Appendix S3: *In Situ* Filters

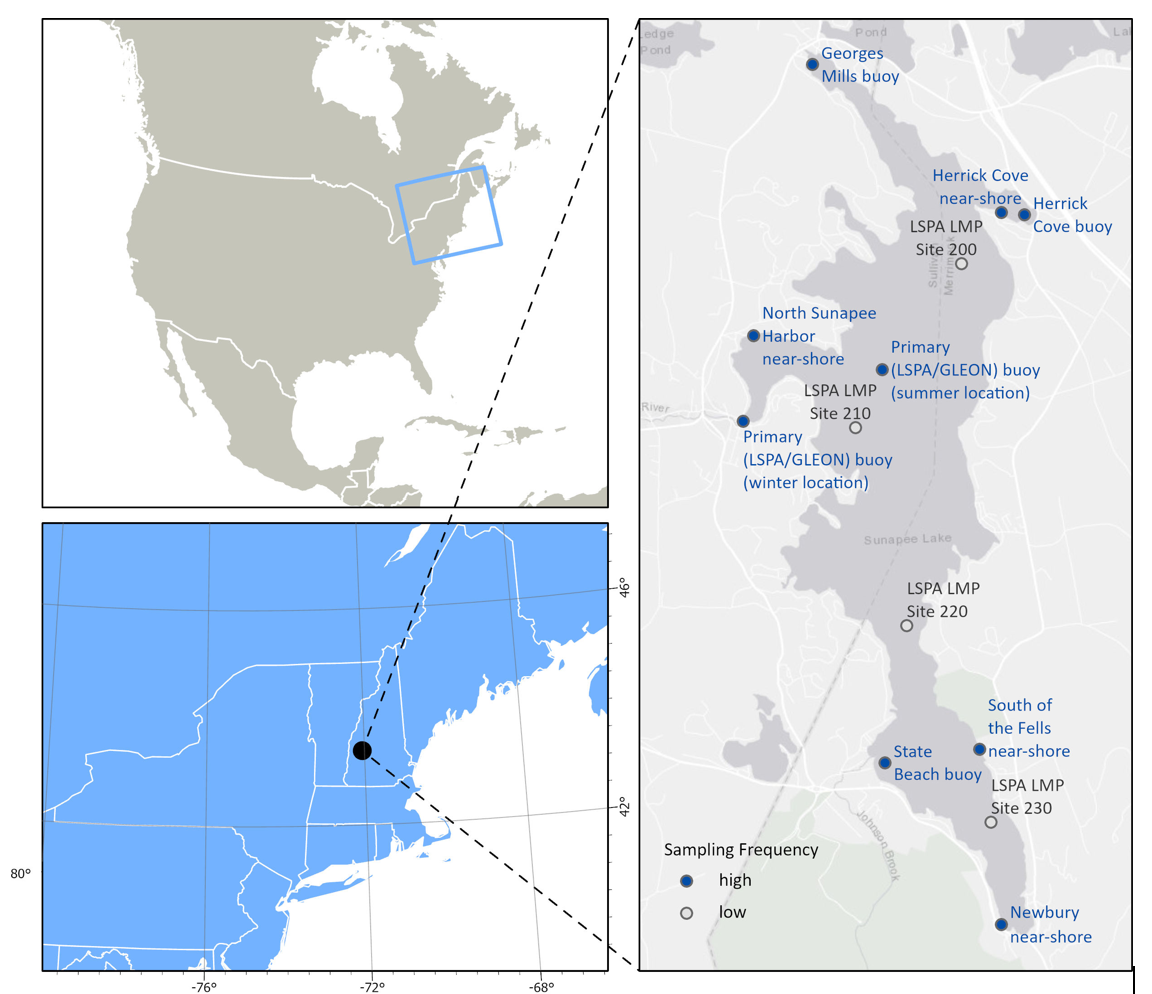
lakeCoSTR: An open-source, interactive retrieval tool to facilitate use of the Landsat Collection 2 surface temperature product to estimate lake surface water temperatures

*Herrick, C, Steele, BG*, Brentrup, JA, Cook, B, Cottingham, KL, Ducey, M, Johnson, K, Lutz, DA, Palace, M, Sullivan, F, Thompson, M, Trout-Haney, JV, Weathers, KC

Ecosphere

# Background

Given that Lake Sunapee is well-instrumented with temperature recording instruments at numerous locations in the lake (Figure S1), we examined the performance of a number of filtered *in situ* data sets for ground-truthing of the Landsat data from the lakeCoSTR tool. In this appendix, we present the results from 6 Landsat-*in situ* data sets from the Lake Sunapee *in situ* temperature database in order to make suggestions about what kind of data may be needed or preferable for validation of the Landsat surface temperature product.



