Weather and Climate Extremes 33 (2021) 100350

Contents lists available at ScienceDirect

Weather and Climate Extremes

journal homepage: www.elsevier.com/locate/wace

# Physical storylines of future European drought events like 2018 based on ensemble climate modelling

Karin van der Wiel *, Geert Lenderink, Hylke de Vries

*Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands*

ARTICLE INFO

*Keywords:*
Climate change
Drought 2018
Extreme weather
Future weather
Large ensembles
Storylines

ABSTRACT

In the aftermath of observed extreme weather events, questions arise on the role of climate change in such events and what future events might look like. We present a method for the development of physical storylines of future events comparable to a chosen observed event, to answer some of these questions. A storyline approach, focusing on physical processes and plausibility rather than probability, improves risk awareness through its relation with our memory of the observed event and contributes to decision making processes through their user focus. The method is showcased by means of a proof-of-concept for the 2018 drought in western Europe. We create analogues of the observed event based on large ensemble climate model simulations representing 2 °C and 3 °C global warming scenarios, and discuss how event severity, event drivers and physical processes are influenced by climate change. We show that future Rhine basin meteorological summer droughts like 2018 will be more severe. Decreased precipitation and increased potential evapotranspiration, caused by higher temperatures and increased incoming solar radiation, lead to higher precipitation deficits and lower plant available soil moisture. Possibly, changes in atmospheric circulation contribute to increased spring drought, amplifying the most severe summer drought events. The spatial extent of the most severe drought impacts increases substantially. The noted changes can partly be explained by changes in mean climate, but for many variables, changes in the relative event severity on top of these mean changes contribute as well.

## 1. Introduction

In the aftermath of an extreme weather or climate event, commonly questions arise concerning the influence of climate change on the event. These questions are often related to event probability: "Has the event become more or less likely due to climate change and will events like this become (even) more (or less) likely in the future?". But also questions related to the physical processes leading up to the event and its severity are often asked: "How did this event come together and how bad will future events be?". Scientific studies seeking to answer these questions, i.e. observed event-based studies, have the advantage that they link future climate change projections to human experiences and memories. The results and implications may therefore be better understood than more abstract climate change projections (Shepherd et al., 2018; Sillmann et al., 2019). Here, we aim to develop a method to answer such societal or user questions. By defining simulated event analogues in present-day and future climate conditions, following a so-called storyline approach (Hazeleger et al., 2015; Shepherd et al., 2018), we compare an observed event to possible future events and identify relevant differences.

Storylines, defined by Shepherd et al. (2018) as *"physically self-consistent unfoldings of past events or of plausible future events"*, improve risk awareness through offering episodic knowledge[1] and strengthen decision making processes through their focus on user-relevant issues (Hazeleger et al., 2015; Shepherd et al., 2018). The strength of storylines lies in the focus on understanding, and plausibility, of the (physical processes leading to the) event of interest. Estimates of probability or likelihood are often not provided, though many users will also have a strong interest in information on event probability. There is some ambiguity regarding the word 'storyline' in the climate science literature. We take an event-based approach, and use IPCC's definition of what is an extreme event: *"An extreme weather event is an event that is rare at a particular place and time of year. ... When a pattern of extreme weather persists for some time, such as a season, it may be classed as an extreme climate event, especially if it yields an average or total that is itself extreme (e.g., drought or heavy rainfall over a season)"*. IPCC (2014). The resulting storylines describe physically plausible future events that are in some pre-defined way comparable to an observed event of interest, these may then be used to inform policies or stress-test systems.

* Corresponding author.
*E-mail address:* wiel@knmi.nl (K. van der Wiel).

[1] One of two classes of memory defined by Tulving (1972): semantic knowledge is factual information, episodic knowledge is memory of past events.

https://doi.org/10.1016/j.wace.2021.100350
Received 11 November 2020; Received in revised form 28 May 2021; Accepted 29 June 2021
Available online 7 July 2021

There are two main methods for creating event-based storylines using climate models: performing simulations in which a model is constrained to reproduce certain aspects of the observed event, or alternatively, selecting analogues of the observed event from unconstrained model simulations. We apply the second method, and construct event analogues by selecting similar simulated events from existing large ensemble climate model simulations. The selection of events will be done based on a quantitative event metric to ensure robust and unbiased selection. Ideally this metric is closely related to observed impacts and relevant to users, which is important because extreme meteorological conditions do not necessarily map onto extreme natural or societal impacts (Van der Wiel et al., 2020). The resulting storylines provide insights into future events like the observed event, but these can potentially be driven by different meteorological (dynamical) conditions or result in more/less severe societal or natural impacts. The alternative method, sometimes referred to as nudging or Pseudo Global Warming experiments (PGW, Schär et al., 1996), ensures dynamic event similarity. However, the best method of perturbation of thermodynamics or boundary and initial conditions is still under debate (Sillmann et al., 2019) and forced changes in event circulation patterns are, by design, not considered (Faranda et al., 2020). Rare event algorithms which apply importance sampling in climate modelling (Ragone et al., 2018; Yiou and Jézéquel, 2020; Ragone and Bouchet, 2020) possibly sit between the above two methods, through their focus on dedicated modelling of extreme events.

A proof of concept for our developed method is provided by means of a case study on the exceptionally hot and dry European summer of 2018 ('Drought'18' hereafter). The summer of 2018 was extremely warm and dry in western Europe (Fig. 1), impacting agriculture, ecosystems, river transport and recreation (Ecorys et al., 2018; Philip et al., 2020b; Zscheischler and Fischer, 2020). A positive phase of the North Atlantic oscillation, global warming and for part of the summer a circumglobal near-stationary Rossby wave led to persistent warm and dry conditions (Drouard et al., 2019; Kornhuber et al., 2019). The likelihood and severity of this event has been shown to increase in response to anthropogenic climate change (Philip et al., 2020b; Wehrli et al., 2020; Zscheischler and Fischer, 2020), which is in line with regional projections of drought under global climate change (IPCC, 2013; Samaniego et al., 2018; Spinoni et al., 2018). Following the storyline rationale, we investigate the physical processes leading to the Drought'18 in the Rhine basin and evaluate how these change in warmer future climates. The analysis is mostly limited to meteorological drought processes, though soil moisture impacts are considered. Users' personal memories of the event (i.e. episodic knowledge), maybe of yellowing lawns, agricultural losses and shipping difficulties (Ecorys et al., 2018), will help them put our results in context. The storylines will contribute to answering user questions like "How much worse can droughts get (e.g. more severe, more wide spread, longer in duration)?", and therewith aid decision making for sustainable climate change adaptation in the region.

In this paper, we describe and apply a method for creating robust event-based storylines using large ensemble climate model simulations. We construct analogues of a chosen observed extreme event using large ensemble simulations of different states of global climate, the event analogues in different climatic states are then compared to provide event-specific climate change information. The analogue construction procedure, including relevant considerations, assumptions and modelling choices, is described in Section 2. The results are split in a comparison of the observed event to Drought'18 analogues in the present-day climate (Section 3.1), a description of Drought'18 analogues in warmer climates (Section 3.2), and a comparison of the trends in our climate model to trends in a multi-model ensemble (Section 3.3). In the Discussion (Section 4) we discuss lessons learned with regards to the large ensemble-based storyline method, finally conclusions on future summers like Drought'18 are provided in Section 5.

**Table 1**
Three metrics for event selection.

| Metric | Time series | Period | Statistic | Selection |
|---|---|---|---|---|
| M1 | PR deficit | August–October | Mean value | Highest |
| M2 | PR deficit | June–August | Regression slope | Highest |
| M3 | PR deficit | April–October | Temporal correlation | Highest |

## 2. Method

### 2.1. Event definition

To allow an unbiased selection of similar simulated events, the observed event of interest and all simulated events have to be compared quantitatively (Sillmann et al., 2019). This quantitative comparison helps to select events in an objective manner and avoids cherry picking of simulated events. We define a metric that describes the event of interest by choosing a variable and its spatial and temporal properties. Once the event definition is formulated, event selection from model simulations is done by means of maximum similarity to the observed event. The metric choices can be somewhat arbitrary, hence the dependence of the results on event definition and metrics should be analysed. The methods of event definition and the considerations involved are similar to those in class-based event attribution studies (Cattiaux and Ribes, 2018; Philip et al., 2020a; Van Oldenborgh et al., 2021).

We defined the Drought'18 event by means of a time series of precipitation deficit in the Rhine basin (Fig. 1a). Precipitation deficit is a common drought indicator in the Netherlands, and is defined to be the cumulative difference between potential evapotranspiration and precipitation in the growing season (April onward), set to zero if smaller than zero (Beersma et al., 2004; Sluijter et al., 2018; Philip et al., 2020b). It can be interpreted as a measure of drought somewhere between meteorological drought (low precipitation) and agricultural drought (low soil moisture). We use the Makkink formula (De Bruin and Lablans, 1998) to compute daily values of potential reference evapotranspiration from 2 m temperature, incoming solar radiation and surface pressure. The ERA5 reanalysis (Hersbach et al., 2020) is used to characterise the event as it occurred in the real world. Within the ERA5 record (1979–2018) 2018 has the highest peak value of precipitation deficit, i.e. it is the most extreme event by this measure.

Next, we defined three metrics for event selection based on this time series (Table 1, Fig. 2), and created future Drought'18 analogues based on these metrics. The metrics capture different aspects of Drought'18: metric 1 selects droughts with high values of end-of-summer precipitation deficit (Fig. 2a), metric 2 focuses on the steep increase of deficit in the months June to August (Fig. 2b), and metric 3 focuses on the precise seasonal progression of drought including the relative late start of extreme drought conditions (Figs. 1a and 2c). Depending on specific user needs or vulnerability, one metric and hence storyline may be more relevant than others. Note that our analogues are 'hazard-analogues', similar in terms of meteorological extreme (i.e. meteorological drought), not analogues in terms of atmospheric circulation.

### 2.2. Large ensemble climate model simulations

Generally, the interest in event-based storylines is for observed extreme events. Since such events are rare by definition, long climate model simulations or large climate model ensembles are required to sample similarly extreme simulated events. Large ensemble experiments, consisting of many repetitions of the same experiment (Deser et al., 2020), allow the direct analysis of extreme events without the need for statistical extrapolation of the tail of the distribution (Van der Wiel et al., 2018, 2019). This is a requirement for our physical based storyline approach, since we are interested in creating storylines describing the physical processes involved in the event, this information cannot be obtained by statistical extrapolation.

