Water temperature projections for England's rivers

Chief Scientist's Group report

January 2025

Version: SC220018/R1

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):

Dr Michael Finney, Dr Matthew Charlton, Dr Judy England, Dr Kieran Khamis, Dr Jermone Masters and Professor David M. Hannah

Keywords:

River water temperature, climate change, projections, modelling, England

Research contractor:

University of Birmingham, Edgbaston, Birmingham, B15 2TT. +44 (0)1214143344

Environment Agency’s Project Manager:

Dr Judy England

Project number:

SC220018/R1

Citation:
Environment Agency (2025) Water temperature projections for England’s Rivers. Environment Agency, Bristol.

Research at the Environment Agency

Scientific research and analysis underpin 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

Acknowledgements

The authors would like to thank the following Environment Agency staff for providing data and advice throughout this project: Mr Simon Barton, Dr Mike Dunbar, Mr Andy Goodwin, Dr Nick Hayes, Dr Alex Lumsdon, Mr Grant McMellin and Dr John Robotham. Special thanks to Dr Kate Heppel from the Chilterns Chalk Stream project and Dr Jon Bass from the Wessex Rivers Trust for providing their river temperature data set.

Executive Summary

How climate change will affect aquatic ecosystems remains uncertain. The strong influence of water temperature on water quality and river ecosystems suggest changes in river water temperature are likely to be of critical importance. Understanding where, when and by how much rivers may warm will inform management activities and target adaption actions to help reduce the adverse effects of potential changes. While projections of river flows under climate change have been developed for England there are currently no future national projections of river water temperature. This limits our understanding of future risks and management choices that could improve resilience to the impacts of climate change.

Environment Agency and third-party water temperature records along with environment variables (air temperature, land cover and river network properties) were used to create models to develop water temperature projections. Most of the data used to generate the models were single observations collected during the working day, night- time temperatures are less well represented. This means that the projections developed were for monthly mean daytime river water temperatures. Sufficient water temperature data were available for 641 sites (≥3 records per month; 60+ monthly means) to generate the model which was then validated for 3441 sites across the English river network (4082 sites in total).

The national model was then used to make monthly river water temperature projections to 2080 using UK climate change projections (UKCP18). Based on a ‘high emissions scenario’, ecologically significant increases in water temperature were projected with summer maxima rising by 0.6 °C per decade above baseline levels (1981 to 2005 inclusive) to 2080. An important temperature threshold of 12 °C for salmonid egg survival during the winter period will likely be exceeded at over 70% of sites by 2080. Adult brown trout will be under threat from high summer temperatures with almost all sites experiencing temperatures that exceed their upper growth/feeding temperature range of 19.5 °C by 2080.

Given England’s rivers cover a wide range of environmental conditions and exhibit significant variability in factors that influence river water temperature, models were developed for subsets of the 4082 sites representing distinct river typologies (based on geology and land use). These models showed no improvement over the national model when estimating future monthly mean temperature. However, differences were seen in the projected temperature responses, with arable/grassland sites on permeable rock experiencing the highest decadal changes in maximum temperatures (0.66 °C per decade) and grass upland sites experiencing the highest summer maxima by 2080 (>28 °C).

This report presents a national perspective on where, when and by how much rivers may warm. The monthly average daytime projections may not reflect local temperature variations and will miss short-term heatwaves, which may have important ecological effects. Future work to explore the value of less spatially extensive but more frequently sampled (e.g. daily) data will be useful. However, these national projections are already proving useful as indicators of how temperature sensitive species may be affected and can be used to explore implications for water quality management.

Introduction

Overview

Climate projections for England suggest there will be an increase of warmer, wetter winters and hotter, drier summers along with more frequent and intense extreme weather events (UKCP18: Met Office, 2019). Understanding the consequences of these changes on water quality and aquatic biota is important to inform the targeting of management actions and adaptation measures to help maintain the integrity of river ecosystems and the services they provide.

Flow (discharge) and water temperature (*Tw*) are considered master variables that influence river ecosystem structure and function (Poff et al., 1997, Woodward et al., 2010). There is a broad consensus regarding historical and likely future changes in climate, and emerging information on how river flows are expected to change (UKCEH, 2023). There is much less clarity about how and where river *Tw* may change. This is particularly problematic as assessments have highlighted the sensitivity of future eutrophication risk to increasing temperatures (Environment Agency, 2019). Hence, better understanding potential changes in thermal regimes will help identify priority areas for management action (Knouft et al., 2021).

River water temperature is controlled by a complex interaction of hydrological, climatological and landscape characteristics, and previous models for predicting water temperature have utilised variables from all three of these categories (e.g., Jackson et al., 2018). Hydrological controls include the relative proportions of surface and groundwater. Groundwater has a stable temperature profile which can modulate the temperature fluctuations observed in surface waters more influenced by atmospheric and climatological controls (Acornley, 1999). Landscape characteristics include riparian trees which create shade and reduce thermal maxima but the thermal benefit of riparian shade depends on physical location within the river network and on prevailing climate conditions (Garner et al., 2017, Wilby and Johnson, 2020).

Quantifying the amount and timing of future warming in rivers will help understand where water quality and ecosystems may be affected and provide more robust evidence of where to target measures to adapt to these changes. To understand how best to develop future projections of river *Tw*, Environment Agency (2021) reviewed potential modelling approaches and produced a robust framework for doing so.

Using chalk streams as a test case, the modelling approach recommended in the scoping project was tested, applied, and refined (Environment Agency, 2022). This study extends that analysis across the whole of England to generate river water temperature projections.

Project aim and outline

The aim of this project was to develop water temperature (*Tw*) projections for England's rivers.

To develop the projections *Tw* readings were collated from various Environment Agency sources (including Water Quality Archive (WQA), Surface Water Temperature Archive (SWTA), continuous monitoring datasets from sonde sampling and gauging station monitoring, field studies from partner organisations) and combined (Stage 1; Figure 1), quality assured (Stage 2), and used to generate monthly mean *Tw* values (Stage 3a). These monthly *Tw* means were combined with corresponding air temperature values (*Ta*) and associated catchment data (Stage 3b/c) and assessed for the potential modelling options they could support (Stage 4a). A global model and landscape typology-specific models were then developed, tested, and compared (Stage 4b/c) before projections of *Tw* under climate change conditions were made (Stage 5).

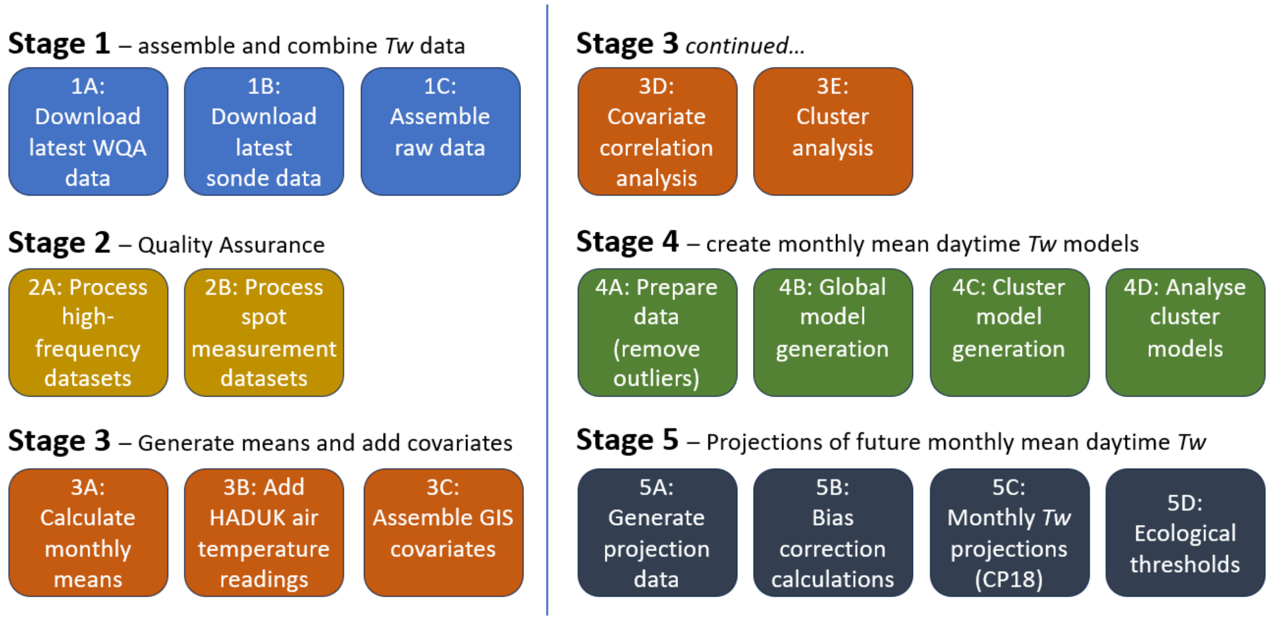


Figure 1: Overview of the steps/stages followed to generate models and predictions of water temperature for English rivers from available data.

Methods

An overview of the data sets used, and the modelling approaches employed are presented here. Further detail is provided in the appendix.

Data for modelling and projections

Observed water temperature readings

A river water temperature (*Tw*) dataset was assembled from various Environment Agency sources: the Surface Water Temperature Archive up to 2007 (Orr et al., 2015), the Water Quality Data archive (WIMS; 2000 to present), as a by-product from flow gauging stations, continuous temperature series from water quality sonde instrumentation, and temperature series from third-party organisations. Before modelling, data were screened using a sequential quality assurance (QA) process and removed if measurements: (1) were greater than ± 3 standard deviation of the mean for each site (324,981 records), (2) exhibited long sequences of unchanging values (1,326,700 records), (3) displayed extreme daily temperature shifts (>10 °C; 433,034 records) or (4) deviated significantly from a nine-value moving average at each site (239,088 records). This resulted in a dataset containing 53,731,931 unique water temperature readings covering the period between November 1952 and April 2023 (median: December 2009).

Monthly mean (n ≥ 3) water temperatures were calculated for each site. Comparisons between monthly means generated from spot samples (manually collected individual readings usually taken during the working day) and daytime monthly means from co-located high frequency sites (automatically collected readings on a 15 minute or hourly basis; n = 137) indicated good correlation ($R^2 = 0.793$), providing confidence in the representation of these monthly mean values. Site locations falling outside the national boundary of England were discarded. Sites with six or more monthly mean values were taken forward as part of the temperature modelling dataset, with those having five or more years' worth of values (60 or more monthly mean values) participating in the development of the water temperature models (734 potential sites). Sites with between 6 and 59 monthly mean values comprised a model validation dataset (3,941 potential sites). Site, landscape, and river channel characteristics for each sampling location in the temperature modelling dataset were extracted from GIS sources and combined with metadata from the various data sources to verify sampling locations. All potential modelling site locations were checked against Ordnance Survey digitised maps to verify water body alignment. Only sites verified as aligning to inland rivers were taken onto the modelling phase; those aligning to lakes, canals, or transitional waters were discarded (649 model development sites and 3505 model validation sites remaining; Figure 2). It should be noted that there is a clear sampling bias in much of the river water temperature dataset with spot sampling measurements almost exclusively collected during working hours (median sample time = 11:00, main range = 06:00 – 18:00, Monday - Friday), with the exception of sites with automated high frequency logging (4.9% of sites, 9.4% of monthly mean values). Hence, the mean values should be considered as estimates of monthly mean daytime water temperature (*MdTw*).

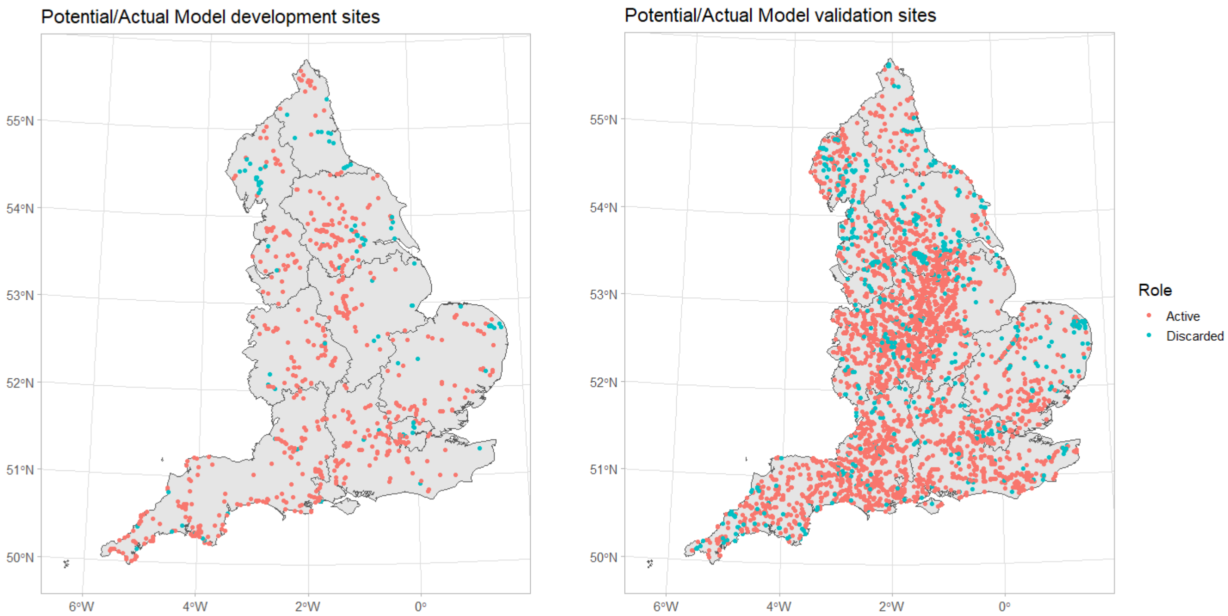


Figure 2: River *Tw* dataset sample sites in England available for model development (left) and validation (right). Sites marked in blue were discarded as they could not be verified as aligning with inland river sites after map analysis (some misaligned, others transitional waters, lakes or canals).

Datasets used for model development

Historical air temperature data were obtained from 1 km gridded datasets from the Met Office (HadUK-Grid products; Met Office 2023). Previous studies (Environment Agency, 2022) have compared the use of monthly mean (temperature at surface; *tas*) and monthly mean of daily maximum (*tasmax*) air temperatures in water temperature modelling and found no significant difference between the two. The monthly mean air temperature at surface (*tas*) was extracted for each sample month and site for use in this study.

To ensure the developed model was practical and suitable for use in a management context, covariates that could be derived solely from available GIS sources were used. Indicators of upstream land-use, geology and sample site characteristics were selected, along with potential proxies for hydrological regime (e.g., river gradient, Strahler stream order). An initial screen of 40 covariates was undertaken, with 20 taken forward to the modelling stage (Table 1; see Appendix for further details).

Landscape characteristics considered included land cover information at 25m resolution obtained from Centre for Ecology and Hydrology (CEH; Land Cover Map 2015 [LCM2015]), topographical information derived from a Digital Elevation Model (Nextmap 50m DEM hydromodel) and LIDAR data, and watercourse characteristics from the Environment Agency’s Detailed River Network GIS layer (Coley et al., 2018).

Table 1. List of covariates selected for water temperature model development. The acronym is displayed along with description and data source.

Covariate acronym	Description	Data source
Altitude_Map	Altitude at site (m)	Environment Agency LIDAR Composite 2022 1m DSM
Strahler	The Strahler river order at site (indicator of river size)	Environment Agency's Detailed River Network GIS layer
Shreve	The Shreve river order at site (indicator of network complexity)	Environment Agency's Detailed River Network GIS layer
Dist2mth	Distance along the main river channel to the river mouth (m)	Environment Agency's Detailed River Network GIS layer
DEM_Aspect	Aspect of river section at site (0-359°)	50m Digital Terrain Model (DTM)
DEM_slope	Gradient of river section at site (m/km)	50m Digital Terrain Model (DTM)
Alkalinity	Acid Buffering capacity at site (indication of upstream calcareous geology) (mg/l)	Environment Agency's Detailed River Network GIS layer
AVE_CATCH_Altitude	Average altitude of upstream catchment (m)	50m Digital Terrain Model (DTM)
AVE_CATCH_Aspect	Average aspect of upstream catchment (0-359°)	50m Digital Terrain Model (DTM)
AVE_CATCH_Slope	Average gradient of upstream catchment (m/km)	50m Digital Terrain Model (DTM)
Geo_Siliceous_PC	Upstream siliceous hard geology (%)	British Geological Survey
Geo_Calcareous_PC	Upstream calcareous hard geology (%)	British Geological Survey
Geo_Chalk_PC	Upstream chalk hard geology (also calcareous; %)	British Geological Survey
LCM_Woodland_PC	Upstream woodland, includes coniferous and broadleaf (%)	CEH; Land Cover Map 2015
LCM_Arable_PC	Upstream arable farmland (%)	CEH; Land Cover Map 2015
LCM_Grassland_PC	Upstream grassland (improved and undisturbed/natural; %)	CEH; Land Cover Map 2015
LCM_Water_wetland_PC	Upstream area containing water/wetland, e.g., pools, reservoirs, etc. (%)	CEH; Land Cover Map 2015
LCM_Built.up_PC	Upstream land dedicated to urban and suburban development (%)	CEH; Land Cover Map 2015
Easting / Northing	Spatial coordinates	From water sample data

Future climate projections

The Met Office UKCP18 datasets provide probabilistic projections of environmental variables (air temperature, rainfall, etc.) based on a range of potential future climate outcomes. A range of future emissions scenarios are considered, and each dataset comprises an ensemble of 12 potential outcomes representing the uncertainty introduced by perturbation of model parameters. The UK regional model projections provide access to spatially coherent 'raw' climate projection data at the highest resolution (12km grid) for the years 1981 to 2080 and were used in this study. These projections are based on the high emissions 'RCP8.5' scenario and are considered a basis for precautionary planning for climate change impacts (Met Office, 2019).

Projected monthly mean air temperature values for the years 1981 to 2080 and for each of the 12 potential outcomes mentioned above were extracted from the 12km RCP8.5 regional UKCP18 gridded dataset for each *Tw* sampling site. Values falling in the date range January 1981 to December 2005 (25-year reference period) were compared to observed air temperature readings for the same period (from HadUK gridded datasets) and differences summarised into a set of monthly bias corrections which were then applied across each Ta projection dataset (Lenderink et al., 2007). This bias correction aligns modelled datasets with a known set of values and improves reliability of subsequent modelling outputs.

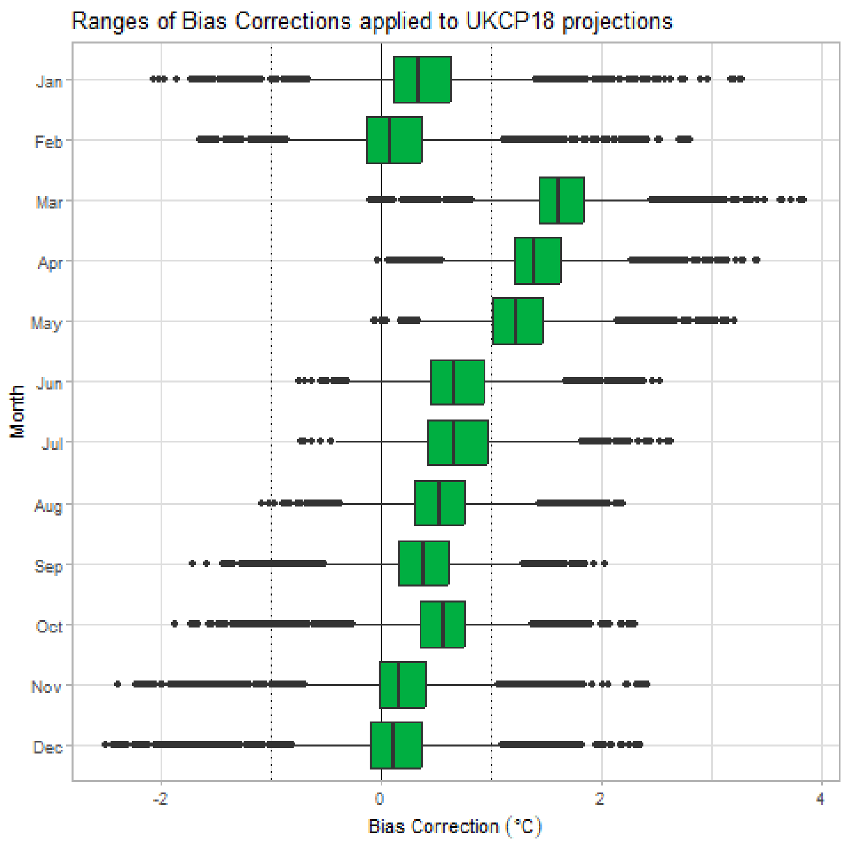


Figure 3: Ranges of bias corrections applied to the UKCP18 monthly mean air temperature projections prior to calculation of water temperature using developed models.

