LIMNOLOGY
and
OCEANOGRAPHY

ASLO

Limnol. Oceanogr. 00, 2017, 00–00

doi: 10.1002/lno.10626

On the predictability of lake surface temperature using air temperature in a changing climate: A case study for Lake Tahoe (U.S.A.)

S. Piccolroaz,[1,2]* **N. C. Healey**,[3] **J. D. Lenters**,[4,5] **S. G. Schladow**,[6,7] **S. J. Hook**,[3] **G. B. Sahoo**,[6,7] **M. Toffolon**[1]

[1]Department of Civil, Environmental and Mechanical Engineering, University of Trento, Trento, Italy
[2]Department of Physics and Astronomy, Institute for Marine and Atmospheric research Utrecht, Utrecht University, Utrecht, The Netherlands
[3]NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California
[4]Department of Geography, University of Colorado at Boulder, Boulder, Colorado
[5]Center for Limnology, University of Wisconsin-Madison, Boulder Junction, Wisconsin
[6]Department of Civil and Environmental Engineering, University of California, Davis, California
[7]Tahoe Environmental Research Center, University of California, Davis, California

Abstract

Can we predict long-term trends of lake surface temperature based on air temperature alone? We explore this question by analyzing the performance of a hybrid model (*air2water*) as a predictive tool for defining scenarios of lake surface temperature in the framework of climate change studies. Employing Lake Tahoe (U.S.A.) as a case study, we apply the model using different air temperature datasets (in situ measurements, gridded observations, and downscaled General Circulation Models). Through a data-driven calibration of the model parameters based on surface water temperature records, we show that *air2water* provides good performance (root mean square error ∼ 0.5°C, on a monthly scale) regardless of the input dataset. The model is able to accurately capture the historical long-term trend and interannual fluctuations over decades (from 1969 to present), using only 7 yr of monthly measurements of surface water temperature for calibration. Additionally, when used to predict future surface water temperature of the lake, *air2water* produces the same projections irrespective of the air temperature dataset used to drive the model. This is certainly desirable, but not immediately expected when using a relatively simple model. Overall, the results suggest the high potential and robustness of *air2water* as a predictive tool for climate change assessment. Lake surface temperature warming of up to 1.1°C (RCP 4.5) and 2.9°C (RCP 8.5) was simulated at the end of the 21st century during summer months in Lake Tahoe. Such a scenario, if realized, would lead to serious consequences on lake water chemistry, primary productivity, plankton community structure, and nutrient cycling.

During recent decades, many lakes throughout the world have undergone rapid warming of their surface water temperature (Quayle et al. 2002; Livingstone 2003; Coats et al. 2006; Austin and Colman 2007; Schneider et al. 2009; Schneider and Hook 2010; O'Reilly et al. 2015; Sharma et al. 2015; Woolway et al. 2016). Since water temperature is a key physical variable controlling a large variety of physical, ecological, and biogeochemical processes, future changes in climate conditions and the resulting meteorological forcing at the air–water interface may significantly alter the overall environmental and water quality status of lakes. Indeed, changes in the thermal behavior of lakes lead to significant consequences for stratification and mixing regimes (Butcher et al. 2015; Kraemer et al. 2015; Piccolroaz et al. 2015; Sahoo et al. 2016; Wood et al. 2016) and in the community structure of many habitats (e.g., Winder et al. 2009; De Senerpont et al. 2013; Schabhüttl et al. 2013), with possible modifications of the biochemical compositions of some algae species (e.g., Flaim et al. 2014). This may have significant implications for local economies, since lakes provide important services such as drinking water, agricultural irrigation, power generation, and cooling water for industries, water transport corridors, recreation and tourism, and support for freshwater fisheries.

For many lakes, an increase of lake surface water temperature (*LSWT*) in summer occurs in conjunction with an increase in the duration and intensity of the stratification

*Correspondence: s.piccolroaz@unitn.it, s.piccolroaz@uu.nl

Additional Supporting Information may be found in the online version of this article.

period, with increased inhibition of vertical exchanges of mass, energy, and momentum between the epilimnion and hypolimnion In turn, this may have important consequences for the oxygenation of deep water layers, possibly jeopardizing the sustainability of aquatic life in the aphotic zone and facilitating the regeneration of soluble phosphorus and metals from lake sediments. The final effect is the degradation of the ecological health of the lake and an increasing risk of summer cyanobacterial blooms (e.g., Gallina et al. 2011; Dokulil 2014; Butcher et al. 2015), including toxic species (Gallina et al. 2013). Additionally, an increase of *LSWT*, especially in winter, may result in a drastic modification of the thermal structure and mixing regime of the lake (Blenckner et al. 2002; Salmaso 2005, 2010; Wood et al. 2016). For example, lakes may change from dimictic to monomictic, from monomictic to oligomictic, or experience a reduced frequency of overturn events in oligomictic lakes, ultimately leading to meromixis. This can be particularly critical in deep lakes (Wood et al. 2016), where a lower frequency of deep mixing events may lead to a progressive reduction of oxygen concentrations in deep layers, eventually reaching hypoxic conditions in the whole water column during turnover events and, possibly, subsequent eutrophication.

Understanding, predicting, and quantifying the thermal response of lakes to evolving climate conditions is critical for future decision making involving water resource management policies. However, trying to make reliable projections of how *LSWT* is likely to evolve in the future is not trivial. In fact, water temperature is a result of a combination of complex thermodynamic fluxes occurring simultaneously, all summing to the net heat flux of a lake (*see* e.g., Henderson-Sellers 1986; Imboden and Wüest 1995). In principle, the accurate prediction of *LSWT* would require the precise estimate of all lake-atmosphere energy fluxes, which can only be accomplished if accurate high-frequency observational data are available (*see* e.g., Woolway et al. 2015 for a MATLAB script to calculate surface energy fluxes in lakes). However, such detailed information is not always available, and this has stimulated the development of *LSWT* prediction models of different types and complexities, from purely regressive/statistical models (McCombie 1959; Webb 1974; Livingstone and Lotter 1998; Kettle et al. 2004; Sharma et al. 2008), to more complex process-based numerical models (e.g., Hamilton and Schladow 1996; Perroud et al. 2009; Martynov et al. 2010; Thiery et al. 2014). As is often the case, both families of models have significant advantages but also substantial shortcomings. Simple regression models usually require few inputs, generally only air temperature (*AT*), but they are not able to address some fundamental physical processes (e.g., thermal stratification), and their use is controversial when applied with *AT* ranges beyond the limits of the time series used for model calibration (e.g., in climate changes studies). Conversely, deterministic models provide a more detailed description of the thermal processes in the lake and their interaction with the surrounding environment, but they generally require a large amount of input data (e.g., detailed and spatially distributed time series of all meteorological variables, geomorphological information, streamflow data), which are often not available for long periods or with sufficient temporal and spatial resolution. Besides these two main model categories, a third approach exists, which is represented by models characterized by a hybrid structure combining a physically based derivation of the governing equations with a statistical calibration of model parameters. This approach is aimed at retaining the simplicity of statistical models while preserving the robustness of deterministic models.

The objective of this study is to assess whether using *AT* alone provides sufficiently reliable estimates of how *LSWT* will respond to changing climate conditions, duly accounting for the limitations that such a simplified approach may introduce. *AT* predictions based on General Circulation Models (GCMs) or Regional Climate Models (RCMs) are usually more reliable than other meteorological variables (e.g., Gleckler et al. 2008), and downscaling is typically associated with smaller uncertainties (Dettinger 2013). Producing reliable projections of the future *LSWT* using only *AT* would certainly be a major advantage for many scientific purposes and practical applications. This is even more attractive if the tool used for *LSWT* prediction is simple, and thus accessible to scientists with different mathematical, physical, and technical backgrounds (e.g., physicists, biologists, engineers, etc). Clearly, this should not be intended in any manner as a justification for the indiscriminate use of *AT* alone to predict *LSWT*, legitimizing the use of modeling tools that are not appropriate for this aim (as is generally the case of purely regression models).

Here, we use a simple hybrid model developed by Piccolroaz et al. (2013) called *air2water*, which predicts *LSWT* and provides a measure of the depth of the well-mixed surface layer, based on *AT* only. The *air2water* model has been shown to provide similar performance to those of more complex, process-based models even though they generally require more extensive inputs (root mean square error [RMSE] on the order of 1°C for daily temperatures, but less for longer averaging windows; *see* Piccolroaz et al. 2016). Additionally, the *air2water* model has been shown to be an effective tool to investigate the role of thermal stratification in *LSWT* response (Piccolroaz et al. 2015), to provide good performance when tested in 14 temperate lakes characterized by different morphologies (Toffolon et al. 2014), and to satisfactorily capture seasonal variability and interannual fluctuations of *LSWT* when using different sources of data (e.g., *LSWT* measured at buoys or retrieved from satellite, as well as *AT* from observations, re-analysis, or GCMs), suggesting high flexibility. This is possible because of the physically based derivation of the model equation, which allows for transfer of information about the study lake from data to

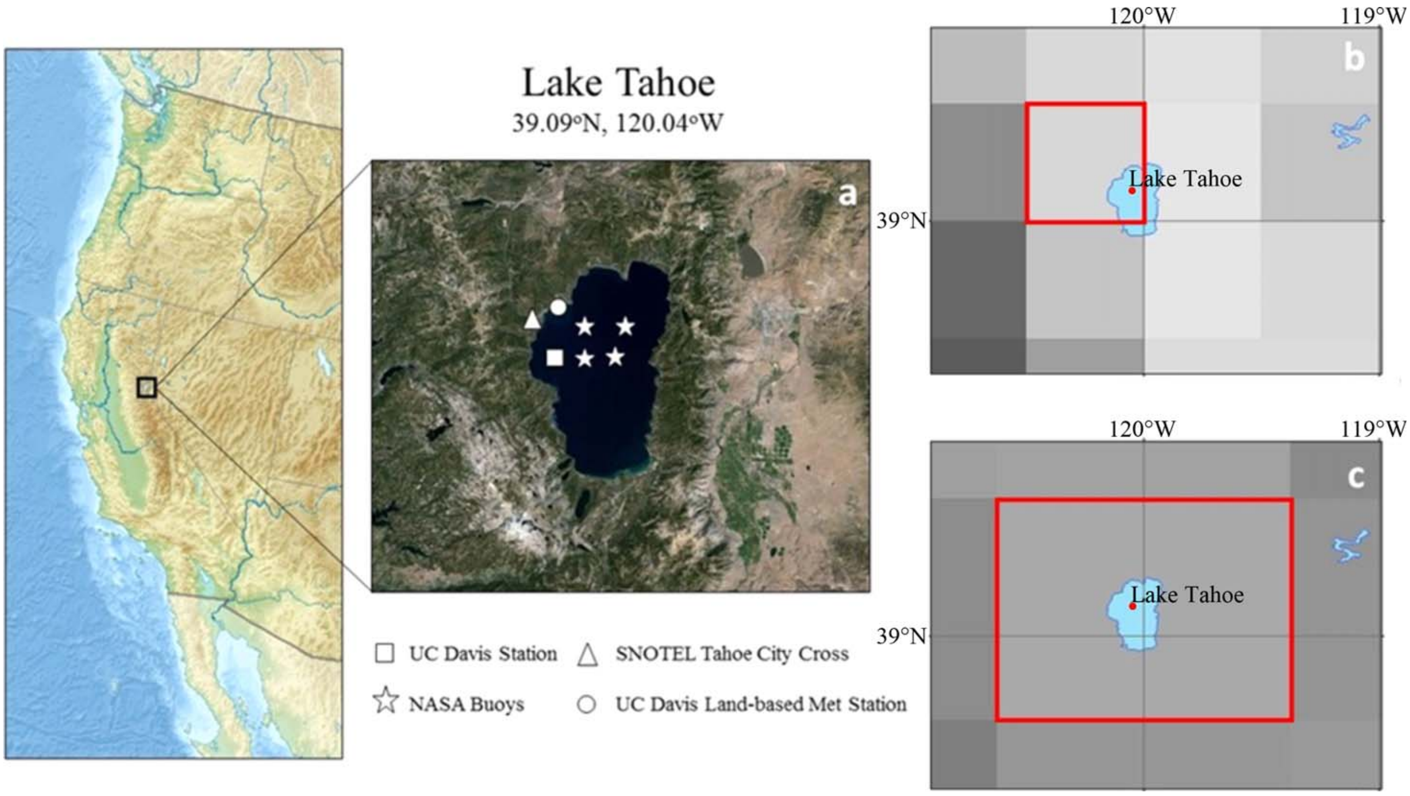

Fig. 1. (**a**) Map of the Lake Tahoe region with locations of onshore weather stations and instrumented buoys; (**b**) example of CRU TS3.21 grid spacing; and (**c**) example of CMIP5-CCSM4 grid spacing. Red boxes outline where data are derived for the gridded datasets. [Color figure can be viewed at wileyonlinelibrary.com]

model parameters through an automatic calibration procedure.

