Annex to the review of stochastic and other approaches in water resources planning

Chief Scientist’s Group report

June 2025

Version: SC220039/R2

We are the Environment Agency. We protect and improve the environment.

We help people and wildlife adapt to climate change and reduce its impacts, including flooding, drought, sea level rise and coastal erosion.

We improve the quality of our water, land and air by tackling pollution. We work with businesses to help them comply with environmental regulations. A healthy and diverse environment enhances people's lives and contributes to economic growth.

We can't do this alone. We work as part of the Defra group (Department for Environment, Food & Rural Affairs), with the rest of government, local councils, businesses, civil society groups and local communities to create a better place for people and wildlife.

Published by:

Environment Agency
Horizon House, Deanery Road,
Bristol BS1 5AH

www.gov.uk/environment-agency

Further copies of this report are available from our publications catalogue: www.gov.uk/government/publications or our National Customer Contact Centre: 03708 506 506

Email: research@environment-agency.gov.uk

Author(s):
Jamie Hannaford[1], Wilson Chan[1], Chris Counsell[2], Mason Durant[2], Amulya Chevuturi[1], Adam Griffin[1], Alison Kay[1], Rosanna Lane[1], Ye Liu[1], Maliko Tanguy[1]

Keywords:
Water Resource Management Planning, WRMP, stochastic methods, climate change

Research contractor:
1. UK Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire. OX10 8BB
2. HR Wallingford, Howbery Park, Wallingford, Oxfordshire. OX10 8BA

Environment Agency's Project Manager:
Stuart Allen

Project number:
SC220039/R2

Citation:
Environment Agency (2025) Annex to the review of stochastic and other approaches in water resources planning. Environment Agency, Bristol.

Research at the Environment Agency

Scientific research and analysis underpins everything the Environment Agency does. It helps us to understand and manage the environment effectively. Our own experts work with leading scientific organisations, universities and other parts of the Defra group to bring the best knowledge to bear on the environmental problems that we face now and in the future. Our scientific work is published as summaries and reports, freely available to all.

This report is the result of research commissioned by the Environment Agency's Chief Scientist's Group.

You can find out more about our current science programmes at https://www.gov.uk/government/organisations/environment-agency/about/research

If you have any comments or questions about this report or the Environment Agency's other scientific work, please contact research@environment-agency.gov.uk.

Dr Robert Bradburne
Chief Scientist

Contents

Introduction

The Environment Agency's Chief Scientist's Group commissioned a review of the use of stochastic and other approaches in water resources planning, including their current use and limitations, potential developments and alternative options.

Stochastic simulation methods have become an established way for water companies to create a larger sample of rainfall inputs for their hydrological and systems models, than from the observed record alone. But there has been little to no assessment of their use in practice, or appropriateness when considering future climate. This project aimed to review the science behind stochastic methods and the performance of existing products, whilst clarifying their strengths, weaknesses and appropriate uses, alongside an equivalent appraisal of alternative approaches.

The project consisted of a review of the stochastic methods currently used by water companies, and the potential alternatives. This was primarily literature-based, including both academic and grey (e.g. water company documents) and the findings summarised in topic focused reviews (e.g. climate change).

An overview of the reviews is presented in Environment Agency (2025). Review of stochastic and other approaches in water resource planning. Environment Agency, Bristol. This report is an annex to the summary report and presents the topic focussed reviews as written by the authors, with only minor editing of the format.

Narrative history of water resources planning in England in relation to droughts (strategic and operational)

Chris Counsell, Mason Durant (with review and input from the author team)

Glossary

Term	Meaning
DO	Deployable Output – the amount of water (Ml/d) that can be obtained from a system under specific conditions
DP	Drought Plan
EVA	Extreme Value Analysis
LoS	Levels of Service
TUB	Temporary Use Ban
UKCP	Deployable Output – the amount of water (Ml/d) that can be obtained from a system under specific conditions
WRMP	Drought Plan

Introduction

Within the latest WRMP process in England, stochastic datasets have emerged as the primary form of assessing water supply system performance under drought conditions. This section of the report charts their development using a narrative approach, combining literature and interview responses, to help understand how they have developed into the datasets that currently exist, the issues and benefits related to their implementation. This understanding is critical in capturing the practicalities of using stochastic datasets. The practicalities are then aligned with statistical issues related to stochastic methods for a number of important themes identified during the literature review and interviews. The report concludes by combining this collected information into a set of characteristics of an idealised dataset against which potential datasets and approaches may be assessed.

Long term water management and drought planning

Water Resource Management Plans (WRMPs) outline the actions that water companies will undertake to ensure the security of water supply within their area over at least the next 25 years. These plans are updated every five years (and reviewed annually) in an iterative, statutory process that was introduced as part of the Water Resources Act 2003. The first statutory plan was introduced in 2004. Prior to this, water companies had produced a voluntary plan in 1999.

Planning within England (and Wales) has historically been done at a relatively local scale (albeit with generally increasing spatial scale over time as water companies have grown in size (Ofwat, 2006) and through regional and national scale planning). The current statutory process is for water companies to submit their WRMPs to the Secretary of State. Often the scale has presented challenges in responding sufficiently to some resource issues where water bodies (both surface water and groundwater) cross water company boundaries, or where there are water transfers between companies. Consequently, regional planning groups have been established to create more formal cooperation, starting in 1996 (Water Resources South East, 2017) and culminating in all 5 regional groups publishing multi-sector regional water resources plans in 2022 (Environment Agency, 2022). Prior to the publication of the latest regional plans, the National Framework for Water Resources (a national scale assessment of England's long-term water needs) was published in 2020 (Environment Agency, 2020).

The publication of Drought Plans (DPs) is also a statutory process that must be undertaken every 5 years (originally every 3 years) under the Water Industry Act 1991. DPs must state how a water company will maintain a secure water supply and protect the environment during dry weather and drought in the nearer term. Historically, drought planning has been undertaken at a water company scale, however more recently, there have been examples of multi-sectoral and regional scale exercises in drought preparedness such as Arica (Thames Water, 2022).

Narrative history of water resource planning in England in relation to droughts

This section briefly summarises the history of water resource planning with respect to strategic and operational planning up to and including the introduction of stochastic weather generators. The development of the weather generators used in water resource planning in England and Wales is summarised, as well as how they have been applied to date.

Prior to 2014, WRMPs largely used the worst historical drought to determine the Deployable Output (DO) of a water supply system against stated Levels of Service (LoS), effectively treating the worst historical drought as a design event (Southern Water, 2014). During this period (2009-2014), stochastic weather generation was expanding rapidly, both within UK adaptation with the UK Climate Projections 2009 (UKCP09) Weather Generator (Jones *et al.*, 2010) and more broadly within the academic literature (see Figure 1). Southern Water felt that the historical record was not sufficient to capture the events that affected their system and, for WRMP14, started to explore the use of stochastic weather generators to improve their assessments (Southern Water, 2014; Interview 1, personal communication, 23rd May 2023).

Southern Water included stochastic methods within their Drought Plan in tandem with their WRMP14. These were used to test drought measures in response to drought events that have different characteristics such as timing or severity than those experienced historically. Since WRMP14, a number of other water companies have also used stochastic approaches for this purpose.

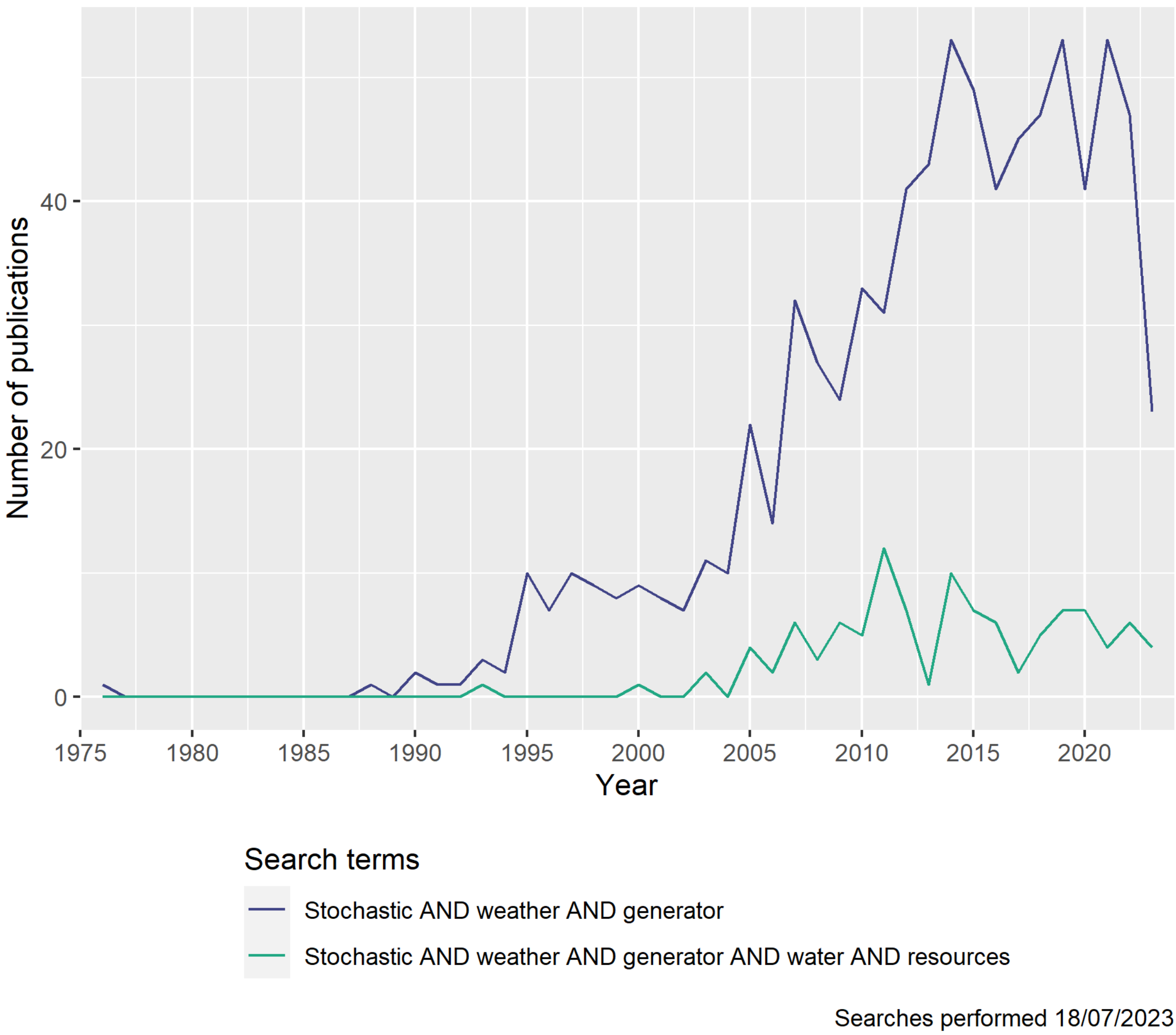

Figure 1. Publication of stochastic related articles over time from the SCOPUS database.

History of stochastic approach development

Stochastic approaches were introduced to water resource planning in the UK through the use of a stochastic weather generator within UKCP09. This was underpinned by a body of work, primarily undertaken by Newcastle University, on stochastic rainfall (e.g., Fowler *et al.*, 2000, 2005; Cowpertwait *et al.*, 2002; Kilsby *et al.*, 2007) that primarily used the Neyman-Scott point process model with a spatial capability. The state of knowledge on climatic drivers and their relationship with UK weather is reviewed in Chapter 4 (see also Hannaford et al. 2023). There were a large number of users of the UKCP09 Weather Generator, and user feedback indicated a need for modelling across large areas with realistic spatial coherence and over long time scales (Interview 3, 2023). Southern Water subsequently procured a stochastic weather generator from Newcastle University, Atkins, and the University of East Anglia (Interview 3, 2023). A new approach was developed based on user requirements that was inherently spatial, closely fitting observations so as to provide credibility through validation, and used a model that could be run in a short time period (Interview 3, 2023). This model produces monthly rainfall at individual sites, which is subsequently downscaled to daily data and paired with potential evapotranspiration data. The stochastic dataset generated for Southern Water's WRMP14 comprised of 17 replicates of the 120 year-long historical record, combined into one single Very Long Time Series (VLTS) of 2,000 years (Southern Water, 2014).

This model was then reviewed and updated by the Met Office in 2016 for the Water Resources East (WRE) region, including the addition of the East Atlantic pattern for the Anglian region and seasonally varying predictors (Dawkins *et al.*, 2022). This model was subsequently used by Atkins to produce stochastic datasets for some water companies for WRMP19, comprising of 200 replicates of 88 years each totalling 17,600 years (Atkins, 2020). For WRMP24, WRSE procured a stochastic weather generator from Atkins that covered the whole of England and Wales for the purposes of regional planning, during which the stochastic generator was updated (Atkins, 2020). This included:

- The addition of a number of climate drivers, including the East Atlantic pattern, which has been shown to have a strong influence on rainfall variability both compared to and in combination with other existing predictors such as the NAO.
- The interaction between the explanatory variables describing mean rainfall at each site.
- The reduction of the period of fitting to 1950-1997 on the basis of the availability of robust climate data prior to 1950 and climate change impacts post 1997.
- The improvement of a curve-fitting (bias correction) approach to improve the final model fit.

The rationale for updating these aspects of the weather generator process was to improve the quality of drivers, improve the fit of the statistical model, unify the model across the country and reduce the impacts of bias correction to avoid implausible droughts (Atkins, 2020). Atkins (2020) states that the (continued) use of bias correction is a debatable but necessary step to ensure that outputs adequately represent and extend the range of droughts within the historical record. The step is included to effectively constrain the uncertainty within the stochastic data to ensure the extremes look plausible (Interview 3, 2023). These improvements were based largely on aspirations of the modellers to improve the model representation of rainfall to known climate drivers and a response to peer review of the previous approach (Atkins, 2020). The stochastic dataset comprises of 400 replicates of a 48 year period, comprising 19,200 years in total (Atkins, 2020).

A number of shortcomings of the approach identified in 2016 (prior to updates for WRMP24) have been attempted to be solved by Dawkins *et al.* (2022) and the development of the Advanced Meteorological Explorer (AME) for the Anglian region. This is largely focused on high spatial and temporal resolution data, plausible long-duration drought characteristics (particularly reflecting droughts in the late 19th and early 20th century) and capturing known behaviour of UK rainfall variability (Dawkins *et al.*, 2022). The strength of fit with the observed data as a result of the additional drivers and the bias correction process were also identified as driving reasons for creating a new approach (Interview 6, 2023). Dawkins *et al.* (2022) could (although not yet implemented) also be used to generate other variables such as potential evapotranspiration (PET), where previous methods have not been able to. This has been a particular limitation (which was traded-off against increased model complexity and run times) of previous methods where PET sequences were developed after the rainfall generation and based solely on the historical record (Atkins, 2020), thereby potentially under sampling the full range of temperatures that might be experienced.

How stochastic datasets have been implemented to date

Long term water resource and drought planning

Stochastic datasets were procured as a means of testing water supply systems to events not seen within the observed historical record. For Southern Water in WRMP14, this meant focussing on the drought frequency and severity (due to not meeting stated Levels of Service), but also exploring the impacts of increased drought duration (including dry winters) and varying characteristics of drought onset (Southern Water, 2014). Southern Water (2014) identified an incoherence between the return period of a severe drought and the return period of drought measures, largely due to the issue of imperfect foresight. Interventions therefore need to be implemented prior to a severe drought hitting it's maximum severity, meaning TUBs need to be implemented more frequently than the frequency of severe droughts. The consequences of using this stochastic approach for Southern Water were that MDO/ADO reduced by 15Ml/d or approximately 2.3% and PDO reduced by 13.8Ml/d or approximately 1.6% across all three of their water resource zones. This represented a change in frequency for TUBs implementation from 1 in 7 to 1 in 10 years and Drought Permit or Order implementation from 1 in 10-15 to 1 in 20 years. For WRMP14, Southern Water did not include stochastic events that were of a greater duration than those in the historical record, suggesting they were not comfortable in allowing the DO of the system to be influenced by such extremes.

Beyond WRMP14, a number of other water companies used stochastic datasets for WRMP19, largely in response to the need to examine system resilience to more extreme drought events than those contained within the historical record. The analyses within these WRMP19 assessments focussed on using the stochastic dataset as a means of testing water supply systems to events with return periods of 1 in 200 years and sometimes 1 in 500 years (e.g., Thames Water, 2020), which also permitted a standardisation of impacts across these water companies (the use of return periods in general when applied to drought is outside the scope of this report). The 1 in 500 year return period (linked to system impacts, rather than return period of rainfall) has been adopted for WRMP24, with water companies required to plan to be resilient to an event of this severity by 2039 at the latest (Environment Agency, 2021a).

In general, water companies currently use stochastic datasets to test their drought plans to more severe drought events than those seen historically. This usually entails more detailed examination of the drought triggers and protocols undertaken during a drought than the WRMP, as well as narrative on the development of drought conditions (e.g., Thames Water (2022); Yorkshire Water (2022b)).

Stochastic demand

As well as using stochastic datasets for supply side modelling, they could also be used to generate estimates of stochastic demand. It is known that demand variability, particularly during heatwaves, can have a significant impact on water supply system performance for some systems, and there are concerns that this is not currently captured within existing, demand profile-based, approaches (Durant and Counsell, 2023). A number of attempts have been made to generate these, with stochastic weather-based models having been developed by organisations including WRSE, Anglian Water, United Utilities, however as yet, none have been implemented within WRMP assessments (Interview 1, 2023; Interview 5, 2023; Interview 6, 2023) but this remains an ambition, with recent years highlighting the significant uncertainties associated with customer water use behaviour during different weather periods. Coherent stochastic demand scenarios have been generated for the Anglian region, but are yet to be incorporated within a WRMP setting (Interview 6, 2023). WRE did consider the effects on agricultural demand as part of their latest regional plan (Knox *et al.*, 2018).

Stochastic methods within bottom-up approaches

Currently, the dominant approach to testing water supply systems within WRMPs is using a top-down, scenario-led approach, where climate data are run through impact models in a traditional modelling chain. In contrast, a number of academic studies (e.g., Borgomeo *et al.*, 2015; Environment Agency, 2015; Prudhomme *et al.*, 2015; Sauquet *et al.*, 2019) have emphasised the potential benefits of adopting a stress testing or vulnerability led (bottom-up) approach to considering drought resilience and inform decision-making. Stress testing involves analysing the potential impacts of a given hazard (in this case drought) on a system, adopting a diverse set of droughts over a range of different severities to reveal thresholds of vulnerability (or tipping points) where the impacts become particularly severe (Environment Agency, 2015). Climate evidence is then added to the outputs from stress testing to inform estimates of likelihood and plausibility (e.g., Prudhomme *et al.*, 2015; Sauquet *et al.*, 2019). Stress testing can be used to test alternative portfolios of interventions over a wide range of drought conditions of different severities, in order to quantify the impacts and potential benefits.

In the latest WRMPs, stochastic datasets have been assimilated to produce a drought response surface akin to the outputs from a vulnerability-led approach (informed by the guidance within the Drought Vulnerability Framework; UKWIR, 2017) but this is not, per se, part of a bottom-up approach to assessing drought resilience and decision making. The resulting drought response surface (DRS) presents water supply system response to increasing stress (in this case intensity and duration of rainfall deficits), with an example presented in Figure 2, but is limited to the range of events present within the stochastic dataset. Further, these outputs have not been taken forward into the wider WRMP process to compare different option portfolios, alternative climate evidence and consequences of future policy options (e.g., impact of Environmental Destinations).

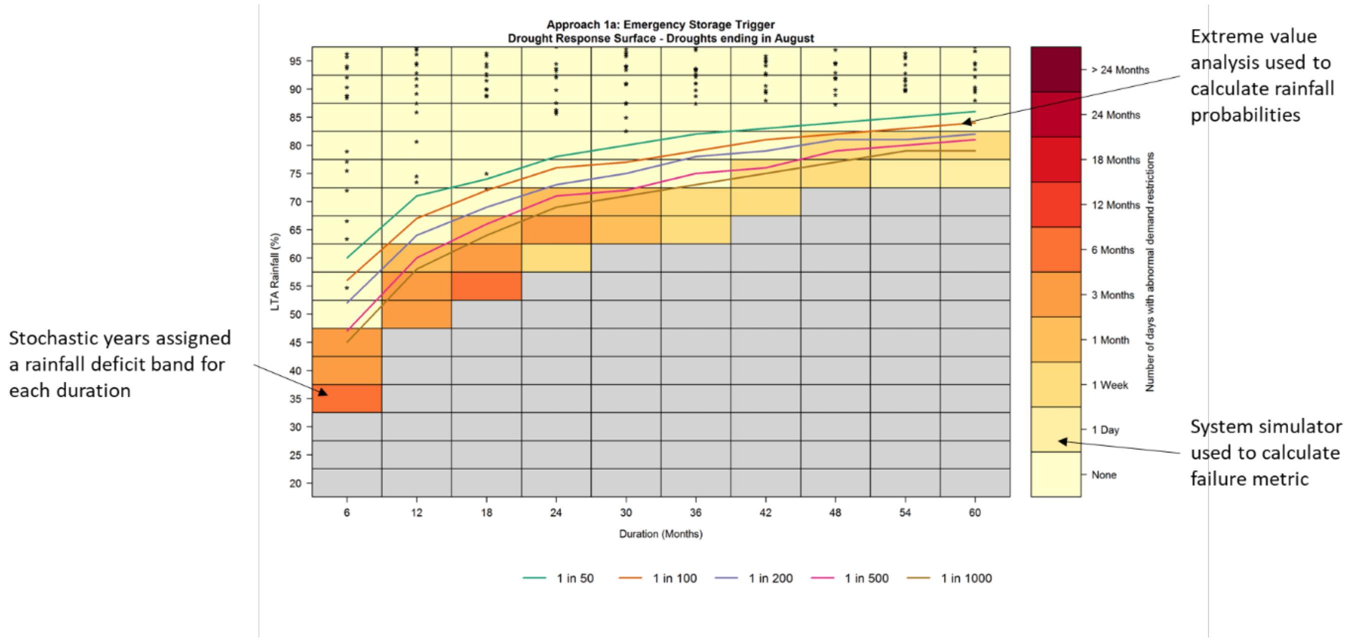

Figure 2. An example Drought Response Surface (UKWIR, 2017), annotated with methods used to generate underlying data.

Stochastic datasets and climate change

Ever since the inception of stochastic approaches within water resources planning, they have been used alongside climate change factors to account for the impacts of future projected change at different time slices (Southern Water, 2014). Change factors are calculated by determining the mean change in a weather variable for each month from a baseline period (e.g., 1981-2000) to a future period (e.g., 2080-2099) and applying these mean monthly factors to the stochastic dataset. This approach has remained unchanged up to WRMP24 where the guidance still recommends the use of change factors in the absence of compelling evidence to move to a different approach despite known drawbacks (Environment Agency, 2021b). The Advanced Meteorological Explorer (AME) has recently developed “future” stochastic data, conditioned on the UKCP18 regional projections that enables climate change to be represented without the use of change factors (Dawkins *et al.*, 2022). As part of this approach within the AME, different climate drivers are also modelled compared with those experienced historically, increasing the natural variability sampled. However, such an approach introduces a complex set of challenges related to climate model credibility, including how much information should be taken from climate models, and how to validate the outputs. At the time of writing, there is no published evidence available that explores these challenges. This should be reviewed once it is made publicly available. Further review of climate change and stochastic datasets is given in Chapter 5.

Extreme Value Analysis

The implementation of stochastic datasets within the current planning framework in England and Wales requires the estimation of return periods of drought events to determine Levels of Service. There are two primary Extreme Value Analysis (EVA) approaches that have been used to compute return periods for both rainfall and system metrics – non-parametric and parametric approaches. The non-parametric approach (e.g., inverse ranking) is underpinned by the concept of the probability of an event being determined by the ranking of that event within a series and the length of the series. There are a number of ways of calculating this (e.g., Cunnane, Gringorten, California, Weibull), which make certain assumptions about the probability of events. Parametric approaches use a pre-determined statistical distribution that satisfies certain mathematical conditions. This distribution is then fitted to the data to determine the return period. In addition to the approach undertaken, there is also a difference in the way the data are processed to extract events on which the return period is assessed. For example, peaks-under-threshold (e.g., river flows below Q95) or block minima (e.g., minimum flow storage over a 12 month period). The use of either inverse ranking or parametric approaches is important because of a number of reasons:

- Inverse ranking produces a single return period for a given value, or a single value for a given return period. This contrasts with parametric approaches in which the fitting process is often undertaken numerous times in a process called bootstrapping, which produces uncertainty associated with the fitting of the distribution. Often the shorter the observed data series, the larger the estimation uncertainty.
- Parametric approaches can be used to determine return period beyond those observed. Inverse ranking cannot.

A number of other issues arise when these reasons are intersected with a stochastic dataset comprised of a number of replicates, where the fitted series (e.g., 48 years) is shorter than the required return period (e.g., 1-in-500). For a comprehensive overview of the differences between stochastic replicates, see Chapter 2. The use of inverse ranking on each stochastic replicate is insufficient (inverse ranking can only give a maximum return period equal to or less than the length of the series) and therefore the stochastic replicates must be joined together to create a longer series and determine more extreme return periods. This is a potential issue where the stochastic series is heavily fitted to the observed series – the frequency of droughts is fixed to that of the historical record and will alter the return period depending on the severity of the events experienced historically (e.g., Yorkshire Water, 2022). Consequently, if each replicate is to be assessed individually, a parametric approach must be used, however where a short series is available, estimation uncertainty will be high and will result in large ranges of possibilities of DO for a given return period (see Figure 3). It is possible that where conditions align, the inverse ranking approach and the median of the parametric approach align (e.g., in Figure 3, the red line fits with the dotted line), however this is location dependent. There are also questions around the assumption that all drought events belong to the same population and therefore can all be treated equally within a frequency analysis (UKWIR, 2017).

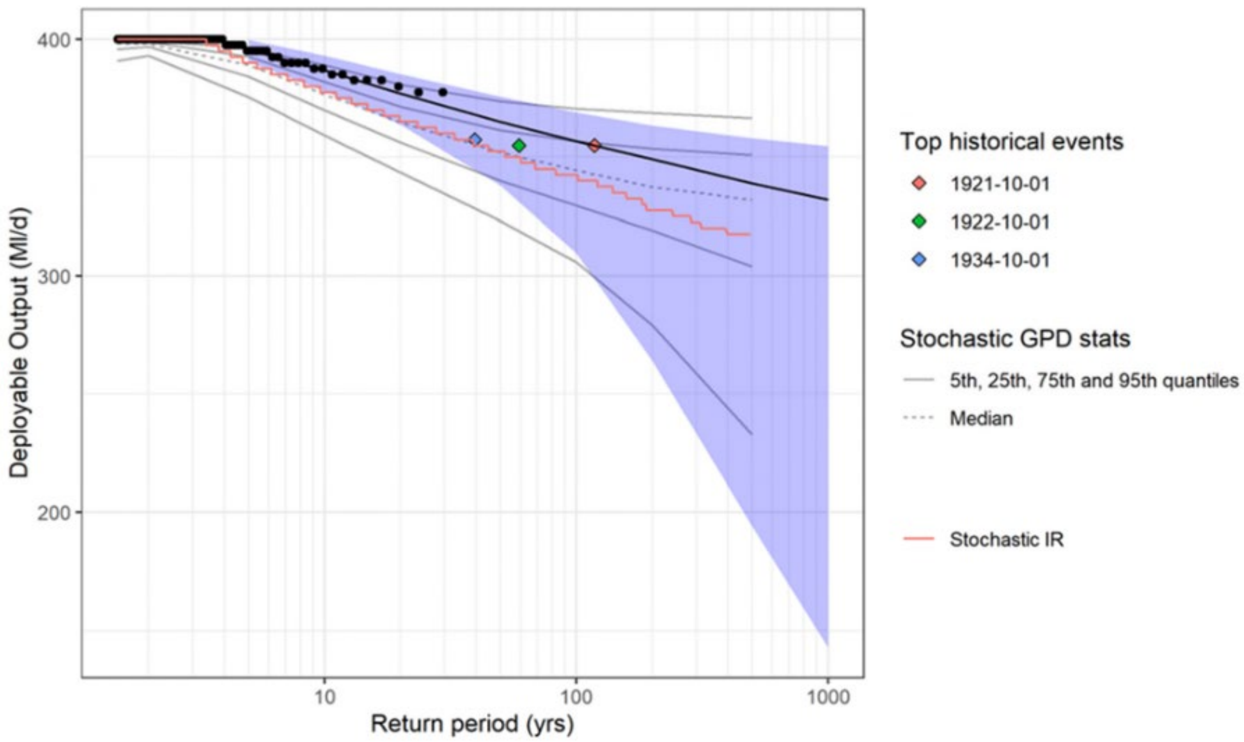

Figure 3. Comparison of EVA and inverse ranking approaches for the Bristol Water supply system. Historical inverse ranking is shown as points, historical EVA at the 95% confidence interval as the purple shaded area, stochastic inverse ranking with 400 replicates combined into one series as the red line, and stochastic EVA undertaken on each replicate as the grey lines. Source: Bristol Water (2022)

For stochastic generators that are explicitly tied to underlying climatological drivers, the replicates produced by stochastic generators are usually treated as independent. This may not always be the case, due to the dependence within drivers across replicates, effectively reducing the sample size, which subsequently reduces the largest return period that can be reliably determined from the dataset, and should also be considered where replicates are combined into one series. This is particularly pertinent when generating data at the national scale, where relationships between drivers and weather variables vary in space, meaning droughts are not spatially coherent between regions. The LoS calculated from a stochastic dataset can therefore vary depending on the stochastic generation method as well as events themselves within the dataset.