These bias correction adjustments are summarised in Figure 3 and, while the majority are small (< 1°C), an underestimate of springtime air temperatures in the UKCP18 data require a greater increase to align with observed values.

Model development and generation of projections

Selection of water temperature model

The river water temperature modelling process was guided by the previously established framework developed by the Environment Agency (2021). While the availability of historical monthly mean water temperature data placed this study outside the recommended bounds (daily to weekly model output), the decision-making process followed a similar series of steps (flowcharts reproduced here in Appendix; Figures A3 and A4). A mixed effect, regression-based approach was selected as most appropriate for multi-site, repeated measurement data. This option also maintained visibility of any influential covariates allowing an assessment of their physical plausibility to take place.

Model development and validation

The strength of the relationship between the observed water temperature values and the mean air temperature values was assessed. The relationship was statistically significant ($r^2 = 0.86$; Figure 4) with a gradient of 0.87. Seasonal variability in the air-water temperature slope was also identified (Figure 5) and retained as a required element in model development (Mohseni and Stefan, 1999; Webb et al., 2008).

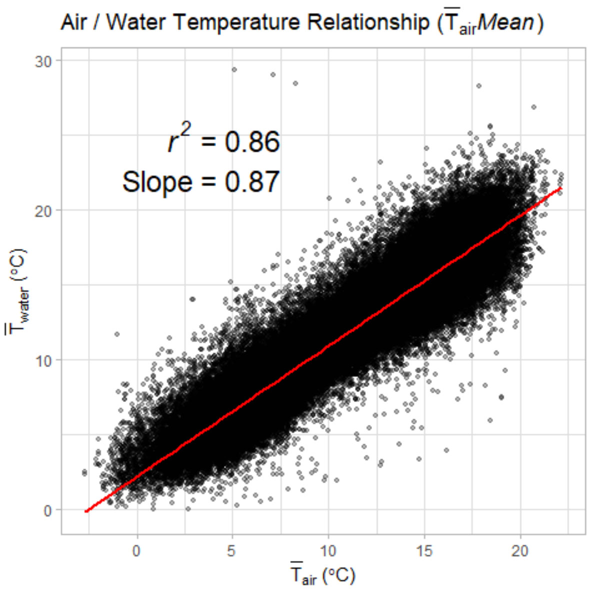


Figure 4: Monthly mean air temperature / water temperature relationship across English *Tw* sampling sites.

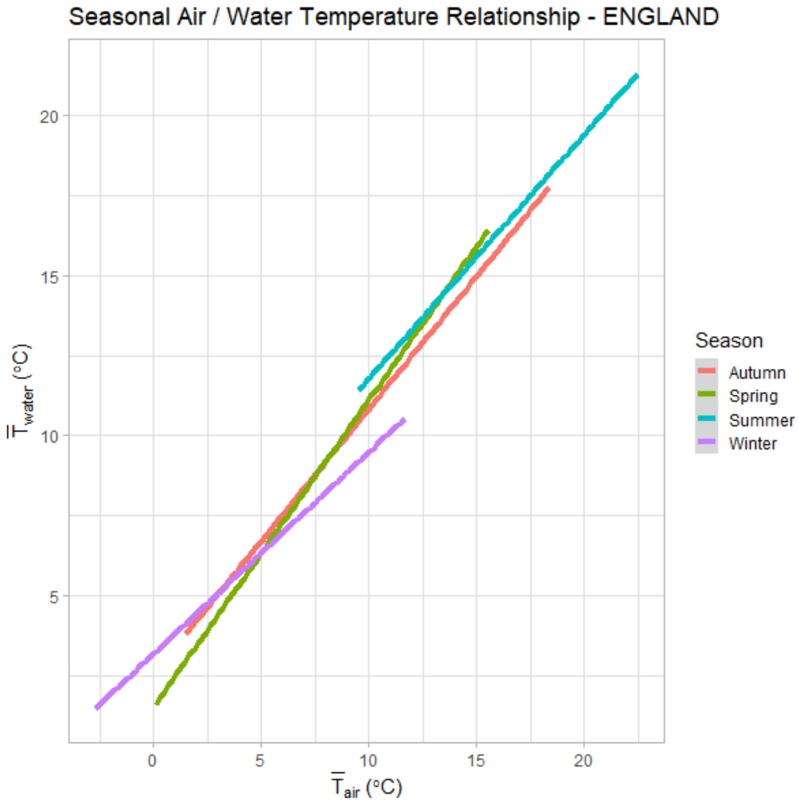


Figure 5: Seasonal variation in monthly mean air temperature / water temperature relationship across English *Tw* sampling sites. Lines fitted using ordinary least squares regression.

Collinearity between predictor variables, the tendency for one variable to change in a very similar way to another, can lead to the development of overly complex models with compounded errors that limit predictive capability (Dormann et al., 2013). Landscape and river channel covariates that displayed strong correlations with each other (Pearson correlation coefficient; $r > 0.7$) were rationalised (Appendix; Figure A5). The covariates taken forward to the modelling stage were:

- Upstream Land cover – percentage upstream arable, urban development, woodland, grassland, and water (LCM_Arable | Built.up | Woodland | Grassland | Water_wetland_PC)
- Upstream hard geology (Geo_Calcareous | Siliceous | Chalk_PC, alkalinity)
- Average catchment Altitude, Aspect and Slope (AVE_CATCH_Altitude | Aspect | Slope)
- Channel characteristics – Altitude at site, Strahler river order, Shreve river order, distance to river mouth, Aspect of site, Slope at site (Altitude_Map, strahler, shreve, dist2mth, DEM_aspect, DEM_slope)

All models were developed using R (R core team, 2023) using functions from the base, lme4 and MuMin packages. A linear mixed-effects model for English river water temperature was developed based on 641 sites and 73,222 observations (note the lower number of sites due to missing covariate data). Based on exploratory data analysis the following features formed the basis of model development:

- The model must include the seasonal air / water temperature relationship
- All selected covariates would initially be included to create a ‘full’ model which would then be simplified to contain only the most influential factors
- Spatial components would be included to represent the west-east and north-south hydro-climatic gradients (Easting and Northing values used)
- Variation between sample sites, catchments and to allow the water/air temperature relationship to vary between sites would be permitted, with the variation following a ‘normal’ distribution (= ‘random effects’ elements)

The ‘full’ model which included all the available covariates selected above was first fitted to available data (‘lmer’ command, lme4 R package) and took the form using the R syntax:

```
MdTw = Tair: Season + Altitude_Map + strahler + shreve+ dist2mth +
DEM_aspect + DEM_slope + alkalinity + AVE_CATCH_Altitude +
AVE_CATCH_Aspect + AVE_CATCH_slope + Geo_Siliceous_PC +
Geo_Calcareous_PC + Geo_Chalk_PC + LCM_Woodland_PC + LCM_Arable_PC
+ LCM_Grassland_PC + LCM_Water_wetland + LCM_Built.up_PC + Easting
+ Northing + (Random effects)
```

‘Random effects’ were applied sequentially as follows (1) Absent, (2) the intercept could vary between sites `(1 | Sites)`, (3) the intercept could vary between catchments `(1 | GIS_catchment_name)`, or (4) the slope and intercept of the *Ta* / *Tw* relationship could vary between sites `(Tair | Sites)`. The different ‘random effects’ outcomes were compared using Bayesian Information Criterion (BIC) and the best performing ‘full’ model reduced to its simplest, statistically significant form to give the following optimised ‘global’ model for monthly mean water temperature:

```
MdTw = Tair: Season + Altitude_Map + dist2mth + Geo_Calcareous_PC +
LCM_Built.up_PC + Northing + (Tair | Site)
```

Ideally, projected *Tw* values generated by a model would be identical to observed *Tw* values in a training dataset so that all points would fall on the $x = y$ identity line when plotting observed against projected *Tw* values. Such an outcome would have an R^2 value of 1 and a Root Mean Squared error (RMSE) of 0. The optimised ‘global’ model the developed here generated good estimates of *MdTw* in the training data ($R^2 = 0.94$, RMSE = 1.058°C; Table 2 and Figure 6-right panel) and performed well with the validation dataset ($R^2 = 0.85$, RMSE = 1.627°C; Figure 7- right panel). These R^2 and RMSE values compared well with those generated by the more complex ‘full’ model (left panels of Figures 6 and 7), indicating that all significant factors had been retained in the optimised ‘global’ model and that it had strong predictive performance for the generation of *MdTw* estimates for inland river sites from limited environmental and air temperature information.

Table 2: Structure and summary statistics of the optimised 'global' model for English rivers monthly mean daytime water temperatures. σ^2 = random effect variance, τ_{00} = random intercept variance, τ_{11} = random slope variance, $\rho01$ = random slope-intercept correlation, ICC = intraclass correlation coefficient, Marginal R^2 = correlation coefficient for fixed effects, Conditional R^2 = correlation coefficient for fixed and random effects

	Mean Value		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	3.45	3.23 – 3.66	**<0.001**
Altitude Map	-0.00	-0.01 – -0.00	**<0.001**
dist2mth	0.01	0.00 – 0.01	**0.001**
Geo Calcareous PC	-0.00	-0.00 – -0.00	**<0.001**
LCM Built up PC	0.02	0.01 – 0.02	**<0.001**
Northing scaled	-0.01	-0.02 – -0.01	**<0.001**
Season [Aut] × Ta mean	0.81	0.80 – 0.82	**<0.001**
Season [Spr] × Ta mean	0.83	0.82 – 0.85	**<0.001**
Season [Sum] × Ta mean	0.85	0.83 – 0.86	**<0.001**
Season [Win] × Ta mean	0.70	0.68 – 0.71	**<0.001**
Random Effects			
σ^2	1.14		
τ_{00} Site_ID	2.19		
τ_{11} Site_ID.Ta_mean	0.02		
ρ_{01} Site_ID	-0.89		
ICC	0.66		
N Site_ID	640		
Observations	73222		
Marginal R^2 / Conditional R^2	0.838 / 0.945		

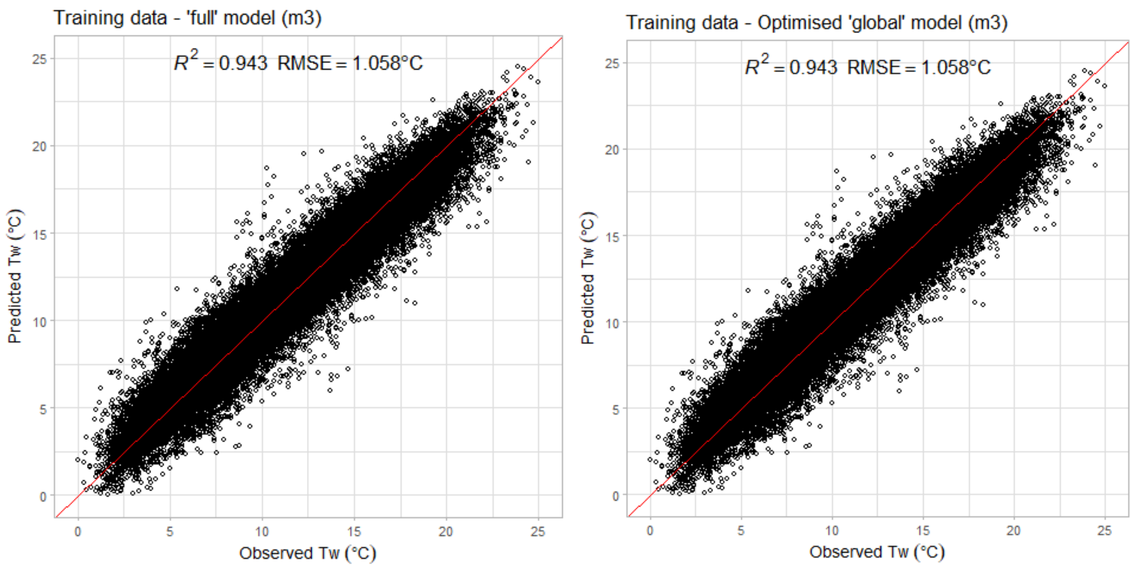


Figure 6: Performance of the 'full' (left) and optimised 'global' model (right) for English river water temperature prediction against the training dataset (3+ value monthly mean *Tw* values from sites with over 5 years of data; $n = 73,222$). The red line indicates the $x = y$ identity line. (The panels appear identical but R^2 and RMSE values differ after 5 significant figures).

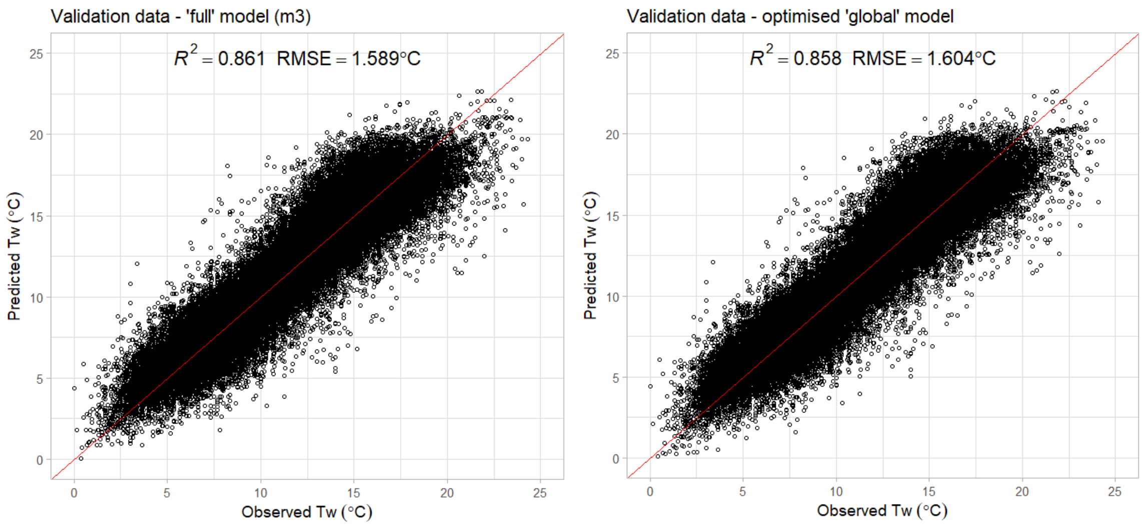


Figure 7: Performance of the 'full' (left) and optimised 'global' model (right) for English river water temperature prediction against the validation dataset (3+ value monthly mean Tw values from sites with less than 5 years of data; $n = 57,409$). The red line indicates the $x = y$ identity line.

Projections of monthly mean daytime river water temperatures

Bias-corrected UKCP18 monthly air temperature values and sample site covariate information (altitude at site, distance to river mouth, site northing, percentage calcareous geology and percentage urban and suburban land use upstream of the sample site) were used in the optimised 'global' model to generate predictions of monthly mean river water temperature at each of the 4,082 sites having at least six-monthly mean water temperature values. Water temperature predictions were generated for each of the twelve UKCP18 potential outcomes and the median value taken for each month between 1981 and 2080 for each site as a summary statistic.

These median values of projected water temperature were used to carry out a high-level investigation into potential temperature changes that could impact on the life cycles of two important salmonid fish species, *Salmo trutta* (brown trout) and *Salmo salar (*Atlantic salmon). Two temperature thresholds were analysed these being 1) the critical growth and feeding temperature range for each species, and 2) the maximum temperatures experienced by eggs. For brown trout a critical growth and feeding temperature range of 3.5 to 19.5 °C was used and for Atlantic salmon 6 to 22.5 °C (Table A12, Soloman and Lightfoot, 2008). Eggs are the life stage that has the lowest temperature tolerance (Table A12, Elliot and Elliot, 2010) and thresholds examined were 13 °C and 16 °C (the upper limits for brown trout and Atlantic salmon egg survival, respectively). 12 °C was also chosen as a threshold, due to reports of increased egg mortality, reduced size at hatching and increased deformity rates when this temperature is exceeded (Solomon and Lightfoot, 2008). The percentage of sites projected to exceed these three temperature thresholds during the period November to February was recorded for each decade across the UKCP18 climate projection range 1981 – 2080. The November to February period was chosen to represent months of the year when salmon and trout eggs could be expected to be present and developing beneath gravel in the rivers. This work is presented only as a demonstration of the kind of investigations that can be facilitated by the model. Factors such as whether sampling points coincide with localities where salmonid eggs would be found, differences in spawning time across the country, or the potential for genetically distinct populations to be adapted to different water temperatures have not been considered. Further work would be required to produce a robust assessment of how temperature changes could impact salmonid life cycles.

Typology-specific models

To identify river typologies, cluster analysis was carried out on the water sampling sites, based on the GIS-derived covariates. Ward's method generated 6 coherent clusters (Figure 8) that aligned with the following typologies: 'Permeable rock / Arable / Grassland' (Cluster 1), 'Harder rock / Grass / Arable' (Cluster 2), 'Urban / generally Calcareous geology' (Cluster 3), 'Grass Upland / mixed geology' (Cluster 4), 'Chalk streams' (Cluster 5), and 'Large rivers / High Strahler order (8-9)' (Cluster 6). The 'full' model with all selected covariates was applied in turn to each cluster's qualifying data and the resultant model optimised to find the simplest solution. The performance of each optimised cluster model was then compared to the performance of the optimised 'global' model in generating predictions for water temperature in each of the clusters.

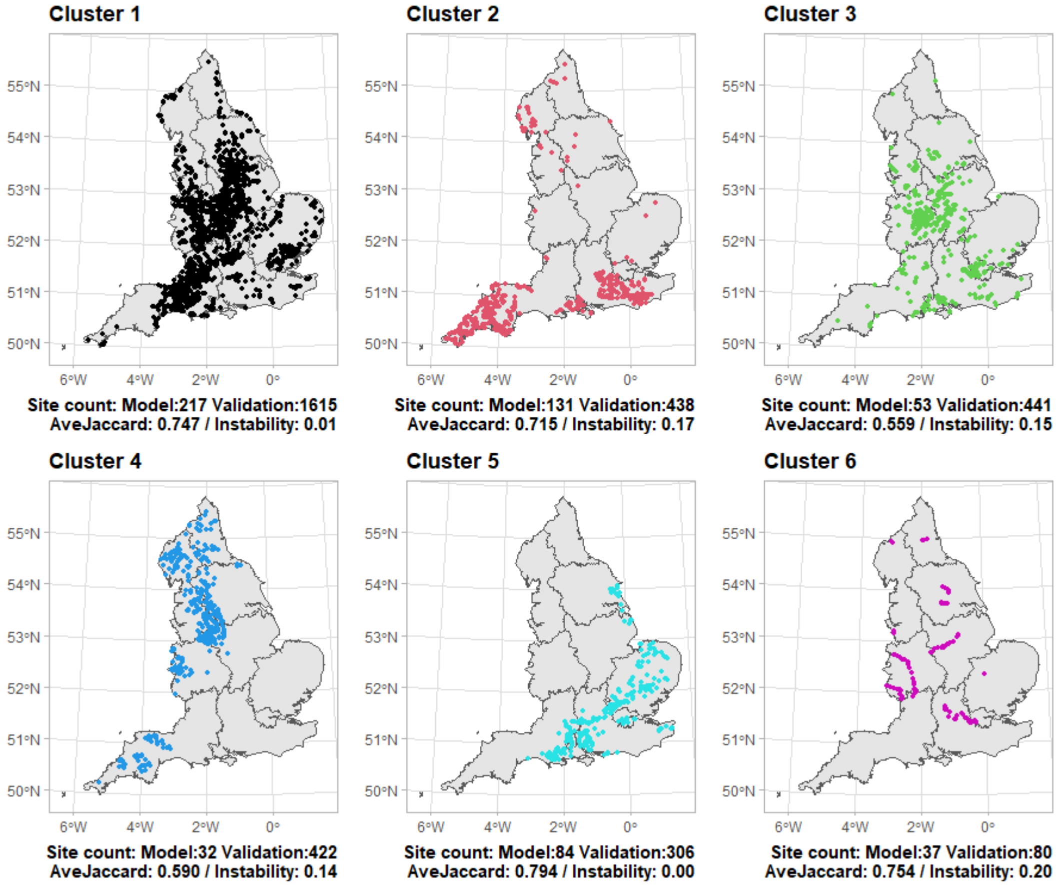


Figure 8: Distribution of sites within each identified typological cluster, 1 = Permeable rock / Arable / Grassland, 2 = Harder rock / Grass / Arable, 3 = Urban / generally Calcareous geology, 4 = Grass Upland / mixed geology, 5 = Chalk streams, 6 = Large rivers. Jaccard Index (AveJaccard; between 0 (dissimilar) and 1 (identical)) gives an indication of similarity between cluster members; Instability Index (between 0 (stable) and 1 (unstable)) gives an indication how coherent the clusters are. Further details in Appendix.

