# Worst-Case European Heat and Drought Storylines generated using Ensemble Boosting

**Laura Suarez-Gutierrez**

laura.suarez@lsce.ipsl.fr

Institute for Atmospheric and Climate Science, ETH Zurich

**Urs Beyerle**

Swiss Federal Institute of Technology https://orcid.org/0000-0002-6464-0838

**Magdalena Mittermeier**

Department of Geography, Ludwig-Maximilians-Universität München

**Robert Vautard**

IPSL

**Erich Fischer**

ETH Zurich https://orcid.org/0000-0003-1931-6737

**Article**

**Keywords:**

**Posted Date:** July 24th, 2025

**DOI:** https://doi.org/10.21203/rs.3.rs-7113183/v1



**Additional Declarations:** There is **NO** Competing Interest.

# Worst-Case European Heat and Drought Storylines generated using Ensemble Boosting

Laura Suarez-Gutierrez[1,2,*], Urs Beyerle[1], Magdalena Mittermeier[3], Robert Vautard[2], Erich M. Fischer[1]

1. Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
2. Institut Pierre-Simon Laplace, CNRS, Paris, France
3. Department of Geography, Ludwig-Maximilians-Universität München, Munich, Germany

*Corresponding author: laura.suarez@lsce.ipsl.fr

## Abstract

Heat and drought extremes pose escalating socio-economic and ecological risks, yet the most severe combinations of these high-impact extremes possible today remain poorly understood. Using thousands of plausible ensemble-boosting current climate storylines, we reveal the risk for more intense drought compounding with far more extreme heat and fire weather than ever experienced over Europe in the recent past. The most extreme boosted heatwaves surpass historical extremes in both intensity and particularly in persistence, and also exceed levels considered extreme in a 3°C warmer world by large margins. Some of the most extreme heatwaves arise under severe soil moisture depletion, while others develop under strong surface temperature gradients in the North Atlantic and extreme heat in the nearby Mediterranean and Atlantic basins, underscoring the diversity of pathways to worst-case conditions. Furthermore, our work reveals an additional risk: worst-case heatwaves occur predominantly after another extreme heatwave. This highlights the potential for aggravated impacts due to decreased recovery times and intensified heat stress on humans, ecosystems and infrastructure made more vulnerable by the first event. Given the scale, intensity, and unprecedented successive and compounding nature of these worst-case storylines, we underscore the urgent need for well-informed adaptation strategies that sufficiently reflect these risks.

## Introduction

In a warming world, extreme heat and drought will become more frequent and more extreme, especially over Europe[1–4]. Extreme heat and drought events in the recent past, such as the compound heat and drought levels in Europe in 2003[5–7] or 2018[8,9], or the Pacific Northwest heatwave in 2021[10,11], shattered previous records by large margins. These unprecedented extremes also caused vast impacts, including widespread crop and ecosystem damages, wildfires, air pollution, and thousands of fatalities[6,8,10,12–14]. Based on observational data and traditional modelling approaches, these record-breaking events appeared to be extremely rare to virtually impossible prior to their occurrence[5,11,15,16].

These unprecedented, high-impact events demonstrate the urgency to better understand worst-case conditions that could unfold not only in a distant much warmer future, but already under current climate conditions. To produce robust and comprehensive risk assessments and well-informed adaptation strategies, it is crucial to shift our focus away from the likely range of extreme conditions and toward the limits of what might be possible[17,18]. For this, an emerging approach is to generate 'storylines', or non-probabilistic but plausible event-based narratives exploring the unfolding of a given set of extreme conditions[11,17,19].

There is an increasing number of techniques used to generate such extreme storylines, including observational-based statistical approaches[3,20] and statistical weather generators[21,22]; or modelling-based approaches such single-model initial condition large ensembles[4,23] and initialized hindcast ensembles for weekly to seasonal predictions[24,25]. A combination of statistical and model-based approaches can be seen for large deviation theory analyses[26] and for rare event sampling algorithms[27,28]. Additionally, other methods explore the extreme conditions that certain atmospheric patterns would produce under higher background warming or drier soils, via statistical sampling in the case of flow analogue approaches[29–31] or via spectrally nudged model simulations[32,33].

Here, we use Ensemble Boosting[11,31,34,35] to generate extreme heat and drought storylines that would be plausible under current climatic conditions. This physically-consistent variability sampling technique is based on rerunning simulated extreme conditions, or *parents*, to create hundred to thousands of alternative realizations of the original parent state, after introducing small rounding-error perturbations in a fully-coupled, global Earth system model, namely the Community Earth System Model version 2 (CESM2[36]). This approach, first introduced to sample pre-industrial climate extremes[34,35], has been demonstrated to generate storylines resembling observed extremes such as the 2021 Pacific Northwest heatwave[11] or the 1963 Central Europe cold spell[31].

Compared to other commonly used methods to produce extreme event storylines, Ensemble Boosting presents some key advantages. First, the minimal perturbations used to sample chaotic climate variability combined with the use of a fully-coupled global climate model ensures that the storylines produced maintain physical consistency across time, space, and variables; including for unprecedented, out-of-sample conditions. This is particularly important when exploring unprecedented extremes and behaviors that cannot be directly evaluated against similar real-world conditions because they exceed observations by large margins, or because they involve combinations of factors never seen in the historical record. It is also a computationally efficient approach for producing large samples of extreme events, with achievable boosted ensemble sizes of thousands of members. This unique combination of large ensemble sizes and physical consistency is crucial to assess unprecedented extreme events, especially those that involve even more rare compounding or successive extreme conditions[4,37].

Furthermore, Ensemble Boosting does not follow a probabilistic approach, nor does it rely on any statistical estimations, assumptions or parametrizations regarding the tail risks or the shape of the distribution of potential outcomes. And, beyond the selection of the initial parent state, it is not targeted nor nudged toward any particular development, as opposed to previous work targeting simulations for extremely warm summers[27,28]. This non-probabilistic experimental set-up ensures that the storylines produced represent freely-emerging sequences of potentially unprecedented conditions and compounding hazards that are, at least within the model world, physically plausible.

Most of the previous work exploring storylines of heat and drought extremes focuses on resembling recent high-impact individual heatwaves[11,26,30,32,33] or droughts[23], or explores univariate extremes on synoptic scales[11,22,30]. The limited body of work exploring extreme storylines that go beyond those resembling past events and beyond synoptic scales focuses on developing simulations targeting high summer mean temperature anomalies[27,28]. These simulations are performed under atmospheric perturbations several orders of magnitude larger than those in the ensemble boosting approach used here, to ensure faster deviations from their initial state. And instead of evolving independently, they are monitored periodically every few simulated days, and re-perturbed and continued only if they exhibit increasing temperature anomalies. Furthermore, they are performed in an atmosphere-land only model settings, with fixed ocean surface conditions and a neutral initial soil moisture state. Although this experimental setup is not completely comparable with the setup used in our work and does not reflect freely-emerging storylines, it shows the potential for summers with mean temperature anomalies up to three times higher than the most extreme observed, which arise due to succeeding heatwaves through the season.

Here, we employ this novel fully-coupled and physically consistent Ensemble Boosting modelling technique to explore freely-emerging, heat and drought storylines spanning through the extended summer season between May and September, and that reflect worst-case extreme conditions plausible in the current climate. We define current climate conditions as those simulated in a world roughly 1˚C warmer on average than preindustrial conditions. Using these storylines, we assess how extreme heat events compound with other system stressors such as meteorological or soil drought or extreme atmospheric evaporative demand to produce the sequences in extreme conditions that yield the most extreme yet physically plausible worst-case storylines of extreme European heat and drought.

## Results

### Heat and drought extremes in CESM2 against ERA5

To demonstrate the ability of the CESM2 large ensemble (CESM2-LE) to capture real-world extreme heat and drought conditions, we evaluate heat and drought characteristics in CESM2-LE against the ERA5[38] reanalysis (Fig. 1 and Supporting Information SI Fig. S1). The CESM2-LE simulations used in this analysis are available

exclusively for the period of 2005-2035 (See methods). Although alternative CESM2 simulations with slightly different setups are available for earlier periods, to avoid inconsistency across the simulations used we choose to use only these simulations for the period of 2005-2035, which is particularly representative of recent past and near future conditions. Therefore, to evaluate the model data against reanalysis data under comparable conditions, we base our supporting evaluation (SI Fig. S1) on the period of 2005-2023 for CESM2-LE and 2000-2023 for ERA5. We extend this period back for ERA5 to include relevant recent extreme events and increase the observed sample size, while maintaining comparable climatic conditions. We also how comparisons against ERA5 for the period of 1970-2023, which yields similar results. Note however that the steep change in background global warming conditions along this longer period in ERA5 may confound its comparison against CESM2-LE simulations for the 2005-2023 period. Beyond this evaluation, all results in the main analysis are based on the same CESM2 simulations for current climate conditions defined by the period of 2010-2028. These current climate conditions are defined by the 2010-2028 period in CESM2-LE, which under historical and SSP3.7 forcing conditions equates to global mean surface temperatures approximately 1˚C higher than preindustrial levels (see Methods).

Our evaluation shows that CESM2-LE adequately captures the variability of ERA5 heat and drought measured as accumulated excess heat (defined as maximum temperature exceedances above the local 95th percentile) and accumulated precipitation deficits during the extended (May-September) summer season (Fig. 1a and SI Fig. S1a and b; see methods for further metric definition details). The two heat and drought metrics exhibit probability distributions, 90th percentile levels, and standard deviations ($\sigma$) that are comparable across model and reanalysis (SI Fig. S1a and b; CESM2-LE $\sigma$ = 6.44, ERA5 $\sigma$ = 6.3 for excess heat; CESM2-LE $\sigma$ = 0.54, ERA5 $\sigma$ = 0.52 for precip. def.). We also find good agreement between CESM2-LE simulations and ERA5 for explicitly defined heatwaves (Fig. 1b; SI Fig. S1c), defined here as periods of a minimum duration of 5 days where at least 30% of the land grid cells of the Central Europe region of study exhibit positive Excess Heat values (see methods). Based on this definition, we characterize heatwaves using a cumulative heatwave intensity metric that reflects the accumulated excess heat summed over the whole duration of the heatwave. We find distributions, 90th percentile levels, and standard deviations for heatwave intensities that are comparable across model and reanalysis (SI Fig. S1c; CESM2-LE $\sigma$ = 2.77, ERA5 $\sigma$ = 3.07).

CESM2-LE shows good agreement with the ERA5 reanalysis samples of accumulated excess heat, precipitation deficits and heatwave intensity. However, this evaluation is limited by the relatively short record of the ERA5 reanalysis and limited sample of extreme heat and drought under current climate conditions, compared to the 35-member CESM2 ensemble. In the CESM2-LE sample, the most extreme extended summer droughts reach levels up to 0.5$\sigma$ more intense than the most extreme droughts in ERA5, namely those in

2018, 2022 and 2003 (Fig. 1b and c). This behaviour is found both for precipitation deficit drought (Fig. 1b) as well as for more strictly defined soil drought metrics using soil moisture (SM) deficits (Fig. 1c; see methods), and for the extended (May-Sept.; Fig. 1c) as well as early (May-June; Fig. 1d) summer seasons. This comparison also shows a clear relationship toward drier and hotter conditions occurring concurrently, both in CESM2-LE and in ERA5. However, the most extreme heat does not necessarily coincide with the most extreme drought in neither sample (Fig. 1a).

For extreme heat, the most extreme events simulated by CESM2-LE in the period of 2010-2028 exceed the most extreme events captured in the 1940-2023 period in ERA5 by large margins of more than 5σ for accumulated excess heat. In the most extreme current-climate heatwaves, CESM2-LE exhibits almost double heatwave intensities than those observed during the 2003 heatwave, the most extreme on record over the Central Europe region assessed here. Although we find the CESM2-LE and ERA5 distributions to have comparable standard deviations and 90th percentile levels (SI Fig. S1c); the large intensities of the far-tail heatwaves in CESM2-LE could indicate a potential overestimation of extreme heat by CESM2-LE. However, the fact that the most extreme event in ERA5 happened in 2003, 25 years earlier than some of the events in the CESM2-LE distribution presented here; combined with the good agreement between both distributions and the relatively short observational period available for ERA5 point to the possibility that events as extreme as those at the tails of the CESM2-LE distribution are plausible, but have not yet occurred, as other lines of evidence also suggest[3].

The 2003 heatwave, which peaked in mid-August, also marks the time of the extended summer season when the most extreme heatwaves tend to occur in CESM2-LE (Fig 1e). Note that this effect is present even for a threshold definition that accounts for seasonality effects as used here, defined as a moving daily heat threshold calculated for each day and grid cell using a centered 15-day window (See Methods for further details). This effect indicates the largest day-to-day deseasonalized variability in this period of the late summer, which could result from potentially compounding factors that increase the likelihood of more intense or more persistent heatwaves, such as heat advection or moisture-temperature feedbacks exacerbated towards the end of the summer.

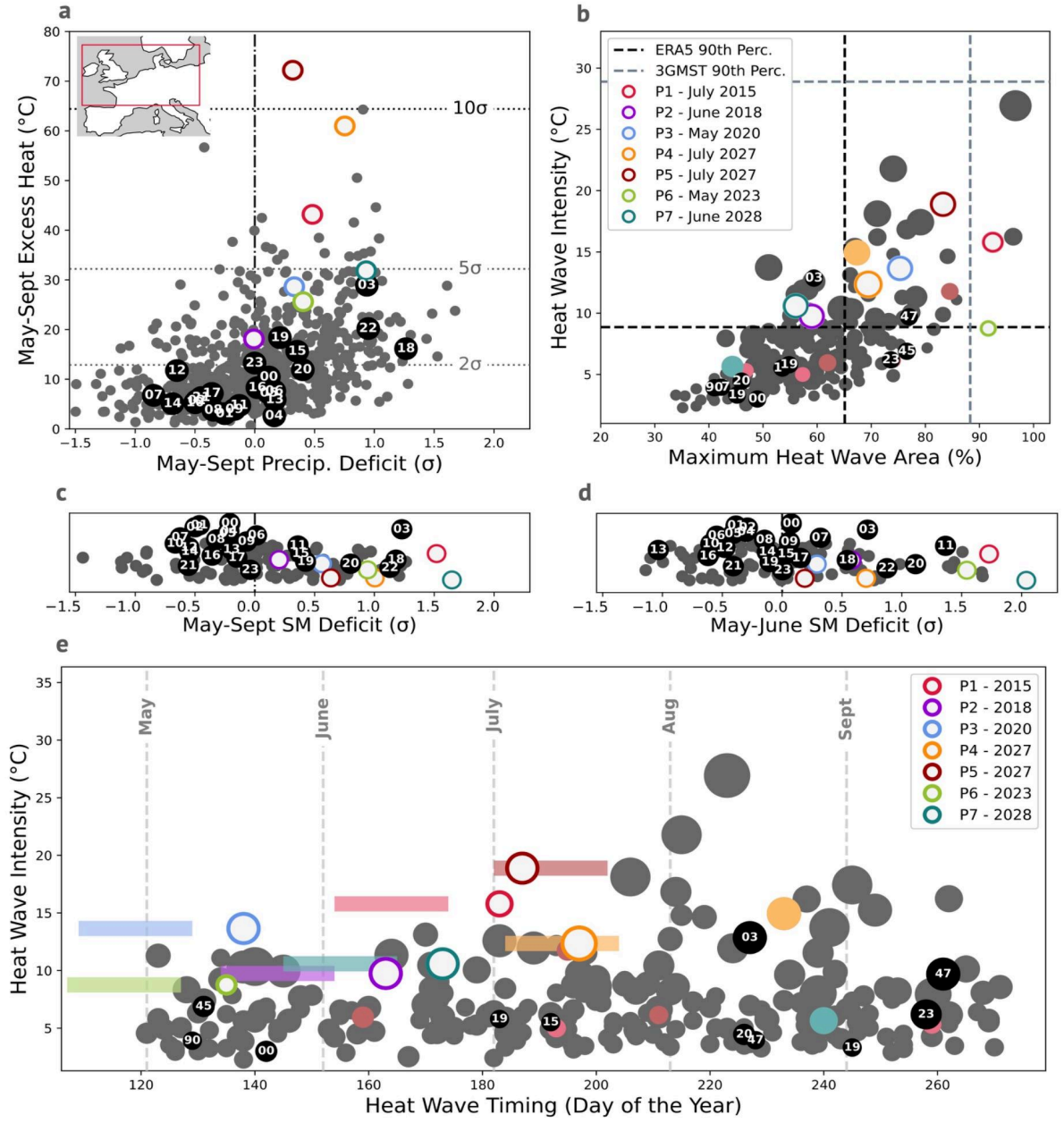



**Figure 1: Heat and Drought conditions in non-boosted ensemble and selected parent storylines vs ERA5. a)** Accumulated Excess Heat and Standardized Precipitation Deficits for the extended summer season (May to September) in ERA5 (black dots, year in white; 2000-2023), 35-member CESM2-LE (gray dots; 2010-2028) and in the 7 parent heatwaves selected for boosting (P1-7; colored, white infilled circles).Dotted lines mark several excess heat standard deviation (σ) levels based on CESM2-LE data. **b)** Same as a) but for Heatwave intensity against maximum heatwave area in ERA5 (black dots, year in white; 1940-2023), 35-member CESM2-LE (gray dots; 2010-2028). Additional heatwaves in the parent simulations before or after the parent heatwaves are also shown (full colored dots). Dot size is proportional to heatwave persistence. Dashed lines mark the 90$^{th}$ percentile across all ERA5 heatwaves in the period of 2000-2023 (black) and for the CESM2-LE projections for a 3˚C warmer world (gray). **c)** Same as a), but for May-September soil moisture (SM) deficit in ERA5-Land (black), 5-member CESM2-LE (grey) and parent events (colored circles). **d)** Same as c), but for May-June soil moisture (SM) deficit. **e)** Same as a) but for heatwave intensity against heat peak timing. Colored horizontal lines mark the 20-day boosting window for each parent case. Additional heatwaves in the parent simulations before or after the parent heatwaves are also shown (full colored dots). Dot size proportional to heatwave persistence. CESM2 simulations are always for the period 2010-2028 under historical and SSP3.7 scenario forcings. All values are aggregated over land grid cells in the Central Europe region marked by the red box. Increasingly positive precipitation and SM deficits indicate increasingly drier conditions.