The case study analyzed here is for Lake Tahoe (U.S.A.), for which especially long-time series of *AT* and *LSWT* data are available from in situ, satellite, and model sources. Lake Tahoe is a particularly relevant case study, as it is world famous for its natural beauty, clarity, and surrounding snow-covered mountains, which make it a popular tourist attraction both in winter and summer seasons. Studies aimed at predicting how the water temperature of this lake will evolve in the future are, therefore, of interest beyond the scientific community.

The article is structured as follows: In data and methods section we describe the case study, available data, and *air2-water* model in detail. The main results are presented in results section, which focuses on evaluating the model performance in predicting *LSWT* during historical and future periods using different sources of *AT* data. Finally, discussion of results and concluding remarks are addressed in discussion and conclusion sections, respectively.

Data and methods

Study site characteristics

Lake Tahoe is located at 39.09° N, 120.04° W (approximately 1898 m above mean sea level) in the Sierra Nevada Mountains on the California-Nevada border (U.S.A., *see* Fig.1). The lake does not freeze in winter, and complete overturn/mixing events occur roughly every 3–4 yr (Sahoo et al. 2013; Tahoe Environmental Research Center 2016). Lake Tahoe is 33 km long and 18 km wide, with an average depth of 330 m, a maximum depth of 501 m, and a total volume of 156 km^3. The lake is fed by 63 streams and drained by one outlet, the Truckee River.

Lake Tahoe is characterized by a combination of great depth, small watershed-to-lake area ratio, and a granitic basin that exhibits relatively cold water with low fertility and high transparency/clarity (Jassby et al. 1994). One of Lake Tahoe's most valued characteristics is the water's deep blue hue, which is related to low algal biomass (Tahoe Environmental Research Center 2016). Between 1968 and 1998 the annual average transparency was observed to be decreasing at Lake Tahoe (Jassby et al. 1999), and in that time period some warm-water (exotic) fish species established populations (Reuter and Miller 2000). Understanding how Lake Tahoe's aquatic ecosystem will respond to climate change will depend on the seasonality of physical (stratification, mixing events, etc.) and biological processes (growth and reproduction of species, phenology, trophic interaction, etc.), both of which are sensitive to alterations in water temperature (Schindler et al. 1990; Magnuson et al. 1997; Livingstone and Dokulil 2001; Straile 2000).

Due to the high frequency of cloud-free periods, Lake Tahoe is an ideal water target for satellite calibration and validation. As such, four instrumented buoys were installed on the lake beginning in 1999 (by the NASA Jet Propulsion Laboratory, California Institute of Technology) to provide data for calibration and validation of absolute radiometric reflectance, temperature, and emissivity data collected by instrumentation onboard various satellites including Landsat (5, 7, and 8), the Moderate Resolution Imaging Spectroradiometer (MODIS: Terra and Aqua), Along Track Scanning Radiometers (ATSR: ATSR-1, ATSR-2, and AATSR), the Advanced Spaceborne Thermal Emission and Reflection (ASTER), the Visible Infrared Imaging Radiometer Suite (VIIRS), and the MODIS/ASTER Airborne Simulator (MASTER). Proper calibration/validation of thermal infrared data collected by spacecraft is critical for accurate, global estimation of lake surface temperatures at locations where in situ observations may be lacking.

Available data

This study primarily focuses on two key variables: *AT* and *LSWT*. To comprehensively understand the predictive capabilities of the hybrid model *air2water*, we employed multiple input datasets from different sources. Historical in situ *AT* data (1967–present) were collected at a NOAA land-based weather station on Lake Tahoe's shoreline (1967–2014; daily minimum and maximum measured 3 m above ground; station ID: USC00048758). Data were also collected at a SNOTEL land-based weather station, in the forest at 170 m above lake level (1989–2014; daily average, measured at 5 m above ground; station ID: 809) and at all four NASA buoys (1999–2014; 5-min intervals, measured at 3 m above water). In addition to these point observations, we also utilize the Climatic Research Unit (CRU) TS3.21 data (0.5° grid spacing; monthly average near surface *AT* at 2 m height) (Harris et al. 2014) as a source of historical (1950–2005), gridded *AT* observations. *LSWT* data were collected at the four NASA buoys (1999–present; 5-min intervals of radiometric skin temperature), as well as at an off-shore station maintained by the University of California, Davis (1967–2014; monthly measurements). In the case of data collected at the NASA buoys, we use the average of all four buoys for both *AT* and *LSWT* throughout the analysis, and we refer to this dataset as the "NASA buoy" data.

Modeled estimates of future *AT* (2006–2100) are derived from GCM data. In particular, we use the Coupled Model Intercomparison Project Phase 5-Community Climate System Model version 4.0 (CMIP5-CCSM4) estimates of gridded future *AT* (1.0° grid spacing; monthly average near surface *AT* at 2 m height) (Taylor et al. 2009). We have chosen to use the Representative Concentration Pathways (RCP) of 4.5 and 8.5 for analysis of intermediate and high emission scenarios, respectively. These same datasets also cover the historical period 1950–2005. The CMIP5-CCSM4 data were downscaled to 30-arcsecond (1 km^2) spatial resolution following methodology outlined in Mosier et al. (2014). According to the fifth IPCC assessment report, the CCSM4 model produces global air temperature estimations within 0 to −0.3°C of observations as compared to the other 42 CMIP5 models that range from greater than 0.5 to −0.5 (Flato et al. 2013), and is noted to have stood out as "best" performers in studies specific to the Pacific Northwest U.S.A. (Rupp et al. 2013) and the Southeast U.S.A. (Rupp 2014). CMIP5 model projections of annual air temperature at Lake Tahoe around year 2100 range from 8.1–11.3°C and 10.5–15.4°C, while the CCSM4 values are 9.7°C and 11.7°C for RCP 4.5 and RCP 8.5, respectively, (to be compared to 7.0°C, evaluated as the mean *AT* for the 1950–2005 historical CCSM4 simulation).

Details of each air and water temperature dataset are presented in Table 1, while the location of the measurement stations and the grid spacing of CRU and CMIP5-CCSM4 are shown in Fig. 1. The period chosen for model calibration is 1999–2005, since this is the longest period with overlap of all relevant measurements, including the NASA buoy *LSWT* data (starting in 1999) and CRU *AT* data (available until 2005). CMIP5-CCSM4 model projections begin after 2005. A shorter period (2002–2005) is used in the case of the NASA buoy *AT* data, due to limitations in temporal coverage from that source.

The long-term mean annual cycles (1999–2005) of the different temperature datasets are shown in Fig. 2 on monthly timescales. Since the temporal resolution of *AT* data differs among the datasets, the coarsest resolution being one month, monthly averages of *AT* and *LSWT* have been considered in all analyses for consistency. Figure 2 shows generally good agreement among the five *AT* datasets, with no significant phase lags in the timing of the annual cycle. However, some differences can be noted in the individual monthly averages (roughly 1–3°C), which are primarily attributable to the different locations of the meteorological stations with respect to the lake, and to the approximations inevitably introduced by interpolation (CRU data) and modeling (CMIP5-CCSM4 data). These differences are used to thoroughly test the robustness of the model, with the desired goal being to obtain the same *LSWT* prediction irrespective of the *AT* data used as input forcing.

Description of the *air2water* model

The *air2water* model (Piccolroaz et al. 2013) is a simple lumped model to predict *LSWT* using *AT* as the only external forcing. The model is derived from the volume-integrated equation of heat applied to the upper volume of the lake that is directly linked to heat exchanges with the atmosphere:

Table 1. Details of air and lake surface water temperature data utilized in this study for Lake Tahoe.

Dataset	Abbreviation	Data type	Height/depth and resolution	Geographic coordinates	Time interval	Frequency
Lake surface water temperature (LSWT)						
NASA buoy radiometric data	NASA buoy	in situ	water surface/skin point location	Buoy 1: 39.155° N 120.004° W Buoy 2: 39.109° N 120.011° W Buoy 3: 39. 110° N 120.075° W Buoy 4: 39.155° N 120.071° W	1999–2014	5-min
UC Davis long-term, off-shore monitoring station	Off-shore UC Davis	in situ	water surface/∼1 m depth point location	39.150° N 120.033° W	1969–2014	Monthly measurements
Air temperature (AT)						
NASA buoy meteorological station	NASA buoy	in situ	∼3 m height (above water) point location	See above	2002–2014	5-min
NOAA meteorological station (USC00048758, Tahoe City)	Shoreline (NOAA)	in situ	∼3 m height (above ground) point location	39.167° N 120.143° W	1969–2014	Daily (maxima and minima)
SNOTEL meteorological station (809, Tahoe City Cross)	Forest (SNOTEL)	in situ	∼5 m height (above ground) point location	39,167° N 120.15° W 2072 m a.s.l.	1999–2014	Daily means
Climate Research Group (CRU) TS3.21	CRU	Gridded Observations	Equivalent height ∼2 m 0.5°×0.5° grid cell	39.25° N 120.25° W	1999–2005	Monthly means
CMIP5-CCSM4 (downscaled)	CMIP5-CCSM4	Gridded GCM	Equivalent height ∼2 m 1.0°×1.0° grid cell	39.110° N 120.000° W	1999–2100	Monthly means

$$\rho c_p V_s \frac{dT_w}{dt} = A\Phi_{\mathrm{net}}, \tag{1}$$

where ρ is water density [M L^{-3}], c_p is the specific heat capacity [L^2 T^{-2} Θ^{-1}], V_s [L^3] is the volume of surface water hereafter referred to as reactive volume, T_w [Θ] is *LSWT*, t [T] is time (hereafter expressed in days), A [L^2] is the surface area of the lake, and Φ_{net} [M T^{-3}] is the net heat flux into the upper water volume. Φ_{net} accounts for the main fluxes entering and exiting the reactive volume V_s, and its dependence on *LSWT* and *AT* (assumed to be a proxy for the integrated effects of the external meteorological forcing, e.g., Livingstone and Padisák 2007) is linearized by a Taylor expansion, giving the model the structure of a simple, ordinary differential equation. Versions characterized by different numbers of parameters were identified by Piccolroaz et al. (2013) and slightly redefined by Toffolon et al. (2014). Here we use the 6-parameter version:

$$\frac{dT_w}{dt} = \frac{1}{\delta}\left\{a_1 + a_2 T_a - a_3 T_w + a_5 \cos\left[2\pi\left(\frac{t}{t_y} - a_6\right)\right]\right\}, \tag{2}$$

where a_i (i=1 − 6) are model parameters, t_y [T] is the duration of the year expressed in days, T_a [Θ] is *AT*, and δ [–] is the dimensionless volume (or depth) defined as the ratio between the reactive volume V_s [L^3] and a reference volume V_r [L^3]. The latter is assumed to be the entire volume of the lake in Piccolroaz et al. (2015) and Toffolon et al. (2014). The following relationship for δ, which is assumed to vary with thermal stratification as a function of the temperature difference between surface and deep water ($T_w - T_h$), has been proposed:

$$\begin{cases} \delta = \exp\left(-\dfrac{T_w - T_h}{a_4}\right) & \text{for } T_w \geq T_h \\ \delta = 1 & \text{for } T_w < T_h \end{cases} \tag{3}$$

where T_h [Θ] is the reference value of deep water temperature. Equations 2 and 3 together are solved numerically by using the Crank-Nicolson numerical scheme, which is implicit, second-order accurate, and unconditionally stable. The equation is solved using a daily time step (i.e., dt = 1 day), so that daily *LSWT* is predicted. The second release of the *air2water* model is available at https://github.com/spiccolroaz/air2water, where the source code (written in Fortran 90/95), the precompiled executable files (Linux/Windows), a readme file, and an example application are freely

