HORIZON 2020

THEME SC5-2017

(Grant Agreement 776613)

European Climate Prediction system (EUCP)

Deliverable D5.5

Storylines of predictions and projections developed and assessed for potential applications

Deliverable Title		*Storylines of predictions and projections developed and assessed for potential applications*
Brief Description		*The deliverable will report on:* *1. Storyline approaches that can potentially inform decisions across prediction and projections timescales using user-informed selection and lines of evidence approaches, and 2. An assessment of the applicability of multiple storyline approaches (investigated in WP2, WP4 as well as WP5) and their suitability for answering user-relevant questions on the two-40-year timescale.*
WP number		5
Lead Beneficiary		*Fai Fung, UKMO*
Contributors		*Fai Fung (UKMO), Christopher R. Goddard (UKMO), Dominic Matte (UCPH), Andrew Ballinger (UEDIN), Gabi Hegerl (UEDIN), Christopher H. O'Reilly (UOXF), Antje Weisheimer (UOXF), Tom Crocker (UKMO), Carol McSweeney (UKMO), Karin van der Wiel (KNMI), Renate Wilcke (SMHI)*
Creation Date		*18/06/2022*
Version Number		*1*
Version Date		*24/03/2022*
Deliverable Due Date		*31/03/2022*
Actual Delivery Date		*01/04/2022*
Nature of the Deliverable	R	*R – Report*
Dissemination Level/ Audience	x	*PU - Public*
		PP - Restricted to other programme participants, including the Commission services
		RE - Restricted to a group specified by the consortium, including the Commission services
		CO - Confidential, only for members of the consortium, including the Commission services

Version	Date	Modified by	Comments
1.0	22/02/2022	All contributors	For EUCP internal review
1.1	09/03/2022	FF, CG, TC, CMS	Edits made during review
2.0	24/03/2022	All contributors	Version submitted to EUCP WP7 following review and updates

1 Executive summary

The research community is currently exploring what storylines mean and how to use them in applications, with numerous definitions and proposed purposes. Here, we broadly define them as narrative-based approaches to creating and communicating physically based information about past, current or future weather and climate. As the storyline approach is still novel, we present work and associated publications across EUCP that look towards bridging (a) storylines as part of the scientific and data production process to (b) storylines as a climate service. The deliverable seeks to answer

- What are climate storylines and where are they useful?
- How could storylines bring together various outputs and products of EUCP science?
- What are the challenges of producing storylines as a service?

The term "storylines" is often also used to refer to steps in the process of sub-selecting the most appropriate climate models for further analysis or reporting. Both meanings are relevant to the European Climate Prediction System proposed in EUCP as a process for the selection and production of data, and bringing together multiple lines of evidence, as well as a product for understanding and building confidence in projections and improving usefulness and usability. In this deliverable, we answer the above questions through studies exploring storylines as a user product, as scientific building blocks and as a multiple lines of evidence assessment.

In the two co-production case studies for the heritage and the water supply management sectors, we find that there is an appetite for storylines across these very different users. We demonstrate that storylines can be used as a communication tool by simplifying the often-overwhelming volume of climate data as well as a tool for better understanding of the characteristics of a climate model ensemble for a climate hazard of interest. We observe that a key part of the storyline construction process is the co-production of knowledge around climate forecasts and projections. We have taken this knowledge and assessed the suitability of each of the scientific building blocks of storylines explored in the deliverable. These scientific building blocks include storylines of hazard events, variability as well as clustering methods.

We present two studies that explore future hazard events sets at pseudo-global warming levels. This includes future flooding storylines for Copenhagen that employ a convection-permitting model and show the climate change dependence of small-scale convective events, which until now have remained outside the scope of current attribution science. We also transpose an observed event (the 2018 European-wide drought) into pseudo-global warming levels, providing the building block of storylines rooted in the recent lived experience of users.

We also present the clustering approach which offers the possibility of reducing large ensembles analyse by selecting representative members which are coherent across variables (from the climate models or potentially user-defined metrics), seasons and regions, while minimising the loss of relevant information. In addition, we have two pieces of work that explore the variability in climate projections which show a tendency to underestimate the contribution of internal variability: we present one study on the North Atlantic Oscillation and another on the extratropics. All of these scientific storyline building blocks could be used to explore different components of uncertainty in climate projections and forecasts.

Finally, we demonstrate a multiple lines of evidence tool where we have gathered multiple climate model projections ensembles to make information accessible to users of projections in Europe. The aim is to provide users of climate data the wider uncertainty context from "multiple lines of evidence" which might "bookend" climate impacts studies or storylines production. That is, at the beginning of a study, the tool could provide the wider uncertainty context to inform initial dialogue with users and help inform the appropriate selection of projection products for analysis. The tool could also be used at the end of the study to interpret the results based on a narrower selection of model projections. We argue that the multiple lines of evidence tool would aid storyline construction as data processing overhead prohibits most individual users from looking at multiple datasets for a wider uncertainty context. It could also offer a useful basis for infilling missing parts of the uncertainty space in downscaled datasets, either with new dynamical downscaling experiments or statistical approaches such as those explored in EUCP D5.4.

As storyline approaches are still novel, further work with users is required to embed them in existing decision-making contexts to demonstrate their value. There are also future challenges to overcome such as how tailored or useful to multiple users they could be as well as how some of the scientific building blocks can be upscaled. However, in this deliverable, we demonstrate the potential power of storylines as a user product and novel scientific building blocks used to constructed them.

2 Project objectives

This deliverable has contributed to the following EUCP objectives (Description of Action, Section 1.1):

No.	Objective	Yes	No
1	Develop an ensembles climate prediction system based on high-resolution climate models for the European region for the near-term (~1-40 years)	x	
2	Use the climate prediction system to produce consistent, authoritative and actionable climate information	x	
3	Demonstrate the value of this climate prediction system through high impact extreme weather events in the near past and near future	x	
4	Develop, and publish, methodologies, good practice and guidance for producing and using EUCP's authoritative climate predictions for 1-40 year timescales	x	

3 Detailed Report

3.1 Introduction

This deliverable outlines the work carried out under EUCP WP5 Task 5.5 on climate storylines, in addition to storylines related work in WP2 T2.3. The term "storylines" related to exploring uncertainty in future climate is gaining traction within climate science, recently featuring in the draft 6th Assessment Report by the Intergovernmental Panel on Climate Change where

> "*The term storyline is used both in connection to scenarios (related to a future trajectory of emissions or socio-economic developments) or to describe plausible trajectories of weather and climate conditions or events, especially those related to high levels of risk*". (Shepherd, 2021).

The research community is currently exploring what storylines mean and how to use them in applications, with numerous definitions and proposed applications (e.g. Shepherd et al. 2018, Jack et al. 2020, Ciullo et al. 2021). **Here, we broadly define them as narrative-based approaches to creating and communicating physically based information about past, current or future weather and climate.** Throughout the deliverable we mention the inclusion of 'drivers' in storylines-related analysis or products, this refers to the physical processes behind events or future change and uncertainty, such as the NAO, jet stream and weather patterns (and changes to them). More broadly in the literature this may also encompass certain causal chains leading to high impact, low likelihood events.

The term storylines is often also used to refer to steps in the process of sub-selecting the most appropriate climate models for further analysis or reporting (see WP2 T2.3, D2.4). Both meanings are relevant to the European Climate Prediction System proposed in EUCP as a process for the selection and production of data, bringing together multiple lines of evidence as well as a product for understanding and building confidence in projections, and improving usefulness and usability. The Venn diagram in Figure 3.1-1 shows one way of conceptualising the storylines-related studies presented in this deliverable, namely: storylines as a user product, scientific storylines, or building blocks and the lines of evidence assessment interface.

As the storyline approach is still novel, there are only a few published examples on how storylines as a user product could be constructed and few demonstrating their real-world application. Here we present work and associated publications across EUCP and aim to take first steps at addressing these gaps. In EUCP, we look towards bridging (a) storylines as part of the scientific and data production process to (b) storylines as a climate service. The deliverable therefore seeks to answer

- What are climate storylines and where are they useful?
- How could storylines bring together various outputs and products of EUCP science?
- What are the challenges of producing them as a service?

The deliverable addresses these questions as follows: in Section 3.2, two case studies are reported in detail, where the end-to-end process for co-producing storylines with users is demonstrated for specific applications, exploring the intersection of storylines as a user product and scientific storylines exploring atmospheric process understanding. In Section 3.3, scientific storylines or analysis which could be building blocks of storylines are explored, such as projecting past events into the future, or analysis of modes of variability. In Section 3.4, we showcase a lines-of-evidence tool that brings together multiple datasets, which could support the construction of storylines. In section 3.5, we discuss how each of the elements reported in Sections 3.2 to 3.4 can be used to potentially develop a storylines climate service or bespoke storylines as a user product.

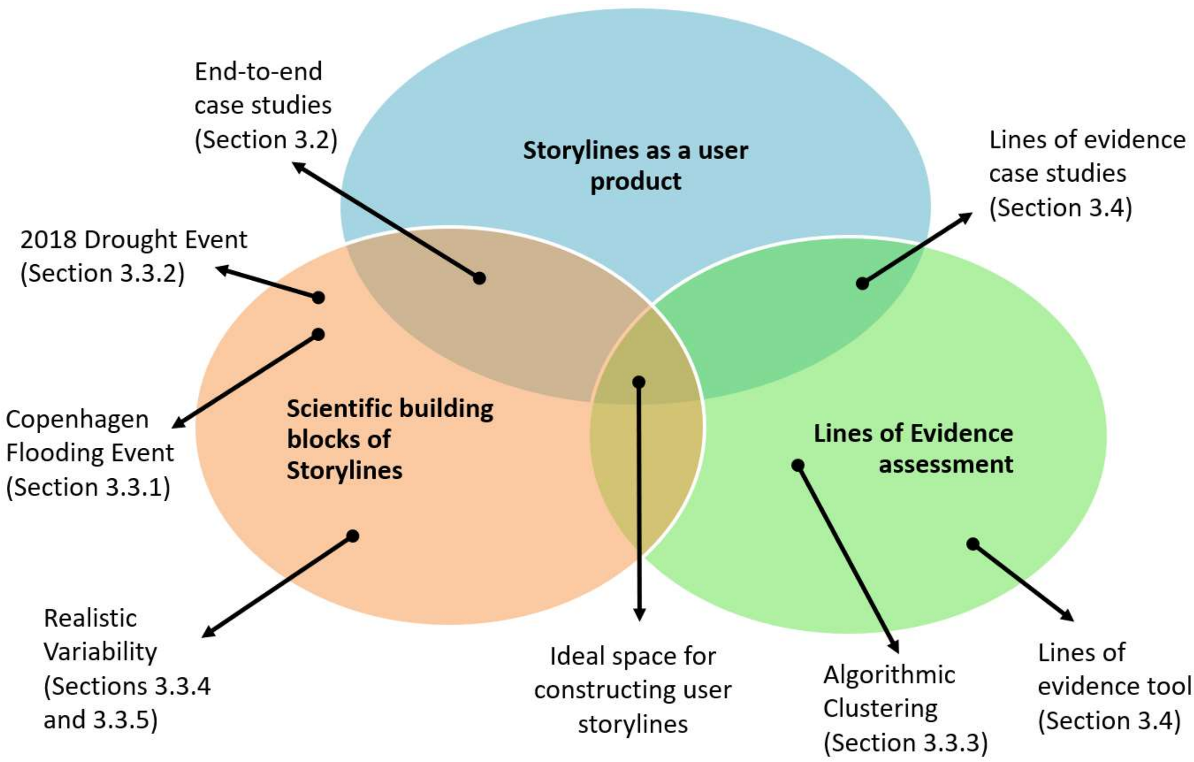


Figure 3.1-1 Venn diagram of the storylines-related elements covered in this deliverable, with an indication of where the different pieces of work sit. *Storylines as a user product* (blue, top) refers to storylines which are provided to, or created by users for a purpose such as improving communication, representing relevant uncertainty or making decisions. *Scientific building blocks of storylines* (orange, left) refers to storylines created to explore atmospheric processes and variability (or related analysis), often for the purpose of scientific understanding, but increasingly as part of producing a user product (intersection with blue). *Lines of evidence assessment* (green, right) refers to the process of examining multiple lines of evidence when producing climate information, which may be undertaken to examine the plausibility and robustness of storylines (intersection with orange and/or blue).

3.2 End-to-end production of storylines for applications (UKMO)

3.2.1 Motivation

In this section we present two examples that explore the potential of storylines in an application context that seek to address the lack of examples of how they can be used by organisations making climate-related decisions. We do this by starting from the user context and using this to inform how the storylines are produced and what they contain, allowing continuous feedback and evaluation (see Figure 3.2-1). When creating storylines as a product we pose that a shared understanding of the purpose from inception is essential, yet this is currently highly variable in the literature. The applications showcased here are opportunistic but also selected to span the range of users interested in climate resilience planning. They are:

* a user from the heritage management sector early in its adaptation journey where efforts are focus on raising awareness within their organisation of the impacts of climate change
* a set of users from the water resource supply sector that, due to regulations, have already been through a few cycles of adaptation planning and are therefore familiar with climate data and risk assessments.

Both sets of users are based in the United Kingdom where climate change adaptation is enshrined in legislature (e.g. Anglian Water, 2020; Historic England, 2022) and national climate change climate change data (Lowe et al., 2018) and risk assessments are available (CCRA3, 2022). We observe in our examples that the volume of information can be an obstacle for those organisations that are new to the adaptation planning (in our case the heritage management sector). Even for organisations that have a wealth of experience (in our case, the water supply management sector), we observe that existing information sources (e.g. the UK Climate Projections science reports and guidance, the underpinning research reports of the CCRA3) do not yield the actionable information or understanding that they are seeking to help inform their decisions.

Storyline approaches can offer simplified, more narrative, outcomes that resonate with users (e.g. Climate Risk Narratives in Jack et al, 2020) as well as the inclusion of physical processes and drivers (e.g. atmospheric drivers of change in Zappa & Shepherd, 2017), enhancing physical plausibility and understanding. In the following examples, we investigate how narrative approaches may enhance the utility and usefulness of climate data. We also expose the importance of information beyond the ensemble mean and trends in climate changes in the analysis as well as to the user, such as the relevant range of uncertainty. The analyses in Sections 3.2.2 and 3.2.3 aim to sub-select representative projection information to explore the relevant range of uncertainty and explores the potential role of physical driver information in analysing and communicating past and future events and changes. Section 3.2.3 also applies this to decadal forecasts. Therefore, we explore the utility and practicalities of using the storylines approach to exploring atmospheric drivers of change to produce storylines for user applications, which interface at the third and fourth steps of the example workflow in Figure 3.2-1, which may be an iterative feedback process.

3.2.2 Heritage management – storylines of winter rainfall and summer temperature hazards

Background

Heritage management organisations in the UK are public or third sector organisations responsible for the care and running of thousands of historic sites, and the wider preservation of its heritage with varying roles and remits. The cultural and built heritage sector is clearly exposed to weather hazards which are evolving under climate change and is responsible for enabling the preservation and enjoyment of an invaluable and non-renewable resource covering natural landscapes, to managed buildings and collections, to ancient monuments. The degree to which current sites, assets and operations are resilient to these changes is currently being explored in the literature. The main stakeholder for this case study is Historic Environment Scotland (HES), which is a public sector organisation responsible for the care and running of more than 300 historic sites, and the wider preservation of Scotland's heritage. Climate information is not routinely used in planning and decision making by this organisation, therefore they represent an organisation who have just embarked on climate change adaptation activities, e.g. their climate change adaptation strategy was published in 2020 (Historic Environment Scotland, 2020) and they have started to raise awareness across the organisation to perform climate risk assessment for their portfolio of assets and operations.

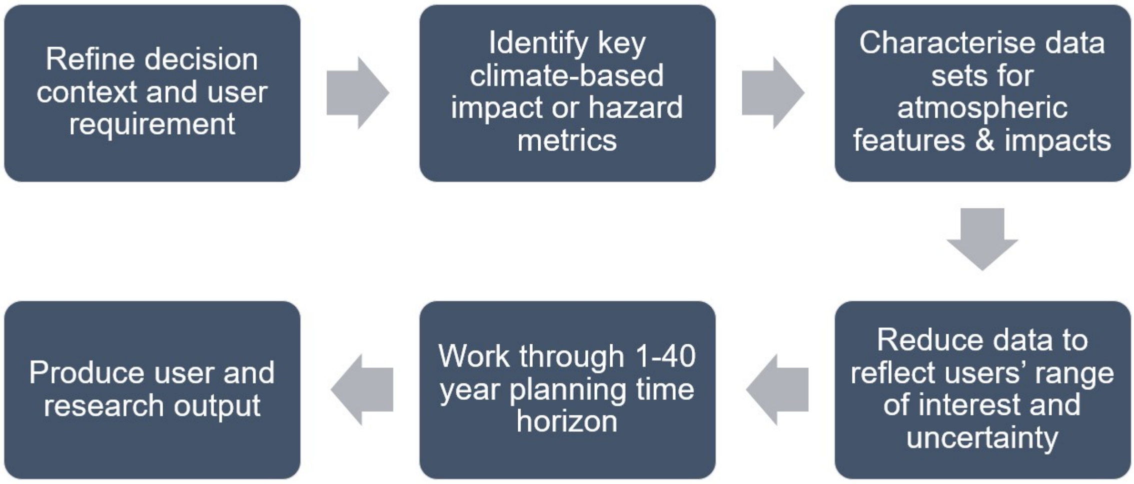


Figure 3.2-1 Proposed workflow for creating end-to-end climate service storylines, co-developed with users. In these case studies, step three refers to the driver analysis using weather patterns, combined with the application of the chosen climate metrics to events and future changes. Step 4 refers to the sub-selection performed in the context of the user requirements, such as the risk tolerance and appetite for uncertainty.

Following initial conversations with their climate change scientist on climate change information needs, and related decisions in the organisation, it was clear there would not be the opportunity to test the utility of example climate storylines on decisions directly, covering step 5 of the workflow in Figure 3.2-1. This also limited the opportunity to explore moving from risk informed hazard metrics to decision-relevant impact metrics. However, the potential usefulness and usability could still be explored with a climate change team member who would further disseminate such information within the organisation. This was done via the production of information from which to create trial storylines, and an example of how this may be provided to HES, with uses from communication to awareness raising to start bringing consideration of future hazards into planning and decision making. A focus group of Heritage sector professional and climate service providers from around Europe provided additional feedback on the example information discussed with HES and added perspectives from around Europe. This approach broadly followed the flowchart shown in Figure 3.2-1.

Objectives

The main aim of this case study is to investigate the usefulness and usability of narrative based information exploring the uncertainty in climate projection ensembles as well as weather and climate drivers of events and future change. We present highlights of the scientific analysis needed to produce the information, alongside the first steps made towards determining whether climate "storylines" or "narratives" are an appropriate means for framing and communicating this information for the chosen application.

The trial application and science goals set the following requirements:

- a spatial domain focussing on Scotland (and variation within it),
- a future period around 40 years into the future which aligns with the end of the EUCP period of interest (we use 2051-2080),
- a focus on mean winter (DJF) rainfall and summer (JJA) temperature changes including a form of 'degree day'/threshold analysis with local thresholds,

- no specific past events to analyse, or impact metrics or operational thresholds to apply,
- the production of a range of storylines showing a physically consistent unfolding of events into the future and the relevant range of uncertainty, underpinned by weather pattern (driver) analysis,
- production of a prototype storylines product for HES to aid in the evaluation of usefulness and usability.

METHODS: USER VARIABLE AND METRIC SELECTION AND CLIMATE DATA

User input: Due to the lack of defined weather impact metrics or damage thresholds in use, the hazard metric used may not be particularly important providing it's possible to use it to understand past events and the degree of change in the future. Winter rainfall and summer temperatures were identified as the main hazard-causing climate variables which have a clear climate signal in future projections.

The analysis used the most recent UK national climate projections (UKCP18, Lowe et al., 2018, Murphy et al., 2019) as this is information the user is already aware of and has easy access to. This includes HadUK-Grid data derived from station observations in the historical period (Hollis et al. 2018), and the climate projections data focuses on the GCM members of the UKCP PPE (15 members, numbered 1-15, PPE-15 hereafter) and selected CMIP5 models (13 members, numbered 16-28, CMIP5-13 hereafter). The RCP8.5 scenario was used to focus on uncertainty beyond scenario uncertainty, which plays a smaller role for the EUCP period of interest, although does cause divergence within our future period of 2051-2080. The main metrics used are the seasonal number of days above the 90th (wet/hot) and 99th (very wet/very hot) percentile thresholds from the baseline period 1981-2010 for summer daily maximum temperature (JJA Tmax) and winter rainfall (DJF Pr), Tmax90d, Pr90d, etc. hereafter, and referred to as 'metrics'. The anomalies in seasonal Tmax averages or Pr totals are also used, as well as changes to the percentile thresholds, so all measures are relative and don't require bias correction. These hazard measures were all calculated on a per-grid cell basis before averaging over the area of interest, ensuring local changes are represented, and satisfy the scientific and user requirements as far as possible within the project scope.

Figure 3.2-2 shows the range of projected uncertainty across all the PPE-15 and CMIP-13 members for the chosen variables in the future period independently for each grid. This is shown to establish the level of uncertainty a storylines approach needs to represent for the chosen region, and to put this in a wider regional context. Over the British Isles there is an anomaly range of around 4 to 5 degrees Celsius in JJA Tmax (difference between the hottest and coldest model) and around 30 to 40 % in winter precipitation (difference between the wettest and driest). Regions of Europe show wider uncertainty ranges, emphasising that this may be even more important to consider for applications in in these parts of Europe, and highlights the difficulty of producing European wide messages and statements, or model sub-selections.

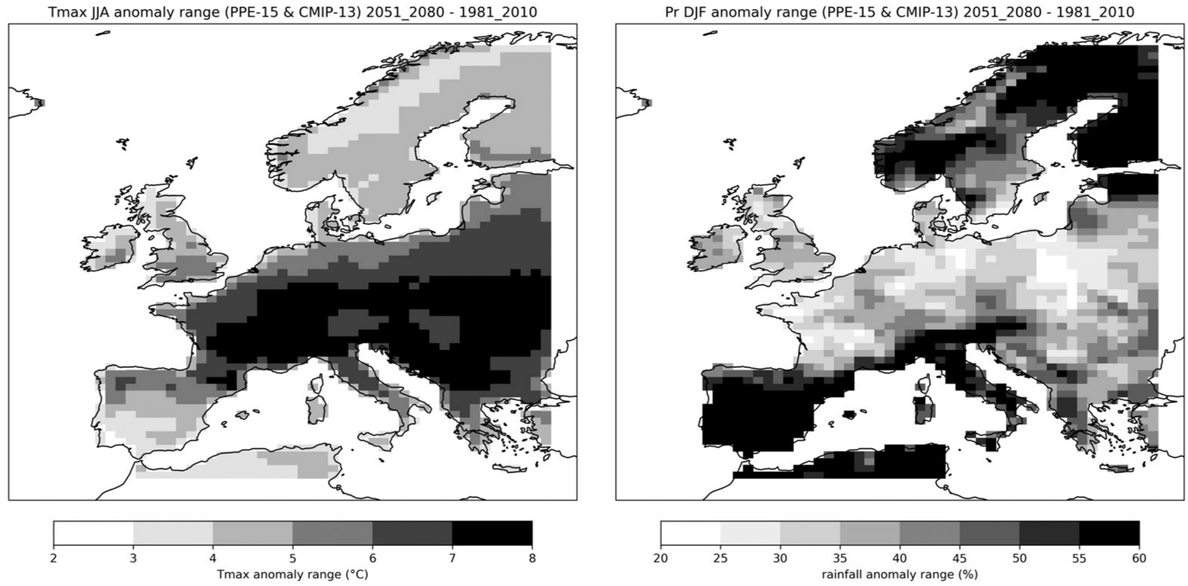


Figure 3.2-2 Full projection anomaly range (maximum change minus minimum change) for each grid cell across the PPE-15 and CMIP5-13 members for the selected hazards, JJA Tmax (left) and DJF Pr (right).The anomaly is the difference between the future period (2051-2080) and baseline period (1981-2010). Note: since the anomaly range for DJF Pr spans wetting and drying in some locations, the percentages given do not indicate rainfall increases of these magnitudes being projected.

METHODS: CLIMATE DRIVER ANALYSIS

User input: The use of weather patterns, or other atmospheric drivers, is an interesting framing of events and future change for the user. This is potentially related to the increased appearance of atmospheric drivers in forecasting and media. The aim of including it in the dialogue is to create a sense of plausibility and scientific grounding, as well as a narrative element related to place and experience which bridges the past and future.

We used the Met Office set of 30 European weather patterns for the weather typing, aggregated to 8 (labelled as weather type, WT, WT1 to WT8 hereafter), this is the driver information used to look at events and future change. The eight weather patterns are described in Figure 3.2-3 (adapted from Neal et al. 2016), with the descriptions corresponding to northwest Europe. Daily pressure field anomalies are weather typed for the ERA5 reanalysis data (C3S, 2017) and the UKCP PPE and CMIP GCM members (McSweeney et al., 2020, Pope et al., 2021), only 11 of the CMIP5-13 members had weather pattern data available (17 & 21 were not included in McSweeney et al., 2020). Weather typing involves hazard metrics or related variables being related to the different weather patterns, allowing them to be linked to events of interest, and events and future changes to be decomposed into anomalies in the occurrence frequency of the patterns (dynamic component) and changes to the climatology of the patterns (thermo-dynamic component, also including other teleconnections and drivers). The lower panels of Figure 3.2-3 show the distinct relative climatologies from the weather pattern analysis for JJA Tmax and DJF Pr, highlighting the differences in the variable/metric climatology and seasonal frequencies of the eight weather patterns. Relative climatologies are calculated as the individual weather pattern mean metric or variable, divided by the mean value of this over all weather patterns. This analysis was not performed for the 99th percentile (very hot/wet days), as doing this on a per ensemble member basis results in too few sample days over a 30-year period for robust statistics.

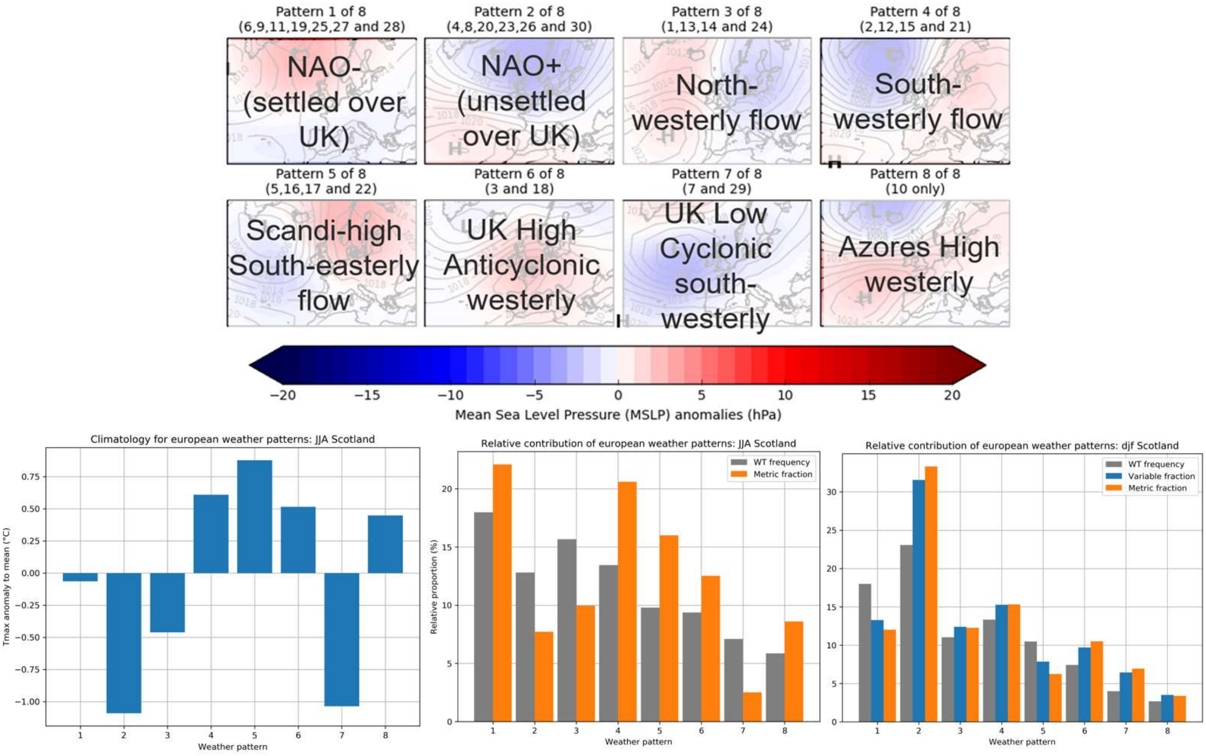


Figure 3.2-3 Top: The 8 weather patterns shown as mean sea level pressure anomalies, labelled with a north-west Europe centric description and the corresponding weather patterns of the full 30 that are included. Adapted from Neal et al. 2016. Bottom: Relative climatologies over the 8 weather patterns during the baseline period using ERA5 data (weather patterns) and HadUK-grid data (metric and variable) for JJA mean Tmax (left), contribution to the (fraction of) hot days (middle), DJF rainfall and wet days(right), the mean seasonal frequency of each weather pattern is also shown in the middle and right panels in grey.

METHODS: BIAS AND PLAUSIBILITY CHECKS

> *User input:* Given the central role of information on drivers of events, variability and future change, bias and plausibility checks allow any caveats or limitations to be communicated alongside the results, beyond the general importance of this step when producing climate service products.

The different aspects assessed were the bias in the underlying variable percentile thresholds and variability, as well as the seasonal metric variability. Additionally, the WT relative climatology, frequency and contribution to variability were investigated. At this stage the findings of metric and driver bias analysis across ensemble members are not fully utilised and becomes more important once sub-selection is performed. Some example plots and findings are discussed in Appendix 6.1.2.

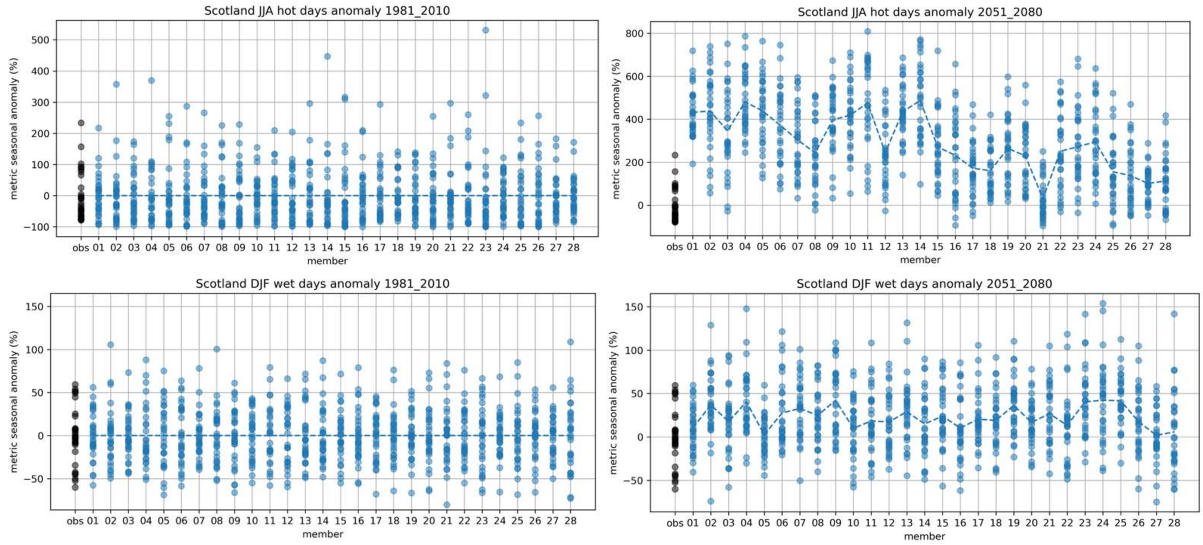


Figure 3.2-4 Top panels: Scatter plots of seasonal anomalies from the PPE-15 (members 1 to 15) and CMIP5-13 (members 16 to 28) during the baseline period & future periods for JJA hot days (Tmax90d) over Scotland (observation in black), the mean for each member is indicated by the dashed line in the future plots. Bottom panels: As above but for DJF wet days (Pr90d).

METHODS: PROJECTED FUTURE CHANGES

User input: For dialogue with the user, we needed to establish the signal and uncertainty range in future projected changes, in the context of past climate and events. The changes to the weather patterns, timing information (the pathway to the future mean) and consistency across regions and metrics were to be discussed as part of the co-production process.

In Figure 3.2-4 changes to the mean and distribution of both the seasonal JJA hot days and DJF wet days metrics are shown, for each ensemble member in the baseline and future periods. Changes to the mean seasonal values as well as the extremes may have relevance to member selection for representative storylines depending on the application. Here the correlation of the 30-year means and top 5 seasons (not shown) was sufficient that the 2051-2080 mean anomalies were used for the selection. Figure 3.2-5 shows how the unfolding of future changes varies across the members, highlighting the high (H), median (M) and low (L) members for the measures chosen to make the selection (selected excluding members with no WT information or with large biases in important features). Hence, the H/M/L members for JJA Tmax90d and DJF Pr are highlighted as this represents an initial option for the sub-selection of representative members for storylines, discussed further in the next section. The right panels show a scatter plot of the 2051-2080 wet/hot days metric and variable means, highlighting the degree to which they scale together and where the H/M/L members lay. Equivalent plots across Scottish regions and the very hot/wet days metrics show the selection is approximately consistent across these for JJA Tmax, also capturing close to the full range of uncertainty, but the full range of DJF changes are not fully captured by these members across regions and for the Pr99d metric (not shown but covered by tables in the appendix). The presence of greater multi-decadal variability in the DJF Pr timeseries plots highlights an additional difficulty in selecting members as representative storylines, as the selection depends heavily on particular 30-year period chosen.

The weather pattern analysis allows future changes to the variables and number of days metrics to be decomposed into a frequency change and a climatology change component, along with a residual largely related to cross-terms. These can then be analysed for each member, individually for each WT or summed into the overall components (Figure 3.2-6, top row), forming part of the storylines product as well as informing the selection process for choosing representative ensemble members. We see that the frequency components for JJA hot days are small across all members, whereas the frequency component plays a larger role in the DJF Pr changes, exceeding the climatology component in some cases (Sexton et. al, in prep). These changes are seen to vary across ensemble members and weather patterns, an example for the low, median and high members for winter rainfall is shown in the bottom panels, showing a wide diversity in the changes from individual weather patterns. Generally, the climatology components of WTs 4, 5 and 6 contributed most significantly to the Tmax90d metric changes, although further investigation is needed to determine if biases in the baseline period contribute to this. For DJF Pr, WTs 2, 4 and 7 generally showed the greatest climatology changes. The mapping between changes to the hazard metrics and the global temperature changes was considered during the selection process.

Changes to the intra-annual (day to day) and inter-annual (year to year) variability of the underlying seasonal variables, as well as the inter-annual variability of the threshold based seasonal metrics, were analysed (not shown). In most cases both measures increased, however this was not clearly correlated with the changes to the main hazard measures to be used for the selection of representative members (DJF Pr anomaly and Tmax90d anomaly). The overall trend of increasing variability, as well as the values for the specific members on which storylines were based were presented in the output (see the tables in Appendix 6.1.1).