To assess how much more extreme worst-case heat and drought conditions could become under current climate conditions, we select seven parent summers simulated by CESM2-LE, all between 2015 to 2028 (Fig. 1). Using Ensemble Boosting, these parent simulations are *boosted,* defined as rerunning the simulations after introducing a small, rounding-error perturbation to the atmospheric state prior to the parent event. The boosted simulations are created in batches of 50 members per boost date, with all realizations perturbed only once on consecutive calendar days for each 50-member set, spanning over a 20-day window (50 batches per perturbation date and 20 perturbation dates, yielding a 1.000 member boosted ensemble perturbed once at the start of the boosted simulation; see Methods for further details). With this boosting setup we generate a thousand realizations or storylines per parent case, sampling the effect of chaotic variability in the climate system on the evolution of extreme heat and drought. To sample a wide range of variability sources and illustrate the effects of different heat and drought parent conditions, we select parent events under a range of states, classified into three categories (Table 1).

First, to assess whether extreme heat in the parent simulation is a required condition to reach the most extreme boosted heatwaves, or whether even moderate heatwaves may reach similarly extreme intensities once atmospheric variability is sampled via ensemble boosting, we choose three parent events showcasing very extreme heat (P1, P4, P5); while the rest exhibit more moderately extreme heat levels, comparable to those reached during past observed extremes (Fig 1a and b). Second, to assess whether extremely dry conditions at the time of boosting are required to reach either the most extreme heatwaves, the most extreme extended summer drought, or the most extreme extended summer compound heat and drought conditions, we choose three parents exhibiting extreme early summer drought, characterized by large positive May-June soil moisture deficits (P1, P6, P7); while the rest exhibit more neutral soil moisture conditions (Fig 1c).

Third, we explore the effect of introducing the initial perturbation at different stages, to sample its effect on atmospheric variability before the heatwave onset (as done in previous ensemble boosting work[11]) versus during an active heatwave. This experimental setup consists of two parent simulations (P4, P5) perturbed following the same set up described above but with a 20-day boosting window that includes the peak in excess heat of the parent heatwave; while the other parent cases are perturbed within a 20-day window that ends 10 days prior to the peak in the corresponding parent heatwave (Fig. 1e). Therefore, the resulting P4 and P5 boosted simulations include the development of an identical parent heatwave in the common parent simulation period before their boosting, allowing us to assess whether previous extreme heat may influence heatwaves and drought the rest of the season. Lastly, to ensure that we sample the variability in heat and drought summer extremes along a sufficiently representative length of the summer season, we choose to start our boosted storylines from parent heat extremes that occur no later than July, and continue the boosted simulations in all cases until the end of October.

| Parent Case | Year | Boosted on | Extreme Heat | Early Drought | Boosted on HW |
|---|---|---|---|---|---|
| P1 | 2015 | June 22nd | x | x | - |
| P2 | 2018 | June 2nd | - | - | - |
| P3 | 2020 | May 8th | - | - | - |
| P4 | 2027 | July 22nd | x | - | x |
| P5 | 2027 | July 20th | x | - | x |
| P6 | 2023 | May 6th | - | x | - |
| P7 | 2028 | June 13th | - | x | - |

**Table 1: Conditions of the parent events selected for boosting.** Parent Case acronym (P1-7); year when the parent simulations occur in CESM2-LE; boosting date, exemplified here by the last date of the 20-day boosting window (see Methods for details). The last three columns mark whether the parent event exhibits extreme excess heat, extreme early summer drought, and if it is boosted during an ongoing heatwave, respectively.

**Worst-case ensemble boosting storylines for extreme heatwaves**

The time-series based comparison of the boosted simulations against their parents shows how minimal atmospheric perturbations can lead to different heatwave evolutions and exacerbate extreme heat in different ways (Fig. 2). For instance, the top three most intense boosted heatwaves for case P1, measured as the three heatwaves with the highest excess heat intensities accumulated over their whole duration (see Methods), exhibit three distinct behaviors (Fig. 2a). From these top three most intense realizations, two develop peak heat levels around the same time as the parent heatwave, roughly 10 days after the last boosting date. At this stage 10-days post-perturbation, heat conditions in the boosted simulations start to differ due to the sampling of atmospheric variability, while maintaining key patterns in the parent climatic and atmospheric state (Table 2). However, from these two of the top three most extreme P1 simulations, one reaches a much higher peak intensity compared to its parent, while the other reaches a similar peak intensity but persists for much longer. In contrast, the third of the top three most intense P1 boosted heatwaves shows a lower initial heat peak, but sustains extreme heat levels for over a month.

How the climate system reacts to this initial perturbation is highly case and event dependent. Right after the initial boosting perturbation is introduced, all realizations remain almost identical, with climate and atmospheric states that have yet not diverged from their parent nor each other. For the first days after the initial perturbation, the effect of chaotic internal variability comes into play mainly through sampling of atmospheric variability, causing atmospheric patterns present in the parent event to unfold differently, over different locations or with different persistence. This can be seen by assessing the

standard deviation across boosted members at different stages after the initial perturbation (Table 2). The initial boosting stage is characterized by the combination of increasingly different atmospheric states that reflect different possible evolutions of the original parent event, and that occur under background parent climate conditions, e.g., sea surface temperatures (SSTs), soil moisture levels (SM), that remain relatively unchanged across all realizations. For example, 10 days post-perturbation, boosted members for all parent cases show ensemble spreads below 15% for SSTs and below 2.5% for SM compared to the spread in independent realizations in the climatological reference period. Therefore, in this initial stage we explore how parent heatwaves may worsen due to a combination of dynamically-driven factors, while maintaining relatively similar conditions across other components of the Earth system. This initial boosting stage, which corresponds to the 2 to 3 weeks window of atmospheric predictability[39–41] sampled in traditional weather prediction, is the stage commonly explored in previous ensemble boosting studies[11,31,34].

In contrast, here we extend the boosted storylines beyond this stage of atmospheric memory conservation and an into a stage where realizations differ both due to atmospheric variability, now evolving beyond the patterns in the parent simulation, as well as due to the effects of sampling variability in other components, such as terrestrial water budgets or sea surface temperatures (Table 2). At the stage 60 days post-perturbation, we find boosted ensemble spreads in atmospheric conditions, measured as spatially averaged Z500 anomalies, almost as large as the spread in the non-boosted CESM2-LE ensemble for most parents; similar to the behaviour seen for SSTs conditions. In contrast, almost all parents exhibit spreads in maximum temperatures and soil moisture deficits below 75% of the climatological reference spread, indicating that boosted simulations remain more similar to each other than an independent sample and exhibit above-average heat and drought conditions throughout the summer (Table 2).

The simulations used here extending well beyond the stage of atmospheric predictability are key to explore multiple sources of variability and how they may compound to produce worst-case heat and drought. And it is particularly during this extended stage where we find the most extreme heat for most parent cases. The exception of this behavior can be seen to some extent for P7 (Fig. 2g), and especially for P1 (Fig. 2a), which exhibits its most intense heatwaves consistently during the initial boosting stage in early July, indicating that the intensification of this particular heatwave is strongly driven by the intensification of its dynamic drivers. For all other cases, the majority of the most intense boosted heatwaves occur from July to August, coinciding with the period when the original non-boosted CESM2-LE simulations also exhibit the most intense heatwaves (Fig. 1e). This timing of the most extreme heatwaves remains largely unchanged regardless of whether this period falls within the initial boosting stage (e.g, P4; Fig. 2d) or well within the extended stage (e.g, P6; Fig. 2f).

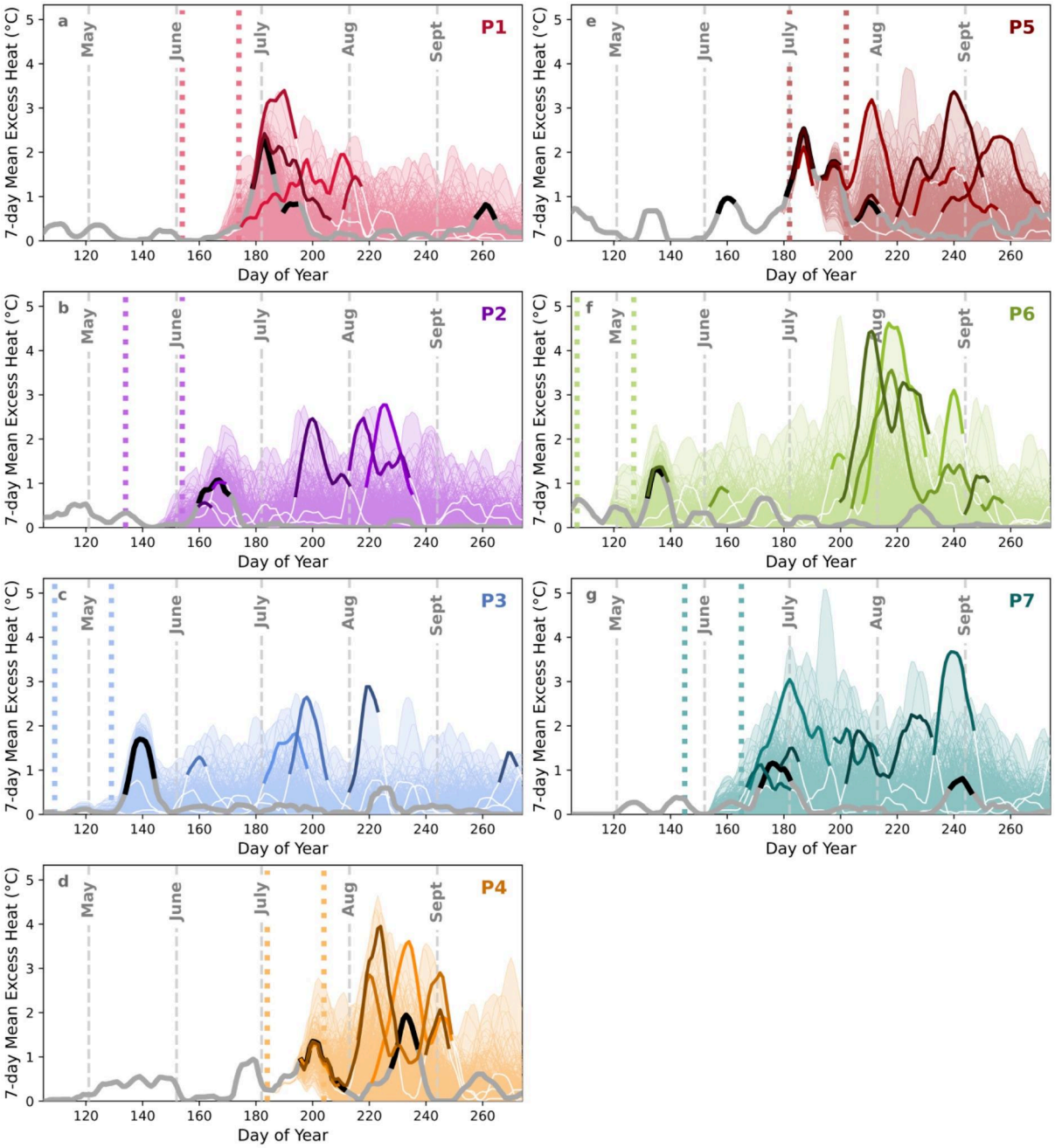



**Figure 2: Time series evolution of Boosted vs. Parent Heatwaves. a-g)** Excess Heat, 7-day centered running mean for boosted simulations (colored thin lines), against parents (black and gray). Dotted vertical lines mark the 20-day boosting window. The realizations with the 3 highest intensity heatwaves for each case, measured as excess heat accumulated for the duration of the heatwave, are highlighted (thick colored lines). Heatwave periods are marked in black for parent and color for boosted simulations; non-heatwave periods are marked gray and white, respectively.

| | σ_Z500 | | σ_Tmax | | σ_SM | | σ_SST | |
|---|---|---|---|---|---|---|---|---|
| | **10 days** | **60 days** | **10 days** | **60 days** | **10 days** | **60 days** | **10 days** | **60 days** |
| **P1** | 13.4 | 79.3 | 5.3 | 54.9 | 0.6 | 56 | 5.9 | 98.2 |
| **P2** | 10.8 | 92.7 | 14.9 | 84.2 | 2.2 | 70 | 11.3 | 95.6 |
| **P3** | 7.1 | 96.5 | 24.5 | 72.7 | 1.2 | 72.8 | 8.3 | 101.5 |
| **P4** | 13.2 | 95.7 | 7.8 | 72.4 | 0.8 | 45.9 | 8.7 | 86.7 |
| **P5** | 12.2 | 79.1 | 11.6 | 68.3 | 0.5 | 69.1 | 9.3 | 100 |
| **P6** | 5.9 | 97.7 | 8.6 | 90.1 | 1.2 | 89.8 | 5.3 | 114.1 |
| **P7** | 9.7 | 65.8 | 11.6 | 60.7 | 0.8 | 48.8 | 14.5 | 80.7 |

**Table 2: Standard deviation (σ) across boosted members at different post-perturbation stages (%).** Standard deviations across boosted simulations for all parent cases (P1-7) at different stages, respectively 10 and 60 days, after the initial perturbation. Values are shown for standard deviations in geopotential height anomalies at the 500 hPa level (z500), maximum temperatures (T_max), soil moisture deficit anomalies (SM), and sea surface temperature anomalies (SST), all for spatially averaged daily values. The values are expressed as a percentage of the standard deviation of the respective climatological reference for each metric, defined as the standard deviation across the 35-member non-boosted CESM2-LE and all years in the period of 2005-2023, for the same day of the year as the respective boosted simulations. The values for the boosted simulations for each parent case correspond to one set of 50 members for the last boosting date in the 20-day boosting window. Therefore, a value of 100% indicates that, at the marked time, the 50 boosted members initialized from identical conditions in each parent event show a spread equals the spread found across the realizations from the non-boosted ensemble, which originate from different initial conditions.

Assessing the characteristics of boosted heatwaves across all parent events (Fig. 3) reveals a clear pattern: moderately hot, moderately dry parent realizations (P2, P3; total sample size of 2000 summers) produce boosted heatwaves that in the most extreme cases are similar to slightly more intense than the most extreme non-boosted heatwaves in the 35-member CESM2-LE (total sample size of 665 summers). This indicates that the variability sampling achieved with ensemble boosting over a sufficiently large boosted ensemble, combined with extending boosted simulations beyond the initial memory-preserving stage, can produce similarly extreme heatwaves as an uninitialized large ensemble of independent members. However, the most extreme heatwaves produced from the more moderate parents are well below the worst-case, most intense heatwaves possible, which result only from the most extreme parent cases.

The parents that most often yield the most extreme boosted heatwaves are P1, P4, P5, P6, P7 (Fig. 3a and b). What these five parent cases have in common is extreme heat in the parent simulation (P1, P4 and P5), extreme early summer drought (P1, P6, and P7), and

being boosted during an ongoing heatwave (P4, P5). In particular, P4 produces large numbers of heatwaves reaching intensities close to its three most intense ones. P6, on the other hand, produces the top three most intense heatwaves overall, which reach accumulated intensity levels that surpass all observed, boosted and non-boosted heatwaves by large margins. The most extreme boosted heatwaves also surpass heatwave levels that would be considered extreme in a 3˚C warmer world (see methods for details), reaching accumulated intensities of more than double than the 90th percentile intensity levels in a 3˚C world (Fig. 3a). The most extreme boosted heatwaves reach accumulated excess-heat intensities (defined as the daily temperature anomaly with respect to the 95$^{th}$ percentile threshold summed over the length of the heatwave, see Methods) well beyond 40˚C, exceeding 80˚C in the most extreme case (Fig. 3a).