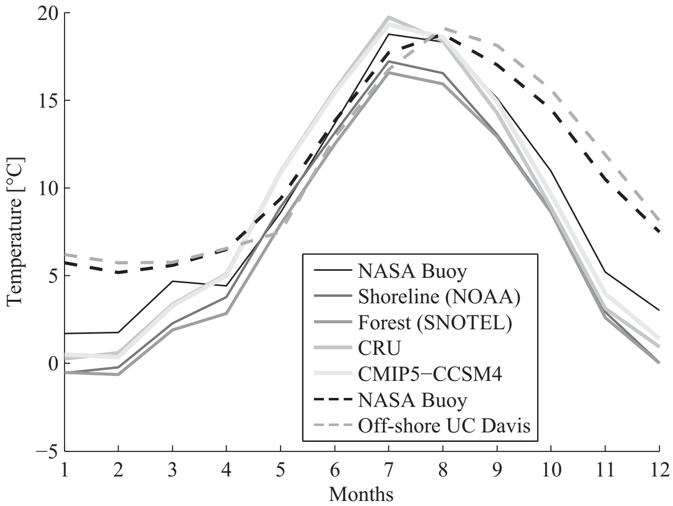

Fig. 2. Long-term monthly mean *AT* (continuous lines) and *LSWT* (dashed lines), averaged across 1999–2005 for the range of datasets used in the analysis (see Table 1).

downloadable (the code is published under the Creative Commons Attribution-ShareAlike 3.0 license).

The *air2water* model can be classified as a hybrid model (Toffolon and Piccolroaz 2015), which combines a physically-based equation with a stochastic calibration of model parameters. In this way, the data directly inform model parameters, whose values can provide insight into the thermal behavior of the lake due to the physically based structure of the governing equation. Model parameters a_1 to a_6 are calibrated using Monte Carlo techniques that exploit an optimization algorithm, with RMSE as the relevant metric. In addition, the following indexes are used for model performance evaluation: Nash-Sutcliffe Efficiency index (NSE) as a normalized metric, the mean error (ME) as a measure of the bias of simulated *LSWT* compared to observations, and the maximum absolute error (MaxAE) as a very stringent index particularly sensitive to outliers (*see* Appendix for details).

In this study, calibration is performed using high-frequency water temperature from NASA buoys aggregated at monthly mean time scales (to be consistent with the resolution of *AT* used in the analysis) for the target *LSWT* time series, resulting in an optimal set of parameters for each of the air station/water station pairs considered. The following analysis is aimed at evaluating the suitability of the model for use in future climate change studies.

Application of the model to Lake Tahoe

Daily *AT* is required by the model, but some of the considered datasets (*see* Table 1) provide only monthly averages. In order to allow for consistent comparisons among the different datasets, all *AT* time series are aggregated to a monthly resolution and then reconstructed to daily resolution (as required by the model) through linear interpolation. An iterative procedure is used, which adjusts the data points while preserving the monthly averages (e.g., Harzallah 1995). We note that linear reconstruction does not provide the most realistic description of *AT* evolution, and more sophisticated, nonlinear interpolation techniques could be used instead. However, the calibration target is monthly mean *LSWT*, so the use of more complex procedures to reconstruct daily *AT* is not necessary. Additionally, the simplest approach is preferred here to test the performance, versatility, and robustness of *air2water* in the most general case.

For each of the considered *AT* datasets, future projections for the period 2006–2100 are determined based on the CMIP5-CCSM4 dataset. This is accomplished by using the "change factor method" or "delta method" (Diaz-Nieto and Wilby 2005; Minville et al. 2008;), which is a downscaling technique that allows for constructing future projections of a climate variable for which observations are available during a historical period (*h*). As such, the change in a climate variable predicted by the climate model is simply added to observed values from the same baseline period, *h*, to obtain the future reconstructed value:

$$T_a^y = \bar{T}_a^h + \left(T_{a,\mathrm{mod}}^y - \bar{T}_{a,\mathrm{mod}}^h \right), \tag{5}$$

where T_a^y and $T_{a,\mathrm{mod}}^y$ are the predicted daily *AT* series in year y for the generic observational dataset and for the climate model, respectively, and $\bar{T}_a^h$ and $\bar{T}_{a,\mathrm{mod}}^h$ are the observed and modeled climatological mean annual cycles (on daily timescales) for the same historical period h, respectively. In this case T_a^y is the projected future *AT* for the generic observational dataset considered in the analysis, y spans the period 2006–2100, the reference historical period h is 1999–2005, and $T_{a,\mathrm{mod}}$ is provided by CMIP5-CCSM4.

The parameters of the *air2water* model are calibrated for different air station/water station pairs. The purpose of applying the model across different *AT* datasets is twofold: (1) to show how different sources of *AT* data affect the performance of the *air2water* model in simulating *LSWT* (*see* model performance during historical periods section), and (2) to examine the robustness of the *air2water* model as a predictive tool for evaluating future projections of *LSWT* (*see* effectiveness of *air2water* as a predictive tool for climate change scenarios section). Additionally, thanks to the existence of a long, historical record of measurements, the ability of the model to capture long-term historical trends and interannual dynamics is also tested (*see* model simulation of interannual variability and long-term trends section).

Results

Model performance during historical periods

The performance of the model in simulating *LSWT* is presented in Fig. 3, which shows the hysteresis curves between

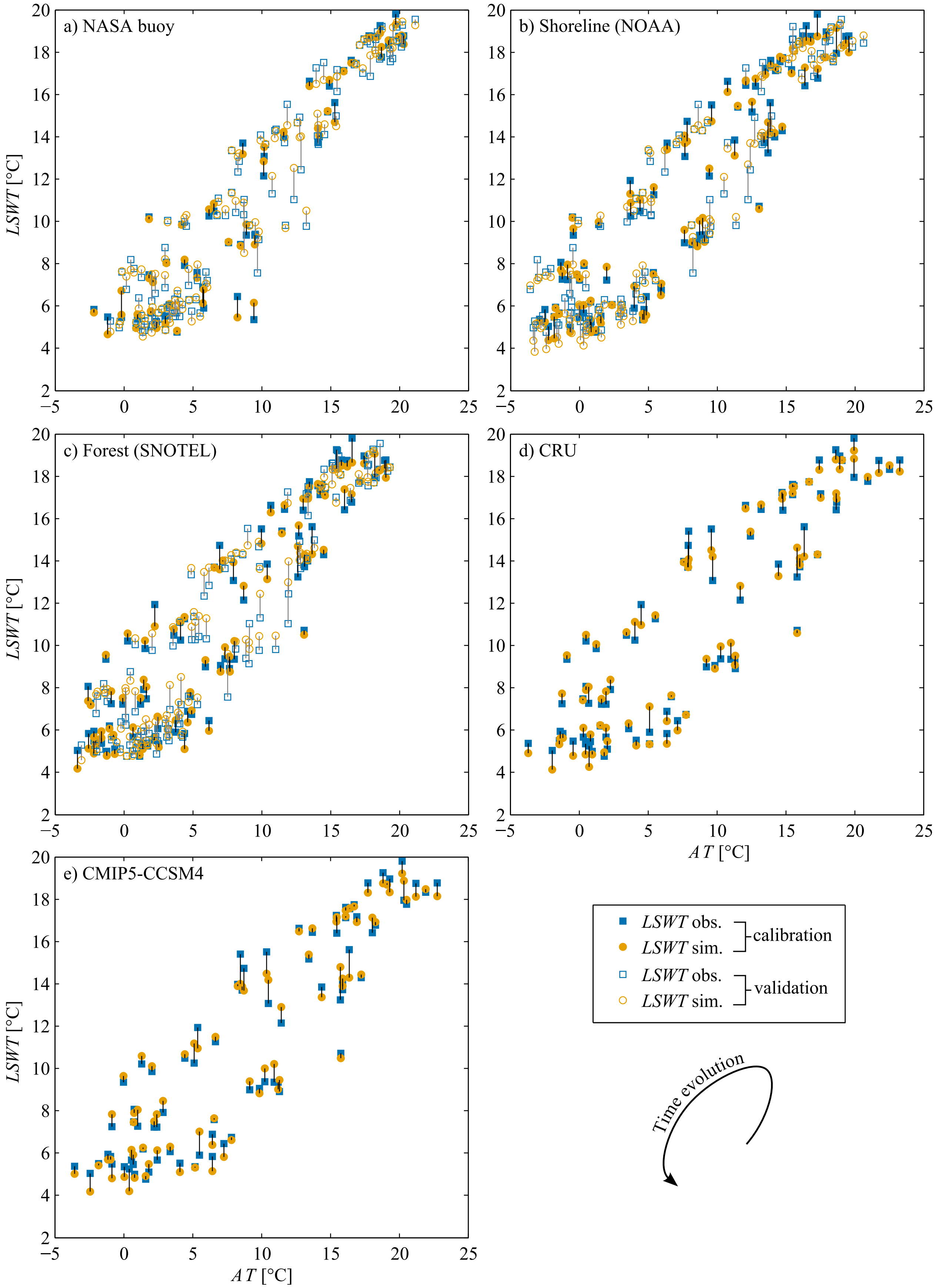

Fig. 3. Hysteresis cycles between observed *AT* and observed (NASA buoy) and simulated (*air2water*) *LSWT*. Observed *AT* belongs to a variety of datasets: (**a**) NASA buoy, (**b**) shoreline (NOAA), (**c**) forest (SNOTEL), (**d**) CRU, and (**e**) CMIP5-CCSM4. Each observed-simulated *LSWT* pair is connected by a segment. [Color figure can be viewed at wileyonlinelibrary.com]

measured *AT* and both measured and simulated *LSWT* (*see* also Fig. S1 in the Supporting Information for the same comparison, but shown as a time series). The different shapes of the hysteresis curves in Fig. 3 further highlight the differences among the various *AT* datasets, which were already discussed above for Fig. 2. Table 2 summarizes the statistics of the model performance and the values of model parameters for different datasets. Note that there is no validation for the last two cases (CRU and historical CMIP5-CCSM4 datasets), because there are no available *AT* data beyond the calibration period (1999–2005). In all cases, the model provides satisfactory results with small values of RMSE, MaxAE, and ME and high values of NSE, both in calibration and validation (Table 2). We emphasize that the calibration/validation exercise in this case study is a fairly demanding test for the model because: (1) only a 7-yr period is used for calibration, and the model is validated over a longer time period (when available), (2) daily *AT* is linearly reconstructed from monthly averages, as described in section Application of the model to Lake Tahoe, and (3) aside from *AT*, no other forcing variables are used to drive the model.

