Journal of Hydrology X 22 (2024) 100172

Contents lists available at ScienceDirect

Journal of Hydrology X

journal homepage: www.sciencedirect.com/journal/journal-of-hydrology-x

Air temperature data source affects inference from statistical stream temperature models in mountainous terrain

Daniel J. Isaak *, Dona L. Horan, Sherry P. Wollrab

US Forest Service, Rocky Mountain Research Station, Boise, ID, USA

ARTICLE INFO

Keywords:
Stream temperature
Air temperature
Thermal sensitivity
Gridded climate data set
Mountains
Statistical model

ABSTRACT

Instream temperatures control numerous biophysical processes and are frequently the subject of modeling efforts to understand and predict responses to watershed conditions, habitat alterations, and climate change. Air temperature (AT) is regularly used in statistical temperature models as a covariate proxy for physical processes and because it correlates strongly with spatiotemporal variability in water temperatures (T_w). Air temperature data are broadly available and sourced from sensors paired with T_w sites, remote weather stations, and gridded climate data sets—often with limited recognition of the tradeoffs these sources present and how microclimatic variation in topographically complex mountain environments could affect model inference. To address these issues, we collected daily T_w records at 13 sites throughout a mountain river network, linked the records to AT data from 11 sources available across much of North America, and fit linear regression models to assess predictive performance and the consistency of parameter estimation. Although the predictive accuracy of these models was generally high, estimates of the AT slope parameter, which is commonly interpreted as thermal sensitivity, varied substantially depending on the AT data source. These results have implications for the comparability of estimates among T_w studies and highlight the challenges that modeling stream temperatures in mountain landscapes presents. Although no AT data source is ideal, some are more advantageous than others for specific use cases and we provide general recommendations on this topic.

1. Introduction

Water temperature (T_w) is a master variable in streams due to its mediation of other water quality attributes (Himmelblau, 1960; Jacobsen, 2020) and direct effects on organisms (Pörtner and Peck, 2010; Verberk et al., 2011). Thermal conditions vary throughout river networks and are responsive to the amount of shade provided by near-stream vegetation (Moore et al., 2005; Nussle et al., 2015), channel morphology (Nichols and Ketcheson, 2013; Weber et al., 2017), groundwater influxes (Briggs et al., 2018; Rey et al., 2023), and meteorological variability (Webb and Nobilis, 2007; Isaak et al., 2012). Models are frequently used to understand the linkages between T_w variation and environmental conditions, and statistical models, in particular, have proliferated in recent decades due to their ease of implementation with growing T_w and covariate databases (Benyahya et al., 2007; Ouellet et al., 2020). Air temperature (AT) is commonly used as a covariate in these models because it represents physical mechanisms such as longwave radiation and sensible heat transfer (Webb and Zhang, 1997; St-Hilaire et al., 2021) and correlates strongly with T_w over a range of spatial and temporal scales (Stefan and Preud'homme, 1993; Morrill et al., 2005; van Vliet et al., 2011). In statistical models that predict spatial variability among stream sites, the ratio between T_w variation and AT variation ($\Delta\ ^\circ C_{Tw}/\Delta\ ^\circ C_{AT}$), estimated as the slope parameter in linear regression models, can be interpreted as a lapse rate and provides information about T_w trends relative to gradients in elevation or latitude (Isaak and Luce, 2023; Isaak and Rieman, 2013; Weller et al., 2023). In statistical models predicting temporal dynamics at individual sites, however, this ratio is commonly referred to as thermal sensitivity (T_{sens}; Kelleher et al., 2012; Luce et al., 2014) and is used to infer the importance of groundwater buffering (e.g., smaller $T_{sens} \sim$ larger groundwater contributions) and to forecast or hindcast changes in T_w by scaling predicted or observed changes in AT (Isaak et al., 2012; Mauger et al., 2017). Similar site-specific temporal applications have been undertaken based on annual amplitude ratios between AT and T_w (Briggs et al., 2018; Johnson et al., 2020), using the popular nonlinear regression model developed by Mohseni et al. (1998,1999), and a growing host of artificial intelligence and hybrid modeling approaches (Gallice et al., 2015; Piccolroaz et al., 2016; Zhu and Piotrowski, 2020).

* Corresponding author.
E-mail address: daniel.isaak@usda.gov (D.J. Isaak).

https://doi.org/10.1016/j.hydroa.2024.100172
Received 28 July 2023; Received in revised form 5 January 2024; Accepted 25 January 2024
Available online 8 February 2024

The AT data sets used in T_w models are obtained from a variety of sources but fall into three basic categories. Early studies usually relied on AT records from long-term weather stations that were located remotely from T_w monitoring sites (e.g., Stefan and Preud'homme, 1993; Webb and Nobilis, 1997; Mohseni et al., 1998). With the advent of miniature temperature sensors in the 1990s, however, the costs of data acquisition decreased substantially and researchers began implementing paired deployments wherein T_w sensors were matched with AT sensors placed in adjacent valleys or near-stream vegetation (e.g., Trumbo et al., 2014; Briggs et al., 2018; Molinos et al., 2022) to improve the accuracy of microclimatic measurements associated with streams (Benyahya et al., 2010). The third source of AT data is gridded climate data sets (e.g., Mohseni et al., 2003; Siegel et al., 2022), which are spatially continuous predictions of thermal conditions derived from an underlying model. These models use statistical approaches to interpolate spatial and temporal AT patterns from data records collected at weather stations and locally dense AT sensor arrays which may be deployed to enhance spatial resolution in complex topographic environments (Lundquist and Cayan, 2007; Holden et al., 2016). The best known of these is the PRISM (Parameter-elevation Relationships on Independent Slopes; Daly et al., 2002,2008) model but several others have been developed and seen use in T_w studies, which are based on different input data sets and interpolation routines that span the coterminous U.S. or the entirety of North America (Table 1).

As the number of T_w models has grown, there has been limited recognition of the potential consequences that the use of AT data from different sources could have on inference. One exception is Mohseni et al. (1998), who compared the predictive accuracy of T_w models developed with AT data from remote stations across a range of distances and found limited performance declines out to spatial lags of 250 km. More recently, Johnson et al. (2020) compared AT measured at stream sites with that predicted for the same locations by PRISM and concluded the values were similar enough that PRISM data could be used at T_w sites which lacked paired AT sensors. However, no study of which we are aware has attempted a broader assessment of AT data from numerous sources or examined the effects of temporal variability in these records, which should be a key determinant of T_{sens} parameter estimates with statistical models. Moreover, AT variability often differs substantially over short distances based on landscape position and local microclimates (Daly et al., 2010; Pepin et al., 2011), especially in mountainous environments where a disproportionately large number of T_w studies have been conducted due to concerns about climatically sensitive cold-water organisms (Gallice et al., 2015; Ouellet et al., 2020). In these environments, AT variability in concave landforms such as the valley bottoms where paired AT sensors are deployed is usually less than convex landforms like the exposed ridgelines where weather monitoring stations are often established or the freely circulating air masses above the terrain which are simulated by global climate models (GCMs; Daly et al., 2010; Pepin et al., 2011).

To explore these issues in more detail, we collected T_w records across a mountain river network, linked these records to AT data from paired, remote, and gridded climate sources, then derived T_{sens} estimates and measures of predictive accuracy using statistical regression models. Our results indicate that nontrivial effects on parameter estimates and predictive accuracy may arise simply due to the source of AT data, a finding that should stimulate a broader discussion within the T_w modeling community about the complexities of using AT as a covariate within mountainous landscapes. Secondary objectives of this study were to broaden awareness regarding some of the AT data sources that are currently available in North America and to describe tradeoffs associated with using different data sets.