To put these accumulated heatwave intensities in context, the most extreme heatwave in the observational record in August 2003 exhibits 10˚C of accumulated heatwave intensity; produced by temperatures more than 1˚C above the local excess heat threshold on average across the duration of the heatwave sustained over a 10-day period (Fig. 3b). The most extreme heatwaves in the non-boosted CESM2-LE reach slightly below 30˚C accumulated excess heat intensities, values roughly equivalent to the 90$^{th}$ percentile heatwaves in a 3˚C warmer world (Fig. 3a; see Methods), and daily mean heat intensities of 1-2˚C above the local heat threshold (Fig. 3b). The most extreme boosted heatwave reaches an accumulated intensity of 80˚C by sustaining daily mean heat exceedances of almost 3˚C above threshold on average for more than 30 consecutive days (Fig. 3b).

The heat intensification present in these worst-case boosted heatwaves occurs due to higher threshold exceedances on average over the whole period of the heatwave, with mean heatwave intensities over 1.5˚C to 3˚C for the most extreme heatwaves (Fig. 3b). These higher mean heatwave intensities (defined as the daily temperature anomaly with respect to the 95$^{th}$ percentile threshold averaged over the length of the heatwave, see Methods) are caused by extremely high absolute temperatures, well over 40˚C in many cases (Fig. 3c). However, compared to the increase in accumulated intensities, the most extreme maximum heatwave temperatures show a relatively moderate increase of 2-3˚C from the most extreme temperatures in the non-boosted or observed heatwaves. Furthermore, daily mean excess heat intensities beyond 2˚C remain relatively rare, even among boosted heatwaves. This indicates that the largest increases in accumulated heatwave intensities do not arise solely from much higher temperatures reached during the heatwave, but rather from heatwaves becoming much more persistent. These heatwave storylines present extreme temperatures and moderately extreme to completely unprecedented daily mean excess heat levels sustained for up to almost 40 days in the most extreme cases (Fig. 3b and c).

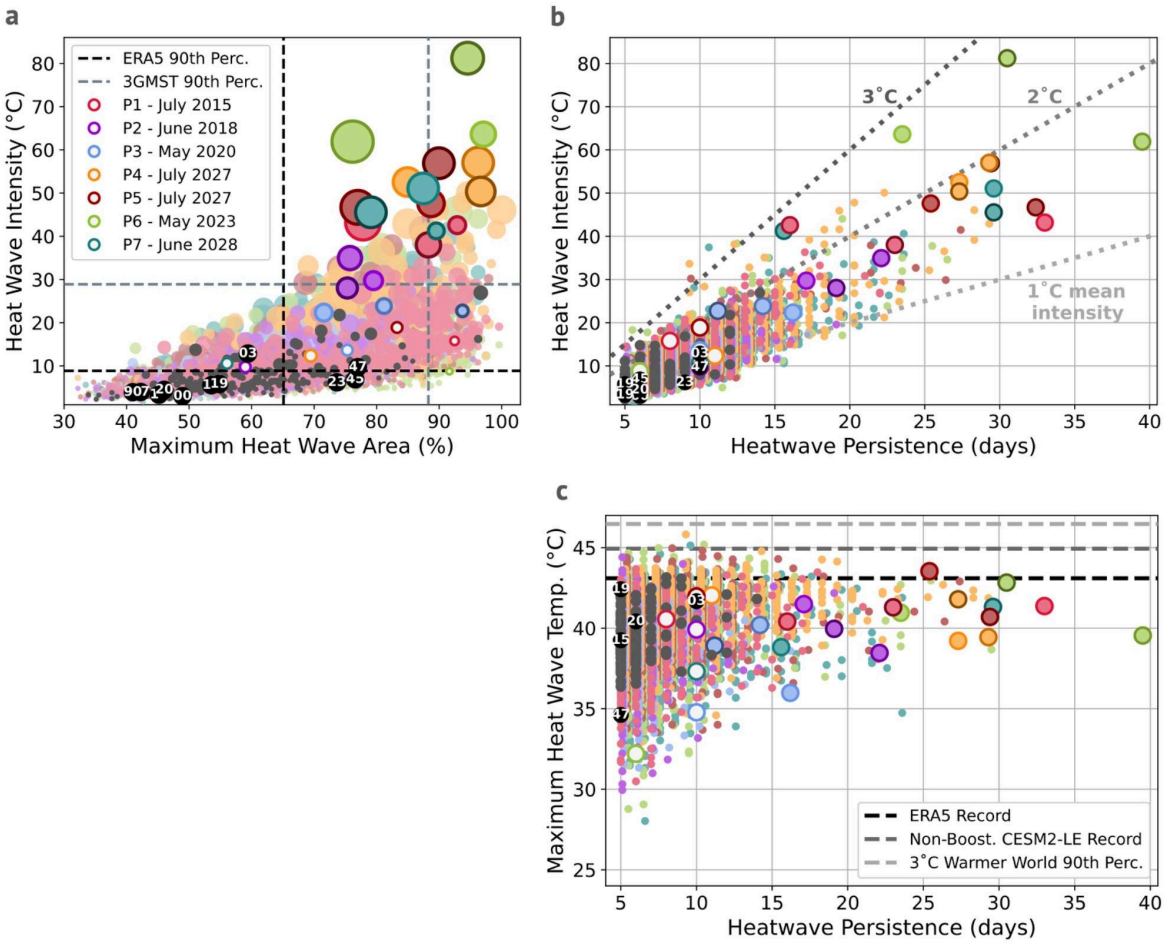



**Figure 3: Boosted vs. Parent Heatwaves. a)** Heatwave intensity against maximum area, same as Fig. 1, b) but now with boosted heatwaves (colored dots), highlighting the 3 most intense heatwaves for each parent case (light, medium and dark colored encircled dots), and including ERA5 (black dots), CESM2-LE (gray dots) and parent heatwaves (P1-7; colored, white infilled circles). Dashed lines mark the 90th percentile across heatwaves in ERA5 heatwaves (black; 2000-2023) and in 3˚C warmer world in CESM2-LE (gray). Dot size proportional to heatwave persistence. **b)** Same as a), but for heatwave intensity against persistence. Dotted lines mark daily mean heatwave intensities of 1˚C (gray), 2˚C (dark gray), 3˚C (black). **c)** Same a b) for maximum heatwave temperature against persistence. Dashed lines mark summer maximum temperature records for non-regridded ERA5 (black), CESM2-LE (gray) and 90$^{th}$ perc. levels for a 3˚C warmer world in CESM2-LE (light gray).

These results are in line with previous work indicating a potential physical limit as to how high temperatures may become in this climatic zone[30,42]. Furthermore, our finding of plausible heatwaves exceeding the most extreme observed levels by large margins is also well in line with previous work indicating that, statistically, the region of Central Europe is globally one of the most at risk of experiencing record-breaking heatwaves due to relatively low heat-levels in the current record[3]. We also underscore the importance of considering heatwave metrics that reflect heat persistence as well as intensity to capture worst-case extremes, such as the cumulative excess-heat metrics[4,43,44] used here, as opposed to limiting heat assessments to traditional metrics based on mean heat levels averaged over the length of the heatwave or over a period of fixed length.

In sum, the worst-case heat storylines presented in our work reveal a worrying picture: the most extreme heatwaves plausible today not only surpass the most extreme heatwave

intensity levels experienced in Europe in the recent past, they also exceed levels that would be considered extreme in a 3°C warmer world by large margins (Fig. 3a), and persist for weeks at a time (Fig. 3b). Furthermore, these unprecedented heat extremes tend to occur predominantly successively after another extreme heatwave (Fig. 4).

**Increasingly growing risk of unprecedented heatwaves happening successively**

An extreme heatwave happening after another one in the same extended summer season is rare, but not unprecedented. Among the sample of 11 ERA5 heatwaves between 1940-2023 and 225 non-boosted heatwaves in the 35-member CESM2-LE for 2010-2028, approximately 20% happen successively (Fig. 4). We find remarkably good agreement in the frequency of successive heatwaves between CESM2-LE and ERA5 (23.5 % successive heatwaves in CESM2-LE, 18.2 in ERA5). In the historical record, successive heatwaves occurred in 1947 and 2019 (Fig. 1e). Interestingly, the most extreme heatwave on record, in August 2003, also occurred successively after another heatwave in June (Fink et al., 2006). However, this June 2003 heatwave occurred mostly further east and south than the central Europe region considered here, and is thus not captured by our analysis.

Among the less extreme boosted heatwaves, the ones exhibiting intensities that are not unpreceded in the ERA5 or in the non-boosted CESM2-LE samples (which coincides roughly with boosted heatwaves below the TOP 300 and TOP 100 levels, respectively), slightly over 50% of them occur successively (Fig. 4a). This tendency towards more frequent successive heatwaves in the boosted sample may firstly arise because this sample consists of summers with conditions that favor heatwave development more than a randomly selected sample. However, exploring this tendency for boosted heatwaves of unprecedented intensity in the ERA5 and non-boosted CESM-LE samples shows that successive heatwaves occur increasingly more likely after another heatwave the higher the heatwave intensity, to the point that the TOP5 most intense boosted heatwaves occur all successively after another extreme heatwave (Fig. 4a).

The most extreme boosted heatwaves come predominantly from the same three parent cases: P4, P5 and P6. Cases P4 and P5 are perturbed during an ongoing heatwave, with P5 realizations starting during one of the most extreme heatwaves in the non-boosted ensemble; therefore, most of their boosted heatwaves are by design successive. However, this pattern of increasingly more intense heatwaves occurring successively with increasing likelihoods persists even when excluding P4 and P5 boosted heatwaves from the sample (Fig. 4b). Note that excluding P4 and P5 simulations removes roughly 5.000 heatwaves from the boosted sample, including the majority of the most extreme events. Without P4 and P5, the remaining top most extreme boosted heatwaves come from the driest parents, P1, P6 and P7. Additionally, excluding P4 and P5 heatwaves also results in a decrease in the frequency of successive heatwaves across the whole sample to roughly 35%, closer to the successive heatwave frequencies of roughly 20% found in ERA5 and non-boosted CESM2-LE simulations.

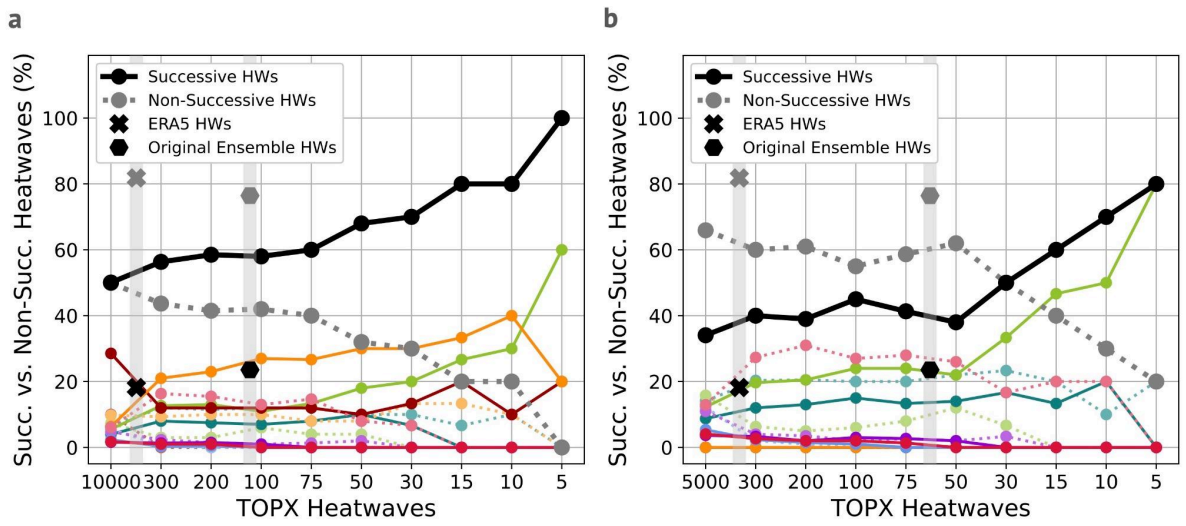



**Figure 4: Frequency of Successive and Non-Successive Heatwaves. a)** Percentage of top most extreme boosted heatwaves (TOP10.000 to TOP5 most intense) occurring successively (solid) versus non-successively (doted lines) for all parents (black and gray), and separated per parent (colors). Percentage of successive (black) and non-successive (gray) heatwaves in ERA5 (crosses; 1940-2023) and in non-boosted CESM2-LE (hexagons; 2010-2028); x-axis location corresponds with the maximum heatwave intensity in each sample during the noted periods. **b)** Same as a), but excluding P4 and P5 boosted heatwaves.

These findings highlight the risk of a potentially devastating high-impact hazard. While successive heatwaves are not unprecedented in the historical record, successive heatwaves of the magnitude presented in these boosted storylines would be indeed unprecedented. Such successive extremes and unprecedentedly long extreme heat periods amplify societal and environmental impacts by increasing experienced heat stress and limiting recovery time between events; intensifying heat stress impacts on populations, terrestrial and aquatic ecosystems, and infrastructure made more vulnerable by the first event[45–49]. Prolonged heat exposure exacerbates drought, further depleting soil moisture and leading to drier soil and atmospheric conditions, which in turn increases the likelihood of cascading impacts such as wildfires, crop and ecosystem damages, and energy system strain.

Therefore, it is crucial to understand the risk of these worst-case successive heatwaves and how they form and intensify. The question arises as to whether the development of successive heatwaves is intrinsically different to non-successive heatwaves. On the one hand, they may happen simply because the parent summer conditions are more prone to heatwave development (e.g., due to pre-existing drought, or sea surface temperature or atmospheric patterns that may favour specific weather systems), thus increasing heatwave occurrence and eventually leading to randomly-occurring extremely intense heatwaves. On the other hand, the occurrence of the first heatwave may affect the system, such as through the evaporation of local soil moisture reservoirs or the warming of nearby ocean basins, and may thus prime the system to develop more uniquely intense heat later on.

**Background conditions during successive vs. non-successive boosted heatwaves**

To explore these questions, we assess the difference in successive and non-successive heatwaves in terms of their intensity and the persistence of associated atmospheric patterns measured as the blocking of geopotential height at the 500 hPa level (z500), as well as the terrestrial water budget measured as soil moisture (SM) deficits and sea surface temperature anomalies (SST) before and after successive and non-successive heatwaves (Fig. 5).

The persistence of boosted successive and non-successive heatwaves relative to the persistence of their associated atmospheric blocking conditions reveals similar overall patterns, especially for less extreme heatwaves (Fig. 5a and b). The vast majority of heatwaves occur under atmospheric blocking and associated persistent anticyclonic conditions. However, the most intense and especially persistent successive heatwaves tend to last longer than their associated atmospheric blocking events. This is most marked for the most intense boosted heatwave, which extends well beyond the relatively short associated period of blocking (Fig. 5a). This indicates the presence of additional mechanisms during successive heatwaves beyond the persistence of atmospheric blocking that sustain extreme heat conditions even after large-scale blocking patterns begin to weaken, such as less-stable sequences of anticyclonic conditions not captured by the blocking metric used here or other thermodynamic mechanisms. In contrast, the most intense and persistent non-successive heatwaves exhibit a stronger coupling between heatwave duration and blocking persistence, indicating a stronger dependence on long-lived large-scale atmospheric blocking patterns (Fig. 5b).

Differences in pre- and post-heatwave moisture deficits reveal that successive heatwaves do not develop from systematically drier conditions than non-successive counterparts (Fig. 5c and d). Beyond successive heatwaves reaching higher intensities, one of the main differences between the two samples arises from non-successive P3 heatwaves starting from similar and abnormally wet states at the start of the initial boosting period (Fig. 5d), something that is rarely seen for successive heatwaves. Similarly, pre- and post-heatwave SST anomalies are also somewhat comparable between successive and non-successive heatwaves, and successive heatwaves do not develop from systematically warmer SST conditions (Fig. 5e and f). The strongest SST anomalies appear in P4 and P5 heatwaves, while the strongest SM deficits occur for P6 and P7 heatwaves, independently of their classification. However, many of these heatwaves under extreme moisture depletion or extremely warm concurrent SSTs do not reach extreme heat intensities, indicating that these factors are not characteristic of one type of heatwave, and they are not sufficient nor required to trigger worst-case heat.