All stochastic generators are underpinned by observed data. There has been a general reduction in the density of rain gauges since the 1970s (Chapter 3) which impacts the ability of all stochastic frameworks to produce robust datasets. Stochastic frameworks need high quality data, that is collected in relevant locations, and has a long record. Greater consideration should be given to the density of the networks, location of instruments and continuous measurement if the quality of stochastic data is to be maintained or improved. Further discussion on the details of input data, including observed rainfall, are included within (Chapter 3).

Models and tools

As the use of stochastic approaches have become more widespread across the industry, there have been conflicts between the quantity of data available within the stochastic dataset and the capability of existing models (particularly water resource models) to run the quantity of data. This was particularly relevant where water resource models were only required to run the relatively short historical period, rather than the equivalent of 19,200 years of stochastic data. This is compounded by the requirement to run a large number of climate change scenarios, as well as other scenarios including sustainability reductions. This conflict between quantity of data and model capacity necessitated the use of sampling approaches (such as Latin Hypercube Sampling) to reduce the length of model runs whilst retaining the range of uncertainty in the larger, unsampled dataset. The need for sampling is reducing as a result of improved rapid simulators that can be run flexibly on multi-core processors, where sampling can be more onerous than running all scenarios through a rapid simulator (Interview 5, 2023).

Review of practical issues related to implementation

Table 1 outlines practical issues related to the implementation of stochastic datasets, based on responses within interviews undertaken as part of this study, enabling a wider view of the issues and benefits of using stochastic datasets in practice. Interviews were undertaken either in person in a semi-structured format or as written responses to questions via email during May, June and July 2023.

Table 1. Overview of key issues with current stochastic approaches

Issue	Statistical issue	Practical implication
Spatial coherence	Increasing the spatial coverage of the dataset comes with trade-offs. More sites are required, which increases processing time. There are issues with modelling spatial correlation over larger areas and retaining plausible extremes (Chapter 2). Interpolation may be a way of getting around this and is generally computationally cheap in comparison but may cause issues for areas with varied orography. In addition, the stochastic method needs to be able to deal with potentially different covariates or covariate relationships from one site or region to another. Further work is needed to understand about trade-offs between increasing the number of sites versus relying more on interpolation.	Spatial coherence is an important requirement for inter-regional planning. Current approaches to match droughts between regions are proportionate, but there may be approaches that can deal with this a more coherent way.
Stochastic outputs as historical replicates	Methods that stochastically vary weather variables based on historical drivers may be under-sampling the historical uncertainty. Of these methods, there is a philosophical decision about how many drivers and how tightly to fit the model to the historical data. There is also a potential trade-off between the number of replicates and the need to fit to the historical weather variables.	Is the historical a fair indicator of the future? Strict dependence on the historical drivers restricts the method of EVA. Not all historical uncertainty is being explored – greater uncertainty could be revealed by varying the underlying historical drivers in some way (e.g., stochastically, outputs from climate models)
Climate change	There are a number of statistical issues related to using change factors as a way of incorporating climate change information into water resource assessments. Climate change has been incorporated into a number of stochastic generating frameworks, some using change factors in a post-processing step, some using bias-corrected	It is difficult to appraise the realism of a 1-in-500 drought plus climate change as a future drought. Assessing the severity of such events is also currently difficult. Notwithstanding this, change factors represent a simple solution that is relatively easy

Issue	Statistical issue	Practical implication
	variables from climate models within the stochastic model, and others using climate drivers (Chapter 5). There are a number of unknowns related to incorporating climate change into stochastic models, including which variables and covariates to include, how much information can and should be taken from the climate models, the representation of climate model uncertainty and whether to produce time slices (e.g., at global warming levels) or transient future projections (for emissions scenarios).	to convey to stakeholders and one that those in the industry understand (Interview 5, 2023). In addition, there is a simplicity in using the same underlying stochastic dataset to understand baseline and future risk.
Plausibility of stochastic events	Plausibility has been assessed using EVA and validating outputs statistically against the historical record. The plausibility of the events is assumed on the basis of the structure of the statistical model and the outcome of this validation step. There is no link back to the physical climate drivers that underly a particular event.	The plausibility of stochastic events can be difficult to assess, but has been achieved using a weight of evidence approach within WRMP24 using the AME stochastic dataset (Interview 6, 2023). This represents a plurality of approaches and datasets that permits qualitative assessment of plausibility. A large dataset means stochastic droughts are not examined in the same way as historical droughts (Interview 7, 2023), however they do permit a more probabilistic risk-based approach (Interview 5, 2023).
Extreme Value Analysis	The EVA method used alters the uncertainty associated with the estimation. Non-parametric approaches generally provide a single estimate, whereas parametric methods usually provide an estimation uncertainty.	The method of EVA to use is dependent on a number of things – the dataset itself and the correlation structures, whether there is a difference between parametric and non-parametric methods within a particular location, and the

Issue	Statistical issue	Practical implication
		methods of sampling or not that are used (Interview 3, 2023). Users are relatively familiar with EVA and methods and a generally setup to deal with large datasets (Interview 2, 2023). More generally, there is a mixed opinion on the use of the 1-in-500 and what this means in terms of both the science (of estimating the 1-in-500) but also the knowing what risks are being adapted to.
Computing power and resource requirements	The statistical model framework has an impact on the computing power required to produce the stochastic outputs.	A few water companies had issues with running a full stochastic dataset through the full modelling chain. This was largely where hydrological models were not available for some areas, or where water resource models or groundwater models were too resource intensive to undertake the required number of runs. There are also issues with re-calibrating models (particularly regional groundwater models (Interview 1, 2023)). Many companies already have, or have moved to, rapid system simulators that permit large numbers of model runs. The computing power to deal with the stochastic datasets as they currently are

Issue	Statistical issue	Practical implication
		implemented appears to be less of an issue.
Drought event diversity	The statistical model design will determine the drought event diversity, largely as a function of the historical observation data used for model calibration (both temporal and spatial components) and the parameters within the statistical model. Including more recent, climate change impacted droughts, may also want to be captured, but represent challenges to incorporating within a framework.	Issues with stochastic methods not capturing issues relevant to a particular water resource system (e.g., spatial rainfall distribution within a large scale drought event, large scale blocking over a region) that underrepresents uncertainty.
Observation data and calibration period	A high density of observational data is required, particularly where gridded datasets are used, to ensure outputs are representative of local weather variability. Uncertainty in the observational data or choice of interpolation method may have an impact on the interpretation of results (Chapter 3).	Practical issues can arise where the stochastic dataset is not based on the data used to calibrate a model, requiring the translation of data to arrive at stochastic impacts. There are issues where the models that require recalibration are large, computationally slow, or where responsibility lies within other organisations (Interview 1, 2023).
Spatial coherence	Increasing the spatial coverage of the dataset comes with trade-offs. More sites are required, which increases processing time. There are issues with modelling spatial correlation over larger areas and retaining plausible extremes (Chapter 2). Interpolation may be a way of getting around this and is generally computationally cheap in comparison but may cause issues for areas with varied orography. In addition, the stochastic method needs to be able to deal with potentially different covariates	Spatial coherence is an important requirement for inter-regional planning. Current approaches to match droughts between regions are proportionate, but there may be approaches that can deal with this a more coherent way.

Issue	Statistical issue	Practical implication
	or covariate relationships from one site or region to another. Further work is needed to understand about trade-offs between increasing the number of sites versus relying more on interpolation.	

As well as points related to the stochastic datasets themselves, wider points were raised by respondents. These included:

- Coherence across guidance documents, primarily the 1-in-500, climate change and stochastic methods documents (Interview 8, 2023).
- Testing of methodologies proposed within guidance documents using worked examples. This would prevent some of the issues that water companies had to solve while delivering WRMPs (Interview 5, 2023; Interview 8, 2023).
- Timely release of guidance documents prior to WRMP work commencing (Interview 5, 2023).
- Skill levels across the sector (including regulators) to deal with stochastic datasets and accompanying methods. This includes poor provision of taught courses at universities. An approach that considers all actors within the sector is needed to meet this challenge (Interview 3, 2023).
- Whether the current complexity of modelling approaches reduce uncertainty sufficiently to be warranted. It is debatable as to whether the current resource intensive approaches improve the decision making position (Interview 8, 2023).
- There is the potential for a collaborative test-bed for stochastic datasets that could permit combining the strengths of different stochastic approaches (Interview 6, 2023).

Summary

In summary, stochastic datasets have been used widely across both operational and long term water resource planning over the last decade and are now well integrated into water company modelling procedures (Interview 2, 2023). Hydrological and water resource modelling capabilities have risen up to the challenge of large datasets, however a number of technical issues related to stochastic datasets and their implementation remain. These issues have been captured in the characteristics of the evidence base required for planning section below.

Characteristics of the evidence base required for planning

There are a vast number of parameters and structures associated with stochastic approaches that may ultimately determine the dataset that is generated. Consequently, instead of considering how each (or combinations) of these may impact water resource modelling outcomes, it is perhaps more useful to consider the ideal characteristics of the evidence base that is required for water resource planning and consider the current and available approaches within this framework.

The list below shows the characteristics of an idealised evidence base for water resource planning. This list can be considered applicable to all datasets, including historical observed, derived and climate change.

- Nationally spatially coherent
- High resolution (space and time)
- Evidence to underpin event plausibility
- Ability to represent "future droughts" (i.e. include a climate change signal)
- Diversity of events (events with contrasting duration, intensity, frequency, spatial and temporal patterns)
- Applicable to multiple sectors (e.g., multiple weather variables, heat waves, include stress events appropriate for agricultural and energy sectors)
- Represent full range of uncertainties
- Assessed prior to use and demonstrated
- Industry consensus as to their application and interpretation
- Can be used in a risk-based approach
- Ability to effectively attribute event likelihood
- Can be practicably applied for all systems and for all water stakeholder / receptors at a national scale across the industry

The reasoning for including these characteristics is given below and is based on information gathered from interviews undertaken as part of this study as well as academic and industry grey literature.

Nationally spatially coherent.

As outlined above, water resource planning in England and Wales is being undertaken at both national and regional scales. Despite the statutory requirement for planning at a water company level, the requirement for water transfers and modelling of existing transfers from one region or water company to another requires that datasets are of a national scale (Interview 1, 2023; Interview 4, 2023). Whilst initial findings of this report indicate that this may be difficult to achieve within stochastic frameworks, it should be explored to understand if trade-offs or costs make it an acceptable proposition, and if it has a material impact on decision making.

High resolution (space and time).

Climate datasets are required to be of at least daily temporal resolution to drive hydrological models that operate on these timescales. Temporal downscaling from monthly to daily is not trivial and the decisions about how this is done may impact drought timing and impacts for some water supply systems, particularly rapidly responding systems. High resolution spatial data are also required for some water resource applications (primarily hydrological modelling). This is more important in areas with small catchments with varied topography. Orographic differences may alter the climate in these areas.

Evidence to underpin event plausibility.

There is a need to demonstrate the plausibility of drought events where these may be driving investment. In this context, “plausibility” is defined as the extent to which droughts or their underlying processes may be validated against observational data.

Ability to represent "future droughts".

The current method of incorporating climate change using mean monthly change factors to perturb stochastic drought events does not capture the changes in frequency, inter-annual variability or spatial extent of drought events (Diaz-Nieto and Wilby, 2005; Fowler *et al.*, 2007). The events created using change factor approaches represent extreme events, however do not represent necessarily what a future drought might look like. Durant and Counsell (2023) outline a number of other issues related to bias correction and incorporation of climate change within water resource modelling.

Diversity of events (duration, intensity, frequency, spatial pattern).

A diversity of different events is required to ensure that all systems are tested to events outside those experienced historically and avoid biasing the resulting investment decisions.

Applicable to multiple sectors.

National, regional and water company planning all require consideration of sectors other than water supply to determine impacts and costs or benefits related to investment (e.g., Environment Agency, 2020; WRSE, 2022). Consequently, datasets used in water supply planning should be available and relevant for other sectors.

Represent full range of uncertainties.

The majority of planning frameworks require an exploration of a large range of uncertainty, regardless of whether a bottom-up or top-down approach is used (Durant and Counsell, 2023). An idealised evidence base would incorporate the full range of uncertainties associated with hydrological variability, however this will need to take into account uncertainty across multiple interconnected spheres, including society, policy and economic. There is a need for coherence across these areas where they intersect, such as the urban heat-island effect increasing temperatures and driving water demand, or increased greenhouse gas concentration as a result of policy decisions driving weather variability. The uncertainties where variables have to be estimated or generated should also be taken into account.

Assessed prior to use and demonstrated.

Interviews undertaken as part of this study highlighted the importance of testing datasets and demonstrating their use within the planning framework prior to use. Ideally, the evidence base used for planning and operational uses will have been rigorously tested to identify issues and demonstrated so that planners may easily recognise how it could be applied to their system or problem within the regulatory framework.

Industry consensus as to their application and interpretation

Prior to their use, a consensus across the industry as to how they should be applied and interpreted in the context of investment planning would encourage consistency and efficiency for future investment plans.

Can be used in a risk-based approach

Risk-based approaches can be classified as those that trade risk off against an expected benefit. In order to be used in a risk-based approach, a dataset doesn’t need to cover all uncertainties, but does need to permit the comparison of risk (this could be the risk conditional on a particular set of events) and benefit across different interventions.

Ability to effectively attribute event likelihood

Event likelihood is an important factor in decision making because it allows an understanding of what events to plan to once a tolerable level of risk is established and because it is relative, it allows comparison between regions. Despite issues related to estimating drought likelihood (summarised in Durant and Counsell (2023), it can also aid in describing differences in severity and frequency between drought events.

Can be practicably applied for all systems and for all water stakeholder / receptors at a national scale across the industry

The practicality of implementing a dataset or technique across multiple stakeholders and receptors at a large scale across the water industry is a key limiting factor that cuts across the characteristics listed above. Regardless of all the above criteria, if the approach is too resource intensive, it is unlikely to be implementable.

The recommendations report summarises current datasets against these idealised criteria, presents a gap analysis and recommendations for progress against the timeline of significant future WRMP planning dates.

References

ATKINS, 2020. Regional climate data tools final report Sutton and East Surrey Water on behalf of WRSE.

BORGOMEO, E., PFLUG, G., HALL, J.W. AND HOCHRAINER-STIGLER, S., 2015. Assessing water resource system vulnerability to unprecedented hydrological drought using copulas to characterize drought duration and deficit. Water Resources Research, 51(11), 8927–8948. Available At: Https://Doi.Org/10.1002/2015WR017324.

BRISTOL WATER, 2022. Bristol Water draft water resources management plan 2024. Available At: Https://7850638.Fs1.Hubspotusercontent-Na1.Net/Hubfs/7850638/Dwrmp2024 Technical Report Finalv2 REDACTED.Pdf.

COWPERTWAIT, P.S.P., KILSBY, C.G. AND O'CONNELL, P.E., 2002. Space-time Neyman-Scott model of rainfall: empirical analysis of extremes. Water Resources Research, 38(8), 6-1-6–14. Available At: Https://Doi.Org/10.1029/2001wr000709.

DAWKINS, L.C., OSBORNE, J.M., ECONOMOU, T., DARCH, G.J.C. AND STONER, O.R., 2022. The advanced meteorology explorer: a novel stochastic, gridded daily rainfall generator. Journal Of Hydrology, 607(February), 127478. Available At: Https://Doi.Org/10.1016/J.Jhydrol.2022.127478.

DIAZ-NIETO, J. AND WILBY, R.L., 2005. A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the river Thames, United Kingdom. Climatic Change, 69(2–3), 245–268. Available At: Https://Doi.Org/10.1007/S10584-005-1157-6.

DURANT, M. AND COUNSELL, C.J.A., 2023. Annex L: Water Supply: Observed and Projected - Review of the state of research on drought. In Environment Agency Review of the research and scientific understanding of drought (pp. 548-600). Available At: https://assets.publishing.service.gov.uk/media/656f10920f12ef07a53e0229/Annex_to_the_review_of_the_research_and_scientific_understanding_of_drought.pdf

ENVIRONMENT AGENCY, 2015. Understanding the performance of water supply systems during mild to extreme droughts.

ENVIRONMENT AGENCy, 2020. Meeting our future water needs: a national framework for water resources, (March). Available At: Https://Assets.Publishing.Service.Gov.Uk/Government/Uploads/System/Uploads/Attachment_Data/File/872759/National_Framework_For_Water_Resources_Main_Report.Pdf.

ENVIRONMENT AGENCY, 2021a. Water resources planning guideline supplementary guidance: 1in 500. March 2021.

ENVIRONMENT AGENCY, 2021b. Water resources planning guideline supplementary guidance: climate change. March 2021.

ENVIRONMENT AGENCY, 2022. Review Of England's emerging regional water resources plans, 1–25. Available At: Https://Www.Gov.Uk/Government/Publications/Review-Of-Englands-Emerging-Regional-Water-Resources-Plans/Review-Of-Englands-Emerging-Regional-Water-Resources-Plans.

FOWLER, H.J., BLENKINSOP, S. AND TEBALDI, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling.

International Journal Of Climatology, 27(12), 1547–1578. Available At: Https://Doi.Org/10.1002/Joc.1556.

FOWLER, H.J., KILSBY, C.G. AND O'CONNELL, P.E., 2000. A stochastic rainfall model for the assessment of regional water resource systems under changed climatic condition. Hydrology And Earth System Sciences, 263–281. Available At: Https://Doi.Org/10.5194/Hess-4-263-2000.

FOWLER, H.J., KILSBY, C.G., O'CONNELL, P.E. AND BURTON, A., 2005. A weather-type conditioned multi-site stochastic rainfall model for the generation of scenarios of climatic variability and change. Journal Of Hydrology, 308(1–4), 50–66. Available At: Https://Doi.Org/10.1016/J.Jhydrol.2004.10.021.

HANNAFORD, TURNER, S., CHEVUTURI, A., CHAN., W., BARKER, L., TANGUY, M., PARRY, S., ALLEN, S. Have river flow droughts become more severe? A review of the evidence from the UK, a data-rich temperate environment. Hydrology and Earth System Science Discussions. Available at: https://hess.copernicus.org/preprints/hess-2024-293/

INTERVIEW 1, 2023. Interview. Conducted By Mason Durant, Chris Counsell and Jamie Hannaford. 23rd May 2023.

INTERVIEW 2, 2023. interview. Conducted By Mason Durant, Chris Counsell and Jamie Hannaford. 24th May 2023.

INTERVIEW 3, 2023. Interview. Conducted By Mason Durant, Chris Counsell and Jamie Hannaford. 27th June 2023.

INTERVIEW 4, 2023. Interview. Conducted By Mason Durant, Chris Counsell and Jamie Hannaford. 21st June 2023.

INTERVIEW 5, 2023. Interview. Conducted By Mason Durant. 22nd June 2023.

INTERVIEW 6, 2023. Interview. Conducted By Mason Durant and Jamie Hannaford. 27th June 2023.

INTERVIEW 7, 2023. Survey Response. 27th July 2023.

INTERVIEW 8, 2023. Survey Response. 20th June 2023.

JONES, P., HARPHAM, C., KILSBY, C., GLENIS, V. AND BURTON, A., 2010. Projections of future daily climate for the UK from the weather generator. Weather, (November), 44. Available At: Http://Ukclimateprojections.Defra.Gov.Uk.

KILSBY, C.G., JONES, P.D., BURTON, A., FORD, A.C., FOWLER, H.J., HARPHAM, C., JAMES, P., SMITH, A. AND WILBY, R.L. 2007. A daily weather generator for use in climate change studies. Environmental Modelling & Software, 22(12), 1705-1719. https://doi.org/10.1016/j.envsoft.2007.02.005.

OFWAT, 2006. The development of the water industry in England and Wales, Office Of Water Services [Preprint]. Available At: Https://Www.Ofwat.Gov.Uk/Wp-Content/Uploads/2015/11/Rpt_Com_Devwatindust270106.Pdf.

PRUDHOMME, C., SAUQUET, E. AND WATTS, G., 2015. Low flow response surfaces for drought decision support: a case study from the UK. Journal Of Extreme Events, 02(02), 1550005. Available At: Https://Doi.Org/10.1142/S2345737615500050.

SAUQUET, E., RICHARD, B., DEVERS, A. AND PRUDHOMME, C., 2019. Water restrictions under climate change: a Rhône–Mediterranean perspective combining bottom-up and top-down approaches. Hydrology And Earth System Sciences, 23(9), 3683–3710. Available At: Https://Doi.Org/10.5194/Hess-23-3683-2019.

SOUTHERN WATER, 2014. Water resources management plan 2015-40 - Technical Report. Available At: Https://Www.Southernwater.Co.Uk/Media/Default/Pdfs/WRMP-Technical-Report.Pdf.

THAMES WATER, 2020. Water resources management plan 2019 - Executive Summary', (April).

THAMES WATER, 2022. Draft drought plan 2022, (128). Available At: Https://Www.Anglianwater.Co.Uk/About-Us/Our-Strategies-And-Plans/Drought-Plan/.

UKWIR, 2017. Drought vulnerability framework. Report Ref. 17/WR/02/12.

WATER RESOURCES SOUTH EAST, 2017. An overview of the WRSE'. Available At: Https://Www.Wrse.Org.Uk/Media/Au1n1gbk/Wrse_File_278_An-Overview-Of-The-Wrse.Pdf.

WRSE, 2022. Futureproofing our water supplies - a consultation on our emerging regional plan for South East England, (November).

YORKSHIRE WATER, 2022a. Draft water resources management plan 2024.

YORKSHIRE WATER, 2022b. Drought plan 2022. Available At: Https://Www.Anglianwater.Co.Uk/About-Us/Our-Strategies-And-Plans/Drought-Plan/.

Existing stochastic methods – A critique of statistical aspects

Adam Griffin and Ye Liu (with review and input from the author team)

Overview

Statistical models generally concern the analysing of random variables that take value under a combined influence of known and unknown factors. For example, the hour-to-hour rainfall variability at a particular location is the combined consequence of the long-term hydroclimatic conditions, the medium-term weather condition, the short-term cloud movement, and other known or unknown processes that generate seemingly random deviation from the expected trend.

There are many names given to the part of the variability that can be attributed to the known and measurable processes and the part that cannot. Here we refer to them as the covariate effect and the residual effect respectively. This concept can also be demonstrated using a standard linear regression model where random variable Y is formulated as $Y = \beta X + \varepsilon$ where

- variable X (also known as the covariate) represents a factor that partially determines the value of Y under some rescaling provided by the linear coefficient β; and
- variable ε (also known as the error term) represents the residual, or the random deviation from the covariate effect of βX.

In a typical use case, a statistical model provides an approach for formulating the mathematical relation between the random variable, its covariate and residual effect, and subsequently performing statistical inferences for the parameters of the formulation. During the formulation, it is common to attribute as much of the variability of the random variable of interest to other known covariates as possible. But it might not always be feasible or desirable to model a known physical relationship as a covariate in a statistical model due to reasons such as lack of data support or the added complexity of parametrising or fitting the statistical model. Often this is left as an intentional choice of a statistical model with consideration of all modelling needs.

The input data (see Chapter 3), climate predictors for stochastic modelling (see Chapter 4) and climate change (see Chapter 5) review provides a useful insight into the importance and availability of relevant hydroclimatic processes that can potentially act as the covariate effect to the rainfall. This part of the review builds upon the insight gathered in the other sections and highlights the utilisation of input data, climate predictors and climate change factors in a statistical modelling framework.

The review of the statistical model starts with some discussion on the temporal resolution of the generated rainfall data in Section 2; Section 3 focuses on choices of model formulation from the approaches under review, including their account for the covariate effects and the structure imposed for the residual effect; Section 4 discusses some key considerations whilst implementing the reviewed approaches; finally Section 5 explains in more detail a selection of the key modelling techniques.

Temporal resolution

Motivation in the choice of temporal resolution

The temporal resolution of the output rainfall data underpins several other modelling choices in the stochastic rainfall generator. Here the temporal resolution refers specifically to the regular time interval over which the average rainfall volume is modelled and simulated.

In general the requirement for fine temporal resolution (e.g., sub-hourly) is most important for flood risk management applications. In contrast, rainfall generation for water resources planning tends to be over a coarser temporal resolution (e.g., daily or monthly). This is because the consequence caused by droughts is usually reflected over a much longer period of rainfall accumulation and variability on a finer temporal resolution does not post any major impact (Barker *et al.*, 2016). Some stochastic rainfall generators will be more suited for either of the two applications; whilst others may be versatile enough to cover both cases.

The choice of temporal resolution is also linked to the fact that the random variable of interest, i.e. rainfall depth over a certain duration, takes only non-negative values. If a probability distribution is used to capture all possible values of rainfall, it is desirable that the probability distribution have a positive probability (denoted p_0) of generating a zero value (i.e. dry spell) and a remaining $(1 - p_0)$ probability of generating some strictly positive value (i.e. the rainfall depth during a wet spell). This type of distribution is referred to in the statistics literature as a mixture model of a discrete (wet versus dry) and a continuous (rainfall depth when wet) subpopulation. The intermittent process model proposed in (Papalexiou, 2022) provides a typical formulation of this mixture distribution such that the rainfall is essentially a product of the binary wet-or-dry process and the continuous rainfall distribution.

When aggregating the short-duration rainfall over a longer time interval, the sum of rainfall deviates further from zero. When the aggregation time interval is sufficiently long (i.e. the temporal resolution sufficiently coarse), the resulting mixture model will carry a negligible probability p_0 of landing on the zero value, such that the aggregated rainfall depth can be practically modelled as a single continuous distribution (instead of a mixture model).

Implication of different temporal resolutions

This differentiation in the temporal resolution broadly divides the approaches under review into two groups:

- High temporal resolution, i.e. daily or sub-daily data – (Chun *et al.,* 2013) and (Dawkins *et al.,* 2022)
- Low temporal resolution, i.e. monthly data – (Serinaldi & Kilsby, 2012), (Atkins & WRSE, 2020)[1]

Incorporating a mixture distribution for the high-resolution rainfall requires additional parameters which contributes to the overall complexity of the stochastic generator. For example, (Dawkins *et al.,* 2022) uses a three-state Markov Chain approach to model the switching between wet and dry spells with two dedicated parameters (i.e. the transient probabilities between wet and dry states in the Markov Chain model). The Neyman-Scott Rectangular Pulses model, as reviewed in (Chun *et al.,* 2013) uses dedicated probability distributions to model the duration of wet or dry spells. This is in contrast to the approaches designed for low-resolution rainfall, e.g., (Serinaldi & Kilsby, 2012), where the intermittency of wet and dry spells are not captured explicitly.

It should be recognised that the added modelling flexibility to account for high-resolution intermittency of rainfall usually comes at a cost of modelling and fitting complexity. With limited computation power, this added complexity might lead to simplification of other parts of the generator, e.g., the daily rainfall generator behind UKCP09, as reviewed in (Chun *et al.*, 2013), does not account for the covariate effect in rainfall due to change in long-term hydroclimatic conditions. In comparison, the low-resolution approaches can be easily formulated to account for the long-term hydroclimatic conditions. (Dawkins *et al.*, 2022) explicitly captures the high-resolution intermittency as well as the long-term hydroclimatic conditions, at the cost of significantly more complex model parametrisation and lengthier model fitting process.

Upscaling or downscaling

On top of the primary temporal resolution, other durations are often of interest. In one direction, upscaling can take daily models and aggregate them into monthly or seasonal summaries. In the other direction, one can downscale to estimate daily or sub-daily time series from monthly outputs.

As mentioned above, (Serinaldi & Kilsby, 2012) focus on monthly time series rather than upscaling daily data to avoid “overdispersion” where monthly variability is underestimated due to differences in the long-term characteristics of daily outputs compared to monthly. Alternatives approaches using daily outputs (e.g., Mehrotra *et al.*, 2006) introduce a lot of extra complexity. Instead (Serinaldi & Kilsby, 2012) simply recommend the use of an alternative monthly model rather than upscaling, which may be more appropriate in many

[1] The final output from Atkins (2020) is daily rainfall data but the stochastic generation is technically over a monthly interval. The daily resolution is achieved through post processing as discussed in Section 2.3.

cases, but complicate the comparison between monthly and daily series as they come from different models.

In the case of downscaling from monthly to daily data as in (Atkins & WRSE, 2020), additional high-frequency data is required to reasonably simulate sub-monthly time series, especially if additional covariates are used. (Atkins & WRSE, 2020) use a small number of sample observed month profiles, which they draw on to simulate the days in a given month, simply by uniformly scaling the daily profile to match the generated month's total. Profiles are drawn which most closely match the monthly statistics (total rainfall depth, for example). In the case of monthly statistics exceeding most observations (in the lowest 20% of totals), one of the closest four months was chosen to introduce some variability in the extremes. No reasoning was given why four was chosen specifically, possibly suggesting it was via trial-and-error or expert judgement.

In all the current approaches, different up- or down-scaling approaches can be applied to the monthly/daily data which comes directly out. (Atkins & WRSE, 2020) suggests just one such approach. However, the different mechanisms (wet/dry days vs months where zero rain is impossible in the model) can cause inconsistencies when trying to harmonise daily and monthly generators.

Model formulation

The main aspect which differs between the different models in operation are the way that covariates are applied and incorporated into the model, and the mechanisms used to get to the final timeseries of rainfall.

