Alternative rainfall storylines for the Western European July 2021 floods from ensemble boosting

communications earth & environment

Article

A Nature Portfolio journal

https://doi.org/10.1038/s43247-025-02386-y

Alternative rainfall storylines for the Western European July 2021 floods from ensemble boosting

Check for updates

Vikki Thompson [1] ✉, Dim Coumou [2], Urs Beyerle [3], Joy Ommer [4,5], Hannah L. Cloke [5,6] & Erich Fischer [3]

In July 2021, a cut-off low-pressure system brought extreme rainfall to Western Europe, leading to flooding that caused loss of life and infrastructure damage. Here, we use ensemble boosting to investigate alternative storylines of the event, given the observed dynamical situation. Using a fully coupled free-running climate model, we identify atmospheric flow analogues of the 2021 event in an initial-condition large ensemble. These analogues are re-initialised with slightly perturbed atmospheric initial conditions to generate physically plausible alternative storylines. The storylines are used to investigate how a potentially worse event could have unfolded given the same large-scale dynamics. We identify rainfall events with longer persistence and larger extent, yet the observed event appears to be towards the upper end of what is plausible in the current climate. Such storylines can be used to prepare for possible future events, helping society to imagine dangerous, but plausible, scenarios.

In July 2021, an extreme rainfall event impacted parts of Germany, Belgium, the Netherlands, and Luxembourg[1,2]. Very high river levels were observed in the Meuse and Rhine catchments[3]. Impacts on society were large—with over 40,000 people affected. The flooding caused at least 220 deaths, mostly in the Ahr catchment, Germany[4]. Infrastructure damage included hospitals, roads, bridges, and utility networks, with an estimated total damage of EUR 46 billion[1].

The extreme rainfall of July 2021 was linked to a persistent cut-off low-pressure system. Rainfall was greatest on July 14th, with daily totals of up to 150 mm recorded in parts of western Germany[1] (Figs. 1a and S1)—twice the climatological mean for the entire month of July (~70 mm). The event was among the heaviest precipitation events in Germany in the past 70 years, with records broken at several observational stations[5]. Due to the stalling of the cut-off low, the rain persisted for several days, which increased ground impacts. Some of the highest rainfall totals hit the densely populated Rhine Valley region, including the cities of Cologne and Bonn. Intense rainfall over different timescales has differing impacts. In this event, intense sub-daily rainfall upstream led to flash floods, which ultimately caused more fatalities. Further downstream, the response is slower; steady rainfall over several days leads to impacts in the larger cities. Rainfall accumulation over different timescales, sub-daily to multiday, is important for different types of flooding.

According to official records and common memory, the event was 'unprecedented' for the region with daily accumulated rainfall and river levels higher than seen in recent years. However, there is evidence of similar events in the past, before official records and beyond the span of human memory (i.e. a lifetime)[6]. Many of those impacted had no prior experience of severe flooding events and did not make any or sufficient preparations[7] Human perception of risk is often biased by past experience[8], leading to preparedness levels below what is required for the most extreme—once in a lifetime—climate events. Planning policy and disaster risk preparedness both require clear information on the risks from heavy rainfall and flooding. Motivating society to think about scenarios beyond what has been experienced will allow better preparation and mitigation, potentially reducing the impacts of future extreme events[9,10].

Thinking back, we often imagine what we could have done differently; thus, the outcome would have been better. In this study, we did not ask ourselves what we could have done differently, but instead we ask whether the rainfall event itself that caused these devastating floods could have been even worse? We know that the dynamical atmospheric situation is plausible as it did actually occur, but could small changes in its evolution have led to

[1]Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands. [2]Institute for Environmental Studies, Vrije Universiteit Amsterdam, Amsterdam, Netherlands. [3]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland. [4]KAJO s.r.o., Bytča, Slovakia. [5]Department of Geography and Environmental Science, University of Reading, Reading, UK. [6]Department of Meteorology, University of Reading, Reading, UK.
✉e-mail: vikki.thompson@knmi.nl

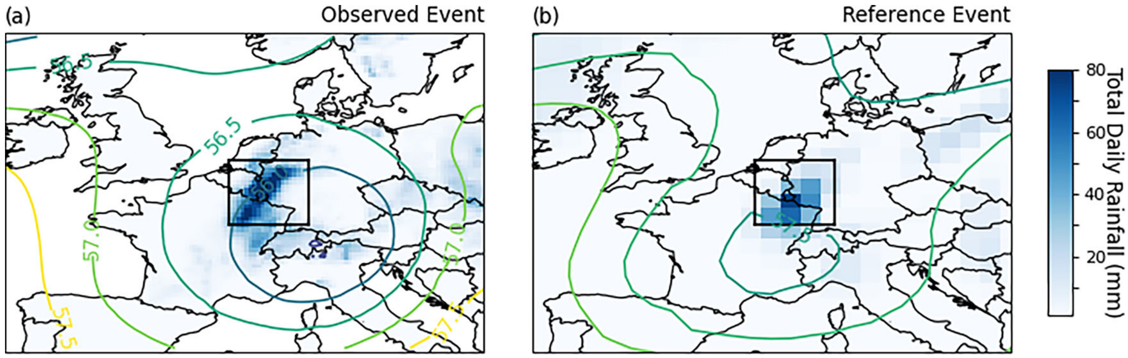

Fig. 1 | The meteorological situations of the observed and reference events. **a** Total 24-h accumulated rainfall from E-OBS and daily mean 500 hPa geopotential height from ERA5[50], over western Europe on 14th July 2021, the day of peak rainfall from the event. Contour lines show the 500 hPa geopotential height (m), shading shows the daily rainfall totals (mm), regions with <1 mm in white. **b** As in (**a**) for the reference event used to create the boosted ensemble, taken from CESM2 large ensemble of the current climate. The box indicates the rainfall region we assess.

greater impacts? We only consider what could have happened, given the dynamical situation. This conditioning leads to the use of the term 'worse' case—rather than worst—as under different dynamical conditions greater rainfall extremes may be able to occur. Often, when investigating unprecedented extreme climate events, magnitude is the primary measure investigated and other characteristics of the event are overlooked[11]. Instead, we investigate three different characteristics related to the extreme rainfall event: given the specific dynamical situation, on a daily timescale could the event have (1) been more persistent, (2) covered a larger area, and (3) impacted a different region? Investigating these different scenarios can better our understanding of possible worse scenarios within the dynamical constraints and hence, increase our knowledge of the characteristics of extreme precipitation events that society should be prepared for.

As we have seen, limited imagination of extreme rainfall scenarios can prevent us from preparing for them. Forecasts and warnings alone may not be enough to prevent impacts[12]. Society may not understand the danger if a forecast suggests rainfall beyond what has previously occurred. People who were affected by the floods in July 2021, may now be able to imagine this kind of severe event and therefore may be able to prepare better in the future. However, it is rather unlikely that the event will happen with the exact same rainfall characteristics as those of July 2021. Therefore, it is important to understand which other scenarios could have taken place to feed our imagination with further possible scenarios. This is especially important for people who were not (severely) affected in July 2021, but actually could have been if the dynamical situation had turned out slightly differently. It is important because these people might think in the future 'that time it did not impact us, so why would it this time?' since they may assume that they already experienced the worst possible event and, hence, may not be able to visualise flooding beyond that of past events.