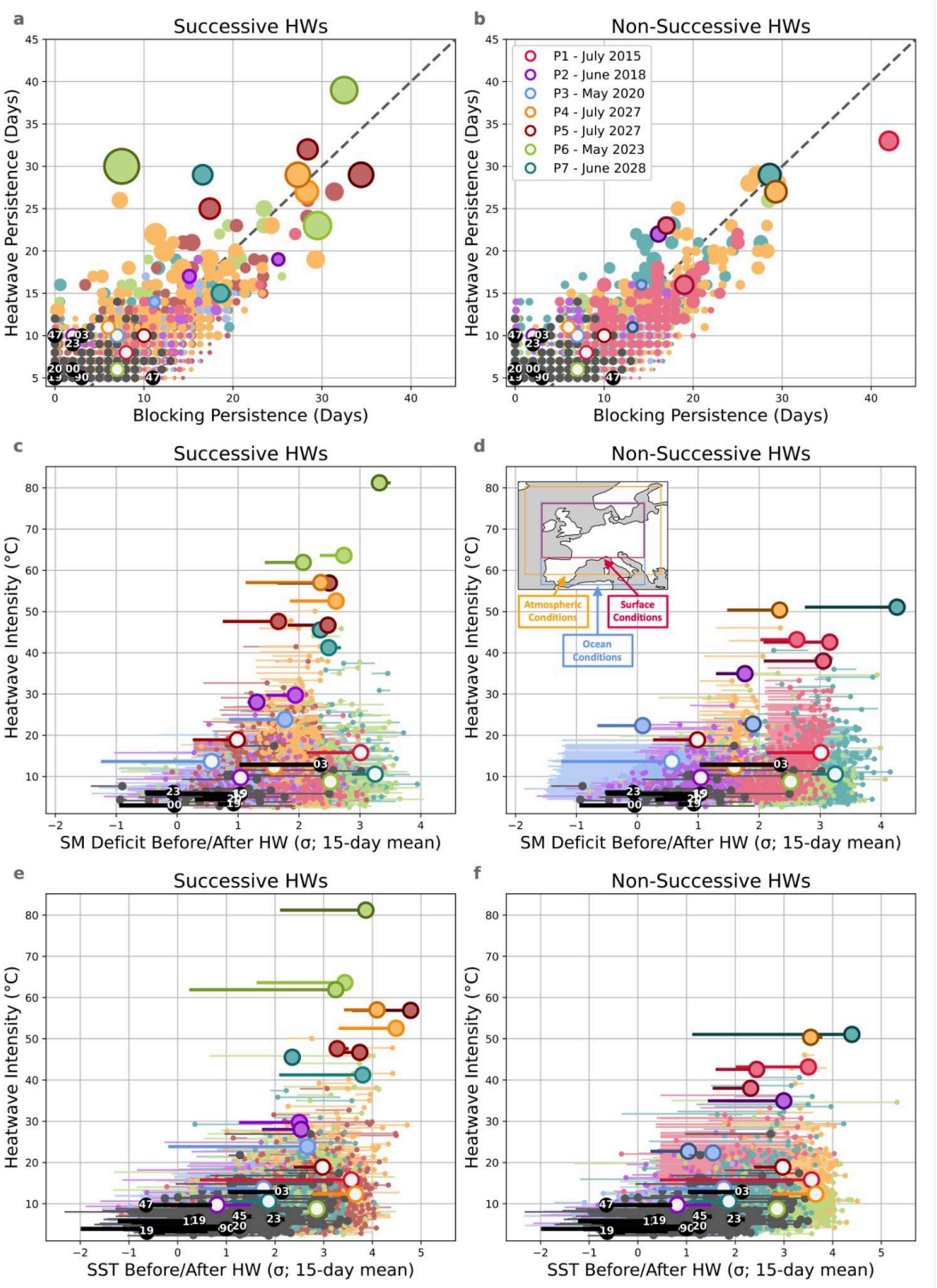



**Figure 5: Conditions during Successive vs. Non-Successive Boosted Heatwaves. a)** Heatwave (HW) persistence against atmospheric blocking persistence for successive heatwaves (colored dots; light, medium and dark colored encircled dots show 3 most intense heatwaves per parent), and ERA5 (black dots), CESM2-LE (gray dots) and parent heatwaves (P1-7; colored, white infilled circles). Dot size proportional to heatwave intensity; dashed line marks 1:1 slope. **b)** Same as a), but for non-successive boosted heatwaves. **c)** Heatwave intensity against 15-day mean standardized soil moisture (SM) deficit starting the day of the heatwave peak (dots) vs. the 15-day period before the peak (line start), for successive boosted heatwaves (colors), ERA5 (black) and CESM2-LE (gray). **d)** Same as c), but for non-successive boosted heatwaves. **e)** Same as c), but for standardized sea surface temperature (SST) anomalies. **f)** Same as e), but for non-successive boosted heatwaves. ERA5 (1940-2023; 1950-2023 for SM in ERA5-LAND) and CESM2-LE (2010-2028; 5 members only for SM) heatwaves are shown for comparison in both columns regardless of whether they occur successively or not. Values are spatially aggregated over the region marked by the red box for heat and SM metrics, over the yellow box for atmospheric metrics, and over the blue box for SSTs.

To disentangle how these factors contribute to the summers with the most intense heatwaves, we assess spatial differences in these conditions for the Top 50 most intense heatwaves (Fig. 6). Among the Top 50, 34 heatwaves occur successively after another heatwave, predominantly from P4, P5, P6 and P7 cases; while the remaining 16 heatwaves are either non-successive or the first heatwave of the season, and come from P1, P2, P4, P6 and P7 cases. Note however that the 4 non-successive P4 heatwaves consist of several weeks of unbroken heatwave conditions with several heat peaks (e.g., Fig. 2d), and may therefore not be generally exemplary of non-successive heatwave behaviour. Note also that to make signals comparable in time we assess these conditions for the July-August period exclusively.

Accumulated over the July-August period, simulations with successive heatwaves also exhibit more accumulated heat conditions that expand further north and particularly east than the region of study over central Europe (Fig. 6a), compared to their non-successive counterparts (Fig. 6b). July-August periods of successive heatwaves also exhibit generally higher Z500 anomalies over a larger area (Fig. 6c) than periods of non-successive heatwaves (Fig. 6d). In contrast, periods of both types of heatwaves exhibit comparable soil moisture deficit drought conditions (Fig. 6e and f). In terms of sea surface temperatures, we find large positive anomalies both along the Mediterranean as well as Atlantic basins in areas close to the continent, and negative anomalies westwards, behaviour that is more intense for composites of successive heatwaves (Fig. 6g) than non-successive ones (Fig. 6h).

These characteristics and the differences between successive and non-successive heatwaves vary widely for different parent cases. For example, the signal of excess heat spanning further north and east of the area of study arises predominantly from successive P5 heatwaves, and is less marked for all types of heatwaves stemming from all other parents (SI Fig. S2). This signal also matches the larger anticyclone spanning north and east of the region of study for P5 heatwaves (SI Fig. S3). Generally, simulations with successive heatwaves showcase stronger z500 anomalies than their non-successive counterparts across all parents, with the exception of P2 and P4 cases (SI Fig. S3).

For SM drought conditions, the cases that have both successive and non-successive TOP50 heatwaves (P4, P6 and P7) exhibit extreme yet similar drought patterns for both types of heatwaves (SI Fig. S4). However, some of these heatwaves (e.g., P1, P4, P5) start from moderate SM deficits that are exacerbated as the heatwaves proceed (Fig. 5c and Fig. 5d), yielding strong evaporation-driven drought conditions. The moderate pre-heatwave moisture depletion and the intensification of moisture deficits post-heatwave in these events indicates that part of the incoming energy is used for the evaporation of moisture, and that latent heat fluxes may somewhat dampen heat intensification in these events. In contrast, the most intense successive heatwaves in P6 and P7 cases start from more extreme SM deficits and exhibit minimal drying, with similar pre- and post-heatwave moisture deficits (Fig. 5c). Soil moisture levels appear already critically depleted before the onset of

the most intense P6 and P7 successive heatwaves, and the limited drying after the peak in heat indicates limited energy is used for moisture evaporation and latent heat fluxes, and therefore more energy is available to be transformed into sensible heat fluxes. This highlights the role that moisture limitation and temperature feedbacks play in the heat intensification for the most intense P6 and P7 successive heatwaves, as seen in the 2003 and other extremely dry heatwaves[7]. Lastly, a drought pattern feature that all TOP50 heatwave simulations have in common is the spread of drought conditions further north- and eastwards of the region of study, suggesting the role that both local and remote upwind drought may play in intensifying heat in central Europe, as seen during the 2003 heatwave[55,56].

SST anomaly patterns also show key differences across the different parent cases. Generally, simulations containing the TOP50 heatwaves all show extreme SST anomalies in the areas surrounding the continent (SI Fig. S5). Many of the less intense simulations, both for successive and non-successive heatwaves, develop these extreme SST conditions in the nearby basins at the same time as they develop heat over land (Fig. 5e and f). These simulations show heatwaves that start from neutral to moderately warm SST anomalies in the surrounding oceans, and which are mostly intensified as the heatwave proceeds. However, P4 and P5 heatwaves start from already extremely warm SST anomalies in the surrounding ocean basins (Fig. 5e and SI Fig. S5), which remain mostly unchanged by the heatwaves. This suggests that these nearby ocean areas could act as heat sources rather than sinks during these heatwaves. The presence of strong positive SSTs suggests enhanced sensible and latent heat fluxes into the lower troposphere and boundary layer. In turn, this process may raise atmospheric temperatures and increase heat advection from the nearby ocean basins towards land areas, strengthening the most intense P4 and P5 heatwaves as seen during the 2003[50] and 2023[51] heatwaves. Successive P6 and P7 heatwaves, although under somewhat less extreme SST anomalies, also exhibit this behaviour of starting from high SST anomalies that remain largely unchanged by the heatwaves (Fig. 5e and SI Fig. S5). This behaviour is not present in their non-successive counterparts (Fig. 5f), and points to heat accumulation in the nearby ocean basins due to the first heatwave as a key factor intensifying these successive heatwaves.

Furthermore, the majority of the TOP50 most intense heatwaves occur under concurrently negative SST anomalies over the North Atlantic further west from Europe. This behaviour is most marked for P4 heatwaves, which occur under a strong negative SST anomaly over the North Atlantic Subpolar Gyre region (SI Fig. S5), suggesting a potentially active role of hemispheric-scale oceanic temperature gradients and associated atmospheric processes for these events. The strong negative North Atlantic SSTs could indicate potential exacerbation of heatwave intensity and prsistency downstream, as seen during the extreme heat and drought conditions in 2015 and 2018 and their concurrent cold subpolar North Atlantic temperature anomalies[52,53], occurring over a region slightly northwards as that

found in our work. This cold subpolar temperature anomaly has been shown to cause a weakening and northward displacement of the jet stream, which in turn strengthens and increases the persistence of anticyclonic conditions over central Europe[52–54], in line with our finding of the P4 parent case producing large numbers of persistent heatwaves (Fig. 3 b and c) and anticyclonic conditions (Fig. 5a and b).

Our work reveals two key ingredients co-occurring with worst-case heat: extreme local moisture depletion, as seen particularly for successive P6 and P7 heatwaves, and extremely warm ocean surface temperatures in the nearby basins, seen particularly for P4 and P5 heatwaves. These extremely warm nearby ocean surfaces occur together with strong negative anomalies westward over the North Atlantic, particularly for P4 heatwaves. We explore potential mechanisms through which these extreme conditions in terrestrial moisture budgets and ocean temperatures can intensify the most extreme heatwaves, as seen in several extreme heatwaves in the historical record. However, neither condition seems to be strictly required nor sufficient to produce worst-case heat in the boosted simulations presented here. Many boosted heatwaves present similarly depleted soil moisture levels or extreme ocean surface temperatures as the top most extreme events, yet do not reach extreme heat intensities (Fig. 5). On the other hand, for the cases where the top most extreme successive and non-successive heatwaves can be directly compared (namely P6 and P7), surface heat accumulation in the nearby ocean basins and moisture depletion reach more extreme levels in successive events as compared to non-successive ones (Fig. 5, SI Fig. S4 and S5). While disentangling these different contributions and assessing the significance of individual processes in driving extreme successive heat needs further investigation, these findings highlight the role that early extreme heat may play in priming the system for more extreme heat later in the season.

Somewhat unsurprisingly, our results reveal that worst-case heat conditions arise from a complex interplay of persistent atmospheric patterns, pre-existing land and ocean conditions and internal variability, rather than from any singular driver. However, our results also reveal the risks of extreme unprecedented heatwaves persisting for unprecedentedly long periods, happening successively after one another, and occurring together with other system stressors such as extreme soil moisture deficits of extremely warm ocean temperatures. This underscores the importance of considering these sources of variability in driving worst-case extreme heat, drought and additional ecological stress, as well as the importance of considering the compound nature of these worst-case conditions for a comprehensive risk assessment.

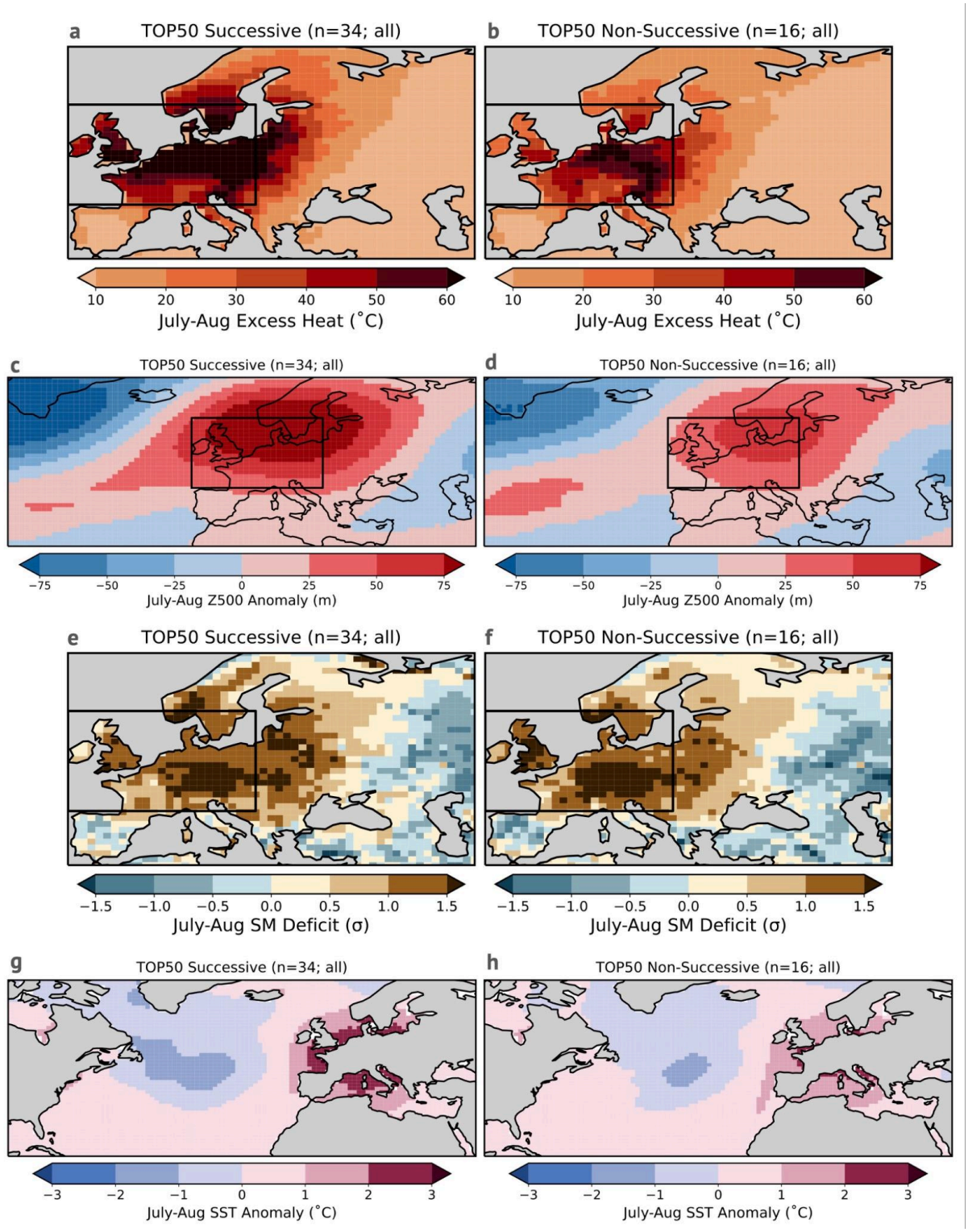



**Figure 6: Conditions during TOP50 Successive vs. Non-Successive Heatwaves. a-b)** Accumulated Excess Heat in July–August, composites for the TOP50 most extreme successive heatwaves (n=34; a) and non-successive (n = 16; b) across all parent cases. Composites for each parent separately can be found in SI Figs. S2-S5. The region used for spatially aggregated heatwave metrics is marked by the black box. **c-d)** Same as (a-b), for mean geopotential height anomalies at the 500 hPa level (z500). **e-f)** Same as (a-b), for mean soil moisture (SM) standardized deficits. **g-h)** Same as (a-b), for mean sea surface temperature (SST) anomalies.