Covariate effect

Some of the older models, such as those used in UKCP09, simply fit the model parameters to observed rainfall series directly, not making use of other covariates in the model. A problem with not sufficiently accounting for possible covariates is the implicit assumption that the observed data is both stationary and representative of the scenario one wishes to generate rainfall for, which is a very questionable assumption given evidence of nonstationarity in the climate system (Kim and Onof, 2020). If the observed records are short, then there is likely an underestimate of the stochastic variability of the precipitation process, and it may lead to an inability to generate events similar to the most extreme storms or droughts observed, leading to an impact on our understanding of variability, hence our understanding of drought severity. The precise period of observations can also lead to biases in the rainfall generator, which present in a lack of drought events or a lack of extreme storm events. Both can be partially mitigated through the use of well-calibrated/modelled future covariates and seasonality metrics when trying to capture multi-decadal climatic variability. Note this is not the case for trying to capture variability of specific observed periods. For the most extreme events, monthly or annual statistics may not capture the full variability of the covariates, and so covariates with a higher temporal resolution are required to better capture the relationship between rainfall and covariates, at the cost of greater computational cost. On the other hand, if only considering monthly timeseries outputs, then there is sometimes lower autocorrelation (Mills, 2005) in the rainfall series, reducing the need to account for it in a model.

Typically, in the discussed generators, covariates are selected *a priori*, rather than through a model selection process (e.g. stepwise regression). Villarini and Serinaldi (2012) do consider this approach using AIC to choose the most appropriate GAMLSS using NAO and SOI (and their lags) as covariates; nearly all seasonal models need some form of each. Dawkins *et al.,* (2022). No specific studies comparing non-covariate (e.g. Gaussian Noise) and covariate models have been found during this study.

As discussed in Chapter 4, NAO (and seasonal variants) and SST are the key covariates that come up in most of the models. The Generalised Linear Model in (Chun *et al.,* 2013) incorporated NAO into both the models of occurrence rate and magnitude of rainfall events, using cubic splines which are very flexible, but can lead to fitted models that are challenging to interpret. The extension in (Serinaldi & Kilsby, 2012) to a GAMLSS for the distribution of monthly total rainfall fits the covariates directly into the distribution parameters (location, scale, and shape) which can help with interpretability, and has the added benefit of being well supported in implementation and guidance for use. The problem with using NAO and SST as covariates for further generation of rainfall is the reliance on how NAO and SST are generated or applied. In (Serinaldi & Kilsby, 2012), NAO and SST were resampled from observations, rather than being parametrically simulated in their own right, which means that extremes beyond those observed are not available to feed into the rainfall generator itself. This is a common issue with this bootstrapping approach, which limits the generated values to those observed. A parametric model (drawing from a fitted extreme-value distribution based on the observed) may extend this range but justification of the extrapolation is required in application.

(Atkins & WRSE, 2020) and (Dawkins *et al.*, 2022) use similar (but many more) covariates, and cross-terms (products of, for example, NAO and SST), using a Bayesian penalised regression method, a systematic way of selecting covariate terms and cross-terms based on statistics of fit, which naturally sets the coefficients of covariates with low importance to zero. Again, due to the use of splines, these terms can become difficult to concisely describe, and neither (Atkins & WRSE, 2020) nor (Dawkins *et al.*, 2022) specify which terms remain in the model after the penalised regression; the code suggests that all the terms are kept (though some coefficients were near-zero). (Dawkins *et al.*, 2022) also makes use of a random effects component which attempts to describe the impact of other unobserved covariates that have an effect on the rainfall magnitude.

Across all the models, most of the covariate effect is constrained to the rainfall magnitude component of the generators. The wet-dry day mechanisms often only use seasonal cyclic patterns, although the GLM (Chun *et al.*, 2013) also uses NAO, and the HMM (Dawkins *et al.*, 2022) used Winter NAO in describing the transitions between wet and dry states. The HMM fits separate models for the wet-dry transition probabilities and the waiting time at each state (dry, wet and very wet).

For more information, we refer to the climate predictors chapter (see Chapter 4) and the climate change chapter (see Chapter 5).

Residual effect

The residual effect is collectively the causes of variability in the generated rainfall data that are not explicitly captured in the covariate effect. As explained earlier, the residual effect is potentially attributable to known physical processes but has been intentionally chosen by the modelling approach to remain as an independent source of variability. It should be noted that the residual effect is not the same as pure random noise. In fact, there is usually a complicated structure imposed on the randomness to capture the observed spatial and temporal coherence in the rainfall data.

Temporal coherence

The temporal coherence in rainfall can be observed and therefore captured by statistical models over different time horizons. As illustrated in the above section on ‘covariate effects’, it is considered relatively easy to attribute medium-to-long-term temporal variability to covariate effects. In most cases where the covariate effect is present, the covariates are used to formulate the probability distribution of the rainfall for each timestep, e.g., the GLM method in (Chun *et al.*, 2013) uses GCM outputs and the North Atlantic Oscillation (NAO) Index to form the probability of raining and the probability distribution of the amount of rain (if it rains) per timestep. The WRSE model (Atkins, 2020) does similarly.

In many cases, the temporal correlation (or the autocorrelation) of the rainfall is captured by constructing the probability distribution based on the rainfall values from the previous timesteps. For example, (Dawkins *et al.*, 2022) uses a first-order Markov Chain to capture the temporal correlation which means the daily rainfall is generated based on the status of the previous day. The GLM approach proposed by (Chandler & Wheater, 2002) and reviewed by (Chun *et al.*, 2013) uses rainfall up to 7 days in advance to determine the distribution of rainfall (occurrence and amount).

On a monthly temporal resolution, (Serinaldi & Kilsby, 2012) found that the temporal correlation is mostly already captured as a covariate effect by the autocorrelation in the hydroclimatic factors. Depending on whether this is the case, they propose a non-parametric bootstrap resampling approach (for no residual autocorrelation) or the GAMLSS model coupled with a Gaussian copula (for residual autocorrelation).

A carefully modelled autocorrelation is important as extreme drought events are a consequence of persistent sampling of zero or low rainfall values over an extended period of time. For high-temporal-resolution approaches (where the number of low-rainfall simulations needs to be high due to the increased number of timesteps) or approaches without a climate-drive covariate effect (where the probability distribution of rainfall does not shrink in dry climate conditions), the capability for model to persistently draw samples from the lower-end of the rainfall distribution becomes paramount.

Spatial coherence

In general, the spatial coherence in the rainfall refers to the fact that the rainfall over multiple spatial locations can be correlated. These locations may be individual rain gauges where the rainfall data is collected, or individual cells in a regular grid over an area. The full extent of the correlation may be the consequence of both the covariate effect (i.e. all locations fall under the same climate or weather system) and residual effect (i.e. some localised factor that affects only nearby locations). The covariate effect is discussed above, so this section only covers the latter.

The spatial coherence is only relevant to models that can be extended to multiple sites, i.e. (Atkins & WRSE, 2020; Dawkins *et al.*, 2022; Serinaldi & Kilsby, 2012). All three models use a variation of copula models (i.e. a multivariate probability distribution with transformed marginal distributions). The two methods as stated and reviewed in (Chun *et al.*, 2013) do not currently cover multiple sites but the GLM method is potentially extensible to do so, as per (Yang *et al.*, 2005). Subsequent developments have been made to the NSRP model (Burton *et al.*, 2008)

Accounting for the spatial coherence is important because a biased estimation of the spatial correlation will cause the areal aggregation to be biased as well. In particular, an overestimated correlation means the areal rainfall is more extreme than reality, because the worst wet or dry patches are generated to occur concurrently more often than expected. By the same logic, an underestimated spatial correlation means underestimated extremes.

There are numerous options for quantifying the correlation between multiple random variables, e.g., the linear correlation coefficient. In a study of extreme values, the extremal dependence is arguably one of the most relevant measures. (Ledford & Tawn, 1996, 1997) highlighted two main types of extremal dependence, i.e. asymptotic dependence and asymptotic independence. This classification underpins the difference in the probability of observing concurrent (or joint) extreme values over multiple locations.

Within the three relevant approaches, (Serinaldi & Kilsby, 2012) and (Atkins & WRSE, 2020) use a Gaussian copula whereas (Dawkins *et al.*, 2022) uses a vine copula. The Gaussian copula is easy to fit and performs relatively well for non-extreme-value-focused analysis. This is due to the fact that the Gaussian copula relies only on the pairwise linear correlation and is most suited for capturing the overall correlation. On the other hand, it makes implicit assumptions that extreme values at all locations are asymptotically independent, which means the probability of concurrent large values at multiple locations is enforced to converge to zero. This assumption could be restrictive and it would be preferable to perform some statistical test of extremal dependence beforehand. Alternatively, a more flexible approach could be to use a copula model that can represent both types of extremal dependence. The Gaussian copula also makes the assumption that the correlation between variables is constant for normal or extreme events, which again would require additional checking.

The vine copula is a more sophisticated copula model and benefits from its capability of modelling both types of extremal dependence. (Simpson *et al.*, 2021) gives a detailed study of the tail behaviour of the vine copula and finds it to be sufficiently flexible on a practical level. The main drawback of the vine copula is associated with the formulation (of the vine dependence structure) and the complex fitting process. When combined with the rest of the stochastic generator model components, the fitting of a vine copula is computationally intensive, and potentially prohibitively so when considering high spatial dimensions (see Section 4.2 for some further discussion).

Key implementation consideration

EVA checking

The extreme value theory (EVT) is a collection of theory and methodology centred around the description of the statistical properties of rare events. These rare events (e.g., severe droughts) are typically characterised by certain associated random variables (e.g., rainfall deficit or cumulative rainfall) reaching an extreme level (e.g., high deficit or low cumulative rainfall).

The extreme value analysis (EVA) refers to the study of observed rare events based on the EVT. A distinct feature of the EVA, as opposed to the standard distribution fitting, is that the EVA focuses on the tail area of the probability distribution and is used to provide a scientific extrapolation to quantiles beyond the range of the data, e.g., the estimation of a 1,000-year event (of 0.001 annual exceedance probability, or AEP) based on 50 years of data. In this example, the lowest AEP that can be estimated empirically is approximately 1/50 (or 0.02) which is still greater than 0.001, so the EVA needs to be performed. This 1/50 estimate possibly a poor estimate.

When a stochastic rainfall generator is validated, it is desirable that the probability of encountering severe droughts within the stochastic series matches the historical data. Ideally the consistency should be observed over different return periods (RPs) or annual exceedance probabilities. Within this aim, an EVA check is observed in most of the reviewed literature as a standard model or outcome validation step. In most cases, the EVA check follows these key steps:

1. Identify individual drought events based on the rainfall or a derived drought-related index
2. Define the key variables (i.e. attributes) of interest for the identified drought events
3. Estimate the extreme values of the key variable over a range of RPs or AEPs
4. Apply step 1 to 3 to both the observed and the stochastic rainfall series and compare the outcome of step 3

Step 1 – event identification

The EVA usually requires the data under analysis to be independently and identically distributed (IID). Therefore, it is important to identify individual drought events from the stochastic or observed rainfall series. The identification of drought events can be performed based on various drought-related indices derived from the rainfall.

- (Chun *et al.*, 2013) and (Serinaldi & Kilsby, 2012) both use the Drought Severity Index (DSI) as proposed by (Phillips and McGregor, 1998) to identify events. In particular (Serinaldi & Kilsby, 2012) uses a range of time scales for the DSI calculation.
- (Dawkins *et al.*, 2020) uses the 36-month rainfall Deficit Drought Index (DDI).
- (Atkins & WRSE, 2020) uses the summer average Rainfall Deficit Index (RDI).

Figure X-1 shows an example of the calculated DSI3 index (DSI over a three-month period) based on the observed or the stochastic rainfall series.

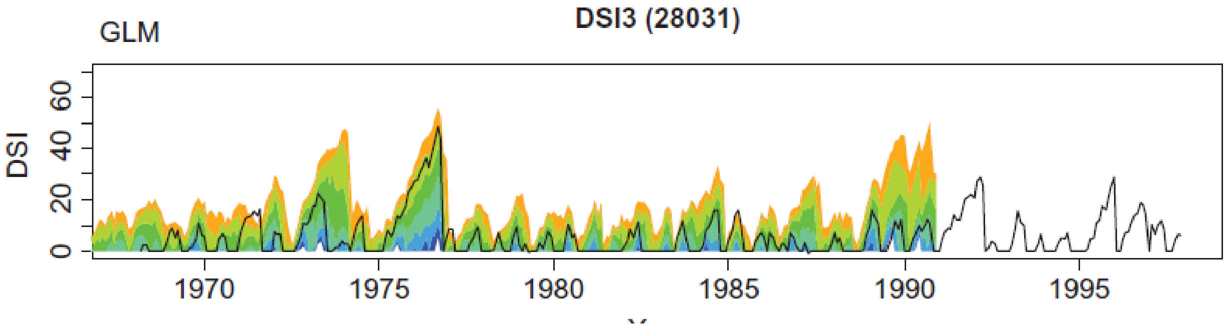

Figure 1: DSI3 series based on the observed (solid line) versus the stochastic (quantiles shown in colours) rainfall at the NRFA station 28031 (source: Figure 6 from Chun *et al.*, 2013).

Typically a drought event is considered to have initiated when the corresponding drought-related index exceeds a certain threshold. Similarly the event is considered to have terminated when the index falls below the threshold. This identification method is similar to the peaks-over-threshold (POT) declustering method seen in the EVT literature such as Smith & Weissman, 1994

Whilst this is a theoretically justified method for identifying IID extreme events, the outcome of the identification can be sensitive to certain parameters used in the process, namely the threshold or the event separation window (also known as the Smith's runs period (Smith & Weissman, 1994)). In particular an event separation window of D days, say, would merge two identified events into a single event if the start of the second event is less than D days away from the end of the first event. The papers which describe current practice have very little discussion around the choice or the implication of the threshold and event separation window. However, more generally several papers have discussed choices of threshold and separation window in the scope of drought indices. (Tallaksen *et al.*, 1997) tests the sensitivity of extracted drought events to variation in the event separation window (or inter-event time as they called it). (Fleig *et al.*, 2006) similarly tests sensitivity of drought deficit volume and duration to choice of different inter-event time windows. (Parry *et al.*, 2016)(Parry *et al.*, 2016)(Parry *et al.*, 2016)(Parry *et al.*, 2016)(Parry *et al.*, 2016)(Parry *et al.*, 2016)(Parry *et al.*, 2016) tests sensitivity of drought termination characteristics to choice of several parameters for event selection.

Step 2 – definition of key variables

Various attributes such as the duration, magnitude and severity of droughts have been used in the literature. In general an EVA check based on a wider variety of statistics would provide better assurance of the quality of the stochastic series (Serinaldi & Kilsby, 2012 and Dawkins *et al.* 2022).

(Chun *et al.*, 013) and (Atkins & WRSE, 2020) only perform the EVA check based on the original drought index. (Dawkins *et al.*, 2022) used three attributes – magnitude, duration, and severity. (Serinaldi & Kilsby, 2012) used four attributes – drought maxima, severity, duration, and interarrival time.

Step 3 – estimation of extreme values

Both non-parametric (empirical) methods and parametric methods have been used in the literature to obtain estimates of extreme levels for the drought variables.

The non-parametric method refers to statistical estimation without making explicit assumptions about the underlying distribution of the studied variable. They include the inverse-ranking method, plotting position (e.g., Weibull method or Weibull plotting position) as used in the reviewed literature. These methods are more reasonable for RPs shorter than the total length of the data but cannot be used for extrapolation to higher RPs.

(Atkins & WRSE, 2020) discusses the option of joining the stochastic series to form a single long series. This is potentially a viable option but the underlying probabilistic assumptions need to be justified, depending on how covariates are used.

On the other hand, parametric methods refer to the use of certain probability distributions to capture the statistical behaviour of the studied variable. These methods allow extrapolation to high RPs (beyond the length of the data) but require additional complex distribution fitting exercise. For example, (Atkins & WRSE, 2020) highlights the fact that multiple fitting methods could potentially be used.

(Atkins &WRSE, 2020) is the only study that presents an EVA including parametric estimation of extreme values. However the distributions used in the study is not in line with the best practice of the EVA, e.g., the only theoretically justified distribution for the POT values is the Generalised Pareto Distribution (GPD), not the Weibull or the Generalised Extreme Value (GEV) distribution as used in that study (Coles, 2001).

Step 4 – comparison of estimated extreme values

The output from step 3 is usually in the form of a curve whereby the x-axis is the AEP (or RP) and the y-axis is the estimated extreme levels of the studied variable or drought attribute. Sometimes the axes may be swapped. For the observed or historical data, there is a single curve per variable. For the stochastic series, there is one curve for each individual stochastic series so the outcome is a collection of curves and quantiles are usually taken to form a confidence interval (CI). A standard operation as seen in the literature is to compare the relative position of the observed curve with the full range or the CI from the stochastic series. In an ideal situation, the observed curve should lie comfortably within the CI whereas the full range of the stochastic series should provide more or less extreme versions of the observed curve.

Figure X-2 gives an example of such a comparison based on extreme level rainfall. Similar figures can be found in (Chun *et al.*, 2013), (Serinaldi & Kilsby, 2012), (Atkins & WRSE, 2020) and (Dawkins *et al.*, 2022).

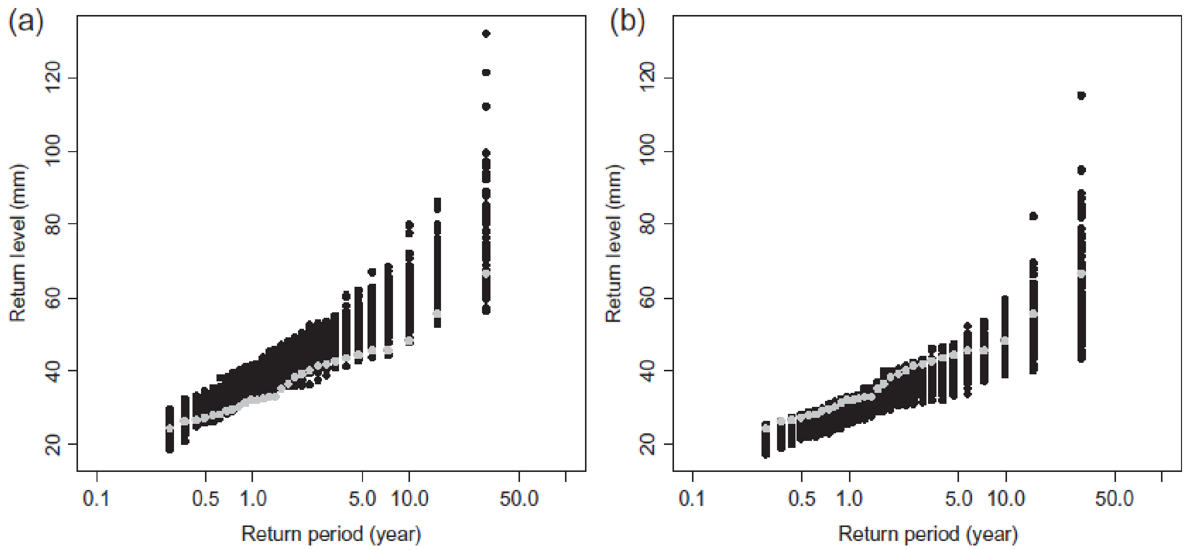

Figure 2: Illustration of an EVA comparison between observed and simulated extreme rainfall levels (source: Fig.5 Chun *et al.* 2013).

Regardless of which method is being used (parametric vs non-parametric) in step 3, it should be recognised that there are uncertainties associated with the estimated values. For example, the inverse-ranking method assigns an N-year RP to the worst event the N-year period, whereas there is approximately 50% chance that the true RP of the worst event is greater than 2N or less than N/2 based on probabilistic calculation. Similar model fitting uncertainty exists with the parametric methods too.

For a robust EVA checking, it is recommended that the output from the extreme value estimation be considered jointly with the underlying estimation uncertainty. So instead of a single curve, a presentation involving a CI for the observed data and each individual stochastic series should be considered. The latter will be challenging to visualise so a simple convolution of distributions can be used to combine the CIs from all the stochastic series. This additional consideration might lead to additional insight given the large uncertainty caused by the relatively few data points per time series.

Computation / implementation costs

A key reason to use stochastic rainfall generators instead of using process-based meteorological models is to reduce the computation time or resources. Many of the models which have been used in practice, (Atkins & WRSE, 2020; Jones *et al.*, 2010; Serinaldi & Kilsby, 2012) involve quite simple models which can be combined with an even simpler random seed generator to generate the stochastic rainfall series. This is beneficial for smaller companies and projects where resources (time, computational power) may be limited. On the other hand, (Dawkins *et al.*, 2022) admits to requiring a large amount of computation time to simulate rainfall at more than a few sites simultaneously, due to high number of parameters being fitted with complex dependencies. For the HMM portion of the (Dawkins *et al.*, 2022) model, it took 40 hours, 8 CPUs and 50GB of memory for each of the 39 sites (over 38000 days per site). For all other methods, the amount of time and memory required is very small, mostly consisting of the size of the input data being used.

On top of this, while there are few decisions required in fitting the GAMLSS of (Serinaldi & Kilsby, 2012), the monthly downscaling of (Atkins & WRSE, 2020), or the GLMs of (Chun *et al.*, 2013, there is additional statistical knowledge or expertise to visually inspect outputs of the Markov Chain Monte Carlo methods of (Dawkins *et al.*, 2022; Jones *et al.*, 2010), where prior distributions and acceptance probabilities must be adjusted to give reasonable results. This should be taken into account depending on the size of the study region in future projects.

Set up of simulations under different climate scenarios

There are three approaches in principle to handle the impact of different climate scenarios on the generation of rainfall and consequently the drought event data.

The first approach post-processes the stochastic model output, while the second approach is based on transforming the rainfall value of an initial data set to match the statistics from a target climate scenario. In both approaches the transformation is guided by a set of change factors derived by calculating the difference between baseline observations with future climate scenarios (need cross-referencing to the CC review).

The third approach explicitly captures the impact of climate scenario in the stochastic generation process through the use of the covariate effect (Section 3.1). For example, (Serinaldi & Kilsby, 2012) uses the NAO and SST index to capture the climate impact and (Atkins & WRSE, 2020) extends this to a set of eight hydroclimatic drivers. (Chun *et al.*, 2013) and (Dawkins *et al.*, 2022) also present explicit modelling mechanism to capture the impact.

Under the third approach, an ensemble of hydroclimatic covariates from the required climate scenario will be used to pre-parameterise the stochastic generator so that the resulting stochastic rainfall series would endorse the desired climate pattern. When the total number of simulations is constrained, there might a trade-off between the size of the ensemble and the number of simulations per ensemble member. However the former is usually pre-determined by the source of the hydroclimatic data.

In contrast to the handling of future climate uncertainty (via the ensemble of hydroclimatic conditions), our review finds that the stochastic generation under the present day scenario does not sufficiently explore the uncertainty embedded in the hydroclimatic condition. This uncertainty arises when the historical hydroclimatic conditions are assumed to be a fair, or in some cases the only, representation of the present day climate.

For example, the AME framework from (Dawkins *et al.*, 2022) reuses the same SNAO, WNAO, EA and SST values as observed for present day rainfall simulation. (Serinaldi & Kilsby, 2012) uses a resampling scheme to draw alternative samples from the observed NAO and SST values, which creates relatively more temporal variability but is still confined to the observed range of values. The consequence from such approaches is essentially an under-sampling of variability, as acknowledged in (Dawkins *et al.*, 2022), and highlighted by some existing reviews such as (Beven, 2021). A potential fix for this might be to echo the treatment of future climate uncertainty and generate an ensemble of present-day hydroclimate conditions allowing for more and less extreme values. Both statistical or physics-based models could be considered for the generation of this present-day ensemble.

For in depth discussion of limitations and recommendations for applying climate change scenarios for hydrological projections, please see Chapter 5.

Recommendation

1. **Select an appropriate stochastic generator with consideration of the modelling mechanism suitable, spatial and temporal resolution of the output, as well as the overall computational complexity of model fitting and simulation.**

Benefit: The choice of the modelling approach is made in an optimal way so that the modelled features are of direct impact on the problem to resolve, and the model complexity is compatible with the modelling and computation resources available.
Effort: *Low.*
Due to the different consideration and model formulation for the daily or monthly models, the users should consider carefully the most appropriate temporal resolution based on the problem they are trying to solve. This is particularly important as it is challenging to ensure the statistical consistency between the original stochastic series and the upscaled or downscaled series. This decision should be made jointly with the data availability and the user's appetite for model complexity. However it could be argued that a monthly timestep would be more suited for water resources planning purposes provided all other factors are indifferent between the two choices.

Additionally, the mechanics of the different stochastic generators are suited to different applications: drought events, long-term average rainfall patterns, flood events. Therefore, the stochastic model should be chosen to most appropriately reflect this. For example, the Markov model of (Dawkins *et al.*, 2022) has a focus on a rain/no-rain daily model, which is more suited to identifying short-duration rainfall events, rather than long-duration drought events.

A comprehensive collection of key features such as high temporal resolution, spatio-temporal coherence, and reliable covariate effects can jointly lead to a high model complexity and consequently prohibitive computation cost. The users are encouraged to take care when specifying their modelling needs and place their resources accordingly.

2. Follow a rigorous evidence-based process to select the most optimal set of covariates.

Benefit: The covariate-driven model is supported by the most relevant set of covariates, which are selected based on a rigorous, justifiable and repeatable process. The influence of any spurious correlation between the covariate and output rainfall is minimised.
Effort: *Moderate.*
Within the reviewed literature, a lot of work has been done to propose suitable covariates for the stochastic generation. A rigorous process should be followed to select the most optimal set of covariates by considering their statistical properties, such as the significance of the correlation between the covariates and the output rainfall, as well as the practical implications, e.g., the reliability of the input data, etc. The user could refer to common statistical model selection techniques for suitable hypothesis testing procedures to use (Davison, 2003).

3. Understand and communicate clearly the implication of resampling the covariates over the same observation period.

Benefit: Unintended biases in the frequency of extreme drought events is avoided.
Effort: *Moderate.*
If a chosen stochastic generation method is driven by covariates, it is of paramount importance to understand the implication of reusing or resampling the covariates over the same observation period. In general, the most extreme covariate value in an N-year period may have a very different natural recurring frequency than once-in-N-years. The users of a covariate-driven stochastic generation method are strongly recommended to consider and test the sensitivity to rare observations, e.g., by placing them on a higher or lower natural frequency. The assumption that the observation period (of the covariates) is an unbiased reflection of the simulated period needs to be backed by compelling evidence.

4. Follow the best practice of the EVA with clear quantification of model fitting uncertainty when estimating the return period or return level of rare events.

Benefit: The extrapolation of extreme (low) rainfall beyond observed range is derived using rigorous and theoretically justifiable approaches.
Effort: *Low to Moderate.*
The review also finds that the best practice of the EVA is not always followed, nor is the uncertainty of the model fitting fully captured. The related statistical methods are mature and abundant, and these two aspects ought to be well covered for any EVA checking of the stochastic series.

5. Consider the use of multivariate statistical models that are more suited for capturing the tail dependence between the modelled variables in the spatio-temporal processes.

Benefit: The estimation bias is minimised for measures that are a function of multiple direct generator outputs (e.g., the total rainfall over a few modelling cells, or the number of consecutive months when the rainfall is below an extremely low threshold) especially when one or more of these generator outputs are extreme.
Effort: *High.*
Over the past decade, several spatio-temporal rainfall generation methods have been proposed. They are typically fitted to the observation data so that a pre-defined set of statistics are kept consistent between the observation and the simulated data. When extreme drought events are concerned, it can be argued that the selection of statistics ought to include those better reflecting the tail dependence of multiple distributions (e.g., rainfall at different spatial locations or at different timesteps).
(Ledford and Tawn, 1996, 1997) provide a good summary of the theory and key statistics around tail dependence. The conditional approach proposed in (Heffernan and Tawn, 2004) or the wider group of suitable copula-based models (Joe, 2014) can be considered for explicitly capturing the tail dependence. It should be noted that most copula models make implicit assumptions about the type of tail dependence so the users should test the tail property of the generated data and ensure its consistency with the observation data.

One key parameter that underpins the joint extrapolation of multivariate (or spatial) random variables to extreme levels is termed the coefficient of tail dependence as defined in (Ledford and Tawn, 1996, 1997). The coefficient of tail dependence measures the relative rate at which two variables converges to the upper limit of their respective marginal distribution. It is recommended that the users verify the supported range of this coefficient of the chosen distribution when selecting a suitable model. This is potentially a challenging task due to the asymmetry of dependence observed in many spatial processes, e.g., variable A is more dependent on B than B on A in the joint tail area of the distribution.

References

ATKINS, & WRSE, 2020. Regional climate data tools final report Sutton and East Surrey Water on behalf of WRSE.

BURTON, A., KILSBY, C.G., FOWLER, H.J., COWPERTWAIT, P.S. P., AND O'CONNELL, P.E., 2008. RainSim: A spatial–temporal stochastic rainfall modelling system. Environmental Modelling & Software, 23(12), 1356–1369. https://doi.org/10.1016/J.ENVSOFT.2008.04.003

CHANDLER, R.E., AND WHEATER, H.S., 2002. Analysis of rainfall variability using generalized linear models: A case study from the west of Ireland. Water Resources Research, 38(10), 10-1-10–11. https://doi.org/10.1029/2001WR000906

CHUN, K.P., WHEATER, H.S., AND ONOF, C., 2013. Comparison of drought projections using two UK weather generators. Hydrological Sciences Journal, 58(2), 295–309. https://doi.org/10.1080/02626667.2012.754544

DAWKINS, L.C., OSBORNE, J.M., ECONOMOU, T., DARCH, G.J.C. AND STONER, O.R., 2022. The Advanced Meteorology Explorer: a novel stochastic, gridded daily rainfall generator. Journal of Hydrology, 607. https://doi.org/10.1016/j.jhydrol.2022.127478

FLEIG, A.K., TALLAKSEN, L.M., HISDAL, H. AND DEMUTH, S., 2006. A global evaluation of streamflow drought characteristics. Hydrology and Earth System Sciences, 10(4), 535–552. https://doi.org/10.5194/hess-10-535-2006

JONES, P., HARPHAM, C., WADE, S., WALLINGFORD, H. AND ROB WILBY, O., 2010. UK Climate Projections science report: Projections of future daily climate for the UK from the Weather Generator Acknowledgements National review group International review group.

