EcoSHEDS Northeast Stream Temperature Model

Jeff Walker, Ben Letcher, and Dan Hocking

v1.4.0 (Jan 3, 2024)

Contents

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.0
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts
```

```
## Warning in readRenviron(version_file): file '../version.sh' cannot be opened
## for reading
```

Chapter 1

Introduction

The SHEDS stream temperature model was developed to predict daily stream temperatures at both gaged and un-gaged catchments across the northeast U.S. based on geospatial characteristics and weather conditions.

The model is based on a linear mixed effects framework that accounts for spatial and temporal correlations using a hierachical Bayesian structure. Letcher et al. (2016) describe the initial development of this model framework, and apply it to a small region in western Massachusetts. See the Model Overview section below for a brief introduction to the model, or the Theory section a more detailed explanation.

The documentation is divided into the following sections:

1. Introduction : provides an overview the stream temperature model and documentation, as well as a snapshot of the current calibration
2. Theory : describes how the model works including the underlying structure and theory
3. Data Sources : describes the datasets used as inputs to the model
4. Data Processing : describes how input datasets are processed prior to model fitting (i.e. QAQC procedures)
5. Calibration and Validation : describes how well the model predicts stream temperature based on observations that were included (calibration) and excluded (validation) from the model fitting process
6. Predictions : describes how predictions are generated after the model is calibrated and describes the various summary metrics that are computed for each catchment
7. Download : provides links to download the model predictions, catchment delineation (shapefiles), and covariates dataset.

The model will be periodically updated and re-calibrated (approximately once every 6 months) to incorporate any new temperature observations, and to make any necessary revisions to the data processing and/or model structure. With each update, a new version will be assigned to the model, and this documentation website will be updated to reflect the most recent performance of the model. A brief summary of the changes associated with each new version is provided in the Change Log below.

1.1 Model Overview

The SHEDS stream temperature model predicts daily mean stream temperature based on geospatial characteristics and daily weather conditions. Predictions are made at a relatively fine spatial resolution based on a customized catchment delineation (average size about 2 km^2) across the northeast U.S. from Maine to Virginia. Predictions are limited to streams and rivers with drainage areas less than 200 km^2. Heavily impounded rivers are also excluded from the model.

The model uses a hierarchical Bayesian structure to account for spatial correlation in temperature between near-by locations through random effects for both the individual catchments and the larger watershed (HUC8) containing that catchment. Therefore, catchments within the same HUC8 watershed share a set of HUC-specific coefficients. Year to year variations in temperature are also accounted for using a random effect for the year.

1.2 Current Snapshot

Table 1.1 provides a snapshot of the calibration and validation performance for the current version of the model (v1.4.0). More details about the model performance can be found in the Calibration and Validation section.

1.3 Model Versioning

The SHEDS stream temperature model uses semantic versioning of the form: `vX.Y.Z`

- `X` is the **major** version, which will be incremented when there is a major change to the model theory, code, or datasets.
- `Y` is the **minor** version, which will be incremented when there is a new set of results due to changes in the model code, datasets, processing procedures, etc.
- `Z` is the **patch** version, which will be incremented only when there is a change to the documentation or code that *does not* yield different results.

Table 1.1: Summary statistics of model calibration and validation

	Calibration	Validation
# Daily Observations	705,509	80,035
# Time Series	7,814	863
# Catchments	2,971	566
# HUC8s	144	109
# Years	31	27
RMSE (degC)	1.061	1.342
Mean Residual (degC)	0.061	0.079
Median Residual (degC)	0.075	0.062
Mean Absolute Residual (degC)	0.813	1.021
Median Absolute Residual (degC)	0.649	0.805
Minimum Residual (degC)	-7.570	-7.665
1st Percentile Residual (degC)	-2.674	-3.225
99th Percentile Residual (degC)	2.659	3.697
Maximum Residual (degC)	9.196	7.411

1.4 Source Code

The source code for the model itself and this documentation is available in the Github repository ecosheds/northeast-temp-model. Each version of the model will be included under the list of Releases.

1.5 Change Log

- **v1.4.0 (Jan 3, 2024)**
  Add 2021-2022 daymet, add stream temperature data from USGS NWIS, re-calibrate full model.
- **v1.3.0 (Sep 7, 2021)**
  Add 2020 daymet, revise data screening procedures, re-calibrate full model.
- **v1.2.0 (Jul 10, 2020)**
  Add 2019 daymet, re-calibrate full model.
- **v1.1.1 (Jan 16, 2020)**
  Add air temperature scenarios (+2, +4, +6 degC), and # days >= 24.9 and 27 degC to prediction derived metrics.
- **v1.1.0 (Dec 2, 2019)**
  Add 2018 daymet data, re-calibrated full model.
- **v1.0.2 (Mar 26, 2019)**
  Update documentation, add Download section containing links to model predictions, catchment delineation, and covariates.
- **v1.0.1 (Dec 10, 2018)**
  Update documentation, remove auto-correlation term from goodness-of-fit

summaries

- **v1.0.0 (Oct 25, 2018)**
  Re-calibrated model, and finished documentation.
- **v0.9.2 (Jul 6, 2018)**
  Major updates to documentation (introduction, theory, calibration and validation sections).
- **v0.9.1 (Jun 6, 2018)**
  Updates to model versioning and configuration framework, and minor updates to documentation (still incomplete).
- **v0.9.0 (May 30, 2018)**
  Preliminary release of the new model framework and documentation.
- **Previous Versions (prior to 2018)**
  Previous versions of the stream temperature model can be found here. That website is now deprecated, but will remain available for future reference. Beginning with v1.0.0 of the new framework and codebase, all model changes and results will be tracked and made available.

Chapter 2

Theory

The stream temperature model is a nested hierarchical Bayesian model that predicts daily stream temperature based on catchment characteristics and climate conditions. An early version of this model can be found in Letcher et al. (2016).

Daily mean stream temperature for each catchment is assumed to be a normally distributed random variable:

$$t_{h,c,y,d} \sim \mathcal{N}(\mu_{h,c,y,d}, \sigma_{[t]})$$

where $t_{h,c,y,d}$ is the mean stream temperature on day d within year y for catchment c, which is located within HUC8 h. This random variable is normally distributed with an expected mean $\mu_{h,c,y,d}$ and standard deviation $\sigma_{[t]}$.

The expected mean is computed as:

$$\mu_{h,c,y,d} = \begin{cases} \omega_{h,c,y,d} + \delta_h(t_{h,c,y,d-1} - \omega_{h,c,y,d-1}) & \text{for } t_{h,c,y,d-1} \text{ is real} \\ \omega_{h,c,y,d} & \text{for } t_{h,c,y,d-1} \text{ is not real} \end{cases}$$

where δ_h is an autoregressive [AR(1)] coefficient and $\omega_{h,c,y,d}$ is the expected temperature before accounting for temporal autocorrelation in the error structure.

The expected temperature is computed as a linear equation with four sets of terms:

$$\omega_{h,c,y,d} = X_{[0]}B_{[0]} + X_{h,c}B_{h,c} + X_hB_h + X_yB_y$$

where

- $B_{[0]}$ is a vector of fixed effect coefficients

- $B_{h,c}$ is a vector of random effect coefficients for catchment c
- B_h is a vector of random effect coefficients for HUC h
- B_y is a vector of random effect coefficients for year y

Each of these vectors is multiplied by a corresponding matrix containing the corresponding predictor values (X) of each catchment c (located within HUC h) and on each day d (within year y).

2.1 Fixed Effects

The fixed effects are shared among all catchments within the model domain. They include the following terms:

Variable	Description
`intercept`	Intercept
`AreaSqKM`	Total Drainage Area (km2)
`impoundArea`	Impounded Drainage Area (km2)
`agriculture`	Agricultural Land Cover (%)
`devel_hi`	High Development Land Cover (%)
`forest`	Riparian (200 ft Buffer) Forest Cover (%)
`prcp2`	2-day Precipitation (mm)
`prcp30`	30-day Precipitation (mm)

The fixed effects also include the following interaction terms.