Climate models can be used to create storylines of rainfall events, allowing possible more impactful events to be investigated[13–15]. We use ensemble boosting to produce alternative storylines of the July 2021 rainfall event—feasible alternative pathways within the current climate. This method has previously been applied to heatwaves, identifying storylines of record-shattering heat in regions yet to experience such[16,17] to create storylines of multi-year droughts[18]. We use daily accumulated rainfall to assess the event. This is a useful measure of flood potential and is used to issue initial warnings. In the July 2021 case, forecasts of daily accumulated rainfall led to the initial warnings issued, first to the Rhine and later the Meuse catchment[19]. Later, higher-resolution information allowed warnings of greater spatial and temporal precision.

When investigating extreme rainfall, the measure most often assessed is how much rain occurs over a given region within a given time frame. For example, ref. 2 carried out an attribution study for the July 2021 event assessing daily rainfall over defined 'tiles' of approximately 140 × 140 km. To calculate the return period of the event in model simulations, ref. 20 also used daily rainfall over a single region defined from the event. Initialised hindcast ensembles have been used to quantify the risk of extreme rainfall totals in other regions in the current climate[21,22]. As well as rainfall magnitude alone, event duration is an important aspect of understanding risk. While the drivers of the dangerous flash floods that affected the upstream valleys in these floods can be captured by extreme daily totals, prolonged rainfall, particularly when this falls on already wet ground can lead to severe flood impacts across the whole river catchment and in particular downstream in the lowland floodplains, which is often where larger population centres reside and most societal and economic damage can occur[23,24]. Spatial coverage and the coincidence of the rainfall with catchment boundaries is also highly relevant for understanding the flood genesis[25]. The impact of rainfall will differ if rain occurs in one catchment or is spread over multiple catchments, leading to the need for different emergency response plans.

Climate change has increased the intensity and frequency of summer rainfall extremes, and this is projected to continue[26–32]. With global warming, the atmosphere can hold more moisture, and this leads to increased extreme rainfall globally. Regional rainfall is also influenced by local dynamical changes, resulting in spatial variability in rainfall trends. Over Europe, there is observational evidence of increased magnitude and intensity of extreme rainfall since the 1950s[33]. Climate projections suggest the most intense rainfall events observed in Europe in the current climate are projected to almost double in frequency for each 1 °C of further global warming[34]. Understanding the range of plausible extreme precipitation in the current climate is essential to enable society, policymakers, and insurers to prepare for the full range of events that could occur.

Results

In the boosted ensemble, we can identify storylines of more extreme rainfall, based on large-scale characteristics (Fig. 1). We assess the boosted ensemble to identify accumulated daily rainfall over the region impacted. We assessed the model's capability to simulate events with similar extreme accumulated daily rainfall within the impacted region and found that the model is capable of simulating such extreme rainfall (Fig. S2). We assess the full distribution of JJA daily rainfall (Fig. S2a, b) showing the model is capable of simulating more extreme events than found in either reanalysis (ERA5) or observational (E-OBS[35]) datasets—as may be expected with a large ensemble sampling further into the extreme tails. We also assess the return period of the annual maximum values (Fig. S2c, d) again showing good agreement between the model, reanalysis, and observational datasets. At the model resolution, the observed event had maximum gridpoint rainfall of 97.4 mm/day, but the greatest local rainfall in the boosted ensemble is 84.0 mm/day (Fig. 2a). Gridpoint level rainfall is >60 mm in a day (DWD alert level 1)[36]. We do find simulations with (slightly) greater regional daily rainfall than observed (Fig. 2b). The observed event had 37.0 mm/day areal average

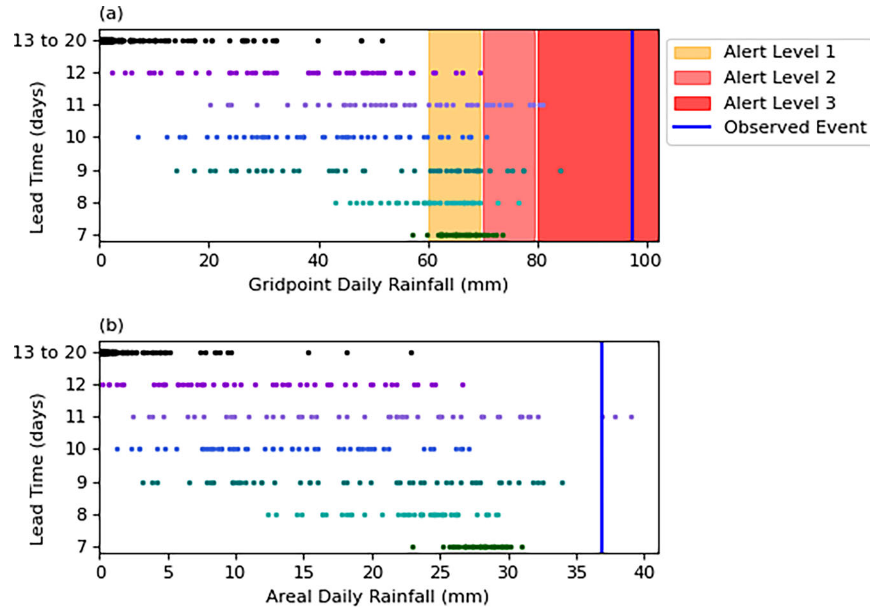

Fig. 2 | Does the boosted ensemble show extreme rainfall? Daily accumulated rainfall over the event region 48–52°N, 4–9°E on the event date, 14 July 2021. **a** The gridpoint maximum daily accumulated rainfall for each simulation in the boosted ensemble (dots) and for the observed event from E-OBS regridded to the boosted ensemble grid (vertical blue line). Background shading indicates the DWD alert levels (Table 1[36]). **b** As in (**a**) but for the areal mean daily accumulated rainfall for each simulation in the boosted ensemble over the event region, with the E-OBS value (vertical blue line).

rainfall over the whole region; the boosted ensemble shows simulations with up to 39.0 mm/day.

The boosted ensemble is initialised 7–20 days before the reference event, but not all the lead times show the extreme rainfall over Western Europe identified in the reference event. At shorter lead times, the rainfall is more consistent across the ensemble as there is less time for the ensemble members to diverge dynamically. Shorter lead times are likely to have greater similarity with each other, and greater similarity with the reference event, thus worse cases are less likely. In the longer lead times, many events show little rain—the dynamics have diverged from the observed event, so we no longer capture extreme rainfall within the ensemble. There is a spread in the precipitation patterns within the ensemble, and to a lesser extent in the dynamical patterns (Fig. S3). This illustrates that the ensemble can capture alternative plausible storylines of daily rainfall from the same initialisation state. To investigate the storylines of worse scenarios, we use only the 6 lead times with simulations showing gridpoint level rainfall >60 mm in a day, 7–12 days—providing 300 simulations of the event (Fig. 2). The range of useful lead times is event dependent.