LEDFORD, A.W. AND TAWN, J.A.,1996. Statistics for near independence in multivariate extreme values. Biometrika, 83(1), 169–187. https://doi.org/10.1093/biomet/83.1.169

LEDFORD, A.W. AND TAWN, J.A. (1997). Modelling Dependence within Joint Tail Regions. Journal of the Royal Statistical Society Series B: Statistical Methodology, 59(2), 475-499. https://doi.org/10.1111/1467-9868.00080

PAPALEXIOU, S.M. (2022). Rainfall Generation Revisited: Introducing CoSMoS-2s and Advancing Copula-Based Intermittent Time Series Modeling. Water Resources Research, 58(6). https://doi.org/10.1029/2021WR031641

PARRY, S., PRUDHOMME, C., WILBY, R.L., AND WOOD, P.J., 2016. Drought termination: Concept and characterisation. Progress in Physical Geography: Earth and Environment, 40(6), 743–767. https://doi.org/10.1177/0309133316652801

SERINALDI, F. AND KILSBY, C.G., 2012. A modular class of multisite monthly rainfall generators for water resource management and impact studies. Journal of Hydrology, 464–465, 528-540. https://doi.org/10.1016/j.jhydrol.2012.07.043

SIMPSON, E.S., WADSWORTH, J.L. AND TAWN, J.A., 2021. A geometric investigation into the tail dependence of vine copulas. Journal of Multivariate Analysis, 184, 104736. https://doi.org/10.1016/j.jmva.2021.104736

SMITH, R. L. AND WEISSMAN, I., 1994. Estimating the Extremal Index. Journal of the Royal Statistical Society: Series B (Methodological), 56(3), 515–528. https://doi.org/10.1111/j.2517-6161.1994.tb01997.x

TALLAKSEN, L.M., MADSEN, H., AND CLAUSEN, B., 1997. On the definition and modelling of streamflow drought duration and deficit volume. Hydrological Sciences Journal, 42(1), 15-33. https://doi.org/10.1080/02626669709492003

YANG, C., CHANDLER, R.E., ISHAM, V.S., & WHEATER, H.S., 2005. Spatial-temporal rainfall simulation using generalized linear models. Water Resources Research, 41(11). https://doi.org/10.1029/2004WR003739

Climate input data

Maliko Tanguy and Amulya Chevuturi (with review and input from the author team)

Introduction

Stochastic weather generators rely on mimicking the behaviour of the observed weather data using statistical relationships and lack any explicit physical basis (Jones *et al.*, 2010). Thus, these approaches are completely dependent on the variability of the input data. The quality of the input data feeding into weather generators is therefore paramount, with larger sample size of input data providing higher confidence in the synthetic time series generated. This requires benchmarking against accurate and long observational records (Barker *et al.*, 2019). In this chapter, we present insights on observational datasets, drawing examples from various sources, including HadUK, CEH-GEAR and rainfall rescue datasets. These examples are used to highlight key issues related to climate input data, even though these datasets may or may not be used for stochastic studies.

Input data quality and homogeneity of long time series

Given the significance of input data in stochastic methods, most studies in this domain meticulously curate their datasets, by comparing different observational datasets and model simulations for selected sites over UK (Atkins, 2020). Extending the observational record can be beneficial to constructing stochastic models. For example, Serinaldi and Kilsby (2012) used rainfall data from 1860-2002 over six catchments. However, the rainfall data (Jones *et al.*, 2004; 2006) originates from two distinct sources: raingauge data from 1860-1980 (Jones, 1984) and the UK Met Office 5km gridded data from 1980-2002 (Perry and Hollis, 2005). Despite previous testing for homogeneity issues prior to merging these datasets (Jones *et al.*, 2004), we have concerns that still persist regarding potential data disparities: the earlier period comes from point observations, whereas the latest part is taken from a gridded dataset obtained from interpolation (please see the detailed discussion below). Some stochastic methods have incorporated long timeseries of rainfall from gauge data (e.g., the Generalized Linear Model (GLM) approach of Leith *et al.* 2005) and this approach was originally tested using data from Heathrow, Birmingham and Manchester airports (Leith, 2005). However, accurate and long observed record of rainfall data are only available over few locations in the UK, and thus the stochastic studies use multiple sources of datasets. This issue becomes even more relevant as recent stochastic studies have shifted from regional analyses to nationwide assessments, increasingly relying on gridded datasets rather than point observations.

Typically, for any hydrological application, any errors or uncertainties associated with the input data propagate in any subsequent analysis, but especially in stochastic methods, which are so dependent on the input data (Jones *et al.*, 2010). Input rainfall data was shown to be one of the most important sources of uncertainty in hydrological applications, accounting for up to 50% of the error in river flow simulations (Bardossy *et al.*, 2022). This is particularly important as there are well-known weaknesses to the input data; observed rainfall derived from weather stations/gauges (Lewis *et al.*, 2018) or radar systems (Chan *et al.*, 2018) have well-documented quality issues. Raingauge data, which are widely considered the closest estimate to 'the truth', can have large uncertainties, mostly from under catch issues which can reach 10% in winter in the UK, resulting in a mismatch between the amount of rain or snowfall collected by the raingauge and the actual amount reaching the surface. In addition, changes in instrumentation and standard installation practice raises concerns over the homogeneity of certain long time series (Muchan and Dixon, 2019). Moreover, under catch of precipitation totals in standard rain gauges has been shown to be even greater with snowfall (e.g., Colli *et al.* 2020). In particular, for the England and Wales Precipitation (EWP) time series, snow was prevalent in the earlier part of the record, before the effects of anthropogenic warming became noticeable. It is therefore believed that winter precipitation is significantly underestimated in the earlier part of the record, widely overestimating winter wetting trends in the 20th century in England and Wales (Murphy *et al.*, 2020). Thus, it is essential to consider the limitations of the input data before using them to train stochastic models, to minimize and estimate uncertainty in our applications.

Stochastic studies expecting to work with long rainfall record need improved data over the earlier period of the rainfall records over the UK, especially before 1910. A recent initiative has significantly improved the rainfall record for the earlier years through the participation of citizen scientists to digitize millions of hand-written observations for hourly to daily rainfall (and other variables) from the UK National Meteorological Archive (Hawkins *et al.*, 2019; Craig and Hawkins, 2020). This project has substantially increased the data points from rainfall gauges before 1960 over the UK (Hawkins *et al.*, 2022; Figure 1), especially compared to the rainfall gauges used for CEH-GEAR (Tanguy *et al.*, 2021; Figure 2) and the HadUK rainfall dataset (Hollis *et al.*, 2019) version 1.0 (Figure 3). The additional data has shown to improve the representation of total monthly rainfall and individual storm representation (Craig and Hawkins, 2020).

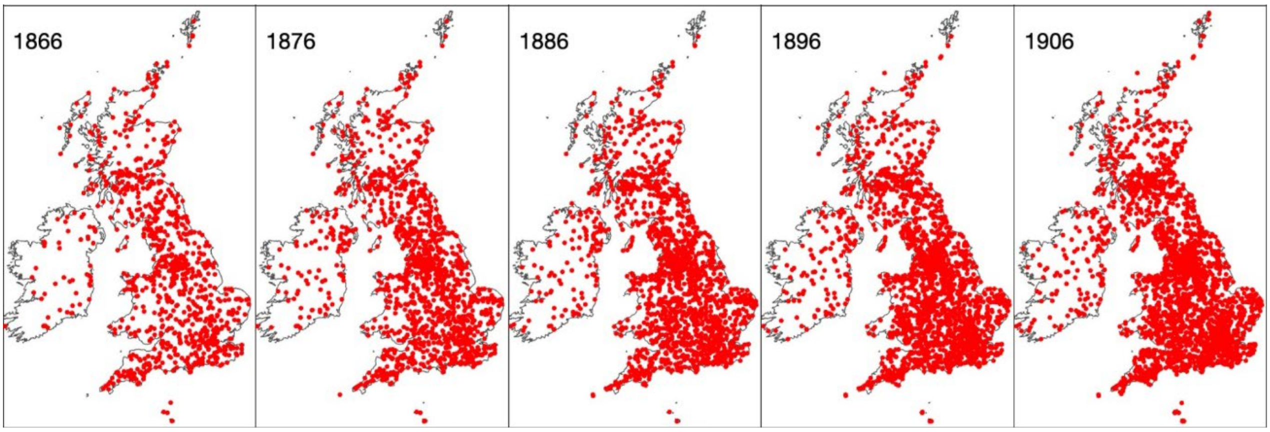

Figure 1: Density and distribution of rainfall gauges in v1.1.0 of the Rainfall Rescue dataset for specific years between 1866 to 1906. [Source: Hawkins *et al.*, 2022].

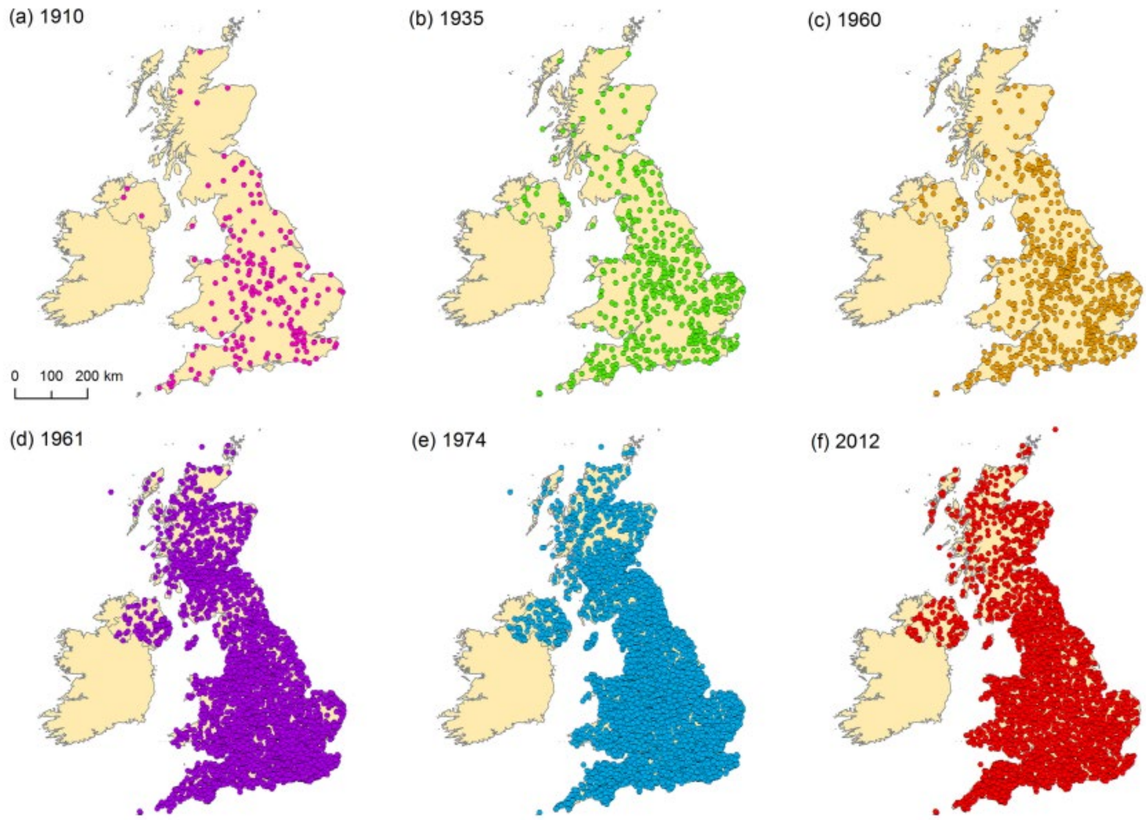

Figure 2: Density and distribution of rainfall gauges used to derive CEH-GEAR daily rainfall on 1st of January of (a) 1910, (b) 1935, (a) 1960, (b) 1961, (a) 1974, (b) 2012 [Source: Keller *et al.*, 2015].

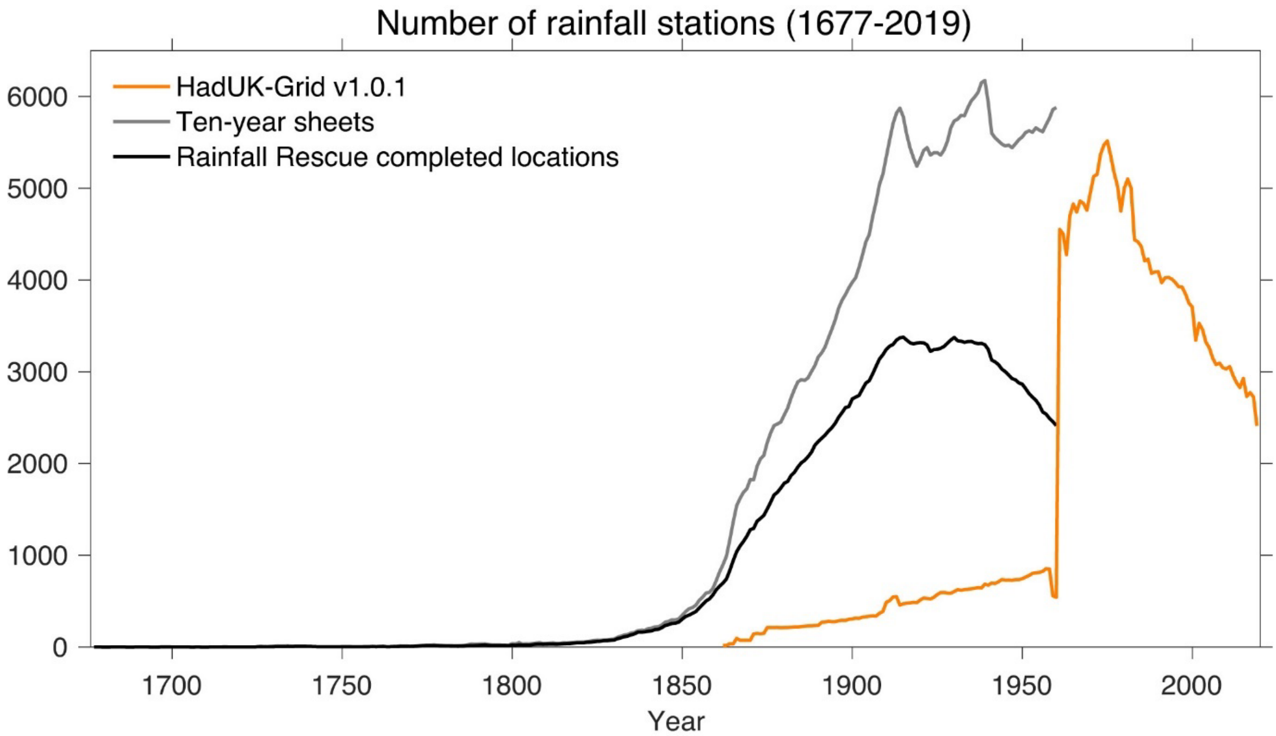

Figure 3: Number of rainfall stations available in HadUK-Grid v1.0 (orange), the raw Rainfall Rescue observations (grey), and v1.1 of the Rainfall Rescue dataset (black) between 1677 and 2019 at monthly timescales. [Source: Hawkins *et al.*, 2022].

Spatial coverage and gridded datasets

Early studies, such as Serinaldi and Kilsby (2012) only used a handful of stations to train their stochastic models, raising concerns about spatial coherence and representativeness across catchments. The latest studies include a greater number of sites to not only improve spatial coverage in each water region within England and Wales, but also focus on locations which are of interest to regional water companies (e.g., high altitude sites). With extensive spatial coverage for all regions, the outputs from these studies can be brought together for national assessments (Atkins, 2020). However, it is important to consider and minimize discontinuities that may arise from merging different data sources or interpolation errors when extracting point data from gridded datasets (discussed more in detail later). In the Water Resources Management Plan 2019 (WRMP19) planning cycle, water companies used catchment average daily precipitation data from gridded products or rainfall station data over 45-60 sites per region, with each region having a bespoke model fitted over 1920-1997. Building on WRMP19, Atkins (2020) used high quality precipitation observations from 1950-1997 over 195 sites (with 50-80 sites per region) extracted from 1km HadUK rainfall data (Hollis *et al.*, 2019) based on stringent site identification criteria. Some studies cover specific regions, for example the Anglian region, which has “complicated water resource situations” (Dawkins *et al.*, 2022) i.e., they are prone to droughts due to low rainfall amounts and their dependency on aquifers and reservoirs, and thus have large implications on the water resources. Dawkins *et al.* (2022) used HadUK 1km gridded rainfall data over 39 sites in the Anglian region to run the

Advanced Meteorology Explorer (AME) model over 105 years from 1914-2018, after strict quality checks.

Although some weather generator studies use only rainfall gauge data as input (e.g., Leith, 2005) and some use the whole gridded dataset as input for their weather generators (e.g., Jones *et al.*, 2010); most stochastic studies reviewed here extract rainfall observations at different study sites from gridded observation datasets (e.g., Atkins 2020; Dawkins *et al.*, 2022). However, site observations derived from gridded rainfall datasets may have differences from actual station data due to the application of data interpolation (Dawkins *et al.*, 2022). Interpolated data is usually not able to represent the observed extreme rainfall, from sources like convective storms, and thus deviates from station data for certain events. This issue is more noticeable when the density and distribution of underlying stations for the gridded dataset is quite sparse (Legg, 2015). Keller *et al.* (2015) showed that the error in rainfall estimate in gridded CEH-GEAR data increases with the distance to the closest raingauge. The number of stations over the highland regions in the UK have always been quite low compared to rest of the country, but it was especially low before 1910 (Keller et al, 2015; Figure 2). The number of UK rain gauges increased steadily until 1960-1970 and then fell in the subsequent years for both CEH-GEAR (Keller *et al.*, 2015; Figure 4) and HadUK rainfall datasets (Hollis *et al.*, 2019). Although CEH-GEAR and HadUK rainfall datasets have inherent differences due to their respective generation methods (discussed in detail later), they also exhibit significant similarities (Figure 5) due to overlaps in the underlying station data.

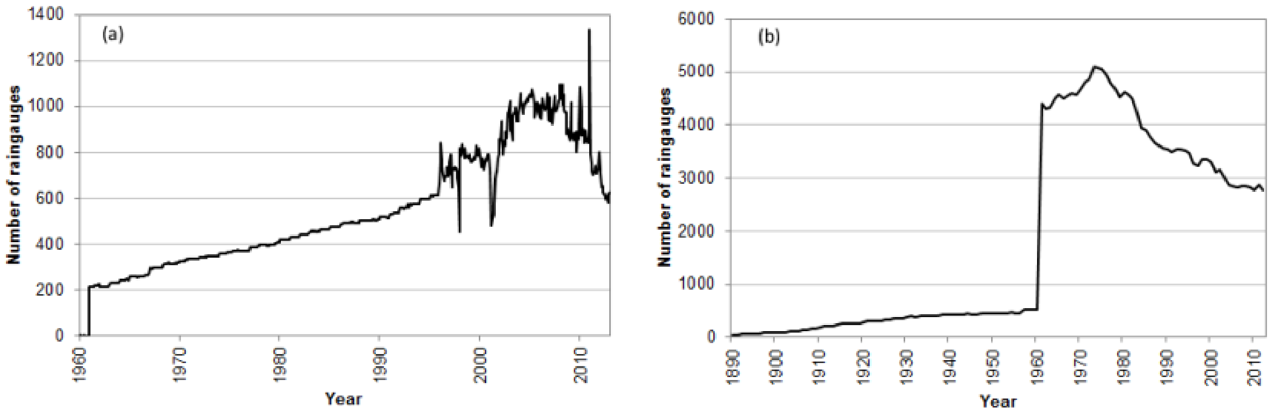

Figure 4: Number of UK rain gauges used to derive CEH-GEAR gridded rainfall for (a) monthly rainfall (prior to 1960 there is only one gauge used); and (b) daily rainfall [Source: Keller *et al.*, 2015].

Regions with sparse gauge distribution for rainfall gridded data relies more on data interpolation and thus, the choice of using gridded datasets rather than station data directly, especially for heterogeneous regions like the highlands, is a strong limitation of stochastic methods. To overcome this, some studies apply stochastic methods directly to the whole grid (i.e., all the grid points) of the dataset rather than interpolating the gridded data to point locations (Jones *et al.*, 2010). The other option is that studies perform stronger quality control check to identify unusual behaviour and large difference between station and gridded datasets over the study sites. For example, in Dawkins *et al.* (2022), quality checks consider the differences in the rainfall distributions related to proportion of zero rainfall days and variance of daily rainfall on an annual basis. Thus, data is only used for the period that passes the stringent checks, which was from 1914-2018, even though the data is available from 1891 onwards. The “limited” study period in this case coincides with periods of good spatial coverage of rain gauges. While prioritizing data quality and truncating series may enhance overall accuracy, it's crucial to acknowledge the risk of overlooking significant drought events and not capturing all the patterns of natural variability. Dawkins *et al.* (2022) missed the extended drought from the 1890s to 1910, highlighting the trade-offs inherent in decisions related to truncating the input data.

An important point to consider is that the choice of gridded datasets may also have an impact on the stochastic weather generator output. Atkins (2020) showed that rainfall datasets like CEH-GEAR and HadUK can have inherent differences arising from the selection of rain gauges, quality control approaches applied, and the interpolation methods used. However, the study noted that differences stemming from the model updates has a greater impact on the output time series compared to the choice of input data. These differences can be larger for high rainfall events (Figure 5a) or regions at higher elevations (Figure 5b), as shown by initial analysis conducted by UKCEH (unpublished). Simpson and McCarthy (2018) also showed that sampling strategies employed to derive different rainfall gridded products lead to significant differences in the resulting estimate of extreme rainfall, which can locally have important impacts.

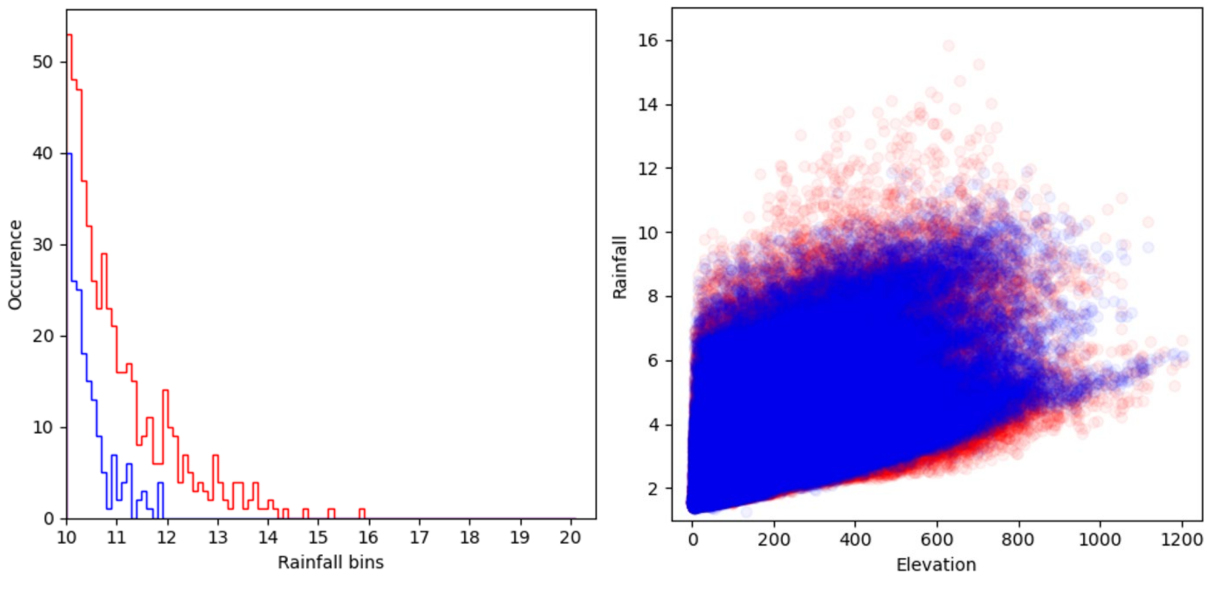

Figure 5: (a) Right-hand tail (with bins of rainfall greater than 10 mm/day) of the histogram for CEH-GEAR (red) and HadUK (blue) rainfall dataset. (b) Scatter plot between elevation and rainfall for each grid point of CEH-GEAR (red) and HadUK (blue) rainfall datasets [Source: Preliminary analysis].

Ingesting the historic rescued rainfall data mentioned in the previous section (Hawkins *et al.*, 2022) has significantly improved UK rainfall gridded data (HadUK rainfall data version 1.2), and global reanalysis datasets (Slivinski *et al.*, 2019; Hawkins *et al.*, 2023). Using gridded datasets that have more observations will be beneficial for future stochastic studies, especially at daily timescales.

In the case of model-based climate gridded datasets being used as input for stochastic weather generators, the considerations have been discussed in the Climate Change chapter (see Chapter 5). Using gridded datasets allows for models to incorporate future scenarios using climate model projections (Chun *et al.*, 2013). Further, alternative input datasets can also be considered to train the stochastic models for future projections like eFLaG and large ensemble runs which have transient climate change simulations (also discussed in Chapter 5 for more details). However, the inclusion of climate projections introduces further climate model-related uncertainties into the stochastic modelling process (for more details please see Chapter 5).

Some studies post-process point-data output from stochastic weather generators to create catchment-scale rainfall and river flow data. The catchment averages are calculated as weighted averages over Thiessen polygons (Atkins, 2020) or HadUK catchment averages (Dawkins *et al.*, 2022). Catchment averages provide a more reliable representation than single grid point data, as they prevent the interpolation errors arising due to the grid point distance from gauge observations. However, applying catchment averages may be partially leading to the biases in river flows of the weather generators (Dawkins *et al.*, 2022), and the choice of interpolation method can have a significant effect on the result (e.g., Yang and Xing, 2021; Antal *et al.*, 2021).

Consideration on PET method

All the studies in the UK using stochastic weather generators discussed so far generally focus mainly on rainfall. Indeed, multiple studies have demonstrated that hydrological models are much more sensitive to errors in rainfall than to errors in potential evapotranspiration (PET), especially in temperate climate such as the UK (Bastola *et al.*, 2011; Guo *et al.*, 2017; Paturel *et al.*, 1995). Although calibrating the rainfall-runoff models should be performed with PET derived from weather generators, it is important to validate the generated PET. Errors in PET have shown to have a potential effect on simulated river discharges especially for high and low flows (Samadi, 2016). For example, the choice of PET calculation method can have a significant impact to modelled runoff and river flows (Oudin *et al.* 2005). PET estimates may vary depending on the methods applied for their calculation (e.g., Kingston *et al.*, 2009), and have variable performance, even after calibration (Tanguy *et al.*, 2018). Moreover, many studies suggest that PET is going to become increasingly important as future warming is certain (Robinson *et al.*, 2022) and the significance of evaporative demand is anticipated to increase in future droughts (Reyniers *et al.*, 2023). Vicente-Serrano *et al.* (2022) shows that a sharp increase in the atmospheric evaporative demand is expected to drive a rise in the severity of future droughts. More generally, the two recent severe droughts to impact the UK, in 2018 (Turner *et al.* 2021) and 2022 (Barker *et al.* 2024) have been associated with very arid summer conditions, with very high evaporation losses during periods of exceptional temperatures. Thus, it is important for future drought studies over the UK to consider changes in both PET and rainfall (Glenis *et al.*, 2015) for better representation of future drought variability and characteristics and that the choice of PET method are tested and validated.

Summary of recommendations

1. Prioritise high-quality, homogeneous input datasets. When using gridded datasets instead of point observations, prioritize those with extensive underlying observations, incorporating historical rescued rainfall data to improve long-term records (Hawkins *et al.*, 2022). In general, ensure the use of consistent rainfall datasets to increase robustness of results.
   Benefit: Enhances the accuracy and reliability of stochastic hydrological simulations, leading to more informed decision-making and planning.
   Effort: *Low to Moderate.* Acquiring and curating extensive datasets demands time and resources, but the long-term benefits outweigh the initial effort. Some of these improvements have already been incorporated in open-source datasets (HadUK: Met Office, 2018; MIDAS: Met Office, 2012).
2. Review the fitness-for-purpose of the rainfall network with a view to its utility for supporting larger-scale regional drought assessments. This could potentially lead to increasing the number of monitoring sites to enhance spatial representation, or prioritising under-sampled areas and high-elevation sites, or exploring alternative methods for regional rainfall measurement. Alternatively, opt for high quality gridded datasets.
   Benefit: Ensures that the data collected aligns with the needs of regional drought assessments, with good spatial representation.
   Effort: *High.* Assessing and potentially expanding monitoring sites can require significant resources, especially if the need for a substantial increase in the number of monitoring sites is identified. However, optimizing existing networks or exploring alternative measurement methods might involve less effort.
3. Estimate the uncertainty arising from irreducible errors in underlying observation data (observational uncertainty), from the choice of input dataset or the choice of interpolation method, and acknowledge the effect of these in the interpretation of results.
   Benefit: Estimating and accounting for uncertainty due to errors in data and methodology improves the credibility of water resources assessments. Acknowledging these uncertainties enhances the transparency and trustworthiness of the results.
   Effort: *Moderate.* The effort for this recommendation primarily involves statistical analysis to quantify uncertainties and developing a clear framework for reporting them. While this analysis demands a level of expertise, it wouldn't necessarily require substantial additional data collection.

