Local-scale and watershed-scale determinants of summertime urban stream temperatures

Derek B. Booth1,2*, Kristin A. Kraseski3, and C. Rhett Jackson3

1Bren School of Environmental Science and Management, University of California Santa Barbara, Santa Barbara, CA 93106-5131, USA
2Quaternary Research Center, University of Washington, Box 351310, Seattle, WA 98195-1310, USA
3Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA 30602-2152, USA

Abstract

The influence of urbanization on the temperature of small streams is widely recognized, but these effects are confounded by the great natural variety of their contributing watersheds. To evaluate the relative importance of local and watershed-scale factors on summer temperatures in urban streams, hundreds of near-instantaneous temperature measurements throughout the central Puget Lowland, western Washington State, were collected during a single two-hour period in August in each of the years 1998-2001. Stream temperatures ranged from 8.9 to 27.5°C, averaging 15.4°C. Pairwise correlation coefficients between stream temperature and four watershed variables (total watershed area and the watershed percentages of urban development, upstream lakes, and permeable glacial-outwash soils as an indicator of groundwater exchange) were uniformly very low. Akaike’s information criterion (AIC) was applied to determine the best supported sets of watershed-scale predictor variables for explaining the variability of stream temperatures. For the full 4-year dataset, the only well-supported model was the global model (using all watershed variables); for the most voluminous single-year (1999) data, AIC showed greatest support for percent outwash (Akaike weight of 0.44), followed closely by percent urban development + percent outwash, percent lake area only, and the global model. Upstream lakes resulted in downstream warming of up to 3°C; variability in riparian shading imposed a similar temperature range. Watershed urbanization itself is not the most important determining factor for summer temperatures in this region; even the long-recognized effects of riparian shading can be no more influential than those imposed by other local and watershed-scale factors. Copyright © 2013 John Wiley & Sons, Ltd.

KEY WORDS stream temperature, urban streams, groundwater, lakes, salmon, western Washington

Received 6 January 2013; Accepted 5 March 2013

INTRODUCTION

Decades of study have investigated the causes, and the consequences, of warmed water in rivers and streams (Hannah et al., 2008). Although the causes of elevated stream temperatures in disturbed landscapes are relatively well understood in principle, their quantification in any given watershed is confounded by channel-network geometry, groundwater inflow and hyporheic exchange, and the interplay of stream orientation and sun angle, canopy cover, and air temperature (Smith, 1972; LeBlanc et al., 1997; Rutherford et al., 1997; Poole and Berman 2001). Individual temperature measurements can characterize the state of a particular stream at a particular time, but they do not provide context to evaluate extreme temperature conditions, whether natural or human-induced, or their possible effects on aquatic ecosystems in the context of an entire channel.
network. Remotely sensed temperature data (e.g., Cherkauer et al., 2005), in contrast, can rapidly generate a system-wide perspective but they also typically lack direct information on the causes of observed patterns. They also require sufficiently wide (or barren) rivers to present an aerial view.

Summertime stream temperatures have particular importance for cold-water fisheries and their associated ecosystem (Hawkins et al., 1997; Kemp and Spotila, 1997; Sponseller et al. 2001; Wang et al., 2003). Although criteria for “optimal,” marginal,” and “lethal” temperature ranges have been established through laboratory studies and fortuitous observations in natural rivers, the utility of such determinations are unclear in natural stream networks where heterogeneity is the rule rather than the exception (Poole, 2002; Benda et al., 2004). Salmonids, commonly the cold-water fish taxon of greatest commercial and regulatory concern in temperate latitudes, are both adaptable and mobile; intolerable temperatures can be localized in space, with highly variable temperature distributions across a channel network as a consequence of riparian cover, topographic aspect, and groundwater inputs (Torgersen et al., 1999; Poole and Berman, 2001, Loheide and Gorelick, 2006). Thus single-point determinations of temperature without spatial context have little meaning, even though cumulative ecological effects of elevated temperatures are nonetheless profound (Holthuysen, 1988; Allen, 1995; National Research Council, 1996; Wang and Kanehl, 2003).

Heat is added to and lost from a stream by radiation, sensible heat from inflows and outflows, latent heat by evaporation or condensation, bed conduction, and friction (e.g., Brown, 1969; Webb et al., 2008). Decades of measurements and models demonstrate that the most important term for stream energy balance is the net radiation, which in turn is determined by the sun angle, stream aspect, and canopy cover (Pluhowski, 1970; Beschta and Taylor, 1988; Poole and Berman, 2001). The least important are generally those of conduction and evaporation, while bed conduction and friction are sometimes ignored altogether.

Of the remaining terms, the types and magnitude of sensible heat inputs are quite variable. The presence and influence of cool groundwater inflows depend on both local and regional variations in subsurface geology, soil thickness and permeability, and upland land cover (e.g., Smith and Lavis, 1975; Brunke and Gonser, 1997; Maude and Di Maio, 1999; Tague et al., 2007), with groundwater flows either inferred from hydrologic records (e.g., Kelleher et al., 2012) or more rarely measured directly (Loheide and Gorelick, 2006). In contrast, prior studies of urban stream temperatures typically have focused on the sensible heat contribution of urban runoff, but they have almost exclusively been conducted in regions where thunderstorms fall on recently sun-warmed pavement surfaces that result in runoff up to 5-10°C warmer than the receiving stream (e.g., Pluhowski, 1970; Galli, 1991; Van Buren et al., 2000; Haq and James, 2002, Thompson et al., 2008, Jones et al., 2012), and with the highest runoff temperatures occurring in the mid-afternoon on sunny days during storm events with low total rainfall amounts (Herb et al., 2008). However, these climatological conditions are not ubiquitous: in particular, much of the west coast of North America and other Mediterranean-type climate regions of the world are rarely subject to warm-season rainfall, and which if it does occur is typically during long-overcast periods. Thus, prior work offers surprisingly little insight into a matter of significant regional environmental concern and regulatory attention.