**Worst-case compound heat and drought stress storylines**

Lastly, we explore worst-case compound storylines of how extreme heat occurs concurrently with other system stressors such soil drought and fire weather conditions driven by extreme atmospheric evaporative demand, defined here as vapour pressure deficit (Fig. 7).

Over the whole extended summer season, boosted storylines exhibit substantially greater variability in accumulated excess heat compared to soil moisture drought (Fig. 7a). The most severe drought conditions reach approximately 0.5 standard deviations beyond their parent events as well as most extreme ERA5 extremes, while accumulated excess heat ranges from 5 to 15 standard deviations higher in the most extreme boosted storylines also with respect to their parent conditions as well as to ERA5 extremes. The driest parents, P1, P6, and P7, produce the most extreme drought conditions, regardless of the extreme accumulated heat levels and associated evapotranspiration reached in other cases. This stems from the lower accumulated precipitation levels across the extended summer season in these parents (SI Fig. S6), combined with the wetter early summer soil moisture conditions in less dry parents (Fig. 1 d) that do not reach as extreme levels of moisture depletion due to evaporation. In particular, P1 and P7 parent simulations reflect the driest extended summer season across the whole CESM2-LE sample for the 2010-2028 time period. Several of the boosted simulations stemming from these two parents result in less soil drought over the May-September season than their parents, and those that reach more extreme drought levels exceed parent conditions by less than 1 standard deviation (Fig. 7a).

In contrast, boosted simulations with the most severe excess heat accumulated over the whole extended summer season stem from parents P4, P5, P6 and P7 (Fig. 7a). The parent simulations for P4 and P5, which reflect the hottest extended summer season across the whole CESM2-LE sample for the 2010-2028 time period, produce boosted simulations with accumulated heat levels over 5 standard deviations more extreme. The parent simulations for P6 and P7 exhibit heat levels comparable to the 2003 summer pre-boosting, which after boosting result in heat levels 10 to 15 standard deviations more extreme than their parents. In the most extreme compound heat and drought boosted storyline (P6, Fig. 7a), which also corresponds with the most extreme heatwave, extreme excess heat levels reach circa 120˚C of excess heat (i.e., accumulated degrees above the 95th percentile threshold) over the 153-day May-September period. Furthermore, parents P6 and P7, also produce the most severe compound heat and drought extremes (Fig. 7a). For these worst-case boosted compound heat and drought events, soil moisture deficits reach levels nearly twice as extreme as the driest historical droughts in 2003 and 2018, occurring concurrently with 4 to 6 times more accumulated excess heat, revealing the potential for more intense drought compounding with far more intense heat than ever recorded in the historical period over Central Europe.

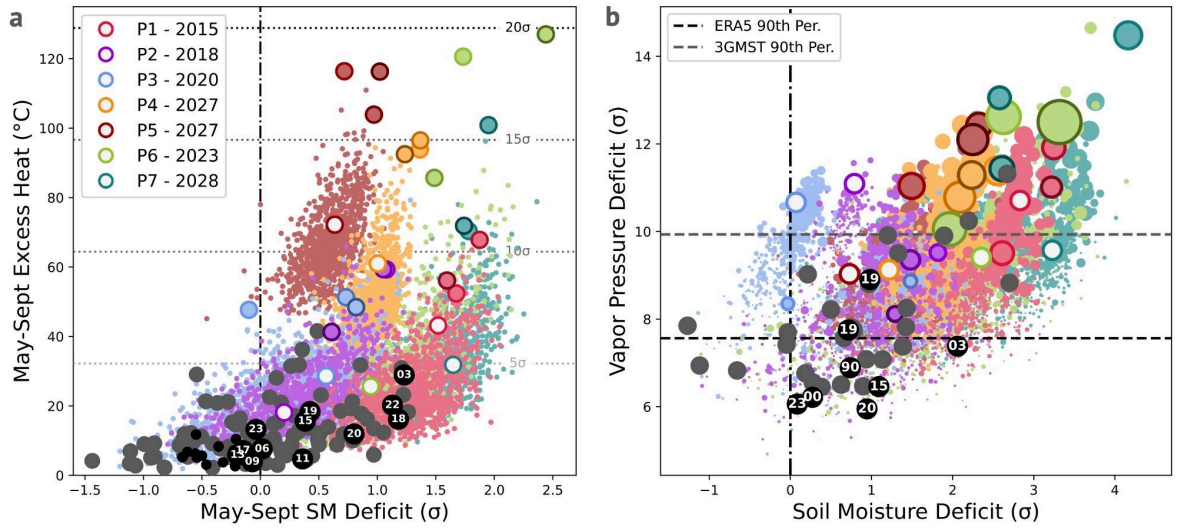



**Figure 7: Boosted Compound Heat and Soil and Atmospheric Drought Storylines. a)** Accumulated Excess Heat against Soil Moisture (SM) deficits over the extended summer season (May to September) for boosted storylines (colored dots), highlighting the storylines with the three most intense heatwaves for each parent case (light, medium and dark colored encircled dots), and including ERA5 (black dots), 5 last members of CESM2-LE (gray dots) and parent summers (P1-7; colored, white infilled circles). Dotted lines mark several excess heat standard deviation levels based on CESM2-LE data. **b)** Same as a) but for standardi Vapor Pressure Deficit (VPD) against Soil Moisture (SM) deficits during concurrent heatwaves. Deficits are averaged over a 15-day period centered around the day of peak heatwave intensity and standardized against all 15-day periods starting on the same day in ERA5, and in CESM2-LE for the model simulations. For boosted events, dot size is proportional to heatwave intensity. Dashed lines mark the 90th percentile VPD across all May-Sept 15-day periods in ERA5 (black; 2000-2023) and in 3˚C warmer world in CESM2-LE (gray).

Beyond compound heat and soil drought over the extended summer season, heatwaves occurring concurrently with extreme soil and extreme atmospheric evaporative demand exemplify shorter-term acute fire weather conditions that act as precursors to flash drought and wildfire risk[57,58]. Here, we show compound soil drought and atmospheric evaporative demand conditions, measured via vapor pressure deficits (VPD; see Methods), occurring concurrently with heatwave conditions (Fig. 7b). These fire weather conditions exhibit extreme levels during historical extreme heatwaves such as 2003. However, in contrast to the relatively moderate extreme heat levels during the two Central Europe 2019 heatwaves, these two events stand out as the two most extreme evaporative pressure extremes in the historical record over Central Europe, particularly in the first heatwave in July 2019. In comparison, boosted heatwaves reach SM deficit and VPD levels up to 2 and 4 standard deviations more extreme than the most extreme historical events, respectively. Boosted heatwaves with the most extreme concurrent SM and VDP extremes come from P6 and P7 cases, two of the driest parents, exemplifying again the limited atmospheric humidity due to the minimal evapotranspiration of soil moisture taking place during these events. Lastly, our findings also highlight that while boosted heatwaves from moderate parents exhibit heat and concurrent soil moisture deficits that remain relatively neutral, they also show potential to reach extreme concurrent fire weather conditions, with VPD levels over 10

standard deviations, in line with some of the most extreme heatwaves, which showcases additional risks from compound events traditionally considered moderate.

Our findings expose the potential for unprecedented compound heat, fire weather and soil drought conditions well beyond historical extremes in the recent past, and well beyond conditions that would be considered extreme even in a world 3˚C warmer than preindustrial levels. These results underscore the need for improved risk assessments and adaptation strategies that capture worst-case conditions plausible under current climate conditions, to mitigate potential unprecedented or cascading impacts on society, agriculture, critical infrastructure and ecosystems, especially over wildfire-prone regions.

## Discussion

Our results show that worst-case heat and drought extremes potentially plausible under today's climate far exceed the most extreme events in the historical record over Europe. Worst-case boosted storylines depict soil moisture deficits nearly twice as severe as those observed in 2003 and 2018, occurring concurrently with extreme accumulated heat levels four to six times more intense. These findings reveal the potential for more intense soil and atmospheric drought compounding with far more extreme heat. Heatwave intensities during these extreme boosted storylines not only exceed historical levels, they also exceed levels that would be considered extreme in a world 3˚C warmer than a preindustrial climate by large margins of more than 50˚C of accumulated heatwave intensity. Beyond reaching extreme intensities, these heatwaves also exhibit extreme persistence, lasting for weeks at a time, and present one additional risk: they tend to occur predominantly successively after another extreme heatwave.

These findings highlight the risk of a potentially devastating high-impact hazard. While successive heatwaves are not unprecedented in the historical record, occurring for example in July and September of 2019, successive heatwaves of the magnitude and especially persistence found in these boosted storylines would be extremely unprecedented. Such successive extremes and unprecedentedly long extreme heat periods amplify societal and environmental impacts by limiting recovery time between events; intensifying heat stress on populations, terrestrial and aquatic ecosystems, and infrastructure made more vulnerable by the first event[45–49]. For example, prolonged heat exposure and compound drought are associated with water availability and quality risks. Drought-driven reduced water levels in rivers and lakes may compound with rising water temperatures and associated decrease in oxygen levels, which may lead to massive fish mortality events, algae blooms and further aquatic ecosystem disruptions, as experienced in Europe in 2018[59,60]. Furthermore, accumulated heat stress exposure of levels well beyond what the local European population has experienced in their lifetimes would yield worsened health outcomes and heat-related mortality and morbidity, as well as labour productivity losses[49,61]. To compensate the effects of heat stress, society may respond with increased use of air

conditioning, which may result in energy price spikes as it compounds with further drought-related strains to the energy production system, such as river-based fuel shipping restrictions and the shutting down of nuclear power plants due of insufficient water for cooling, as experienced in 2018 in Europe[59,62]. The impacts resulting from these heat and drought stotylines may also span beyond those previously experienced, as surpassing previously unprecedented climatic conditions may lead to the exceedance of certain safety thresholds or tip systems into further disruptions, leading to previously unseen compounding or cascading impacts across different humans and environmental systems.

In terms of the driving mechanisms behind the most extreme heat and drought conditions, the picture is also complex. The most extreme droughts arise from a combination of extreme early-summer moisture depletion, precipitation deficits and extreme heat levels experienced over the extended summer season. However, extreme heat and associated evapotranspiration is not sufficient to reach the most extreme soil moisture depletion levels, with early summer drought arising as a necessary factor, at least in the model world of CESM2. The most extreme heatwaves in these boosted storylines arise in contrast from an interplay of persistent atmospheric patterns, pre-existing extreme soil moisture and ocean surface heat, and internal variability, with neither factor being strictly necessary nor sufficient for the development of worst-case heat. While some of the most intense heatwaves occur under concurrent severe soil drought, others develop under relatively neutral drought conditions yet under exceptionally warm ocean temperatures and strong negative anomalies in the North Atlantic subpolar gyre region.

However, one aspect the vast majority of the most intense heatwaves have in common is that they occur successively after another extreme heatwave. This suggests the first event may prime the system for even more extreme heat to follow, either through soil moisture depletion or through the heat accumulation in nearby ocean basins, or a combination of these and potentially other factors. The extent to which the first heatwave actively leads to the intensification of successive heatwaves, and how this may compound with other factors making these summers more prone to extreme heat successions, such as quasi-stationary atmospheric wave patterns as those found in simulations targeting the development of extreme mean summer temperatures[28], requires further investigation and entails critical implications for understanding and anticipating the most devastating heatwave sequences.

Our findings highlight the diversity of pathways to worst-case heat and drought, and the importance of sampling the full range of climate variability sources in the coupled climate system to robustly explore worst-case conditions and their exacerbating factors. Furthermore, even though the worst-case storylines presented here involve combinations of factors that may appear rare or even unlikely, past unprecedented extremes such as the 2021 Pacific North West heatwave demonstrate that events once deemed extremely rare to virtually impossible can happen, especially as the world continues to warm[3,4,11,16]. Given the scale, intensity, and unprecedented nature of these heat and drought storylines, both

individually and in their potential for successive and compounding stressors, we underscore the urgent need for well-informed adaptation and resilience strategies that account for these compounding and cascading risks, to sufficiently protect vulnerable populations, critical infrastructure and ecosystems.

# Methods

### Observational and model data

The model simulations used in this study are based on a 35-member ensemble of the Community Earth System Model version 2.1.2 (CESM2.1.2[36]). These simulations span from 2005 to 2035 under historical (2005-2014) and SSP3-7.0 (2015-2035) forcing conditions. This set of simulations has been used in previous work using ensemble boosting[11,31], and share the same initial conditions from a transient historical simulation with a random round-off perturbation in atmospheric conditions at year 2005. The model spread in atmospheric conditions saturates after several months, while the upper ocean and soil levels might not be fully independent until after a few years[11]. For this reason, we constrain our analysis here to simulated years after 2010. Additionally, for comparisons with future conditions under further global warming, we use a 100-member CESM2 large ensemble[36] covering the period from 1850–2100 also under historical and SSP3-7.0 forcing conditions.

Reanalysis data from the ERA5[38] and, for soil moisture, ERA5-Land[63] datasets are used for the period of 1940-2023 and 1950-2023, respectively, for the evaluation of the CESM2 simulations and for comparison with heat and drought conditions over the historical record. Both reanalysis datasets are regridded to the coarser resolution of CESM2 using conservative remapping Climate Data Operators (CDO[64]).

### Ensemble boosting setup

Ensemble boosting is performed by re-initializing CESM2 between about 10 to 30 days (referred to as boost date) before the peak in excess heat conditions in the corresponding parent events (except for cases P4 and P5 for which this boost dates are shifted to cover the onset of the parent heatwave). For this technique, bit-by-bit reproducibility on the high-performance computing environment where the runs are performed is required, ensuring that parent conditions occurring in an existing simulation can be exactly reproduced. For this reason, we base these storylines on the 35-member CESM2-LE performed in the same computing environment as the boosted ensembles.

The *boost* perturbation involved in this technique consist exclusively of an instantaneous random perturbation to the specific humidity field at each grid point in the order of $10^{-13}$. The global average of the perturbation is equal to zero, and designed to be as small as possible to ensure that mass, energy and momentum are conserved up to the precision of a round-off error. This ensures that the simulations evolve under different trajectories while maintaining full physical consistency across time, space, and variables. These perturbations

are introduced to produce sets of 50 members each, one for each of the 20 consecutive boost dates in the selected 20-day boosting window, yielding a thousand-member ensemble in total for each parent (similar to the setup in Fischer et al., 2023). After this initial perturbation the fully coupled model is run freely for the whole extended summer season until the end of October (further details on this experimental setup and the effect of the initial perturbation can be seen in ref. [11] ).

**Excess Heat and heat-related metrics and definitions**

Heat-related metrics are based on daily Excess Heat[4,43]. Excess Heat is defined, for each grid cell and day, as the difference between the actual daily maximum 2m air temperature minus the given heat threshold for said grid cell and day. We consider Excess Heat over land grid cells only, and only for days with positive Excess Heat anomalies. Days when this difference yields a negative value are considered to exhibit no excess heat, and thus null, to not confound this Excess Heat metric with potential negative values elsewhere in the region of study. Heat thresholds are defined as the 95$^{th}$ percentile in daily maximum 2m air temperature for each day and grid cell, pooled over a 15-day window centered on each day, for the period of 2005-2023 to allow a comparison between CESM2-LE and ERA5. These Excess Heat daily values are then summed, or accumulated, either over the whole extended summer season defined as the period of May to September, or for the duration of the heatwave for heatwave intensity, and averaged over all land grid cells in the Central European region defined by the [44–60 N, 10W–20E] latitude-longitude domain.

We choose cumulative heat metrics based on exceedances above thresholds specific to the region and time of the year to better capture the effect of both intensity and persistence of the extreme conditions and associated adverse impacts[4,43]. While different heatwave measures based on daily or fixed-period average intensities are most commonly used, metrics based on cumulative values of exceedances above a given threshold offer a more unequivocal approach and a better comparison of events of different length. Furthermore, heat stress impacts are primarily determined by the accumulation of heat levels that exceed a certain level experienced over a specific time period, rather than by averaged values[43,44]. To give additional context to these cumulative heat metrics and make them more comparable to traditional average-based measures, we also include daily mean heatwave intensity values, defined as the accumulated Excess Heat intensity summed across the whole duration of the heatwave, divided by the duration of the heatwave in days.

