 1979 to 2018[30].

Predictability of selected teleconnection drivers in ML model

The light gradient boosting model (LGBM) algorithm[31] is known for its strengths in growing complex decision trees with high precision and preventing overfitting even with a small sample size[32,33]. We hypothesized that LGBM would excel in predicting HF and tested a total of nine regression algorithms, including both the machine learning and linear regression techniques (Supplementary Fig. 1). LGBM demonstrated superior performance in terms of root mean square error (RMSE) and correlation coefficient (R) among the algorithms tested, we constructed the HF prediction model using LGBM. The predictability for HF using teleconnection drivers in the LGBM model was evaluated and compared with simple statistical and climate models, MLR, and the Pusan National University Coupled GCM (PNU CGCM). We used leave-one-year-out cross-validation (LOYOCV) to evaluate the prediction performance, using as many samples as possible and hindcast validation to compare prediction performance with the operational climate model, PNU CGCM (see Methods for details). LGBM exhibited the most superior predictability for HF in South Korea. LGBM consistently outperformed MLR in RMSE, R, and mean square skill score (MSSS) for both validation methods (Table 1). LGBM also achieved a lower RMSE (4.303 days), higher R (0.756), and higher MSSS (0.488), outperforming PNU CGCM, which had an RMSE of 6.586 days, an R of 0.264, and an MSSS of −0.200.

Table 1 | Comparison of prediction performance for the Light Gradient Boosting Model (LGBM), Multi-Linear Regression (MLR), and Pusan National University Coupled General Circulation Model (PNU CGCM) using leave-one-year-out cross-validation (LOYOCV) and hindcast validation

Model	LOYOCV			Hindcast validation		
	RMSE	R	MSSS	RMSE	R	MSSS
LGBM	3.151	0.804	0.644	4.303	0.756	0.488
MLR	3.740	0.731	0.500	4.447	0.678	0.453
PNU CGCM	-	-	-	6.586	0.264	−0.200

Performance metrics—root mean square error (RMSE), correlation coefficient (R), and mean square skill score (MSSS)—were computed by aggregating the outputs from individual annual predictions for each model.

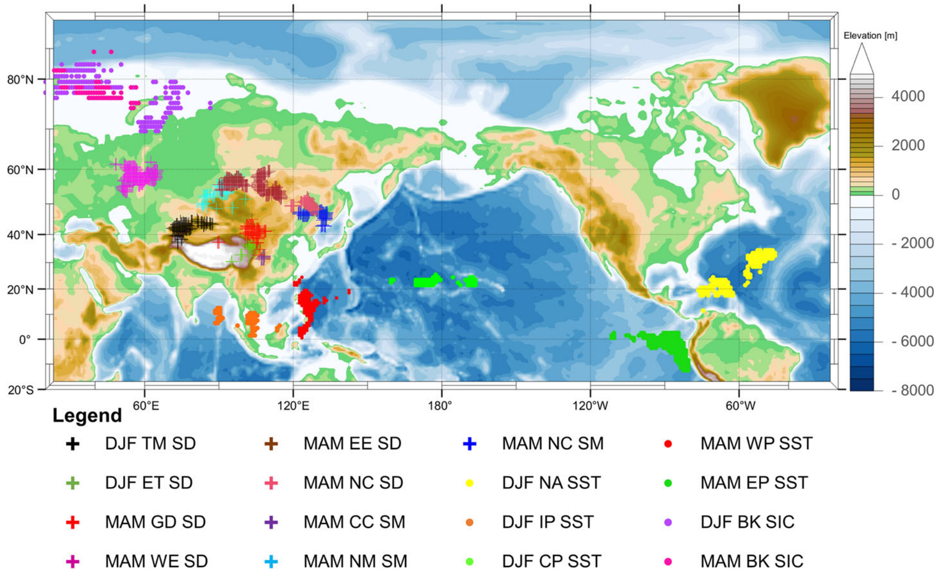


Fig. 1 | Geographical distribution of selected teleconnection drivers for heatwave frequency in South Korea. The map shows the location of selected teleconnection drivers for heatwave frequency in South Korea. Climate components include snow depth (SD), soil moisture (SM), sea surface temperature (SST), and sea-ice concentration (SIC) during winter (December-January-February) and spring (March–April–May). Clusters representing teleconnection drivers are marked by color-coded symbols corresponding to the legend; the background shows elevation. The full names of teleconnection drivers denoted by acronyms in the legend are listed in Supplementary Table 1.

Even though the MLR utilized same input variables as the LGBM model, it showed inferior prediction skills. The differences in RMSE, R, and MSSS between MLR and LGBM highlight the advantages of using advanced ML techniques for HF prediction. Notably, LGBM's ability to capture nonlinear relationships and complex interactions among teleconnection drivers likely enhances its performance. In contrast, MLR establishes a linear relationship between multiple teleconnection drivers and HF, which may not fully capture the underlying complexity. LGBM's various hyperparameters were meticulously optimized by evaluating a range of parameter combinations and selecting the optimal one based on the lowest mean squared error. We tested various parameters (see Methods for details), and examples of the distribution of mean squared errors for each parameter setting can be seen in Supplementary Fig. 2.

In the annual time series of LOYOCV (red solid line in Fig. 2a), we observed that the HF predicted by the LGBM model closely matched the reference HF calculated from the in situ daily maximum temperature (Tmax). In the annual time series of hindcast validation (red solid line in Fig. 2b), we observed that the predicted HF showed a slight increase in the deviation from the reference HF. This higher accuracy of LOYOCV could be due to the use of all available data for training, except for the predicted year. In hindcast validation, annual HF predictions were created by integrating that year's data into the training sample for the following year without using future data (see Methods). The inability to learn the most recent patterns in hindcast validation is thought to contribute to lower accuracy. As demonstrated in the gray area of Fig. 2b, the calibration performance was impacted by the iterative updates to the training samples, leading to a wide range of calibrated outputs that improved as the diversity and quantity of samples rose.

Even under hindcast validation, LGBM outperformed PNU CGCM (blue solid line in Fig. 2b) for years with significant heatwaves, such as 2012, 2013, and 2016. Moreover, LGBM maintained close alignment with the observed HF in the recent years of 2020 and 2021, when the PNU CGCM predictions showed large divergences. This underscores the usefulness and robustness of the LGBM model in capturing years with extreme cases.

A crucial component of predictability evaluation goes beyond R and RMSE computations and pertains to the model's capacity to forecast interannual variation within classification criteria (less normal, normal, or above normal). To assess the categorical predictability, we constructed a contingency table using ±1 standard deviation from the observed HF as the threshold (Table 2). The "Above normal" category exhibited a precision of 0.833 and a recall of 0.714, indicating that the LGBM model correctly predicted "Above normal" in 83% of instances and captured 71% of actual extreme heatwave years. Notably, LGBM correctly predicted the extreme heatwaves in 1973, 1994, 2013, 2016, and 2018. The "Normal" category exhibited a precision of 0.896 and a recall of 0.876, indicating that the LGBM model accurately predicted and captured the year having heatwaves as normal. The LGBM model performance in the "Less normal" category was lower, with the tendency of underrating years with normal HF as "Less normal" and overrating years with low HF as "Normal."

Contribution of drivers to HF predictions within the ML model

SHAP values represented whether each selected teleconnection driver had an increasing or decreasing impact on the HF predictions within the LGBM

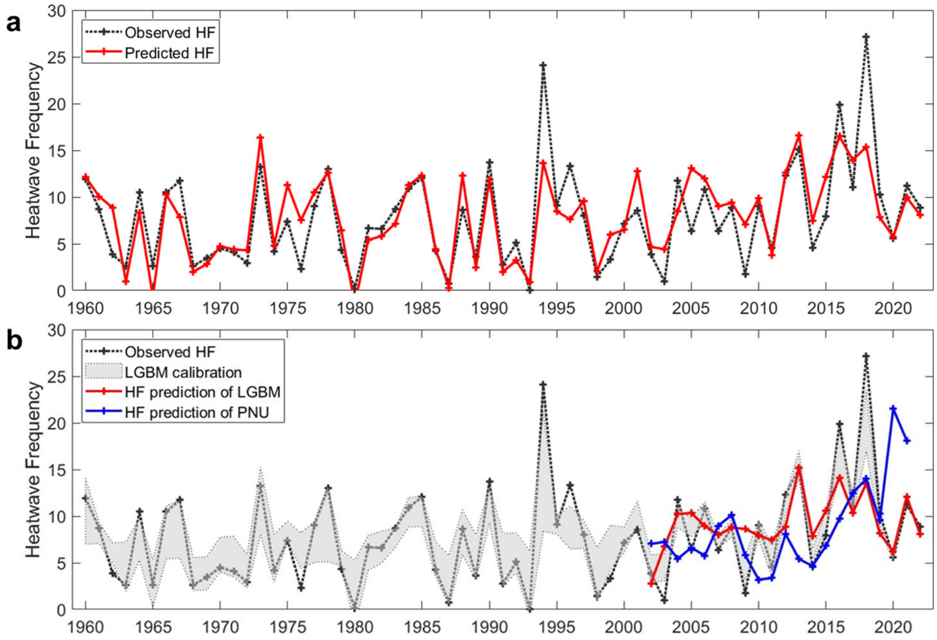


Fig. 2 | Time series of observed and predicted heatwave frequency (HF) in South Korea. **a** Observed HF (black dotted line) and predicted HF from the light gradient boosting model (LGBM) model in the leave-one-year-out-cross-validation (LOYOCV) approach (red solid line). **b** Observed HF (black dotted line), predicted HF (red solid line), range of calibration outputs from the LGBM model in the hindcast approach (gray shaded area), and predicted HF from Pusan National University (PNU) Coupled General Circulation Model (blue solid line).