Existing studies, both empirical and model-based, suggest the likely magnitude of stream-temperature changes resulting from human activity, particularly as a result of increased solar radiation on the water surface. Hewlett and Fortson (1982) reported typical water-temperature increases in the southeastern Piedmont of about 3°C (± 3°C) from riparian clearing (and up to about 7°C during the hottest days of a Georgia summer). A pre- and post-clearcutting investigation of a small headwater stream in Pennsylvania (Rishel et al., 1982) showed the average monthly maximum stream-temperature increase to be 4.4°C. Burton and Likens (1973) found increases of 4-5°C in riparian-cleared areas of Hubbard Brook experimental forest, New Hampshire, a similar magnitude to the measured and modeled influence of shading in western Oregon (Risley et al., 2002). Pluhowski (1970) reported urban-induced increases in summertime stream temperatures on Long Island by 5°C to as much as 8°C. LeBlanc and others (1997) investigated various human-induced changes via a calibrated temperature
model for a temperate mid-latitude site; they found that typical simulated temperature increases from vegetation removal to be 2°C from direct solar radiation augmented by increased channel width (resulting from urban-increased discharges) and baseflow reduction.

These studies leave several important questions unanswered. Can the interplay of watershed-scale and more localized influences on stream temperature be unraveled? What is the magnitude of urban influence on summertime stream temperature compared to other factors? Finally, what are the channel-network and regional contexts for a measurement, or set of measurements, from which we might evaluate the biological significance of the resulting temperature data? We therefore conducted the following study to test the hypothesis that streams in urban watersheds experience urban heat island effects, tend to have less riparian shading, and thus have higher maximum temperatures. We also hypothesized that flow from the epilimnion in summer-stratified lakes would raise water temperatures at stream locations with nearby upstream lakes, and that the wider channels and lower elevations associated with increasing watershed area would result in warmer stream temperatures within larger watersheds. Finally, we hypothesized that the deeper flow paths and greater groundwater exchange associated with deep, permeable soils would result in lower stream temperatures wherever the watershed hosts a greater proportion of glacial outwash deposits, the dominant geology in this region associated with permeable soils and abundant groundwater flux.

To evaluate these hypotheses, this study was designed to answer three questions regarding the determinants of maximum summer stream temperatures:

1) Do watershed-scale factors exert a significant, demonstrable influence on summertime stream temperatures in urban watersheds? Specific factors of anticipated potential importance are (a) urban land use, (b) basin area, (c) groundwater-supporting geologic substrate, and (d) upstream lakes.

2) What is the magnitude of the local riparian shading effect on stream temperature relative to the effects of the aforementioned watershed-scale factors?

3) Can simultaneous, spatially widespread measurements of stream temperature provide useful insight into the causes of spatial variability of summertime stream temperatures?

METHODS

Establishing a truly regional context for measured stream temperatures requires broad spatial coverage. A model could nominally accomplish this goal, but the need for calibration data (or, in the absence of calibration, the resulting uncertainty) spurred a search for an alternative approach. At first blush, however, the difficulties of such an alternative also appear daunting—how can enough spatially separated measurements be collected under “equivalent” conditions to define a truly regional pattern of stream temperatures? The challenge is not in taking a temperature measurement itself, but in taking many such measurements without influence of the diurnal temperature cycle, whose intra- and interday variability at any one site may be similar to the spatial patterns of interest over the region as a whole; and to take those measurements with protocols that provide good (or at least known) precision and accuracy. Addressing the research questions of this study therefore required a methodology that could collect a very large number of simultaneous (or near-simultaneous) temperature measurements, using identical protocols, at broadly distributed locations that span a sufficient variety of topographic, geologic, and human influences to represent an entire region.

Data Collection

A heretofore unreported approach was developed to provide this context of regional stream temperature, whose methods we describe in some detail in light of its apparent novelty. Over 100 individuals, representing approximately 20 different agencies and community groups, were organized to collect temperature data at multiple sites within a single geographic region (western Washington) on the same day, at the same time, using the same protocols. Sites were selected to provide coverage of both a broadly distributed region and whole stream systems on a watershed-
wide basis, with drainage areas ranging from over 200 km² down to somewhat less than 1 km², approximately the lower limit of perennial flow in this region (Konrad and Booth, 2002). Reflecting an overriding interest in quantifying human influences, watersheds with primarily urban and suburban land uses were targeted but some rural and forested basins were included as controls (Figure 1).

Generating a truly “regional” data set presented several scientific and logistical problems. The changeability, and unpredictability, of western Washington weather required same-day measurements. Diurnal stream-temperature variations (typically 3-5°C or more; Preud’homme and Stefan, 1992) further narrowed the time interval available to collect “equivalent” data. Yet no agency or institution could install a sufficient number of recording temperature gages, nor field a sufficient number of staff or volunteers, to provide anything approaching the breadth of near-instantaneous spatial coverage envisioned.

