Research papers

Paired air-water annual temperature patterns reveal hydrogeological controls on stream thermal regimes at watershed to continental scales

Zachary C. Johnson[a],*, Brittany G. Johnson[a], Martin A. Briggs[b], Warren D. Devine[c],
Craig D. Snyder[d], Nathaniel P. Hitt[d], Danielle K. Hare[e], Teodora V. Minkova[c]

[a] School of Environmental and Forest Sciences, University of Washington, 3715 West Stevens Way NE, Seattle, WA 98195, USA
[b] U.S. Geological Survey, Earth System Processes Division, Hydrogeophysics Branch, 11 Sherman Place, Unit 5015, Storrs, CT 06269, USA
[c] Washington State Department of Natural Resources, Forest Resources Division, 1111 Washington Street SE, Olympia, WA 98504, USA
[d] U.S. Geological Survey, Leetown Science Center, 11649 Leetown Road, Kearneysville, WV 25430, USA
[e] Department of Natural Resources and the Environment, University of Connecticut, 1376 Storrs Road, Unit 4087, Storrs, CT 06269, USA

ABSTRACT

Despite decades of research into air and stream temperature dynamics, paired air-water annual temperature signals have been underutilized to characterize watershed processes. Annual stream temperature dynamics are useful in classifying fundamental thermal regimes and can enhance process-based interpretation of stream temperature controls, including deep and shallow groundwater discharge, when paired with air signals. In this study, we investigated multi-scale variability in annual paired air-water temperature patterns using sine-wave linear regressions of multi-year daily temperature data from streams of various sizes. A total of 311 sites from two spatially intensive regional datasets (Shenandoah National Park and Olympic Experimental State Forest) and a spatially extensive national dataset spanning the contiguous United States (U.S. Geological Survey gages) were evaluated. We calculated three annual air-water thermal metrics (mean ratio, phase lag, and amplitude ratio) to deduce the influence of groundwater and other watershed processes on stream thermal regimes at multiple spatial scales. Site-specific values of the three annual air-water thermal metrics ranged from 0.69 to 5.29 (mean ratio), −9 to 40 days (phase lag), and 0.29 to 1.12 (amplitude ratio). Regional patterns in the annual thermal metrics revealed persistent yet spatially variable influences of shallow groundwater discharge and high levels of thermal variability within watersheds, indicating the importance of local hydrogeological controls on stream temperature. Furthermore, annual thermal metric patterns from the regional datasets were generally concordant with the national dataset suggesting the utility of these annual thermal metrics for analysis at multiple scales. Analysis of the national dataset showed that previously defined thermal regimes based on water temperature alone could be further refined using air-water metrics and these metrics were related to physiographic watershed characteristics such as contributing area, elevation, and slope. This research demonstrates the importance of spatial scale and heterogeneity for inferring hydrological process in streams and provides guidance for the interpretation of annual air-water temperature metrics that can be efficiently applied to the growing database of multi-year temperature records. Results from this research can aid in the prediction of future thermal habitat suitability for coldwater-adapted species at ecologically and management-relevant spatial scales with readily available data.

1. Introduction

As climate change and other anthropogenic alterations to watersheds transform the natural thermal regime of streams and rivers (Bassar et al., 2016; Isaak et al., 2012; Kaushal et al., 2010; Kędra and Wiejaczka, 2018), understanding the sensitivity and vulnerability of stream segments to these changes is increasingly imperative for long-term ecological management. Temperature is one of the most important properties controlling the water quality of streams (Caissie, 2006). Almost all physical, chemical, and biological processes of the stream corridor are influenced by temperature, including fish development and metabolism, dissolved oxygen concentration, biogeochemical cycling, and organic matter decomposition. Absent hydrologic alteration, channel water temperature is driven by meteorological and

---

Nomenclature

A, A_a, A_w	Annual temperature amplitude (air or water)
A_R	Annual temperature amplitude ratio
a, b	Linear regression coefficients
M_R	Annual temperature mean ratio
T, T_a, T_w	Mean daily temperature (air or water)
T_0, $T_{0,a}$, $T_{0,w}$	Annual temperature mean (air or water)
t	Time
ϵ	Linear regression error term
ϕ, ϕ_a, ϕ_w	Annual temperature phase (air or water)
$\Delta\phi$	Annual temperature phase lag

ω	Angular frequency
ANOVA	Analysis of variance
BFI	Base-flow Index
GW	Groundwater
NOAA	National Oceanic and Atmospheric Administration
OESF	Olympic Experimental State Forest
PRISM	Parameter-elevation regressions on independent slopes model
SHEN	Shenandoah National Park
USGS	U.S. Geological Survey
WADNR	Washington State Department of Natural Resources

---

hydrogeological factors, such as incoming solar radiation, outgoing longwave radiation, air temperature, wind speed, humidity, stream channel dimensions (depth, width, and flow volume), and groundwater (GW) inputs (Caissie, 2006; Cluis, 1972; Westhoff et al., 2007). Because of its effect on ecosystem processes, stream temperature has received abundant attention in the scientific literature for more than a half-century and numerous modeling approaches have been used to predict different facets of its regime (key reviews include Anderson, 2005; Benyahya et al., 2007a; Caissie, 2006; Gallice et al., 2015; Kurylyk et al., 2019; Webb et al., 2008). Short-term air-water temperature relations have been used to map spatially variable stream water thermal sensitivity in summer (Kelleher et al., 2012). However, longer-term annual water temperature patterns, and how these relate to local air temperature patterns and watershed processes, have been under-explored.

Stream temperature models fall into deterministic or statistical groups (Benyahya et al., 2007a; Caissie, 2006). Deterministic models are based on the balance of energy (heat) and mass (flow) fluxes in a water body (Boyd and Kasper, 2003; Glose et al., 2017). These models are best for conducting impact studies that assess changes to one or more components of the stream heat budget or when exploring changes in temperature at multiple spatial scales (Benyahya et al., 2007a; Caissie, 2006; Westhoff et al., 2007). However, total heat budget models are complex, computationally expensive, and require numerous hydrological, physiographic, and meteorological inputs that may be excluded from typical measurement protocols, poorly defined at the spatial scale needed for ecological applications, or difficult to quantify. In contrast, statistical models are computationally simpler with minimal data requirements (Benyahya et al., 2007a) facilitating prediction at ecologically relevant spatial grains and extents. However, statistical models currently lack a clear understanding of the relationships between derived model coefficients and important watershed processes potentially limiting their utility. Further, unsampled spatial heterogeneity can lead to overly simplistic predictions at multiple scales. For instance, the development of national-scale models, such as The National Water Model (NOAA, 2019), requires understanding of intra- and inter-regional spatial variation for meaningful downscaled predictions.

Statistical models typically use stochastic methods to assess relationships between water and air temperatures for small time-steps (e.g., daily). These models result in the derivation of a long-term periodic component of water temperature data, which is commonly assumed to be temporally stable (Caissie et al., 1998). However, inter-annual and long-term trends in water temperature may be driven by temporally variable watershed processes, including solar radiation, GW discharge dynamics from adjacent aquifers and temporal runoff patterns. While tracing heat signatures has long been used to monitor the activity of these and other contributing factors (Anderson, 2005; Constantz, 2008; Halloran et al., 2016), few studies have explored their influence on patterns and properties of the long-term water temperature signal (i.e., mean, phase, and amplitude). Together, these

properties can indicate the presence of stream thermal inertia (Letcher et al., 2016), influence of shallow (~upper 6 m) versus deeper GW discharge (Briggs et al., 2018b), riparian shading (Fabris et al., 2018; Johnson and Jones, 2000; Wawrzyniak et al., 2017), or dam operations (Rounds, 2007). However, it is unclear whether these observations are applicable across locations and climates and if annual stream temperature signals are relatively stable over time. Further, air temperature is often assumed to be a dominant control on stream temperature at seasonal timescales, but recent research indicates uncoupling of the annual signal amplitude and phase relations between air and water temperature may be used to infer other watershed controls on the stream heat budget (Briggs et al., 2018a).

Here, we assess the utility of simple statistical modeling approaches using multi-year paired air-water temperature data for developing ecologically relevant thermal metrics, and to evaluate how these diagnostic metrics vary across multiple spatial scales. The specific objectives are to: (1) compare annual air and water temperature signal parameters and combined air-water metrics across watershed, regional, and national spatial scales; and (2) provide guidance for the interpretation of three paired air-water annual thermal metrics (mean ratio, phase lag, and amplitude ratio) derived from sine-wave linear regressions with respect to physical watershed processes. These objectives are addressed using paired air and water temperature data collected at relatively high spatial resolution in mountain headwater streams from two different climatic regions of the United States (U.S.) (Pacific Northwest and Mid-Atlantic), along with data from streams of generally larger size distributed across the contiguous U.S. The use of the latter, national dataset, allows for a comparison between the paired air-water annual temperature signals method presented here and another multi-year thermal classification method that used the same sites (Maheu et al., 2016). This previously developed stream thermal regime classification system, using identified key environmental drivers, was based exclusively on water temperature data. Their results showed that sites could be clustered based on differences in the annual water temperature mean, amplitude, and phase parameters, and that these clusters could be predicted by air temperature and flow characteristics. We hypothesize that the inclusion of the comparison of annual water temperature patterns to local annual air temperature patterns in the present analysis will improve the ability to predict processes such as GW exchange and riparian vegetation characteristics that may vary across steep gradients or spatial discontinuities.

2. Data and methods

In this section, we describe the site characteristics of the regional and national datasets used in this study and methods for analyzing annual air and water temperature sine-wave signals with linear regression.
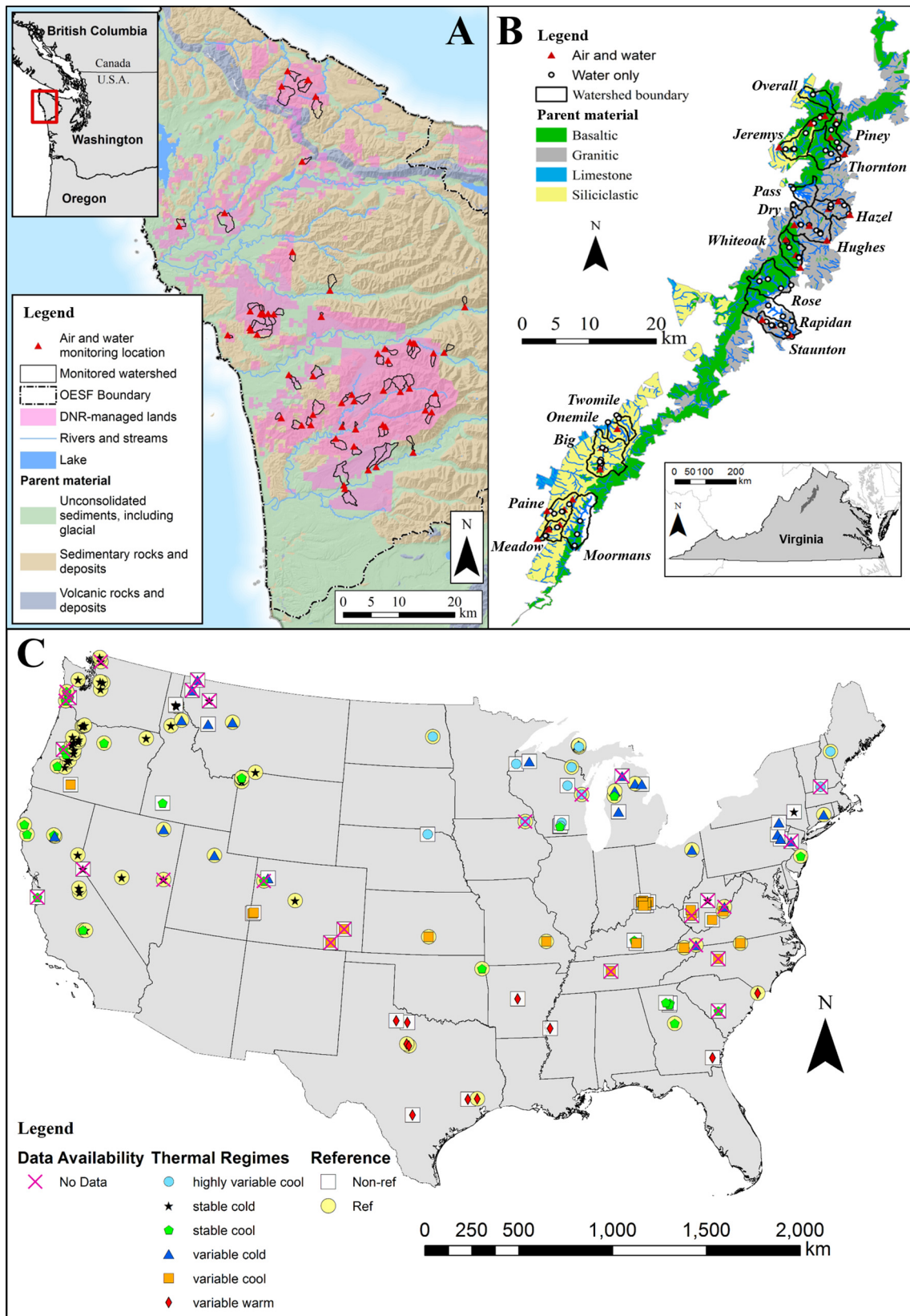
Fig. 1. Locations of the (A) OESF (n = 56), (B) SHEN (n = 120), and (C) USGS (n = 135) sites. "DNR" in (A) stands for Department of Natural Resources. The thermal regime classifications displayed in (C) are from Maheu et al. (2016).

Table 1

Summary of median upstream basin characteristics for the OESF, SHEN, and USGS sites. Statistical differences between the datasets were determined using a one-way ANOVA followed by a Tukey honestly significant difference post-hoc test (Graves et al., 2015) and represented as italicized letters following the mean values.

Dataset	Statistic	Basin Area (ha)	Basin Elevation (m)	Basin Slope (%)
OESF[1]	Mean	210*a*	320*a*	32.3*b*
	Range	15.4–789	76.8–854	4.00–69.0
	Std. Dev.	169	187	19.2
SHEN[2]	Mean	800*a*	781*b*	34.0*b*
	Range	27.4–3630	584–1110	13.7–52.2
	Std. Dev.	723	120	8.60
USGS[3]	Mean	$1.41 \times 10^5 b$	914*b*	16.8*a*
	Range	$432–3.32 \times 10^6$	17.0–3680	0.10–53.1
	Std. Dev.	3.73×10^5	774	14.7

[1] Basin area and elevation values were calculated using a 0.9 m (3 ft) LiDAR-derived digital elevation model (DEM) (WADNR, 2019) and basin slope values were calculated using USGS 10-m DEM.
[2] Values were calculated using USGS 10-m DEM. [3] Values were extracted from the GAGES-II database (U.S. Geological Survey, 2011).