4. Due to the growing importance of potential evapotranspiration (PET) in future climate scenarios, there is a need to test and validate various PET calculation methods and its impact on calibration of models and subsequently the results.
   Benefit: Using the best PET method after testing and validating different calculation methods will offer a more accurate representation of hydrological simulations, especially with growing trends of evaporative demand driven by increasing temperature.
   Effort: *Low to Moderate.* Testing various calculation methods demands time for analysis and validation, but is generally quite simple.

Conclusion

In summary, the quality and selection of rainfall input data are pivotal for the reliability of stochastic weather generators. Ensuring data accuracy, consistency, and coverage, especially for historical records, is essential. Moreover, future studies should also consider the choice of method to estimate potential evapotranspiration for a more comprehensive understanding of future drought dynamics in a warming world.

References:

ANTAL, A., GUERREIRO, P.M.P. AND CHEVAL, S., 2021. Comparison of spatial interpolation methods for estimating the precipitation distribution in Portugal. Theor Appl Climatol 145, 1193-1206. https://doi.org/10.1007/s00704-021-03675-0

ATKINS, 2020. Regional Climate Data Tools: Sutton and East Surrey Water on behalf of WRSE. URL: https://www.wrse.org.uk/media/ok1mtsoq/wrse_file_1338_regional-climate-data-tools.pdf

BÁRDOSSY, A., KILSBY, C., BIRKINSHAW, S., WANG, N. AND ANWAR, F., 2022. Is Precipitation Responsible for the Most Hydrological Model Uncertainty? Front. Water 4:836554. https://doi.org/10.3389/frwa.2022.836554

BARKER, L.J., HANNAFORD, J., PARRY, S., SMITH, K.A., TANGUY, M., AND PRUDHOMME, C. 2019. Historic hydrological droughts 1891-2015: systematic characterisation for a diverse set of catchments across the UK. Hydrology and Earth System Sciences 23(11): 4583-4602.

BARKER, L.J., HANNAFORD, J., MAGEE, E., TURNER, S., SEFTON, C., PARRY, S., EVANS, J., SZCZYKULSKA, M. AND HAXTON, T. 2024. An appraisal of the severity of the 2022 drought and its impacts. Weather, 79(7): 208-219. https://doi.org/10.1002/wea.4531

BASTOLA, S., MURPHY, C. AND SWEENEY, J., 2011. The sensitivity of fluvial flood risk in Irish catchments to the range of IPCC AR4 climate change scenarios. Sci. Total Environ., 409, 5403–5415, https://doi.org/10.1016/j.scitotenv.2011.08.042

CHAN, S.C., KAHANA, R., KENDON, E.J. *ET AL.*, 2018. Projected changes in extreme precipitation over Scotland and Northern England using a high-resolution regional climate model. Clim Dyn 51, 3559–3577. https://doi.org/10.1007/s00382-018-4096-4

CHUN, K.P., WHEATER, H.S. AND ONOF, C., 2013. Comparison of drought projections using two UK weather generators. Hydrological sciences journal 58(2): 295-309.

CRAIG, P.M. AND HAWKINS, E., 2020. Digitizing observations from the Met Office Daily Weather Reports for 1900–1910 using citizen scientist volunteers. Geoscience Data Journal, 7(2), 116-134.

DAWKINS, L.C., OSBORNE, J.M., ECONOMOU, T., DARCH, G.J. AND STONER, O.R., 2022. The advanced meteorology explorer: a novel stochastic, gridded daily rainfall generator. Journal of Hydrology, 607, 127478.

GLENIS, V., PINAMONTI, V., HALL, J.W. AND KILSBY, C.G., 2015. A transient stochastic weather generator incorporating climate model uncertainty. Advances in Water Resources 85, 14-26.

GUO, D., WESTRA, S., AND MAIER, H.R., 2017. Use of a scenario-neutral approach to identify the key hydro-meteorological attributes that impact runoff from a natural catchment, J. Hydrol., 554, 317-330, https://doi.org/10.1016/j.jhydrol.2017.09.021

HAWKINS, E., BURT, S., BROHAN, P., LOCKWOOD, M., RICHARDSON, H., ROY, M. AND THOMAS, S., 2019. Hourly weather observations from the Scottish Highlands (1883-1904) rescued by volunteer citizen scientists, Geoscience Data Journal, 6, 160–173.

HAWKINS, E., BURT, S., MCCARTHY, M., MURPHY, C., ROSS, C., BALDOCK, M., BRAZIER, J., HERSEE, G., HUNTLEY, J., MEATS, R., O'GRADY, J., SCRIMGEOUR, I. AND SILK, T., 2022. Millions of historical monthly rainfall observations taken in the UK and Ireland rescued by citizen scientists. Geoscience Data Journal, 10(2), 246-261.

HAWKINS, E., BROHAN, P., BURGESS, S.N., BURT, S., COMPO, G.P., GRAY, S.L., HAIGH, I.D., HERSBACH, H., KUIJJER, K., MARTINEZ-ALVARADO, O., MCCOLL, C., SCHURER, A.P., SLIVINSKI, L. AND WILLIAMS, J., 2023. Rescuing historical weather observations improves quantification of severe windstorm risks. Natural Hazards and Earth System Sciences, 23(4), 1465-1482.

HOLLIS, D., MCCARTHY, M., KENDON, M., LEGG, T., AND SIMPSON, I., 2019. HadUK-Grid - A new UK dataset of gridded climate observations. Geoscience Data Journal, 6(2), 151-159.

JONES, P.D., LISTER, D.H., KOSTOPOULOU, E., 2004. Reconstructed River Flow Series from 1860s to Present. Tech. Rep. SC040052/SR, Environment Agency, Rio House, Waterside Drive, Aztec West, Almondsbury, Bristol, BS32 4UD. http://www.cru.uea.ac.uk/cru/data/riverflow/flowreconstruction/ReconstructedFlows-TR.pdf.

JONES, P.D., 1984. Riverflow reconstruction from precipitation data. Journal of Climatology 4: 171–185.

JONES, P.D., LISTER, D.H., WILBY, R.L., KOSTOPOULOU, E., 2006. Extended riverflow reconstructions for England and Wales, 1865–2002. International Journal of Climatology. 26(2), 219–231.

JONES, P.D., KILSBY, C.G., HARPHAM, C., GLENIS, V. AND BURTON, A., 2010. UK Climate Projections science report: Projections of future daily climate for the UK from the Weather Generator. University of Newcastle, UK. URL: https://eprints.ncl.ac.uk/file_store/production/217986/B1B47084-F4B6-446C-A750-8E47554D8B7F.pdf

KAUR, K. AND KAUR, N., 2023. Comparison of bias correction methods for climate change projections in the lower Shivaliks of Punjab. Journal of Water and Climate Change; 14(8): 2606–2625. doi: https://doi.org/10.2166/wcc.2023.503

KELLER, V.D.J., TANGUY, M., PROSDOCIMI, I., TERRY, J.A., HITT, O., COLE, S.J., FRY, M., MORRIS, D.G. AND DIXON, H., 2015. CEH-GEAR: 1 km resolution daily and monthly areal rainfall estimates for the UK for hydrological and other applications. Earth System Science Data, 7(1), 143-155.

KINGSTON, D.G., TODD, M.C., TAYLOR, R.G., THOMPSON, J.R. AND ARNELL, N.W., 2009. Uncertainty in the estimation of potential evapotranspiration under climate change. Geophysical Research Letters, 36(20).

LEITH, N., 2005. Using GLMs to simulate daily rainfall under scenarios of climate change. London: University College London. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=652b33601d2365b0258df7b3325e2ea00714ed93

LEGG, T., 2015. Uncertainties in gridded area-average monthly temperature, precipitation and sunshine for the United Kingdom. International Journal of Climatology, 35(7), 1367-1378.

LEWIS, E., QUINN, N., BLENKINSOP, S., FOWLER, H.J., FREER, J., TANGUY, M., HITT, O., COXON, G., BATES, P. AND WOODS, R., 2018. A rule based quality control method for hourly rainfall data and a 1km resolution gridded hourly rainfall dataset for Great Britain: CEH-GEAR1hr. Journal of Hydrology, 564, 930-943, https://doi.org/10.1016/j.jhydrol.2018.07.034

LUO, M., LIU, T., MENG, F., DUAN, Y., FRANKL, A., BAO, A. AND DE MAEYER, P., 2018. Comparing Bias Correction Methods Used in Downscaling Precipitation and Temperature from Regional Climate Models: A Case Study from the Kaidu River Basin in Western China. Water, 10, 1046. https://doi.org/10.3390/w10081046

MET OFFICE, 2012. Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853-current). NCAS British Atmospheric Data Centre, date of citation. http://catalogue.ceda.ac.uk/uuid/220a65615218d5c9cc9e4785a3234bd0

MET OFFICE, HOLLIS, D., MCCARTHY, M., KENDON, M., LEGG, T. AND SIMPSON, I., 2018: HadUK-Grid gridded and regional average climate observations for the UK. Centre for Environmental Data Analysis, [last accessed: 11/11/2023]. http://catalogue.ceda.ac.uk/uuid/4dc8450d889a491ebb20e724debe2dfb

MUCHAN, K. AND DIXON, H., 2019. Insights into rainfall undercatch for differing raingauge rim heights. Hydrology Research, 50 (6): 1564–1576. doi: https://doi.org/10.2166/nh.2019.024

MURPHY, C., WILBY, R.L., MATTHEWS, T.K.R, *ET AL.*, 2020. Multi-century trends to wetter winters and drier summers in the England and Wales precipitation series explained by observational and sampling bias in early records. Int J Climatol. 40: 610–619. https://doi.org/10.1002/joc.6208

PATUREL, J.E., SERVAT, E. AND VASSILIADIS, A., 1995. Sensitivity of conceptual rainfall-runoff algorithms to errors in input data – case of the GR2M model, J. Hydrol., 168, 111–125, https://doi.org/10.1016/0022-1694(94)02654-T

PERRY, M. AND HOLLIS, D., 2005. The generation of monthly gridded datasets for a range of climatic variables over the UK. International Journal of Climatology 25: 1041–1054.

REYNIERS, N., OSBORN, T.J., ADDOR, N. AND DARCH, G., 2023. Projected changes in droughts and extreme droughts in Great Britain strongly influenced by the choice of drought index. Hydrology and Earth System Sciences, 27(5), 1151-1171.

ROBINSON, E.L., BROWN, M.J., KAY, A.L., LANE, R.A., CHAPMAN, R., BELL, V., AND BLYTH,E.M., 2022. Hydro-PE gridded datasets of historical and future Penman-Monteith potential evaporation for the United Kingdom. Earth Syst. Sci. Data Discuss. [preprint], https://doi.org/10.5194/essd-2022-288, in review

SAMADI, S.Z., 2016. Assessing the sensitivity of SWAT physical parameters to potential evapotranspiration estimation methods over a coastal plain wate4rshed in the southeastern United States, Hydrol. Res., 48, 395–415.

SERINALDI, F. AND KILSBY, C.G., 2012. A modular class of multisite monthly rainfall generators for water resource management and impact studies. Journal of Hydrology 464: 528-540.

SIMPSON, I.R. AND MCCARTHY, M.P., 2019. Structural and sampling uncertainty in observed UK daily precipitation extremes derived from an intercomparison of gridded data sets. Int J Climatol., 39, 128–142. https://doi.org/10.1002/joc.5789

SLIVINSKI, L.C., COMPO, G.P., WHITAKER, J.S., SARDESHMUKH, P.D., GIESE, B.S., MCCOLL, C., … AND WYSZYŃSKI, P., 2019. Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. Quarterly Journal of the Royal Meteorological Society, 145(724), 2876-2908.

TANGUY, M., PRUDHOMME, C., SMITH, K., HANNAFORD, J., 2018. Historical gridded reconstruction of potential evapotranspiration for the UK, Earth System Science Data, 10, 951–968. https://doi.org/10.5194/essd-10-951-2018

TANGUY, M., DIXON, H., PROSDOCIMI, I., MORRIS, D.G., KELLER, V.D.J., 2021. Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890-2019) [CEH-GEAR]. NERC EDS Environmental Information Data Centre. https://doi.org/10.5285/dbf13dd5-90cd-457a-a986-f2f9dd97e93c

TURNER, S., BARKER, L.J., HANNAFORD, J., MUCHAN, K., PARRY, S. AND SEFTON, C., 2021. The 2018/2019 drought in the UK: a hydrological appraisal. Weather, 76(8), 248-253.

VICENTE-SERRANO, S.M., PEÑA-ANGULO, D., BEGUERÍA, S., DOMÍNGUEZ-CASTRO, F., TOMÁS-BURGUERA, M., NOGUERA, I., GIMENO-SOTELO, L. AND EL KENAWY, A.., 2022. Global drought trends and future projections. Philos Trans A Math Phys Eng Sci. Dec 12;380(2238):20210285. doi: 10.1098/rsta.2021.0285. Epub 2022 Oct 24. PMID: 36300353; PMCID: PMC9588423.

YANG, R. AND XING, B., 2021. A Comparison of the Performance of Different Interpolation Methods in Replicating Rainfall Magnitudes under Different Climatic Conditions in Chongqing Province (China). Atmosphere 2021, 12, 1318. https://doi.org/10.3390/atmos12101318

Climate Predictors

Amulya Chevuturi and Maliko Tanguy (with review and input from the author team)

Introduction

Stochastic weather generators, for water resources applications in England, typically use the relationship between climate predictors and local rainfall patterns as a basis for stochastic simulation. Originally, these generators utilized observed indices such as the North Atlantic Oscillation (NAO; Jones *et al.*, 1997) and sea surface temperatures (SSTs) of specific areas of the North Atlantic Ocean (Serinaldi and Kilsby, 2012) to generate synthetic rainfall datasets, which were employed for water resources planning in the UK (Water UK, 2016; Atkins, 2020). These climate predictors are selected based on their strong relationships with rainfall in the specific region of interest. For instance, Sea Surface Temperature (SST) domains within the North Atlantic adjacent to the UK (e.g., Dawkins *et al.*, 2022), and associated indices such as NAO, Atlantic Multi-decadal Oscillation (AMO) (Brown, 2008) have been found to exhibit significant correlations with UK rainfall. Beyond the Atlantic, there are weak influences from the Pacific Ocean indices such as El Niño-Southern Oscillation (ENSO), the Pacific Decadal Oscillation (PDO) (Brown, 2008). In addition to oceanic predictors (e.g., NAO index, SSTs, sea level pressure), some studies also use local atmospheric variables from the North Atlantic region (e.g., relative humidity and temperature) and related indices of East Atlantic (EA), East Atlantic West Russian (EAWR) and Scandinavia (SCA) patterns (West *et al.*, 2022) as predictors in their stochastic weather generators (Chun *et al.*, 2013).

Atlantic Predictors of UK Rainfall

The climate indicators over the Atlantic have varying influence on the regional rainfall variability over UK; for example, NAO has stronger influence on western UK winter rainfall (West *et al.*, 2019). Please refer to Figure 1, sourced from West *et al.* (2019), illustrating the diverse regional effects of the NAO on winter and summer rainfall in the UK, summarising the findings from historical literature. Therefore, it is crucial for stochastic weather generators to incorporate the strongest predictor-precipitation relationships available (Lister, *et al.* 2018). In line with this principle, Atkins and the UK Met Office enhanced the original Serinaldi and Kilsby (2012) model by including EA index for generating rainfall for Water Resources East (previously East Anglian) region (Chapter 1) and this predictor has been adopted in other studies focused on the same region, such as Lister *et al.* (2018) and Dawkins *et al.* (2022The same model was further improved by including several additional climate predictors (EAWR, SCA, AMO), which were selected after rigorous testing of correlation strengths between the predictor and precipitation over different stations using different sources of predictor datasets, for different time periods and seasons (Atkins, 2020). It should be noted that the precise implementation of these models and the predictor combinations is not made clear in the reports, and has not been published in the peer-reviewed literature. This refinement resulted in a more accurate model fit, i.e., reduced mean absolute error in the stochastic rainfall output, particularly for regions where the NAO exerts a weaker influence, ultimately enabling a single model to represent the temporal and spatial variations of rainfall across the entire country. Recent stochastic methods have combined the model selection step, based on different climate predictors, along with model fitting to gain additional flexibility in identifying influential climate drivers for specific regional rainfall within UK (Dawkins *et al.*, 2022). However, excessive number of predictors result in complex, computationally intensive models, which due to overfitting lead to inaccurate representation of rainfall variability (e.g., Semenov, 2008). In stochastic models, the definite causal relationship between predictor and rainfall is important, rather than just the strong statistical connection between the two, as that lends confidence to the multiple rainfall realisations that are produced using this relationship.

The impacts of seasonally varying climate predictors, such as North Atlantic SSTs, on rainfall variability, particularly in the European context, have been demonstrated (e.g., Sutton and Hodson, 2005; Dunstone, 2018; Smith *et al.*, 2018). Therefore, incorporating seasonally varying indices into stochastic weather generators can enhance their performance. In Dawkins *et al.* (2022) models were fitted with different predictors for different seasons. The EA index and regional SST anomalies were used for all seasons, while the NAO indices derived from mean sea level pressure varied depending on whether it was summer or winter.

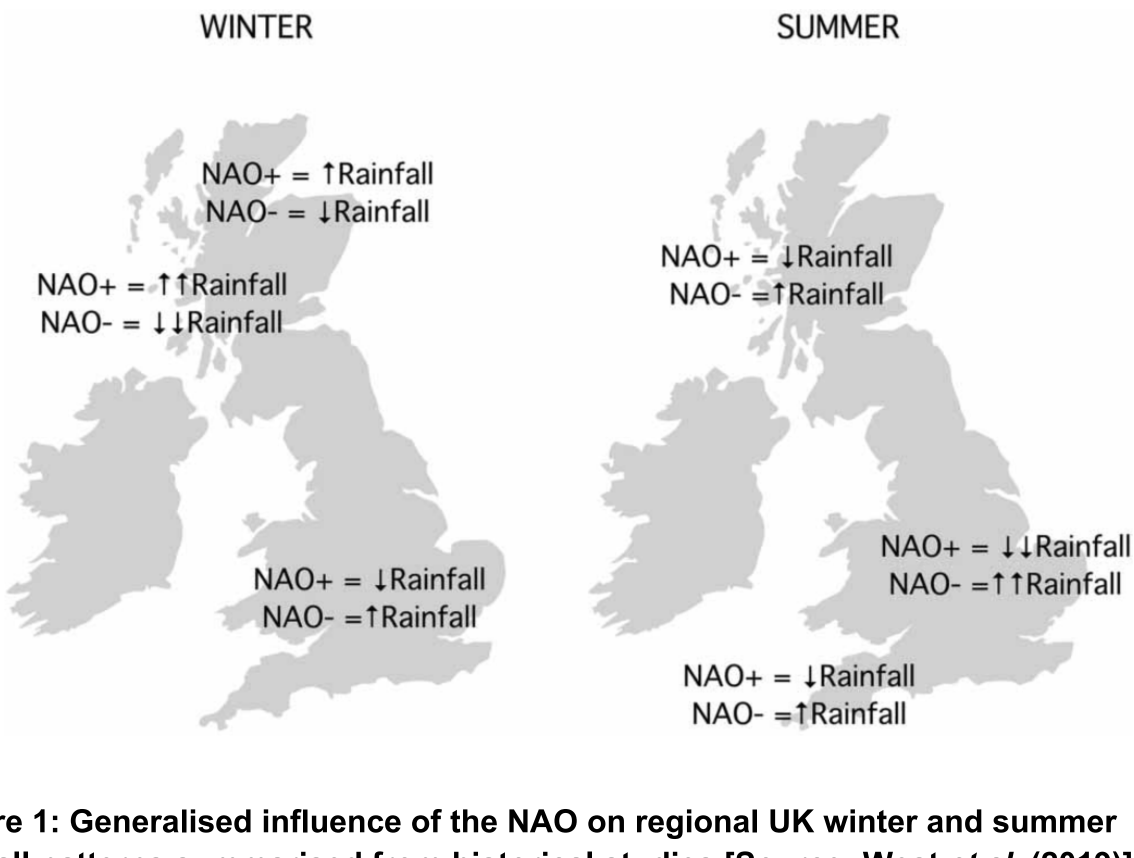

Figure 1: Generalised influence of the NAO on regional UK winter and summer rainfall patterns summarised from historical studies [Source: West *et al.* (2019)].

To ensure skilful generation of rainfall, stochastic generators must incorporate the most influential climate drivers available as predictors, as demonstrated in Atkins (2020). North

Atlantic climate variability has been widely recognized as a key contributor to European precipitation, including that of the UK (e.g., Marshall *et al.*, 2001). Particularly during the winter season, the NAO plays a significant role (e.g., López-Moreno and Vicente-Serrano, 2008), but to a much lesser extent during the summer (Folland *et al.*, 2009). Beyond the NAO, other atmospheric patterns in the North Atlantic, such as the EA pattern (Hall and Hanna, 2018), EAWR pattern (e.g., Parry *et al.*, 2012), and SCA pattern (e.g., Hannaford *et al.*, 2011), significantly impact regional precipitation in the UK. Additionally, rainfall responses may be represented more accurately by considering the interactions between different atmospheric circulation patterns. For example, studies have shown that the combined influence of the NAO and EA patterns modulates the position of pressure systems across the British Isles. Studies have shown that for some parts of southern UK, the phases of the EA pattern could dampen or even reverse the rainfall signal expected from a given NAO phase (Mellado-Moore and Renfrew 2009; Cano *et al.* 2019; West *et al.* 2022). Some low frequency modes of variability such as the decadal variability in the North Atlantic SSTs, known as the AMO also influence European rainfall during the summer season (e.g., Dong *et al.*, 2013). These lower frequency modes of variability at multi-annual and decadal timescales can improve the representation of rainfall variability for models operating on longer time steps, due to the slowly varying nature of such indictors.

The climate forcings originating from the North Atlantic exert their influence on European climate by modulating the Atlantic storm track (e.g., Seierstad *et al.*, 2007; Vallis and Gerber, 2008). As a result, these drivers have been extensively employed in stochastic weather generators, as mentioned earlier in this work. However, the excessive utilization of predictors can lead to overfitting the model, as evidenced by studies conducted for other regions (e.g., Qian *et al.*, 2008; Semenov, 2008) as mentioned before. Therefore, it is advisable for researchers and practitioners to exercise caution in employing an excessive number of predictors to fit their models, particularly considering the limited investigation of this issue in the context of stochastic generators applied to UK rainfall.

Teleconnections from beyond the Atlantic

While the North Atlantic region has been a primary focus of research, recent studies have highlighted the importance of considering teleconnections from beyond the region to gain a comprehensive understanding of UK drought drivers. Exploring such teleconnections with remote regions will allow researchers to unravel the intricate interactions and climate dynamics that influence drought occurrences in the UK. Many studies reviewed in Brönnimann *et al.* (2007) show impacts of ENSO on the European climate, with some focusing specifically on UK (Wilby, 1993; Svensson and Prudhomme, 2005). The correlation between polar Eurasian patterns and the North Atlantic Oscillation (NAO) has shown significant influence on UK droughts (Wedgbrow *et al.*, 2002). Folland *et al.* (2015) demonstrated the impact of teleconnections from the Pacific Ocean, specifically La Niña events, and the quasi-biennial oscillation of stratospheric winds on droughts in the English Lowlands. Conversely, El Niño patterns have been strongly correlated with high spring precipitation in Southern England (Van Oldenborgh *et al.*, 2000). The Pacific Decadal Oscillation (PDO) causes a three-month delayed impact on north-west and south-east rainfall regions within the UK (Figure 2a), modulated by the different phases of AMO

(Figure 2b). This lagged influence is more pronounced for streamflow compared to rainfall (Svensson and Hannaford, 2019).

Non-stationarity in teleconnection between rainfall and climate predictors

It is important to note that the influence of teleconnections under investigation can vary depending on the input data period due to inherent long-term variabilities in the relationships between the large-scale climate predictors and UK hydrology (Rust *et al.*, 2021). For instance, studies have shown non-stationarity in the relationship between NAO and European precipitation due to decadal shifts in the NAO pressure centre (e.g., Vicente-Serrano and López-Moreno, 2008). Links between North Atlantic circulation patterns and European streamflows show a 7-years cycle post mid 1980’s not seen over the earlier periods (Lorenzo-Lacruz *et al.*, 2022). This non-stationarity in the context of a changing climate further introduces significant uncertainty (Woolings and Blackburn, 2012), which will have implications for using weather generators to make decisions regarding future water resource management (e.g., Kilsby *et al.*, 2007). While correlations between UK rainfall and climate predictors tend to be stronger over shorter time periods compared to longer ones (Lister *et al.*, 2018), utilizing a shorter observational period to train weather generators may result in the omission of significant historical drought events, despite achieving a better model fit (Atkins, 2020). However, the utilization of longer periods of teleconnections in stochastic models faces challenges due to poor quality and lack of availability of historical climate predictor and rainfall datasets prior to 1950s (Atkins, 2020). To address this limitation, researchers often merge datasets from various sources to create a more comprehensive record (e.g., Dawkins *et al.*, 2020). However, it is important to note that this merging process can introduce additional uncertainties to the model, stemming from inherent differences in the datasets and potential data quality issues during the historical period (please see Chapter 3 for detailed discussion). Another approach that stochastic modelling studies use is to draw alternative samples of the climate predictors from the observed values by resampling (Serinaldi and Kilsby, 2012) or parametric modelling (Chun *et al.*, 2013) (please see Chapter 2 for detailed discussion).

Recommendations

Our findings underscore the evolving nature of the study of UK rainfall drivers and subsequent droughts. Recent research by our group has unveiled a substantial impact of freshwater incursion and sea salinity changes in the North Atlantic Ocean on both UK rainfall and streamflow with long lead times (Chevuturi *et al.*, under review.). These factors play a crucial role in modulating North Atlantic Ocean temperatures, thereby influencing the hydrological patterns in the region. In light of these developments, future studies focusing on stochastic models for UK rainfall should consider incorporating additional predictors from the latest scientific research, to enhance the accuracy of rainfall projections. By exploring a broader range of drivers, researchers and practitioners alike can improve the quality and reliability of generated rainfall data. Nevertheless, when increasing the number of climate predictors in the stochastic model it is crucial to strike the right balance between improving the predictive power of the stochastic model and the inherent risk of model overfitting. It's worth noting that these drivers may exhibit regional varying influences on UK rainfall, thereby introducing the possibility of region dependant overfitting concerns. This suggests that different predictors or combinations of predictors might be needed for different regions, which would lead to regionally inconsistent simulations. The best approach might be to have a master set of predictors, for the stochastic modelling of a region, in which the different predictors get switched on and off based on the correlations between the predictor subset and rainfall for different sub-regions.

Furthermore, even if recent developments and emerging predictors are incorporated into newer generations of stochastic simulators, there will always be a proliferation of new predictors (or improvements in existing predictors) as our process understanding of the drivers of UK hydroclimatic variability continues to grow. There is an ongoing need for communication and collaboration between industry practitioners and hydroclimate researchers to ensure regular updates to operational practices.

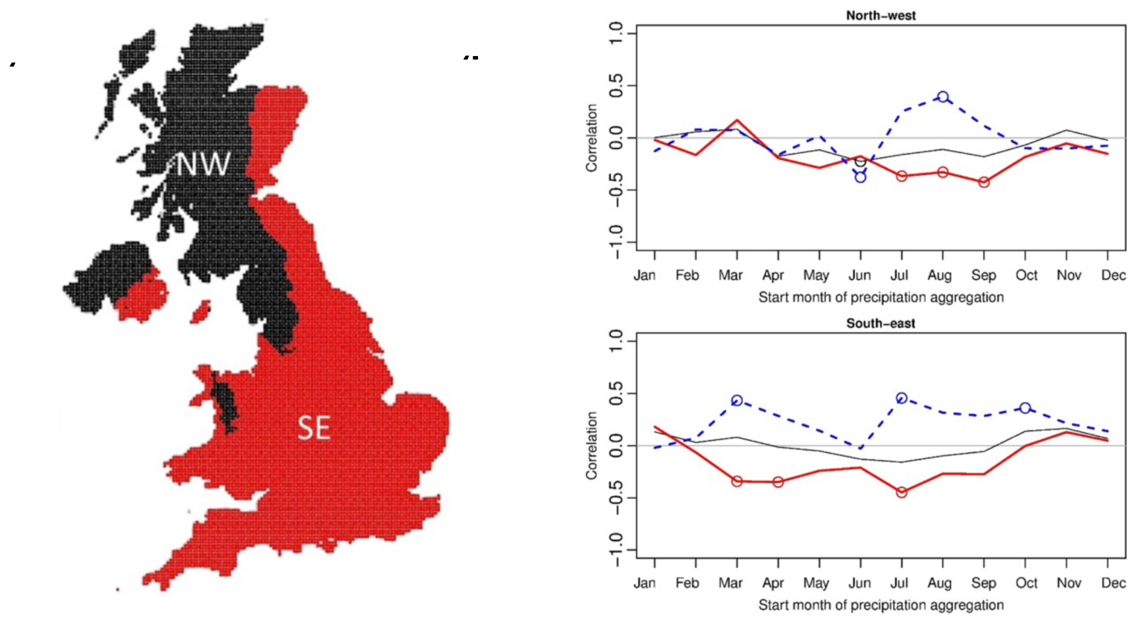

Figure 2: The UK divided into (a) two precipitation regions. (b) Correlations between three-month aggregations of PDO and two rainfall regions from 1961–2016, with rainfall lagged three months after the PDO. The correlations are stratified on AMO phase: AMO < −0.05 (dashed blue line); AMO ≥ −0.05 (thick solid red line); all AMO regardless of phase (thin solid black line). Significant correlations are circled [Source: Svensson and Hannaford, 2019].