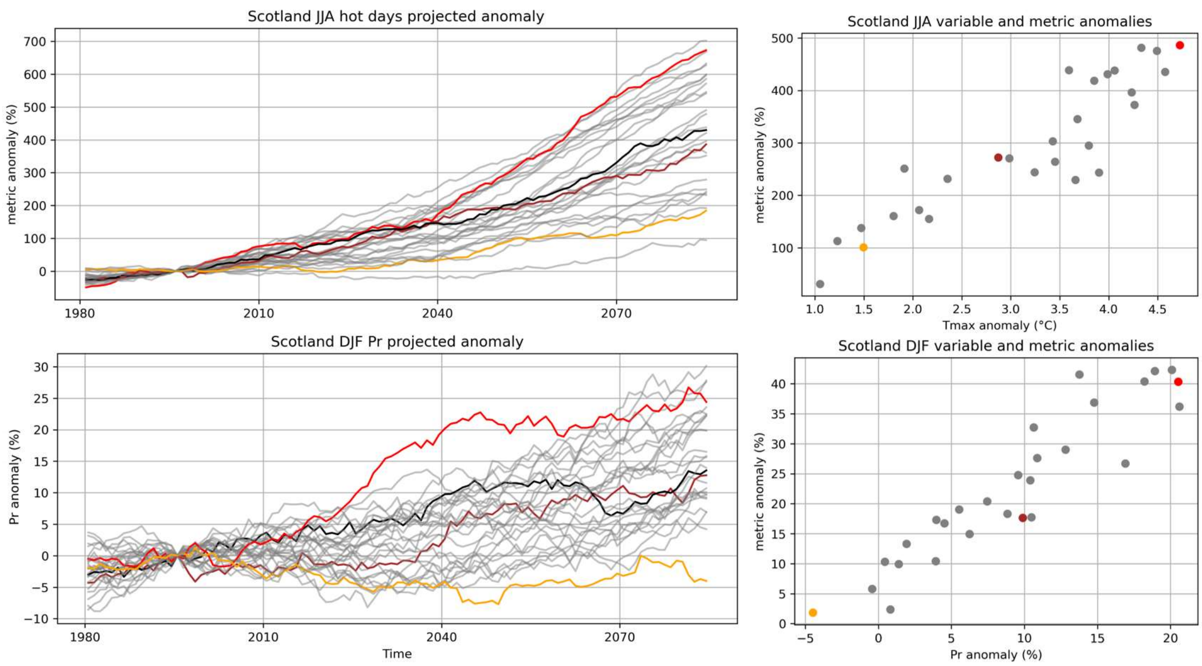


Figure 3.2-5 Seasonal metric 30-year mean time series used for member selection, and scatter plots of the values for 2051-2080. JJA number of hot days (metric) and Tmax anomalies are shown in the top panels, DJF Pr and number of wet day anomalies (metric) are shown in the lower panels. High, median and low members are highlighted in red, brown and yellow, with the closest to the mean in black (members missing WT data are present but ignored in the selection).

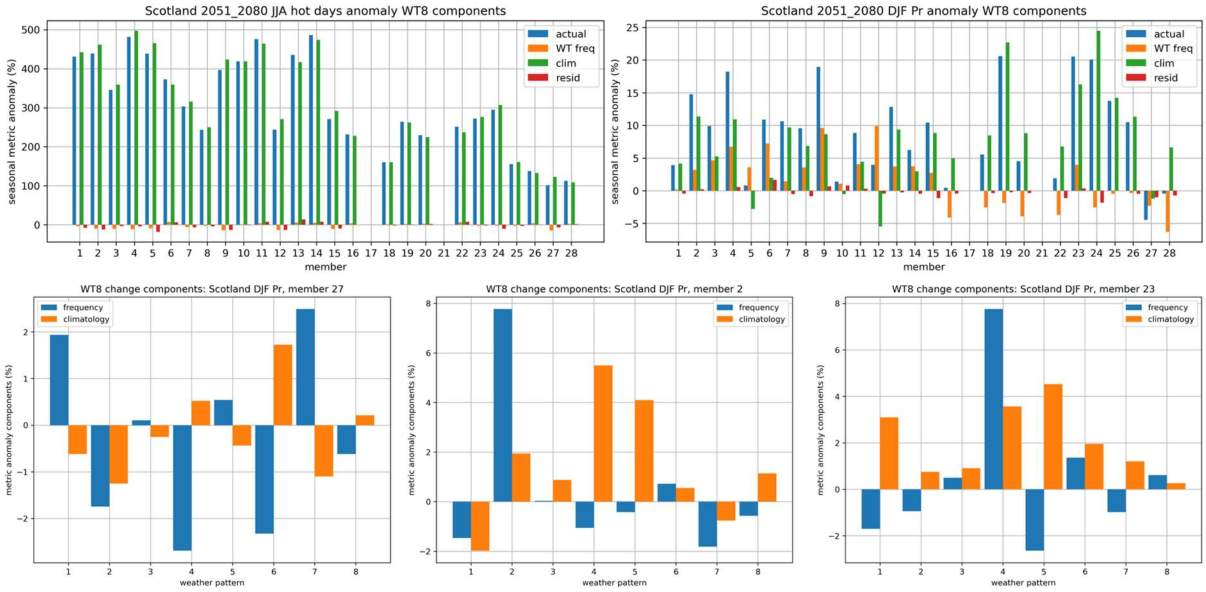


Figure 3.2-6 Upper panels: weather pattern frequency and climatology components of future change for the seasonal JJA hot days and DJF Pr metrics, for each PPE-15 and CMIP-11 member (1 to 28 along the x-axis). The four bars refer to the 'actual' change projected by a given member, the change due to the frequency component ('WT freq'), the change due to the weather pattern climatology component ('clim'), and a residual ('resid'). Bottom panels: weather pattern frequency and climatology components for individual members, decomposing the DJF Pr90d metric change in the upper panels into the contributions of individual patterns (1 to 8 along the x-axis) for the high, median and low DJF Pr members (23, 2 and 27), as a percentage anomaly in the metric.

METHODS: INFORMATION SELECTION OPTIONS

Here potential selection options to produce storylines from the available information are discussed, as summarised in the table below.

Option	Pros	Cons
Hazard metric full uncertainty range	• Test the relevance of the full uncertainty range in the hazard metric • Simple to understand • Physical plausibility may still be assessed • No confidence or likelihood assessment	• Not physically based • Wide uncertainty range • Based on a single hazard or impact metric
Hazard metric clustering	• Potential to include multiple metrics • Possibility to introduce likelihood via 'number of votes'	• Full range of individual metrics not explored • Requires large ensembles • Not physically based
Driver-based clustering	• Physically based • Possibility to introduce likelihood via 'number of votes'	• Full range of hazard metrics not explored • Requires larger ensembles
Pan-European selection	• Consistent set of information across Europe • Could be driver based	• Limited hazard and driver coverage at regional levels

Table 1 Potential options to select representative ensemble members to act as the basis for storylines for a user application which explore uncertainty, and associated pros and cons. Pan-European selection refers to the possibility that a project such as EUCP provides a representative set of storylines selected at a European level, to be used across many applications.

For this application a starting point for selecting representative members is to assume the full range of uncertainty may be relevant (unless considered implausible), so to begin by selecting high, median and low (H/M/L) members independently for each hazard to present to users, and work through an iterative co-production process to reduce this range where appropriate. The DJF Pr and JJA Tmax90d anomalies were used for the selection, as shown in Figure 3.2-5. This is because 'easy to communicate' selections showing the full range were requested, and this still samples the uncertainty range of most other hazard measures well. In the end, this was the final sub-selection approach used for the prototype information as representing the full range of both future hazards was considered application relevant, with some of the other options discussed further below. A full table of values from the selected simulations, alongside the full range, is provided in Appendix 6.1.1, and this selection was used for the prototype product evaluated in the next section. Selection was performed based on Scotland-wide averages; however, the regional values and ensemble ranges were also included in the DJF rainfall information as there is significant regional variation in baseline values, future anomalies and the associated weather patterns. Other selection options, and related plots, are presented in Appendix 6.1.3.

Figure 3.2-7 shows that the selected storylines for this case study may approximately represent the uncertainty range in a wider European context for summer temperatures, but not for winter precipitation, emphasising the difficulty in performing Europe-wide selection of representative ensemble members for applications beyond the domain-wide mean changes (Europe-wide anomaly maps for the selected GCM simulations for both hazards are shown). This illustrates pan-European narratives for similar applications would need their own selection against wider geographical evaluation and may also suggest that any pan-European narratives are likely to be less effective at spanning impact-relevant projection outcomes in smaller geographical regions. There would be a need to balance narratives tailored to local scales against identifying narrative changes that provide physical consistency across larger scales, depending on the application (that would enable impacts, such as food security, to be joined up).

METHODS: SUMMARY

The goal of producing storylines which explore uncertainty, tailored to an application relevant range but selected to represent different behaviours of physical drivers is currently difficult to achieve. Due to the constraints of ensemble size, and the lack of clearly distinct driver behaviours this was not used as the basis for sub-selection in the user product here. However, the physical driver-related information from the weather pattern analysis was still utilised, and the selection of projection information was application relevant in terms of spanning the full range of uncertainty available. The information produced for the H/M/L members was tested with the user, and summarised in Table 7 and Table 8 in the appendix, assessing the impact of sub-selection in this manner is important for scientific understanding, as well as allowing the user to judge if the selection is appropriate across different hazard measures and regions.

In Appendix 6.2 UK-wide plots for JJA Tas and DJF Pr from Section 3.4 (Lines of Evidence assessment) are shown, which help put the anomalies in the underlying variables seen in the UKCP projections in a wider context, as well as shedding light on the potential results from using higher resolution downscaling and use of constraints. The next section summarises the work on understanding the usefulness and useability of the information provided to the user.

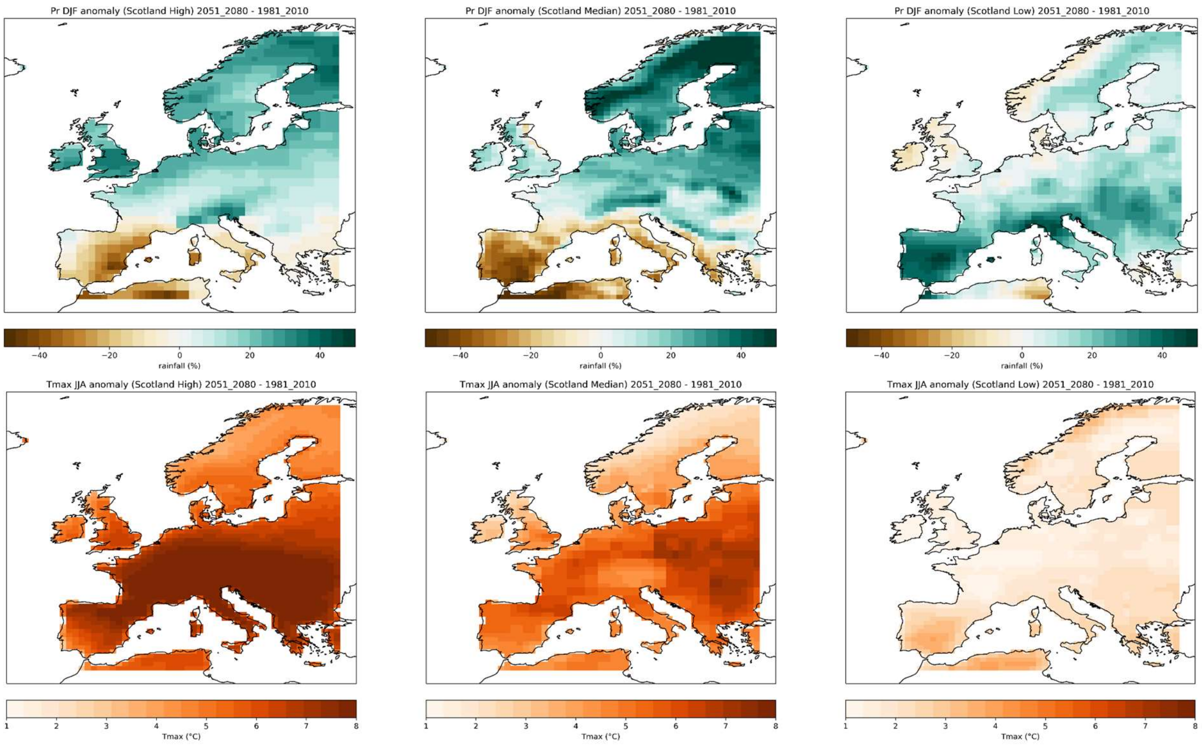
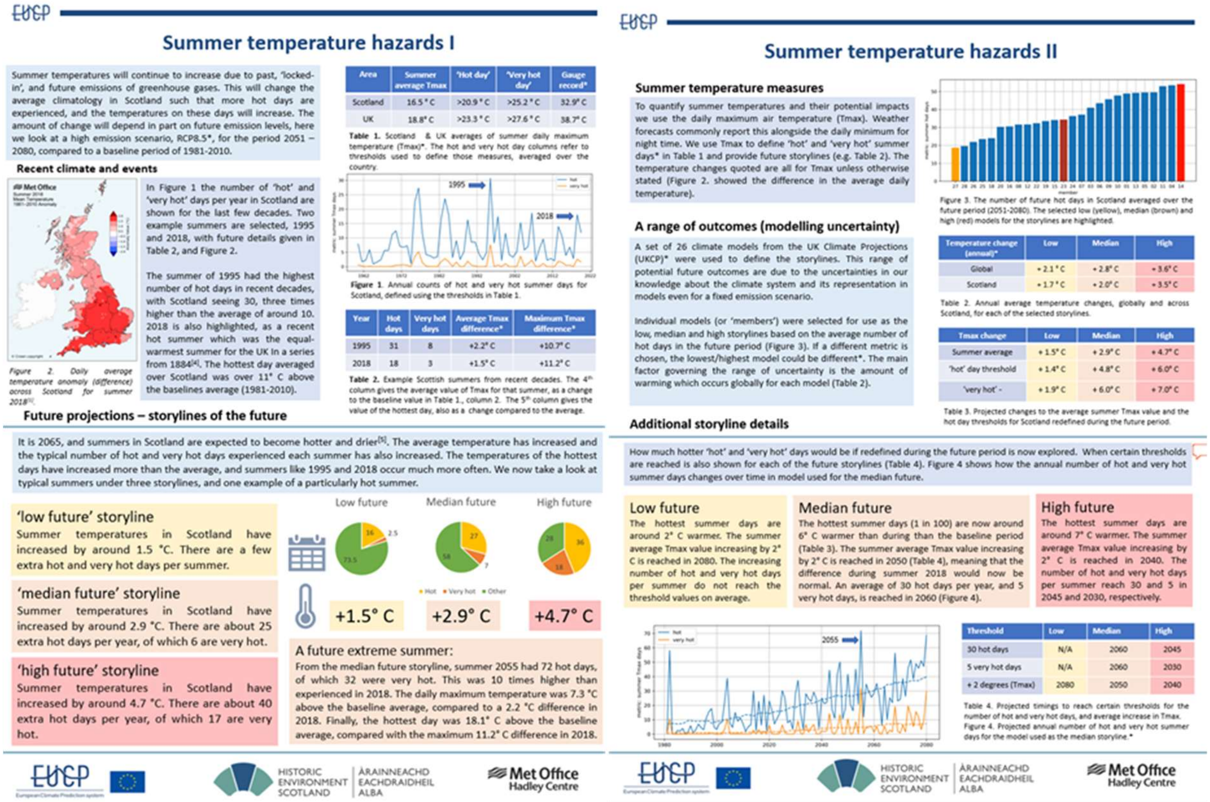


Figure 3.2-7 Top panels: European DJF Pr anomalies for the high, median and low members (left to right, members 23, 2, & 27) selected based on mean Scotland DJF Pr anomaly. Bottom panels: as above, but JJA Tmax for members selected (left to right, members 14, 23, & 27) based on mean Scotland JJA Tmax90d anomaly.



Summer temperature hazards I

Summer temperatures will continue to increase due to past, 'locked-in', and future emissions of greenhouse gases. This will change the average climatology in Scotland such that more hot days are experienced, and the temperatures on these days will increase. The amount of change will depend in part on future emission levels, here we look at a high emission scenario, RCP8.5*, for the period 2051 – 2080, compared to a baseline period of 1981-2010.

Area	Summer average Tmax	'Hot day'	'Very hot day'	Gauge record*
Scotland	16.5 °C	>20.9 °C	>25.2 °C	32.9° C
UK	18.8° C	>23.3 °C	>27.6 °C	38.7° C

Table 1. Scotland & UK averages of summer daily maximum temperature (Tmax)*. The hot and very hot day columns refer to thresholds used to define those measures, averaged over the country.

Recent climate and events

In Figure 1 the number of 'hot' and 'very hot' days per year in Scotland are shown for the last few decades. Two example summers are selected, 1995 and 2018, with future details given in Table 2, and Figure 2.

The summer of 1995 had the highest number of hot days in recent decades, with Scotland seeing 30, three times higher than the average of around 10. 2018 is also highlighted, as a recent hot summer which was the equal-warmest summer for the UK in a series from 1884[4]. The hottest day averaged over Scotland was over 11° C above the baselines average (1981-2010).

Figure 2. Daily average temperature anomaly (difference) across Scotland for summer 2018[1].

Figure 1. Annual counts of hot and very hot summer days for Scotland, defined using the thresholds in Table 1.

Year	Hot days	Very hot days	Average Tmax difference*	Maximum Tmax difference*
1995	31	8	+2.2° C	+10.7° C
2018	18	3	+1.5° C	+11.2° C

Table 2. Example Scottish summers from recent decades. The 4th column gives the average value of Tmax for that summer, as a change to the baseline value in Table 1., column 2. The 5th column gives the value of the hottest day, also as a change compared to the average.

Future projections – storylines of the future

It is 2065, and summers in Scotland are expected to become hotter and drier[5]. The average temperature has increased and the typical number of hot and very hot days experienced each summer has also increased. The temperatures of the hottest days have increased more than the average, and summers like 1995 and 2018 occur much more often. We now take a look at typical summers under three storylines, and one example of a particularly hot summer.

'low future' storyline
Summer temperatures in Scotland have increased by around 1.5 °C. There are a few extra hot and very hot days per summer.

'median future' storyline
Summer temperatures in Scotland have increased by around 2.9 °C. There are about 25 extra hot days per year, of which 6 are very hot.

'high future' storyline
Summer temperatures in Scotland have increased by around 4.7 °C. There are about 40 extra hot days per year, of which 17 are very hot.

Low future | Median future | High future

Hot | Very hot | Other

+1.5° C | +2.9° C | +4.7° C

A future extreme summer:
From the median future storyline, summer 2055 had 72 hot days, of which 32 were very hot. This was 10 times higher than experienced in 2018. The daily maximum temperature was 7.3 °C above the baseline average, compared to a 2.2 °C difference in 2018. Finally, the hottest day was 18.1° C above the baseline average, compared with the maximum 11.2° C difference in 2018.

HISTORIC ENVIRONMENT SCOTLAND | ÀRAINNEACHD EACHDRAIDHEIL ALBA
Met Office Hadley Centre

Summer temperature hazards II

Summer temperature measures

To quantify summer temperatures and their potential impacts we use the daily maximum air temperature (Tmax). Weather forecasts commonly report this alongside the daily minimum for night time. We use Tmax to define 'hot' and 'very hot' summer days* in Table 1 and provide future storylines (e.g. Table 2). The temperature changes quoted are all for Tmax unless otherwise stated (Figure 2. showed the difference in the average daily temperature).

Figure 3. The number of future hot days in Scotland averaged over the future period (2051-2080). The selected low (yellow), median (brown) and high (red) models for the storylines are highlighted.

A range of outcomes (modelling uncertainty)

A set of 26 climate models from the UK Climate Projections (UKCP)* were used to define the storylines. This range of potential future outcomes are due to the uncertainties in our knowledge about the climate system and its representation in models even for a fixed emission scenario.

Individual models (or 'members') were selected for use as the low, median and high storylines based on the average number of hot days in the future period (Figure 3). If a different metric is chosen, the lowest/highest model could be different*. The main factor governing the range of uncertainty is the amount of warming which occurs globally for each model (Table 2).

Temperature change (annual)*	Low	Median	High
Global	+2.1 °C	+2.8° C	+3.6° C
Scotland	+1.7 °C	+2.0° C	+3.5° C

Table 2. Annual average temperature changes, globally and across Scotland, for each of the selected storylines.

Tmax change	Low	Median	High
Summer average	+1.5° C	+2.9° C	+4.7° C
'hot' day threshold	+1.4° C	+4.8° C	+6.0° C
'very hot' -	+1.9° C	+6.0° C	+7.0° C

Table 3. Projected changes to the average summer Tmax value and the hot day thresholds for Scotland redefined during the future period.

Additional storyline details

How much hotter 'hot' and 'very hot' days would be if redefined during the future period is now explored. When certain thresholds are reached is also shown for each of the future storylines (Table 4). Figure 4 shows how the annual number of hot and very hot summer days changes over time in model used for the median future.

Low future
The hottest summer days are around 2° C warmer. The summer average Tmax value increasing by 2° C is reached in 2080. The increasing number of hot and very hot days per summer do not reach the threshold values on average.

Median future
The hottest summer days (1 in 100) are now around 6° C warmer than during than the baseline period (Table 3). The summer average Tmax value increasing by 2° C is reached in 2050 (Table 4), meaning that the difference during summer 2018 would now be normal. An average of 30 hot days per year, and 5 very hot days, is reached in 2060 (Figure 4).

High future
The hottest summer days are around 7° C warmer. The summer average Tmax value increasing by 2° C is reached in 2040. The number of hot and very hot days per summer reach 30 and 5 in 2045 and 2030, respectively.

Threshold	Low	Median	High
30 hot days	N/A	2060	2045
5 very hot days	N/A	2060	2030
+ 2 degrees (Tmax)	2080	2050	2040

Table 4. Projected timings to reach certain thresholds for the number of hot and very hot days, and average increase in Tmax.
Figure 4. Projected annual number of hot and very hot summer days for the model used as the median storyline.*

HISTORIC ENVIRONMENT SCOTLAND | ÀRAINNEACHD EACHDRAIDHEIL ALBA
Met Office Hadley Centre

Figure 3.2-8 Pages from a draft prototype product on the summer temperature hazard storylines which acted as a discussion piece with the stakeholder in HES, as well as the wider sector audience from the focus group session. This provides an idea of the final prototype, and the detail is not intended to be important here.

USER EVALUATION OF THE INFORMATION AND STORYLINES PROTOTYPE

Following the above analysis, a prototype product to communicate climate hazard information from the observations, the selected uncertainty spanning members, ('storylines') and driver analysis was developed. This was to aid in conversations assessing usefulness and usability of the information and this manner of presentation, and involved iteration and tailoring based on input from the user. JJA Tmax hazards were covered first, and an example excerpt is shown in Figure 3.2-8 (just for context, the detail not important). The same underpinning information was produced for DJF Pr, to allow for discussion of the different hazard and added complexities (not shown in the prototype, full data table in Appendix 6.1.1).

The prototype example consisted of an introduction, three pages of tiered information on the storylines with increasing detail, a placeholder section on what this means for the organisation with suggested themes, and a glossary and further details section. The tiered, or layered, provision of the information from headline messages through to regional variation and driver information aims to increase the flexibility of the product and allow it to be disseminated for different purposes withing HES. This was also used as a discussion piece with the wider sector in a focus group session attended by eight additional sector representatives from across Europe. Some key findings on the information presented from HES and the focus group participants are summarised below, representing individual points made, as interpreted by the information providers (further details of the included information, focus group participants and feedback are given in Appendix 6.1.4).

- There are uses of worst- and best- case 'scenarios'. The finite resource of Heritage and the potential loss from taking action requires minimising adaptation, hence planning for the best case. This contrasts with the need to know what the worst case looks like and to raise awareness in the sector to enable action (mitigation and adaptation).
- A focus on past events and impacts, and example decisions (including what could have been different) will maximise take-up. Risk and 'unseen' events in the current climate should be explored further in the sector.
- There is an appetite to move to more regional climate hazard messages, with more detail, rather than national headline messages, as demonstrated in the included information across the Scottish regions (e.g. differences in the drivers and patterns of rainfall in east vs west Scotland).
- Storylines could aim to help address underutilisation of available climate information, pull through and test new science, and introduce new concepts to users. This may be done in a tiered manner, layering information to provide consistent storylines across headline messages to increasingly detailed information.
- A clear decision framework is important to allow true testing of storylines for decisions, likely requiring an organisation that already uses other forms of climate information routinely in decisions and planning.

Some more specific findings on the information and example prototype can be found in the appendix, some of the main points which emerged were:

- Framing individual ensemble members as alternative futures, consistent over different pieces of information, was said to increase the understandability of uncertainty and the tangibility of the future hazards.
- The information about weather patterns and variability was well received and perceived to strengthen the link between the past events and future projections and understanding of uncertainty and regional variation.
- Supplementing likelihood focused analysis and headline messages with more narrative approaches appears to be important for sectors outside those typically considered in climate services research. However, holistic decisions are made, therefore it is more likely climate information would form part of the justification for a decision, rather than the reason for it.
- To go beyond the present example (fully tailoring the uncertainty range to the application) requires a deep understanding from both the science or service provider and the user. In future work there is a need for deep specialists in user engagement and decision science.
- A single emission scenario was used, RCP8.5, and have instead explored modelling uncertainty (including climate sensitivity) and natural variability. The choice of 'high, median, low' nomenclature was pointed out to be potentially confusing, with many readers likely to assume this relates directly to emission scenarios or warming levels.
- An interesting comment was that the actual metric/measures may not be particularly important at this stage, and what was important to communicate was the degree of change, in a manner making it possible to link to past events through comparisons and impacts.

Some potential benefits & uses for similar climate narratives or storylines in the sector and beyond were discussed during the interaction with HES and the focus group, where the aim could be;

- co-production and 'meeting in the middle', allowing feedback in both directions and the potential to aid in understanding complicated aspects such as the uncertainty, and to create a sense of shared ownership,
- to enable and encourage uptake of climate information (exploring what is possible),
- a way to communicate potential impacts to raise awareness internally and externally,
- to help bring climate information into decisions and planning,
- an aid in thinking about risk appetite, tolerance, and change management in adaptation processes.

Naturally, there are also downsides to this approach, from the resources required to pursue it, to the potential of storylines or other narrative user products to draw attention away from other information sources which may be equally or more relevant to a given application.

3.2.3 Water supply management - storylines of future drought

BACKGROUND

The water resources sector across Europe is one which is very familiar with using climate change information and have pan-European level information available to them (e.g. European Drought Observatory at https://edo.jrc.ec.europa.eu) but also within each of EU member state. Much work has already been done as part of the Copernicus Climate Change Service to provide drought information for users (e.g. and the Drought impact on water resources forecasting tool at https://ecmwf-ukrain-demo.hrwallingford.com). Like other European countries, water resources are a regulated in the UK; in England, the regulators include the governmental department responsible for the environment, environmental protection agencies. As the water supply was privatised in the 1990s, England also includes a consumer regulator that regulates the water companies that manage public water supply. Every 5 years, water companies need to present a revised drought plan (Defra, 2021) focused on operational management and a long-term water resources management plan for the Periodic Review (Ofwat, 2020) focused on long-term investments that build resilience in water supply for the next 100 years. Here we present a storyline production process which involved representatives from three potential users of climate storylines: the environmental protection agency (Environment Agency), a water company (Anglian Water) and a consultancy that often provides the climate services for water resources risk assessment (HR Wallingford).

The User Requirement for Storylines and Decadal Forecasts

Following the process described in Figure 3.2.1, a set of interviews were carried out with the stakeholders at the start of the project to (i) understand their current data needs over and beyond the information that they already have available (ii) if and where decadal forecasts could be embedded in decision-making (iii) what the potential for physically based climate storylines could be in their decision contexts. The outcome of this consultation is summarised in Figure 3.2-9.

Embedding decadal forecasts in decision-making

- The key question is whether the data helps answer "How do we invest?"
- 1-10 years is "a bit of a blind spot for the industry"
- How can the industry deal with yearly changing forecasts? (or how can you make a decision in anticipation of yearly changing forecasts?)
- Currently no climate information used beyond historical resampling and climate projections

Potential of storylines for informing decisions

- Would be useful to help inform investment pathways
- There may be issues moving from strict planning timetable to adaptive pathways
- Management plans generally use drought scenarios for designing and testing resilience of supplies. High potential for use of narratives to inform these.
- Process understanding is key to understand what governs the changes in drought.
- Any additional analysis is helpful to build portfolio of evidence to back investment decisions

Storyline content

- Do not forget the probabilistic projections approach, i.e. using the UKCP09 Probabilistic Projections*
- Range of uncertainty even without signal is important
- Historical information for context
- What can storylines of decadal forecasts bring to the table compared to probabilistic distribution functions?
- Data needs to be spatially coherent and daily
- If not using hydrological models, need daily drought index as existing monthly ones not helpful

Figure 3.2-9 Overview of user requirements of representatives of the public water supply management sector in the UK. *UKCP09 were the previous set of UK Climate Projections published in 2009 and focused on the providing Probabilistic Projections of future climate.

We note that in the storyline content that our stakeholders would like to understand how their existing approaches may be changed using storylines. In the last Periodic Review in 2019, UK water companies based their climate change information on UKCP09 (Murphy et al, 2009) which focused on Probabilistic Projections of future climate as well as a future river flows and groundwater levels dataset based on an ensemble of 11 regional climate models for one emissions scenario (von

Christierson et al, 2013). Water companies used either or both a probabilistic and small climate model ensemble approach to assessing water resources.

During the past two years, water companies have been preparing their water resources management plans (WRMPs) to be completed by 2024. While the overall water supply assessment method approach is unlikely to change there are additional factors that are being taken into account. These are the introduction of the UK Climate Projections released in 2018, UKCP18 (Lowe et al, 2018) and the move towards meeting the water supply deficit through bulk water transfers across the UK (Defra, 2018; Environment Agency, 2020). This calls for spatially coherent information so that regions across the UK can assess water resources consistently. This isn't possible using the UKCP18 Probabilistic Projections. Our stakeholders believe that the potential for physically based storylines even without probabilistic information can potentially provide additional information on how droughts may occur and potentially confidence in the climate models' ability to simulate drought. In fact, the industry is currently using a combination of global climate model and probabilistic information to inform their assessment. In addition, some water companies use climate model information as the basis for building stochastic weather simulators to capture statistics to feed into their "level of service" assessment (i.e. whether they are able to meet the water demand). In this study, we focus the analysis using deterministic climate models and discuss the potential for probabilistic information to be included in storylines.

We did not explore in detail the potential of results from temporal and spatial merging given the novelty of decadal forecasts and the limitations of current planning frameworks that work on operational (the next year) and long-term investment timescales. The requirements summarised in Figure 3.2-9 provided the basis of the investigation into physically based storylines of future droughts for the UK. This includes:

- Investigating the use of a daily drought index to provide storylines without the need for further downstream modelling. We deemed this necessary to focus the research on how useful storylines based solely on climate models were.
- Linking the drought index to weather patterns to investigate future droughts to understand the role of atmospheric processes in storyline construction for the water supply sector.
- Exploring where the useful storylines are by constructing different storyline typologies using the existing climate model data at the decadal and projections timescales.

The climate change assessment in the WRMPs is underpinned by water resources systems modelling with each water company running their own. These require physical information such as precipitation, potential evapotranspiration and resulting river flows and groundwater level as well as scenarios of water demand. In this task we investigate drought indices that provides a method for characterising a climate model ensemble, aiding the identification of drought events and/or ensemble members of concern so that stakeholders can focus their detailed simulations and analysis rather than investigating all possible future climates. The latter can be a large burden for water companies who often reduce the number of simulations (von Christierson et al, 2012; von Christierson et al, 2013).

THE HISTORICAL DROUGHT CONTEXT: DROUGHT INDICES AND WEATHER PATTERNS

In Figure 3.2-10, we show three self-calibrating (i.e. no need for bias-correction) precipitation-based drought indices for a drought event often used to test the resilience of water resources infrastructure, i.e. the Standpipe drought of 1976 (Rodda et al, 2011). We compare two monthly drought indices used in the peer-reviewed literature: the Standardised Precipitation Index (SPI) (McKee et al, 1993) and the Drought Severity Index (DSI) for three accumulation periods (12, 24 and 36 months) (Phillips and MacGregor, 1998). In addition, we show the daily Effective Drought Index (Buhn and Wilhite, 1999) which has been included to address the stakeholders' requirement for an index with higher temporal resolution. All drought indices are able to track the drought based solely on accumulated rainfall and its peak for two large UK drought events: the summer 1976 Standpipe drought (as named in the Drought Inventory available at https://historicdroughts.ceh.ac.uk/content/drought-inventory) and summer 2018.

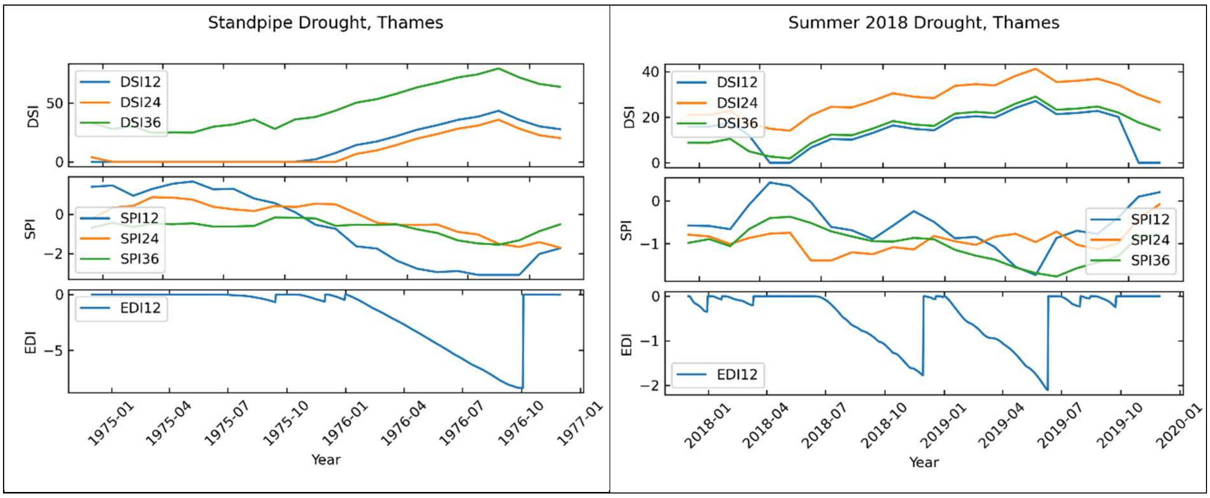


Figure 3.2-10 Time evolution of the 1976 Standpipe Drought (left panel) and the Summer 2018 Drought (right panel) for the Thames basin region using precipitation-based drought indices where, SPI-n is the n-month standarized precipitation index (McKee et al, 1993), DSI-n is the n-month drought severity index (Phillips and McGregor, 1998) and EDI is the Effective Drought Index (Buhn and Wilhite, 1999). Indices calculated using HadUK-Grid (Hollis et al, 2018).

Using the weather patterns described in Section 3.2.2, we show the weather patterns' monthly frequency anomalies during each of the droughts in Figure 3.2-11. Focusing on the winter recharge season (October to March) we see that the frequency anomalies of the weather patterns are very different between each of the droughts. For the Standpipe Drought, we see a clear decrease in WT1 (NAO-, where we should see more settled, drier conditions) and increase in WT4 (Southwesterly flow, slightly wetter conditions). For the Summer 2018 drought, we see an increase in WT1 and a small increase in WT4 which seems more intuitive. For the Standpipe Drought, the results are counter-intuitive and indicates further work required to understand the relationships as seen in Figure 3.2-6 between the weather patterns and precipitation patterns.

