Communications Earth & Environment https://doi.org/10.1038/s43247-026-03341-1

Article in Press

Precipitation and soil moisture coupling constrains subseasonal predictability of a prolonged extreme heatwave

Received: 18 July 2025

Accepted: 16 February 2026

Cite this article as: Lv, B., Wang, S., Chen, G. *et al.* Precipitation and soil moisture coupling constrains subseasonal predictability of a prolonged extreme heatwave. *Commun Earth Environ* (2026). https://doi.org/10.1038/s43247-026-03341-1

Bingjie Lv, Shuguang Wang, Gang Chen & Baoqiang Xiang

Precipitation and soil moisture coupling constrains subseasonal predictability of a prolonged extreme heatwave

Bingjie Lv[1,2,3], Shuguang Wang[1,2,3], Gang Chen[4], Baoqiang Xiang[5,6]

[1]State Key Laboratory of Severe Weather Meteorological Science and Technology, Nanjing University, Nanjing, China
[2]Key Laboratory of Mesoscale Severe Weather, Nanjing University, Nanjing, China
[3]School of Atmospheric Sciences, Nanjing University, Nanjing, China
[4]Dept of Atmospheric and Oceanic Sciences, University of California-Los Angeles, USA
[5]NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA
[6]Cooperative Programs for the Advancement of Earth System Science, University Corporation for Atmospheric Research, Boulder, Colorado

Corresponding author: Shuguang Wang (wangsg@outlook.com)

Revised
December 2025

Submitted to *Communications Earth & Environment* for publication

Abstract:

Demand for accurate forecasts of long-lasting heatwaves is increasing, yet the processes that fundamentally limit their predictability remain poorly understood. Here we study the predictability of the August 2022 Yangtze River Valley (YRV) heatwave using real-time subseasonal to seasonal (S2S) forecast ensembles, which systematically underestimate this event's intensity. To identify factors limiting predictability, we analyze processes across multiple scales, including large-scale circulation anomalies, local convection, and soil-moisture conditions. Analysis shows that regional precipitation effectively separates high- and low-skill forecasts of surface maximum temperature. Quantitative attribution analysis indicates that precipitation, rather than the local highs, explains most of the ensemble spread in surface temperature, suggesting precipitation–soil moisture coupling as the dominant limiting factor. Numerical experiments using a simple heatwave model demonstrate that increased precipitation reduces surface warming by up to ~4°C, whereas reduced precipitation amplifies warming by ~2°C. These findings underscore the importance of improving convective representation to enhance heatwave predictability.

1. Introduction

Global warming is driving an increase in the frequency and severity of extreme heatwaves worldwide[1,2], posing a serious threat to society. The Yangtze River Valley (YRV), a densely populated region with numerous megacities located in the East Asian monsoon zone[3], is especially vulnerable to heatwave risks[4]. In the summer of 2022 (July–August), the YRV experienced an exceptionally severe heatwave characterized by exceptional intensity, prolonged duration, and widespread impacts[5-7]. This event broke many records that had stood since 1961[7] (National Climate Center, *http://www.ncc-cma.net*).

The severe societal and environmental impacts of the 2022 YRV heatwave have motivated extensive investigations into its physical mechanisms [7-17]. Many studies have identified multiple contributing factors associated with remote large-scale conditions, including remote forcing associated with record-breaking heavy precipitation in Pakistan[8,11-13,15,18], which intensified the heatwave through the generation of a Rossby wave train, the westward extension of the Western Pacific Subtropical High (WNPSH)[7,15,16], and the eastward strengthening of the South Asia High (SAH)[7,14,16]. Concurrent heatwaves in Europe—linked to the Silk Road Pattern (SRP)[7,11-13,19,20]— may also have contributed.

In addition to remote influences, several studies have emphasized local drivers. These include the intensification of regional high-pressure systems and associated subsidence over the YRV[7], as well as abnormal easterly winds in the upper troposphere[8,11,12], which may be regarded as part of these circulation anomalies. Prolonged drought conditions further exacerbated the heatwave by reducing soil moisture, strengthening land–atmosphere coupling, and amplifying surface warming[10,21,22]. Together, these findings indicate that the 2022 YRV heatwave resulted from multiscale interactions among remote teleconnections, regional circulation anomalies, and local feedbacks.

Accurately prediction of the intensity and duration of extreme heatwave events, such as the 2022 YRV heatwave, is crucial for effective disaster prevention and mitigation. Sub-seasonal to seasonal (S2S) forecasting[23] offers considerable potential in this regard[24], providing predictions weeks to months in advance. However, state-of-the-art S2S models have the predictive skill for extreme heatwaves limited to about one week[25-33], and their accuracy varies substantially across individual events, reducing reliability and delay disaster responses[33].

For the 2022 YRV heatwave, numerical weather prediction models substantially underestimated its intensity on the synoptic timescale[6]. Given the multiple interacting drivers described above, it remains unclear which processes most strongly limited the predictability of this event and how individual physical mechanisms contributed to the forecast errors. The primary objective of this study is therefore to evaluate the S2S forecasts of the 2022 YRV heatwave and to identify key factors responsible for the prediction errors. By addressing these issues, we seek to provide

insights that may help improve S2S forecasts of extreme heatwave events.

2. Results

2.1 Local Characteristic of Heatwaves

Figures 1a and 1b show the average daily maximum 2 m air temperature (T2max) anomalies for July and August 2022, respectively. A pronounced positive anomaly is evident in the YRV region (27–33 °N, 100–120 °E), indicating a significant occurrence of exceptionally high temperatures during both months. The heat wave reached its peak in August, with the largest T2m anomalies exceeding 6°C in the Sichuan Basin. As shown in Figure S1, during the mature stage of the heatwave in August 2022 (August 10th to 23rd, Supplementary Figure 1), the average T2max anomaly over YRV reached 5.8 °C, with a cumulative probability of 99.9%, exceeding three standard deviations (σ), which is statistically extremely rare.

In addition to the extreme heat, an extreme drought occurred in the YRV region. Precipitation anomaly for August 2022 (Figure 1c) reached nearly -100 mm, underscoring the severity of drought. This evidence suggests that the YRV region experienced a typical compound extreme event during the summer of 2022, with exceptionally high temperatures and dry conditions, especially in August.