An acceptable period of near-maximum and relatively unchanging stream temperatures was determined from existing records. Seasonal patterns showed that early August had the greatest likelihood of yielding annual maximum water temperatures (Figure 2); late afternoon generally experienced the daily temperature peak. The final choice was for a two-hour interval of data collection, balancing the goal of collecting “instantaneous” data unaffected by diurnal changes with the intention of including a large amount of data, over a broad area, with sufficient spatial resolution to discern patterns and trends within individual channel networks. In the end, over 500 temperature measurements across the south-central Puget Lowland were collected in a two-hour period from 3:00 to 5:00 PM on each of four years: 19 August 1998 (n = 555), 3 August 1999 (n = 792), 2 August 2000 (n = 671), and 1 August 2001 (n = 508). In each year of the survey, about 100 one- or two-person teams dispersed throughout the region on predetermined routes during the same two-hour interval. The logistics of such a number of teams required that the date for the effort be selected far in advance of any useable weather forecast.

The sites all lay on lowland streams with most drainage areas under 100 km² (Figure 3); although temperature conditions on the region’s larger rivers are also of concern to fisheries managers, the differing scales would have imposed irreconcilable differences in the choice of date and time of day.

**Figure 1.** Study area and sample sites (open circles). Shading approximates the magnitude of urban land cover, with darkest areas being most urban and unshaded areas being rural, agricultural, and/or forested. Dark X in south-central map area marks National Weather Service station at Seattle-Tacoma International Airport (Table 1).

**Figure 2.** 60-year daily averages of air temperature and precipitation at Seattle-Tacoma International Airport (see Figure 1; Western Regional Climate Center; [http://www.wrccl.dri.edu](http://www.wrccl.dri.edu)) with the range of sampling dates highlighted.
for maximum temperatures (Smith, 1972). Furthermore, the smaller systems that were the study focus have been affected most directly and severely by urban development of their watersheds, which in some cases has resulted in 100 percent urban development in their contributing area (with nearly 75% associated impervious-area coverage). Logistics and sampling efficiency required that all sites lie on or very close to road crossings; given this requirement and density of coverage, locations were generally spaced 1-3 km apart on the streams included in this survey. Most routes revisited their first site for a repeat measurement at the end of the two-hour period; nearly one-quarter of the routes also sampled a site that had been (or would be) sampled by another team in the course of the afternoon, unbeknownst to either group (except where the two groups accidentally met). These two types of semi- and wholly independent measurements have provided a critical measure of the accuracy and precision of the collected data, and they allowed good characterization of the limits of the data’s utility.

Field data collection was designed to be rapid. The requested measurements and observations were specified on a one-page “site form” that could be filled out in less than five minutes and included the following information:

- Basic information to specify the site and to evaluate the route logistics (site number, time of day, directions to the sampling point)
- The basic data needed for the study (air temperature, water temperature)
- Local flow conditions, characterized by four categorical descriptors (free-flowing, sluggish, stagnant, dry)
- Conditions of the nearby riparian canopy, characterized by four categorical descriptors (fully shaded, partially shaded predominantly trees, partly shaded predominantly shrubs, full sun)

To minimize variability in the results, standardized forms and pre-determined routes were mailed to all volunteers (or their agency coordinators). Every volunteer was first asked to calibrate their (typically student lab-grade) thermometer in an ice-water bath in the week before the sampling day, and to note the temperature registered by their thermometer after 10 minutes to the nearest degree (°F) or half-degree (°C). About 60 percent of the volunteers reported a “true” temperature, within the limits of this precision (i.e., ice-water reading less than ±0.5°C); only a few percent in each year had a correction greater than 1°C. The reported calibration value was used to adjust all subsequently reported temperatures from that thermometer.

Each route was planned in accord with the principles of regional coverage, channel-network density, and transportation efficiency. In the first year (1998), routes were assembled from local knowledge and promising road crossings of streams suggested by topographic maps. In subsequent years, reports back from volunteers improved routes by eliminating mis-mapped channels and sites for which private property or heavy brush made access infeasible. For all years, the desired sites were plotted on a road map and sent to the volunteers along with a thermometer-calibration form and individual site forms. Limited time, changing land use, erroneous plotting, and creative volunteers continued to produce final data sets that differed from the planned set and some unusable data that were not included in subsequent analyses. Over the four years of this effort, however, the “yield” of readily located, well-distributed sampling points was over 90 percent of those assigned.

Figure 3. Range of watershed areas draining to temperature sites (based on 1999 survey, excluding replicate measurements). Note x-axis scale change at dashed vertical line.
Data Entry

The field forms were reviewed for any ambiguities in location or obvious errors in recording (e.g., water temperature reported for a “dry” channel). Measurement points were then entered into a GIS system to determine corresponding quantitative watershed characteristics of importance to be used in analysis.

Every point of the 1999 survey (the most voluminous) was first plotted on digital raster graphics (DRG’s) of 7.5-minute topographic maps from the U.S. Geological Survey (USGS), and a script in ArcView 3.2 was executed to determine the UTM coordinates of each plotted point. The coordinates of the remaining points (i.e., not part of the 1999 dataset) were later determined using GoogleEarth and the location descriptions provided by the volunteers. The watershed area draining to each sampling point was determined using the spatial analyst hydrology tools within ArcMap 10 and a 10-m digital elevation model (DEM), compiled from over 100 7.5’ quadrangles accessed from public-domain USGS data locally hosted by the Geomorphological Research Group at the University of Washington (http://gis.ess.washington.edu, accessed April 2012).