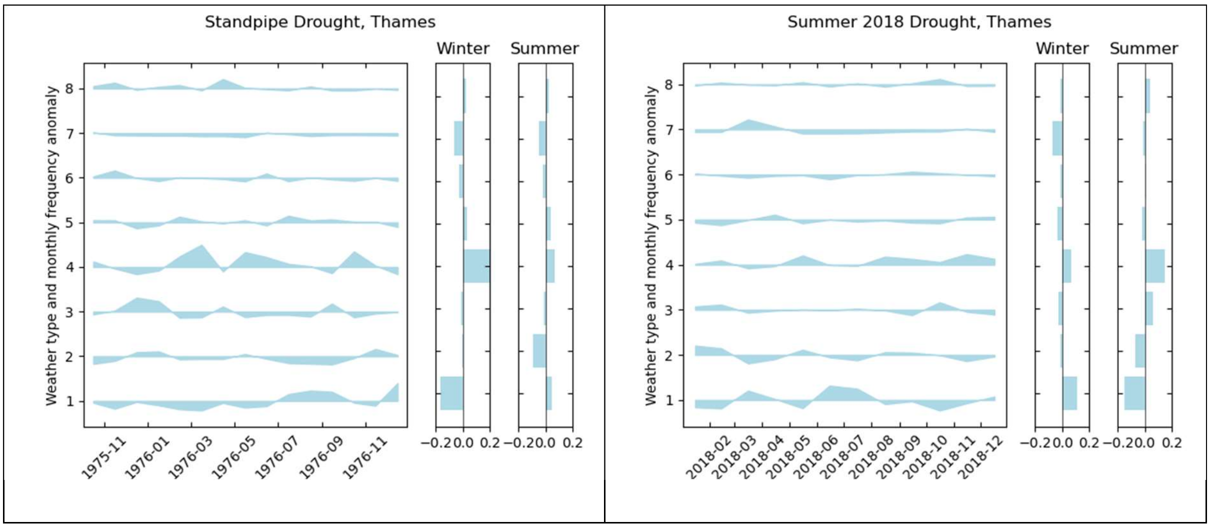


Figure 3.2-11 Eight weather pattern monthly frequency anomalies compared to 1981-2010 during the Standpipe Drought in 1976 for the Thames basin region (left panel) and the Summer 2018 Drought (right panel). Plots on the left are the time evolution of the weather patterns through the drought period defined by the Effective Drought Index (Buhn and Wilhite, 1999) and plots on the right are the aggregated anomalies for the winter recharge period (October to March) and summer period (April to September).

STORYLINES OF 21ST CENTURY DROUGHTS

The UK water sector uses the UK Climate Projections for planning and we use them here with a focus on the results of the global climate model, UKCP Global, rather than UKCP Regional as it (i) provides a diversity of future outcomes (larger ensemble size and two emissions scenarios) unavailable with higher-resolution datasets (ii) precipitation drought indices in the UK will be driven mostly by large-scale circulation patterns rather than local effects: a comparison of cumulative frequency of EDI between the PPE-15 and their downscaled versions (not shown) indicates similar behaviour for EDI.

We analysed the EDI as projected by the PPE-15 and CMIP5-13 ensembles (see Section 3.2.2) with the objective of seeking drought events that could be used to test the resilience of water supply systems and provide additional information on the dynamic climate drivers of drought using weather patterns. As the water supply sector is particularly interested in the extremes, in Figure 3.2-12 (left panel), we select the top ten 5-year droughts (i.e. the 5-year mean EDI) seen across the PPE-15 and CMIP5-13 ensemble for each year through the 21st century. In Figure 3.2-12 (right panel) we select the top 10 droughts for each of the PPE-15 and CMIP13 ensembles to explore the weather patterns' frequency anomaly that may be causing the projected droughts. In both plots we see that the PPE-15 is generally drier with larger future drought events in the latter part of the 21st century, although one of the members of the CMIP5-13 is also similarly dry. Interestingly, we see that for the 1-year droughts in Figure 3.2-12 (right panel) weather patterns anomalies are similar, for longer droughts, there is a distinct behaviour change with the CMIP5-13 ensemble showing more "intuitive" behaviour with an increase in WT1 anomalies which generally bring settled drier conditions. Storylines that could explore the diversity of weather patterns that may provide the conditions for different types of future droughts is an approach that one of our stakeholders was particularly interested in.

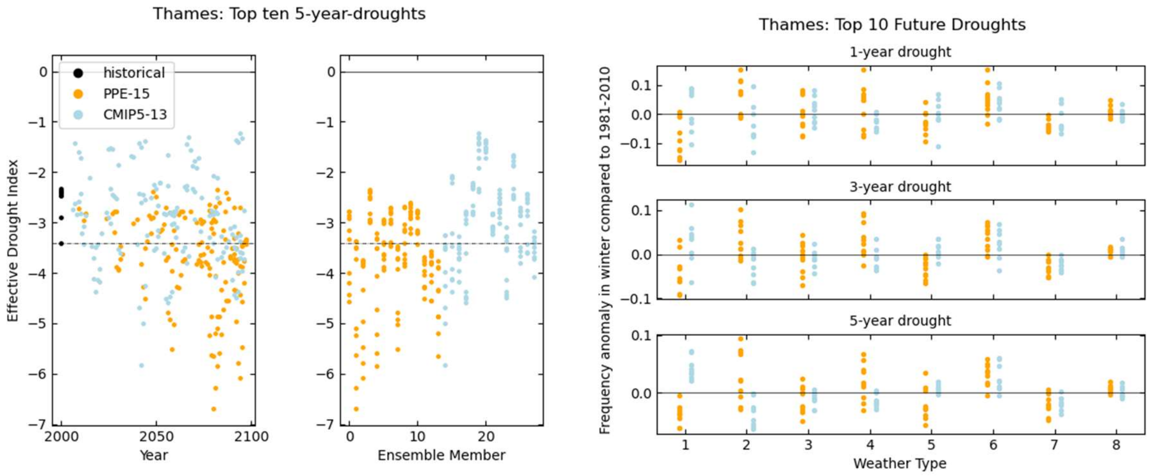


Figure 3.2-12 (a) Top ten 5-year droughts per year for the Thames basin region based on the Effective Drought Index (EDI), (Buhn and Wilhite, 1999) left panel: evolution of top 10 droughts across both PPE-15 and CMIP5 in the 21st Century for RCP 8.5, right panel: as left panel but for each ensemble member. Black dots are for observations (based on HadUK-Grid), orange dots are the PPE-15 and light blue dots are the CMIP5-13. The dashed line indicates the EDI value for the largest drought observed in the historical period based on HadUK-Grid (b) Eight weather pattern monthly frequency anomalies for the top 10 n-year droughts for the whole of the 21st century across the whole PPE-15 (orange dots) and CMIP5 ensemble (light blue dots).

During our user consultation, it became clear that as in Section 3.2.2, storylines of future drought would be tiered in detail. Where the headline story would be of increasing dryness that would lead to multi-year dry conditions and droughts: a narrative that is already well established in water sector but is most likely to be of interest around regional and national scale strategic adaptation options and comparing how different portfolios would perform. But further sub-storylines are required that could be:

1. Storylines that span the range of drought futures, i.e. two detailed storylines: one based on a climate model member that is at the dry-end and one that is at the wet-end (similar to the high/medium/low storylines of Section 3.2.2
2. Storylines of specific future drought events, i.e. extracting extreme drought events from the whole ensemble. These would provide events for stress-testing water resources systems.
3. Storylines that span the range of weather patterns that may drive droughts, i.e. extracting a CMIP-13 and a PPE-15 member that exhibits the 5-year drought behaviour seen in Figure 3.2-12.
4. Storylines of future drought events based on weather patterns similar to important historical droughts that are already used to stress-test systems (e.g. the Standpipe Drought), i.e. event analogues similar to those described in Section 3.3.2 for the summer 2018 drought event experienced across Europe.

While we have yet to consult all our stakeholders, there seems to be preference for option 4 as the water industry could potentially train statistical models or examine other products or even generate synthetic alternatives. Option 2 is also considered to be potentially useful to underpin the case for Strategic Regional Options which are local and regional infrastructure investment schemes being considered by the water regulators and do not require likelihood information (Ofwat, 2021); further work is required to investigate this application which would suit the storyline products considered here.

At the present, the industry does not use seasonal forecasts although there is some interest in their application. But the main hurdle is the requirement of high skill in the forecasts given that water companies are heavily penalised if droughts occur. At the moment, the water industry uses an envelope of forecasts based on historical years. Decadal forecasts is not known to have been discussed apart from the engagement as part of the EUCP project. However, out stakeholders anticipate that the decadal forecasts could potentially be used as part of their reporting as multiple sources of information are used for making decisions. However, spatial accuracy is very important given the kilometre-scale catchments in some parts of the UK where the accurate location of rainfall can be critical). Given the novelty of seasonal and decadal forecasts for the water industry, engaging with them on storylines of future drought at the decadal timescales proved to be more challenging.

In Figure 3.2-13, we show the 2020 drought forecast using the EDI for the Anglian region for ten members of the Met Office's Decadal Prediction System (Smith et al, 2013). Here both the magnitude and the likelihood of a drought based on the ensemble appear to be important to our stakeholder group. Two droughts of similar magnitude to the summer 2018 drought occur in the forecast. However, the number of ensemble members that show this is small. Further work is still required to (a) confirm with users that the range of event sets shown in a decadal forecast context is as important as the ensemble mean forecast as indicated in Figure 3.2-9 – this could pave the way for storyline construction similar to the climate projections (b) understand whether confidence in the forecasts are built should they agree year on year into the future (c) whether the appetite for using weather pattern-based storylines would be similar to that for climate projections. As with the climate projections, our user group is primarily interested in the credibility of the drought forecast before being able to use them in anger.

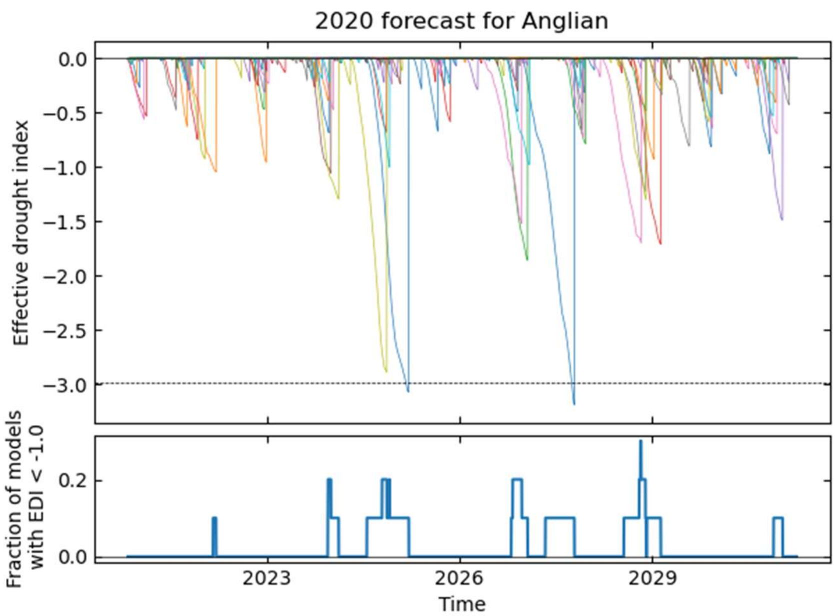


Figure 3.2-13 Top panel: Effective Drought Index (Buhn and Wilhite, 1999) for the Anglian basin region based on the Met Office Decadal Prediction System 10-member ensemble. Colours represent each model member. Dashed line represents the minimum EDI value during the summer 2018 drought. Bottom panel: fraction of 10-member ensemble that have a value of EDI less than -1.

In conclusion, through our user consultation, we observe that the industry is unlikely to be able to use the storylines directly at the moment, particularly weather-pattern based storylines. However, what it does provide is a potential way forward to understand better some of the methods that are currently being used to inform water resources such as the stochastic approaches dependent on circulation indices. The industry uses multiple sources of evidence to make the case for long-term investments. One of our stakeholders believes that the storylines construction process here builds on the evidence base that the sector is at risk to drought events and that these events are exacerbated by climate change, i.e. they help further to build the case that companies need to invest sooner rather than later.

CHALLENGES IN STORYLINE CONSTRUCTION

While constructing the storylines, there were a number of challenges that we faced when applying them to in a decision context. These included:

- The weather patterns driving extreme local drought conditions can vary across the UK as shown in Figure 3.2-14, where weather pattern anomalies driving the top ten droughts in northwest UK differ greatly to those in the east. Therefore, defining a set of consistent storylines that apply for all geographical regions may be a challenge.
- When constructing storylines to water supply requires further investigation to fully assess the drought risk, i.e. water resources managers will likely need to feed the storylines into their systems models to build confidence the forecasts/projections of hydrological drought. The challenge in this case is how much more modelling and information is required in the storylines to make them more decision relevant. The stakeholders engaged here were a group with a detailed understanding of climate science. Other organisations may need some capacity building in order to use them.

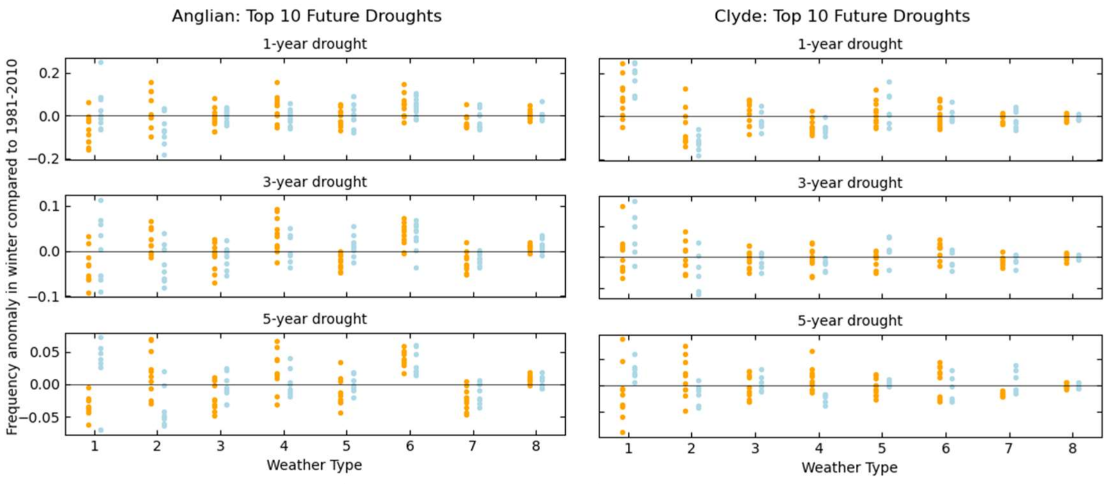


Figure 3.2-14 Weather pattern frequency anomalies for the top 10 n-year droughts in the 21st century for the Anglian basin region in eastern UK and the Clyde basin region in northwestern UK where orange dots represent the members of the PPE-15 and blue dots members of the CMIP5-13.

3.3 Building blocks of storylines

In this section, we present explorations carried out in EUCP to understand different ways in which storylines can be constructed, including event-based storylines (Sections 3.3.1 and 3.3.2), a clustering approach (Section 3.3.3) as well as including more realistic decadal variability (Sections 3.3.4 and Section 3.3.5). In framing the work in this deliverable, we consider the event-based storylines and analysis of variability as sitting in the 'scientific building blocks of storylines' segment of the Venn diagram in Figure 3.1-1, and the clustering study as a form of 'lines of evidence assessment', as indicated. Using the outputs to create a 'storylines product' for a particular sector or user would then sit under the overlapping segments.

3.3.1 Event-based storylines of heavy convective rainfall using a PGW approach (UCPH)

INTRODUCTION

In the evening on July 2, 2011, a severe cloudburst occurred over Copenhagen, Denmark. Between 90 and 135 mm of precipitation was recorded in less than 2 hours, which flooded the city and caused hundreds of millions of Euros in insured damages. Figure 3.3-1 shows the radar reflectivity indicating the evolution of the precipitation pattern. The precipitation started to develop around 14 UTC over southern Sweden and intensified south of Sweden and over the Copenhagen area (indicated by the black arrow in Figure 3.3-1 panel a) by the end of afternoon. High reflectivity is clearly seen over Copenhagen around 16 UTC and 17 UTC (panels d-e). It is worth noting the spatial pattern of the precipitation system, i.e. locally intense precipitation covering rather small areas which is a signature of convective precipitation. The impact on society of such rare events is important and must be understood in the context of global warming. Simulating such an event accurately is still challenging. Using a forecast-ensemble based method with a convection permitting model where observations are assimilated, we, here, assess the likelihood of exceeding high precipitation rates (hereafter just called Likelihood) under the present climate conditions. In this study, we also added a new methodology within the toolbox of attribution's science for severe convective storms. Indeed, we are using an adapted Pseudo-Global warming approach (Schar et al., 1996) to investigate the likelihood of this event under pre-industrial period and different future warming levels. Such information is presented as event-based storylines or narratives, describing how a past impactful event may unfold under different levels of future warming, allowing users such as city managers and policy makers to explore changes in impact, and potential adaptation options.

SCIENTIFIC FRAMEWORK

Simulation set-up Due to a combination of multiple factors such as the high sensitivity of initial conditions and the difficulties to explicitly resolve convective mechanisms (Prein et al., 2015; Coppola et al., 2020; Prein et al., 2021), it is well known that simulating in a deterministic way the exact location and intensity of convective precipitation is one of the blind spots of climate and weather modelling (Bachmann et al., 2020; Hagelin et al., 2017; Schellander-Gorgas et al., 2017). This is especially true for climate modelling where the experimental set-up challenges even further the issue (Copolla et al. 2020). In this study, some attempts were made using the cycle 38 of HARMONIE-Climate model to reproduce the Copenhagen event using a typical climate set-up (i.e. using ERA5 as initial and lateral boundary conditions (IC, LBC); single nesting and no data assimilation cycle activated), which all fail to even reproduce medium precipitation rates. Therefore, we used the Danish operational HARMONIE-AROME limited-area numerical weather prediction model at 2.5 km grid mesh (Bengtsson et al., 2017) where data were assimilated for our study. To overcome the "initial conditions" challenge, we used a sophisticated ensemble approach called Scaled lagged average forecasting (Ebisuzaki et al., 1991) which used the forecasting error extracted from a control forecast to perturb members where initialization is at the same time as the control forecast. Figure 3.3-2 is showing that our approach was able to capture the main characteristics (small and intense precipitation systems) of the observed event (Figure 3.3-1) which peaks around 16 UTC.

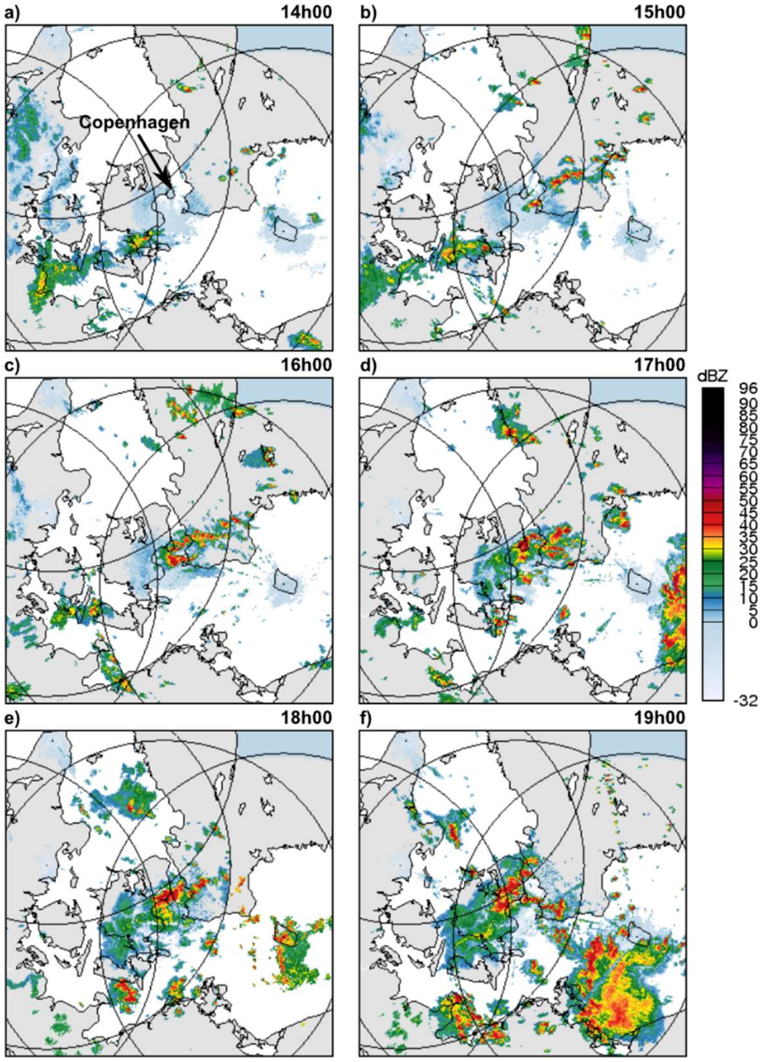


Figure 3.3-1 Temporal evolution of the radar reflectivity indicating the evolution of the precipitation pattern during the observed event.

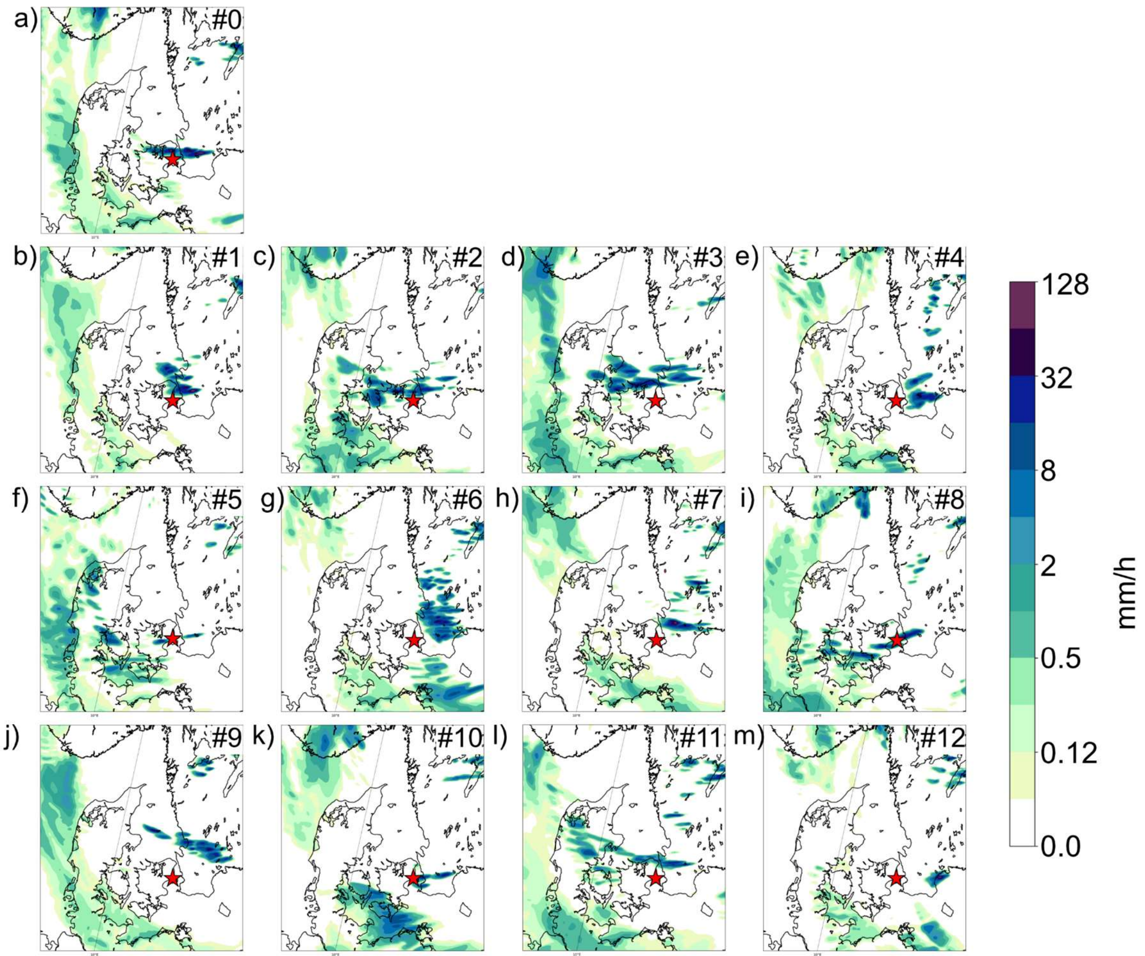


Figure 3.3-2 Hourly precipitation at 16 UTC for all members (a-m) for the control ensemble for July 2, 2011. The red star indicates the location of Copenhagen and the domain of interest is ~400 km on the west-east axis and ~600 km on the north-south axis.

Likelihood of exceeding high precipitation rate In order to take into account the model's incapacity to precisely simulate the location of the precipitation due to the inherent chaotic behaviour of the convective system, we have relaxed the spatial criteria. So, instead of simply computing the number of members that produced the selected precipitation rate, we also took into account the spatial neighbourhood of the specific grid point as discussed/proposed in a previous study (Ben Bouallègue, Z. & Theis, 2014). This is done by using a negative exponential law such as the one shown at the second term of Eq.2 where l is the distance around the selected grid-point (limited to a radius of 40 km around the selected grid point [i,j]) and L is the e-folding distance defined at 150 km. Furthermore, a hard precipitation threshold might also dampen artificially the risk for heavy precipitation, which is why if none of the grid points within the 40 km of radius has a precipitation rate of T mm/h (for example the value used for Fig.2 is 60mm/h since it was close to the observed rate) then the same exercise was done using a new threshold of T-2 mm/h and repeated for T-4 mm/h, T-6 mm/h, T-8 mm/h and T-10 mm/h defined as the varying threshold t in Eq.2 and T the e-folding threshold. Specifically, the Likelihood is computed as follow:

$$Likelihood[i,j] = \frac{1}{M}\sum_{m=0}^{M} l_m[i,j] \qquad \text{Equation 1}$$

Where Likelihood[i,j] is the so-called upscaled Likelihood located at i,j; $l_m[i,j]$ is the likelihood value per member (m) computed as:

$$l_m[i,j] = \begin{cases} 1, & \text{if } pr \geq T \text{ mm/h} \\ e^{-l/L}\, e^{-t/T}. & \text{if } pr \leq T \text{ mm/h} \end{cases}$$ Equation 2

Due to the highly chaotic aspect of such small-scale heavy precipitation systems, all ensembles, with no exception, could accidentally render a very high precipitation rate. This means that one system could exhibit a very high amount of precipitation within the same ensemble while not representing the overall behaviour and produce an inconsistent or noisier risk pattern (not shown). In order to overcome this issue, the risks have been computed using only 12 out of the 13 members, leaving out the member with the highest precipitation rate close to Copenhagen.

Pseudo-global warming set-up. In order to investigate how the Likelihood will evolve under different conditions, we applied an adapted pseudo-global warming approach (Schar et al., 1996). In a nutshell, cold (-1°C) and warm (+1°C, +2°C and +3°C) anomalies were added to the IC and LBC from the driving data and the specific humidity adjusted while making sure that the relative humidity remains unchanged from the reference set. This very specific set-up allowed us to investigate the risk of the event occurring under different warming levels including mimicking conditions from a pre-industrial period (e.g. using 1°C colder conditions). It is worth mentioning that the anomalies added should be considered as local temperature perturbations to the present day, not as global warming levels which may be different.

RESULTS

This specific set-up has allowed us to not only reproduce adequately the Likelihood of the observed event but also the one associated with different warming levels. The second row of Figure 3.3-3 is showing the evolution of the Likelihood using a threshold of 60 mm/h, under present day conditions. One can see that the Likelihood from 15 UTC increases for 16 UTC and then vanishes toward the end of the day, which is quite plausible with what has been observed on July 2 for this region. The first row of Figure 3.3-3, is showing the same event under colder conditions of -1°C which approximately represented the pre-industrial period (Masson-Delmotte et al., 2021). We can see that the cold conditions did not completely mitigate the event but reduced the rainfall intensity by around 50 %. This means that even under colder conditions, the likelihood of flooding would have been non-negligible, but the Likelihood would have been overall less than in present conditions.

On the other hand, when simulated using warmer conditions we see the opposite trend. That not only the overall Likelihood is increasing but seems to be a longer duration and more spatially spread. As expected by a Clausius-Clapeyron scaling argument (i.e. an increase of ~7 % of precipitation per degree of warming; Trenberth 2003), is also impacting the Likelihood at higher precipitation rate (see Matte et al., submitted to GRL) that was not existing in the control ensemble produced under present-day conditions. For example, rates higher than 90 mm/h have not been observed and simulated in the reference ensemble which occurred within the two warmest ensembles.

OUTCOMES

The outcome of this study is important as the methodology allows us to bring scientific insight on the climate change dependence of small-scale convective events, which until now have remained outside the scope of current attribution science. Such small-scale events are caused by a unique combination of several mechanisms (Otto and members of the Climate Science Communications Group, 2019) from local to synoptic scale that limits the possibility of finding a complete analogue in long, continuous simulations. Because of these reasons, studying these types of events has been a blind spot of typical attribution science until now. In order to respond to the public concerns linked to such a disaster, understanding the impact of climate change is a milestone in any adaptation planning.

The information produced by this study could be presented as event-based storylines, covering the event and impacts which occurred, and how the details of the event vary with different levels of warming. Attributing present-day events to observed climate change improves user engagement and utility due to social memory and data existing around experienced impacts, but also may help ensure physical consistency. In this study, the analysis took one more step than only attributing an event to the observed climate change by also analysing the impact of future local climate change upon the observed events (conditional on no changes to large-scale atmospheric circulation). The outcome of this analysis becomes a very strong dissemination tool for developing climate resilience and mitigation strategies.

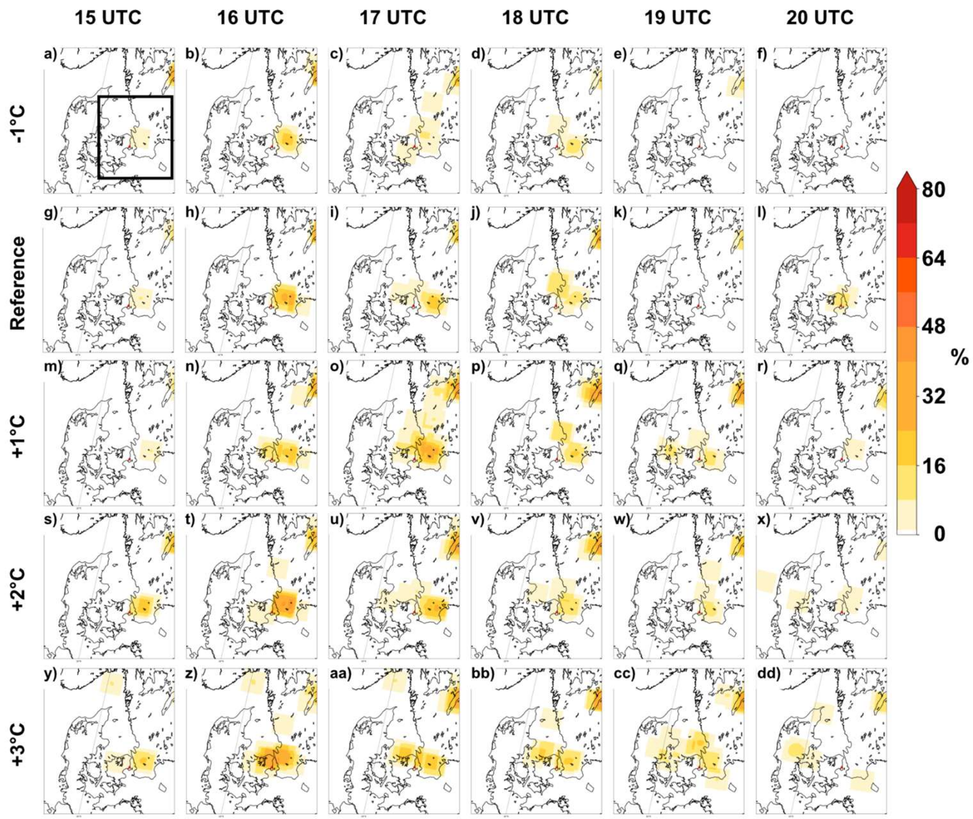


Figure 3.3-3 Risk using a precipitation threshold of 60 mm/h from 15 UTC to 20 UTC (columns) for all warming levels (a-f, m-r, s-x and y-dd for the -1 ◦ C, +1 ◦ C, +2 ◦ C and +3 ◦ C, respectively) and the reference case (g-l).

3.3.2 Event-based storylines for future drought events, using large ensembles (KNMI)

INTRODUCTION

Here an event analogues method presented in sections 3.7 and 3.8 of EUCP Deliverable 2.4, and the associated publications, is summarised as this adds to the discourse on storylines in this task and deliverable. As discussed in the previous section, event-based information can be presented as storylines, centred on a past event of interest, and there are indications of the utility of this for a range of applications that should be more widely explored.

Knowledge on the intensity and probability of future weather extremes is important to societies for the planning of adaptation policies. Such information is provided in climate projections, though research has shown that the abstract nature of general projections sometimes hinders the uptake by policy makers and the general public. Event-based storyline approaches can help to resolve some of this disconnect, in the way they relate climate projections to real-world events and human memories (Shepherd et al. 2018, Sillmann et al. 2019).

In the related EUCP subproject (WP2, T2.3), we developed a new method to transpose real-world extreme weather events into a warmer future. We use readily available large ensemble climate model simulations and take advantage of natural variability: if one generates enough simulations of a given climatic state, there will be simulated events that are comparable to the observed event. We use these simulated analogues to create event-based storylines that show how the observed event of interest might present itself in a warmer climate (Van der Wiel et al. 2021). The new method is complementary to so-called PGW-experiments, in which a regional climate model is run twice: once to reproduce an observed extreme event, and a second time in 'Pseudo-Global Warming' mode in which the background climatology is representative of a warmer future (Schrär et al. 1996). A PGW experiment for the vent in question was also performed for comparison. The PGW-approach has been shown to be very useful, and is well-appreciated by users of climate projection information (e.g. as shown by (use of) extreme rainfall case in KNMI climate scenarios 2014). If successful, we therefore expect the new method to be appreciated as well.

The new method is tested and showcased for the observed extreme drought of 2018 in western Europe. The summer of 2018 was extremely dry and warm (e.g. Philip et al. 2020, Zscheischler and Fischer, 2020), leading to widespread societal and natural impacts.

SCIENTIFIC FRAMEWORK

To find simulated analogues to the observed event of interest, we define a quantitative event metric. This metric is calculated for both the observed record (here ERA5) and the large ensemble data set (here simulations with EC-Earth v2.3, Van der Wiel et al. 2019), after which comparable events are picked from the distribution. We take a composite mean over a few simulated events (here 20 events) to improve the signal to noise ratio, i.e. limit the influence of natural variability on the estimate of the impact of climate change on the event. This composite mean is referred to as the 'simulated analogue'. To transpose the event to future warmer climates, we repeat this procedure for large ensemble datasets that reflect warmer climates (here we use large ensembles for present-day climate, and two future climates: pre-industrial+2C warming and pre-industrial+3C warming).

To verify if the simulated analogues reflect the observed event, we compared time series of different drought-related variables between the real event and the analogue. These variables were not part of the event-selection methodology here. If this is satisfactory, we trust that the selection method selects relevant events, and we compare the analogues across climatic states.

RESULTS

For the 2018 drought event, we take a time series of precipitation deficit starting from April 1 over the Rhine basin. We then compute the distributions of, for example, the mean value in the months August-October of this time series (Figure 3.3-4, panel a), and select the 20 most extreme simulated events. No bias-correction was performed on the data, but we analyse anomalies (departures from observed or model climatology) to remove climatological mean biases. Three out of these 20 events were more extreme than the observed event. The composite mean over these events is taken as our simulated analogue of the observed event. The selection based on this time series and metric was assessed to be satisfactorily as the analogue drought developed in a comparable manner to the observed event (e.g. in precipitation deficit, soil moisture availability, precipitation, temperature, radiation, and heat fluxes, Figure 3.3-4, panel b). The largest discrepancy between analogue and observed event is the fact that in 2018 the drought continued into November, whereas in the analogue recovery starts approximately mid-October.

In warmer climates the analogues develop a larger precipitation deficit and lower soil moisture, indicating future droughts like 2018 are more severe (Figure 3.3-4, panel c). This is caused by lower precipitation in the spring and summer, combined with higher temperatures and increasing atmospheric evaporative demand. These changes can be explained by a mean climatic change, though forced changes in variability add to the total change in drought severity.

The same drought event was also investigated by means of the more traditional PGW-experimental setup. The results are comparable, though of course within the limits of the two experimental designs. The PGW experiment reported in D2.4 follows the observed event more closely than the ensemble-based analogues, but the interpretation of the climate change response is somewhat limited (only mean-state changes, not drought-specific changes).