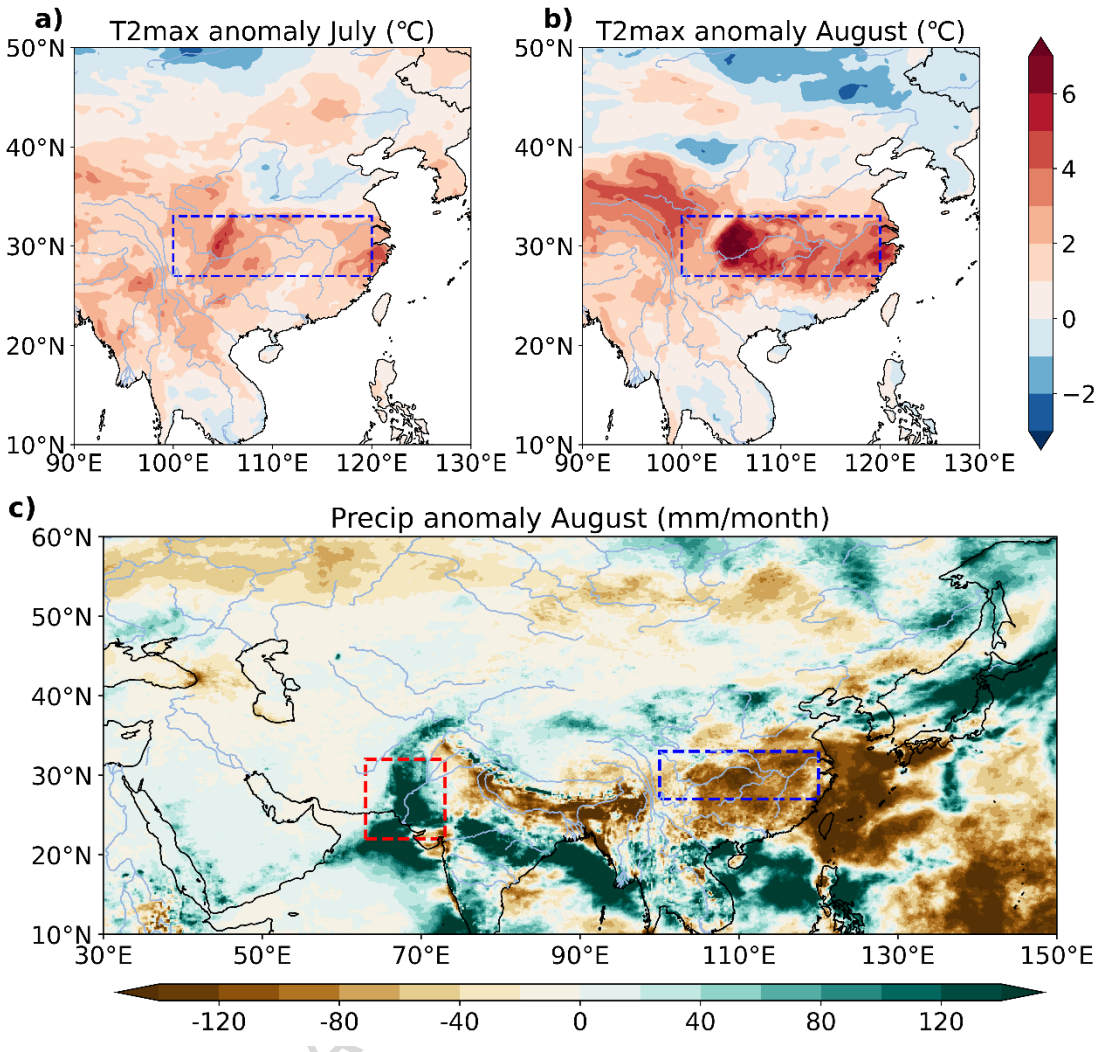

Figure 1. Temperature and precipitation anomalies during the 2022 YRV heatwave. Observed anomaly of monthly average T2max (shading, °C) in a) July and b) August 2022, based on ERA5 reanalysis. The blue dashed box indicates the location of the Yangtze River valley region (27–33 °N, 100–120 °E). c) Monthly precipitation anomaly (shading, mm month^{-1}) in August 2022, based on MSWEP V2. The blue dashed box indicates the YRV region, red dashed box, Pakistan (22–32 °N, 63–73 °E).

2.2 S2S Temperature forecast

The forecasted ensemble mean T2max for the YRV region (blue boxes in Figure 1) is shown as chiclet chart in Figure 2a and Figure 2b, which visualizes forecast performance across different lead times and initialization dates, including forecasts from both ECMWF and CMA models (see Methods). By comparing the chiclet chart with the orange line chart below, the forecast performance of S2S models for temperature can be preliminarily evaluated. During the heatwave, both models predicted significantly above-normal temperatures 2–5 days before the event, although the predicted amplitudes were slightly smaller (by 1–2 °C) than the observed

values. For forecast lead time of 6–9 days, both models predicted weaker T2max values compared to those with 2–5 days lead time. However, as the lead time exceeds two weeks, the models' forecast skills deteriorated significantly. This result aligns with previous studies, one of which highlighted the limited predictive skill of S2S models for the intensity of the 2021 western North American heatwave at a one-week lead time[29], and a latest study[33], which reported similar challenges for the 2023 North China heatwaves.

The ECMWF model is widely recognized as one of the most accurate forecasting systems available[28] and employs a large ensemble size[23]. Our subsequent analysis will primarily focus on the ECMWF model. The time series of T2max in Figure 2a and Figure 2b show that the heatwave was most intense during the two weeks in August (August 10th - 23rd) and the mature period of the heatwave will be used for the quantitative analysis. Among real-time forecasts with different initialization dates, we have analyzed two initialization dates, - 2022-07-28 and 2022-08-08 -, that produce T2max values closest to the ERA5 reanalysis. We then grouped all 51 ensemble members of these two initialization dates respectively into three categories - Warm, Medium, and Cold - based on the T2max values, among which the 'Warm' group is the most accurate (Figure 2c, 2d). Using the ERA5 reanalysis data as a reference, the results show that for the initialization date of 2022-07-28, even the Warm group underestimates the magnitude by more than 2°C. In contrast, for the 2022-08-08 initialization, the T2max values from August 10 to 16 were relatively accurate across all groups, while larger deviations occurred during the following week (August 17–23, Figure 2d).

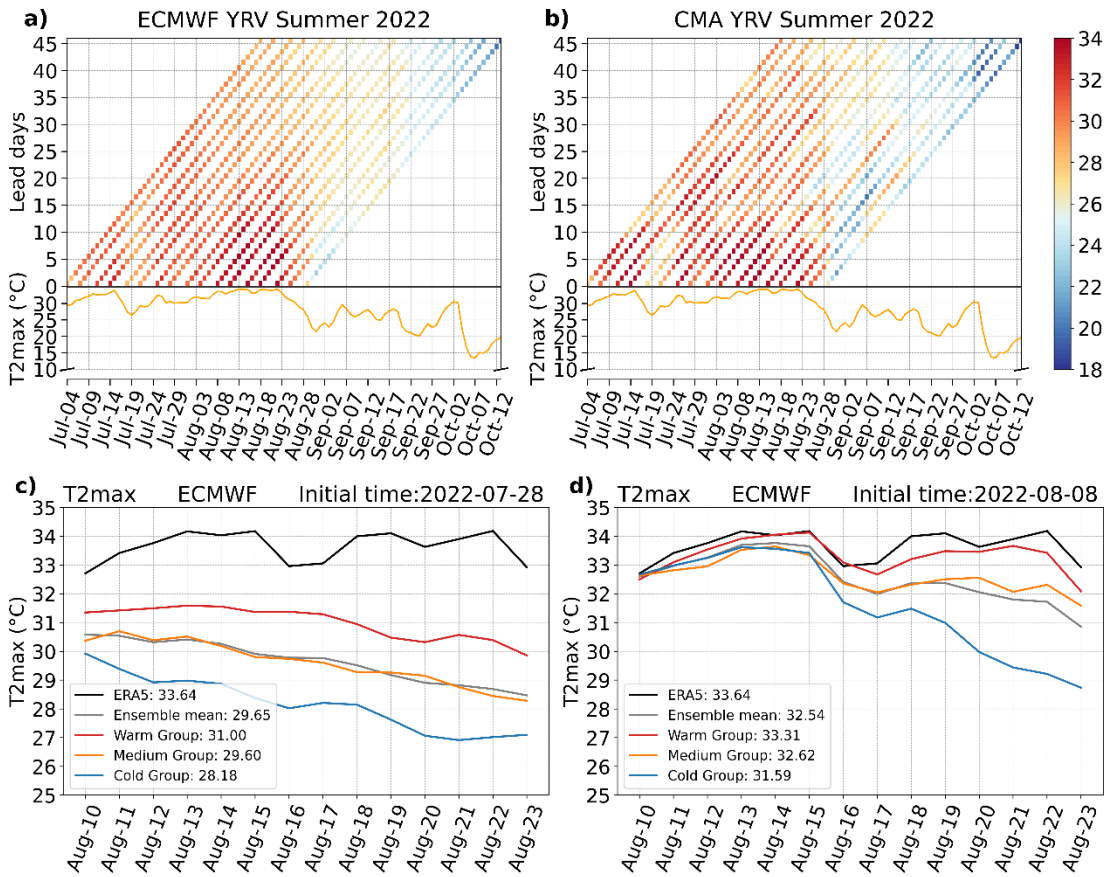

Figure 2. Subseasonal forecasts of T2max over the YRV during summer 2022 by the ECMWF and CMA models. Chiclet diagram of daily ensemble-mean T2max forecasts from ECMWF a) and CMA b), as a function of the verification date (horizontal axis) and lead time (vertical axis). The orange curves below the chiclet chart in (a) and (b) represents the ERA5 reanalysis daily T2max (°C). c) Ensemble mean (gray) of ECMWF T2max forecast initialized on 28 July 2022. Composite of ensemble members in three groups is categorized based on the average T2max over August 10th - 23rd: Warm (red line), Medium (orange line), and Cold (blue line). The ERA5 T2max is also shown in black. d) Same as c), but with the initialization date set to 8 August 2022.