Interaction Term	Description
`prcp2.da`	2-day Precipitaation x Drainage Area
`prcp30.da`	30-day Precipitaation x Drainage Area
`airTemp.da`	Air Temperature x Total Drainage Area
`airTemp.impoundArea`	Air Temperature x Impounded Drainage Area
`airTemp.agriculture`	Air Temperature x Agricultural Land Cover
`airTemp.forest`	Air Temperature x Riparian (200 ft Buffer) Forest Cover
`airTemp.devel_hi`	Air Temperature x High Development Land Cover
`airTemp.prcp2`	Air Temperature x 2-day Precipitation
`airTemp.prcp30`	Air Temperature x 30-day Precipitation
`airTemp.prcp2.da`	Air Temperature x 2-day Precipitation x Drainage Area
`airTemp.prcp30.da`	Air Temperature x 30-day Precipitation x Drainage Area

2.2 Catchment Random Effects

The random effects for each catchment (c) include the following variables:

Variable	Description
`intercept`	Intercept
`airTemp`	Air Temperature (degC)
`temp7p`	7-day Mean Air Temperature (degC)

2.3 HUC Random Effects

The random effects for each HUC (h) include the following variables:

Variable	Description
`intercept`	Intercept
`airTemp`	Air Temperature (degC)
`temp7p`	7-day Mean Air Temperature (degC)

2.4 Year Random Effects

The random effects for each year (y) include the following variables:

Variable	Description
`intercept`	Intercept

Chapter 3

Data Sources

3.1 Covariate Datasets

3.1.1 Catchment Delineation

The stream temperature model uses the National Hydrography Dataset High Resolution Delineation Version 2 (NHDHRDV2) catchment delineation.

Shapefiles of the catchments can be downloaded from: https://ecosheds.org/assets/nhdhrd/v2/

The `spatial_XX.zip` files contain pre-staged datasets for each HUC2 region (e.g. spatial_01.zip contains the catchments for the HUC2 region 01).

3.1.2 Climate Data

Daily air temperature and precipitation were obtained from the Daymet V3 gridded estimates of daily weather parameters for North America. The raw Daymet datasets (NetCDF4 files) were processed using the code in `walkerjeffd/sheds-daymet`.

3.1.3 Geospatial Characteristics

Geospatial characteristics of each catchment were generated using the code in basinCharacteristics. More information can be found on the SHEDS GIS Data page.

3.2 Stream Temperature Observations

Temperature observations extractd from the SHEDS Stream Temperature Database, which contains data uploaded from numerous state and federal agencies, universities, and non-profit organizations.

3.2.1 Data Summary

The following table summarizes the number of monitoring stations, deployments, and total observations for each agency and organization.

Agency ID	Agency Description	# Statio
ARWC	Androscoggin River Watershed Council	
AsRA	Ausable River Association	
BCDEPS	Baltimore County Department of Environmental Protection and Sustainability	
BSP	Baxter State Park	1
CTDEEP	CT Department of Energy and Environmental Protection	5,0
CTVOLMON	CT Volunteer Monitoring	8
FRITSCHIE	Keith Fritschie, Dartmouth	
GOMRHN	Gulf of Maine River Herring Network	
HAYDEN	Michael Hayden	
HBM	Houlton Band of Malisset	
HCSWCD	Hancock County Soil and Water Conservation District	
HMF	Hopkins Memorial Forest	
HVA_BERKSHIRE	Housatonic Valley Association, Berkshires	
JMU	James Madison University	
MADEP	MA Department of Environmental Protection	1,7
MAFW	MA Fish and Wildlife	
MAFW_RQ	MA Division of Fish and Wildlife, Rebecca Quinones	
MBSS	Maryland Biological Stream Survey	4
MEDEP	ME Department of Environmental Protection	1,3
MEDMR	Maine Department of Marine Resources	
MEDOT	ME Department of Transportation	
MEIFW	ME Inland Fisheries and Wildlife	2
ME_MC	Midcoast Conservancy	
MRWC	Miller's River Watershed Council	
NHDES	NH Department of Environmental Services	
NHFG	NH Department of Fish and Game	
NJDEP	New Jersey Department of Environmental Protection	
NJDFW	NJ Dept of Environmental Protection, Div of Fish and Wildlife	1
NOAA	National Oceanic and Atmospheric Administration	2
NPS_SNHA	National Park Service, Susquehanna National Heritage Area	
NPS_UDR	National Park Service, Upper Delaware River	
NRWA	Nashua River Watershed Association	
NYDEC	NY Department of Environmental Conservation	1,1
OARS	OARS	
PINDNR	Penobscot Indian Nation Department of Natural Resources	
PSU	Plymouth State University	1
RIDEM	Rhode Island Department of Environmental Management	2
SRBC	Susquehanna River Basin Commission	3
SRCC	Saco River Corridor Commission	
SVWA	Southeastern Vermont Watershed Alliance	
TNC_NJ	The Nature Conservancy, New Jersey	
TU	Trout Unlimited	
TU_ADK	Trout Unlimited, Adirondacks	
TU_AF	Trout Unlimited, Art Flick Chapter	
TU_CLEARWATER	Trout Unlimited, Clearwater Chapter	
TU_DW	Trout Unlimited, Deerfield Watershed	
TU_HOME	Trout Unlimited, Homewaters Chapter	
TU_HR	Trout Unlimited, Housatonic River	
TU_KV	Trout Unlimited, Kennebec Valley	
TU_MMB	Trout Unlimited, Merry Meeting Bay	
TU_NJ	Trout Unlimited, New Jersey Chapter	
TU_NY	Trout Unlimited, New York Chapter	
UMAINE	University of Maine	

Chapter 4

Data Processing

A series of automated data screening and QA/QC procedures are used to process the raw stream temperature observations, and generate the final input dataset of mean daily stream temperatures, which is then used to calibrate and validate the model.

Processing of the raw stream temperature observations involves the following steps:

1. Screen monitoring stations based on location and drainage area characteristics
2. Aggregate instantaneous observations to daily time steps
3. Screen daily mean temperature values and time series according to various QA/QC tests
4. Aggregate daily time series measured at multiple stations within a single catchment
5. Screen time series to exclude cold weather conditions when air temperature and water temperature are de-coupled
6. Split the final dataset into subsets for calibration (90% of time series) and validation (10% of time series)

4.1 Screen Monitoring Stations

The full set of monitoring stations are first screened to exclude those that meet the following criteria:

1. Located within 100 m of an impoundment
2. Located within a tidal zone
3. Located within a catchment having a drainage area greater than 200 sq km

4. Located within a catchment having a drainage area with greater than 50% coverage by open water or wetlands

4.2 Screen Daily Mean Values

After screening for the monitoring stations, the remaining time series are screened according to the following criteria:

1. Remove daily values that have user-defined flag stored in the database
2. Remove daily values with mean temp < -25 degC
3. Remove daily values with mean or max temp > 35 degC
4. Remove time series with a high correlation between daily mean water and air temperature, which are suspected to be measurements of air temperature
5. Remove time series with fewer than 5 days of observations
6. Remove first and/or last day of each timeseries if the number of observations on those days is less than the median number of observations per day for the whole series (i.e. the first and/or last day are incomplete)

4.3 Combine Overlapping Timeseries

If there are two or more overlapping time series at a single monitoring station, then those time series are combined by:

1. Compute the overall mean value on each date of the overlapping time series.
2. Remove any daily values having more than one time series and a difference in mean observed temperature greater than 5 deg C between those time series

4.4 Screen Multiple Locations within Single Catchment

If there are multiple stations within a single catchment, then only one station is retained to represent that catchment based on the following criteria:

1. Exclude stations located more than 60 m from the main flowline of the catchment
2. If more than one station remains, choose the stations with greatest number of observations between April and October
3. If more than one station still remains, choose the station located nearest to the pour point of the catchment

4.5 Remove Cold Weather Periods

Lastly, each time series are truncated to the seasonal period during which water temperature and air temperature are coupled to exclude observations during cold weather periods.