Results

National temperature projections

Projected summer maximum water temperatures for four decades up to 2080 suggest substantial increases are likely by 2080 (Figures 9, 10). Regional differences are evident with sites in the northwest of England projected to experience the lowest increases in hottest month water temperatures whereas those in central / southern regions, and potentially associated with urban centres are projected to experience the greatest increase. Some ‘hotspots’ (2080 *maxTw* >27 °C: River Thames, 8 sites; River Trent, 3 sites; River Ouse, 5 sites; single sites on rivers Aire, Weaver, Welland, Rother and Anglian Stour) as well as ‘cold spots’ (2080 *maxTw* <17°C: River Tees, 4 sites; River Teign, 2 sites; single sites on River Cerne and Beer Stream) can also be identified amongst the sites.

While the median monthly averages of Tw, of the 12 potential outcomes derived from the UKCP18 climate change projections, provide a suitable summary statistic, the range of responses encompassed by these 12 outcomes is also of interest. Figure 11 presents the range of minimum (blue points), mean (black points) and maximum (red points) values projected for each of the twelve UKCP18 potential outcomes across all sites for the UKCP18 projection period 1981 to 2080. While all ranges indicate an increasing Tw trend, the rate of change is highest for the maximum monthly averages (22.31±1.39 °C in the 1981-2005 reference period to 28.26±1.56 °C in decade 2070-2079; a change of +5.94 °C, approximately 0.6 °C per decade) and lowest for the minimum monthly averages (-0.56±1.06 °C in the 1981-2005 reference period to 1.89±0.95 °C in decade 2070-2079; a change of +2.45 °C).

There is also a slight shift in the timing of maximum and minimum temperatures across the projection period. During the baseline period (1981 to 2005 inclusive) the average timing of maximum temperatures during the year occurs at month ‘7.46’ (mid-July), with the minimum temperature occurring at month ‘1.24’ (early January). By the decade 2070-2079, this has shifted to month ‘7.96’ (late July) for the maximum and month ‘1.63’ (mid-January) for the minimum.

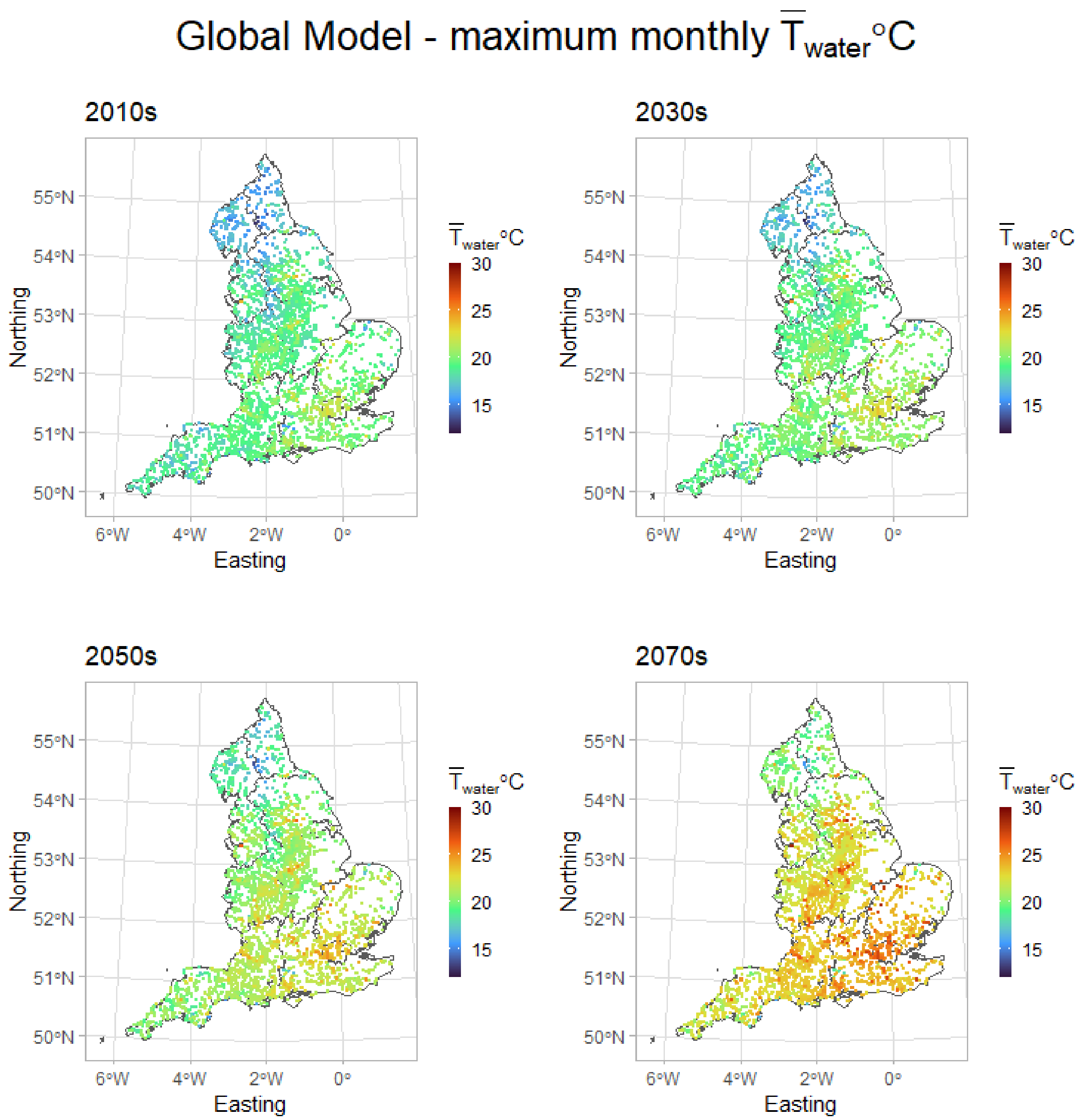


Figure 9: Projected warmest monthly water temperatures at English river sample sites for the decades 2010-2019, 2030-2039, 2050-2059 and 2070-2079, based on bias corrected UKCP18 RCP8.5 air temperature projections and the 'global' model for monthly mean water temperatures. Values represent the maximum median Tw across all UKCP18 ensemble members for each site in each decade. Larger versions of the figures are presented in the Appendix (Figure A9).

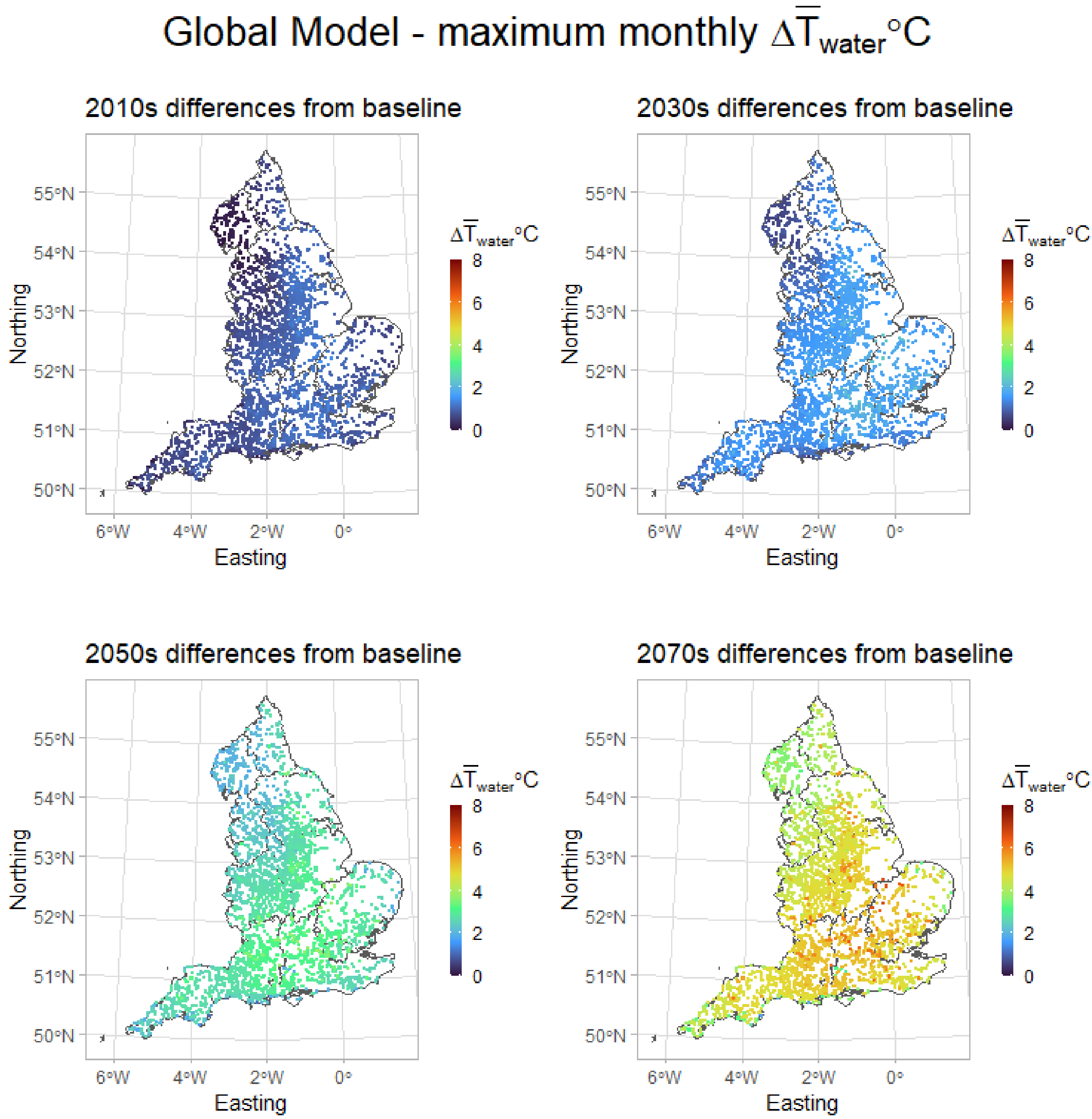


Figure 10: Changes in the warmest month at English river sampling sites over baseline levels (1981-2005 inclusive) for the decades 2010-2019, 2030-2039, 2050-2059 and 2070-2079, based on bias corrected UKCP18 RCP8.5 air temperature projections and the 'global' model for monthly mean water temperatures. Larger versions of the figures are presented in the Appendix (Figure A10).

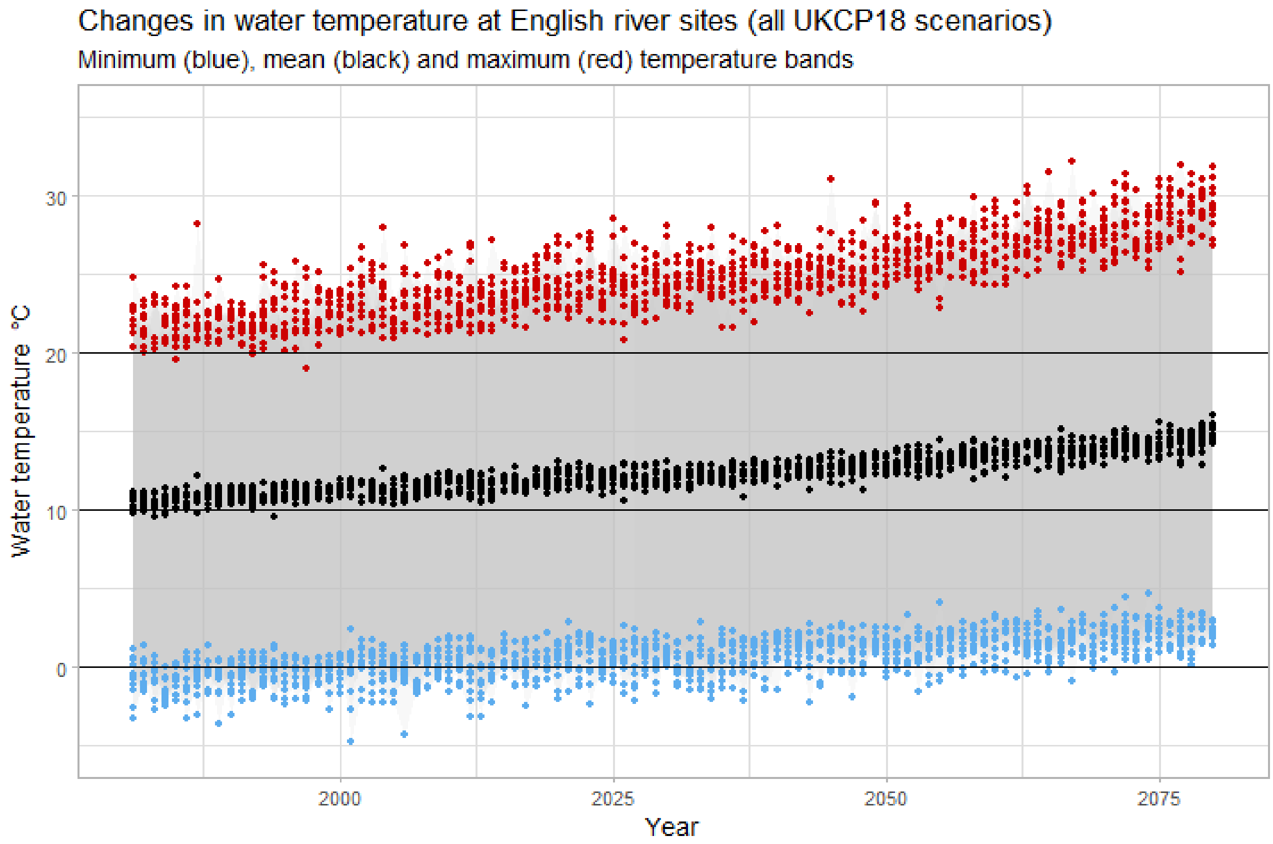


Figure 11: Projected change in monthly water temperature in England's rivers over the UKCP18 projection period 1981 to 2080. Each series of coloured points represent the values projected for each of the 12 ensemble members: blue points = minimum values, black points = mean values, and red points = maximum values.

Critical temperature thresholds for brown trout and Atlantic salmon

Brown trout and Atlantic salmon are two important salmonid fish species found in English rivers, with salmon rivers mainly being restricted to the northeast, western draining coastlines, and southern chalk streams, and brown trout being distributed more widely (CEFAS, Environment Agency, Natural Resources Wales, 2023; Figure 12). The percentage of sites projected to exceed the upper values of the feeding and growth temperature ranges (19.5 °C for brown trout and 22.5 °C for Atlantic salmon, Table A12) is presented in Figure 13. For brown trout we considered all sites for which we had projections and found that monthly mean water temperatures are projected to exceed the 19.5 °C upper limit at the majority of sites by 2070. For salmon we considered all sites where we had projections on the principal salmon rivers (CEFAS, Environment Agency, Natural Resources Wales, 2023). The higher salmon threshold of 22.5 °C is projected to start being exceeded after 2050, increasing to 50% of these sites by 2080.

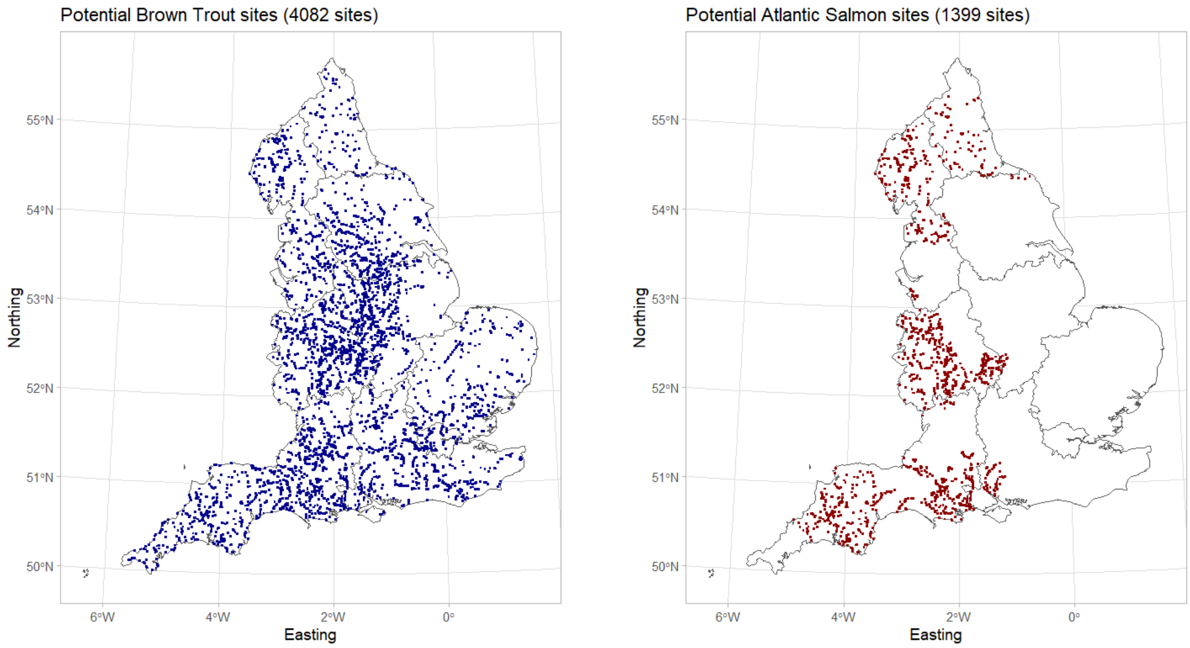


Figure 12: distribution of potential brown trout (left panel) and Atlantic salmon (right panel) sites in England for which we have water temperature projections. Salmon sites are restricted to principal salmon rivers (CEFAS, Environment Agency, Natural Resources Wales, 2023).

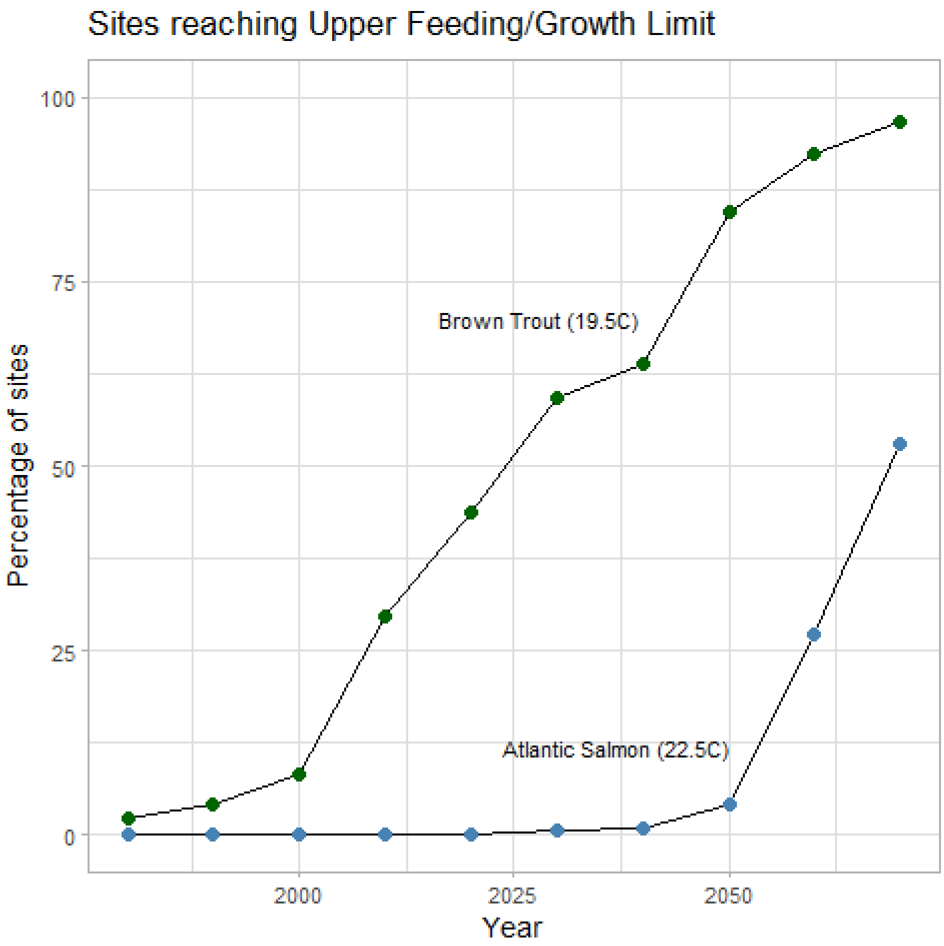


Figure 13: Percentage of sites in English rivers with monthly average projections expected to reach upper critical temperatures for two salmonid fish species.