2. Materials and methods

2.1. Study area

The study was conducted in the upper Boise River Basin of central Idaho in the northwestern U.S. (Fig. 1a). The river basin encompasses

Table 1
Air temperature data sets with daily temporal resolution that are broadly available for the United States and North America.

Source type	Data source[1]	Spatial extent and resolution	Temporal extent	Data description	References
Weather monitoring station	SNOTEL	–	1980–present	Site records of daily temperatures, snowpack, precipitation, and other climatic conditions collected by a network of 900 automated monitoring stations in high-elevation mountain watersheds across the western U.S.	USDA, 1988; Schaefer and Johnson, 1992; Serreze et al., 1999
	NOAA COOP	–	1895–present	Site records of daily minimum and maximum temperatures from a network of 1,219 monitoring stations, which are maintained by volunteers across the coterminous U.S. Record lengths vary by site but many are long-term and the most consistent subset of records comprises the U.S. Historical Climatology Network.	Menne et al., 2009
Gridded climate data set	PRISM[2]	Coterminous U.S. at 4 km^2	1980–present	Gridded data sets of air temperature and precipitation that were predictively interpolated from records at ~10,000 monitoring sites, which include records from SNOTEL, COOP, and other sources.	Daly et al., 2002, 2008
	TopoWx	Coterminous U.S. at 0.8 km^2	1948–2016	Gridded data sets of air temperature that were predictively interpolated from records at ~14,000 monitoring sites, which include records from SNOTEL, COOP, and other sources. Interpolation routines also incorporate remotely sensed satellite estimates of land surface temperature as a covariate.	Oyler et al., 2014, 2016
	DayMet	North America at 1 km^2	1980–present	Gridded data sets of air temperature, precipitation, and other climate variables that were predictively interpolated from records at ~12,000 monitoring sites, which include records from SNOTEL, COOP, and other sources.	Thornton et al., 1997, 2012, 2021
	GridMet	Coterminous U.S. at 4 km^2	1979–present	Gridded data sets of air temperature, precipitation, and other climate variables derived by integrating PRISM spatial data sets with temporal data from the North American Land Data Assimilation System Phase 2 (Mitchell et al., 2004).	Abatzoglou, 2013

[1] Websites hosting data: SNOTEL (Snow Telemetry): https://www.nrcs.usda.gov/wps/portal/wcc/home/aboutUs/monitoringPrograms/automatedSnowMonitoring; NOAA COOP (National Oceanic and Atmospheric Administration Cooperative Observer Program): https://www.ncei.noaa.gov/products/land-based-station/cooperative-observer-network; PRISM (Parameter-elevation Relationships on Independent Slopes Model): https://prism.oregonstate.edu/; TopoWx (Topography Weather): https://www.scrim.psu.edu/resources/topowx/; DayMet (Daily Meteorology): https://daymet.ornl.gov/; GridMet (Gridded Meteorology): https://www.climatologylab.org/gridmet.html.

[2] Higher-resolution data at 0.8 km^2 are available for a fee but were not used in this study.

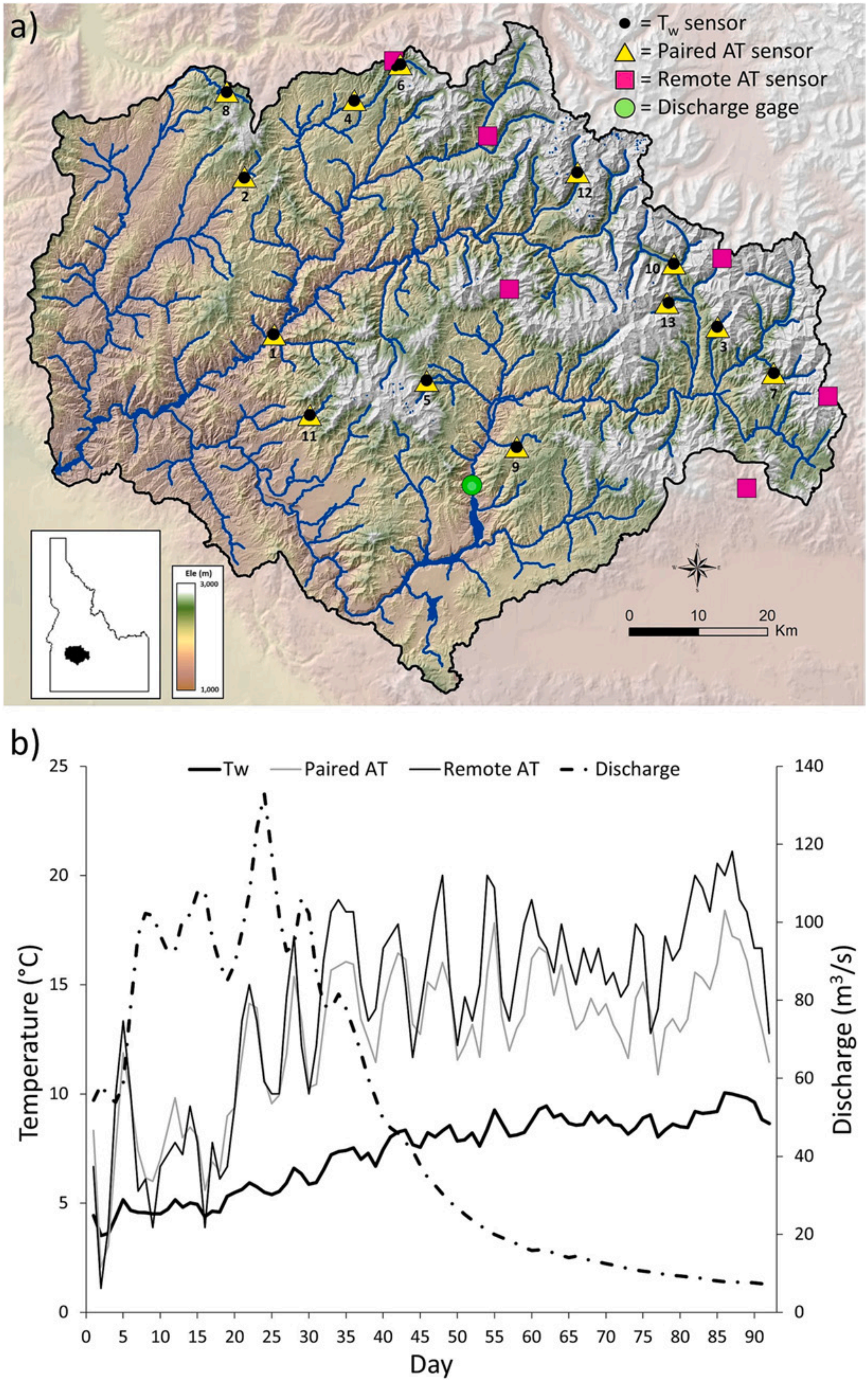

Fig. 1. Monitoring sites in the Boise River basin of central Idaho where temperature and stream discharge records were collected from June 1 through August 31, 2011 (panel a). Mean daily T_w, AT, and Q records at site 7 on Big Peak Creek, which were indicative of general trends at all stream sites during the study period (panel b).