Table 2 | Contingency table assessing the categorical predictability of Heatwave Frequency (HF) using the Light Gradient Boosting Model (LGBM) model

		Observed HF			
		Less normal	Normal	Above normal	Precision
Predicted HF	Less normal	4	5	0	0.444
	Normal	3	43	2	0.896
	Above normal	0	1	5	0.833
	Recall	0.571	0.876	0.714	

The table categorizes HF predictions into three categories: less normal, normal, and above normal, based on ±1 standard deviation from observed HF.

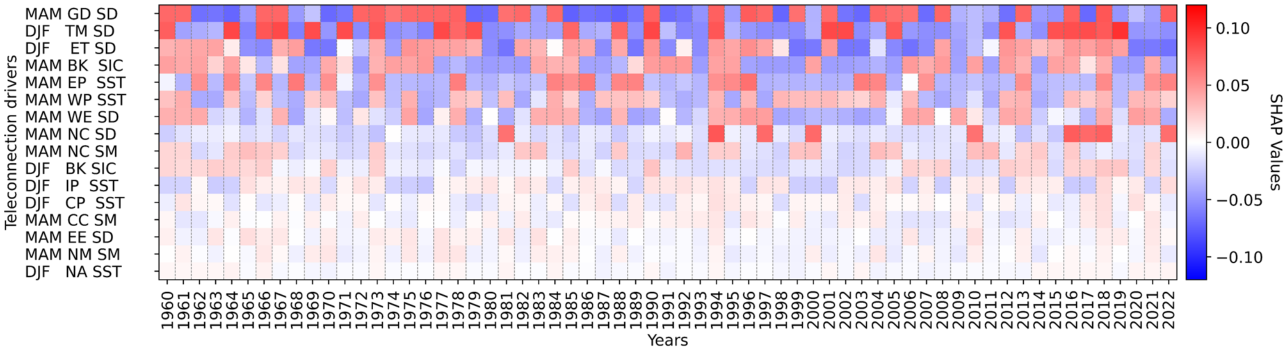


Fig. 3 | Time series of SHAP values for teleconnection drivers. The x-axis represents the years 1960 to 2022. The teleconnection drivers are arranged on the y-axis in descending order of their mean absolute SHapley Additive exPlanation (SHAP) values, meaning that drivers higher on the y-axis more strongly influence heatwave frequency (HF) prediction. Red (positive SHAP values) denotes teleconnection drivers that increase the HF prediction output, while blue (negative SHAP values) denotes teleconnection drivers that decrease the HF prediction output. The color gradient represents the intensity of the SHAP values, with darker shades indicating stronger influences.

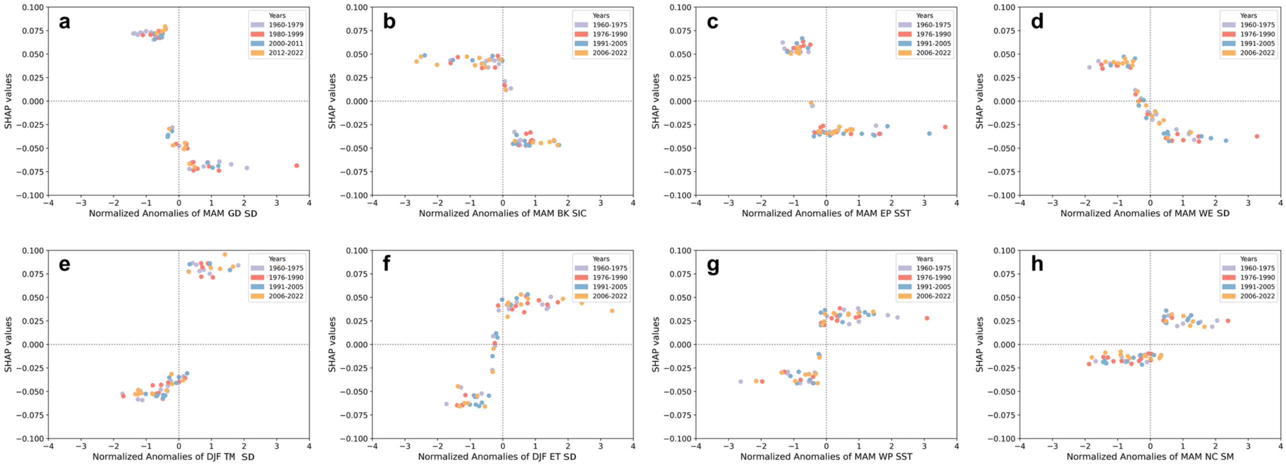
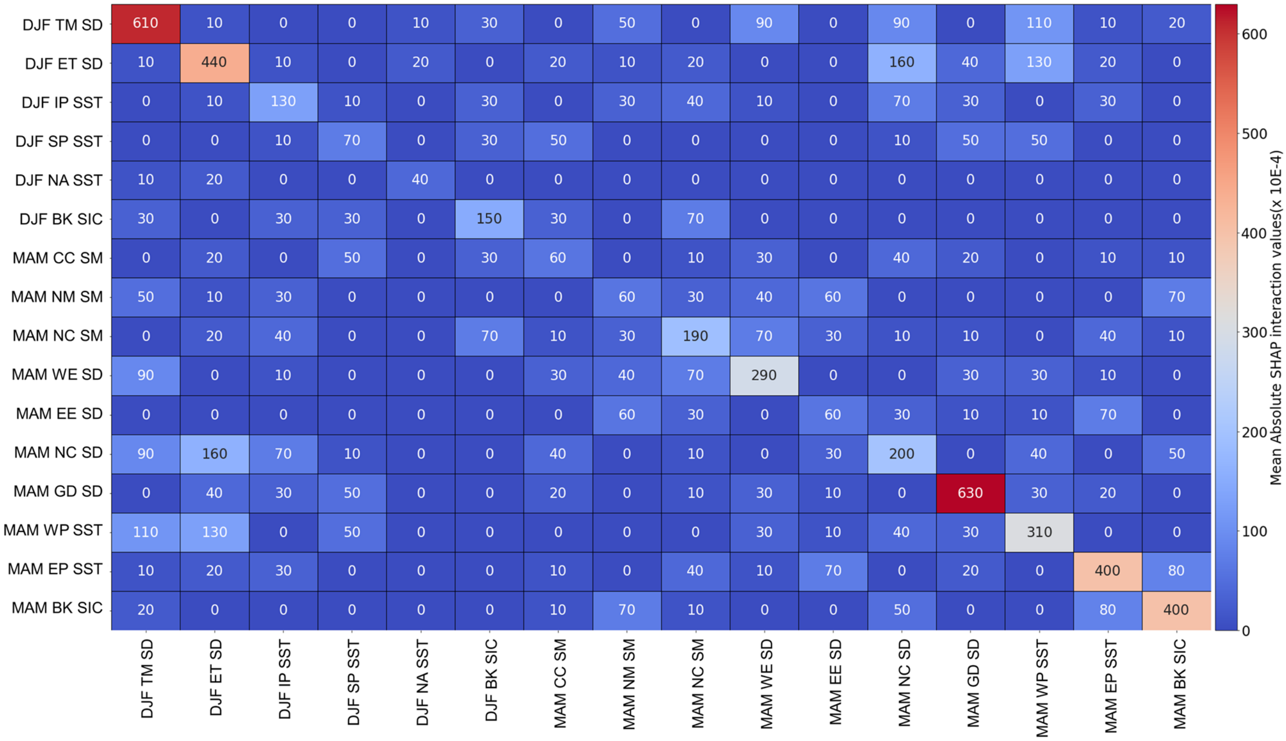