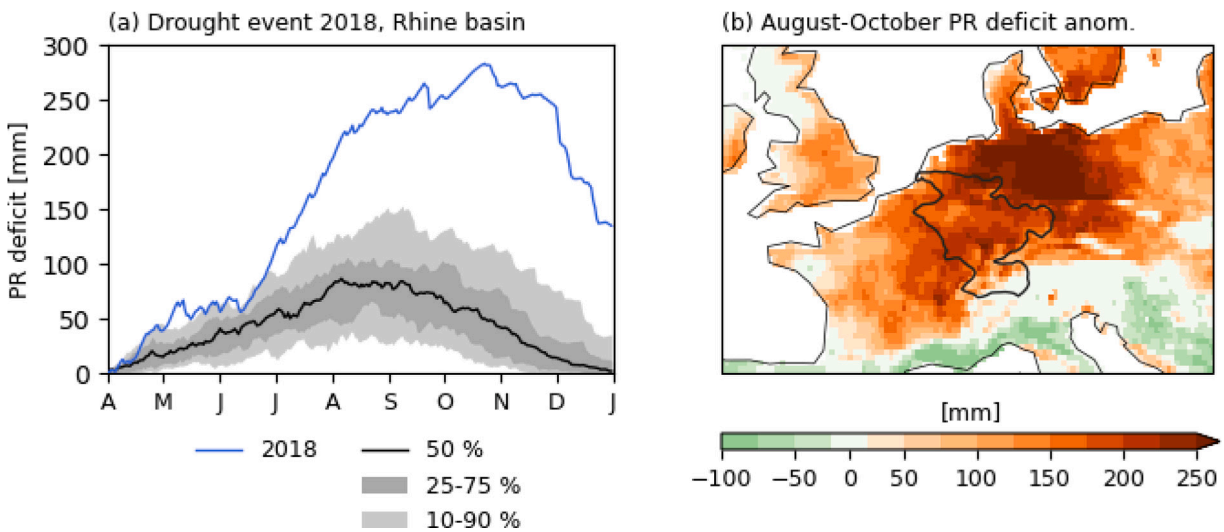


**Fig. 1.** Precipitation deficit [mm] during the 2018 drought. (a) Rhine basin mean time series in blue and associated climatology in grey/black (10-25-50-75-90th percentiles). (b) Map of August–October mean precipitation deficit anomaly [mm] in shaded colours, Rhine basin outlined in grey. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

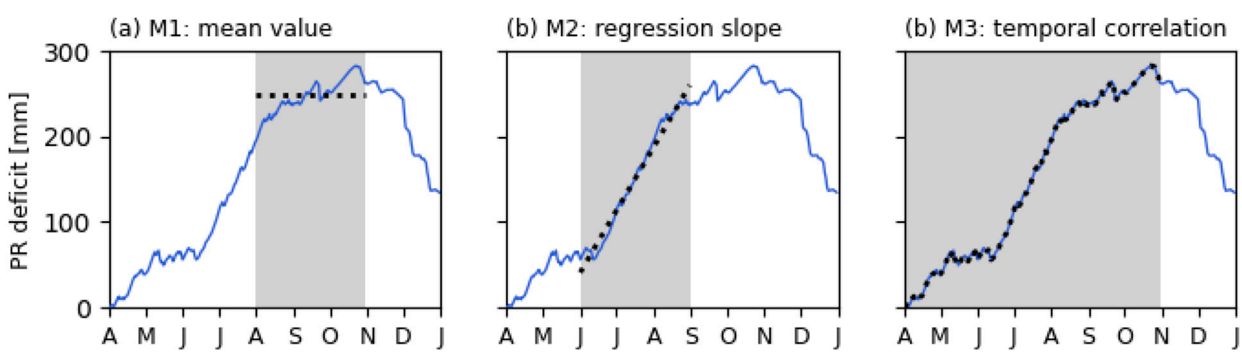


**Fig. 2.** Schematic diagram of the three metrics for event selection. The Rhine basin mean time series of precipitation deficit [mm] during the 2018 drought in blue, grey shading shows the period of evaluation, black dotted line the measure considered by the metric. (a) Metric 1, mean value of precipitation deficit over August–October, (b) metric 2, regression slope of precipitation deficit from June–August, (c) metric 3, temporal correlation (Pearson's $r$) of the precipitation deficit time series from April–October. All metrics computed based on daily data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

A slight advantage of the selection method over nudged or PGW simulations, is that it can be based on existing climate model data sets. Several large ensemble data sets are available, for example in the Multi-Model Large Ensemble Archive (Deser et al., 2020). The event selection method is therefore less computationally expensive, and can thus be applied more routinely, and potentially within a shorter time frame if analysis time is of relevance (as in rapid attribution studies, Stone and Hansen, 2016; Van der Wiel et al., 2017; Van Oldenborgh et al., 2021).

Here we used large ensemble simulations (Van der Wiel et al., 2019) created using the EC-Earth global coupled climate model (v2.3, ~100 km horizontal resolution; Hazeleger et al., 2012). Three ensembles of 2000 model years each are available, each representing a different state of global climate. The present-day experiment has an absolute global mean surface temperature (GMST) similar to that observed in 2011–2015; the 2C and 3C-warming experiments have an absolute GMST 2 and 3 °C warmer than observed pre-industrial temperatures (1851–1899), respectively. A more detailed description of experimental setup is provided in Van der Wiel et al. (2019).

We construct simulated *analogues* of the observed event by selecting the 20-most similar simulated events from each large ensemble and taking the composite mean. For each of the 2000 simulated years we compute the metrics defined in Section 2.1 and select the 20 highest scoring events. These are the years with highest August to October mean precipitation deficit (metric 1, Fig. 2a), the years with steepest increase of precipitation deficit between June and August (metric 2, Fig. 2b), or the years that have the most similar seasonal progression, i.e. highest temporal correlation coefficient, (metric 3, Fig. 2c). We investigate composite means to bring out the climate change signal and limit the influence of random weather noise. Our choice to build the analogues from 20 events is a compromise between increasing event return period and reducing noise. We performed a sensitivity analysis to show the considerations involved in this choice (Appendix A). In general, the larger the ensemble of simulated data, the higher the chances of finding multiple comparable simulated events to the observed event (in terms of event characteristics, severity and return period). This then allows the building of an analogue using more events, which improves the chance of finding a robust, coherent and interpretable climate change signal. The composite analogues for the three metrics are analysed individually, this allows us to investigate the dependence of the result on the method of event selection.

The analogues are used to investigate how event severity, event drivers and physical processes are influenced by further climate change. Before doing so, we checked if EC-Earth was fit for purpose by analysing model biases in the seasonal cycle of relevant variables. It is not possible to directly evaluate model bias in the extreme events of interest given the short observational record with a limited number of such events (Wagener et al., 2010; Van der Wiel et al., 2018; Kelder et al., 2020). Regionally in Europe, EC-Earth is too cold and too wet in spring and summer (Hazeleger et al., 2012), which results in an underestimation of precipitation deficits in the Rhine basin (Fig. 3a).

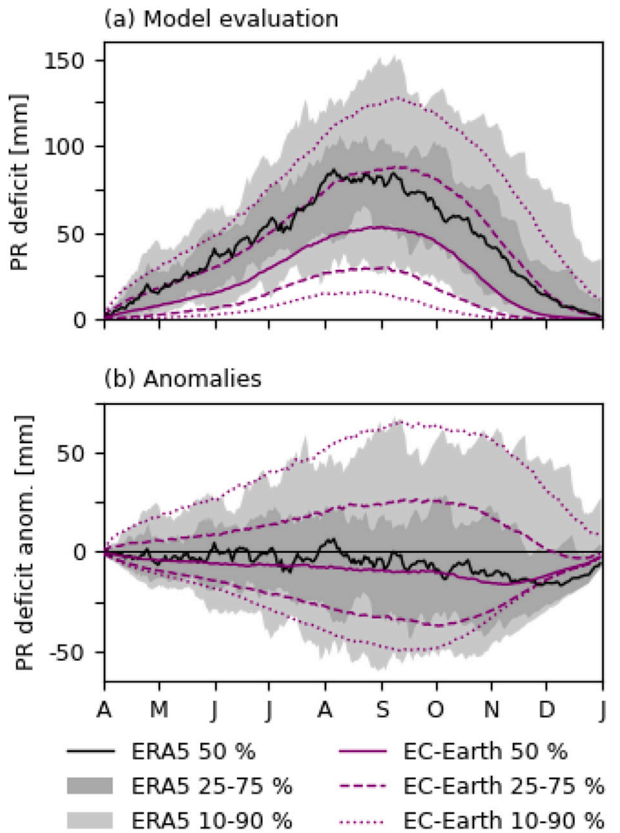


**Fig. 3.** EC-Earth model evaluation based on 10-25-50-75-90% percentiles. In grey/black shading ERA5 data (40 years), in purple lines EC-Earth present-day model simulations (2000 model years). (a) Precipitation deficit [mm] in the Rhine basin, (b) the same but showing anomalies rather than actual values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

When removing the climatological mean bias, model variability in precipitation deficit is close to observed variability (Fig. 3b). We will therefore select events and analyse physical processes based on daily anomaly values, these are computed using the observed climatology for observational data and the model climatology for model data. Climatologies are computed using all available years (1979–2018 for ERA5, or 2000 years for the model data) and by applying a 15-day running mean.

### 2.3. CMIP6 comparison

Generally, multi-model experiments perform better than single-model experiments (Reichler and Kim, 2008), and hence it would be good to repeat the event-analogue study using the models in the CMIP6 archive. Unfortunately, this is not possible since CMIP6 does not contain the large ensemble experiments required to select enough comparable events. As an alternative, we compared the climate change response of the EC-Earth model to that of the CMIP6 archive. Because the EC-Earth ensembles were defined by means of absolute GMST (Section 2.2), we compared similar GMST-based periods in the CMIP6 archive using 35 historical+SSP5-8.5 experiments[2] (Kriegler et al., 2017) available on the ETH server (Brunner et al., 2020). For each model, we selected the 30-year period of present-day GMST and the 30-year period of 3C-warming GMST, and computed the changes in the distribution of monthly temperatures and precipitation. We then compared the changes in the EC-Earth ensembles to the range of changes found for the 35 CMIP6 models.

## 3. Results

### 3.1. Comparison of present-day analogue and observed event

We calculate the event metrics for all years in ERA5 and in the EC-Earth present-day large ensemble (Fig. 4a,b). Measured by metrics 1 and 2, the Drought'18 event is the most extreme in the reanalysis dataset (metric 3 does not aim to capture drought severity). For these metrics, three summers in the EC-Earth present-day large ensemble are more extreme than the observed event, this shows that large ensemble data is indeed needed to capture events of comparable severity or return period to the observed event. We select the 20 simulated events that score highest or are most extreme in each metric and compute the composite mean *analogue* of the observed event. All analyses will be done for the three metrics separately, to allow a sensitivity analysis into the dependence of the results on the method of event selection.

Fig. 4c,d show time series of precipitation deficit anomalies for the observed event (ERA5) and the 20 individual selected simulated events (EC-Earth). For both metrics (and metric 2, not shown here) the selected model events follow the observed seasonal progression of Drought'18. As may be expected, events selected by means of metric 1 have higher absolute precipitation deficit anomalies and events selected by means of metric 3 remain closer to the observed seasonal progression, for example by capturing the relatively low deficits from April to mid-June. Each of the metrics captures a unique part of the observed Drought'18 event and there is no metric that outperforms the others in an obvious way. Note that there is quite some overlap in the selection of simulated summers for the different metrics: comparing two metrics nine or ten of the selected summers are the same, five summers were selected for all three metrics.

Next we compare time series of different drought-related variables of the observed event and the simulated composite analogue. The Drought'18 event was characterised by below normal precipitation (Fig. 5c) and above normal potential evapotranspiration (Fig. 5d). The latter is the result of higher than normal 2 m temperatures and higher than normal incoming solar radiation (Fig. 5e,f). The combination of low precipitation and high potential evapotranspirations leads to a higher than normal precipitation deficit (Fig. 5a) and lower than normal plant available soil moisture, here shown by means of the soil wetness index[3] (SWI, Fig. 5b). This has consequences for the surface heat fluxes, in the summer months the latent heat flux is lower than normal, resulting in a higher than normal sensible heat flux (Fig. 5g,h). The analogue shows the same characteristics as the observed event, though the time series of the observed event are more variable than the time series of the analogue. This is a result of the composite method that is used for the creation of the analogue, natural variability for 20 events (analogue, composite mean) is smaller than for one event (observed event). For all variables the observed event lies within the envelope of variability of the 20 individual selected simulated events, taking into account some random temporal shifts for e.g. temperature

[2] Models included (one member each): ACCESS-CM2, ACCESS-ESM1-5, AWI-CM-1-1-MR, BCC-CSM2-MR, CAMS-CSM1-0, CanESM5-CanOE, CanESM5, CESM2, CESM2-WACCM, CIESM, CMCC-CM2-SR5, CNRM-CM6-1, CNRM-ESM2-1, EC-Earth3, EC-Earth3-Veg, FGOALS-f3-L, FGOALS-g3, FIO-ESM-2-0, GFDL-CM4, GFDL-ESM4, GISS-E2-1-G, HadGEM3-GC31-LL, HadGEM3-GC31-MM, INM-CM4-8, INM-CM5-0, IPSL-CM6A-LR, KACE-1-0-G, MIROC6, MIROC-ES2L, MPI-ESM1-2-HR, MPI-ESM1-2-LR, MRI-ESM2-0, NESM3, NorESM2-MM, UKESM1-0-LL.