4.6 Split Dataset for Calibration and Validation

The final dataset of mean daily stream temperature observations are then split for calibration and validation. The split is performed by randomly selecting 90% of the time series for calibration, and using the remaining 10% for validation.

Chapter 5

Calibration and Validation

After processing all of the data, the model was fitted using jags. The observation dataset was split into 80% for calibration, 20% for validation.

5.1 Parameter Estimates

5.1.1 Fixed Effects

Figure 5.1 and Table 5.1.1 present the estimated mean and 95% credible region interval (CRI) of each fixed effect parameter. The intercept term is not shown in the figure because the values are much larger than the other parameters, and would thus skew the scale.

\begin{table}

\caption{Estimated Mean and 95% CRI of Fixed Effects}

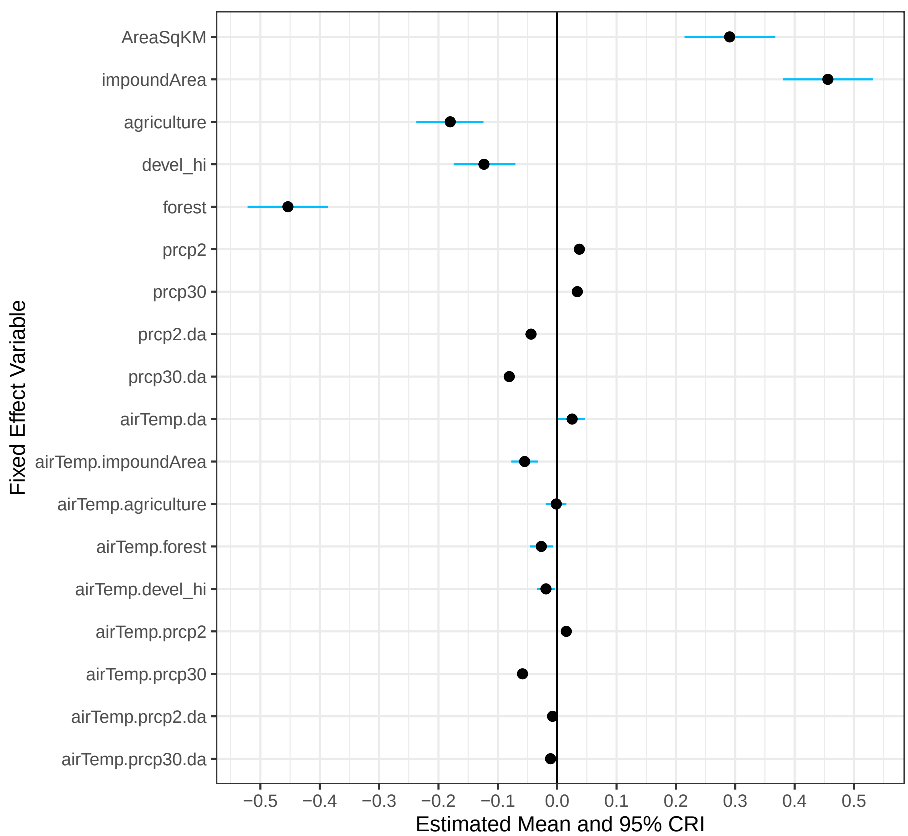


Figure 5.1: Estimated Mean and 95% CRI of Fixed Effects

Variable	Mean	Lower CRI	Upper CRI
intercept	16.831668438	16.682524667	16.975449785
AreaSqKM	0.290520111	0.214532987	0.367534804
impoundArea	0.456016628	0.380017105	0.532435569
agriculture	-0.180179569	-0.237304212	-0.124162093
devel_hi	-0.123342802	-0.174409953	-0.070354748
forest	-0.453606158	-0.521601617	-0.385848596
prcp2	0.037368478	0.035442637	0.039230802
prcp30	0.033889432	0.027601220	0.039870142
prcp2.da	-0.044082207	-0.045919083	-0.042220414
prcp30.da	-0.080723453	-0.086901209	-0.074180702
airTemp.da	0.025152719	0.001480242	0.047233232
airTemp.impoundArea	-0.054816312	-0.077282009	-0.031984213
airTemp.agriculture	-0.001544660	-0.019411269	0.015453974
airTemp.forest	-0.026747107	-0.046134745	-0.006951219
airTemp.devel_hi	-0.018879464	-0.033975352	-0.003815090
airTemp.prcp2	0.015308325	0.013553468	0.017116532
airTemp.prcp30	-0.058429824	-0.061922872	-0.054993205
airTemp.prcp2.da	-0.007801793	-0.009506454	-0.006054510
airTemp.prcp30.da	-0.011195140	-0.014792829	-0.007682154

\end{table}

5.1.2 HUC8 Random Effects

Figure 5.2 shows the estimated mean and 95% credible region interval (CRI) for each random effect and HUC8. Table 5.1.2 lists the estimated mean and 95% CRI of each parameter averaged over all HUC8s (mean value with standard deviation in parentheses).

\begin{table}

\caption{Mean and 95% CRI of HUC8 Random Effects Averaged Over All HUC8s (Mean Value and Std. Dev. in Parentheses)}

Variable	Count	Mean	Lower CRI	Upper CRI
intercept.huc	144	-0.001 (0.467)	-0.680 (0.538)	0.677 (0.512)
airTemp	144	2.009 (0.169)	1.781 (0.210)	2.236 (0.180)
temp7p	144	1.552 (0.335)	1.200 (0.371)	1.906 (0.377)

\end{table}

5.1.3 Catchment Random Effects

Figure 5.3 shows the distribution of the estimated mean for each random effect term over all catchments. CRIs are not shown due to the large number of

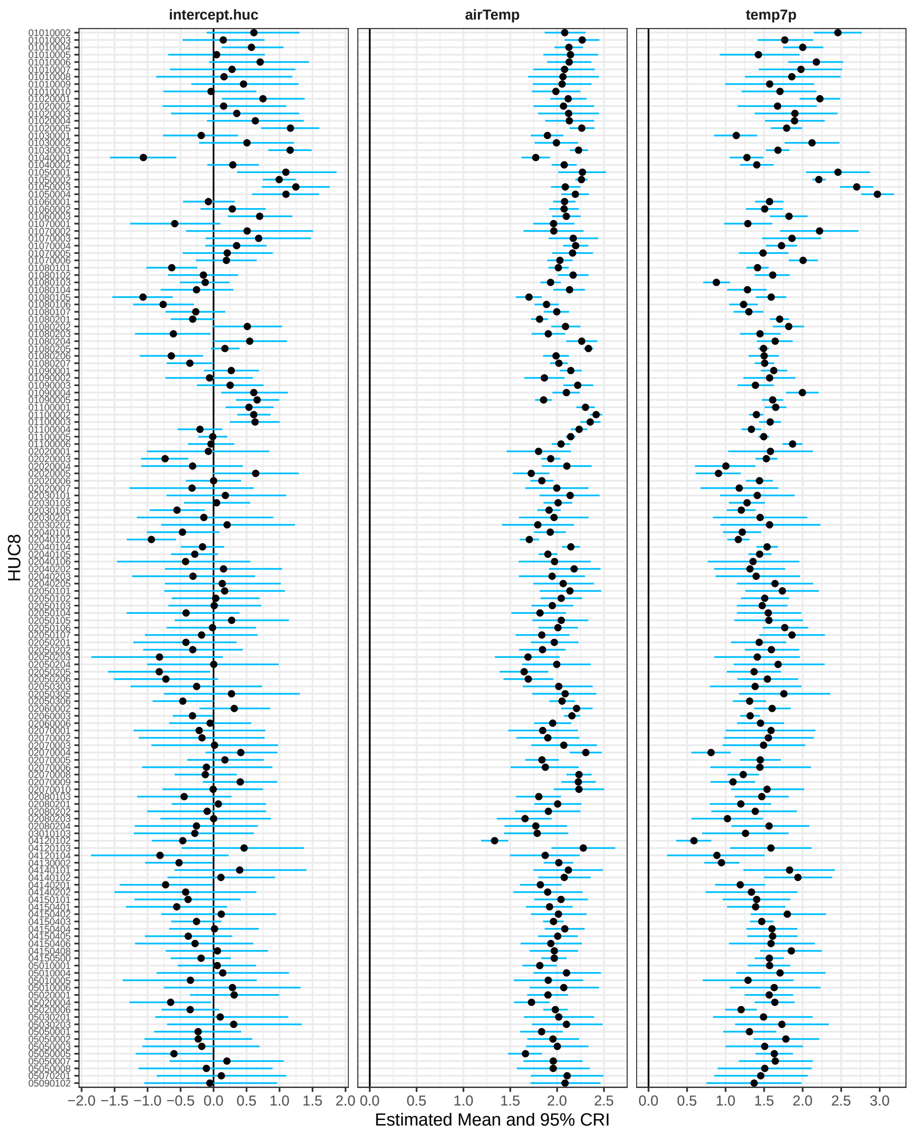


Figure 5.2: Estimated Mean and 95% CRI of HUC Random Effects for Each HUC8

individual catchments (8913). Table 5.1.3 lists the estimated mean and 95% CRI of each parameter averaged over all catchments (mean value with standard deviation in parentheses).