Summary of limitations

- The most influential predictor varies depending on the studied region and missing influential predictors could potentially result in a loss of accuracy in the stochastic model output.

- Non-stationarity of teleconnections in the face of current and future climate conditions, posing challenges for accurately modelling their impacts.
- Restricted availability of predictors in historical datasets, hampering the ability to encompass key extreme events.
- Prevalence of regional-scale studies and associated predictors, limiting the generalizability of findings to broader contexts or larger spatial scales (see Spatial Coherence Section in Chapter 2).
- Require balancing the risk of overfitting due to inclusion of an excessive number of predictors with the need to include the most influential predictors for a given region to prevent compromising the reliability and robustness of results

Summary of recommendations

- Broadening the range of predictors by incorporating insights from emerging research on climate drivers (e.g., Lorenzo-Lacruz *et al.*, 2022; Chevuturi *et al.*, under review)
  Benefit: By keeping up with the latest research in the field, the operational stochastic models can be improved to have the best predictors representing the rainfall variability, which would ultimately improve stochastic rainfall output.
  Effort: *Moderate to High.* Thorough evaluation and validation through extensive peer-reviewed research establishes widely accepted climate predictors to be used. For example, the slowly varying lower frequency modes of variability in the SSTs can be used to improve the representation of rainfall variability at multi-annual and decadal timescales. Operational models would need to keep up with the ever-evolving landscape of hydroclimatic research and would need to constantly test and then upgrade their models with any new research findings. However, additional predictors and other complexities, like non-linear relationships, should only be introduced to stochastic models if there is clear and well demonstrated benefits to the applications of interest.
- Balancing improving the predictive power of the stochastic model vs. the inherent risk of model overfitting when increasing the number of climate predictors being used
  Benefit: This would prevent model overfitting while capturing the maximum range of rainfall variability in the stochastic simulations.
  Effort: *Low to Moderate.* To weigh the benefits of using large number of climate predictors against the overfitting of the stochastic model requires stringent testing via multiple stochastic runs as done in Atkins (2020). This can be computationally intensive but should be done at the testing phase of the study. Based on the response of sub-region rainfall variability to different predictors, it might be best to use only the relevant sub-set of predictors for each region. Further, it should be considered that depending on the water resources planning use cases, perhaps alternative simpler models are more than up to the task rather than increasingly complex stochastic models.
- Enhancing the representation of seasonal variation in rainfall through the inclusion of seasonally varying indices as in Dawkins *et al.* (2022)
  Benefit: Using seasonally varying climate predictors allows for better representation of rainfall variability in the stochastic output, as the climate teleconnections vary with seasons (e.g., West *et al.*, 2019).

Effort: *Low to Moderate.* There is existing research on seasonally varying teleconnections of UK rainfall (e.g., Sutton and Hodson, 2005; Dunstone, 2018; Smith *et al.*, 2018) which can be integrated into the stochastic weather generators. However, the challenge lies in testing and implementing this approach for every climate predictor, given the complexities involved in running seasonally varying stochastic simulations. Alternatively, as in Atkins (2020), the month of the year can be used as an input predictor to represent the inter-seasonal variability.

- Considering the effects of non-stationarity, including decadal shifts and long-term variability, particularly in the context of a changing climate (see Chapter 5)
  Benefit: Understanding changes in teleconnections over time is crucial for representing the uncertainty in stochastic rainfall simulations and ensuring the most accurate representation of rainfall variability.
  Effort: *High.* Incorporating all aspects of teleconnection variability into a stochastic weather generator may require a time-varying stochastic model for each region, which can be resource intensive.
- Improving the availability and quality of predictor data by selecting the best time period, best sources of datasets (see Chapter 3) or using approaches to derive alternative samples of predictors from observations (see Chapter 2)
  Benefit: Improving quality of climate predictor data will provide accurate representation of rainfall teleconnections and ultimately the simulated rainfall output.
  Effort: *Low to Moderate.* Atkins (2020) suggested that climate predictor data for UK rainfall has higher quality mid 1950s onwards and specifically tested multiple sources of climate predictor datasets before using them for stochastic models. Further, rather than just using fixed predictor dataset, alternative samples of the climate predictors can be derived from the observed values using resampling (Serinaldi and Kilsby, 2012) or modelling (Chun *et al.*, 2013). However, operational stochastic modelling for water resource planning should always research/test the climate predictors datasets for accuracy before implementing them operationally.

References

ATKINS, 2020. Regional Climate Data Tools: Sutton and East Surrey Water on behalf of WRSE. URL: https://www.wrse.org.uk/media/ok1mtsoq/wrse_file_1338_regional-climate-data-tools.pdf

BRÖNNIMANN, S., 2007. Impact of El Niño–southern oscillation on European climate. Reviews of Geophysics 45(3).

CHEVUTURI, A., OLTMANNS, M., TANGUY, M., HARVEY, B., SVENSSON, C. AND HANNAFORD, J., under review. Oceanic drivers of UK summer droughts. Communications Earth & Environment.

CHUN, K.P., WHEATER, H.S. AND ONOF, C., 2013. Comparison of drought projections using two UK weather generators. Hydrological sciences journal 58(2), 295-309.

DAWKINS, L.C., OSBORNE, J.M., ECONOMOU, T., DARCH, G.J. AND STONER, O.R., 2022. The advanced meteorology explorer: a novel stochastic, gridded daily rainfall generator. Journal of Hydrology, 607, 127478.

DONG, B., SUTTON, R.T., WOOLLINGS, T. AND HODGES, K., 2013. Variability of the North Atlantic summer storm track: mechanisms and impacts on European climate. Environmental Research Letters 8(3), 034037.

DUNSTONE, N., SMITH, D., SCAIFE, A., HERMANSON, L., FEREDAY, D., O'REILLY, C., STIRLING, A., EADE, R., GORDON, M., MACLACHLAN, C., WOOLLINGS, T., SHEEN, K. AND BELCHER, S., 2018. Skilful seasonal predictions of summer European rainfall. Geophysical Research Letters 45(7), 3246-3254.

FOLLAND, C.K., KNIGHT, J., LINDERHOLM, H.W., FEREDAY, D., INESON, S. AND HURRELL, J.W., 2009. The summer North Atlantic Oscillation: past, present, and future. Journal of Climate, 22(5), 1082-1103.

FOLLAND, C.K., HANNAFORD, J., BLOOMFIELD, J.P., KENDON, M., SVENSSON, C., MARCHANT, B.P., PRIOR, J. AND WALLACE, E., 2015. Multi-annual droughts in the English Lowlands: a review of their characteristics and climate drivers in the winter half-year. Hydrology and Earth System Sciences 19(5), 2353-2375.

HALL, R.J. AND HANNA, E., 2018. North Atlantic circulation indices: links with summer and winter UK temperature and precipitation and implications for seasonal forecasting. International Journal of Climatology 38, e660-e677.

HANNAFORD,J., LLOYD-HUGHES, B., KEEF, C., PARRY, S. AND PRUDHOMME, C., 2011. Examining the large-scale spatial coherence of European drought using regional indicators of precipitation and streamflow deficit. Hydrological Processes 25(7), 1146-1162.

JONES, P.D., JONSSON, T. AND WHEELER, D., 1997. Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. International Journal of Climatology 17(13), 1433-1450.

KILSBY, C.G., JONES, P.D., BURTON, A., FORD, A.C., FOWLER, H.J., HARPHAM, C., JAMES, P., SMITH, A. AND WILBY, R.L., 2007. A daily weather generator for use in climate change studies. Environmental Modelling & Software 22(12), 1705-1719.

LISTER, D., OSBORN, T., JONES, P. AND DARCH, G. 2018. An assessment of extreme drought rainfall sequences in stochastic weather generator output for the greater Anglian region. URL: https://www.uea.ac.uk/documents/96135/5613032/ENV-CRU-RP23-Reduced-2019.pdf/b4d70894-df9a-3c79-d90d-4c7da201676c?t=1630500281521

LÓPEZ-MORENO,J.I. AND VICENTE-SERRANO, S.M., 2008. Positive and negative phases of the wintertime North Atlantic Oscillation and drought occurrence over Europe: a multitemporal-scale approach. Journal of Climate 21(6), 1220-1243.

LORENZO-LACRUZ, J., MORÁN-TEJEDA, E., VICENTE-SERRANO, S.M., HANNAFORD, J., GARCÍA, C., PEÑA-ANGULO, D. AND MURPHY, C., 2022. Streamflow frequency changes across western Europe and interactions with North Atlantic atmospheric circulation patterns. Global and Planetary Change, 212, 103797.

MARSHALL, J., KUSHNIR, Y., BATTISTI, D., CHANG, P., CZAJA, A., DICKSON, R., HURRELL, J., MCCARTNEY, M., SARAVANAN, R. AND VISBECK, M., 2001. North Atlantic climate variability: phenomena, impacts and mechanisms. International Journal of Climatology 21(15), 1863-1898.

MELLADO-CANO, J., BARRIOPEDRO, D., GARCÍA-HERRERA, R., TRIGO, R.M. AND HERNÁNDEZ, A., 2019. Examining the North Atlantic Oscillation, East Atlantic Pattern, and Jet Variability since 1685. Journal of Climate 32, 6285–6298.

MOORE, G.W.K. AND RENFREW, I.A., 2012. Cold European winters: interplay between the NAO and the East Atlantic mode. Atmospheric Science Letters 13, 1–8.

QIAN, B., GAMEDA, S. AND HAYHOE, H., 2008. Performance of stochastic weather generators LARS-WG and AAFC-WG for reproducing daily extremes of diverse Canadian climates. Climate Research 37(1), 17-33.

VAN OLDENBORGH, G.J., BURGERS, G. AND TANK, A.K., 2000. On the El Niño teleconnection to spring precipitation in Europe. International Journal of Climatology 20(5), 565-574.

PARRY, S., HANNAFORD. J., LLOYD-HUGHES, B. AND PRUDHOMME, C., 2012. Multi-year droughts in Europe: analysis of development and causes. Hydrology Research 43(5), 689-706.

SEMENOV, M.A., 2008. Simulation of extreme weather events by a stochastic weather generator. Climate Research 35: 203–212

SERINALDI, F. AND KILSBY, C.G., 2012. A modular class of multisite monthly rainfall generators for water resource management and impact studies. Journal of Hydrology 464, 528-540.

SUTTON, R.T. AND HODSON, D.L., 2005. Atlantic Ocean forcing of North American and European summer climate. Science 309(5731), 115-118.

SVENSSON, C. AND HANNAFORD, J., 2019 Oceanic conditions associated with Euro-Atlantic high pressure and UK drought. Environmental Research Communications, 1(10), 101001.

RUST, W., CUTHBERT, M., BLOOMFIELD, J., CORSTANJE, R., HOWDEN, N. AND HOLMAN, I., 2021. Exploring the role of hydrological pathways in modulating multi-annual climate teleconnection periodicities from UK rainfall to streamflow. Hydrology and Earth System Sciences 25(4), 2223-2237.

SEIERSTAD, I.A., STEPHENSON, D.B. AND KVAMSTØ, N.G., 2007. How useful are teleconnection patterns for explaining variability in extratropical storminess? Tellus A: Dynamic Meteorology and Oceanography 59(2), 170-181.

SVENSSON, C. AND PRUDHOMME, C. 2005. Prediction of British summer river flows using winter predictors. Theoretical and Applied Climatology 82, 1-15.

VALLIS, G.K. AND GERBER, E.P., 2008. Local and hemispheric dynamics of the North Atlantic Oscillation, annular patterns and the zonal index. Dynamics of atmospheres and oceans 44(3-4), 184-212.

VICENTE-SERRANO, S.M. AND LÓPEZ-MORENO, J.I., 2008. Nonstationary influence of the North Atlantic Oscillation on European precipitation. Journal of Geophysical Research: Atmospheres 113(D20).

WATER UK, 2016. Water resources long term planning framework (2015-2065). Water UK. URL: https://www.water.org.uk/wp-content/uploads/2018/11/WaterUK-WRLTPF_Final-Report_FINAL-PUBLISHED-min.pdf

WEDGBROW, C.S., WILBY, R.L., FOX, H.R. AND O'HARE, G., 2002. Prospects for seasonal forecasting of summer drought and low river flow anomalies in England and Wales. International Journal of Climatology 22(2), 219-236.

WEST, H., QUINN, N. AND HORSWELL, M., 2019. Regional rainfall response to the North Atlantic oscillation (NAO) across Great Britain. Hydrology Research 50(6): 1549-1563.

WEST, H., QUINN, N. AND HORSWELL, M., 2022. The Influence of the North Atlantic Oscillation and East Atlantic Pattern on Drought in British Catchments. Frontiers in Environmental Science 10.

WILBY, R., 1993. Evidence of ENSO in the synoptic climate of the British Isles since 1880. Weather 48(8), 234-239.

WOOLLINGS, T. AND BLACKBURN, M., 2012. The North Atlantic jet stream under climate change and its relation to the NAO and EA patterns. Journal of Climate 25(3): 886-902.

Climate Change

Amulya Chevuturi and Maliko Tanguy (with review and input from the author team)

Introduction

Climate is undergoing significant changes due to anthropogenic warming, leading to an increase in hydrological extreme event frequency, and these events are expected to be more severe than rare “worst historic” events (Water UK, 2016). To effectively address these challenges, it is imperative to make long-term investments in infrastructure development for climate change adaptation. There is a high degree of certainty over future temperature increase from warming. However, the response of atmospheric circulation to the future warming is uncertain, leading to significant uncertainties in the potential rainfall changes. There is, therefore, a need to incorporate climate projections and the associated uncertainty in water resources planning, as more extreme events are expected due to anthropogenic warming in the future climate. At the same time, even in the absence of anthropogenic warming, there could be a risk of being exposed to more extreme events than witnessed in the historic record (e.g., Thompson *et al.* 2017) - any record-breaking events expected by chance may not be represented in the limited sample of historic observations in the current climate (Water UK, 2016).

To include the extreme events when planning for future water resources over UK, stochastic approaches are widely implemented, such as to model droughts with intensities surpassing the historical record (Anderton *et al.*, 2015) or to create ensemble of rainfall simulations for different subregions of UK (Atkins, 2020). However, with the changing climate it has become imperative that we include future climate variability into these stochastic simulations.

Stochastic approaches implementing the impact of climate change

Climate change scenarios can be incorporated into stochastic weather generators through three approaches: post-processing of output, stochastic model fitting, and input of climate drivers. In the <u>post-processing method,</u> the approach most widely used in practice in the water industry, uses change factor methodology. Change factors for different climate scenarios, are derived from future flow projections (Water UK, 2016) or rainfall projections (Atkins, 2020),and applied to stochastically generated output, i.e., the sequences derived from stochastic generation applied to observational records. Alternatively, in the <u>stochastic model fitting</u> approach, future projections of the stochastic output are produced by fitting a stochastic model to perturbed observations generated by applying the change factors derived from climate model simulations to the observations (e.g., Chun *et al.*, 2013; Manning *et al.*, 2009). A final approach is to directly use <u>drivers derived from future climate projections</u> as predictors in stochastic models, as suggested in Serinaldi & Kilsby (2012). Alternatively, a new unpublished study is exploring the use of UK climate projections based on specific global mean surface temperature thresholds (Chapter 1), as also outlined in Brown *et al.* (2014). For more details of stochastic simulations under different climate scenarios using the third approach, please see Chapter 2. In practice, the last two approaches have not been widely implemented.

Change factor method

As previously discussed, the incorporation of climate change scenarios into stochastic weather generators commonly involves the use of the change factor methodology (e.g., Wilks 2010) and can be used in both post-processing or stochastic model fitting approach. This methodology derives change factors for specific variables by comparing baseline observations with future climate projections from various climate model scenarios at different time slices (e.g., Kilsby *et al.*, 2007). Multiplicative factors are typically used for variables like rainfall, while additive factors are employed for other climate variables such as temperature (Glenis *et al.*, 2015). To capture the spatial and temporal variations of climate change impacts, the change factors may vary across months and catchments (Atkins, 2020). Stochastic models generally focus on the mean climate and thus can only use change factors associated with the mean (see also Chapter 1). Meanwhile UKCP09 weather generator work with higher order statistics such as variance, skewness, autocorrelation, and proportion of dry days, and thus, they can incorporate perturbations related to not only the mean but also all these higher order statistics (Jones *et al.*, 2010). However, the change factors with higher order statistics cannot be applied to all stochastic models.

Benefits of the climate change methods

Stochastic methods provide a somewhat reliable representation of the impact of global climate change on local-scale rainfall, generating consistent and reasonably realistic precipitation patterns (Chun *et al.*, 2013) at any suitable temporal and spatial resolutions (e.g., Kilsby *et al.*, 2007; Harris *et al.*, 2014). The change factor methodology, employed in stochastic weather generators, offers computational efficiency (Jones *et al.*, 2010) and flexibility to generate precipitation time series representing different climate scenarios or climate models (Kilsby *et al.*, 2007). While some approaches discussed earlier may not provide the transient evolution of future climate in their output (e.g., UKCP09 in Chun *et al.*, 2013; AME in Dawkins *et al.*, 2022), certain methods incorporate climate change over time by utilizing non-stationary change factors (e.g., UKCP09 in Glenis *et al.*, 2015) or incorporating global climate covariates that capture long-term variability (e.g., GLM in Chun *et al.*, 2013). Stochastic models can also estimate uncertainties arising from climate models and natural climatic variability through the sampling of multiple change factors in the weather generator and repeated realizations of climate projections, respectively (Glenis *et al.*, 2015).

Limitation of the climate change methods

Despite the advantages of employing stochastic methods in climate change studies, there are several limitations to the use of change factors. Firstly, the assumption of similarity between past and future climates in terms of variability and seasonality is not valid (Diaz-Neito and Wilby, 2005). Moreover, most weather generators lack basis in physical climate processes, further compromising their ability to accurately represent the changing climate (Jones *et al.*, 2010). The empirical relationships utilized by stochastic models to mimic the physical world are not guaranteed to remain valid under future climate conditions (Diaz-Neito and Wilby, 2005) and has potential implications with adequately reflecting changes to future drought frequency, magnitude, persistence and spatial extents. For more details on the non-stationarity of the predictor-to-rainfall relationship please see Chapter 4. Additionally, the inadequate representation of observed relationships between climate drivers and UK rainfall in climate models raises doubts about the skill and accuracy of the resulting stochastic climate simulation (Chun *et al.*, 2013).

Although certain studies mention the suitability of weather generators for climate change analysis, they do not evaluate and implement these methods (e.g., Serenaldi and Kilsby). Dawkins *et al.* (2022) mentions incorporating UK climate projections into Advanced Meteorology Explorer (AME) stochastic rainfall weather generator as part of future work, to generate rainfall simulations at any global mean temperature However, this work is yet to be published, although interviews given to our group suggest that the work in this direction is ongoing (Chapter 1). Further, review of the existing literature suggests that the few studies which explicitly evaluated limitations or quantified the associated uncertainties for the stochastically generated rainfall projections over UK, use the change factor method rather than directly applying projections to stochastic weather generators (e.g., Glenis *et al.*, 2015).

The uncertainties associated with stochastic weather generators may be exacerbated when used to generate events in a future climate. For instance, the UKCP09 weather generator's inability to accurately simulate very extreme events poses a problem, especially considering the expected worsening of extreme events in the future (Jones *et al.*, 2010). Additionally, biases present in the models at different sites and the lack of spatial relationships between sites during the current period can persist in future projections, introducing further potential errors (Chun *et al.*, 2013). Climate models and their projections exhibit biases stemming from factors such as model dynamics (Berthou *et al.*, 2020), resolution, parametrization (Chan *et al.*, 2013), domain extent, and driving data (Chan *et al.*, 2018), among other intricacies. Consequently, projections, especially for precipitation, have a large spread between climate models when downscaled to the catchment-scale, as the models cannot accurately represent local-scale dynamics (Smith *et al.*, 2013).

Recommendations

To enhance the reliability of stochastic models to explore the hydrological impacts of climate change, a comprehensive assessment and mitigation of inherent biases in these models are essential (Chun *et al.*, 2013). To improve reliability in climate projections, studies could implement approaches to aid climate model selection like machine learning techniques (e.g., Jewson and Hawkins, 2018) and correct the biases in the projections by comparing model projections with observational data using bias correction methods (e.g., Lafon *et al.*, 2012; Cloke *et al.*, 2012; Klein *et al.*, 2021). However, it should be noted that there is a wide range of bias correction methods which corrects for biases in different aspects of the model data (e.g., in mean, spread, trend or extremes). For bias-correcting climate projections, trend preserving methods are considered very important, which represent the long-term mean well, but may unduly influence extremes and small-scale features (e.g., Hempel *et al.*, 2013). While applying bias-correction, it should also be considered that the use of different bias-correction methods could increase the overall range of uncertainty (see Maraun, 2016). Therefore, it is crucial to robustly and systematically assess whether bias-correction methods are suitable for the intended needs before implementation.

The simple change factor approach can be complemented by comparing results with the direct use of bias-corrected projections (Cloke *et al.*, 2012; Smith *et al.*, 2013) or by employing perturbed weather generator methods (Fatichi *et al.*, 2011). Instead of a single set of change factors, the effects of natural climate variability can be estimated by re-sampling the observed time series to generate a wide range of change factor sets (e.g., Ledbetter *et al.* 2011) which can be used to perturb stochastic weather generator output or fit stochastic models. Model biases may be improved if output is benchmarked against observed or reconstructed historical data from different sites (Barker *et al.*, 2019), thereby underscoring the importance of a robust and representative observational record (Chapter 3). Further, rather than focusing on only improving the accuracy of climate projections, quantifying and framing uncertainties is equally crucial in climate change studies (e.g., Glenis *et al.*, 2015), as it enables informed decision-making (Harris *et al.*, 2014). Consequently, incorporating large ensembles of transient scenarios under a changing climate proves beneficial for stochastic models (e.g., Manning *et al.*, 2009; Chun *et al.*, 2013), enabling the development of a probabilistic framework to aid decision-making (Kilsby *et al.*, 2007). Furthermore, alternative methods for water resource projections should be considered. For example, the use of emulators applying extreme value distributions to generate rainfall simulations at various global carbon dioxide concentration levels (Brown *et al.*, 2014), or employing a storyline approach to navigate uncertainty in future climate risks (Sillmann *et al.*, 2020); please see Chapter 6 for details on alternative methods.

Summary of limitations

- Inaccurate representation of future climate due to lack of physical basis in weather generators.
- Sources of uncertainty include the assumption of constant variability in the change factor method and limited knowledge of how drivers of UK rainfall respond to climate change. Insufficient quantification of uncertainties in climate projections.
- Exacerbation of inherent biases and limitations of models with future climate applications may lead to cascade of uncertainty.

Summary of recommendations

- Prioritise integration of climate change influence for future methodological refinements and improvements for stochastic modelling
  Benefit: Incorporating climate change impacts into stochastic modelling will enhance

the resilience of water resource infrastructure decisions by ensuring their relevance and effectiveness in future conditions.
Effort: *Moderate to high.* While climate change is presently integrated into stochastic-based methodologies, typically through simplistic techniques like change factors (e.g., Jones *et al.*, 2010), fully incorporating climate change methodologies into stochastic models, encompassing non-stationarity and diverse uncertainties, presents a significant challenge. However, to maintain the applicability of stochastic outputs in future climate scenarios, it is crucial to assess various approaches to climate change adaptation within these frameworks.

- Improve climate model biases by benchmarking climate model output against good quality data.
  Benefit: Evaluating and benchmarking climate models against good quality observations allows for the best quality climate projections to be used as input for stochastic models.
  Effort: *Low to Moderate.* Many methods and frameworks are available for climate model evaluation and benchmarking (e.g., Watson-Parris *et al.*, 2022) that should be applied before using the climate model projections to calculate change factors. However, finding good quality source of observations over the correct time period to benchmark models can be challenging (see Chapter 3).
- Thoroughly assess and address the biases inherent in stochastic models to enhance their reliability in future climate studies.
  Benefit: Reducing biases in stochastic model output would allow for improved future projections by preventing from errors propagating through the stochastic modelling chain.
  Effort: *Low to Moderate.* Bias-corrected climate projections, widely used in climate studies, are readily available or can be computed with existing packages (see the point above). However, selecting the optimal bias-correction method for stochastic modelling output, presents a challenge, as different methods may address only specific bias features, such as in mean, spread, trend or extremes, rather than dealing with all types of biases comprehensively. Bias-correction methods should be systematically and robustly evaluated before implementation. For optimal results, it is recommended to apply bias-correction to both climate model projections and stochastic model output.
- Accurately quantify uncertainties in stochastic future projections to support informed decision-making.
  Benefit: Quantifying uncertainty enhances the credibility of the climate projections and might ultimately help in narrowing uncertainty by using methods such as emergent constraints (Hall *et al.*, 2019).
  Effort: *Moderate to high.* There are existing studies that provide methods to quantifying uncertainties in stochastic weather generators (e.g., Glenis *et al.*, 2015). However, implementing such methods while dealing with different kinds of uncertainties: model, scenario, and internal variability; and ultimately aiming to narrow the band of uncertainty (Hawkins and Sutton, 2009) can be quite challenging. This challenge is particularly relevant when transitioning these methodologies and findings into practical applications, where pragmatic considerations demand a careful balance between costs and benefits.

- Explore the use of large ensembles of different transient scenarios under changing climate in stochastic models to evaluate their robustness and effectiveness.
  Benefit: Using large ensemble projections allows for the quantification of uncertainties, while transient runs capture the dynamic evolution of climate variables, effectively representing the non-stationarity inherent in climate projections.
  Effort: *Moderate to high.* Many multi-model, large ensemble future climate projections of transient scenarios, that are bias-corrected, are readily available (e.g., CMIP6 simulations), and can be directly applied to stochastic weather generators. However, this process can escalate as high effort, particularly when considering the multitude of combinations between climate projections with large ensembles and many stochastic realizations.
- Consider the appropriateness of alternative methods to complement existing approaches to enhance future water resource planning (see Chapter 6)
  Benefit: Considering diverse approaches to future water resource planning allows for a more comprehensive understanding of future scenarios and ultimately results in more robust and adaptable strategies.
  Effort: *High.* Evaluating multiple alternative methods would need a lot of resources as it would require coordinated planning to compare different methods on a standardized basis for variety of criteria.

References

ANDERTON, S., LEDBETTER, R. AND PRUDHOMME, C., 2015. Understanding the performance of water supply systems during mild to extreme droughts, Environment Agency, Report – SC120048/R. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/484944/Performance_of_water_supply_systems_during_mild_to_extreme_droughts_report.pdf

ATKINS, 2020. Regional Climate Data Tools: Sutton and East Surrey Water on behalf of WRSE. URL: https://www.wrse.org.uk/media/ok1mtsoq/wrse_file_1338_regional-climate-data-tools.pdf

BARKER, L.J., HANNAFORD, J., PARRY, S., SMITH, K.A., TANGUY, M. AND PRUDHOMME, C., 2019. Historic hydrological droughts 1891–2015: systematic characterisation for a diverse set of catchments across the UK. Hydrology and Earth System Sciences 23(11), 4583-4602.

BERTHOU, S., KENDON, E.J., CHAN, S.C., BAN, N., LEUTWYLER, D., SCHÄR, C. AND FOSSER, G., 2020. Pan-European climate at convection-permitting scale: a model intercomparison study. Climate Dynamics 55, 35-59.

BROWN, S.J., MURPHY, J.M., SEXTON, D.M. AND HARRIS, G.R., 2014. Climate projections of future extreme events accounting for modelling uncertainties and historical simulation biases. Climate Dynamics 43: 2681-2705.

CHAN, S.C., KENDON, E.J., FOWLER, H.J., BLENKINSOP, S., FERRO, C.A. AND STEPHENSON, D.B., 2013. Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? Climate Dynamics 41, 1475-1495.

CHAN, S.C., KAHANA, R., KENDON, E.J. AND FOWLER, H.J., 2018. Projected changes in extreme precipitation over Scotland and Northern England using a high-resolution regional climate model. Climate Dynamics, 51, 3559-3577.

CHEN, J., BRISSETTE, F.P. AND LECONTE, R., 2012. Downscaling of weather generator parameters to quantify hydrological impacts of climate change. Climate Research 51(3), 185-200.

CHUN, K.P., WHEATER, H.S. AND ONOF, C., 2013. Comparison of drought projections using two UK weather generators. Hydrological sciences journal 58(2): 295-309.

CLOKE, H.L., WETTERHALL, F., HE, Y., FREER, J.E. AND PAPPENBERGER, F., 2013. Modelling climate impact on floods with ensemble climate projections. Quarterly Journal of the Royal Meteorological Society 139(671), 282-297.

DAWKINS, L.C., OSBORNE, J.M., ECONOMOU, T., DARCH, G.J. AND STONER, O.R., 2022. The advanced meteorology explorer: a novel stochastic, gridded daily rainfall generator. Journal of Hydrology, 607, 127478.

DIAZ-NIETO, J. AND WILBY, R.L., 2005. A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom. Climatic Change 69(2-3): 245-268.