2.3 Local and large-scale conditions

To understand the physical processes contributing to the occurrence of the studied event, here we present the large-scale and local environmental conditions. Figure 3 shows the local correlation between T2max and precipitation, as well as between precipitation and soil moisture, using reforecast data from the past 20 years together with the real-time forecast for 2022. The correlations associated with the 2022 event are notably stronger than those derived from the ensemble of reforecasts over the two-decade period. In addition, relative to the 20-year reference period (2002–2021), T2max reaches exceptionally high values, whereas precipitation and soil

moisture are unusually low in 2022 (ensemble means of 32.7 °C, 1.8 mm day^{-1}, and 252.2 kg m^{-3}, respectively). Correlations computed for individual years (Supplementary Figure 11) indicate that August 2022 ranks among the strongest T2max–precipitation relationships in the past two decades, highlighting the role of precipitation deficits in amplifying extreme temperatures. These correlations primarily reflect strong land–atmosphere coupling during the 2022 event (see details regarding energy partitioning in Supplementary Figure 2 and Discussion 1) and do not, by themselves, establish causality or confirm feedback mechanisms, particularly given the short analysis window of this case study.

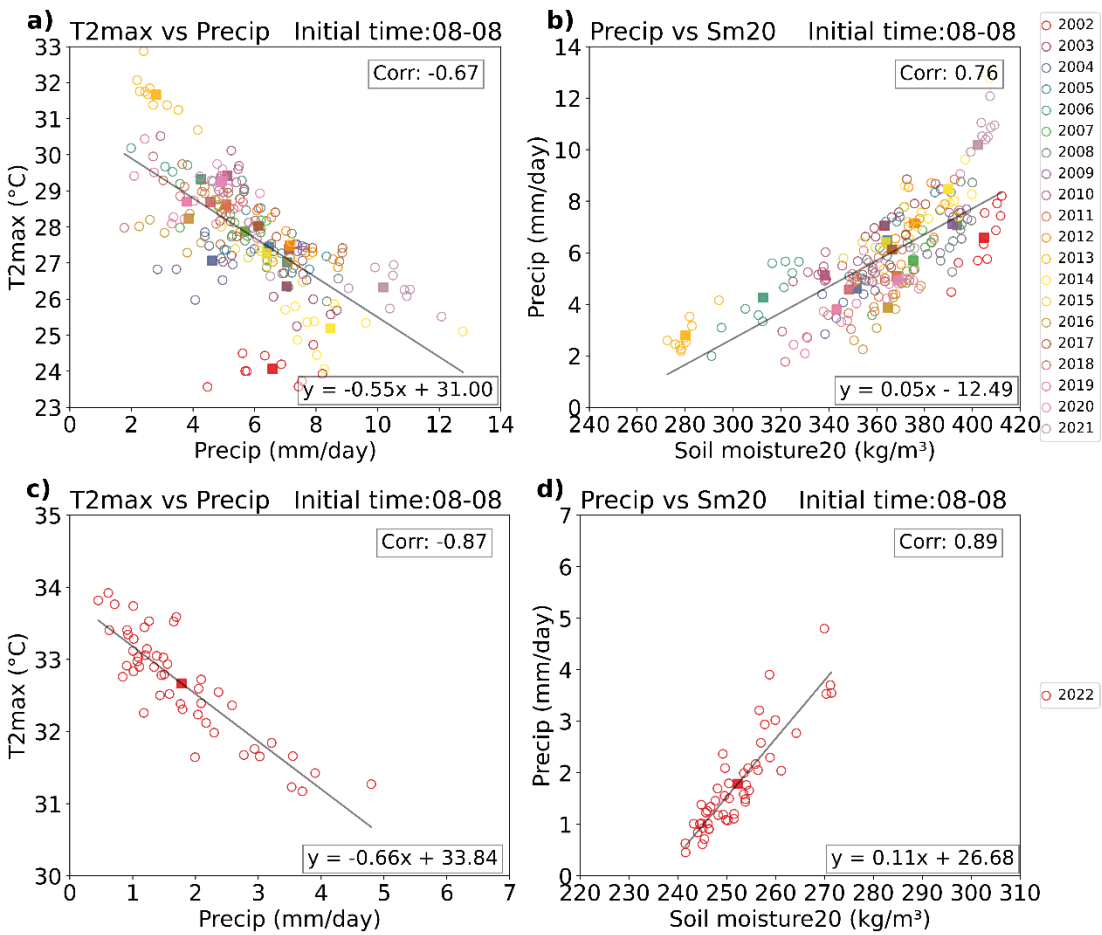

Figure 3. Relationships among T2max, precipitation, and soil moisture in Reforecast and Real-Time August Forecasts. Forecasts are initialized on 8 August, and evaluations are for the average over the period from August 10th - 23rd. a), b) Correlation for reforecast 20 years (2002–2021): a) T2max versus precipitation, and b) precipitation versus soil moisture. c), d) same as a), b), but for real-time forecast (2022). Hollow scattered dots represent individual ensemble members, while solid squares denote the ensemble mean. The correlation coefficients and regression equations are provided in the top right and bottom right of each panel, respectively. The scatterplots show area-averaged values over the YRV region for each year.

Large-scale circulation sets the stage for this long-lasting heatwave event. During its mature stage, 500-hPa geopotential height anomalies exhibit a Eurasian high–low–high pattern, with a blocking high over northern Europe, a low over northeast Asia, and a high centered over the YRV (Figure 4). The ECMWF real-time forecasts capture this structure reasonably well, with minor biases in the positions and intensities of the centers. At 850 hPa (Supplementary Figure 5), a similar pattern is evident, including a local high over the YRV exceeding 20 gpm in both forecasts and observations. Given the short forecast lead time, biases in large-scale circulation contribute little to forecast uncertainty. Instead, ensemble spread is primarily driven by local land–atmosphere coupling over the YRV (Section 2.4).