6,900 km^2 and is drained by 2,500 km of perennial streams that range in elevation from 900 to 2,500 m (Isaak et al., 2010). The terrain is complex and most stream valleys are v-shaped, as is characteristic of landscapes underlain by the granitic Idaho Batholith (Hyndman, 1983; Meyer et al., 2001). Stream-side vegetation is dominated by coniferous trees and willow shrubs whereas hillslopes are covered by conifers, grasses, and exposed rock areas depending on elevation and aspect. The climate is strongly seasonal with warm, dry summers and cold winters during which deep snowpacks accumulate at the highest elevations. Runoff patterns are typical of snowmelt-driven streams and consist of a pronounced peak in discharge during April–July and lower baseflows during the remainder of the year (Fig. 1b). Human population densities are low within the basin and most of the area is publicly owned and federally administered by the U.S. Forest Service for a variety of land-use, recreational, and conservation purposes. Unpaved road networks have been developed in some drainages for timber harvest but many drainages are unroaded and have minimal anthropogenic effects.

2.2. T_w and air temperature data sets

Water temperature records were obtained from 13 stream sites throughout the Boise River network using miniature sensors (Onset Computer Corporation TidbiT sensor, accuracy of ±0.2 °C) that were attached to the downstream surfaces of large boulders with underwater epoxy (Isaak et al., 2013; Fig. 1a). Sensors were shielded to eliminate the bias that incident sunlight causes and were set to record temperatures 18 times daily during a three-month period from June 1 through August 31, 2011. Sampling encompassed the warmest portion of the year when most T_w monitoring is done in western North America and the effects of maximum temperatures on aquatic communities are of greatest concern (Isaak et al., 2017; Weller et al., 2023). At each stream site, a paired AT sensor (ThermoWorks Logtag® sensor TRIX-8, accuracy of ±0.5 °C) was attached to a large tree within 10–100 m of the T_w sensor, shielded from sunlight following the protocol developed by Holden et al. (2013), and set to record 18 daily measurements. All sensors were downloaded in September of 2011 and the temperature records were subject to quality control measures as described elsewhere (Chandler et al., 2016) before the multiple daily recordings were averaged to create mean daily temperature values for the 92 days composing the study period.

Additional daily AT records for the study period were obtained from six remote weather stations (1–105 km from T_w sites) that were in or adjacent to the basin and are part of the SNOTEL (Snowpack Telemetry) monitoring network in mountainous areas of the western U.S. (Table 1). Two of these stations occurred in valley locations (SNOTEL station IDs 496 and 769) and four occurred on ridgelines (SNOTEL station IDs 306, 450, 550, and 845). Mean daily AT values in these records were aligned with each of the 13 T_w records to create data sets for analysis. Mean daily ATs for the stream sites were also obtained from four gridded climate data sets that are freely available across much of North America and included PRISM, TopoWx (Topography Weather model), DayMet (Daily Meteorology model), and GridMet (Gridded Meteorology model). To match T_w records with those data, AT values were assigned from the grid cell that encompassed a T_w sensor site. Additional details concerning the gridded climate data sets, their spatiotemporal extents and resolutions, and supporting references are provided in Table 1.

Because stream discharge (Q) exhibited a strong trend during the study period (Fig. 1b) and often mediates the relationship between T_w and AT (Hockey et al., 1982; Webb et al., 2003; van Vliet 2011), we obtained a Q record from a flow gage on the South Fork of the Boise River (gage ID 13186000 in the National Water Information System https://waterdata.usgs.gov/nwis). Mean daily Q values were calculated from the record and aligned with the T_w and AT records described above so that Q could be used as an additional regression model covariate. The final data sets used for analysis, therefore, consisted of 13 T_w records, each of which was aligned with data from one Q source and 11 AT sources (a paired sensor, six remote stations, and four gridded climate data sets).

2.3. Analysis

Similarities among the AT, Q, and T_w records were described by calculating a pairwise correlation matrix. Linear regression models were used to estimate the effects of AT on T_w with models of two forms, one that included AT as the sole covariate ($T_w = b_0 + b_1 \cdot AT$) and one that also included Q as a covariate ($T_w = b_0 + b_1 \cdot AT + b_2 \cdot Q$) where b_0 was the intercept, b_1 was the slope parameter for AT, and b_2 was the slope parameter for Q. The predictive performance of the models was summarized using r^2 and the root mean square error (RMSE). The b_1 parameters were interpreted as estimates of T_{sens} as is common practice (Morrill et al., 2005; Hilderbrand et al., 2014; Luce et al., 2014). To further illustrate how temporal variability in AT records may have related to the T_{sens} estimates, the estimates were plotted against the standard deviations of the AT records at individual stream sites and correlations calculated.

3. Results

Stream sites were distributed throughout the river network across a range of elevations (1,054–2,099 m), reach slopes (0.44–13.9 %), and watershed sizes (6.34–1,978 km^2), and had mean T_w of 6.19 to 13.3 °C during the study (Table 2). The trend in daily T_w at the Big Peak Creek site, which was indicative of the other sites, showed warming in concert with AT during the first half of the study before reaching a plateau characterized by daily variability in the latter half (Fig. 1b). Stream Q was high during the first 30 days, dropped five-fold over the subsequent 30 days, and persisted at low levels thereafter.

Pairwise correlations among the AT records from the 11 sources were strong overall although several sub-patterns were notable (Table 3). Correlations among the remote weather station records were especially strong ($r \geq 0.94$) and suggested a general coherence of AT dynamics across the basin. However, these records were more weakly associated with measurements at AT sensors paired with stream sites ($r = 0.82$), which indicated a degree of near-stream microclimatic decoupling. Across all AT sources, TopoWx model predictions correlated most strongly with paired sensor records ($r = 0.88$) and GridMet predictions were weakest ($r = 0.78$). The correlation strength between T_w and AT records was greatest for the paired sensors, TopoWx, and PRISM, and weakest for the remote stations.

The T_w models developed with AT records resulted in good predictive performance, although models including Q were substantially better than those which omitted this covariate (Table 4). Models built with AT data from paired sensors only slightly outperformed those using data from PRISM, TopoWx, and DayMet, and this group of models was consistently better than those relying on remote station data. Estimates of T_{sens} from the AT + Q models are summarized in Fig. 2 and show two notable patterns. First, there was consistency across stream sites in terms of relative magnitudes, i.e., sites with low T_{sens} estimates using one AT source also had low estimates with the other AT sources. Second, substantial variability existed within sites where estimates could span two-fold differences depending on the AT source. Similar patterns within and among stream sites occurred in the T_{sens} estimates derived from the AT-only regression models (not shown in Fig. 2) except that these estimates were systematically higher when averaged across sites as shown in Fig. 3. Consistent across both model types, however, estimates derived from remote ridge stations, as well as PRISM and GridMet, were always lower than those from paired AT records. Estimates from TopoWx and DayMet were most similar to estimates from paired AT sensors, while those derived using AT data from the two remote valley stations showed mixed results (higher than the paired sensor estimates in the AT-only models, but similar to paired sensor estimates in the AT + Q models). Finally, plots of the relationships between T_{sens} estimates and the standard deviation of AT records used in regression model development

Table 2
Summary of environmental conditions and mean daily T_w records at 13 stream sites in the Boise River basin of central Idaho for June 1 through August 31, 2011.