Fatichi S, Ivanov VY, Caporali E (2011) Simulation of future climate scenarios with a weather generator. Advances in Water Resources 34(4), 448-467.

GLENIS, V., PINAMONTI, V., HALL, J.W. AND KILSBY, C.G., 2015. A transient stochastic weather generator incorporating climate model uncertainty. Advances in Water Resources 85, 14-26.

HALL, A., COX, P., HUNTINGFORD, C. AND KLEIN, S., 2019. Progressing emergent constraints on future climate change. Nature Climate Change, 9(4), 269-278.

HARRIS, C.N.P., QUINN, A.D. AND BRIDGEMAN, J., 2014. The use of probabilistic weather generator information for climate change adaptation in the UK water sector. Meteorological Applications 21(2), 129-140.

HAWKINS, E. AND SUTTON, R., 2009. The potential to narrow uncertainty in regional climate predictions. Bulletin of the American Meteorological Society, 90(8), 1095-1108.

HEMPEL, S., FRIELER, K., WARSZAWSKI, L., SCHEWE, J. AND PIONTEK, F., 2013. A trend-preserving bias correction–the ISI-MIP approach. Earth System Dynamics, 4(2), 219-236.

JEWSON, S. AND HAWKINS, E., 2018. Uncertain Climate Forecasts From Multimodel Ensembles: When to Use Them and When to Ignore Them. arXiv preprint arXiv:1006.5327.

JONES, P.D., KILSBY, C.G., HARPHAM, C., GLENIS, V. AND BURTON, A., 2010. UK Climate Projections science report: Projections of future daily climate for the UK from the Weather Generator. University of Newcastle, UK. URL: https://eprints.ncl.ac.uk/file_store/production/217986/B1B47084-F4B6-446C-A750-8E47554D8B7F.pdf

KILSBY, C.G., JONES, P.D., BURTON, A., FORD, A.C., FOWLER, H.J., HARPHAM, C., JAMES, P., SMITH, A. AND WILBY, R.L., 2007. A daily weather generator for use in climate change studies. Environmental Modelling & Software 22(12): 1705-1719.

KLEIN, C., JACKSON, L.S., PARKER, D.J., MARSHAM, J.H., TAYLOR, C.M., ROWELL, D.P., GUICHARD, F., VISCHEL, T., FAMIEN, A.M. AND DIEDHIOU, A., 2021. Combining CMIP data with a regional convection-permitting model and observations to project extreme rainfall under climate change. Environmental Research Letters 16(10), 104023.

LEDBETTER, R, PRUDHOMME, C., ARNELL, N., 2012. A method for incorporating climate variability in climate change impact assessments: Sensitivity of river flows in the Eden catchment to precipitation scenarios. Climatic Change 113, 803–823.

MANNING, L.J., HALL, J.W., FOWLER, H.J., KILSBY, C.G. AND TEBALDI, C., 2009. Using probabilistic climate change information from a multimodel ensemble for water resources assessment. Water Resources Research 45: W11411.

MARAUN, D., 2016. Bias Correcting Climate Change Simulations - a Critical Review. Current Climate Change Reports 2, 211–220.

SERINALDI, F. AND KILSBY, C.G., 2012. A modular class of multisite monthly rainfall generators for water resource management and impact studies. Journal of Hydrology 464, 528-540.

SILLMANN, J., SHEPHERD, T.G., VAN DEN HURK, B., HAZELEGER, W., MARTIUS, O., SLINGO, J. AND ZSCHEISCHLER, J., 2021. Event-based storylines to address climate risk. Earth's Future 9(2): e2020EF001783.

SMITH, A., BATES, P., FREER, J. AND WETTERHALL, F., 2014. Investigating the application of climate models in flood projection across the UK. Hydrological Processes, 28(5), 2810-2823.

THOMPSON, V., DUNSTONE, N.J., SCAIFE, A.A., SMITH, D.M., SLINGO, J.M., BROWN, S. AND BELCHER, S.E., 2017. High risk of unprecedented UK rainfall in the current climate. Nature Communications, 8(1), 107.

WATER UK, 2016. Water resources long term planning framework (2015-2065). Water UK. URL: https://www.water.org.uk/wp-content/uploads/2018/11/WaterUK-WRLTPF_Final-Report_FINAL-PUBLISHED-min.pdf

WATSON-PARRIS, D., RAO,Y., OLIVIÉ, D., SELAND, Ø., NOWACK, P., CAMPS-VALLS, G., STIER, P., BOUABID, S., DEWEY, M., FONS, E., GONZALEZ, J., HARDER, P., JEGGLE, K., LENHARDT, J., MANSHAUSEN, P., NOVITASARI, M., RICARD, L. AND ROESCH, C., 2022. ClimateBench v1. 0: A Benchmark for Data-Driven Climate Projections. Journal of Advances in Modeling Earth Systems, 14(10), e2021MS002954.

WILKS, D.S., 2010. Use of stochastic weather generators for precipitation downscaling. Wiley Interdisciplinary Reviews: Climate Change 1(6), 898-907.

Alternative Approaches

Rosanna Lane and Alison Kay (with review and input from the author team)

Introduction

In addition to stochastic datasets, a number of other approaches have been implemented to do a similar job to stochastic approaches, i.e., to provide a way of producing drought events more severe than those observed in the historical record. As such, these provide 'alternative approaches' that could be applied in a UK water resources planning context, to test the resilience of supply systems – and some have in fact already been used in an applied setting. Here, in line with the project brief, we survey the literature to understand (and classify) the range of approaches that could be used as credible alternatives to stochastic methods.

We have classified these other possible cross-cutting approaches to water resource planning into five broad categories:

1. **Extending the historical record**: Relatively short instrumental/observational records contain a limited number of drought events. Extending the record back in time can increase the number of recognised events, and include events more extreme than those experienced in the recent past. This is perhaps the most obvious way to produce droughts more severe than current observational records, by expanding the historical coverage of the records themselves.
2. **Synthetic droughts:** synthetically-created droughts, that have no physical basis, or which can be guided by historical experience/physical reasoning, can be used to test water supply systems.
3. **Scenario-neutral methods and sensitivity- / stress-testing frameworks:** Scenario-neutral methods (e.g., response surfaces) and sensitivity- / stress-testing frameworks, aim to assess the sensitivity of a system to a systematic set of changes in key drivers.
4. **Large ensembles**: By providing multiple climate simulations of historical (or future) periods, large ensembles provide a larger number of drought events than available from limited observational records, and can help to identify plausible extremes.
5. **Storylines / HILL events / H++ scenarios**: Storylines are physically self-consistent and plausible scenarios, which could be based on alternative realisations of past events or possible future events. They often focus on high-impact events, often referred to as High Impact Low Likelihood (HILL) events, but typically do not have probabilities attached to them. Similarly, H++ scenarios represent extreme climate change scenarios that typically lie beyond the 10^{th}-90^{th} percentile ranges often presented from model ensembles.

It should be noted that other categories, and different boundaries between them, may have been used in the literature, as there is overlap between the approaches and it is not necessarily the case that a particular study falls clearly into one category rather than another.

The approaches can be considered complementary, and multiple approaches are often used together. For example, storylines can be a useful way to distil results from large ensembles into a manageable yet representative set of scenarios for practitioners / policymakers (Clark *et al.*, 2016), and both methods to extend the historical record and large ensembles can be used to define the bounds of scenario-neutral methods (Chan *et al.*, 2023; Quinn *et al.*, 2020). These approaches should therefore not be considered in isolation, but as a set of alternative approaches that can be complementary in helping to inform water resource planning.

The following section describes each of the categories above in more detail, using examples from the literature. Note that large ensembles are differentiated from stochastic approaches here as they typically use physically-based climate models with rare outcomes that are spatially and internally consistent, rather than statistical models (Chan *et al.*, 2023; Kelder *et al.*, 2022).

Other approaches for water resource planning

Extending the historical record

A key limitation for water resource planning is the length of the observed time-series, with a limited number of observed droughts. Short observational records mean there are few drought events to test water supply systems against, the observations do not represent the full range of events that could arise from natural climate variability, and there are large uncertainties when trying to define extreme (e.g., 100-year return period) events. One approach to overcome this is to extend the historical record back in time, with options including: the use of stations with long observational records, recovering/rescuing previously lost observations, infilling/ statistical / modelling methods to extend observed time-series, documentary sources (such as newspapers and journals) to confirm past drought events, and paleo-climatology approaches.

A recent rainfall data rescue and recovery initiative has been successful in improving the UK rainfall record over earlier years. Hawkins *et al.* (2023) used a citizen science approach to digitise, quality control and make available over 3 million rainfall observations from 1677 to 1960, that were previously stored as hand written observations on paper sheets. This has significantly improved and extended UK rainfall records (Chapter 3).

A number of studies have extended rainfall and temperature observations back through the 1800s, to re-construct historic rainfall or streamflow droughts across the British Isles (Barker *et al.*, 2019; Hanel *et al.*, 2018; Lennard *et al.*, 2016; Murphy *et al.*, 2020; Noone *et al.*, 2017; Spraggs *et al.*, 2015; Watts *et al.*, 2012). For example, Watts *et al.* (2012) aimed to use long severe droughts of the 19th Century to test water companies' water supply and drought plans for events outside of recent experience. They used river flows that had been reconstructed back to 1800, based on a statistical model with inputs of monthly rainfall records and long-term average evapotranspiration (Jones *et al.*, 2006; Wade *et al.*, 2006). In an innovative modelling approach, they invited water managers and regulators from two case study systems to be part of the modelling process. This involved water managers responding to developing drought conditions, and applying different drought measures with respect to the drought management plans, with no prior knowledge of the drought magnitude or duration. For this exercise, the major advantage of using artificial droughts based on historical events (the droughts were created by stacking together past episodes) over entirely synthetically generated droughts was that events were more plausible and realistic – so water managers could fully engage without questions over whether the drought events were credible.

More recently, Murphy *et al.* (2020) and Noone *et al.* (2017) explored past droughts for Britain and Ireland using the standardised precipitation index (SPI), and Barker *et al.* (2019a) used reconstructed flow time series to characterise UK droughts back to 1891. Noone *et al.* (2017a) developed a drought catalogue for Ireland extending back to 1765. They used infilled/extended rainfall time-series back to 1850, and monthly precipitation reconstructions back to 1765 (Pauling *et al.*, 2006), with key drought events confirmed through historical sources such as newspaper archives. This revealed that recent decades (1990s onwards) are unrepresentative of the longer-term drought climatology, and identified a number of major drought events with diverse characteristics to strengthen the evidence base for future drought and water resource planning (Figure). Murphy *et al.* (2020) similarly reconstruct droughts for the British Isles, identifying a major 'forgotten' drought in 1765-1768, which offers an extreme benchmark scenario for stress-testing the resilience of water systems. Barker *et al.* (2019) further focused on streamflow droughts, identified using the standardised streamflow index (SSI), shedding light on regionally important events as well as events that had been poorly documented (such as during the post-war years in the 1940s). Studies have also used reconstructed historical streamflow drought events to test water supply systems (Lennard *et al.*, 2016; Spraggs *et al.*, 2015).

Whilst reconstructions are based on the historical record, they are often derived using a model which will likely contain some uncertainty. The size and nature of this uncertainty will depend on the method employed, but should be recognised when comparing historical reconstructions with other methods.

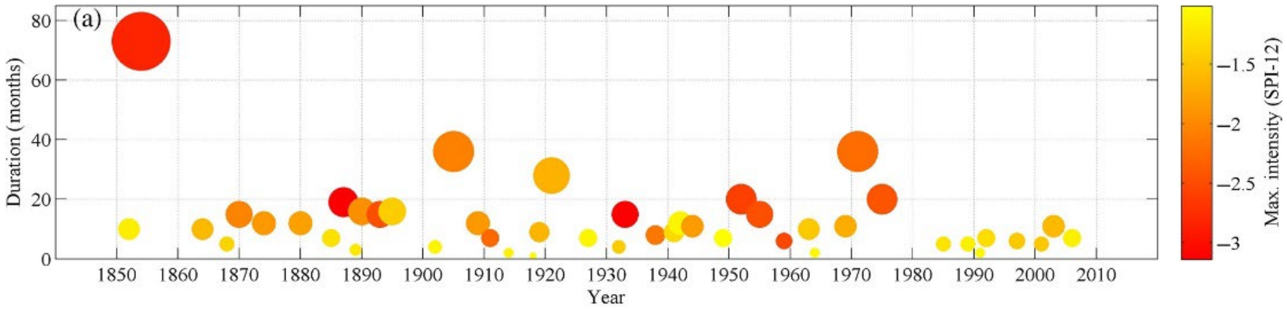

Figure 1: Drought duration against maximum intensity for the droughts identified by Noone *et al.* (2017) based on infilled/extended precipitation time-series. Circle size represents duration, colour relates to intensity. Figure reproduced from Noone *et al.* (2017).

The use of paleo-records provides the opportunity to extend records even further back in time, but with considerable assumptions and uncertainties. This includes reconstructions from climate proxies, such as tree ring widths (Cooper *et al.*, 2013; Wilson *et al.*, 2013), tree ring stable isotopes (Loader *et al.* 2020), and crop harvest date (Pribyl, 2020; Pribyl *et al.*, 2012). While most research on climate proxies focuses on temperature/ precipitation, some studies have also used proxy information to reconstruct monthly/annual streamflow (Stagge *et al.*, 2018; Viorica *et al.*, 2023). Despite very large uncertainties, these can help to better understand long term streamflow variability and trends.

A number of studies have used tree ring information as a proxy for spring-summer precipitation totals, to produce reconstructions extending back over 1000 years. For example, Wilson *et al.* (2013) and (Cooper *et al.*, 2013) use tree ring-width data from oak trees to reconstruct spring/summer precipitation for southern-central England, and East Anglia, respectively. Their results were broadly consistent, with similarity in the most prominent wet/dry periods, but more recent studies have highlighted inconsistencies between the statistical properties of observations and these reconstructions (Bothe *et al.*, 2019). Studies using tree ring stable isotopes produced more promising results, as demonstrated by Loader *et al.* (2020) in their reconstruction of May-August precipitation totals for England and Wales back to 1201. While these very long precipitation time-series provide opportunities to explore extreme droughts of the past, a key issue is the lack of winter precipitation information. An additional consideration regarding paleoclimate reconstruction is apparent in (Wilson *et al.*, 2013), where critical droughts in the calibration period are not well represented in the model used to convert tree ring width to precipitation. Involvement of those that will use the dataset is important to ensure the paleoclimate model reflects the phenomena of interest.

For reconstruction of temperature, one possible proxy is the starting date of grain harvest. (Pribyl *et al.* (2012) present a reconstruction of medieval April-July temperatures for

East Anglia, based on dates of the onset of grain harvest from manorial accounts. They used a linear regression to derive temperature from grain harvest dates for 1256-1431, calibrated on grain harvest dates for 1768-1816 when instrumental temperature records were available. Reconstructed temperature time-series could be used alongside reconstructed precipitation to explore extreme droughts of the past, but again there is a lack of data available for autumn/winter periods and large uncertainties.

For all methods that involve reconstruction or extension of records, the climate change signal means that previous periods, before the coverage of most instrumental records over the last c.50 years, may not be representative of the current climate. This applies especially in terms of temperature, although appraisals of nonstationarity in precipitation have become increasingly confident, especially in winter, and in northern and western areas of the UK.

Put another way, in a warming world, ‘the past is the key to the future’ breaks down as an argument. However, this is only a limitation if we try and use historical drought events directly, under an assumption of future recurrence. In reality, this constraint is well known in historical climate/paleo communities, but such reconstructions continue to be pursued, because knowledge of past extremes is still valuable information on past variability, that can be used to test the resilience of our current systems. Historical data is only one strand of evidence for assessing risk, along with climate model chains, extreme value analysis and so on. To return to the theme of this section on the use of such information in planning: past events derived through such approaches could be used for example, in ‘stress tests’, storylines or the various other approaches outlined in this report, alongside events from stochastic methods, future climate runs and so on. It is worth re-emphasising that this limitation also applies with stochastic estimation approaches – they only provide many more realisations of the current climate, and hence also have to account for climate change when applied in future contexts (Chapter 5).

Finally, as noted in Hannaford et al. 2024 “While the past may not be so readily a guide to the future in a warming world, at the same time observed historical droughts represent an important benchmark of drought risk, given that these events have actually unfolded – they also offer the opportunity to learn from past experiences in drought management”. Reconstruction of past events enables learning from real-world events that have been influential in shaping the evolution of water supply systems and management practices (e.g. Taylor et al. 2006), and which have impacted the environment and societies over very long timeframes (e.g. Pribyl et al. 2020). Given the anticipated magnitudes of future climate change over even the 21st Century timeframes, such learning has important benefits beyond the more immediate technical implementation of water resource planning.

Synthetic Droughts

Synthetic droughts (droughts created manually but systematically, that have no formal physical basis) have been used in both long-term water resources planning and operational drought management settings. Synthetic droughts have been used within the Drought Vulnerability Framework, in combination with Extreme Value Analysis, to underpin the Drought Response Surface (UKWIR, 2017), which is present in both WRMPs and Drought Plans, including assimilation of stochastic datasets where available. Drought events are generated by taking percentages of the long-term average precipitation (/temp) over specific durations covering alternative timings (start and end of the event). The response surface is built up by considering all permutations of these combinations of long-term average, event start, event end, etc – as illustrated in the following section. A similar approach is often taken to build scenarios of near-term future climate both within a Drought Plan (Thames Water, 2017) and within a drought event itself (Environment Agency, 2022), and statistics of the percentage of the long term average are used to describe river flows and precipitation in near-real time (e.g., Environment Agency, (2023)). This type of event generation is relatively simple and provides a means of altering the severity and duration of droughts on an event basis that is linked to the historical climatology of an area. It is therefore very useful for building stress tests, sampling across a range of different event types, as described in the following section.

Scenario neutral methods and sensitivity- / stress-testing frameworks

Scenario-neutral methods offer a bottom-up approach to water resource planning, focusing on understanding system sensitivity (Broderick *et al.*, 2019).

One scenario-neutral approach is the use of response surfaces to characterise system sensitivity to incremental changes in temperature/ precipitation (Anderton *et al.*, 2015; Prudhomme *et al.*, 2015; Whateley *et al.*, 2014), as shown in the examples given in Figure 1. Response surfaces have been used to characterise runoff sensitivity to climate variation (Arnell, 1996; Němec and Schaake, 1982), and were further developed by Prudhomme *et al.* (2010) in the context of flood risk. (Prudhomme *et al.* (2010) carried out thousands of hydrological model simulations per catchment to explore flood peak sensitivity to changes in the mean annual values and seasonality of precipitation and temperature. They then overlaid an ensemble of climate projections on the catchment response surfaces, to calculate how many scenarios the Government’s 20% allowance for climate change was appropriate for. This demonstrates a key advantage of the approach – that once response surfaces have been generated for a catchment, it is possible to rapidly assess different climate scenarios/ policy options without having to repeat hydrological model simulations.

The response surface approach has since been used to assess water resource system vulnerability to droughts and changing low flows (Anderton *et al.*, 2015; Prudhomme *et al.*, 2015; Sauquet *et al.*, 2019; Whateley *et al.*, 2014). Prudhomme *et al.* (2015) developed response surfaces looking at how low flows in two contrasting UK catchments responded to changes in rainfall and temperature, demonstrating the utility of the method in understanding future pressures on water resources. (Whateley *et al.*, 2014) generated response surfaces looking at the cost of water supply shortfalls, and how this relates to changing temperature and precipitation. They demonstrate how these can be used to identify the 'climate robustness' of different water supply systems and/or management options, based on the proportion of the response surface space over which the system provides acceptable performance. (Anderton *et al.*, 2015) present response surfaces for the impact of variations in drought duration (months) and intensity (rainfall deficit) against water supply system metrics (such as total unfilled demand), as shown in Figure 1. They highlighted that these response surfaces could be a useful screening tool, to estimate a water supply system's response to untested droughts of the past or future.

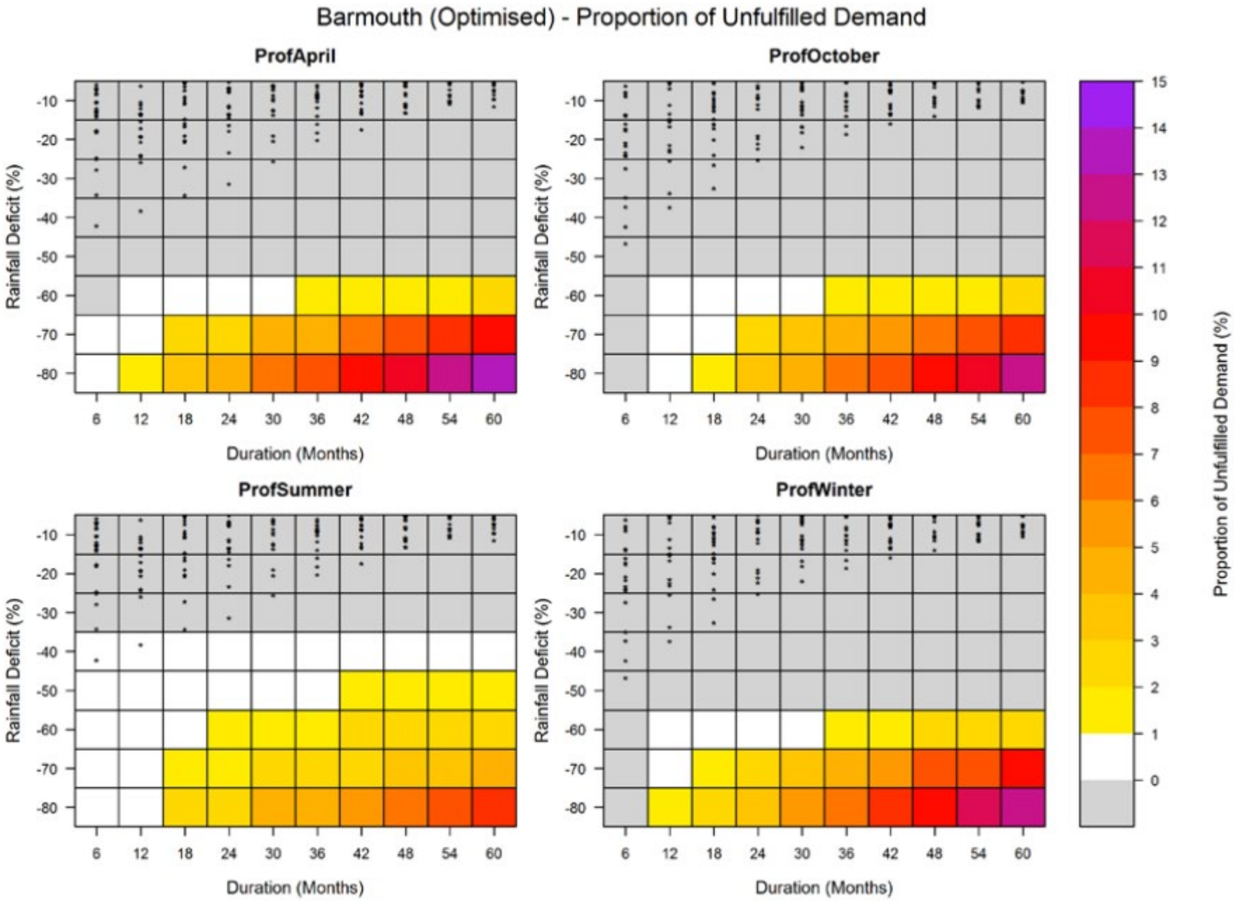

Figure 1: Example response surfaces, showing proportion of unfilled water resources demand in Barmouth given variations in drought duration (x axes) and drought intensity (rainfall deficit, y axes). Reproduced from Anderton *et al.*, 2015).

Another complementary scenario-neutral approach is to carry out systematic stress tests to understand drought sensitivity to changes in driving variables (Anderton *et al.*, 2015; Hellwig *et al.*, 2021). For example, Hellwig *et al.* (2021) carried out a stress test to explore drought sensitivity to precipitation seasonality. They increased winter and decreased summer precipitation by +/- 5%, 10%, 15%, 20% and 30% while increasing temperatures over the whole year. This helped to improve understanding of how a seasonal shift of recharge would impact groundwater and baseflow drought across Germany.

Response surfaces have been implemented within water resource planning through the Drought Vulnerability Framework (UKWIR, 2017) and have been included within water companies Drought Plans (e.g., (Thames Water, 2022)) to describe water supply system vulnerability. These surfaces can have climate change superimposed upon them, and may be derived from either synthetic (systematically altered climate variables) or stochastic data (UKWIR, 2017).

Large Ensembles

By providing multiple climate model simulations of historical (or future) periods, large ensembles provide a larger sample of extreme events than available from limited observational records – even in the context of very long rescued or reconstructed records (2.1). They can therefore help to identify plausible extremes that have no precedent in the observed record. Large ensembles can comprise; a) multiple climate model structures, b) perturbations to model physical parameters, c) sampling of initial conditions. Options b) and c) often derive from a single GCM, whilst option a) includes multi-model ensembles such as CMIP5/CMIP6. Climate model and perturbed physics ensembles (a & b) are generally used to understand climate modelling uncertainties, for example in the simulation of future low flow or drought events (Lane and Kay, 2021). Initial condition ensembles (c) can be useful for separating the effects of natural variability from other climate modelling uncertainties (Deser *et al.*, 2020).

A number of studies have used transient climate model output from multi-model GCM ensembles or single model perturbed-parameter ensembles (PPEs) to assess changing flows and drought risk across the UK (Dobson *et al.*, 2020; Fung *et al.*, 2013; Kay *et al.*, 2018; Lane *et al.*, 2022; Lane and Kay, 2021; Lopez *et al.*, 2009; Marx *et al.*, 2018; Rudd *et al.*, 2019). As one of the first studies to apply a PPE to drought risk, (Lopez *et al.*, 2009) demonstrate that the large ensemble approach provides a better understanding of the possible ranges of future outcomes and enables decision makers to more easily compare the merits of different management options. Borgomeo *et al.* (2018) show how large ensembles can be used to support risk-based decision-making and identify investments that are resilient to future uncertainties, within the context of London's water supply system, using the MaRIUS dataset. This dataset has been used in other water supply settings in the UK, including the National Framework (Environment Agency, 2020), and by Thames Water as a means of testing present and future drought coherence between the Thames and Severn catchments (Thames Water, 2020). This does not represent a use of the dataset as a central part of the planning process, however as a means of testing a specific characteristic of a particular water transfer option.

Initial condition ensembles are based on a single climate or weather forecasting model, with perturbations made to the initial conditions of each ensemble member (often referred to as Single Model Initial-condition Large Ensembles – SMILEs). A key advantage of SMILEs is that they focus on the aleatoric uncertainties (due to randomness arising from internal climate variability), as opposed to PPEs where aleatoric and epistemic uncertainties are combined (Chan *et al.*, 2023; Shepherd, 2019). They therefore offer an opportunity to more robustly sample the range of extreme events that are possible within climate variability (Mankin *et al.*, 2020).

An example is the UNSEEN (UNprecedented Simulated Extremes using ENsembles) technique, which uses the Met Office near-term climate prediction system to provide multiple simulations of the current climate (Thompson *et al.*, 2017). Analysis of these data showed a high chance of exceeding monthly rainfall records in many parts of the UK (Thompson *et al.*, 2017). An extension of the approach used a nested convection-permitting model to investigate intense summer daily rainfall, and estimated that a damaging storm in July 2007 could plausibly have had 50-100% more rainfall in the day (Kent *et al.*, 2022). Brunner and Slater, (2022) took a similar approach to directly investigate extreme floods by pooling an ensemble of reforecasts from the European Flood Awareness System. They conclude that such ensemble pooling is an efficient approach to increase sample size to derive more robust local and regional flood estimates, provided hydrological model performance is good. However, use of such forecast-system datasets for meteorological or hydrological drought is less straightforward because of the more prolonged nature of droughts, given the way the ensembles are performed (Kelder *et al.*, 2022). Recently, (Chan *et al.*, 2023) have used the EC-Earth time-slice large ensemble output to generate a large set of plausible droughts, and subsequently to estimate the chance of unprecedented drought events. These initial condition ensemble approaches therefore are of great value for vastly increasing our sample of events that could have been observed; improving understanding of plausible extremes within current climate variability even without climate change (which can be explored separately).

Storylines

Storylines offer another complementary approach to support water resources planning, that have been applied recently in the UK (Chan *et al.*, 2023, 2022; Fung *et al.*, 2022) and elsewhere (Gessner *et al.*, 2022; van der Wiel *et al.*, 2021; van Tiel *et al.*, 2023; Zhang *et al.*, 2020). Storylines are physically self-consistent and plausible scenarios, which could be based on alternative realisations of past events or plausible future pathways (Shepherd *et al.*, 2018). Generally, storylines focus on plausibility rather than probability, often focusing on high-impact events which could pose significant risks to society (Sillmann *et al.*, 2021).

The storylines approach can be used to identify plausible low-likelihood but high impact drought events, or ‘climate surprises’, which can then be used to stress-test water resource systems (Sillmann *et al.*, 2021; Woo, 2021). One example of this is Chan *et al.* (2022) who identifies physical climate storylines based on alternative realisations of past extreme droughts. They do this by looking at different ways that the 2010-2012 UK drought could have unfolded, given different initial conditions, seasonal precipitation, an extra dry year and climate change. Their results confirmed the importance of dry winters in the development of multi-year UK droughts, highlighting vulnerability to a ‘third dry winter’ storyline. They show that perturbing an observed event to create downward counterfactuals of how the event could have turned out worse could be beneficial from a risk awareness perspective. The results suggest there was considerable scope for a worse drought to have occurred, and conditions for all storylines could have resulted in more severe drought conditions than other past droughts (1975-76 and 1989-93). The third dry winter scenario has not been used by every water company, but by those that may have a vulnerability to such long droughts where other datasets used (e.g., stochastic) potentially do not sufficiently test such events (South East Water, 2019).

