# Partitioning model uncertainty in multi-model ensemble river flow projections

Gordon Aitken[1] · Lindsay Beevers[1,2] · Simon Parry[3] · Katie Facer-Childs[3]

## Abstract

Floods are the largest natural disaster currently facing the UK, whilst the incidents of droughts have increased in recent years. Floods and droughts can have devastating consequences on society, resulting in significant financial damage to the economy. Climate models suggest that precipitation and temperature changes will exacerbate future hydrological extremes (i.e., floods and droughts). Such events are likely to become more frequent and intense in the future; thus to develop adaptation plans climate model projections feed hydrological models to provide future water resource projections. 'eFLaG' is one set of future river flow projections produced for the UK driven by UKCP18 climate projections from the UK Met Office. The UKCP18-derived eFLaG dataset provides state-of-the-art projections for a single GCM driven by RCP 8.5 across the entire UK. A QE-ANOVA approach has been used to partition contributing sources of uncertainty for two flow quantiles (Q5 high flows and Q95 low flows), at near and far future time scales, for each of the 186 GB catchments in the eFLaG dataset. Results suggest a larger hydrological model uncertainty associated with low flows and greater regional climate model uncertainty for high flows which remains stationary between flow indicators. Total uncertainty increases from near to far future and highly uncertain catchments have been identified with a high concentration in South-East England.

**Keywords** UKCP18 · Regional climate model · QE-ANOVA

## 1 Introduction

Floods and droughts can have devastating consequences on society. Recent events across Europe such as the floods in continental Europe in the summer of 2021, followed by widespread drought in the summer of 2022 highlight the real impact such events have on people. The July 2021 event devastated large areas of Belgium and Germany resulting from

✉ Gordon Aitken
  ga41@hw.ac.uk

1  Water Resilient Cities, Institute for Infrastructure and Environment, Heriot-Watt University, Edinburgh, UK

2  Institute for Infrastructure and Environment, University of Edinburgh, Edinburgh, UK

3  UK Centre for Ecology & Hydrology, Wallingford, UK

recording breaking rainfall across the region and the highest recorded river discharge in the Meuse and Rhine catchments (Copernicus 2022). A combination of factors including a slow moving low pressure which stalled over Northern Europe, and saturated soils led to devastating flooding which killed over 200 people and caused billions of euros in damages (AXA 2022). The following year, in the summer of 2022 widespread drought across Europe resulted in exceptional water and heat stress and impacting crop yields and energy generation (Toreti et al. 2022). The Po river basin was particularly affected, with the lowest water levels on record and saline intrusion inland from the delta. The hot and dry conditions across the region additionally exacerbated the wildfire threat. These are clearly severe events with catastrophic impacts to lives, livelihoods, economies and the environment. With the impact of anthropogenic climate change affecting weather patterns, the likelihood of encountering such events more frequently is increasing (IPCC 2020).

As natural hazards such as floods and droughts are set to intensify (Visser et al. 2019; Hannaford et al. 2022), alongside increasing global urbanisation, significant challenges arise in the way we live with and manage our future water resources (Beevers et al. 2022; Watts et al. 2015; Parry et al. 2023). Societies need to adapt to the changing climate, increasing the resilience of both communities and infrastructure to such shocks. In order to adapt efficiently, there is a need to understand future changes in climate variability and how that may impact river flows in terms of their extremes.

Future projections of climate are available (Aitken et al. 2022a). These are often the result of a modelling chain where global climate models are used to drive regional climate models giving higher resolution projections of future climate variables. These outputs can be post-processed (bias correction and/or downscaling) and used as inputs to hydrological models in order to deliver future of river flow time series projections. This is a resource intensive process, particularly when multi-model ensembles are considered, and require long, complex modelling chains. The modelling chain is subject to cascading uncertainties as one model feeds another. Different uncertainties manifest through the chain arising from parameter choice, input uncertainty and model structure to name a few (Beevers et al. 2020). An ensemble approach to model results aims to quantify some of the uncertainties within the chain (Smith et al. 2019).

The latest generation of UK climate projections (UKCP18; Lowe et al. 2018) provides an updated state-of-the-art assessment of the future changes in the climatology of the UK. The 'eFLaG' dataset capitalised on these newly available climate projections to derive corresponding river flow, groundwater level and recharge projections for the UK (Hannaford et al. 2023). The previous generation of flow projections — the 'Future flows and groundwater levels' dataset (Prudhomme et al. 2013) derived from UKCP09 (Jenkins et al. 2010) — has been used for a wide range of applications (e.g., Collet et al. 2017). The most important aspects of the eFLaG dataset for understanding future change are its spatial consistency, transient attributes, and the inclusion of multiple RCM-HM combinations. Assessments of future water resources are often undertaken independently by around 30 water companies across the UK. In addressing questions of future water resource availability, companies adopt a range of approaches that are not readily comparable (despite recent attempts to improve regional and national cohesion). This has important implications when considering potential approaches for enhancing the resilience of regional and national water resources, such as inter-regional water transfers. A spatially coherent dataset such as eFLaG is required to better understand drought occurrence and characteristics in different regions prior to expensive investment decisions. A further advantage of datasets such as eFLaG are the transient nature of data which highlight the importance of inter-decadal variability in a manner which

traditional time slice data cannot provide. Parry et al. (2023) have undertaken the first analysis of changing future drought for the UK based on the eFLaG dataset, but only discuss RCM and HM uncertainty in a qualitative manner.

As with all future projections the model chains required result in cascading uncertainties. These uncertainties need careful navigation in order to ensure that the important information tracks through to their use in decision-making (Smith et al. 2018). Picking out the necessary detail requires an understanding of the projections. For example, which uncertainties are the most important to consider and quantify in a modelling chain; how can these be narrowed down; which need most attention and where should more effort be focussed? Until recently, the common perception was that climate model uncertainty was the dominant mode of variability in climate impact studies (Wilby and Harris 2006; Thober et al. 2018; Zhang et al. 2022). However, over the last decade, there has been increasing evidence that hydrological model uncertainty is at least as important and in some instances more important than climate model uncertainty (Melsen et al. 2018; Collet et al. 2017, 2018; Visser-Quinn et al. 2019). Thus in model cascades which typically includes global and regional climate models, alongside scenario choice and hydrological models, there is a pressing need to untangle the relative components which may be contributing to the total projection uncertainty. The relevance of such studies is critical to feed into the decision making context (Smith et al. 2018). For example, in the UK, there are critical questions for water managers around what the future entails for water resources and hydrological extremes. The answers to these questions will inform significant investment decisions with ramifications for water companies and their customers.

Partitioning the cascading model uncertainty provides critical insight into internal variability of future climate projections. Analysing the interdependencies of modelling choices highlights the contributions of model choices on uncertainties. Typically, this is performed through empirical statistical analysis or a formal analysis of variance (ANOVA) method (Hingray and Saïd 2014). Quantification of uncertainty contributions has seen significant interest recently with a number of studies applying ANOVA to large multi-model climate ensembles (Bosshard et al. 2013; Addor et al. 2014; Visser-Quinn et al. 2019). Examining the partitioned uncertainty provides key insight into catchment behaviours and allocation of research efforts along the multi-stage climate modelling process.

Different ANOVA methods are available with QE-ANOVA selected for the work herein. A quasi-ergodic ANOVA (QE-ANOVA) framework for partitioning climate ensemble variability was first introduced by Hawkins and Sutton (2011) before being formalised by Hingray and Saïd (2014). It provides an analysis of variance under the ergodic assumption for climate projections. It considers variations over large time scales to be smooth (ergodic assumption) with internal variations creating any observed high-frequency changes. QE-ANOVA has been applied to large multi-model ensembles to identify uncertainties associated with choices in general circulation models, regional climate models and hydrological models (Kay et al. 2009; Vidal et al. 2016; Visser-Quinn et al. 2019). The close overlap between work allows similar application of QE-ANOVA to the eFLaG dataset. An alternative approach is single time ANOVA (STANOVA) method. This considers ANOVA for a select time step $t$, disregarding all other data. However, a comparison of both methods has identified QE-ANOVA as a more appropriate assessment of model variations (Hingray et al. 2019).

Consequently the aim of this paper is to understand and quantify uncertainty in the newly available eFLAG dataset which has been developed for the UK as a pilot climate service to support enhanced resilience of water resources. The analysis addresses a number of specific questions:

1. How does total uncertainty across the ensemble of the streamflow projections for Great Britain vary across the country?
2. How do the relative contributions of the modelling chain regional climate model (RCM) uncertainty and hydrological model (HM) uncertainty vary regionally?
3. How do the answers to these questions vary spatially (across GB) and temporally (2050s and 2080s), at both high (Q5) and low (Q95) flows?

## 2 Methods

The eFlaG dataset contains 48 modelling chains for UK flow series derived from the UKCP18 climate projections. UKCP18 uses a single GCM (HadGEM3-GC3.05) driven by one emissions scenario (RCP 8.5) but applies multiple perturbations of the HadREM3-GA705 RCM and multiple HM's to assess uncertainty in the latter two modelling stages. Twelve regional climate models are combined with four hydrological models in a balanced matrix. Analysing uncertainty contributions requires multiple climate models to create ensemble time series. This provides model variability which can be quantified to calculate the contribution from the two sources (RCM and HM), without considering GCM and emissions scenario uncertainty (several studies have investigated GCM uncertainty (Addor et al. 2014; Bosshard et al. 2013; Lane et al. 2022; Prudhomme and Davies 2009)). Flow quantiles can then be extrapolated from the bias corrected precipitation outputs of the eFLaG's dataset.

Assessing the relative contributions requires an analysis of variance (ANOVA). In this case, a quasi-ergodic ANOVA assessment has been performed which assumes that all states are captured when time periods are sufficiently long (Hingray and Saïd 2014). Applying QE-ANOVA to river flow series can separate the relative contributions of the multi-stage modelling chain, providing a deeper understanding of the sources of variability. Figure 1 demonstrates the overview of the methodology, which is explained in more detail in Section 2.1–2.4.

### 2.1 River flow projections for the UK

#### 2.1.1 UKCP18 climate projections

The eFLaG hydrological projections were derived from the UKCP18 'Regional' PPE dataset representing the state-of-the-art in future climate projections for the UK. The regional PPE contains 12 ensemble members representing variations in the boundary conditions of the Met Office Hadley Centre climate model. The original UKCP18 precipitation data were bias-corrected to ensure the resulting daily time series of precipitation and potential evapotranspiration (PET) required for hydrological modelling better reflect observations. The bias-correction followed an approach previously used in the UK and is described briefly herein with full details presented in Hannaford et al. (2023). A monthly mean approach was used with change factors calculated between RCM and observed data for every grid cell over the period 1981–2010. Change factor grid smoothing was applied using weighted combinations of the original value and neighbouring cells to prevent discontinuities. Simulated RCM precipitation time series were then multiplied by the factor grid for each month. PET was calculated using input variables from the regional PPE and according to the approach used by CHESS (Robinson et al. 2016). The 12-km regional PPE data were also

**Fig. 1** Method framework for the separation of uncertainties in the eFLaG's dataset
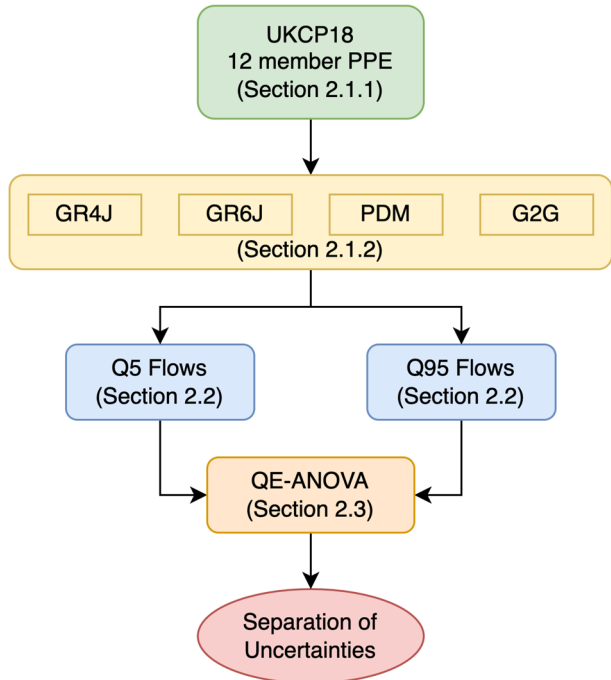


downscaled spatially to the 1-km resolution necessary as input to hydrological models for simulating flows in some relatively small catchments. Downscaling was performed using the distribution of standard-period average annual rainfall (SAAR) (Bell et al. 2007; Kay and Crooks 2014).