Urbanization was quantified from the 2001 National Land Cover Database (30-m raster dataset) accessed from the USGS seamless server (http://gisdata.usgs.gov/website/mrlc/viewer.htm, accessed April 2012) and classified into different levels of urban development to calculate the percentage of developed area within each delineated watershed. Geologic shapefile data were accessed from the Washington State Department of Natural Resources, Division of Geology and Earth Resources (converted to updated shapefiles by the Washington State Geospatial Data Archive, http://wagda.lib.washinton.edu/data/geography/wa_state/wageol, accessed April 2012). This 1:24,000-scale geology data layer was queried for all classifications of glacial outwash and more recent river alluvium, and the percentage of each watershed underlain by these deposits was calculated. Outwash deposits typically have high porosity and conductivity and were thus chosen as an index of groundwater–streamflow interactions. Finally, lake polygons were downloaded from the Washington State Department of Ecology (www.ecy.wa.gov/services/gis/data/data.htm, accessed April 2012); the surface area of each lake was determined, with the presence of lakes (virtually all with surface spillways) considered to be a possible downstream contributor of warm water during the summer. From these data, the percentages of urbanized, outwash, and lake areas were calculated for the watershed area draining to each sampling point.

Data Quality

A major focus of this study was the evaluation of data quality, not only to determine whether the study results carried any useful information about stream temperature but also to evaluate the general utility of volunteer efforts in otherwise impossibly large data-collection efforts. Three approaches were used: (1) evaluation of systematic temperature trends (if any), (2) repeated sites (same observer, different times), and (3) replicate sites (different observers, different times). The first approach evaluated whether the 2-hr measurement window violated the assumption of “uniform” conditions; the second investigated the influence of random fluctuations in temperature and in thermometer reliability. The third was most critical, because it showed not only the variability in data to expect from multiple observers but also the minimum recorded temperature difference that would be meaningful in any subsequent analyses.

Data Analysis

We calculated correlation coefficients between water temperature, watershed area, percent urbanization, percent outwash soils, and percent lake area at each sample point to test the strength and direction of individual associations between temperature and basin-scale factors as well as among the basin factors themselves. For both this correlation analysis and the AIC analysis (below) we evaluated the full dataset (i.e., all four years) and the data from 1999 alone to assess any effects of record length or year-on-year differences. We did not examine the other individual years (1998, 2000, and 2001) because we did not want to confound the analyses with substantial differences in sampled locations. The number of sites was substantially greater in 1999 than in the other years (for example, the 2001 sample set had only 64% of number of sites of 1999), and the 1999 dataset
included more than 90% of the sites in the “all years” dataset.

We plotted the distributions of temperatures in the full shade and open sun sites to quantify the maximum variability in temperature due to local riparian effects. We used the non-parametric Mann-Whitney U test after testing for normality with the Shapiro-Wilk test to test for significance of the difference in the distributions.

Nine candidate models were developed for analysis of the watershed-scale determinants of stream temperature using four predictor variables: total watershed area and the watershed fractions of urban development, upstream lakes, and permeable glacial-outwash soils. The dependent variable was water temperature (specifically, the difference of a site’s water temperature from the year’s average for all sites). To analyze the influence of these variables on measured stream temperature, an information-theoretics approach (“Akaike’s information criterion,” as described in Burnham and Anderson 2002) was used. Akaike’s information criterion (AIC) is an objective means for selecting an estimated “best approximating model” for the data. The tested models included a global model, containing all variables, and 8 submodels with different combinations of the four variables: the four individual factors, and all combinations that included percent urban development. AIC was calculated using SAS, and the relative strength of each candidate model was determined by calculating Akaike’s weights, which can range from 0 to 1. The Akaike weights (w_i) can be interpreted as the probability that a model is the best approximating model based on the dataset and the set of proposed candidate models. All predictor variables were tested using Pearson correlations (see above) to avoid multicollinearity. Because we did not have reliable quantitative data on shade or flow (which are known to affect stream energy budgets), and because the basin-scale correlation coefficients were all very weak (see below), we did not fit models to the “best” sets of variables. The AIC analysis was conducted only to determine which sets of the basin-scale variables had the strongest support for explaining the variability in stream temperatures.

The AIC analysis was run in two ways: with all data pooled, and using only the year with the most data points (1999). The sample size is large (with n/K [1640/4] = 410 for the four-year dataset and [512/4] = 128 for the 1999 data), but Burnham and Anderson (2004) suggest using the second-order bias correction, AIC_c, as a convention even when the sample size (n) is much larger than the number of parameters (K) because AIC_c and AIC values converge as n gets large. As a check, the analysis was run using both AIC and AIC_c: the results were nearly identical and the models were ranked in the same order, so only the results of AIC_c are reported here.

In addition to the AIC analysis of these watershed-scale factors, the data were queried to evaluate the downstream distance (if any) over which the warming effect of a lake had a discernible influence. Sites more than about 3 km downstream were ignored, based on the “buffer reach” length of Sridhar et al. (2004) for dissipating the influence of upstream temperatures (and whose predictions are supported by field measurements of Ham et al. [2006] below detention ponds). This distance also typically resulted in a reduction of “lake-influenced” water to significantly less than half of the total contributing area by the accumulation of non-lake-influenced tributary areas.