It is important to note that the values of model parameters are different for the different datasets, because they implicitly adjust to differences in the sources of information. Importantly, the predictive performance of the model is also quite similar in the five cases considered (*see* Table 2), with the RMSE between simulated and observed monthly *LSWT* ranging from 0.45°C to 0.58°C during the calibration period. As expected, slightly higher performance is achieved when using *AT* measured at the same location as where *LSWT* is measured (e.g., NASA buoy RMSE = 0.45°C). Conversely, larger errors (albeit still relatively small RMSE = 0.58°C), are obtained when the model is forced with the *AT* time series simulated by the CMIP5-CCSM4 model. In all cases, ME is negligible (i.e., lower than about 1E-3 in absolute value), indicating that there is no bias between simulated and observed *LSWT*. RMSE and ME increase in the validation period, but are still consistently lower than roughly 1°C and 0.1°C, respectively. NSE is always close to 1, and the MaxAE is on the order of 1°C, both in calibration and validation., These values are fully comparable with those of more complex deterministic models (Toffolon et al. 2014).

For comparison, we evaluated the performance of a linear regression model using the different *AT* datasets considered here, and the same calibration and validation periods as above. Model performances are listed in Table S1 in the Supporting Information, and show an evident worsening, with RMSE and MaxAE increasing up to about 2°C and 4°C, respectively, and NSE decreasing to about 0.8. These values suggest the inadequacy of simple linear regression models to properly simulate *LSWT*, confirming previous results by Piccolroaz et al. (2016).

Model simulation of interannual variability and long-term trends

The suitability of the *air2water* model as a simple and reliable predictive tool to capture interannual variability and

Table 2. Model parameters and statistics of model performance at monthly scale (daily for the last case: off-shore UC Davis) obtained in calibration and in validation considering different sources of observed air and water temperature.

								Calibration				Validation			
Water temperature data source	Air temperature data source	a_1 (°C day^{-1})	a_2 (day^{-1})	a_3 (day^{-1})	a_4 (°C)	a_5 (°C day^{-1})	a_6	RMSE (°C)	NSE	ME (°C)	MaxAE (°C)	RMSE (°C)	NSE	ME (°C)	MaxAE (°C)
NASA buoy	NASA buoy	0.171	0.0124	0.0254	5.97	0.081	0.594	0.45	0.99	3.00E-3	1.27	0.59	0.99	0.12	2.52
	Shoreline (NOAA)	0.345	0.0243	0.0470	7.79	0.130	0.670	0.53	0.99	3.25E-6	1.46	0.63	0.98	−0.06	1.70
	Forest (SNOTEL)	0.351	0.0199	0.0441	8.03	0.144	0.639	0.55	0.99	2.05E-5	1.47	0.68	0.98	0.26	2.02
	CRU	0.390	0.0256	0.0550	11.79	0.164	0.672	0.56	0.99	−7.91E-5	1.42	n.a.	n.a.	n.a.	n.a.
	CMIP5-CCSM4	0.286	0.0190	0.0409	9.42	0.117	0.650	0.58	0.99	5.45E-5	1.56	n.a.	n.a.	n.a.	n.a.
Off-shore UC Davis	Shoreline (NOAA)	0.246	0.0196	0.0340	8.24	0.074	0.663	0.82	0.97	−1.30E-5	2.93	0.93	0.97	−0.07	5.10

MaxAE,; maximum absolute error; ME,; mean error (i.e., bias); NSE,; Nash-Sutcliffe Efficiency index; RMSE,; root mean square error.

long-term trends in *LSWT* is examined here using the longest datasets of *LSWT*. Since 1969, the University of California at Davis has measured *LSWT* at an offshore site (*see* Fig. 1) at monthly frequencies. We calibrate the model using this long-term time series of monthly *LSWT* as the target series, but with the model forced by daily *AT* measured at NOAA land-based weather station on Lake Tahoe's shoreline (station ID: USC00048758). Since *LSWT* measurements are only available at monthly resolution (spot measurements), it is not possible to calibrate the model to match the monthly means as done for the previous simulations. Model parameters are therefore calibrated by minimizing the RMSE between simulated and observed *LSWT*, but only on days of the month when measurements are available. Consistent with the previous analysis (*see* model performance during historical periods section), the model is calibrated during the 7-yr period 1999–2005 and validated during 1969–1998 and 2006–2014. Calibrated model parameters and model performance are shown in Table 2. Although NSE and ME values are comparable to those obtained in the previous analysis (both in calibration and validation) a worsening of RMSE and MaxAE is noticeable. This is mainly due to two reasons: (1) errors in this case are not evaluated on monthly means of *LSWT* but on a few, daily values (84 values total during the 7-yr calibration period, i.e., one measurement per month); (2) *AT* and predicted *LSWT* are daily means, while observed *LSWT* corresponds to a series of sporadic samples that may differ from the daily means depending on the time of day when the measurement has been taken. It should also be noted that *air2water* may be limited when dealing with changes in water transparency (as has been the case for Lake Tahoe during the last 50 yr, Jassby et al. 1999; Tahoe Environmental Research Center 2016), which likely modify the parameter governing the stratification dynamics (i.e., δ). In addition, the effect of other factors such as changes in solar brightening or wind forcing during the validation period (compared to the calibration period) cannot be directly captured by the simple structure of the model (except, indirectly, through their potential effect on *AT*), despite the fact that they have been shown to be important in some lakes (Schmid and Köster 2016; Woolway et al. 2017). These considerations undoubtedly challenge the application of the model. However, the performance of the model is still satisfactory (RMSE lower than 1°C, both in calibration and validation), confirming that the main thermal dynamics are well captured and that the model can be reasonably used to predict *LSWT* in periods other than the one used for calibration.

Results presented in Fig. 4 show interannual to interdecadal variability, as well as long-term trends in seasonal-mean *LSWT* for the seasons of January-February-March (JFM, approximately corresponding to northern hemisphere winter), April-May-June (AMJ, spring), July-August-September (JAS, summer), and October-November-December (OND, autumn). Given the monthly resolution of the measurements, seasonal means for observed *LSWT* are evaluated only if at least three values are available (i.e., one spot measurement per month in each month of the season). The corresponding seasonal means of simulated *LSWT* are evaluated by only averaging over the days when observed *LSWT* is available. Therefore, the resulting averaged values of *LSWT* should not be considered as true "seasonal means" in a strict sense. However, they suffice in providing an indication of interannual and long-term variability, which is what the model is being tested to reproduce.

Figure 4 shows that both interannual variability and long-term trends in seasonal-mean *LSWT* are generally well captured, although some discrepancies can be noted. Examples of discrepancies include (1) differences between observed and predicted AMJ long-term trends, and (2) generally stronger interdecadal variability in the observed LSWT time series, compared to *air2water* model simulations. On the other hand, although the model is driven by *AT* alone, it is able to effectively reproduce the long-term trend of *LSWT* even when it is substantially different from or opposite in sign to that of *AT* (e.g., AMJ). This is made possible by the fact that *air2water* includes all the major physical processes driving *LSWT* dynamics (albeit in an implicit, simplified form), most importantly the role of thermal stratification.

The use of long-term linear trends may obscure the identification and interpretation of interannual fluctuations, anomalous interdecadal warming/cooling periods, and regime shifts (North et al. 2013; Van Cleave et al. 2014; Woolway et al., 2017). This suggests that linear regressions should be used with caution to avoid the risk of oversimplifying the true temperature dynamics. For this reason, the 10-yr running means of observed *AT*, observed *LSWT*, and modeled *LSWT* are also shown in Fig. 4, which provides evidence of the existence of significant interannual fluctuations and decadal-scale variability. The figure shows that *air2water* is able to capture most of these interdecadal dynamics, which are substantially different in the four seasons. Importantly, RMSE is relatively constant among the different seasons and on the order of roughly 0.5°C, suggesting the absence of any significant seasonal biases in *LSWT* modeling.

For comparison, Fig. 4 also shows the results of a linear regression model, calibrated and validated with the same data and over the same periods as for the *air2water* model. This model clearly fails to describe *LSWT* dynamics, as indicated by the significantly poor performance summarized in Supporting Information Table S1 (i.e., RMSE of more than 2°C, NSE around 0.7, large biases, and MaxAE of about 6°C) and noticeably illustrated in Fig. 4. By definition, the linear regression model is not able to reproduce the *AT-LSWT* hysteresis cycle, providing substantial overestimates and underestimates of *LSWT* during the warming (JFM and AMJ) and cooling (JAS and OND) periods of *AT*, respectively. In addition, long-term trends and interannual fluctuations of *LSWT*

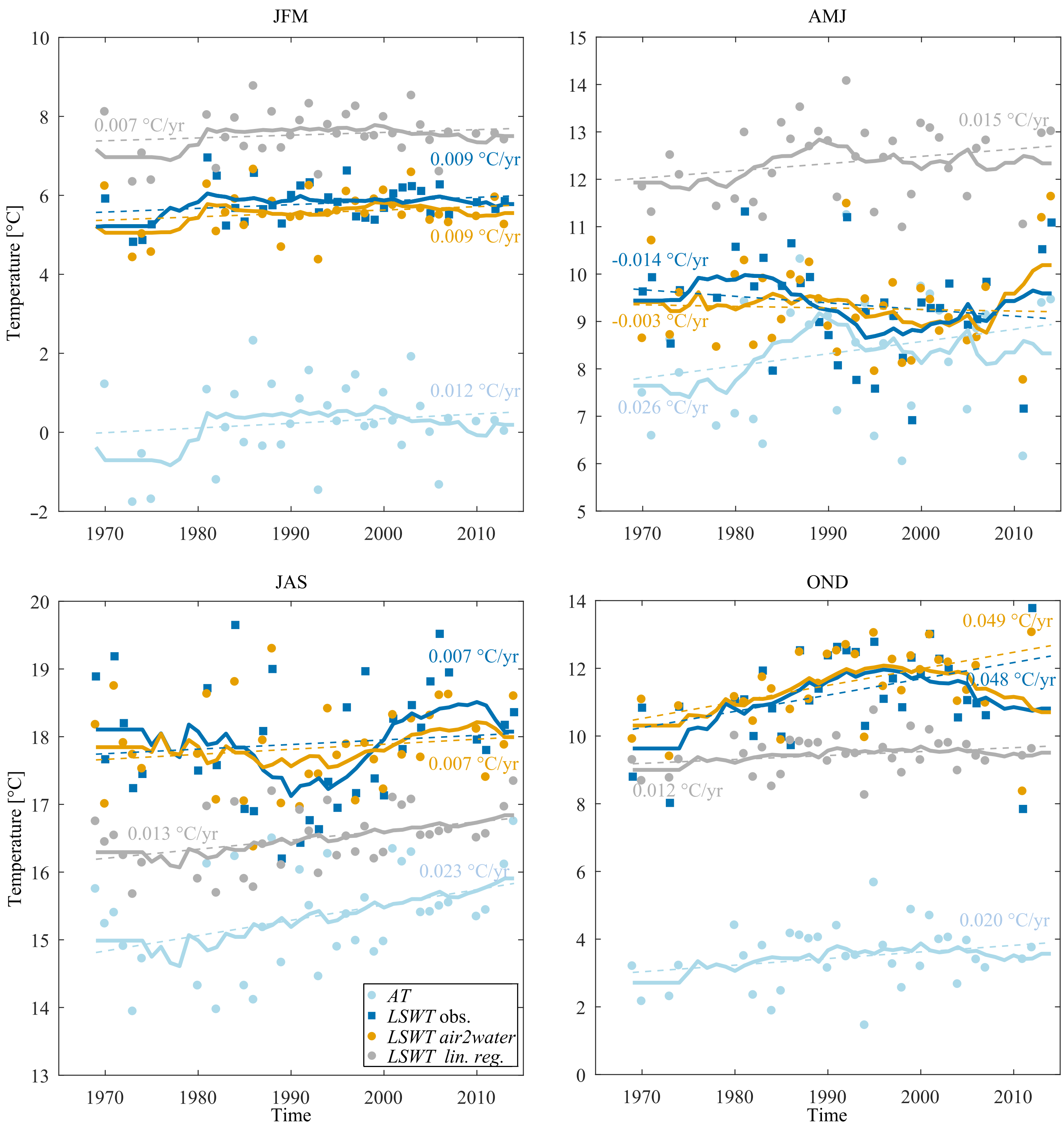

Fig. 4. Comparison of mean JFM, AMJ, JAS, and OND *AT* and *LSWT* (observed at the off-shore station maintained by the University of California, Davis and simulated by the *air2water* model and by a linear regression model) during the period 1969–2014. Also shown are the 10-yr moving averages (solid lines) and long-term trends (linear regression; dashed lines). [Color figure can be viewed at wileyonlinelibrary.com]

inevitably follow those of *AT*. Although the undeniable inadequacy of the linear regression model is probably obvious to many, we believe that it is important to explicitly show the comparison here. Deep flaws of such a simple but widely used approach, compared to the performance of a similarly simple but not oversimplified model (the *air2water* model),