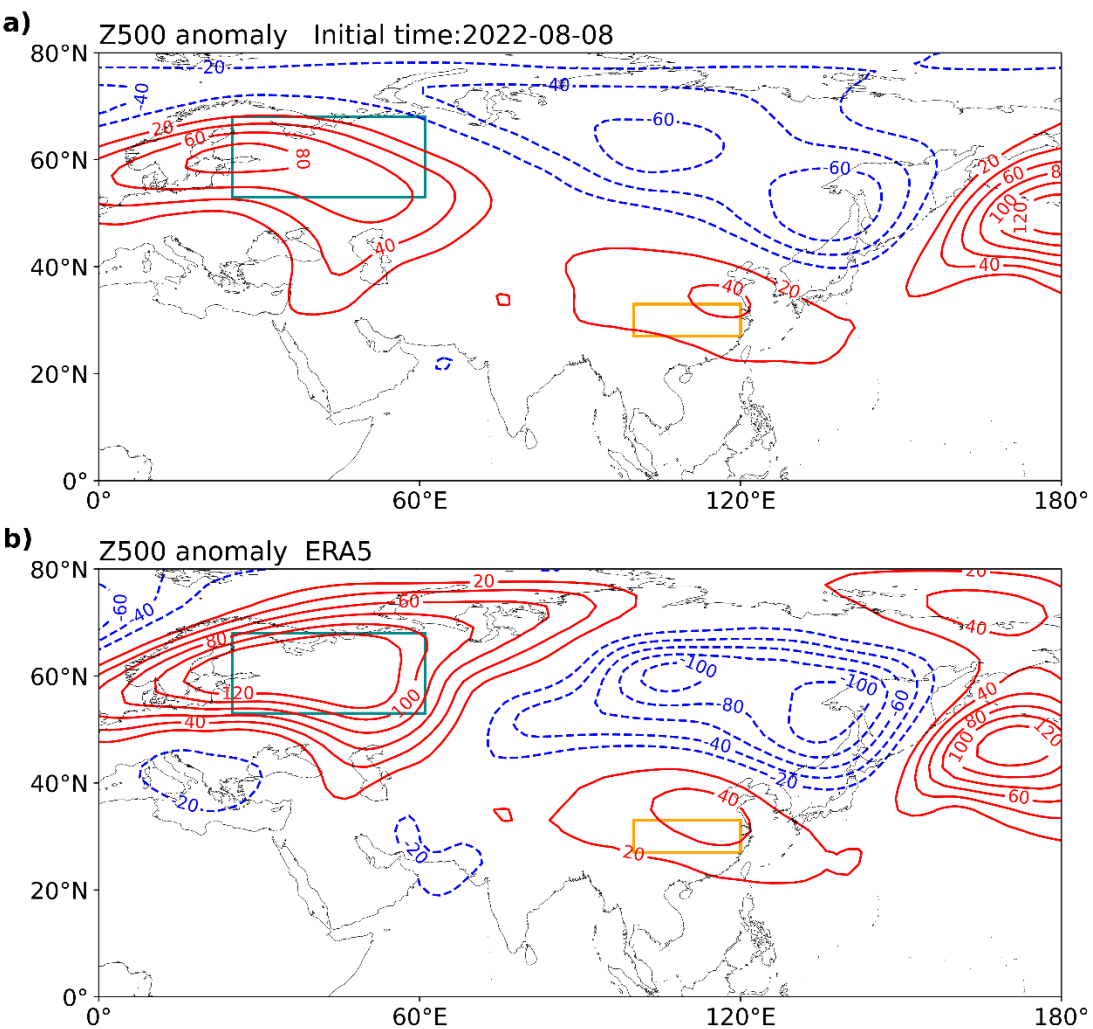

Figure 4. Large-scale circulation anomalies during the 10–23 August 2022 YRV heatwave. Contours of 500-hPa geopotential height anomaly (gpm; red, positive, blue, negative) from the ECMWF forecast initialized on 8 August 2022 a) and from ERA5 reanalysis b). The orange rectangle indicates the YRV region. The dark-cyan rectangle indicates the North Europe region (53–68 °N, 25–61 °E).

2.4 Forecast error attribution

We next explore the predictability of this event through a detailed error attribution analysis. Building on the preceding discussion, we explore six potentially important factors, from both local conditions (e.g., soil moisture, precipitation) and remote large-scale circulation (e.g., teleconnection). Specifically, these factors are soil moisture, YRV precipitation, Pakistan precipitation, YRV geopotential height anomaly, North Europe geopotential height anomaly and WNPSH. These factors were categorized into three groups according to the surface maximum temperature at 2 meters (T2max) —Warm, Medium, and Cold. Time evolution of these factors is shown in Figure .

Before August 14, the six factors exhibit minimal differentiation across groups. However, these variables evolve differently in the three groups of T2max. We first examine the two remote conditions. The strength of the high-pressure systems over Northern Europe (Figure 4) shows similar progression in all three T2max groups. As shown in Figure 5e, all groups substantially underestimated 500-hPa geopotential height over North Europe compared to ERA5 after Aug 14. For Pakistan precipitation, difference between the Warm and Medium groups remains small (1-2 mm day^{-1}), though both are distinctly separated from the Cold group by 2 mm day^{-1} throughout the period. For the WNPSH (Figure 5f), time evolution of the three groups is also similar, with a difference smaller than 8 gpm among the three groups. Furthermore, after August 20, WNPSH in all three groups shows considerable deviations from the ERA5 values. Overall, the absence of clear separation among the three YRV T2max groups for these remote variables suggests that they do not systematically covary with YRV T2max in the forecast ensembles. In other words, forecast errors in YRV T2max show little dependence on these remote drivers.

On the other hand, soil moisture, YRV precipitation, and YRV geopotential height anomaly after Aug 14 are well separated in the T2max groups. The Warm group exhibits lower soil moisture, less precipitation, and stronger local high. Among all factors, YRV precipitation displays the most pronounced group separation. A comparison of daily mean precipitation anomalies between the Warm and Cold groups is provided in Supplementary Figure 4, highlighting significant differences in YRV precipitation (lower in Warm) and moderate differences in Pakistan precipitation (higher in Warm).

To further examine the temporal evolution of the large-scale circulation, we composited Z500 anomalies every three days for the Warm and Cold groups (Supplementary Figure 6). While both groups reproduce the main circulation features reasonably well during the first forecast week, notable differences emerge in the second week: the YRV ridge and the northern European high decay more rapidly in the Cold group, and both features are weaker in amplitude compared to ERA5 (see Supplementary Figures 8 and 9 and Discussion 4). These differences in circulation evolution likely contribute to the differences in YRV precipitation and T2max between the

ensemble groups.

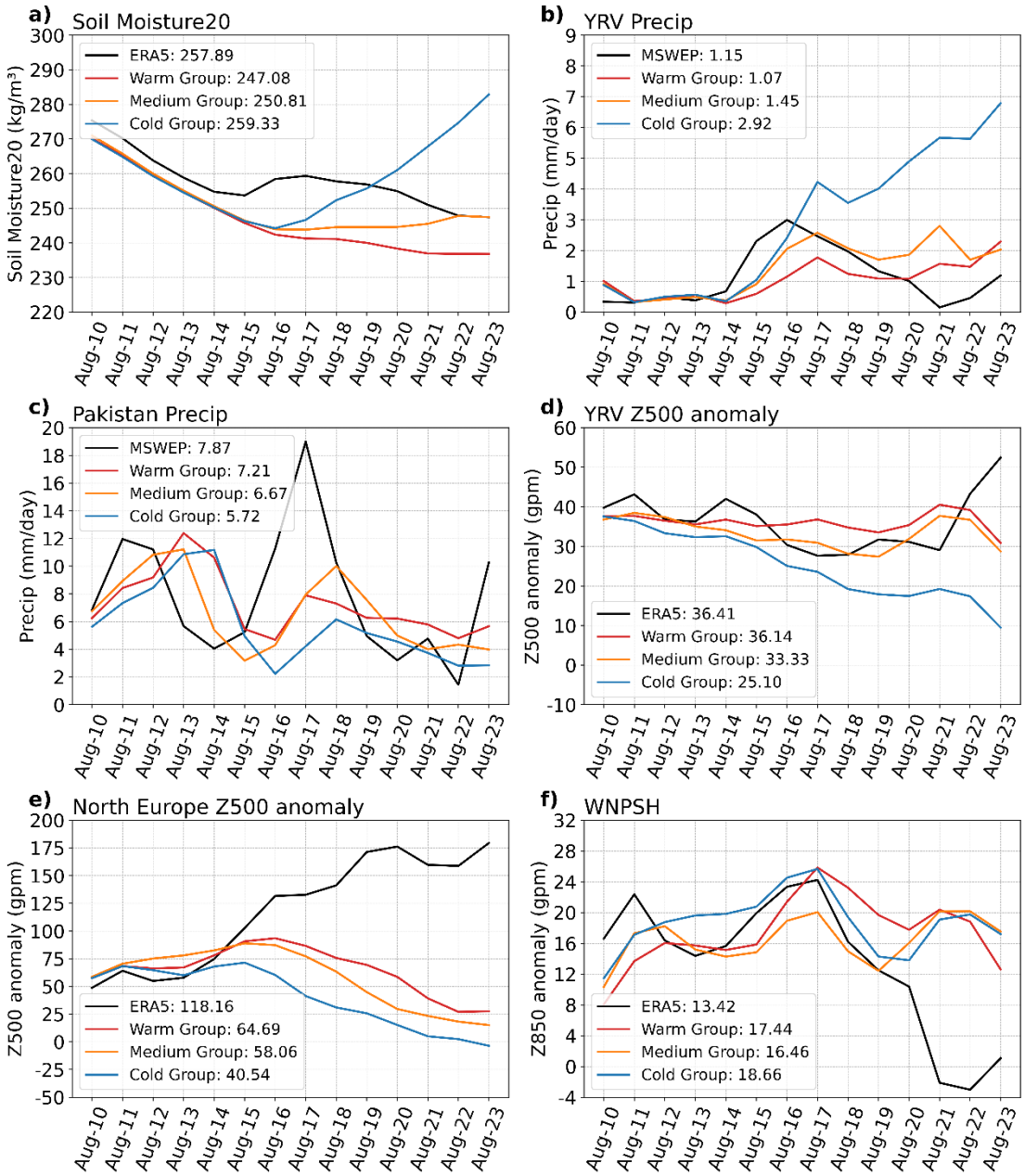

Figure 5. Comparison of key land–atmosphere factors during the YRV extreme heatwave across ensemble groups. Ensemble members are categorized into Warm (red line), Medium (orange line), and Cold (blue line) according to their 2-week predictions of T2max, while the reanalysis values are shown as a unified reference (black line; ERA5 for all variables except precipitation, which is from MSWEP V2). a) Soil moisture at 20 cm depth (Soil Moisture20, kgm^{-3}); b) Precipitation over YRV ($mm\ day^{-1}$); c) Precipitation over Pakistan ($mm\ day^{-1}$); d) Z500 anomaly over YRV (gpm); e) Z500 anomaly over North Europe (gpm); f) WNPSH intensity: Z850 anomaly over (15–25 °N, 115–150 °E) (gpm). Two-week average values are displayed in the top left or bottom left of each subplot.