RESULTS

Weather Conditions

Conditions in the four sampling years differed only modestly (Table I). The summer of 1998 was somewhat drier than usual, but the sample date (August 19) had exactly the same maximum temperature as for the entire month’s average (23.9°C) at Seattle-Tacoma (“SeaTac”) International Airport, the regional station with the longest record. This day was too late in the summer for what turned out to have been the hottest interval of 1998, and so the results are representative of "normal" but not "extreme" conditions. In contrast, 1999 was a wetter summer than average (July rainfall was 30 mm), and 5.3 mm of rain were recorded on the morning of the sampling day at SeaTac Airport. The day’s maximum air temperature was almost identical to that in 1998. In 2000, monthly precipitation (6 mm in July) and the air temperature on August 2nd (25.6°C) were both drier and warmer than long-term seasonal norms, and no measurable rain had fallen for the 11 preceding days. August 1st 2001 was the coolest day in the four-year set, with a
maximum SeaTac temperature of 21.1°C. July rainfall was 26 mm, with the last precipitation occurring four days before the sampling.

Table I. Summary of weather conditions on the days of sampling

<table>
<thead>
<tr>
<th>Date</th>
<th>Daily max air, $T$ (°C)</th>
<th>July rainfall (mm)</th>
<th>Antecedent dry period (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>19 August 1998</td>
<td>23.9</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>3 August 1999</td>
<td>23.3</td>
<td>30</td>
<td>&lt;1</td>
</tr>
<tr>
<td>2 August 2000</td>
<td>25.6</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>1 August 2001</td>
<td>21.1</td>
<td>26</td>
<td>4</td>
</tr>
</tbody>
</table>

Data are from the National Weather Service records (Seattle–Tacoma International Airport; Figure 1).

Regional Coverage

Overall, the use of multiple volunteer teams over a two-hour measurement period was very successful at achieving a primary goal of the effort, namely the collection of a voluminous and near-synchronous data set (Table II). Of the 2526 site measurements over the four years, about 80% provided unique site-temperature data (i.e., individual sites with flowing water). Over 10% were repeated sites that allowed for semi-independent evaluation of the quality of the collected data. Watershed total impervious areas (TIA) ranged from 4% (rural) to 72% (highly urbanized), with a median TIA for all sites of 28%. Using the National Land Cover Dataset definitions (http://www.epa.gov/mrlc/definitions.html), half of the sites were “Developed, low-intensity” (25–49% TIA), one-third of sites had a watershed TIA <25% (“Developed, open-space” and “Undeveloped”), and the remainder (about 1/6th) were above 50% TIA (“Developed, medium-intensity”). The majority of sites would typically be considered “suburban” but spanning a broad range of urbanization both above and below this category (see also Smith et al. 2010); no sites, however, were fully high-intensity urban.

Data Quality

Our three chosen attributes of data quality were explored by direct comparisons between sites and between observers. The plot of temperature as a function of measurement time (Figure 4) shows no obvious pattern; the (non-significant) least-squares regression line has a slope of less than 0.1°C per hour during the two-hour period. Multiple temperature measurements at the same site were more consistent with the same observer than with a different team, but for both cases the majority of measurements lie within 0.5°C of one another. Over 95 percent of the duplicated readings lie within ±1°C (Table III), which characterizes the useful precision of the data (and which is nearly one order of magnitude less than the range of measured water temperatures: 8.9–22.5°C in 1998, 8.9–24.5°C in 1999, 9.7–26.4°C in 2000, and 10.5–23.5°C in 2001).

Flow and canopy conditions were also recorded at both repeated and replicated sites. Generally (but not invariably), the same team described the same site in the same way, but observations by multiple observers were significantly less reliable. Whereas 97% (244 of 251) repeated measurements by a single observer, for example, did not have any differences in their characterizations of either flow or canopy, only 82% and 51% of flow and canopy observations, respectively, were identically categorized by two different observers. More than 90% between-observer agreement in canopy characterization was achieved by lumping all “shaded” categories together, suggesting that accurate single-site evaluation of shade conditions by multiple observers is little better than “open sun” vs. “not open sun” if criteria are descriptive and locations are not rigidly controlled. In contrast, flow conditions are much more persistent along a reach.
of channel, so minor differences in location apparently produce only modest disagreements.

Figure 4. Variation in air and water temperatures over the 2-hour measurement period (1999 data). Air temperatures declined by about 1°C during the sampling intervals; water temperatures showed no systematic variation at all.

Table III. Summary statistics on repeated water temperature measurements

<table>
<thead>
<tr>
<th>Year</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of T pairs</td>
<td>42</td>
<td>122</td>
<td>97</td>
<td>73</td>
<td>334</td>
</tr>
<tr>
<td>Standard deviation (°C)</td>
<td>0.17</td>
<td>0.53</td>
<td>0.53</td>
<td>0.27</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Interannual Differences

Weather conditions differed between all four years of sampling and were reflected in the resulting stream-temperature data. Because the set of measured sites differed modestly between years, comparative averages are best displayed from only those sites (115, in total) visited in every one of the four years. The annual averages of measured temperatures at these sites spanned about 2°C between the warmest (2000) and coolest (2001) year (Table IV); measured afternoon air temperatures displayed about twice this range. Frequency distributions for water temperature measurements (Figure 5) spanned a similar range in all years at these sites, with nearly all temperatures between 11 and 23°C.

Although continuous temperature gages demonstrate the importance of air temperature and insolation through cyclical diurnal variation patterns, we found no correlation between measured air temperature and measured water temperature across the sites, a result echoed by other studies that show good correlations between air and water temperatures only for the same stream over time (e.g., Mayer, 2012), but weak between-stream associations, particularly for short time periods or simultaneous measurements (Sinokrot and Stefan, 1993; Stefan and Preud’homme, 1993; Webb and Nobilis, 1997; Chen et al., 1998; Pilgrim et al., 1998). Thus any simple correlation between these two parameters, even if developed on a “representative” set of streams, will likely lack predictive value when applied to another stream.