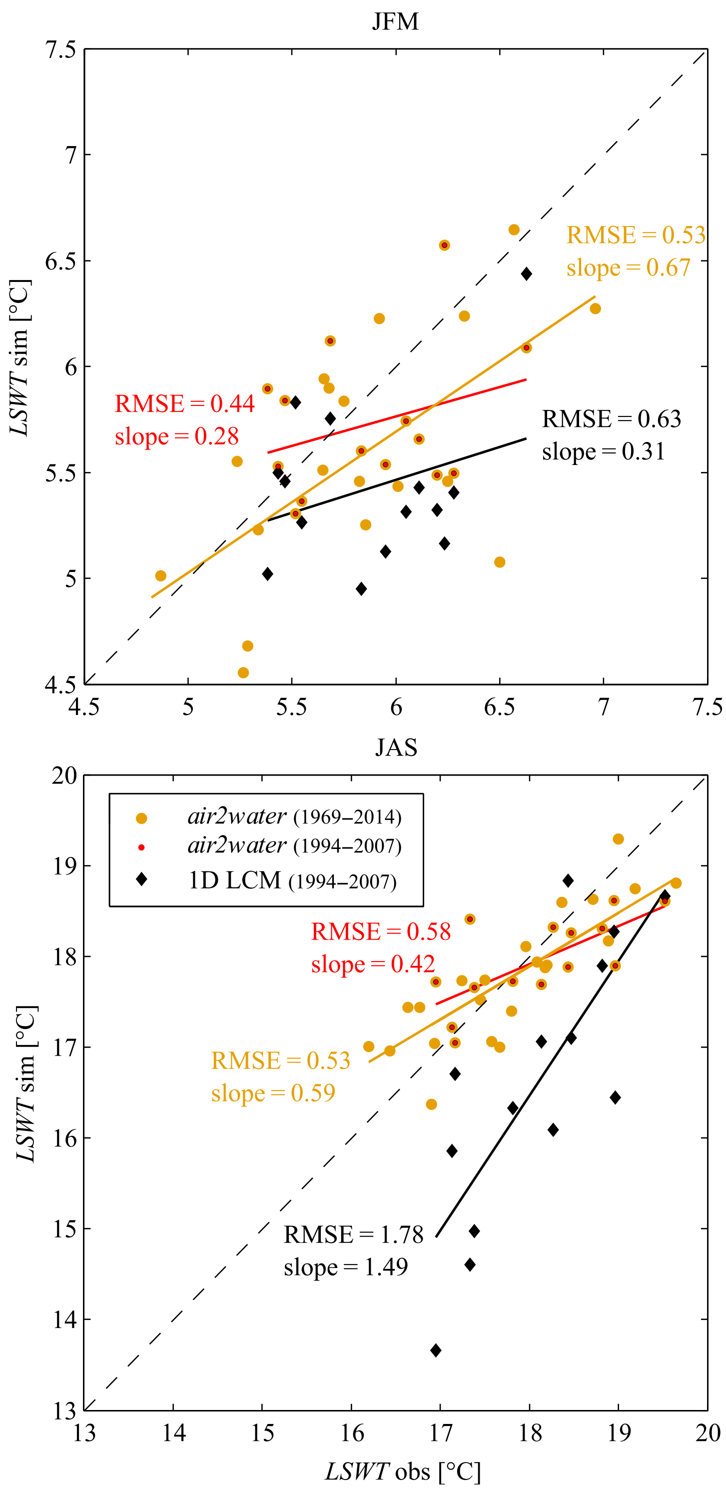

Fig. 5. Scatterplot between observed *LSWT* (off-shore UC Davis) and *LSWT* simulated using the *air2water* and the 1D Lake Clarity Model (1D LCM Sahoo et al. 2013) in the different seasons. Continuous lines identify linear regressions. Dashed lines identify perfect agreement (1 : 1 line). [Color figure can be viewed at wileyonlinelibrary.com]

should definitely discourage the use of purely regression-based approaches and support the use of more robust, physically based models.

The *air2water* model is also compared with the process-based one-dimensional hydrodynamic Lake Clarity Model (1D-LCM, black dots, Sahoo et al. 2013) in Fig. 5, which

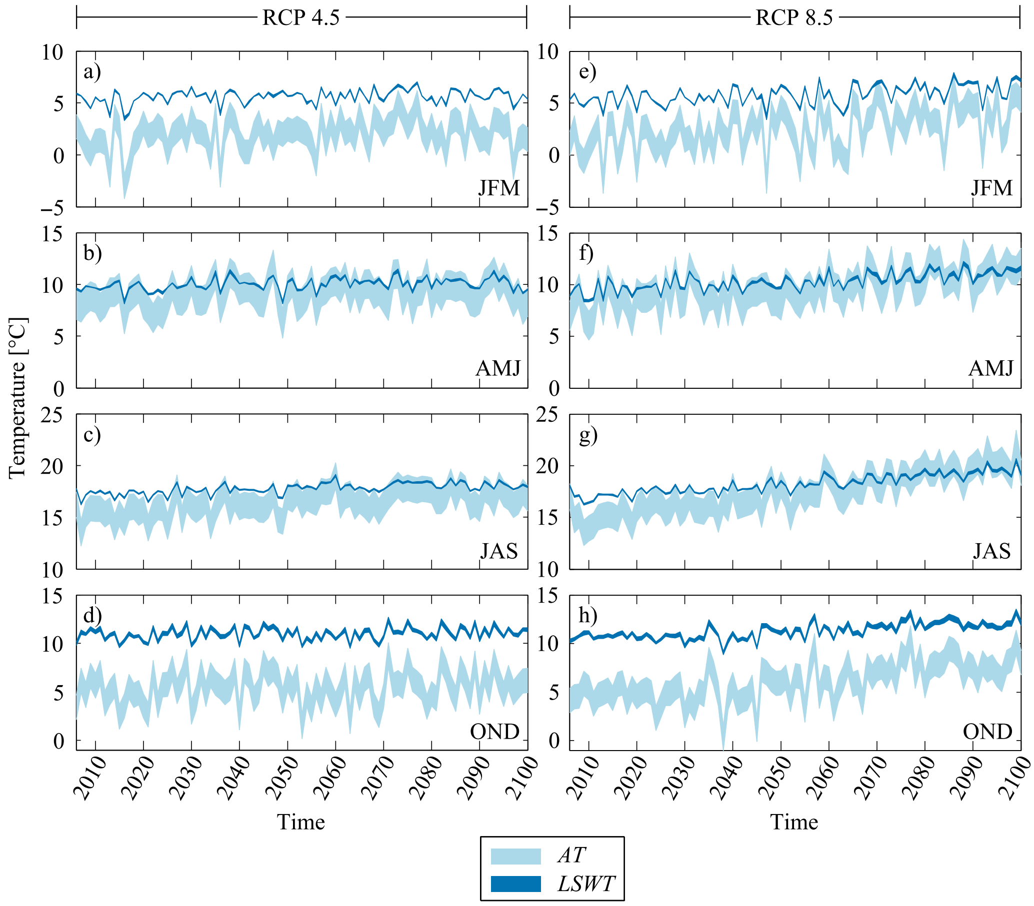

Fig. 6. Projected seasonal averages of *AT* and *LSWT* for the period 2006–2100 under the scenarios RCP 4.5 (**a–d**) and RCP 8.5 (**e–h**). Thickness of the curves represents the interval of variability of *AT* and *LSWT* corresponding to the different air temperature datasets: NASA buoy, shoreline (NOAA), forest (SNOTEL), CRU, and CMIP5-CCSM4. [Color figure can be viewed at wileyonlinelibrary.com]

shows the scatterplot between observed and predicted *LSWT* in the two cases. The 1D-LCM model solves the heat and hydrologic budgets of the lake and requires that all incoming and outgoing quantities be estimated with sufficient accuracy. Since this model has been successfully applied to Lake Tahoe (*see* e.g., Sahoo et al. 2010; Sahoo et al. 2013), it has been identified as particularly relevant and appropriate for a comparison in this study. To allow for a fair comparison, *air2water* results over the same simulation period are shown with different color and symbols. The results of *air2water* generally show less scatter and better alignment along the 1:1 line than 1D-LCM, despite the fact that the latter is relatively more complex and requires additional input variables (e.g., precipitation, shortwave radiation, wind speed, longwave radiation, and vapor pressure). However, we should note that 1D-LCM is calibrated not just for *LSWT* but also to reproduce the entire temperature profile, thermocline depth, water budget, water clarity, etc. In both cases, JFM seems to be the most difficult season for the two models to predict *LSWT*.

The results presented in Figs 4, 5 provide objective evidence of the predictive ability of the *air2water* model and of its potential to be used to analyze the long-term response of *LSWT* to evolving external conditions using only *AT* as input information.

Effectiveness of *air2water* as a predictive tool for climate change scenarios

In this section, we extend the previous historical analysis by testing the suitability of *air2water* to be used for predicting future *LSWT*. To this end, we run the model using climate change scenarios (RCP4.5 and RCP 8.5) defined for each *AT* dataset in Table 1 by means of the change factor method described in section Application of the model to Lake Tahoe. The model is run using different sets of calibration parameters (*see* model performance during historical periods section and Table 2) for the different *AT* datasets. Since model parameters are always calibrated assuming the NASA buoy series of *LSWT* as the target (*see* Table 2), the expectation is that *LSWT* prediction under future climate scenarios is not dependent on the *AT* dataset used to force the model. This result is far from trivial and, if achieved, would provide confidence in the suitability of *air2water* to

Table 3. Average width of the interval of variability of *AT* (*wAT*) and *LSWT* (*wLSWT*) projected for the period 2006–2100 under the two climate change scenarios (RCP 4.5 and RCP 8.5) and shown in Fig. 6 (*see* also Supporting information Fig. S3 for the same analysis but considering annual averages of *AT* and *LSWT*). *wAT* and *wLSWT* are evaluated for the four seasons and for the whole year. Linear trends of *AT* (*tAT*) and *LSWT* (*tLSWT*) are also reported.