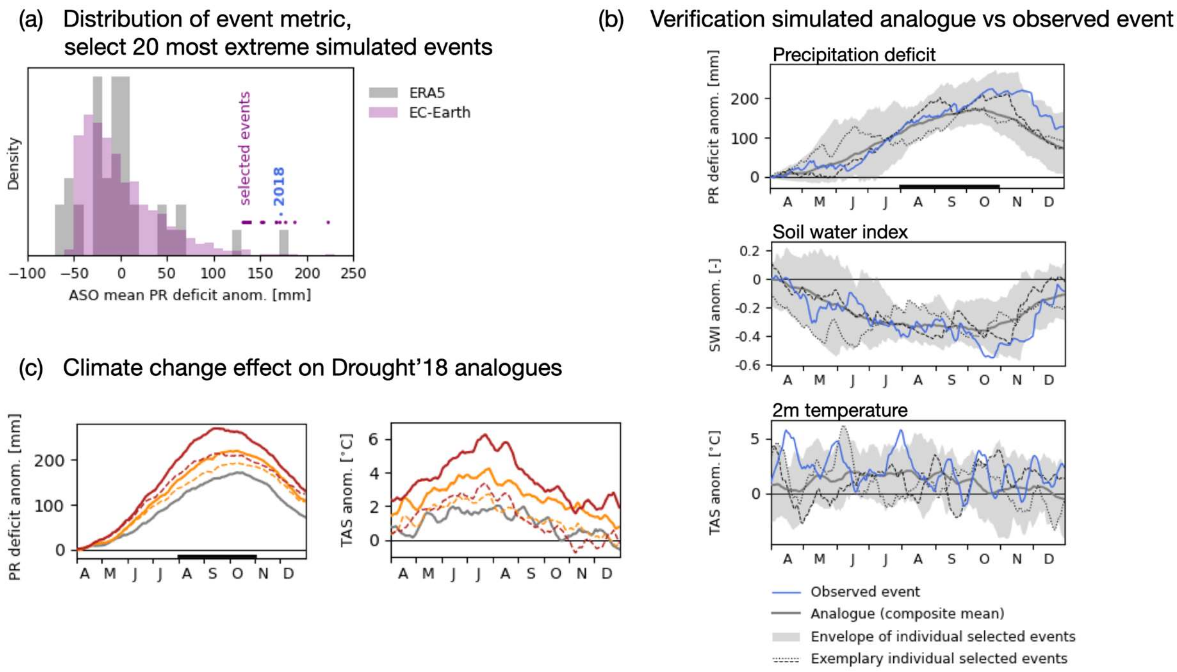


Figure 3.3-4 Schematic of steps in Drought'18 storyline development. (a) Distributions of event metric in grey for observed ERA5 data, in purple for simulated EC-Earth present-day ensemble. The blue dot and label shows the observed 2018 event, selected simulated events shown in purple dots. (b) Verification of simulated analogue (composite mean over 20 simulated events, solid grey dark line) against observed 2018 event (blue line). Envelope of individual simulated events in grey shading, examples of two example simulated events in dashed/dotted grey lines. (c) Effect of climate change on the Drought'18 analogues, grey lines show the present-day analogue, yellow and red lines respectively PI+2C and PI+3C warming. Solid lines show anomalies relative to the present-day climatology, dashed lines are anomalies relative to the future climatology (PI+2C, PI+3C; showing changes in variability on top of changes in mean climate). Figure after Van der Wiel et al. (2021).

OUTCOME

The main result of this subproject is the newly developed method for the creation of physical event-based storylines, that allow one to transpose an observed event into another climatic base state. This has mostly been an academic exercise, but KNMI has wide experience with PGW-experiments and users have been very interested in those outcomes. We therefore expect that this event-analogues method, shown to compare well with a PGW approach, would be met with interest from the climate information user community, alongside the more traditional probabilistic information on future climates. Further details of the comparison between the PGW and analogues-based methods, including assessment of the event drivers, can be found in EUCP D2.4 and the associated publications.

3.3.3 Application of metric-based sub-selection via clustering (SMHI, UKMO)

INTRODUCTION

This section reflects and discusses the clustering approach to regional ensemble member sub-selection presented in D2.4 section 3.5 (and references therein), aided by presenting additional plots in the climate variable space for the application case studies. This could be an appropriate means of selecting representative simulations as storylines for a given application, or part of the information generation process (as discussed in D2.4), sub-selecting to downscale or perform impact studies. Therefore, the method itself sits under the 'lines of evidence' part of the Venn diagram in Figure 3.1-1, and two case studies show example application of this, which could further be used to generate storylines as a user product. Proposed benefits of the approach (Wilcke and Bärring, 2016) include minimising information loss for a specific application, as there is strong sensitivity to the choice of variables/information input, but there remains a strong need for a careful and thoughtful selection process.

METHOD

This form of sub-selection based on clustering may be useful where analysing the full ensemble is intractable, such as for impact studies or further downscaling, and aims to create a sub-selection which still captures the range of behaviour in the projections. This is done in a multi-variate, multi-regional space as most applications necessitated the sub-selection captures the behaviour of a diverse range of variables across regions, and the algorithmic approach chosen Is one way of achieving this. The approach uses singular vector decomposition of a matrix formed from the variables of interest, and clustering is performed across regions, variables, seasons and timescales, as required by the application. The closeness of the models in this parameter space is used to perform the clustering, using no physical relationship, just statistical measures regarding variability within the ensemble. The cluster means (in the entire parameter space) are then used to make the final model selection, although there are also other possibilities, such as representing the cluster range. Further details of the method are given in EUCP deliverable D2.4, and Wilcke and Bärring (2016).

CASE STUDY RESULTS

In case study 1 in D2.4, regional climate model (RCM) selection for hydrological modelling is described. The trial input indices were climate change signals (2035-2060 minus 1980-2014) for precipitation over Northern, Central and Southern Europe, annually and seasonally, under RCP8.5. The projection data used were 32 RCMs from EURO-CORDEX (Jacob et al. 2014). The approach resulted in 7 clusters and selected models. Figure 3.3-7 shows the clusters and selected simulations for Northern and Southern Europe with DJF Pr anomalies plotted against JJA Pr anomalies in arbitrary units (the corresponding dendrogram included in D2.4 is shown in Figure 3.3-5). We see that in some cases clusters are single valued (only include one model), and in the case of the larger clusters (blue/red) the representatives labelled do not lay at the centre of the included models for the variable space plotted, due to the high-dimensional space of multiple variables being included in the clustering and mean selection. The selected members almost entirely cover the uncertainty space, indicating that representing clusters by more than the closest model to the mean, such as the cluster ranges, may not be beneficial in this case. One exception is that the drier end of DJF Pr for Northern Europe is not represented (top left panel), which may have relevance to hydrological applications, as this is when resource recharge occurs. This illustrates how using this algorithmic approach to sub-selection may be augmented by considering other lines of evidence, and a manual selection based on expert elicitation and the application at hand (the intersection of lines of evidence and storylines for application in Figure 3.1-1).

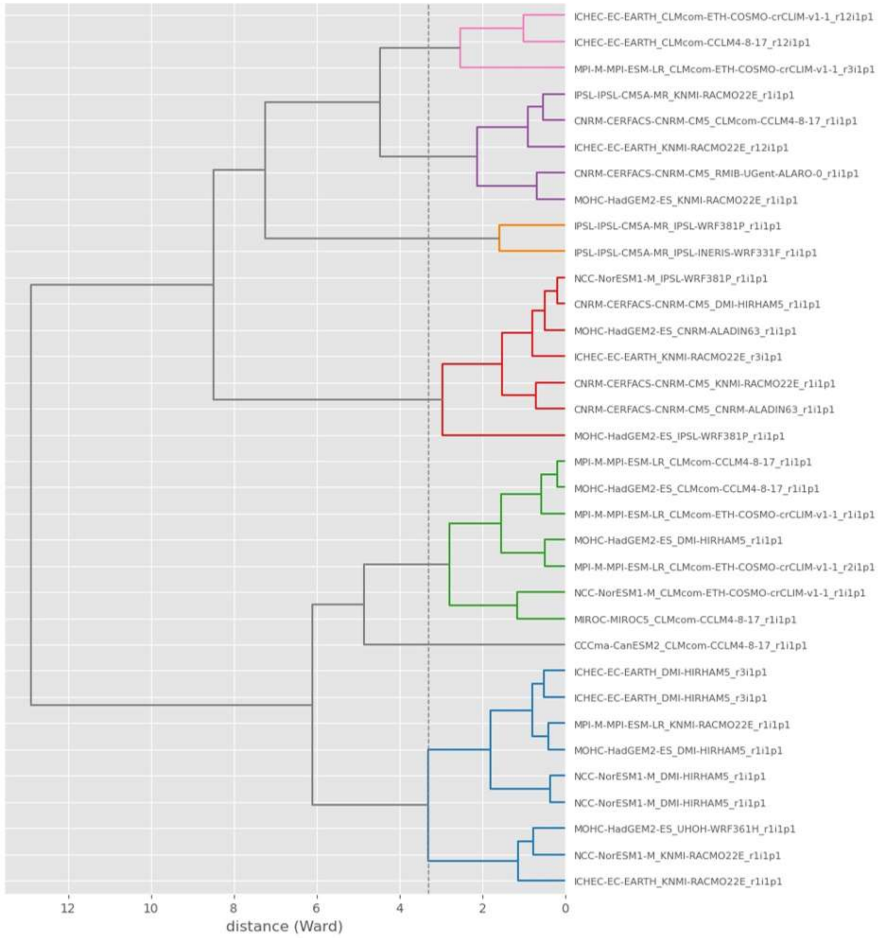


Figure 3.3-6 Tree-diagram (dendrogram) for case study 1 showing the colour-coded clusters and model IDs, the closeness of relation (no physical relation, just statistical relation regarding the variability within the ensemble) is indicated by the Ward distance on the x-axis.

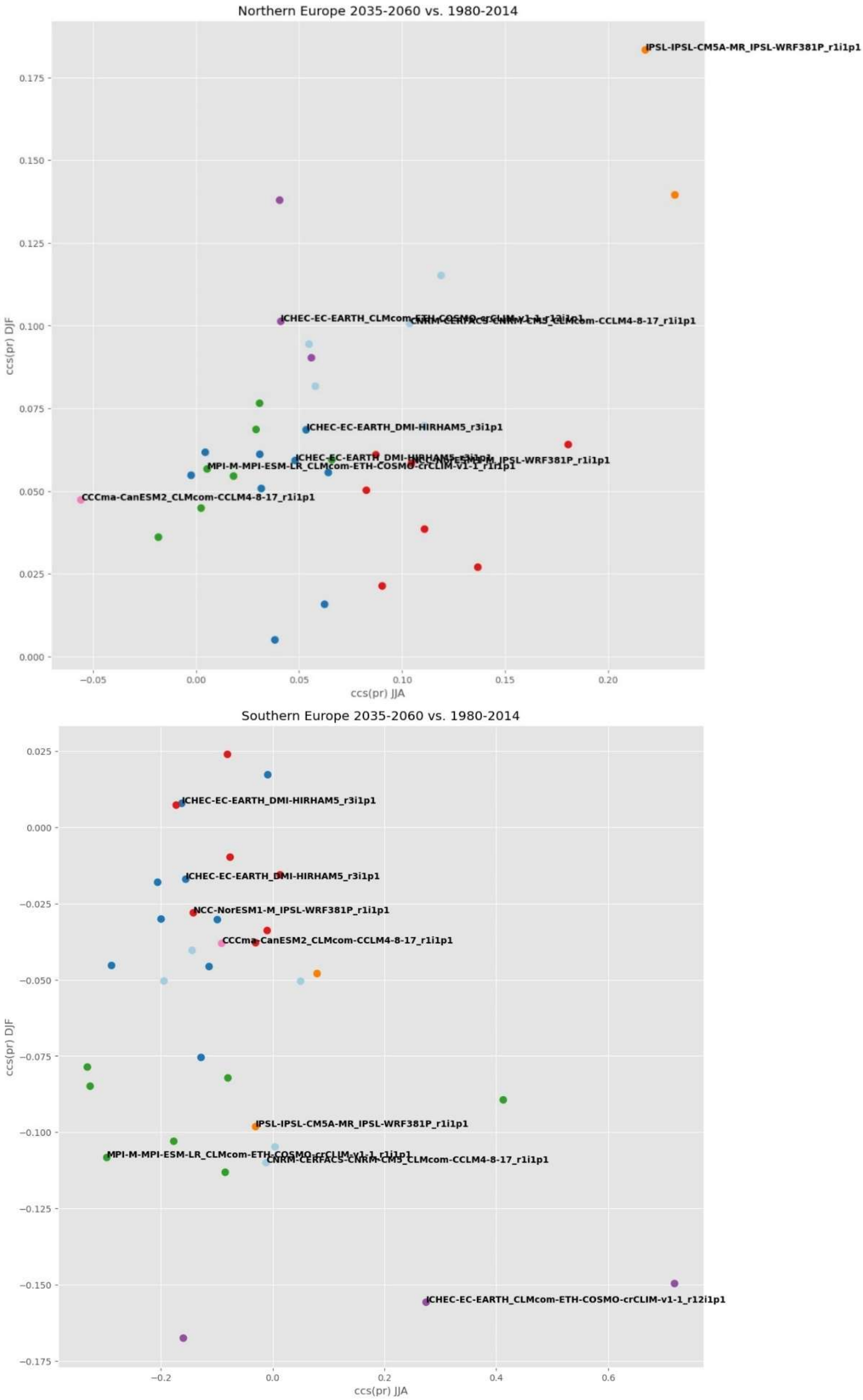


Figure 3.3-7 Colour-coded (as Figure 3.3-6) clusters in the climate variable space for case study 1, with DJF Pr fractional percentage anomalies (vertical axis) plotted against JJA Pr (horizontal axis) for Northern and Southern Europe, and the chosen member closest to the cluster centre marked by the model ID.

The second clustering case study aimed to reduce a large single model ensemble (SLENS, Wyser et al. 2021) for high resolution downscaling over Sweden. In particular the focus was to pre-select members which may trigger different types of convective precipitation events, covering the range of relevant uncertainty. The parameters included in the SVD analysis were the seasonal precipitation and temperature anomalies over northern and southern Sweden, for 2040-2070 minus 1980-2010. The approach resulted in 12 clusters. Figure 3.3-8 shows the dendrogram with coloured clusters (left), and the clusters and selected members for Pr and surface air temperature (tas) over southern Sweden for DJF and JJA (right). Again, the selected members cover a large region of the uncertainty space, apart from the lowest end of summer rainfall and temperature change for southern Sweden. If the downscaled simulations would be used for multiple applications, this indicates a situation where there may be value in applying expert judgement to override some of the cluster-based selection to ensure this is part of the uncertainty space is represented, or by increasing the weighting of summer rainfall in the clustering.

OUTCOME

These case studies show the potential utility of multivariate methods of clustering and sub-selection applied to projection ensembles. In particular, the clustering approach offers the possibility of reducing the number of members to analyse by selecting representative members which are coherent across variables, seasons, and regions, while minimising the loss of relevant information. These could then be used as the basis of storylines as a user product, or as part of the data production process. However, the lack of a physical basis for the clusters does limit the potential explanatory power and understanding of the drivers of uncertainty behind the sub-selection, however utilising other lines of evidence and analysis could fill this gap where required. Other potentially complementary methods of sub-selection and ensemble reduction were discussed in WP2 D2.3 & D2.4 (uncertainty quantification, atmospheric driver analysis and model weighting), as well as WP5 D5.3 (constraints).

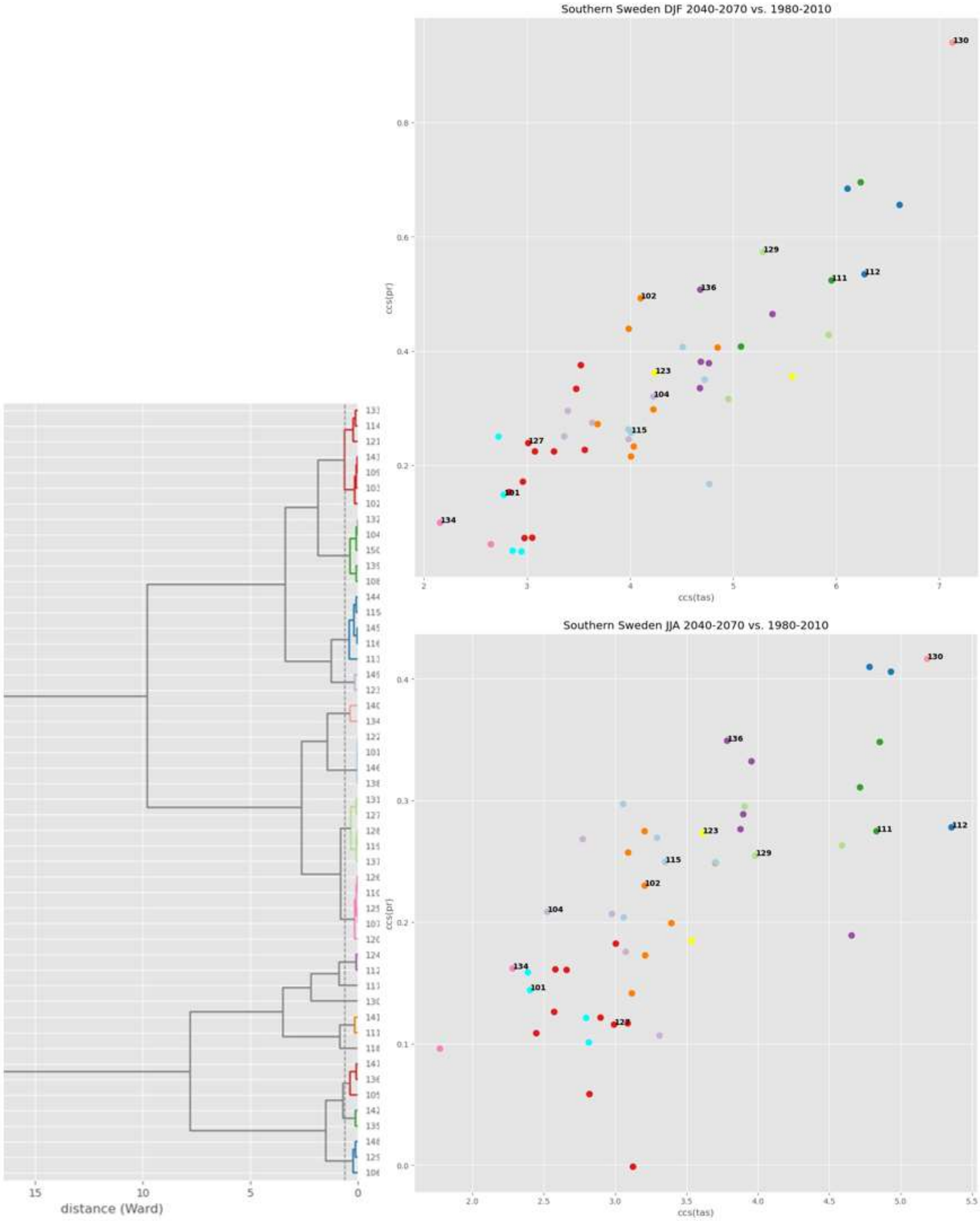


Figure 3.3-8 Left: Tree-diagram for case study 2 showing the clusters and model IDs, the closeness of relation is indicated by the Ward distance on the x-axis. Right: clusters in the climate variable space for Southern Sweden, with Tas anomaly in degrees Celsius vs Pr anomaly as a fractional percentage for DJF and JJA in arbitrary units, the clusters are colour coded, and the chosen representative member is marked by the model ID.

3.3.4 NAO analysis on multiple timescales and application prospects (UEDIN)

Other sections of this deliverable show the importance of understanding atmospheric processes and how they are represented in climate models, whether this is for the purpose of forming climate driver-led storylines, understanding the robustness and limits of prediction and projection information, or understanding and producing event sets. This becomes increasingly important as users and the general public become more aware of the climate processes which lead to impactful weather events. This section presents work which builds on some analysis of the North Atlantic Oscillation presented in EUCP deliverable 5.2, and in Section 3.3.5, looks at the representation of internal variability in projections, and the resulting impact on uncertainty estimation and quantification.

The North Atlantic Oscillation undergoes strong variability on all timescales, which can influence climate trends in the past and future (Deser et al., 2015; Iles and Hegerl, 2014), which can enhance or mask the effect of greenhouse gas induced changes in past and future. However, the trends in the NAO simulated in CMIP class models appear to show insufficient NAO variability on decadal timescales (O'Reilly et al., 2021; Schurer et al., in prep.). Since the NAO leads to increased precipitation in Scandinavia and reduced precipitation in the Mediterranean, particularly in winter, the low frequency variability could have an appreciable impact on rainfall trends. Thus, users may experience periods with a stronger increase in precipitation than indicated by CMIP type models, along with periods of counteracting variability, and they should prepare for such possibilities.

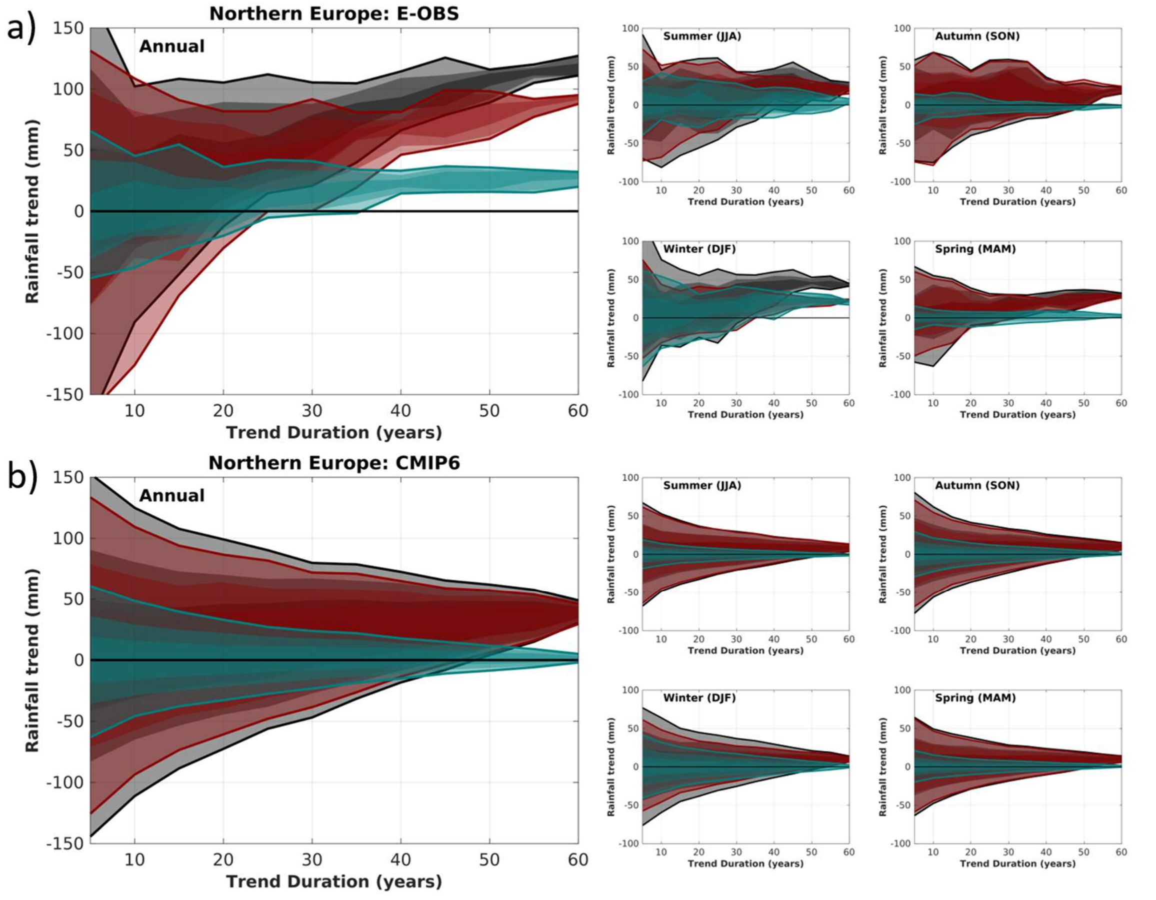


Figure 3.3-9 Distributions of the linear trends (of 5-,10-,…,60-yr duration, shown along the x-axis) in northern European annual (left panels) and seasonal (right panels) rainfall, sampled from: a) E-OBS v19e observations, and b) CMIP6 historical simulations (41 models, 163 ensemble members). Grey shading indicates the trends in the raw time-series; turquoise shading indicates the trends in the component of rainfall associated with the NAO; red shading indicates the trends of the time-series with effect of the NAO removed. The lightest shading spans the minimum to maximum trends, and the darker levels of shading indicates the 10th-90th and 25th-75th percentile ranges of sampled trends. Trend distributions are randomly sampled over the period 1950-2014, and the CMIP6 panels display the multi-model mean of ensemble means. Units are given as total accumulated (annual or seasonal) rainfall (in mm) per trend duration (in years, along the x-axis). From Ballinger et al., in prep.

Figure 3.3-9 illustrates the strong observed decadal variability in annual trends in Northern European precipitation. The black / grey range indicates the range of trends of annual rainfall in Northern Europe over the historical period (1950 to 2014) and shows that the observed trends over timescales of 50-60 years are not within the range simulated by the CMIP6 historical simulations considered here (Eyring et al., 2016).

The turquoise range in Figure 3.3-9 indicates the component of the decadal trends in rainfall that are associated purely with the evolution of the NAO over the same time periods. Here the NAO is characterised by the first EOF of the Atlantic sector SLP (see Ballinger et al., 2022 for detail), calculated for observations and each model simulation individually. The NAO effect on precipitation is then calculated by regressing the NAO index on averages of Northern European precipitation for each month of the year separately (with stronger influences in winter months and weaker, different relationships in summer), and showing the trend associated with that regression coefficient in turquoise in Figure 3.3-9. Results show a tendency for the observed NAO having caused a strong positive contribution for multi-decadal rainfall trends that is a lot stronger than the range shown in CMIP6 models which favours only slightly positive trends in a narrow range for multi-decadal trends. If this NAO-induced precipitation change is removed from the E-OBS rainfall data, the red ranges indicate a better agreement between climate model simulated and observed NAO trends, with the remaining discrepancy being consistent with low frequency variability.

This discrepancy is further illustrated in Figure 3.3-10, where rainfall changes (%) are plotted against temperature changes for raw EOBs data (circles), and compare the constrained range based on attribution without removing the NAO (red) and after removing the NAO (turquoise). This again indicates that the NAO trend has caused a substantial fraction of multidecadal winter warming and wetting of Northern Europe, which is far outside the range of individual climate model simulations (small squares). After subtracting the NAO's effects, the uncertainty range around the attributed signal spans many of the individual model simulated ranges, indicating good hindcases of the combined change in temperature and precipitation in Northern Europe.

It is presently unclear what caused this low-frequency NAO variability, which is well measured and consistent across observed datasets. If this variability is (partly) forced by unique features of the 20th century, such as aerosol forcing, future trends may be closer to the model range, yet presently this is not supported by analyses (e.g., Undorf et al., 2018). If it is due to internal variability, at least to a large fraction, then this discrepancy between observed and simulated decadal NAO variability needs to be considered when predicting future changes. Storylines for future rainfall changes in the winter season, particularly in Northern Europe and to some extent also in the Mediterranean region (see Ballinger et al., 2022), need to account for the possibility of long-term NAO changes enhancing or masking the human climate change signal. This can be done, for example, by adding low-frequency observed NAO variability on the NAO-removed change in model simulations. These findings could also form part of a lines of evidence assessment, assessing the robustness and plausibility of a given set of prediction or projection information.

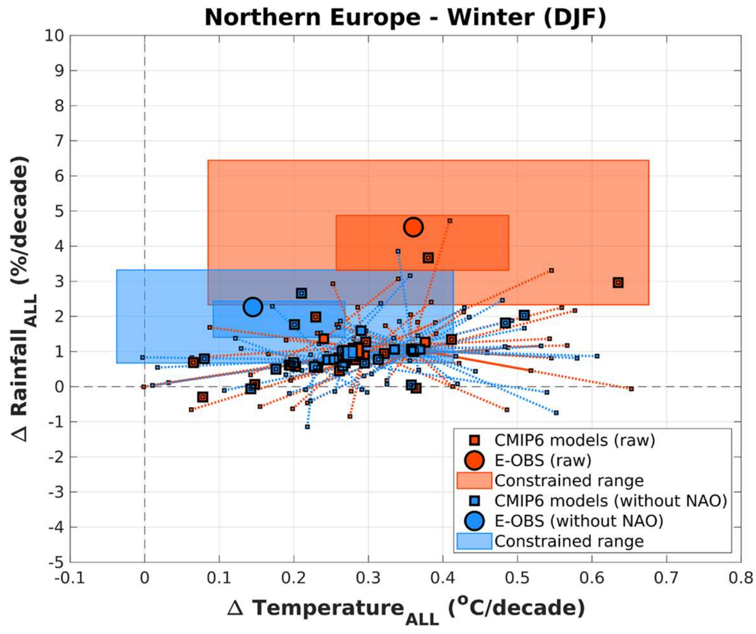


Figure 3.3-10 The change in mean temperature (°C decade-1) and rainfall (% decade-1) from 1950-1969 to 1995-2014, for winter (DJF) computed over Northern Europe. Square markers show the CMIP6 models: small – individual ensemble members (n=66); medium – model ensemble means (n=24); large – multi-model mean. Individual ensemble members have been linked to the associated model ensemble mean via dashed lines (for those models with more than one ensemble member). Circle markers show the observations (E-OBS, v19). The shaded regions indicate the range of the estimated constraint (ALL) applied to the CMIP6 multi-model mean, displaying the 25th-75th percentile (inner shaded region) and 5th-95th percentile (outer shaded region). The orange colour shows the differences computed from the raw models and observations, while blue shows the differences computed from the time series of temperature and rainfall after first regressing out the influence of the NAO.

3.3.5 Projections of northern hemisphere extratropical climate underestimate internal variability and associated uncertainty (UOXF, UEDIN)

The internal variability of the large-scale atmospheric circulation exhibits a dominant influence on the uncertainty of the continental climate on decadal-to-multidecadal timescales. For example, the extratropical warming over land during the Northern Hemisphere winter over the later part of the twentieth century was enhanced substantially by anomalies in the large-scale atmospheric circulation and their associated impact on surface-air temperature (e.g. Wallace et al., 2012). However, coupled climate models have been shown to show too little variability on decadal timescales (e.g. Bracegirdle et al., 2018). The aim of this study was to characterise the large-scale circulation variability in coupled models and produce observationally constrained projections of extratropical climate that include more realistic estimates of internal atmospheric circulation variability. To investigate the influence of the underestimation of large-scale circulation variability on decadal timescales, we generated synthetic temperature and precipitation projections that are consistent with the observed large-scale circulation variability. The methodology is outlined here but discussed in more detail in O'Reilly et al. (2021).

Methods

First, the signature of SLP variability was subtracted from temperature and precipitation fields in the raw 99-member MPI-GE ensemble (MPI-GE-raw hereafter) using linear regression. Only the first three EOFs of SLP were used, as these were found to make the dominant contributions to the multidecadal SLP variability. A random member is selected from MPI-GE-raw and the temperature/precipitation variability associated with the first three EOFs of SLP are removed. This variability is then replaced with the same patterns of temperature/precipitation anomalies but multiplied by a random surrogate PC time series that is constructed to have the same spectral characteristics as the observations. The surrogate PC time series, therefore, tends to have more power on multidecadal timescales, though the overall standard deviation is unchanged on average, across the ensemble. This process was repeated 10,000 times to produce a synthetic observationally constrained 10,000-member ensemble (MPI-GE-obs hereafter). To include the influence of observational uncertainty, the surrogate PC time series were calculated from four different observational data sets, with each contributing equally to produce the 10,000 members in MPI-GE-obs. In the analysis that follows, we compare the raw 99-member ensemble, MPI-GE-raw, with the synthetic 10,000-member observationally constrained ensemble, MPI-GE-obs.

RESULTS

To analyse the influence the observationally constrained large-scale circulation variability on future climate projections, we examine the changes in surface-air temperature and precipitation for the mid-century period (2041–2060) with respect to a present-day baseline period (1995–2014) in MPI-GE-raw and MPI-GE-obs. Distributions of the projected regional change of temperature and precipitation for the winter and summer seasons are summarised in Figure 3.3-11. The median projected changes under the Representative Concentration Pathway (RCP) 4.5 scenario for the boreal winter season (panel a) consist of widespread warming across the extratropical regions and are very similar in MPI-GE-raw and MPI-GE-obs in most regions, demonstrating that changes in the large-scale atmospheric circulation are not responsible for the distribution of average temperature changes. There are, however, substantial differences in interquartile range (i.e. 25–75%) of the projected changes, with MPI-GE-obs exhibiting a larger range over most regions in the Northern extratropics. The difference in the interquartile range in MPI-GE-obs increases by over 50% compared to MPI-GE-raw in regions within Northern Europe, North America and the Mediterranean, almost doubling in some areas. For winter precipitation, the MPI-GE-obs shows substantial increases in the interquartile range compared with MPI-GE-raw, most notably over Northern Europe and Mediterranean regions (panel b). For the Mediterranean region, there is more than a doubling of the likely range, with more substantial drying becoming much more likely in MPI-GE-obs. The broadening of the distributions is also clear in the tails of the distributions, where mid-century changes in the winter temperature and precipitation that would be deemed highly unlikely are now well within the range of likely outcomes in the presence of observationally constrained large-scale circulation variability (i.e. MPI-GE-obs).

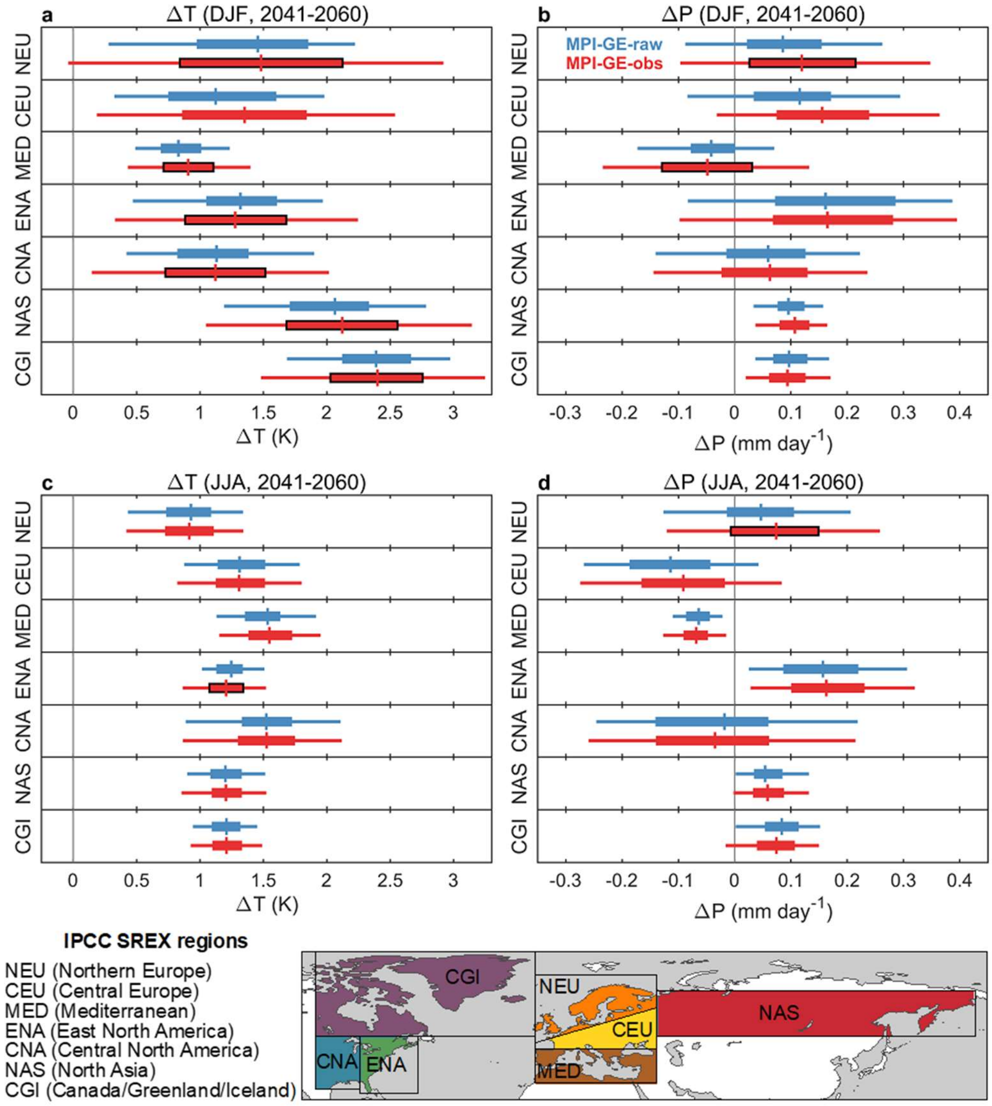


Figure 3.3-11 Distributions of projected regional changes for the 2041–2060 mean from a 1995–2014 baseline period for: a Winter surface-air temperature, b winter precipitation, c summer surface-air temperature and d summer precipitation. The thin horizontal lines show the 5–95% range, the rectangular boxes show the interquartile range (i.e. 25–75%) and the vertical lines show the median change. Distributions are shown for the MPI-GE-raw ensemble (in blue) and the MPI-GE-obs ensemble (in red). The black boxes show where the interquartile range of the 2041–2060 projection in the MPI-GE-obs ensemble is significantly different from the respective interquartile range in the MPI-GE-raw ensemble (at the 95% level, based on a Monte Carlo resampling.