Temperature relationships with watershed-scale variables and riparian condition

All correlation coefficients between temperature and the watershed-scale variables were low (Table V). The highest correlation coefficients with temperature were those of percent outwash and percent lake, 0.14 and 0.10 (all years) and 0.10 and 0.07 (1999), respectively. The correlations were weak but matched the predictions of our hypotheses for all variables except percent urban, which showed temperature virtually (but inversely) uncorrelated with percent watershed urbanization (-0.03 for all years; -0.04 for 1999).

In contrast, canopy conditions did exert a strong effect on stream temperature measurements (Figure 6). Although the replicate results demonstrated that intermediate levels of canopy cover were not reliably discriminated by different observers, “open sun” and “full shade” were consistently categorized and provide a reliable basis to evaluate the relative importance of this factor on the adjacent stream temperature. Median temperatures in shaded sites were two degrees Celsius cooler than in open sun sites (15°C versus 17°C, p<0.0001). Furthermore, the distribution of temperatures in the full shade sites was tighter with less variance. Although this simple, local measurement did not incorporate the thermal inertia imposed by unshaded upstream reaches draining into a locally shaded one (Poole and
Berman, 2001; Mayer, 2012), the effect of local canopy cover and its magnitude was nonetheless substantial and statistically significant.

Table IV. Annual means and range of water temperature only for the 115 stations visited in each of the four years of the study.

<table>
<thead>
<tr>
<th>Date</th>
<th>SeaTac daily max air $T$ (°C)</th>
<th>Average of all sites</th>
<th>Max reported</th>
<th>Min reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>19 August 1998</td>
<td>23.9</td>
<td>14.8</td>
<td>22.5</td>
<td>8.9</td>
</tr>
<tr>
<td>3 August 1999</td>
<td>23.3</td>
<td>15.8</td>
<td>24.5</td>
<td>8.9</td>
</tr>
<tr>
<td>2 August 2000</td>
<td>25.6</td>
<td>15.9</td>
<td>26.4</td>
<td>9.7</td>
</tr>
<tr>
<td>1 August 2001</td>
<td>21.1</td>
<td>13.8</td>
<td>23.5</td>
<td>10.5</td>
</tr>
</tbody>
</table>

Pearson correlation coefficients demonstrated that predictor variables used in the AIC analysis were uncorrelated or weakly correlated ($r<0.500$ for positively correlated variables and $r>-0.500$ for negatively correlated values), suggesting that they were appropriate for use in the candidate models. For the pooled data from all four years, only the global model is supported, with a weight of 0.99 (Table VI). In contrast, analysis of the 1999 data alone yields Akaike weights less than 0.5 (Table...
VII) for all models. The single best supported model for the 1999 data is percent outwash only, and the next best supported model (but with a weight less than half of the best model) is percent urban development + percent outwash. Other models that meet Royall’s (1997) strength of evidence criterion (> 12% of the maximum) are percent lake coverage, the global model, and percent urban development + watershed area + percent outwash. Percent outwash, the watershed-scale indicator of groundwater flux, is a near-ubiquitous predictor variable for the 1999 data, appearing in four of the top five best supported models.

**Figure 6.** Influence of canopy conditions on the distribution of temperatures (1999 data). Box spans the 25th to 75th percentile of the data (median line included); whiskers span the 10th through 90th percentile of the data. The medians differ by 2.0°C (15 vs. 17°C) and are significant at P = <0.001. Quartiles for the full shade sites are 14.0 and 17.0°C, and quartiles for the open sun sites are 15.0 and 19.5°C.

<table>
<thead>
<tr>
<th>Table V. Pearson correlation coefficients for pairwise comparisons of the entire (1998-2001) dataset and the 1999 data alone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water temp difference</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>Entire dataset</td>
</tr>
<tr>
<td>Water temp difference</td>
</tr>
<tr>
<td>Area (km²)</td>
</tr>
<tr>
<td>%Urban</td>
</tr>
<tr>
<td>%Outwash</td>
</tr>
<tr>
<td>%Lake</td>
</tr>
<tr>
<td>1999 only</td>
</tr>
<tr>
<td>Water temp difference</td>
</tr>
<tr>
<td>Area (km²)</td>
</tr>
<tr>
<td>%Urban</td>
</tr>
<tr>
<td>%Outwash</td>
</tr>
<tr>
<td>%Lake</td>
</tr>
</tbody>
</table>

Water temperature was characterized as the difference between the site value and the year’s average value for all sites.

**Temperature Criteria for Salmonid Health**

Temperature thresholds set by the Washington State Department of Ecology to identify thermal stresses to the taxon of greatest concern in Pacific Northwest rivers and streams, members of the Salmonid family, range from 16°C (core summer salmonid habitat) to 17.5°C (salmonid rearing and migration) for the 7-day average of the daily maximum temperatures (Table 200 (1)(c) of WAC 173-201A-200). For the combined all-years dataset, 28% of the measured temperatures were greater than 16°C, and 15% were above 17.5°C, suggesting that a modest fraction of our survey sites could be prone to exceedances were similar weather conditions to those encountered during the surveys persist for a week or more. More directly comparable to our data is the one-day temperature threshold for acute lethality of 23°C; 1% of the sites (5 individual measurements in total) exceeded this value.
DISCUSSION

It is widely assumed that urbanization leads to higher summer stream temperatures (e.g., Walsh et al., 2005; Wenger et al., 2007), given that urban streams feature less canopy cover, urban areas experience higher air temperatures due to heat island effects, and runoff from heated pavement flows directly into stream channels. However, this assumption has rarely been tested at whole-watershed scales (cf. Wang et al. 2003). For this study, we used voluminous, widespread measurements under identical conditions of weather and insolation to explore the correlation of summertime stream temperatures with urbanization in the context of other credible local and watershed-scale drivers.