Scenario	Period	*wAT* (°C)	*wLSWT* (°C)	*tAT* (°C decade $^{-1}$)	*tLSWT* (°C decade $^{-1}$)
RCP 4.5	JFM	2.30	0.14	0.10	0.04
	AMJ	2.79	0.17	0.12	0.07
	JAS	2.46	0.13	0.23	0.11
	OND	2.52	0.40	0.09	0.04
	Year	1.97	0.07	0.14	0.07
RCP 8.5	JFM	2.30	0.18	0.35	0.16
	AMJ	2.79	0.22	0.39	0.21
	JAS	2.46	0.17	0.62	0.29
	OND	2.52	0.42	0.42	0.19
	Year	1.97	0.10	0.45	0.21

analyze scenarios where the external forcing (i.e., *AT*) goes beyond the range of variability typical of historical/current conditions.

Figure 6 presents future projections of *AT* and *LSWT* for each of the four seasons and for the two climate change scenarios (RCP 4.5 and RCP 8.5). (Annual mean results are shown in Supporting Information Fig. S4.) In all cases, future projections of *LSWT* show a narrow interval of variability, regardless of the *AT* dataset used. In fact, the interval of variability (~ 0.1°C) is much narrower than that of the five *AT* datasets used to force the model (~ 2.5°C; *see* Table 3). A slightly larger uncertainty is obtained when predicting *LSWT* under the RCP 8.5 scenario compared to the RCP 4.5 (*see* Table 3), but the interval of variability is still on the order of just a few tenths of a degree Celsius. This result can only be attained because *air2water* is able to provide a physically reasonable description of the main processes involved and succeeds in reproducing the actual thermal behavior of the system.

Conversely, purely regression-based models (linear or nonlinear) that are not based on physical phenomena are likely to produce relatively large uncertainty bands in the predicted *LSWT* under future conditions. For example, projections of *LSWT* obtained from the same linear regression models discussed in model performance during historical periods section show a much wider interval of variability, especially in AMJ and JAS, despite the fact that the model is always forced to match the same, observed *LSWT* dataset during the calibration process. In addition, the future long-term trends of *LSWT* are significantly different than those obtained with *air2water*, consistent with the results discussed in model simulation of interannual variability and long-term trends section and shown in Fig. 4. Results are summarized in Supporting Information Figs. S5, S6, as well as in Table S2 in the Supporting Information, indicating that a linear regression model cannot provide a reliable description of the future evolution of *LSWT* in Lake Tahoe. We refer the reader to Piccolroaz et al. (2016) for a similar exercise focused on river water temperature modeling, where a cross-validation exercise was employed to compare the performance of standard regression models with those of *air2stream*, a hybrid model similar to *air2water*.

We note that the width of the interval of variability of *AT* is the same for the two climate scenarios (RCP 4.5 and 8.5) since, by definition, the change factor method used to construct the future scenarios (*see* section 2.5) preserves the variance of the five historical datasets. The lowest uncertainty (narrow interval of variability) in the prediction of *LSWT* is achieved during the JAS season. One possible explanation for this is that the lake is strongly stratified during JAS (Tahoe Environmental Research Center 2016), reducing the lake's effective thermal inertia such that *LSWT* responds quickly to external forcing (Toffolon et al. 2014; Piccolroaz et al. 2015). Overall, this reduces the complexity of the thermal dynamics of the lake, making the modeling of *LSWT* easier, to the extent that some studies opt to use the equilibrium *LSWT* (i.e., the temperature at which the net heat flux at the lake surface is zero) as a good alternative for predicting the actual *LSWT* (*see*, e.g., Schmid et al. 2014). Conversely, the largest uncertainty in the width of the confidence band of *LSWT* is in the autumn period (OND), when it is nearly three and two times larger than that of the other seasons for the RCP 4.5 and RCP 8.5 scenarios, respectively. Part of the reason for this is that the lake rapidly moves from strongly stratified to non-stratified conditions in autumn (Tahoe Environmental Research Center 2016), introducing significant nonlinearities to its thermal behavior and making *LSWT* modeling very sensitive to the external forcing. Similar considerations, though to a lesser extent, are also valid when the lake moves from weak to strong stratification during the AMJ period.

Discussion

The results presented here can be used to gain insight into the expected thermal behavior of Lake Tahoe in the future. In response to an annual *AT* warming rate of 0.14 and 0.45°C decade^{-1} during the 21st century for the RCP 4.5 and RCP 8.5 emission scenarios, respectively, the *air2water* model predicts mean annual *LSWT* to increase by 0.07 and 0.21°C decade^{-1} (i.e., roughly half the rate of *AT* warming). JFM *LSWT* is expected to increase the least (0.04 and 0.16°C per decade^{-1} for the RCP 4.5 and RCP 8.5 emission scenarios, respectively), while JAS *LSWT* will increase the most (0.11 and 0.29°C decade^{-1} for the RCP 4.5 and RCP 8.5 emission scenarios, respectively). JAS *AT* is expected to increase at a rate of 0.23 and 0.62°C decade^{-1} for the RCP 4.5 and RCP 8.5 emission scenarios, respectively (Table 3; Fig. 6). Interestingly, JAS shows *AT* warming more rapidly than *LSWT*, during both the historical period (1967–2014, *see* Fig. 4) and the future period (2005–2100, *see* Fig. 6; Table 3).

These results are in contrast to recent findings by O'Reilly et al. (2015) that performed a global analysis of in situ and satellite-derived LSWT data, observing that Lake Tahoe LSWT warmed at a much more rapid rate of 0.54 to 0.71°C decade^{-1} in summer. Schneider et al. (2009) found an even larger rate of summer LSWT warming for Lake Tahoe of 1.3°C decade^{-1}, which is approximately twice as fast as the rate of AT warming. However, they considered nighttime *LSWT* and minimum daily *AT*, and, more importantly, analyzed a different (and shorter) time period from 1991 to 2008. This period was indeed characterized by a rapid warming of *LSWT* that was much stronger than the concurrent warming of *AT*. However, this was an exceptional warming episode not representative of the longer time period, as is clearly shown in Fig. 4.

Overall, these varying results lead to the conclusion (initially discussed in model simulation of interannual variability and long-term trends section) that the use of linear trends to describe temperature dynamics must always be contextualized and analyzed with critical awareness to avoid misinterpretation. Linear trends evaluated over a historical period may not necessarily be generalized to a different period, and they do not necessarily provide a comprehensive overview of the actual dynamics. For example, linear fitting is generally not able to detect the occurrence of interannual fluctuations or regime shifts, and is likely to provide substantially different trends depending on the period considered in the analysis. While these considerations may seem trivial, it is critical to bear this point in mind, especially when comparing results from different studies. In this sense, one should be cautious when stating, for example, that a lake is "warming faster than air temperature," as such a statement may be an overgeneralization based on a relatively limited time period of study.

As a consequence of the larger long-term warming trend of *AT* relative to *LSWT* during JAS (in the current study), it is interesting to note that *air2water* predicts a progressive convergence of *LSWT* towards *AT*, which is more pronounced in the RCP 8.5 scenario (*see* Fig. 6). The two temperature time series tend to converge starting in the 2050s, and eventually *AT* is expected to become coincident and even exceed LSWT on a regular basis (in the RCP 8.5 scenario) by the end of the 21st century. This behavior follows from the fact that summer *AT* warming is rapid compared to the other seasons, while summer *LSWT* warming is restricted by an increase in evaporative cooling (Lenters et al. 2005)—a feedback that eventually leads to compensation via a downward flux of sensible heat (i.e., air warmer than water; Fig. 6). In addition, the equilibration of *LSWT* to *AT* is relatively easy in summertime due to strong stratification (i.e., low thermal inertia of the upper mixed layer; Toffolon et al. 2014; Piccolroaz et al. 2015), the duration of which gets progressively longer as *LSWT* increases and the onset of thermal stratification occurs earlier (as a consequence of a warming *AT*).

The main limitation of the *air2water* model, embodied in its structure, is that it does not explicitly account for the effects of external forcing other than *AT* (e.g., wind speed, solar radiation, and air humidity). Rather, such effects are only indirectly included through their potential influence on *AT*. Changes in water clarity, which may affect the surface mixed layer depth, are not considered in the model. As a result, the values of model parameters inevitably depend on the lake and climate conditions during the calibration period and in principle may (slightly) change if considering a different calibration period. All these aspects necessarily introduce some uncertainties, which can be discussed via reference to Fig. 4, together with the main strengths and limitations of the *air2water* model. Despite the short calibration period (7 yr, from 1999 to 2005) and the fact that the model is required to reproduce sparse spot measurements sampled at monthly frequency, interannual variability and long-term trends in *LSWT* are well captured. This is also true during the validation period, when water clarity was significantly higher relative to the calibration period (Tahoe Environmental Research Center 2016). This is a direct confirmation that the model is able to explain much of the *LSWT* variability using *AT* alone, including cases when lake and climate conditions differ from those in the calibration period. The strong warming of *LSWT* that occurred in the 1990s is also captured, but with an evident underestimate of the intensity of warming. This indicates that other factors besides the mere warming of *AT* may have contributed to this intense *LSWT* warming period, whose effect cannot be simulated by the model. Likely changes include solar brightening (e.g., changes in cloud cover) and the consistent degradation of water clarity during summer (Tahoe Environmental Research Center 2016).

However, one may still wonder whether model performance may have been different, possibly lower, if we had considered a different calibration period. In order to dispel any misgiving about this point, the same historical period is simulated considering four different 12-yr calibration windows (1969–1980, 1981–1992, 1993–2004, and the shorter 2005–2014). The four simulations are fully comparable with each other, the maximum difference of RMSEs evaluated over the whole 46-yr period being very small and equal to 0.05°C, and the mean RMSE being equal to 0.92°C, coherent with the values in Table 2. A Fig. 4, but for the four calibration windows, is shown in the Supporting Information (Fig. S7). The figure shows some bias between the four model runs, which however is smaller (about one half) than the RMSE between simulated and observed *LSWT*. More importantly, besides this bias, the shape and slope of the curves are essentially coincident, suggesting that the simulation of *LSWT* dynamics is basically independent of the calibration period and thus supporting the robustness of the model.

A natural extension of the model, which could partially overcome some of the aforementioned limitations, would be coupling with atmospheric circulation or weather forecasting models. This would allow one to directly simulate fundamental lake-climate interactions, thus improving the simulation of both lake processes and regional climate. Recent attempts in this direction have been made to adopt one-dimensional lake models (e.g., Goyette and Perroud, 2012; Zhong et al. 2016; Sugiyama et al. 2017), which however require an entire set of meteorological variables, causing a potentially larger propagation of uncertainties and errors typical of climate models. An example is given by the comparison between *air2water* (Piccolroaz et al. 2015) and a coupled RCM-lake model (Zhong et al. 2016) for the case of Lake Superior (U.S.A.-Canada), where the latter model is characterized by larger errors, with many cases being statistical significant.

A more in-depth quantification of climate change effects on *LSWT* would require a detailed analysis on at least monthly time scales in order to properly describe the complex system of feedbacks among *AT*, thermal stratification, and *LSWT*. Furthermore, an analysis of additional climate scenarios and the use of an ensemble of GCMs/RCMs and lake models (e.g., lakeMIP project, Thiery et al. 2014) would be recommended. This would allow for evaluating the propagation of uncertainty stemming from different levels of the modeling chain, coherent with the concept of the "cascade of uncertainties" introduced by Wilby and Dessai (2010). Although this goes beyond the scope of the present study, these new results provide an important first step toward such an approach. Having demonstrated that *air2water* is suitable for predicting the response of *LSWT* to climate change, future work will focus on the application of this tool for climate change predictions in other lakes where interesting dynamics have been observed (e.g., Laurentian Great Lakes, Western lakes in the U.S.A., North European Lakes, Alpine Lakes, etc.; *see* e.g., O'Reilly et al. 2015).