Longer persistence

Although it has been shown that a majority of the rainfall leading to flooding in the observed event occurred within the 24 h of 14th July[1], rainfall persisted over several days due to the stalling of the cut-off low. Further downstream, where the river is bigger, it will react more slowly to rainfall—and persistent rainfall over several days may lead to larger impacts. For the observed event, much of the impacts, and particularly loss of life, occurred in the upper parts of the catchment where runoff response is more rapid, and ideally we would assess shorter duration rainfall, but this is unavailable for the model ensemble we use. Downstream the catchment, longer rainfall persistence could be more impactful, and this aligns with the locations of larger cities. Rainfall accumulations over a specific time period will lead to a specific impacts in a specific catchment[25,37]—thus we assess a range of thresholds.

We assess the persistence of rainfall in the boosted ensemble over the four time periods over which DWD issues rainfall warnings[36]. We identify the greatest 1-, 2-, 3-, and 4-day accumulated rainfall within the event region. The boosted ensemble shows no 1-day rainfall accumulations as extreme as was observed (Figs. 1a and 2a), though many members do reach warning levels. For 2-day rainfall accumulation, the boosted ensemble extremes are closer to the observed, but still not as great (Figs. 3b and S4). The apparent lack of more extreme modelled events at both 1- and 2-day scales may be due to the use of daily resolution data. Greater 24- or 48-h events that span multiple calendar days will not be included. For both 3- and 4-day rainfall accumulations, the boosted ensemble shows simulations with more extreme rainfall (Fig. 3c, d). For 4-day rainfall accumulation, the observed event did not reach alert levels, 1 member of the boosted ensemble is above the 150 mm in 4-day threshold, and 2 other simulations show greater rainfall than observed (Fig. S4).

Although single-day rainfall accumulation is not as extreme as observed, the boosted ensemble shows worse scenarios for multi-day rainfall accumulation. These storylines—149 mm in 3 days and 158 mm in 4 days—could lead to different flood impacts than occurred. Imagining the consequences of such storylines is important to reduce possible impacts should such an event actually occur. Continuous heavy rainfall over multiple days can increase flood extents in the lower reaches of large rivers, such as the Rhine[25]. In regions where weirs and other artificial measures have been engineered, persistent rainfall can be particularly impactful in adjusting the rate of flood propagation; this can increase flood extents upstream as the water is slowed leading to larger flood extents[38].

Larger rainfall extent

Next, we investigate the plausible spatial extent of the rainfall—could the rainfall have been spread over a larger area, potentially increasing the impacts? If rain is spread within the same catchment, it could potentially lead to higher total run-off. Rain spread over multiple catchments—though less in each individual catchment—may cause more impacts as coordination of emergency response may be harder as equipment is stretched over a larger region. Spatial extent of European floods has been shown to increase over the past 70 years—in part due to larger spatial extent of rainfall[39]. We assess only the current climate; if the observed trends continue, it is possible that greater extents than we find would occur.

We take two accumulated daily rainfall thresholds, 40 and 60 mm, we count the number of gridboxes over each threshold on the event day. The model grid is 1.25 × 0.9°, approximately 140 × 60 km at 50°N. The spatial resolution prevents assessment of catchment-level rainfall, which would allow us to better determine where impacts would occur, but we can assess differences in spatial extent more generally. In the observed event, 7 gridboxes show rainfall over 40 mm in a day, and 4 gridboxes with over 60 mm in a day. As was the case for persistence, the majority of simulations show a less extreme event in terms of spatial extent. Many simulations do not show any gridboxes above the threshold (Fig. 4a, c). Although few model simulations show a greater spatial scale than the observed event—but there are 4 events with 7 and 8 gridboxes of daily accumulated rainfall >40 mm,

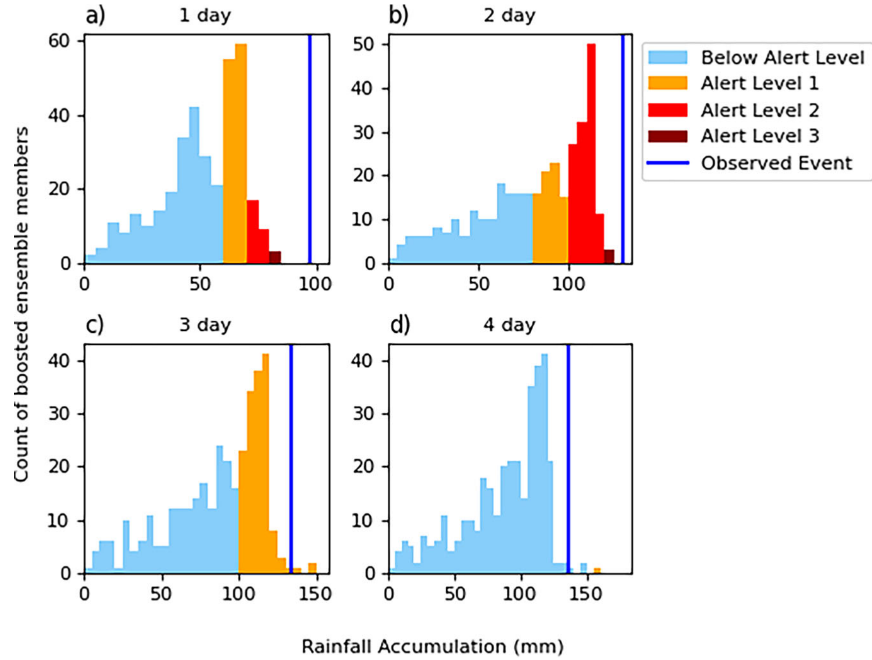

Fig. 3 | Could the event have persisted longer? **a** Histogram of 1-day rainfall accumulation in mm, for the 300 members of the boosted ensemble, for the maximum gridbox within the region 48–52°N, 4–9°E. Histogram colours indicate the DWD rainfall alert levels (Table 1[36]). The vertical blue line shows the accumulated rainfall of the observed event gridbox maxima (E-OBS, regridded to the model grid). **b**–**d** As in (**a**), for 2-, 3-, and 4-day rainfall accumulation.

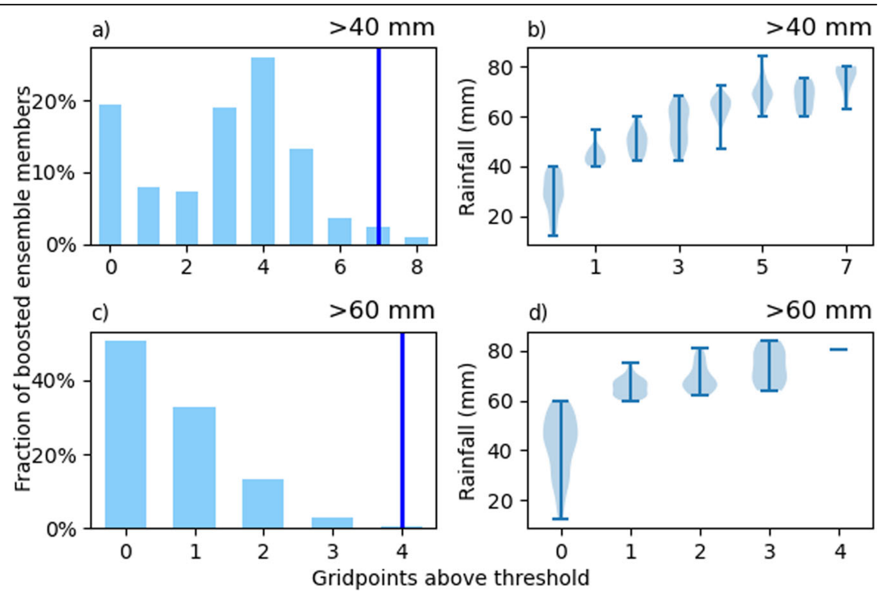

Fig. 4 | Could the event have covered a larger area? **a** Histogram showing the fraction of the 300 boosted ensemble members displaying each number of gridpoints of accumulated daily rainfall >40 mm. Vertical line indicates the observed event from E-OBS. **b** Violin plots of gridpoint maximum daily rainfall (mm) for each simulation, separated by the number of gridpoints with rainfall >40 mm. Note—the width of the violin is not indicative of the number of events, it is uniform. **c**, **d** As in (**a**, **b**), for an accumulated daily rainfall threshold of >60 mm.

corresponding to an area of 60,000–70,000 km^2. We show that gridbox maximum daily rainfall generally increases with the scale of the event—so events covering a larger area are likely to also be more intense locally, at least on the spatial and temporal scale of the model (Fig. 4b, d).