The processing described above resulted in 1 km grids of daily precipitation and PET for the whole of the UK spanning a timeframe of December 1980 to November 2080 for each of the 12 ensemble members. Whilst these gridded data were used to drive the distributed G2G model, the lumped catchment models (GR4J, GR6J, PDM) required catchment average driving data. Following a spatial averaging process using catchment shapefiles, one precipitation and one PET time series were generated for each catchment and each of the 12 ensemble members.

Of particular relevance for this study contrasting climate model uncertainty with hydrological model uncertainty, it is important to note that neither the Regional PPE nor its bias-corrected product comprehensively represent the full range of climate model uncertainty. This is because the 12 ensemble members only represent boundary condition variability within the same climate model, rather than a range of model structures or processes. Similarly, the climate model data are based only on the RCP8.5 emissions scenario, a relatively high emissions scenario which does not allow for the possibility of emissions reductions in future.

### 2.1.2 Hydrological modelling

Four hydrological models were driven by the UKCP18 climate data in order to generate a multi-model ensemble of past, present and future river flows (the 'simrcm').

The distributed 'Grid-2-Grid' model (G2G) is a gridded model at 1 km spatial resolution providing coverage of Great Britain. G2G does not cover catchments in Northern Ireland thus all results for this model are run across the 186 GB catchments in the eFLaGs dataset. In contrast to the other models, G2G simulates naturalised river flows since it is not calibrated to observations (therefore implicitly incorporating artificial influences into flow estimates). In addition, of the four models applied herein, G2G is the only one to produce gridded estimates of flows, although for comparability with the other three models in this study, flows are extracted at the catchment outlet locations. G2G has been demonstrated to perform acceptably in a range of previous drought and low flow assessments in the UK (e.g., Bell et al. 2018; Rudd et al. 2019).

The lumped catchment 'Probability Distributed Model' (PDM) is a very flexible modelling framework that can adopt a wide variety of different model structures. These configurations were explored as part of the setup process and the most appropriate one selected from a range of options as identified following multiple rounds of calibration.

Two models from the 'airGR' family (Coron et al. 2017) were selected for use in the eFLaG multi-model ensemble: GR4J and GR6J, with the number identifying how many parameters are available to calibrate. Simple to understand and relatively easy to calibrate, they have previously been applied on numerous occasions to simulate the full range of river flows (e.g., Harrigan et al. 2018, Anglian Water Drought Plan, 2021).

Comprehensive information on hydrological model setup for the eFLaG projections are provided by Hannaford et al. (2023).

## 2.2 High and low flow metrics and hydrological model performance

Flow quantiles are used as simple metrics of high and low flows. Flow exceedance percentiles (Qxx) quantify the flow that is exceeded xx% of the time and are considered to be robust when calculated from a sufficient record length of data, typically a 30-year timeframe at a minimum. A range of different high and low flow quantiles have been applied in high and low flows studies, and there is no consensus on the most appropriate metric, which tends to be application- and system-specific. Here, Q5 is used for high flows and Q95 is used for low flows. Q5 and Q95 are calculated from daily data within 30-year time windows by ordering daily flows from highest to lowest and extracting the 5th and 95th percentiles to give the flows exceeded 5% and 95% of the time. Capitalising on the benefits of the transient nature of the river flow projections, Q5 and Q95 is calculated for every simrcm run (48 time series for each catchment) for all 30-year time windows across the 1983–2079 timeframe. The Q5 and Q95 values in 2012 represent Q5 and Q95 calculated across 1983–2012, in 2013 represent 1984–2013, and so on to 2079 (2050–2079).

The hydrological models employed in eFLaG were evaluated using a range of metrics to assess the models' capabilities in simulating high, median and low flows. These metrics included absolute percent error in Q95 (Q95_APE), among others, for low flows, whilst high flows were evaluated more generically in eFLaG using the Nash Sutcliffe efficiency (NSE) and the modified Kling Gupta Efficiency (KGE2) (Hannaford et al. 2023, Figs. 4 and S2). For this study, absolute percent error in Q5 was also calculated to assess high flows more rigorously. Results suggest reasonable agreement between observed and simulated values across all four hydrological models (Hannaford et al. 2023). The lumped catchment models showed very similar performance, with KGE and NSE scores above 0.8, Q95_APE for GR6J and PDM was below 20%, and Q5_APE below 5% for most catchments (Supplementary material S1 and S2). GR6J performed slightly better than GR4J, and PDM showed the best results. G2G had

worse performance, with KGE and NSE scores above 0.6 for most catchments, though median Q95_APE for G2G was 48% and median Q05_APE was 11%. The poorer performance from G2G is likely due to its simulations of naturalised flows as opposed to the other models that are directly calibrated against observations, and occurs mainly in the south-east region where catchments have significant groundwater interactions and water transfer schemes, giving flows more heavily influenced by abstractions which are more difficult to explicitly simulate.

## 2.3 QUALYPSO

Climate model uncertainty can be partitioned into contributing sources using quasi-ergodic analysis of variance (QE-ANOVA) (Hingray & Saïd 2014; Vidal et al. 2016). This method estimates the uncertainty associated with climate model projections, and partitions the specific sources in a statistical framework. QE-ANOVA highlights contributing sources of uncertainty and provides a deeper understanding of dominant uncertainties across the dataset.

The first step to performing QE-ANOVA is to estimate the noise-free signal (NFS) for the change variable $X$ (e.g., flow percentile). This is performed using a trend model of the raw projections $Y$, giving

$$\widehat{NFS}(g,h,t) = \frac{y(g,h,t)}{y(g,h,b)} - 1 \tag{1}$$

where $g$ represents the RCM, $h$ represents the hydrological model with $y(g,h,t)$ and $y(g,h,b)$ the trend estimates of the raw projections across the future time period ($t$) and baseline period ($b$), respectively. Having calculated the NFS, it is assumed that the noise-free change response can be partitioned as

$$\widehat{NFS}(g,h,t) = \mu(t) + \alpha(g,t) + \beta(h,t) + \gamma(g,h,t) \tag{2}$$

where $\mu(t)$ is the overall climate response which represents the complete ensemble mean at time $t$, $\alpha(g,t)$ and $\beta(h,t)$ are the mean deviations of RCM $g$ and HM $h$ from the ensemble mean $\mu(t)$, and $\gamma(g,h,t)$ is the residual. These parameters can be estimated using a classical two-way ANOVA approach with no interactions [Berrington de González and Cox (2007); Searle (1971)].

Firstly, define the means across each variable (e.g., $g$, $h$), with the dot symbol representing averaging over the index, as

$$\widehat{NFS}(g,\bullet,t) = \frac{1}{N_h} \sum_{h=1}^{N_h} \widehat{NFS}(g,h,t), \tag{3}$$

$$\widehat{NFS}(\bullet,h,t) = \frac{1}{N_g} \sum_{g=1}^{N_g} \widehat{NFS}(g,h,t), \tag{4}$$

$$\widehat{NFS}(\bullet,\bullet,t) = \frac{1}{N_g N_h} \sum_{g=1}^{N_g} \sum_{h=1}^{N_h} \widehat{NFS}(g,h,t) \tag{5}$$

Then, from a least squares estimation using the constraints $\sum_{g=1}^{N_g} \widehat{\alpha}(g,t) = 0$ and $\sum_{h=1}^{N_h} \beta(h,t) = 0$, the parameter estimations in Eq. (2) can be given as

$$\hat{\alpha}(g, t) = \widehat{NFS}(g, \bullet, t) - \widehat{NFS}(\bullet, \bullet, t), \tag{6}$$

$$\hat{\beta}(h, t) = \widehat{NFS}(\bullet, h, t) - \widehat{NFS}(\bullet, \bullet, t), \tag{7}$$

$$\hat{\mu}(t) = \widehat{NFS}(\bullet, \bullet, t), \tag{8}$$

$$\hat{\gamma}(g, h, t) = \widehat{NFS}(g, h, t) - \hat{\mu}(g, t) - \hat{\alpha}(g, t) - \hat{\beta}(h, t). \tag{9}$$

The total uncertainty associated with the NFS can then defined as

$$\mathbb{V}ar\left(\widehat{NFS}(g, h, t)\right) = \mathbb{V}ar\left(\hat{\alpha}(g, t)\right) + \mathbb{V}ar\left(\hat{\beta}(h, t)\right) + \mathbb{V}ar\left(\hat{\gamma}(g, h, t)\right). \tag{10}$$

This was performed using the "*QUALYPSO*" function in the *QUALYPSO* R-package (Evin 2022). This function runs QE-ANOVA across the input ensemble dataset with contributions extracted from the figure data. The QUALYPSO function requires a smoothing factor for fitting the trendlines to the baseline and future projections using cubic smoothing splines. A smoothing factor of 1 (default value) was used following the functions authors Evin et al. (2019).

## 2.4  Catchment selection

River flow projections were derived for a representative set of 186 catchments in Great Britain (excluding Northern Ireland catchments due to G2G not covering these). Ensuring an acceptable geographical coverage was an important selection criterion to maximise the inclusion of the breadth of hydrometeorological and hydrogeological variability in GB. As an additional layer of validation, the representativeness of catchment characteristics within the subset was checked against the distribution of values in the full set of catchments on the UK National River Flow Archive. The majority of UK catchments are subject to some level of artificial influence, and the catchment selection was not limited to only those that cannot be empirically quantified. An effort to identify near-natural "benchmark network" catchments (Harrigan et al. (2018) identified 80 such catchments (shaded blue in Fig. 2), of which 11 were flagged to be used with caution due to minor influence). However, the catchment selection for the eFLaG simulations, and this study, was not limited to these benchmark catchments as the assessment of hydrometric networks with national importance was critical to determine the implications of uncertainty for water resources management in GB. Catchments were also chosen to reassess those included in the previous study of GB river flow projections by Prudhomme et al. (2013). Finally, catchments were selected according to criteria of record length (covering the entire baseline period (1981–2012) in Hannaford et al. (2023)) and data completeness, as well as expert visual judgement on data quality. Feedback was sought from both the water industry and research scientists to maximise utility of the river flow projections. This selection includes catchments in the south-east which have very complex geology (chalk aquifers) and are heavily exploited for drinking water and agricultural abstractions, thus representing catchments with significant human influence.