Conclusions

The main objective of the paper is to test the suitability of a simple, but mechanistically based model (*air2water*) as a predictive tool for studying the impacts of climate change on lake temperature and related thermal structure. As shown by the results in results section, the predictive ability of the model is tested (explicitly) for Lake Tahoe by comparing simulated and observed long-term surface water temperature during a historical period. The consistency of the model's ability to analyze future scenarios of climate change is also validated by analyzing results obtained from different sources of input *AT* data. In the first case, our results show that *air2water* is able to accurately predict long-term trends of *LSWT*, and reasonably captures interannual and interdecadal fluctuations, even when seasonal trends and variability differ substantially from one season to another. In the latter case, our results indicate that the model is able to provide coherent predictions of *LSWT* independent of the *AT* dataset used to force the model, including when applied to *AT* ranges beyond the limits of the series used for model calibration. Although this is a desirable result, it is not entirely expected, and is an indirect confirmation that *air2water* provides a physically realistic description of the lake's thermal dynamics, preventing the propagation and amplification of errors when used as a prognostic tool.

The results of this study indicate a warming of summer (JAS) *LSWT* of up to 1.1°C (RCP 4.5) and 2.9°C (RCP 8.5) by the end of the 21st century, with summer being the season that is expected to warm the most (for both *AT* and *LSWT*). In contrast, *LSWT* in winter is expected to experience the least amount of warming. This follows from the fact that *AT* will undergo a milder warming in winter, but it is also due to the higher thermal inertia of the lake during this season (the water column is well mixed), which causes winter *LSWT* to respond slowly to changes in winter *AT*, thereby propagating some of the thermal response into subsequent seasons.

We conclude by returning to the question "can we predict long-term trends of lake surface temperature based on *AT* alone?" The answer to this question is inherently dependent on the modeling tool used for the analysis, and one can expect reasonable predictions of *LSWT* from *AT* only when using a tool that is able to encapsulate most of the fundamental physical processes controlling the thermal dynamics of a lake. In this regard, we posit that the *air2water* model certainly represents an attractive option. We also propose, however, that a hierarchy of additional, more complex, and physically based mechanistic models be used to fully evaluate the capabilities (and weaknesses) of simpler, hybrid models such as *air2water*.

Appendix
Model calibration and performance evaluation

Model calibration is based on Monte Carlo simulation techniques in which a large number of parameter combinations is evaluated in terms of a given metric of model performance. In this work, the evolutionary and self-adaptive Particle Swarm Optimization algorithm (Kennedy and Eberhart 1995) is used as a stochastic optimization technique for automatic model calibration. According to the physically based derivation of Eq. 3, reasonable a priori ranges of variation of model parameters can be defined (*see* Piccolroaz et al. 2013 and Piccolroaz 2016 for details about the derivation of the model, the definition of model parameters, and the Particle Swarm Optimization algorithm).

Automatic model calibration is performed using the Root Mean Square error (RMSE) between simulated (T_w) and observed ($\hat{T}_w$) *LSWT* as optimization metric:

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(T_{w,i}-\hat{T}_{w,i}\right)^2}, \tag{4}$$

where n is the length of the observational time series. In this way, the best set of parameters is identified as the one that minimizes RMSE. Model performance is also tested using other statistical indices:

The Nash Sutcliffe Efficiency index (NSE, Nash and Sutcliffe 1970)

$$\mathrm{NSE}=1-\frac{RMSE^2}{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{T}_{w,i}-\bar{T}_w\right)^2}, \tag{5}$$

where $\bar{T}_w$ is the mean of observed *LSWT*. *NSE* provides a normalized measure (-∞,1) of model performance by evaluating the relative magnitude of the residual variance compared to the variance of observations. In particular NSE=1 indicates perfect model fitting, while NSE=0 indicates that the model performs as good as assuming the overall mean of observed *LSWT* as a predictor.

The Maximum Absolute Error (MaxAE)

$$\mathrm{MaxAE}=\max|T_{w,i}-\hat{T}_{w,i}| \tag{6}$$

where the vertical bars | | denote the absolute value operator. This metric is extremely stringent as it is sensitive to the presence of outliers.

The Mean Error or bias (ME)

$$\mathrm{ME}=\frac{1}{n}\sum_{i=1}^{n}\left(T_{w,i}-\hat{T}_{w,i}\right), \tag{7}$$

which can be used to estimate whether the model produces under- or over-estimation.

References

Austin, J. A., and S. M. Colman. 2007. Lake Superior summer water temperatures are increasing more rapidly than air temperatures: A positive ice-albedo feedback. Geophys. Res. Lett. **34**: L06604. doi:10.1029/2006GL029021

Blenckner, T., A. Omstedt, and M. Rummukainen. 2002. A Swedish case study of contemporary and possible future consequences of climate change on lake function. Aquat. Sci. **64**: 171–184. doi:1015-1621/02/020171-14

Butcher, J. B., D. Nover, T. E. Johnson, and C. M. Clark. 2015. Sensitivity of lake thermal and mixing dynamics to climate change. Clim. Change **129**: 295–305. doi: 10.1007/s10584-015-1326-1

Coats, R., J. Perez-Losada, S. G. Schladow, R. C. Richards, and C. R. Goldman. 2006. The warming of Lake Tahoe. Clim. Change **76**: 121–148. doi:10.1007/s10584-005-9006-1

De Senerpont, D., L. N., and others. 2013. Plankton dynamics under different climatic conditions in space and time. Freshw. Biol. **58**: 463–482. doi:10.1111/fwb.12053

Dettinger, M. D. 2013. Projections and downscaling of 21st century temperatures, precipitation, radiative fluxes and winds for the Southwestern US, with focus on Lake Tahoe. Clim. Change **116**: 17–33. doi:10.1007/s10584-012-0501-x

Diaz-Nieto, J., and R. L. Wilby. 2005. A comparison of statistical downscaling and climate change factor methods - impacts on low flows in the River Thames, United Kingdom. Clim. Change **69**: 245–268. doi:10.1007/s10584-005-1157-6

Dokulil, M. T. 2014. Impact of climate warming on European inland waters. Inland Waters **40**: 27–40. doi: 10.5268/IW-4.1.705

Flaim, G., U. Obertegger, A. Anesi, and G. Guella. 2014. Temperature induced changes in lipid biomarkers and mycosporine-like amino acids in the psychrophilic dinoflagellate *Peridinium aciculiferum*. Freshw. Biol. **59**: 985–997. doi:10.1111/fwb.12321

Flato, G. J., and others. 2013. Evaluation of climate models, p. 741–866. *In* T. F. Stocker and others [eds.], Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. Cambridge Univ. Press.

Gallina, N., O. Anneville, and M. Beniston. 2011. Impacts of extreme air temperatures on cyanobacteria in five deep peri-Alpine lakes. J. Limnol. **70**: 186–196. doi:10.4081/jlimnol.2011.186 doi:10.4081/jlimnol.2011.186

Gallina, N., N. Salmaso, G. Morabito, and M. Beniston. 2013. Phytoplankton configuration in six deep lakes in the peri-Alpine region: Are the key drivers related to eutrophication and climate?. Aquat. Ecol. **47**: 177–193. doi:10.1007/s10452-013-9433-4

Gleckler, P. J., K. E. Taylor, and C. Doutriaux. 2008. Performance metrics for climate models. J. Geophys. Res. **113**: D06104. doi:10.1029/2007JD008972

Goyette, S., and M. Perroud. 2012. Interfacing a one-dimensional lake model with a single-column atmospheric model: Application to the deep Lake Geneva, Switzerland. Water Resour. Res. **48**: W04507. doi:10.1029/2011WR011223

Hamilton, D. P., and S. G. Schladow. 1996. Prediction of water quality in lakes and reservoirs. Part I - Model Description. Ecol Modell. **96**: 91–110. doi:10.1016/S0304-3800(96)00062-2

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister. 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. Int. J. Climatol. **34**: 623–642. doi:10.1002/joc.3711

Harzallah, A. 1995. The interpolation of data series using a constrained iterating technique. Mon. Weather Rev. **123**: 2251–2254. doi:10.1175/1520-0493(1995)123<2251:TIODSU>2.0.CO;2

Henderson-Sellers, B. 1986. Calculating the surface energy balance for lake and reservoir modeling: A review. Rev. Geophys. **24**: 625–649. doi:10.1029/RG024i003p00625

Imboden, D. M., and A. Wüest. 1995. Physics and chemistry of lakes, p. 83–138. chap. 4 *In* A. Lerman, D. M. Imboden, and J. R. Gat [eds.], Mixing mechanisms in lakes. Springer-Verlag.

Jassby, A. D., J. E. Reuter, R. P. Axler, C. R. Goldman, and S. H. Hackley. 1994. Atmospheric deposition of nitrogen and phosphorus in the annual nutrient load of Lake Tahoe (California–Nevada). Water Resour. Res. **30**: 2207–2216. doi:10.1029/94WR00754

Jassby, A. D., C. R. Goldman, J. E. Reuter, and R. C. Richards. 1999. Origins and scale dependence of temporal variability in the transparency of Lake Tahoe, California-Nevada. Limnol. Oceanogr. **44**: 282–294. doi:10.4319/lo.1999.44.2.0282

Kennedy, J., and R. C. Eberhart. 1995. Particle swarm optimization, p. 1942–1948. *In* Proceedings of IEEE International Conference on Neural Networks, Institute of Electrical & Electronics Engineering, University of Western Australia, Perth, Western Australia.

Kettle, H., R. Thompson, N. J. Anderson, and D. M. Livingstone. 2004. Empirical modeling of summer lake surface temperatures in southwest Greenland. Limnol. Oceanogr. **49**: 271–282. doi:10.4319/lo.2004.49.1.0271

Kraemer, B. M., S., and others. 2015. Century-long warming trends in the upper water column of Lake Tanganyika. PLoS One **10**: e0132490. doi:10.1371/journal.pone.0132490

Lenters, J. D., T. K. Kratz, and C. J. Bowser. 2005. Effects of climate variability on lake evaporation: Results from a long-term energy budget study of Sparkling Lake, northern Wisconsin (USA). J. Hydrol. **308**: 168–195. doi:10.1016/j.jhydrol.2004.10.028

Livingstone, D. M. 2003. Impact of secular climate change on the thermal structure of a large temperate central European lake. Clim. Change **57**: 205–225. doi:10.1023/A:1022119503144

Livingstone, D. M., and A. F. Lotter. 1998. The relationship between air and water temperatures in lakes of the Swiss Plateau: A case study with palaeolimnological implications. J. Paleolimnol. **19**: 181–198. doi:10.1023/A:1007904817619

Livingstone, D. M., and M. T. Dokulil. 2001. Eighty years of spatially coherent Austrian lake surface temperatures and their relationship to regional air temperature and the North Atlantic Oscillation. Limnol. Oceanogr. **46**: 1220–1227. doi:10.4319/lo.2001.46.5.1220

Livingstone, D. M., and J. Padisák. 2007. Large-scale coherence in the response of lake surface-water temperatures to synopticscale climate forcing during summer. Limnol. Oceanogr. **52**: 896–902. doi:10.4319/lo.2007.52.2.0896

Magnuson, J. J., and others. 1997. Potential effects of climate changes on aquatic systems: Laurentian Great Lakes and Precambrian Shield Region. Hydrol. Process. **11**: 825–871. doi:10.1002/(SICI)1099-1085(19970630)11:8<825::AID-HYP509>3.0.CO;2-G

Martynov, A., L. Sushama, and R. Laprise. 2010. Simulation of temperate freezing lakes by one-dimensional lake models: Performance assessment for interactive coupling with regional climate models. Boreal Environ. Res. **15**: 143–164.