The differences between the projections for changes in summer climate are relatively muted compared to the winter season. The reason for this is that the SLP EOFs exhibit a stronger relationship with temperature and precipitation anomalies in the summer season compared to the winter season. Nonetheless, there are still significant increases in the interquartile range of the projected summertime precipitation changes over Northern Europe. There are also significant changes in the distribution of projected temperatures over the East North America region.

We also examined the influence of the observational constraint on future extreme season occurrence. Here we define an extreme season as the highest or lowest seasonal mean value over the baseline climate period, 1995–2014, representing a 1/20 year event based on a present-day climate period, which one could also estimate in the observational record. The number of extreme seasons in a future climate period is then calculated in each ensemble member. An example of the occurrence rate of extreme seasons for Northern European winters are shown in Figure 3.3-12 for the mid-century period, 2041–2060. The occurrence of a greater number of extreme seasons within the 20-year window is larger in MPI-GE-obs than in MPI-GE-raw in a number of instances, particularly in the tails of the distribution, whereas the occurrence of relatively few events tends to be higher in MPI-GE-raw. In the Northern Europe regions shown here there is a >10% probability of exceeding 8 seasons with extreme high temperatures in the MPI-GE-obs data set, whereas the probability of an equally extreme realisation occurring in MPI-GE-raw is about 1% (i.e. panel a). Similarly, increases in the numbers of extremely wet winters over the mid-century period are found to be significantly more likely in MPI-GE-obs (i.e. panel b). Another notable feature is that the occurrence of having a number of extremely dry Mediterranean winters over the mid-century period in the future is significantly higher in MPI-GE-obs. In the summer and the other extratropical regions, there are fewer clear differences in the occurrence of extreme seasons.

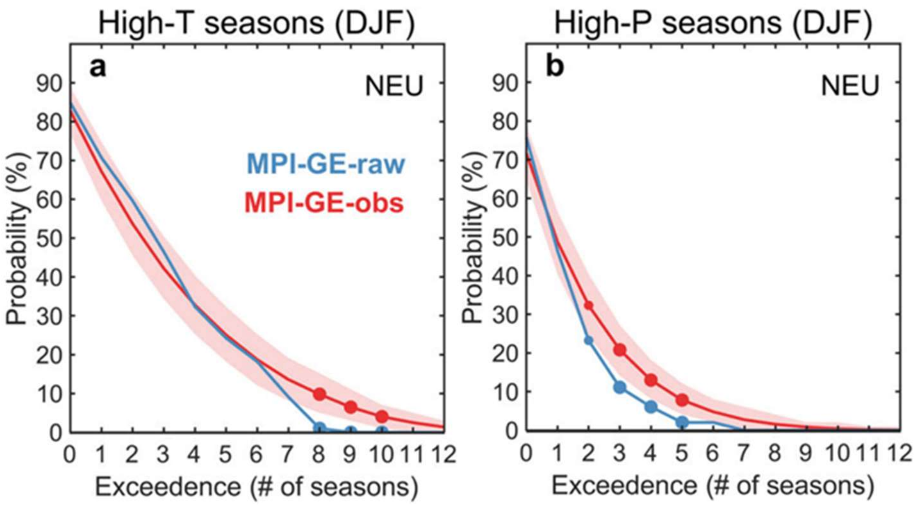


Figure 3.3-12 The probability of exceeding a number of extreme seasons over the period 2041–2060, where extreme seasons are defined as the highest seasonal mean value over the baseline period 1995–2014. The panels show probability of exceeding a number of extreme seasons in the Northern Europe region for a high temperatures and b high precipitation. Blue lines show the probability of exceeding a given number of extreme seasons in the 99-member MPI-GE-raw ensemble, while the red lines show the same for the MPI-GE-obs ensemble. The shaded regions around the MPI-GE-obs ensemble shows the 5–95% range of the MPI-GE-obs when only 99-members are resampled at random (10,000 times). Large dots show where the probabilities in the MPI-GE-raw and MPI-GE-obs ensembles are significantly different at the 95% level and the small dots show significance at the 90% level (based on a Monte Carlo resampling). (From O'Reilly et al., 2021)

The higher probability of a large number of extreme winter seasons occurring in a future period is related to the relatively large variability on multidecadal timescales in the MPI-GE-obs, which is absent in MPI-GE-raw. An explanation for this is that the influence of low-frequency variability in the large-scale circulation can set a relatively high background anomaly over a 20-year period, meaning that the year-to-year variability superimposed onto this can produce clusters of extreme seasons. In MPI-GE-raw, however, there is relatively little low-frequency variability so the occurrence of future extreme seasons in a given year is largely independent of the surrounding years.

SUMMARY

The analysis presented here demonstrates that factoring the influence of the observed variability of the large-scale atmospheric circulation into future climate projections substantially increases the uncertainty arising from internal variability. The current generation of coupled climate models, which are used to produce future climate projections, are therefore likely to underestimate the contribution of internal variability in the extratropics. There are some significant differences in the projections of the MPI-GE-obs and MPI-GE-raw ensembles in the summer season around the North Atlantic sector but the influence of the observed large-scale atmospheric circulation on future projections is largest during the winter season, influencing most regions in the Northern Extratropics. For future twenty-first century periods, the underestimation of the uncertainty due to large-scale atmospheric circulation is comparable with the structural uncertainty in the forced response. An example of where this underestimation could be important is the recent literature considering the differing impacts of 1.5 and 2 °C of global warming; the underestimation of internal variability in the extratropics implies that regional differences between 1.5 and 2 °C warming are likely to be somewhat overconfident. Furthermore, the increased uncertainty also raises questions about the treatment of internal variability in regional model projections. The EURO-CORDEX ensemble, for example, uses a relatively small subset of global coupled climate model simulations that, as has shown here, themselves underestimate the contribution of internal variability and this will be compounded in projections made using regional model ensembles.

The increased projection uncertainty may be important to factor in producing information for future risk assessment and decision-making exercises, such as storylines exploring hazard or impact metrics for a user application. This could include selecting models runs with enhanced variability, that may be more likely than previously thought, or applying a similar procedure as that described above to create the synthetic results to bias correct, or boost, variability before performing further analysis or impact modelling. However, this would require careful treatment to ensure physical consistency and plausibility. These points may have relevance to performing lines of evidence assessments, as well as sub-selection of members for data production or creating storylines, and assessing the robustness of the included information.

3.4 Regional projections for Europe from Multiple Lines of Evidence (UKMO)

3.4.1 Introduction / Motivation

In sections 3.2 and 3.3 we have explored a number of components of storylines which might be used to explore ‘snapshots’ of plausible regional future climate or climate events. Understanding the wider uncertainty context of regional climate change is key to both the design or selection of storylines, as well as their application and interpretation, as visualised in the Venn diagram in Figure 3.1-1.

Here we explore how that uncertainty context might differ depending on which projection product we are using, and what we can learn about the wider uncertainty context and the robustness of projected uncertainty ranges by treating the products as multiple ‘lines of evidence’. EUCP has made available a number of new projection products for Europe. These new products include multi-convection-permitting regional climate model (CPM, or CP-RCM) projections of future climate for a set of European domains (WP3), offering very high-resolution projections with significant benefits for the representation of realistic weather features such as extreme rainfall. EUCP has also provided new products that constrain regional projection ranges based on observations and model dependencies to offer ‘added value’ to raw projection data (WP2). These new projection products clearly offer new opportunities for a wide range of users. However, they also add to an increasingly complex data landscape for Europe by adding to the existing array of projection products which will each offer the user a different future projection range. Comparison of the CMIP5 and CMIP6 global projections for European regions already indicates differences in their regional projection ranges (Palmer et al. 2021). Downscaled projections via EURO-CORDEX offer higher resolution information driven by CMIP5 global models, and will offer different projection ranges from that of the parent CMIP5 ensemble for ‘physical’ reasons as a result of higher model resolution, but also because of the incomplete sampling of the driving model uncertainties. This sampling issue is particularly relevant to the new CPM projections because the number of members is particularly small. Additionally, other projection ranges are available via the global and downscaled regional projections from a perturbed-parameter ensemble (PPE) as part of the UKCP simulations (Murphy et al. 2019).

This work aims to bring together these datasets amongst other ‘lines of evidence’ (Figure 3.4-1) to explore the wider uncertainty context of regional climate projections across multiple products. Firstly, comparing the different datasets will expose where there are important differences in projection range, or magnitude of projected changes between projection products, which could lead to different choices of storylines and differences in estimated impacts of regional climate change. By showing how multiple products compare, we can offer a means of understanding the wider context of any one dataset, scenario or storyline (e.g. by showing the wider uncertainty context for the small ensemble of high-resolution CPM simulations). Secondly, we look for evidence that explains why projections may diverge or agree; in the first instance by showing the impact of sampling the parent GCM vs the physical differences impact of downscaling, and subsequently drawing on other lines of evidence related to our physical understanding of the different projection sets such as how they resolve relevant key processes. Finally, we consider how this information can be usefully communicated to users. The intended outcome of this process is to provide a more robust basis for estimating risk and a greater confidence base to the users. Providing this broader information on the evidence base (types of sources, consistency, quality) and the level of agreement provides the important context that helps inform how and where this information can be used, including in the selection or production of climate storylines.

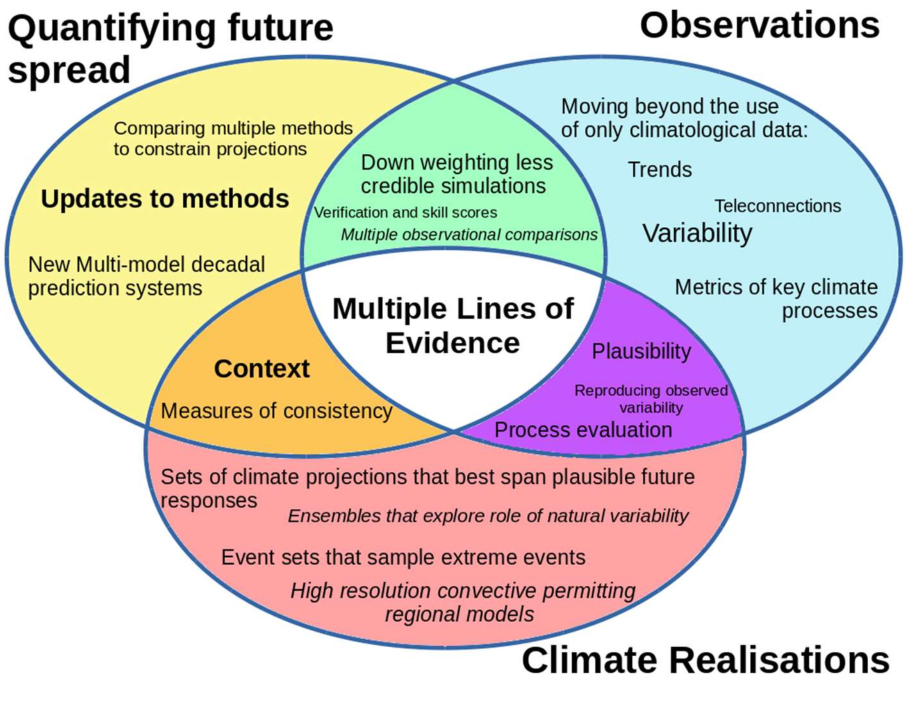


Figure 3.4-1: Venn diagram illustrating how the Multiple Lines of evidence drawn from EUCP's WP strands will combine to contribute to the overall picture of confidence in the projection data.

3.4.2 Building summary plots across multiple projection datasets

The projection datasets that we draw on can be broadly categorised into 3 types:

- Data from global climate models (GCMs)
- Results from different methods of constraining or weighting climate model projections
- Data from high resolution downscaled climate model projections (RCMs and CP-RCMs)

Figure 3.4-2 and Figure 3.4-3 demonstrate the presentation of the quantitative comparison of these methods for selected case study examples. A full set of similar figures will be available for each of mean temperature and precipitation changes at https://zenodo.org/record/6046762, for regions defined by the EUCP WP3 CP-RCM domains (as detailed in Appendix 6.2.2).

The analysis in this deliverable is based primarily on anomalies of the time period 2041-2060 with respect to 1995-2014, although with some exceptions, mainly for the CPM models. A summary of the time periods used, including details of these exceptions are given in Appendix 6.2.2. Each dataset is linearly regridded onto a common 1.5° x 1.5° regular latitude and longitude grid and seasonal and area means are calculated before then computing the anomaly of the historical period to the mid-century period. In addition, spatial maps (not shown) to aid in understanding of spatial variation of multi model ensemble means, and also individual models, have been produced of the seasonal mean anomalies and are also made available at https://zenodo.org/record/6046762/

REGIONAL PROJECTIONS FROM GLOBAL MODELS

CMIP5 and CMIP6 global multi-model ensembles underpin the majority of global and regional assessments of climate change, either through direct use or via downscaling. Importantly, the multi-model uncertainty ranges inferred from the ensembles inform international IPCC assessment reports. Other global model ensembles also contribute important information about parts of the total uncertainty range; the perturbed parameter ensemble (PPE) method explores a different source of modelling uncertainty through a range of possible settings within one model, and large single model ensembles (LSMEs) provide information about natural variability by varying the initial conditions. We include in this comparison the CMIP5, CMIP6 multi-model ensembles and UKCP-Global PPE, but we note that a number of LSMEs could usefully be added as additional lines of evidence from global models.

Differences between European regional projection ranges from global models in CMIP5, CMIP6 and UKCP-Global can be partially attributed to differences in forcings (Table 2) and different global responses due to climate sensitivity. It has been well documented that CMIP6 includes a larger number of 'high sensitivity' models compared with CMIP5 (e.g. Zelinka et al. 2020) and which lead to a higher global temperature responses. However, a number of modelling groups have also highlighted that these high temperature responses are partly related to difference in the forcing applied to the two experiments which mean that ssp585 has a higher effective radiative forcing (ERF) than RCP85 (Tebaldi et al., 2021)). Wyser et al. (2020), for example, find that around 50% of the difference in temperature response by 2100 in equivalent CMIP5 and CMIP6 experiments for one GCM can be attributed to the differences in forcings between RCP85 and ssp585. The implications of these global scale differences for Europe are explored by Palmer et al. (2021) who find that that the warmer ranges projected by

CMIP6 for European regions are largely due to the higher global temperature changes in CMIP6 (a combination of higher sensitivities and forcing differences between RCP85 and ssp585), rather than differences in regional response. The exception to this is the central European region (CEU), for which the study found differences between CMIP5 and CMIP6 projections for CEU summer rainfall arise due to a combination of both the global scale differences and a change in regional response. Additional information is available from the CMIP6 HiResMIP experiment which compares standard and high resolution versions of global models in consistent formulations (Haarsma et al. 2016).

Ensemble	Forcing Scenario	Ensemble type	Members	Reference
CMIP6	ssp585 (concentration driven)	MME	34	Eyring et al. 2016
CMIP5	RCP8.5 (concentration-driven)	MME	35	Taylor et al. 2012
UKCP-Global	Varying emissions pathways consistent with RCP8.5 sampling carbon cycle uncertainties	PPE	15	Murphy et al. 2019
HighResMIP	ssp585 (concentration driven)	MME	6 “high” resolution. 6 “standard” resolution	(Haarsma et al. 2016)

Table 2: Data from global models used in EUCP lines of evidence comparison. The full list of models used from the CMIP5 and CMIP6 experiments is detailed in appendix 6.2.2.

The perturbed parameter models in UKCP-Global models tend to have larger global climate responses due to a combination of two factors. Firstly, the UKCP_global members, like other models in the HadGEM3 family, have relatively high climate sensitivities compared with the CMIP5 models (Yamazaki et al. 2021). Secondly, the external forcings applied to the UKCP-global simulations are higher, as a result of the use of an experimental design that accounts for carbon cycle uncertainty by sampling a range of CO2 emissions pathways that are consistent with RCP8.5 but lie above the standard RCP8.5 pathway due to carbon feedbacks (Murphy et al, 2018).

REGIONAL PROJECTIONS FROM CONSTRAINING OR WEIGHTING GLOBAL MODEL PROJECTIONS

Approaches to weighting or constraining model projections based on model performance, and/or independence between ensemble members have the potential to improve the information about projection uncertainty by 1) adding additional information from observations and 2) avoiding ‘double counting’ of outcomes from closely related models in probability estimates.

At the global scale, these approaches have indicated that the higher-end global responses in CMIP6 are less consistent with observed recent climate change and can therefore be ‘downweighted’ in estimates of uncertainty in future warming (e.g. Ribes et al. 2017; Brunner et al. 2020b). Further, these methods have informed the IPCC AR6 process to assess multiple lines of evidence in the assessment of the ‘very likely’ range of global climate sensitivity at 2-5K, which has important implications for interpreting the projections from CMIP6 which lie outside of this range. By applying constraints at the regional scale, we explore the implications of down-weighting high-end global responses of regional projections.

EUCP WP2 compared a number of approaches to weighting and constraining the range of regional projections using observational data (Brunner et al. 2020a; Booth 2021) and have made available online in an atlas of constrained projections (Liu et al. 2021). The comparisons of multiple approaches to applying weighting and constraints applied regionally for Europe in WP2 have demonstrated some diversity in the impact on projection range depending on the method selected suggesting a lack of robustness in these methods (Brunner et al. 2020a). Further analysis testing the degree of skill across the methods using an out-of-sample testing framework however suggest that for summer temperature, all methods were found to add value and tend to bring down the upper end of the summer temperature projections range. However, for precipitation, there is little evidence of added value by these methods (Booth *et al*., 2021).

We draw on the results made available through the online atlas of constrained projections (Liu et al. 2021), listed in Table 3. These methods themselves are diverse. Importantly, the UKCP-probabilistic projections are fundamentally quite different from the CMIP5 and CMIP6 based methods. UKCP probabilistic estimates capture a wider range of model uncertainties by including both parameter and structural uncertainty sources, and account for carbon cycle feedback uncertainty, the combined impact of which is a wider uncertainty range, with carbon cycle impacts leading to a notably higher 'top end' of the projection range.

Method	Ensemble	Variables	Reference
Climate Model Weighting by Independence and Performance (ClimWIP)	CMIP5, CMIP6	pr, tas	Brunner et al. 2019
Allen–Stott–Kettleborough (ASK)	CMIP6	tas	Hegerl et al. 2021
Reliability ensemble averaging (REA)	CMIP5, CMIP6	pr, tas	(Giorgi and Mearns 2002)Brunner et al. 2020
UKCP Bayesian probability estimation	Hadley Centre PPEs and statistical emulation	pr, tas	Murphy et al. 2019
KCC	CMIP6	tas	Ribes et al. 2021

Table 3: Constraint and weighting approaches applied to European projections in EUCP.

REGIONAL PROJECTIONS FROM HIGH RESOLUTION DOWNSCALED CLIMATE MODEL PROJECTIONS

Dynamically downscaled projections offer improved interactions between atmosphere and land as a result of the more detailed representation of the land surface (including coastlines and topography), and therefore offer projections which are more relevant for impacts studies. EURO-CORDEX offer a relatively large ensemble allowing for uncertainty estimates, however, these simulations sample a relatively small number of driving GCMs from CMIP5. CPMs have been demonstrated to offer significant added value over GCMs and RCMs in the representation of important high-impact weather features such as extreme rainfall (e.g. Kendon et al. 2014)), but in most cases the small ensemble sizes mean that we cannot expect those members to adequately represent model uncertainty in either the large scale characteristics inherited by driving GCMs or finer scale characteristics of the CP-RCM..

Statistically downscaled products offer an alternative way to add this spatial detail to projections and a number of statistically downscaled projection products based on different methods are also used across Europe.

Ensemble	Driving models	RCM / CP-RCM members	Ensemble Size	Resolution	Reference
EURO-CORDEX	6 members from CMIP5 (RCP8.5)	11	48	~12.5 km	(Jacob et al. 2014)
UKCP-Regional	12 members of UKCP-Global (RCP8.5)	1	12	~12 km	Murphy et al. (2019)

EUCP WP3 CP-RCM	3 CORDEX RCMs, 4 CORDEX-like RCMs	Up to 8	Up to 15 (maximum ens size is 15 for ALP-3, minimum 1 for NEU)	~3 km	Belušić et al. (2021) - All domains. Coppola et al. 2020, Ban et al. (2021); Pichelli et al. (2021) - *ALP-3* domain.

Table 4 Downscaled Projections for European domain. Full details are provided in appendix 6.2.2

3.4.3 Interpreting projections from multiple lines of evidence: Case Studies

As a result of this analysis, we have produced a series of figures summarising the data from these data sources. Figure 3.4-2 represents an example of these figures, with representation of the projections from the datasets mentioned in the previous section all contained in one figure. All box and whisker representations in these figures represent the inter quartile range (25^{th} to 75^{th} percentiles), median, and 10^{th} and 90^{th} percentiles.

The upper left-hand panel in each summary plot shows a direct comparison between the regional projections from the three GCM ensembles, there is also an additional box and whisker representing data from all the GCMs together. Horizontal lines representing the 10^{th}, 25^{th}, 75^{th} and 90^{th} percentiles from this distribution are overlaid across all panels in the top left as an aid to interpretation of the information, and their use is discussed in more detail in Section 3.4.4.

The middle panel of the top row represents probability estimates from the weighting and constraint methods in EUCP work package 2, with the unconstrained ranges represented in semi-transparent colours, and the constrained or weighted ranges represented with opaque colours. Finally on the top row, the top right-hand panel represents data from downscaled climate models, with the driving models for each multi model ensemble represented by triangles immediately to their left.

The bottom row presents information intended to aid in further understanding of these lines of evidence. The bottom left panel represents projections from the GCMs after normalising by global temperature response, to assist in determining the significance of the global sensitivity of models in their regional results. The two right hand panels on the bottom row use scatter plots intended to reveal information on the relationships in responses between downscaled models and their parent driving models. The middle panel represents outputs from the RCM data (CORDEX) and the corresponding CMIP5 driving models, while the right-hand panel represents the EUCP WP3 CP-RCM models and their driving CORDEX models.

CASE STUDY 1 – SUMMER ALPINE TEMPERATURES

The alpine region has been a focus of recent developments in convection permitting regional climate modelling, and a much larger amount of data from convection permitting models is available for this region than others in Europe, our analysis includes 8 CP-RCMs from this region, whereas most of the other EUCP WP3 regions are only covered by 1 or 2 CP-RCMs. This relatively reduced number of models compared to the larger multi-model ensembles of CMIP5/6 and CORDEX can present challenges to users of climate information who require data from high resolution models for impact studies as the data from these models may not capture the full range of uncertainty from climate projections provided from other data sources.

Figure 3.4-2 shows information from the multiple lines of evidence over the ALP-3 EUCP WP3 domain. Despite the relatively large number of CP-RCM models compared to other domains, the projections from these models are concentrated towards the lower range of the projected outcomes, and do not represent possible warmer outcomes of approximately 3K or more, that are indicated by the UKCP models and the upper end of the CMIP6 and CMIP5 ensembles.

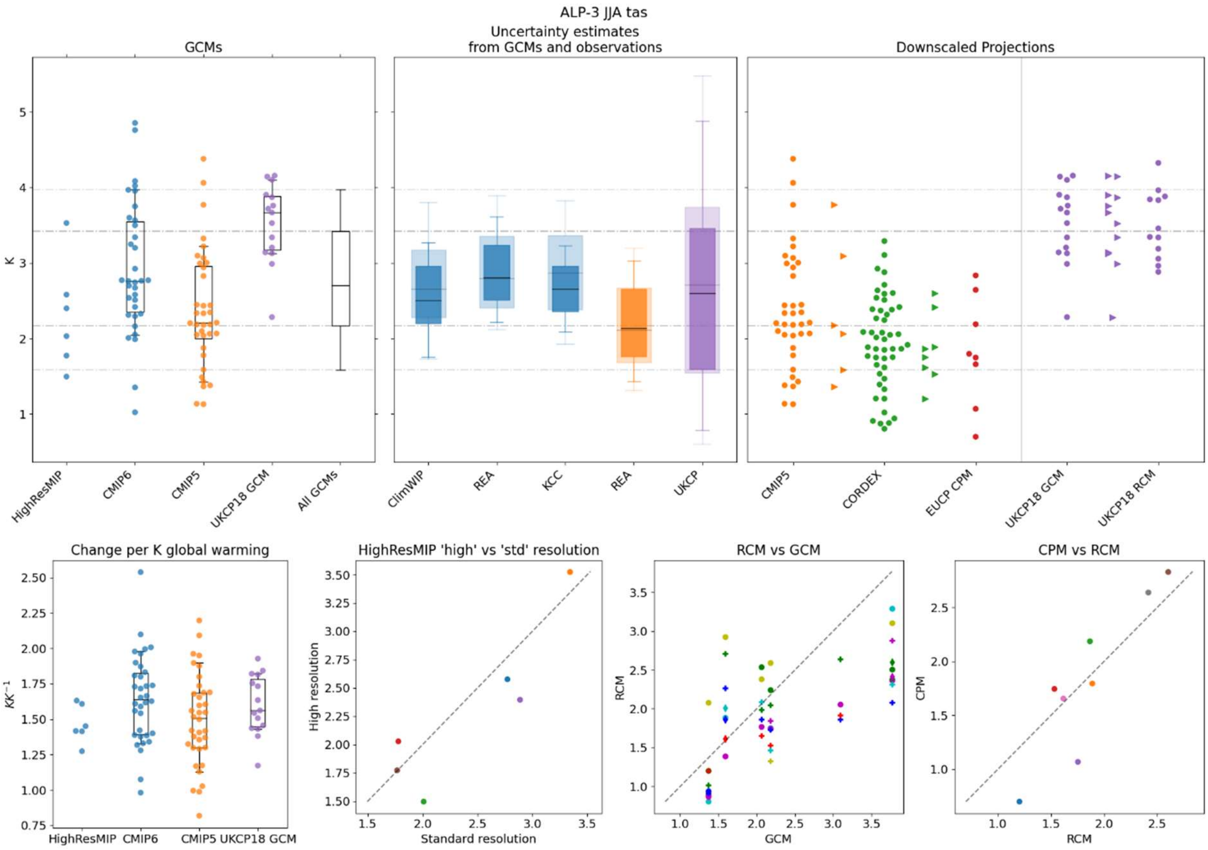


Figure 3.4-2 Change in summer alpine temperatures. Projected mean temperature anomalies between 1995-2014 and 2041-2060. Semi transparent box (interquartile range and median) and whiskers (10th and 90th percentile) in top middle panel represents unconstrained ranges, opaque box and whiskers represent constrained ranges. Triangles in top right hand panel represent driving models of respective downscaled model experiments.

The bottom panels in the figure help to provide some insight into reasons for differences between the different lines of evidence, for example, the bottom left panel demonstrates that the warmer responses exhibited by the UKCP Global models in the top left panel are due to the higher global sensitivity of the UKCP Global models.

The bottom right panels demonstrate the impact on projections of downscaling, showing that for many (but not all) of the GCM/RCM combinations in CORDEX downscaling produces a cooler projection than from the driving model.

This case study is explored further in 3.4.4 and Table 5.

Case study 2 – Summer Precipitation in Romania

Landslides are a common hazard in Romania with rainfall being a key triggering factor for these events. A recent study (Niculiţă 2020) finds that climate projections of rainfall in Romania imply an increase in landslide risk in the mid-21st century. However, this study is based on a subset of the EURO-CORDEX ensemble and it is useful to examine data from other lines of evidence in order to put the results into the context of potentially wider uncertainties.

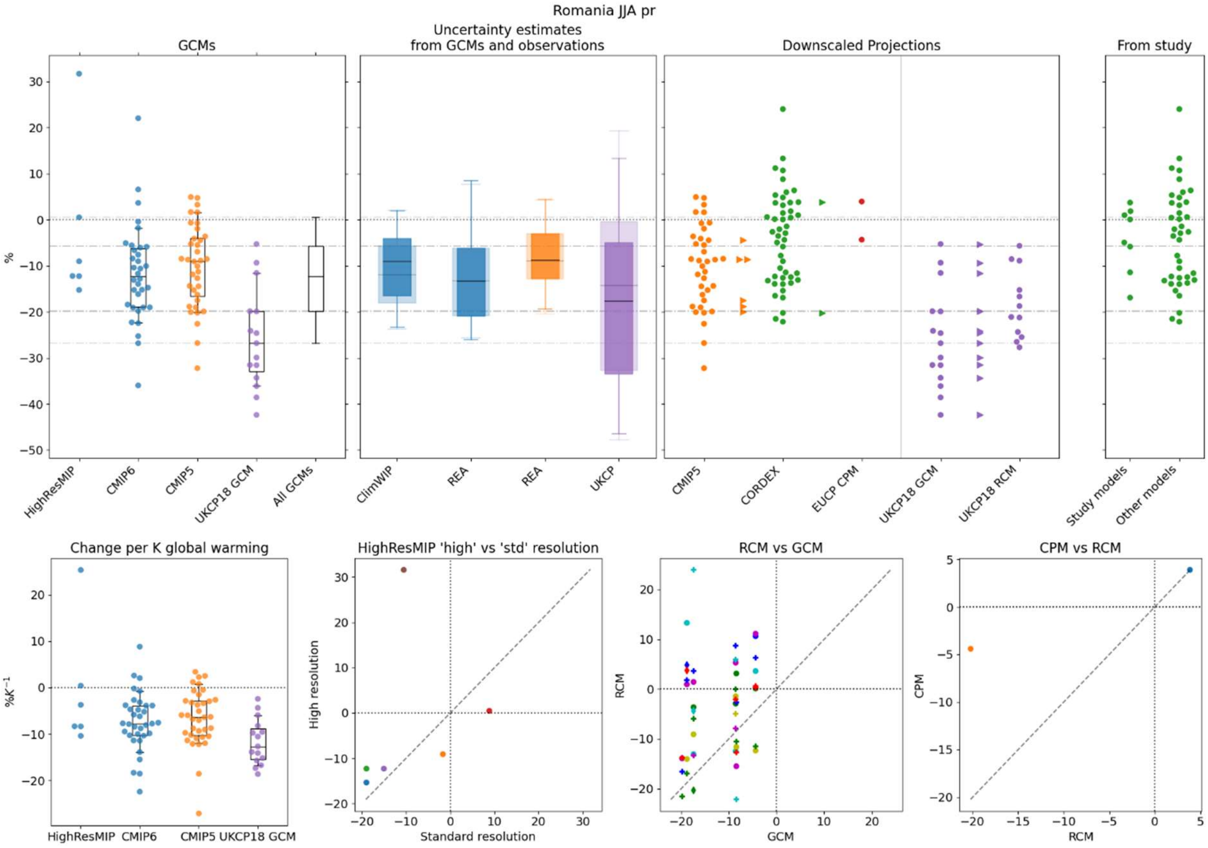


Figure 3.4-3 Different lines of evidence for Romania pr. As Figure 3.4-2 but with a further panel in the top right that shows the subset of CORDEX models used in (Niculiţă 2020) in the left column, with remaining CORDEX models in the right column.

Figure 3.4-3 presents the information from the different lines of evidence in a similar way to Figure 3.4-2, but with an extra panel on the right highlighting the results from the models used in the (Niculiţă 2020) study. It is evident that the chosen models are a fair representation of the centre of the uncertainty range across the lines of evidence, particularly from the CORDEX multi-model ensemble that they are drawn from, although there are significant possible drier projections (e.g. CMIP5/6 and UKCP GCMs and RCMs), as well as some wetter projections, mainly from CORDEX and CMIP6. In interpreting the results of the study, it is therefore useful to consider the evidence for projections of rainfall beyond the range of those analysed in the study.

The middle bottom panel also demonstrates that the impact of downscaling from CMIP5 tends to lead to wetter outcomes from the downscaled models. The information in the top panels then shows that if different, wetter driving models had been chosen the resulting range of projections from CORDEX might have been even wetter than the current range. The possibility of much wetter scenarios is also supported by the HiResMIP experiment where one member suggests increases of more than 30% are plausible.

While the Alpine region benefits from a relatively large number of CPM simulations, other regions of Europe (and indeed the globe) are more sparsely catered for. In this case, while the 2 CPMs available may be able to provide very high-resolution data that can provide more detailed information for impact studies, users should be aware that on the broader scale their projections only capture a relatively narrow range of the full range of plausible future climate projections.

CASE STUDY 3: MULTIPLE LINES OF EVIDENCE FOR THE EU'S OUTERMOST REGIONS – DJF SEASON RAINFALL IN FRENCH GUIANA.

For many of the EU's outermost regions (OMRs), the available climate projection information is less extensive than for mainland Europe. However, even with fewer projection products, and smaller ensembles typically available, those projections products can still be expected to offer different projection ranges for a given region, season and timescale for similar reasons to those mainland Europe, and therefore the provision of information across multiple lines of evidence approach offer similar benefits in terms for understanding the wider uncertainty context of any individual simulation or product.

Here (Figure 3.4-4) we demonstrate the value of this approach for French Guiana, for which downscaled projections are available from the CORDEX Central America experiment. The use of CPM simulations for the region has been applied in Pseudo-Global warming experiments, which are one of the wider 'lines of evidence' on which evidence would be drawn but are not directly quantitatively comparable. We also draw on constrained projections produced for this region using one of the EUCP WP2 methods, the UKCP methodology. Note that several of the methods applied for mainland Europe in D2.2 and D2.3 (also Brunner et al. 2020b) depend on constraints which are demonstrated to apply the climate of mainland Europe and are not transferable to other climates experienced by many of the OMRs.

In this example, global model products largely agree that reductions on mean rainfall are likely or very likely, with just a few individual models from CMIP5 and 6 suggesting increased mean rainfall. However, additional lines of evidence from UKCP constrained projections, as well as higher resolution information from CORDEX, support the possibility of increased mean rainfall. CORDEX simulations do undersample the drier driving global models, and so users should note that the dataset may exclude plausible drier scenarios.