2.1. Site descriptions

The Pacific Northwest Olympic Experimental State Forest (OESF) and Mid-Atlantic Shenandoah National Park (SHEN) sites (Fig. 1 and Table 1) represent two regions of the U.S. with watersheds of differing climate, elevation, vegetation, geology, logging history, and glaciation history. In these disparate regions similarly sized mountainous watersheds (n = 56 sites in OESF and 120 in SHEN) were instrumented with paired air-water temperature sensors for multiple years. These sites provide an opportunity to compare spatial and temporal temperature patterns from opposite coastal areas of the U.S. as well as examine within-region patterns. In addition, stream temperature data were also analyzed from 135 U.S. Geological Survey (USGS) stream gage sites previously classified by Maheu et al. (2016), which are spatially dispersed across the contiguous U.S. (Fig. 1 and Table 1). This regional- and national-scale analysis expands the spatio-temporal analysis of thermal regime across different biomes with unique climates, hydrogeology, and land-use histories.

2.1.1. Regional datasets: OESF and SHEN

The OESF is a 523,000-ha mostly forested planning region that contains 110,000 ha of state trust lands on the Olympic Peninsula in western Washington, U.S. (Martens et al., 2019) (Fig. 1A). SHEN is a 77,700-ha mostly forested protected area located along the spine of the Blue Ridge Mountains in northern Virginia, U.S. (Fig. 1B). Elevation in the OESF ranges from sea level to 1,155 m above sea level (a.s.l.). Mean annual precipitation in this maritime climate ranges between 203 and 355 cm with the majority falling as rain primarily during the fall and winter months (approximately October to March). Elevation in SHEN ranges from 162 to 1,235 m a.s.l. and receives an average annual precipitation amount between 100 and 150 cm (Jastram et al., 2013), with the majority falling as rain at all but the highest elevations. A comparison of cumulative monthly precipitation amounts and mean air temperatures between OESF and SHEN is shown in Appendix A.

OESF stream flows rise with the onset of the rainy season in autumn, peak during winter to early spring (approximately December through March), and decline to reach base-flow conditions in summer (approximately June through August) based on discharge data from USGS gages 12041200 and 12043000 (U.S. Geological Survey, 2019). Average monthly base-flow index (BFI) values for these two gages, calculated using the USGS Groundwater Toolbox (Barlow et al., 2017) HySEP-Fixed method over the same time period as temperature measurements (see below), were 0.70 and 0.68, respectively. Maximums (> 0.87) typically occurred in July and minimums (< 0.48) in October

for both gages. Standard deviation (SD) between annual means (2013–2018 water years, i.e., 1 October to 30 September) were < 0.03 for both gages.

Although SHEN experiences less dramatic seasonal variation in rainfall amounts than OESF, late spring through summer months generally have slightly more precipitation than autumn through winter months. However, because peak precipitation rates occur during the growing season in these primarily deciduous forests and much of the incoming water is utilized by flora during transpiration, stream flows typically peak during winter to spring months (approximately December through May) with a slow rise starting in autumn and a slightly faster decline to reach base-flow conditions during late summer to early autumn months (approximately August through September) based on discharge data from USGS gages 01662800 and 01663500 (U.S. Geological Survey, 2019). Average monthly BFI values for these two gages were 0.72 and 0.59, respectively. Annual BFI patterns were approximately bi-modal, with maximums (> 0.80 and > 0.65, respectively) occurring from December to April and in August and minimums (< 0.62 and < 0.50, respectively) occurring in July and from September to October. SD of annual means (2013–2016 water years) were < 0.04 for both gages. BFI values for all gages near OESF and SHEN generally indicate that GW is an important contribution to overall stream flow in these regions.

A total of 56 flow-disconnected OESF subwatersheds were used in this study that fall within the Coast Range level III ecoregion (see Section 2.1.2 for a more detailed description of these classifications) (Omernik, 1987; Omernik and Griffith, 2014). Paired air and water temperature measurements were collected at these sites between 1 October 2012 and 31 December 2018. These sites have been used in recent publications regarding stream and forest conditions of the OESF (Martens et al., 2019; Minkova et al., in press). The SHEN sites used in this study are within the Blue Ridge level III ecoregion (Omernik, 1987; Omernik and Griffith, 2014) and include 120 subwatersheds within 18 flow-disconnected watersheds that all lie within the boundaries of the park. Hourly water temperature data were collected at all 120 sites, while air temperature was collected at a subset of the sites and modeled using latitude and elevation as predictors (Johnson et al., 2017). For this study, data collected between 23 June 2012 and 30 September 2016 were used, and the details of the temperature data collection can be found in Snyder et al. (2015) and Johnson et al. (2017). Various subsets of these sites and time period have been the foci of previous stream temperature research (Briggs et al., 2018b,a; Johnson et al., 2017; Snyder et al., 2015) and the data are publicly available (Briggs et al., 2017; Snyder et al., 2018).

The OESF encompasses more than 17,000 km of streams (stream order 1–4, excludes the mouth of the Quillayute River), including portions of several major regional rivers (WADNR, 2016). The 56 sites represent tributaries to these regional rivers, whose waters reach the Pacific Ocean via the Clallam River to the north and the Quillayute, Hoh, and Queets Rivers to the west, as well as coastal streams, such as Goodman Creek and Mosquito Creek to the west. Collectively, these small fish-bearing streams contain various populations of several coldwater salmonid species (Martens, 2016). SHEN contains over 1000 km of streams (stream order 1–3) that ultimately drain into Chesapeake Bay through the Rappahannock, Potomac, and James Rivers. SHEN's native brook trout (*Salvelinus fontinalis*) coldwater habitat is widely recognized as being among the best in the middle Appalachians and streams in the park collectively provide habitat for at least 41 fish species (Jastram et al., 2013; National Park Service, 2019a).

Two general types of geologic substrate dominate within the OESF: sedimentary rocks and deposits (58.2% of the area) at higher elevations and unconsolidated sediments including glacial and alluvial deposits (39.0% of area) at lower elevations. Volcanic rocks and deposits also exist within the OESF but represent a minor fraction of the area (2.8%). Three principal bedrock types exist within SHEN and each represents

approximately one-third of the total park area: granitic, metabasaltic, and siliciclastic (Gathright, 1976; Southworth et al., 2009). Limestone is also present but in less than 2% of the park's area. Colluvium dominates higher elevation surficial deposits and debris-fan deposits dominate lower elevations. Currently, there is limited known information regarding GW sources within the OESF (Safeeq et al., 2014b,a, 2013) but much is known in SHEN. Perennial springs are prevalent in SHEN and thus GW inputs are a major driver of stream flow (DeKay, 1972; Lynch, 1987; Snyder et al., 2013), the source of which originates from generally shallow (< 10 m depth) layers of residuum and colluvium that overlie low-permeability bedrock (Busenberg and Plummer, 2014; Lynch, 1987; Plummer et al., 2001). In one SHEN watershed, spatially distributed passive seismic geophysical measurements in the riparian zone indicated an average depth to bedrock of only 2.6 m (Briggs et al., 2018b).

All but four of the 56 OESF subwatersheds used in this study are managed by the Washington State Department of Natural Resources. The remaining four sites serve as reference sites for the area because they are ecologically similar to the OESF state lands sites but are located in the adjacent Olympic National Park and never experienced timber harvesting. Natural ecosystem disturbance is dominated by major wind events and wildfire (WADNR, 2016). Riparian forest conditions in the OESF state lands are primarily in the earlier stages of forest development (< 80 years), with 70% of riparian areas dominated by hardwoods or young conifers (WADNR, 1997). Outside of riparian areas, OESF forests are primarily coniferous (Franklin and Dyrness, 1973; Henderson et al., 1989). SHEN's federally protected forests have not experienced any timber harvesting since the establishment of the park in 1935. Today, SHEN's forests cover approximately 95% of the park's land and natural ecosystem disturbance is dominated by invasive insects and wildfire. The majority of the forest is made up of deciduous species (National Park Service, 2019b), and a recent outbreak of the invasive hemlock woolly adelgid (*Adelges tsugae*) has resulted in a 95% loss of Eastern hemlock (*Tsuga canadensis*), which has significantly decreased riparian shading in the affected areas (National Park Service, 2019b).

2.1.2. National dataset: USGS gage sites

Maheu et al. (2016) selected 135 sites for which daily water temperature data were available from the USGS Geospatial Attributes of Gages for Evaluating Streamflow (GAGES-II) database (U.S. Geological Survey, 2011) (Fig. 1C). These sites span seven level I, 14 level II, and 45 level III ecoregions (Omernik, 1987; Omernik and Griffith, 2014). Ecoregion levels are hierarchical, grouping locations on different spatial scales (I = continental, II = national and subcontinental, III = regional) based on patterns of biotic and abiotic phenomena including geology, vegetation, climate, and land use. Median basin elevations range from 17 to 3680 m a.s.l. and basin areas are generally larger than the OESF and SHEN sites (Table 1). Of the 135 USGS sites, 76 are classified as reference sites (i.e., represent watersheds with minimal hydrological disturbance) (U.S. Geological Survey, 2011) and the remaining 59 were included to expand the spatial coverage despite experiencing some form of hydrological disturbance. The degree of hydrologic disturbance is measured as an index in the GAGES-II database, ranging from 1 (minimally disturbed) to 42 (highly disturbed). Hydrologic disturbance index values were generally only slightly lower in the reference sites (range = 1 to 18, mean = 7.7, SD = 3.4) than the non-reference sites (range = 4 to 25, mean = 12.2, SD = 4.0) used for this study. Therefore, the sites used in this study generally represent less disturbed hydrologic regimes. Site average BFI values extracted from the GAGES-II database (divided by 100) were slightly greater for reference sites (range = 0.07 to 0.85, mean = 0.57, SD = 0.16) than non-reference sites (range = 0.14 to 0.88, mean = 0.47, SD = 0.20).

In this study, a date range of 1 January 2010 to 1 January 2019 was chosen to both maximize the possible amount of temperature data and to generally correspond to the time period of data collection in SHEN and OESF. Of the original 135 sites, 26 were removed due to the absence of mean daily water temperature data within this time frame. This resulted in a loss in coverage of three level III ecoregions but no loss in level I or level II ecoregions. Because a large number of USGS sites exhibited long durations of zero and sub-zero water temperatures, which decouples the air-water relationship (Letcher et al., 2016), any data points ≤ 0 °C were removed before performing the sine-wave regressions.

Of the remaining 109 sites, 67 are reference sites and 42 are non-reference sites. Where mean daily water temperature was not available for a given site, but daily minimum and maximum temperatures were, the mean daily water temperature was calculated as the average of
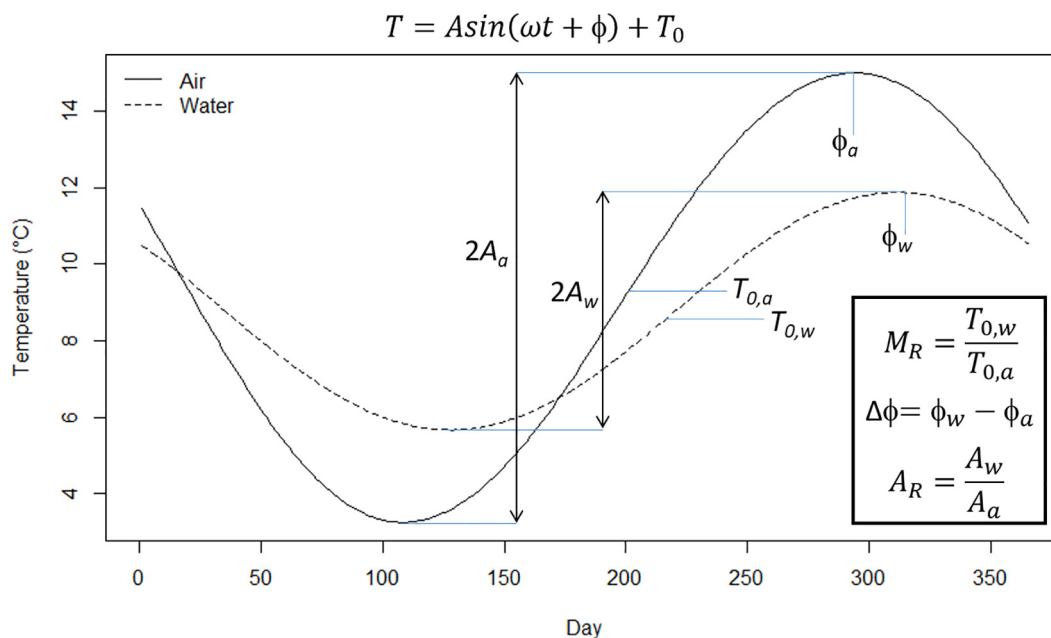
$$T = A\sin(\omega t + \phi) + T_0$$

Fig. 2. Conceptual figure of an example air and water temperature annual sinusoidal curve (equation shown on top). Symbols represent the various air and water annual thermal parameters and the equations represent the combined air-water annual thermal metrics.

these two temperatures. As local air temperature data were only available for two years at two sites, PRISM 4-km grid cell resolution mean daily air temperatures (PRISM Climate Group, 2019) were extracted from the pixel nearest the site coordinates for each of the 109 sites over the same time period as the water temperature data using R (Hart and Bell, 2015). Accuracy of the PRISM mean daily air temperature values were tested by comparing the observed local streamside air temperature values from the two USGS sites and a subset of OESF and SHEN sites with the PRISM values extracted from the nearest pixel. This comparison is summarized at the end of Section 2.2.

2.2. Annual sine-wave regression

Linear regressions of sine-waves were used to analyze observed multi-year air and water temperature data from across the contiguous U.S. First, a watershed- to regional-scale comparison was undertaken to explore intra- and inter-watershed to intra- and inter-regional thermal variability using paired air-water data collected at relatively high spatial resolution from the OESF and SHEN focal regions. Second, a lower-resolution but spatially distributed comparison was performed to explore air-water thermal variability across many regions of the contiguous U.S. using the USGS national dataset.