The spawning season for both Atlantic salmon and brown trout is during the late autumn and winter months, and we have examined the November to February period as representing the time when eggs could be expected to be developing within the nests (redds) in river gravel. Egg survival is temperature dependent with critical ranges of 0 to 13 °C for trout and 0 to 16 °C for salmon (Table A12). An additional threshold of 12 °C has been selected for consideration due to reports of increased egg mortality, smaller size in hatching and increased levels of deformity for both species (Solomon and Lightfoot, 2008). Sites projected to have a monthly mean that exceeds this 12 °C threshold between November and February in different decades up to 2080 are shown in Figures 14 and 15, for brown trout and Atlantic salmon respectively. While the initial sites affected appear few and mostly in the south of England (Figures 14 and 15, 2010 panel), with affected catchments including greater numbers of more northerly sites as the decades progress. Around 70% of sites (70.5% of potential brown trout sites and 69.5% of salmon sites) are projected to be impacted by 2080. Figure 16 shows the percentage of sites projected to experience temperatures at or above these three important spawning threshold temperatures of 12 °C, 13 °C and 16 °C to 2080 and indicates a significant number of sites reaching 12 °C and 13 °C across this period, particularly after 2040. When considering these results, it must be remembered that this is a high-level exercise only and further factors would need to be considered for a robust assessment of future changes, including ensuring sampling points are in localities where spawning may occur.

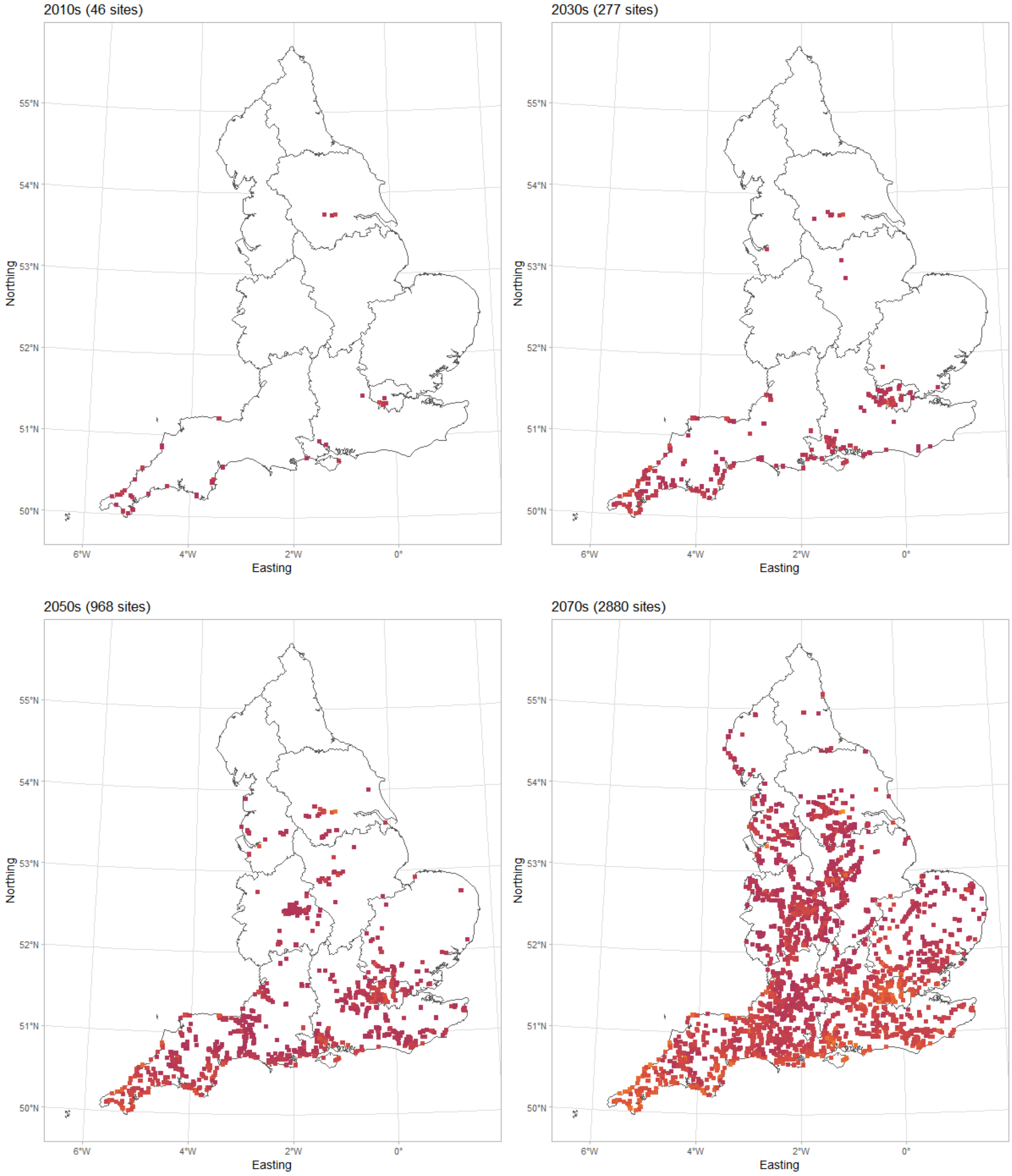


Figure 14: Brown trout egg thermal tolerance: Sites across England projected to exceed a monthly mean water temperature of 12°C in November to February when eggs could be expected to be within river gravels in the decades 2010-2019, 2030-2039, 2050-2059 and 2070-2079. Sites exceeding threshold marked in red/orange (total number of sites modelled: 4082).

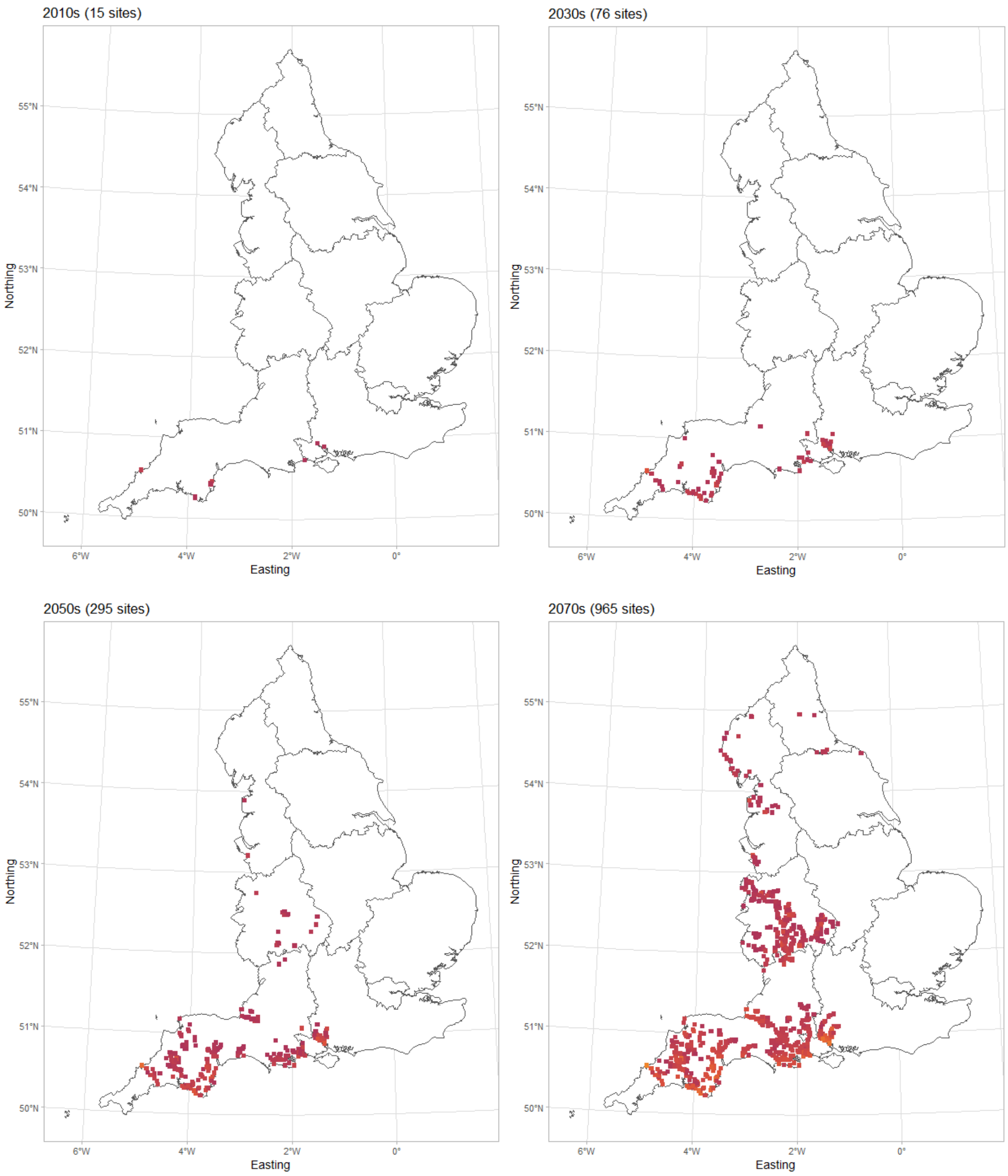


Figure 15: Atlantic salmon egg thermal tolerance: Sites across England's salmon rivers projected to exceed a monthly mean water temperature of 12°C in November to February when eggs could be expected to be within river gravels in the decades 2010-2019, 2030-2039, 2050-2059 and 2070-2079. Sites exceeding threshold marked in red/orange (total number of sites modelled: 1399. Note: these sites are on principal salmon rivers but do not necessarily coincide with salmon spawning locations (CEFAS, Environment Agency, Natural Resources Wales, 2023).

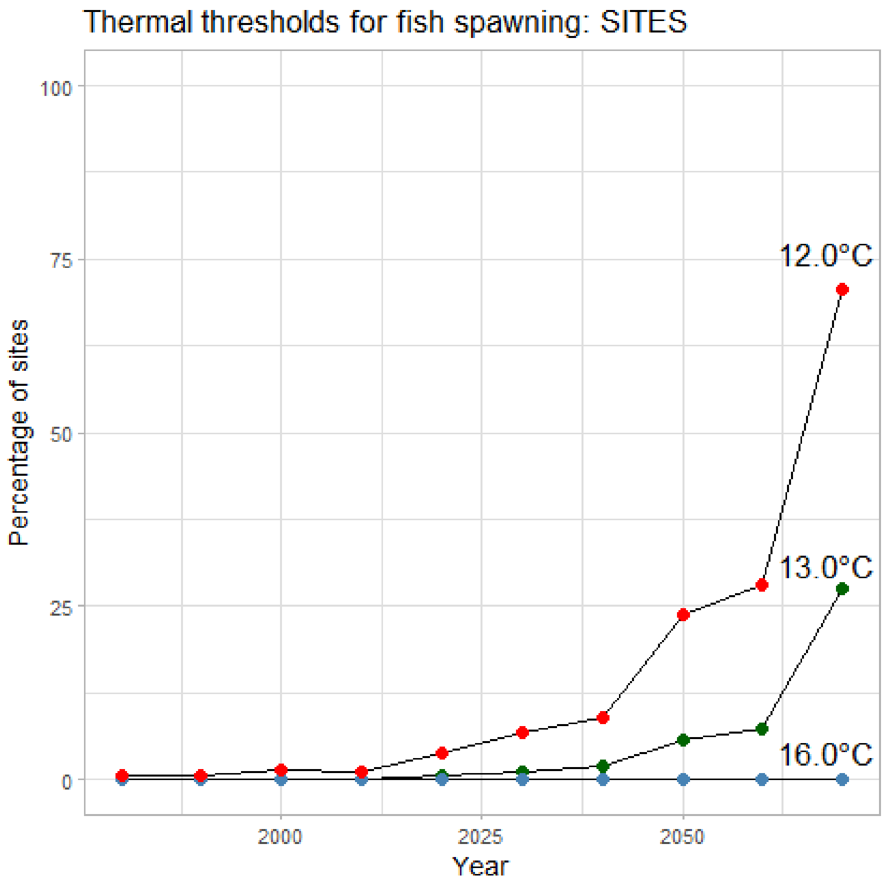


Figure 16: Projected changes in thermal thresholds important for fish egg survival during November to February from 1981 to 2080. The 13 °C threshold is the upper limit for brown trout egg survival, the 16 °C threshold relates to Atlantic salmon egg survival, and the 12 °C threshold is associated with increased egg mortality and increased deformity rates in both species (Solomon and Lightfoot, 2008).

Typology specific models

The 'full' model (see page 16) was applied to each of the six identified cluster groups and the resultant models simplified to retain only significant covariate factors. Projected water temperature values from each of these models was compared with those generated by the optimised 'global' model and none were found to differ significantly (ANOVA; $p<0.05$). Demonstrating that these models showed no improvement over the national model when estimating future monthly mean temperature. Further details are available in the Appendix. The projected temperature responses of sites within these clusters were analysed and differences were seen in the expected decadal rates of temperature change (Figure 17, Table 3). Differences were seen in the projected temperature responses of these groups, with arable/grassland sites on permeable rock experiencing the highest decadal changes in maximum temperatures (0.66 °C per decade) and chalk the lowest (0.51 °C per decade). Grass upland sites are projected to experience the highest summer maxima by 2080 (>28 °C) and chalk the lowest summer maxima (<26 °C).

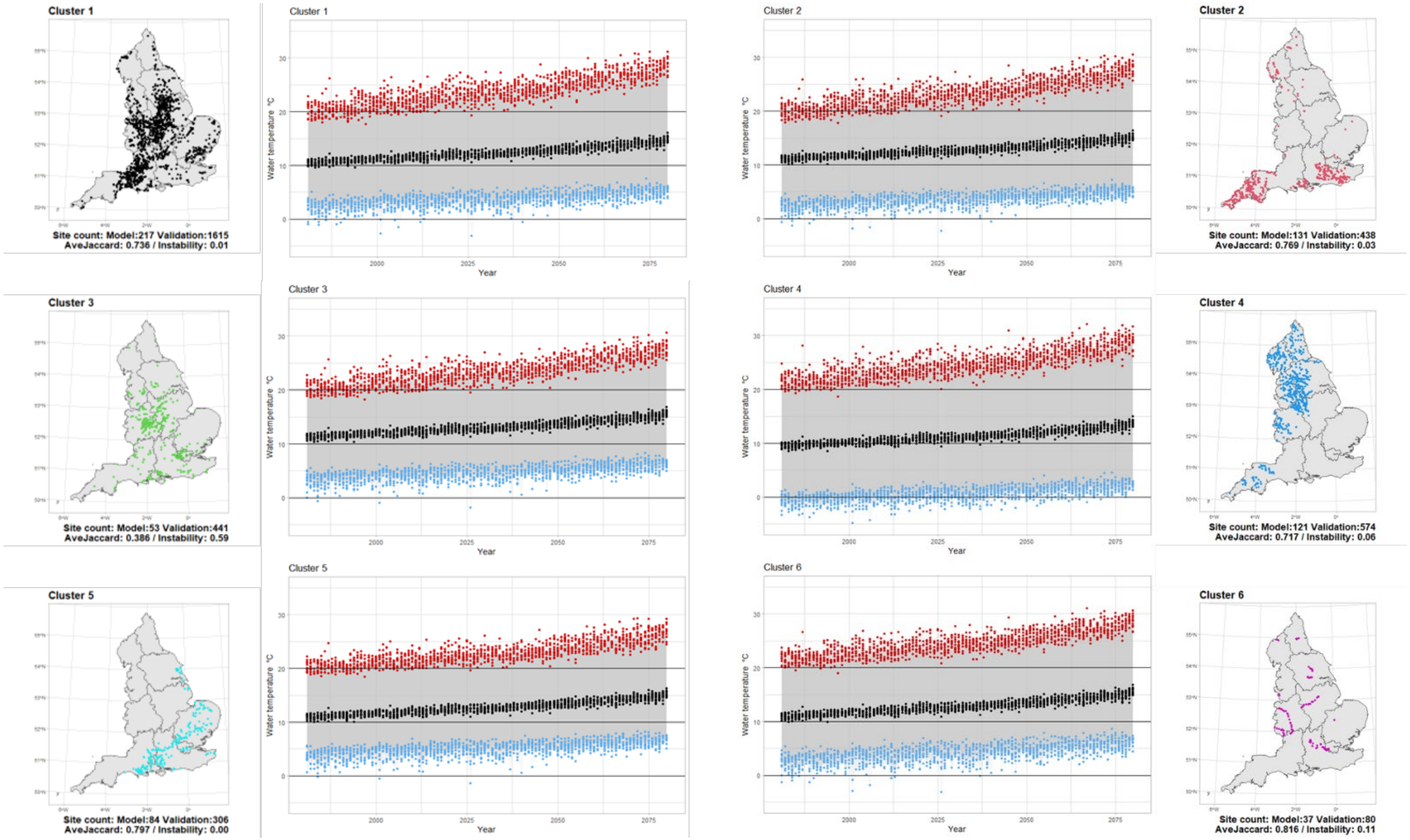


Figure 17: Projected change in monthly water temperature in England's rivers over the UKCP18 projection period 1981 to 2080 for each cluster / typological group. Each series of coloured points represent the values projected for each of the 12 ensemble members: blue points = minimum values, black points = mean values, and red points = maximum values.

Table 3: Trends in projected monthly mean water temperature predictions, including average decadal change in temperature for each cluster / typological group.

Cluster	Typology	Range	Baseline mean (°C; 1981 – 2005)	2070-2079 mean (°C)	Decadal change (°C)
1	Permeable rock, arable / grass	Maximum	20.93	27.52	0.66
		Mean	10.88	14.14	0.33
		Minimum	2.47	4.90	0.24
2	Harder rock, grass / arable	Maximum	20.69	27.01	0.63
		Mean	11.46	14.57	0.31
		Minimum	2.45	4.80	0.24
3	Urban, generally permeable	Maximum	20.88	26.83	0.60
		Mean	11.64	14.90	0.33
		Minimum	3.50	5.93	0.24
4	Grass upland	Maximum	22.07	28.11	0.60
		Mean	9.85	12.99	0.31
		Minimum	-0.56	1.89	0.25
5	Chalk streams	Maximum	20.80	25.92	0.51
		Mean	11.29	14.51	0.32
		Minimum	3.84	6.36	0.25
6	Large rivers	Maximum	22.18	27.74	0.56
		Mean	11.31	14.77	0.35
		Minimum	2.75	5.77	0.30

Discussion

Temperature projections

Validation of the developed 'global' model for English river water temperatures showed strong predictive performance (R^2=0.85). Application of this model to predict monthly mean water temperature values at sites under the influence of climate change showed significant water temperature increases of ecological significance. Regional variation in the level of *Tw* rise is seen with sites around London showing the highest increases and those in northern England the lowest.

It should be noted that the projections are based on a simple future scenario where only air temperature is changing. More complex future scenarios incorporating changes in rainfall frequency and intensity, surface water/groundwater balance (including changes in abstractions), evapotranspiration, as well as changes in water-temperature-influencing environmental and landscape characteristics would help increase confidence in water temperature projections. Future changes would also be associated with wider changes in human activity (such as altering land use) and climate change such as river flows, rainfall intensity and heat waves.

Changes in a monthly mean value will disguise potentially significant fluctuations in water temperature that may occur throughout that monthly period. An increased frequency of summer heat waves may only influence the monthly mean temperature by a small amount, yet may cause significant stress on ecological habitats, including aquatic ecosystems. Model sensitivity to such fluctuations would require a move to higher frequency *Tw* data collection in the sub-daily range.

The ecological significance of potential increases in river water temperatures is illustrated by the effect on important temperature thresholds for salmonid fish. Thermal boundaries are known to affect salmonid at different stages of their life cycles. The increasing number of sites projected to experience water temperatures above critical ranges for brown trout (19.5°C) is of particular concern (Solomon and Lightfoot, 2008; Elliott and Elliott, 2010). The challenge to salmonid egg survival is also projected to increase throughout the UKCP18 projection period, with the important 12°C threshold (Solomon and Lightfoot, 2008) over which increased deformity rates and decreased egg survival are experienced, being breached extensively by 2080. Further work is needed to increase the robustness of these assessments.