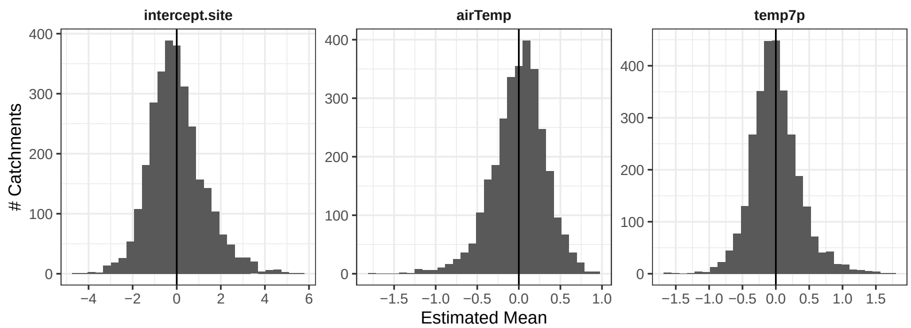


Figure 5.3: Distribution of estimated mean for each random effect over all catchments

\begin{table}

\caption{Estimated mean and 95% CRI for each random effect averaged over all catchments (mean value with std. dev. in parentheses)}

Variable	Count	Mean	Lower CRI	Upper CRI
intercept.site	2,971	-0.000 (1.254)	-0.721 (1.267)	0.721 (1.281)
airTemp	2,971	0.000 (0.319)	-0.283 (0.331)	0.283 (0.336)
temp7p	2,971	-0.000 (0.366)	-0.500 (0.433)	0.499 (0.370)

\end{table}

5.1.4 Year Random Effects

Figure 5.4 and Table 5.1.4 present the mean and 95% CRI of the intercept term for each year. Recall that there are no random effects for years other than the intercept.

\begin{table}

\caption{Estimated Mean and 95% CRI of Intercept Random Effect for Each Year}

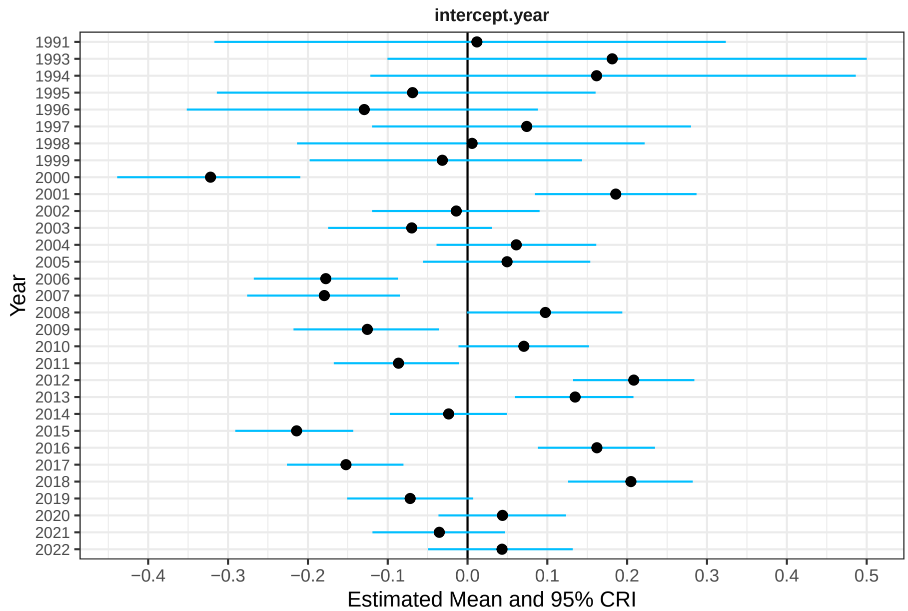


Figure 5.4: Estimated Mean and 95% CRI of Intercept Random Effect for Each Year

Year	Mean	Lower CRI	Upper CRI
1991	0.012	-0.317	0.324
1993	0.181	-0.100	0.500
1994	0.162	-0.122	0.486
1995	-0.069	-0.314	0.160
1996	-0.129	-0.352	0.088
1997	0.074	-0.119	0.280
1998	0.006	-0.214	0.222
1999	-0.032	-0.198	0.143
2000	-0.322	-0.439	-0.209
2001	0.186	0.084	0.287
2002	-0.014	-0.119	0.090
2003	-0.070	-0.174	0.031
2004	0.061	-0.039	0.161
2005	0.050	-0.056	0.154
2006	-0.178	-0.268	-0.087
2007	-0.179	-0.276	-0.085
2008	0.098	-0.001	0.194
2009	-0.126	-0.218	-0.036
2010	0.071	-0.011	0.152
2011	-0.086	-0.168	-0.011
2012	0.208	0.132	0.284
2013	0.135	0.059	0.208
2014	-0.024	-0.097	0.049
2015	-0.214	-0.291	-0.143
2016	0.162	0.088	0.235
2017	-0.152	-0.226	-0.080
2018	0.205	0.126	0.282
2019	-0.072	-0.151	0.007
2020	0.044	-0.036	0.123

Table 5.1: Summary statistics of model calibration and validation

	Calibration	Validation
# Daily Observations	705,509	80,035
# Time Series	7,814	863
# Catchments	2,971	566
# HUC8s	144	109
# Years	31	27
RMSE (degC)	1.061	1.342
Mean Residual (degC)	0.061	0.079
Median Residual (degC)	0.075	0.062
Mean Absolute Residual (degC)	0.813	1.021
Median Absolute Residual (degC)	0.649	0.805
Minimum Residual (degC)	-7.570	-7.665
1st Percentile Residual (degC)	-2.674	-3.225
99th Percentile Residual (degC)	2.659	3.697
Maximum Residual (degC)	9.196	7.411

\end{table}

5.2 Goodness-of-Fit

5.2.1 All Observations

Table 5.1 summarizes the model goodness-of-fit for all observations in the calibration and validation datasets.

Figure 5.5 presents scatterplots of predicted vs. observed daily mean temperature for the calibration and validation datasets. The black line is the 1:1 line of equality. The red line is a linear regression trend line.

5.2.2 Deployments

Table 5.2 summarises the mean, median, minimum and maximum RMSE for each deployment (i.e. continuous timeseries of observations at a single location) in the calibration and validation datasets.

Figure 5.6 shows the distribution of deployment RMSE.

5.2.2.1 Calibration Deployment Examples

Figures 5.7 to 5.8 show example deployments from the calibration dataset with the highest and lowest RMSE.