Defining annual signals is common for stochastic modeling of water temperature. In a stochastic approach, water temperature is separated into two components: a long-term (i.e., seasonal or annual) periodic component and a short-term nonseasonal residual component (Caissie et al., 1998; Kothandaraman, 1971; Stefan and Preud'homme, 1993). The annual component of water temperatures is defined by a Fourier series or sinusoidal function, which is then subtracted from the observed temperature data, leaving the residual component. Box-Jenkins methods or a second-order Markov process are common methods to predict the residual values (Benyahya et al., 2007a; Caissie et al., 1998; Caissie, 2006; Cluis, 1972). These modeled residuals are then added back to the modeled long-term component to predict water temperature. Many studies exist that examine the residuals component (Ahmadi-Nedushan et al., 2007; Benyahya et al., 2007b; Caissie et al., 2001, 1998; Cluis, 1972; Hague and Patterson, 2014; Salas, 1992), but the annual component is largely ignored in the context of process-based analysis. The characteristics of annual stream temperature signals can be used to deduce which hydrogeological processes are driving stream thermal regimes (Johnson, 1971; Maheu et al., 2016), particularly when compared to local air temperature dynamics (Briggs et al., 2018b,a). In this study, the annual signal of mean daily air and water temperatures for each site were approximated as sine waves using only the 1st harmonic (Kothandaraman, 1971):

$$T(t) = A * \sin(\omega t + \phi) + T_0 + \varepsilon(t) \tag{1}$$

where T is daily mean temperature (°C), A is the amplitude (°C), ω is the angular frequency (rad d^{-1}), t is time (d), ϕ is the phase (rad), T_0 is the mean temperature (°C), and ε is the error term (°C). The amplitude, phase, and mean represent the core parameters of a sine-wave signal—the variation, horizontal displacement, and vertical displacement, respectively (Fig. 2). The annual time period was defined based on the water year from $t = 1$ (1 October) to $t = 365$ (30 September, $t = 366$ in leap years), resulting in annual vectors of ω and t each containing daily values. Although the use of higher harmonics improves the regression fit by accounting for sub-annual variation, this study focuses on the fundamental annual signals and thus the first harmonic is appropriate (Kothandaraman, 1971). The use of mean daily temperatures eliminates the need to account for lags between air and water temperatures on the scale of hours (Stefan and Preud'homme, 1993; Webb et al., 2003).

The sine-wave equation (Eq. (1)) was linearized using Ptolemy's theorem to enable linear regression (Kothandaraman, 1971):

$$T(t) = a * \sin(\omega t) + b * \cos(\omega t) + T_0 + \varepsilon(t) \tag{2}$$

where a and b are the regression coefficients equal to $A * \cos(\phi)$ and

$A * \sin(\phi)$, respectively. The annual air and water temperature signal amplitudes are then determined as $\sqrt{a^2 + b^2}$ and phase for each signal is calculated as $\tan^{-1}(b/a)$. Amplitude is approximately half of the annual range in mean daily temperatures (Ward, 1985), when disregarding anomalously warm or cold days. Phase was converted into units of days by multiplying by the length of the year divided by 2π. To control for the influence of local weather patterns on annual water temperature signals, annual air and water temperature parameters were combined into three annual metrics based on the core features of a sinusoid curve (Fig. 2). Amplitude ratio (A_R), also known as magnitude ratio (Kothandaraman, 1972) or damping factor (Kurylyk et al., 2015), is the water amplitude (A_w) divided by the air amplitude (A_a) (Briggs et al., 2018b,a). Phase lag ($\Delta\phi$) was calculated as the difference between water temperature phase (ϕ_w) and air temperature phase (ϕ_a) (Briggs et al., 2018b,a; Kothandaraman, 1972). Mean ratio (M_R), also known as ratio of means (Kothandaraman, 1972), is the water mean temperature ($T_{0,w}$) divided by the air mean temperature ($T_{0,a}$). Comparing multiple metrics simultaneously can give insights into potential GW contributions and effective depth of that contribution. Shallow GW (within approximately 8 m of land surface) has an annual temperature signal that is both lagged and damped compared to local air temperature, while deeper GW typically shows little annual temperature variation (Constantz, 2008). Therefore, shallow and deep GW discharges introduce unique annual temperature signals into stream water. An extended $\Delta\phi$ (> ~10 d) corresponding with a low A_R (i.e., negative relation) has been shown to be the result of shallow GW influence, whereas deeper GW influence would induce similar or even lower A_R with minimal $\Delta\phi$ (Briggs et al., 2018a).

Using a moving window in the linear regression (Eq. (2)), where the best-fit sinusoidal parameters and combined metrics can change at each step through time, can indicate the temporal stability of the sinusoid fits (Zivot and Wang, 2006). Many iterations of the linear regression were calculated for each site by shifting the 365-day window by 1 day for each iteration. This resulted in a possible maximum of 1919, 1463, and 2924 separate regressions for each OESF, SHEN, and USGS site, respectively, and the three annual thermal parameters and combined metrics were calculated for each regression iteration. However, in most cases, the median parameter and median combined metric values from these regressions for each site are reported, hereafter "site median", in order to compare values between sites.

Regression fit was assessed using root mean squared error (RMSE). Temporal variation of RMSE and each annual thermal parameter and combined metric was assessed by calculating the standard deviation of values from the many regression iterations for each site. Up to 49 days of missing data within each 365-day window were allowed for each site, as we found through simulation (data not shown) would still yield accurate estimates of the three annual thermal parameters. One OESF site, three SHEN sites, and 11 USGS sites were removed due to no 365-day periods with ≤49-days of missing data. Five of the 11 USGS sites removed in this way were a result of removing temperatures ≤0 °C mentioned earlier. This left a total of 55, 117, and 98 sites for the OESF, SHEN, and USGS datasets respectively. All regressions and statistical calculations were carried out using R (R Core Team, 2017). Spearman's rank correlation coefficient (ρ) between various parameters and combined metrics was calculated using the *cor.test* function. One-way ANOVA followed by a Tukey honestly significant difference post-hoc test was used to evaluate differences among datasets using the *multcompView* package in R (Graves et al., 2015). The relationship of annual thermal parameters and combined metrics with selected watershed characteristics was summarized for each site. These characteristics included median upstream basin elevation, median upstream basin slope, and accumulated upstream basin area (Table 1).

For a subset of sites, PRISM air temperature records were compared to locally measured streamside temperatures. Overall, the difference between the PRISM air temperature values and observed values was

minor. Therefore, while error surrounding the annual thermal parameters estimated with PRISM data is slightly higher than with locally measured data, the overall patterns and relationships among the annual thermal parameters and combined metrics are still valid. On average, PRISM slightly overestimated the air temperature annual mean, phase, and amplitude when compared to sites with locally measured air temperature data. Therefore, these differences would generally result in a modest underestimation of annual mean ratio, phase lag, and amplitude ratio. However, the differences between PRISM predictions and observed air temperatures at a specific site may lead to small overestimations of the annual thermal parameters as well. Comparing PRISM air temperature predictions with air temperature from the only two USGS sites in the dataset with observed data, on average PRISM overestimated the air temperature annual mean (by 0.07 ˚C) and phase (by 3.9 d) and underestimated the annual amplitude (by 0.57 ˚C) in one site but underestimated the annual mean (by 0.78 ˚C) and overestimated the annual phase (by 1.4 d) and amplitude (by 0.37 ˚C) in the other site. For the comparison with a subset of OESF and SHEN sites, PRISM tended to overestimate the annual mean by an average of 0.58 ˚C (difference range of −0.75 to 2.12 ˚C), overestimate annual phase by an average of 1.8 d (difference range of −3.7 to 4.1 d), and overestimate annual amplitude by an average of 0.68 ˚C (difference range of −0.2 to 1.35 ˚C).

3. Results

In summary, OESF and SHEN thermal regimes were dramatically different when annual air and water temperature signal parameters were compared separately. However, when air and water temperature signals were combined (i.e., M_R, $\Delta\phi$, and A_R), OESF and SHEN sites generally exhibited similar patterns of shallow GW discharge dominance, as indicated by a negative $\Delta\phi$-A_R relation. The generally larger USGS sites exhibited greater spatial variation in the annual air and water temperature parameters among its sites, which led to high variability in the combined air-water annual thermal metric values. Some USGS sites exhibited potentially deeper GW influence than the OESF and SHEN sites (low A_R with lower $\Delta\phi$). The Maheu et al. (2016) thermal regime classifications generally clustered together when plotted in M_R-$\Delta\phi$-A_R metric space despite only considering annual water temperature patterns during their development. However, some regimes exhibited a large range in metric values, which led to some overlap between regimes. Basin area, elevation, and slope were variably, but significantly ($p < 0.05$), correlated with the three annual thermal metrics among the three datasets.

3.1. Annual sine-wave regressions

Air temperature regression fits were much better (i.e., lower RMSE) for OESF sites than SHEN and USGS sites (Table 2 and Appendix Figure B.1), which is likely related to greater sub-annual temperature variation (high and low) at the SHEN and USGS sites that are not well captured by the fundamental annual sinusoid. The USGS sites exhibited substantially higher variability in air temperature RMSE among its sites (SD = 0.57 ˚C) than that of either the OESF (SD = 0.12 ˚C) or SHEN (SD = 0.02 ˚C) sites. Water temperature regressions showed stronger fits and were more similar between the three datasets than for air temperature regressions (Fig. 3), though the USGS sites again showed higher variability among sites.

3.2. Annual parameter and metric patterns

Despite large differences in annual precipitation, air temperature, and water temperature regimes between the two regional datasets, the observed patterns in combined air-water annual thermal metrics were similar for the SHEN and OESF regions. Overall, OESF sites exhibited lower air and water annual means ($T_{0,a}$ and $T_{0,w}$ respectively), lower

amplitudes (A_a and A_w, respectively), and higher phases (i.e., later peak temperatures; ϕ_a and ϕ_w, respectively) than SHEN sites (Fig. 4). However, the majority of site-specific combined air-water annual metric values overlap between the two regions (Fig. 4, panels G-I). In general, there were slightly higher M_R and $\Delta\phi$ values, and slightly lower A_R values for OESF sites than SHEN sites (Table 2 and Fig. 4). Interestingly, the SHEN M_R values are mostly < 1.0 (i.e., water temperatures generally cooler than air temperatures) despite all of the sites experiencing sub-zero air temperatures on at least 10% of the days. The USGS sites exhibited a larger range in air and water temperature annual mean, phase, and amplitude values than OESF and SHEN sites (Fig. 4, panels A-F). This large parameter range also resulted in greater variability in the M_R, $\Delta\phi$, and A_R metric values with generally lower $\Delta\phi$ and higher A_R than the OESF and SHEN sites (Table 2 and Fig. 4, panels G-I). The three extreme USGS values of M_R (Fig. 4G) come from sites on the geothermally influenced Firehole and Gibbon Rivers in Yellowstone National Park, Wyoming.

The days of the year corresponding to the fitted annual maximum air and water temperatures were typically later for OESF sites than SHEN sites (Fig. 4E and Appendix Table B.1), but water temperatures generally reached their maximum less than two weeks after the maximum air temperatures for both regions (Fig. 4B and Appendix Table B.1). The fraction of days with water temperatures cooler than air temperatures was typically higher for SHEN sites than OESF sites (Table B.1). The USGS sites exhibited a similar range in the timing of the annual air maxima, but the annual water maxima were generally earlier than the OESF and SHEN sites. The range in the fraction of days with water temperatures cooler than air temperatures was much larger for the USGS sites, but the mean of the fraction was between that of the OESF and SHEN sites. A comparison of the three annual metric values between OESF sites on state (n = 51) and federal (i.e., reference sites) (n = 4) lands is shown in Appendix C.

Variation in the three annual thermal metrics between regressions for a given site (i.e., temporal variation in the regressions) was present and similar for both regional datasets but of minor magnitude, especially for M_R (Fig. 5 and Appendix Table B.2). This temporal variation is less than the spatial variation in median metric values among OESF and SHEN sites (i.e., Appendix Table B.2 versus Table 2) for all three annual thermal metrics. In general, the time series for a given annual thermal metric within the OESF or SHEN datasets exhibited similar shapes among sites but were vertically displaced from each other. Temporal patterns were generally not consistent between the three annual metrics within these regions, suggesting that temporally variable factors

Table 2

Summary statistics for the site median air and water temperature regression fits and combined air-water annual thermal metrics for OESF, SHEN, and USGS sites. Statistical differences between the datasets were determined using a one-way ANOVA followed by a Tukey honestly significant difference post-hoc test (Graves et al., 2015) and represented as italicized letters following the mean values.

| Metric	Dataset	Mean	Range	Std. Dev.
Air RMSE (˚C)	OESF	2.38*a*	2.07–2.66	0.12
	SHEN	4.09*b*	4.07–4.14	0.02
	USGS	4.07*b*	2.75–5.29	0.57
Water RMSE (˚C)	OESF	1.12*a*	0.60–1.33	0.15
	SHEN	1.55*b*	0.56–2.12	0.25
	USGS	1.84*c*	0.63–3.32	0.50
M_R (–)	OESF	1.03*ab*	0.90–1.10	0.04
	SHEN	0.98*a*	0.89–1.09	0.03
	USGS	1.12*b*	0.69–5.29	0.61
$\Delta\phi$ (d)	OESF	12.7*c*	6.8–39.8	4.4
	SHEN	10.4*b*	6.3–30.6	3.4
	USGS	3.3*a*	−9.3–31.8	6.6
A_R (–)	OESF	0.612*a*	0.293–0.744	0.097
	SHEN	0.683*b*	0.445–0.805	0.069
	USGS	0.758*c*	0.312–1.123	0.168
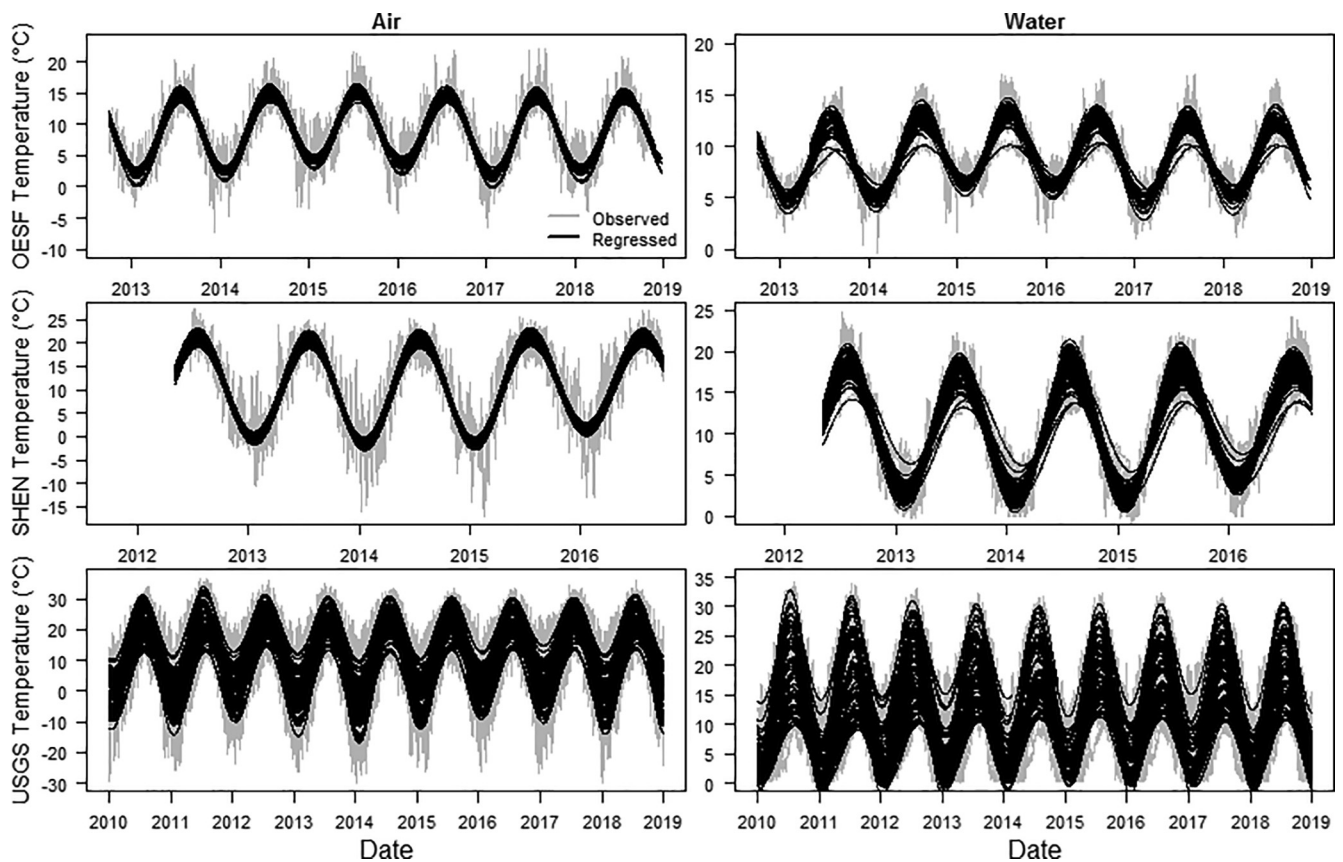
Fig. 3. Observed (gray) and regressed (black) air (left column) and water (right column) temperature time series for all OESF (top row), SHEN (middle row), and USGS (bottom row) sites. Each line (gray or black) represents a single site. Each black line represents a site's mean (n ≤ 365) of the moving window regression fits for each date. Note that the x-and y-axes scales are different for each dataset.