Although our findings should be relevant to temperate regions worldwide that lack frequent summertime rainfall, the Pacific Northwest is particularly well-suited to explore the effects of human-induced changes on stream temperature—undisturbed watersheds are characterized by extensive forest cover, relatively cool summertime air temperatures, and a hydrologic regime where groundwater generally supplies abundant baseflow during the hottest times of year. Human disturbance of these watersheds has changed those conditions, particularly through the clearing of riparian vegetation and altered hydrologic processes that result from upland soil compaction (Poole and Berman, 2001) and, in the case of urban development, the construction of impervious surfaces (Booth and Jackson, 1997, Cuo et al., 2009).

Although a variety of urban-related factors have been previously suggested by others to influence stream temperatures, our data show that the fraction of developed land in the upstream watershed, on its own, does not provide much predictive value, a result echoed by some but not all prior studies (e.g., Wang and Kanehl 2003). These results do not contradict the measured, and long-recognized, influence of riparian canopy clearing, a condition not requiring “urban development.” Indeed, some of the most consistently cleared systems are in agricultural settings with very little imperviousness in their

<table>
<thead>
<tr>
<th>Candidate model</th>
<th>$AIC_c$</th>
<th>$\Delta AIC_c$</th>
<th>$w_i$</th>
<th>% maximum $w_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>6351.42</td>
<td>0.0000</td>
<td>0.99718</td>
<td>100</td>
</tr>
<tr>
<td>Per cent urban development, watershed area, and per cent outwash</td>
<td>6364.57</td>
<td>13.1591</td>
<td>0.00138</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Per cent urban development and per cent outwash</td>
<td>6365.70</td>
<td>14.2822</td>
<td>0.00079</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Per cent outwash</td>
<td>6366.11</td>
<td>14.6933</td>
<td>0.00064</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Per cent lake coverage</td>
<td>6382.86</td>
<td>31.4404</td>
<td>0.00000</td>
<td>0</td>
</tr>
<tr>
<td>Per cent urban development and per cent lake coverage</td>
<td>6383.21</td>
<td>31.7975</td>
<td>0.00000</td>
<td>0</td>
</tr>
<tr>
<td>Watershed area</td>
<td>6394.17</td>
<td>42.7572</td>
<td>0.00000</td>
<td>0</td>
</tr>
<tr>
<td>Per cent urban development and watershed area</td>
<td>6394.34</td>
<td>42.9235</td>
<td>0.00000</td>
<td>0</td>
</tr>
<tr>
<td>Per cent urban development</td>
<td>6397.14</td>
<td>45.7206</td>
<td>0.00000</td>
<td>0</td>
</tr>
</tbody>
</table>

$AIC_c$, Akaike's information criterion with bias correction; $w_i$, Akaike weights.

<table>
<thead>
<tr>
<th>Candidate model</th>
<th>$AIC_c$</th>
<th>$\Delta AIC_c$</th>
<th>$w_i$</th>
<th>% maximum $w_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per cent outwash</td>
<td>2069.50</td>
<td>0.0000</td>
<td>0.44131</td>
<td>100</td>
</tr>
<tr>
<td>Per cent urban development and per cent outwash</td>
<td>2071.17</td>
<td>1.6698</td>
<td>0.19148</td>
<td>43</td>
</tr>
<tr>
<td>Per cent lake coverage</td>
<td>2072.44</td>
<td>2.93736</td>
<td>0.10160</td>
<td>23</td>
</tr>
<tr>
<td>Global</td>
<td>2072.83</td>
<td>3.32655</td>
<td>0.08364</td>
<td>19</td>
</tr>
<tr>
<td>Per cent urban development, watershed area, and per cent outwash</td>
<td>2073.17</td>
<td>3.66970</td>
<td>0.07045</td>
<td>16</td>
</tr>
<tr>
<td>Per cent urban development and per cent lake coverage</td>
<td>2074.31</td>
<td>4.80946</td>
<td>0.03985</td>
<td>9</td>
</tr>
<tr>
<td>Per cent urban development</td>
<td>2074.84</td>
<td>5.34294</td>
<td>0.03052</td>
<td>7</td>
</tr>
<tr>
<td>Watershed area</td>
<td>2074.94</td>
<td>5.43984</td>
<td>0.02907</td>
<td>7</td>
</tr>
<tr>
<td>Per cent urban development and area</td>
<td>2076.70</td>
<td>7.19480</td>
<td>0.01209</td>
<td>3</td>
</tr>
</tbody>
</table>

$AIC_c$, Akaike's information criterion with bias correction; $w_i$, Akaike weights.
watershed. More generally, the land cover adjacent to any particular site along a stream is only weakly correlated to the land cover of the watershed as a whole (e.g., Morley and Karr, 2002; Carroll and Jackson, 2008) and so the importance of the former does not guarantee that is will be well-characterized by the latter.