Of these 186 catchments, a subset of 12 (Fig. 2) were selected to illustrate a range of examples of uncertainty partitioning. Trios of catchments (Figs. 2 and 3) were identified to represent four general categories of HM uncertainty:
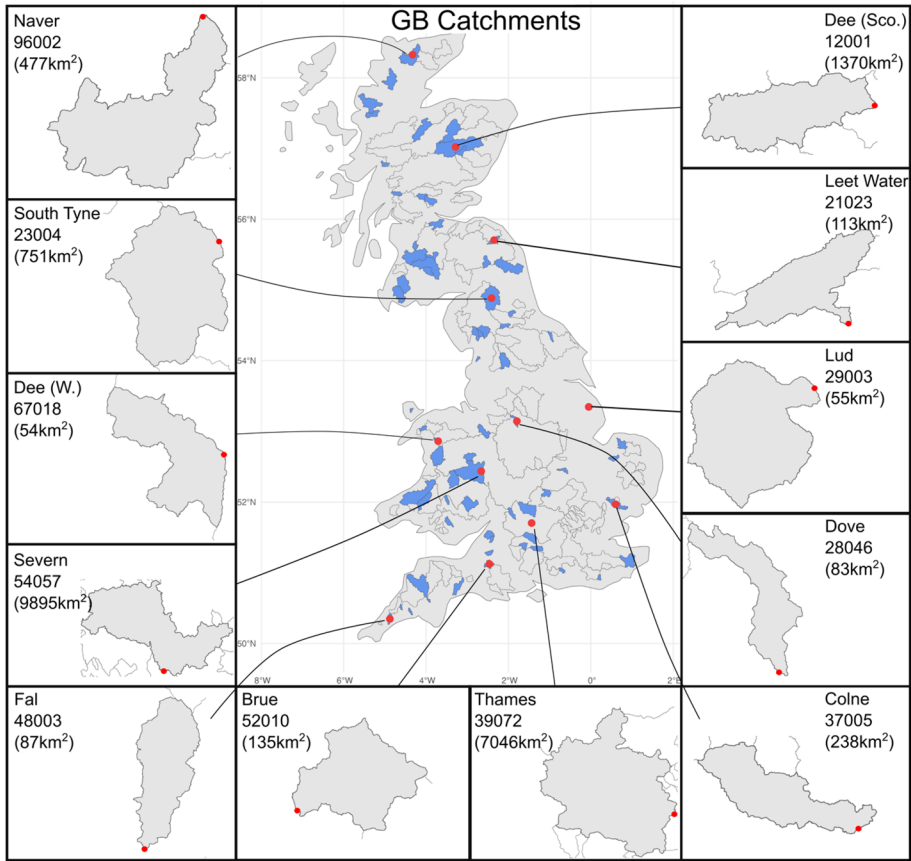
**Fig. 2** eFLaG catchments with groupings — all catchments included in the analysis are depicted by their outline. Sample catchments for comparison and analysis purposes are identified and indicated in the figure by red dots. Benchmark network catchments are shown in blue

A. Similar projections across all four HMs;
B. Similar projections across three of four HMs, with one outlier;
C. Different projections between each of two pairs of HMs;
D. Different projections for each of the four HMs.

Three catchments were selected to represent each of these four categories as defined for transient Q95 data in the far future (2080). When sub-sampling these 12 catchments, consideration was made to ensure a reasonable geographical spread across the UK as well as a representative cross-section of catchment characteristics (Table 1). Catchment areas range from 53 to 9885km$^2$ (median of 187km$^2$), median catchment elevations range from 68 to 509 m (median of 118 m), catchment average annual rainfall ranges from 566 to 2022 mm (median of 982 mm), base flow index values range from 0.27 to 0.90 (median of 0.53), and urban proportion of catchment ranges from 0.3 to 10.4% (median of 3.5%).

## 3 Results

In order to examine the results of the analysis, a few different types of results were extracted. Firstly, the paper explores the results of the high and low flow metrics across catchments, which are a direct output of the eFLaG modelling. From there, the results then explore the original research questions posed in Section 1. The total uncertainty in the model cascade as quantified using the QE-ANOVA method set out in Section 2.3 for
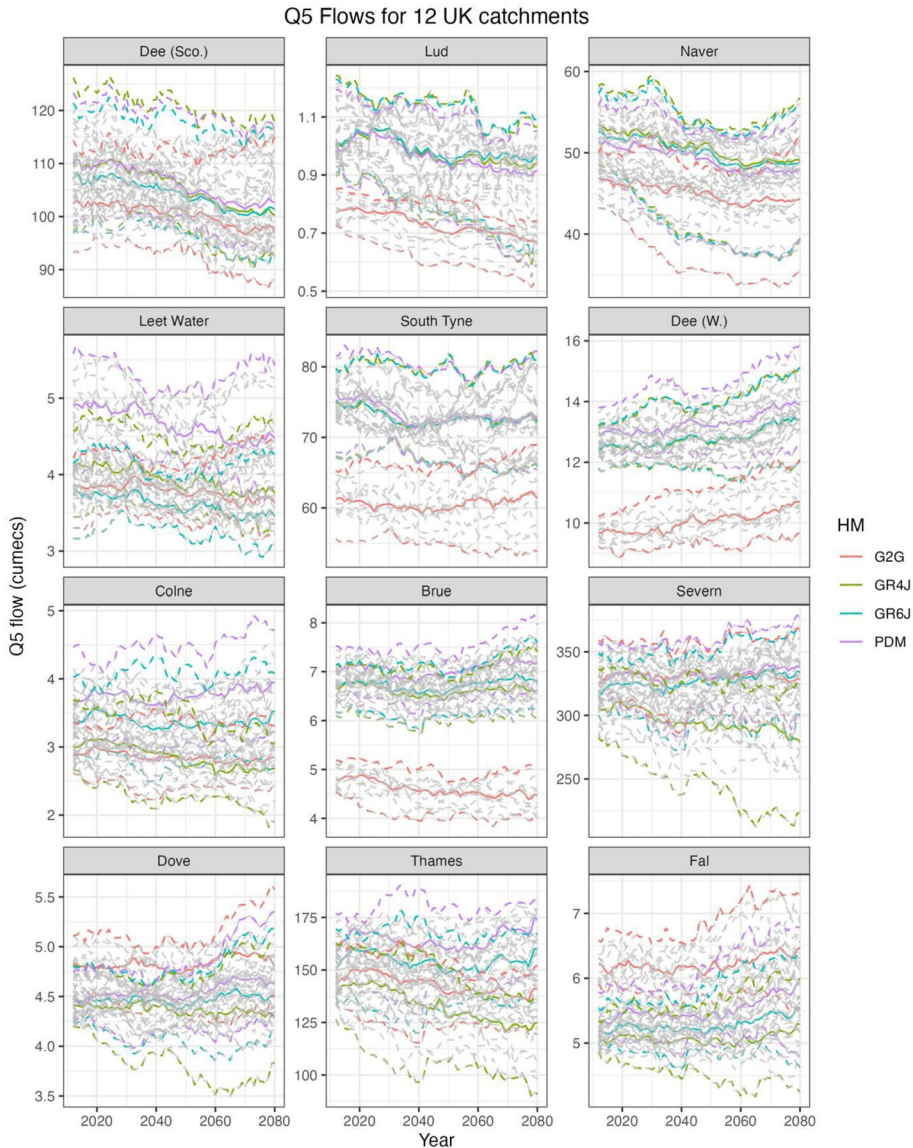


**Fig. 3** Raw Inflow values for the 12 selected catchments **a** Q5 flow and **b** Q95 flows. Coloured dashed lines represent the upper and lower bounds for each HM with grey lines showing the individual projections
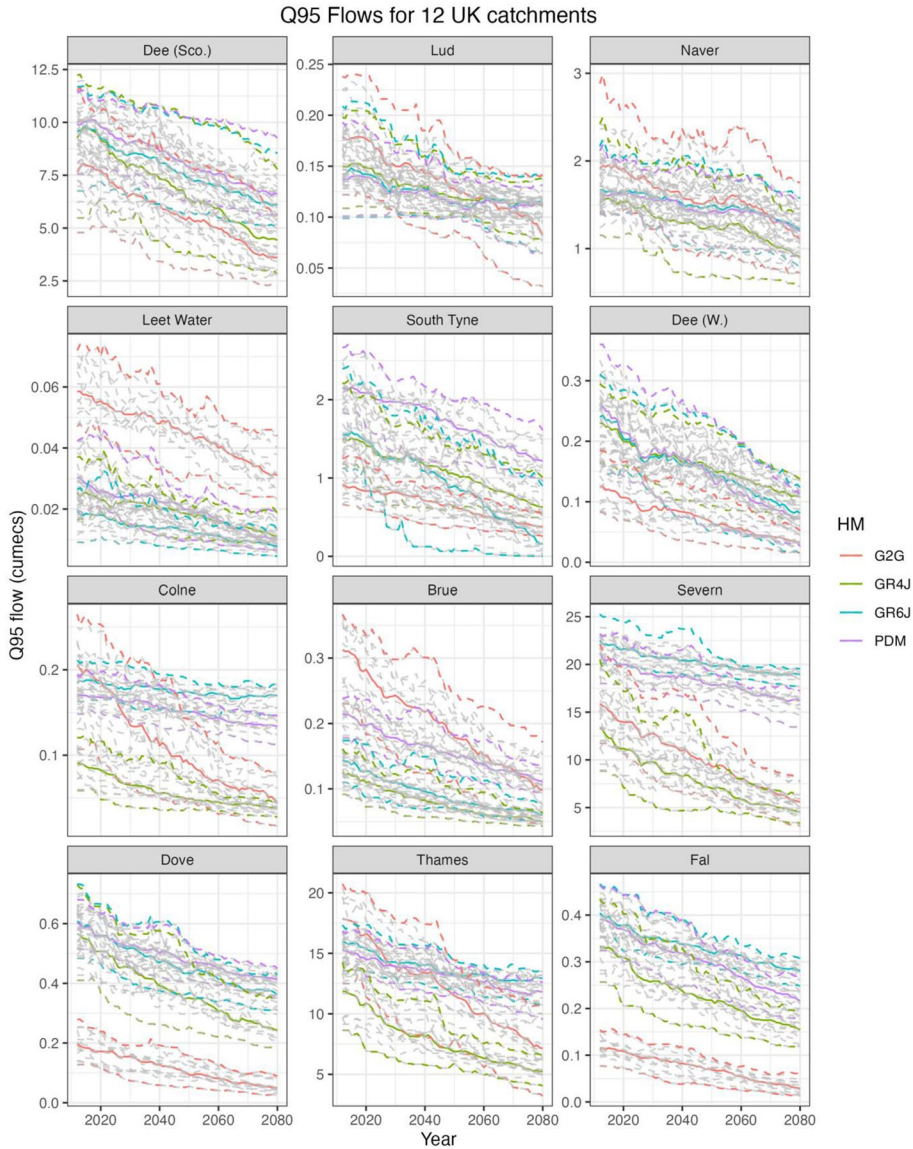
**Fig. 3** (continued)

the noise-free signal (NFS) for the change variable *X*, and the partitioned components of the main sources of uncertainty. Using a sub-set of catchments set out in Section 2.5, the results are then explored in more detail for trends and explanations. Finally the cluster analysis performed highlights whether trends in the results can be attributed to particular physical characteristics.