Stream	Elevation (m)	Reach Slope (%)	Watershed size (km^2)	Temperature (°C) Mean	SD	Minimum	Maximum
1. MFK Boise River	1,054	0.44	1,978	13.3	3.64	7.79	19.3
2. Mores Creek	1,499	5.28	27.2	8.56	2.37	3.77	12.03
3. Paradise Creek	2,063	7.35	14.6	6.90	1.74	3.44	9.48
4. Pikes Fork	1,701	2.09	23.4	8.53	2.36	3.46	12.0
5. Trinity Cr	2,099	8.77	6.34	7.54	3.70	1.52	11.9
6. Crooked River	1,943	2.10	36.5	6.44	2.18	2.18	9.59
7. Big Peak Creek	1,928	3.83	45.8	7.31	1.77	3.52	10.0
8. Grimes Creek	2,008	1.44	17.1	7.17	3.37	1.01	11.6
9. Grouse Creek	1,666	2.94	28.2	11.9	2.25	5.36	15.2
10. Little Bear Creek	1,993	13.90	6.5	6.19	1.79	3.25	8.95
11. Rattlesnake Creek	1,527	5.50	24.6	9.07	1.87	4.49	12.1
12. Queens River	1,990	4.64	29.8	6.30	2.91	2.00	10.9
13. Bear Creek	2,079	3.02	20.5	6.30	2.34	2.54	9.67

Table 3
Average pairwise correlations among mean daily T_w, AT, and Q records that were collected from June 1 through August 31, 2011 for 13 stream sites in the Boise River basin.

	Paired sensor	Remote valley 1	Remote valley 2	Remote ridge 1	Remote ridge 2	Remote ridge 3	Remote ridge 4	PRISM	TopoWx	DayMet	GridMet	Q
Paired sensor	–											
Remote valley 1	0.82	–										
Remote valley 2	0.82	0.96	–									
Remote ridge 1	0.82	0.96	0.98	–								
Remote ridge 2	0.82	0.94	0.98	0.99	–							
Remote ridge 3	0.82	0.96	0.98	1.00	0.99	–						
Remote ridge 4	0.82	0.96	0.97	0.99	0.98	0.99	–					
PRISM	0.82	0.78	0.83	0.80	0.82	0.80	0.79	–				
TopoWx	0.88	0.84	0.86	0.86	0.86	0.86	0.86	0.89	–			
DayMet	0.83	0.89	0.89	0.90	0.89	0.90	0.90	0.84	0.87	–		
GridMet	0.78	0.78	0.77	0.80	0.79	0.81	0.82	0.68	0.88	0.78	–	
Q	−0.52	−0.55	−0.59	−0.63	−0.64	−0.64	−0.64	−0.59	−0.58	−0.58	−0.56	–
T_w	0.80	0.58	0.60	0.61	0.62	0.61	0.61	0.79	0.79	0.74	0.68	−0.63

Table 4
Performance metrics across 13 stream sites for T_w regression models developed using AT data from 11 sources.

Air temperature data source	Model (T_w = AT) r^2	RMSE (°C)	Model (T_w = AT + Q) r^2	RMSE (°C)
Paired	0.74	1.27	0.92	0.70
Remote valley 1	0.62	1.55	0.88	0.86
Remote valley 2	0.67	1.44	0.89	0.81
Remote ridge 1	0.69	1.40	0.88	0.86
Remote ridge 2	0.71	1.35	0.89	0.81
Remote ridge 3	0.70	1.39	0.88	0.85
Remote ridge 4	0.69	1.40	0.88	0.87
PRISM	0.73	1.32	0.91	0.76
TopoWx	0.75	1.24	0.91	0.74
DayMet	0.72	1.34	0.89	0.80
GridMet	0.59	1.61	0.83	1.02
Average:	0.69	1.39	0.89	0.83

showed strong negative correlations indicating that, as expected, T_{sens} estimates were strongly affected by the variability in AT records (Fig. 4).

4. Discussion

Our results highlight the important effects that the choice of an AT data source can have on T_w model results. Differences in predictive performance were expected because these sources vary in resolution, locations relative to stream monitoring sites, and ability to accurately represent near-stream environments. However, if it is assumed that data from paired AT sensors were most accurate, then T_{sens} estimates derived from some sources appear to be biased. If these biases are consistent among sites within a study, it would not degrade comparative inference but inter-study comparisons could be affected when AT sources differ. Also potentially problematic are individual studies which employ mixtures of AT data sources. These mixtures could occur spatially, for example, when AT predictions based on a gridded climate data set are used at a subset of T_w sites in lieu of paired sensors which are deployed at the remainder of sites. Problematic mixtures could also occur temporally, as when site-specific T_w models are calibrated to historical AT variability measured at monitoring stations and the resulting T_{sens} estimates are used to scale AT forecasts from GCMs that may exhibit different variability (Daly et al., 2010; MRI, 2015). In this instance, contextualizing T_w forecasts by conditioning on the expectation that a specific AT increase is realized at the stream monitoring sites is probably most appropriate. An alternative may be sourcing AT data from historical reanalysis datasets that are generated by the same dynamic climate models used in GCM forecasts (Compo et al., 2011; Hersbach et al., 2020). Such an approach would provide a consistent AT data context but accuracy assessments would first be required to understand the utility of reanalysis data sets for T_w modeling given their relatively coarse grid cell sizes.

Those latter points highlight the challenges of modeling T_w in topographically complex terrains which exhibit a diversity of microclimates related to landscape position. Because mountain streams are located in valleys adjacent to steep hillslopes, they often experience cold-air drainage, persistent thermal inversions and a degree of decoupling from these phenomena and broader synoptic weather patterns due to the insulating effects of riparian vegetation canopies. The end result is that AT is usually less variable near streams than other locations on or above mountain landscapes. As our results indicate, this dampened variability translates to T_{sens} estimates which are higher when derived