affecting stream temperature do so differently for each annual thermal metric. M_R (1.9% and 1.7%) showed the least temporal variation and $\Delta\phi$ (12.4% and 9.2%) showed the greatest for both OESF and SHEN respectively, when comparing the site temporal SD of the three annual thermal metrics to the site median value of each metric (Appendix Table B.2). The USGS sites also exhibited lower temporal variation than between site spatial variation. However, temporal signals in the three annual thermal metrics were more variable between sites in the USGS national dataset because of the spatial extent of this dataset in comparison to OESF and SHEN regional datasets. Similar to the OESF and SHEN sites, though, M_R (4.6%) generally showed the least temporal variation and $\Delta\phi$ (206%) the most temporal variation when compared to the site median value for the USGS sites. These same trends were also observed in the annual air and water temperature parameters for each dataset (Appendix Table B.2), with slightly larger variation magnitude than the annual thermal metrics when compared to the site median value for each. The site time series for air temperature parameters mirror those of water temperature in each dataset, with vertical displacement among sites (Appendix Fig. B.2).

A strong negative correlation was observed between $\Delta\phi$ and A_R for both OESF ($\rho = -0.671$, $p < 0.01$) and SHEN ($\rho = -0.588$, $p < 0.01$) sites (Table 3 and Fig. 6A), implying variable-strength shallow GW influence in both datasets. There is also a moderate correlation between M_R and A_R for OESF ($\rho = 0.341$, $p < 0.01$) sites and a significant but weak correlation between M_R and $\Delta\phi$ for SHEN ($\rho = 0.181$, $p < 0.10$) sites. Both regions have a few sites exhibiting signs of strong shallow GW influence (low A_R with high $\Delta\phi$). For those SHEN sites, M_R varies above and below 1.0, similar to SHEN sites having seemingly weaker GW influence (higher A_R with lower $\Delta\phi$). However, for those OESF sites, M_R values are consistently approximately 0.9 and represent the lowest M_R values in the region. Additionally, several OESF sites show signs of

somewhat deeper GW influence than the SHEN sites. These can be seen in Fig. 6 (panels A or C) as the OESF sites with A_R values < ~0.51 and with $\Delta\phi$ values < ~17 d. One of these is a reference site, a tributary to the Hoh River.

The USGS sites show significant ($p < 0.01$) correlation for every pairing of the three annual thermal metrics (Table 3 and Fig. 6B), though the negative correlation between $\Delta\phi$ and A_R ($\rho = -0.354$) is weaker than in the OESF and SHEN datasets. A few USGS sites also exhibited signs of strong shallow GW influence while several other sites showed signs of strong deep GW influence (low $\Delta\phi$ and low A_R) that were not observed in OESF or SHEN sites. The majority of sites exhibiting relatively strong shallow to deep GW influence were reference sites (i.e., minimal hydrological disturbance). Negative phase lag values were observed in several reference and non-reference sites (Fig. 6B), which is unusual as minimum phase lags are typically expected to approach zero. For reference sites, the negative phase lag values may have been due to error associated with using modeled rather than locally measured air temperatures. For non-reference sites, negative phase lags could also have been due to the degree of hydrologic disturbance at the site in addition to using modeled air temperatures. Reference sites generally exhibited lower M_R (mean of 1.10 versus 1.16) and A_R (mean of 0.713 versus 0.825) and higher $\Delta\phi$ (mean of 4.4 d versus 1.8 d) than the non-reference sites (Fig. 6B). The hydrologic disturbance index was positively correlated with M_R ($\rho = 0.393$) and A_R ($\rho = 0.479$) and negatively correlated with $\Delta\phi$ ($\rho = -0.320$).

The Maheu et al. (2016) thermal regime classifications tended to group together in $\Delta\phi$-A_R space (Fig. 6C), though with some spread in values especially in the "stable cold", "stable cool", and "variable cold" regimes, which led to overlap between some regime groupings. In general, the "variable warm" and "variable cool" regimes exhibited moderate M_R (means of 1.04 and 1.13) with low $\Delta\phi$ (means of 0.0 d and
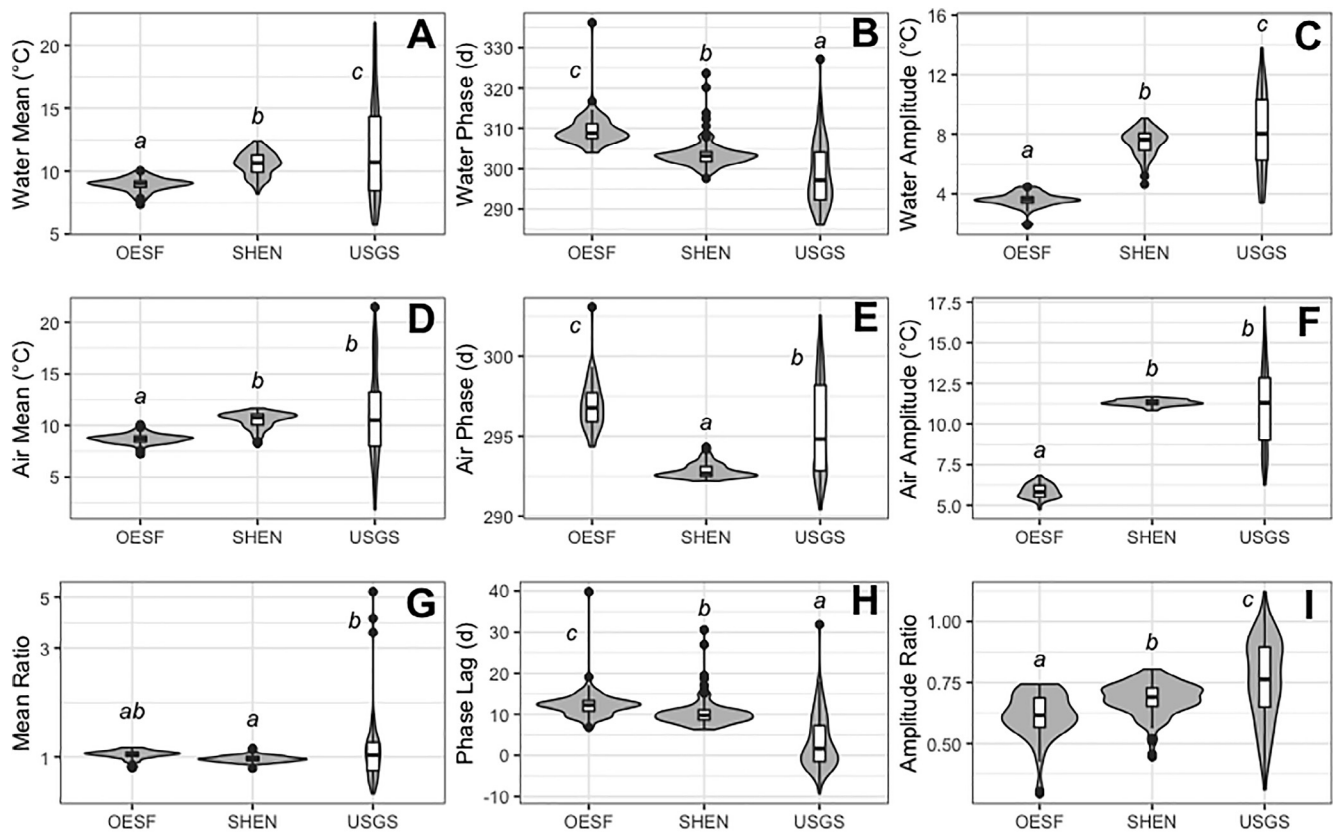
Fig. 4. Violin plots (i.e., combination of boxplot and density plot) of site median annual temperature mean (A,D,G), phase (B,E,H), and amplitude (C,F,I) parameters for OESF, SHEN, and USGS sites. Water temperature parameters are shown on top (A,B,C), air temperature parameters in the middle (D,E,F), and combined metrics (i.e., M_R, $\Delta\phi$, and A_R) on the bottom (G,H,I respectively). Water (B) and air (E) phase values are given in day of the water year (1 October to 30 September). Statistical differences between the datasets were determined using a one-way ANOVA followed by a Tukey honestly significant difference post-hoc test (Graves et al., 2015) and represented as italicized lowercase letters.

−1.2 d) and high A_R (means of 0.920 and 0.919), while the "stable cold" regime exhibited low to moderate M_R (mean = 0.96) with moderate to high $\Delta\phi$ (mean = 9.2 d) and low to high A_R (mean = 0.654). The "highly variable cool" regime generally exhibited high M_R (mean = 1.36) with low $\Delta\phi$ (mean = −0.1 d) and high A_R (mean = 0.888). The "variable cold" regime generally exhibited moderate M_R (mean = 1.14) with low $\Delta\phi$ (mean = 0.6 d) and moderate A_R (mean = 0.724). The "stable cool" regime was the most widespread, exhibiting anywhere from low to extremely high M_R (mean = 1.35) with low to moderate $\Delta\phi$ (mean = 2.7 d) and low to high A_R (mean = 0.704).

Derived annual thermal metrics showed variable associations with physical characteristics of contributing watersheds. Basin area was most strongly correlated with the three annual thermal metrics in the OESF and SHEN datasets in comparison with basin elevation and slope (Fig. 7 and Appendix Table C.2). Basin area, elevation, and slope were all significantly correlated with the three annual thermal metrics in the USGS dataset, with basin slope having the strongest correlations. Basin area differences between OESF and SHEN sites was related to some of the separation between the two regions noted in Fig. 4 for the three annual thermal metrics. The larger SHEN basins skewed the overall M_R and $\Delta\phi$ values lower and the overall A_R values higher. When excluding SHEN sites with basin area > 850 ha, the distributions of M_R (mean = 0.99), $\Delta\phi$ (mean = 11.0 d), and A_R (mean = 0.661) overlap with those of OESF to a greater degree. Basin elevation had a greater effect on OESF sites than SHEN sites, where it was negatively correlated with A_R in both regions but negatively correlated with M_R and positively correlated with $\Delta\phi$ only in OESF sites (Appendix Table C.2). Basin slope was significantly correlated with $\Delta\phi$ (positively) and A_R

(negatively) in OESF and SHEN sites. When aggregating all three datasets, basin area exhibited the strongest correlation with $\Delta\phi$ (negatively) and A_R (positively) among the basin characteristics but was not correlated with M_R (Appendix Table C.1). However, basin elevation was significantly correlated with M_R (negatively) and A_R (negatively), and basin slope was significantly correlated with M_R (negatively), $\Delta\phi$ (positively), and A_R (negatively). The strength of correlations among the three basin characteristics is summarized in Appendix C (Appendix Table C.2).