The storylines with the largest spatial extent cover a region greater than the area of Belgium. An event of this magnitude would require coordination between nations. The European Emergency Coordination centre has struggled with previous events, including the 2021 observed event, as the sharing of pumps between countries is not feasible when both countries are flooding simultaneously[40]. Later, we discuss the impacts of a storyline with both a larger extent and a shifted location.

Different locations

The observed event of July 2021 impacted a densely populated area of the Rhine valley, including the cities of Dusseldorf and Bonn. Had the rainfall been shifted, it could have impacted less populated regions, or the densely populated regions of the Netherlands to the north west or Paris to the south west. Past extreme rainfall events in Germany have occurred over cities—for example, in June 2017 rainfall was more intense than the July 2021 event hit Berlin[5]. Despite extensive flooding, including to the Berlin metro system, the June 2017 event led to no fatalities and much lower insured losses than July 2021[41].

Other studies have considered the impacts of the rainfall occurring in neighbouring locations[9,42]. The method we use allows us to investigate if such imagined events are plausible scenarios, given the dynamical constraints. An attribution study of the event pooled tiles over a larger region than impacted by the event to allow more robust calculation of return levels to be calculated[2]. The larger region was chosen for testing statistical similarity of extreme rainfall characteristics—not dynamical similarity. The range of extreme rainfall locations in the boosted ensemble could better guide regions over which such pooling is dynamically suitable.

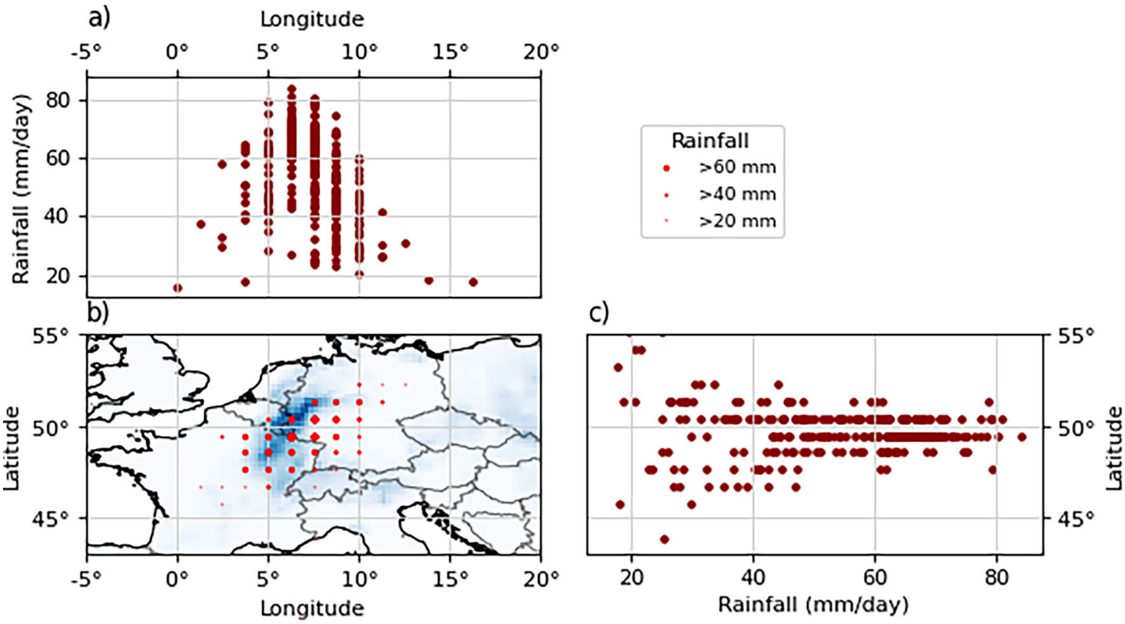

Fig. 5 | Could the event have impacted a different region? a The longitudinal position of the maximum rainfall plotted against the total daily rainfall. **b** As in (**a**), but showing latitudinal position. **c** Map of the maximum rainfall locations, scaled by the amount of rainfall, with the observed rainfall from E-OBS shaded.

We show that in the boosted ensemble events with the greatest rainfall totals are constrained to a relatively small region, centred on 50°N, 7°E (Fig. 5). Assessing the proportion of simulations with rainfall above specified thresholds also shows that the most intense rainfall in the boosted ensemble occurs in the same region (Fig. S6). This will be partly due to the enforced dynamical constraints, as all boosted ensemble members are initiated from the same reference event. The observed event is towards the northern limits of the plausible region suggested by the boosted ensemble. The plausible region is centred on the German state of Rheinland Pfalz, close to France and Luxembourg. No events in the boosted ensemble occur further to the north west than the observed event, suggesting that—given the dynamical constraints—the event may not be plausible over the Netherlands. There are events that are centred further south west including the Paris region, but these events show smaller rainfall totals.

Given the dynamical conditions, the observed event corresponded with the most densely populated region plausible. Rheinland Pfalz, where the ensemble events are centred, is less populated than where the observed event was centred, but impacts could still be large depending on whether that region has systems in place to prepare them for such an event. Impacts will depend on how prepared a region is—in July 2021, Luxembourg had relatively large impacts due to the lack of an integrated warning system[43]. The model does not represent the complexity of the orography, which could affect both the rainfall location and accumulated rainfall totals[44]. Orography has been shown to be important for the observed event—with the highest rainfall intensities found in the Eifel mountain range, where multiple rivers originate[1]. Further research investigating whether the same dynamical conditions are equally probable in another location would be valuable. Likewise, it would be valuable to repeat the experiment with a convection-permitting model that accounts for the complexity of the orography and further resolves the convection.

The alternative storylines are not chosen to simply have greater rainfall, but changes to characteristics that would require different preparation and emergency responses to reduce impact on society. Assessing each characteristic in turn allows us to think about the possible consequences of such a rainfall storyline—but we can identify storylines that incorporate multiple changes. Some examples of the daily rainfall accumulation from the wettest days of both the observed event and model events are shown in Fig. 6. Some rainfall storylines look similar to the observed event (Fig. 6e), others show changes in location and/or spatial extent. Figure 6d shows a rainfall storyline where the extent is larger, and coincides with the French mountainous region of the Vosges—another region with topography susceptible to flashy floods. Such an event would require clear communication and understanding of the risks in the flashy upstream region, and then an understanding of which catchments will get the most run-off, to reduce downstream impacts. We also show a storyline with greater daily accumulated rainfall over Luxembourg (Fig. 6c). This region struggled with flood warning systems during the July 2021 event[43]; this study suggests that policymakers need to consider the impacts of even greater rainfall.