**Table 1** Subset of 12 catchments used for in-depth analysis of QUALYPSO results

| Name | NRFA ID | Location | Category | Area (km$^2$) |
|------|---------|----------|----------|---------------|
| Dee (Sco.) | 12001 | E Scot | A | 1370 |
| Lud | 29003 | E Eng | A | 55 |
| Naver | 96002 | NW Scot | A | 477 |
| Leet Water | 21023 | SE Scot | B | 113 |
| South Tyne | 23004 | NE Eng | B | 751 |
| Dee (W.) | 67018 | N Wal | B | 54 |
| Colne | 37005 | E Anglia | C | 238 |
| Brue | 52010 | SW Eng | C | 135 |
| Severn | 54057 | Mid. Wal | C | 9895 |
| Dove | 28046 | Central Eng | D | 83 |
| Thames | 39072 | SE Eng | D | 7046 |
| Fal | 48,003 | SW Eng | D | 87 |

## 3.1 Median change signal

Figure 4 shows the median flow metric results for each catchment, and the median change across the ensembles into the 2050s and 2080s. It is clear from the figure that there are both positive changes (i.e., increases) and negative (i.e., decreases) to high flow (Q$_5$) estimates in both the 2050 and 2080 time periods. Similar magnitude changes are observed for both time periods across both time periods with the maximum changes concentrated along the west coast. Changes generally tend to increase marginally between 2050 and 2080 for the majority of catchments. Some exceptions are observed for example the Thames where reductions in high flows are projected after 2050. All median changes range between $-69.1$ and 13.9 with 79.5% of these less than zero (100% of the low flow values and 59% of the high flows). The distributions of change values for each flow quantile are given in Supplementary materials S3.

A general trend of decreasing median flow values is identified for low flow projections. The median estimate for Q95 is set to decrease further from the 2050s to the 2080s. The reduction is observed most acutely in the south of the country, with catchments in the south west having a 40% reduction in the 2050s, and further reducing to $-60\%$ of the 2012 median flows by 2080. All catchments show a decreasing trend in Q95 median values from the 2050s to the 2080s.

## 3.2 Uncertainty components

Figure 5 shows the total uncertainty as calculated using the QUALYPSO package across all study catchments in Great Britain. There is a clear trend of increasing uncertainty between near and far future time periods, as well as between high and low flows. There is greater uncertainty in the far future as a result of diverging results from the modelling chain. This is in line with previous studies conducted such as Evin et al. (2019).

What is interesting however is that there is a trend of greater uncertainty between high and low flows with low flows having increased total uncertainty in comparison. Examining the raw results from the eFLaG dataset, this appears to arise from the calibration of the hydrological models. Three of the four hydrological models (G2G being the exception) are
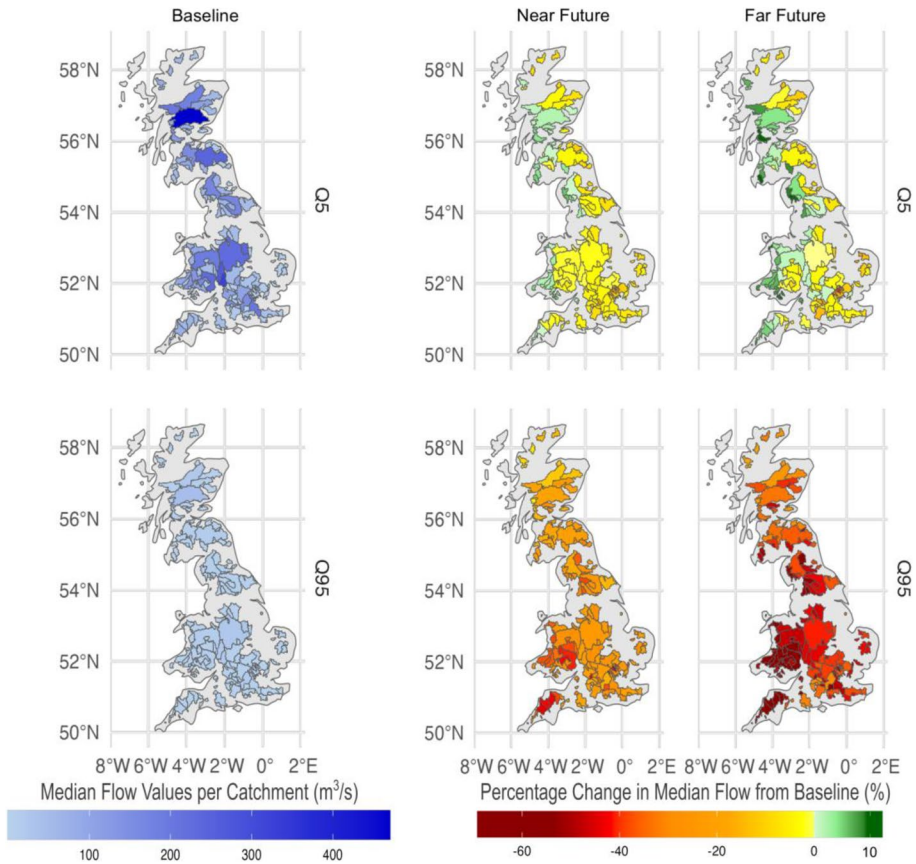
**Fig. 4** Median catchment results across Great Britain. The top panel displays the results for the high flow metric ($Q_5$), whilst the lower panel shows low flow metric ($Q_{95}$). The first column shows the median catchment flows for the baseline period (2012), whilst the middle column shows the percentage change in median flows for the 2050s and the third column shows the change in the 2080s

calibrated to gauged historical flow records which include some level of artificial influence (depending on catchment). Whilst these three models (GR4J, GR6J, PDM) are calibrated to the same historical flow records, despite attempts to ensure as much consistency as possible, there are differences in the calibration procedures for PDM and GR4J/GR6J. Since the impact of artificial influences on the flow regime is more significant (in relative terms as a proportion of flows) at low flows rather than high flows, there are greater differences between the lumped models calibrated to observations and the uncalibrated G2G which simulates naturalised flows. This yields more HM uncertainty at low flows than high flows.

Figure 6 maps the evolution of the untangled dominant uncertainty components for both low and high flow indicators across the time periods. There is a clear pattern for high flows of dominant uncertainty arising from the RCM. This trend intensifies in the far future. A few key catchments do not follow this trend (e.g., 54,057 — Severn) where the uncertainty components are more balanced between HM and RCM; these catchments are generally located in Southern England. For low flow indicators, the opposite is true. It appears that the dominant source of uncertainty across both time periods is the hydrological model,

and this trend intensifies in the far future. Regional climate model uncertainty remains sta-
tionary between flows, the additional uncertainty in low flows is a direct consequence of
increased hydrological model uncertainty.

## 3.3  Catchment analysis

QUALYPSO results have been produced for each trio of catchments defined in Section 2.4.
Comparing the separation of uncertainty identifies the largest source of variability and
what may need to be considered when working with catchments with similar flow relation-
ships to those examined herein.

### 3.3.1  Category A

The first of these groups to be investigated include catchments with similar Q95 flow series
for each of the 4 hydrological models in the eFLaG dataset.

Across each of the catchments, the RCM uncertainty is seen to dominate within the high
flow (Q5) series. Partitioning uncertainties shows that the variability between 12 RCM
ensemble members accounts for over 80% of the total uncertainty in the 2080s (Fig. 7a).
The remaining uncertainty is typically made up of internal variability with only the Scot-
tish Dee producing only moderate HM uncertainty (4%). As such, the selection of hydro-
logical model is seen to have less influence on the modelled flows than the 12 RCM ensem-
ble members. Interestingly, this occurs also for the Lud even when one hydrological model
(usually G2G) produces significantly different high flow series. This may be a consequence
of the similarities seen in the other HM's with QUALYPSO designed to calculate the HM
mean deviations from the ensemble mean.

Uncertainty partitioning for low flows (Q95) show a larger proportion of hydrological
model uncertainty compared to the high flow results. The Lud and Naver catchments are
characterised by equivalent HM uncertainty partitioning ranging from ~15 to ~45% for the
Scottish Dee (Fig. 7a). All three HM values are significantly larger than those for Q5 flows,
highlighting an increased importance of HM selection at lower flows. This importance
increases through time at all three locations with more uncertainty in the far future. The
larger proportion of HM uncertainty occurs even in the case of Lud whereby the modelled
Q95 flow series from the 4 HMs are more similar to one another than the Q5.

### 3.3.2  Category B

The second group of catchments have been identified as having 3 hydrological models with
similar Q95 flow series and one outlier. G2G is the outlier for both the Leet Water and
Welsh Dee catchments with PDM producing different flows for the South Tyne.