4. Discussion

In this study, paired air and water annual temperature signals were analyzed and summarized at watershed, regional, and national spatial scales. These annual patterns are also discussed in the context of upstream physical watershed characteristics. Our results demonstrate the utility of paired air-water annual temperature data for inferring hydrogeological processes in streams. Specifically, we show that (1) sine-wave regression yields useful statistical parameters for interpreting GW influence on stream temperature; (2) regions within North America can be highly spatially patchy in this regard, indicating the importance of localized influences such as GW discharge at the reach to watershed scale; and (3) while results from a national dataset exhibited wider parameter ranges, patterns were largely consistent with those of the regional datasets, suggesting the general utility of our approach. These findings suggest that heat-as-a-tracer approaches, such as those used here, may provide a foundation for assessing the relative importance of local processes on stream thermal regimes and in predicting climate change effects at ecologically relevant spatial scales.
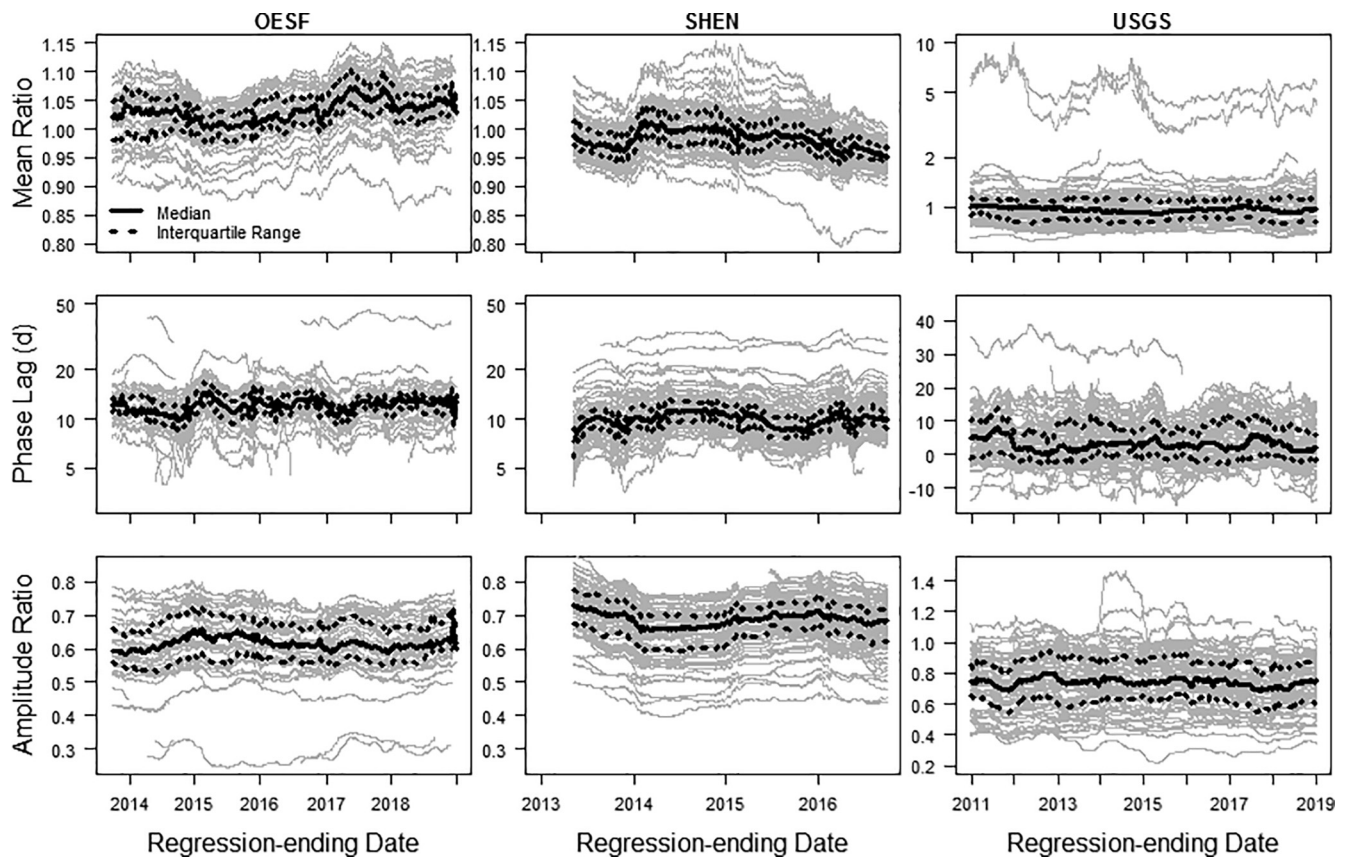
Fig. 5. Time series of the annual mean ratio (M_R, top row), phase lag ($\Delta\phi$, middle row), and amplitude ratio (A_R, bottom row) values for OESF (left column), SHEN (middle column), and USGS (right column) sites. Each gray line represents a time series for a single site. The median (solid black line) and interquartile range (dashed black line) values are calculated from all sites for each regression window. The time series for each annual metric is produced by using a 365-day moving window linear regression (see Methods) and values are reported on the last day of that period.

Table 3

Spearman correlation coefficient (ρ) values between the site median annual thermal metrics from the OESF, SHEN, and USGS sites. Statistical significance (p) level is shown with asterisks (*) after the ρ values. *** represents $p < 0.01$, ** represents $0.01 < p < 0.05$, and * represents $0.05 < p < 0.10$.

Dataset	Metric	M_R	$\Delta\phi$	A_R
OESF	M_R	1		
	$\Delta\phi$	−0.211	1	
	A_R	0.341***	−0.671***	1
SHEN	M_R	1		
	$\Delta\phi$	0.181*	1	
	A_R	−0.032	−0.588***	1
USGS	M_R	1		
	$\Delta\phi$	−0.425***	1	
	A_R	0.292***	−0.354***	1
OESF-SHEN-USGS Aggregate	M_R	1		
	$\Delta\phi$	−0.159***	1	
	A_R	0.178***	−0.597***	1

The high spatial resolution datasets (i.e., OESF and SHEN) allowed for intra- and inter-regional comparisons. Combined annual air-water temperature patterns were found to differ more within these two regions on opposite coasts of the U.S. than between them despite moderate to large differences in both annual air and water temperature signals between regions. This suggests that differences at watershed and sub-regional spatial scales may be more important for drivers of stream temperature variation than differences at the regional scale. Evidence of strong shallow GW influence in SHEN sites was reported in a previous study (Briggs et al., 2018a), with extended $\Delta\phi$ up to approximately one month associated with low A_R. Interestingly, the OESF sites generally

exhibited similar $\Delta\phi$-A_R relationships, suggesting that these sites are also influenced by strong shallow GW. This finding is consistent with the conceptual hydrogeological model developed by Nelms and Moberg (2010) for the Blue Ridge Physiographic Province that indicates a dominance of shallow steep hillslope GW flowing over low permeability bedrock and sourcing stream water, compared to deeper potential GW flow paths. Additionally, the OESF sites generally showed greater influence of somewhat deeper GW than SHEN sites as indicated by smaller $\Delta\phi$ associated with similarly low A_R.

The (mostly) larger basins assessed in the national-scale USGS dataset showed greater ranges in the three annual thermal metrics but there was considerable overlap with metrics observed for regional-scale sites. However, the large-scale spatial variation in the USGS dataset likely contains only a subset of the local variation in any given region due to the low spatial resolution of site locations. Therefore, watershed- and regional-scale variation in stream temperature patterns will be important considerations for coldwater fish habitat prediction, which is dependent on network-level variability in stream temperature. Additionally, while a few of the USGS sites also exhibited strong shallow GW influence, several other sites showed signs of strong deep GW influence (low $\Delta\phi$ with low A_R), which is expected to provide more resilient thermal refuges for coldwater-adapted aquatic species (Briggs et al., 2018b,a; Kurylyk et al., 2015, 2014; Menberg et al., 2014). This deep GW signal was missing in the two regional datasets, whose hydrogeology may limit deep GW influence (Busenberg and Plummer, 2014; Lynch, 1987; Plummer et al., 2001). Characteristics of the contributing basin such as area, elevation, and slope generally corresponded more strongly with the annual air-water temperature metric patterns when comparing sites in the USGS dataset, which suggests that these characteristics are generally correlated with upstream processes
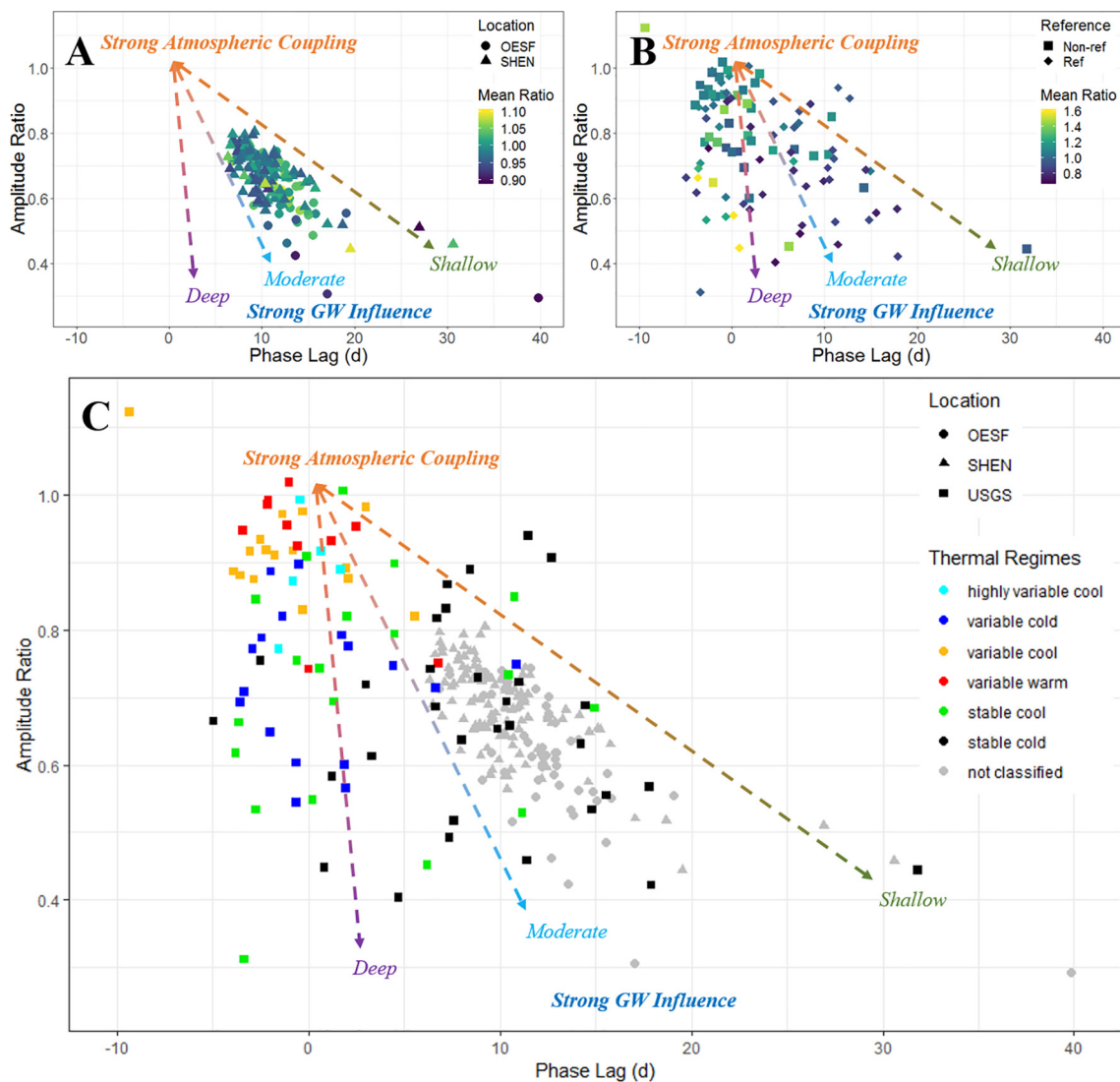
Fig. 6. Site median phase lag versus amplitude ratio colored by (A) mean ratio for OESF (circles) and SHEN (triangles) sites, (B) mean ratio for USGS non-reference (squares) and reference (diamonds) sites, and (C) Maheu et al. (2016) thermal regime classifications for OESF (circles), SHEN (triangles), and USGS (squares) sites. Colored text annotations and dashed arrows (disregards M_R values) show approximate bounding regions of the plots in the continuum between expected atmospheric coupling and GW dominance, assuming no hydrological alteration. These should not be interpreted as hard boundaries for classification purposes. M_R values in (B) were truncated above 1.5.

such as varied-depth GW discharge at a national scale. However, when comparing sites at a higher spatial resolution (i.e., sub-regional), these correlations are more variable and generally weaker, which further suggests that watershed-scale processes are important for determining habitat-level conditions for aquatic biota. The approaches outlined in this study, utilizing metrics calculated from annual temperature sine-wave fits to paired air-water temperature data, show promise for improving classifications of stream thermal regimes on a continental scale and our understanding of current and potential future controls on stream temperature patterns.

Linear regression of a sinusoidal curve represents a relatively simple but powerful method to summarize and compare annual air and stream temperature signals. Prediction of daily stream temperatures was not the goal of this study, but the performance of these simple sinusoidal regressions, with site median RMSE values between 0.56 °C and 2.12 °C and a mean RMSE among sites of 1.41 °C for the regional datasets, was comparable to other methods of stream temperature prediction, some using much more complex methods (Benyahya et al., 2008; Caissie et al., 2001; Laanaya et al., 2017; Zhu et al., 2018). The national-scale USGS dataset experienced a wide range of site median RMSE values

(0.63 to 3.32 °C) and a high mean RMSE (1.85 °C) because of the large spatial coverage of the sites. Locally observed air temperature is commonly lacking in many stream temperature datasets (e.g., USGS dataset). However, predicted local air temperature data, such as those in the PRISM dataset, can be used to accurately calculate annual air temperature parameters. This could greatly expand the spatial coverage of paired air-water temperature datasets and provide a useful management tool for predicting suitable thermal habitat.

4.1. Interpretation of annual parameter and metric patterns

Results reported here indicate that paired air-water annual signals are a promising tool for efficiently diagnosing the influence of local drivers of stream thermal regimes, particularly the pervasive influence of GW in small streams. Attenuation of the annual water signal when compared to the annual air signal (i.e., lower values of A_R) can result from multiple physical watershed factors such as GW discharge, heavy shading, increased channel water volume, presence of a nearby upstream reservoir/lake, or snowmelt input. However, more research is needed to identify the unique signature of each factor in order to
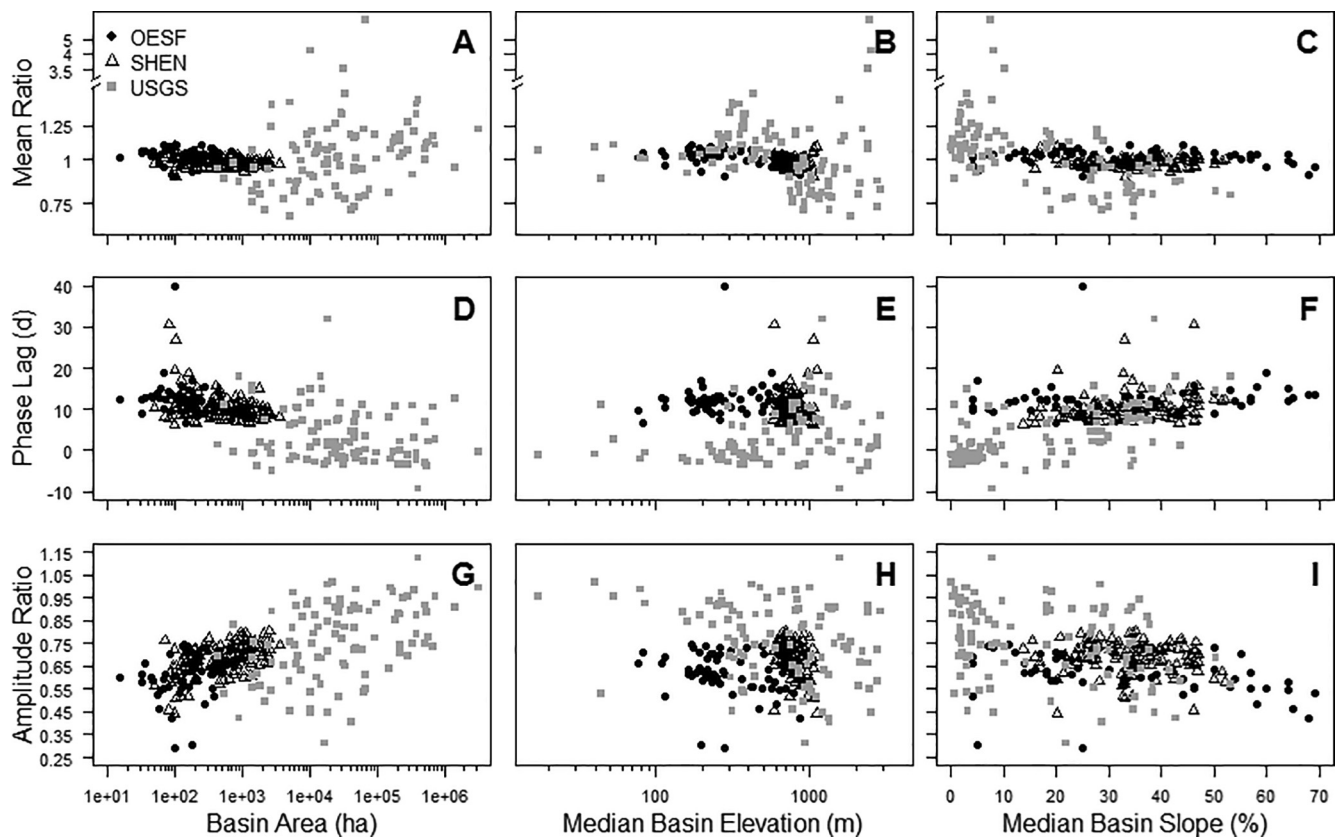
Fig. 7. Scatter plots of site median mean ratio (A,B,C), phase lag (D,E,F), and amplitude ratio (G,H,I) versus basin area (A,D,G), elevation (B,E,H), and slope (C,F,I) for OESF (black circles) SHEN (white triangles), and USGS (gray squares) sites.