Fig. 4 | Scatter plots of the top eight important input features. The top eight teleconnection drivers with high mean absolute SHapley Additive exPlanation (SHAP) values are shown: (**a**) MAM GD SD; (**b**) MAM BK SIC; (**c**) MAM EP SST; (**d**) MAM WE SD; (**e**) DJF TM SD; (**f**) DJF ET SD; (**g**) MAM WP SST; (**h**) MAM NC SM. The full names of teleconnection drivers denoted by acronyms in the legend are listed in Supplementary Table 1. The x-axis indicates the anomaly values of input features, which are normalized to a range between −4 and 4. This normalization facilitates the comparison of features with different scales and units by bringing them to a common range. The y-axis indicates the SHAP values. To discern temporal changes in the anomalies and SHAP values, the data points are color-coded (purple, red, blue, and yellow), corresponding to four periods: 1960–1975, 1976–1990, 1991–2005, and 2006–2022, respectively.

model, and how much of an impact. A closer examination of Fig. 3 revealed that MAM GD SD was the most influential driver. The order of the y-axis in Fig. 3 suggests that MAM variables had more substantial influences than DJF variables. However, there are a few exceptions; DJF TM SD and DJF ET SD were placed near the top, whereas MAM CC SM, MAM EE SD, and MAM MR SM were around the bottom. Higher-ranked drivers generally had a strong positive or negative correlation with HF; however, some exceptions were observed. For example, despite having one of the lowest R-values, MAM EP SST was the fifth most important variable. In addition, MAM EE SD had an R-value of 0.43 and MAM WE SD had an R-value of −0.44; however, MAM EE SD was the 14th most important variable, and MAM WE SD was the 7th most important variable.

After analyzing the overall impact ranking of teleconnection drivers based on their mean absolute SHAP values, we further investigated the directionality of their impact on HF predictions. The scatter plots in Fig. 4 illustrate the relationship between the anomalies in the input features and their SHAP values. Many scatter plots have a discrete relationship rather than a linear one because the characteristics of LGBMs that construct decision trees to make predictions can cause the SHAP value to remain the same over a range of feature values. The variables in the first row negatively contributed to the model output, indicating that their positive anomalies corresponded to a decrease in the predicted HF and vice versa. The variables in the second row had positive contributions, indicating that their positive anomalies corresponded to an increase in the predicted HF. For example, in the LGBM model, the below-normal SD in the MAM GD SD and MAM WE SD regions, the below-normal SIC in the MAM BK SIC, and the below-normal SST in the MAM EP SST resulted in higher predictions of HF. In contrast, the above-normal SD in the DJF TM SD and DJF ET SD regions, the above-normal SST in the MAM WP SST region, and the above-normal SM in the MAM NC SM region resulted in higher predictions of HF. The remaining eight drivers also exhibited negative or positive relationships; however, their magnitudes were not as large as those of the top eight variables, as shown in Supplementary Fig. 3.

We divided the study period into quarters and colored them differently to visualize how the influence of these drivers evolved. A notable shift in the behavior of the MAM GD SD over time was evident (Fig. 4a). It showed high fluctuations in past data, but its variability was reduced and it exhibited predominantly negative anomalies in recent data. This trend toward more negative anomalies and positive SHAP values in recent years, as highlighted by the yellow dots, was similarly observed for other variables, such as MAM EP SST (Fig. 4c) and MAM WE SD (Fig. 4d). This implies that the influence



	DJF TM SD	DJF ET SD	DJF IP SST	DJF SP SST	DJF NA SST	DJF BK SIC	MAM CC SM	MAM NM SM	MAM NC SM	MAM WE SD	MAM EE SD	MAM NC SD	MAM GD SD	MAM WP SST	MAM EP SST	MAM BK SIC
DJF TM SD	610	10	0	0	10	30	0	50	0	90	0	90	0	110	10	20
DJF ET SD	10	440	10	0	20	0	20	10	20	0	0	160	40	130	20	0
DJF IP SST	0	10	130	10	0	30	0	30	40	10	0	70	30	0	30	0
DJF SP SST	0	0	10	70	0	30	50	0	0	0	0	10	50	50	0	0
DJF NA SST	10	20	0	0	40	0	0	0	0	0	0	0	0	0	0	0
DJF BK SIC	30	0	30	30	0	150	30	0	70	0	0	0	0	0	0	0
MAM CC SM	0	20	0	50	0	30	60	0	10	30	0	40	20	0	10	10
MAM NM SM	50	10	30	0	0	0	0	60	30	40	60	0	0	0	0	70
MAM NC SM	0	20	40	0	0	70	10	30	190	70	30	10	10	0	40	10
MAM WE SD	90	0	10	0	0	0	30	40	70	290	0	0	30	30	10	0
MAM EE SD	0	0	0	0	0	0	0	60	30	0	60	30	10	10	70	0
MAM NC SD	90	160	70	10	0	0	40	0	10	0	30	200	0	40	0	50
MAM GD SD	0	40	30	50	0	0	20	0	10	30	10	0	630	30	20	0
MAM WP SST	110	130	0	50	0	0	0	0	0	30	10	40	30	310	0	0
MAM EP SST	10	20	30	0	0	0	10	0	40	10	70	0	20	0	400	80
MAM BK SIC	20	0	0	0	0	0	10	70	10	0	0	50	0	0	80	400

Fig. 5 | Heatmap of SHAP interaction values. Each entry represents the mean absolute SHapley Additive exPlanation (SHAP) interaction value of a pair of features. Diagonal entries represent the independent impact on its own; off-diagonal entries represent the interaction between a pair of features.

of the variable on model predictions changed significantly, suggesting a greater and more stable influence on increasing HF predictions in recent years.

SHAP interaction values can show the magnitude of the pairwise contribution between input features on the model output, as shown in Fig. 5. Although most pairs exhibited zero or small values, indicating the predominance of individual contributions, specific interactions were prominent. First, the MAM variables tended to have higher SHAP interaction values, implying that within-season interactions among MAM drivers are more influential than those among DJF drivers. In addition, SHAP interaction values were observed between the MAM and DJF drivers, mainly driven by DJF TM SD and DJF ET SD. Such interaction values support the inter-seasonal interaction between snow conditions over the TP and various MAM conditions, and synergistically impact Korean HFs. For example, DJF TM SD strongly interacted with MAM WP SST, implying that their interaction has a considerable impact on HF predictions. Their thermodynamic linkages and possible contributions to Korean summer climate variability have been addressed in many previous studies[34,35]. MAM GD SD had a substantial individual influence on HF prediction (Fig. 3) but with minimal interaction effects with the other drivers, except for a slight inter-seasonal interaction with DJF ET SD and DJF CP SST.

Impact of important SD drivers on actual summer climate in South Korea

To understand the connection between the best predictive drivers of HF determined through XAI and the summer climate system, we conducted a composite analysis. This involved comparing the average conditions during the positive and negative phases of two important SD drivers: MAM GD SD and DJF TM SD. Positive and negative phases were defined as years when each SD driver's anomaly value was above 0.5 standard deviations or below −0.5 standard deviations. Notably, MAM GD SD was linked to increased HF during its negative anomalies (Fig. 4a). Consequently, we examined the average summer climatic conditions during the negative phase of MAM GD SD. We found that these conditions during negative phase of MAM GD SD was closely aligned with those favorable for heatwaves in South Korea. Therefore, we created composite maps by subtracting the average conditions of the positive phase from the negative phase for MAM GD SD (Fig. 6), and the reverse for DJF TM SD based on Fig. 4e. (Fig. 7).

During the negative phase of MAM GD SD, the composite maps of geopotential heights at 850-hPa, 500-hPa, and 200-hPa (Z850, Z500, and Z200) showed positive values over South Korea (Fig. 6a–c). This indicates the formation of a strong ridge in the lower troposphere and extension to higher altitudes, implying a stable atmosphere. Such atmospheric stability is typically associated with hot and dry summers. The observed high 2-m temperature (T2M) and low precipitation in South Korea supported these conditions (Fig. 6e, f). Specifically, the strong ridge results in higher atmospheric pressure and subsidence, which suppress cloud formation and precipitation, leading to increased temperatures and reduced rainfall. Moreover, during this phase, similar stable atmospheric conditions were observed in Northern Europe and the North Pacific. This pattern suggests a potential connection between MAM GD SD and the occurrence of heatwaves in these regions, indicating that the influence of MAM GD SD might extend beyond South Korea.