The YRV precipitation and soil moisture are not independent: their correlation within the ensemble members during the two weeks reaches 0.89 (Figure d). The results indicate that local soil moisture and precipitation may form a feedback process that cannot be linearly separated. Accounting for this, we remove the precipitation-related component of soil moisture and use soil moisture residuals in the subsequent analysis.

To identify the primary sources of forecast bias and the spread among different ensemble members, we perform a MLR analysis using the six factors discussed above, with soil moisture replaced by its residuals (see Supplementary Method 2). Figure 6 compares the linear regression model predicted T2max and ECMWF forecasted values. The regression model yields an R^2 value of 0.8, suggesting that these six factors collectively explain a substantial portion of the ensemble spread of T2max.

To further assess the contributions of these local and remote factors, we systematically remove each factor and examine reduction in R^2 (Figure 6b). The largest reduction in R^2 (~0.20) occurs when YRV precipitation is removed, followed by soil moisture residuals (~0.04) and Pakistan precipitation (~0.02). In contrast, the contribution from the northern European high (~0.000) and WNPSH (~0.002) is negligible. Similarly, removing YRV geopotential height anomaly results in only a minor reduction in R^2 (0.004), indicating that the local high-pressure system does not significantly contribute to forecast ensemble spread.

Results from the attribution analysis are not unique to the initial date (August 8). The same analysis is repeated for different initial dates (from July 25 to August 4). As shown in Supplementary Table 2, reduction in R^2 after removing YRV local precipitation in these cases almost consistently exceeds 0.20, while the reduction in R^2 associated with the soil moisture residuals is smaller but remains greater than other factors. These findings highlight the pivotal role of atmospheric convection, particularly through its coupling with local soil moisture, in limiting predictability of this heat wave event.

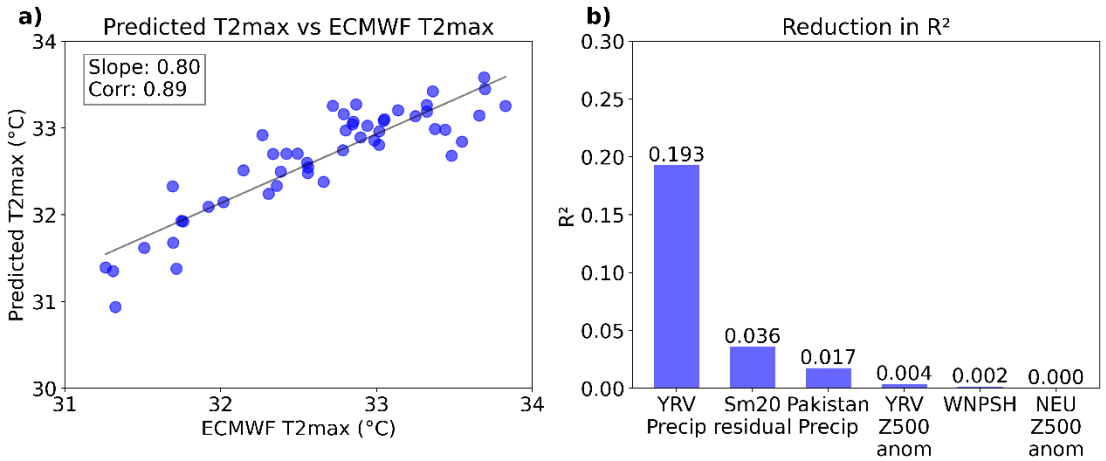

Figure 6. Prediction of T2max using a multiple linear regression model with six key factors. a) Scatter plot comparing T2max predictions from the linear regression model with forecasted T2max by ECMWF forecasts. Each dot indicates one forecast ensemble member. Values of slope and correlation coefficient are displayed in the top left of the subplot. b) Reduction in R^2 after each factor is removed. Soil moisture was replaced with its residual for the analysis.

2.5 Effect of Precipitation forcing in a heatwave model

The core heatwave dynamics can be conceptualized by the soil-moisture-precipitation coupling as represented in a simple heatwave model, the Z-model (see Methods), which focuses on co-dependance of surface temperature and soil moisture driven by precipitation $P(t)$ and radiative forcing $F_{SW}(t)$. We take $P(t)$ and $F_{SW}(t)$ from the S2S ensemble forecast and conduct ensemble integrations. Figure 7a shows the output of T2max versus P, both averaged during the two-week period (Aug 10 – 23). Despite the inherent nonlinearity in the soil-moisture feedback as represented in the Z-model, negative correlation between T2max and P is evident with the correlation coefficient *r* reaching –0.93 and the slope –2.03 °C (mm day^{-1})$^{-1}$. This slope is steeper than the slope derived from the ECMWF ensemble (–0.66 °C (mm day^{-1})$^{-1}$, see Figure c), indicating that T2max in the Z-model is perhaps too sensitive to changes in precipitation forcing.

The experiments are repeated with *P* increased to 200% or decreased it to 50% of its original values for all the ensembles. Figure 7b shows the mean T2max and its variability under different precipitation scenarios. The Z-model's ensemble-mean T2max is within 0.5 °C of the ECMWF forecast, suggesting overall reasonable accuracy. When precipitation is doubled, T2max decreases by approximately 4 °C, whereas halving precipitation leads to an increase of about 2 °C. These results offer additional evidence that precipitation controls the magnitude of temperature for this heatwave event, and also underscore the asymmetry of the temperature response to precipitation anomalies.

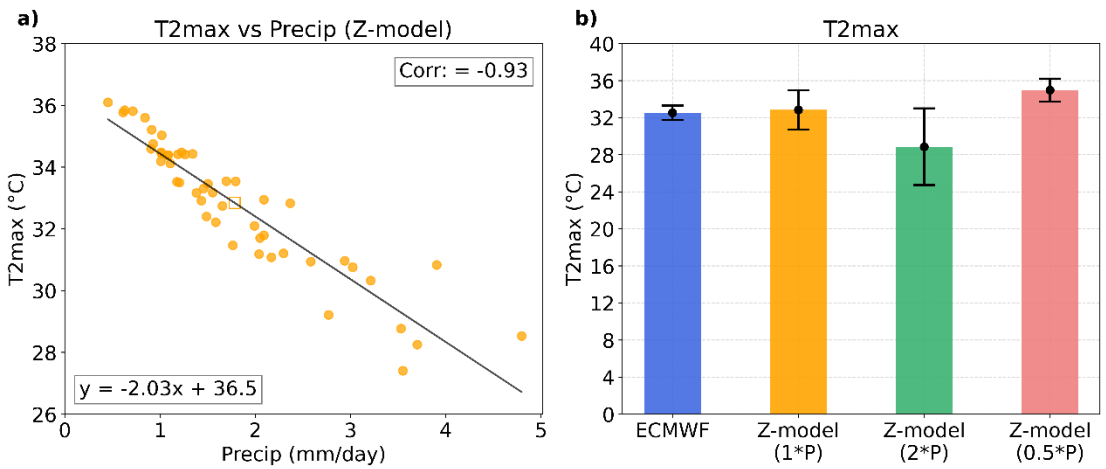

Figure 7. Relationship between T2max and precipitation in the Z-model and sensitivity of T2max to precipitation forcing. a) Scatter plot of daily maximum temperature (T2max) versus precipitation simulated by the Z-model. The hollow square indicates the ensemble mean. The fitted linear regression line and Pearson correlation coefficient (Corr.) are shown in the panel. b) Mean and standard deviation (1σ error bars) of T2max from ECMWF and Z-model simulations using different precipitation forcing: the default (1×P), doubled (2×P), and halved (0.5×P) precipitation.