separate their individual influence. For example, GW discharge may influence $\Delta\phi$ and A_R, but not M_R, to a greater degree than riparian shading. Predominant shifts in water phase (i.e., $\Delta\phi > \sim 10$ d) have been shown to be mainly the result of shallow GW discharge or the influence of upstream lakes or impoundments whereas factors such as streambed conduction and flow volume do not influence $\Delta\phi$ appreciably (Briggs et al., 2018b,a; Ward, 1963, 1985). However, the examination of $\Delta\phi$ and A_R in combination provides a diagnostic indicator of the relative strength and depth of GW influence, where strong deeper sources are associated with low $\Delta\phi$ and low A_R while strong shallower sources are associated with high $\Delta\phi$ and low to moderate A_R. Shallow GW was shown to primarily influence the heavily shaded SHEN streams in previous studies (Briggs et al., 2018a,b), which should be less resilient to climate and land use changes than deeper GW sources because shallower GW sources are more sensitive to air temperature and GW thermal properties more closely reflect surface conditions. Thus, these measures are useful for characterizing thermal resilience of streams.

Multiple physical watershed factors can also affect patterns of M_R, however, ascertaining GW influence from M_R patterns in tandem with $\Delta\phi$ or A_R is less straightforward than with $\Delta\phi$-A_R patterns. The effect of GW input on M_R may be minor if the common, but not necessarily accurate, assumption holds that mean annual GW temperature is within ~ 1–2 °C of the mean annual air temperature (Anderson, 2005; Benz et al., 2017; Ward, 1985). Therefore, other watershed factors may be better construed from M_R patterns. For example, in some areas, water temperatures are bounded at low and high extremes, uncoupling their behavior from that of air temperature (Letcher et al., 2016; Mohseni and Stefan, 1999). Therefore, colder regions may be expected to exhibit M_R values > 1.0, whereas warmer regions may be expected to exhibit M_R values < 1.0. The seasonality of flow volume and thermal inertia may affect M_R values because more energy is required to change the temperature of a larger flow volume. Snow- or glacial-melt may also

induce M_R values < 1.0. Geothermal influences likely result in anomalously high M_R values, which was observed in three of the USGS sites. Additional research is needed to understand how other factors such as shading impact M_R.

Hydrological alteration, such as dams or diversions, may complicate efforts to quantify GW influence. For instance, greater hydrological alteration in the non-reference USGS sites generally resulted in higher M_R and A_R values and lower $\Delta\phi$ values, which could be interpreted as weaker GW influence. However, GW inputs could be important at sites experiencing hydrological disturbance as well. Another complicating factor is the influence of freezing temperatures on the annual water signal. Sites that freeze for a portion of each year exhibit a sinusoid curve that is cut off at the lower temperatures, which affects the estimate of annual amplitude and mean. Thus, the more days spent at or below 0 °C results in more of the lower end of the sinusoid curve that is cut off, which produces further changes to the annual amplitude and mean and less confidence in inferences regarding GW influence. In these cases, it may be more appropriate to calculate annual median temperature rather than annual means.

Few studies have investigated the applicability of comparing sinusoidal temperature signals of different periods (i.e., ≤365 days) (Caissie et al., 1998) or sinusoidal temperature signals with varying amounts of missing data. Therefore, more work is needed to provide guidance for summarizing the annual signals at such sites. More research on an international scale is also needed to classify where different types of streams currently fit in the annual air-water temperature metric space (Fig. 6C) and to identify which watershed processes primarily influence these patterns. Identifying the main drivers of these patterns is important for our understanding of current thermal regimes and spatial variation, but it may also aid in predicting how these thermal regimes will shift given future changes to air temperature and precipitation patterns. This knowledge is crucial for the management of

coldwater-adapted species. Therefore, future work should also focus on developing new thermal regime classifications that account for local air temperature patterns to further our understanding of current and potential future thermal regimes.

4.2. Regional and national patterns

When air temperature patterns are taken into account, variation in OESF and SHEN water temperature patterns was greater within these regions than between them and potentially driven by similar GW-surface water exchange dynamics. This suggests that local processes such as GW-surface water interactions, riparian shading, and flow volumes are more important than regional differences in climatic and landscape variables such as air temperature, precipitation, and forest type (deciduous versus coniferous) in driving thermal characteristics of headwater streams. Moreover, the large range in A_R values exhibited among sites in both regions despite consistently heavy riparian shading, which has been shown to be an important factor for stream heat budgets (Dugdale et al., 2018; Groom et al., 2017; Johnson and Jones, 2000), solidifies attribution of these thermal patterns to variable-strength GW damping of seasonal water temperatures.

On average, OESF sites had slightly lower A_R and slightly longer $\Delta\phi$ than SHEN sites, suggesting greater GW influence overall. Both regions also included some sites that exhibited strong shallow GW influence (low to moderate A_R with moderate to high $\Delta\phi$). But OESF had some sites with patterns indicative of deeper GW (low to moderate A_R with low to moderate $\Delta\phi$). These deeper GW-influenced sites could represent areas that are more resistant to changes in air temperature over the longer term, potentially providing durable climate refugia for coldwater-adapted fish (Briggs et al., 2018a; Kurylyk et al., 2015; Menberg et al., 2014; Taylor and Stefan, 2009). Additionally, deeper GW flow paths will presumably provide more stable base-flow conditions during droughts than shallow GW flow paths.

The effect of advective heat exchanges associated with GW inputs and channel flow volumes on M_R appear to be more consistent in OESF than SHEN. Since OESF sites do not experience many sub-zero air temperature days, high GW influence acts to dampen summer temperature highs more than winter temperature lows leading to lower M_R values. Further, because of relatively higher winter flow volumes, OESF sites without significant GW influence will dampen winter temperature lows more than summer base-flow temperature highs, leading to higher M_R values. In contrast, SHEN correlations between M_R and GW influence were weaker and not directly related to factors such as basin area, elevation, apparent GW depth, or relative GW discharge volume. These findings suggest that other local factors such as aspect or differences in the temporal variability of flow volume may be responsible.

Analysis of thermal metrics derived from the coarse-scale temperature data collected from USGS gages across the country indicated that larger stream sites generally clustered into groups based upon annual air-water thermal metrics that were remarkably similar to thermal regime classes developed by Maheu et al. (2016), which were based solely on water temperature data (Fig. 6C). These patterns suggest that summaries of water temperature data alone can explain substantial variation in thermal regimes over large regions. However, we found that some of the site classifications based upon the Maheu et al. (2016) method exhibited wide variation in annual air-water thermal metric space, suggesting that classifications could be improved by incorporating local air temperature data. We argue that incorporating air temperature data would provide more discriminatory power within regions where local processes like GW may vary greatly but weather patterns, including precipitation, air temperature, and surface flow patterns, are relatively consistent.

The USGS national dataset demonstrated an even greater range in air-water temperature patterns than what was observed in the regional datasets. However, only a few sites exist within any region of the USGS dataset and likely only represent a subset of the variability in air-water temperature patterns within any one region. Still, similar to some OESF sites, some of the USGS sites exhibited strongly reduced M_R and A_R values along with longer $\Delta\phi$ values. These sites fell within the "stable cold" regime, which may be influenced by GW discharge that has been shown to be spatially variable at regional and watershed spatial scales (Johnson et al., 2017; Lowry et al., 2007; Snyder et al., 2015; Wawrzyniak et al., 2017). Several USGS sites exhibited potentially even deeper GW influence than the OESF sites, verifying the existence of strong GW influence even in streams with relatively larger contributing areas. More work is needed to identify the prevalence of shallow to deep GW influence in small to large streams. Slightly negative $\Delta\phi$ values observed in the USGS dataset are likely a result of error in the estimate of air temperature phase when using PRISM data instead of locally observed data. However, hydrologic alteration and seasonally varying flow volumes could also account for these negative values. For the site with the greatest negative $\Delta\phi$ value (gage 11501000), quickly decreasing discharge volumes in spring, which dramatically increase the stream's susceptibility to warming trends, likely explains why the water temperature appears to reach an annual maximum prior to the air temperature. These discharge trends are also likely, in part, due to minor upstream diversions.

4.3. Watershed factors

Previous studies have shown stream temperature to be correlated with watershed characteristics (Daigle et al., 2010; Dugdale et al., 2015; Kelleher et al., 2012; Mayer, 2012; Scott et al., 2002). In this study, the relationships of the annual air-water thermal metrics with three basic watershed properties were explored: median upstream basin elevation, median upstream basin slope, and accumulated upstream basin area. Basin area appeared to have the greatest influence on the three annual thermal metrics, consistently negatively correlated with $\Delta\phi$ and positively correlated with A_R. This suggests that, in general, as streams increase in size, they become more coupled to air temperature dynamics. Correlation with M_R was both positive and negative depending on which dataset was used. Therefore, it may be implied that, in general, larger basins were associated with less relative GW influence. However, there was significant variation in the three annual thermal metrics for small and large basins, with many basins on either end of this spectrum not conforming to this general trend. This apparent contradiction could account for some of the disparity in predictions of the sensitivity of water temperatures to changing air temperatures (Bogan et al., 2003; Isaak et al., 2016; Leach and Moore, 2019; Snyder et al., 2015; Stefan and Preud'homme, 1993). Basin slope was negatively correlated with area and consequently reflected the opposite correlations to the three annual thermal metrics. Negative correlation between basin elevation and M_R is consistent with relatively cool water temperatures at high elevations progressing toward the equilibrium temperature as elevation decreases. Therefore, GW, shading, or other watershed processes may be cooling water temperatures relative to air temperatures at higher elevations. Further exploration of the relationships between the annual thermal metrics and watershed characteristics is needed.

Investigating watershed controls of annual stream temperature patterns on a site level is important, however understanding how the annual stream temperature patterns evolve in the longitudinal direction is crucial for predicting the network extent of suitable thermal habitat for coldwater-adapted species (Fausch et al., 2002; Fullerton et al., 2015). To the best of our knowledge, minimal research has focused on how annual stream temperature signals evolve in a downstream direction. Stream temperature has been shown to exhibit positive spatial autocorrelation tendencies (Isaak et al., 2014). In other words, stream sites that are flow-connected will to some degree exhibit the same influence of their shared upstream processes, but the downstream site will reflect influences that are added between the sites. All of the OESF sites are flow-disconnected, which increases the uncertainty when

interpreting differences in the annual stream temperature metrics between sites due to decreased spatial autocorrelation and the many potential differences in contributing watersheds. Most of the SHEN sites are flow-disconnected but many are also flow-connected, representing both main-stem and tributaries to main-stem sites.

Longitudinal flow connection allows for tracking of these annual stream temperature signals as they travel downstream. Creating a network of temperature dataloggers (e.g., Marsha et al., 2018) could help address critical questions regarding the primary drivers of various facets of the annual stream thermal regime, while still allowing for comparison between separate watersheds or regions. This analysis could be taken further by utilizing the moving window approach presented here to examine the temporal variation in longitudinal evolution of stream temperature and monitor temporal changes in the primary drivers of longitudinal annual stream temperature pattern evolution. As an illustration, graphics and interpretation of the evolution of the annual stream temperature metrics as they migrate downstream in the SHEN watershed of Jeremys Run are provided in Appendix D. However, more research is needed to relate upstream watershed processes to observed changes in the annual stream temperature metrics as the water travels longitudinally downstream.

4.4. Temporal variation

Temporal variation (i.e., within-site or among-regression variation) was found to be relatively small for all three annual thermal metrics derived from both regional datasets as well as the national dataset. In general, $\Delta\phi$ exhibited the greatest variability followed by A_R then M_R. Changes over time to the annual thermal metrics likely correspond to temporal changes in heat fluxes, flow volume, and GW input. Diagnosing the primary drivers of temporal change may be possible by analyzing the simultaneous change between A_R, $\Delta\phi$, and M_R. While this is outside the scope of the current study and warrants further attention, what follows is a brief and hypothetical discussion of these relationships.

Increased flow volume, assuming all else is equal, should result in a decrease in A_R due to water's high heat capacity. However, depending on the source of the increased flow volume (e.g., runoff, snowmelt, GW seepage, industrial waste, etc.), A_R could also increase. Snowmelt, dam operations, and GW seepage may be the only sources of flow volume that could induce a simultaneous large increase in $\Delta\phi$. M_R would likely not change with an increase in flow volume unless the source of the extra water exhibited a significant difference in its mean temperature from that of the air, such as snowmelt.

Unfortunately, it is difficult to interpret sub-annual effects on annual stream temperature signals. For example, increased incoming short-wave radiation due to decreased riparian shading is generally expected to increase A_R and M_R and decrease $\Delta\phi$ as the equilibrium temperature's annual amplitude is increased and stream temperature is pushed towards it (Caissie et al., 2005). Yet, a decrease in riparian shading could result from relatively short-term changes, such as the loss of deciduous riparian cover in autumn, or a longer-term change, such as a large removal of riparian vegetation. Changes that persist for more than approximately one year will be more discernible within the annual parameters. However, sub-annual variation about the annual signal (i.e., residuals) corresponds to characteristics of the annual signal. For example, the water temperature RMSE values are strongly correlated with $\Delta\phi$ ($\rho = -0.709$) and A_R ($\rho = 0.773$) when aggregating the three datasets in this study. Therefore, stronger GW influence will result in less sub-annual variation.