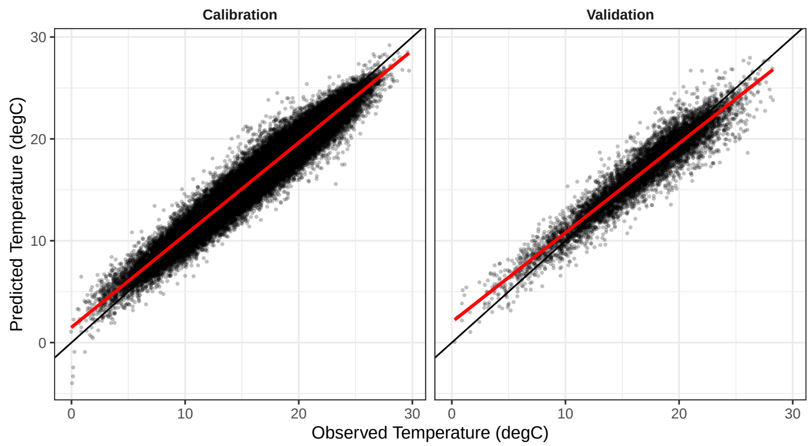


Figure 5.5: Predicted versus Observed Daily Mean Temperature (degC) for Calibration and Validation Datasets

Table 5.2: Summary statistics of model calibration and validation RMSE for each deployment

	Calibration	Validation
# Time Series	7814	863
Mean RMSE (degC)	0.995	1.216
Median RMSE (degC)	0.915	1.087
Minimum RMSE (degC)	0.250	0.338
Maximum RMSE (degC)	5.303	5.421

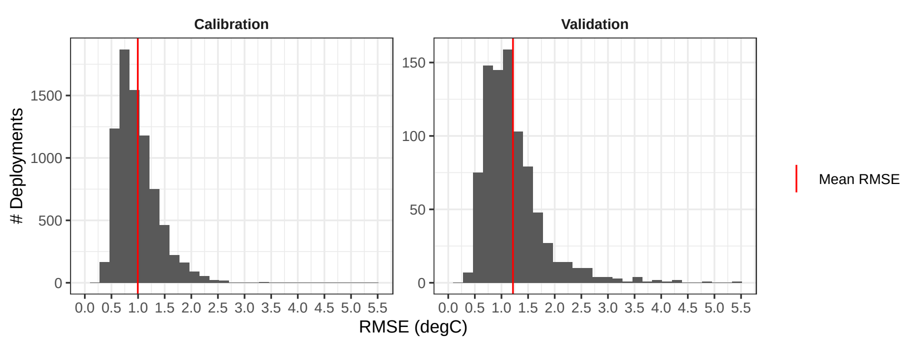


Figure 5.6: Distribution of deployment RMSE

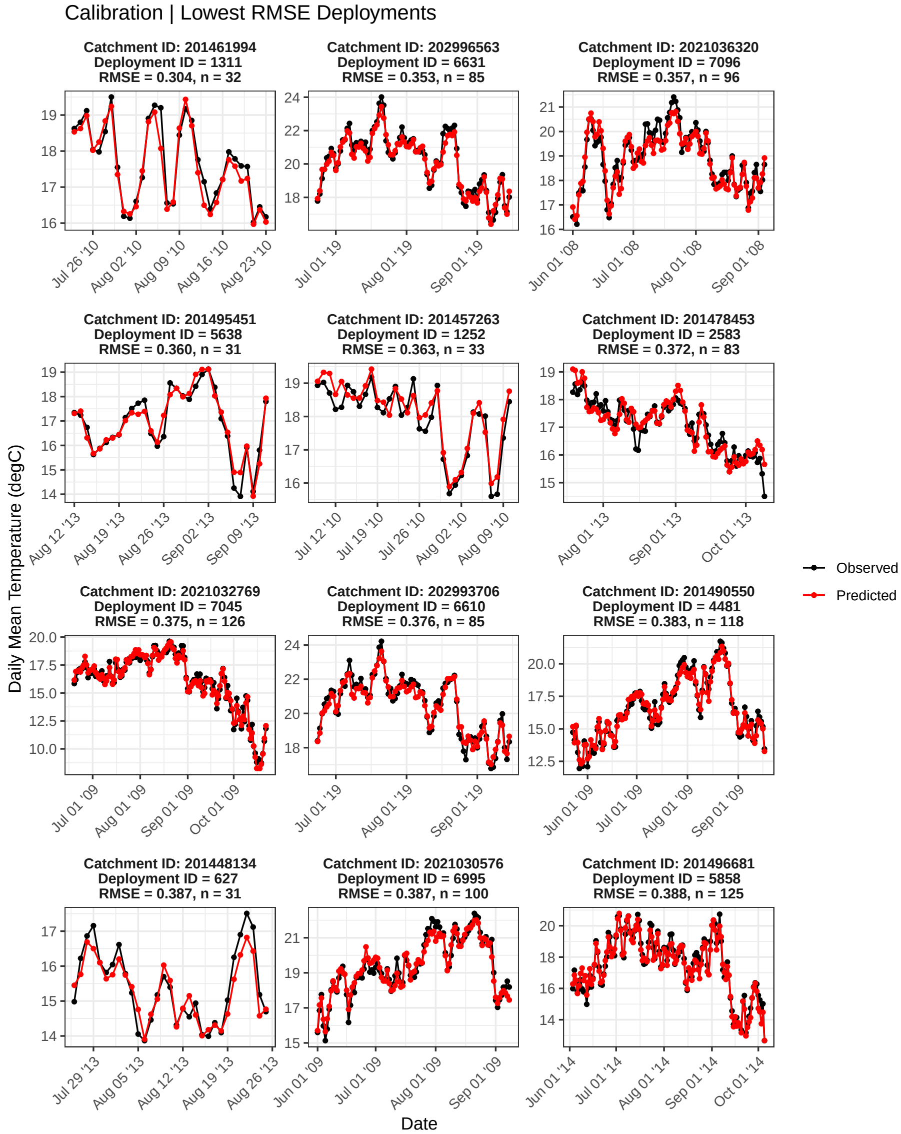


Figure 5.7: Deployments with lowest RMSE in calibration dataset and n >= 30

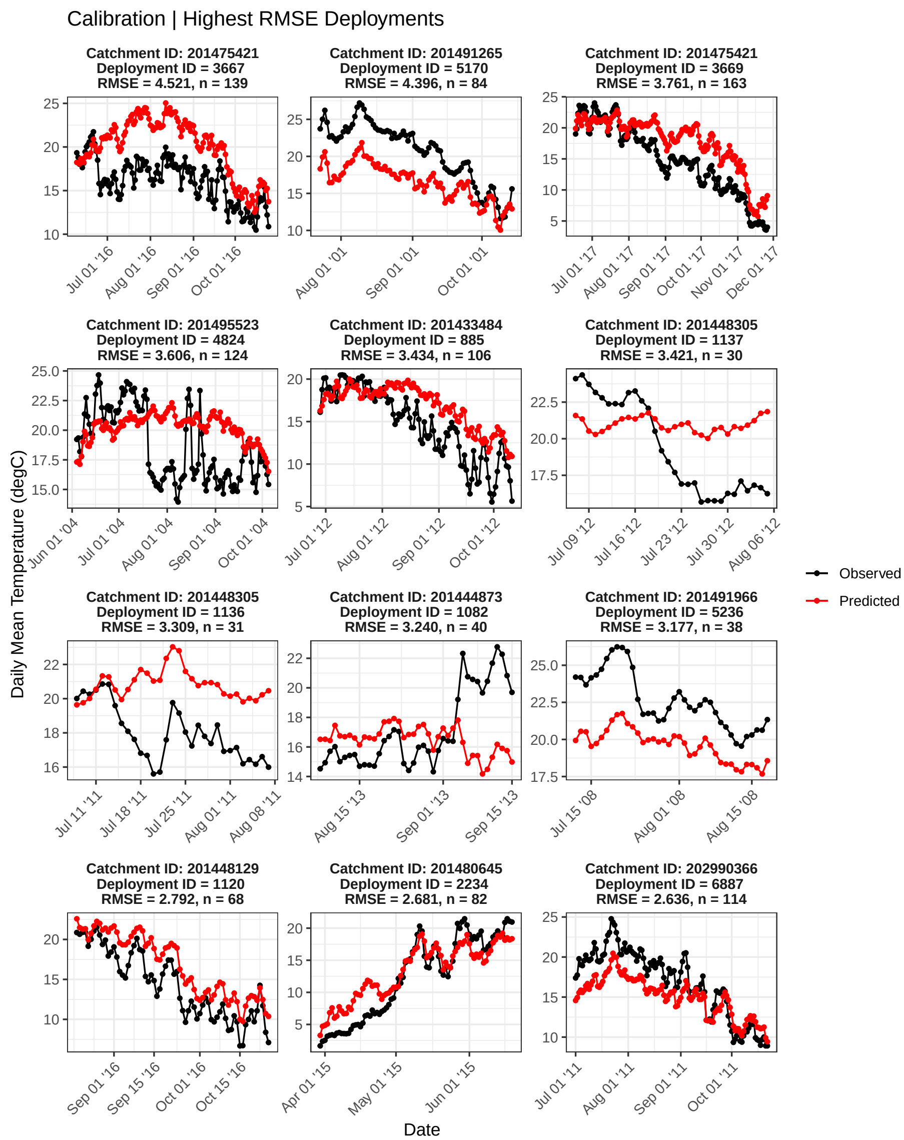


Figure 5.8: Deployments with highest RMSE in calibration dataset and n >= 30

Table 5.3: Summary of catchment RMSE values for calibration and validation datasets

	Calibration	Validation
# Time Series	2971	566
Mean RMSE (degC)	0.918	1.267
Median RMSE (degC)	0.842	1.112
Minimum RMSE (degC)	0.278	0.490
Maximum RMSE (degC)	3.240	5.421

5.2.2.2 Validation Deployment Examples

Figures 5.9 to 5.10 show example deployments from the validation dataset with the highest and lowest RMSE.