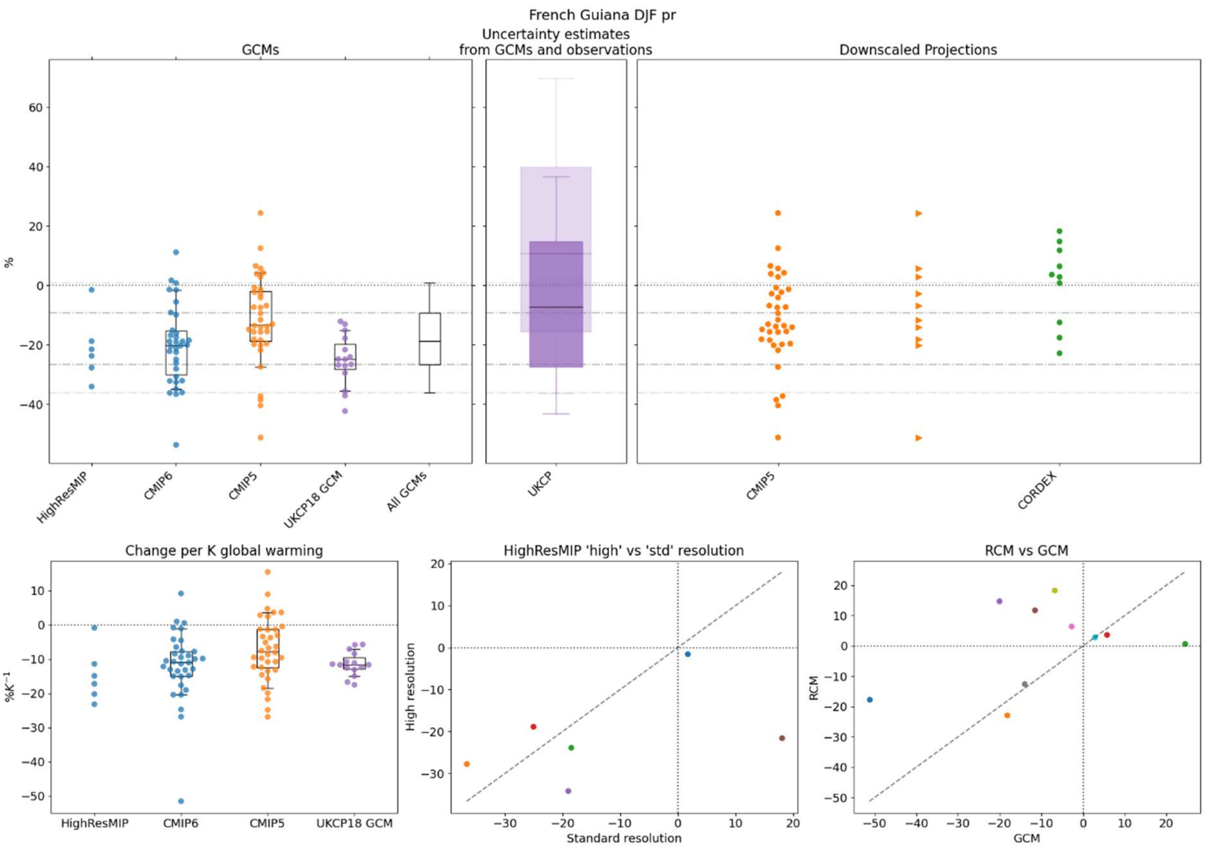


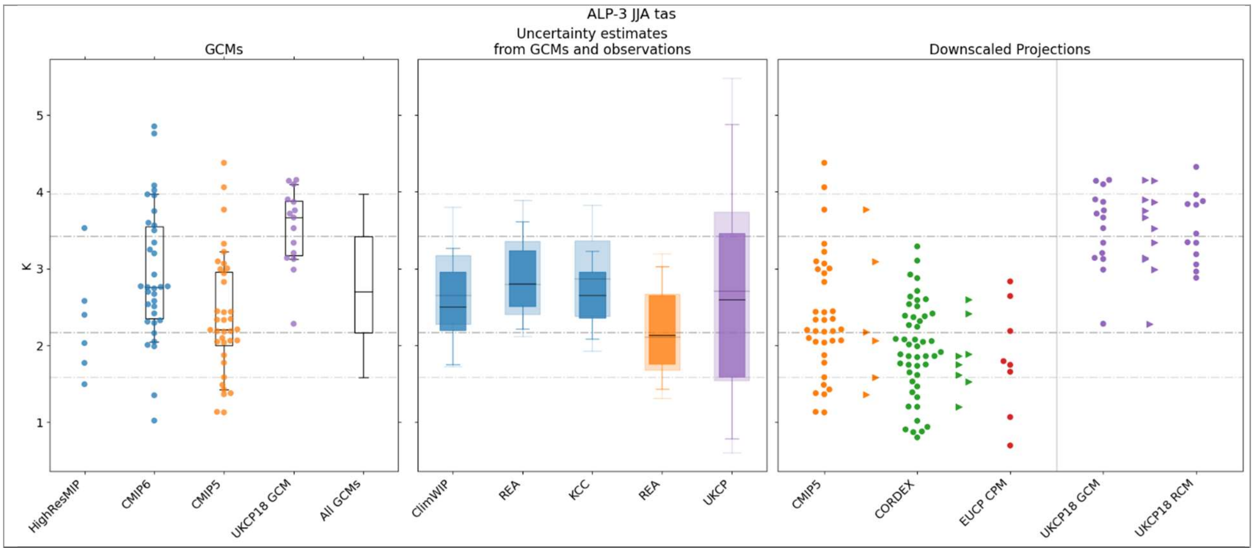
Figure 3.4-4 Change in DJF mean precipitation for French Guiana. Projected mean precipitation anomalies between 1995-2014 and 2041-2060. Semi transparent box (interquartile range and median) and whiskers (10th and 90th percentile) in top middle panel represents unconstrained ranges, opaque box and whiskers represent constrained ranges. Triangles in top right hand panel represent driving models of respective downscaled model experiments.

3.4.4 Synthesizing and communicating projection confidence based on multiple lines of evidence

In this analysis we have demonstrated the benefits of looking across multiple projections datasets to gain wider uncertainty context for individual products. The data remains complex, and difficult to interpret without familiarity with the characteristics of different products. While the quantitative comparison of available datasets offers a valuable reference for those working with projections data in the region, further supporting evidence about each of those ranges may help to explain where there are conflicts, and even offer additional understanding that might lead to higher or lower confidence in some datasets or parts of the projection range.

The most recent IPCC assessment report has demonstrated the value of using 'multiple lines of evidence' to provide more robust estimates of global climate sensitivity, drawing not only on projection datasets, but also on additional evidence from observational constraints and paleoclimate. Here we have started to look towards a more comprehensive assessment of 'multiple lines of evidence' at the regional scale through a comparison of some of the key projection datasets, with a view towards integrating further lines of evidence and the synthesis of robustness or uncertainty statements around regional climate change.

Here we explore a potential framework for providing supporting narrative information to projection ranges as well as supporting users in selecting appropriate products. The distribution of raw GCM members across CMIP5, HighResMIP, CMIP6 and UKCP-Global is used to approximate three parts of the projection range: ‘Warmest projection range’, ‘mid projection range’ and ‘lowest projection range’. We use IPCC uncertainty language to discuss the likelihood of projections based on the different estimates and products within these three broad parts of the range (Unlikely = < 33%, Very unlikely = <10%, Likely = >66%, More likely than not >50%).



Projected outcome	Supporting evidence	Which products/members?
Warmest projection range (~+4 – 5.5K)	Scenarios above 4K may be plausible but *unlikely* or *very unlikely* according to the weighted estimates. While these changes appear in models the upper tails of raw CMIP5/6, and the upper half of the UKCP-Global, these warmest scenarios tend to arise in models with high global sensitivity which are considered less plausible and tend to be downweighted by constraints or weighting approaches.	Explore the warmest projection range using highest responding members of CMIP5, CMIP6, UKCP-Global and UKCP-Regional. This ‘high end’ is not captured by EURO-CORDEX or EUCP CPMs.
Mid projection range (~+2– 4K)	Changes within this range are more *likely than not* according to most methods based on global models. However, within this range, there remain significant differences between the projections from different products. CMIP5-based downscaled products (EURO-CORDEX and EUCP-CP-RCMs) favour lower responses than the UKCP Global and regional products. These differences arise for a combination of two reasons: 1) The driving global models in UKCP global typically capture larger global warming responses than CMIP5 which are reflected in the ALP-3 region 2) Downscaling further enhances the differences between ensembles; EURO-CORDEX tends to cooler outcomes than driving CMIP5 GCMs.	Explore the mid projection range using any of the GCM products, and/or the EURO-CORDEX or UKCP-regional products. However, members from both the EURO-CORDEX and UKCP-regional downscaled products will be needed to capture the full +2-4K range. Very high-resolution projections from CPMs do not represent changes above +3K.
Lowest projection range (~+0.5 – 2K)	A number of global models indicate changes less than +2K. The UKCP probabilistic range, as well as the and the reduced temperatures in EURO-CORDEX downscaled projections relative to driving CMIP models support lower end projections as part of the *likely* range.	Explore the lowest projection range using low-response members of EURO-CORDEX or EUCP-CPM. This part of the projection range (<+2K) is not captured by UKCP Global or Regional members.

Table 5: Summarising multiple lines of evidence for summer temperature changes in the ALP-3 domain.

3.4.5 Discussion

Multiple projection products offer different projection ranges for European regions, which may lead to different conclusions when applying data to impacts or adaptations studies. The data processing overhead prohibits most individual users from looking at multiple datasets for a wider uncertainty context. Here we have gathered projections from a wide range of products in order to make such information accessible to users of projections in Europe.

The figures produced as part of this analysis are being made available at https://zenodo.org/record/6046762 along with the raw data used to compute them. In addition, the scripts used to compute these data are made available at https://github.com/eucp-project/Lines-of-evidence-catalog . This could form the basis for a more flexible interactive tool that would allow a user to produce a bespoke figure for specific regions, seasons and time periods.

This wider uncertainty context from 'multiple lines of evidence' offers users valuable context which might 'bookend' climate impacts studies. At the outset of a study, this information might provide an initial indication of the wider uncertainty context to inform a dialogue with users, and inform the selection of projection products to ensure that the study captures key parts of the uncertainty space. For some studies this may mean capturing a representative spread by combining ensemble members from more than one dataset (e.g. combining EURO-CORDEX with UKCP-regional to span a wider range of uncertainties), while for others this may mean targeting specific part of the uncertainty range (e.g. driest, wettest, hottest as 'worst case' scenarios or storylines). The wider uncertainty context then provides further information with which to interpret the results of an impacts study that is based on a specific set of models, or storylines, as we have demonstrated in case study 2. This information can also add more general background information around mean climate change projections to support event-based storylines studies.

The analysis also offers a useful basis for infilling missing parts of the uncertainty space in downscaled datasets, either with new dynamical downscaling experiments or statistical approaches such as those explored in EUCP D5.4.

This information might be further augmented by more diverse lines of evidence including internal variability from LSEMs, and statistical downscaling methods, physical process understanding and model evaluation, and physical evidence from other experiments such as pseudo-global warming (PGW) experiments. Grainger et al. (2022) for example demonstrate an expert elicitation process by which such alternative lines of evidence can be synthesized into quantitative uncertainty ranges which can be compared with projection products such as CMIP5 and 6. Alternatively, qualitative lines of evidence could be included by annotating the lines of evidence figures, or through the framework for building confidence statements proposed in Table 4.

Some further useful steps that could contribute to a such a synthesis include:

1. A simplified format for the combined figures capturing the 'key' or combined lines of evidence in order to provide a more digestible assessment suitable for a wider range of users
2. Annotations to combined figures could be used to highlight additional qualitative evidence, and which parts of the uncertainty range have elevated or reduced

3. Reducing or combining the number of weighted uncertainty ranges may further simplify the complex figures in e.g. figure 2 and 3. Booth et al. (2021) suggests that a 'combined' methods offer improved skill over any one approach, and would complement the UKCP pdf as a 'most comprehensive' uncertainty range.

There are a number of directions in which this work could very usefully be further developed. We have focussed on seasonal mean temperature and precipitation, but extending to include indices of extremes or impacts drivers such as those explored in Coppola et al. 2021 would offer new information about the wider uncertainty context and the development of robustness statements for these more impact relevant measures. Currently the constrained ranges offered in the WP2 atlas and explored in intercomparisons do not extend to extremes indices, but a number of studies have applied methods to extreme indices e.g. (Murphy et al. 2020)).

IPCC AR6 has demonstrated the benefit of using 'multiple lines of evidence' approaches to arrive at robust confidence estimates. A number of developing research areas will help to arrive at combined uncertainty range for regions of Europe that take in these multiple lines of evidence:

- Published code for multiple methods of constraint so that a combined method can be applied to specific regions and ensembles, including downscaled datasets.
- Improved sampling of GCM-RCM-CPM matrix will help to provide more complete uncertainty information in downscaled products – this could be partially achieved with improved experimental design (e.g. for CMIP6 generation EURO-CORDEX) as well as emulated downscaled projections from a larger number of GCM-RCM(-CPM) combinations
- Weighted/constrained downscaled projections – including information from performance and model dependency, and making use of emulated projections to fill gaps.

3.5 Constructing storylines for real-world applications (All)

We have presented the studies carried out as part of T5.5 that seek to address some of the gaps in the growing area of climate storylines. The breadth of storylines-related science and applications in this deliverable are summarised in Figure 3.1-1, the sections covered:

- Examples of end-to-end production of storylines for specific user applications, where hazard metrics and weather pattern analysis are combined to investigate their construction and use in dialogue with users,
- Two studies that explore future hazard events sets at pseudo-global warming levels, providing the building block of storylines rooted in the recent lived experience of users via recent impactful events,
- An algorithmic clustering approach which offers the possibility of reducing large ensembles by selecting representative members which are coherent across the regions, timescales and metrics of interest
- Two studies that explore the variability in climate projections which show a tendency to underestimate the contribution of internal variability, which storylines may be able to aid in addressing,
- A multiple lines of evidence assessment tool where climate model projections ensembles are made easily accessible to users interested in European projections and need to decide which information to use.

At the beginning of this deliverable, we posed three questions which we attempt to answer, bringing in the results and learning from Sections 3.2 to 3.4 as well as the work carried out in other EUCP work packages. In this discussion section on storylines for applications we focus on storylines as a user product, rather than as part of the data production process or a scientific exercise.

3.5.1 What are climate storylines and where are they useful?

We have shown different approaches to constructing storylines for a variety of purposes from data production and scientific understanding to impact assessments. We have also demonstrated the potential of storylines as a user product in some real word applications, adding to the body of published work which shows their use in generating scientific understanding. While it remains difficult to generalise what storylines as a user product should comprise, as this is dependent on the application, there are a number of features which appear to be common in our as well as other published work. We summarise below some the key findings from the reported studies, and we note that many of them are not exclusive to the storyline approach. However, storylines provide another tool or source of evidence for users to add to their portfolio to inform climate risk assessments and adaptation planning:

The value of storylines as a product is the co-production process where the knowledge produced through the collaboration is as valuable to building the credibility and confidence in using the climate data as well as ensuring that the output is useful and usable. See Sections 3.2.2 and 3.2.3 as well as Jack et al (2020). In particular, understanding the purpose of the storylines, whether for informing decisions, driving engagement or communication is important to establish at the beginning and inform the format of the storyline.

Contextualising the future using past events is essential for engaging any audience and to promote better understanding. We observe this in both sets of users who want to use this information to provide future narratives that are associated with lived experiences (Section 3.2.2) as well as embedding in existing analysis that investigate current vulnerability of systems (Section 3.2.3). In Sections 3.3.1 and 3.3.2, we show that event-based information can allow deeper interrogation of changes to the severity and drivers of such an event in the future, focussing on physically plausible analogues and limiting the need for interpretation of uncertainty. Interestingly, we see that pseudo-warming approaches to producing event-based data on climate hazards produce information which may be used in a similar manner to selecting event analogues from ensembles. These ‘‘scientific storylines’’ based on past events would likely form the basis of a powerful user product for the right applications. Using past events to engage an audience is also used in other arenas such as climate change event attribution as well as climate analogues that storyline approaches could learn from.

At some point in the storyline construction process, a decision needs to be made on which aspect(s) of uncertainty to focus on. This could be the range of outcomes for a specific impact metric (Sections 3.3.3, 3.3.1 and 3.3.2), variability (Sections 3.3.4 and 3.3.5), other climate drivers (Sections 3.2.2 and 3.2.3) or a combination of the above. This is unlikely to be known a priori when forming storylines for a specific application. The focus (foci) of the uncertainty analysis also informs the sub-selection method which storyline methods ultimately require. This point applies to any climate risk assessment; however, storylines may offer explicit ways to expose uncertainties, tailoring the messaging to the user and their application.

There is value in the use of "driver information" in characterising events of interest for a wide range of users, acting as a link between past events and future projections and explaining uncertainty using storylines. In the application examples (Sections 3.2.2, 3.2.3), we observe that both organisations new to and those experienced in climate risk assessment valued the information contained in climate drivers. Those in the water supply sector were keen on understanding models' ability to capture droughts and their drivers: this credibility evaluation does not need to be part of a storyline production process, but these discussions were helpful in selecting the building blocks of potential storylines. While not explored in T5.5, there is a clear need to extend the analysis to include storylines of decadal variability explored in Sections 3.3.4 and 3.3.5 for drought applications.

Scientific analyses or output need to be carefully chosen to align with a particular application. While storylines or building blocks generated as part of a scientific process may be standalone products relevant to certain users with minimal tailoring, ideally the production of storylines, or related information, for user applications should be demand driven. In Table 6 potential pros and cons of the different scientific building blocks of storylines covered in this deliverable are listed against the two example applications from Section 3.2. It is worth noting that the types of information and methods listed do not only apply to storylines, but also the production of user relevant climate information more broadly.

<table>
<tr><th colspan="2">Storyline building blocks</th><th>Heritage management: awareness raising</th><th>Water supply management: operational planning and investment decisions</th></tr>
<tr><td>Driver-based selection</td><td>Pros</td><td colspan="2">Allows consideration of the unfolding of events and timing of changes in a self-consistent manner – requirement of impacts models. Provides a narrative bridging past events and future changes.</td></tr>
<tr><td></td><td>Cons</td><td>Difficult to understand and may limit the range of hazard metrics explored.</td><td>Cannot replace existing methods but could add to the knowledge base.</td></tr>
<tr><td>Event-based using PGW</td><td>Pros</td><td>Events and example impacts seen as key to engaging with climate information.</td><td>Can be embedded in existing analytical frameworks.
Relaxes one axis of uncertainty.</td></tr>
<tr><td></td><td>Cons</td><td>Lack of historical events which show extreme heat. Difficult to upscale.</td><td>Timing of event is extremely important.</td></tr>
<tr><td>Impact-based clustering</td><td>Pros</td><td colspan="2">Consistent selection across multiple hazard metrics (variables, seasons, indicies) and regions.</td></tr>
<tr><td></td><td>Cons</td><td colspan="2">Difficult to communicate, and no physical explanation for impacts.
limits the uncertainty range to cluster representatives.
Potential lack of diversity given the small ensemble sizes.</td></tr>
<tr><td>Enhancing variability</td><td>Pros</td><td colspan="2">Provides more realistic information on the key hazard metrics of interest where variability is vital.</td></tr>
<tr><td></td><td>Cons</td><td>Difficult to upscale or re-produce in the sector.</td><td>No likelihood information.</td></tr>
</table>

Table 6 Anticipated pros and cons of different storyline building blocks for two example applications

3.5.2 How could storylines bring together various outputs and products of EUCP science?

We have shown how different climate datasets and methods across EUCP could be used for storyline construction. But we have not explored how additional developments in EUCP may be incorporated into storyline construction and consequently build on the key characteristics of a storyline service such as credibility, salience and legitimacy (Cash, 2003). These developments include:

- The impact of enhanced skill in the decadal projections in WP1 (Smith et al, 2020).
- Using weighting methods and constraints investigated in WP2 and WP5.
- Applying methods to produce seamless storylines of predictions and projections based on the work in WP5.

The decadal predictions and the groundwork for comparing and merging them with projections presents the opportunity to bring this emerging source of shorter timescale information into climate services products. Decadal predictions have the potential to constrain projection information, or form storylines over the next 10 years, utilising the behaviour of ensemble members as well as the skilful mean where appropriate. Further exploration of this requires a test application which can benefit from climate information over both timescales, which did not emerge in the user case studies in Section 3.2.

The weighting and constraints methods provide sub-selection options and weightings to form the basis of 'Europe-wide' storylines, to be analysed further at country or regional levels, or to be used for downscaling. This would aim at reducing the range of uncertainty in the raw climate model output where possible with constraints, and ensuring the remaining uncertainty space is covered by smart sub-selection. These methods may also be used to constrain the uncertainty space before forming storylines as a bespoke user product, such as the examples in Section 3.2.

As with most user requirement studies, our application examples also find that users carrying out climate risk assessments and making related decisions require more information than the climate hazard, in particular the risk related impacts relevant to a given decision. A storyline service would benefit greatly if they included impacts information such as those generated in WP4, e.g. the urban flood hazard and wind-drought. The storylines could then be constructed using the impact hazard, e.g. fluvial flooding and overheating in building for Section 3.2.2 and fluvial flows and groundwater levels for Section 3.2.3. The physically based storylines could go further towards the risk and provide climate risk storylines which could be very different. This is because there are likely to be non-linearities as we cascade down the modelling chain from climate model, bias-correction, impact, systems and risk modelling. The analysis could reveal system thresholds and parts of the uncertainty space that are of greater interest when the full hazard-vulnerability-exposure matrix is involved. The purpose of these storyline would also be somewhat different compared to those explored in this deliverable, i.e. they would provide storylines of the whole system as explored in Hazeleger et al (2015) rather than providing the basis of further investigation and risk assessment.

3.5.3 What are the challenges of producing storylines as a climate service?

Climate services is a growing area of science and application research and converting scientific analysis and information to a service requires considering other elements beyond the veracity and robustness of the science. We have shown in this deliverable the opportunities that storylines provide for users; here we reflect on the challenges faced with converting storylines production into a climate service:

Availability of relevant uncertainty information One of the important achievements of T5.5 has been the development of a tool to allow the assessment of multiple lines of evidence. Such a tool would aid a storyline production service as well as informing the building blocks of storyline construction. For example, understanding the spread of the uncertainty across the multiple EUCP datasets would have supported early discussions with users to explore how best to construct storylines to explore uncertainty. This may 'bookend' the development of storylines for applications. Appendix 6.2.1 shows this assessment for UK climate variables, related to the applications in Sections 3.2.2 and 3.2.3. This may be used by service providers at the beginning of a project to assess the available information and degree of uncertainty and conflict therein. It may also be used in the final stages to make statements about the information used in a full or constrained uncertainty context, or provided directly to technical users, forming part of the co-production dialogue.

Tailoring the storylines As discussed in Sections 3.2 and 3.3.3, sub-selection is required to construct climate storylines and can be performed to produce a pan-European set of storylines. Examples include the latest EUROCORDEX discussions around key drivers and metrics to determine global climate models for downscaling to sample the uncertainty space. However, as discussed in Section 3.5.1, the format and content of storylines as a user product will depend on the audience and purpose. They are likely to be location, hazard and application-specific and as shown in Section 3.2, can require bespoke storylines that require a lot of resource. Pan-European storylines may not be useful foundation when generating storylines for these audiences. On the other hand, pan-European storylines may still be valued, even given evident trade-offs and compromises with how well they can capture individual sector, region and application demands. They allow impacts that cross sectors and geographical regions to be explored. Examples like flood management and food security rely on being able to explore the impact of particular climate narratives that cross national and sectorial boundaries. In the context of potential future European climate service frameworks, exploring the trade-offs between tailored sectorial/local climate narratives and necessarily imperfect pan-European narrative approaches, would need to be considered.

Upscaling resource intensive model experiments and analysis We show that the study of large-scale, long-duration events such as heatwaves and droughts can be carried out using coarse resolution models, and show potential for use in storylines products. However, this is after resource intensive physical consistency and plausibility checks which are also important for sub-selection, such as in section 3.2.2. These depend on the user metric of interest – expanding the lines of evidence tool could help in this regard. In WP3 and WP4, the use of high-resolution climate convection-permitting models provide key climate hazard information (i.e. surface water flooding) and also shown in T5.5 to be a useful tool for event-based storyline construction. However, the computer resource requirements could limit their application given the different types of events that could be of interest to an organisation. Efforts in T5.4 on spatial merging and emulating CPMs offers an opportunity for a less resource-intensive method to explore storylines that cover a larger part of the GCM-CPM uncertainty space should CPM-related hazard information be required. This may also compliment the role of sub-selection when ensuring downscaling activities or impact studies cover an adequate range of uncertainty more generally, such as EUROCORDEX mentioned above.

Future work should focus on testing storylines in decision making, risk assessment and communication applications, incorporating relevant scientific building blocks and framed with a lines of evidence assessment at the appropriate level of detail. This is the intersection of these elements in the Venn diagram in Figure 3.1-1, posed as the ideal space for constructing storylines as a user product. While this deliverable has not directly addressed this via a single example, the breadth of storylines-related science and applications presented has helped shed light on this gap, and what is required to fill it.

4 Lessons Learnt and links built, deviations and additional activities

4.1 Lessons learnt and links built

As one of the tasks that included user engagement and depended on other tasks within the work package and across the project, a number of procedural lessons were learnt through the project:

- The step change in the methods for engagement due to the COVID-19 pandemic introduced obstacles in terms of cross-collaboration and traditional user engagement methods. However, virtual meetings were more inclusive allowing us to arrange a pan-European heritage management storylines workshop which would not have happened otherwise.
- The original proposal made a number of assumptions about early availability of data from the experiments from WP1 and WP2 as well as results from other tasks in WP5. This could possibly have been anticipated and T5.5 objectives modified earlier. This may have improved the coordination of activities.
- Earlier engagement with the Multi-User Forum (MUF) and its setup may have helped to draw together the storylines work for pan-European examples.

As demonstrated in this deliverable, Task 5.5 involved elements across multiple EUCP work packages as well as working with external organisations. Links built include:

- Expanding our understanding of storyline construction methods, particularly with WP2 authors who have contributed to this deliverable.
- Through the application examples, we were able to foster new collaborations with heritage management sector as well as expand the MUF with the majority of collaborators able to attend the MUF workshops.
- Collaborators in the water supply example have also reached out to other work packages to understand how best to use the novel research. Through the drought example application, we were able to join up with the Horizon 2020 project B4est to improve their understanding and application of climate data across Europe.

4.2 Deviations in activities

The original ambition of the task was to construct seamless storylines from decadal forecasts to climate projections. As explained in Section 4.1, this was deemed too challenging in the timeframe of the EUCP project, particularly due to much of the emerging science on these aspects being conducted in parallel (temporal comparison, temporal and spatial merging, constraints). While WP5 has advanced these topics, the conclusions have only just become available and there are several open questions about the application of this novel science. Therefore the objectives (and deliverable title) were pre-emptively revised to those set out in Section 2. Focussing instead on example applications of storylines, and using these case studies to reflect on various pieces of science from across EUCP has resulted in an arguably more impactful body of work which begins to fill in various gaps in this emerging area of climate science and services.

4.3 Activities in support of Task 5.5

Storylines seminar (external): A seminar for the EUCP project on storylines was organised by UKMO on 25/03/2021 with external speakers who have recently published work on different storylines-related aspects. There was a discussion session following the talks and the recording was made available to project participants on the EUCP internal wiki.

Storylines workshop (internal): EUCP scientists participated in an internal min-workshop on 29/09/2021 to present and discuss the range of scientific research and applications related to storylines across the project with the aim of promoting discussions and future collaboration, as well as finalising the content of this deliverable.

User engagement: As described in Section 3, and the associated appendices, user engagement activities were performed with the water resources and heritage management sectors to support two case studies of end-to-end storyline production.

EUCP published and planned articles

- *Van der Wiel et al. 2021* - K van der Wiel, G Lenderink, H de Vries (2021): Physical storylines of future European drought events like 2018 based on ensemble climate modelling. Weather and Climate Extremes, 33, pp. 100350.
- Matte, D., J.H. Christensen, H. Feddersen, R.A. Pedersen, H. Vedel, and N.W. Nielsen, 2022: Event attribution with a convective permitting ensemble forecast model, Geophys. Res. Lett. [In review]
- Ballinger, A.P., A.P. Schurer, and G.C. Hegerl (in prep, 2022), Accounting for the NAO when applying observational constraints to future European climate projections, Environ. Res. Lett. (in prep)
- O'Reilly, C.H., Befort, D.J., Weisheimer, A. et al. Projections of northern hemisphere extratropical climate underestimate internal variability and associated uncertainty. Commun Earth Environ 2, 194 (2021). https://doi.org/10.1038/s43247-021-00268-7
- Wilcke, R. A. I., Bärring, L. (2016). Selecting regional climate scenarios for impact modelling studies. Environmental Modelling & Software, 78, 191-201. https://doi.org/10.1016/j.envsoft.2016.01.002
- Planned: EUCP storylines position paper on '*Towards a storylines climate service*'
- Planned: Article on section 3.2.2 '*Role of climate science in developing storylines for the heritage sector*'
- Planned: Article on Section 3.2.3, '*Weather pattern-based storylines of future drought*'

Other resources

EUCP Storyboards have been produced on Sections 3.3.1, 3.3.2 and 3.4, with more planned before the end of the project (https://eucp-project.github.io/storyboards/).

The data and tool used in the lines of evidence assessment in Section 3.4 is being made available. The figures produced as part of this analysis are being made available at https://zenodo.org/record/6046762 along with the raw data used to compute them. In addition, the scripts used to compute these data are made available at https://github.com/eucp-project/Lines-of-evidence-catalog .

5 References

Anglian Water, Climate change adaptation report 2020, https://www.anglianwater.co.uk/siteassets/household/in-the-community/climate-change-adaptation-report-2020.pdf, accessed 23/03/2022

Bachmann K, Keil C, Craig GC, Weissmann M & Welzbacher CA (2020). Predictability of deep convection in idealized and operational forecasts: Effects of radar data assimilation, orography, and synoptic weather regime. Mon. Weather Rev. 148, 63–81.

Ban N, Caillaud C, Coppola E, et al (2021). The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part I: evaluation of precipitation. Clim Dyn. https://doi.org/10.1007/s00382-021-05708-w.

Belušić D, Lind P, Adinolfi M, et al (2021). HORIZON 2020 THEME SC5-2017 European Climate Prediction system (EUCP) Regional Climate Model (CPM) simulations for the historical and near future.

Ben Bouallègue, Z & Theis SE (2014). Spatial techniques applied to precipitation ensemble forecasts: From verification results to probabilistic products. Meteorol. Appl. 21, 922–929.

Bengtsson L, Andrae, U, Aspelien T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Hansen-Sass, B., Homleid, M., and Hortal, M. (2017). The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system. Mon. Wea. Rev., 145(5):1919–1935.

Booth B et al (2021) Production of uncertainty quantifications / PDFs , including separation into different components .

Brunner L, Lorenz R, Zumwald M, Knutti R (2019) Quantifying uncertainty in European climate projections using combined performance-independence weighting. Environ Res Lett 14. https://doi.org/10.1088/1748-9326/ab492f.

Brunner L, Mcsweeney C, Ballinger AP, et al (2020a) Comparing Methods to Constrain Future European Climate Projections Using a Consistent Framework. J Clim 33:8671–8692. https://doi.org/10.1175/jcli-d-19-0953.1.

Brunner L, McSweeney C, Ballinger AP, et al (2020b) Comparing Methods to Constrain Future European Climate Projections Using a Consistent Framework. J Clim 33:8671–8692. https://doi.org/10.1175/JCLI-D-19-0953.1.

Brunner L, Pendergrass AG, Lehner F, et al (2020c) Reduced global warming from CMIP6 projections when weighting models by performance and independence. Earth Syst Dyn 11:995–1012. https://doi.org/10.5194/esd-11-995-2020.

Byun H-R and Wilhite DA (1999). Objective quantification of drought severity and duration, Journal of Climate, 12(9), 2747-2756. https://doi.org/10.1175/1520-0442(1999)012%3C2747:OQODSA%3E2.0.CO;2.

Carley Iles and Gabriele Hegerl (2017), Role of the North Atlantic Oscillation in decadal temperature trends, Environ. Res. Lett. 12 114010

CCRA3, UK's Third Climate Change Risk Assessment, Technical Report for the CCRA3. https://www.ukclimaterisk.org/independent-assessment-ccra3/technicalreport, accessed 23/03/2022.

von Christierson B, Vidal J-P and Wade S (2012) Using UKCP09 probabilistic climate information for UK water resource planning, Journal of Hydrology 424-425:48-67, DOI:10.1016/j.jhydrol.2011.12.020.

von Christierson B, Wade S, Counsell C, Arnell A, Charlton M, Prudhomme C, Hannaford J, Lawson R, Tattersall C, Fenn C and Ball A (2013) Climate change approaches in water resources planning – overview of new methods, Report – SC090017/R3, Environment Agency. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/291598/LIT_7764_ea1e43.pdf.

Ciullo A, Martius O, Strobl E, Bresch DN (2021). *A framework for building climate storylines based on downward counterfactuals: The case of the European Union Solidarity fund*, Climate Risk Management, Volume 33, 2021, 100349, ISSN 2212-0963. https://doi.org/10.1016/j.crm.2021.100349.

Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate . Copernicus Climate Change Service Climate Data Store (CDS), date of access. https://cds.climate.copernicus.eu/cdsapp#!/home.

Coppola, E. et al. (2020) A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean. Clim. Dyn. 55, 3–34.

Defra (2018) Water industry: Guidance to Ofwat for water bulk supply and discharge charges, Defra, April 2018. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/696389/ofwat-guidance-water-bulk-supply-discharge-charges.pdf.

Defra (2021) Drought: how water companies plan for dry weather and drought, Policy Paper, 24 March 2021. https://www.gov.uk/government/publications/drought-managing-water-supply/drought-how-water-companies-plan-for-dry-weather-and-drought.

Deser, C., Hurrell, J.W. & Phillips, A.S. The role of the North Atlantic Oscillation in European climate projections. Clim Dyn 49, 3141–3157 (2017). https://doi.org/10.1007/s00382-016-3502-z.

Ebisuzaki, W. L. N. A. Y. & Kalnay, E. (1991) Ensemble experiments with a new lagged average forecasting scheme. WMO Research Activities in Atmospheric and Oceanic Modeling Rep 15, 308.

Environment Agency (2020). Meeting our Future Water Needs: a National Framework for Water Resources. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/873100/National_Framework_for_water_resources_summary.pdf.

Eyring V, Bony S, Meehl GA, et al (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 9:1937–1958. https://doi.org/10.5194/gmd-9-1937-2016.

Hagelin, S. et al. (2017) The met office convective scale ensemble, MOGREPSUK. Q. J. Roy. Meteor. Soc. 143, 2846–2861.

Hazeleger, W., van den Hurk, B., Min, E. et al. Tales of future weather. Nature Clim Change 5, 107–113 (2015). https://doi.org/10.1038/nclimate2450.

Historic England, Climate adaptation report, 2022, https://historicengland.org.uk/research/results/reports/8614/ClimateChangeAdaptationReport, accessed 23/03/2022.

Historic Environment Scotland climate action plan 2020-2025, https://www.historicenvironment.scot/archives-and-research/publications/publication/?publicationId=94dd22c9-5d32-4e91-9a46-ab6600b6c1dd, accessed 24/03/2022.

Jack CD, Jones R, Burgin L, Daron J, Climate risk narratives: *An iterative reflective process for co-producing and integrating climate knowledge*, Climate Risk Management, Volume 29, 2020, 100239, ISSN 2212-0963, https://doi.org/10.1016/j.crm.2020.100239.

Jacob, D., Petersen, J., Eggert, B. et al. EURO-CORDEX (2014). New high-resolution climate change projections for European impact research. Reg Environ Change 14, 563–578 (2014). https://doi.org/10.1007/s10113-013-0499-2 .

Liu Y, Kalverla P, Alidoost F, et al (2021) Pre-processed data of atlas in EUCP-WP2.

Lowe, J. A., Bernie, D., Bett, P., Bricheno, L., Brown, S., Calvert, D., Clark, R., Eagle, K., Edwards, T. & Fosser, G. 2018. UKCP18 science overview report. Met Office Hadley Centre: Exeter, UK.

Masson-Delmotte, V. et al. IPCC, 2021: Climate change 2021: The physical science basis contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change.