Once more, the total uncertainty at high flows (Q5) is dominated by the RCM
uncertainty which remains above 85% from 2050 through to 2080. In contrast, HM
uncertainty is negligible by comparison (< 5%) throughout the twenty-first century.
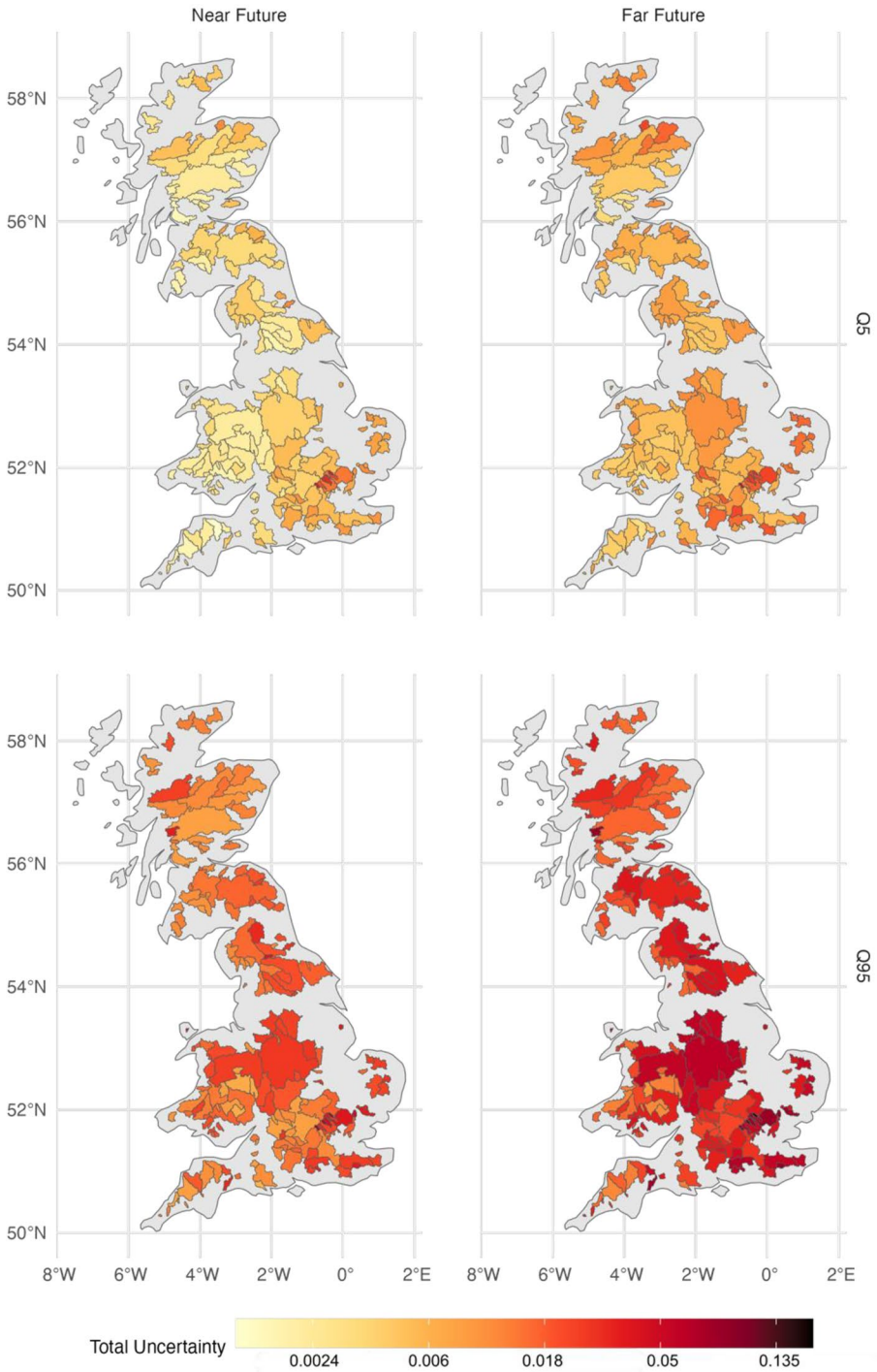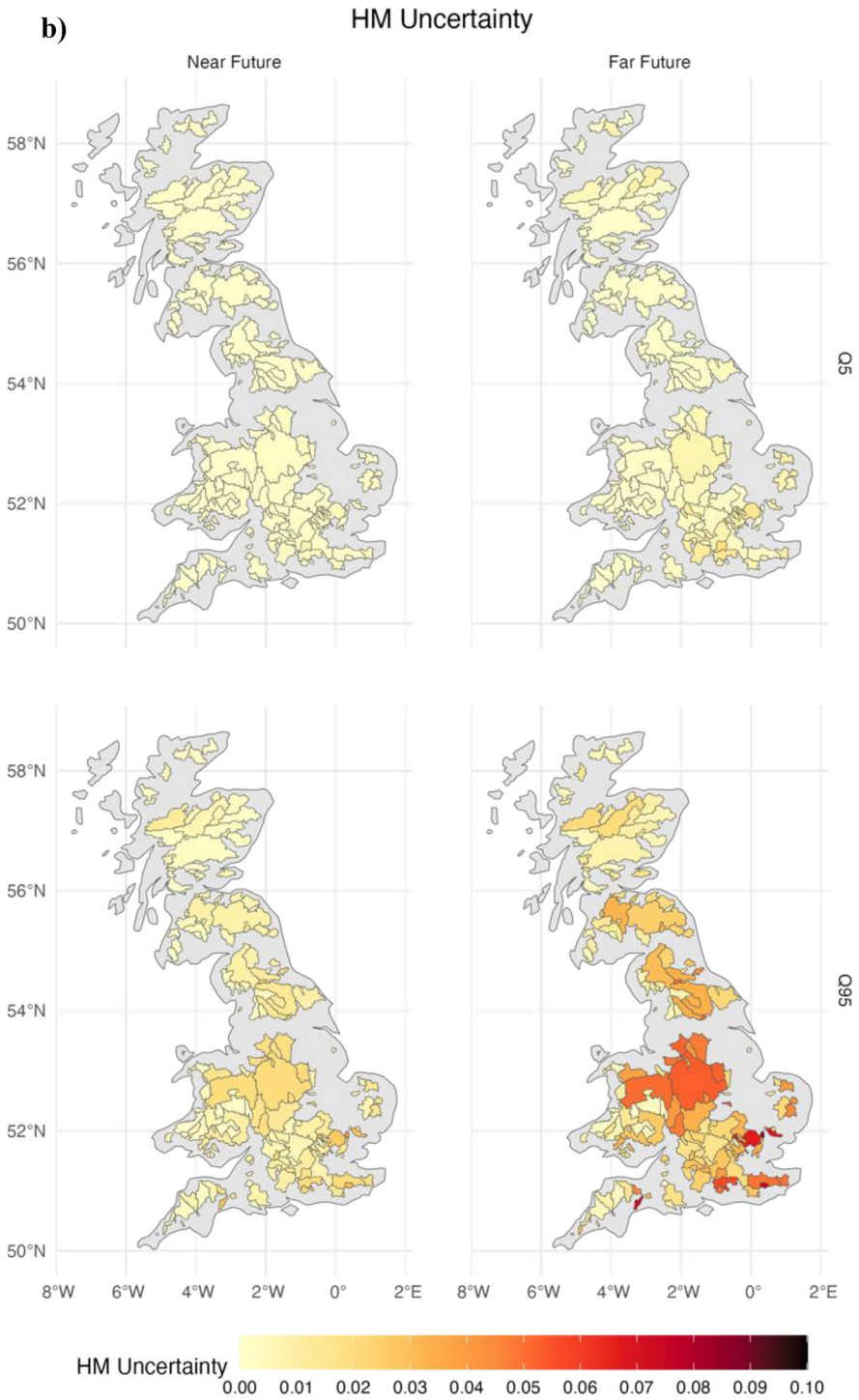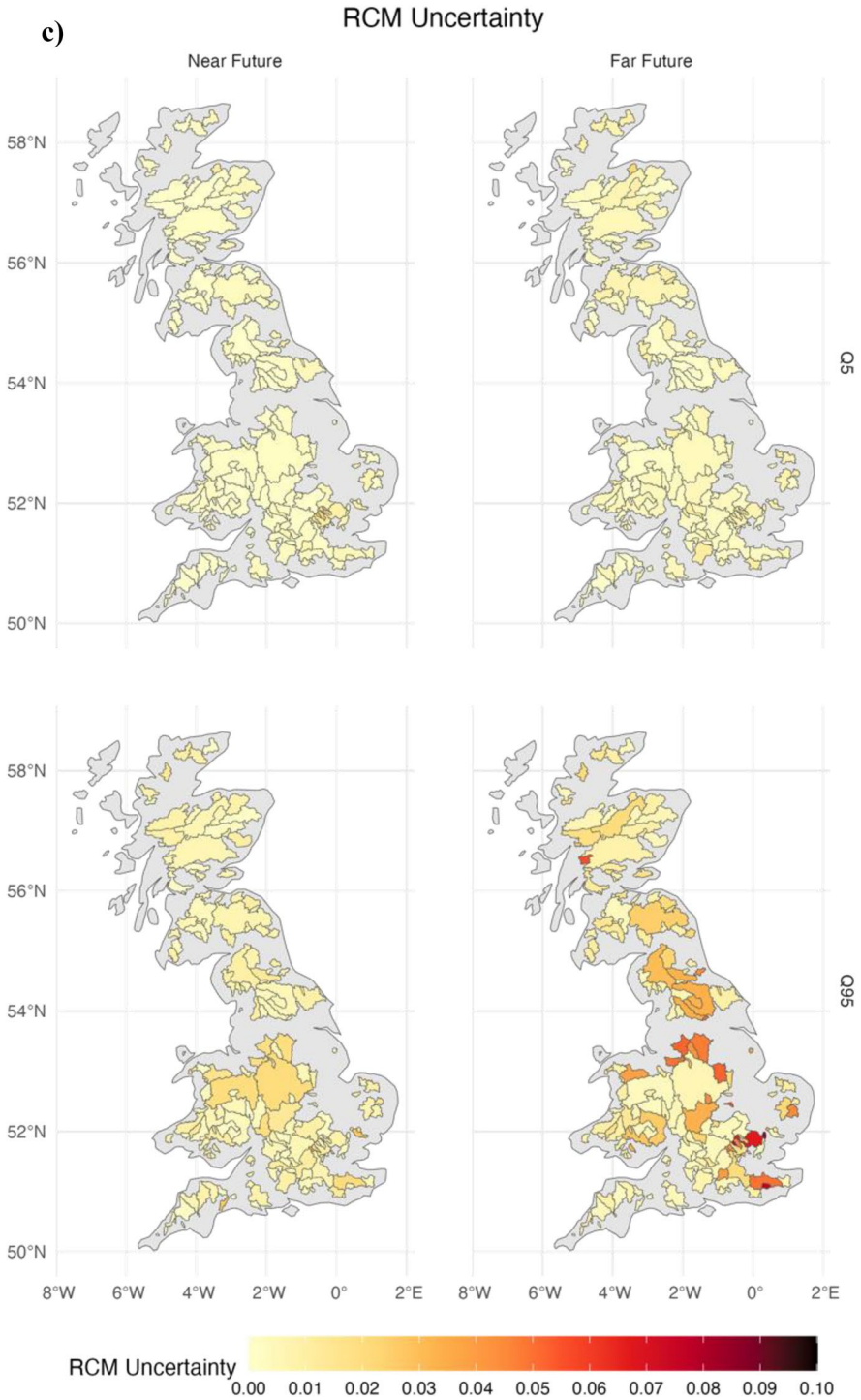
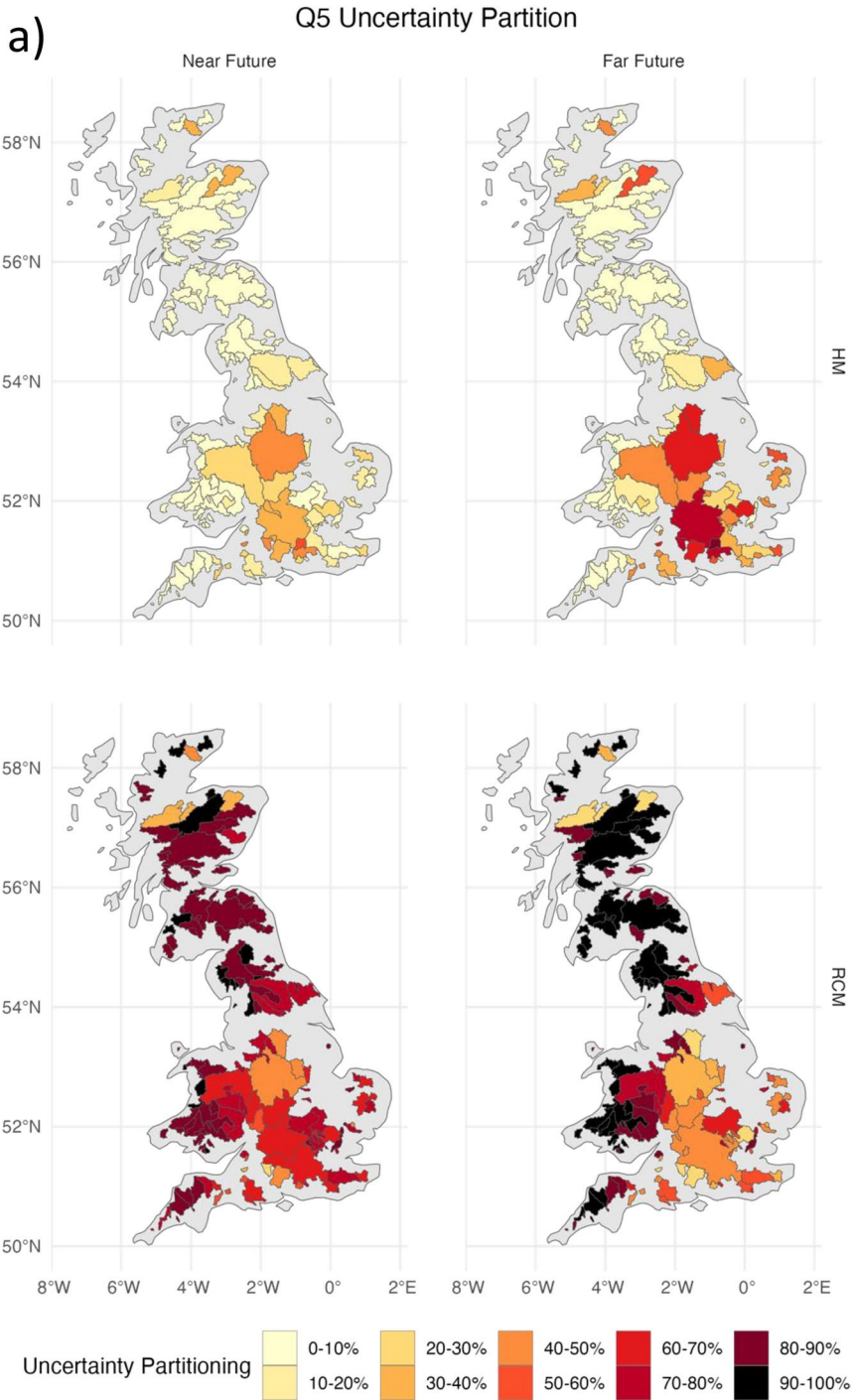**Fig. 5** (continued)