5.2.3 Catchments

Table 5.3 summarises the mean, median, minimum and maximum RMSE of all catchments in the calibration and validation datasets.

Figure 5.11 shows the distribution of catchment RMSE.

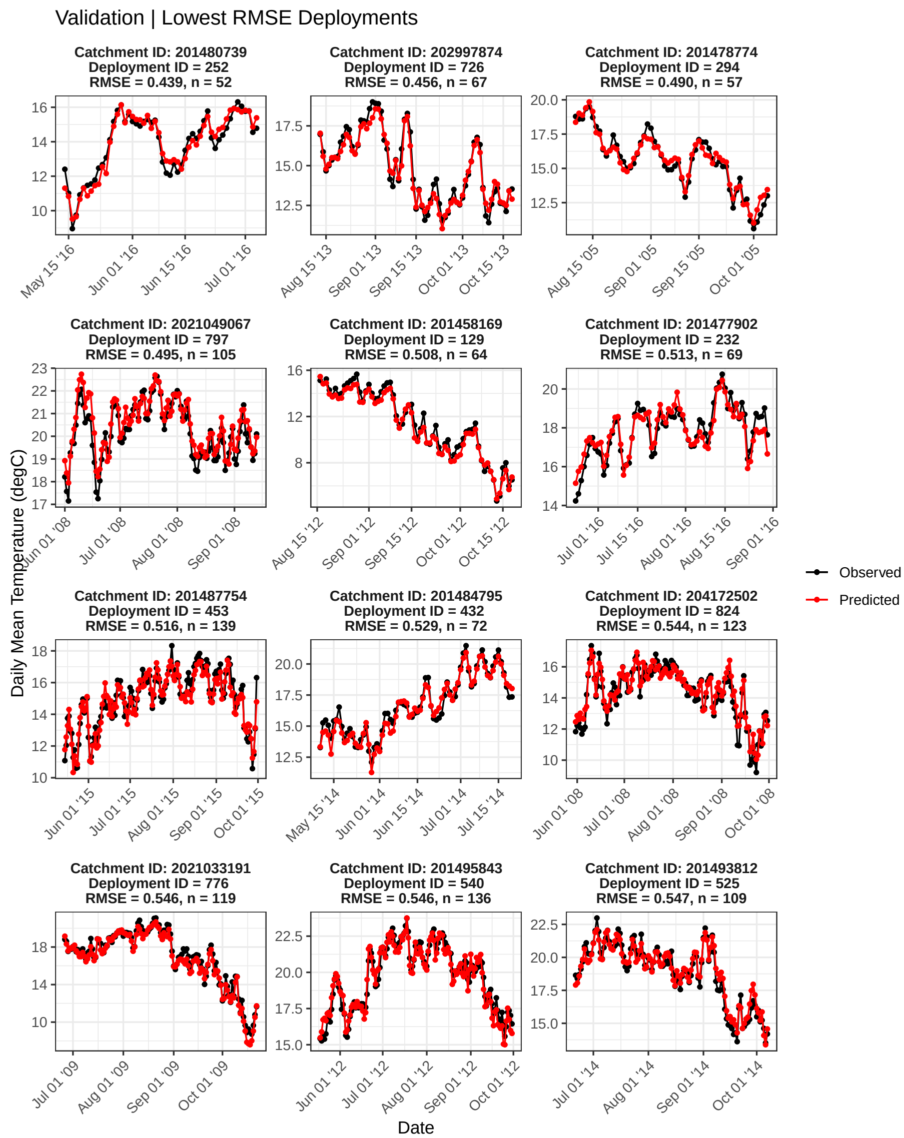


Figure 5.9: Deployments with lowest RMSE in validation dataset and n >= 30

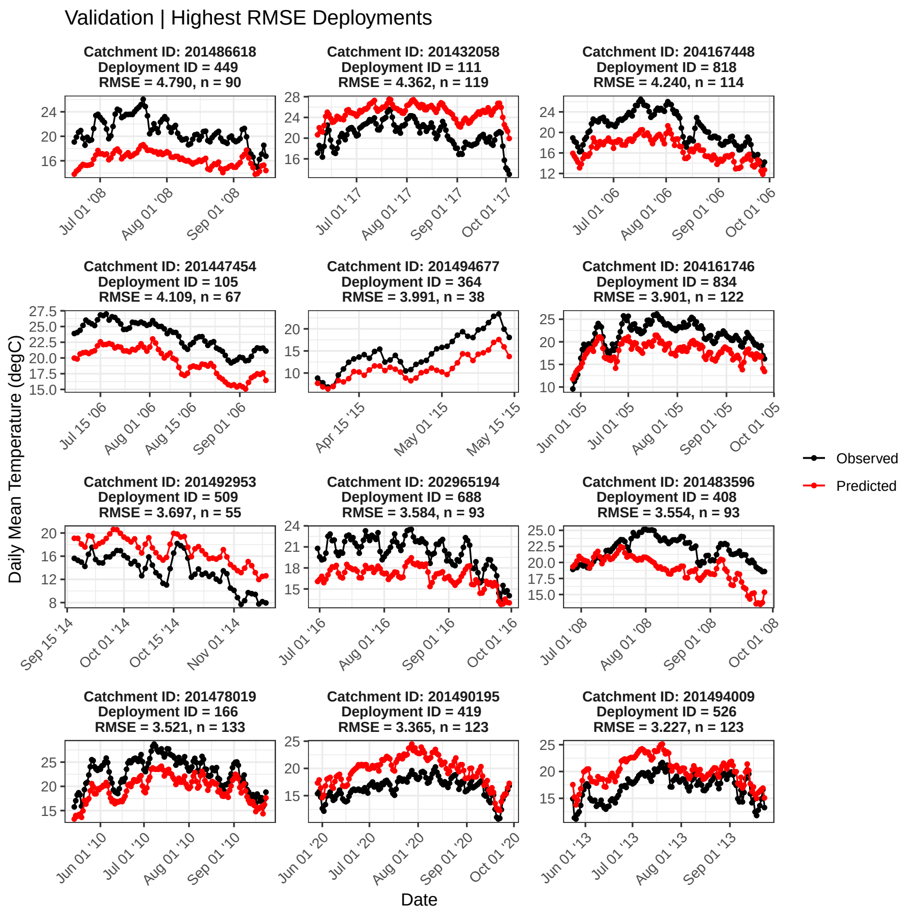


Figure 5.10: Deployments with highest RMSE in validation dataset and n >= 30

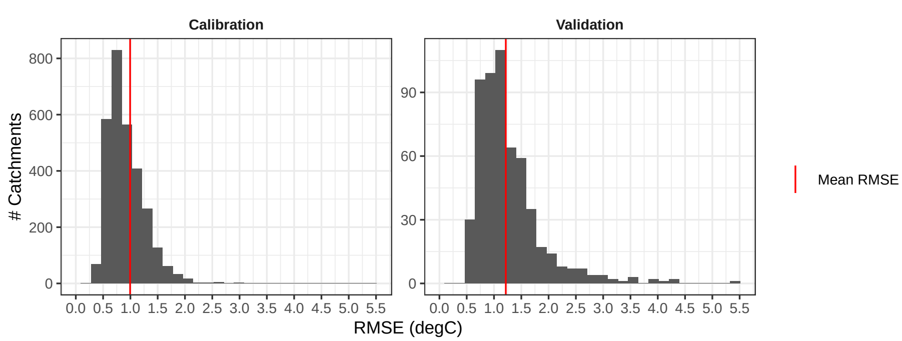


Figure 5.11: Distribution of catchment RMSE

Chapter 6

Predictions

Once calibrated, the model is used to generate stream temperature predictions for all catchments in the region having drainage areas less than 200 sq km and less than 70% open water coverage.