McKee T B, Doeskin NJ and Kieist J (1993). The Relationship ofDrought Frequency and Duration to Time Scales. Proc. 8thConf. on Applied Climatology, January 17-22, 1993, AmericanMeteorological Society, Boston, Massachusetts, pp. 179-184.

McSweeney, C. and Thornton, H. (2020) *UKCP European Circulation Indices: Weather Patterns*. UKCP Factsheet. Met Office. https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/research/ukcp/ukcp18_factsheet_weather_patterns.pdf

Met Office; Hollis, D.; McCarthy, M.; Kendon, M.; Legg, T.; Simpson, I. (2018): *HadUK-Grid gridded and regional average climate observations for the UK*. Centre for Environmental Data Analysis, 01/06/2021.

Murphy, J., Harris, G., Sexton, D., Kendon, E., Bett, P., Clark, R. & Yamazaki, K. (2019). UKCP18 land projections: Science report. Met Office. https://www.metoffice.gov.uk/pub/data/weather/uk/ukcp18/sciencereports/UKCP18-Land-report.pdf

Neal, R., Fereday, D., Crocker, R. and Comer, R.E. (2016), A flexible approach to defining weather patterns and their application in weather forecasting over Europe. Met. Apps, 23: 389-400. https://doi.org/10.1002/met.1563.

Niculiţă M (2020) Landslide Hazard Induced by Climate Changes in North-Eastern Romania. In: Climate Change Management. pp 245–265.

Ofwat (2020) PR24 and beyond: Future challenges and opportunities for the water sector, Ofwat, Birmingham. https://www.ofwat.gov.uk/wp-content/uploads/2020/12/PR24-and-beyond-Future-challenges-and-opportunities-for-the-water-sector.pdf.

Ofwat, Environment Agency and DWI (2021) The regulatory and commercial framework for strategic water resource solutions – a consultation, Regulators' Alliance for Progressing Infrastructure Development. https://www.ofwat.gov.uk/wp-content/uploads/2021/12/RAPID-Autumn-2021-condoc.pdf.

Palmer TE, Booth BBB, McSweeney CF (2021) How does the CMIP6 ensemble change the picture for European climate projections? Environ Res Lett 16:094042. https://doi.org/10.1088/1748-9326/ac1ed9

Philip SY, SF Kew, K van der Wiel, N Wanders, GJ van Oldenborgh (2020). Regional differentiation in climate change induced drought trends in the Netherlands. Environmental Research Letters, 15, pp. 094081.

Philip, S. Y., Kew, S. F., van Oldenborgh, G. J., Anslow, F. S., Seneviratne, S. I., Vautard, R., Coumou, D., Ebi, K. L., Arrighi, J., Singh, R., van Aalst, M., Pereira Marghidan, C., Wehner, M., Yang, W., Li, S., Schumacher, D. L., Hauser, M., Bonnet, R., Luu, L. N., Lehner, F., Gillett, N., Tradowsky, J., Vecchi, G. A., Rodell, C., Stull, R. B., Howard, R., and Otto, F. E. L.: Rapid attribution analysis of the extraordinary heatwave on the Pacific Coast of the US and Canada June 2021, Earth Syst. Dynam. Discuss. [preprint], https://doi.org/10.5194/esd-2021-90, in review, 2021.

Phillips ID and Mcgregor GR (1998). The utility of a drought index for assessing the drought hazard in Devon and Cornwall, South West England, Meteorological Applications, 5 (1998), pp. 359-372.

Pichelli E, Coppola E, Sobolowski S, et al (2021) The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2: historical and future simulations of precipitation. Clim Dyn 56:3581–3602. https://doi.org/10.1007/s00382-021-05657-4.

Prein, A. F. et al. (2015) A review on regional convection permitting climate modeling: demonstrations, prospects, and challenges. Rev. Geophys. 323–361.

Prein, A. F., Rasmussen, R. M., Wang, D. & Giangrande, S. E. (2021) Sensitivity of organized convective storms to model grid spacing in current and future climates. Phil. Trans. R. Soc. 379, 20190546.

Ribes A, Qasmi S, Gillett NP (2021) Making climate projections conditional on historical observations. Sci Adv 7:1–10. https://doi.org/10.1126/sciadv.abc0671.

Ribes A, Zwiers FW, Azaïs JM, Naveau P (2017) A new statistical approach to climate change detection and attribution. Clim Dyn 48:367–386. https://doi.org/10.1007/s00382-016-3079-6.

Rodda JC and Marsh TJ, 2011.The 1975-76 Drought - a contemporary and retrospective review. Wallingford, UK, NERC/Centre for Ecology & Hydrology, 58pp.

Schär, Christoph, et al. (1996) Surrogate climate change scenarios for regional climate models. Geophys. Res. Lett. 23.6: 669-672.

Schellander-Gorgas, T. et al. (2017) On the forecast skill of a convection-permitting ensemble. Geosci. Model Dev. 10, 35–56.

Schurer A., et al. (2022) Role of long term variability of the North Atlantic Oscillation on Northern Hemisphere climate, in prep.

Sexton et al., *Describing future UK winter precipitation in terms of changes in local circulation patterns*, In review.

Shepherd, T.G., Boyd, E., Calel, R.A. et al. (2018). Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. Climatic Change 151, 555–571. https://doi.org/10.1007/s10584-018-2317-9.

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

Smith DM, Eade R and Pohlmann H (2013) A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction. Clim Dyn 41, 3325–3338 (2013). https://doi.org/10.1007/s00382-013-1683-2.

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93:485–498. https://doi.org/10.1175/BAMS-D-11-00094.1.

Undorf S., Bollasina M. A., Hegerl G.C. (2018): Impacts of the 1900-74 Increase in Anthropogenic Aerosol Emissions from North America and Europe on Eurasian Summer Climate. J Climate 31, 8381-8399.van der Wiel K, N Wanders, FM Selten, MFP Bierkens (2019): Added value of large ensemble simulations for assessing extreme river discharge in a 2 °C warmer world. Geophysical Research *Letters*, 46, pp. 2093-2102.

Wyser, K., Koenigk, T., Fladrich, U., Fuentes-Franco, R., Karami, M. P., and Kruschke, T.: The SMHI Large Ensemble (SMHI-LENS) with EC-Earth3.3.1, Geosci. Model Dev., 14, 4781–4796. https://doi.org/10.5194/gmd-14-4781-2021.

Zappa, G., Shepherd T. G., (2017) Storylines of atmospheric circulation change for European Regional climate impact assessment. J Clim 30:6561–6577. doi: https://doi.org/10.1175/JCLI-D-16-0807.1.

Zelinka MD, Myers TA, McCoy DT, et al (2020) Causes of Higher Climate Sensitivity in CMIP6 Models. Geophys Res Lett 47:1–12. https://doi.org/10.1029/2019GL085782.

Zscheischler J, EM Fischer (2020): The record-breaking compound hot and dry 2018 growing season in Germany. Weather Climate Extremes, pp. 100270.

6 Appendices to individual sections

6.1 Appendix to 3.2.2 (Heritage storylines case study)

6.1.1 Table of values for the high, median and low storylines (and full range).

The tables below show various measures of future change for the JJA Tmax and DJF Pr hazards. The values shown are for the low, median and high members for each hazard (selected based on the anomaly in JJA Tmax90d and DJF Pr in Scotland). The full range of anomalies in each measure across all the 15 PPE and 11 CMIP5 models is also shown (the two CMIP5 models without WT data were not included). This allows the impacts of the selection being made on the Tmax90d anomaly and Pr anomaly to be seen on the other measures via comparison to the full range. For JJA Tmax all values are for Scotland, for DJF Pr the regional variation within Scotland is also shown, for West Scotland, East Scotland and North Scotland (WS, ES, NS). The time to reach certain thresholds was not included for DJF Pr as the increased multi-decadal variability in the hazard measures compared to JJA Tmax makes this less meaningful. The size of the weather pattern frequency and climatology components are only shown for DJF Pr as the climatology component dominates for JJA Tmax.

Cases where the full range of some of the metrics is not captured by the selected storylines are apparent. For example, due to different regions of Scotland seeing maximum DJF Pr anomalies in different ensemble members the Scotland wide selection misses an additional 5% anomaly in the west and east. Another example is the anomaly in the Pr99d threshold, where the upper and lower ends are not captured by the selection, independent of location. Assessing the impact of sub-selection in this manner is important for scientific understanding, as well as allowing the user to judge if the selection is appropriate.

Measure (JJA Tmax)	Measure units	Region	Low	Median	High	Full range
Member			27	23	14	
Anomaly	° C	Scotland	+ 1.5	+ 2.9	+ 4.7	+ 1.2 to + 4.7
Seasonal maximum anomaly	° C	-	+ 1.3	+ 5.6	+ 6.2	+ 0.5 to + 7.2
Tmax90d anomaly	Days	-	+ 9.3	+ 25.0	+ 44.8	+ 9 to + 45
Tmax99d anomaly	Days	-	+ 1.5	+ 6.0	+ 17.4	+ 0.5 to + 18.5
Tmax90 anomaly	° C	-	+ 1.4	+ 4.8	+ 6.0	+ 1.3 to + 6.2
Tmax99 anomaly	° C	-	+ 1.9	+ 6.0	+ 7.0	+ 0.23 to + 7.9
Intra-season variance anomaly	%	-	+ 1	+ 100	+ 62	-14 to + 100
Inter-season variance anomaly (detrended)	%	-	- 41	+ 65	+ 56	- 41 to + 202
Inter-season variance anomaly	%	-	- 40	+ 78	+ 98	- 40 to + 229
Annual Tas anomaly	° C	Scotland	+ 1.7	+ 2	+ 3.5	+ 1.2 to + 3.8
Annual Tas anomaly	° C	Global	+ 2.1	+ 2.8	+ 3.6	+ 1.79 to + 4.0
Timing thresholds						
30 hot days per summer		Scotland	N/A	2060	2045	
5 very hot days per summer		Scotland	N/A	2060	2030	
+ 2 ° C (Tmax)		Scotland	2080	2050	2040	

Table 7 Values for various measures of JJA Tmax hazards in the projections, including the H/M/L values (selected based on the Tmax90d anomaly), and the full range across all members independent of the selection.

Measure (DJF Pr)	Measure units	Region	Low	Median	High	Full range
Member			27	3	23	
Anomaly	%	Scotland	- 4.5	+ 9.9	+ 20.5	- 4.5 to + 20.5
		WS	- 3.6	+ 15.5	+ 20.9	- 3.6 to + 26.5
		ES	- 3.5	+ 2	+ 19.9	- 3.5 to + 24.8
		NS	- 6.3	+ 13.4	+ 20.8	- 6.3 to + 22.4
Pr90d anomaly	Days	Scotland	+ 0.2	+ 1.6	+ 3.6	+ 0.2 to + 3.8
		WS	+ 0.5	+ 2.3	+ 4.1	+ 0.5 to + 5.0
		ES	+ 0.1	+ 0.6	+ 3.5	+ 0.1 to + 4.2
		NS	- 0.1	+ 2.1	+ 3.5	- 0.5 to + 4.0
Pr99d anomaly	Days	Scotland	+ 0.7	+ 0.6	+ 0.9	+ 0.1 to + 1.3
		WS	+ 0.6	+ 0.9	+ 0.8	+ 0.2 to + 1.8
		ES	+ 0.8	+ 0.4	+ 0.9	+ 0.1 to + 1.4
		NS	+ 0.7	+ 0.5	+ 1.2	- 0.2 to + 1.3
Pr90 anomaly	%	Scotland	+ 0.8	+ 9.3	+ 20.4	+ 1.2 to + 24.2
		WS	+ 2.3	+ 11.8	+ 22.3	+ 2.3 to + 24.2
		ES	+ 0.6	+ 4.1	+ 20.3	+ 0.6 to + 26.2
		NS	- 2.7	+ 11.2	+ 18.8	- 2.7 to + 27.2
Pr99 anomaly	%	Scotland	+ 14.2	+ 12.6	+ 16.3	+ 1.7 to + 24
		WS	+ 10.7	+ 18	+ 12.1	+ 3.7 to + 26.5
		ES	+ 16.8	+ 8.4	+ 14.5	+ 2.6 to + 27.7
		NS	+ 14.9	+ 11.1	+ 23	-5.6 to + 26.4
Intra-season variance anomaly	%	Scotland	+ 22.6	+ 23	+ 35.4	+ 7.4 to + 49.1
Inter-season variance anomaly (detrended)	%	-	+ 67.5	+ 37.0	- 2.7	-52.4 to 102.9
Inter-season variance anomaly	%	-	+ 65.2	+ 33.4	- 0.5	-50.1 to + 97.8
Tas anomaly	° C	Scotland	+ 1.7	+ 2.8	+ 2.0	+ 1.2 to + 3.8
Tas anomaly	° C	Global	+ 2.1	+ 3.8	+ 2.8	+ 1.8 to + 4.0
Change components						
WT frequency	%	Scotland	- 2.3	+ 4.7	+ 4.0	- 6.3 to + 9.9
WT climatology	%	Scotland	-1.2	+ 5.2	+ 16.3	- 5.5 to + 24.5

Table 8 Values for various measures of DJF Pr hazards in the projections, including the H/M/L values (selected based on the Pr anomaly averaged over Scotland), and the full range across all members independent of the selection. Here the different Scottish regions are included as there is significant regional variation.

6.1.2 Further details of the bias and process analysis

Metric/variable representation:

It is important that representative model runs chosen from the ensembles can represent relevant climatological features in the region of interest. One example is the difference in winter rainfall totals and thresholds between the north/west of Scotland and the east (Figure 3.2-2, middle and right), which is due to the interaction of orography with prevailing rain bringing westerly weather patterns. All CMIP5-13 models included in UKCP were regridded to the 60km grid of the PPE GCMs, however the underlying data often has a far lower resolution, which may contribute to their inability to capture this Scottish rainfall pattern, as well as larger scale drivers. Of potential concern for our application are members 16, 19 and 25. This is less of an issue for JJA Tmax analysis where there is a more uniform climatology (Figure 3.2-2, top left) and pattern of future change. Biases in the percentile thresholds are shown in the top row of Figure 6.1-2. There are no systematic differences between the PPE and CMIP members for JJA Tmax over Scotland, and most models show a cool bias in the thresholds. DJF Pr thresholds show a consistently dryer bias in CMIP members than the PPE members, with the PPE bias improving when west Scotland values are plotted. Since all analysis will use relative thresholds and anomalies the importance of these underlying biases is not clear.

The variability of the seasonal metrics, the anomaly to baseline and number of days above the percentile thresholds, is also important to consider. In Figure 3.2-4 the anomalies for the 30 seasons in the baseline period are shown for our measures of DJF Pr90d and JJA Tmax90d. These, and other regions and measures, show that the spread of seasonal anomaly values is generally well represented across the ensembles. This information may be useful when assessing the robustness of individual members and their projected changes. Events of interest and extremes in this context are seasons that deviate from the baseline strongly by these measures. Recent examples are 2018 for JJA Tmax, where the seasonal Tmax anomaly was +2.2 C, and the number of Tmax90d and Tmax99d days were 18 and 3 respectively. Other past summers far exceeded these measures, with 1995 having the most Tmax90d days, however a more recent summer was deemed a better choice by the user. For DJF rainfall, 2016 had a Pr anomaly of + 59 %, and 18 and 4 Pr90d and Pr99d days across Scotland.

Weather pattern representation: frequency (and variance) biases.

The seasonal DJF WT frequencies, and the inter-annual variance from the UKCP GCMs and the ERA5 reanalysis were compared, see Figure 6.1-1. There are no clear outliers across ensemble members, with different members being further from the main cluster for different weather patterns. The inter-annual DJF variance of WT2 is lower than ERA5 for all members, which may be linked to an underrepresentation of NAO variability. Most members have a positive frequency bias for WT7 (responsible for a large relative proportion for winter precipitation), which is generally stronger in the CMIP5-11 members. Additionally, most of the CMIP5-11 members have a negative frequency bias for WT8. The plots for JJA show the PPE-15 appears to have a stronger negative bias for WT6 (UK centred high), and a positive rather than negative bias for WT3, compared to the CMIP-11 members. There are no clear differences in the variance, and so the seasonal variability of the weather patterns, except where it appears to correlate, or scale, loosely with the frequency.

Weather pattern representation: Relative climatology and variability

To analyse and understand future changes using weather pattern analysis requires an awareness of how well the relative climatology (frequency normalised contributions) between different patterns is represented in the baseline period. In the lower panels of Figure 6.1-2 an example for DJF rainfall over Scotland is shown. We see that there are notable differences depending on the weather type, which also vary spatially. For example, WT1, 3, and 6 are significantly drier relative to the other weather patterns than in the observations, while WT4 is significantly wetter, especially the CMIP5-11 members. There are some regional differences, with the dry bias mentioned above being higher for west Scotland for example. The same analysis of JJA Tmax anomalies and hot days shows WT5 contributing up to three times more hot days than in the observations, and up to a +2 degree bias in the relative temperature. WT2, 3, and 8 all contribute a notably lower proportion of summer hot days. While much of the rainfall climatology biases may be due to orographic and resolution/parameterisation effects, the summer temperature climatologies may also be affected by biases in the persistence of weather patterns.

The weather type frequency and its seasonal anomaly, and the relative climatology can be used to assess to what extent weather pattern variability contributes to the variability of seasonal variables and metrics, as well as individual seasons. This can also be used to assess the representation of weather patterns and variability of these measures in the models, and individual ensemble members. Using ratios of standard deviations, weather type frequency variability accounts for 34% of DJF Pr variability in Scotland using the observed data and ERA5 weather patterns, and 33% of the Pr90d metric. The value for JJA hot days is only around 10%. For the GCM members during the baseline period, the fraction of precipitation from WT frequency variability is generally much higher, up to a factor of two in some cases. However, the overall variability of the seasonal variables and metrics is well represented, as discussed above and shown in Figure 3.2-4, meaning this is likely countered due to other sources of variability being under represented.

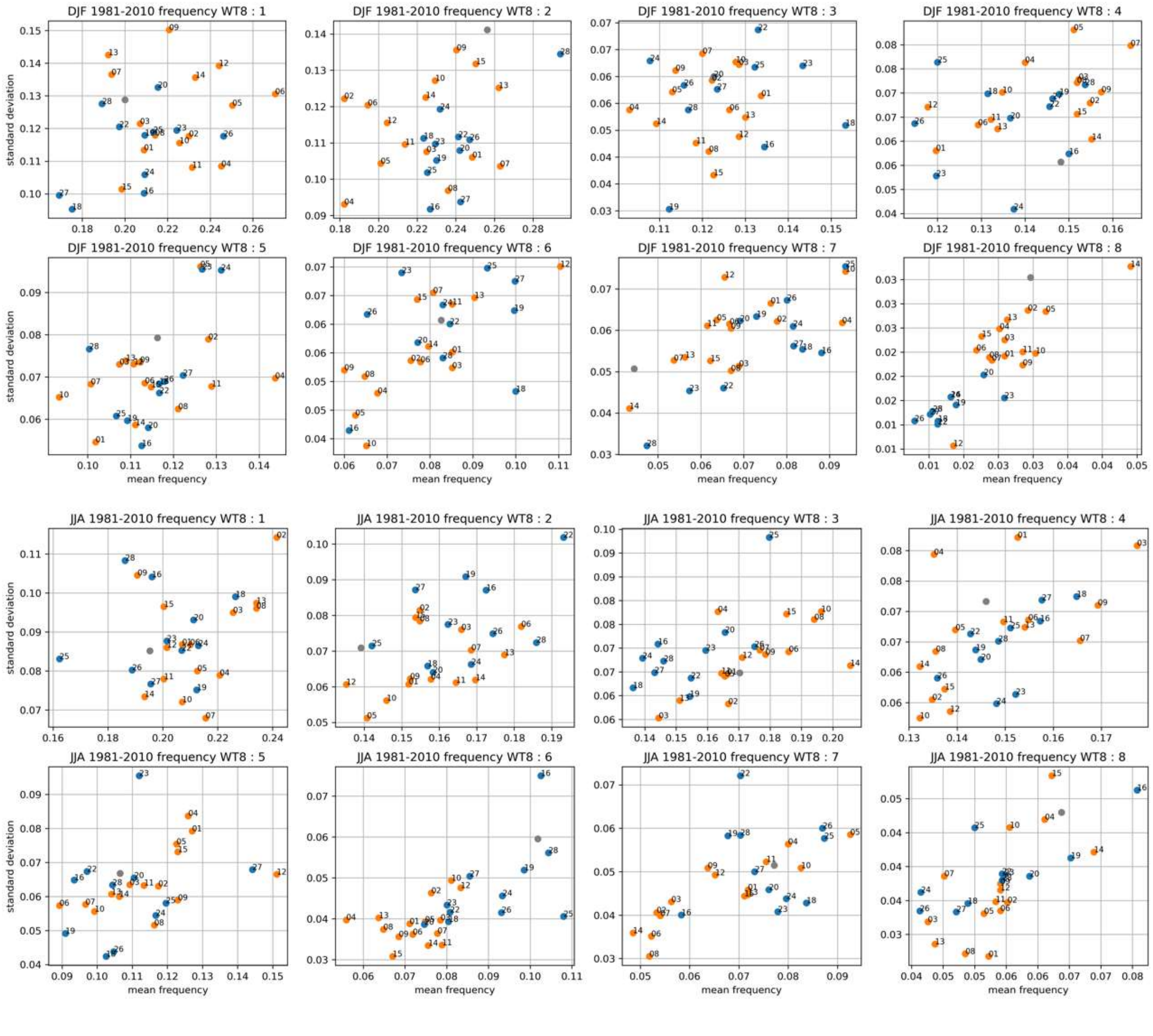


Figure 6.1-1 Top two rows: DJF WT8 baseline standard deviation against mean frequency for each weather pattern (panels), CMIP-13 MME members in blue, UKCP PPE-15 members in orange and HadUK-grid values in grey (points), allowing the bias to be discerned for each ensemble member. Top bottom rows: as above for JJA.

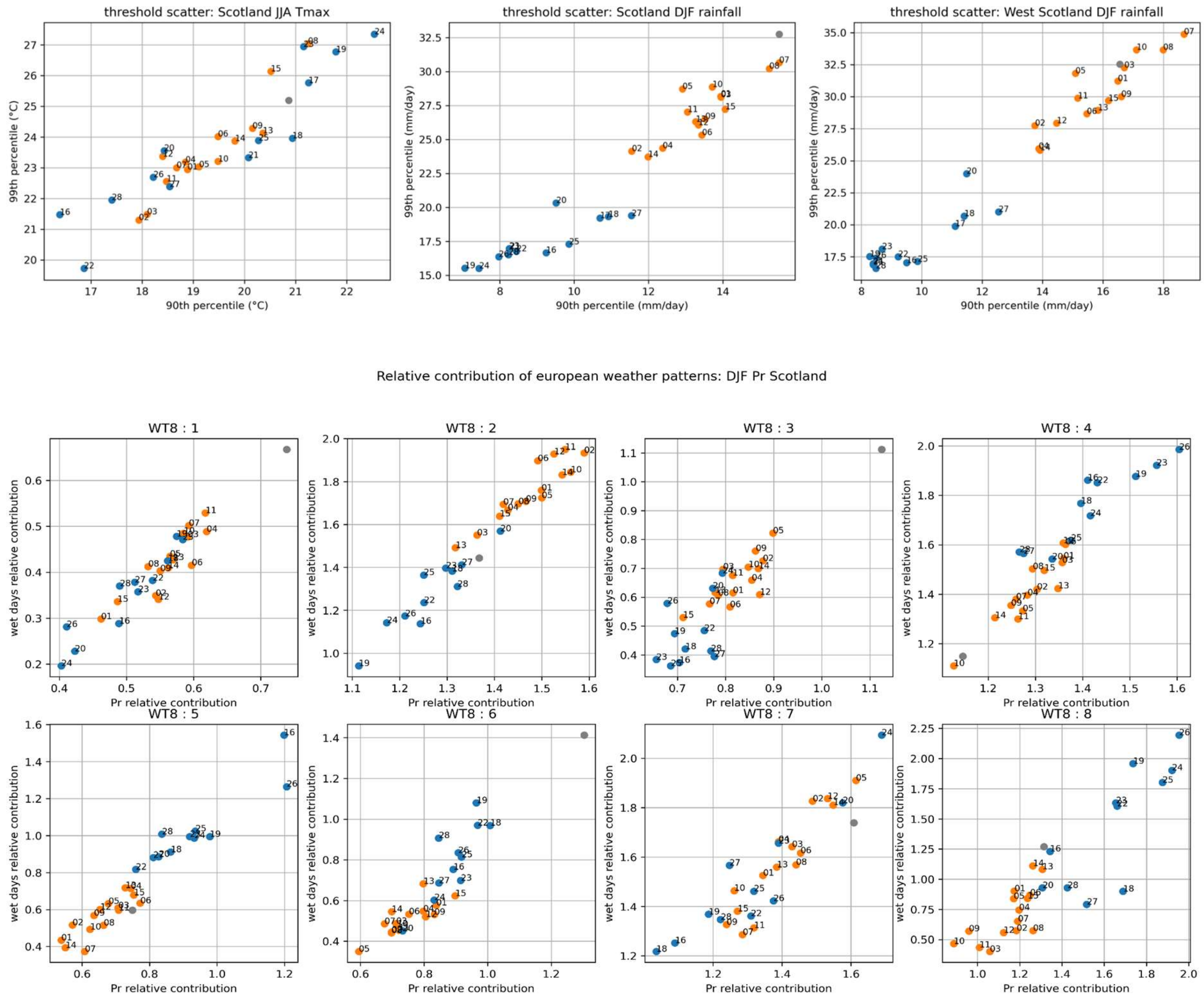


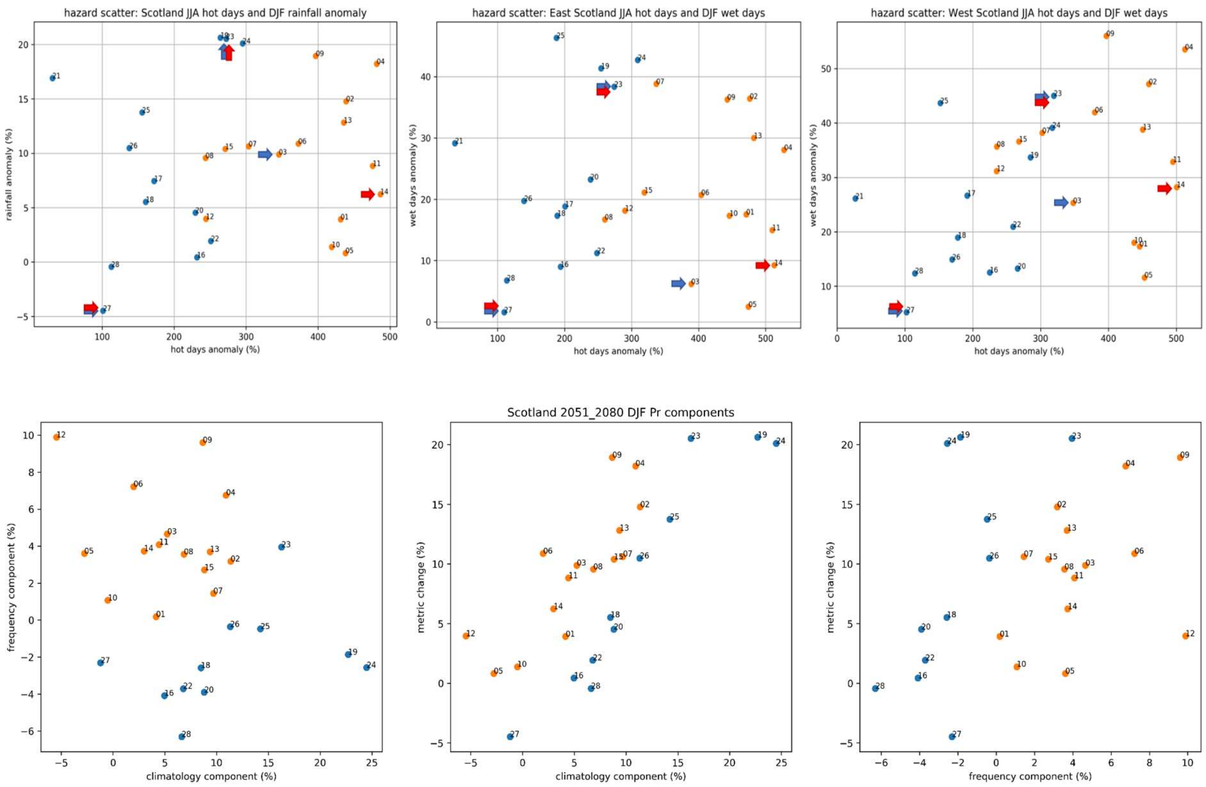
Figure 6.1-2 Top panels: 90th and 99th percentile threshold values for JJA Tmax (Scotland), and DJF Pr (Scotland, West Scotland). CMIP-13 MME members in blue, UKCP PPE-15 members in orange and HadUK-grid values in grey (allowing the bias for each ensemble member to be seen). Bottom 8 panels: WT baseline relative climatologies (frequency normalised contribution to seasonal totals) for DJF rainfall and the wet days (Pr90d) metric. The coloured symbols are as above, with the observations shown again to allow the bias to be assessed.

6.1.3 Further details projection information selection

Further comments on individual ensemble members:

As discussed above, various biases were analysed related to the spatial representation, variability and weather pattern climatology of the hazard metrics used. On a per-model or ensemble member basis these become important to understanding differences in process representation, and potentially flag models to be avoided in the selection. For example, A few CMIP members were avoided in the DJF Pr selection due to their low spatial resolution (16, 19 and 25) meaning differences between East and West Scotland were not represented but didn't significantly change the range presented (and the full range was still communicated). Since these models have been extensively analysed elsewhere there are also other lines of evidence to draw on, such as PPE-15 members 11 and 12 being identified as examples of early AMOC collapse (Murphy et al., 2019). This information is also important when identifying and explaining differences between the selected members in their future behaviour. For example, for JJA Tmax in Figure 3.2-6, member 27 (low storyline) is unique in that it has a relatively high negative frequency component, but low degree of global warming (many PPE members have a similar frequency component, but higher warming and so the frequency component is insignificant). The lower panel of Figure 3.2-6 showed how the weather pattern component changes were not at all consistent across the selected models.

Initial analysis of the 12 UKCP-PPE members for which downscaled RCM simulations are available was also performed. For the percentile-based metrics this indicated that the increases in wet and very wet days are generally slightly lower following the downscaling, in line with the lower rainfall anomalies in winter (Murphy et al., 2019). The hot and very hot day metrics generally showed a one-to-one scaling, although for both sets of hazards there was scatter which would 'shuffle' the order of the ensemble members, so member selection across the 12 GCM and downscaled RCM members would not be consistent, however the range of uncertainty covered would be similar. These findings are in line with the results of the lines of evidence assessment for the UK and Scotland reported in Appendix 6.2.1. The 12 RCMs and parent GCMs cover most of the uncertainty space in the winter rainfall metrics from the full 28 models, however this is not the case for summer temperatures due to the lower climate sensitivities in the CMIP-13 being missed.



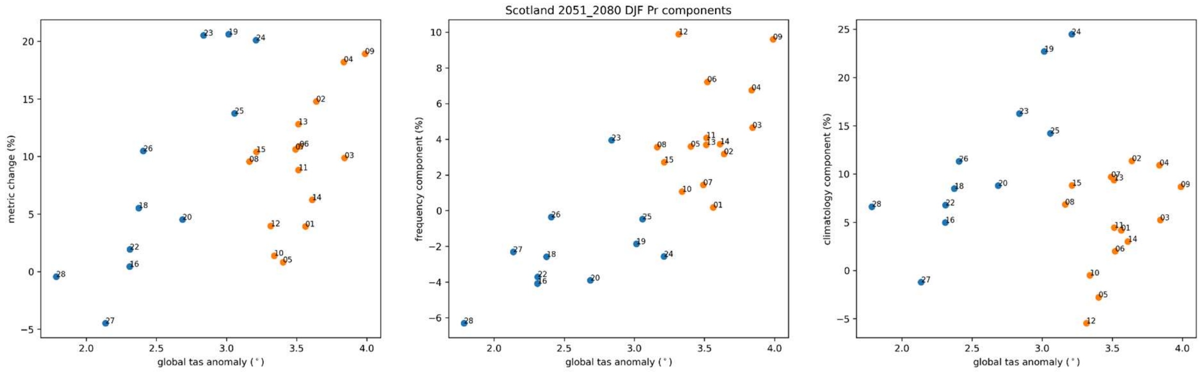


Figure 6.1-3 Top panels: cross-hazard plots for JJA hot days and DJF wet days over Scotland, East Scotland and West Scotland, with the high, median, and low members selected for each highlighted by the arrows. Middle panels: the weather pattern climatology and frequency components of future change for DJF Pr over Scotland, and how the total change scales with both. Bottom panels: as above showing how the metric change and components scale with the global surface temperature anomaly of each member.

Further comments on selection options:

For this application a starting point for selecting representative members is to assume the full range of uncertainty may be relevant (unless considered implausible), so to begin by selecting high, median and low members for each hazard to present to users, and work through the iterative co-production process. This approach still requires decisions to be made on the time horizon, and the particular hazard metric on which to make this selection. In this case testing the relevance of the uncertainty is an aim, and the DJF Pr and JJA Tmax90d anomalies were used for the selection. This is because 'easy to communicate' selections showing the full range were requested, and this still samples the uncertainty range of most other hazard measures well (except very wet days and thresholds, Pr99d). Regional differences in DJF Pr responses are seen, however there was no desire to generate multiple storylines at this stage, so whole-Scotland values were used to make the selection and full ranges for each region are reported alongside those for the selected members. A full table of values from the selected simulations, alongside the full range, is provided in Appendix 6.1.1.

Selection across both DJF Pr and JJA Tmax hazards is explored in the top panels of Figure 6.1-3. The H/M/L members are shown by the member IDs and additionally by the red and blue arrows. The expected inconsistency is apparent, such as member 23 being the high DJF Pr member, but median JJA Tmax member, and there is regional variation on top of this. However, it would be possible to make a consistent selection across both hazards maximising the range of each if required for a given application, however this was not desired in this case. The bottom panels of explore the potential to group the members by the weather pattern components of future change, i.e. by the dynamic and thermodynamic drivers of change. For the JJA Tmax hazard metrics over Scotland, driver-led storyline formation would involve grouping by global temperature anomalies as the WT frequency component is not significant. Conversely, the middle-left panel shows that for DJF Pr the WT frequency component is significant, however, taking the mean or median of each group would not lead to values that are sufficiently distinct in the hazard metric space, especially when the goal is to communicate and explore the relevance of the full uncertainty range, and so this option was not taken forwards.

6.1.4 Further details of the application example, feedback and findings

A prototype product was produced to allow discussion with HES and the wider focus group as it was found that making progress in discussion was extremely difficult without a concrete example. Some details were given in Section 3.2.2, with an example excerpt in Figure 3.2.9. The information provided in the future storylines was based on the information in Table 7 and Table 8, alongside information from the observations and weather patterns from the ERA5 reanalysis. A full list of the utilised information products is as follows;

Information used in the prototype

- Past time series of metrics and variables from HadUK-grid.
- Observed thresholds and variable values during the baseline period (maps and area average).
- Individual past seasons (events of interest), as anomalies to baseline.
- Future changes to variables and metrics. Example future seasons.
- Full uncertainty range in measure used for the selection.
- Future global and Scotland Tas anomalies.
- Redefined thresholds and variable means for future period as anomalies.
- An example time series for the median storylines covering the past and future period.
- Timing to reach certain metric thresholds, such as a mean of 5 extra very hot days per summer.
- Weather pattern baseline climatologies (DJF Pr would also use the future changes).
- Full range of Global Tas uncertainty over time for context.
- Changes to daily and seasonal variability of the underlying variable.
- Regional variation and full ranges of the value presented (independent of the storylines).
- Difference in the importance of WT frequency changes between the two hazards.