Unlike urban land cover, however, geology was identified as an important watershed-scale factor in our analysis, particularly for the single-year (1999) analysis. A well-sampled watershed (Soos Creek, a 180-km$^2$ catchment in the east-central part of the study area with 74 individual site measurements in 1999) provides a useful case study (Figure 7). The northern part, which encompasses the upper Soos Creek drainage and the headwaters of one of its two major tributaries, Jenkins Creek, is a rolling till-mantled upland surface with shallow perched groundwater, small headwater channels that dry rapidly in the summer, and abundant seasonal wetlands that reflect the shallow perched water table. The highest recorded temperatures were found in the highest headwater reaches, perched on till and with presumably little or no groundwater influence. The southeastern part of the watershed, in contrast, is a broad plain of outwash, punctuated by till and bedrock hills around which the main stream channels of Jenkins and Covington creeks flow towards their confluences with Soos Creek. Streams, associated wetlands, and lakes are all expressions of the regional water table, which fluctuates only slowly throughout the year and produces both steady discharges in the summertime and slow response to rainfall throughout the year. Stream temperatures here are high only downstream of one lake, but they otherwise show a strong moderating thermal influence, presumably of deep groundwater and particularly along the axis of the upper Soos Creek watershed and in the lower parts of both Jenkins and Covington creeks.

Figure 7. Stream temperatures in the Soos Creek basin (1999 data), displayed spatially with respect to dominant geologic deposits and inferred groundwater flow. Shallow subsurface flow perched over glacial till (shaded region) correlates with relatively warm temperatures (typically 17°C or more) or dry channels; to the southeast, deep groundwater flow through glacial outwash sands and gravels correlates with a marked ~3°C cooling of water, particularly in the lower mainstem channels of Covington and Jenkins creeks. Cool temperatures are also found in the main channel of Soos Creek, which flows through a broad, outwash-filled valley.

Figure 8. Trend in water cooling downstream of a lake outlet. A consistent trend is present for about 2 km downstream of an outlet and suggests a range of 2-3°C warming associated with the lake.
Lakes would be expected to raise downstream temperatures, given that lakes in this region are unshaded and typically drain from the surface (Stefan and Preud’homme 1993). This inference is not well-supported by the AIC analysis, although displaying these lake-influenced water temperatures against their distance downstream from a lake displays a noisy (but suggestive) negative trend (Figure 8). “Lake-influenced” water is on average about 2–3°C warmer than would otherwise be anticipated. Parenthetically, these results suggest that the local strategy of constructing permanent open-water ponds for water quality improvement (e.g., Comings et al., 2000) may simply be trading one pollution problem (elevated metal and phosphorus) for another (elevated temperatures) (Lieb and Carline, 2000).

The results confirm many theoretical expectations and past empiricisms. Human influence is a noteworthy determinant of stream temperature, particularly not exclusively through the clearing of riparian vegetation. Watershed-scale changes in land cover are much less influential, however, at least under conditions of summertime low flows. If urbanization reduces groundwater recharge during the rainy season and so results in lower water tables, the effects in this region on summertime stream temperature are apparently either inconsequential or roughly balanced by the additional baseflow contribution from landscape watering or septic systems (Konrad and Booth, 2005).

CONCLUSIONS
In aggregate, a variety of different local and watershed-scale drivers, operating at many different scales and locations, interact such that only the most consistently influential impose any systematic trends in summertime stream temperature. For the Pacific Northwest under the typical climatic conditions that give rise to annual maximum water temperature, these key effects are watershed geology, the effects of riparian vegetation, and the warming from upstream lakes. Each of these factors can influence potential stream temperatures by about 2–3°C. Watershed land use, considered independent of these other effects, does not approach this degree of systematic influence.

These results also display both the value, and the limitation, of volunteer data. Replicability is good but by no means perfect. Inconsistencies render certain types of information nearly useless and place irreducible limits on the precision of others. Yet such efforts can generate information obtainable in no other way, and the results can well justify the effort involved and the inescapable limitations. Here, they have provided an unusually voluminous, synoptic characterization of a particular attribute of the region’s streams, setting a context to interpret local stream temperature measurements and to guide present-day restoration efforts. These results also serve to remind that an urban landscape does not support fully functioning natural watershed processes and conditions, but that the cause(s) of that impairment can be quite variable with respect to both the drivers of degradation and the environmental response.

ACKNOWLEDGEMENTS
Many agencies, tribes, and citizen groups contributed personnel and logistical support to this effort. These included King County, Snohomish County, Pierce County, Kitsap County, Tulalip Tribe, Stillaguamish Tribe, Puyallup Tribe, Muckleshoot Tribe, Washington Department of Fish and Wildlife, City of Bellevue, City of Seattle, City of Renton, City of Tukwila, Washington Trout, Adopt-a-Stream, Thornton Creek Alliance, and the Ravenna Creek Alliance. Financial support was provided as part of the United States Environmental Protection Agency and National Science Foundation Water and Watersheds Program, EPA STAR Grant R82-5284-010. Additional support came from the Coweeta Long Term Ecological Research program (NSF Award DEB-0823293). Special thanks are due Leslie Wall of the Center for Water and Watershed Studies, University of Washington, who provided critical organizational assistance throughout the four years of the sampling project. Jeffrey Hepinstall-Cymerman for his help with planning GIS data acquisition, and Robert Bahn for his valuable contribution to the statistical analysis.
REFERENCES

Kemp SJ, Spotila JR. 1997. Effects of urbanization on brown trout Salmo trutta other fishes and macroinvertebrates in Valley Creek Valley Forge