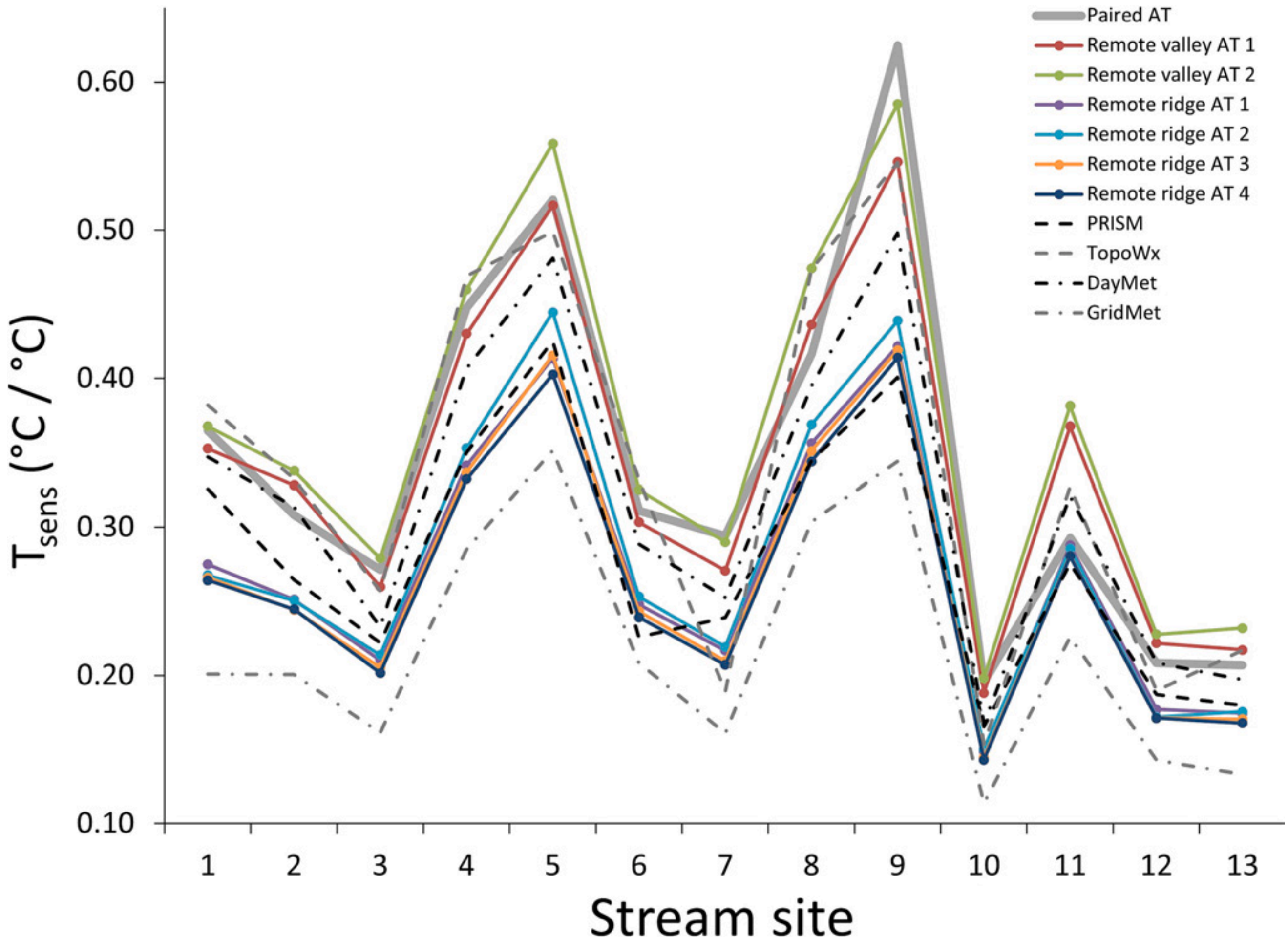

Fig. 2. Water temperature sensitivity estimates at 13 stream sites that were derived from linear regression models using 11 sources of AT data. Regression models also included a covariate for stream Q to control for trends during the study period.

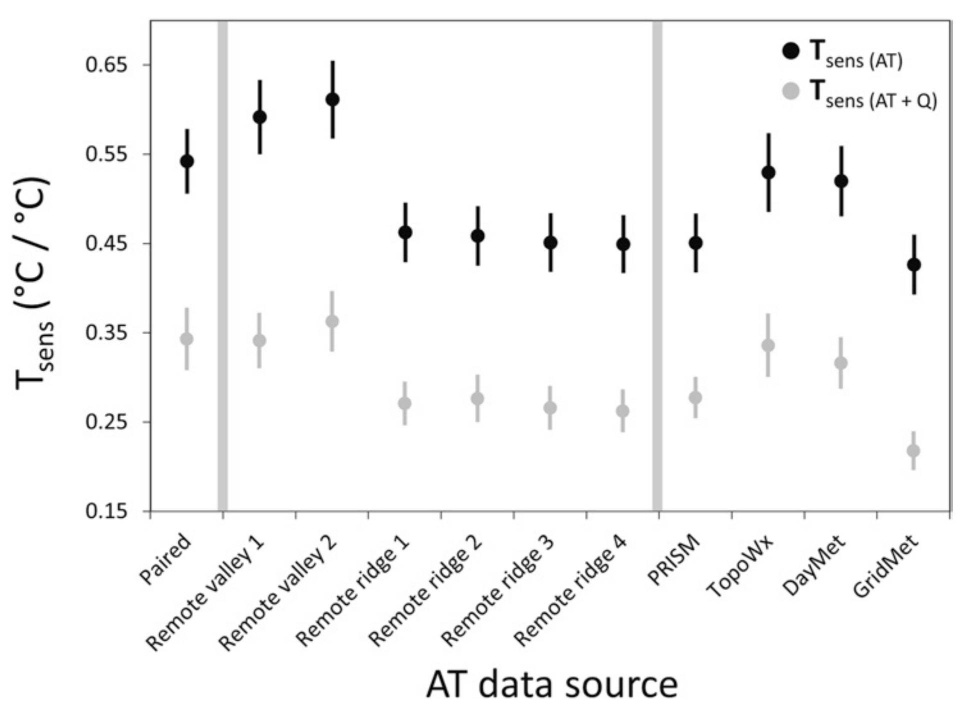

Fig. 3. Mean T_w sensitivity estimates (+/−1 standard error) across 13 stream sites from 11 regression models using different sources of AT data. Grey bars demarcate classes of AT data sources.

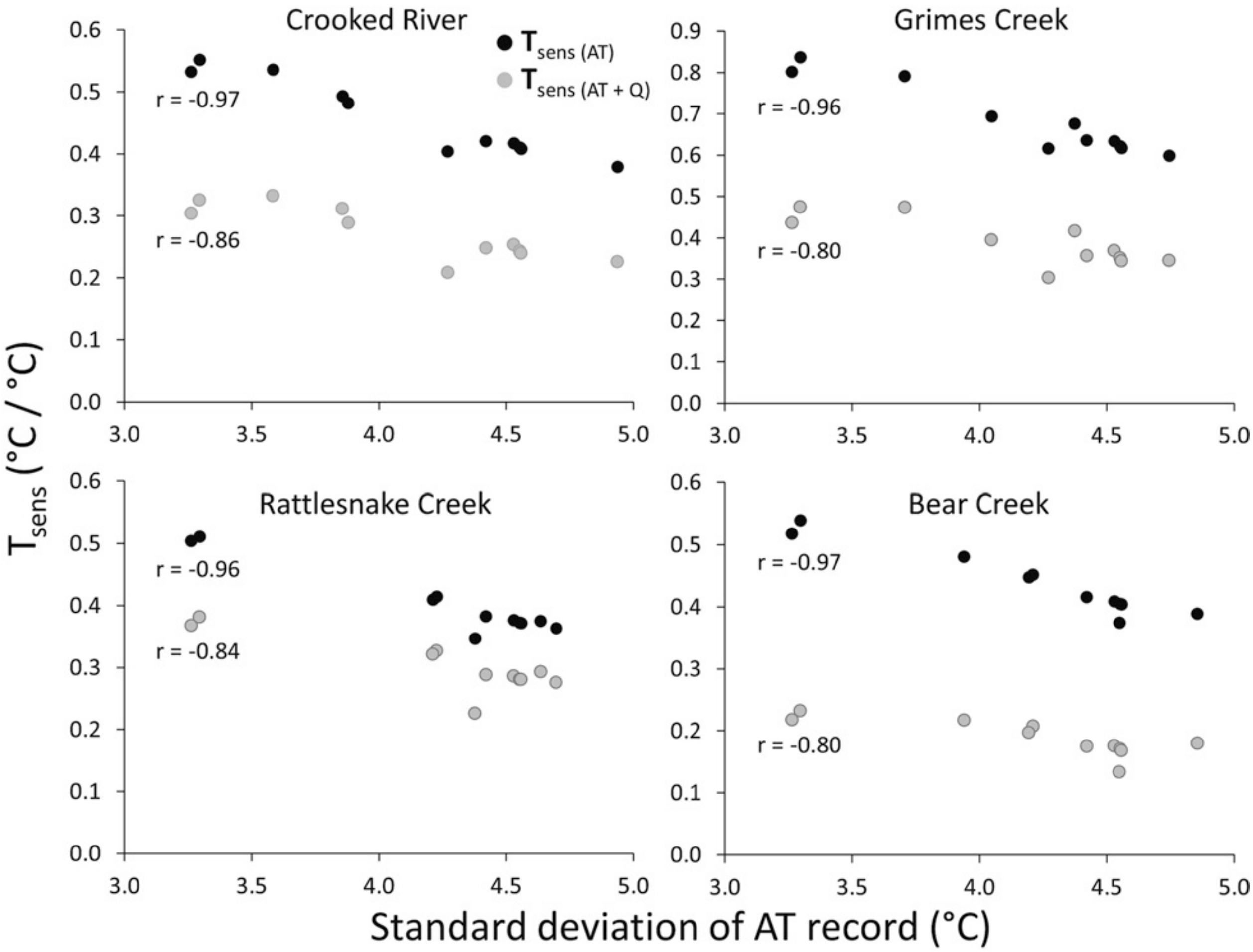

Fig. 4. Correlations between T_w sensitivity estimates at four stream sites and the temporal variability in AT records from 11 sources used to derive estimates. Similarly strong negative correlations between T_w sensitivity and AT variability occurred at the other nine stream sites not shown.

from paired AT sources compared to some alternative sources. Although we documented this pattern based on records with a daily resolution from a single summer, the expectation is for similar results with T_w models that predict longer timesteps (e.g., weekly, monthly, annual) or those calibrated to longer records based on persistent spatial differences in AT variability which are common to mountains (Lundquist and Cayan, 2007; Pepin et al., 2011), and which we also observed in our AT sources when those records were examined over a longer twelve-year period in the study basin (unpublished analysis).