Discussion

We present plausible storylines for alternative extreme rainfall events similar to that observed in Western Europe in July 2021, conditioned on the dynamical situation observed. We use climate model simulations to show storylines with larger rainfall extents, longer persistence, and different locations (Fig. 6). We find the observed event is towards the upper limit of the storylines suggested by the model ensemble in terms of spatial extent and persistence. Despite this, the climate model simulations include many events that could cause disastrous surface impacts. We find that events that are more intense locally tend to also show a greater spatial extent. Society may not be prepared for such events, as they have not occurred in the historical record. Similar events may have happened in the past but have gone unrecorded or been forgotten, or it may be that they have never happened, and our changing climate has only now made them a plausible possibility. The model events follow the real-life dynamical situation that occurred in July 2021, giving us confidence that they are plausible events that we need to be prepared for.

In 2021, the German early warning system failed in various ways[40]. This resulted in many citizens not receiving any warnings until the water arrived at their doorstep. This was especially noticed in the worst-affected areas. In this regard, the lack of understanding, imagination and belief played a role in the dissemination of warnings through the top-down warning chain. This lack was further identified at the individual level when warnings were received[7]. In particular, the lack of understanding refers to the translation of received warnings into potential flooding extents and impacts, the imagination of what this would look like in reality, and the cognitive processes that make us believe something or the bias that prevents us from believing something to happen (wishful thinking). Hence, using these storylines can

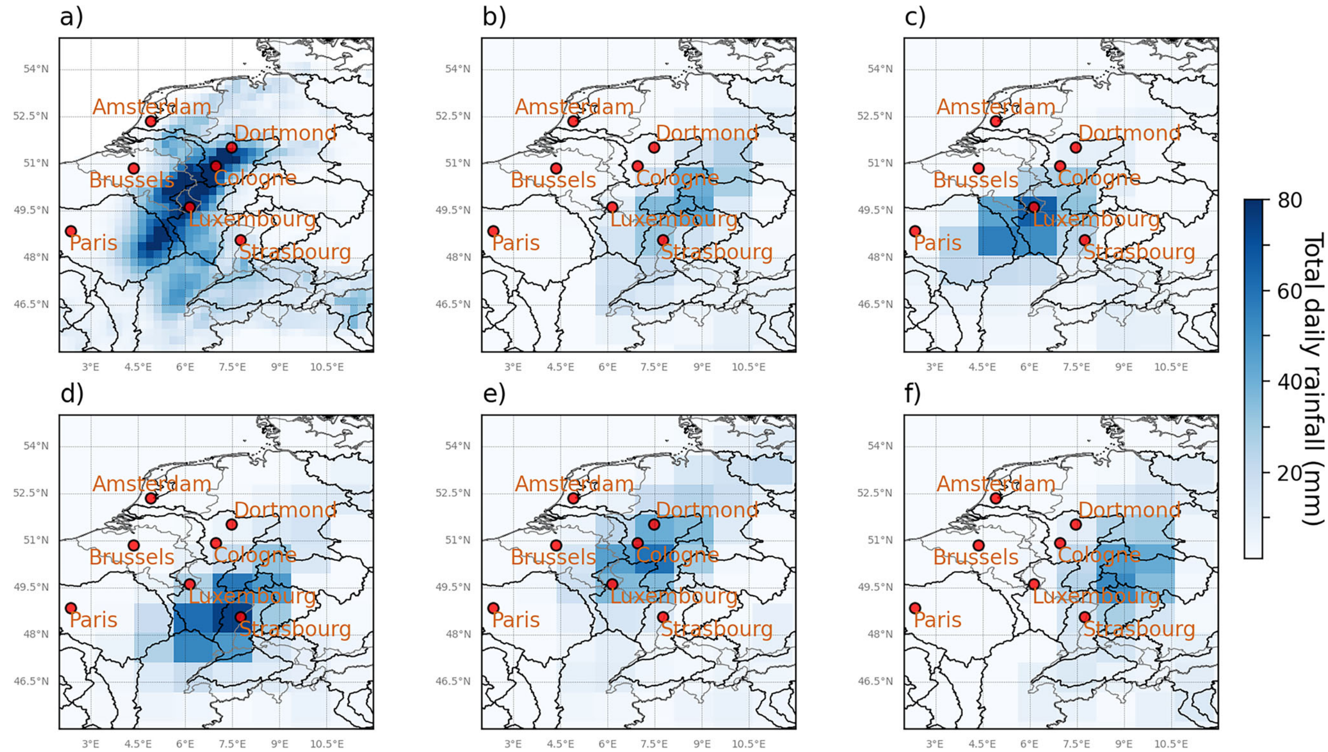

Fig. 6 | Plausible alternative storylines. a The observed event, with E-OBS rainfall data, and **b–f** five examples of possible alternative storylines found in the boosted ensemble. River basins are outlined in black, national borders in grey, and key cities in red/orange.

help to address this, and building on a real-life case can support the impact of communication campaigns or workshops.

Being able to understand the impact of forecasted events, to imagine these, and actually believe that they can happen is of high importance for disaster preparedness. It triggers action, both in a long-term (from now onwards) and short-term perspective (when a warning is received). In this context, individual preparedness includes both physical preparedness, such as moving valuables upstairs, and psychological preparedness. Hence, if storylines can help us to imagine, we can not only physically but also psychologically prepare ourselves for future events. Psychological preparedness is foremost important for us to be able to cope with disastrous and shocking events and recover from them.

The developed physical storylines in this study are not only of interest to climate scientists but also need to be communicated to disaster risk management practitioners, decision-makers as well as to the public. We can use these storylines to communicate, not only the plausibility of these events, but also their danger. Further development of such storylines using high-resolution models will enable society to consider what actions need to be taken to better prepare in terms of, for example, planning, insurance, disaster risk management, and societal wellbeing[45]. Bridging the knowledge gained from the storylines to the imagination of people, different communication tools can be applied, ranging from storytelling, participatory workshops, serious games, virtual reality visualisations and more[46]. One example of storytelling is The Weather Channel's FloodFX technology[47], which can help us imagine how an extreme weather event would impact us personally and how to physically and mentally prepare and adapt, reducing impacts. The worse cases we identify can be used to produce storylines of plausible events on a local scale, to encourage visualisation of the knock-on effects of such rainfall totals enabling the reduction of impacts should it actually occur.

There are alternative methods for creating rainfall storylines, such as relocating an observed event[9,42] or the use of artificial intelligence (AI). The advantage of using a climate model is that we obtain physically plausible events. Simply moving an event geographically, or multiplying the accumulated rainfall totals, does not tell us anything about whether the event is dynamically possible. AI-generated extreme events may not follow the rules of physics, and may be limited by training data, which cannot include record-shattering future events[48]. Storylines could also be taken from an ensemble weather forecast. Such ensembles are generated every day, providing a range of physically plausible possibilities for the future weather. For this event, the extreme rainfall was predictable from a 3-day lead time[49], thus impactful storylines show less variation than the boosted ensemble—where we assess lead times up to 12 days.