[3] The soil wetness index (SWI) represents the fraction of total plant available water, it is the scaled soil moisture amount between permanent wilting point and field capacity (Aalbers et al., 2021).

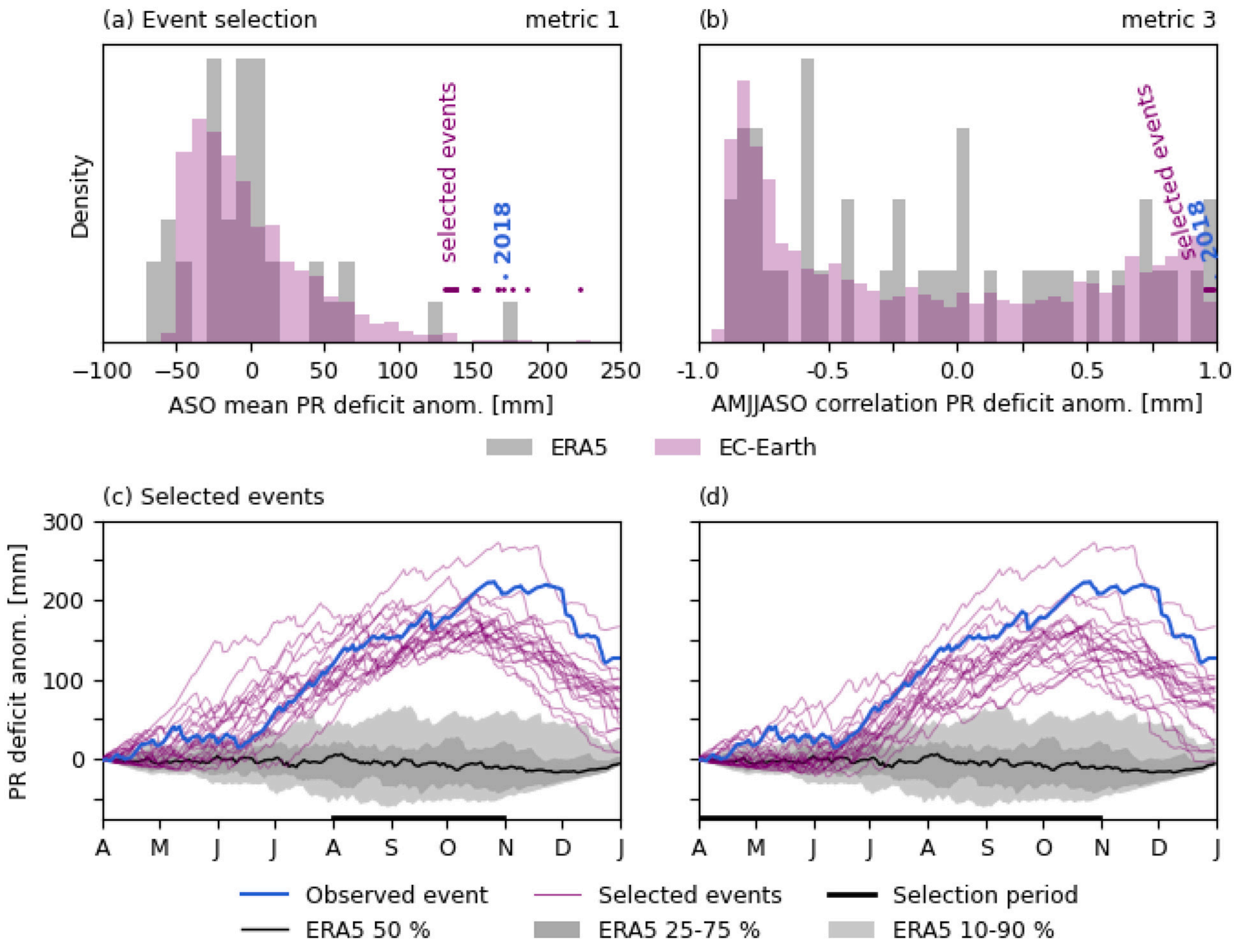


**Fig. 4.** (a,b) Histograms of the selection metrics for ERA5 in grey and the EC-Earth present-day ensemble in purple. The blue dot shows the 2018 event in ERA5, purple dots show the 20 selected simulated events. (c,d) Time series of precipitation deficit anomalies for the observed event in blue and the 20 selected simulated events in purple. Grey/black shading shows the observed climatology, as in Fig. 3b. The horizontal black bar shows the selection period. Each column shows the result of a different event selection metric, left for metric 1 and right for metric 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and radiation, and temporal variability of the observed event is comparable to the temporal variability of individual selected model events. Fig. 5 shows the analogue created using selection metric 1, similar agreement between observed event and analogues is found for the other event selection metrics not shown here.

Given the high degree of similarity between the Drought'18 simulated analogues and the observed event, we have confidence that we can use the analogue approach in conjunction with the EC-Earth large ensembles for a study of large-scale droughts in western Europe. We accounted for model biases by investigating anomalies (Section 2.2), indeed our results show that the climatological state of the model is not completely relevant, rather it is the variability around this state that matters. However, if there are thresholds in the system, e.g. for soil wetness which cannot go below a certain value, the climatological bias may negatively impact the temporal development of anomalies. We will continue to investigate the effects of climate change on event severity, event drivers and the physical processes causing and contributing to the event.

### 3.2. Drought'18 under 2 °C and 3 °C global climate change

We repeat the event selection procedure for the 2C- and 3C-warming large ensembles of the EC-Earth model. As for the present-day ensemble, we compute the event metric for all years in each ensemble, select the 20 most similar or most extreme simulated events and compute the composite mean analogue.

The Drought'18 analogue has a higher absolute precipitation deficit in the 2C- and 3C-warmer climates than in the present-day climate (Fig. 6a–c). This intensification of future drought is found independent of the event selection metric. The timing and magnitude of largest changes do however depend on the method of event selection. For future Drought'18 analogues selected by means of metric 1, lower precipitation and higher potential evapotranspiration from April through to August lead to a significant increase of precipitation deficit (Figs. 6a,d,g and S3). The dry conditions in April and May are due to a high pressure system centred over the United Kingdom (Fig. 7a). For the future events this high pressure system is slightly larger, amplifying drought conditions: the spring months are drier, warmer and sunnier (Figs. 6d,g,j and S3). This dry start of the event is then amplified by warm and dry summer months (Fig. 6d,g,j). The precipitation deficit increases, and does so at a higher rate in the future analogues than in the present-day analogue (Fig. 6a). This summertime enhanced drying cannot be related to a strengthening of the high pressure system (Fig. 7d), we only find a slight southeastward displacement of the high pressure system. From September onward the future events are wetter than the present-day event, this may be related to weaker high pressure conditions at the end of summer (Fig. 7g).

The described climate change signal in the spring months for analogues constructed using metric 1 is not found when analogues are selected by means of metrics 2 or 3 (Fig. 6b,c). This can, at least partly, be explained by the absence of an high pressure system (Fig. 7b,c), consequently both anomalies of precipitation and potential evapotranspiration and the changes therein are small (Figs. 6e,f,h,i and

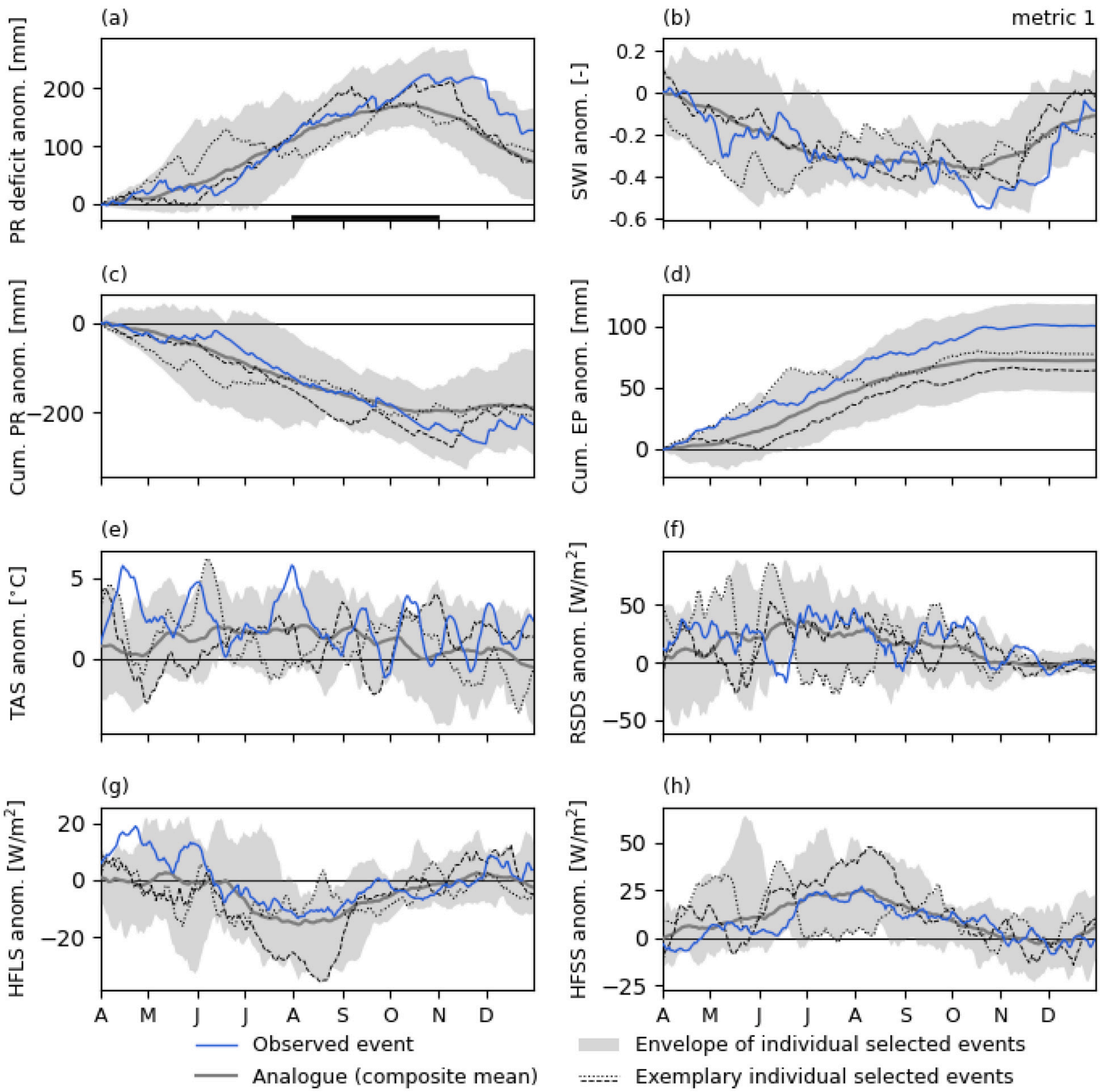


**Fig. 5.** Time series of drought-related variables for the observed event in blue and for the analogue in grey. Grey shading indicates the envelope of the 20 individual selected simulated events, black dashed and dotted lines show time series of 2 exemplary individual selected simulated events. Event selection by means of metric 1. Variables shown, anomalies of: (a) precipitation deficit [mm], (b) soil wetness index [unitless], (c) cumulative precipitation [mm], (d) cumulative potential evapotranspiration [mm], (e) 2 m temperature [°C], (f) incoming solar radiation [W $m^{-2}$], (g) latent heat flux [W $m^{-2}$], (h) sensible heat flux [W $m^{-2}$]. The horizontal black bar in (a) shows the selection period. A 15-day running mean was applied in panels (e–h). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

S3). Instead, future Drought'18 analogues selected by metric 2 are characterised by very warm and sunny conditions in June, July, and August (Figs. 6k and S3). Higher values of potential evapotranspiration (Figs. 6h and S3), lead to a faster increase of the precipitation deficit in the future analogues than in the present-day analogue (steeper line in Fig. 6b), resulting in a higher absolute precipitation deficit and an earlier peak of precipitation deficit. This change is not due to a change in circulation patterns (Fig. 7e,h), a possible explanation may be found in a positive feedback through high values of sensible heat flux (Fig. S3).