Information used in the analysis

- WT mean occurrence frequency and variance bias
- WT metric climatology bias
- Spatial bias of underlying hazard-related variable
- WT climatology & frequency changes and resulting changes to hazard metrics

Specific application feedback

Direct feedback to summary questions from HES is quoted below, which confirmed and underpinned some of the application related findings reported in Section 3.2.2;

What are your thoughts on the potential for this style of climate information provision? '*I'd say there is high potential for this to be used as an engagement tool – showcasing the complexity of climate change projections in a way that logically teases out that complexity. It starts from a strong place, using an observed event and setting that as the baseline. This means there's an opportunity to start from a 'place' that is meaningful to users. It equips someone who already has a decent foundational understanding of climate science / projections etc with tools/data to better communicate the uncertainty.*'

What have been the drawbacks, benefits, or lessons learned from the process of producing the information? '*The process, from my perspective, has been informative / educational. The direct* *contact with colleagues at MO and being led through the process has been really valuable. It feels that the products coming out toward the end of the process have been shaped by input right at the beginning of the process. It was difficult to picture what the final outputs would be during the early stages, but it's become easier as the process has gone on. The only major drawback I can see is that for an* [organisation] *to be able to take on a storyline/narrative produced like this, they need to have people with that foundational knowledge of climate science/projections.*'

Have there been any changes to your perception or thoughts on climate information and its application during the process? '[Absolutely], *I knew a limited amount about uncertainty in the climate projections, and the ranges of possible future climate, coming into the process. But as more of the data has been presented, and as we've talked through how the data was produced, the difficulties in predicting future climate have become somewhat clearer, as has some of the science behind the projections.*'

Following a co-production process ensured there was an understanding of the potential application of the storylines product for HES from the start of the analysis and prototyping, and also allowed this to evolve over time as the users' understanding of the project and potential outcomes evolved. The dialogue appears to have been important in creating a sense of shared ownership over the outcomes, and also equipping the stakeholder to further communicate the information and engage with climate services in the future. The ongoing dialogue also prompted further work in the organisation, HES, including the development of a climate risk appetite statement or strategy, and a desire for more research into current risk and decision relevant metrics and trigger points. Climate storylines also feature in the recently developed 'Climate Ready HES' action plan.

Heritage Sector Focus group participants and questions

The focus group participants and some details of their role in the cultural heritage sector are listed in the table below. In general, this group confirmed the wider applicability of the findings from working with HES, however some differences also emerged, particularly in how the sector is organised in different European countries. HES is quite unique as they act as the governmental advisory, research and policy organisations, as well as having a role in maintaining and running a large number of properties directly. Focus group feedback contributed to the findings summarised in the application part of Section 3.2.2, and the further detail given below.

Organisation	***Organisation's role***	***Participant's role***
Historic Environment Scotland (HES)	Public body investigating, caring for and promoting the historic environment, running sites.	Climate change team member. Stakeholder for the prototype storylines.
Historic England	Government's statutory adviser and a statutory consultee, preservation and enhancement of Historic Environment.	Senior policy advisor, Heritage at risk strategy.
Swedish National Heritage Board	Government administrative agency, cultural heritage protection and management policy and advice.	Advisor, guidance and supervision.
Norwegian Institute for	Cultural heritage research institute.	Paintings Conservator and researcher.

Cultural Heritage Research		
Directorate for Cultural Heritage Norway	Governmental management on the national level.	Senior Climate Adviser.
SPSG Berlin-Brandenburg	Preserving and running a group of Heritage assets.	Scientific Associate, Horticulture and Gardens. Stakeholder involved in the KERES project (Heritage impacts in Germany).
GERICS/Fraunhofer institute	Climate services/research institute.	Climate scientist. Service provider in the KERES project.
Adaptation Scotland/Sniffer	Government programme for national level resilience and adaptation policy and guidance.	Climate adaptation services specialist.
ISAC-CNR/Copernicus	Academic research on Cultural Heritage.	Cultural Heritage impacts research/project lead.

Table 9 Organisations involved in the focus group discussions to review the storylines information and prototype developed for HES

Specific feedback on the information and choices in both sets of hazard storylines that was explore included;

- Are the hazard metrics used appropriate?
- Is the hazard signal presented relevant in planning and decisions? What about the difference between the presented storylines?
- Is the final selection appropriate? Does it limit the range of other metrics, or regional change, too much?
- Is the representation of uncertainty effective?
- Is the inclusion of driver information useful?
- Does the presentation of the information as 'storylines' or alternative futures, increase it's usefulness and useability?

Detailed notes, recordings and other information gathering exercises from HES and the focus group began to answer these questions and formed the basis for the findings reported in Section 3.2.

6.2 Appendix to section 3.4 (Lines of Evidence)

6.2.1 UK plots for the application case studies in Section 3.2.

Figure 6.2-1 shows multiple lines of evidence for UK 2041-2060 anomalies in JJA Tas, JJA Pr and DJF Pr, including downscaled projection plotted against the values from the parent GCM or RCM. These plots are relevant to the case studies in Section 3.2, as well as the lines of evidence section and discussion, where further description of the data sources, methods and plots are given in Section 3.4.

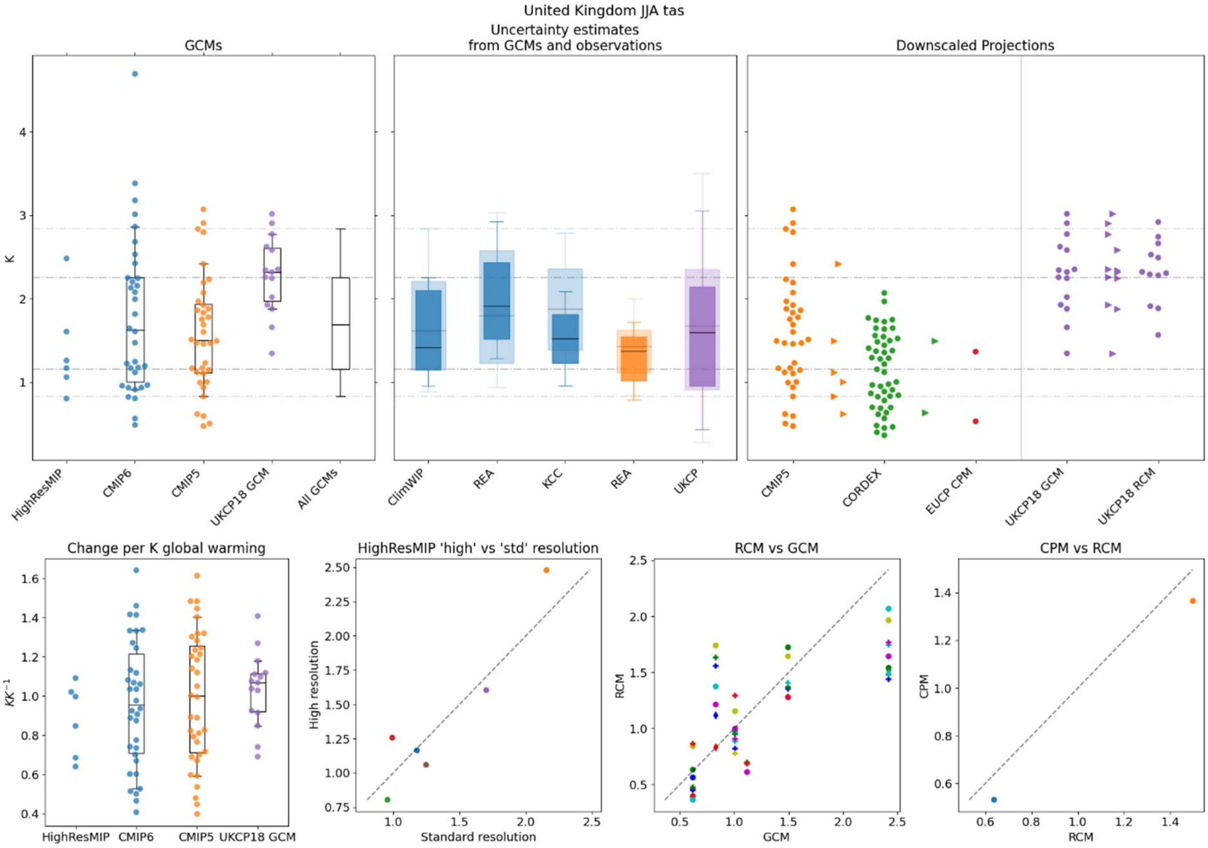
The top panels show that the available CORDEX-CPM members from EUCP span the lower end of the projected JJA tas changes for the UK, sitting below all of the UKCP RCM members. The GCM RCM scatter plot shows a degree of scatter, but centred about the 1 to 1 line without a significantly changed uncertainty range, although this varies for individual parent GCMs. The middle panels show that the UKCP GCM and RCM members don't cover the wetter end of the uncertainty space in CMIP5 and CORDEX, JJA Pr. The EUCP CORDEX-CPM members, however, span both ends of this uncertainty space.

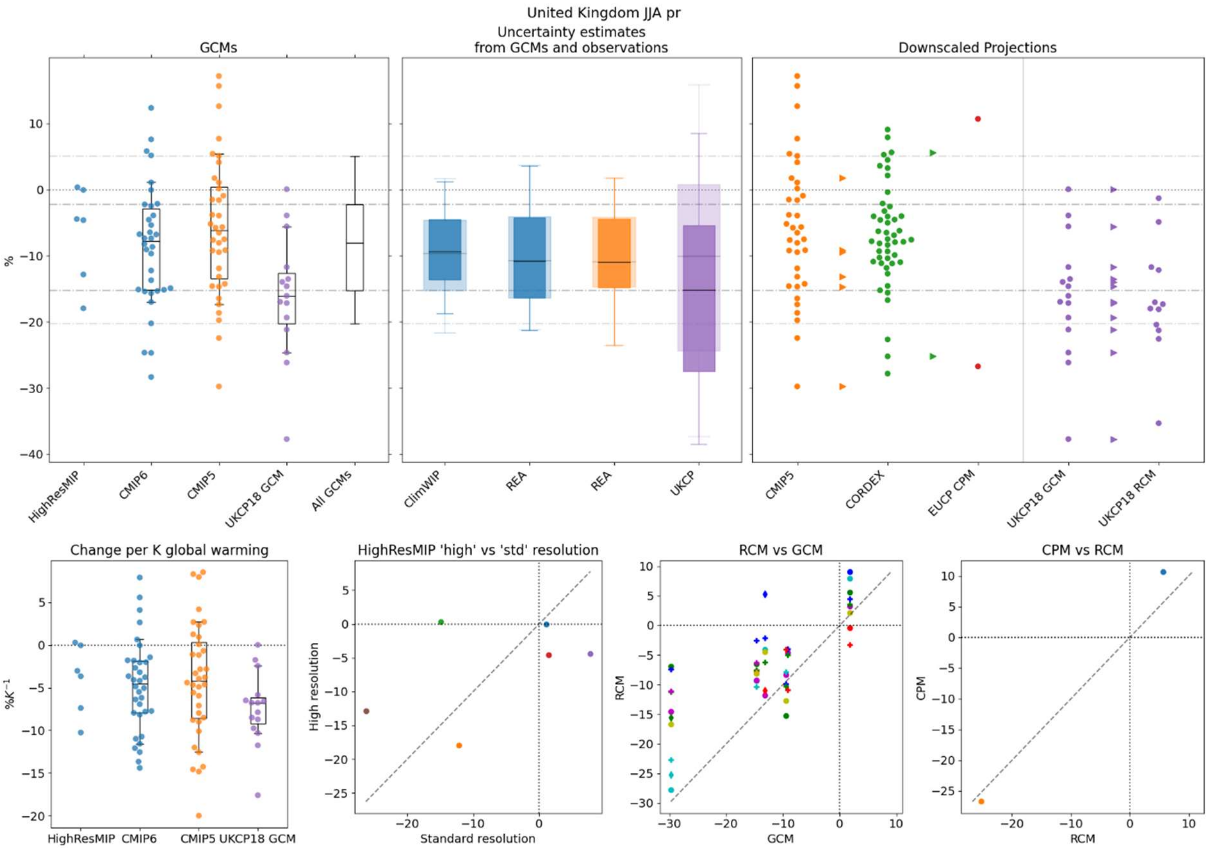
The GCM RCM scatter plot confirms that RCMs downscaled from GCMs generally give enhanced rainfall anomalies over the UK in summer. For DJF Pr the bottom panel shows that the UKCP GCM and RCM cover all but the wetter end of the changes in CMIP5 and CORDEX, and the two CORDEX-CPM runs both lay towards the upper end of the projected changes, and the scatter plot indicates RCMs downscaled from GCMs generally have reduced anomalies in winter.

In all cases the spread in values for each of the parent GCMs also shows the impact of using different models for downscaling, even with the same input boundary conditions. In general the constraints shown do not have a large impact on projected Pr ranges (see WP2 for further details), although the upper range of JJA tas is limited by some approaches. Finally, the sub-plots showing the anomalies scaled by the degree of global warming indicate that the different behaviour of the UKCP PPE GCMs is due to the higher climate sensitivity, rather than any differences in regionally relevant drivers, potentially with the exception of JJA Pr.

For the examples in Sections 3.2.2 and 3.2.3, the UK-wide lines of evidence plots set the UKCP PPE and CMIP5 GCMs in a wider uncertainty context in terms of the underlying climate variables, including the impact of using constraints. This kind of information may be used retrospectively to communicate this alongside the main data provision. The results show that the UKCP PPE-15 are at the upper end of summer temperatures changes projected, with the CMIP5-13 members filling in the lower end of the uncertainty space. The constraints methods are seen to slightly reduce the upper end of this range, shown via the 90th percentile. For winter rainfall all but the largest anomalies in CMIP5 and CORDEX are covered by the UKCP GCM/RCM range, whereas for summer rainfall the potential for increased mean rainfall is not represented by the UKCP GCM/RCM. The scatter plots comparing the downscaled models to their parent GCMs/RCMs confirms that at the country-scale the uncertainty range is not strongly affected, however downscaling does tend to lead to reduced winter rainfall anomalies, and increased summer rainfall anomalies. Finally, the panel showing the GCM anomaly scatter scaled by the degree of warming indicates whether conflicting information is due to differing regional responses, or climate sensitivity. The differences in the UKCP GCM response appear to be largely to do with higher climate sensitivities, perhaps with the exception of JJA Pr.

For future work this information may also be valuable at the initial stages of a climate service project, when selecting information sources for impact modelling or downscaling. This is discussed further in section 3.4, and in the discussion in 3.5, as it has relevance across several pieces of work in this deliverable. Further potential could be realised if combined with expert elicitation and physical analysis to understand the underlying causes behind the differences seen, and whether the uncertainty space can be reduced for a given location and application.


United Kingdom JJA tas
Uncertainty estimates
from GCMs and observations
GCMs
Downscaled Projections
K
HighResMIP
CMIP6
CMIP5
UKCP18 GCM
All GCMs
ClimWIP
REA
KCC
REA
UKCP
CMIP5
CORDEX
EUCP CPM
UKCP18 GCM
UKCP18 RCM
Change per K global warming
KK^{-1}
HighResMIP
CMIP6
CMIP5
UKCP18 GCM
HighResMIP 'high' vs 'std' resolution
High resolution
Standard resolution
RCM vs GCM
RCM
GCM
CPM vs RCM
CPM
RCM


United Kingdom JJA pr
Uncertainty estimates
from GCMs and observations
GCMs
Downscaled Projections
%
HighResMIP
CMIP6
CMIP5
UKCP18 GCM
All GCMs
ClimWIP
REA
REA
UKCP
CMIP5
CORDEX
EUCP CPM
UKCP18 GCM
UKCP18 RCM
Change per K global warming
$\%K^{-1}$
HighResMIP
CMIP6
CMIP5
UKCP18 GCM
HighResMIP 'high' vs 'std' resolution
High resolution
Standard resolution
RCM vs GCM
RCM
GCM
CPM vs RCM
CPM
RCM

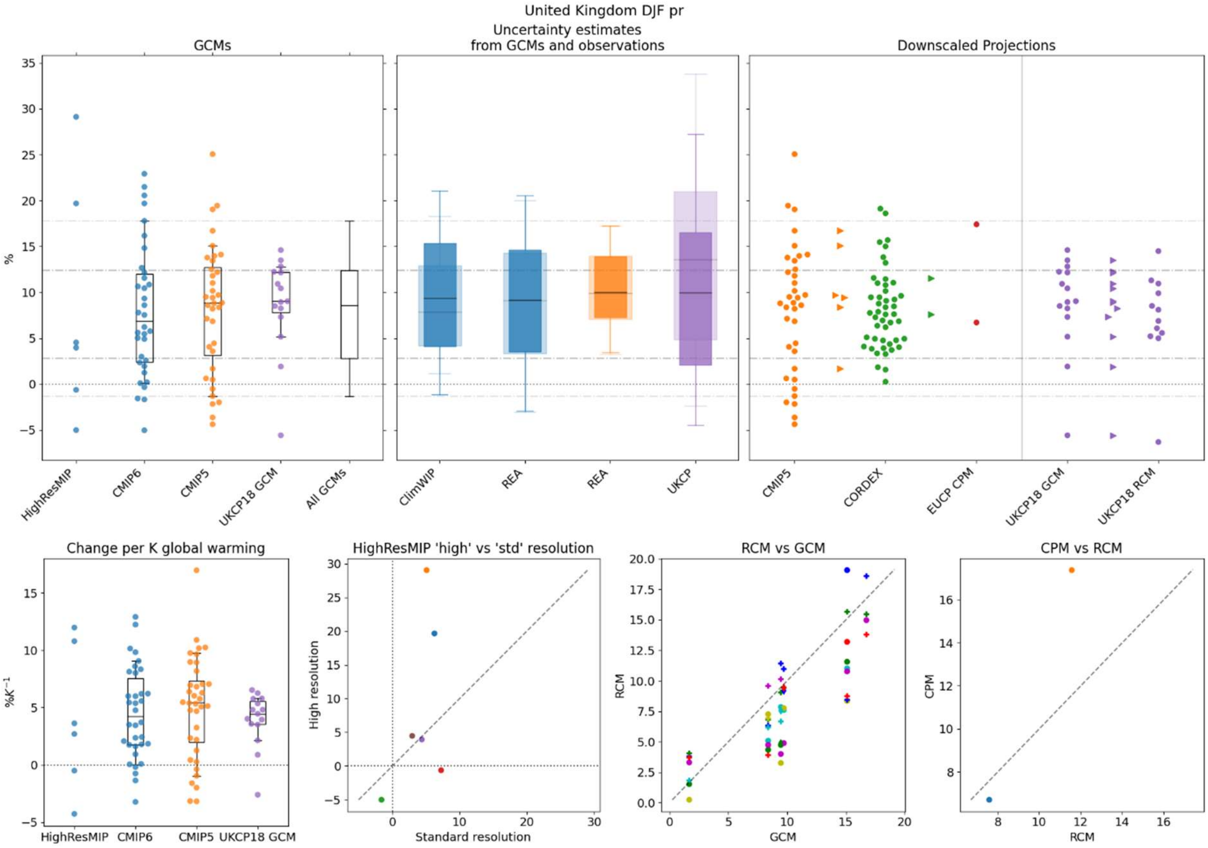


Figure 6.2-1 UK JJA Tas, JJA Pr and DJF Pr as anomalies for 2041-2060. The different panels shown are as explained in Section 3.4, showing the spread of the projected anomalies across multiple data sources, the impact of constraints, the anomalies per degree of warming and the effect of downscaling. Further details are given in the corresponding section.

6.2.2 Additional information

Tables of the full model details used in this section follow.

CMIP5 Models

Model name	Ensemble member	Historical and future periods
ACCESS1-0	r1i1p1	1995-2014, 2041-2060
ACCESS1-3	r1i1p1	1995-2014, 2041-2060
BNU-ESM	r1i1p1	1995-2014, 2041-2060
CCSM4	r1i1p1	1995-2014, 2041-2060
CESM1-BGC	r1i1p1	1995-2014, 2041-2060
CESM1-CAM5	r1i1p1	1995-2014, 2041-2060
CMCC-CESM	r1i1p1	1995-2014, 2041-2060
CMCC-CMS	r1i1p1	1995-2014, 2041-2060
CMCC-CM	r1i1p1	1995-2014, 2041-2060
CNRM-CM5	r1i1p1	1995-2014, 2041-2060
CSIRO-Mk3-6-0	r1i1p1	1995-2014, 2041-2060
CanESM2	r1i1p1	1995-2014, 2041-2060
EC-EARTH	r12i1p1	1995-2014, 2041-2060
FGOALS-g2	r1i1p1	1995-2014, 2041-2060
FIO-ESM	r1i1p1	1995-2014, 2041-2060
GFDL-CM3	r1i1p1	1995-2014, 2041-2060
GFDL-ESM2G	r1i1p1	1995-2014, 2041-2060
GFDL-ESM2M	r1i1p1	1995-2014, 2041-2060
GISS-E2-H	r1i1p1	1995-2014, 2041-2060
GISS-E2-R	r1i1p1	1995-2014, 2041-2060
HadGEM2-CC	r1i1p1	1995-2014, 2041-2060

HadGEM2-ES	r1i1p1	1995-2014, 2041-2060
IPSL-CM5A-LR	r1i1p1	1995-2014, 2041-2060
IPSL-CM5A-MR	r1i1p1	1995-2014, 2041-2060
IPSL-CM5B-LR	r1i1p1	1995-2014, 2041-2060
MIROC-ESM-CHEM	r1i1p1	1995-2014, 2041-2060
MIROC-ESM	r1i1p1	1995-2014, 2041-2060
MIROC5	r1i1p1	1995-2014, 2041-2060
MPI-ESM-LR	r1i1p1	1995-2014, 2041-2060
MPI-ESM-MR	r1i1p1	1995-2014, 2041-2060
MRI-CGCM3	r1i1p1	1995-2014, 2041-2060
NorESM1-ME	r1i1p1	1995-2014, 2041-2060
NorESM1-M	r1i1p1	1995-2014, 2041-2060
bcc-csm1-1-m	r1i1p1	1995-2014, 2041-2060
bcc-csm1-1	r1i1p1	1995-2014, 2041-2060
inmcm4	r1i1p1	1995-2014, 2041-2060
KNMI-EC-EARTH (To drive KNMI-RACMO23E) *	r14i1p1 (historical) r13i1p1 (future)	1996-2005, 2041-2050

* This model simulation was performed as part of EUCP and is not part of the original CMIP5 set of experiments available on ESGF.

CMIP6 Models

Model name	Ensemble member	Historical and future periods
ACCESS-CM2	r1i1p1f1	1995-2014, 2041-2060
ACCESS-ESM1-5	r1i1p1f1	1995-2014, 2041-2060
AWI-CM-1-1-MR	r1i1p1f1	1995-2014, 2041-2060
BCC-CSM2-MR	r1i1p1f1	1995-2014, 2041-2060
CAMS-CSM1-0	r1i1p1f1	1995-2014, 2041-2060
CESM2-WACCM	r1i1p1f1	1995-2014, 2041-2060
CMCC-CM2-SR5	r1i1p1f1	1995-2014, 2041-2060
CNRM-CM6-1-HR	r1i1p1f2	1995-2014, 2041-2060
CNRM-CM6-1	r1i1p1f2	1995-2014, 2041-2060
CNRM-ESM2-1	r1i1p1f2	1995-2014, 2041-2060
CanESM5-CanOE	r1i1p2f1	1995-2014, 2041-2060
CanESM5	r1i1p1f1	1995-2014, 2041-2060
E3SM-1-1	r1i1p1f1	1995-2014, 2041-2060
EC-Earth3-Veg	r1i1p1f1	1995-2014, 2041-2060
FGOALS-g3	r1i1p1f1	1995-2014, 2041-2060
FIO-ESM-2-0	r1i1p1f1	1995-2014, 2041-2060
GFDL-CM4	r1i1p1f1	1995-2014, 2041-2060
GFDL-ESM4	r1i1p1f1	1995-2014, 2041-2060
GISS-E2-1-G	r1i1p1f2	1995-2014, 2041-2060
HadGEM3-GC31-LL	r1i1p1f3	1995-2014, 2041-2060
INM-CM4-8	r1i1p1f1	1995-2014, 2041-2060
INM-CM5-0	r1i1p1f1	1995-2014, 2041-2060
IPSL-CM6A-LR	r1i1p1f1	1995-2014, 2041-2060
KACE-1-0-G	r1i1p1f1	1995-2014, 2041-2060
MIROC-ES2L	r1i1p1f2	1995-2014, 2041-2060
MIROC6	r1i1p1f1	1995-2014, 2041-2060
MPI-ESM1-2-HR	r1i1p1f1	1995-2014, 2041-2060
MPI-ESM1-2-LR	r1i1p1f1	1995-2014, 2041-2060
MRI-ESM2-0	r1i1p1f1	1995-2014, 2041-2060
NESM3	r1i1p1f1	1995-2014, 2041-2060
NorESM2-LM	r1i1p1f1	1995-2014, 2041-2060
NorESM2-MM	r1i1p1f1	1995-2014, 2041-2060
TaiESM1	r1i1p1f1	1995-2014, 2041-2060
UKESM1-0-LL	r1i1p1f2	1995-2014, 2041-2060

EUCP
European Climate Prediction system

CORDEX MODELS

Model name	Driving CMIP5 Model	Ensemble member	Historical and future periods	Domain
ALADIN53	CNRM-CERFACS-CNRM-CM5	r1i1p1	1995-2014, 2041-2060	EUR-11
ALADIN63	CNRM-CERFACS-CNRM-CM5	r1i1p1	1995-2014, 2041-2060	EUR-11
ALADIN63	MOHC-HadGEM2-ES	r1i1p1	1995-2014, 2041-2060	EUR-11
ALADIN63	MPI-M-MPI-ESM-LR	r1i1p1	1995-2014, 2041-2060	EUR-11
ALADIN63	NCC-NorESM1-M	r1i1p1	1995-2014, 2041-2060	EUR-11
CCLM4-8-17	CNRM-CERFACS-CNRM-CM5	r1i1p1	1995-2014, 2041-2060	EUR-11
CCLM4-8-17	MOHC-HadGEM2-ES	r1i1p1	1995-2014, 2041-2060	EUR-11
CCLM4-8-17	ICHEC-EC-EARTH	r12i1p1	1995-2014, 2041-2060	EUR-11
CCLM4-8-17	MPI-M-MPI-ESM-LR	r1i1p1	1995-2014, 2041-2060	EUR-11
COSMO-crCLIM-v1-1	MOHC-HadGEM2-ES	r1i1p1	1995-2014, 2041-2060	EUR-11
COSMO-crCLIM-v1-1	MPI-M-MPI-ESM-LR	r1i1p1	1995-2014, 2041-2060	EUR-11
COSMO-crCLIM-v1-1	NCC-NorESM1-M	r1i1p1	1995-2014, 2041-2060	EUR-11
RCA4	CNRM-CERFACS-CNRM-CM5	r1i1p1	1995-2014, 2041-2060	EUR-11
RCA4	MOHC-HadGEM2-ES	r1i1p1	1995-2014, 2041-2060	EUR-11
RCA4	MPI-M-MPI-ESM-LR	r1i1p1	1995-2014, 2041-2060	EUR-11
RCA4	IPSL-IPSL-CM5A-MR	r1i1p1	1995-2014, 2041-2060	EUR-11
RCA4	ICHEC-EC-EARTH	r12i1p1	1995-2014, 2041-2060	EUR-11
RCA4	NCC-NorESM1-M	r1i1p1	1995-2014, 2041-2060	EUR-11
RACMO22E	CNRM-CERFACS-CNRM-CM5	r1i1p1	1995-2014, 2041-2060	EUR-11
RACMO22E	MOHC-HadGEM2-ES	r1i1p1	1995-2014, 2041-2060	EUR-11
RACMO22E	MPI-M-MPI-ESM-LR	r1i1p1	1995-2014, 2041-2060	EUR-11
RACMO22E	IPSL-IPSL-CM5A-MR	r1i1p1	1995-2014, 2041-2060	EUR-11
RACMO22E	ICHEC-EC-EARTH	r12i1p1	1995-2014, 2041-2060	EUR-11
RACMO22E	NCC-NorESM1-M	r1i1p1	1995-2014, 2041-2060	EUR-11
HIRHAM5	CNRM-CERFACS-CNRM-CM5	r1i1p1	1995-2014, 2041-2060	EUR-11
HIRHAM5	MOHC-HadGEM2-ES	r1i1p1	1995-2014, 2041-2060	EUR-11
HIRHAM5	MPI-M-MPI-ESM-LR	r1i1p1	1995-2014, 2041-2060	EUR-11
HIRHAM5	IPSL-IPSL-CM5A-MR	r1i1p1	1995-2014, 2041-2060	EUR-11
HIRHAM5	ICHEC-EC-EARTH	r3i1p1	1995-2014, 2041-2060	EUR-11
HIRHAM5	NCC-NorESM1-M	r1i1p1	1995-2014, 2041-2060	EUR-11
REMO2009	MPI-M-MPI-ESM-LR	r1i1p1	1995-2014, 2041-2060	EUR-11
REMO2015	NCC-NorESM1-M	r1i1p1	1995-2014, 2041-2060	EUR-11
REMO2015	CNRM-CERFACS-CNRM-CM5	r1i1p1	1995-2014, 2041-2060	EUR-11
REMO2015	IPSL-IPSL-CM5A-MR	r1i1p1	1995-2014, 2041-2060	EUR-11
REMO2015	MPI-M-MPI-ESM-LR	r3i1p1	1995-2014, 2041-2060	EUR-11
RegCM4-6	MPI-M-MPI-ESM-LR	r1i1p1	1995-2014, 2041-2060	EUR-11
RegCM4-6	MOHC-HadGEM2-ES	r1i1p1	1995-2014, 2041-2060	EUR-11
ALARO-0	CNRM-CERFACS-CNRM-CM5	r1i1p1	1995-2014, 2041-2060	EUR-11
HadREM3-GA7-05	CNRM-CERFACS-CNRM-CM5	r1i1p1	1995-2014, 2041-2060	EUR-11
HadREM3-GA7-05	MOHC-HadGEM2-ES	r1i1p1	1995-2014, 2041-2060	EUR-11
HadREM3-GA7-05	MPI-M-MPI-ESM-LR	r1i1p1	1995-2014, 2041-2060	EUR-11
HadREM3-GA7-05	ICHEC-EC-EARTH	r12i1p1	1995-2014, 2041-2060	EUR-11
HadREM3-GA7-05	NCC-NorESM1-M	r1i1p1	1995-2014, 2041-2060	EUR-11
HCLIMcom-HCLIM38-ALADIN*	ICHEC-EC-EARTH	r12i1p1	1996-2005, 2040-2049	ALP-12
ICTP-RegCM4-7-0*	MOHC-HadGEM2-ES	r1i1p1	1996-2005, 2040-2049	EUR-11
KNMI-RACMO23E*	KNMI-EC-EARTH	r14i1p1 (historical) r13i1p1 (future)	1996-2005, 2041-2050	EUR-11
SMHI-HCLIM38-ALADIN*	ICHEC-EC-EARTH	r12i1p1	1996-2005, 2041-2050	CEE-12
MOHC-HadGEM3-GC3.1-N512*	MOHC-HadGEM2-ES	r1i1p1	1998-2007, 2040-2049	REU-25
RCA4	CCCma-CanESM2	r1i1p1	1995-2014, 2041-2060	CAM-44
RCA4	CNRM-CERFACS-CNRM-CM5	r1i1p1	1995-2014, 2041-2060	CAM-44
RCA4	CSIRO-QCCCE-CSIRO-Mk3-6-0	r1i1p1	1995-2014, 2041-2060	CAM-44
RCA4	ICHEC-EC-EARTH	r12i1p1	1995-2014, 2041-2060	CAM-44

RCA4	IPSL-IPSL-CM5A-MR	r1i1p1	1995-2014, 2041-2060	CAM-44
RCA4	MIROC-MIROC5	r1i1p1	1995-2014, 2041-2060	CAM-44
RCA4	MOHC-HadGEM2-ES	r1i1p1	1995-2014, 2041-2060	CAM-44
RCA4	MPI-M-MPI-ESM-LR	r1i1p1	1995-2014, 2041-2060	CAM-44
RCA4	NCC-NorESM1-M	r1i1p1	1995-2014, 2041-2060	CAM-44
RCA4	NOAA-GFDL-GFDL-ESM2M	r1i1p1	1995-2014, 2041-2060	CAM-44

* These model simulations were performed as part of EUCP and are not part of the standard CORDEX set of experiments available on ESGF.

HIGHRESMIP

Model name	Ensemble member	Historical and future periods
CMCC-CM2-HR4	r1i1p1f1	1995-2014, 2040-2049
CNRM-CM6-1	r1i1p1f2	1995-2014, 2040-2049
EC-Earth3P	r1i1p2f1	1995-2014, 2040-2049
HiRAM-SIT-LR	r1i1p1f1	1995-2014, 2040-2049
HadGEM3-GC31-LL	r1i1p1f1	1995-2014, 2040-2049
MPI-ESM1-2-HR	r1i1p1f1	1995-2014, 2040-2049
CMCC-CM2-VHR4	r1i1p1f1	1995-2014, 2040-2049
CNRM-CM6-1-HR	r1i1p1f2	1995-2014, 2040-2049
EC-Earth3P-HR	r1i1p2f1	1995-2014, 2040-2049
HiRAM-SIT-HR	r1i1p1f1	1995-2014, 2040-2049
HadGEM3-GC31-HH	r1i1p1f1	1995-2014, 2040-2049
MPI-ESM1-2-XR	r1i1p1f1	1995-2014, 2040-2049

EUCP CP-RCMS

Model name	Driving CORDEX Model	Ensemble member	Historical and future periods	Domain
CLMcom-CMCC-CCLM5-0-9	CCLM4-8-17 ICHEC-EC-EARTH	r12i1p1	1996-2005, 2041-2050	ALP-3
HCLIMcom-HCLIM38-AROME	HCLIMcom-HCLIM38-ALADIN ICHEC-EC-EARTH	r12i1p1	1996-2005, 2041-2050	ALP-3
COSMO-pompa	CCLM4-8-17 MPI-M-MPI-ESM-LR	r1i1p1	1996-2005, 2041-2050	ALP-3
CNRM-AROME41t1	ALADIN63 CNRM-CERFACS-CNRM-CM5	r1i1p1	1996-2005, 2041-2050	ALP-3
GERICS-REMO2015	REMO2015 MPI-M-MPI-ESM-LR	r1i1p1	1996-2005, 2041-2050	ALP-3
ICTP-RegCM4-7-0	ICTP-RegCM4-7-0 MOHC-HadGEM2-ES	r1i1p1	1996-2005, 2040-2049	ALP-3
KNMI-HCLIM38h1-AROME	KNMI-RACMO23E KNMI-EC-EARTH	r14i1p1 (historical) r13i1p1 (future)	1996-2005, 2041-2050	ALP-3
SMHI-HCLIM38-AROME	SMHI-HCLIM38-ALADIN ICHEC-EC-EARTH	r12i1p1	1996-2005, 2041-2050	CEE-3
ICTP-RegCM4-7-0	ICTP-RegCM4-7-0 MOHC-HadGEM2-ES	r1i1p1	1996-2005, 2040-2049	CEE-3
GERICS-REMO2015	REMO2015 MPI-M-MPI-ESM-LR	r1i1p1	1996-2005, 2041-2050	CEU-3
GERICS-REMO2015	REMO2015 MPI-M-MPI-ESM-LR	r1i1p1	1996-2005, 2041-2050	NEU-3
CNRM-AROME41t1	ALADIN63 CNRM-CERFACS-CNRM-CM5	r1i1p1	1996-2005, 2041-2050	NWE-3
MOHC-HadREM3-RA-UM10.1	MOHC-HadGEM3-GC3.1-N512 MOHC-HadGEM2-ES	r1i1p1	1998-2007, 2040-2049	REU-2
ICTP-RegCM4-7-0	ICTP-RegCM4-7-0 MOHC-HadGEM2-ES	r1i1p1	1996-2005, 2040-2049	SEE-3
CLMcom-CMCC-CCLM5-0-9	CCLM4-8-17 ICHEC-EC-EARTH	r12i1p1	1996-2005, 2041-2050	SWE-3