First, daily time series of predicted mean temperature are generated for each catchment. The daily predictions are then aggregated by year by calculated various metrics such as mean monthly temperatures and frequencies of exceeding various thresholds. Lastly, the annual temperature metrics are aggregated by catchment over all years in the historical period of record.

For each catchment, predictions are generated by the following steps:

1. Calculate the predicted daily mean stream temperature for each year in the historical period of record (1980 - 2022) during the months of June, July and August
2. For each month and year, calculate the following metrics:
    - Mean monthly temperature (June, July, August)
    - Mean and maximum summer temperature (June 1 - August 31)
    - Maximum 30-day rolling mean temperature
    - Number of days with temperature exceeding 18, 20, and 22 deg C
    - Resistivity computed as the sum of the absolute difference between the predicted stream temperature and the daily mean air temperature (June 1 to August 31)
3. For each catchment, the annual metrics from the previous step are aggregated by calculating the following metrics over all years:
    - Mean monthly (June, July, August) and summer (June 1 - August 31) temperature
    - Mean and maximum of the annual maximum summer temperature
    - Mean of the annual maximum 30-dy rolling mean temperature
    - Mean number of days per year exceeding 18, 20, and 22 deg C
    - Mean resistivity

Figure 6.1 shows the distribution (histogram) of each final metric for all catchments.

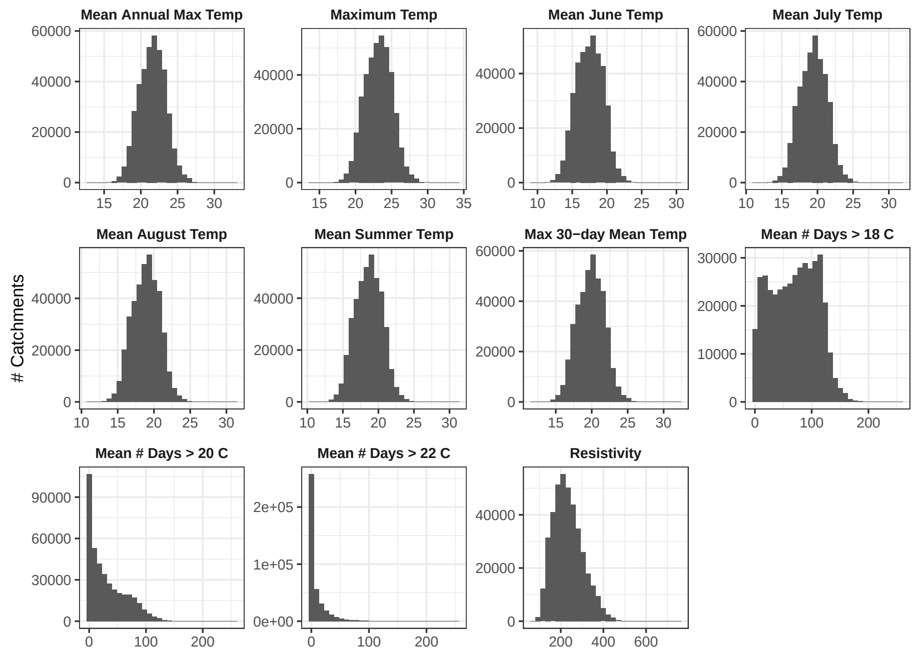


Figure 6.1: Distributions of Catchment Derived Metrics

6.1 Air Temperature Scenarios

Beginning with `v1.1.1`, the prediction metrics described above are generated for a series of climate change scenarios with air temperatures increasing by 2, 4, and 6 degC. These scenarios are calculated by increasing each daily air temperature by the respective amount. In other words, these scenarios assume a uniform increase in air temperature, which does not vary seasonally.

Figure 6.2 shows the distribution of mean July stream temperature among all catchments for each air temperature scenario.

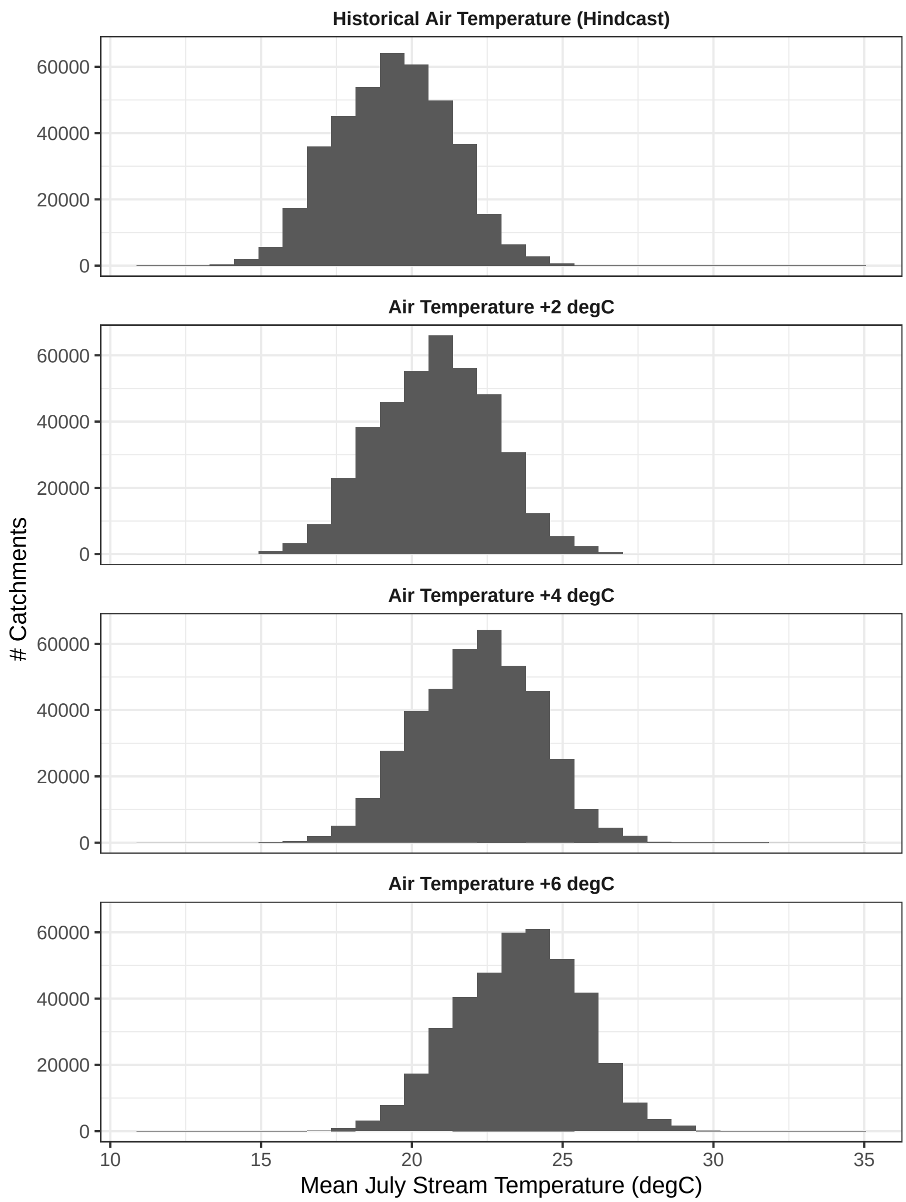


Figure 6.2: Distributions of Mean July Stream Temperature with Varying Air Temperature Increases

Chapter 7

Download

7.1 Stream Temperature Model Predictions Dataset

The stream temperature model predictions can be downloaded as a static CSV file from the following link.