Future Drought'18 analogues selected by metric 3 show an absolute increase in precipitation deficit (solid lines in Fig. 6c) due to climatological mean changes (Fig. S4), we find no drought intensification beyond that change (dashed lines in Fig. 6c). This may be caused by the method of event selection; for metric 3 we selected analogues based on the temporal correlation between simulated precipitation deficit anomalies and the observed event, faster increases of precipitation deficit, as found for the other analogues, are penalised and are therefore less likely to be selected. This metric and analogue may therefore not be as suitable for providing information on future changes in physical processes and event severity. The long-lasting drought conditions of the observed Drought'18 event (until November, Fig. 5a) is captured best by metric 3. The analogue selected by metric 3 remains driest in September and October (Fig. S1), this dry end of the summer is amplified in the 2C- and 3C-warming analogues (Figs. 6c and S2).

We have described the effects of future climate change on the variables leading to meteorological drought conditions, next we shortly analyse the impact of these future Drought'18 events on plant available soil moisture. Increased future precipitation deficits lead to an absolute decrease of available soil moisture (Fig. 8a–c). A large part of this

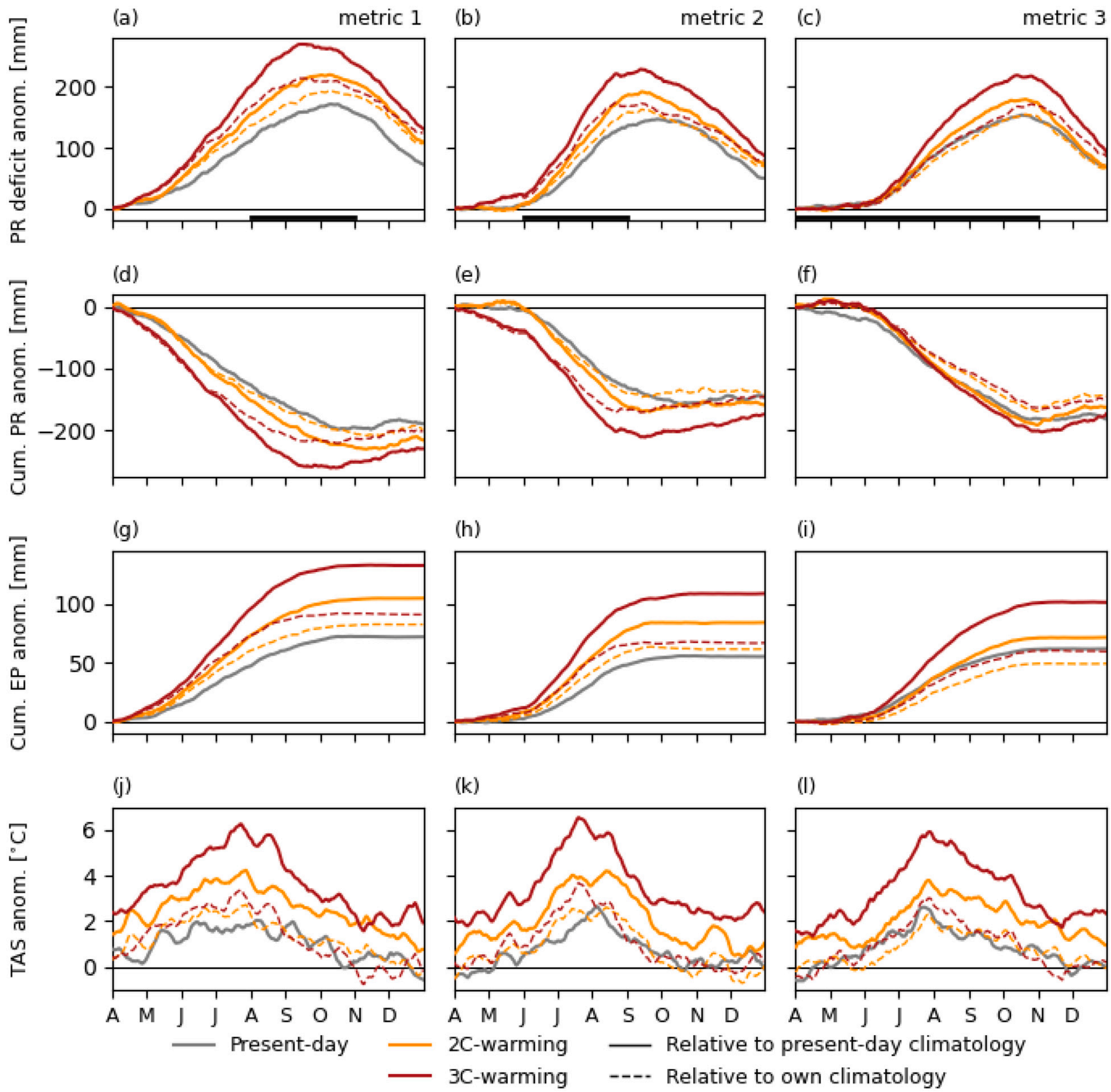


**Fig. 6.** Time series of drought-related variables for the Drought'18 analogues in different states of global climate. Present-day analogue in grey, 2C-warming in yellow, 3C-warming in red. Solid lines show anomalies relative to the present-day model climatology, dashed lines show anomalies relative to the new climatology (2C- or 3C-warming). Each column shows the result of a different event selection metric, from left to right metrics 1, 2 and 3. Variables shown, anomalies of: (a–c) precipitation deficit [mm], (d–f) cumulative precipitation [mm], (g–i) cumulative potential evapotranspiration [mm], (j–l) 2 m temperature [°C]. The horizontal black bar in (a–c) shows the selection period. A 15-day running mean was applied for 2 m temperature. Similar graphs for additional variables are included in Fig. S3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

change is due to a climatological mean change to lower SWI values (from June onward, Fig. S4). Lower soil moisture values lead to a reduced latent heat flux and increased sensible heat flux (Fig. S3), potentially contributing to increased probabilities of heatwaves and enhancing drought conditions. If we compare SWI values of future Drought'18 analogues to their 'own' future climatology (grey line and dashed coloured lines in Fig. 8a–c), we find that the analogues selected with metrics 1 and 2 lead to relatively more intensive soil moisture drought conditions in spring, but relatively less intensive soil moisture droughts at the end of summer. The distinction between changes in mean climate and changes in event severity on top of this change in mean climate can help inform the design of adaptation policies. The dry end of summer captured in the analogue of metric 3 means soil moisture recovery is slower, the soil moisture drought conditions last longer and are more severe in the future analogue.

The spatial extent of area experiencing drought conditions in the present-day Drought'18 event is not limited to the Rhine basin, though the most extreme SWI values are found there (Fig. 8d–f). This area, e.g. quantified by SWI values lower than the 10th percentile, increases substantially for future Drought'18 events (Fig. 8g–i). Again, this change can for a large part be explained by a climatological mean change towards lower SWI values.

### 3.3. Intermodel comparison of mean climate change response

The results of the analyses in Section 3.2 depend on the specific climate change response of the EC-Earth model during drought events. To put these results in context of other global climate models, we compare the EC-Earth climate change response to that from 35 models

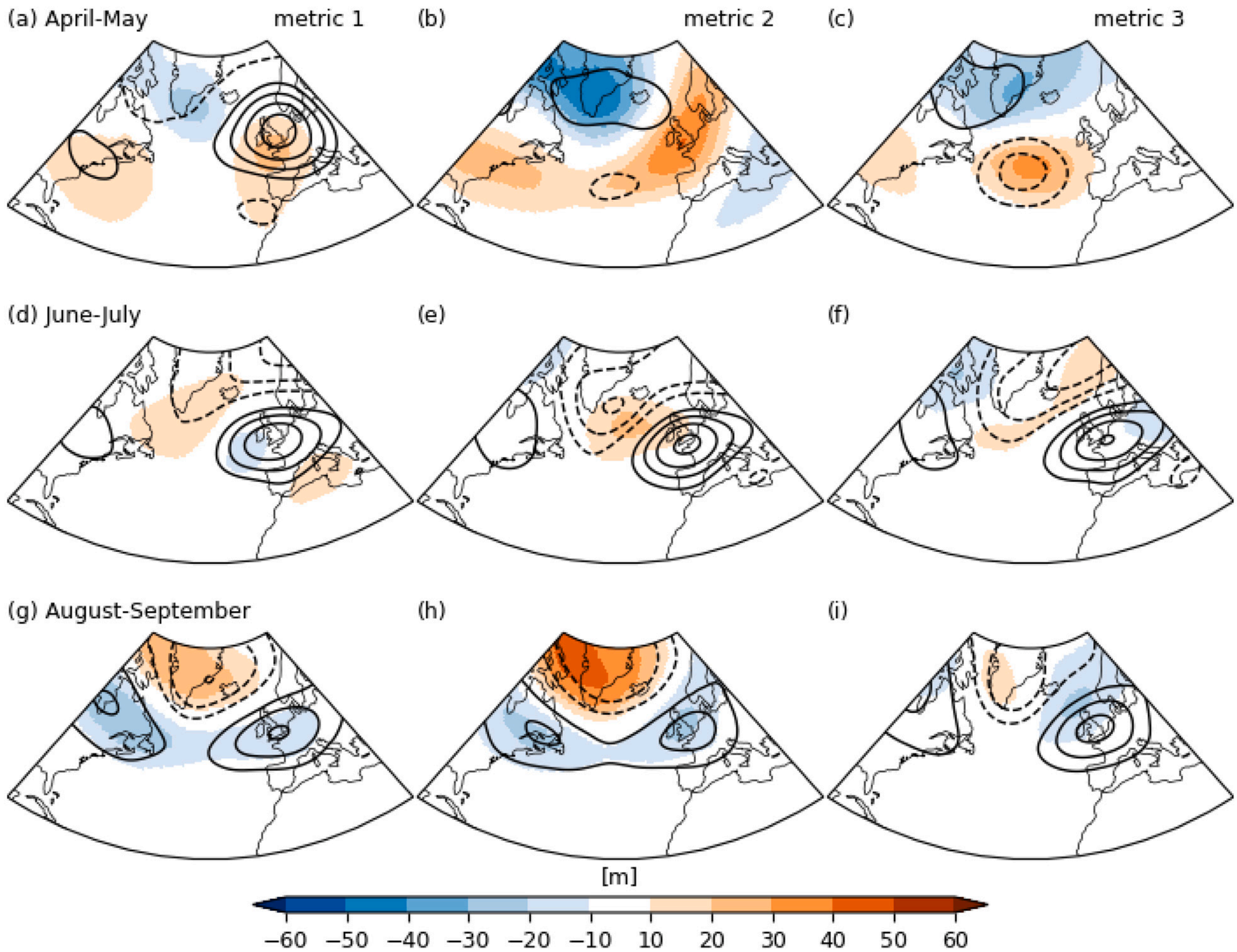


**Fig. 7.** Maps of 500 hPa geopotential height anomalies [m] for the Drought'18 analogues. Contour lines show the circulation anomalies for the present-day analogue (contour interval 10 m, positive anomalies solid, negative anomalies dashed, zero contour omitted). Shaded colours show the difference in circulation anomalies between the 3C-warming analogue and the present-day analogue (3C minus PD), colour interval same as contours. Each column shows the result of a different event selection metric, from left to right metrics 1, 2 and 3. Time periods shown: (a–c) April–May, (d–f) June–July, (g–i) August–September. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in the CMIP6 archive. We focus on spring and summer 2 m temperature and precipitation in the Rhine basin as these have been shown to be most relevant for the Drought'18 summers (Fig. 9). For these variables and months the EC-Earth large ensembles fall within the 25 to 75% percentile range of the CMIP6 response, indicating that our ensembles do not display unique or unexpected behaviour. This gives us confidence that our results, though based on a single climate model, provide a fair look into future Drought'18 events under additional climate change forcing. It is not feasible to repeat the analogue creation procedure for all models in the CMIP6 archive, the method requires a large ensemble of simulations to be available for each individual model, rather than an ensemble of different models.