Previous use of ensemble boosting has selected events based on temperature or rainfall maxima, we select based on large-scale dynamics. We find that not all simulated events with high similarity in the 500 hPa geopotential height show extreme rainfall—but some do. In principle, the dynamical analogue conditions could be more narrowly defined to also ensure consistency in the moisture convergence or the vertical structure. We specifically aimed at using a simple definition for which it is not a given that the model would show any rainfall events from dynamical conditioning alone. One limitation is the coarse resolution of the model, which does not represent the complex topography and does not resolve the convective processes potentially embedded in the large-scale weather system. The impacts of the observed event were driven by sub-daily extreme rainfall, which we cannot assess with only a daily temporal resolution. Convection-permitting models are computationally very expensive, and the currently available simulations are not yet long enough to specifically probe the tail of the rainfall distribution in pre-selected regions. Furthermore, we find that a global model can represent key characteristics of such an event. It is promising that the model generates comparable rainfall amounts under similar dynamical conditions. For this particular event, it has been shown by observational studies that the large-scale rainfall, rather than convective, plays the greater role[1]. Other extreme rainfall events with a greater proportion of convective rainfall would likely be harder for a boosted ensemble to reproduce. It would be valuable to repeat these experiments with other global models and to extend the method to use regional convective permitting models. However, it may be challenging to ensure that the most extreme boosted boundary conditions yield the most extreme local precipitation. We see this study as a first step towards looking at a wider range of research questions assessing plausible worse cases—which would require sub-daily rainfall and the floods themselves.

We are limited by ensemble size; simply by increasing the size of the ensemble, we would likely capture more extreme events. For example, some ensemble members are close to the 2-day rainfall total of the observed event, and a larger ensemble could sample further into the extreme tails and perhaps include a more extreme simulation in terms of 2-day rainfall. Likewise, selecting analogues based on the dynamical conditions prior to the observed event could also lead to more extreme members. Further work could investigate the distribution of the boosted ensemble, compared to the observations, to assess whether extremes appear constrained by ensemble size.

The boosted ensemble used in this study could be used to investigate other research questions. For example, what causes some of the simulations to show a good representation of the rainfall of an observed event, when some show very little rainfall over that region? All simulations start from very similar initial states, so when and why do the rainfall patterns diverge? The reference event cannot be a perfect analogue of the observed event—identifying reference events with deliberate choice in the difference could help understand event development. For example, the reference event we select has a smaller cut-off low and more anticyclonically tilted large-scale flow. If we were to compare with an ensemble from a reference event with, for example, less large-scale tilting, we may find systematic shifts in persistence. Perhaps ensemble boosting can aid attribution studies by helping to understand more about the causes of specific observed events. Further work exploring the differences between the wet and dry ensemble members is planned. Boosted ensembles could also help understand more about the causes of specific observed events, or whether specific dynamical situations depend upon antecedent conditions.

It is important to remember that flooding is much more than just rainfall, yet we only present storylines in terms of the rain, merely speculating about its impacts. The severity of the floods will depend on many variables, such as antecedent conditions, local level rainfall (beyond the resolution of the model), and land use. Flood generation varies greatly depending on the soil moisture levels, river conditions, and dam management state. The impacts, particularly the storyline of flood events, will depend on many more factors than the rainfall alone, and the exposure and vulnerability of communities are also important considerations. We can think through how flooding may play out in the rainfall storylines presented, but only by downscaling the rainfall to use as inputs for flood models—with a range of initial states—can we quantify the plausible impacts. Such a framework enables us to assess different climate counterfactuals but also different adaptation counterfactuals.

Society must prepare for future extreme rainfall events, but to do so we need to understand and imagine the possible characteristics of such events—not solely the rainfall magnitude. Using ensemble boosting, we have shown that, conditioned to dynamics similar to the observed event, we were perhaps unlucky in July 2021. The event was towards the upper range of rainfall persistence and spatial coverage that we find plausible, and impacted one of the more populated regions. We do find storylines that could result in greater accumulation in different regions, and we can use these to imagine some of the storylines we need to prepare for. But it is important to note that we do not assess the full range of possible extreme rainfall events over Western Europe, and only assess in the current climate. Different dynamical conditions could lead to more extreme rainfall events over western Europe, and as the climate warms we should expect extreme rainfall, and the flood disasters that follow, to occur more frequently. It is therefore essential that we can imagine and prepare for this more extreme future.

Methods

Data

We use the fifth-generation ECMWF atmospheric reanalysis of the global climate (ERA5) as a proxy for observations for 500 hPa geopotential height[50]. This reanalysis provides spatially complete gridded climate data by combining observational records with forecasting models, data assimilation systems filling gaps where direct observations are unavailable or unreliable. ERA5 is comparable to the observational dataset E-OBS[35] for the total 24-h accumulated rainfall of this event (Fig. S1). We use E-OBS for rainfall data given its dense station coverage in the study area, and the known limited ability of ERA5 to capture rainfall extremes[51]. To identify analogues of the event in the climate model, a single date is required. The event date is chosen as the day with the greatest rainfall over the region—July 14th, 2021[1]. We use daily (24-h) accumulated rainfall from E-OBS and ERA5 daily mean 500 hPa geopotential height fields.

Generating the boosted ensemble

We use the Community Earth System Model Version 2 (CESM2)[52]—starting with a 30-member CESM2 initial condition large ensemble for the current climate, 2005–2035, with historical forcings for 2005–2014 and SSP3-7.0 for 2015–2035[17]. The spatial resolution of the model is 1.25° × 0.9°, using linear regression, and the observational data is regridded to the model grid. We assess the model's ability to simulate the observed event, to ensure events with similar daily accumulated rainfall over the impacted region are possible in the model (Fig. S2). Return curves for ERA5 and E-OBS are calculated by ranking the annual maximum daily maximum rainfall, both over the study area and the maximum gridpoint within it (Fig. S2c, d). The model data is resampled to the length of the observations 10,000 times, and the 2.5–97.5% range of these is shown.

To generate the boosted ensemble, first we search for a 'reference event' from which to create the ensemble. In the CESM2 large ensemble we identify events with similar dynamical situations over Western Europe to that in ERA5 for this event; these are our atmospheric flow analogues[53–55]. We calculate analogues using pointwise Euclidean Distance between the 500 hPa geopotential height of the event day in ERA5 and each individual summer (JJA) day in the CESM2 large ensemble. Two nested domains are used—a larger domain: 30–70°N, 30°W to 30°E and a smaller domain: 42–62°N, 20°W to 10°E. Over the full ensemble, we identify the 30 closest analogues for each domain, these have a spatial correlation with the observed event of 0.78–0.90 for the larger domain and 0.89–0.96 for the smaller domain. We retain model days that feature in the 30 closest analogues of both, resulting in 7 model days with the closest dynamical patterns over the larger region and more locally. We assess the daily accumulated rainfall field of these seven most similar model days to identify the day with rainfall most like the observed event. This is used as the reference event (Figs. 1 and S5).