**Fig. 5** (continued)

**Fig. 6** Uncertainty components per catchment as a percentage of total uncertainty: hydrological model (HM) uncertainty versus regional climate model (RCM) uncertainty
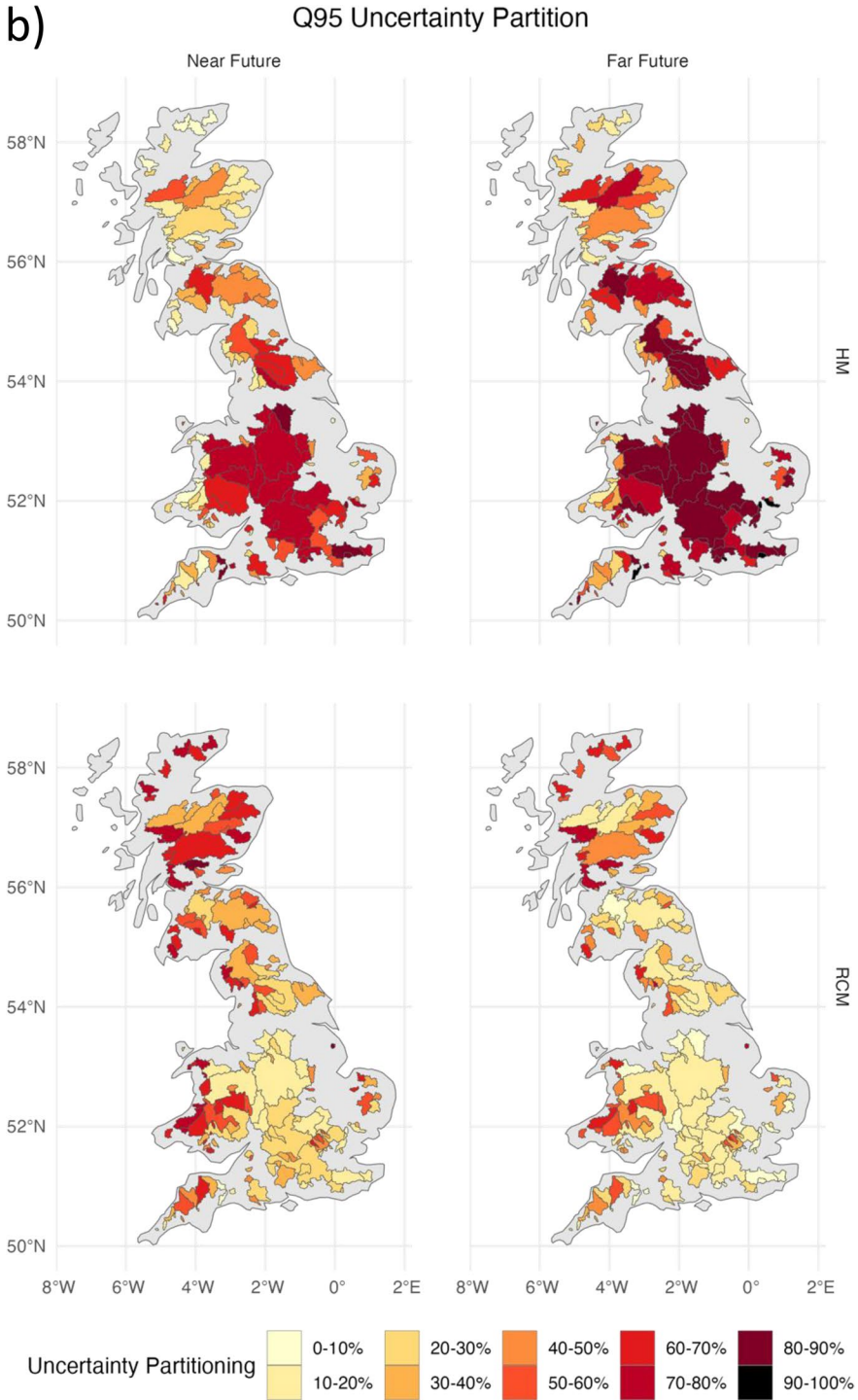
**Fig. 6** (continued)

**Fig. 7 a** Category A: Catchments where the hydrological model results demonstrate similar behaviour ▶ (Scottish Dee, Lud, Naver). **b** Category B: Catchments where three hydrological model results demonstrate similar behaviour and one is an outlier (Leet Water, South Tyne, Welsh Dee). **c** Category C: Catchments where two sets of hydrological model results demonstrate similar behaviour (Colne, Brue, Severn). **d** Category D: Catchments where all four hydrological model results demonstrate different behaviour (Thames, Dove, Fal). Internal variability is coloured red; hydraulic model uncertainty is coloured light blue; regional climate model is coloured dark blue; residual uncertainty is coloured orange

This trend is consistent for all catchments, including where modelled flows from the HMs are significantly different (e.g., Fig. 7b — Welsh Dee). Higher RCM uncertainty indicates greater importance on selection of RCM ensemble members compared to influence of the HM.

HM uncertainty partitions increase through time for low flows (Q95), becoming the major contributor to overall uncertainty for the South Tyne by 2080. RCM uncertainty continues to dominate in the near future but the choice of hydrological model becomes more important when analysing far future low flows. However, HM uncertainty percentages remain smaller than the equivalent for RCM uncertainty for Leet Water and Welsh Dee (Fig. 7b). The relative proportions of uncertainty contributions vary across years and catchments for Q95 flows compared to the more stable Q5 partitions.

### 3.3.3  Category C

The third subset of catchments is those for which modelled low flows in 2080 fall into two distinct pairs of HMs. For the Colne and Severn, far future low flows are similar for G2G and GR4J and for GR6J and PDM (Fig. 7c), whereas for the Brue the G2G and PDM modelled flows are similar (as are those modelled by GR4J and GR6J).

Uncertainty at high flows (Q5) for this third subset of catchments (Fig. 7c) is partitioned far more equally between RCM and HM relative to the previous two subsets (Fig. 7a and b). Whilst RCM uncertainty remains the majority component in the near future (2050) — 50–70% of total uncertainty across the three catchments — the importance of HM uncertainty continues to increase into the far future. By 2080 and consistently across all three catchments, RCM and HM uncertainty are approximately similar.

Once again, there is less consistency in the uncertainty partitioning between catchments for low flows (Q95) in this subset. Unlike for those catchments in preceding subsets, HM uncertainty is the dominant component of total uncertainty for the Colne and Severn, accounting for more than 70% of total uncertainty in both the near (2050) and far future (2080). The relatively narrow envelopes representing the RCM variability for each HM illustrate that differences in flows attributable to RCM ensemble members are much less significant than differences between HMs. The Brue is an exception to this general pattern, with RCM uncertainty larger than HM uncertainty throughout the twenty-first century (though not to the same extent as for low flows in the preceding subsets of catchments; Fig. 7a and b). For the Brue, RCM and HM uncertainty are relatively consistent through time at ~ 50% and ~ 30%, respectively, with residual variability comprising a larger proportion than for other catchments (~ 10%).
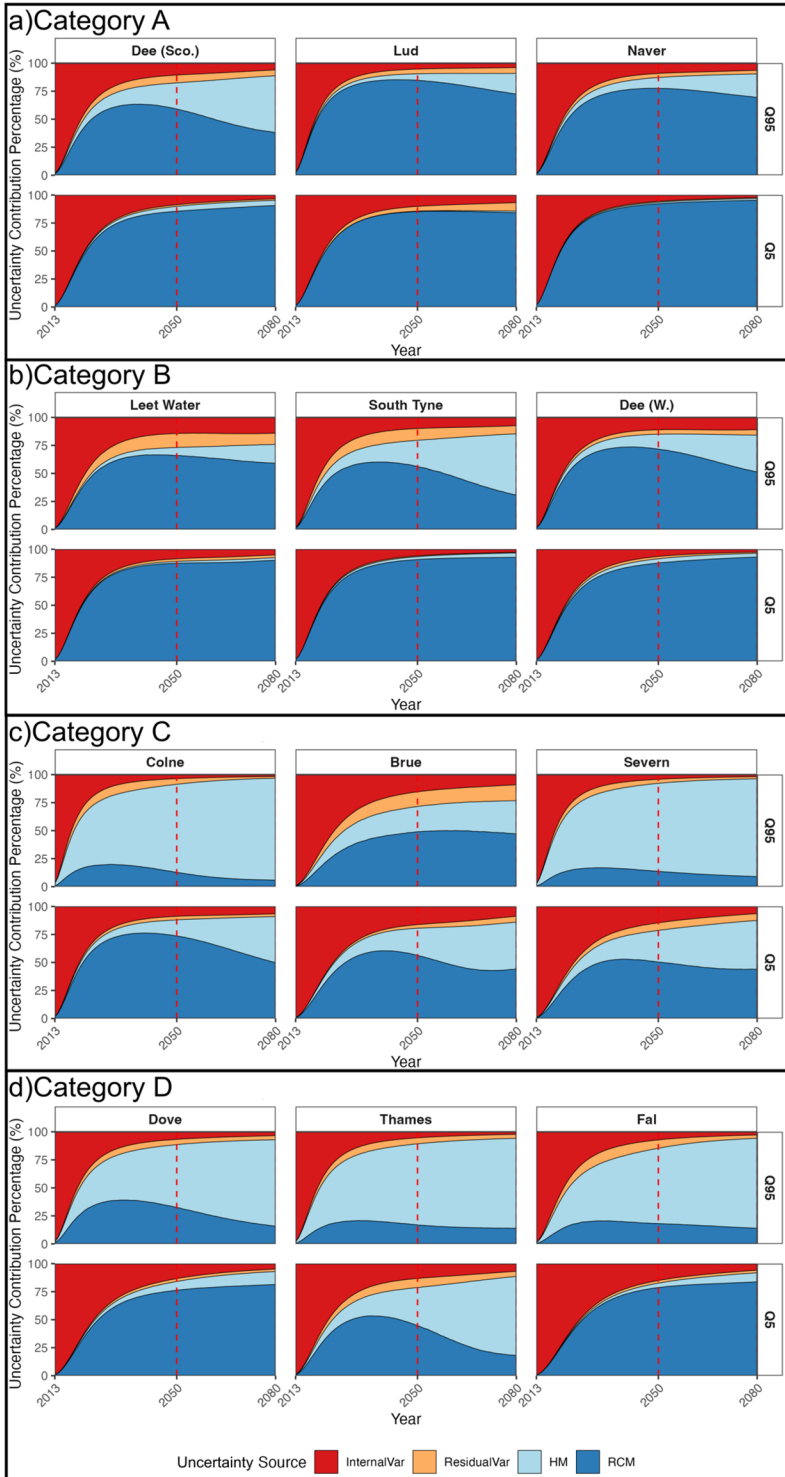
### 3.3.4 Category D

The final subset comprises catchments for which modelled low flows (Q95) in the eFLaG dataset are generally different for each of the four HMs. A range of catchment sizes are represented, from the Fal and increasing in catchment size to the Thames catchment. Whilst the four subsets have been determined based on the relative similarity of modelled low flows in 2080, in the case of this subset, modelled high flows (Q5) also generally differ between each of the four HMs (this is not necessarily true for the other subsets).