3. Discussion

This study investigates the S2S predictability of the record-breaking heatwave that affected the Yangtze River Valley (YRV) region in July and August 2022, with a particular focus on its middle August peak. This extreme heatwave in the YRV was a compound event, marked by persistent high temperatures, severe drought, unusually broad spatial extent, and prolonged duration. During the peak phase of the heatwave, the YRV was dominated by a local high-pressure system and influenced by the WNPSH. Additionally, upstream influences, including the European high (associated with the Silk Road Pattern, SRP) and Pakistan precipitation, exerted teleconnection impacts.

The performance of S2S models in predicting this event is closely tied to their ability to represent large-scale circulation patterns. While these models can generally predict heatwave occurrences with a lead time of 2–5 days, they often underestimate the intensity of such events. As lead time increases, biases in maximum temperature (T2max) forecasts and other variables become more pronounced, with prediction skills deteriorating significantly beyond two weeks.

To understand predictability, the role of both the remote and local circulation conditions in the ECMWF real-time forecast ensembles are explored. Composite analysis highlights the critical factors limiting predictability of this extreme heatwave event, demonstrating that YRV

precipitation serves as a key factor in S2S model forecasts among the six factors considered. Quantitative attribution analysis further indicates that YRV precipitation-soil moisture coupling and soil moisture (represented by its residual from precipitation) are the primary contributors to the heatwave's intensity in S2S models, with YRV precipitation being the most influential. The negligible contribution of the YRV geopotential height anomaly to ensemble spread (R^2 reduction of only 0.004) may appear surprising. While the persistent ridge over the YRV was clearly a necessary precondition for the heatwave, its amplitude did not emerge as a source of forecast uncertainty. This result indicates that the ensemble spread was not primarily controlled by large-scale circulation features. Instead, precipitation and soil moisture processes dominated the forecast uncertainty of the heatwave event on the subseasonal time scale.

Sensitivity experiments using the simple heatwave model proposed by Zeppetello *et al.*[34] further highlight the critical role of local precipitation in heatwave prediction. These simple model experiments demonstrate that precipitation strongly influences both the predicted peak temperature and the associated uncertainty (see further sensitivity of Z-model in Supplementary Discussion 3). Specifically, variations in precipitation substantially affect the ensemble spread (Supplementary Figure 7) and the magnitude of forecasted extremes, underscoring the importance of accurately representing local precipitation processes in heatwave forecasting frameworks. The sensitivity to precipitation differs in the Z-model (-2.03 °C (mm day^{-1})$^{-1}$) and the ECMWF model (-0.66 °C (mm day^{-1})$^{-1}$). This might be caused by some important physical processes in the comprehensive GCM but absent in the simple model (more details in Supplementary Discussion 5). For example, cloud–radiation feedback, horizontal advection, and boundary-layer turbulent mixing. The absence of these physical processes likely leads to biases in the variability of temperature anomalies and exaggerate the sensitivity to precipitation forcing.

4. Conclusions

Our study aims to understanding the sources of limited predictability in the prolonged 2022 YRV heatwave by combining composite analysis, MLR, and sensitivity experiments. The composite analysis reveals that YRV precipitation is the most influential factor among the six examined, significantly affecting forecast skill in ECMWF S2S ensembles. The regression analysis further quantifies the contributions of YRV precipitation and soil moisture, showing that the precipitation–soil moisture coupling, particularly the component of soil moisture not directly explained by precipitation, plays a key role in modulating heatwave intensity. Sensitivity experiments using a simplified heatwave model reinforce these findings, illustrating that variations in precipitation lead to substantial changes in both predicted peak temperature and ensemble spread, thereby confirming the dominant role of local precipitation in heatwave predictability.

This study provides a new perspective for subseasonal predictability of prolonged heat wave. Our results show that precipitation-soil coupling critically limits predictability of the long-lasting record-breaking YRV heatwave in summer 2022. While most analyses focus on this single event, applying the same framework to three additional heatwave cases confirms the broader relevance of our findings (see Supplementary Figures 3 and 10 and Discussion 2). These findings underscore the need to improve the representation of convection and teleconnections in S2S models to enhance their predictive skill for similar extreme events in the future.

5. Methods

5.1 Reanalysis data

We use the ERA5 dataset[35] with a 0.25° × 0.25° horizontal grid spacing, including hourly 2-m temperature (T2m), surface sensible heat flux, surface latent heat flux, and daily soil moisture at 0–28 cm. These high-resolution fields are employed in the heatwave analyses to capture regional-scale variability in the Yangtze River Valley. To diagnose the large-scale atmospheric circulation background, we use daily ERA5 geopotential height fields at standard pressure levels, provided with a coarser horizontal resolution of 2° × 2°. Because these variables primarily represent large-scale circulation patterns, the coarser resolution is adequate. The MSWEP V2[36] precipitation data, with a horizontal resolution of 0.1° × 0.1°, are also utilized. MSWEP V2 combines gauge observations, satellite retrievals, and reanalysis fields, and has been shown to provide improved accuracy over East Asia compared with single-source products such as ERA5[37]. In particular, it better captures summer precipitation intensity and spatial structure, making it more suitable for diagnosing precipitation–temperature–soil moisture relationships in the YRV region.

5.2 S2S data

The S2S prediction data[23] from two operational centers, including the European Centre for Medium-Range Weather Forecasting (ECMWF) and the China Meteorological Administration (CMA), are used to evaluate the sub-seasonal predictability of the 2022 YRV heatwave. The analysis involves assessing real-time forecast and historical reforecast variables, including T2m, total precipitation, convective precipitation, large scale precipitation, surface sensible heat flux, surface latent heat flux, soil moisture at 0-20 cm (hereafter referred as the surface layer), 500-hPa and 850-hPa geopotential height, surface net solar radiation and surface net long wave radiation. Here, lead time refers to the forecast time horizon, defined as the time interval between the forecast initialization date and the target date of prediction. Anomalies of real-time forecast data are calibrated with the reforecast data by subtracting the reforecast climatology from the real-time forecast computed as a function of both lead time and initialization date. Climatology in models was defined using the dates from the reforecast data that matched the initialization frequency in the real-time data[33]. Detailed information regarding the ensemble member and

initialization frequency of each model is listed in Supplementary Table 1.

To assess the reliability of the ensemble predictions, we examine the consistency between ensemble spread and forecast error. Following previous studies[38,39], this relationship is quantified using a spread–skill ratio (SSR), defined as the ratio of ensemble spread to forecast error (Supplementary Discussion 2).

To elucidate the temporal evolution of the physical processes potentially driving the heatwave, composite analysis of the ECMWF ensemble forecast is adopted. We aggregate all the 51 ensemble members into three distinct groups based on the daily maximum temperature at 2 meters (T2max) averaged over the two-week period from August 10 to 23, which we define as the peak stage of the heatwave. This stage corresponds to the period of highest temperature anomalies, as shown in Figure. 2a and Figure. 2b, and is thus used for further quantitative analysis. Our definition follows the approach of earlier research[40], which identified August 14–25 as the mature phase of the same heatwave, but we slightly extend the period to August 10–23 to better capture the overall peak intensity of the event.

Specifically, the ensemble members were ranked according to the area-averaged T2max over this period and then divided into three groups: the warm group (top 17 members with the highest T2max), the cold group (bottom 17 members with the lowest T2max), and the medium group (the remaining 17 members). Importantly, this classification is determined only by T2max, while all other variables (including soil moisture, precipitation, and geopotential height anomalies) were grouped following the same T2max-based classification, ensuring consistency in the comparison of physical processes across groups.

The S2S real-time forecasts of T2max were bias-corrected to remove systematic model errors (see Supplementary Method 1 for details).

5.3 Multiple linear regression

We adopted a multiple linear regression (MLR) approach to analyze contributions of different factors to forecast errors. Based on previous studies, six key predictors are selected to construct a MLR model, including two remote factors (Europe geopotential high, Pakistan precipitation), and four local factors (YRV precipitation, YRV soil moisture, YRV geopotential high and WNPSH), as discussed in the introduction section.