There is also the opportunity to combine probabilistic approaches (such as the UNSEEN estimates of the chance of extremes) with the storyline approach to sample for plausible worst cases of specific concern within the large sample of plausible events generated using a large ensemble. For example, van der Wiel *et al.* (2021) used a storyline approach to help understand how the 2018 European Drought could have been more severe under global warming, based on event analogues selected from a large climate ensemble. Similarly, (Chan *et al.*, 2023) sampled for drought storylines characterized by specific conditions (such as dry spring followed by dry summer and dry autumn followed by dry winter). Extreme event storylines could be used to stress test water resource systems.

Another example of a storylines approach is given in Hellwig *et al.* (2021), who carried out a sequence of stress tests that altered known droughts from the past. These included 1) exploring drought sensitivity to antecedent recharge conditions, by modelling known historic drought years with systematic reductions in the 3-, 9-, and 24-month antecedent recharge conditions; and 2) exploring system recovery from drought, with a simulation starting with the most severe historic drought conditions and testing how long it would take to recover under different recharge conditions (**Error! Reference source not found.**). The results improved understanding of groundwater drought sensitivity to meteorological forcing across Germany, highlighting the spatially heterogeneous groundwater response to changes in recharge.

Storylines can also be used to help interpret climate model output – distilling a large number of climate simulations into a manageable set of storylines that encompass key features and the range of plausible events (Chan *et al.*, 2023; Clark *et al.*, 2016; Shepherd *et al.*, 2018; Sillmann *et al.*, 2021). It can be difficult to interpret ensemble climate output, especially if projected changes for certain variables (such as rainfall) can span a wide range of uncertainty in both magnitude and direction. The use of an ensemble mean smooths out extreme events and results in features of events between individual model runs being merged, leading to a washed-out response that no longer relates to any individual model simulation (Shepherd *et al.*, 2018; Zappa *et al.*, 2021). Taking a more probabilistic approach, such as looking at the most likely future river flow/ drought changes across multi-model ensembles, may risk paying insufficient attention to plausible high-impact but low likelihood events that are of most concern to decision-makers (Sutton, 2019). Developing storylines from climate simulations to help communicate the uncertainties from large ensembles overcomes these issues as the storyline approach seeks to understand the relevant casual factors that have led to an extreme event and ask how those factors could have made the event worse (such as with climate change). Thus, individual storylines need not be associated with probabilities as they represent alternative plausible hypotheses, recognizing that the lack of statistical significance (i.e. wide uncertainty ranges) does not dismiss the potential for significant changes in risk (Rodrigues and Shepherd, 2022; Zappa *et al.*, 2021).

Another key advantage of the storylines approach is its use for evaluating risk from compound events, going beyond just climate change and looking at other potential stresses to water systems (Bevacqua *et al.*, 2023, 2021).

Shepherd *et al.* (2018) summarises four key reasons for taking a storylines approach in climate impact analyses. First, it improves risk awareness by providing information in a way that is intuitive; storylines draw on our natural tendency to value information on lived-events rather than statistical likelihoods when evaluating risk. Second, storylines can be framed in a context that is relevant to specific decision-makers, and is particularly effective for considering risks from compound events which may not arise from a purely statistical method. Third, it can help with partitioning model uncertainties, by considering storylines for different aspects of model uncertainty instead of generalising across a large ensemble where all uncertainties are combined. Finally, it can be used to explore the bounds of plausibility, exploring mechanisms that are not currently captured by GCMs (for example due to incomplete physics, insufficient sampling or insufficient resolution).

Advantages and limitations

A summary of the key advantages / issues with each of the discussed alternative methods for water resource planning is given in Table 1.

Table 1: Summary of the key advantages/ problems with alternative methods for water resources planning.

Approach	Advantages	Problems
Extending the historical record	1. Provides a larger sample of extreme events than the observational time-series 2. Events are plausible and realistic 3. Improves understanding of climate variability	1. Limited numbers of new events 2. Large uncertainties in historical data 3. Does not incorporate climate change
Scenario-neutral / sensitivity frameworks	1. Can identify management options that are robust to a wide range of possible futures 2. Separate to climate projections, but can be combined 3. Improves understanding of system vulnerabilities	1. Choice of the variables and magnitude of changes shown on response surfaces has a large impact on the results 2. Can only show two variables at once: simplification of a complex problem 3. Some perturbations not realistic
Large ensembles	1. Give a much greater sample of extreme events 2. More physically-based than statistical or stochastic methods for generating extremes 3. Simulations are spatially and internally consistent – can look at climate drivers of rare events 4. SMILEs isolate effects of internal climate variability	1. Climate models may contain biases/ missing processes for simulation of extremes 2. Large volumes of data – challenge to distil into useful information, run models and visualise 3. An individual SMILE does not consider uncertainties associated with different climate model structures
Storylines	1. Basing storylines around real events gives them plausibility 2. Can be combined with scenario-neutral approaches and large ensembles to explore low-likelihood but high impact events, like a third dry winter, to stress-test systems 3. Helpful in planning to have options distilled into a discrete set of storylines	1. Event storylines focus on individual events and it may be difficult to generalize 2. Hard to assign likelihoods to individual storylines 3. Cannot be used to assess the likely characteristics of extreme future droughts

Summary of recommendations

- Create and utilise long time-series of plausible flows to test the resilience of water resource systems to climate variability. These could be developed through a range of different approaches, including 1) extending the historical record back in time using modelling approaches and/or proxy data, 2) collating output from large climate ensembles and using them to drive hydrological models.
  Benefit: Extended time-series better represent the range of drought events that are plausible within our current climate, ensuring that risk assessments are not biased by the limited observational record.
  Effort: Moderate to high. Historic flow reconstructions and flow derived from climate ensembles are freely available to download for selected sites (Hannaford *et al.*, 2022; Smith *et al.*, 2018). Further development of longer-term reconstructions e.g., using longer rescued climate data would be a significant investment. Work on proxies is advanced, but developing these into useable time series for risk assessment (and appraisal of their utility via proof-of-concepts) would require more significant investment.
- Develop plausible high-impact but low likelihood future drought scenarios to explore possible extreme droughts. This could be through the use of drought storylines developed around past events, or rare events extracted from large climate model ensembles.
  Benefit: A set of high-impact storylines can be an effective way to stress-test water systems, and highlight potential failures.
  Effort: Moderate. Use of a discrete set of storylines reduces computational demands while focusing on the most extreme events. There is currently significant planned work on developing storylines for extreme drought under various projects (CANARI, Climate+) that could be leveraged to support such investments.
- Use scenario-neutral methods (such as response surfaces and stress testing frameworks) to assess the sensitivity of water resources systems.
  Benefit: Scenario-neutral methods can help to identify management options that are robust to a range of possible futures, and improve understanding of system vulnerabilities. They improve understanding of potential future pressures on systems while remaining independent from specific climate projections.
  Effort: Moderate. Large numbers of model runs may be required to develop response surfaces, but modelling doesn't need to be repeated when new climate projections are released.
- Combine multiple alternative approaches for a more complete understanding of the robustness of water resource systems.
  Benefit: These alternate approaches are complementary, and there are similarities in approach between the storyline, scenario-neutral and stress testing approaches. A set of alternative approaches would be complementary in helping to inform water resource planning. It would be beneficial to explore ways to combined and 'mix and match' these approaches, developing a modular toolkit to allow multiple approaches to be applied, according to the particular water resource system in question, in a consistent way (that is, with consistent datasets and methods for combination). Such an approach could be explored through demonstrators/case studies that could

pave the way for the development of guidance that advises on which methods to use in what circumstances, and how they can be effectively combined (and when they should not).

Effort: High. Some of these approaches have been developed and applied in relative isolation and while there have been some efforts to integrate these, significant new work would be required to develop and test these approaches in combination, and apply them to water supply systems. Pilot/demonstration approaches in a limited number of cases/systems would be more tractable.

References

ANDERTON, S., LEDBETTER, R. AND PRUDHOMME, C., 2015. Understanding the performance of water supply systems during mild to extreme droughts. Horizon House, Bristol.

ARNELL, N.W., 1996. Global Warming, River Flows and Water Resources. Wiley, Chichester.

BARKER, L.J., HANNAFORD, J., PARRY, S., SMITH, K.A., TANGUY, M. AND PRUDHOMME, C., 2019. Historic hydrological droughts 1891–2015: systematic characterisation for a diverse set of catchments across the UK. Hydrol Earth Syst Sci 23, 4583–4602. https://doi.org/10.5194/hess-23-4583-2019

BEVACQUA, E., DE MICHELE, C., MANNING, C., COUASNON, A., RIBEIRO, A.F.S., RAMOS, A.M., VIGNOTTO, E., BASTOS, A., BLESIĆ, S., DURANTE, F., HILLIER, J., OLIVEIRA, S.C., PINTO, J.G., RAGNO, E., RIVOIRE, P., SAUNDERS, K., VAN DER WIEL, K., WU, W., ZHANG, T. AND ZSCHEISCHLER, J., 2021. Guidelines for Studying Diverse Types of Compound Weather and Climate Events. Earths Future 9. https://doi.org/10.1029/2021EF002340

BEVACQUA, E., SUAREZ-GUTIERREZ, L., JÉZÉQUEL, A., LEHNER, F., VRAC, M., YIOU, P. AND ZSCHEISCHLER, J., 2023. Advancing research on compound weather and climate events via large ensemble model simulations. Nat Commun 14, 2145. https://doi.org/10.1038/s41467-023-37847-5

BORGOMEO, E., MORTAZAVI-NAEINI, M., HALL, J.W. AND GUILLOD, B.P., 2018. Risk, Robustness and Water Resources Planning Under Uncertainty. Earths Future 6, 468–487. https://doi.org/10.1002/2017EF000730

BOTHE, O., WAGNER, S. AND ZORITA, E., 2019. Inconsistencies between observed, reconstructed, and simulated precipitation indices for England since the year 1650 CE. Climate of the Past 15, 307–334. https://doi.org/10.5194/cp-15-307-2019

BRODERICK, C., MURPHY, C., WILBY, R.L., MATTHEWS, T., PRUDHOMME, C. AND ADAMSON, M., 2019. Using a Scenario-Neutral Framework to Avoid Potential Maladaptation to Future Flood Risk. Water Resour Res 55, 1079–1104. https://doi.org/10.1029/2018WR023623

BRUNNER, M.I. AND SLATER, L.J., 2022. Extreme floods in Europe: going beyond observations using reforecast ensemble pooling. Hydrol Earth Syst Sci 26, 469–482. https://doi.org/10.5194/hess-26-469-2022

CHAN, W.C.H., ARNELL, N.W., DARCH, G., FACER-CHILDS, K., SHEPHERD, T.G., TANGUY, M. AND VAN DER WIEL, K., 2023. Current and future risk of unprecedented hydrological droughts in Great Britain. J Hydrol (Amst) 625, 130074. https://doi.org/10.1016/j.jhydrol.2023.130074

CHAN, W.C.H., SHEPHERD, T.G., FACER-CHILDS, K., DARCH, G. AND ARNELL, N.W., 2022. Storylines of UK drought based on the 2010–2012 event. Hydrol Earth Syst Sci 26, 1755–1777. https://doi.org/10.5194/hess-26-1755-2022

CLARK, M.P., WILBY, R.L., GUTMANN, E.D., VANO, J.A., GANGOPADHYAY, S., WOOD, A.W., FOWLER, H.J., PRUDHOMME, C., ARNOLD, J.R. AND BREKKE, L.D., 2016. Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. Curr Clim Change Rep 2, 55–64. https://doi.org/10.1007/s40641-016-0034-x

COOPER, R.J., MELVIN, T.M., TYERS, I., WILSON, R.J.S. AND BRIFFA, K.R., 2013. A tree-ring reconstruction of East Anglian (UK) hydroclimate variability over the last millennium. Clim Dyn 40, 1019–1039. https://doi.org/10.1007/s00382-012-1328-x

DESER, C., LEHNER, F., RODGERS, K.B., AULT, T., DELWORTH, T.L., DINEZIO, P.N., FIORE, A., FRANKIGNOUL, C., FYFE, J.C., HORTON, D.E., KAY, J.E., KNUTTI, R., LOVENDUSKI, N.S., MAROTZKE, J., MCKINNON, K.A., MINOBE, S., RANDERSON, J., SCREEN, J.A., SIMPSON, I.R. AND TING, M., 2020. Insights from Earth system model initial-condition large ensembles and future prospects. Nat Clim Chang 10, 277–286. https://doi.org/10.1038/s41558-020-0731-2

DOBSON, B., COXON, G., FREER, J., GAVIN, H., MORTAZAVI‐NAEINI, M. AND HALL, J.W., 2020. The Spatial Dynamics of Droughts and Water Scarcity in England and Wales. Water Resour Res 56. https://doi.org/10.1029/2020WR027187

ENVIRONMENT AGENCY, 2023. Monthly Water Situation Report 1–23.

ENVIRONMENT AGENCY, 2022. National Drought Group projections for winter water resources [WWW Document].

ENVIRONMENT AGENCY, 2020. Meeting our future water needs: a national framework for water resources.

FUNG, F., GODDARD, C., MATTE, D., BALLINGER, A., HEGERL, G., O'REILLLY, C., WEISHEIMER, A., CROKER, T., MCSWEENEY, C., VAN DER WIEL, K. AND WILCKE, R., 2022. Storylines of predictions and projections developed and assessed for potential applications. European Climate Prediction system (EUCP).

FUNG, F., WATTS, G., LOPEZ, A., ORR, H.G., NEW, M. AND EXTENCE, C., 2013. Using Large Climate Ensembles to Plan for the Hydrological Impact of Climate Change in the Freshwater Environment. Water Resources Management 27, 1063–1084. https://doi.org/10.1007/s11269-012-0080-7

GESSNER, C., FISCHER, E.M., BEYERLE, U. AND KNUTTI, R., 2022. Multi-year drought storylines for Europe and North America from an iteratively perturbed global climate model. Weather Clim Extrem 38, 100512. https://doi.org/10.1016/j.wace.2022.100512

HANEL, M., RAKOVEC, O., MARKONIS, Y., MÁCA, P., SAMANIEGO, L., KYSELÝ, J. AND KUMAR, R., 2018. Revisiting the recent European droughts from a long-term perspective. Sci Rep 8, 9499. https://doi.org/10.1038/s41598-018-27464-4

HANNAFORD, J., MACKAY, J.D., ASCOTT, M., VICTORIA, A., CHITSON, T., COLE, S., COUNSELL, C., DURANT, M., JACKSON, C.R., KAY, A.L., LANE, R.A., MOORE, R., PARRY, S., RUDD, A.C., FACER-CHILDS, K., TURNER, S. AND WALLBANK, J.R., 2022. eFLaG: enhanced future FLows and Groundwater. A national dataset of hydrological projections based on UKCP18. Earth System Science Data Discussions. https://doi.org/https://doi.org/10.5194/essd-2022-40

HANNAFORD, TURNER, S., CHEVUTURI, A., CHAN., W., BARKER, L., TANGUY, M., PARRY, S., ALLEN, S. Have river flow droughts become more severe? A review of the evidence from the UK, a data-rich temperate environment. Hydrology and Earth System Science Discussions. Available at: https://hess.copernicus.org/preprints/hess-2024-293/

HAWKINS, E., BURT, S., MCCARTHY, M., MURPHY, C., ROSS, C., BALDOCK, M., BRAZIER, J., HERSEE, G., HUNTLEY, J., MEATS, R., O'GRADY, J., SCRIMGEOUR, I. AND SILK, T.,

2023. Millions of historical monthly rainfall observations taken in the UK and Ireland rescued by citizen scientists. Geosci Data J 10, 246–261. https://doi.org/10.1002/gdj3.157

HELLWIG, J., STOELZLE, M. AND STAHL, K., 2021. Groundwater and baseflow drought responses to synthetic recharge stress tests. Hydrol Earth Syst Sci 25, 1053–1068. https://doi.org/10.5194/hess-25-1053-2021

JONES, P.D., LISTER, D.H., WILBY, R.L. AND Kostopoulou, E., 2006. Extended riverflow reconstructions for England and Wales, 1865–2002. International Journal of Climatology 26, 219–231. https://doi.org/10.1002/joc.1252

KAY, A.L., BELL, V.A., GUILLOD, B.P., JONES, R.G. AND RUDD, A.C., 2018. National-scale analysis of low flow frequency: historical trends and potential future changes. Clim Change 147, 585–599. https://doi.org/10.1007/s10584-018-2145-y

KELDER, T., MARJORIBANKS, T.I., SLATER, L.J., PRUDHOMME, C., WILBY, R.L., WAGEMANN, J. AND DUNSTONE, N., 2022. An open workflow to gain insights about low-likelihood high-impact weather events from initialized predictions. Meteorological Applications 29. https://doi.org/10.1002/met.2065

KENT, C., DUNSTONE, N., TUCKER, S., SCAIFE, A.A., BROWN, S., KENDON, E., SMITH, D., MCLEAN, L. AND GREENWOOD, S., 2022. Estimating unprecedented extremes in UK summer daily rainfall. Environmental Research Letters 17, 014041. https://doi.org/10.1088/1748-9326/ac42fb

KIM, D., ONOF, C., 2020. A stochastic rainfall model that can reproduce important rainfall properties across the timescales from several minutes to a decade, Journal of Hydrology 589, 125150. https://doi.org/10.1016/j.jhydrol.2020.125150

LANE, R.A., COXON, G., FREER, J., SEIBERT, J. AND WAGENER, T., 2022. A large-sample investigation into uncertain climate change impacts on high flows across Great Britain. Hydrol Earth Syst Sci 26, 5535–5554. https://doi.org/10.5194/hess-26-5535-2022

LANE, R.A. AND KAY, A.L., 2021. Climate Change Impact on the Magnitude and Timing of Hydrological Extremes Across Great Britain. Frontiers in Water 3. https://doi.org/10.3389/frwa.2021.684982

LENNARD, A.T., MACDONALD, N., CLARK, S. AND HOOKE, J.M., 2016. The application of a drought reconstruction in water resource management. Hydrology Research 47, 646–659. https://doi.org/10.2166/nh.2015.090

LOADER, N.J., YOUNG, G.H.F., MCCARROLL, D., DAVIES, D., MILES, D. AND BRONK RAMSEY, C., 2020. Summer precipitation for the England and Wales region, 1201–2000 <scp>ce</scp> , from stable oxygen isotopes in oak tree rings. J Quat Sci 35, 731–736. https://doi.org/10.1002/jqs.3226

LOPEZ, A., FUNG, F., NEW, M., WATTS, G., WESTON, A. AND WILBY, R.L., 2009. From climate model ensembles to climate change impacts and adaptation: A case study of water resource management in the southwest of England. Water Resour Res 45. https://doi.org/10.1029/2008WR007499

MANKIN, J.S., LEHNER, F., COATS, S. AND MCKINNON, K.A., 2020. The Value of Initial Condition Large Ensembles to Robust Adaptation Decision-Making. Earths Future 8. https://doi.org/10.1029/2020EF001610

MARX, A., KUMAR, R., THOBER, S., RAKOVEC, O., WANDERS, N., ZINK, M., WOOD, E.F., PAN, M., SHEFFIELD, J. AND SAMANIEGO, L., 2018. Climate change alters low flows in Europe under global warming of 1.5, 2, and 3 °C. Hydrol Earth Syst Sci 22, 1017–1032. https://doi.org/10.5194/hess-22-1017-2018

MILLS, T.C., 2005. Modelling precipitation trends in England and Wales. Meteorological Applications. 12(2):169-176. https://doi.org/10.1017/S1350482705001611

MURPHY, C., WILBY, R.L., MATTHEWS, T., HORVATH, C., CRAMPSIE, A., LUDLOW, F., NOONE, S., BRANNIGAN, J., HANNAFORD, J., MCLEMAN, R. AND JOBBOVA, E., 2020. The forgotten drought of 1765–1768: Reconstructing and re-evaluating historical droughts in the British and Irish Isles. International Journal of Climatology 40, 5329–5351. https://doi.org/10.1002/joc.6521

NĚMEC, J. AND SCHAAKE, J., 1982. Sensitivity of water resource systems to climate variation. Hydrological Sciences Journal 27, 327–343. https://doi.org/10.1080/02626668209491113

NOONE, S., BRODERICK, C., DUFFY, C., MATTHEWS, T., WILBY, R.L. AND MURPHY, C., 2017. A 250-year drought catalogue for the island of Ireland (1765-2015). International Journal of Climatology 37, 239–254. https://doi.org/10.1002/joc.4999

PAULING, A., LUTERBACHER, J., CASTY, C. AND WANNER, H., 2006. Five hundred years of gridded high-resolution precipitation reconstructions over Europe and the connection to large-scale circulation. Clim Dyn 26, 387–405. https://doi.org/10.1007/s00382-005-0090-8

PRIBYL, K., 2020. A survey of the impact of summer droughts in southern and eastern England, 1200–1700. Climate of the Past 16, 1027–1041. https://doi.org/10.5194/cp-16-1027-2020

PRIBYL, K., CORNES, R.C. AND PFISTER, C., 2012. Reconstructing medieval April-July mean temperatures in East Anglia, 1256–1431. Clim Change 113, 393–412. https://doi.org/10.1007/s10584-011-0327-y

PRUDHOMME, C., SAUQUET, E. AND WATTS, G., 2015. Low Flow Response Surfaces for Drought Decision Support: A Case Study from the UK. Journal of Extreme Events 02, 1550005. https://doi.org/10.1142/S2345737615500050

PRUDHOMME, C., WILBY, R.L., CROOKS, S., KAY, A.L. AND REYNARD, N.S., 2010. Scenario-neutral approach to climate change impact studies: Application to flood risk. J Hydrol (Amst) 390, 198–209. https://doi.org/10.1016/j.jhydrol.2010.06.043

QUINN, J.D., HADJIMICHAEL, A., REED, P.M. AND STEINSCHNEIDER, S., 2020. Can Exploratory Modeling of Water Scarcity Vulnerabilities and Robustness Be Scenario Neutral? Earths Future 8. https://doi.org/10.1029/2020EF001650

RODRIGUES, R.R. AND SHEPHERD, T.G., 2022. Small is beautiful: climate-change science as if people mattered. PNAS Nexus 1. https://doi.org/10.1093/pnasnexus/pgac009

RUDD, A.C., KAY, A.L. AND BELL, V.A., 2019. National-scale analysis of future river flow and soil moisture droughts: potential changes in drought characteristics. Clim Change 156, 323-340. https://doi.org/10.1007/s10584-019-02528-0

SAUQUET, E., RICHARD, B., DEVERS, A. AND PRUDHOMME, C., 2019. Water restrictions under climate change: a Rhône–Mediterranean perspective combining bottom-up and top-down approaches. Hydrol Earth Syst Sci 23, 3683–3710. https://doi.org/10.5194/hess-23-3683-2019

SHEPHERD, T.G., 2019. Storyline approach to the construction of regional climate change information. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 475, 20190013. https://doi.org/10.1098/rspa.2019.0013

SHEPHERD, T.G., BOYD, E., CALEL, R.A., CHAPMAN, S.C., DESSAI, S., DIMA-WEST, I.M., FOWLER, H.J., JAMES, R., MARAUN, D., MARTIUS, O., SENIOR, C.A., SOBEL, A.H., STAINFORTH, D.A., TETT, S.F.B., TRENBERTH, K.E., VAN DEN HURK, B.J.J.M., WATKINS, N.W., WILBY, R.L. AND ZENGHELIS, D.A., 2018. Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. Clim Change 151, 555–571. https://doi.org/10.1007/s10584-018-2317-9

SILLMANN, J., SHEPHERD, T.G., VAN DEN HURK, B., HAZELEGER, W., MARTIUS, O., SLINGO, J. AND ZSCHEISCHLER, J., 2021. Event-Based Storylines to Address Climate Risk. Earths Future 9. https://doi.org/10.1029/2020EF001783

SMITH, K.A., TANGUY, M., HANNAFORD, J. AND PRUDHOMME, C., 2018. Historic reconstructions of daily river flow for 303 UK catchments (1891–2015). https://doi.org/10.5285/f710bed1-e564-47bf-b82c-4c2a2fe2810e,.

SPRAGGS, G., PEAVER, L., JONES, P. AND EDE, P., 2015. Re-construction of historic drought in the Anglian Region (UK) over the period 1798–2010 and the implications for water resources and drought management. J Hydrol (Amst) 526, 231–252. https://doi.org/10.1016/j.jhydrol.2015.01.015

STAGGE, J.H., ROSENBERG, D.E., DEROSE, R.J. AND RITTENOUR, T.M., 2018. Monthly paleostreamflow reconstruction from annual tree-ring chronologies. J Hydrol (Amst) 557, 791–804. https://doi.org/10.1016/j.jhydrol.2017.12.057

SUTTON, R.T., 2019. Climate Science Needs to Take Risk Assessment Much More Seriously. Bull Am Meteorol Soc 100, 1637–1642. https://doi.org/10.1175/BAMS-D-18-0280.1

TAYLOR, V., CHAPPELS, H, MEDD, W., TRENTMANN, F. 2009. Drought is normal: the socio-technical evolution of drought and water demand in England and Wales, 1893 – 2006. J. Hist. Geog. 35, 3, 568 – 591. https://doi.org/10.1016/j.jhg.2008.09.004

THAMES WATER, 2022. Draft Drought Plan 2022.

THAMES WATER, 2020. Water Resources Management Plan 2019 - Executive Summary.

THAMES WATER, 2017. Final Drought Plan April 2017 Main Report.

THOMPSON, V., DUNSTONE, N.J., SCAIFE, A.A., SMITH, D.M., SLINGO, J.M., BROWN, S. AND BELCHER, S.E., 2017. High risk of unprecedented UK rainfall in the current climate. Nat Commun 8, 107. https://doi.org/10.1038/s41467-017-00275-3

UKWIR, 2017. Drought Vulnerability Framework. Report Ref. 17/WR/02/12.

VAN DER WIEL, K., LENDERINK, G.AND DE VRIES, H., 2021. Physical storylines of future European drought events like 2018 based on ensemble climate modelling. Weather Clim Extrem 33, 100350. https://doi.org/10.1016/j.wace.2021.100350

VAN TIEL, M., WEILER, M., FREUDIGER, D., MORETTI, G., KOHN, I., GERLINGER, K. AND STAHL, K., 2023. Melting Alpine Water Towers Aggravate Downstream Low Flows: A Stress-Test Storyline Approach. Earths Future 11. https://doi.org/10.1029/2022EF003408

VIORICA, N., CĂTĂLIN-CONSTANTIN, R., ANDREI, M., MARIAN-IONUȚ, ȘTIRBU, IONEL, P. AND MONICA, I., 2023. The first tree-ring reconstrruction of streamflow variability over the last ~250 years in the Lower Danube. J Hydrol (Amst) 617, 129150. https://doi.org/10.1016/j.jhydrol.2023.129150

VILLARINI, G. and SERINALDI, F. (2012), Development of statistical models for at-site probabilistic seasonal rainfall forecast. Int. J. Climatol., 32: 2197-2212. https://doi.org/10.1002/joc.3393

WADE, S.D., JONES, P.D. AND OSBORN, T., 2006. The Impacts of Climate Change on Severe Droughts: Implications for Decision Making. Environment Agency Science Report: SC040068/SR.

WATTS, G., CHRISTIERSON, B. VON, HANNAFORD, J. AND LONSDALE, K., 2012. Testing the resilience of water supply systems to long droughts. J Hydrol (Amst) 414–415, 255–267. https://doi.org/10.1016/j.jhydrol.2011.10.038

WHATELEY, S., STEINSCHNEIDER, S. AND BROWN, C., 2014. A climate change range-based method for estimating robustness for water resources supply. Water Resour Res 50, 8944-8961. https://doi.org/10.1002/2014WR015956

WILSON, R., MILES, D., LOADER, N.J., MELVIN, T., CUNNINGHAM, L., COOPER, R. AND BRIFFA, K., 2013. A millennial long March-July precipitation reconstruction for southern-central England. Clim Dyn 40, 997–1017. https://doi.org/10.1007/s00382-012-1318-z

WOO, G., 2021. A counterfactual perspective on compound weather risk. Weather Clim Extrem 32, 100314. https://doi.org/10.1016/j.wace.2021.100314

ZAPPA, G., BEVACQUA, E. AND SHEPHERD, T.G., 2021. Communicating potentially large but non-robust changes in multi-model projections of future climate. International Journal of Climatology 41, 3657–3669. https://doi.org/10.1002/joc.7041

ZHANG, L., ZHENG, H.X., TENG, J., CHIEW, F.H.S. POST, D.A., 2020. Plausible Hydroclimate Futures for the Murray-Darling Basin. A report for the Murray–Darling Basin Authority. CSIRO, Australia.

Would you like to find out more about us or your environment?

Then call us on

03708 506 506 (Monday to Friday, 8am to 6pm)

Email: enquiries@environment-agency.gov.uk

Or visit our website

www.gov.uk/environment-agency

Incident hotline

0800 807060 **(24 hours)**

Floodline

0345 988 1188 **(24 hours)**

Find out about call charges (https://www.gov.uk/call-charges)

Environment first

Are you viewing this onscreen? Please consider the environment and only print if absolutely necessary. If you are reading a paper copy, please don't forget to reuse and recycle.