5. Conclusions

Paired air-water annual temperature signals are a promising tool for efficiently diagnosing GW influence and other major controls on stream water thermal regimes (e.g., dam operation, riparian shade, etc.) from readily available data sources. Previous thermal regime classifications based on seasonal water temperatures alone tended to cluster when plotted in air-water annual temperature metric space in expected ways, however, including the paired air-water annual signal analysis adds critical information such as local controls. In this study, two high spatial resolution regional datasets and a coarser spatial resolution national dataset were used to describe subwatershed- to continental-scale variability in derived annual thermal parameters. Patterns in combined air-water annual thermal metrics were discussed in relation to upstream watershed processes and characteristics. High spatial resolution data collection (e.g., SHEN and OESF) can be used to identify major sub-regional patterns in annual air-water temperature metrics, such as the negative relation between A_R and $\Delta\phi$ indicating dominance of shallow GW discharge, but also reach-scale variability within watersheds. Differences in riparian shade can lead to differences in A_R values, but both SHEN and OESF are heavily shaded systems throughout. Therefore, the large range in observed A_R values within each region can be more confidently attributed to variable GW damping of seasonal water temperatures. OESF shows a greater cluster of somewhat deeper GW discharge as indicated by lower A_R without extended $\Delta\phi$.

The national-scale annual air-water temperature metric analysis using USGS GAGES-II data shows broader total variability than the regional-scale analysis, where the metrics appear to be related to physical watershed characteristics such as contributing area, elevation, and slope. This dataset also includes sites with apparent deep GW influence, as indicated by lower A_R with minimal $\Delta\phi$ that is lacking in the regional datasets. However, relatively high spatial resolution tracking of air-water temperature signals reveals strong spatial variation within and between subwatersheds, demonstrating an overwhelming influence of local controls on thermal patterns. Overall, these patterns are surprisingly similar between OESF and SHEN and imply similar controlling hydrogeological processes such as shallow GW discharge. In contrast, the spatially extensive nationwide dataset is likely only sampling a subset of the local variation in any one region. The complex nature of ecosystems and the importance of spatial scale for understanding ecosystem pattern and process has long been a focus in ecology (Cheruvelil et al., 2013; Fahrig, 1992; Fausch et al., 2002; King et al., 2019; Levin, 1992; Stendera et al., 2012) and, more recently, recognizing the connections among and within terrestrial and aquatic ecosystems. Our results demonstrate the importance of spatial scale for inferring hydrological process in streams from annual air-water temperature metrics, and that analysis based on coarse spatial sampling is likely to miss important intra-regional patterns. Further research into watershed- to continental-scale variability in the annual air-water temperature patterns within the U.S. and beyond, including the prevalence of shallow versus deep GW influence, are needed to aid in the prediction of thermal habitat suitability for coldwater-adapted species at ecologically relevant spatial scales.

CRediT authorship contribution statement

Zachary C. Johnson: Conceptualization, Methodology, Writing - review & editing, Data curation. **Brittany G. Johnson:** Conceptualization, Methodology, Writing - review & editing. **Martin A. Briggs:** Conceptualization, Writing - review & editing. **Warren D. Devine:** Data curation, Writing - review & editing. **Craig D. Snyder:** Conceptualization, Data curation, Writing - review & editing. **Nathaniel P. Hitt:** Conceptualization, Data curation, Writing - review & editing. **Danielle K. Hare:** Data curation, Writing - review & editing. **Teodora V. Minkova:** Data curation, 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.

Appendices A-D. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2020.124929.

References

Ahmadi-Nedushan, B., St-Hilaire, A., Ouarda, T.B.M.J., Bilodeau, L., Robichaud, É., Thiémonge, N., Bobée, B., 2007. Predicting river water temperatures using stochastic models: case study of the Moisie River (Québec, Canada). Hydrol. Process. 21, 21–34. https://doi.org/10.1002/hyp.6353.

Anderson, M.P., 2005. Heat as a ground water tracer. Groundwater 43, 951–968. https://doi.org/10.1111/j.1745-6584.2005.00052.x.

Barlow, P.M., Cunningham, W.L., Zhai, T., and Gray, M., 2017. U.S. Geological Survey Groundwater Toolbox version 1.3.1, a graphical and mapping interface for analysis of hydrologic data: U.S. Geological Survey Software Release, 26 May 2017, http://dx.doi.org/10.5066/F7R78C9G.

Bassar, R.D., Letcher, B.H., Nislow, K.H., Whiteley, A.R., 2016. Changes in seasonal climate outpace compensatory density-dependence in eastern brook trout. Glob. Change Biol. 22, 577–593. https://doi.org/10.1111/gcb.13135.

Benyahya, L., Caissie, D., St-Hilaire, A., Ouarda, T.B.M.J., Bobée, B., 2007a. A Review of statistical water temperature models. Can. Water Resour J. Rev. Can. Ressour. Hydr. 32, 179–192. https://doi.org/10.4296/cwrj3203179.

Benyahya, L., St-Hilaire, A., Ouarda, T.B.M.J., Bobée, B., Dumas, J., 2008. Comparison of non-parametric and parametric water temperature models on the Nivelle River. France. Hydrol. Sci. J. 53, 640–655. https://doi.org/10.1623/hysj.53.3.640.

Benyahya, L., St-Hilaire, A., Quarda, T.B.M.J., Bobée, B., Ahmadi-Nedushan, B., 2007b. Modeling of water temperatures based on stochastic approaches: case study of the Deschutes River. J. Environ. Eng. Sci. 6, 437–448. https://doi.org/10.1139/s06-067.

Benz, S.A., Bayer, P., Blum, P., 2017. Global patterns of shallow groundwater temperatures. Environ. Res. Lett. 12, 034005. https://doi.org/10.1088/1748-9326/aa5fb0.

Bogan, T., Mohseni, O., Stefan, H.G., 2003. Stream temperature-equilibrium temperature relationship. Water Resour. Res. 39. https://doi.org/10.1029/2003WR002034.

Boyd, M., Kasper, B., 2003. Analytical methods for dynamic open channel heat and mass transfer: Methodology for heat source model Version 7.0.

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., 2018a. Inferring watershed hydraulics and cold-water habitat persistence using multi-year air and stream temperature signals. Sci. Total Environ. 636, 1117–1127. https://doi.org/10.1016/j.scitotenv.2018.04.344.

Briggs, M.A., Lane, J.W., Snyder, C.D., White, E.A., Johnson, Z.C., Nelms, D.L., Hitt, N.P., 2018b. Shallow bedrock limits groundwater seepage-based headwater climate refugia. Limnologica, Special Issue on Aquatic interfaces and linkages: An emerging topic of interdisciplinary research 68, 142–156. https://doi.org/10.1016/j.limno.2017.02.005.

Briggs, M.A., Lane, J.W., Snyder, C.D., White, E.A., Johnson, Z.C., Nelms, D.L., Hitt, N.P., 2017. Modeled temperature data developed for study of shallow mountain bedrock limits seepage-based climate refugia, Shenandoah National Park, Virginia. U.S. Geological Survey data release. https://dx.doi.org/10.5066/F7F47M8Q.

Busenberg, E., Plummer, L.N., 2014. A 17-Year record of environmental tracers in spring discharge, Shenandoah National Park, Virginia, USA: Use of climatic data and environmental conditions to interpret discharge, dissolved solutes, and tracer concentrations. Aquat. Geochem. 20, 267–290. https://doi.org/10.1007/s10498-013-9202-y.

Caissie, D., 2006. The thermal regime of rivers: a review. Freshw. Biol. 51, 1389–1406. https://doi.org/10.1111/j.1365-2427.2006.01597.x.

Caissie, D., El-Jabi, N., Satish, M.G., 2001. Modelling of maximum daily water temperatures in a small stream using air temperatures. J. Hydrol. 251, 14–28. https://doi.org/10.1016/S0022-1694(01)00427-9.

Caissie, D., El-Jabi, N., St-Hilaire, A., 1998. Stochastic modelling of water temperatures in a small stream using air to water relations. Can. J. Civ. Eng. 25, 250–260. https://doi.org/10.1139/l97-091.

Caissie, D., Satish, M.G., El-Jabi, N., 2005. Predicting river water temperatures using the equilibrium temperature concept with application on Miramichi River catchments (New Brunswick, Canada). Hydrol. Process. 19, 2137–2159. https://doi.org/10.1002/hyp.5684.

Cheruvelil, K.S., Soranno, P.A., Webster, K.E., Bremigan, M.T., 2013. Multi-scaled drivers of ecosystem state: quantifying the importance of the regional spatial scale. Ecol. Appl. 23, 1603–1618. https://doi.org/10.1890/12-1872.1.

Cluis, D.A., 1972. Relationship between stream water temperature and ambient air temperatures: a simple autoregressive model for mean daily stream water temperature fluctuations. Hydrol. Res. 3, 65–71. https://doi.org/10.2166/nh.1972.0004.

Constantz, J., 2008. Heat as a tracer to determine streambed water exchanges. Water Resour. Res. 44. https://doi.org/10.1029/2008WR006996.

Daigle, A., St-Hilaire, A., Peters, D., Baird, D., 2010. Multivariate modelling of water temperature in the Okanagan Watershed. Can. Water Resour J. Rev. Can. Ressour. Hydr. 35, 237–258. https://doi.org/10.4296/cwrj3503237.

DeKay, R.H., 1972. Development of Ground-water Supplies in Shenandoah National Park, Virginia: Charlottesville, VA (Mineral Resources Report 10). Virginia Division of Mineral Resources, Charlottesville, VA.

Dugdale, S.J., Bergeron, N.E., St-Hilaire, A., 2015. Spatial distribution of thermal refuges analysed in relation to riverscape hydromorphology using airborne thermal infrared imagery. Remote Sens. Environ. 160, 43–55. https://doi.org/10.1016/j.rse.2014.12.021.

Dugdale, S.J., Malcolm, I.A., Kantola, K., Hannah, D.M., 2018. Stream temperature under contrasting riparian forest cover: understanding thermal dynamics and heat exchange processes. Sci. Total Environ. 610–611, 1375–1389. https://doi.org/10.1016/j.scitotenv.2017.08.198.

Fabris, L., Malcolm, I.A., Buddendorf, W.B., Soulsby, C., 2018. Integrating process-based flow and temperature models to assess riparian forests and temperature amelioration in salmon streams. Hydrol. Process. 32, 776–791. https://doi.org/10.1002/hyp.11454.

Fahrig, L., 1992. Relative importance of spatial and temporal scales in a patchy environment. Theor. Popul. Biol. 41, 300–314. https://doi.org/10.1016/0040-5809(92)90031-N.

Fausch, K.D., Torgersen, C.E., Baxter, C.V., Li, H.W., 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes: A continuous view of the river is needed to understand how processes interacting among scales set the context for stream fishes and their habitat. Bioscience 52, 483–498. https://doi.org/10.1641/0006-3568(2002)052[0483:LTRBTG]2.0.CO;2.

Franklin, J.F., Dyrness, C.T., 1973. Natural vegetation of Oregon and Washington. (Gen. Tech. Rep. No. PNW-GTR-008). U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR.

Fullerton, A.H., Torgersen, C.E., Lawler, J.J., Faux, R.N., Steel, E.A., Beechie, T.J., Ebersole, J.L., Leibowitz, S.G., 2015. Rethinking the longitudinal stream temperature paradigm: region-wide comparison of thermal infrared imagery reveals unexpected complexity of river temperatures. Hydrol. Process. 29, 4719–4737. https://doi.org/10.1002/hyp.10506.

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. https://doi.org/10.5194/hess-19-3727-2015.

Gathright II, T.M., 1976. Geology of the Shenandoah National Park, Virginia (No. Bulletin 86). Virginia Division of Mineral Resources, Charlottesville, VA.

Glose, A., Lautz, L.K., Baker, E.A., 2017. Stream heat budget modeling with HFLUX: Model development, evaluation, and applications across contrasting sites and seasons. Environ. Model. Softw. 92, 213–228. https://doi.org/10.1016/j.envsoft.2017.02.021.

Graves, S., Piepho, H.-P., Selzer, L., Dorai-Raj, S., 2015. multcompView: Visualizations of paired comparisons. R package version 0.1-7. https://CRAN.R-project.org/package=multcompView.

Groom, J.D., Johnson, S.L., Seeds, J.D., Ice, G.G., 2017. Evaluating links between forest harvest and stream temperature threshold exceedances: The value of spatial and temporal data. JAWRA J. Am. Water Resour. Assoc. 53, 761–773. https://doi.org/10.1111/1752-1688.12529.

Hague, M.J., Patterson, D.A., 2014. Evaluation of statistical river temperature forecast models for fisheries management. North Am. J. Fish. Manag. 34, 132–146. https://doi.org/10.1080/02755947.2013.847879.

Halloran, L.J.S., Rau, G.C., Andersen, M.S., 2016. Heat as a tracer to quantify processes and properties in the vadose zone: a review. Earth-Sci. Rev. 159, 358–373. https://doi.org/10.1016/j.earscirev.2016.06.009.

Hart, E.M., Bell, K., 2015. prism: Download data from the Oregon prism project. R package version 0.0.6. http://github.com/ropensci/prism.

Henderson, J.A., Peter, D.H., Lesher, R.D., Shaw, D.C., 1989. Forested plant associations of the Olympic National Forest (No. R6 Ecol Tech. Paper 001-88). U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR.

Isaak, D.J., Peterson, E.E., Hoef, J.M.V., Wenger, S.J., Falke, J.A., Torgersen, C.E., Sowder, C., Steel, E.A., Fortin, M.-J., Jordan, C.E., Ruesch, A.S., Som, N., Monestiez, P., 2014. Applications of spatial statistical network models to stream data. Wiley Interdiscip. Rev. Water 1, 277–294. https://doi.org/10.1002/wat2.1023.

Isaak, D.J., Wollrab, S., Horan, D., Chandler, G., 2012. Climate change effects on stream and river temperatures across the northwest U.S. from 1980–2009 and implications for salmonid fishes. Clim. Change 113, 499–524. https://doi.org/10.1007/s10584-011-0326-z.

Isaak, D.J., Young, M.K., Luce, C.H., Hostetler, S.W., Wenger, S.J., Peterson, E.E., Hoef, J.M.V., Groce, M.C., Horan, D.L., Nagel, D.E., 2016. Slow climate velocities of mountain streams portend their role as refugia for cold-water biodiversity. Proc. Natl. Acad. Sci. 113, 4374–4379. https://doi.org/10.1073/pnas.1522429113.