Data availability

Effective modelling of water temperature in complex fluvial systems requires a baseline of water temperature records at a sufficient sampling frequency and over a sufficient period to reflect the temporal dynamics of in-channel temperature variability (Environment Agency, 2021). Much of the water temperature data collected under the Environment

Agency’s long-term environmental monitoring programme has been linked to statutory monitoring requirements to inform an understanding of river health rather than to explore river water temperature dynamics. Hence, the data available for this study was collected at coarse timesteps that supported the development of a monthly daytime mean water temperature model for England’s rivers. Using a monthly mean value as an indicator of a temporally and spatially dynamic variable such as water temperature is useful in assessing long-term trends and relative changes. However, highly dynamic events, such as heat waves, may be missed.

The spot-sample water temperature readings in the combined *Tw* dataset provide a ‘snapshot’ temperature value for water from various sources (e.g., upstream flow, precipitation, surface water runoff, groundwater supply; Leach et al., 2022) all of which will have different temperature profiles and volume contributions to the water in the river. They will also contain a reflection of the physical environment surrounding the river channel and the dynamics of water exchange between different compartments within the channel. These contributions will vary throughout the year and across climatic cycles and model development requires a strong baseline of measurements which encompass a significant amount of this variation. The qualifying threshold of sites to have a minimum of five years monthly mean water temperature data before inclusion in the model development process was a pragmatic choice given the data available but was regarded as an absolute minimum to reflect meaningful interannual *Tw* variation. Modelling at this temporal scale can, however, limit the model’s sensitivity to environmental and landscape characteristics known to influence river water temperature (Jackson et al., 2018).

Targeted placement of temperature dataloggers with 15-minute monitoring resolution across catchments and regions at sites representative of different temperature-influencing landscape and environmental features is driving water quality management decisions in Scotland (Jackson et al., 2016; 2018; 2020) and offers an alternative to staff-intensive synoptic monitoring approaches (Webb et al., 2008). High frequency water temperature information from existing Environment Agency and partner organisations provide potential datasets that could facilitate higher resolution projections. Whichever data are used in modelling the outputs will be limited by where the input observations are recorded, reducing the potential to apply any generated water temperature models to unmonitored sites.

Additional improvements could be made through incorporation of a national dataset of groundwater level and flow data and incorporating a shading model. Riparian shading along the river channel is a factor known to influence water temperature and identified as a future management technique to mitigate adverse river temperatures (Garner et al., 2017).

Modelling process

The model development process previously described by the Environment Agency (2021) was successfully implemented to develop a model for English river water temperature based around monthly mean daytime values, despite the limitations of the data available.

The selection of suitable covariates known to influence in-channel water temperatures was guided by previous studies (Jackson et al., 2016, 2018) and the pragmatic choice to limit these to values that could be derived from GIS sources was taken with practical management application in mind. Establishing values for all selected covariates using this method was not entirely successful as sample site locations did not always coincide with river sections where such values could be derived. Covariate values for river width could not be calculated for multi-channel reaches or where channel modification (culverts, pipes, underground sections) had taken place, and ‘distance’ metrics (distance to mouth of river, Strahler river order) could not be calculated for sites associated with secondary river channels rather than the main channel. This reduced the number of sites that could participate in model development and could potentially be mitigated by closer inspection of site/GIS alignments and/or the development of automated processing to establish suitable values from nearby sites.

The developed models displayed a strong reliance on the air temperature / water temperature relationship. The seasonal nature of this relationship (Mohseni and Stefan, 1999; Webb et al., 2008) was reflected in the monthly mean water temperature measurements derived from the combined *Tw* dataset and a series of seasonal linear relationships were included to approximate the ‘S’ shaped curve usually used to describe this relationship in river and open water systems. The ‘summer’ *Ta* / *Tw* relationship is of particular significance when modelling maximum water temperatures as too steep a gradient in this section of the relationship will over-estimate water temperatures from the corresponding air temperature values. Including this seasonal aspect of the *Ta* / *Tw* relationship was judged essential in model development and was therefore present in all model solutions.

The ‘global’ model for monthly mean daytime water temperatures for English rivers presented here is a relatively simple one. Apart from the seasonal *Ta* / *Tw* relationship, the model is only dependent on five covariate values, the percentage upstream calcareous hard geology, the percentage urban and suburban land use upstream, an ‘x-y-z’ coordinate type combination of ‘Distance to mouth’, ‘Northing’, and site altitude, and a ‘random effects’ component allowing the seasonal *Ta* / *Tw* relationship to vary between sites. This concise group of dependencies is seen as a function of the monthly timestep employed and the large spatial scale of the modelled sites. The influence of landscape and river channel covariates would be expected to increase if the modelling timestep, or spatial scale were reduced. Typology-specific models did select different groups of covariates than the ‘global’ model (see Appendix) but the limited influence of these covariates against the overwhelming influence of the *Ta* / *Tw* relationship meant that the optimised ‘global’ model produced equally effective water temperature estimates.

The hydrological basis behind the model’s selected covariates indicates broad landscape controls to river water temperature at this temporal scale. The percentage of upstream calcareous hard geology gives an indication of potential water sources entering the river channel, with low percentages reflecting impermeable underlying geology and surface water dominated systems, and greater groundwater contributions with increasing rock permeability (Berrie, 1992). Groundwater exhibits a stable temperature profile and can modulate the temperature fluctuations seen in surface water where atmospheric and climatological influences prevail (Acornley, 1999). Levels of urbanisation have direct and indirect influences on water temperature by increasing the surface water component in river channels due to hard surface runoff, increased industrial and water processing inputs, modifying groundwater/surface water interactions due to river channel modification (lining, straightening/canalising, culverting, for example), and indirectly through urban heat island effects (Ficklin et al., 2023). Distance to river mouth can be considered an indication of continentality and the different thermal properties (specific heat capacity and heat exchange processes) between land and water, and site altitude and Northing indicate differences in the vertical and north-south air temperature gradients (Jackson et al., 2016).

The ‘global’ model outputs presented here are based on the high emissions RCP8.5 scenarios of the UKCP18 climate change projections. While these are seen as a suitable basis for contingency planning there is some debate whether they present a realistic projection given their reliance on continued rises in CO_2 levels due to fossil fuel burning well into the 21st century. The ‘global’ model is not dependent on this particular projection of future *Ta* and so Tw could be projected using different scenarios as they are developed and refined.

Conclusions

The future impacts of climate change on aquatic ecosystems remain uncertain. Future changes in river flows and temperature increases are expected to be of critical importance. Quantifying the magnitude and timing of future warming in rivers is needed to understand where and when changes may happen and inform potential adaptation measures and management actions.

This project has generated the first national water temperature projections for English rivers. Based on the UKCP18 ‘high emissions scenario’, ecologically significant increases in water temperature were projected with the warmest monthly average rising by 0.6 °C per decade above reference levels to 2080. An important temperature threshold of 12 °C for salmonid egg survival during the winter will likely be exceeded at over 70% of sites by 2080. Adult brown trout will be under threat from high summer temperatures with almost all sites experiencing temperatures that exceed their upper growth/feeding temperature range of 19.5 °C by 2080. Further work is needed to increase the robustness of these assessments.

Whilst the projected increases in monthly mean river water temperatures will underestimate impacts of short-term events such as heat waves, they are the first indication of the magnitude of change which may happen and where and when it may occur. These results will enable more informed decisions about potential management actions and adaptation measures. The temporal and spatial resolution of the projections may be improved as more higher resolution data become available.

References

ACORNLEY, R.M., 1999. Water temperatures within spawning beds in two chalk streams and implications for salmonid egg development. Hydrological Processes, 13(3), 439-446

BARTOŃ, K., 2023. MuMIn: Multi-Model Inference. R package version 1.47.5. https://CRAN.R-project.org/package=MuMIn

BATES, D., MAECHLER, M., BOLKER, B., AND WALKER, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48.

BERRIE, A.D., 1992. The chalk-stream environment. Hydrobiologia, 248, 3-9.

BIVAND, R.S., PEBESMA, E. AND GOMEZ-RUBIO, V., 2013. Applied spatial data analysis with R, Second edition. Springer, NY.

CEFAS, ENVIRONMENT AGENCY, NATURAL RESOURCES WALES, 2023. Salmon stocks and fisheries in England and Wales, 2022. Preliminary assessment prepared for ICES, March 2023.

COLEY, A., PETTIT, L., McMELLIN, G., CROUCH, A. AND CLEMENTS, A., 2018. Detailed River Network (DRN) Data Description and Guide Document (Version 6). Bristol: Environment Agency

DORMANN, C.F, ELITH, J., BACHER, S., BUCHMANN, C., CARL, G., CARRÉ, G., GARCÍA MARQUÉZ, J.R., GRUBER, B., LAFOURCADE, B., LEITÃO, P.J., MÜNKEMÜLLER, T., McCLEAN, C., OSBORNE, P.E., REINEKING, B., SCHRÖDER, B., SKIDMORE, A.K., ZURELL, D., AND LAUTENBACH, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography, 36(1), 27-46

DOWLE, M., AND SRINIVASAN, A., 2023. data.table: Extension of `data.frame`. R package version 1.14.8. https://CRAN.R-project.org/package=data.table

ELLIOTT, J.M. AND ELLIOTT, J.A., 2010. Temperature requirements of Atlantic salmon *Salmo salar*, brown trout *Salmo trutta* and Arctic charr *Salvelinus alpinus*: predicting the effects of climate change. Journal of Fish Biology, 77, 1793-1817.

ENVIRONMENT AGENCY, 2019. Climate change and eutrophication risk thresholds in English rivers. Report SC140013/R2. Bristol: Environment Agency.

ENVIRONMENT AGENCY, 2021. Scoping a flexible framework for producing river water temperature projections. Report SC200008/R. Bristol: Environment Agency.

ENVIRONMENT AGENCY, 2022. River water temperature projections for English Chalk streams. Report SC210011/R. Bristol: Environment Agency.

ENVIRONMENT AGENCY, 2023. Salmon stocks and fisheries in England and Wales 2022. Report . Bristol: Environment Agency.

FICKLIN, D.L., HANNAH, D.M., WANDERS, N., DUGDALE, S.J., ENGLAND, J., KLAUS, J., KELLEHER, C., KHAMIS, K. AND CHARLTON, M.B., 2023. Rethinking river water temperature in a changing human-dominated world. Nature Water, 1, 125-128

FOX, J., and WEISBERG, S., 2019. An R Companion to Applied Regression, 3rd Edition. Thousand Oaks, CA.

GARNER, G., MALCOLM, I.A., SADLER, J.P. AND HANNAH, D.M., 2017. The role of riparian vegetation density, channel orientation and water velocity in determining river temperature dynamics. Journal of Hydrology, 553, 471-485.

GROLEMUND, G., AND WICKHAM, H., 2011. Dates and Times Made Easy with lubridate. Journal of Statistical Software, 40(3), 1-25.

JACKSON, F.L., MALCOLM, I.A. AND HANNAH, D.M., 2016 A novel approach for designing large-scale river temperature monitoring networks. Hydrology Research, 47(3), 569-590.

JACKSON, F.L., FRYER, R.J., HANNAH, D.M., MILLAR, C.P. AND MALCOLM, I.A., 2018. A spatio-temporal statistical model of maximum daily river temperatures to inform the management of Scotland’s Atlantic salmon rivers under climate change. The Science of the Total Environment, 612, 1543–1558.

JACKSON, F.L., HANNAH, D.M., OUELLET, V. AND MALCOLM I.A., 2020. A deterministic river temperature model to prioritize management of riparian woodlands and reduce summer maximum river temperatures. Hydrological Processes, 35(8), e14314.

KNOUFT, J.H., BOTERO-ACOSTA, A., WU, C.-L., CHARRY, B., CHU, M.L., DELL, A.I., HALL, D.M. & HERRINGTON, S.J., 2021. Forested riparian buffers as climate adaptation tools for management of riverine flow and thermal regimes: A case study in the Meramec River basin. Sustainability: Science Practice and Policy, 13, 1877.

KUHN, M., 2008. Building Predictive Models in R using the caret Package. Journal of Statistical Software, 28(5), 1-26.

KUZNETSOVA, A., BROCKHOFF, P.B., AND CHRISTENSEN, R.H.B., 2017. lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software, 82(13), 1-26.

LEACH, J.A., HUDSON, D.T. AND MOORE, R.D., 2022. Assessing stream temperature response and recovery for different harvesting systems in northern hardwood forests using 40 years of spot measurements. Hydrological Processes, 36(11), e14753

LENDERINK, G., BUISHAND, A. AND VAN DEURSEN, W., 2007. Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. Hydrology and Earth System Sciences, 11(3), 1145-1159.

LÜDECKE, D., 2022. sjstats: Statistical Functions for Regression Models (Version 0.18.2). https://CRAN.R-project.org/package=sjstats

MACKEY, A. P. AND BERRIE, A. D., 1991. The prediction of water temperatures in chalk streams from air temperatures. Hydrobiologia, 210, 183-189.

MET OFFICE; HOLLIS, D., MCCARTHY, M., KENDON, M., AND LEGG, T. 2023: HadUK-Grid Gridded Climate Observations on a 1km grid over the UK, v1.2.0.ceda (1836-2022). NERC EDS Centre for Environmental Data Analysis, 30 August 2023.

MET OFFICE, 2019. UKCP18 Science Overview Report November 2018 (Updated March 2019).

MOHSENI, O. AND STEFAN, H.G., 1999. Stream temperature/air temperature relationship: a physical interpretation. Journal of Hydrology, 218, 128-141.

ORR, H.G., SIMPSON, G.L., DES CLERS, S., WATTS, G., HUGHES, M., HANNAFORD, J., DUNBAR, M.J., LAIZÉ, C.L.R., WILBY, R.L., BATTARBEE, R.W. AND EVANS, R., 2015. Detecting changes river temperatures in England and Wales. Hydrological Processes, 29, 752-766.

PEBESMA, E.J., BIVAND, R.S., 2005. Classes and methods for spatial data in R. R News, 5(2).

PEBESMA, E. AND BIVAND, R., 2023. Spatial Data Science: With Applications in R. Chapman and Hall/CRC.

PEBESMA, E., 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal, 10(1), 439-446.

POFF, N.L., ALLAN, J.D., BAIN, M.B., KARR, J.R., PRESTEGAARD, K.L., RICHTER, B.D., SPARKS, R.E. AND STROMBERG, J.C., 1997. The natural flow regime. Bioscience, 47(11), 769-784.

R CORE TEAM 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

RANGELEY-WILSON, C., 2021 Chalk Stream Restoration Strategy 2021 Main Report and Appendices. Catchment Based Approach https://catchmentbasedapproach.org/learn/chalk-stream-strategy/

SEAR, D.A., ARMITAGE, P.D. AND DAWSON, F.H., 1999. Groundwater dominated rivers. Hydrological processes, 13, 255-276.

SOLOMON, D.J. AND LIGHTFOOT, G.W., 2008. The thermal biology of brown trout and Atlantic salmon. Report SCHO0808BOLV-E-P. Bristol: Environment Agency.

UKCEH. 2023. eFLaG: Enhanced future flows and groundwater. https://www.ceh.ac.uk/our-science/projects/eflag-enhanced-future-flows-and-groundwater

WARING, E., QUINN, M., McNAMARA, A., ARINO de la RUBIA, E., ZHU, H. and ELLIS, S., 2021. skimr: Compact and Flexible Summaries of Data. R package version 2.1.5. https://CRAN.R-project.org/package=skimr

WEBB, B.W., HANNAH, D.M., MOORE, D., BROWN, L.E. AND NOBILIS, F., 2008. Recent advances in stream and river temperature research. Hydrological Processes, 22, 902-918

WICKHAM, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

WICKHAM, H. et al., 2019. Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686.

WICKHAM, H., AND BRYAN, J., 2023. readxl: Read Excel Files. R package version 1.4.3. https://CRAN.R-project.org/package=readxl

WILBY, R.L., AND JOHNSON, M.F., 2020. Climate variability and implications for keeping rivers cool in England. Climate Risk Management, 30, 100259.

WOODWARD, G., PERKINS, D.M., AND BROWN, L.E., 2010. Climate change and freshwater ecosystems: impacts across multiple levels of organisation. Philosophical Transactions of the Royal Society of London B, 365, 2093-2106.

List of abbreviations

BIC	Bayesian Information Criterion
CEH	Centre for Ecology and Hydrology
DRN	Digital River Network
GAM	Generalized Additive Model
GAMM	Generalized Additive Mixed effect Model
GIS	Geographical Information Systems
LCM	Land Cover Map
LIDAR	Light detection and ranging
LM	Linear Model
LMM	Linear Mixed effect Model
MdTw	Monthly mean daytime water temperature
RMSE	Root Mean Squared Error
RNS	River Network Smoother
SSN	Spatial Statistical Model
SWTA	Surface Water Temperature Archive (Environment Agency)
Ta	Temperature of the air at the land surface
Tw	Temperature of the water
UCR	Upper Critical Range
WIMS	Water quality data archive (Environment Agency)
WFD	Water Framework Directive

Appendix

Detailed methodology

All data analysis presented herein was conducted using R (R Core Team, 2021). In addition to the base functions in R the following packages were used: the 'tidyverse' ecosystem (Wickham et al., 2019), 'readxl' (Wickham and Bryan, 2023), 'lubridate' (Grolemund and Wickham, 2011), 'data.table' (Dowle and Srinivasan, 2023) and 'skimr' (Waring et al., 2022) for data import, manipulation, generating summaries and plotting data; 'car' (Fox and Weisberg, 2019) for correlation analysis; 'sp' (Pebesma and Bivand, 2005, Bivand et al., 2013) and 'sf' (Pebesma, 2018, Pebesma and Bivand, 2023) for spatial data analysis; 'lme4' (Bates et al., 2015), 'lmerTest' (Kuznetsova et al., 2017), 'MuMIn' (Barton, 2023), 'caret' (Kuhn, 2008), 'effects' (Fox and Weisberg, 2019) and 'sjstats' (Lüdecke, 2022) for developing mixed effects models and aiding model selection.

Data preparation

Observed water temperature data were acquired from 5 main sources: the Environment Agency's Surface Water Temperature Archive up to 2007 (Orr et al., 2015), Water Quality Data (WIMS; 2000 to present), water temperature data rescued from river flow monitoring stations, water quality monitoring studies using sonde probes, and ecological studies from partner organisations. Details regarding the creation of the former can be found in Orr et al., 2015 (and on data.gov.uk). It consists of 7 Microsoft Access databases, one containing data for each of the previous 7 Environment Agency regions in England. All water temperature data were filtered to select only those samples collected from "RIVER / RUNNING SURFACE WATER". These were then extracted and saved as CSV files for each region. Data on OpenWIMS is available from 2000. All samples between the years 2000 and 2023 were downloaded for the determinand "0076" (Temperature of Water) and filtered for those where the sampled Material Type was "RIVER / RUNNING SURFACE WATER" and saved as an additional CSV file. Data rescued from flow monitoring stations was provided by the Environment Agency as an R dataframe file (.rds type) or discrete comma delimited files (CSV) for each site, and sonde monitoring datasets downloaded as CSV files from the Epimorphics web site (https://hydrology-uat.epimorphics.net/hydrology/explore). Temperature datasets from partner organisation ecological studies were provided in either Microsoft Excel/xlsx or comma delimited/CSV formats (Lyd and Chess catchments).

The resultant dataset of 60,059,841 records comprised a mixture of spot sample and high frequency (sub hourly) water temperature readings often in a raw/unchecked state. The data for identifiable sites was passed through five quality-assurance (QA) 'filters' to identify and remove outliers and suspect data points:

1. Points outside a robust (Winsorized) mean ± 3 x standard deviations range were removed (flagged as 'Trimmed'; high frequency and spot sample data series)

2. Points exhibiting extreme daily temperature ranges (>10°C) were removed (flagged 'Air'; high frequency data series only)
3. Extreme deviations from a moving average were removed (flagged 'ExtremeSD'; high frequency data series only)
4. Non-changing sequences of values (> 1 day) were removed (flagged 'FixedRuns'; high frequency and spot sample data series)
5. Daily temperature fluctuations modelled to a sine function and fit assessed (data points with poor fit flagged 'PoorFit' rather than removed; high frequency data series only)

An example of output from this QA process is presented in Figure A1. This process was completed for each of the 38,142 identifiable sites with 6,327,910 data points being removed, leaving 53,731,931 to take forward in the study.