In addition, the anomalous large-scale circulation patterns were observed during the negative phase of MAM GD SD compared to climatological mean patterns depicted with contours in composite maps. For example, the extension of the Tibetan High into northern East Asia, including South Korea, is evident in the Z200 composite map (Fig. 6c). Similarly, the Z500 composite map (Fig. 6b) shows a northward shift of the Western North Pacific Subtropical High (WNPSH) extending to South Korea. The composite map of the zonal wind at 200-hPa (U200) (Fig. 6d) showed strong easterlies dominating East Asia, coupled with strong westerlies across its northern extent of East Asia (Fig. 6d). The meridional tri-pole patterns in both U200 and Z200 and a northward displacement of the East Asian Subtropical Jet (EASJ) were observed. This tri-pole pattern was also observed in the U200 patterns of 1994 and 2018 when South Korea experienced extreme heatwaves[10], supporting the potential of MAM GD SD as an indicator of extreme heatwaves. The

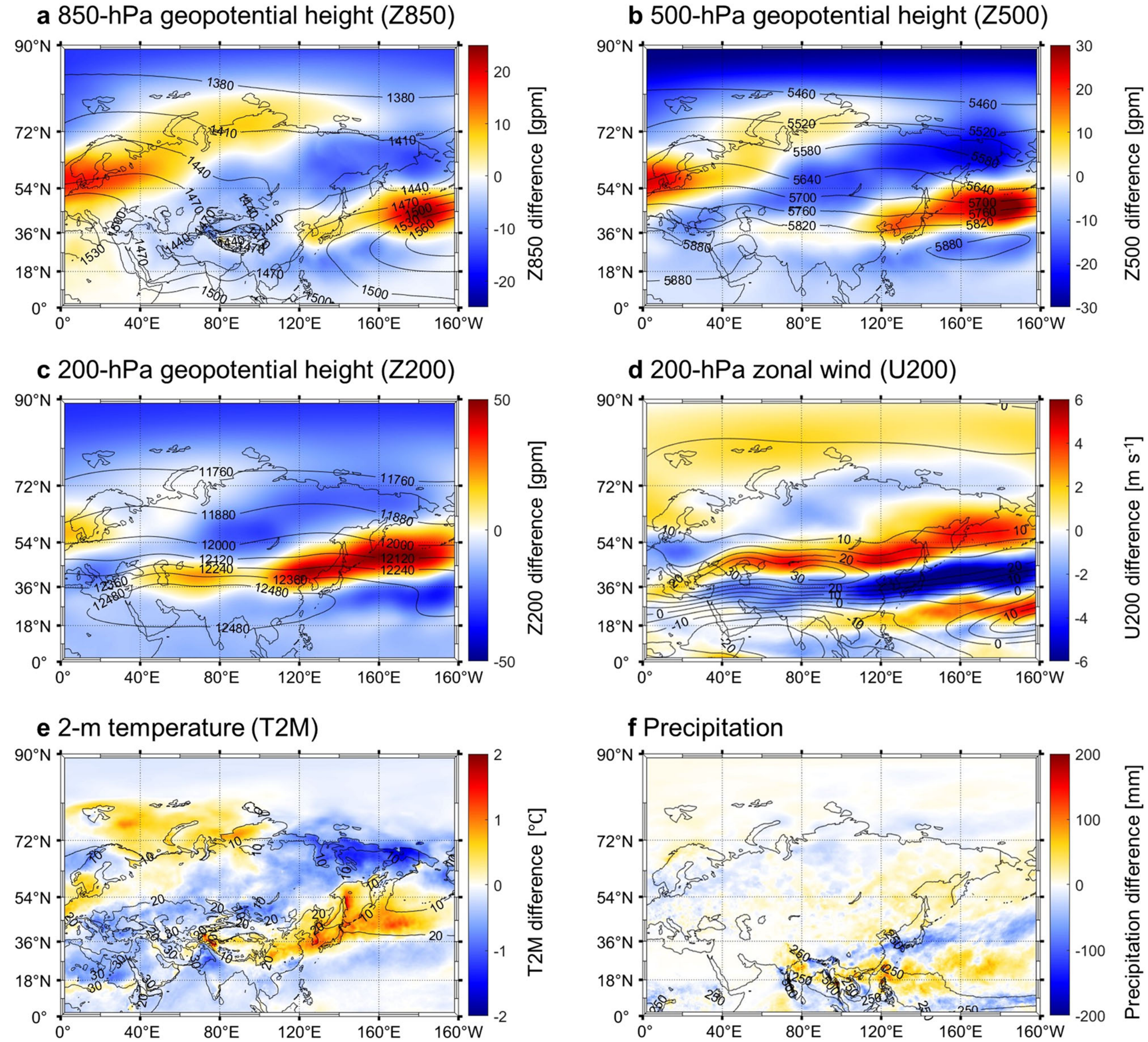


Fig. 6 | Composite maps of summer tropospheric conditions between the negative and positive phases of the MAM GD SD. The positive phase indicates when MAM GD SD's anomalies were higher than 0.5 standard deviation, and the negative phase indicates when its anomalies were lower than −0.5 standard deviation. The negative phase of MAM GD SD is associated with heatwaves over South Korea. To highlight the impact of negative phase, composite maps are created by subtracting the mean condition during the positive phase from that during the negative phase. Six tropospheric variables are analyzed: (**a**) 850-hPa geopotential height, (**b**) 500-hPa geopotential height, (**c**) 200-hPa geopotential height, (**d**) 200-hPa zonal wind, (**e**) 2-m temperature, and (**f**) precipitation. Warmer colors (e.g., red) highlight higher values during the negative phase, whereas cooler colors (e.g., blue) indicate lower values during the negative phase compared to those of the positive phase. Black lines denote the climatological mean conditions for the period 1979–2020.

northward displacement of the EASJ and the inflow of humid easterlies from the Pacific Ocean are vital synoptic conditions for the formation and northwestward extension of the WNPSH, a critical element in East Asian summer climate variablity[10,36].

During the positive phase of DJF TM SD, the composite maps of Z850, Z500, and Z200 showed positive values over South Korea, accompanied by high T2M and low precipitation (Fig. 7). Although these patterns are less intense than those of MAM GD SD over South Korea (Fig. 6), they still indicate stable atmospheric conditions conducive to hot and dry summers. Additionally, distinct zonal wave patterns characterized by alternating positive and negative values in geopotential heights were observed over the Eurasian continent (Fig. 7a–c). These patterns represent vertically coherent atmosphere structures that can significantly impact local weather conditions, as evidenced by the corresponding T2M (Fig. 7e). The di-pole pattern over the western North Pacific, likely extending from Z850 to Z500, could be related to the positioning of ridges and troughs in the lower troposphere, which influences the paths of moisture transport (Fig. 7f). In the upper troposphere, easterlies were evident over East Asia and westerlies to the north, indicating a northward shift of EASJ compared to its climatological mean location depicted by the contour lines (Fig. 7d). Furthermore, the northward extensions of both the WNPSH and the Tibet High were also observed (Fig. 7b, c).

Discussion

Through comprehensive SHAP analysis, we identified SD variability over the Gobi Desert (MAM GD SD) and Tianshan Mountains (DJF TM SD) as the most significant teleconnection drivers influencing HF predictions in South Korea. To verify the predictability of HF by these two drivers in the LGBM model, we conducted LOYOCV under three conditions: without MAM GD