Jastram, J.D., Snyder, C.D., Hitt, N.P., Rice, K.C., 2013. Synthesis and interpretation of surface-water quality and aquatic biota data collected in Shenandoah National Park, Virginia, 1979-2009 (USGS Numbered Series No. 2013–5157), Scientific Investigations Report. U. S. Geological Survey, Reston, VA.

Johnson, F.A., 1971. Stream temperatures in an Alpine area. J. Hydrol. 14, 322–336. https://doi.org/10.1016/0022-1694(71)90042-4.

Johnson, S.L., Jones, J.A., 2000. Stream temperature responses to forest harvest and debris flows in western Cascades, Oregon. Can. J. Fish. Aquat. Sci. 57, 30–39. https://doi.org/10.1139/f00-109.

Johnson, Z.C., Snyder, C.D., Hitt, N.P., 2017. Landform features and seasonal precipitation predict shallow groundwater influence on temperature in headwater streams. Water Resour. Res. 53, 5788–5812. https://doi.org/10.1002/2017WR020455.

Kaushal, S.S., Likens, G.E., Jaworski, N.A., Pace, M.L., Sides, A.M., Seekell, D., Belt, K.T., Secor, D.H., Wingate, R.L., 2010. Rising stream and river temperatures in the United States. Front. Ecol. Environ. 8, 461–466. https://doi.org/10.1890/090037.

Kędra, M., Wiejaczka, Ł., 2018. Climatic and dam-induced impacts on river water temperature: assessment and management implications. Sci. Total Environ. 626, 1474–1483. https://doi.org/10.1016/j.scitotenv.2017.10.044.

Kelleher, C., Wagener, T., Gooseff, M., McGlynn, B., McGuire, K., Marshall, L., 2012. Investigating controls on the thermal sensitivity of Pennsylvania streams. Hydrol. Process. 26, 771–785. https://doi.org/10.1002/hyp.8186.

King, K., Cheruvelil, K.S., Pollard, A., 2019. Drivers and spatial structure of abiotic and biotic properties of lakes, wetlands, and streams at the national scale. Ecol. Appl. 29, e01957. https://doi.org/10.1002/eap.1957.

Kothandaraman, V., 1972. Air-water temperature relationship in Illinois River. JAWRA J. Am. Water Resour. Assoc. 8, 38–45. https://doi.org/10.1111/j.1752-1688.1972.tb05091.x.

Kothandaraman, V., 1971. Analysis of water temperature variations in large river. J. Sanit. Eng. Div. 97, 19–31.

Kurylyk, B.L., Irvine, D.J., Bense, V.F., 2019. Theory, tools, and multidisciplinary applications for tracing groundwater fluxes from temperature profiles. Wiley Interdiscip. Rev. Water 6, e1329. https://doi.org/10.1002/wat2.1329.

Kurylyk, B.L., MacQuarrie, K.T.B., Linnansaari, T., Cunjak, R.A., Curry, R.A., 2015. Preserving, augmenting, and creating cold-water thermal refugia in rivers: concepts derived from research on the Miramichi River, New Brunswick (Canada). Ecohydrology 8, 1095–1108. https://doi.org/10.1002/eco.1566.

Kurylyk, B.L., MacQuarrie, K.T.B., Voss, C.I., 2014. Climate change impacts on the temperature and magnitude of groundwater discharge from shallow, unconfined aquifers. Water Resour. Res. 50, 3253–3274. https://doi.org/10.1002/2013WR014588.

Laanaya, F., St-Hilaire, A., Gloaguen, E., 2017. Water temperature modelling: comparison between the generalized additive model, logistic, residuals regression and linear regression models. Hydrol. Sci. J. 62, 1078–1093. https://doi.org/10.1080/02626667.2016.1246799.

Leach, J.A., Moore, R.D., 2019. Empirical stream thermal sensitivities may underestimate stream temperature response to climate warming. Water Resour. Res. 55, 5453–5467. https://doi.org/10.1029/2018WR024236.

Letcher, B.H., Hocking, D.J., O'Neil, K., Whiteley, A.R., Nislow, K.H., O'Donnell, M.J., 2016. A hierarchical model of daily stream temperature using air-water temperature synchronization, autocorrelation, and time lags. PeerJ 4, e1727. https://doi.org/10.7717/peerj.1727.

Levin, S.A., 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur Award Lecture. Ecology 73, 1943–1967. https://doi.org/10.2307/1941447.

Lowry, C.S., Walker, J.F., Hunt, R.J., Anderson, M.P., 2007. Identifying spatial variability of groundwater discharge in a wetland stream using a distributed temperature sensor. Water Resour. Res. 43. https://doi.org/10.1029/2007WR006145.

Lynch, D.D., 1987. Hydrologic conditions and trends in Shenandoah National Park, Virginia, 1983-84 (USGS Numbered Series No. 87–4131), Water-Resources Investigations Report. U.S. Geological Survey, Reston, VA.

Maheu, A., Poff, N.L., St-Hilaire, A., 2016. A classification of stream water temperature regimes in the conterminous USA. River Res. Appl. 32, 896–906. https://doi.org/10.1002/rra.2906.

Marsha, A., Steel, E.A., Fullerton, A.H., Sowder, C., 2018. Monitoring riverine thermal regimes on stream networks: insights into spatial sampling designs from the Snoqualmie River. WA. Ecol. Indic. 84, 11–26. https://doi.org/10.1016/j.ecolind.2017.08.028.

Martens, K.D., 2016. Washington State Department of Natural Resources' Riparian Validation Monitoring Program for salmonids on the Olympic Experimental State Forest—study plan. Washington Department of Natural Resources, Forest Resources Division, Olympia, WA.

Martens, K.D., Devine, W.D., Minkova, T.V., Foster, A.D., 2019. Stream conditions after 18 years of passive riparian restoration in small fish-bearing watersheds. Environ. Manage. 63, 673–690. https://doi.org/10.1007/s00267-019-01146-x.

Mayer, T.D., 2012. Controls of summer stream temperature in the Pacific Northwest. J. Hydrol. 475, 323–335. https://doi.org/10.1016/j.jhydrol.2012.10.012.

Menberg, K., Blum, P., Kurylyk, B.L., Bayer, P., 2014. Observed groundwater temperature response to recent climate change. Hydrol. Earth Syst. Sci. 18, 4453–4466. https://doi.org/10.5194/hess-18-4453-2014.

Minkova, T.L., Hicks, M., Martens, K.D., in press. Ecoregion 7.1.8 Coast Range: Olympic Experimental State Forest, Washington, in: Ryan, D.F., Ed. Biological Responses to Stream Nutrients: A Synthesis of Science from Experimental Forests and Ranges, Gen. Tech. Rep. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR.

Mohseni, O., Stefan, H.G., 1999. Stream temperature/air temperature relationship: a physical interpretation. J. Hydrol. 218, 128–141. https://doi.org/10.1016/S0022-1694(99)00034-7.

National Park Service, 2019a. Fish—Shenandoah National Park, accessed 1 October 2019, at https://www.nps.gov/shen/learn/nature/fish.htm.

National Park Service, 2019b. Trees and Shrubs—Shenandoah National Park, accessed 1 October 2019, at https://www.nps.gov/shen/learn/nature/treesandshrubs.htm.

Nelms, D.L. and Moberg, R.M., Jr., 2010. Hydrogeology and groundwater availability in Clarke County, Virginia. U.S. Geological Survey Scientific Investigations Report 2010-5112.

NOAA, 2019. The National Water Model, accessed 1 October 2019, at https://water.noaa.gov/about/nwm.

Omernik, J.M., 1987. Ecoregions of the conterminous United States. Ann. Assoc. Am. Geogr. 77, 118–125. https://doi.org/10.1111/j.1467-8306.1987.tb00149.x.

Omernik, J.M., Griffith, G.E., 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environ. Manage. 54, 1249–1266. https://doi.org/10.1007/s00267-014-0364-1.

Plummer, L.N., Busenberg, E., Böhlke, J.K., Nelms, D.L., Michel, R.L., Schlosser, P., 2001. Groundwater residence times in Shenandoah National Park, Blue Ridge Mountains, Virginia, USA: a multi-tracer approach. Chem. Geol Hydrochemistry of Springs 179, 93–111. https://doi.org/10.1016/S0009-2541(01)00317-5.

PRISM Climate Group, 2019. Oregon State University. http://prism.oregonstate.edu (downloaded 10.1.19).

Core Team, R., 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rounds, S.A., 2007. Temperature effects of point sources, riparian shading, and dam operations on the Willamette River, Oregon. U.S. Geological Survey Scientific Investigations Report 2007-5185.

Safeeq, M., Grant, G.E., Lewis, S.L., Kramer, M.G., Staab, B., 2014a. A hydrogeologic framework for characterizing summer streamflow sensitivity to climate warming in the Pacific Northwest, USA. Hydrol. Earth Syst. Sci. 18, 3693–3710. https://doi.org/10.5194/hess-18-3693-2014.

Safeeq, M., Mauger, G.S., Grant, G.E., Arismendi, I., Hamlet, A.F., Lee, S.Y., 2014b. Comparing large-scale hydrological model predictions with observed streamflow in the pacific northwest: effects of climate and groundwater. J. Hydrometeorol. 15, 2501–2521. https://doi.org/10.1175/JHM-D-13-0198.1.

Safeeq, M., Grant, G.E., Lewis, S.L., Tague, C.L., 2013. Coupling snowpack and groundwater dynamics to interpret historical streamflow trends in the western United States. Hydrol. Process. 27, 655–668. https://doi.org/10.1002/hyp.9628.

Salas, J.D., 1992. Analysis and modeling of hydrologic time series, in: Maidment, D.R., Ed., Handbook of Hydrology. McGraw-Hill, New York, NY, pp. 19.1-19.72.

Scott, M.C., Helfman, G.S., McTammany, M.E., Benfield, E.F., Bolstad, P.V., 2002. Multiscale influences on physical and chemical stream conditions across blue ridge landscapes. JAWRA J. Am. Water Resour. Assoc. 38, 1379–1392. https://doi.org/10.1111/j.1752-1688.2002.tb04353.x.

Snyder, C.D., Hitt, N.P., Johnson, Z.C., 2018. Air-water temperature data for the study of groundwater influence on stream thermal regimes in Shenandoah National Park, Virginia (ver. 2.0, May 3, 2018). U.S. Geological Survey data release. https://doi.org/10.5066/F7B56H72.

Snyder, C.D., Hitt, N.P., Young, J.A., 2015. Accounting for groundwater in stream fish thermal habitat responses to climate change. Ecol. Appl. 25, 1397–1419. https://doi.org/10.1890/14-1354.1.

Snyder, C.D., Webb, J.R., Young, J.A., Johnson, Z.B., 2013. Significance of headwater streams and perennial springs in ecological monitoring in Shenandoah National Park. USGS Open-File Report 2013–1178. U.S. Geological Survey, Reston, VA.

Southworth, S.J., Aleinikoff, J.N., Bailey, C.M., Burton, W.C., Crider, E.A., Hackley, P.C., Smoot, J.P., Tollo, R.P., 2009. Geologic map of the Shenandoah National Park region, Virginia. USGS Open-File Report 2009-1153. U.S. Geological Survey, Reston, VA.

Stefan, H.G., Preud'homme, E.B., 1993. Stream Temperature Estimation from Air Temperature1. JAWRA J. Am. Water Resour. Assoc. 29, 27–45. https://doi.org/10.1111/j.1752-1688.1993.tb01502.x.

Stendera, S., Adrian, R., Bonada, N., Cañedo-Argüelles, M., Hugueny, B., Januschke, K., Pletterbauer, F., Hering, D., 2012. Drivers and stressors of freshwater biodiversity patterns across different ecosystems and scales: a review. Hydrobiologia 696, 1–28. https://doi.org/10.1007/s10750-012-1183-0.

Taylor, C.A., Stefan, H.G., 2009. Shallow groundwater temperature response to climate change and urbanization. J. Hydrol. 375, 601–612. https://doi.org/10.1016/j.jhydrol.2009.07.009.

U.S. Geological Survey, 2019, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed 1 October 2019, at https://doi.org/10.5066/F7P55KJN.

U.S. Geological Survey, 2011, GAGES-II—Geospatial attributes of gages for evaluating streamflow: U.S. Geological Survey database, accessed 1 October 2019, at https://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml.

Ward, J.C., 1963. Annual variation of stream water temperature. J. Sanit. Eng. Div. 89, 1–16.

Ward, J.V., 1985. Thermal characteristics of running waters. Hydrobiologia 125, 31–46. https://doi.org/10.1007/BF00045924.

Washington State Department of Natural Resources (WADNR), 2019. LiDAR portal, accessed 1 October 2019, at https://lidarportal.dnr.wa.gov.

Washington State Department of Natural Resources (WADNR), 2016. Olympic Experimental State Forest HCP planning unit forest land plan: Final environmental impact statement. Olympia, WA.

Washington State Department of Natural Resources (WADNR), 1997. Final habitat conservation plan. Washington State Department of Natural Resources, Olympia, WA.

Wawrzyniak, V., Allemand, P., Bailly, S., Lejot, J., Piégay, H., 2017. Coupling LiDAR and thermal imagery to model the effects of riparian vegetation shade and groundwater inputs on summer river temperature. Sci. Total Environ. 592, 616–626. https://doi.org/10.1016/j.scitotenv.2017.03.019.

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. Process. 17, 3069–3084. https://doi.org/10.1002/hyp.1280.

Webb, B.W., Hannah, D.M., Moore, R.D., Brown, L.E., Nobilis, F., 2008. Recent advances in stream and river temperature research. Hydrol. Process. 22, 902–918. https://doi.

org/10.1002/hyp.6994.

Westhoff, M.C., Savenije, H.H.G., Luxemburg, W.M.J., Stelling, G.S., van de Giesen, N.C., Selker, J.S., Pfister, L., Uhlenbrook, S., 2007. A distributed stream temperature model using high resolution temperature observations. Hydrol. Earth Syst. Sci. 11, 1469–1480. https://doi.org/10.5194/hess-11-1469-2007.

Zhu, S., Nyarko, E.K., Hadzima-Nyarko, M., 2018. Modelling daily water temperature from air temperature for the Missouri River. PeerJ 6, e4894. https://doi.org/10.7717/peerj.4894.

Zivot, E., Wang, J. (Eds.), 2006. Rolling Analysis of Time Series, in: Modeling Financial Time Series with S-PLUS®. Springer, New York, NY, pp. 313–360. https://doi.org/10.1007/978-0-387-32348-0_9.