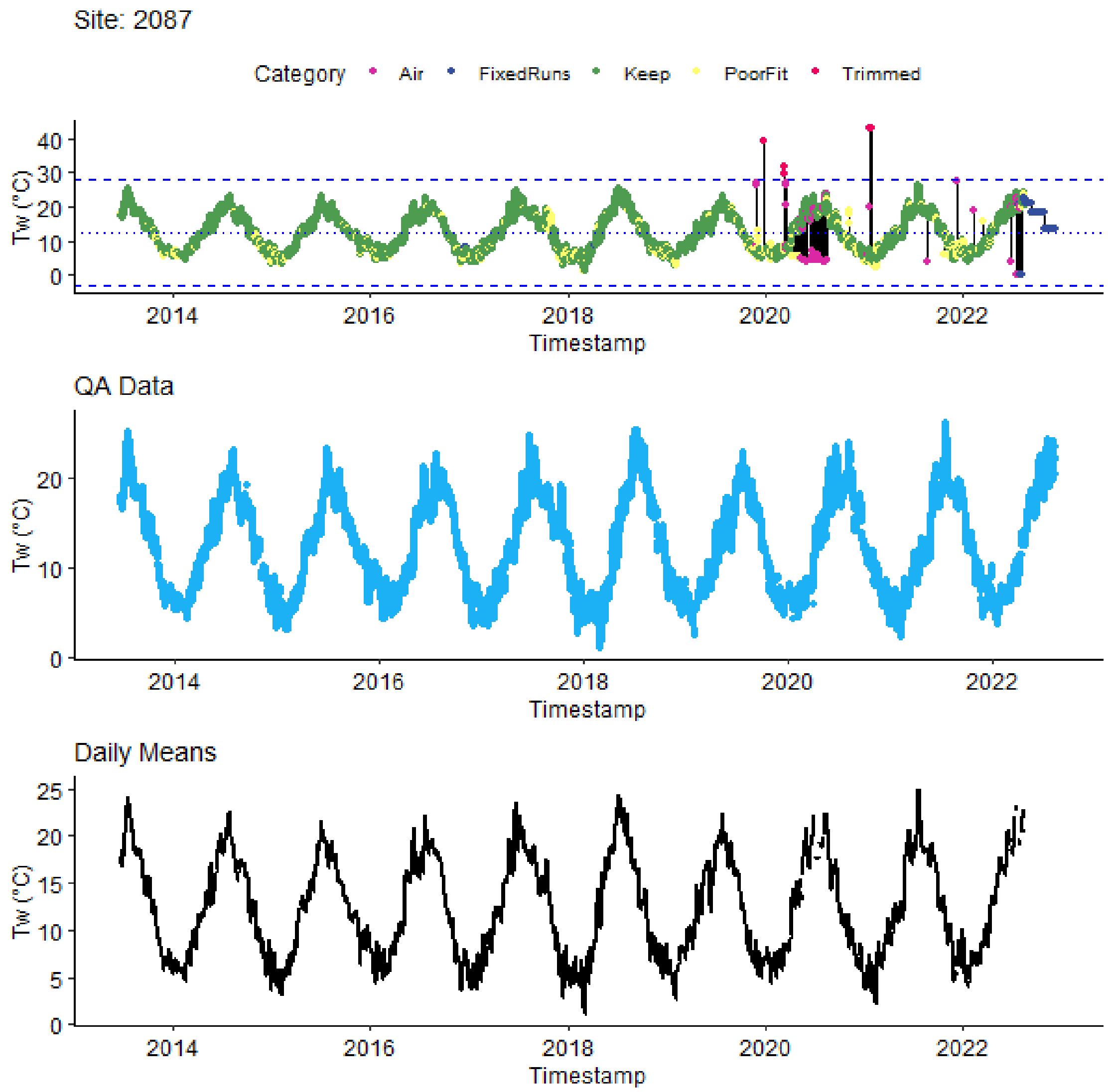


Figure A1: An example of the output from Quality-Assurance processes used to clean water temperature data. The top panel indicates some of the categories assigned to questionable data points, with the cleaned raw data presented in the middle panel. The lower panel show the resultant daily means generated from the raw data.

It should be noted that these QA filters did not capture all suspect data points and a further round of ‘Outlier checking’ was required before *Tw* modelling was undertaken.

Gridded air temperature values were downloaded from the Centre for Environmental Data Analysis (CEDA) website https://archive.ceda.ac.uk. Historical air temperature values were obtained from the HADUK 1km gridded datasets to cover the same date range as seen in the *Tw* samples (1952 to present day). Monthly mean air temperature (*tas*) datasets were downloaded in NetCDF format (one file per year). *‘tas’* datasets of monthly data were also obtained for the UKCP18 regional projections at 12km resolution for 1980-2080 for all RCP8.5 scenarios from the same site (NetCDF format).

R code data folder structure

```
<DataRoot folder> ─ External_data ─ CEDA ─ Air_temp ─ monthly ─ Ta_mean_1km (HADUK tas)
        │                 │               │
        │                 │               └─── ukcp18 ─ Ta_mean (UKCP18 tas)
        │                 │
        │                 ├──────── CEH_LCM_DRN ─ EA_GIS (Site DRN data)
        │                 ├──────── EA ─ Water_temperature data
        │                 │              ├────── Continuous_monitoring
        │                 │              ├────── epimorphics
        │                 │              ├────── GaugingStations
        │                 │              ├────── SWTA
        │                 │              └────── WQA
        │                 └──────── shapefiles
        │                                ├────── EA_Operational_Catchments (Eng. Regions)
        │                                ├────── EA_River_Water_Bodies (WFD catchments)
        │                                └────── England_Rivers (EA Main river map)
        │
        ├────────── Generated_data (R code generated output)
        │
        ├────────── England-R_Code (stage-based R scripts)
        │
        └────────── ENG_figures (folders containing R-generated figures)
```

Observed water temperature readings: sampling frequency considerations

While there was a deliberate effort to assemble higher frequency *Tw* data series for use in this study, the initial aim was to extend the previous ‘chalk stream’ study (Environment Agency, 2022) to the whole of England and, therefore, the monthly time step was retained for mean *Tw* generation. A comparison between the distribution of sites available for daily and monthly mean modelling is shown in figure A2. It should also be noted that data series duration was significantly more extensive for the monthly mean model development data (start February 1965, end April 2023, median time series duration 28.5 years) versus the equivalent daily data (start July 1986, end April 2023, median time series duration 5.4 years).

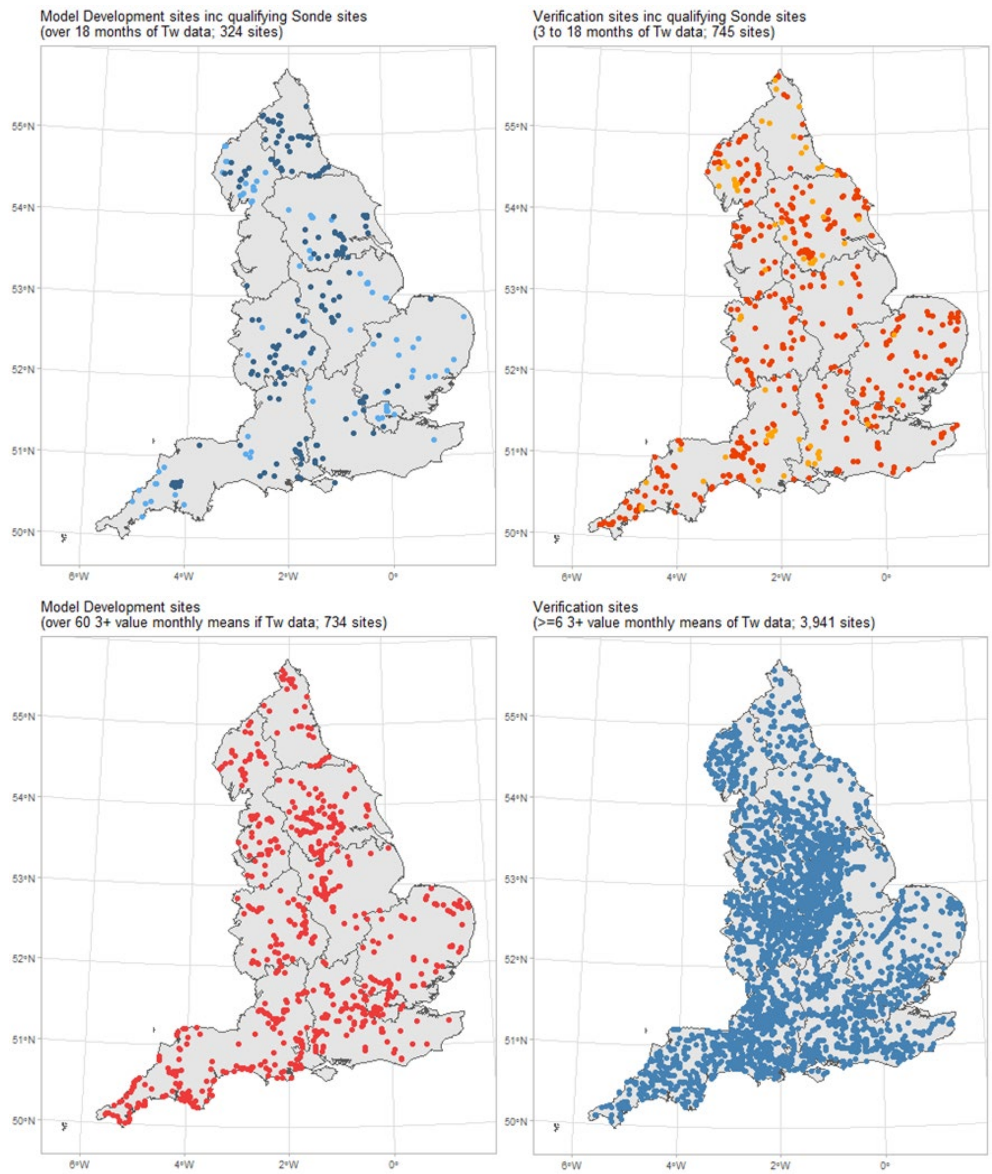


Figure A2: Sites available for modelling: Top panels show high frequency sites that could contribute to a daily mean *Tw* model (L: model development; R: model validation) and Lower panels show sites available from a monthly mean *Tw* model (L; model development; R: model validation)

Selection of water temperature model

Similar to the chalk stream study (Environment Agency, 20223), the river water temperature modelling process was guided by the previously established framework developed by Environment Agency (2021) and summarised in Figures A3 and A4. The first stage involved identification of the most suitable modelling approaches given the data sets available. An initial assessment highlighted that adequate data was not available to build a 'process-based' or 'hybrid' model to predict water temperature. A regression-based approach was favoured over a machine-learning model or black-box model as the ability to assess the physical plausibility of coefficients was deemed essential. The schematic in Figure A4 was adopted to identify the most suitable regression approach. A temporally dynamic model was required and given the data available at multiple sites, a mixed model was identified as the most suitable approach.

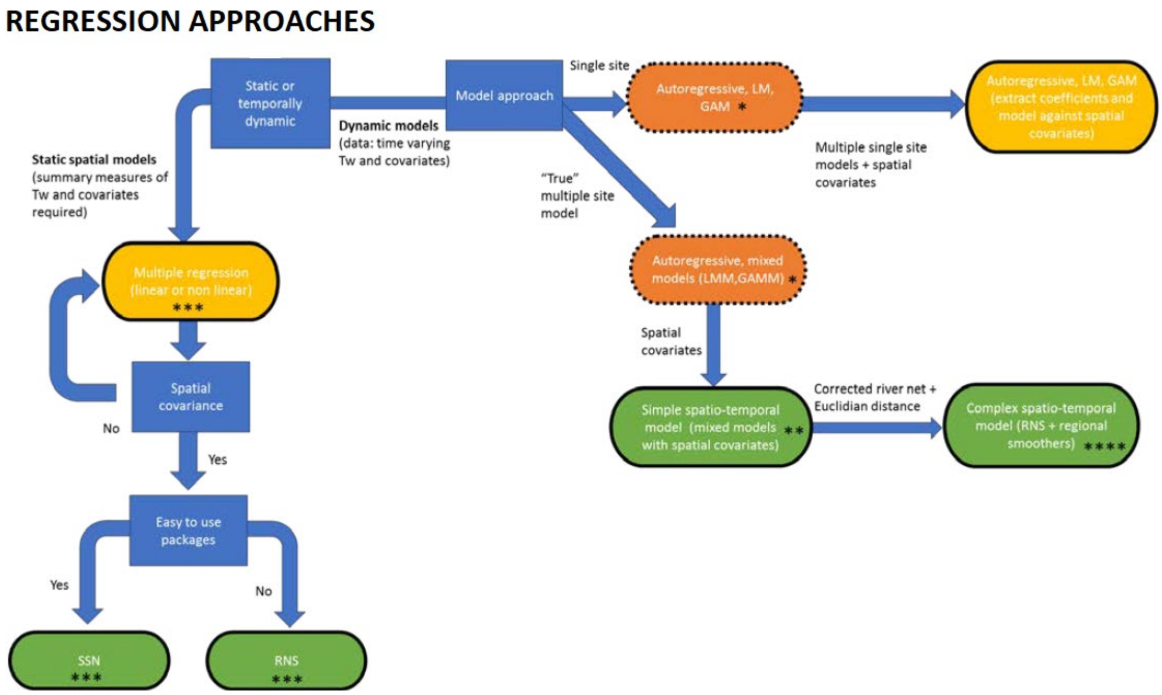


Figure A3: Decision tree for selecting the appropriate regression-based method for water temperature modelling. Taken from Environment Agency (2021)

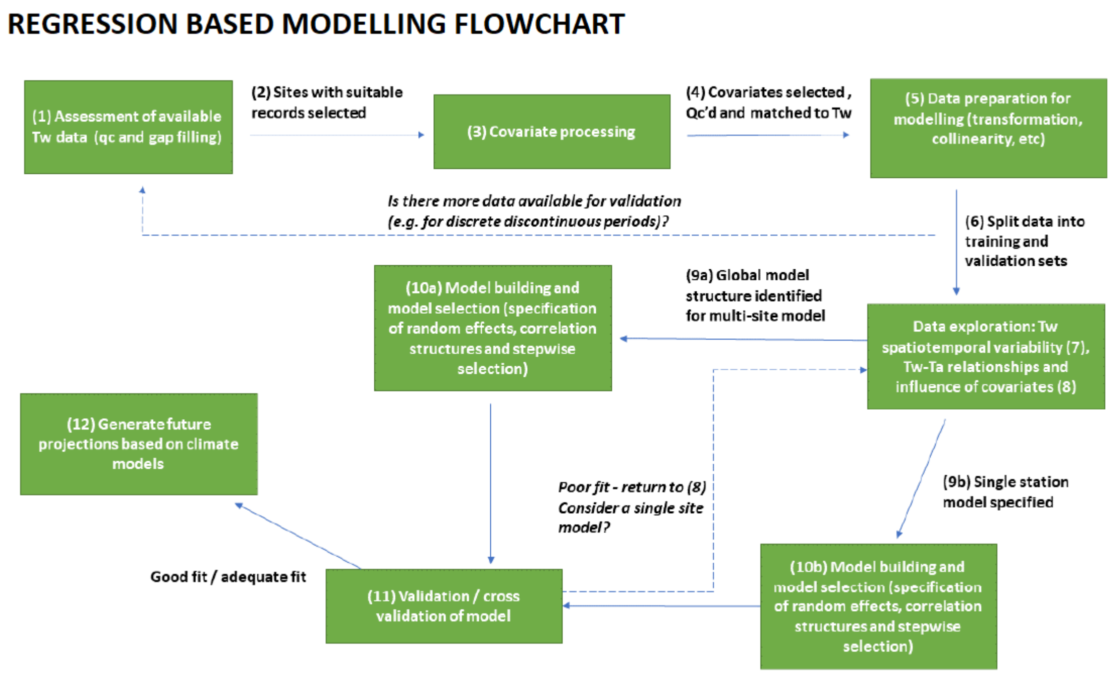


Figure A4: High-level schematic representation of the steps required when developing regression-based models for site specific water temperature predictions. Taken from Environment Agency (2021).

Model development and validation: Covariate correlation analysis

Relationships between covariates were assessed and those displaying strong correlations (Pearson correlation coefficient; r > 0.7) were rationalised (Figure A5). The site altitude covariates (Altitude_Map, Altitude_SWTA & DEM_Altitude) were rationalised to 'Altitude_Map', network complexity covariates (Shreve, us_accum, wet_are_us, dist_src and catchment_sqm) rationalised to 'Shreve', and the catchment altitude covariates (AVE_CATCH_Altitude and src_alt) rationalised to 'AVE_CATCH_Altitude. Covariates with significant numbers of missing values were also excluded (seg_length, seg_width and seg_gradient). Hard geology and land-use variables with limited/low ranges (Geo_Unspecified_PC, Geo_salt_PC, Geo_Peat_PC, LCM_Unspecified_PC and LCM_Other_PC) were also dropped. This left 18 covariates plus 2 spatial coordinates (Easting and Northing) to take forward to the modelling stage.

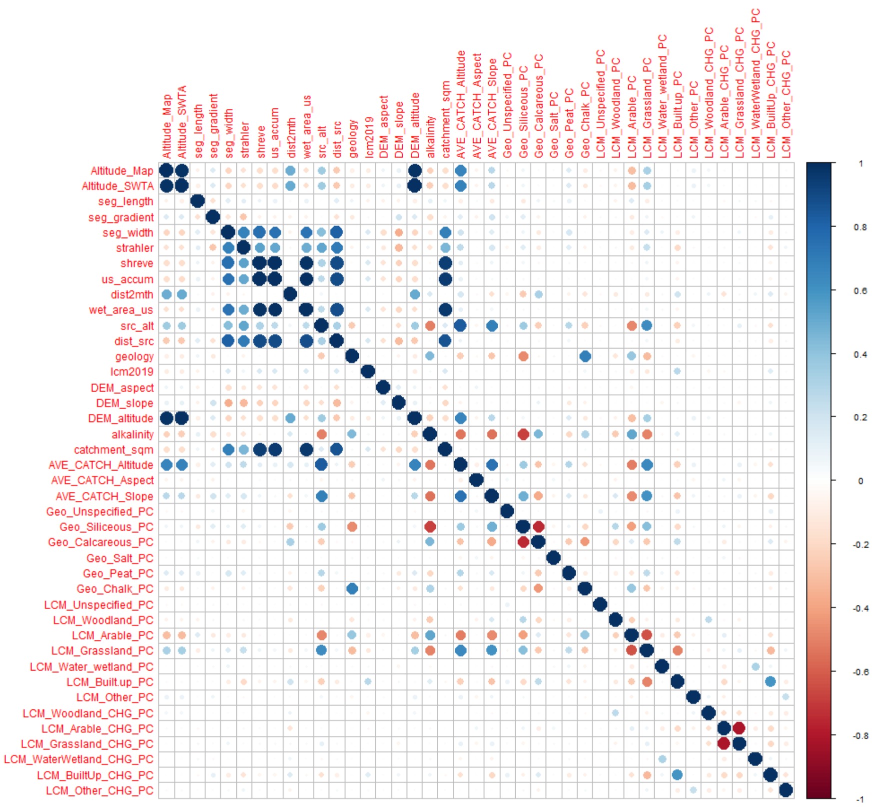
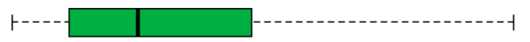
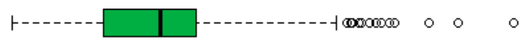
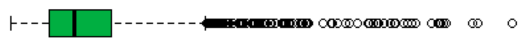
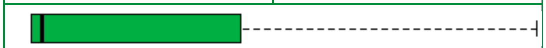
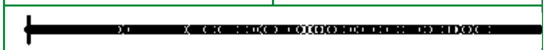
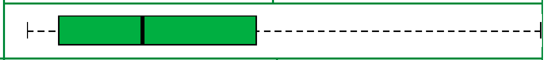
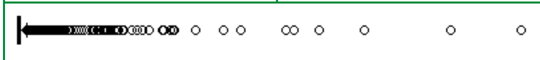


Figure A5: Correlation matrix of covariates indicating Pearson correlation coefficients. Blue circles indicate positive correlations and red circles indicate negative correlations. Circle size is proportional to the Pearson correlation coefficient.

