Journal of Hydrology 517 (2014) 187–200

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

A regional neural network ensemble for predicting mean daily river water temperature

Jefferson Tyrell DeWeber[a,*], Tyler Wagner[b]

[a] Pennsylvania Cooperative Fish and Wildlife Research Unit, The Pennsylvania State University, 413 Forest Resources Building, University Park, PA 16802, USA
[b] U.S. Geological Survey, Pennsylvania Cooperative Fish and Wildlife Research Unit, The Pennsylvania State University, 402 Forest Resources Building, University Park, PA 16802, USA

ARTICLE INFO

Article history:
Received 2 October 2013
Received in revised form 17 March 2014
Accepted 5 May 2014
Available online 22 May 2014
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Fi-John Chang, Associate Editor

Keywords:
Water temperature
Neural network
River
Stream
Ensemble
Regional prediction

SUMMARY

Water temperature is a fundamental property of river habitat and often a key aspect of river resource management, but measurements to characterize thermal regimes are not available for most streams and rivers. As such, we developed an artificial neural network (ANN) ensemble model to predict mean daily water temperature in 197,402 individual stream reaches during the warm season (May–October) throughout the native range of brook trout *Salvelinus fontinalis* in the eastern U.S. We compared four models with different groups of predictors to determine how well water temperature could be predicted by climatic, landform, and land cover attributes, and used the median prediction from an ensemble of 100 ANNs as our final prediction for each model. The final model included air temperature, landform attributes and forested land cover and predicted mean daily water temperatures with moderate accuracy as determined by root mean squared error (RMSE) at 886 training sites with data from 1980 to 2009 (RMSE = 1.91 °C). Based on validation at 96 sites (RMSE = 1.82) and separately for data from 2010 (RMSE = 1.93), a year with relatively warmer conditions, the model was able to generalize to new stream reaches and years. The most important predictors were mean daily air temperature, prior 7 day mean air temperature, and network catchment area according to sensitivity analyses. Forest land cover at both riparian and catchment extents had relatively weak but clear negative effects. Predicted daily water temperature averaged for the month of July matched expected spatial trends with cooler temperatures in headwaters and at higher elevations and latitudes. Our ANN ensemble is unique in predicting daily temperatures throughout a large region, while other regional efforts have predicted at relatively coarse time steps. The model may prove a useful tool for predicting water temperatures in sampled and unsampled rivers under current conditions and future projections of climate and land use changes, thereby providing information that is valuable to management of river ecosystems and biota such as brook trout.

1. Introduction

Water temperature is a fundamental property of river habitat that shapes biological communities and determines ecosystem services. Water temperature can limit the distribution of species through physiological constraints and thus is an important factor in structuring aquatic assemblages (Caissie, 2006; Magnuson et al., 1979). River water temperature also places constraints on river metabolism and ecosystem services that depend upon energy transfers (Demars et al., 2011). Human activities that alter rivers directly (e.g., dams; reviewed in Olden and Naiman, 2010) or indirectly through changes to the landscape (e.g., land use; reviewed in Poole and Berman, 2001) can alter water temperatures. Global climate change is also expected to result in warmer river water temperatures (e.g., Mohseni et al., 1999; Nelson and Palmer, 2007; van Vliet et al., 2013) primarily as a result of increased air temperatures, and reduced summer flows may further exacerbate water temperature increases (Isaak et al., 2010; van Vliet et al., 2013). These changes are likely to affect riverine biota and may act independently or in conjunction with other abiotic or biotic factors to render river habitat unsuitable for some species (Ficke et al., 2007; Rahel and Olden, 2008). For example, stream warming due to climate change is predicted to have negative effects on coldwater fish species, such as Pacific salmon (*Oncorhynchus* spp.; Ruesch et al., 2012), but may also result in the upstream expansion of an introduced predator (smallmouth bass *Micropterus dolomieu*; Lawrence et al., 2012). Thus, the combined effects of physiological stress and expanding ranges of introduced predators could interact to have large negative effects on native coldwater fish populations. Because of its importance to biota and susceptibility to human

---

* Corresponding author. Tel.: +1 814 317 6591.
E-mail addresses: jtdeweber@gmail.com (J.T. DeWeber), txw19@psu.edu (T. Wagner).

http://dx.doi.org/10.1016/j.jhydrol.2014.05.035
0022-1694/

activities and climate change, river water temperature and anticipated changes resulting from climate and land use changes are of great interest for resource management and biodiversity conservation.

Although technological advances have made monitoring river water temperature comparatively feasible and inexpensive in recent years (Webb et al., 2008), it is still logistically infeasible to measure, and difficult to obtain existing data, for a significant portion of river reaches across large basins or regions due to limited fiscal resources for monitoring and a lack of coordination among various research programs (Isaak, 2011). As a result, models predicting river water temperature characteristics for unsampled time periods, in unsampled rivers or under alternative management or environmental scenarios have become common in recent years (e.g., Hill et al., 2013; Isaak et al., 2010; Mohseni et al., 1998; Nelson and Palmer, 2007; Wehrly et al., 2009). For example, models are useful for making predictions of water temperature under future climate (Isaak et al., 2010; Mohseni et al., 1999), alternative land use scenarios (Hill et al., 2013; Nelson and Palmer, 2007; Sugimoto et al., 1997), or various water release scenarios from impoundments (Olden and Naiman, 2010; Wright et al., 2009). Models are also useful for understanding the processes that control river water temperature (e.g., Johnson, 2004; Story et al., 2003). Models predicting river water temperature range from deterministic models that require detailed meteorological and hydrological data used to solve heat budget equations (e.g., Johnson, 2004; Story et al., 2003) to empirical models with varying degrees of spatial complexity (e.g., Ruesch et al., 2012) that rely upon relationships between water temperature observations and relatively easy to collect climatic and landscape variables (e.g., Chenard and Caissie, 2008; Hill et al., 2013; Isaak et al., 2010; Mohseni et al., 1998). Although deterministic models can perform well and are physically based, the detailed data on river-specific energy transfers that are required to develop these models makes transferability to other rivers difficult. By contrast, empirical models are often more easily transferable and thus more useful for predicting river water temperatures at unmonitored locations throughout large watersheds or regions to support local and transboundary management efforts (Caissie, 2006).