The reference event is used to create the boosted ensemble by re-initialising the model between 7 and 20 days before the reference event—referred to as the lead time. Bit-by-bit reproducibility on the high-performance computing environment ensures that an extreme event, which is part of an existing long simulation, can be exactly reproduced and that perturbed ensembles can be produced for the corresponding events with different lead times. At the time of the initialisation, the specific humidity q is randomly perturbed at each gridpoint by a factor of the order of 10^{-13} to generate 50 ensemble members for each lead time[17]. The boosted ensemble is output with a spatial resolution of 1.25 × 0.9° (approximately 140 × 60 km at 50°N) and daily temporal resolution.

Assessing storylines

The boosted ensemble provides physically plausible storylines. These are not necessarily the worst possibilities as they are constrained to the large-scale atmospheric conditions we impose—those similar to the observed event. There may be different dynamical set-ups that could cause even greater rainfall totals. We use the boosted ensemble to assess three specific worse scenario criteria relating to the extreme rainfall as outlined in the introduction—could the event have (1) persisted longer, (2) had a larger spatial extent, or (3) impacted a different region. For this purpose, we investigate the rainfall totals over the event region, defined as 48–52°N, 4–9°E (shown in Fig. 1). We assess the daily accumulated rainfall areal mean and the daily accumulated maximum gridbox rainfall

Table 1 | German Weather Service (DWD) rainfall warning levels[36]

Rainfall duration	Alert level 1	Alert level 2	Alert level 3	Alert level 4
Daily	>60 mm	>70 mm	>80 mm	>90 mm
2-day	>80 mm	>100 mm	>120 mm	>140 mm
3-day	>100 mm	>150 mm	>200 mm	>250 mm
4-day	>150 mm	>200 mm	>250 mm	>300 mm

for the boosted ensemble. We compare to the observed event, using the regridded reanalysis. We identify which lead times have rainfall >60 mm in a day in at least one gridpoint and ensemble member and choose only these lead times to investigate further. This level corresponds to the German Weather Service (DWD) 24 h alert level 1 (Table 1[36]). We assess three specific alternative storylines:

1. Longer persistence: To assess rainfall persistence, we calculate the maximum accumulated rainfall for 1-, 2-, 3-, and 4-day periods, over the model gridpoints in the event region. We compare these to the observed event, regridding E-OBS data to the model grid. We assess if the boosted ensemble shows more extreme rainfall over multiple days than was observed. We compare the values to the DWD alert levels for rainfall over 24–96 h, shown in Table 1.
2. Larger rainfall extent: To investigate the size of the region impacted we count the number of gridboxes in the observed event in E-OBS and the boosted ensemble with daily accumulated rainfall above two thresholds: 40 and 60 mm. We also assess the relationship between maximum gridpoint daily accumulated rainfall totals and extent.
3. Different Location: To identify whether other regions could have been impacted, we identify the gridpoint of the greatest accumulated daily rainfall in each simulation and assess the geographic spread.

Data availability

ERA5 data is available from the European Centre for Medium-Range Weather Forecasts (ECMWF), Copernicus Climate Change Service (C3S) at Climate Data Store (CDS; https://cds.climate.copernicus.eu/). We acknowledge the E-OBS dataset from the Copernicus Climate Change Service (C3S, https://surfobs.climate.copernicus.eu) and the data providers in the ECA&D project (https://www.ecad.eu).

Code availability

The code used to generate the figures in this paper and the Supplementary Materials is available from github and zenodo: https://github.com/vikki-thompson/LimburgBoosting and https://doi.org/10.5281/zenodo.15494701. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Received: 4 November 2024; Accepted: 15 May 2025;
Published online: 08 July 2025

References

1. Mohr, S. et al. A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe. Part 1: Event description and analysis. *Nat. Hazards Earth Syst. Sci*. **23**, 525–551 (2023).
2. Tradowsky, J. S. et al. Attribution of the heavy rainfall events leading to severe flooding in Western Europe during July 2021. *Clim. Change* **176**, 90 (2023).
3. Journée, M., Goudenhoofdt, E., Vannitsem, S. & Delobbe, L. Quantitative rainfall analysis of the 2021 mid-July flood event in Belgium. *Hydrol. Earth Syst. Sci.* **27**, 3169–3189 (2023).
4. Koks, E. E., van Ginkel, K. C. H., van Marle, M. J. E. & Lemnitzer, A. Brief communication: Critical infrastructure impacts of the 2021 mid-July western European flood event. *Nat. Hazards Earth Syst. Sci.* **22**, 3831–3838 (2022).
5. Ludwig, P. et al. A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe – Part 2: Historical context and relation to climate change. *Nat. Hazards Earth Syst. Sci.***23**, 1287–1311 (2023).
6. Roggenkamp, T. & Herget, J. Reconstructing peak discharges of historic floods of the River Ahr, Germany. *Erdkunde* **68**, 49–59 (2014).
7. Ommer, J., Neumann, J., Kalas, M., Blackburn, S. & Cloke, H. L. Surprise floods: the role of our imagination in preparing for disasters. *Nat. Hazards Earth Syst. Sci.***24**, 2633–2646 (2024).
8. Merz, B., Vorogushyn, S., Lall, U., Viglione, A. & Blöschl, G. Charting unknown waters—on the role of surprise in flood risk assessment and management. *Water Resour. Res.***51**, 6399–6416 (2015).
9. Merz, B. et al. Spatial counterfactuals to explore disastrous flooding. *Environ. Res. Lett.* **19**, 044022 (2024).
10. Cloke, H. Science needs to address its imagination problem – lives depend on it. *New Scientist* (23 February 2022).
11. Heinrich, D., Stephens, E. & de Perez, E. C. More than magnitude: towards a multidimensional understanding of unprecedented weather to better support disaster management. *Water Secur.* **23**, 100181 (2024).
12. Coughlan De Perez, E. C. et al. Learning from the past in moving to the future: Invest in communication and response to weather early warnings to reduce death and damage. *Clim. Risk Manag.* **38**, 100461 (2022).
13. Shepherd, T. G. et al. Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. *Clim. Change* **151**, 555–571 (2018).
14. Bevacqua, E. et al. Larger spatial footprint of wintertime total precipitation extremes in a warmer climate. *Geophys. Res. Lett.* **48**, e2020GL091990 (2021).
15. Leach, N. J., Watson, P. A., Sparrow, S. N., Wallom, D. C. & Sexton, D. M. Generating samples of extreme winters to support climate adaptation. *Weather Clim. Extremes* **36**, 100419 (2022).
16. 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).
17. Fischer, E. M. et al. Storylines for unprecedented heatwaves based on ensemble boosting. *Nat. Commun.***14**, 4643 (2023).
18. Gessner, C., Fischer, E. M., Beyerle, U. & Knutti, R. Very rare heat extremes: quantifying and understanding using ensemble reinitialization. *J. Clim.* **34**, 1–46 (2021).
19. European Flood Awareness System. EFAS Bulletin June-July 2021. Issue 2021(4). https://european-flood.emergency.copernicus.eu/sites/default/files/efasBulletins/2021/EFAS_Bimonthly_Bulletin_Jun_Jul2021.pdf (2021).
20. de Vries, I., Sippel, S., Zeder, J., Fischer, E. & Knutti, R. Increasing extreme precipitation variability plays a key role in future record-shattering event probability. *Commun. Earth Environ.* **5**, 482 (2024).
21. Thompson, V. et al. High risk of unprecedented UK rainfall in the current climate. *Nat. Commun.* **8**, 40 (2017).
22. Kelder, T. et al. and Nipon, T. Using UNSEEN trends to detect decadal changes in 100-year precipitation extremes. *npj Clim. Atmos. Sci.* **3**, 4 (2020).
23. Serinaldi, F. & Kilsby, C. G. Understanding persistence to avoid underestimation of collective flood risk. *Water* **8**, 152 (2016).
24. Serinaldi, F. & Kilsby, C. G. A blueprint for full collective flood risk estimation: demonstration for European river flooding. *Risk Anal.* **37**, 1958–1976 (2017).
25. Rottler, E., Bronstert, A., Bürger, G. & Rakovec, O. Rhine flood stories: Spatio-temporal analysis of historic and projected flood genesis in the Rhine River basin. *Hydrol. Process.* **37**, e14918 (2023).
26. van de Vyver, H., Van Schaeybroeck, B., De Cruz, L., Hamdi, R. & Termonia, P. Bias-adjustment methods for future sub-daily