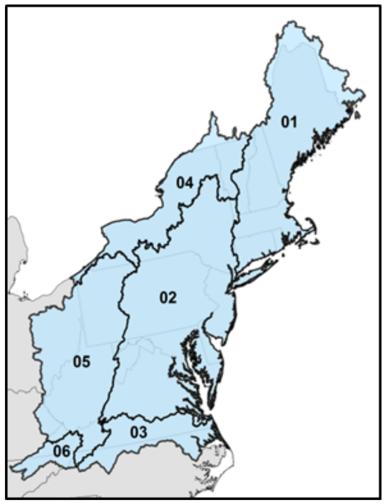
Stream Temperature Predictions v1.4.0 (csv)

This file contains the following columns:

CSV Column	Description
featureid	Catchment ID
max_max_temp	Maximum Temp (degC)
max_max_temp_air2	Maximum Temp (degC) w/ Air Temp +2 degC
max_max_temp_air4	Maximum Temp (degC) w/ Air Temp +4 degC
max_max_temp_air6	Maximum Temp (degC) w/ Air Temp +6 degC
max_temp_30d	Max 30-day Mean Temp (degC)
max_temp_30d_air2	Max 30-day Mean Temp (degC) w/ Air Temp +2 degC
max_temp_30d_air4	Max 30-day Mean Temp (degC) w/ Air Temp +4 degC
max_temp_30d_air6	Max 30-day Mean Temp (degC) w/ Air Temp +6 degC
mean_aug_temp	Mean August Temp (degC)
mean_aug_temp_air2	Mean August Temp (degC) w/ Air Temp +2 degC
mean_aug_temp_air4	Mean August Temp (degC) w/ Air Temp +4 degC
mean_aug_temp_air6	Mean August Temp (degC) w/ Air Temp +6 degC
mean_jul_temp	Mean July Temp (degC)
mean_jul_temp_air2	Mean July Temp (degC) w/ Air Temp +2 degC
mean_jul_temp_air4	Mean July Temp (degC) w/ Air Temp +4 degC
mean_jul_temp_air6	Mean July Temp (degC) w/ Air Temp +6 degC
mean_jun_temp	Mean June Temp (degC)
mean_jun_temp_air2	Mean June Temp (degC) w/ Air Temp +2 degC
mean_jun_temp_air4	Mean June Temp (degC) w/ Air Temp +4 degC
mean_jun_temp_air6	Mean June Temp (degC) w/ Air Temp +6 degC
mean_max_temp	Mean Annual Max Temp (degC)
mean_max_temp_air2	Mean Annual Max Temp (degC) w/ Air Temp +2 degC
mean_max_temp_air4	Mean Annual Max Temp (degC) w/ Air Temp +4 degC
mean_max_temp_air6	Mean Annual Max Temp (degC) w/ Air Temp +6 degC
mean_summer_temp	Mean Summer Temp (degC)
mean_summer_temp_air2	Mean Summer Temp (degC) w/ Air Temp +2 degC
mean_summer_temp_air4	Mean Summer Temp (degC) w/ Air Temp +4 degC
mean_summer_temp_air6	Mean Summer Temp (degC) w/ Air Temp +6 degC
n_day_temp_gt_18	Mean # Days per Year Temp > 18 C
n_day_temp_gt_18_air2	Mean # Days per Year Temp > 18 C w/ Air Temp +2 degC
n_day_temp_gt_18_air4	Mean # Days per Year Temp > 18 C w/ Air Temp +4 degC
n_day_temp_gt_18_air6	Mean # Days per Year Temp > 18 C w/ Air Temp +6 degC
n_day_temp_gt_20	Mean # Days per Year Temp > 20 C
n_day_temp_gt_20_air2	Mean # Days per Year Temp > 20 C w/ Air Temp +2 degC
n_day_temp_gt_20_air4	Mean # Days per Year Temp > 20 C w/ Air Temp +4 degC
n_day_temp_gt_20_air6	Mean # Days per Year Temp > 20 C w/ Air Temp +6 degC
n_day_temp_gt_22	Mean # Days per Year Temp > 22 C
n_day_temp_gt_22_air2	Mean # Days per Year Temp > 22 C w/ Air Temp +2 degC
n_day_temp_gt_22_air4	Mean # Days per Year Temp > 22 C w/ Air Temp +4 degC
n_day_temp_gt_22_air6	Mean # Days per Year Temp > 22 C w/ Air Temp +6 degC
n_day_temp_gte_24_9	Mean # Days per Year Temp >= 24.9 C
n_day_temp_gte_24_9_air2	Mean # Days per Year Temp >= 24.9 C w/ Air Temp +2 degC
n_day_temp_gte_24_9_air4	Mean # Days per Year Temp >= 24.9 C w/ Air Temp +4 degC
n_day_temp_gte_24_9_air6	Mean # Days per Year Temp >= 24.9 C w/ Air Temp +6 degC
n_day_temp_gte_27	Mean # Days per Year Temp >= 27 C
n_day_temp_gte_27_air2	Mean # Days per Year Temp >= 27 C w/ Air Temp +2 degC
n_day_temp_gte_27_air4	Mean # Days per Year Temp >= 27 C w/ Air Temp +4 degC
n_day_temp_gte_27_air6	Mean # Days per Year Temp >= 27 C w/ Air Temp +6 degC
resist	Resistivity
resist_air2	Resistivity w/ Air Temp +2 degC
resist_air4	Resistivity w/ Air Temp +4 degC
resist_air6	Resistivity w/ Air Temp +6 degC

7.2 Catchment Delineation Shapefiles

The EcoSHEDS Northeast Catchment Delineation (NECD) is available as a series of shapefiles, pre-staged by 2-digit hydrologic unit codes (HUCs). The model predictions and covariates CSV files can be joined to these shapefiles using the mutual `featureid` column.



- Region 01 Catchments (zipped shp)
- Region 02 Catchments (zipped shp)
- Region 03 Catchments (zipped shp)
- Region 04 Catchments (zipped shp)
- Region 05 Catchments (zipped shp)
- Region 06 Catchments (zipped shp)

The documentation for the catchment delineation is also available:

Catchment Delineation (NHDHRDV2) Documentation (docx)

7.3 Catchment Covariates Dataset

The SHEDS catchment covariates are available as a series of CSV files, pre-staged by 2-digit hydrologic unit codes (HUCs). The covariates contain the catchment characteristics that are used as input variables to the stream temperature model.

- Region 01 Covariates (zipped csv)
- Region 02 Covariates (zipped csv)
- Region 03 Covariates (zipped csv)
- Region 04 Covariates (zipped csv)
- Region 05 Covariates (zipped csv)

- Region 06 Covariates (zipped csv)

The documentation for catchment covariates is also available:

Catchment Covariates Documentation (docx)

References

Bibliography

Letcher, B. H., Hocking, D. J., O'Neil, K., Whiteley, A. R., Nislow, K. H., and O'Donnell, M. J. (2016). A hierarchical model of daily stream temperature using air-water temperature synchronization, autocorrelation, and time lags. *PeerJ*, 4:e1727.