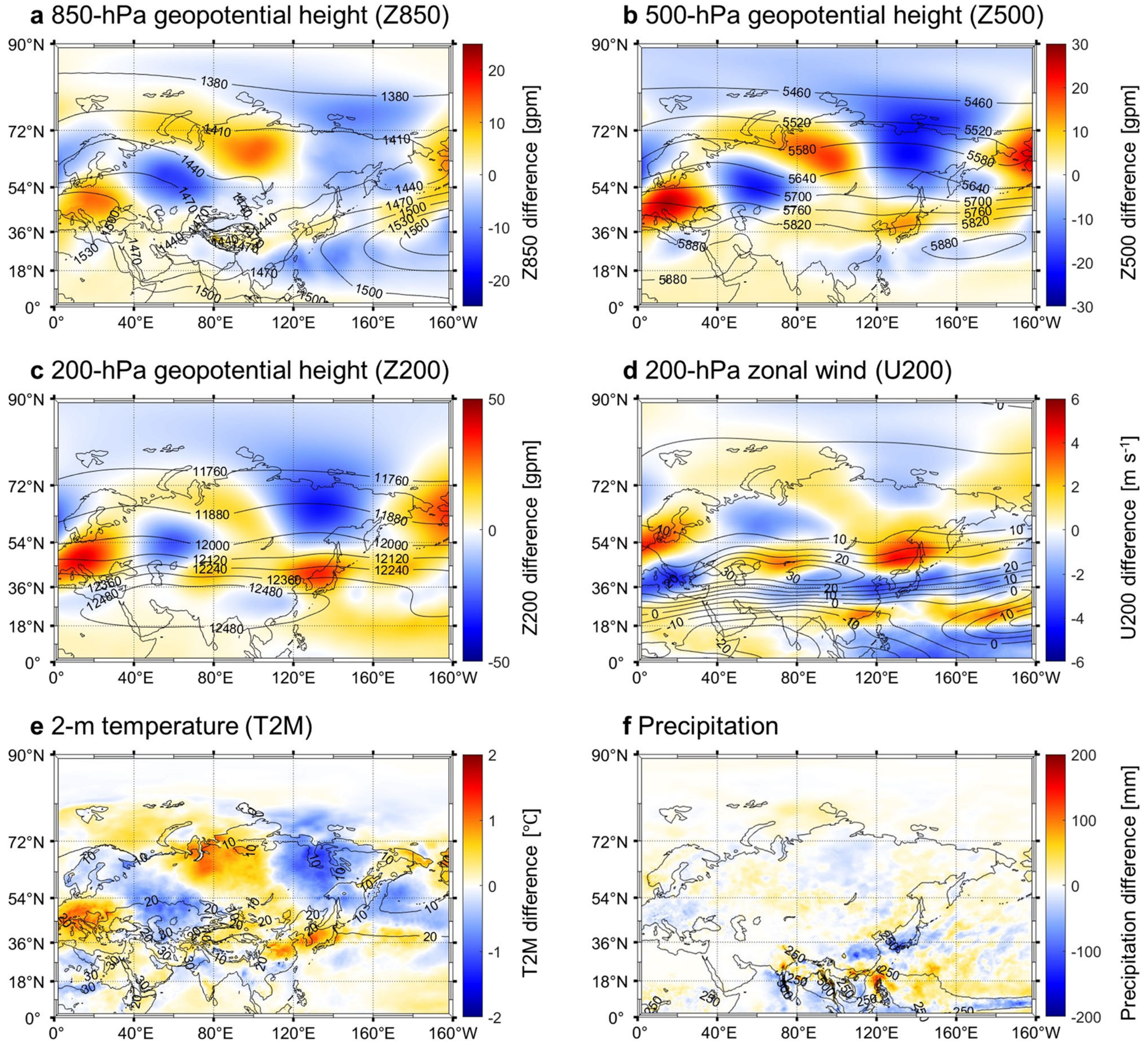


Fig. 7 | Composite maps of summer tropospheric conditions between the positive and negative phases of the DJF TM SD. The definition of positive and negative phase is same with MAM GD SD. The positive phase of DJF TM SD is related to the heatwaves over South Korea, so composite maps are drawn by subtracting the mean condition during the negative phase from that during the positive phase. Six tropospheric variables are analyzed similar to Fig. 6: (**a**) 850-hPa geopotential height, (**b**) 500-hPa geopotential height, (**c**) 200-hPa geopotential height, (**d**) 200-hPa zonal wind, (**e**) 2-m temperature, and (**f**) precipitation. Warmer colors (e.g., red) highlight higher values during the positive phase, whereas cooler colors (e.g., blue) indicate lower values during the positive phase compared to those of the negative phase. Black lines denote the climatological mean conditions for the period 1979–2020.

SD, without DJF TM SD, and without both MAM GD SD and DJF TM SD (Supplementary Table 2 and Supplementary Fig. 4). When MAM GD SD was removed, the R of the LGBM model decreased from 0.804 to 0.687, and an increase in the RMSE from 3.15 days to 3.96 days. When DJF TM SD was removed, the R decreased to 0.756 and RMSE increased to 3.48 days. Removing both drivers yielded R and RMSE values nearly identical to those observed with the removal of MAM GD SD alone. The time series of predicted HF under these three conditions show significant misalignment without MAM GD SD (blue line in Supplementary Fig. 4) and relatively smaller discrepancies without DJF TM SD (green line in Supplementary Fig. 4). Further analysis of the absolute differences between the original prediction errors and those following the removal of MAM GD SD and DJF TM SD confirmed that errors significantly increased when MAM GD SD was removed (blue bars in Supplementary Fig. 5), and also increased when DJF TM SD was removed (green bars in Supplementary Fig. 5), indicating its important impact on HF prediction. These findings emphasize the critical roles of both MAM GD SD and DJF TM SD in accurate HF prediction, though the impact of DJF TM SD impact is less pronounced than that of MAM GD SD.

Our composite analysis revealed a significant correlation between these SD drivers and the formation of robust vertically coherent atmospheric ridges, along with positive anomalies in T2M and negative anomalies in precipitation (Figs. 6 and 7). The average troposphere conditions when the two SD drivers highly contributed to the HF predictions were similar to those when extreme heatwaves occurred in South Korea[10], such as northward movement of the EASJ, moisture inflow from easterlies, and descending air with clear sky over South Korea (Figs. 6 and 7).

The region of MAM GD SD, the Gobi Desert, has experienced notable aridification, driven by a positive feedback loop between SM depletion and temperature increase[37–39]. This environmental change correlates with regional increases in surface solar radiation, air temperature, and Z200 (Supplementary Fig. 6). Notably, MAM GD SD showed a dramatic decrease in recent years, displaying a clear inverse trend with South Korean HF (blue line in Supplementary Fig. 7). These associations support the potential role of MAM GD SD in influencing large-scale circulation, as other studies have emphasized the influence of land–atmosphere interactions on upper atmospheric circulation[20,34,40]. In addition, the teleconnection pattern of MAM GD SD was similar to the SCAND pattern, characterized by a zonally oriented EASJ. We generated regression maps of Z200 for MAM GD SD and SCAND (Supplementary Fig. 8a, c) and calculated the Structural Similarity Index Measure (SSIM), which measures the similarity between images in range of 0 to 1 (see Method in detail). Two Z200 regression maps for MAM GD SD and SCAND showed high similarity, particularly in the mid-latitudes from Europe to East Asia, with an overall SSIM of 0.859 (Supplementary Fig. 8e). The SCAND pattern was consistent with the Z200 pattern regressed on recent heatwave patterns in East Asia (Fig. 5a and Fig. 9a of Choi et al.[9]), which is known to increase wave propagation from Eurasia to East Asia, leading to upper-level stationary waves and prolonged heatwaves in East Asia[9]. The similarities of Z200 patterns regressed on MAM GD SD, SCAND, and recent East Asian heatwaves suggest that MAM GD SD is intricately connected with large-scale atmospheric responses associated with heatwaves in South Korea.

The teleconnection pattern of DJF TM SD was similar to the double-trough type (DT-type) of summertime anomalous circulation over Eurasia analyzed by Yang and Wang[41] (Fig. 7a–c). To compare these patterns, the June–July–August (JJA) Z500 pattern representing the DT-type was constructed following the description in Yang and Wang[41] (Supplementary Fig. 8d). In addition, the JJA Z500 pattern for DJF TM SD to calculate SSIM between them is depicted in Supplementary Fig. 8b. The similarities, including the alternation of troughs and ridges over Eurasia, were evident with high SSIM values across Eurasia and an overall SSIM of 0.901 (Supplementary Fig. 8f). Notably, years corresponding to the DT-type of Yang and Wang[41] were characterized by a strongly extended WPSH, and were also the years when DJF TM SD had noticeably high positive SHAP values. This is consistent with DJF TM SD contributing substantially to the model's performance in predicting high HF and summer conditions favorable for heatwaves. Snowfall in the Tianshan Mountains, a vital component of DJF TM SD, is strongly influenced by moisture inflow from the North Atlantic with mid-latitude westerlies over Eurasia[42]. Although studies investigating the direct impact of DJF TM SD on Eurasian summer circulation are still sparse, the fact that these wind systems are major contributors to Eurasian summer circulation supports the hypothesis that the DJF TM SD is deeply connected to this existing global teleconnection system.

However, the precise mechanisms by which MAM GD SD and DJF TM SD affect heatwaves in South Korea remain to be fully elucidated. This study provides important insights into the complex interactions between local land surface changes and their impact on heatwaves in South Korea through large-scale atmospheric patterns. We believe that this study lays a foundation for future research to delve deeper into the underlying mechanisms of land surface variability over the Gobi Desert and Tianshan Mountains, and particularly how they contribute to heatwave patterns in South Korea. A further in-depth analysis of their mechanisms could significantly improve prediction abilities for extreme weather events, including heatwaves.