precipitation extremes consistent across durations. *Earth Space Sci.* **10**, e2022EA002798 (2023).
27. Seneviratne, S. I. et al. *Weather and Climate Extreme Events in a Changing Climate. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, 2021).
28. Fowler, H. J. et al. Anthropogenic intensification of short-duration rainfall extremes. *Nat. Rev. Earth Environ.***2**, 107–122 (2021).
29. Fischer, E. M. & Knutti, R. Observed heavy precipitation increase confirms theory and early models. *Nat. Clim. Change* **6**, 986–991 (2016).
30. Rajczak, J. & Schär, C. Projections of future precipitation extremes over Europe: a multimodel assessment of climate simulations. *J. Geophys. Res. Atmos.***122**, 10773–10800 (2017).
31. Masson-Delmotte, V. et al. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, 2021).
32. Kahraman, A., Kendon, E. J., Chan, S. C. & Fowler, H. J. Quasi-stationary intense rainstorms spread across Europe under climate change. *Geophys. Res. Lett.***48**, e2020GL092361 (2021).
33. Sun, Q., Zhang, X., Zwiers, F., Westra, S. & Alexander, L. V. A global, continental, and regional analysis of changes in extreme precipitation. *J. Clim.* **34**, 243–258 (2021).
34. Myhre, G. et al. Frequency of extreme precipitation increases extensively with event rareness under global warming. *Sci. Rep.* **9**, 16063 (2019).
35. Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. & Jones, P. D. An ensemble version of the E-OBS temperature and precipitation data sets. *J. Geophys. Res. Atmos.* **123**, 9391–9409 (2018).
36. Wettergefhren Fruewarnung. www.wettergefahren-fruehwarnung.de (2024).
37. Beven, K. J. A history of the concept of time of concentration. *Hydrol. Earth Syst. Sci.* **24**, 2655–2670 (2020).
38. Strijker, B., Asselman, N., de Jong, J. & Barneveld, H. The 2021 flood event in the Dutch Meuse and tributaries from a hydraulic and morphological perspective. *J. Coastal Riverine Flood Risk* **2**, 6 (2023).
39. Fang, B., Bevacqua, E., Rakovec, O. & Zscheischler, J. An increase in the spatial extent of European floods over the last 70 years. *Hydrol. Earth Syst. Sci.* **28**, 3755–3775 (2024).
40. Thieken, A. H. et al. Performance of the flood warning system in Germany in July 2021 – insights from affected residents. *Nat. Hazards Earth Syst. Sci.* **23**, 973–990 (2023).
41. Caldas-Alvarez, A. et al. Meteorological, impact and climate perspectives of the 29 June 2017 heavy precipitation event in the Berlin metropolitan area. *Nat. Hazards Earth Syst. Sci.* **22**, 3701–3724 (2022).
42. de Bruijn, K. M. et al. Storylines of the impacts in the Netherlands of alternative realizations of the Western Europe July 2021 floods. *J. Coastal Riverine Flood Risk* **2**, 8 (2023).
43. da Costa, J., Cloke, H., Neumann, J. & Robinson, S. Unprecedented but not unexpected: a case study of the European Flood disaster 2021 in Luxembourg using a value chain approach. In *EMS Annual Meeting 2023* EMS2023-435 (EMS, 2023).
44. Roe, G. H. Orographic precipitation. *Annu. Rev. Earth Planet. Sci.* **33**, 645–671 (2005).
45. WMO. WMO and Early Warnings for All initiative. https://wmo.int/activities/early-warnings-all/wmo-and-early-warnings-all-initiative (2024).
46. Wang, H. & Coren, E. In *Storytelling to Accelerate Climate Solutions* (eds Coren, E. & Wang, H.) 1–16 (Springer, 2024).
47. Weather Channel. FloodFX witnessing the terrifying reality of storm surge. https://www.youtube.com/watch?v=Jrvt75pVePM (2024).
48. WMO. United in Science 2024. https://wmo.int/publication-series/united-science-2024 (2024).
49. Magnusson, L., Simmons, A., Harrigan, S. & Pappenberger, F. Extreme rain in Germany and Belgium in July 2021. ECWMF newsletter. https://www.ecmwf.int/en/newsletter/169/news/extreme-rain-germany-and-belgium-july-2021 (2021).
50. Bell, B. et al. The ERA5 global reanalysis: preliminary extension to 1950. *Q. J. R. Meteorol. Soc.* **147**, 4186–4227 (2021).
51. Lavers, D. A., Simmons, A., Vamborg, F. & Rodwell, M. J. An evaluation of ERA5 precipitation for climate monitoring. *Q. J. R. Meteorol. Soc.* **148**, 3152–3165 (2022).
52. Danabasoglu, G. et al. The community earth system model version 2 (CESM2). *J. Adv. Model. Earth Syst.* **12**, e2019MS001916 (2020).
53. Yiou, P. et al. Ensemble reconstruction of the atmospheric column from surface pressure using analogues. *Clim. Dyn.***41**, 1333–1344 (2013).
54. Faranda, D. et al. A climate-change attribution retrospective of some impactful weather extremes of 2021. *Weather Clim. Dyn.* **3**, 1311–1340 (2022).
55. Thompson, V. et al. Changing dynamics of Western European summertime cut-off lows: a case study of the July 2021 flood event. *Atmos. Sci. Lett.* **25**, e1260 (2024).

Acknowledgements

This research has been supported by the KNMI multi-year strategic research funding (grant name MSO-ExtremeWeather). This study was partly supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003469 (XAIDA project). Hannah Cloke acknowledges funding from the U.K.'s Natural Environment Research Council (NERC) The Evolution of Global Flood Risk (EVOFLOOD) project Grant NE/S015590/1.

Author contributions
V.T., D.C., and E.F. designed the study. U.B. performed the climate model experiments. V.T. performed the data analysis, produced the figures, and led the writing. V.T., D.C., J.O., H.L.C., and E.F. contributed to the interpretation of the results and the writing of the paper.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-025-02386-y.

Correspondence and requests for materials should be addressed to Vikki Thompson.

Peer review information *Communications Earth & Environment* thanks Pauline Rivoire and the other anonymous reviewer(s) for their contribution to the peer review of this work. Primary handling editors: Akintomide Akinsanola and Alireza Bahadori. A peer review file is available.