Furthermore, we define heatwaves as periods of a minimum duration of 5 days where at least 30% of the land grid cells of the Central Europe region of study exhibit positive Excess Heat values. Heatwaves are broken by 3 consecutive days where these conditions are not fulfilled, period after which a new heatwave may start. If these heatwave conditions occur, but then stop being fulfilled for a period of 1-2 days and continue to be fulfilled after, the heatwave continues. In this case, the accumulated heatwave intensity is summed across all days since the first heatwave day; but heatwave persistence reflects only days within the

heatwave period that fulfill these heatwave conditions. The timing of the peak heat intensity within the heatwave is defined as the day of maximum daily Excess Heat averaged over the region of study. Maximum heatwave temperatures are the maxima over the whole region and heatwave period.

Successive heatwaves are defined as heatwaves preceded by at least one additional heatwave in the May-Sept. extended summer season; while non-successive heatwaves are defined as either the first or the only heatwave in the May-Sept. extended summer season.

**Soil and atmospheric drought metrics and definitions**

We use soil moisture (SM) and accumulated precipitation deficits to illustrate soil drought conditions. SM deficits are based on daily SM values vertically accumulated over a soil layer spanning roughly over the first meter of soil depth. This upper soil layer corresponds to the first 3 layers in ERA5-Land (corresponding with the first 100 cm of soil[63]) and to the first 7 layers in CESM2 (corresponding with the first 92 cm of soil in the land component of CESM2, CLM5[65]). SM Deficits are calculated as the daily difference with respect to the daily SM climatological reference, averaged for each day and grid cell for the period of 2005-2023. These SM deficits are then averaged in time and in space across all land grid cells in the Central Europe region. SM deficits are then standardized by dividing them by their standard deviations in the period of 2005-2023. Increasingly large positive deficits reflect increasingly lower soil moisture levels and increasingly drier conditions.

To sufficiently capture soil moisture variability as heatwave conditions develop, this analysis requires daily soil moisture data. This level of temporal resolution is only available for the last 5 members of the 35-member CESM2-LE ensemble, as well as for all boosted simulations and for the seven parent simulations. Due to this limitation, for all simulated SM values we use a climatological reference defined as the daily mean across all 5 members available, for the period of 2005-2023. These multi-year averages exhibit only minor differences across individual members (not shown), for this reason we believe this limitation in the calculation of a climatological reference does not largely affect our conclusions.

To explore the effect of ensemble size on soil drought variability and give further context to these SM deficits, we also use accumulated precipitation deficits to illustrate soil drought conditions. These precipitation deficits, available for all simulations used, are calculated based on daily precipitation accumulated over the period of May to September. Deficits are then defined as the difference between the accumulated precipitation and the climatological reference accumulated over the same period of May to September and averaged for the period of 2005-2023. These precipitation deficits are then averaged over all land grid cells in the Central Europe region of study and standardized. Similarly as with SM deficits, increasingly large positive deficits reflect increasingly lower accumulated precipitation levels and increasingly drier conditions.

Lastly, we use Vapor Pressure Deficit (VPD) to illustrate atmospheric drought conditions. VPD represents the difference between the saturation vapor pressure and the actual vapor pressure in the atmosphere, illustrating atmospheric evaporative demand and heightened wildfire risk conditions commonly referred to as fire weather[66,67]. We calculate daily VPD from daily mean 2m air temperature and relative humidity using the following expressions for the VPD[68] and for the saturated vapor pressure[69]:

$$VPD = e_s - e_a = (1 - RH/100) \times e_s,$$

$$e_s = 6.1078 \times \exp((17.2694 \times T)/(237.3+T))$$

Where $e_s$ is the saturation vapor pressure, $e_a$ is the actual vapor pressure, T is the surface air temperature in ˚C, and RH is the relative humidity in percentage. All soil and atmospheric drought metrics are calculated over land-grid cells only over the same spatial region used to define heat metrics representing Central Europe and defined by the [44–60 N, 10W–20E] latitude-longitude domain.

**Atmospheric patterns and persistence metrics and definitions**

Atmospheric conditions are illustrated by geopotential height anomalies at the 500 hPa level (z500). These values are presented as daily mean anomalies with respect to the 2005-2023 climatological reference. Spatially aggregated anomalies are calculated over all grid cells of a region covering a slightly larger area as the Central Europe region of study, defined by the [39–65 N, 15W–25E] latitude-longitude domain, to ensure that relevant atmospheric patterns not directly collocated over the peak in surface heat conditions are captured.

We assess the persistence of atmospheric patterns using an atmospheric blocking metric based on z500 anomalies. For this metric calculation, anomalies of the z500 field are calculated after a 2-day running mean temporal smoothing and relative to the 31-day running mean climatology in the period 2005-2034 and across the 35 members of CESM2-LE. Using the ConTrack-package[70] contours of anticyclones are tracked along their trajectory. The anticyclones are identified as z500 anomalies above the 90th percentile of the z500 anomaly distribution in the mid-latitudes (50-80°N). If the movement of an anticyclone fulfills the following blocking criteria, it is classified as "blocked": a) the contours have a spatial overlap of at least 50% and, b) the anticyclone persists for at least five days. A limitation of this type of blocking index is that it does not consider column-averaged potential vorticity and does not require a reversal of the zonal flow. It is instead based on closed contours of the 500hPa geopotential height level[71]. This means it can potentially classify strong, persistent ridges as blocking even if the jet stream remains mostly intact. In a next step, blocking persistence is counted as the sum of days fulfilling these conditions over at least 10% of the grid cells in the larger Central Europe area defined by the [39–65 N, 15W–25E] latitude-longitude domain, for the duration of the heatwave as

well as the three days before and after the heatwave. Note that the smaller fraction of grid cells used as threshold (10%, ~ 90 grid cells) in this case compared to the heatwave area threshold (30%, ~70 grid cells) is to account for the larger area considered as well as considering both land and ocean grid cells used for spatially aggregated atmospheric metrics. Slightly higher or lower spatial thresholds (not shown) do not substantially affect our conclusions.

**Ocean surface temperature metrics and definitions**

To represent ocean surface temperature conditions, we use a combination of sea surface and 2m air temperature values over ocean-only grid cells. These values are presented as anomalies with respect to the 2005-2023 climatological reference. Strictly defined ocean surface temperatures are not available for the boosted simulations used here. Due to this limitation, for all comparisons involving boosted simulated data, sea surface temperature (SST) is represented by 2m air temperatures over ocean grid cells. A comparison between strictly defined SSTs and sea surface 2m air temperatures (SSATs) for the 35-member CESM2-LE and ERA5 data (SI Fig. S7) reveals that these two metrics are highly related, with the largest differences among the two metrics around ±0.5˚C. Absolute SST values are slightly lower than their SSAT counterparts, both for strong negative and positive anomalies. This indicates that although the proxy SSAT values presented in the main analysis may somewhat overestimate absolute anomalies for strictly defined ocean surface temperatures, they offer a reasonable representation of strictly defined SSTs. All spatially aggregated ocean metrics averages over ocean-only grid cells in the [36–60 N, 10W–20E] latitude-longitude domain.

**Metric standardization**

To account for potential biases in the representation of the mean state and variability in the distribution of heat, drought and other related climatic conditions in CESM2-LE, we base our analysis almost exclusively on high-percentile threshold exceedances or standardized metrics, particularly when comparing simulated and observed events. For all metrics used in this study, standardized values are obtained by dividing by the metric's standard deviation calculated in the period of 2000-2023 for ERA5 and 2005-2023 for for the available CESM2-LE simulations, and always based on quantities averaged or accumulated over the same period length and starting at the same time of the year to account for potential intra-seasonal variability changes.

**Comparison to conditions under 3˚C of global warming**

To assess how the current climate extreme storylines presented here compare with conditions under further global warming, we compare them with extreme conditions typical of a climate 3˚C warmer than the preindustrial climate in CESM2-LE. For this analysis we use exclusively the 100-member CESM2-LE simulations covering the period 1850-2100 under historical and SSP3-7.0 forcing conditions. This 3˚C warmer world climate is illustrated by a 15-year period around the first year where the 100-member CESM2-LE

ensemble mean global mean surface 2m air temperature (GMST) anomalies exceed 3˚C above preindustrial conditions, defined as the average GMST in the 1850-1875 period. The first year where the ensemble mean surpasses this 3˚C anomaly is 2069, and thus the period reflecting the 3˚C warmer world climate is defined as 2061-2076.

All metrics presented for this 3˚C warmer world climate are calculated with respect to the same thresholds, climatological references, and standardization periods as the metrics for current climate conditions. For reference, according to these definitions, the current climate period represented here by 2010-2028 corresponds in CESM2-LE to a world roughly 1˚C warmer than preindustrial conditions, with an average GMST anomaly of 1.17˚C above preindustrial levels. The 35-member CESM2-LE simulations use in the main analysis yield over this 19-year period a total of 665 summers, while the 15-year period defining a 3˚C warmer world in the 100-member CESM2-LE ensemble yields a sample of 1500 summers.

## References

1. S. I. Seneviratne *et al.* Weather and Climate Extreme Events in a Changing Climate. in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* 1513–1766 (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021).
2. De Luca, P. & Donat, M. G. Projected changes in hot, dry, and compound hot-dry extremes over global land regions. *Geophys. Res. Lett.* **50**, e2022GL102493 (2023) doi: 10.1029/2022GL102493.
3. Thompson, V. *et al.* The most at-risk regions in the world for high-impact heatwaves. *Nat. Commun.* **14**, 2152 (2023) doi: 10.1038/s41467-023-37554-1.
4. Suarez-Gutierrez, L., Müller, W. A. & Marotzke, J. Extreme heat and drought typical of an end-of-century climate could occur over Europe soon and repeatedly. *Commun. Earth Environ.* **4**, 415 (2023) doi:

10.1038/s43247-023-01075-y.

5. Schär, C. *et al.* The role of increasing temperature variability in European summer heatwaves. *Nature* **427**, 332–336 (2004) doi: 10.1038/nature02300.
6. Fink, A. H. *et al.* The 2003 European summer heatwaves and drought –synoptic diagnosis and impacts. *Weather* **59**, 209–216 (2004) doi: 10.1256/wea.73.04.
7. Miralles, D. G., Teuling, A. J., van Heerwaarden, C. C. & Vilà-Guerau de Arellano, J. Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nat. Geosci.* **7**, 345–349 (2014) doi: 10.1038/ngeo2141.
8. Toreti, A. *et al.* The Exceptional 2018 European Water Seesaw Calls for Action on Adaptation. *Earths Future* **7**, 652–663 (2019) doi: 10.1029/2019EF001170.
9. Rousi, E. *et al.* The extremely hot and dry 2018 summer in central and northern Europe from a multi-faceted weather and climate perspective. *Nat. Hazards Earth Syst. Sci.* **23**, 1699–1718 (2023) doi: 10.5194/nhess-23-1699-2023.
10. White, R. H. *et al.* The unprecedented Pacific Northwest heatwave of June 2021. *Nat. Commun.* **14**, 727 (2023) doi: 10.1038/s41467-023-36289-3.
11. Fischer, E. M. *et al.* Storylines for unprecedented heatwaves based on ensemble boosting. *Nat. Commun.* **14**, 4643 (2023) doi: 10.1038/s41467-023-40112-4.
12. Vautard, R., Honore, C., Beekmann, M. & Rouil, L. Simulation of ozone during the August 2003 heat wave and emission control scenarios. *Atmos. Environ.* **39**, 2957–2967 (2005) doi: 10.1016/j.atmosenv.2005.01.039.
13. Robine, J.-M. *et al.* Death toll exceeded 70,000 in Europe during the summer of

2003. *C. R. Biol.* **331**, 171–178 (2007) doi: 10.1016/j.crvi.2007.12.001.

14. Knutzen, F. *et al.* Impacts on and damage to European forests from the 2018–2022 heat and drought events. *Nat. Hazards Earth Syst. Sci.* **25**, 77–117 (2025) doi: 10.5194/nhess-25-77-2025.
15. Zscheischler, J. *et al.* A typology of compound weather and climate events. *Nat. Rev. Earth Environ.* **1**, 333–347 (2020) doi: 10.1038/s43017-020-0060-z.
16. Fischer, E. M., Sippel, S. & Knutti, R. Increasing probability of record-shattering climate extremes. *Nat. Clim. Change* **11**, 689–695 (2021) doi: 10.1038/s41558-021-01092-9.
17. Sillmann, J. *et al.* Event-Based Storylines to Address Climate Risk. *Earths Future* **9**, e2020EF001783 (2021) doi: 10.1029/2020EF001783.
18. Kelder, T. *et al.* How to stop being surprised by unprecedented weather. *Nat. Commun.* **16**, 2382 (2025) doi: 10.1038/s41467-025-57450-0.
19. Shepherd, T. G. *et al.* Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. *Clim. Change* **151**, 555–571 (2018) doi: 10.1007/s10584-018-2317-9.
20. Zwiers, F. W., Zhang, X. & Feng, Y. Anthropogenic influence on long return period daily temperature extremes at regional scales. *J. Clim.* **24**, 881–892 (2011) doi: 10.1175/2010JCLI3908.1.
21. Yiou, P. & Jézéquel, A. Simulation of extreme heat waves with empirical importance sampling. *Geosci. Model Dev.* **13**, 763–781 (2020) doi: 10.5194/gmd-13-763-2020.

22. Yiou, P. *et al.* Ensembles of climate simulations to anticipate worst case heatwaves during the Paris 2024 Olympics. *Npj Clim. Atmospheric Sci.* **6**, 188 (2023) doi: 10.1038/s41612-023-00500-5.

23. Van Der Wiel, K., Lenderink, G. & De Vries, H. Physical storylines of future European drought events like 2018 based on ensemble climate modelling. *Weather Clim. Extrem.* **33**, 100350 (2021) doi: 10.1016/j.wace.2021.100350.

24. Thompson, V. *et al.* Risk and dynamics of unprecedented hot months in South East China. *Clim. Dyn.* **52**, 2585–2596 (2019) doi: 10.1007/s00382-018-4281-5.

25. Kelder, T. *et al.* Using UNSEEN trends to detect decadal changes in 100-year precipitation extremes. *Npj Clim. Atmospheric Sci.* **3**, 47 (2020) doi: 10.1038/s41612-020-00149-4.

26. Lucarini, V., Melinda Galfi, V., Riboldi, J. & Messori, G. Typicality of the 2021 Western North America summer heatwave. *Environ. Res. Lett.* **18**, 015004 (2023) doi: 10.1088/1748-9326/acab77.

27. Ragone, F., Wouters, J. & Bouchet, F. Computation of extreme heat waves in climate models using a large deviation algorithm. *Proc. Natl. Acad. Sci.* **115**, 24–29 (2018) doi: 10.1073/pnas.1712645115.

28. Ragone, F. & Bouchet, F. Rare event algorithm study of extreme warm summers and heatwaves over europe. *Geophys. Res. Lett.* **48**, e2020GL091197 (2021) doi: 10.1029/2020GL091197.

29. Vautard, R. *et al.* Attribution of human-induced dynamical and thermodynamical contributions in extreme weather events. *Environ. Res. Lett.* **11**, 114009 (2016)

doi: 10.1088/1748-9326/11/11/114009.

30. Noyelle, R., Zhang, Y., Yiou, P. & Faranda, D. Maximal reachable temperatures for Western Europe in current climate. *Environ. Res. Lett.* **18**, 094061 (2023) doi: 10.1088/1748-9326/acf679.
31. Sippel, S. *et al.* Could an extremely cold central European winter such as 1963 happen again despite climate change? *Weather Clim. Dyn.* **5**, 943–957 (2024) doi: 10.5194/wcd-5-943-2024.
32. Van Garderen, L., Feser, F. & Shepherd, T. G. A methodology for attributing the role of climate change in extreme events: a global spectrally nudged storyline. *Nat. Hazards Earth Syst. Sci.* **21**, 171–186 (2021) doi: 10.5194/nhess-21-171-2021.
33. Sánchez-Benítez, A., Goessling, H., Pithan, F., Semmler, T. & Jung, T. The july 2019 european heat wave in a warmer climate: storyline scenarios with a coupled model using spectral nudging. *J. Clim.* **35**, 2373–2390 (2022) doi: 10.1175/JCLI-D-21-0573.1.
34. Gessner, C., Fischer, E. M., Beyerle, U. & Knutti, R. Very rare heat extremes: quantifying and understanding using ensemble re-initialization. *J. Clim.* 1–46 (2021) doi:10.1175/JCLI-D-20-0916.1 doi: 10.1175/JCLI-D-20-0916.1.
35. Gessner, C., Fischer, E. M., Beyerle, U. & Knutti, R. Multi-year drought storylines for Europe and North America from an iteratively perturbed global climate model. *Weather Clim. Extrem.* **38**, 100512 (2022) doi: 10.1016/j.wace.2022.100512.