McCombie, A. M. 1959. Some relations between air temperatures and the surface water temperatures of lakes. Limnol. Oceanogr. **4**: 252–258. doi:10.4319/lo.1959.4.3.0252

Minville, M., F. Brissette, and R. Leconte. 2008. Uncertainty of the impact of climate change on the hydrology of a Nordic watershed. J. Hydrol. **358**: 70–83. doi:10.1016/j.jhydrol.2008.05.033

Mosier, T. M., D. F. Hill, and K. V. Sharp. 2014. 30-Arcsecond monthly climate surfaces with global land coverage. Int. J. Climatol. **34**: 2175–2188. doi:10.1002/joc.3829

Nash, J. E., and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models part I - a discussion of principles. J. Hydrol. **10**: 282–290. doi:10.1016/0022-1694(70)90255-6

North, R. P., D. M. Livingstone, R. E. Hari, O. Köster, P. Niederhauser, and R. Kipfer. 2013. The physical impact of the late 1980s climate regime shift on Swiss rivers and lakes. Inland Waters **3**: 341–350. doi:10.5268%2FIW-3.3.560

O'Reilly, C. M., S., and others. 2015. Rapid and highly variable warming of lake surface waters around the globe. Geophys. Res. Lett. **42**: 773–781. doi:10.1002/2015GL066235

Perroud, M. S., S. Goyette, A. Martynov, M. Beniston, and O. Anneville. 2009. Simulation of multiannual thermal profiles in deep Lake Geneva: A comparison of one-dimensional lake models. Limnol. Oceanogr. **54**: 1574–1594. doi:10.4319/lo.2009.54.5.1574

Piccolroaz, S. 2016. Prediction of lake surface temperature using the air2water model: Guidelines, challenges, and future perspectives. Adv. Oceanogr. Limnol. **7**: 36–50. doi:10.4081/aiol.2016.5791

Piccolroaz, S., M. Toffolon, and B. Majone. 2013. A simple lumped model to convert air temperature into surface water temperature in lakes. Hydrol. Earth Syst. Sci. **17**: 3323–3338. doi:10.5194/hess-17-3323-2013

Piccolroaz, S., M. Toffolon, and B. Majone. 2015. The role of stratification on lakes' thermal response: The case of Lake Superior. Water Resour. Res. **51**: 7878–7894. doi:10.1002/2014WR016555

Piccolroaz, S., E. Calamita, B. Majone, A. Gallice, A. Siviglia, and M. Toffolon. 2016. Prediction of river water temperature: A comparison between a new family of hybrid models and statistical approaches. Hydrol. Process. **30**: 3901–3917. doi:10.1002/hyp.10913

Quayle, W. C., L. S. Peck, H. Peat, J. C. Ellis-Evans, and P. R. Harrigan. 2002. Extreme responses to climate change in Antarctic lakes. Science **295**: 645. doi:10.1126/science.1064074

Reuter, J. E., and W. W. Miller. 2000. Aquatic resources, water quality, and limnology of Lake Tahoe and its upland watershed; Lake Tahoe watershed assessment, p. 373–375. *In* D. D. Murphy and C. M. Knopp [eds.], Lake Tahoe watershed assessment. Volume 2. Forest Service General Technical Report, PSW-GTR-175. U.S. Department of Agriculture Forest Service, Pacific Southwest Research Station, Albany, California.

Rupp, D. E. 2014. An evaluation of CMIP5 20th century climate simulations for the Southeast USA, p. 1–49. Oregon Climate Change Research Institute, USGS Southeast Climate Science Center.

Rupp, D. E., J. T. Abatzoglou, K. C. Hegewisch, and P. W. Mote. 2013. Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA. J. Geophys. Res. Atmos. **118**: 10884–10906. doi:10.1002/jgrd.50843

Sahoo, G. B., S. G. Schladow, and J. E. Reuter. 2010. Effect of sediment and nutrient loading on Lake Tahoe optical conditions and restoration opportunities using a newly developed lake clarity model. Water Resour. Res. **46**: W10505. doi:10.1029/2009WR008447

Sahoo, G. B., S. G. Schladow, and J. E. Reuter. 2013. Hydrologic budget and dynamics of a large oligotrophic lake related to hydro-meteorological inputs. J. Hydrol. **500**: 127–143. doi:10.1016/j.jhydrol.2013.07.024

Sahoo, G. B., A. L. Forrest, S. G. Schladow, J. E. Reuter, R. Coats, and M. Dettinger. 2016. Climate change impacts on lake thermal dynamics and ecosystem vulnerabilities. Limnol. Oceanogr. **61**: 496–507. doi:10.1002/lno.10228

Salmaso, N. 2005. Effects of climatic fluctuations and vertical mixing on the interannual trophic variability of Lake Garda, Italy. Limnol. Oceanogr. **2**: 553–565. doi:10.4319/lo.2005.50.2.0553.

Salmaso, N. 2010. Long-term phytoplankton community changes in a deep subalpine lake: Responses to nutrient availability and climatic fluctuations. Freshw. Biol. **55**: 825–846. doi:10.1111/j.1365-2427.2009.02325.x

Schabhüttl, S., P. Hingsamer, G. Weigelhofer, T. Hein, A. Weigert, and M. Striebel. 2013. Temperature and species richness effects in phytoplankton communities. Oecologia **171**: 527. doi:10.1007/s00442-012-2419-4

Schindler, D. W., and others. 1990. Effects of climatic warming on Lakes of the Central Boreal Forest. Science **250**: 967–970. doi:10.1126/science.250.4983.967

Schmid, M., S. Hunziker, and A. Wüest. 2014. Lake surface temperatures in a changing climate: A global sensitivity analysis. Clim. Change **124**: 301–315. doi:10.1007/s10584-014-1087-2

Schmid, M., and O. Köster. 2016. Excess warming of a Central European lake driven by solar brightening. Water Resour. Res. **52**: 8103–8116. doi:10.1002/2016WR018651

Schneider, P., S. J. Hook, R. G. Radocinski, G. K. Corlett, G. C. Hulley, S. G. Schladow, and T. E. Steissberg. 2009. Satellite observations indicate rapid warming trend for lakes in California and Nevada. Geophys. Res. Lett. **36**: L22402. doi:10.1029/2009GL040846

Schneider, P., and S. J. Hook. 2010. Space observations of inland water bodies show rapid surface warming since 1985. Geophys Res Lett. **37**: L22405. doi:10.1029/2010GL045059

Sharma, S., S. C. Walker, and D. A. Jackson. 2008. Empirical modelling of lake water-temperature relationships: A comparison of approaches. Freshw. Biol. **53**: 897–911. doi:10.1111/j.1365-2427.2008.01943.x

Sharma, S., and others. 2015. A global database of lake surface temperatures collected by in situ and satellite methods from 1985–2009. Sci. Data **2**: 150008. doi:10.1038/sdata.2015.8

Straile, D. 2000. Meteorological forcing of plankton dynamics in a large and deep continental European lake. Oecologia **122**: 44–50. doi:10.1007/PL00008834

Sugiyama, N., S. Kravtsov, and P. Roebber. 2017. Multiple climate regimes in an idealized lake–ice–atmosphere model. Clim. Dyn. doi:10.1007/s00382-017-3633-x

Tahoe Environmental Research Center (TERC), University of California Davis. 2016. Tahoe: State of the Lake Report 2016. Incline Village, NV, [accessed January 2017]. Available from http://terc.ucdavis.edu/stateofthelake/sotl-reports/2016/2016sotl_complete.pdf

Taylor, K. E., R. J. Stouffer, and G. A. Meehl. 2009. A summary of the CMIP5 experiment design. World Climate Research Program. 32 p. [accessed 25 May 2015]. Available from http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf

Thiery, W., V. M., and others. 2014. LakeMIP Kivu: Evaluating the representation of a large, deep tropical lake by a set of 1-dimensional lake models. Tellus Ser. A **66**: 21390. doi:10.3402/tellusa.v66.21390

Toffolon, M., S. Piccolroaz, B. Majone, A.-M. Soja, F. Peeters, M. Schmid, and A. Wüest. 2014. Prediction of surface temperature in lakes with different morphology using air temperature. Limnol. Oceanogr. **59**: 2182–2202. doi:10.4319/lo.2014.59.6.2185

Toffolon, M., and S. Piccolroaz. 2015. A hybrid model for river water temperature as a function of air temperature and discharge. Environ. Res. Lett. **10**: 114011. doi: 10.1088/1748-9326/10/11/114011

Van Cleave, K., J. D. Lenters, J. Wang, and E. M. Verhamme. 2014. A regime shift in Lake Superior ice cover, evaporation, and water temperature following the warm El Niño winter of 1997–1998. Limnol. Oceanogr. **59**: 1889–1898. doi:10.4319/lo.2014.59.6.1889

Webb, M. S. 1974. Surface temperatures of Lake Erie. Water Resour. Res. **10**: 199–210. doi:10.1029/WR010i002p00199

Wilby, R. L., and S. Dessai. 2010. Robust adaptation to climate change. Weather **65**: 180–185. doi:10.1002/wea.543

Winder, M., J. E. Reuter, and S. G. Schladow. 2009. Lake warming favors small-sized planktonic diatom species. Proc. R Soc B **276**: 427–435. doi:10.1098/rspb.2008.1200

Wood, T. M., S. A. Wherry, S. Piccolroaz, and S. F. Girdner. 2016. Simulation of deep ventilation in Crater Lake, Oregon, 1951–2099: U.S. Geological Survey Scientific Investigations Report 2016–5046. 43 p. [accessed 20 January 2017]. Available from https://doi.org/10.3133/sir20165046

Woolway, R. I., I. D. Jones, D. P. Hamilton, S. C. Maberly, K. Muraoka, J. S. Read, R. L. Smyth, and L. A. Winslow. 2015. Automated calculation of surface energy fluxes with high-frequency lake buoy data. Environ. Model. Softw. **70**: 191–198. doi:10.1016/j.envsoft.2015.04.013

Woolway, R. I., K., and others. 2016. Lake surface temperatures [in "State of the Climate in 2015"]. Bull. Am. Meteorol. Soc. **97**: S17–S18. doi:10.1175/2016BAMSStateoftheClimate.1

Woolway, R. I., P. Meinson, P. Nõges, I. D. Jones, and A. Laas. 2017. Atmospheric stilling leads to prolonged thermal stratification in a large shallow polymictic lake. Clim. Change **141**: 759–773. doi:10.1007/s10584-017-1909-0

Woolway, R. I., M. T. Dokulil, W. Marszelewski, et al. 2017. Climatic Change. **142**: 505–520. doi:10.1007/s10584-017-1966-4

Zhong, Y., M. Notaro, S. J. Vavrus, and M. J. Foster. 2016. Recent accelerated warming of the Laurentian Great Lakes: Physical drivers. Limnol. Oceanogr. **61**: 1762–1786. doi:10.1002/lno.10331

Acknowledgments

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling group at the National Center for Atmospheric Research (NCAR) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We thank the anonymous Reviewers for their constructive comments, which helped us to improve the manuscript. The research described in this paper was carried out, in part, at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Conflict of Interest

None declared.

Submitted 25 January 2017
Revised 24 April 2017
Accepted 06 June 2017

Associate editor: Craig Stevens