The CMIP6 models show warming in all seasons, with stronger warming in summer than in spring (Fig. 9a–c). Furthermore, the warm tail of the distribution warms more than the median and cool tail. EC-Earth warming falls between the 25th and 50th percentile of CMIP6, from June to September it also shows higher warming in the warm tail, this agreement is not found in the spring months. The sign of spring precipitation change is uncertain, EC-Earth very closely follows the CMIP6 median response (Fig. 9d–f). In June–July the dry tail of precipitation dries by about 20% in CMIP6, the August–September response is comparable though slightly more modest. For both periods EC-Earth very closely follows the median response.

The temperature change in the Drought'18 analogues corresponds most closely with the warm tail of the CMIP6 changes, as might have been expected (Fig. 9a–c). As noted earlier, metric 3 warms the least in the early summer, though it shows slightly stronger warming at the end of summer than the other metrics. In terms of precipitation change, metrics 1 and 2 fall at the lower end of the CMIP6 range in April–May (Fig. 9d) and correspond to the CMIP6 median change of the dry tail in June–July (Fig. 9e). The dry end of summer for metric 3 corresponds to the dry tail of CMIP6. The 3C-warming analogues based on metrics 1 and 2 have close to normal precipitation in September (Figs. 6d,e and S3), it is therefore not expected that the climate change response agrees with the dry tail of the distribution. Compared to the CMIP6 ensemble, EC-Earth is a middle-of-the-road model in terms of its climate change response in these variables. Likely therefore there are models that have stronger responses in their Drought'18 analogues, with larger increases of precipitation deficits and larger decreases in available soil moisture than noted in Section 3.2.

## 4. Discussion

We have presented a method for the creation of simulated analogues of observed events from large ensemble climate model data, and used it to create physical storylines of future Drought'18 events. This method of storyline development complements the alternative method of dedicated nudging or PGW experiments (as for example Wehrli et al. (2020) and Aalbers et al. (2021) have done for the summer of 2018). In this section we will discuss the strengths and limitations of our method,

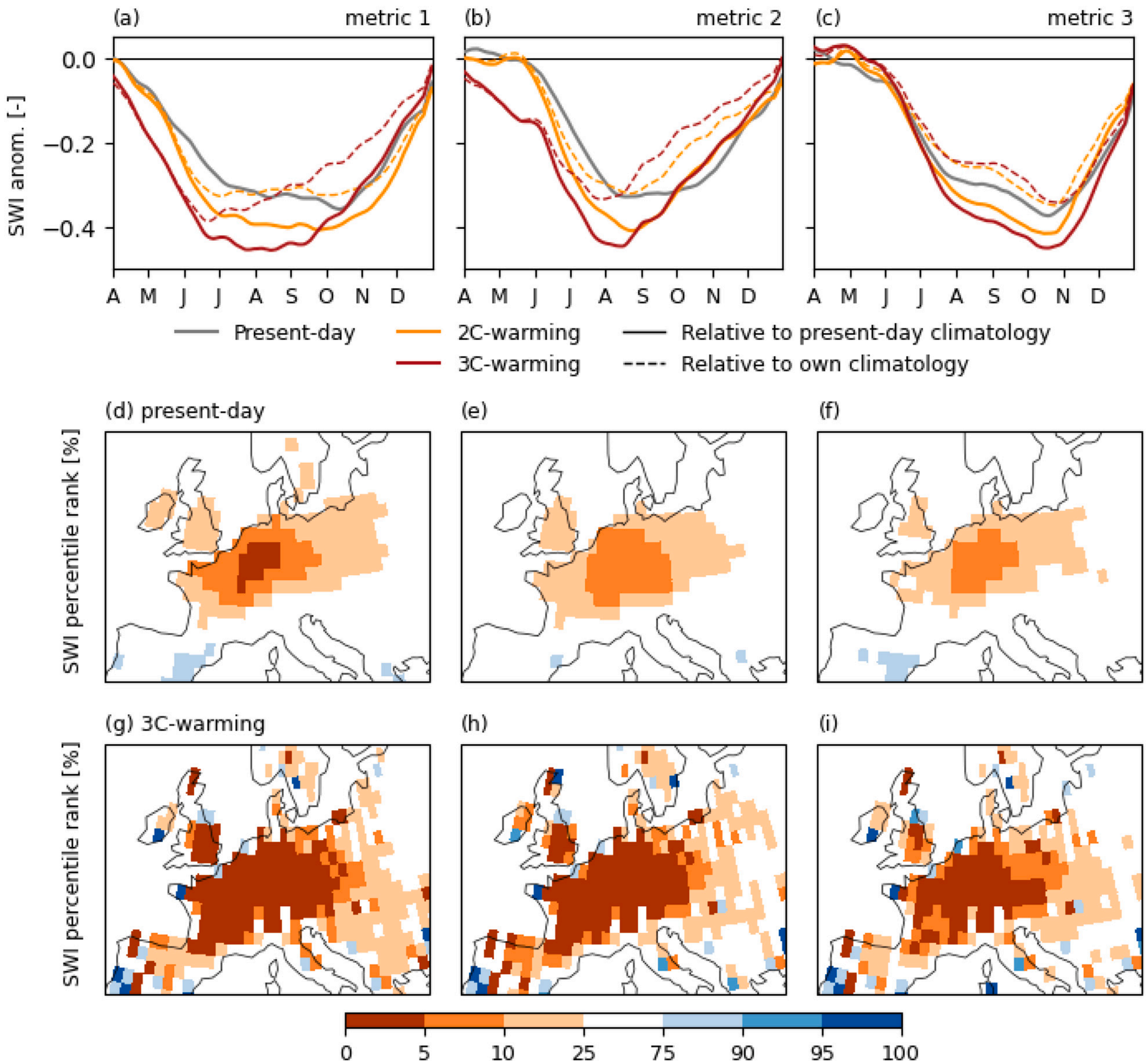


**Fig. 8.** (a–c) Time series of anomalies of soil wetness index (SWI) of the top 1 m of the soil, as in Fig. 6. (d–f) Maps of drought severity in August, expressed as the percentile rank of SWI of the top 1 m of the soil, shown here for the present-day analogues. (g–i) Same, but now for 3C-warming analogues relative to the present-day climatology. Each column shows the result of a different event selection metric, from left to right metrics 1, 2 and 3. Note irregular colour bar in (d–i). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

focusing on event definition, event return period and the physical self-consistency and plausibility of the storylines.

The present-day analogues of Drought'18 share many of the same features as the observed event, which provides us with confidence that the method works satisfactorily. The analogue events are warm, dry and sunny, resulting in high precipitation deficits and low soil water availability. When the different selection metrics are compared, we note subtle differences in event timing or severity, which also propagate into subtle differences in the climate change response (e.g. the intensification of spring drought found for metric 1 only). As noted in Section 2.1 this dependence on event definition is unavoidable and should be investigated and noted. We have chosen here to discuss three storylines, each based on a different selection metric, and note the similarities, the differences and the origin of these differences. We have interpreted similarities in climate change response between storylines as robust changes, and interpreted differences in the response as areas where further research is needed.

It is not obvious whether one of the three storylines presented here is more plausible or more relevant than the others, or whether there are theoretical restrictions to the choice of selection metric. Though metric 3, based on a temporal correlation, does not allow as much freedom in finding changes in event severity as the other metrics. Metrics close to societal or natural impacts increase user relevance (Hazeleger et al., 2015; Van der Wiel et al., 2020), the metrics in this study capture similar meteorological hazards. These storylines provide information on future hazards and associated risk and investigate the variety of weather conditions leading to the impact of interest (Van der Wiel et al., 2020). It is up to a potential user which metric is most suitable: a farmer with a crop particularly sensitive to drought conditions in high summer would likely choose metric 2, someone interested in vegetation states, which respond to the most extreme meteorological droughts, would likely choose metric 1. If instead one chooses for metrics that maximise, for example, (meteorological) similarity in time, as metric 3 or in nudging/PGW experiments, the climate change effects in terms of changing hazards and risk may be underestimated. This discrepancy is likely to be larger when event circulation patterns are subject to change (for droughts e.g. changes in persistence of weather systems,

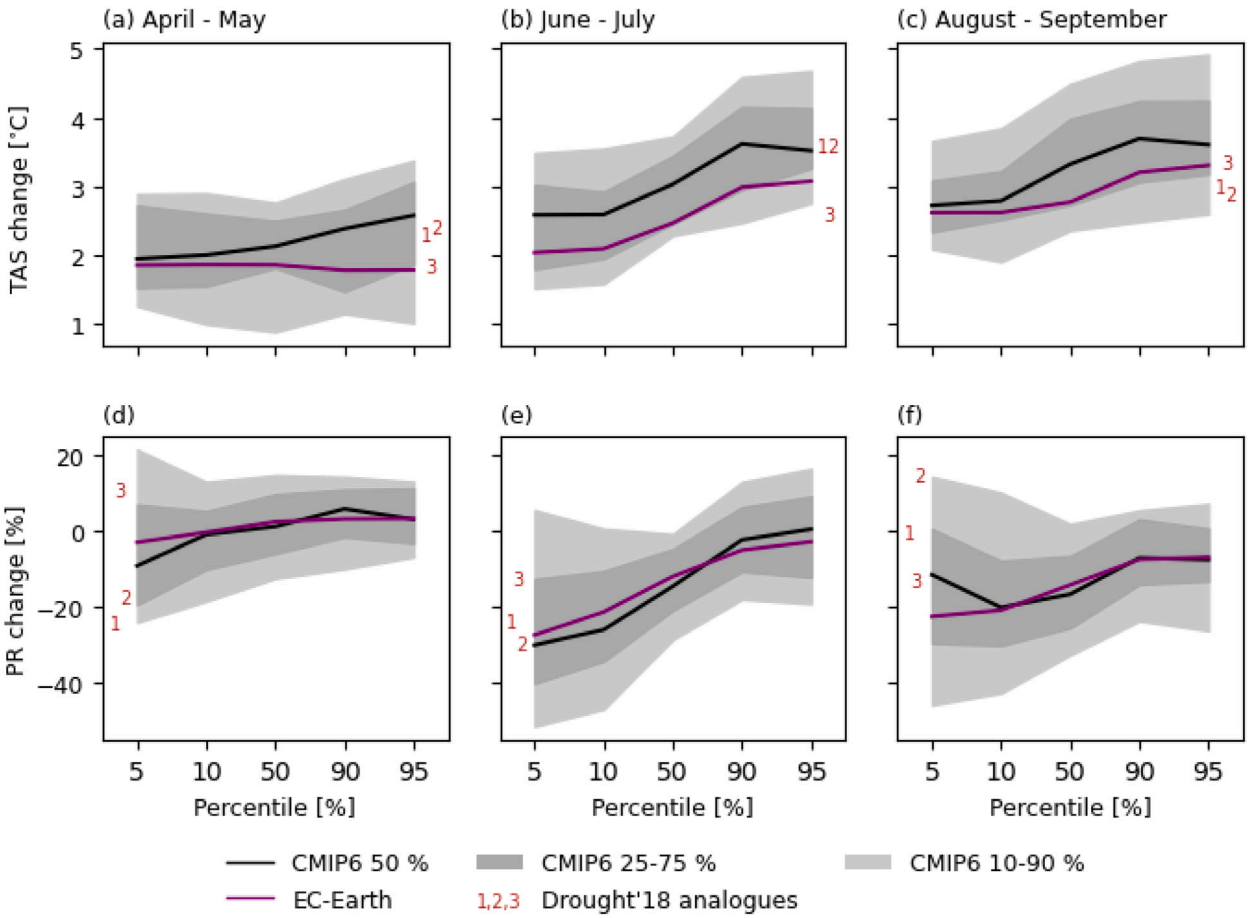


**Fig. 9.** Change in 5-10-50-90-95% percentiles of monthly mean (a–c) 2 m temperature [°C] and (d–f) precipitation [%] in the Rhine basin for 35 CMIP6 models (grey/black shading) and the EC-Earth ensembles (purple line), see Section 2.3 for details. Changes computed over a 30-year period/large ensemble with present-day Global Mean Surface Temperature (GMST) and a 30-year period/large ensemble with 3C-warming GMST under SSP5-8.5/RCP8.5 forcing. Red marks 1, 2 and 3 show the difference between the Drought'18 analogues in the 3C-warming and present-day climates for metrics 1, 2 and 3 respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Mann et al., 2018; Pfleiderer et al., 2019), or if there are positive feedback mechanisms that may intensify future events (for droughts e.g. through the sensible heat flux, Fischer et al., 2007; Schumacher et al., 2019).