Some limitations of the study are as follows: One of the primary limitations is the small number of training samples. The target variable, annual HF, inherently restricts the sample size. Generally, machine learning models rely heavily on sample size and may not generalize well beyond the trained data patterns. To assess the impact of sample size on our LGBM model's performance, we conducted a bootstrapping analysis, training and predicting 100 times with 70–90% random sampling from the dataset (Supplementary Fig. 9). Fortunately, the average prediction across the 100 bootstrap samples remained stable, even as the sample size decreased. This suggests that the model's predictive behavior is consistent, although the variance in predictions increases as sample size decreases, leading to increased uncertainty and decreased precision. Another limitation is the sensitivity of the LGBM model to hyperparameters. The performance of LGBM can vary significantly based on the parameter settings, requiring extensive parameter tuning to identify the optimal combination. In this study, we addressed this issue by carefully optimizing the hyperparameters through grid search and cross-validation, ensuring optimal predictive performance (Supplementary Fig. 2). Despite these efforts, this process is computationally intensive and time-consuming, limiting our ability to test all parameters. Consequently, inherent uncertainty remains in determining the optimal parameters, potentially affecting the robustness and reliability of the model's predictions

Additionally, we developed a predictor screening strategy by clustering statistically significant points related to HF for each climate component (SST, SIC, SD, and SM) into teleconnection drivers using the density-based spatial clustering of applications with noise (DBSCAN) algorithm (Fig. 1). Based on a comprehensive review of previous studies, we defined the potential regions and seasons for each climate component (Supplementary Table 3). This clustering method effectively groups densely located, highly correlated points into arbitrary shapes. However, it may not detect newly emerging drivers that could significantly influence South Korean heatwaves. Future research should consider alternative approaches for identifying new drivers, such as globally analyzing land and ocean variables to uncover additional significant predictors.

Methods

Land and ocean variables

We used global monthly SST and SIC data at a resolution of 1° × 1° from the Hadley Center Sea Ice and Sea Surface Temperature[43]. The amount of water in the soil ("volumetric soil water layer 1") and the extent and amount of snow over land ("snow depth") from the European Center for Medium-Range Weather Forecasts Reanalysis (ERA5) monthly averaged data[44], which provided a fine spatial resolution of 0.25° × 0.25°, were used.

In situ temperature data and HF

The daily maximum temperatures measured at 103 in situ stations were obtained from the Korean Meteorological Administration's (KMA) Automated Synoptic Observing System. As a dependent variable, the total number of days that a heatwave occurred in July and August was calculated for the period 1960–2022 using these in situ temperature data and referred to as the HF. We chose 33 °C as the threshold of heatwave occurrence; this is the temperature at which the KMA issues a heatwave warning.

Selection of teleconnection drivers

We comprehensively reviewed previous studies and selected the potential regions and seasons for each of the land and ocean variables (SST, SIC, SD, and SM) to be explored (Supplementary Table 3 and Supplementary Fig. 10). Subsequently, we computed the nonlinear correlation between HF and each climate component using Spearman's rank correlation coefficient (R) and identified the statistically significant points ($|R| > 0.3$, $p < 0.05$). Adjacent significant points were clustered into teleconnection drivers using DBSCAN. DBSCAN is an algorithm that clusters large spatial datasets into arbitrary shapes based on the radius and minimum number of points in the neighborhood without a predetermined number of clusters or extensive domain expertise[45]. The overall flow of selection teleconnection drivers is illustrated in Supplementary Fig. 11. The anomaly values of the pixels within each cluster were averaged and detrended and then used as input variables for the HF prediction models.

LGBM algorithm

LGBM is an advanced gradient-boosting algorithm that utilizes decision trees[31]. It can grow complex trees by using a leaf-wise strategy that sequentially generates new trees from an existing tree, resulting in high precision[32]. However, it is prone to overfitting; therefore, LGBM provides several additional settings to prevent overfitting. One of the key components of LGBM is gradient-based one-side sampling (GOSS). GOSS strategically focuses on data instances with larger gradients, which indicate higher errors. LGBM can maintain robust and effective performance with smaller sample sizes than other ML algorithms because incorrectly predicted instances are typically more informative and contribute significantly to model improvements[33].

For the LGBM model, parameter tuning was performed using the 'GridSearchCV' function in the Python scikit-learn library. It evaluated all possible combinations of given hyperparameters and selected the optimal parameter set based on achieving the lowest mean squared error in cross-validation. The hyperparameters fine-tuned included the number of leaves ('num_leaves'), learning rate ('learning_rate'), number of estimators ('n_estimators'), bagging fraction ('bagging_fraction'), and two lambda parameters ('lanbda_l1' and 'lambda_l2'). The 'num_leaves' controls the complexity of the model, where more leaves potentially increasing accuracy of model but also increase the risk of overfitting. The 'learning_rate' determines the step size at each iteration, with smaller values providing robustness but requiring more boosting rounds. The 'n_estimators' specifies the number of boosting rounds or trees, where more estimators generally improve performance but increase training time. The 'bagging_fraction' introduces randomness by specifying the fraction of data used in each iteration, aiding in overfitting control. Finally, 'lambda_l1' and 'lambda_l2' are regularization terms that penalize model complexity to prevent overfitting.

HF prediction modeling and evaluation

Using the selected teleconnection drivers as input features, we predicted the annual HF in South Korea using the LGBM regression model. Model performance was assessed using the RMSE, R, and MSSS. R and RMSE are the most popular metrics[46–49], and MSSS is widely used particularly in atmospheric forecasting[50,51]. Two validation approaches were adopted: LOYOCV and hindcast validation. LOYOCV evaluated the model performance by constructing a total of 63 models from 1960 to 2022, each with one case as the test year and the remaining 62 cases as training samples. Hindcast evaluated the model performance by constructing 21 models from 2002 to 2022, each with one target year to predict, and then updating that sample as the training sample. A diagram of the two validation methods is shown in Supplementary Fig. 12. The hyperparameters of the LGBM model were optimized through grid search across both validation approaches. For comparative analysis, we utilized MLR and PNU CGCM, representing a traditional statistical model and an operational climate model, respectively.

MLR

MLR is a traditional data-driven method widely used to predict a dependent variable based on linear relationships with multiple independent variables. MLR models these relationships by fitting a linear equation to observed data. The general formula of an MLR model is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon$$

where Y is a dependent variable, $X_1, X_2, \cdots, X_n$ are independent variables, β_0 is a y-intercept, $\beta_1, \beta_2, \cdots, \beta_n$ are coefficients of corresponding independent variables, and ϵ is an error term. In this study, MLR served as a comparative model for LGBM model, using the same independent variables (teleconnection drivers) and dependent variable (annual HF) to ensure consistency. Additionally, MLR underwent the same validation approaches (LOYOCV and hindcast validation) and evaluation metrics (RMSE, R, and MSSS) as the LGBM model

PNU CGCM

PNU CGCM participates in the multimodel ensemble system for the long-term prediction of the Asia-Pacific Economic Cooperation Climate Center[52,53]. The model is composed of the National Centers for Environmental Prediction Community Climate Model version 3[54] for atmosphere GCM, Geophysical Fluid Dynamics Laboratory Modular Ocean Model version 3[55] for ocean GCM, and Los Alamos National Laboratory Elastic–Viscous–Plastic[52] for Sea-Ice Model. The seasonal prediction skills with atmosphere–ocean interaction in PNU CGCM are comparable to other state-of-the-art models, especially the notable high prediction performance in East Asia[52,53]. Additionally, it is conducive to examining extreme events owing to the production of seasonal predictions at an hourly resolution. In this study, HF predictions were produced for five different initial months (December to May), and five ensemble members for each initial month were generated using the time-lag method[56]. The production of initial states and ensemble members for the prediction is explained in detail by Kim and Ahn[52] and Lee et al.[53].

SHAP analysis

SHAP analysis has been widely adopted to interpret the impact of input variables on ML model outputs[57–59]. The SHAP value was determined by averaging the changes in the model output over all possible combinations of input variables[29]. Generally, the average absolute SHAP value across all instances describes the strength of the impact of the input variables on the model output, with a greater value indicating a greater influence of the variable. As SHAP values represent both the independent and interaction effects of the input variables, we also used SHAP interaction values, which can provide additional insights into complex pairwise contributions to the model output beyond individual contributions. The SHAP interaction value is calculated as the difference in the SHAP value for one feature in the presence or absence of another feature[60]. First, we plotted a time series of the SHAP values, and a teleconnection driver was drawn to determine the order of variable importance and temporal evolution. Next, we plotted scatterplots of the anomalies in the input features and their corresponding SHAP values to determine whether the impact of each input feature was positive or negative. Finally, we plotted a heat map of the SHAP interaction values to determine whether there were additional significant interactions that affected the model output.