Table A6: Covariates selected for water temperature model generation. Boxplots indicate the median, interquartile range, and distribution of the variables within the monitoring sites

Identifier	Description	Minimum	Maximum
Altitude_Map	Altitude at site (m)	-9.21	535.23
			
Strahler	The Strahler river order at site (indicator of river size)	1	9
			
Shreve	The Shreve river order at site (indicator of network complexity)	1	26,860
			
Dist2mth	Distance along the main river channel to the river mouth (m)	2.35	359,185
			
DEM_aspect	Aspect of river section at site (0-359°)	0	359
			
DEM_slope	Gradient of river section at site (m/km)	0	54
			
Alkalinity	Acid Buffering capacity at site (indication of upstream calcareous geology) (mg/l)	0.18	597.4
			
AVE_CATCH_Altitude	Average altitude of upstream catchment (m)	-0.12	655.8
			
AVE_CATCH_Aspect	Average aspect of upstream catchment (0-359°)	24	341
			
AVE_CATCH_Slope	Average gradient of upstream catchment	0.1	23.5
			
Geo_Siliceous_PC	Upstream siliceous hard geology (%)	0	100
			
Geo_Calcareous_PC	Upstream calcareous hard geology (%)	0	100
			
Geo_Chalk_PC	Upstream chalk hard geology (also calcareous; %)	0	100
			
LCM_Woodland_PC	Upstream woodland (%)	0	100
			

Identifier	Description	Minimum	Maximum
LCM_Arable_PC	Upstream arable farmland (%)	0	100
			
LCM_Grassland_PC	Upstream grassland (improved and undisturbed/natural; %)	0	100
			
LCM_Water_wetland_PC	Upstream area containing water/wetland, e.g., pools, reservoirs, etc. (%)	0	83
			
LCM_Built.up_PC	Upstream land dedicated to urban and suburban development (%)	0	100
			
Easting / Northing	Spatial coordinates		

Review of outliers

Additional outliers that had survived the QA process were identified while model preparation was underway. An iterative process was undertaken to identify the nature of these outliers and assess whether it was justifiable to remove some or all of them.

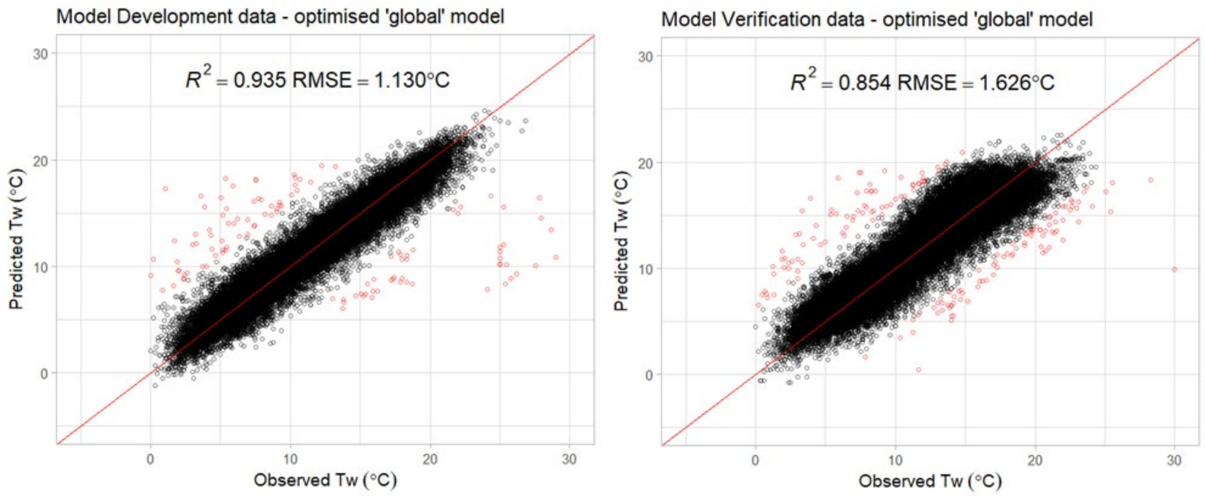


Figure A7: Potential outliers (red) in the model development (left) and verification (right) datasets.

Two sites were completely removed from the verification dataset (W4007 and 26M06) and thirteen sites were identified that had sequences of spurious Tw values (example in Figure A8) which were selectively deleted. Individual readings at fourteen sites were removed and 241 means with high standard deviations (>6) were taken out.

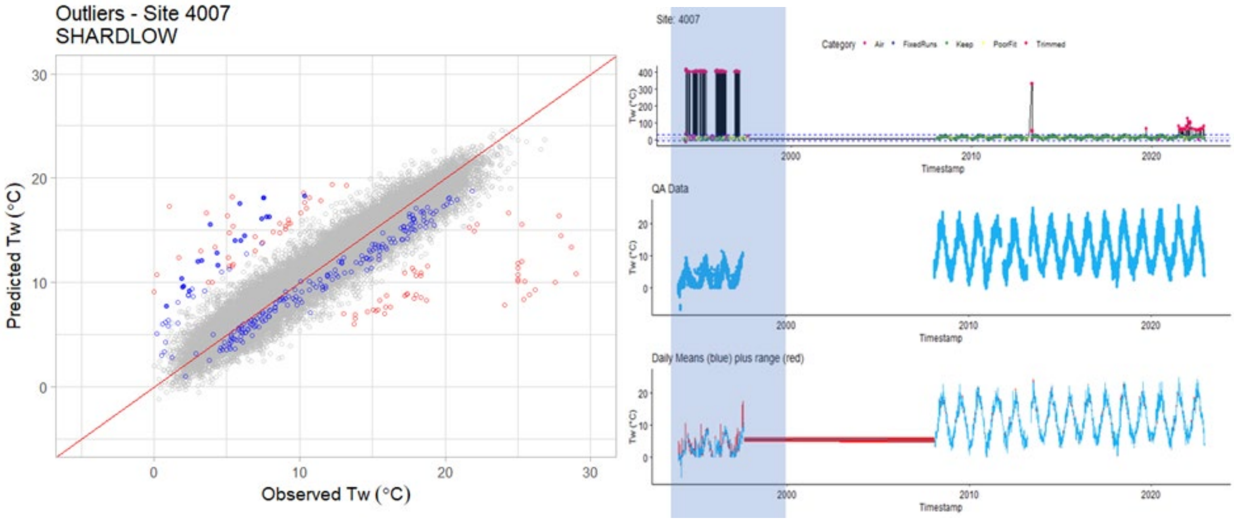


Figure A8: Outliers at Site 4007 / Shardlow. The QA plot (right, blue highlight) indicates a sequence of readings that are significantly dissimilar to the rest of the time series and constitute all the filled blue outliers in the left panel. These values were removed from the Tw dataset ahead of the modelling process.

Global model temperature predictions

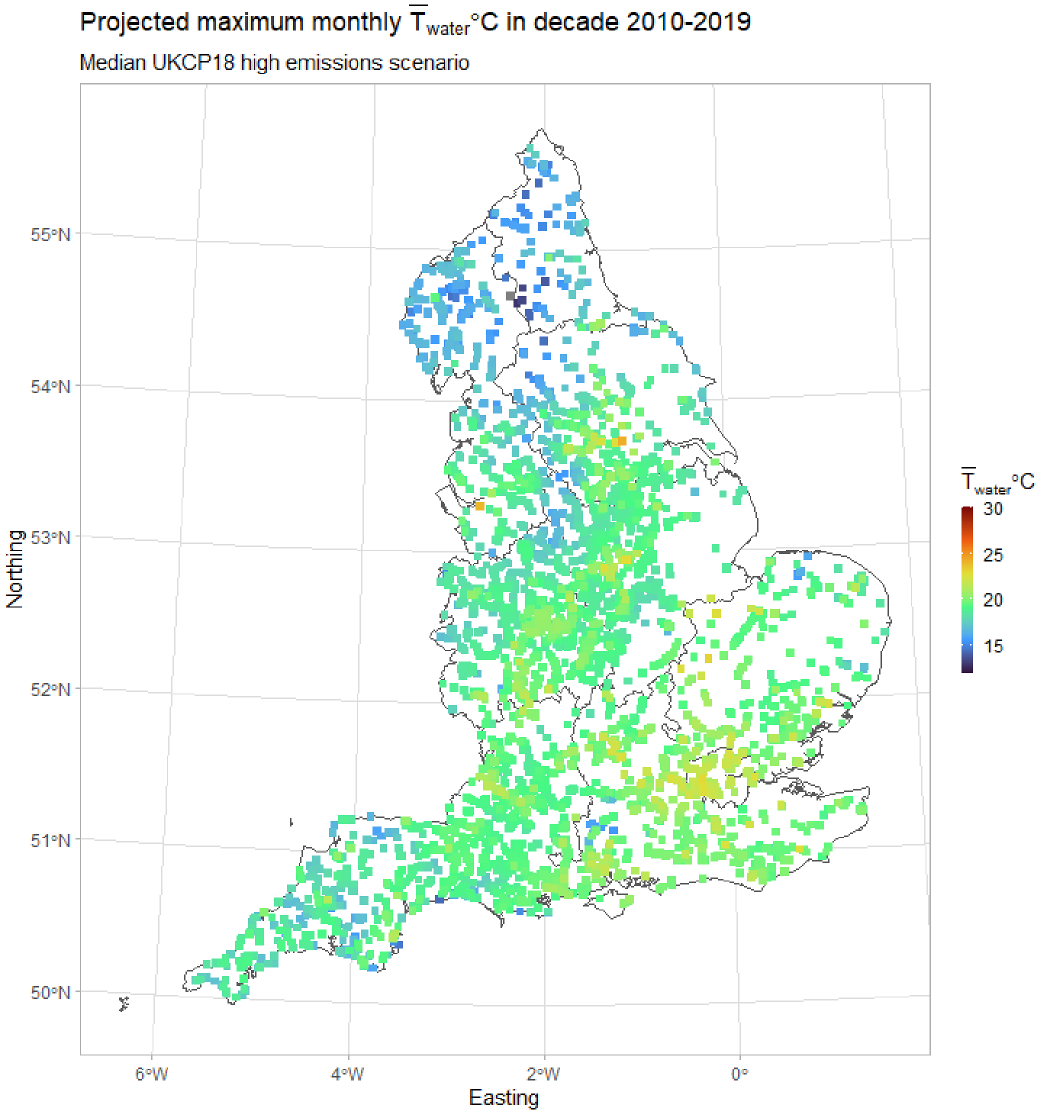


Figure A9(i): Enlarged version of Figure 9, top-left panel showing the maximum *MdTw* projected at each qualifying site in the decade 2010 to 2019.

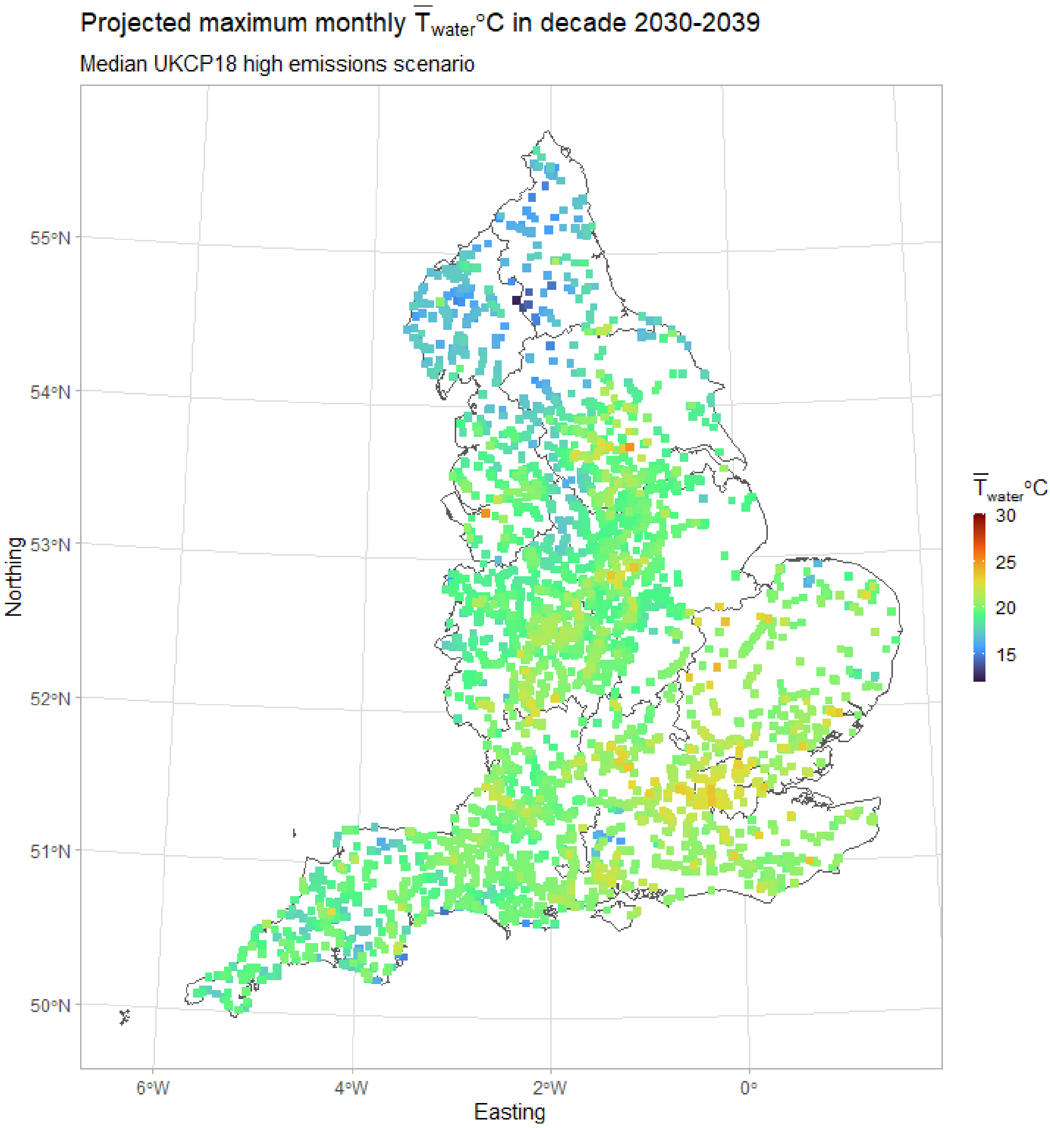


Figure A9(ii): Enlarged version of Figure 9, top-right panel showing the maximum *MdTw* projected at each qualifying site in the decade 2030 to 2039.

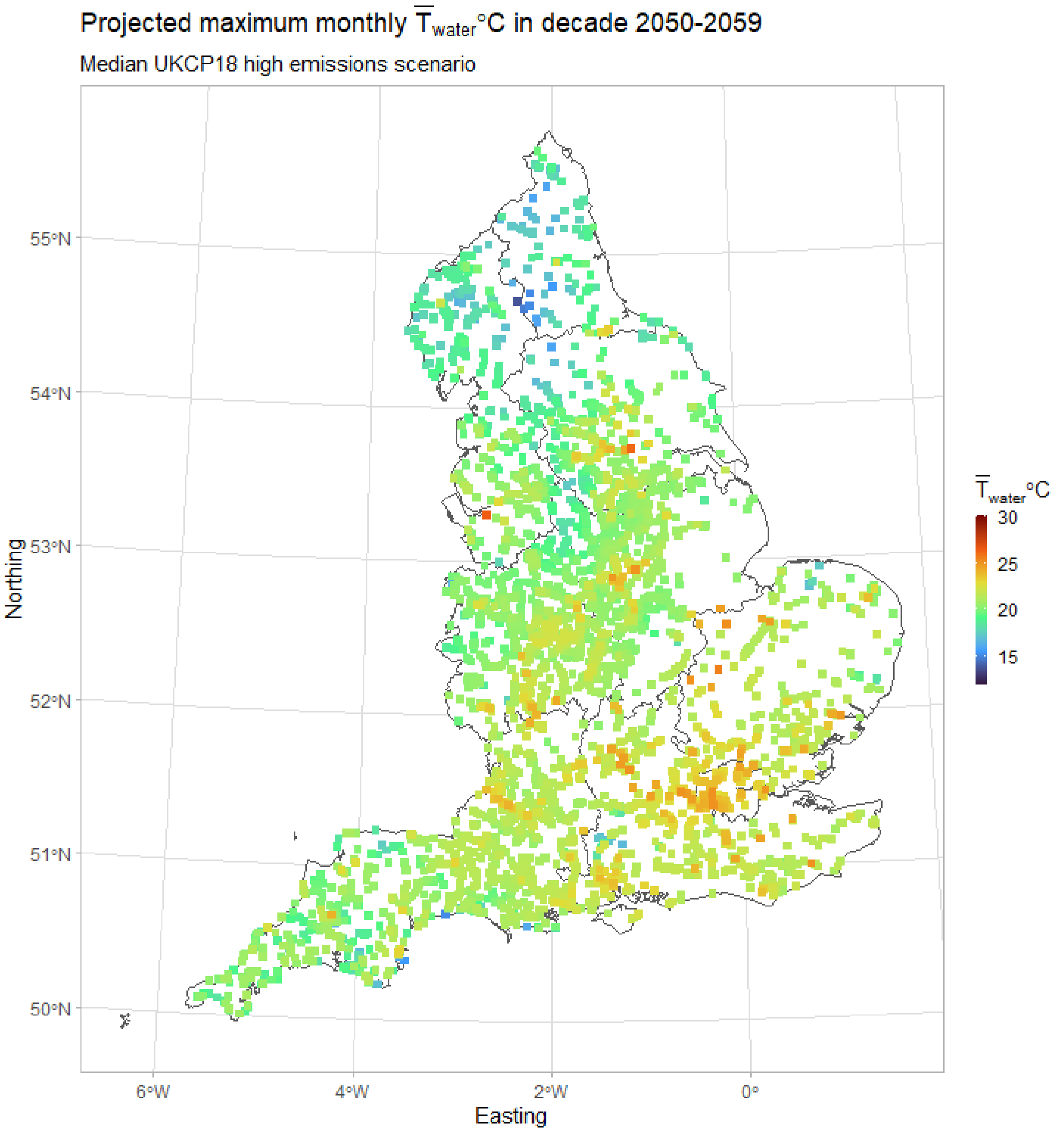


Figure A9(iii): Enlarged version of Figure 9, bottom-left panel showing the maximum *MdTw* projected at each qualifying site in the decade 2050 to 2059.

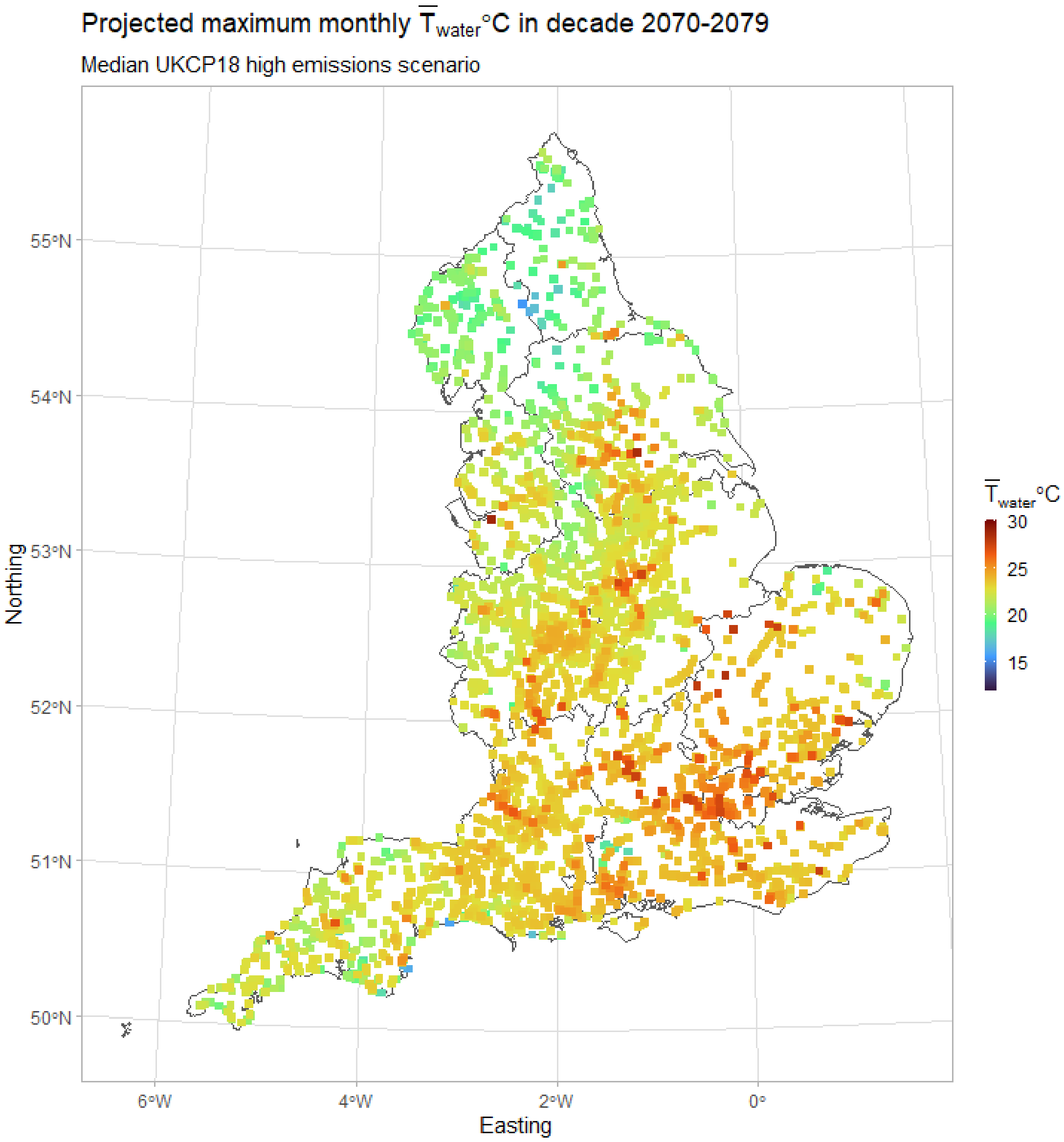


Figure A9(iv): Enlarged version of Figure 9, bottom-right panel showing the maximum *MdTw* projected at each qualifying site in the decade 2070 to 2079.