The created analogues should be of comparable return period or severity as the observed extreme event. The climate change response of extreme events with high return period is not always the same as that of events of lower return period (Appendix A, Van der Wiel et al., 2018, 2019). Hence a difference in return period between analogue and observed event may negatively influence the quality of the results. In practise this is why large ensemble data are required, as these are more likely to contain the relevant extreme events. Here we selected 20 simulated events and took the composite mean to create an analogue. Selecting more events would have reduced the influence of natural variability in event specifics (noise from unrelated weather), but meant we would have lower return period and may have impacted the results negatively. Vice versa, selecting less events would have increased the return period (better matching the extreme event of interest), but would have increased the noise due to natural variability and thus decreased the robustness and interpretability of the results. This balance between reducing noise from natural variability and increasing return period should be carefully considered, sensitivity analyses such as those in Appendix A can provide insight in the influence of these choices on the results. The sensitivity analysis here emphasises that large ensemble simulations are a strict requirement for the proposed storyline method by event-selection, they are necessary to adequately capture the event return period and sample natural variability. Note that the variability of individual events within the analogues can also be of interest. Information of different pathways/weather leading to a common impact or matching a common selection metric (e.g. Van der Wiel et al., 2020), and the changes therein due to climate change, may inform user preparation or adaptation.

The physical-consistency of the presented storylines is ensured by their source. Global coupled climate models describe the physical relations between different components of the climate system and are physically self-consistent by definition. However, in cases where climate models show discrepancies with observed processes or observed trends, the plausibility or accuracy of the storylines cannot be as easily assured. Relevant to our study, further research and possibly model development is needed for example to align observed and modelled trends in heatwaves in Northwestern Europe (Min et al., 2013; Vautard et al., 2020), and to align observed and modelled tail dependence for correlated hot-and-dry compound events (Zscheischler and Fischer, 2020). Improvements in the realism of simulated summer weather conditions would improve the accuracy of our storylines. Furthermore, the climate simulations did not include dynamic vegetation, drought feedbacks involving vegetation (Wramneby et al., 2010) are thus not included. Including dynamic vegetation in climate simulations would improve the accuracy of the results.

Finally, we note that the presented method may also be applied to other extreme weather events than summer droughts (e.g. heatwaves, cold waves, or large-scale extreme rainfall events). Since droughts develop over time, we fixed our selection procedure to certain calendar months, but analogue selection can be broadened by allowing shifts in time or space, therewith reducing the necessary ensemble size. The main scientific advantage of the selection method over the nudging/PGW methods is the fact that possible positive feedback mechanisms and event circulation changes are taken into account and the results are therefore more generally applicable. An additional practical advantage is the fact that no dedicated experiments have to be created and the selection method can be applied with limited additional computational costs. Do note that global coupled climate models have a relatively coarse resolution, meaning not all extreme events are resolved adequately (e.g. Pascale et al., 2016; Van der Wiel et al., 2016; Pilo et al., 2019). Storyline creation for such events are therefore not possible using the analogue selection method on global coupled model data. The PGW method, using regional climate models, may allow such analyses.

## 5. Conclusions

The aim of this work was to develop a method for the creation of physical storylines of future extreme events similar to observed events based on large ensemble climate model data. A storyline approach, focusing on physical processes and plausibility rather than probability, improves risk awareness through its relation with our memory of the observed event (episodic knowledge) and contributes to decision making processes through their user focus (Hazeleger et al., 2015; Shepherd et al., 2018). As a proof of concept, we created analogues of the 2018 western European drought event for three different states of global climate, one like the present-day climate and two warmer future scenarios (2 °C- and 3 °C-warmer than pre-industrial). These were used to describe future event characteristics and compare these to the event as it occurred, and contribute to answering user questions like "Will droughts take a different form in the future?" or "How much worse can such droughts get?".

We have shown that future events like Drought'18 in the Rhine basin are likely to become more severe in a warmer future, in agreement with existing general projections of drought (e.g. Samaniego et al., 2018; Spinoni et al., 2018). Increased drought severity is caused by decreased precipitation and increased potential evapotranspiration (due to higher temperatures and increased incoming solar radiation). The timing and strength of relevant future changes are somewhat dependent on the precise event definition. For example, when focusing on high end-of-summer precipitation deficit, an amplified high pressure system in spring kick-starts the drought, a driver of drought that may increase in a warmer future (see also Haarsma et al., 2015; Van der Linden et al., 2019). Instead, when focusing on the temporal development of precipitation deficit during the entire growing season, no circulation effects are found and future drought intensification is the result of climatological mean changes alone. The latter storyline is probably closest to the results of Aalbers et al. (2021) who performed a PGW experiment for the Drought'18 event. By design a PGW experiment has no circulation response beyond climatological mean changes, and both methods find, for example, a small increase in precipitation in spring.

The storylines further show that future events like Drought'18 lead to significantly lower plant available soil moisture. Between June and November this can for a large part be explained by a climatological mean reduction of available soil moisture. Taking into account these mean changes, two of our storylines show that future droughts lead to relatively lower available soil moisture in the spring months (i.e. increased spring agricultural drought), but relatively higher available soil moisture at the end of summer. The analogue with dry conditions continuing into autumn shows comparable soil moisture drought intensities in the future also at the end of summer. The spatial extent of future events is significantly larger than the event in its present-day form, also here we find that this is mostly the result of changes in the soil moisture mean climatology.

The changes in drought severity and extent are larger in a world with 3 °C warming than in a world with 2 °C warming. Climate change mitigation will thus help to limit the societal and natural impact of future drought events somewhat. Societal adaptation to on average drier conditions (higher mean precipitation deficits and lower mean plant available soil moisture) and more severe drought events (additionally higher peak precipitation deficits and lower peak soil moisture availability) is inevitable if impacts are to be limited to present-day levels.

For users, it is the societal and natural impacts rather than the hazard of future droughts that are relevant. Hence studies investigating such future impacts, including options for adaptation or mitigation, are required. From a meteorological perspective, the role of future circulation changes in drought severity (Haarsma et al., 2015; Van der Linden et al., 2019), especially during the spring months, remains unclear. Finally a technical comparison between the PGW method and the analogue method used here is of interest. A PGW experiment for Drought'18 gives qualitatively comparable results as presented here (Aalbers et al., 2021), but a quantitative comparison would provide insights in the strengths and weaknesses of the two methods.

## CRediT authorship contribution statement

**Karin van der Wiel:** Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Visualization, Writing - original draft, Writing - review & editing. **Geert Lenderink:** Conceptualization, Writing - review & editing, Funding acquisition. **Hylke de Vries:** Writing - review and editing, Project administration.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

For all analysed variables, Rhine-basin mean time series for each large ensemble (i.e. 3 × 2000 years) are available at DOI https://doi.org/10.5281/zenodo.5083160. These include the data of the constructed Drought'18 analogues. Further EC-Earth data are available from the corresponding author upon reasonable request.

## Acknowledgements

The authors thank Emma Aalbers and Erik van Meijgaard for advise on calculating the soil wetness index. This work was carried out as part of the EUCP project, which received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 776613. KW was further funded by the RECEIPT project, a European Union's Horizon 2020 research and innovation programme under grant agreement No. 870712. The authors are grateful to the EC-Earth consortium for their contribution to the development of the Earth System Model EC-Earth. The large ensemble data sets were created as part of the HiWAVES3 project, for which funding was supplied by the Netherlands Organisation of Scientific Research (NWO) under Grant Number ALWCL.2016.2. Acknowledgement is made for the use of ECMWF's computing and archive facilities in producing these data sets, and to the individual climate modelling groups for producing and making available their CMIP6 model output.