5.4 The heatwave model

To further investigate the role of precipitation in heatwaves, we employ the simple heatwave model[34] (hereafter referred to as the Z-model) for the 2022 YRV heatwave. The governing equation for time variation of the coupled energy and moisture balance of the land surface can be written as:

$$C\frac{dT}{dt} = F_{SW}(t) - \alpha(T - T_D) - L_{vL}m[q_s(T) - \overline{q}] \quad (1)$$

$$\mu\frac{dm}{dt} = P(t) - v_L m[q_s(T) - \overline{q}] \quad (2)$$

In Equation (1), T denotes the land surface temperature as a function of time *t*, C is the surface heat capacity per unit area, $F_{SW}(t)$ represents the absorbed shortwave radiation. The parameter α is an effective heat exchange coefficient $(W\,m^{-2}K^{-1})$, accounting for the combined effects of sensible heat flux and longwave radiative cooling between the land surface and the atmosphere. T_D is the climatological local dewpoint temperature, used as a proxy for the thermal state of near-surface air. m is the fraction of soil pore space occupied by liquid water, varying between 0 (dry) and 1 (saturated), L_{vL} is the latent heat of vaporization, $q_s(T)$ denotes the saturation specific humidity at temperature T, determined by the Clausius–Clapeyron relationship, while $\overline{q}$ is the climatological near-surface specific humidity. Equation (2) governs the evolution of soil moisture, where μ is the water holding capacity of the soil layer (kg/m^2). $P(t)$ is the precipitation rate, and $v_L = \rho_a/r_s$ is the effective moisture transfer coefficient, with ρ_a representing the density of near-surface air and r_s the effective surface resistance to moisture flux. The second term $v_L m[q_s(T) - \overline{q}]$ is the latent heat flux due to evapotranspiration.

This model takes precipitation forcing and radiative forcing as inputs, along with initial conditions of temperature and soil moisture, to simulate the subsequent evolution of temperature over time. In this study, the input *P* and F_{SW} and the initial values of T_0 and m_0 are taken from the S2S ECMWF ensemble forecast, with T_0 = 305.5 K and m_0 = 0.3. To explore the impact of precipitation on T2max, we vary *P* to 50% and 200% of its original values in S2S forecasts, while adjusting F_{SW} according to the linear relationship between the two variables as derived from the S2S forecast ensemble.

5.5 Region definition

The regions analyzed in this study are defined as follows: the Yangtze River Valley (YRV) region corresponds to (27–33 °N, 100–120 °E); the Pakistan region is represented by (22–32 °N, 63–73 °E); and the Northern Europe region encompasses (53–68 °N, 25–61 °E). Following an earlier study[41], we define the intensity of the Western North Pacific Subtropical High (WNPSH) using the mean 850-hPa geopotential height anomaly over the region (15–25 °N, 115–150 °E).

Data Availability Statement

ERA5 data is acquired from *https://cds.climate.copernicus.eu/datasets*. The S2S database is

available online at *https://apps.ecmwf.int/datasets/data/s2s*. The MSWEP V2 precipitation is available at *https://www.gloh2o.org/mswep/*.

Code Availability Statement

All figures in this article are produced by the Python version 3.10, and the source codes can be obtained upon request from the corresponding authors. Z-model can be downloaded from *https://github.com/vargaszeppetello/*.

References

1. Perkins-Kirkpatrick, S. E. & Lewis, S. C. Increasing trends in regional heatwaves. *Nat Commun* **11**, 3357 (2020).
2. Vautard, R. et al. Heat extremes linearly shift with global warming, with frequency doubling per decade since 1979. *Environmental Research Letters* **19** (2024).
3. Li, H., Wei, Y. D. & Swerts, E. Spatial inequality in the city-regions in the Yangtze River Valley, China. *Urban Studies* **57**, 672-689 (2020).
4. Wei, J., Ting, D., Hui, G. & Zhuozhuo, L. China's Yangtze River basin is becoming the super heatwave centre in the East Asian monsoon regions. *International Journal of Climatology* **44**, 5028-5038 (2024).
5. Lu, R. et al. Heat waves in summer 2022 and increasing concern regarding heat waves in general. *Atmospheric and Oceanic Science Letters* **16** (2023).
6. Mallapaty, S. China's extreme weather challenges scientists trying to study it. *Nature* **609**, 888-888 (2022).
7. Zhang, D., Chen, L., Yuan, Y., Zuo, J. & Ke, Z. Why was the heat wave in the Yangtze River valley abnormally intensified in late summer 2022? *Environmental Research Letters* **18** (2023).
8. He, C., Zhou, T., Zhang, L., Chen, X. & Zhang, W. Extremely hot East Asia and flooding western South Asia in the summer of 2022 tied to reversed flow over Tibetan Plateau. *Climate Dynamics* **61**, 2103-2119 (2023).
9. Huang, Z., Tan, X. & Liu, B. Relative Contributions of Large‐Scale Atmospheric Circulation Dynamics and Anthropogenic Warming to the Unprecedented 2022 Yangtze River Basin Heatwave. *Journal of Geophysical Research: Atmospheres* **129** (2024).
10. Jiang, J., Liu, Y., Mao, J. & Wu, G. Extreme heatwave over Eastern China in summer 2022: the role of three oceans and local soil moisture feedback. *Environmental Research Letters* **18** (2023).
11. Wang, Z., Luo, H. & Yang, S. Different mechanisms for the extremely hot central-eastern China in July–August 2022 from a Eurasian large-scale circulation perspective. *Environmental Research Letters* **18** (2023).
12. Hong, C.-C. et al. Causes of 2022 Pakistan flooding and its linkage with China and Europe heatwaves. *npj Climate and Atmospheric Science* **6** (2023).
13. Zhang, J., Chen, H., Fang, X., Yin, Z. & Hu, R. Warming-induced hydrothermal anomaly over the Earth's three Poles amplifies concurrent extremes in 2022. *npj Climate and Atmospheric Science* **7** (2024).
14. Zhou, B. et al. The extreme heat wave in China in August 2022 related to extreme northward movement of

the eastern center of SAH. *Atmospheric Research* **293** (2023).
15. Tang, S. et al. Linkages of unprecedented 2022 Yangtze River Valley heatwaves to Pakistan flood and triple-dip La Niña. *npj Climate and Atmospheric Science* **6** (2023).
16. Zhang, D., Huang, Y., Zhou, B., Wang, H. & Sun, B. Who is the major player for 2022 China extreme heat wave? Western Pacific Subtropical high or South Asian high? *Weather and Climate Extremes* **43** (2024).
17. Yuan, S., Sun, X., Zhang, X., Xu, S. & Yang, X. Q. The Contrast Precipitation Patterns in Yangtze River Valley Between the Two La Niña Decaying Summers in 2021 and 2022. *Journal of Geophysical Research: Atmospheres* **129** (2024).
18. Ma, Q. et al. Multiscale interaction underlying 2022 concurrent extreme precipitation in Pakistan and heatwave in Yangtze River Valley. *npj Climate and Atmospheric Science* **7** (2024).
19. Lu, R. Y., Oh, J. H. & Kim, B. J. A teleconnection pattern in upper-level meridional wind over the North African and Eurasian continent in summer. *Tellus A* **54**, 44-55 (2002).
20. Enomoto, T., Hoskins, B. J. & Matsuda, Y. The formation mechanism of the Bonin high in August. *Quarterly Journal of the Royal Meteorological Society* **129**, 157-178 (2003).
21. Ni, Y. et al. Shift of soil moisture-temperature coupling exacerbated 2022 compound hot-dry event in eastern China. *Environmental Research Letters* **19** (2024).
22. Chen, X. et al. Land-atmosphere feedback exacerbated the mega heatwave and drought over the Yangtze River Basin of China during summer 2022. *Agricultural and Forest Meteorology* **361** (2025).
23. Vitart, F. et al. The Subseasonal to Seasonal (S2S) Prediction Project Database. *Bulletin of the American Meteorological Society* **98**, 163-173 (2017).
24. Vitart, F. & Robertson, A. W. The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *npj Climate and Atmospheric Science* **1** (2018).
25. Ardilouze, C., Batté, L. & Déqué, M. Subseasonal-to-seasonal (S2S) forecasts with CNRM-CM: a case study on the July 2015 West-European heat wave. *Advances in Science and Research* **14**, 115-121 (2017).
26. Wang, S., Anichowski, A., Tippett, M. K. & Sobel, A. H. Seasonal Noise Versus Subseasonal Signal: Forecasts of California Precipitation During the Unusual Winters of 2015–2016 and 2016–2017. *Geophysical Research Letters* **44**, 9513-9520 (2017).
27. Gibson, P. B. et al. Subseasonal-to-Seasonal Hindcast Skill Assessment of Ridging Events Related to Drought Over the Western United States. *Journal of Geophysical Research: Atmospheres* **125** (2020).
28. Li, X., Chen, R. & Qiao, Y. Assessing the extended-range forecast skills of the extreme heat events over South China based on three S2S models. *Atmospheric Science Letters* **25** (2024).
29. Lin, H., Mo, R. & Vitart, F. The 2021 Western North American Heatwave and Its Subseasonal Predictions. *Geophysical Research Letters* **49** (2022).
30. Mitropoulos, D., Pytharoulis, I., Zanis, P. & Anagnostopoulou, C. Subseasonal-to-Seasonal Predictability Assessment of an Early Heat Wave in the Eastern Mediterranean in May 2020. *Comecap* **2023** (2023).
31. Qi, X. & Yang, J. Extended-range prediction of a heat wave event over the Yangtze River Valley: role of intraseasonal signals. *Atmospheric and Oceanic Science Letters* **12**, 451-457 (2019).
32. White, R. H. et al. The unprecedented Pacific Northwest heatwave of June 2021. *Nat Commun* **14**, 727 (2023).
33. Xiao, H., Xu, P. & Wang, L. The Unprecedented 2023 North China Heatwaves and Their S2S Predictability. *Geophysical Research Letters* **51** (2024).