Hourly or daily variation in water temperature can be important for stream ecosystem functioning, and some models have predicted daily water temperature with moderate accuracy in individual streams using only air temperature (e.g., Caissie et al., 2001). However, because water temperature variability generally increases with the number of streams, empirical models for predicting in multiple streams and across regions usually predict at weekly, monthly or seasonal time steps to achieve reasonable accuracy (Caissie, 2006). The loss of temporal variation in predictions is undesirable because daily predictions could provide more information and can be summarized to yield weekly, monthly or seasonal metrics as needed. Prediction in geographically diverse basins and over large spatial extents is also improved by including landform, geological, and stream attributes that are directly or indirectly related to water temperature as predictors (e.g., Hill et al., 2013; Isaak et al., 2010; Wehrly et al., 2009). There are a growing number of empirical modeling techniques that allow for multiple predictors and have been used for predicting water temperature (e.g., regression, stochastic models with time series decomposition, geospatial models, machine learning). Artificial neural networks (ANNs) are a particularly promising machine learning method because they are able to model nonlinear relationships, handle interactions among predictors, and often have high predictive power (Lek and Guégan, 1999; Olden et al., 2008). ANNs have been used widely and often outperformed other methods for predicting streamflow (e.g., Besaw et al., 2010; Chen et al., 2013; Huo et al., 2012), dissolved oxygen (e.g., Antanasijević et al., 2013; Wen et al., 2013), fish species distributions (Olden and Jackson, 2002) and richness (Chang et al., 2013), and water temperature (e.g., Chenard and Caissie, 2008; Risley et al., 2003; Westenbroek et al., 2010).

Although predicting river water temperature is of importance for the management and conservation of many aquatic species (Domisch et al., 2011; Xenopoulos et al., 2005), it is of particular importance for the conservation of cold-water salmonids (Almodóvar et al., 2012; Jones et al., 2006; Isaak et al., 2010; McKenna et al., 2010; Ruesch et al., 2012,), including brook trout *Salvelinus fontinalis*. Brook trout is a species of management concern throughout much of its native range in the eastern U.S., and the Eastern Brook Trout Joint Venture (EBTJV, http://easternbrooktrout.org/) was formed to promote regional, transboundary management and conservation. Brook trout are limited physiologically to coldwater (mean July water temperature <~22 °C) streams, rivers and lakes and are sensitive to habitat and biotic disturbances (MacCrimmon and Campbell, 1969). An EBTJV assessment concluded that brook trout populations were extirpated or reduced (>50% of previously suitable habitat lost) in >71% of subwatersheds, and these losses were attributed to human activities, which include historical forestry practices, habitat alterations, nonnative species introductions and recent land use changes (Hudy et al., 2008). Future water temperature increases as a result of global climate warming are expected to result in further losses of brook trout habitat throughout their native range in eastern North America (Clark et al., 2001; Flebbe et al., 2006; Meisner, 1990). Even where temperatures rise but remain suitable, brook trout growth could be reduced unless food availability and consumption increase with temperature (Ries and Perry, 1995). Past predictions of brook trout range shifts in the eastern U.S. due to climate change were made by identifying thermally suitable habitat based on surrogates of river water temperature (e.g., elevation, groundwater temperature as determined by mean annual air temperature), and overlaying projected air temperature changes to determine potential habitat losses (Flebbe et al., 2006; Meisner, 1990). Combining predicted river water temperature with thermal limits represents a more direct route for characterizing current thermally suitable habitat and future changes due to climate change.

To assist in the management of rivers and brook trout in the eastern U.S., we developed an ensemble model of 100 ANNs to predict mean daily river water temperature for the majority of streams throughout the brook trout's native range in the eastern U.S. We first compared four models of increasing complexity to determine how well daily water temperatures could be predicted by the following sets of predictors: (1) air temperature only, (2) air temperature and landform attributes, (3) air temperature, landform attributes and forested land cover, and (4) air temperature, landform attributes, and forest, agricultural and developed land covers. We then select a final model and demonstrate its utility by mapping predicted water temperatures averaged for the month of July across the 1980–2010 modeling period. Our ensemble approach proves useful for understanding the importance of predictor variables and we are not aware of other models described in the peer-reviewed literature that predict daily water temperatures in individual stream reaches throughout a similarly large region.

2. Study area

The study region included the native range of brook trout in the eastern U.S. as defined by the EBTJV, and represents approximately 30% of the worldwide native range of brook trout and 70% of its range in the U.S. (Fig. 1; Hudy et al., 2008). We modified the EBTJV region slightly to align with the boundaries of local catchments from the National Hydrography Dataset Plus Version 1.0