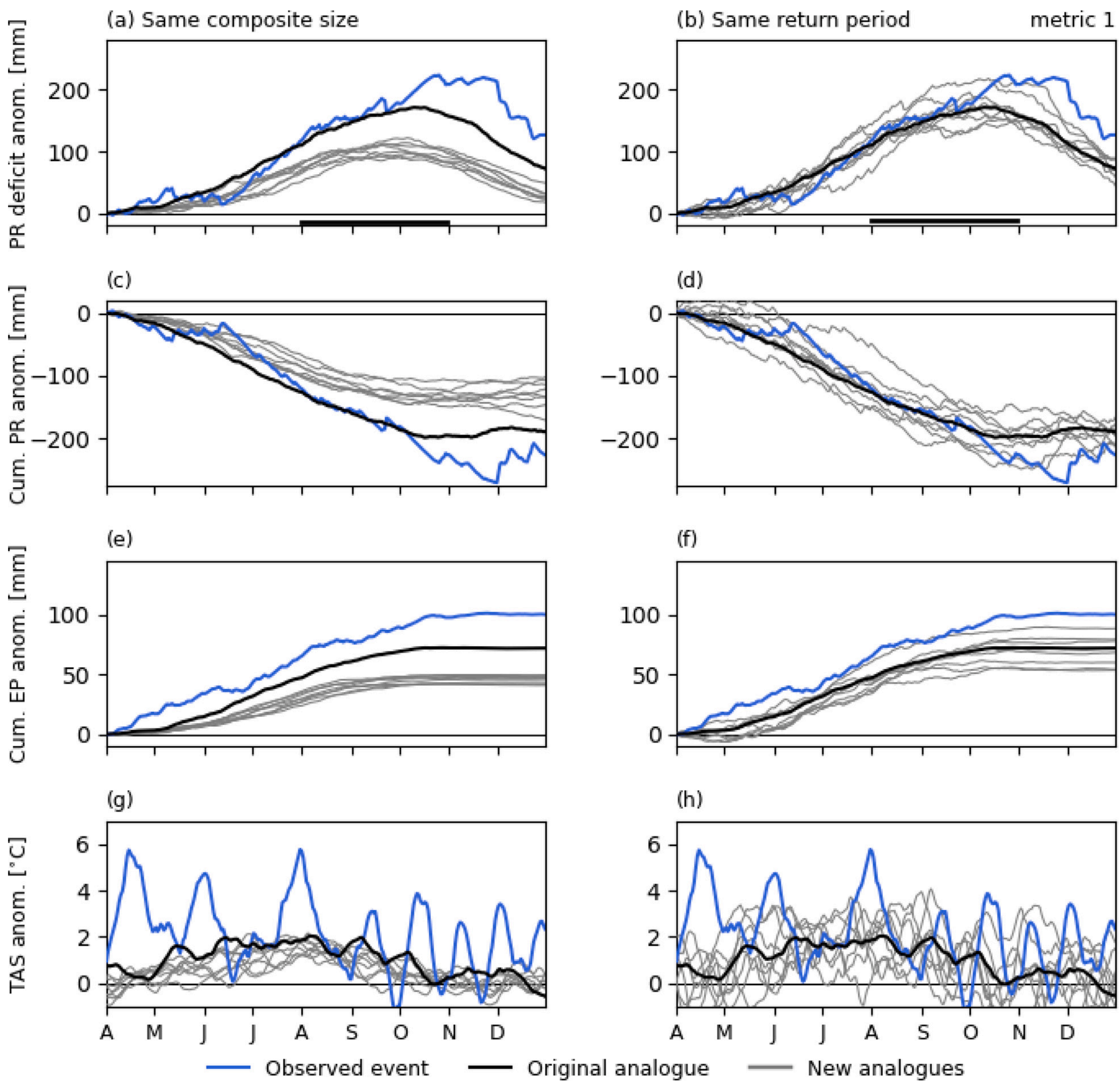


**Fig. A.1.** Present-day analogue time series for two sensitivity experiments. Shown are the analogues from the eight small ensembles in grey, the original analogue in black (as in Fig. 5 of the main text), and the observed event in blue. (Left) same composite size experiment, (right) same return period experiment. Variables shown, anomalies of: (a–b) precipitation deficit [mm], (c–d) cumulative precipitation [mm], (e–f) cumulative potential evapotranspiration [mm], (g–h) 2 m temperature [°C]. The horizontal black bar in (a–b) shows the selection period, selection by means of metric 1. A 15-day running mean was applied for 2 m temperature. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

## Appendix A. Ensemble size, analogue composite size and analogue return period

The storylines of future Drought'18 events discussed in this study are based on analogue events computed from large ensemble climate model simulations. These analogues were generated by taking the composite mean over 20 selected simulated events that were most similar to the observed event by some metric. Two of the metrics in the main text selected the most extreme events, this essentially constrains the return period of the analogues (taking 20 events out of 2000 years means these represent roughly 1-in-100 year return period events). In the main text we noted that *"Our choice to build the analogues from 20 events is a compromise between increasing event return period and reducing noise."*. By means of a sensitivity experiment we demonstrate this compromise and the consequences of different choices. It highlights the importance of having a sufficiently large ensemble to select events from, only then is it possible to create reliable and robust climate change information.

We split each full large ensemble of 2000 years in eight smaller ensembles of 250 years each. Using these eight small ensembles we repeat the analysis of the main text, we create analogues and investigate changes in event severity and drivers. We do this twice, each time making a different choice in the aforementioned compromise:

- **Same composite size** For each of the small ensembles we created analogues based on the same number of simulated events as in the main study, i.e. 20 events. This implies we are looking at less extreme events here, a rough estimate gives a 1-in-12.5 years return period event rather than 1-in-100 years.
- **Same return period** For each of the small ensembles we create analogues based on 2 events, these analogues have approximately the same return period as those in the main text (roughly 1-in-125 versus 1-in-100 years).

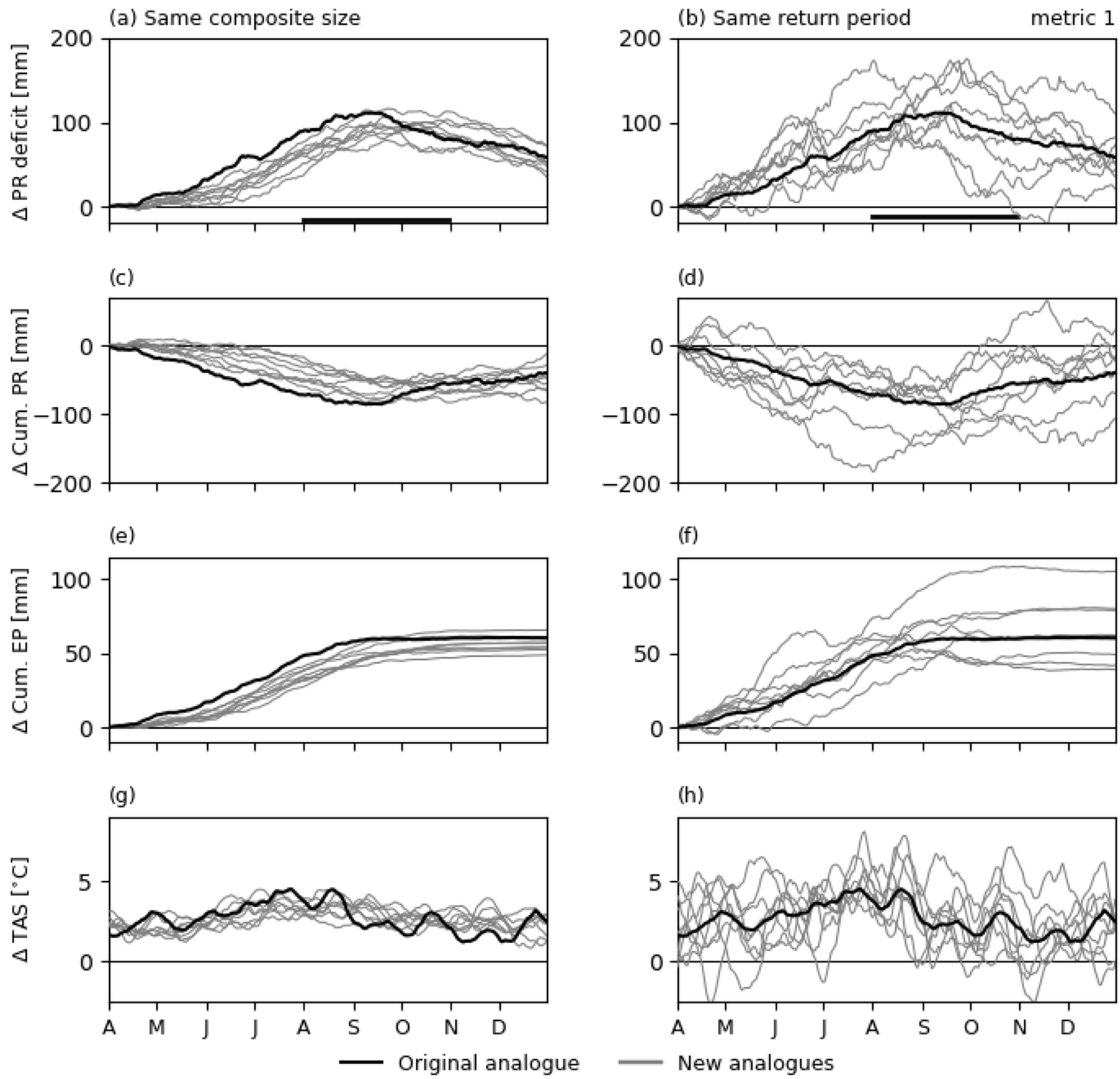


**Fig. A.2.** As Fig. A.1 but here showing the climate change signal, 3C-warming analogue minus present-day analogue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Note that the provided estimates of return period are too simplistic, and that storylines do not have a probabilistic focus, hence we have not computed or noted probabilities in the main text.

As expected the 'same composite size' experiment results in well defined analogues, with limited influence of random weather noise in single events (left column in Fig. A.1). However, the analogue drought is less severe than the observed event, the original analogue is in better agreement in this perspective. The 'same return period' experiment does not have this problem, the maximum precipitation deficit of the new analogues is comparable to that of the original analogue and close to that of the observed event (right column in Fig. A.1). The new analogues do however suffer somewhat from noise, which is especially obvious in the time series of 2 m temperature (Fig. A.1h). This is not necessarily problem, since the observed event also contains a form of this noise, but makes comparisons is slightly more difficult. Based on this analysis we would prioritise analogue return period over analogue composite size.

If we look at the computed effect of climate change on the event analogues, we find that the 'same composite size' experiment provides robust climate change signals (left column Fig. A.2). The signal found for the different small ensembles is robust, though it is lower than what was found for the original analogue. This latter effect may be caused by a non-linear climate change response, where events of low return period respond differently than events of high return period (Van der Wiel et al., 2018, 2019). This robustness of the climate change signal is not found in the 'same analogue return period' experiment, here we find large variability in the climate change signal of different small ensembles (right column Fig. A.2). This is the result of random weather noise, which clouds the real climate change signal, and makes it impossible to make reliable statements on the influence of climate change on the event. Based on this analysis we would prioritise analogue composite size over analogue return period.

This sensitivity experiment shows the consequences of the different choices in the development of simulated analogues from large ensemble climate modelling data. It showcases that a balance, or compromise, is needed between analogue composite size and analogue return period. Having a sufficiently large ensemble, such that robust analogues can be created for the observed event under consideration, is essential to provide reliable and robust climate change information.

## Appendix B. Supplementary figures

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.wace.2021.100350.

## References


Aalbers, E., Meijgaard, E., Lenderink, G., Van den Hurk, B., 2021. Pseudo global warming experiment of European drought 2018. Environ. Res. Lett. In preparation.

Beersma, J.J., Buishand, T.A., Buiteveld, H., 2004. Droog, droger, droogst. Tech. Rep. 199-I, Koninklijk Nederlands Meteorologisch Instituut.

Brunner, L., Hauser, M., Lorenz, R., Beyerle, U., 2020. The ETH Zurich CMIP6 next generation archive: technical documentation. Tech. rep., http://dx.doi.org/10.5281/zenodo.3734128.

Cattiaux, J., Ribes, A., 2018. Defining single extreme weather events in a climate perspective. Bull. Am. Meteorol. Soc. 99 (8), 1557–1568. http://dx.doi.org/10.1175/BAMS-D-17-0281.1.

De Bruin, H., Lablans, W., 1998. Reference crop evapotranspiration determined with a modified makkink equation. Hydrol. Process. 12 (7), 1053–1062. http://dx.doi.org/10.1002/(SICI)1099-1085(19980615)12:7<1053::AID-HYP639>3.0.CO;2-E.

Deser, C., Lehner, F., Rodgers, K., Ault, T., Delworth, T., DiNezio, P., Fiore, A., Frankignoul, C., Fyfe, J., Horton, D., et al., 2020. Insights from earth system model initial-condition large ensembles and future prospects. Nature Clim. Change 1–10. http://dx.doi.org/10.1038/s41558-020-0731-2.

Drouard, M., Kornhuber, K., Woollings, T., 2019. Disentangling dynamic contributions to summer 2018 anomalous weather over europe. Geophys. Res. Lett. 46 (21), 12537–12546. http://dx.doi.org/10.1029/2019GL084601.

Ecorys, Van Hussen, K., Van de Velde, I., Lakamp, R., Van der Kooij, S., 2018. Economische schade door droogte in 2018. Tech. rep., https://www.ecorys.com/sites/default/files/2019-10/2019022120door.

Faranda, D., Vrac, M., Yiou, P., Jézéquel, A., Thao, S., 2020. Changes in future synoptic circulation patterns: consequences for extreme event attribution. Geophys. Res. Lett. 47 (15), e2020GL088002. http://dx.doi.org/10.1029/2020GL088002.

Fischer, E.M., Seneviratne, S.I., Lüthi, D., Schär, C., 2007. Contribution of land-atmosphere coupling to recent European summer heat waves. Geophys. Res. Lett. 34 (6), http://dx.doi.org/10.1029/2006GL029068.

Haarsma, R.J., Selten, F.M., Drijfhout, S.S., 2015. Decelerating atlantic meridional overturning circulation main cause of future west European summer atmospheric circulation changes. Environ. Res. Lett. 10 (9), 094007. http://dx.doi.org/10.1088/1748-9326/10/9/094007.