34. Zeppetello, L. R. V., Battisti, D. S. & Baker, M. B. The Physics of Heat Waves: What Causes Extremely High Summertime Temperatures? *Journal of Climate* **35**, 2231-2251 (2022).
35. Hersbach, H. et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **146**, 1999-2049 (2020).
36. Beck, H. E. et al. MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society* **100**, 473-500 (2019).
37. Tian, W., Wu, Y.-l., Lin, C., Zhang, J.-g. & Lim Kam Sian Kenny, T. C. Assessing the Applicability of Multi-Source Precipitation Products over the Chinese Mainland and Its Seven Regions. *Journal of Tropical Meteorology* **30**, 275-288 (2024).
38. Buizza, R. Potential Forecast Skill of Ensemble Prediction and Spread and Skill Distributions of the ECMWF Ensemble Prediction System. *Monthly Weather Review* **125**, 99-119 (1997).
39. Scherrer, S. C., Appenzeller, C., Eckert, P. & Cattani, D. Analysis of the Spread–Skill Relations Using the ECMWF Ensemble Prediction System over Europe. *Weather and Forecasting* **19**, 552-565 (2004).
40. Zhang, T., Deng, Y., Chen, J., Yang, S. & Dai, Y. An energetics tale of the 2022 mega-heatwave over central-eastern China. *npj Climate and Atmospheric Science* **6** (2023).
41. Xiang, B., Wang, B., Yu, W. & Xu, S. How can anomalous western North Pacific Subtropical High intensify in late summer? *Geophysical Research Letters* **40**, 2349-2354 (2013).

Figure Captions

Figure 3. Temperature and precipitation anomalies during the 2022 YRV heatwave. Observed anomaly of monthly average T2max (shading, °C) in a) July and b) August 2022, based on ERA5 reanalysis. The blue dashed box indicates the location of the Yangtze River valley region (27–33 °N, 100–120 °E). c) Monthly precipitation anomaly (shading, mm month^{-1}) in August 2022, based on MSWEP V2. The blue dashed box indicates the YRV region, red dashed box, Pakistan (22–32 °N, 63–73 °E).

Figure 4. Subseasonal forecasts of T2max over the YRV during summer 2022 by the ECMWF and CMA models. Chiclet diagram of daily ensemble-mean T2max forecasts from ECMWF a) and CMA b), as a function of the verification date (horizontal axis) and lead time (vertical axis). The orange curves below the chiclet chart in (a) and (b) represents the ERA5 reanalysis daily T2max (°C). c) Ensemble mean (gray) of ECMWF T2max forecast initialized on 28 July 2022. Composite of ensemble members in three groups is categorized based on the average T2max over August 10th - 23rd: Warm (red line), Medium (orange line), and Cold (blue line). The ERA5 T2max is also shown in black. d) Same as c), but with the initialization date at 8 August 2022.

Figure 3. Relationships among T2max, precipitation, and soil moisture in Reforecast and Real-Time August Forecasts. Forecasts are initialized on 8 August, and evaluations are for the average over the period from August 10th - 23rd. a), b) Correlation for reforecast 20 years (2002–2021): a) T2max versus precipitation, and b) precipitation versus soil moisture. c), d) same as a), b), but for real-time forecast (2022). Hollow scattered dots represent individual ensemble members, while solid squares denote the ensemble mean. The correlation coefficients and regression equations are provided in the top right and bottom right of each panel,

respectively. The scatterplots show area-averaged values over the YRV region for each year.

Figure 4. Large-scale circulation anomalies during the 10–23 August 2022 YRV heatwave. Contours of 500-hPa geopotential height anomaly (gpm; red, positive, blue, negative) from the ECMWF forecast initialized on 8 August 2022 a) and from ERA5 reanalysis b). The orange rectangle indicates the YRV region. The dark-cyan rectangle indicates the North Europe region (53–68 °N, 25–61 °E).

Figure 5. Comparison of key land–atmosphere factors during the YRV extreme heatwave across ensemble groups. Ensemble members are categorized into Warm (red line), Medium (orange line), and Cold (blue line) according to their 2-week predictions of T2max, while the reanalysis values are shown as a unified reference (black line; ERA5 for all variables except precipitation, which is from MSWEP V2). a) Soil moisture at 20 cm depth (Soil Moisture20, kg m^{-3}); b) Precipitation over YRV (mm day^{-1}); c) Precipitation over Pakistan (mm day^{-1}); d) Z500 anomaly over YRV (gpm); e) Z500 anomaly over North Europe (gpm); f) WNPSH intensity: Z850 anomaly over (15–25 °N, 115–150 °E) (gpm). Two-week average values are displayed in the top left or bottom left of each subplot.

Figure 6. Prediction of T2max using a multiple linear regression model with six key factors. a) Scatter plot comparing T2max predictions from the linear regression model with forecasted T2max by ECMWF forecasts. Each dot indicates one forecast ensemble member. Values of slope and correlation coefficient are displayed in the top left of the subplot. b) Reduction in R^2 after each factor is removed. Soil moisture was replaced with its residual for the analysis.

Figure 7. Relationship between T2max and precipitation in the Z-model and sensitivity of T2max to precipitation forcing. a) Scatter plot of daily maximum temperature (T2max) versus precipitation simulated by the Z-model. The hollow square indicates the ensemble mean. The fitted linear regression line and Pearson correlation coefficient (Corr.) are shown in the panel. b) Mean and standard deviation (1σ error bars) of T2max from ECMWF and Z-model simulations using different precipitation forcing: the default (1×P), doubled (2×P), and halved (0.5×P) precipitation.

Acknowledgements

B.J.L and S.W. acknowledge the funding support of National Natural Science Foundation of China No. 42275055, and the National Key R&D Program of China, Project No. 2024YFC3013100.

Author Contributions

S.W. and B.J.L. conceived and designed the study. B.J.L and S.W. performed the analysis. B.J.L. generated the figures. B.J.L., S.W., G.C., and B.X wrote the manuscript and contributed to the interpretation of the results, as well as to the review and editing of the manuscript.

Competing Interests

All authors declare no financial or non-financial competing interests.

Editorial summary:
Precipitation–soil moisture coupling plays a key role in modulating the intensity of a prolonged Yangtze River Valley heatwave, confirming the dominant influence of local precipitation in limiting its subseasonal predictability, according to high-resolution heatwave analyses.

Peer review information:
Communications Earth and Environment thanks Yamin Qing, Mahmoud Osman and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Primary Handling Editors: Yi Jiao and Somaparna Ghosh
[A peer review file is available.]