Global model temperature predictions: differences from 1981-2025 baseline

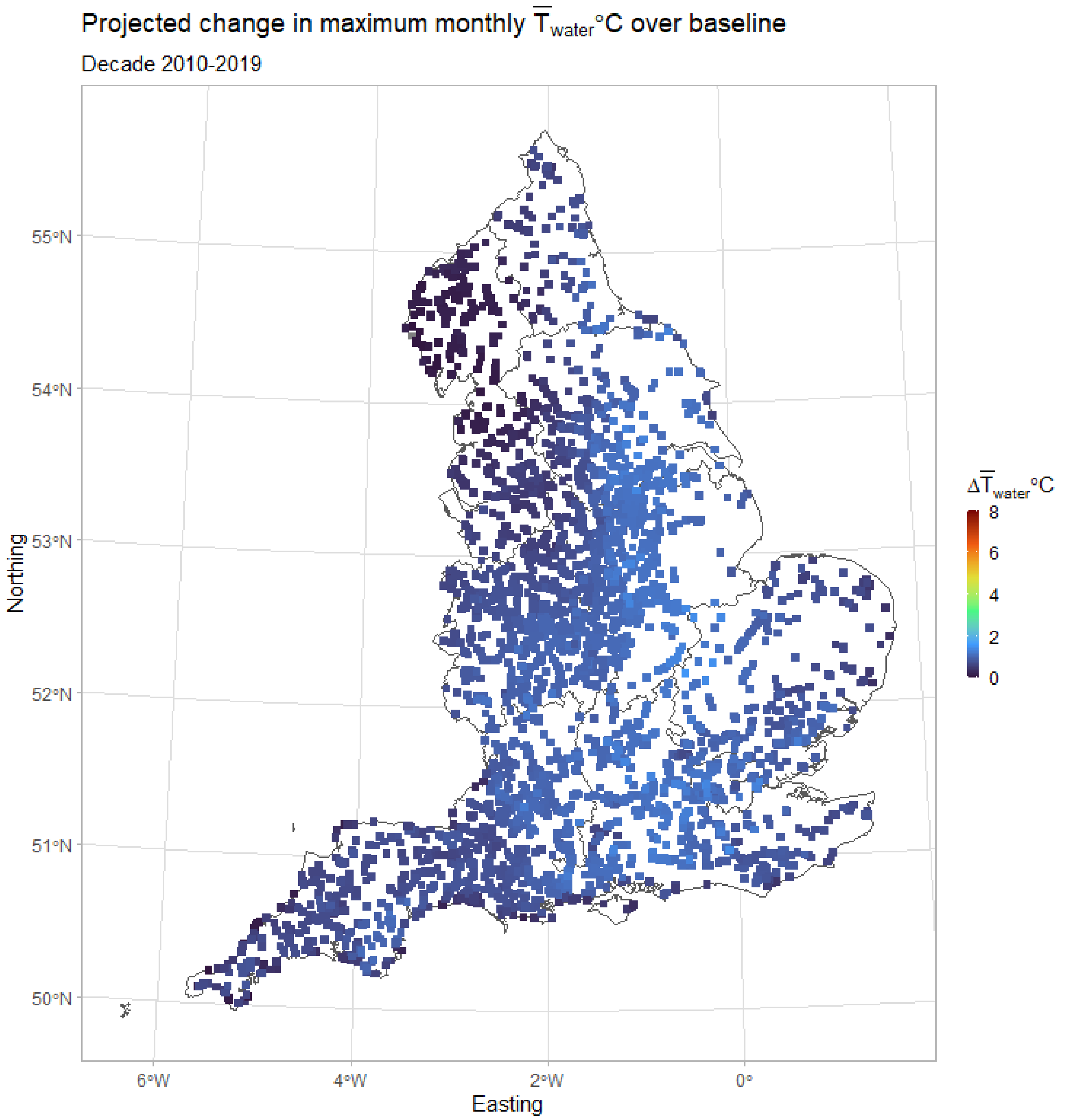


Figure A10(i): Enlarged version of Figure 10, top-left panel showing the change in maximum *MdTw* projected at each qualifying site over baseline (1981 to 2005 inclusive) levels in the decade 2010 to 2019.

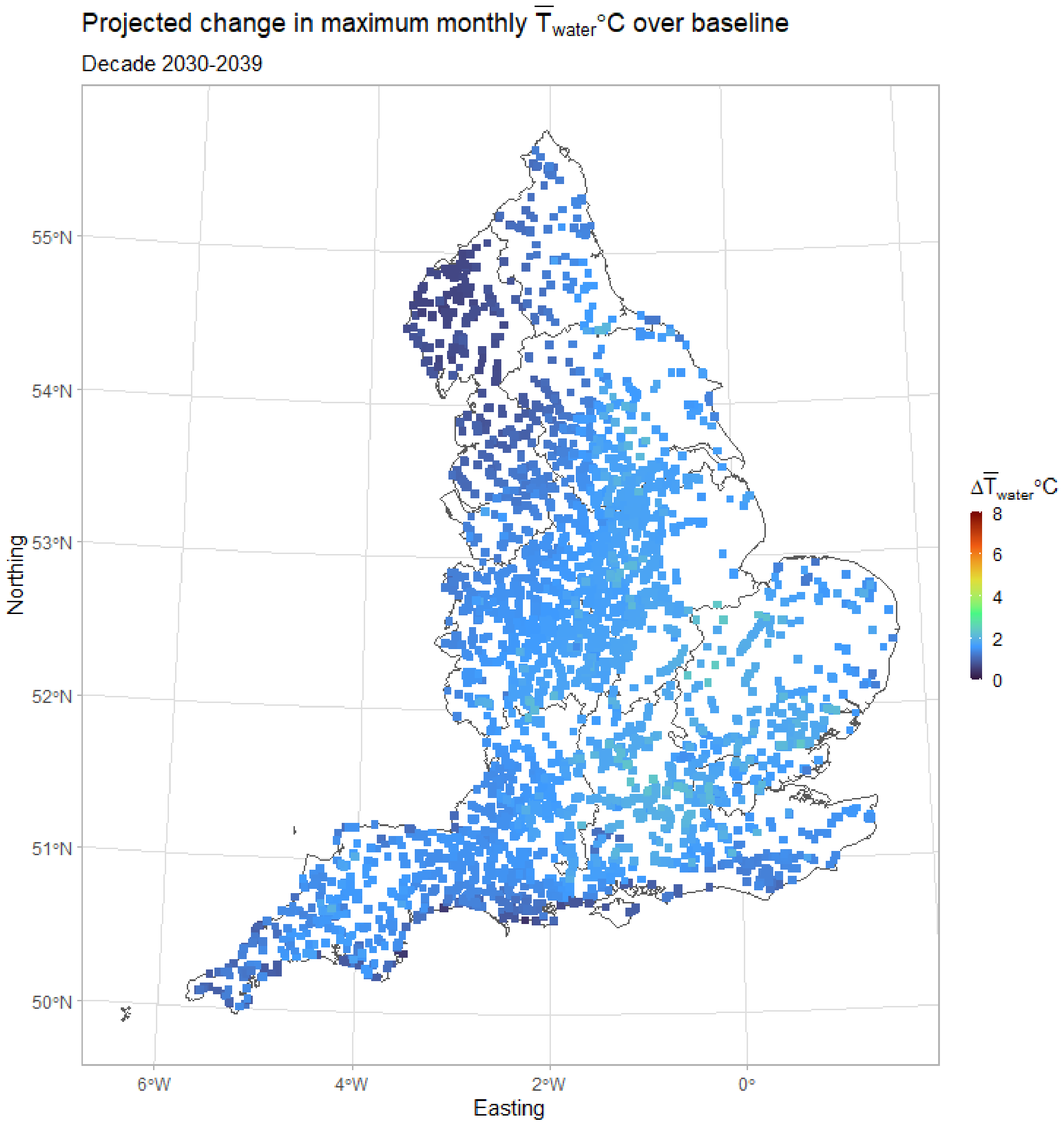


Figure A10(ii): Enlarged version of Figure 10, top-right panel showing the change in maximum *MdTw* projected at each qualifying site over baseline (1981 to 2005 inclusive) levels in the decade 2030 to 2039.

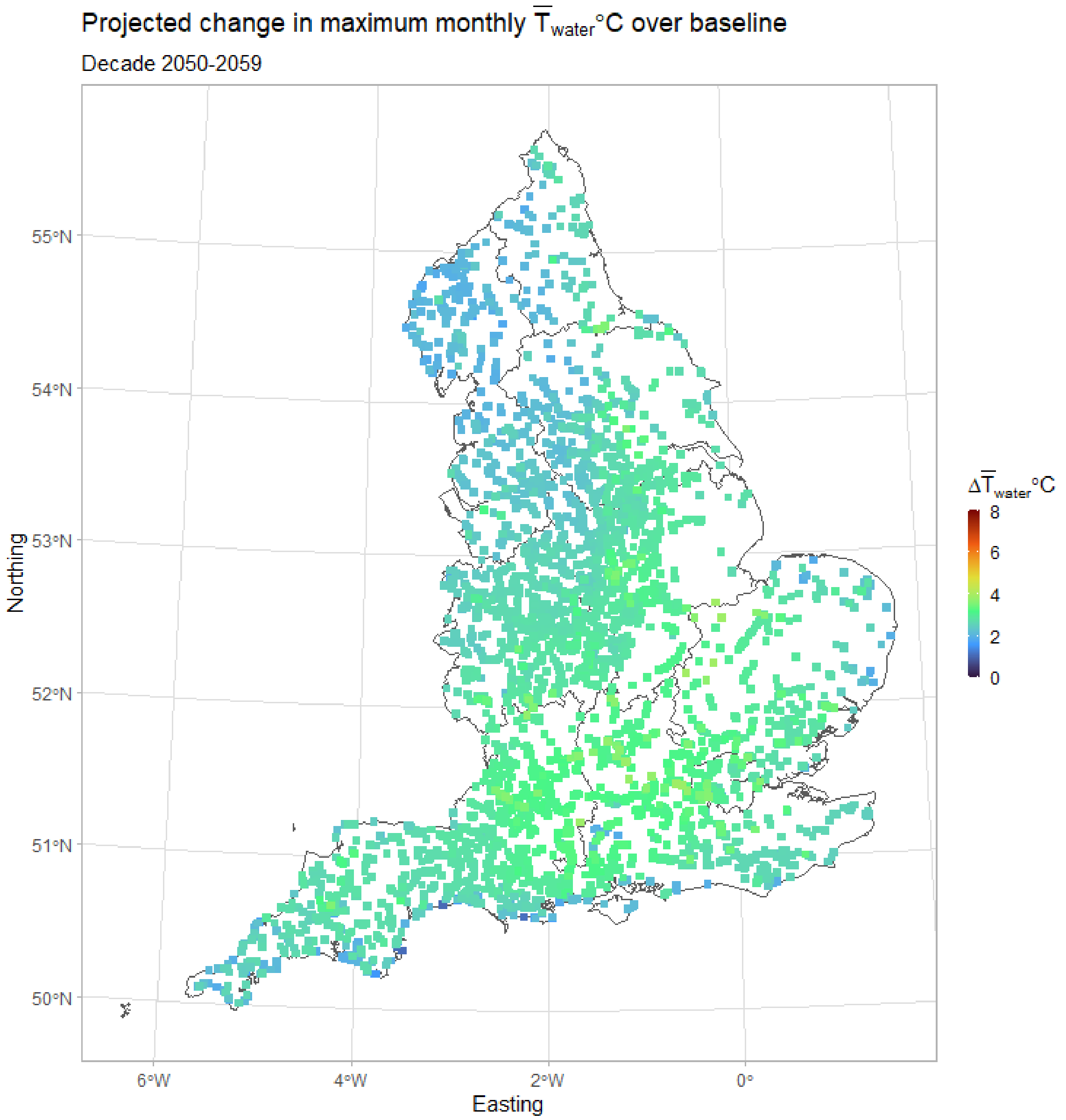


Figure A10(iii): Enlarged version of Figure 10, bottom-left panel showing the change in maximum *MdTw* projected at each qualifying site over baseline (1981 to 2005 inclusive) levels in the decade 2050 to 2059.

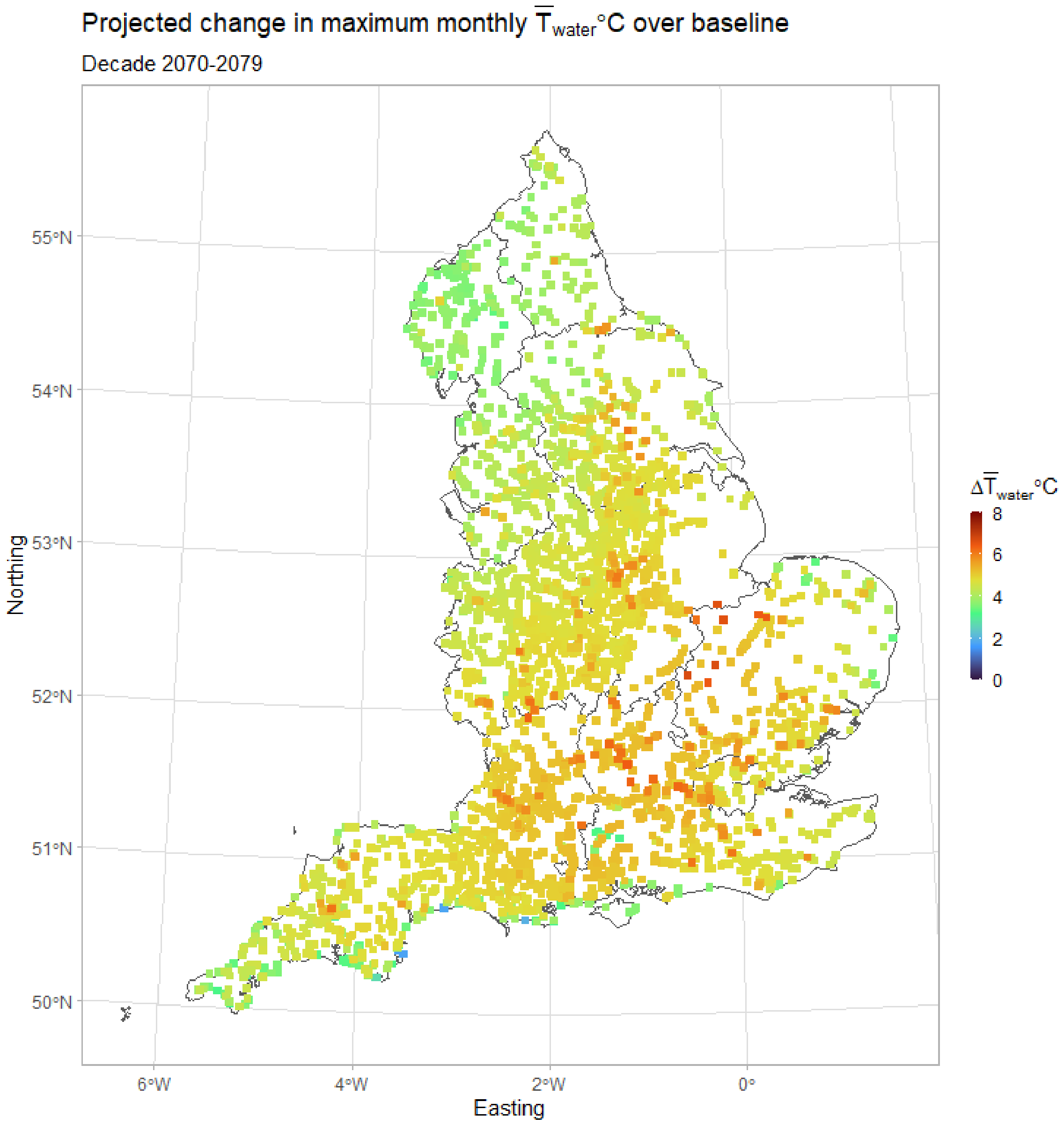


Figure A10(iv): Enlarged version of Figure 10, bottom-right panel showing the change in maximum *MdTw* projected at each qualifying site over baseline (1981 to 2005 inclusive) levels in the decade 2070 to 2079.

Typology specific models

Typological clusters were identified (Ward's method) based on site covariate values. An average Jaccard value (values between 0 and 1 with 1 being most robust) and Instability Index (values between 0 and 1, with 0 being most stable) were used to identify coherent clusters and the pivot point fell in the region of 6 or 7 clusters. Cluster group 4 from the '6 cluster' analysis (Grass Upland, mixed geology) split into two when generating 7 clusters, with a 'higher altitude' and a 'lower altitude' grouping. The Instability Index for this 'higher altitude' grouping rose to 0.93 and little value was seen in maintaining this split, thus arriving at 6 cluster groups. The 'full' model was applied to each of these clusters and resulting models optimised to identify influential covariates. The projected *Tw* values obtained from these cluster models was compared to those obtained from the optimised 'global' model and no significant difference could be seen ($R^2 = [0.9989,1]$).

Table A11: Cluster model characteristics and comparison to the 'global' model

	Global Model	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Typology	All England	Permeable rock, arable / grass	Harder rock, grass / arable	Urban, gen. permeable	Grass upland	Chalk streams	Large rivers
Top model	Site altitude Dist to mouth Calc Geol % Urbanised % Northing	Shreve Urbanised % Northing	Chalk % Arable % Urbanised % Northing	Northing	Shreve Urbanised %	Alkalinity Dist to mouth Arable % Grassland % Northing	Ave. Slope Calc Geol % Silic Geol % Chalk % Arable % Grassland % Woodland % Urbanised %
Model averaged (δ<=2)	Site altitude -Dist to mouth Calc Geol % Urbanised % Northing	- Shreve Urbanised % Northing	- Chalk % Arable % Urbanised % Northing	Northing	Shreve Urbanised %	-Alkalinity -Dist to mouth -Arable % -Grassland % -Northing	Ave. Slope Calc Geol % Silic Geol % Chalk % Arable % Grassland % Woodland % Urbanised %
R^2 / RMSE **Model dev Validation**	0.935 / 1.13 0.854 / 1.62	0.938 / 1.13 0.858 / 1.63	0.941 / 0.95 0.864 / 1.46	0.918 / 1.25 0.828 / 1.79	0.927 / 1.24 0.875 / 1.50	0.948 / 0.89 0.830 / 1.58	0.921 / 1.44 0.726 / 2.87
Global comparison	--- / --- --- / ---	0.938 / 1.13 0.860 / 1.61	0.939 / 0.95 0.865 / 1.52	0.918 / 1.25 0.826 / 1.89	0.927 / 1.24 0.869 / 1.49	0.948 / 0.89 0.849 / 1.60	0.921 / 1.45 0.855 / 2.07
Fitted values R^2	1.0000	0.9999	0.9989	0.9993	0.9998	1.0000	0.9995

Fish temperature thresholds

Table A12: Thermal thresholds for *Salmo salar* (Atlantic salmon) and *Salmo trutta* (brown trout), from Solomon and Lightfoot, 2008 and Elliott and Elliott, 2010. All values in degrees Celsius. The 'Stress Zone' for egg survival correlates with increased mortality and deformity of resultant fry.

	Egg survival		**Feeding / growth range**		**Adult survival critical range**	
	S. salar	*S. trutta*	*S. salar*	*S. trutta*	*S. salar*	*S. trutta*
Low limit	0	0	6	3.5	2.5	0.7
Optimal	4 to 7	1 to 8	15.9	13.1	2.5 to 22	2.5 to 22
Stress zone	12	12	--	--	24-26	24
High limit	16	13	22.5	19.5	28	25

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.