Hazeleger, W., van den Hurk, B.J., Min, E., van Oldenborgh, G.J., Petersen, A.C., Stainforth, D.A., Vasileiadou, E., Smith, L.A., 2015. Tales of future weather. Nature Clim. Change 5 (2), 107. http://dx.doi.org/10.1038/nclimate2450.

Hazeleger, W., Wang, X., Severijns, C., Ştefănescu, S., Bintanja, R., Sterl, A., Wyser, K., Semmler, T., Yang, S., Van den Hurk, B., et al., 2012. EC-earth V2.2: description and validation of a new seamless earth system prediction model. Clim. Dynam. 39 (11), 2611–2629. http://dx.doi.org/10.1007/s00382-011-1228-5.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. http://dx.doi.org/10.1002/qj.3803.

IPCC, 2013. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, New York, USA.

IPCC, 2014. Annex II: Glossary. In: Agard, J., Schipper, E.L.F., Birkmann, J., Campos, M., Dubeux, C., Nojiri, Y., Olsson, L., Osman-Elasha, B., Pelling, M., Prather, M.J., Rivera-Ferre, M.G., Ruppel, O.C., Sallenger, A., Smith, K.R., Clair, A.L.S., Mach, K.J., Mastrandrea, M.D., Bilir, T.E. (Eds.), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, New York, USA, pp. 1757–1776.

Kelder, T., Muller, M., Slater, L., Marjoribanks, T., Wilby, R.L., Prudhomme, C., Bohlinger, P., Ferranti, L., Nipen, T., 2020. UNSEEN trends: Detecting decadal changes in 100-year precipitation extremes. http://dx.doi.org/10.31223/osf.io/hyxeq, preprint.

Kornhuber, K., Osprey, S., Coumou, D., Petri, S., Petoukhov, V., Rahmstorf, S., Gray, L., 2019. Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern. Environ. Res. Lett. 14 (5), 054002. http://dx.doi.org/10.1088/1748-9326/ab13bf.

Kriegler, E., Bauer, N., Popp, A., Humpenöder, F., Leimbach, M., Strefler, J., Baumstark, L., Bodirsky, B.L., Hilaire, J., Klein, D., et al., 2017. Fossil-fueled development (SSP5): an energy and resource intensive scenario for the 21st century. Global Environ. Change 42, 297–315. http://dx.doi.org/10.1016/j.gloenvcha.2016.05.015.

Van der Linden, E.C., Haarsma, R.J., Van der Schrier, G., 2019. Impact of climate model resolution on soil moisture projections in central-western europe.. Hydrol. Earth Syst. Sci. 23 (1), http://dx.doi.org/10.5194/hess-23-191-2019.

Mann, M.E., Rahmstorf, S., Kornhuber, K., Steinman, B.A., Miller, S.K., Petri, S., Coumou, D., 2018. Projected changes in persistent extreme summer weather events: The role of quasi-resonant amplification. Sci. Adv. 4 (10), eaat3272. http://dx.doi.org/10.1126/sciadv.aat3272.

Min, E., Hazeleger, W., Van Oldenborgh, G., Sterl, A., 2013. Evaluation of trends in high temperature extremes in north-western europe in regional climate models. Environ. Res. Lett. 8 (1), 014011. http://dx.doi.org/10.1088/1748-9326/8/1/014011.

Pascale, S., Bordoni, S., Kapnick, S.B., Vecchi, G.A., Jia, L., Delworth, T.L., Underwood, S., Anderson, W., 2016. The impact of horizontal resolution on North American monsoon gulf of california moisture surges in a suite of coupled global climate models. J. Clim. 29 (21), 7911–7936. http://dx.doi.org/10.1175/JCLI-D-16-0199.1.

Pfleiderer, P., Schleussner, C.-F., Kornhuber, K., Coumou, D., 2019. Summer weather becomes more persistent in a 2 °C world. Nature Clim. Change 9 (9), 666–671. http://dx.doi.org/10.1038/s41558-019-0555-0.

Philip, S., Kew, S., Van Oldenborgh, G.J., Otto, F., Vautard, R., Van der Wiel, K., King, A., Lott, F., Arrighi, J., Singh, R., Van Aalst, M., 2020a. A protocol for probabilistic extreme event attribution analyses. Adv. Stat. Climatol. Meteorol. Oceanogr. 6, 177–203. http://dx.doi.org/10.5194/ascmo-6-177-2020.

Philip, S., Kew, S., Van der Wiel, K., Wanders, N., Van Oldenborgh, G., 2020b. Regional differentiation in climate change induced drought trends in the netherlands. Environ. Res. Lett. http://dx.doi.org/10.1088/1748-9326/ab97ca.

Pilo, G.S., Holbrook, N.J., Kiss, A.E., Hogg, A.M., 2019. Sensitivity of marine heatwave metrics to ocean model resolution. Geophys. Res. Lett. 46 (24), 14604–14612. http://dx.doi.org/10.1029/2019GL084928.

Ragone, F., Bouchet, F., 2020. Rare event algorithm study of extreme warm summers and heat waves over europe. ArXiv Preprint arXiv:2009.02519.

Ragone, F., Wouters, J., Bouchet, F., 2018. Computation of extreme heat waves in climate models using a large deviation algorithm. Proc. Natl. Acad. Sci. 115 (1), 24–29. http://dx.doi.org/10.1073/pnas.1712645115.

Reichler, T., Kim, J., 2008. How well do coupled models simulate today's climate? Bull. Am. Meteorol. Soc. 89 (3), 303–312. http://dx.doi.org/10.1175/BAMS-89-3-303.

Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E.F., Marx, A., 2018. Anthropogenic warming exacerbates European soil moisture droughts. Nature Clim. Change 8 (5), 421–426. http://dx.doi.org/10.1038/s41558-018-0138-5.

Schär, C., Frei, C., Lüthi, D., Davies, H.C., 1996. Surrogate climate-change scenarios for regional climate models. Geophys. Res. Lett. 23 (6), 669–672. http://dx.doi.org/10.1029/96GL00265.

Schumacher, D.L., Keune, J., Van Heerwaarden, C.C., de Arellano, J.V.-G., Teuling, A.J., Miralles, D.G., 2019. Amplification of mega-heatwaves through heat torrents fuelled by upwind drought. Nat. Geosci. 12 (9), 712–717. http://dx.doi.org/10.1038/s41561-019-0431-6.

Shepherd, T.G., Boyd, E., Calel, R.A., Chapman, S.C., Dessai, S., Dima-West, I.M., Fowler, H.J., James, R., Maraun, D., Martius, O., et al., 2018. Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. Clim. Change 151 (3–4), 555–571. http://dx.doi.org/10.1007/s10584-018-2317-9.

Sillmann, J., Shepherd, T., van den Hurk, B., Hazeleger, W., Romppainen-Martius, O., Zscheischler, J., 2019. Physical modeling supporting a storyline approach. Tech. Rep. 2019:01, CICERO Senter for klimaforskning, http://dx.doi.org/10.7892/boris.138767.

Sluijter, R., Plieger, M., Van Oldenborgh, G., Beersma, J., De Vries, H., 2018. De droogte van 2018: Een analyse op basis van het potentiële neerslagtekort. Tech. Rep. 117162, Koninklijk Nederlands Meteorologisch Instituut.

Spinoni, J., Vogt, J.V., Naumann, G., Barbosa, P., Dosio, A., 2018. Will drought events become more frequent and severe in Europe? Int. J. Climatol. 38 (4), 1718–1736. http://dx.doi.org/10.1002/joc.5291.

Stone, D.A., Hansen, G., 2016. Rapid systematic assessment of the detection and attribution of regional anthropogenic climate change. Clim. Dynam. 47 (5–6), 1399–1415. http://dx.doi.org/10.1007/s00382-015-2909-2.

Tulving, E., 1972. Episodic and semantic memory. In: Tulving, E., Donaldson, W. (Eds.), Organization of Memory. Academic Press, New York, pp. 381–403.

Van Oldenborgh, G.J., Van der Wiel, K., Kew, S., Philip, S., Otto, F., Vautard, R., King, A., Lott, F., Arrighi, J., Singh, R., Van Aalst, M., 2021. Pathways and pitfalls in extreme event attribution. Clim. Change 166 (13), 1–27. http://dx.doi.org/10.1007/s10584-021-03071-7.

Vautard, R., van Aalst, M., Boucher, O., Drouin, A., Haustein, K., Kreienkamp, F., van Oldenborgh, G.J., Otto, F.E., Ribes, A., Robin, Y., et al., 2020. Human contribution to the record-breaking June and July 2019 heat waves in Western Europe. Environ. Res. Lett. http://dx.doi.org/10.1088/1748-9326/aba3d4.

Wagener, T., Sivapalan, M., Troch, P.A., McGlynn, B.L., Harman, C.J., Gupta, H.V., Kumar, P., Rao, P.S.C., Basu, N.B., Wilson, J.S., 2010. The future of hydrology: An evolving science for a changing world. Water Resour. Res. 46 (5), http://dx.doi.org/10.1029/2009WR008906.

Wehrli, K., Hauser, M., Seneviratne, S.I., 2020. Storylines of the 2018 northern hemisphere heat wave at pre-industrial and higher global warming levels. Earth Syst. Dyn. Discuss. 1–25. http://dx.doi.org/10.5194/esd-2019-91.

Van der Wiel, K., Kapnick, S.B., Oldenborgh, G.J.v., Whan, K., Philip, S., Vecchi, G.A., Singh, R.K., Arrighi, J., Cullen, H., 2017. Rapid attribution of the August 2016 flood-inducing extreme precipitation in south Louisiana to climate change. Hydrol. Earth Syst. Sci. 21 (2), 897–921. http://dx.doi.org/10.5194/hess-21-897-2017.

Van der Wiel, K., Kapnick, S.B., Vecchi, G.A., Cooke, W.F., Delworth, T.L., Jia, L., Murakami, H., Underwood, S., Zeng, F., 2016. The resolution dependence of contiguous US precipitation extremes in response to CO2 forcing. J. Clim. 29 (22), 7991–8012. http://dx.doi.org/10.1175/JCLI-D-16-0307.1.

Van der Wiel, K., Kapnick, S.B., Vecchi, G.A., Smith, J.A., Milly, P.C., Jia, L., 2018. 100-year lower mississippi floods in a global climate model: Characteristics and future changes. J. Hydrometeorol. 19 (10), 1547–1563. http://dx.doi.org/10.1175/JHM-D-18-0018.1.

Van der Wiel, K., Selten, F.M., Bintanja, R., Blackport, R., Screen, J.A., 2020. Ensemble climate-impact modelling: extreme impacts from moderate meteorological conditions. Environ. Res. Lett. 15 (3), 034050. http://dx.doi.org/10.1088/1748-9326/ab7668.

Van der Wiel, K., Wanders, N., Selten, F., Bierkens, M., 2019. Added value of large ensemble simulations for assessing extreme river discharge in a 2 °C warmer world. Geophys. Res. Lett. 46 (4), 2093–2102. http://dx.doi.org/10.1029/2019GL081967.

Wramneby, A., Smith, B., Samuelsson, P., 2010. Hot spots of vegetation-climate feedbacks under future greenhouse forcing in europe. J. Geophys. Res.: Atmos. 115 (D21), http://dx.doi.org/10.1029/2010JD014307.

Yiou, P., Jézéquel, A., 2020. Simulation of extreme heat waves with empirical importance sampling. Geosci. Model Dev. 13 (2), 763–781. http://dx.doi.org/10.5194/gmd-13-763-2020.

Zscheischler, J., Fischer, E.M., 2020. The record-breaking compound hot and dry 2018 growing season in Germany. Weather Clim. Extremes 100270. http://dx.doi.org/10.1016/j.wace.2020.100270.