The uncertainty partitioning for high flows (Q5) results in a range of different characteristics across the three catchments (Fig. 7d). Total uncertainty for both the Dove and Fal is dominated by RCM uncertainty (~80% in both the near and far future), akin to those catchments in the first two subsets (Fig. 7a and b) although marginally less dominant. For the Dove, the widths of the RCM ensemble envelopes in high flows are much larger than differences between HMs. In contrast, high flow uncertainty for the Thames is dominated by HM variability in the far future (~70%), although the balance between RCM and HM uncertainty in the near future is almost equal.

At low flows (Q95), uncertainty partitioning results across all three catchments are consistent: HM uncertainty is the dominant component, ~70% of the total uncertainty in the far future for all catchments (though having increased from ~55% in the near future for the Dove).

## 4 Discussion

### 4.1 Trends across the UK

A north-west/south-east divide is observed across the UK with greater total uncertainty observed in southern England. River flows in catchments of the south-east have particularly high uncertainty, and there are many contributing factors at play in this region. The south-east has a very complex geology of chalk aquifers that are exploited by abstractions due to high water supply demands (including public water supply and irrigated agriculture) that are not met by the relatively dry conditions of the region, and the dominance of arable agriculture. Water transfer schemes are, though not limited to this region, also employed here increasing the effect of artificial influence. Results are in agreement with previous QE-ANOVA studies on different climate ensembles (Collet et al. 2017; Visser-Quinn et al. 2019; Lane and Kay 2021).

### 4.2 Trends in dominant uncertainty

Regional differences in total and partitioned uncertainty highlight the need to assess ensembles at catchment scales. As part of this assessment, any relationship between catchment characteristics might explain these regional differences; however, no clear correlation was found between NRFA descriptors and total uncertainty (Supplementary Material S4). South-eastern catchments are shown to have a higher degree of uncertainty with more variability in Q5 and Q95 flow estimates. This may be due to G2G producing outlier values in these regions as a result of anthropogenic changes. Similar variations are not observed in the north-west as G2G is more likely to align with lumped models for more natural catchments (fewer abstractions occur in these catchments). RCM uncertainty dominates for high

flows but is not the controlling uncertainty source. The percentage of uncertainty associated with the RCM has been shown to be large for high flows, reaching 90% in some locations. However, examining raw uncertainty values shows a different picture with smaller increases in RCM uncertainty compared to HM for low flows.

On the other hand, low flow uncertainty is dominated by uncertainty from the HM. This is in agreement with several studies in the UK (Chegwidden et al. 2019; De Niel et al. 2019; Vetter et al. 2017) and may be a consequence of the flow calibration to observed data including abstractions which has a larger impact on Q95 values. Furthermore, hydrological model performance at low flows is poorer than high flows, likely a consequence of the calibration procedure and catchment characteristics.

The results presented herein suggest that RCM uncertainty is larger for high flows and HM uncertainty more significant at low flows. However, it is important to note that this is not to claim that intrinsically there is more uncertainty in RCM output during wetter spells triggering higher flows. Absolute values of RCM uncertainty are similar at both high and low flows and in the near and far future (Fig. 6). This highlights that in relative terms the RCM ensemble members are as similar to one another regardless of the hydrometeorological situation. It is important to note these similarities are likely due to RCM uncertainty being assessed across a PPE with different parameters of a single model being modified. Therefore, RCM model structure uncertainty is not considered — rather it provides the uncertainty associated with one RCM. However, the corresponding absolute values of HM uncertainty (Fig. 5b) span a much larger range than for RCM uncertainty. This highlights that, whilst there is no significant difference in the RCM during high and low flow periods, it is HM uncertainty that dominates the relative contributions illustrated in Fig. 6. Correspondingly, the extent of dominance of RCM uncertainty at high flows only occurs because the HM uncertainty is the controlling factor in the relative contributions.

The finding that HM uncertainty is the dominant factor which drives the whole uncertainty partitioning process raises the important question of why there is nothing intrinsically different in the RCM at high and low flows, but there are identifiable differences between the HMs at high and low flows. In the first instance, the use of G2G (simulating naturalised flows) alongside three lumped catchment models (calibrated to gauged observations which include artificial influences) will introduce differences. In addition, amongst the three calibrated models, whilst attempts were made to maximise consistency between models, there were differences in approaches. However, it is worth reflecting on whether there are explicit differences between the ability of models to calibrate to high and low flows. Whilst an early application of the eFLaG dataset was to support enhanced resilience of water resources in the UK, the hope is that the future projections will be useful for a range of applications, in the same manner that the previous iteration of 'future flows' for the UK (Prudhomme et al. 2013). As such, the metrics that were used as objective functions during the calibration process were chosen to simulate a range of flows across the regime (rather than favouring low flows as has been undertaken previously; e.g., Smith et al. 2019). The calibration process attempts to minimise errors between observed and simulated flow, and the most efficient target is large differences in high flows (which yields the greatest reductions in errors). Correspondingly, errors in low flows following calibration are likely to be larger (in relative terms) than at high flows. This is a potentially important explanation for why low flows might naturally be more uncertain than high flows when calibrations are undertaken using general purpose flow metrics, even if ensuring the calibration procedure was identical for all models.

## 4.3 Understanding uncertainty

Application of QE-ANOVA to the eFLaG dataset has provided insight into an important UK dataset. Untangling the uncertainty of the 12-member PPE, 4 HM dataset identifies potential considerations for future studies across 186 UK catchments. Researchers can thus use eFLaG whilst fully aware of regional and modelling uncertainties.

The assessment of uncertainties across UK catchments has highlighted regions with more uncertain flow projections. Large total uncertainties are a consequence of increased variations between the flow projections. Application of eFLaG projections in these regions are likely to produce considerable uncertainties. Therefore, such catchments demand more attention for both high and low flow water management assessments. Additional assessment of catchment characteristics found no clear correlation between NRFA descriptor (BFIHOST) and total uncertainty (Supplementary Material S4).

QE-ANOVA has been applied to a state-of-the-art dataset and can be performed on larger ensembles with additional GCM members. Untangling the eFLaG uncertainties provides more insight into the RCM and HM modelling chains, as well as regional variations. Having assessed uncertainties this dataset can be used for flood and drought estimations with confidence. However, as a PPE the GCM/RCM structural uncertainty has not been assessed. Inclusion of additional RCM models will capture variations in underlying equations and assumptions. Application of QE-ANOVA to further multi-model ensemble climate projection datasets should be considered before their use in water management.

## 4.4 Implications for water resources management in the UK

Separation of climate projection uncertainties has identified dominant sources. However, if costs are not an issue, then we advocate for robust sampling of both RCM and HM uncertainties. This paper has isolated uncertainty contributions relating to RCM and HM choices within the eFLaG dataset. This has provided additional insight into model interactions and importance to specific regions. However, the temptation to focus on one uncertainty should be avoided. Future assessments must continue to consider all cost- and time permitting uncertainties to ensure robust conclusions.

Practitioners can use the QE-ANOVA results presented herein to determine how computational budgets should be assigned under resource constraints. Catchments identified as having large HM uncertainties should prioritise multiple HMs over RCM ensembles to best capture variations in flow estimates. Conversely, RCM dominant catchments could preferentially consider RCM ensembles at the expense of HMs, reducing computational costs. Tailoring ensemble flow series to catchments will ensure computational feasibility targets are met, increasing the likelihood of probabilistic approaches. However, it should be noted that including multiple RCM's and HM's is encouraged to reduce the risk of biased conclusions.

Future water resource management planning currently considers climate change through climate projections run through only a single hydrological model (e.g., United Utilities 2019). The work herein has shown that this is likely to be unsuitable for climate impact studies due to the lack of HM uncertainty quantification, particularly with far future low flows. Significant hydrological model uncertainty is observed across the south-east where population densities are large. A single hydrological model will not account for flow uncertainty in this region and will underestimate flow uncertainty. This has the potential to

impact investment decisions required to provide water to increasing populations in the face of climate change. The same is true for assessing floods in deterministic manner runs the risk of under adaptation thus probabilistic approaches to flood hazard assessment are necessary (Aitken et al. 2022b).

## 5 Conclusions

This paper has quantified uncertainties relating to state-of-the-art eFLaG's climate dataset. Total uncertainties have been assessed before separating into regional climate model and hydrological model contributions across 186 GB river catchments with the aim of identifying national trends in dominant source via a QE-ANOVA approach.

Results have identified large uncertainties in the southeast of England particularly with low flow projections. A national assessment of total uncertainties observed greater variability in projections for southern catchments, highlighting the need for increased care and attention in these domains. High flow river projections produce smaller uncertainties than the equivalent Q95 values and thus may be used with more confidence than their low flow counterparts.

Uncertainty contributions from RCM and HM choices have been analysed for every catchment included in the eFLaG dataset for the single GCM-emissions scenario presented in UKCP18. Regions which are more sensitive to hydrological or regional climate model choices have been highlighted, providing additional information for water resource management in these areas. This allows cost and time limited studies to prioritise the most impactful source of flow variability.

An in-depth analysis of uncertainty partitions and temporal changes has been performed at 12 UK catchments. Four categories (defining similarities in Q95 flows) have been used to differentiate tendencies in hydrological model variability. This analysis has identified trends in dominant uncertainties as a consequence of variability in flows.

Future studies must perform an analysis of variance to ensure robust assessment of uncertainties and prevent unnecessary social, environmental, and economic losses. QE-ANOVA provides a method to understand uncertainties and dependencies which should be applied to new case studies. This will ensure uncertainties are quantified with accurate conclusions for future flood and drought events, increasing national resilience to climatic changes.