Because of the differences among AT sources, some are more advantageous than others for specific T_w modeling efforts and factors such as accuracy, cost, convenience, and flexibility should be considered in their application. Paired AT sensors have emerged as a popular choice and provide accurate descriptions of the AT variability which occurs at T_w monitoring sites (Benyahya et al., 2010). However, the dedication of two sensors to each stream site also means that half as many sites are monitored and the amount of variation encompassed by data sets of this type are reduced. The use of data from remote weather stations avoids this limitation but may bias T_{sens} estimates depending on station location. In particular, it seems best to avoid the use of ridgeline AT stations if a primary goal is estimation of the T_{sens} parameter as underestimates are likely. Nonetheless, the predictive accuracy of T_w models developed with remote weather station data may not be markedly worse than other AT sources, which is attractive because these stations enable the use of long-term records with shorter term T_w records in historical reconstructions and trend assessments (Moatar and Gailhard, 2006; Isaak et al., 2012). These reconstructions should be unbiased as long as the predictive accuracy of the T_w model is high and calibration occurs over a period long enough to encompass a breadth of hydroclimatological variability (Jones and Schmidt, 2018).

The use of AT data from gridded climate data sets potentially offers the fewest compromises, although researchers should be aware of the nuances among these sources. Two of these data sets, TopoWx and DayMet, resulted in T_w models with unbiased T_{sens} estimates and predictive accuracy close to that of models developed with paired AT data. Daily AT values are available from both sources at ≤ 1 km^2 resolution dating back to at least 1980, although only DayMet spans the entirety of North American and is regularly updated to maintain currency. PRISM data yielded models with good predictive accuracy but low T_{sens} estimates, perhaps an artefact of the coarser resolution 4 km^2 data grids we used that less accurately represented conditions in narrow mountain valleys. It should be noted that 0.8 km^2 PRISM data sets are available for a fee but were not considered here. GridMet data resulted in T_w models with the poorest predictive accuracy and lowest T_{sens} estimates. However, this data set also had lower spatial resolution (4 km^2) and its temporal variability is derived from dynamic climate models (Abatzoglou, 2013) rather than the interpolation routines applied by other models to observed AT measurements.

Although not the primary focus of our study, the use of Q as a covariate in our site-specific T_w models substantially affected their predictive accuracy and the magnitude of T_{sens} estimates. Other researchers have consistently observed T_w model improvements when using a Q covariate (Webb et al., 2003; van Vliet et al., 2011; Hilderbrand et al., 2014), yet it remains common for it to be excluded from T_w models (Trumbo et al., 2014; Caldwell et al., 2015; Johnson et al., 2020; Molinos et al., 2022). This likely results in another avoidable source of inter-study variability because exclusion of covariates that are strongly correlated with a response is known to affect parameter estimates and is referred to as the Omitted Variable Bias problem in the statistical literature (Wilms et al., 2021). Use of a Q covariate could be less important where variability in runoff is limited but this circumstance may also be rare for mountain river networks where snowmelt is usually a key hydrological component and T_w models are calibrated during transitional summer flow periods. Unfortunately, the expense of operating Q gages means that co-location with T_w sensors will remain unlikely at most of the places where stream temperatures are monitored. As our results and those from previous studies demonstrate, however, Q data from remote gages can be a viable substitute (Isaak et al., 2010, 2017; Luce et al., 2014; Hilderbrand et al., 2014). Alternatively, Q predictions from spatially distributed hydrologic models such as the National Water Model are becoming increasingly common (Salas et al., 2018; Office of Water Prediction, 2021) and these could be matched with T_w sites where available (e.g., Siegel et al., 2022).

5. Conclusions

T_w modelers have many options when sourcing AT data sets and should carefully consider their strengths and weaknesses when choosing among them. Although no single source of AT data is ideal, some result in greater predictive accuracy and less estimation bias than others. In our view, AT values from gridded climate data sets at fine spatial resolutions are particularly attractive because of reduced site instrumentation costs (i.e., paired sensors are unnecessary) and the availability of long-term historical record reconstructions across large swaths of North America. If broadly adopted by T_w modelers, the best of these—TopoWx and DayMet—could provide a consistent foundation for representing spatiotemporal variability in AT that would enable more robust and consistent inference regarding T_w dynamics and regulating factors in mountain landscapes. Although our results are derived from a single river basin and therefore should be replicated elsewhere to confirm their generality, they are also based on well understood and broadly documented patterns of AT variability in mountainous terrain (Whiteman, 2000; Daly et al., 2010; Pepin et al., 2011). With the rapid growth and availability of AT and T_w data from mountainous areas globally (Gallice et al., 2015; Ouellet et al., 2020), such replication should be straightforward and presents additional learning opportunities.

CRediT authorship contribution statement

Daniel J. Isaak: Conceptualization, Methodology, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Dona L. Horan:** Data curation, Formal analysis, Resources, Visualization, Writing – review & editing. **Sherry P. Wollrab:** Data curation, Investigation, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors acknowledge the Rocky Mountain Research Station of the U.S. Forest Service for provision of funding to collect study data sets and salary during preparation of this manuscript. The authors also thank Zack Holden for technical support and training associated with the protocol for AT data collection at stream sites.

References

Abatzoglou, J.T., 2013. Development of gridded surface meteorological data for ecological applications and modelling. Int. J. Climatol. 33, 121–131.

Benyahya, L., Caissie, D., St-Hilaire, A., Ouarda, T.B., Bobee, B., 2007. A review of statistical water temperature models. Canadian Water Resour. J. 32, 179–192.

Benyahya, L., Caissie, D., El-Jabi, N., Satish, M.G., 2010. Comparison of microclimate vs. remote meteorological data and results applied to a water temperature model (Miramichi River, Canada). J. Hydrol. 380, 247–259.

Briggs, M.A., Johnson, Z.C., Snyder, C.D., Hitt, N.P., Kurylyk, B.L., Lautz, L., Irvine, D.J., Hurley, S.T., Lane, J.W., 2018. Inferring watershed hydraulics and cold-water habitat persistence using multi-year air and stream temperature signals. Sci. Total Environ. 636, 1117–1127.

Caldwell, P., Segura, C., Gull, L.S., Sun, G., McNulty, S.G., Sandercock, M., Boggs, J., Vose, J.M., 2015. Short-term stream water temperature observations permit rapid assessment of potential climate change impacts. Hydrol. Process. 29 (9), 2196–2211.