Composite analysis

While SHAP analysis provided valuable interpretations of teleconnection drivers in ML models, it is important to understand these drivers within the context of actual summer climate dynamics. To determine if the MAM GD SD and DJF TM SD were indeed key drivers of Korean heatwaves, we performed a composite analysis, a widely used method for analyzing the influence of triggers on weather events[61,62]. The composite map was produced by subtracting the average conditions of the years when the anomalies of SD variables exceeded 0.5 standard deviations (positive phase) from those when the anomalies of SD variables were lower than −0.5 standard deviations (negative phase). This approach helps highlight differences in climatic conditions associated with the different phases of SD drivers. Typically, a threshold of 1 standard deviation has been used for such analyses, but in this study, a threshold of 0.5 standard deviation was selected to ensure a sufficient sample size for meaningful analysis.

We focused on several well-known summer climate dynamics affecting heatwaves in South Korea, including the WNPSH, the Tibet high (also known as the South Asian high), and the EASJ. Accordingly, we selected the Z500, Z200, and U200 for our analysis. Additionally, we examined conditions in the lower troposphere, including T2M, precipitation, and Z850. The Z200, Z850, U200, and T2M were obtained from the ERA5 reanalysis dataset, and precipitation data were from Multi-Source Weighted-Ensemble Precipitation (MSWEP). These variables were averaged over July–August, when heatwaves intensify in South Korea, and were prepared for the period 1979 to 2022, given the availability of data from MSWEP. In addition, the climatological mean values of each atmospheric circulation pattern are depicted with black contour lines (Figs. 6 and 7). Especially, we adjusted the level of contours to give additional information of Tibetan High (12,480 gpm) in 200-hPa geopotential height composite maps and the Western North Pacific Subtropical High (5880 gpm) in 500-hPa geopotential height composite maps.

SSIM

The SSIM is a widely used metric to compare the similarity between images[63,64]. SSIM values range from 0 to 1, and higher values indicate higher similarity. In this study, we calculated SSIM values between teleconnection patterns using the 'ssim' function in MATLAB. The pixel-wise SSIM was calculated using a sliding window of size 11 × 11 under the default setting, and the overall SSIM was calculated by averaging all pixel-wise SSIM values.

To calculate SSIM between teleconnection patterns related to DJF TM SD and DT-type circulation, the 'circulation index (CI)' defined by Yang and Wang[41] was calculated. The CI describes the anomalous characteristics of summertime circulation over the Eurasian mid-high latitude using JJA mean Z500 anomalies. They classified double-ridge type (DR-type) when normalized CI was greater than +1 standard deviation and DT-type when the normalized CI was less than –1 standard deviation.

Data availability

All data used in this study are publicly available. Monthly SST and SIC data were downloaded from https://www.metoffice.gov.uk/hadobs/hadisst/. Monthly SD, SM, T2M, and solar radiation data were provided in 'ERA5 monthly averaged data on single levels from 1940 to present' (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels-monthly-means?tab=form). Monthly Z850, Z200, and U200 data were provided in 'ERA5 monthly averaged data on pressure levels from 1940 to present'(https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels-monthly-means?tab=overview). Precipitation data can be downloaded from the MSWEP website via a Google Drive link after approval (https://www.gloh2o.org/mswep/). In situ daily temperatures measured by KMA's ASOS in South Korea can be downloaded from the KMA Weather Data Service (https://data.kma.go.kr/data/grnd/selectAsosRltmList.do?pgmNo=36). The SCAND index was downloaded from the homepage sharing teleconnection indices operated by the National Oceanic and Atmospheric Administration Climate Prediction Center (https://www.cpc.ncep.noaa.gov/data/teledoc/scand.shtml).

Code availability

The MATLAB and python codes used in this study are available from the first author or corresponding author upon reasonable request. Maps shown in this paper and Supplementary Information were drawn using M_Map[65], a mapping package for MATLAB, available online at www.eoas.ubc.ca/~rich/map.html.

Received: 2 February 2024; Accepted: 23 July 2024;

References

1. Zachariah, M. et al. Extreme heat in North America, Europe and China in July 2023 made much more likely by climate change. *Imperial College London.* https://spiral.imperial.ac.uk/handle/10044/1/105549 (2023).
2. Åström, D. O., Bertil, F. & Joacim, R. Heat wave impact on morbidity and mortality in the elderly population: a review of recent studies. *Maturitas* **69**, 99–105 (2011).
3. Brás, T. A., Seixas, J., Carvalhais, N. & Jägermeyr, J. Severity of drought and heatwave crop losses tripled over the last five decades in Europe. *Environ. Res. Lett.* **16**, 065012 (2021).
4. Ballester, J. et al. Heat-related mortality in Europe during the summer of 2022. *Nat. Med.* **29**, 1857–1866 (2023).
5. Copernicus Climate Change Service. Copernicus: 2023 is the hottest year on record, with global temperatures close to the 1.5°C limit. https://climate.copernicus.eu/copernicus-2023-hottest-year-record (2024).
6. The Government of the Republic of Korea. Adaptation Communication. https://unfccc.int/sites/default/files/ACR/2023-03/The%20Republic%20of%20Koreas%20Adaptation%20Communication.pdf (2023).
7. Lee, W. S. & Lee, M. I. Interannual variability of heat waves in South Korea and their connection with large-scale atmospheric circulation patterns. *Int. J. Climatol.* **36**, 4815–4830 (2016).
8. Yeh, S.-W. et al. The record-breaking heat wave in 2016 over South Korea and its physical mechanism. *Monthly Weather Rev.* **146**, 1463–1474 (2018).
9. Choi, N., Lee, M.-I., Cha, D.-H., Lim, Y.-K. & Kim, K.-M. Decadal changes in the interannual variability of heat waves in East Asia caused by atmospheric teleconnection changes. *J. Clim.* **33**, 1505–1522 (2020).
10. Ha, K.-J. et al. What caused the extraordinarily hot 2018 summer in Korea? *J. Meteorol. Soc. Jpn. Ser. II* **98**, 153–167 (2020).
11. Zhou, Z. Q., Xie, S. P., Zheng, X. T., Liu, Q. Y. & Wang, H. Global warming-induced changes in El Nino teleconnections over the North Pacific and North America. *J. Clim.* **27**, 9050–9064 (2014).
12. Shi, L. M. et al. What caused the interdecadal shift in the El Nino-Southern Oscillation (ENSO) impact on dust mass concentration over northwestern South Asia? *Atmos. Chem. Phys.* **22**, 11255–11274 (2022).
13. Yim, S. Y., Jhun, J. G., Lu, R. Y. & Wang, B. Two distinct patterns of spring Eurasian snow cover anomaly and their impacts on the East Asian summer monsoon. *J. Geophys. Res.* **115**, D22113 (2010).
14. Wang, Z. Q., Yang, S., Lau, N. C. & Duan, A. M. Teleconnection between Summer NAO and East China rainfall variations: a bridge effect of the Tibetan plateau. *J. Clim.* **31**, 6433–6444 (2018).
15. You, Q. L. et al. Review of snow cover variation over the Tibetan Plateau and its influence on the broad climate system. *Earth Sci. Rev.* **201**, 103043 (2020).
16. Seneviratne, S. I. et al. Investigating soil moisture-climate interactions in a changing climate: a review. *Earth Sci. Rev.* **99**, 125–161 (2010).
17. Whan, K. et al. Impact of soil moisture on extreme maximum temperatures in Europe. *Weather Clim. Extremes* **9**, 57–67 (2015).
18. Felsche, E., Böhnisch, A. & Ludwig, R. Inter-seasonal connection of typical European heatwave patterns to soil moisture. *npj Clim. Atmos. Sci.* **6**, 1 (2023).
19. Kug, J.-S. et al. Statistical relationship between two types of El Niño events and climate variation over the Korean Peninsula. *Asia Pac. J. Atmos. Sci.* **46**, 467–474 (2010).
20. Choi, Y.-W. & Ahn, J.-B. Possible mechanisms for the coupling between late spring sea surface temperature anomalies over tropical Atlantic and East Asian summer monsoon. *Clim. Dyn.* **53**, 6995–7009 (2019).
21. Liu, Q., Zhou, T., Mao, H. & Fu, C. Decadal variations in the relationship between the western Pacific subtropical high and summer heat waves in East China. *J. Clim.* **32**, 1627–1640 (2019).
22. Guo, D. et al. Mechanism on how the spring Arctic sea ice impacts the East Asian summer monsoon. *Theor. Appl. Climatol.* **115**, 107–119 (2014).
23. Streffing, J., Semmler, T., Zampieri, L. & Jung, T. Response of Northern Hemisphere weather and climate to Arctic Sea Ice decline: resolution Independence in polar amplification model intercomparison project (PAMIP) simulations. *J. Clim.* **34**, 8445–8457 (2021).
24. Hoskins, B. The potential for skill across the range of the seamless weather-climate prediction problem: a stimulus for our science. *Q. J. R. Meteorol. Soc.* **139**, 573–584 (2013).
25. Chang, N.-B., Yang, Y. J., Imen, S. & Mullon, L. Multi-scale quantitative precipitation forecasting using nonlinear and nonstationary teleconnection signals and artificial neural network models. *J. Hydrol.* **548**, 305–321 (2017).
26. Lim, W.-I. & Seo, K.-H. Physical–statistical model for summer extreme temperature events over South Korea. *J. Clim.* **32**, 1725–1742 (2019).
27. Maity, R. et al. Contrasting features of hydroclimatic teleconnections and the predictability of seasonal rainfall over east and west Japan. *Meteorol. Appl.* **27**, e1881 (2020).