36. Rodgers, K. B. *et al.* Ubiquity of human-induced changes in climate variability. *Earth Syst. Dyn.* **12**, 1393–1411 (2021) doi: 10.5194/esd-12-1393-2021.

37. Bevacqua, E. *et al.* Advancing research on compound weather and climate events via large ensemble model simulations. *Nat. Commun.* **14**, 2145 (2023) doi: 10.1038/s41467-023-37847-5.

38. Hersbach, H. *et al.* The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020) doi: 10.1002/qj.3803.

39. Lorenz, E. N. The predictability of a flow which possesses many scales of motion. *Tellus Dyn. Meteorol. Oceanogr.* **21**, (1969) doi: 10.3402/tellusa.v21i3.10086.

40. Zhang, F. *et al.* What Is the Predictability Limit of Midlatitude Weather? *J. Atmospheric Sci.* **76**, 1077–1091 (2019) doi: 10.1175/JAS-D-18-0269.1.

41. Domeisen, D. I. V. *et al.* Prediction and projection of heatwaves. *Nat. Rev. Earth Environ.* **4**, 36–50 (2023) doi: 10.1038/s43017-022-00371-z.

42. Zhang, Y. & Boos, W. R. An upper bound for extreme temperatures over midlatitude land. *Proc. Natl. Acad. Sci.* **120**, e2215278120 (2023) doi: 10.1073/pnas.2215278120.

43. Perkins-Kirkpatrick, S. E. & Lewis, S. C. Increasing trends in regional heatwaves. *Nat. Commun.* **11**, 3357 (2020) doi: 10.1038/s41467-020-16970-7.

44. Russo, E. & Domeisen, D. I. V. Increasing intensity of extreme heatwaves: the crucial role of metrics. *Geophys. Res. Lett.* **50**, e2023GL103540 (2023) doi: 10.1029/2023GL103540.

45. Bérard, A., Bouchet, T., Sévenier, G., Pablo, A. L. & Gros, R. Resilience of soil

microbial communities impacted by severe drought and high temperature in the context of Mediterranean heat waves. *Eur. J. Soil Biol.* **47**, 333–342 (2011) doi: 10.1016/j.ejsobi.2011.08.004.

46. Amengual, A. *et al.* Projections of heat waves with high impact on human health in Europe. *Glob. Planet. Change* **119**, 71–84 (2014) doi: 10.1016/j.gloplacha.2014.05.006.
47. Błażejczyk, K., Twardosz, R., Wałach, P., Czarnecka, K. & Błażejczyk, A. Heat strain and mortality effects of prolonged central European heat wave—an example of June 2019 in Poland. *Int. J. Biometeorol.* **66**, 149–161 (2022) doi: 10.1007/s00484-021-02202-0.
48. Johnson, M. F. *et al.* Rising water temperature in rivers: Ecological impacts and future resilience. *WIREs Water* **11**, e1724 (2024) doi: 10.1002/wat2.1724.
49. Lüthi, S. *et al.* Storylines for month-long heatwaves and associated heat-related mortality impacts over Western Europe. Preprint at https://doi.org/10.21203/rs.3.rs-5356341/v1 (2024) doi: 10.21203/rs.3.rs-5356341/v1.
50. Feudale, L. & Shukla, J. Role of Mediterranean SST in enhancing the European heat wave of summer 2003. *Geophys. Res. Lett.* **34**, 2006GL027991 (2007) doi: 10.1029/2006GL027991.
51. Berthou, S. *et al.* Exceptional atmospheric conditions in June 2023 generated a northwest European marine heatwave which contributed to breaking land temperature records. *Commun. Earth Environ.* **5**, 287 (2024) doi:

10.1038/s43247-024-01413-8.

52. Duchez, A. *et al.* Drivers of exceptionally cold North Atlantic Ocean temperatures and their link to the 2015 European heat wave. *Environ. Res. Lett.* **11**, 074004 (2016) doi: 10.1088/1748-9326/11/7/074004.
53. Bischof, S., Pilch Kedzierski, R., Hänsch, M., Wahl, S. & Matthes, K. The role of the north atlantic for heat wave characteristics in europe, an echam6 study. *Geophys. Res. Lett.* **50**, e2023GL105280 (2023) doi: 10.1029/2023GL105280.
54. Wallberg, L., Suarez-Gutierrez, L., Matei, D. & Müller, W. A. Extremely warm European summers preceded by sub-decadal North Atlantic ocean heat accumulation. *Earth Syst. Dyn.* **15**, 1–14 (2024) doi: 10.5194/esd-15-1-2024.
55. Schumacher, D. L. *et al.* Amplification of mega-heatwaves through heat torrents fuelled by upwind drought. *Nat. Geosci.* **12**, 712–717 (2019) doi: 10.1038/s41561-019-0431-6.
56. Vautard, R. *et al.* Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit. *Geophys. Res. Lett.* **34**, (2007) doi: 10.1029/2006GL028001.
57. Noguera, I., Vicente-Serrano, S. M. & Domínguez-Castro, F. The Rise of Atmospheric Evaporative Demand Is Increasing Flash Droughts in Spain During the Warm Season. *Geophys. Res. Lett.* **49**, e2021GL097703 (2022) doi: 10.1029/2021GL097703.
58. Richardson, D. *et al.* Global increase in wildfire potential from compound fire weather and drought. *Npj Clim. Atmospheric Sci.* **5**, 23 (2022) doi:

10.1038/s41612-022-00248-4.

59. Niggli, L., Huggel, C., Muccione, V., Neukom, R. & Salzmann, N. Towards improved understanding of cascading and interconnected risks from concurrent weather extremes: Analysis of historical heat and drought extreme events. *PLOS Clim.* **1**, e0000057 (2022) doi: 10.1371/journal.pclm.0000057.
60. Conradt, T. *et al.* Cross-sectoral impacts of the 2018–2019 Central European drought and climate resilience in the German part of the Elbe River basin. *Reg. Environ. Change* **23**, 32 (2023) doi: 10.1007/s10113-023-02032-3.
61. Szewczyk, W., Mongelli, I. & Ciscar, J.-C. Heat stress, labour productivity and adaptation in Europe—a regional and occupational analysis. *Environ. Res. Lett.* **16**, 105002 (2021) doi: 10.1088/1748-9326/ac24cf.
62. Erfurt, M., Glaser, R. & Blauhut, V. Changing impacts and societal responses to drought in southwestern Germany since 1800. *Reg. Environ. Change* **19**, 2311–2323 (2019) doi: 10.1007/s10113-019-01522-7.
63. Muñoz-Sabater, J. *et al.* ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* **13**, 4349–4383 (2021) doi: 10.5194/essd-13-4349-2021.
64. Schulzweida, U. Cdo user guide. (2023) doi:10.5281/ZENODO.1435454 doi: 10.5281/ZENODO.1435454.
65. Lawrence, D. M. *et al.* The community land model version 5: description of new features, benchmarking, and impact of forcing uncertainty. *J. Adv. Model. Earth Syst.* **11**, 4245–4287 (2019) doi: 10.1029/2018MS001583.

66. Sedano, F. & Randerson, J. T. Multi-scale influence of vapor pressure deficit on fire ignition and spread in boreal forest ecosystems. *Biogeosciences* **11**, 3739–3755 (2014) doi: 10.5194/bg-11-3739-2014.

67. Hermann, M., Wernli, H. & Röthlisberger, M. Drastic increase in the magnitude of very rare summer-mean vapor pressure deficit extremes. *Nat. Commun.* **15**, 7022 (2024) doi: 10.1038/s41467-024-51305-w.

68. Bonan, G. B. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* **320**, 1444–1449 (2008) doi: 10.1126/science.1155121.

69. Murray, F. W. On the computation of saturation vapor pressure. *J. Appl. Meteorol.* **6**, 203–204 (1967) doi: 10.1175/1520-0450(1967)006<0203:OTCOSV>2.0.CO;2.

70. Steinfeld, D. GitHub - steidani/ConTrack: Contour Tracking of circulation anomalies (atmospheric blocking, cyclones and anticyclones) in weather and climate data. https://github.com/steidani/ConTrack (2020).

71. Dole, R. M. & Gordon, N. D. Persistent Anomalies of the Extratropical Northern Hemisphere Wintertime Circulation: Geographical Distribution and Regional Persistence Characteristics. *Mon. Weather Rev.* **111**, 1567–1586 (1983) doi: 10.1175/1520-0493(1983)111<1567:PAOTEN>2.0.CO;2.

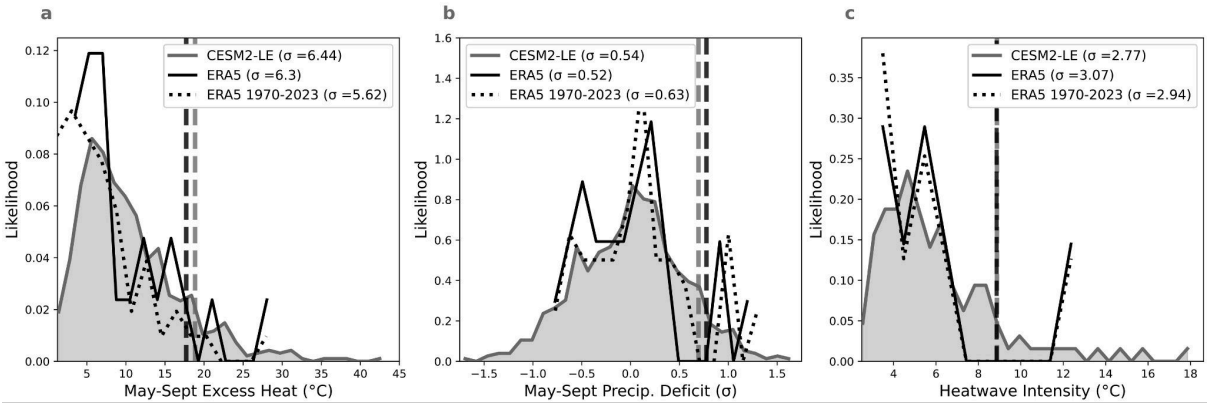
**Acknowledgements**

This project and L.S.G. received funding from the European Union's Horizon Europe Framework Programme under the Marie Skłodowska-Curie Action grant agreement No 101064940. We acknowledge the Swiss National Computing Centre (CSCS) for providing the necessary computational resources. We acknowledge the CESM2 Large Ensemble Community Project and supercomputing resources provided by the IBS Center for Climate Physics for producing the original 100-member CESM2 Large Ensemble.

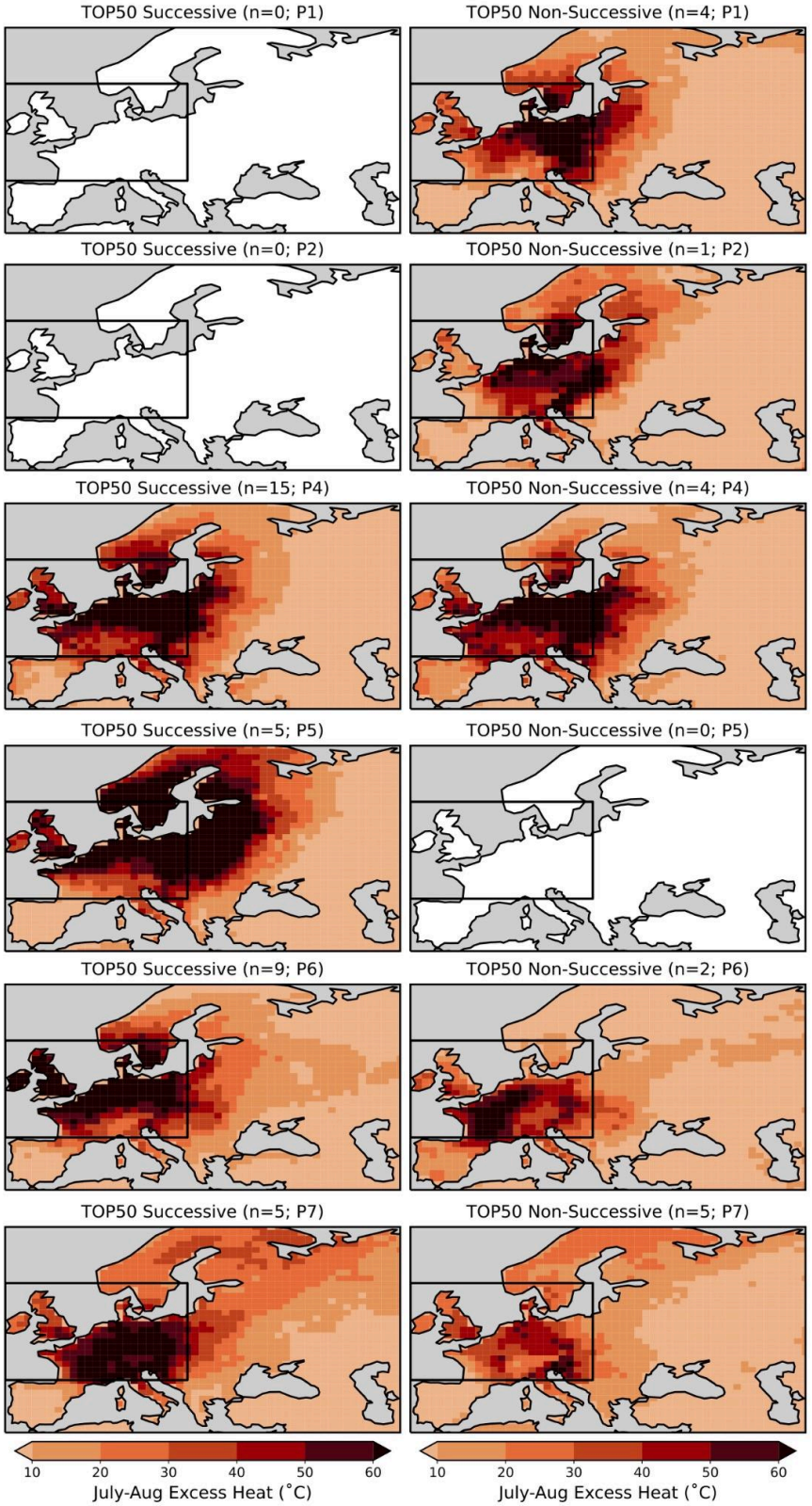
**Author Contributions**

L.S.G. designed the study, performed the majority of the analysis, figures and writing. U.B. performed the climate model experiments. M.M. performed the atmospheric blocking analysis. U.B., M.M., E.M.F and R.V. contributed to the interpretation of the results, and provided feedback on the manuscript.