Chandler, G.L., Wollrab, S., Horan, D., Nagel, D., Parkes, S., Isaak, D.J., Wenger, S.J., Peterson, E.E., Ver Hoef, J.M., Hostetler, S.W., Luce, C.H., Dunham, J.B., Kershner, J. L., Roper, B.B. 2016. NorWeST stream temperature data summaries for the western U.S. Fort Collins, CO: U.S. Forest Service, Rocky Mountain Research Station Research Data Archive. https://doi.org/10.2737/RDS-2016-0032.

Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Matsui, N., Allan, R.J., Yin, X., Gleason, B.E., Vose, R.S., Rutledge, G., Bessemoulin, P., Brönnimann, S., 2011. The twentieth century reanalysis project. Q. J. R. Meteorolog. Soc. 137 (654), 1–28.

Daly, C., Gibson, W.P., Taylor, G.H., Johnson, G.L., Pasteris, P., 2002. A knowledge-based approach to the statistical mapping of climate. Clim. Res. 22, 99–113.

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. Int. J. Climatol. 28, 2031–2064.

Daly, C., Conklin, D.R., Unsworth, M.H., 2010. Local atmospheric decoupling in complex topography alters climate change impacts. Int. J. Climatol. 30, 1857–1864.

Gallice, A., Schaefli, B., Lehning, M., Parlange, M.B., Huwald, H., 2015. Stream temperature prediction in ungauged basins: Review of recent approaches and description of a new physics-derived statistical model. Hydrol. Earth Syst. Sci. 19, 3727–3753.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., 2020. The ERA5 global reanalysis. Q. J. R. Meteorolog. Soc. 146 (730), 1999–2049. https://doi.org/10.1002/qj.3803.

Hilderbrand, R.H., Kashiwagi, M.T., Prochaska, A.P., 2014. Regional and local scale modeling of stream temperatures and spatio-temporal variation in thermal sensitivities. Environ. Manag. 54, 14–22.

Himmelblau, D.M., 1960. Solubilities of inert gases in water: 0°C to near the critical point of water. J. Chem. Eng. Data 5, 10–15.

Hockey, J.B., Owens, I.F., Tapper, N.J., 1982. Empirical and theoretical models to isolate the effect of discharge on summer water temperatures in the Hurunui River. J. Hydrol. (New Zealand) 21, 1–12.

Holden, Z.A., Klene, A.E., Keefe, R.F., Moisen, G.G., 2013. Design and evaluation of an inexpensive radiation shield for monitoring surface air temperatures. Agric. For. Meteorol. 180, 281–286.

Holden, Z.A., Swanson, A., Klene, A.E., Abatzoglou, J.T., Dobrowski, S.Z., Cushman, S.A., Squires, J., Moisen, G.G., Oyler, J.W., 2016. Development of high-resolution (250 m) historical daily gridded air temperature data using reanalysis and distributed sensor networks for the US northern Rocky Mountains. Int. J. Climatol. 36, 3620–3632.

Hyndman, D.W., 1983. The Idaho batholith and associated plutons, Idaho and Western Montana. Circum-Pacific Plutonic Terranes 159, 213.

Isaak, D.J., Horan, D.L., Wollrab, S. 2013. A simple protocol using underwater epoxy to install annual temperature monitoring sites in rivers and streams. U.S. Forest Service, Rocky Mountain Research Station, GTR-RMRS-314, Fort Collins, CO.

Isaak, D.J., Luce, C.H., 2023. Elevation-dependent warming of streams in mountainous regions: Implications for temperature modeling and headwater climate refugia. Canadian Water Resour. J. 48, 167–188.

Isaak, D.J., Luce, C.H., Rieman, B.E., Nagel, D.E., Peterson, E.E., Horan, D.L., Parkes, S., Chandler, G.L., 2010. Effects of climate change and recent wildfires on stream temperature and thermal habitat for two salmonids in a mountain river network. Ecol. Appl. 20, 1350–1371.

Isaak, D.J., Rieman, B.E., 2013. Stream isotherm shifts from climate change and implications for distributions of ectothermic organisms. Glob. Chang. Biol. 19, 742–751.

Isaak, D.J., Wollrab, S., Horan, D., Chandler, G., 2012. Climate change effects on stream and river temperatures across the northwest US from 1980–2009 and implications for salmonid fishes. Clim. Change 113, 499–524.

Isaak, D.J., Wenger, S.J., Peterson, E.E., Ver Hoef, J.M., Nagel, D.E., Luce, C.H., Hostetler, S.W., Dunham, J.B., Roper, B.B., Wollrab, S.P., Chandler, G.L., 2017. The NorWeST summer stream temperature model and scenarios for the western US: A crowd-sourced database and new geospatial tools foster a user community and predict broad climate warming of rivers and streams. Water Resour. Res. 53, 9181–9205.

Jacobsen, D., 2020. The dilemma of altitudinal shifts: Caught between high temperature and low oxygen. Front. Ecol. Environ. 18 (4), 211–218.

Johnson, Z.C., Johnson, B.G., Briggs, M.A., Devine, W.D., Snyder, C.D., Hitt, N.P., Hare, D.K., Minkova, T.V., 2020. Paired air-water annual temperature patterns reveal hydrogeological controls on stream thermal regimes at watershed to continental scales. J. Hydrol. 587, 124929.