28. Galelli, S. et al. An evaluation framework for input variable selection algorithms for environmental data-driven models. *Environ. Model. Softw.* **62**, 33–51 (2014).
29. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inform. Process. Syst.* **30**, (2017).
30. Kumar, A., Yadav, J. & Mohan, R. Spatio-temporal change and variability of Barents-Kara sea ice, in the Arctic: Ocean and atmospheric implications. *Sci. Total Environ.* **753**, 142046 (2021).
31. Ke, G. et al. Lightgbm: a highly efficient gradient boosting decision tree. *Adv. Neural Inform. Process. Syst.* **30**, (2017).
32. Pham, T. D. et al. Improvement of mangrove soil carbon stocks estimation in North Vietnam using Sentinel-2 data and machine learning approach. *GIScience Remote Sens.* **58**, 68–87 (2021).
33. Guo, S. Y. et al. Interpretable machine-learning and big data mining to predict gas diffusivity in metal-organic frameworks. *Adv. Sci.* **10**, 2301461 (2023).
34. Zhao, P., Zhou, Z. J. & Liu, J. P. Variability of Tibetan spring snow and its associations with the hemispheric extratropical circulation and East Asian summer monsoon rainfall: An observational investigation. *J. Clim.* **20**, 3942–3955 (2007).
35. Nan, S. L., Zhao, P., Yang, S. & Chen, J. M. Springtime tropospheric temperature over the Tibetan Plateau and evolutions of the tropical Pacific SST. *J. Geophys. Res.* **114**, D10104 (2009).
36. Tang, S. K. et al. Linkages of unprecedented 2022 Yangtze river valley heatwaves to Pakistan flood and triple-dip La Nina. *npj Clim. Atmos. Sci.* **6**, 44 (2023).
37. Dashkhuu, D., Kim, J. P., Chun, J. A. & Lee, W.-S. Long-term trends in daily temperature extremes over Mongolia. *Weather Clim. Extremes* **8**, 26–33 (2015).
38. Chen, F., Yuan, Y. J., Zhang, T. W. & Linderholm, H. W. Annual precipitation variation for the southern edge of the Gobi Desert (China) inferred from tree rings: linkages to climatic warming of twentieth century. *Nat. Hazards* **81**, 939–955 (2016).
39. Han, J., Dai, H. & Gu, Z. L. Sandstorms and desertification in Mongolia, an example of future climate events: a review. *Environ. Chem. Lett.* **19**, 4063–4073 (2021).
40. Matsumura, S. & Yamazaki, K. Eurasian Subarctic summer climate in response to anomalous snow cover. *J. Clim.* **25**, 1305–1317 (2012).
41. Yang, D. & Wang, L. J. The summertime circulation types over Eurasia and their connections with the North Atlantic oscillation modulated by North Atlantic SST. *Atmosphere* **13**, 2093 (2022).
42. Yang, T. et al. Snowfall climatology in the Tianshan Mountains based on 36 cold seasons of WRF dynamical downscaling simulation. *Atmos. Res.* **270**, 106057 (2022).
43. Rayner, N. et al. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.* **108**, 4407 (2003).
44. Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
45. Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining,* 226–231 (1996).
46. Lee, S. et al. A hybrid machine learning approach to investigate the changing urban thermal environment by dynamic land cover transformation: a case study of Suwon, republic of Korea. *Int. J. Appl. Earth Observ. Geoinf*. **122**, 103408 (2023).
47. Han, D. et al. Precipitation nowcasting using ground radar data and simpler yet better video prediction deep learning. *GIScience Remote Sens*. **60**, 2203363 (2023).
48. Kang, Y., Kim, Y., Im, J. & Lim, J. Machine learning-based atmospheric correction based on radiative transfer modeling using sentinel-2 msi data and its validation focusing on forest. *Korean J. Remote Sens.* **39**, 891–907 (2023).
49. Hao, X. et al. A robust gap-filling method for predicting missing observations in daily Black Marble nighttime light data. *GIScience Remote Sens*. **60**, 2282238 (2023).
50. Scaife, A. A. & Smith, D. A signal-to-noise paradox in climate science. *npj Clim. Atmos. Sci*. **1**, 28 (2018).
51. Choi, J. & Son, S.-W. Seasonal-to-decadal prediction of El niño–southern oscillation and pacific decadal oscillation. *npj Clim. Atmos. Sci*. **5**, 29 (2022).
52. Kim, H. J. & Ahn, J. B. Improvement in prediction of the Arctic oscillation with a realistic ocean initial condition in a CGCM. *J. Clim*. **28**, 8951–8967 (2015).
53. Lee, J., Lee, M. I. & Ahn, J. B. Importance of ocean initial conditions of late autumn on winter seasonal prediction skill in atmosphere-land-ocean-sea ice coupled forecast system. *Clim. Dyn*. **58**, 3427–3440 (2022).
54. Kiehl, J. T. et al. Description of the NCAR community climate model (CCM3). (No. NCAR/TN-420+STR). University Corporation for Atmospheric Research. https://doi.org/10.5065/D6FF3Q99 (1996).
55. Pacanowski, R. C. MOM 3.0 Manual. http://www.gfdl.gov/~smg/MOM/web/guide_parent/guide_parent.html (1999).
56. Brankovic, C., Palmer, T. N., Molteni, F., Tibaldi, S. & Cubasch, U. Extended-range predictions with Ecmwf Models - Time-Lagged Ensemble Forecasting. *Q. J. R. Meteorol. Soc*. **116**, 867–912 (1990).
57. Gao, J., Meng, Q., Zhang, L. & Hu, D. How does the ambient environment respond to the industrial heat island effects? An innovative and comprehensive methodological paradigm for quantifying the varied cooling effects of different landscapes. *GIScience Remote Sens.* **59**, 1643–1659 (2022).
58. Kang, Y., Kim, M., Kang, E., Cho, D. & Im, J. Improved retrievals of aerosol optical depth and fine mode fraction from GOCI geostationary satellite data using machine learning over East Asia. *ISPRS J. Photogramm. Remote Sens.* **183**, 253–268 (2022).
59. Lv, Z., Nunez, K., Brewer, E. & Runfola, D. Mapping the tidal marshes of coastal Virginia: a hierarchical transfer learning approach. *GIScience Remote Sens.* **61**, 2287291 (2024).
60. Lundberg, S. M., Erion, G. G. & Lee, S.-I. Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888* (2018).
61. Zhang, R., Sun, C., Zhu, J., Zhang, R. & Li, W. Increased European heat waves in recent decades in response to shrinking Arctic sea ice and Eurasian snow cover. *npj Clim. Atmos. Sci.* **3**, 7 (2020).
62. Jiang, X., Holbrook, N. J., Marshall, A. G. & Love, P. T. Quasi-Biennial oscillation influence on Australian summer rainfall. *npj Clim. Atmos. Sci.* **7**, 19 (2024).
63. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. image Process.* **13**, 600–612 (2004).
64. Wang, Z. et al. Virtual image-based cloud removal for Landsat images. *GIScience Remote Sens.* **60**, 2160411 (2023).
65. Pawlowicz, R. M_Map: a mapping package for MATLAB, version 1.4 m, [Computer software], available online at www.eoas.ubc.ca/~rich/map.html (2020).

Acknowledgements

This research was supported by the National Research Foundation of Korea (NRF) (NRF-2021R1A2C2008561), by the Korea Meteorological Administration Research and Development Program under Grant KMIPA 2017-7010, and by Korea Institute of Marine Science and Technology Promotion funded by the Ministry of Oceans and Fisheries (RS-2023-00256330).

Author contributions

Yeonsu Lee: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft. Dongjin Cho: Methodology, Investigation, Writing – review & editing. Jungho Im: Conceptualization, Methodology, Supervision, Investigation, Methodology, Funding

Acquisition, Writing – review & editing. Cheolhee Yoo and Joonlee Lee: Investigation, Writing – review & editing. Yoo-Geun Ham and Myong-In Lee: Writing – review & editing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41612-024-00722-1.

Correspondence and requests for materials should be addressed to Jungho Im.