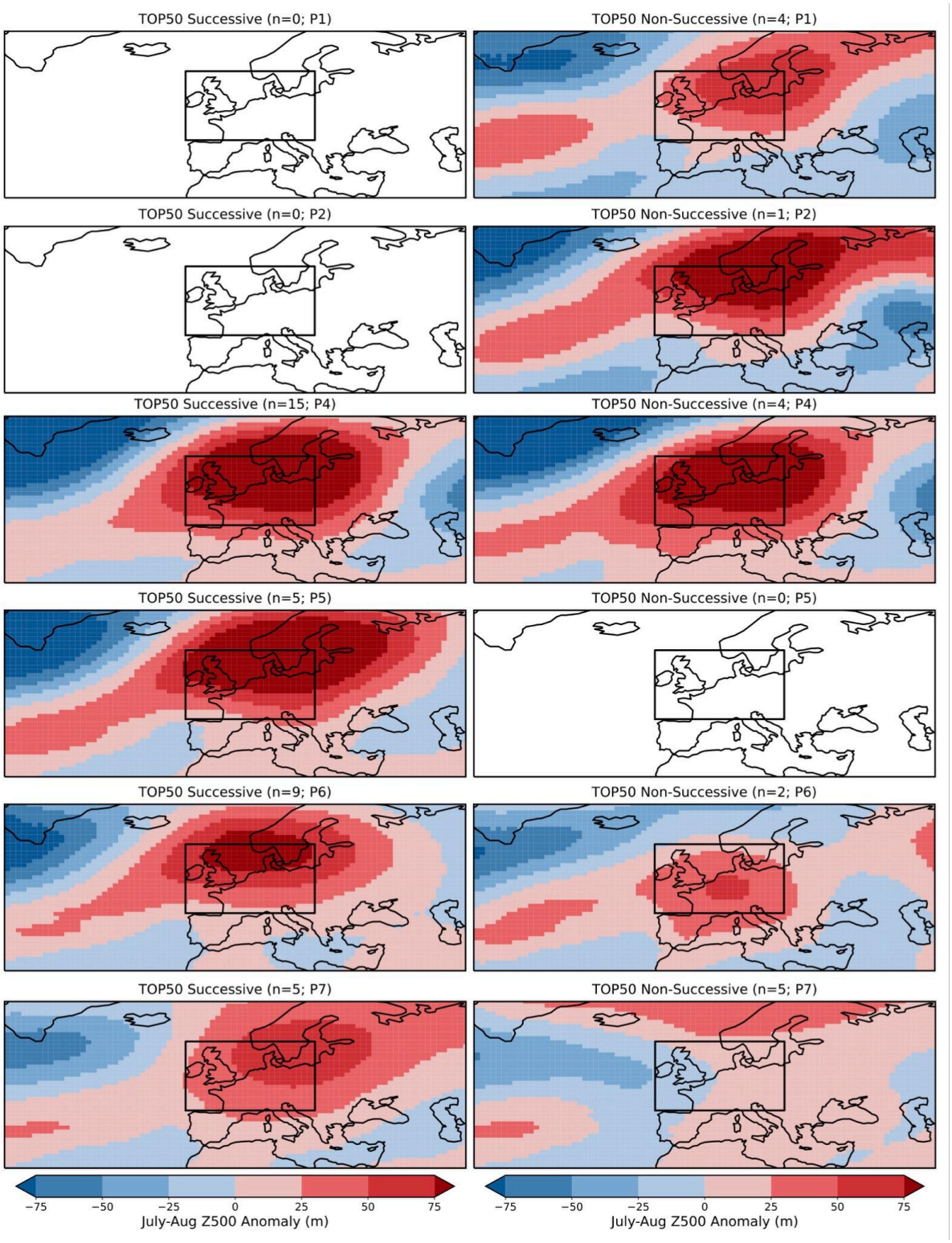
# Supporting Figures



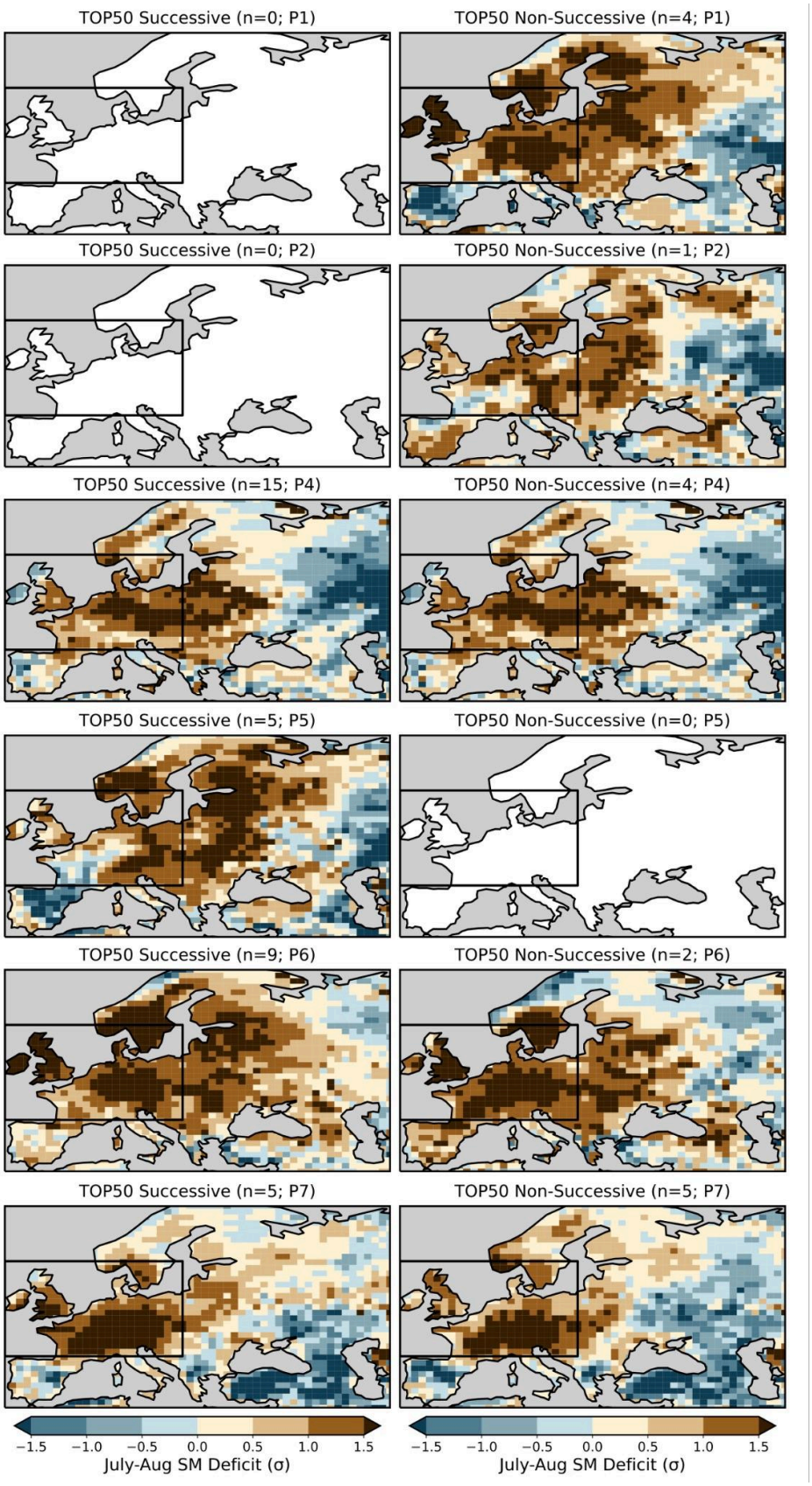

**SI Figure S1: Evaluation of Heat and Drought conditions in CESM2-LE against ERA5. a)** Probability density functions (PDF) of May–September accumulated Excess Heat as in Fig. 1a, for the 35-member CESM2-LE (gray; 2005–2023) and ERA5 (solid black line for 2000–2023 and dotted black line for 1970–2023). Vertical dashed lines mark the 90th percentile in ERA5 (black; 2000–2023) and CESM2-LE (gray; 2005–2023). σ values show the standard deviations of the different samples.**b)** Same as a, but for standardized May-September accumulated Precipitation Deficits**.** Increasingly positive precipitation deficits indicate increasingly drier conditions. **b)** Same as a), but for heatwave intensity as in Fig. 1b. CESM2-LE simulations follow historical (2005-2014) and SSP3.7 (2015-2023) forcings.



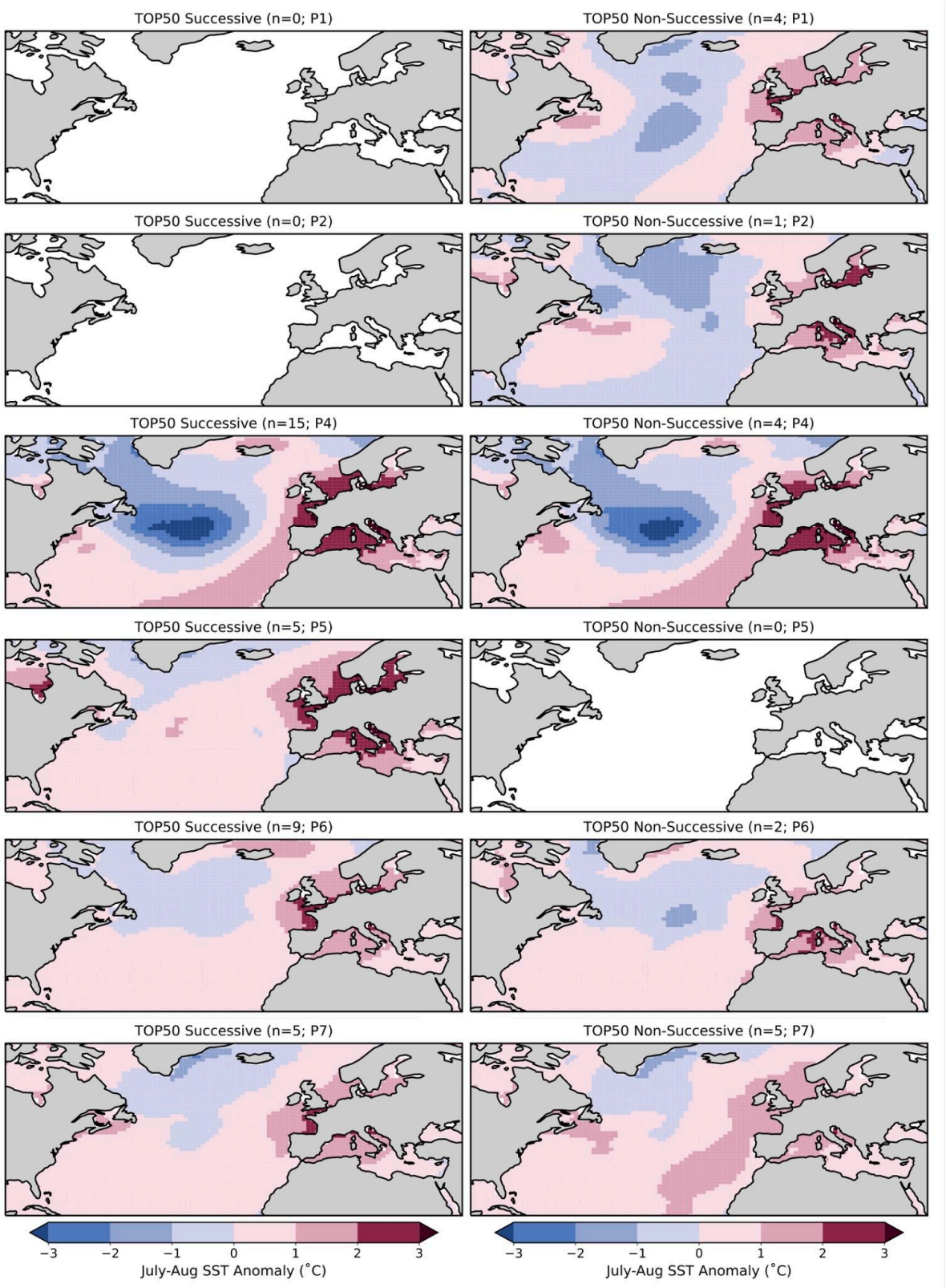

**SI Figure S2: Excess Heat during TOP50 Successive vs. Non-Successive Heatwaves per parent.** Accumulated Excess Heat in July–August, composites for the TOP50 most extreme successive heatwaves for different parent cases. The region used for spatially aggregated heatwave metrics is marked by the black box.



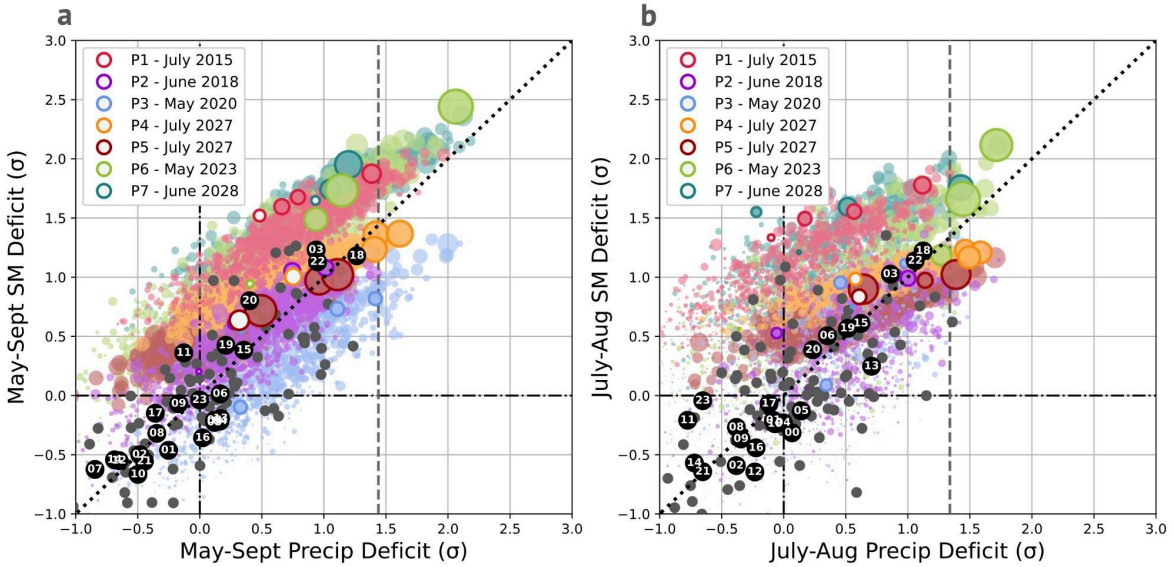

**SI Figure S3: Atmospheric conditions during TOP50 Successive vs. Non-Successive Heatwaves per parent.** Mean geopotential height anomalies at the 500 hPa level (z500) averaged in July–August composites for the TOP50 most extreme successive heatwaves for different parent cases. The region used for spatially aggregated heatwave metrics is marked by the black box.



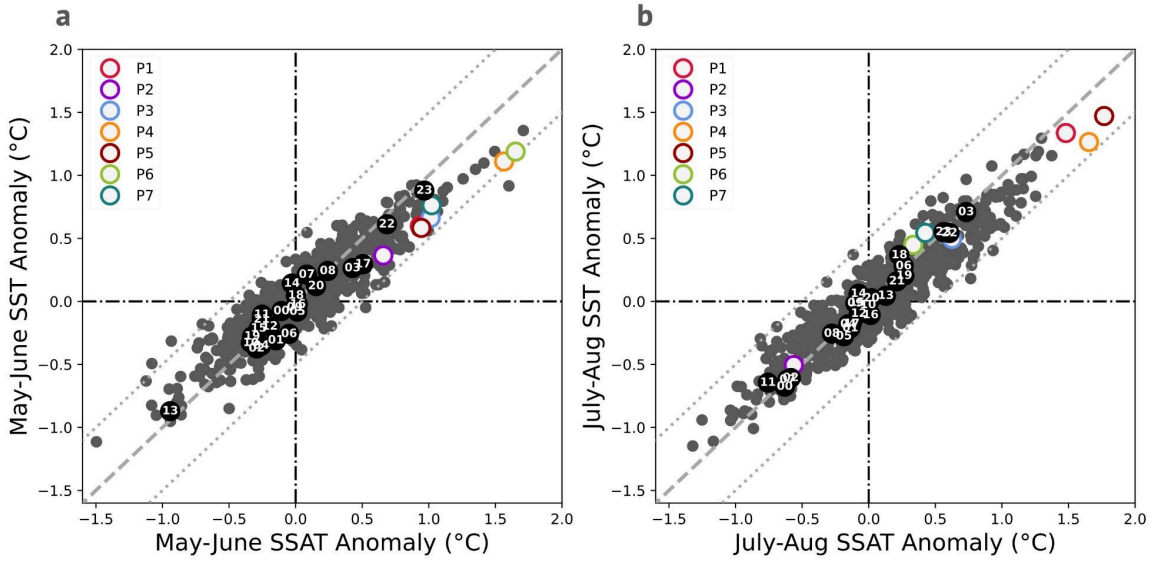

**SI Figure S4: Drought conditions during TOP50 Successive vs. Non-Successive Heatwaves per parent.** Mean soil moisture (SM) standardized deficits averaged in July–August composites for the TOP50 most extreme successive heatwaves for different parent cases. The region used for spatially aggregated heatwave metrics is marked by the black box.




**SI Figure S5: Sea surface temperature conditions during TOP50 Successive vs. Non-Successive Heatwaves per parent.** Mean sea surface temperature (SST) anomalies averaged in July–August composites for the TOP50 most extreme successive heatwaves for different parent cases.




**SI Figure S6: Soil Drought conditions. a)** Standardized Soil Moisture (SM) vs. precipitation (Precip) deficits in the last 5 members of the 35-member non-boosted CESM2-LE (gray dots), against ERA5 (black dots, year in white; 1940-2023), seven parent simulations (P1-7; white infilled colored circles), and boosted storylines (colored dots) averaged for the May-Sept. period. **b)** Same as a), but averaged for July-August. For parent and boosted simulations, dot size is proportional to Excess Heat accumulated over concurrent period. Dotted lines mark the 1:1 slope, dashed gray lines mark the 90$^{th}$ perc. in precipitation deficits in a 3˚C warmer world.




**SI Figure S7: Sea Surface Temperatures a)** Sea Surface Temperature (SST) vs. Sea Surface Air Temperature (SSAT) anomalies in the 35-member non-boosted CESM2-LE (gray dots), against ERA5 (black dots, year in white; 1940-2023), and against seven parent simulations (P1-7; white infilled colored circles), averaged for the May-June period. **b)** Same as a), but averaged for the July-August period. Dashed gray lines mark the 1:1 line, while dotted gray lines mark a ±0.5˚C range.