Jones, N.E., Schmidt, B.J., 2018. Thermal regime metrics and quantifying their uncertainty for North American streams. River Res. Appl. 34, 382–393.
Kelleher, C., Wagener, T., Gooseff, M., McGlynn, B., McGuire, K., Marshall, L., 2012. Investigating controls on the thermal sensitivity of Pennsylvania streams. Hydrol. Processes 26, 771–785.
Luce, C.H., Staab, B., Kramer, M., Wenger, S., Isaak, D., McConnell, C., 2014. Sensitivity of summer stream temperatures to climate variability in the Pacific Northwest. Water Resour Res. 50, 3428–3443.
Lundquist, J.E., Cayan, D., 2007. Surface temperature patterns in complex terrain: daily variations and long-term change in the central Sierra Nevada, California. J. Geophys. Res., 112: D11124.
Mauger, S., Shaftel, R., Leppi, J.C., Rinella, D.J., 2017. Summer temperature regimes in southcentral Alaska streams: watershed drivers of variation and potential implications for Pacific salmon. Can. J. Fish. Aquat. Sci. 74 (5), 702–715.
Menne, M.J., Williams Jr, C.N., Vose, R.S., 2009. The US historical climatology network monthly temperature data, version 2. Bull. Am. Meteorol. Soc., 90, 993-1008.
Meyer, G.A., Pierce, J.L., Wood, S.H., Jull, A.J.T., 2001. Fire, storms, and erosional events in the Idaho batholith. Hydrol. Process. 15, 3025–3038.
Mitchell, K.E., et al., 2004. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. J. Geophys. Res. 109, D07S90.
Moatar, F., Gailhard, J., 2006. Water temperature behaviour in the River Loire since 1976 and 1881. C. R. Geosci. 338 (5), 319–328.
Mohseni, O., Stefan, H.G., Erickson, T.R., 1998. A nonlinear regression model for weekly stream temperatures. Water Resour. Res. 34, 2685–2692.
Mohseni, O., Erickson, T.R., Stefan, H.G., 1999. Sensitivity of stream temperatures in the United States to air temperatures projected under a global warming scenario. Water Resour. Res. 35 (12), 3723–3733.
Mohseni, O., Stefan, H.G., Eaton, J.G., 2003. Global warming and potential changes in fish habitat in US streams. Clim. Change 59 (3), 389–409.
Molinos, J.G., Nobuo, I., Sueyoshi, M., Nakamura, F., 2022. Timescale mediates the effects of environmental controls on water temperature in mid-to low-order streams. Sci. Rep. 12, 12248.
Moore, R.D., Spittlehouse, D.L., Story, A., 2005. Riparian microclimate and stream temperature response to forest harvesting: a review. J Am Water Resour Ass. 41, 813–834.
Morrill, J., Bales, R., Conklin, M., 2005. Estimating stream temperature from air temperature: Implications for future water quality. J. Environ. Eng. 131, 139–146.
MRI (Mountain Research Initiative), 2015. Elevation-dependent warming in mountain regions of the world. Nature. Clim. Change 5, 424–430.
Nichols, R.A., Ketcheson, G.L., 2013. A two-decade watershed approach to stream restoration log jam design and stream recovery monitoring: Finney Creek, Washington. J. Am. Water Resour. Assoc. 49, 1367–1384.
Nussle, S., Matthews, K.R., Carlson, S.M., 2015. Mediating water temperature increases due to livestock and global change in high elevation meadow streams of the Golden Trout Wilderness. PLoS One 10, e0142426.
Office of Water Prediction, 2021. The National Water Model. National Oceanic and Atmospheric Administration. https://water.noaa.gov/about/nwm.
Ouellet, V., St-Hilaire, A., Dugdale, S.J., Hannah, D.M., Krause, S., Proulx-Ouellet, S., 2020. River temperature research and practice: Recent challenges and emerging opportunities for managing thermal habitat conditions in stream ecosystems. Sci. Total Environ. 736, 139679.
Oyler, J.W., Ballantyne, A., Jencso, K., Sweet, M., Running, S.W., 2014. Creating a topoclimatic daily air temperature dataset for the conterminous United States using homogenized station data and remotely sensed land skin temperature. Int. J. Climatol. https://doi.org/10.1002/joc.4127.
Oyler, J.W., Dobrowski, S.Z., Holden, Z.A., Running, S.W., 2016. Remotely sensed land skin temperature as a spatial predictor of air temperature across the conterminous United States. J. Appl. Meteorol. Climatol. https://doi.org/10.1175/JAMC-D-15-0276.1.
Pepin, N.C., Daly, C., Lundquist, J., 2011. The influence of surface versus free-air decoupling on temperature trend patterns in the western United States. J. Geophys. Res. 116, D10109.
Piccolroaz, S., Calamita, E., Majone, B., Gallice, A., Siviglia, A., Toffolon, M., 2016. Prediction of river water temperature: a comparison between a new family of hybrid models and statistical approaches. Hydrol. Process. 30 (21), 3901–3917.
Pörtner, H.O., Peck, M.A., 2010. Climate change effects on fishes and fisheries: towards a cause-and-effect understanding. J. Fish Biol. 77 (8), 1745–1779.
Rey, D.M., Briggs, M.A., Walvoord, M.A., Ebel, B.A., 2023. Wildfire-induced shifts in groundwater discharge to streams identified with paired air and stream water temperature analyses. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2023.129272.
Salas, F.R., Somos-Valenzuela, M.A., Dugger, A., Maidment, D.R., Gochis, D.J., David, C. H., Yu, W., Ding, D., Clark, E.P., Noman, N., 2018. Towards real-time continental scale streamflow simulation in continuous and discrete space. J. Am. Water Resour. Assoc. 54 (1), 7–27.
Schaefer, G.L., Johnson, D.E., 1992. Development and operation of the SNOTEL system in the western United States. Pages 29-48 in United States/People's Republic of China Flood Forecasting Symposium Portland, Oregon March 29-April 4, 1989, vol. 1. Office of Hydrology, National Weather Service, Washington D.C.
Serreze, M.C., Clark, M.P., Armstrong, R.L., McGinnis, D.A., Pulwarty, R.S., 1999. Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. Water Resour. Res., 35, 2145-2160.
Siegel, J.E., Fullerton, A.H., Jordan, C.E., 2022. Accounting for snowpack and time-varying lags in statistical models of stream temperature. J. Hydrol. X 17, 100136.
Stefan, H.G., Preud'homme, E.B., 1993. Stream temperature estimation from air temperature. J. Am. Water Resour. Assoc. 29, 27–45.
St-Hilaire, A., Caissie, D., Bergeron, N.E., Ouarda, T.B., Boyer, C., 2021. Climate change and extreme river temperature. In: Climate Change and Extreme Events. Elsevier, Amsterdam, pp. 25–37.
Thornton, P.E., Running, S.W., White, M.A., 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. J. Hydrol., 190, 214–251.
Thornton, P.E., Thornton, M.M., Mayer, B.W., Wilhelmi, N., Wei, Y., Cook, R.B., 2012. Daymet: Daily Surface Weather on A 1 km Grid for North America, 1980–2012. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, TN doi: 10.3334/ORNLDAAC/Daymet_V2. http://daymet.ornl.gov/ (accessed 19 November 2013).
Thornton, P.E., Shrestha, R., Thornton, M., Kao, S.C., Wei, Y., Wilson, B.E., 2021. Gridded daily weather data for North America with comprehensive uncertainty quantification. Sci. Data 8, 1–17.
Trumbo, B.A., Nislow, K.H., Stallings, J., Hudy, M., Smith, E.P., Kim, D.Y., Wiggins, B., Dolloff, C.A., 2014. Ranking site vulnerability to increasing temperatures in southern Appalachian brook trout streams in Virginia: an exposure-sensitivity approach. Trans. Am. Fish. Soc. 143, 173–187.
USDA Soil Conservation Service, 1988. Snow Survey and Water Supply Products Reference. West National Technical Center, Snow Survey Program, Portland, Oregon, 12 pp.
van Vliet, M.T.H., Ludwig, F., Zwolsman, J.J.G., Weedon, G.P., Kabat, P., 2011. Global river temperatures and sensitivity to atmospheric warming and changes in river flow. Water Resour. Res. 47 (2), W02544.
Verberk, W.C., Bilton, D.T., Calosi, P., Spicer, J.I., 2011. Oxygen supply in aquatic ectotherms: partial pressure and solubility together explain biodiversity and size patterns. Ecology 92 (8), 1565–1572.
Webb, B.W., Clack, P.D., Walling, D.E., 2003. Water-air temperature relationships in a Devon river system and the role of flow. Hydrol. Proc. 17, 3069–3084.
Webb, B.W., Nobilis, F., 1997. Long-term perspective on the nature of the air–water temperature relationship: a case study. Hydrol. Process. 11 (2), 137–147.
Webb, B.W., Nobilis, F., 2007. Long-term changes in river temperature and the influence of climatic and hydrological factors. Hydrol. Sci. J. 52, 74–85.
Webb, B.W., Zhang, Y., 1997. Spatial and seasonal variability in the components of the river heat budget. Hydrol Proc. 11, 79–101.
Weber, N., Bouwes, N., Pollock, M.M., Volk, C., Wheaton, J.M., Wathen, G., et al., 2017. Alteration of stream temperature by natural and artificial beaver dams. PLoS ONE 12 (5), e0176313.
Weller, J.D., Moore, R.D., Iacarella, J.C., 2023. Stream thermalscape scenarios for British Columbia, Canada. Can. Water Resour. J. https://doi.org/10.1080/07011784.2023.2267028.
Whiteman, C.D., 2000. Mountain Meteorology: Fundamentals and Applications. Oxford University Press.
Wilms, R., Mäthner, E., Winnen, L., Lanwehr, R., 2021. Omitted variable bias: a threat to estimating causal relationships. Methods Psychol. 5, 100075.
Zhu, S., Piotrowski, A.P., 2020. River/stream water temperature forecasting using artificial intelligence models: a systematic review. Acta Geophys. 68, 1433–1442.