# Declarations

**Ethics approval** Not applicable.

**Consent for publication** Not applicable.

**Competing interests** The authors declare no competing interests.

# References

Addor N, Rössler O, Köplin N, Huss M, Weingartner R, Seibert J (2014) Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments, Water Resour. Res 50:7541–7562. https://doi.org/10.1002/2014WR015549

Aitken G, Visser-Quinn A, Beevers L (2022a) EURO-CORDEX: a multi-model ensemble fit for assessing future hydrological change? Front Water 4:804146. https://doi.org/10.3389/frwa.2022.804146

Aitken G, Beevers L, Christie MA (2022) Multi-level Monte Carlo models for flood inundation uncertainty quantification. Water Resour Res 58:e2022WR032599. https://doi.org/10.1029/2022WR032599

Anglian Water: Anglian Water DRAFT Drought Plan. https://www.anglianwater.co.uk/siteassets/household/about-us/draft-drought-plan-2022.pdf. Last accessed 5th Nov 2022.

AXA 2022. 2021 European Summer Floods: A warning about the climate-induced increase in flood risk? https://www.climate.axa/articles/2021-european-summer-floods-a-warning-about-the-climate-induced-increase-in-flood-risk last access September 2022

Beevers L, Collet L, Aitken G, Maravat C (2020) The influence of climate model uncertainty on fluvial flood hazard estimation. Nat Hazards 104:2489–2510. https://doi.org/10.1007/s11069-020-04282-

Beevers L, Popescu I, Pregnolato M, Liu Y, Wright N (2022) Identifying hotspots of hydro-hazards under global change: a worldwide review. Front Water 4:879536. https://doi.org/10.3389/frwa.2022.879536

Bell VA, Kay AL, Jones RG, Moore RJ (2007) Development of a high resolution grid-based river flow model for use with regional climate model output. Hydrol Earth Syst Sci 11(1):532–549

Bell VA, Kay AL, Rudd AC, Davies HN (2018) The MaRIUS-G2G datasets: grid-to-grid model estimates of flow and soil moisture for Great Britain using observed and climate model driving data. Geosci Data J 5(2):63–72

Bosshard T, Carambia M, Goergen K, Kotlarski S, Krahe P, Zappa M, Schär C (2013) Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections, Water Resour. Res 49:1523–1536. https://doi.org/10.1029/2011WR011533

Chegwidden OS, Nijssen B, Rupp DE, Arnold JR, Clark MP, Hamman JJ et al (2019) How do modeling decisions affect the spread among hydrologic climate change projections? Exploring a large ensemble of simulations across a diversity of hydroclimates. Earth's Future 7:623–637. https://doi.org/10.1029/2018EF001047

Collet L, Beevers L, Prudhomme C (2017) Assessing the impact of climate change on extreme flows across Great Britain. Water 9(2):103. https://doi.org/10.3390/w9020103

Collet L, Formetta G, Harrigan S, Prudhomme C, Beevers L (2018) Future hot-spots for hydro-hazards in Great Britain: a probabilistic assessment. Hydrol Earth Syst Sci 22(10):5387–5401

Copernicus (2022). Flooding in Europe | Copernicus https://climate.copernicus.eu/esotc/2021/flooding-july accessed September 2022.

Coron L, Thirel G, Delaigue O, Perrin C, Andréassian V (2017) The suite of lumped GR hydrological models in an R package. Environ Model Softw 94:166–171

de Berrington González A, Cox DR (2007) Interpretation of interaction: a review. Ann Appl Stat 1(2):371–385. https://doi.org/10.1214/07-AOAS124

De Niel J, Van Uytven E, Willems P (2019) Uncertainty analysis of climate change impact on river flow extremes based on a large multi-model ensemble. Water Resour Manage 33:1–15. https://doi.org/10.1007/s11269-019-02370-0

Evin G, Hingray B, Blanchet J, Eckert N, Morin S, Verfaillie D (2019) Partitioning uncertainty components of an incomplete ensemble of climate projections using data augmentation. J Clim 32(8):2423–2440. https://doi.org/10.1175/JCLI-D-18-0606.1

Evin, G. (2022). QUALYPSO: Partitioning uncertainty components of an incomplete ensemble of climate projections. R package, version 2.0. https://CRAN.R-project.org/package=QUALYPSO

Hannaford J, Mackay JD, Ascott M, Bell VA, Chitson T, Cole S, Counsell C, Durant M, Jackson CR, Kay AL, Lane RA, Mansour M, Moore R, Parry S, Rudd AC, Simpson M, Facer-Childs K, Turner S, Wallbank JR, Wells S, Wilcox A (2023) The enhanced future flows and groundwater dataset: development and evaluation of nationally consistent hydrological projections based on UKCP18. Earth Syst Sci Data 15:2391–2415. https://doi.org/10.5194/essd-15-2391-2023

Harrigan S, Prudhomme C, Parry S, Smith K, Tanguy M (2018) Benchmarking ensemble streamflow prediction skill in the UK. Hydrol Earth Syst Sci 22(3). https://doi.org/10.5194/hess-22-2023-2018

Hawkins E, Sutton R (2011) The potential to narrow uncertainty in projections of regional precipitation change. Clim Dyn 37(1–2):407–418. https://doi.org/10.1007/s00382-010-0810-6

Hingray B, Blanchet J, Evin G (2019) Uncertainty component estimates in transient climate projections. Clim Dyn 53:2501–2516. https://doi.org/10.1007/s00382-019-04635-1

Hingray B, Saïd M (2014) Partitioning internal variability and model uncertainty components in a multimember multimodel ensemble of climate projections. J Clim 27(17):6779–6798. https://journals.ametsoc.org/view/journals/clim/27/17/jcli-d-13-00629.1.xml

IPCC (2020) Sixth assessment report. https://www.ipcc.ch/assessment-report/ar6/. Accessed Aug 2022

Jenkins GJ, Murphy JM, Sexton DM, Lowe JA, Jones P, Kilsby CG (2010) UK climate projections briefing report. Met Office Hadley Centre, Exeter

Kay AL, Crooks SM (2014) An investigation of the effect of transient climate change on snowmelt, flood frequency and timing in northern Britain. Int J Climatol 34(12):3368–3381. https://doi.org/10.1002/joc.3913

Kay AL, Davies HN, Bell VA, Jones RG (2009) Comparison of uncertainty sources for climate change impacts: flood frequency in England. Clim Chang 92:41–63. https://doi.org/10.1007/s10584-008-9471-4

Lane RA, Kay AL (2021) Climate change impact on the magnitude and timing of hydrological extremes across Great Britain. Front Water 3:684982. https://doi.org/10.3389/frwa.2021.684982

Lane RA, Coxon G, Freer J, Seibert J, 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

Lowe JA, Bernie D, Bett P, Bricheno L, Brown S, Calvert D, Clark R, Eagle K, Edwards T, Fosser G, Fung F (2018) UKCP18 science overview report. Exeter, UK, Met Office Hadley Centre

Melsen LA, Addor N, Mizukami N, Newman AJ, Torfs PJJF, Clark MP, Uijlenhoet R, Teuling AJ (2018) Mapping (dis)agreement in hydrologic projections. Hydrol Earth Syst Sci 22:1775–1791

Parry S, Mackay JD, Chitson T, Hannaford J, Bell VA, Facer-Childs K, Kay A, Lane R, Moore RJ, Turner S, Wallbank J (2023) Divergent future drought projections in UK river flows and groundwater levels. Hydrol Earth Syst Sci Discuss 1–36. https://doi.org/10.5194/hess-2023-59

Prudhomme C, Davies H (2009) Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: future climate. Clim Change 93:197–222. https://doi.org/10.1007/s10584-008-9461-6

Prudhomme C, Haxton T, Crooks S, Jackson C, Barkwith A, Williamson J, Kelvin J, Mackay J, Wang L, Young A, Watts G (2013) Future Flows Hydrology: an ensemble of daily river flow and monthly groundwater levels for use for climate change impact assessment across Great Britain. Earth Syst Sci Data 5(1):101–107

Robinson EL, Blyth E, Clark DB, Comyn-Platt E, Finch J, Rudd AC (2016) Climate hydrology and ecology research support system potential evapotranspiration dataset for Great Britain (1961–2015) [CHESS-PE]. NERC Environ Inf Data Centre. https://doi.org/10.5285/8baf805d-39ce-4dac-b224-c926ada353b7

Rudd AC, Kay AL, Bell VA (2019) National-scale analysis of future river flow and soil moisture droughts: potential changes in drought characteristics. Clim Change 156(3):323–340

Searle SR (1971) Topics in variance components estimation. Biometrics 27:1–76. https://doi.org/10.2307/2528928

Smith KA, Wilby RL, Broderick C, Prudhomme C, Matthews T, Harrigan S, Murphy C (2018) Navigating cascades of uncertainty — as easy as ABC? Not Quite. J Extreme Events 5(1):1–8

Smith KA, Barker LJ, Tanguy M, Parry S, Harrigan S, Legg TP, Hannaford J (2019) A multi objective ensemble approach to hydrological modelling in the UK: an application to historic drought reconstruction. Hydrol Earth Syst Sci 23(8):3247–3268

Thober S, Kumar R, Wanders N, Marx A, Pan M, Rakovec O, Samaniego L, Sheffield J, Wood EF, Zink M (2018) Multi-model ensemble projections of European river floods and high flows at 1.5, 2, and 3 degrees global warming. Environ Res Lett 13:014003

Toreti A, Bavera D, Acosta Navarro J, Cammalleri C, de Jager A, Di Ciollo C, Hrast Essenfelder A, Maetens W, Magni D, Masante D, Mazzeschi M, Niemeyer S, Spinoni J (2022) Drought in Europe August 2022. Publications Office of the European Union, Luxembourg. https://doi.org/10.2760/264241,JRC130493

United Utilities (2019) Water resources management plan. https://www.unitedutilities.com/corporate/about-us/our-future-plans/water-resources/water-resources-management-plan/. Accessed Aug 2022

Vetter T, Reinhardt J, Flörke M et al (2017) Evaluation of sources of uncertainty in projected hydrological changes under climate change in 12 large-scale river basins. Clim Change 141:419–433. https://doi.org/10.1007/s10584-016-1794-y

Vidal J-P, Hingray B, Magand C, Sauquet E, Ducharne A (2016) Hierarchy of climate and hydrological uncertainties in transient low-flow projections. Hydrol Earth Syst Sci 20:3651–3672. https://doi.org/10.5194/hess-20-3651-2016

Visser-Quinn AG, Beevers L, Collet L, Formetta G, Smith K, Wanders N, Thober S, Pan M, Kumar R (2019) Spatio-temporal analysis of compound hydro-hazard extremes across the UK. Adv Water Resour. https://doi.org/10.1016/j.advwatres.2019.05.019

Watts G, Battarbee RW, Bloomfield JP, Crossman J, Daccache A, Durance I, Elliott JA, Garner G, Hannaford J, Hannah DM, Hess T, Jackson CR, Kay AL, Kernan M, Knox J, Mackay J, Monteith DT, Ormerod SJ, Rance J, Stuart ME, Wade AJ, Wade SD, Weatherhead K, Whitehead PG, Wilby RL (2015) Climate change and water in the UK – past changes and future prospects. Progress in Physical Geography: Earth and Environment 39(1):6–28. https://doi.org/10.1177/0309133314542957

Wilby R, Harris I (2006) A framework for assessing uncertainties in climate change impacts: low-flow scenarios for the River Thames. UK Water Resour Res 42:W02419

Zhang S, Chen J, Gu L (2022) Overall uncertainty of climate change impacts on watershed hydrology in China. Int J Climatol 42(1):507–520. https://doi.org/10.1002/joc.7257