Lake Sunapee with locations labelled

**Figure S1. The study area, Lake Sunapee, on the right with all *in situ* sites labeled.**

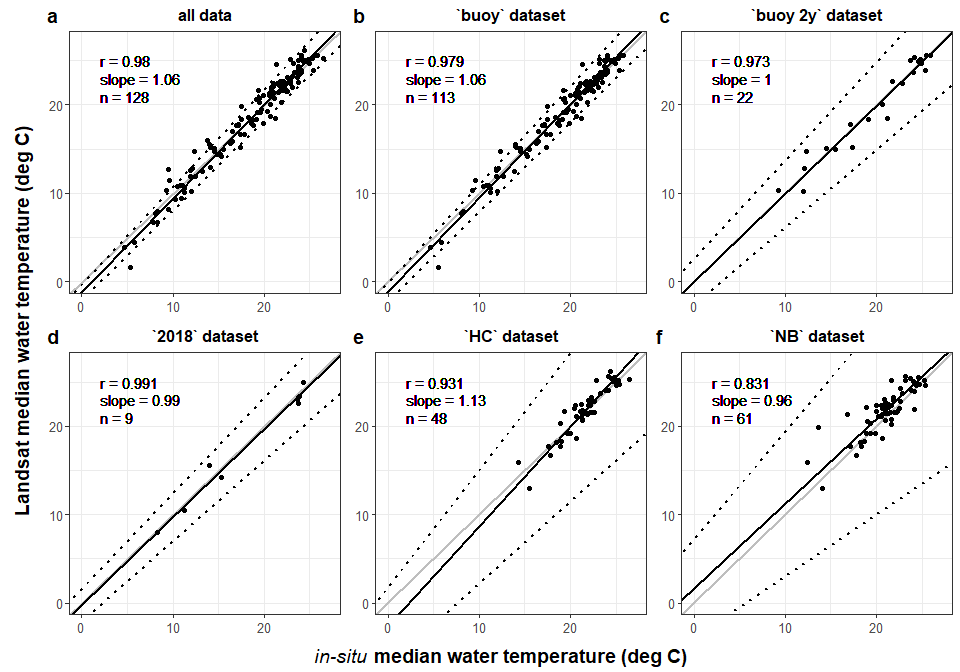
# Methods and Results

All Landsat data were filtered using the kurtosis filter described in section 3.1.1 of the manuscript and Appendix S2. The paired Landsat-(in situ\* data sets presented here were filtered from the complete validation data set described in section 3.1.2 in the manuscript. All analyses here use the median Landsat surface temperature value and the median *in situ* value from the subset of data described below. Deming regression (Deming 1943), as described in section 3.1.2, was used to analyze performance of each of the validation data sets.

Description of Landsat-*in situ* paired data sets presented in this appendix: \* all data: all Landsat-*in situ* pairs (as presented in manuscript) \* buoy: all available data from the primary LSPA/GLEON buoy only \* buoy 2y: two years (2019-2020) of data from the primary LSPA/GLEON buoy \* 2018: all data from 2018, when temperature data were collected from 9 sites \* HC: data from Herrick Cove near-shore sensor only \* NB: data from Newbury near-shore sensor only

These subsets of data were chosen to mimic the kinds of data availability that may be present for other systems. The Herrick Cove and Newbury sites were selected due to the length of their data records.

Regressions for each of these data subsets exhibited strong positive correlations with slopes close to 1 (Figure S2). These analyses suggest that, at Lake Sunapee, two years of temperature data from a single buoy at a central location (panel c) or data from multiple locations in a single year (panel d) both provide adequate data to ground-truth Landsat-derived temperature. Moreover, even when limiting the *in situ* dataset to singular near-shore sites (panels e, f), the *in situ* and Landsat-derived temperatures seem to generally agree, though more prediction error is associated with those models.



**Figure S2. Deming regressions for 6 paired Landsat - in situ datasets: a) all data, b) the ‘buoy’ subset, c) the ‘buoy 2y’ subset for 2018-2019, d) the ‘2018’ data for 9 sites, e) the Herrick Cove ‘HC’ subset, and f) the Newbury ‘NB’ subset. Within each panel, the gray line is the 1:1 line, the black line is the Deming regression, and the dotted black line is the 95% confidence interval for the regression line. The Pearson correlation, slope, and n are indicated in the upper left corner.**

# Discussion and Validtion Dataset Suggestions

These analyses suggest that, generally speaking, any high-frequency temperature data can be used to evaluate the Landsat surface temperature data, but using a median value from multiple locations or from a single, centrally-located location may perform the best and give the user the best insight into potential biases in the Landsat data. Even the near-shore sites (Figure S2, e and f) show that the data obtained using lakeCoSTR are in the correct range - and given that the slope slope is near 1, no calibration of the Collection 2 product is required. In general, we suggest that best practice would be to evaluate the lakeCoSTR data with *in situ* values from at least one centrally-located high-frequency temperature sensor. Further, the Landsat-*in situ* pairs should cover a large range of water temperature values in order to assess biases in the Landsat data.

It is important to note that the data collected from the cove buoys in 2018 and all of the near-shore sites were obtained using low-cost sensors that did not require an instrumented buoy with external power. HOBO and Onset temperature sensors, like those used to gather data at these locations, could be deployed for a few ice-free seasons in order to evaluate the Landsat temperature product for specific lakes. Given the additional QA/QC filtering described in Appendix S2 that was used for this study, we would advise having enough *in situ* data to validate the lakeCoSTR data product for individual lakes before applying the data from the product. Additionally, data that span the time period of interest may be especially important, given the bias observed in early-season data at Lake Sunapee.

# Suggestions for Data Collection

When new *in situ* data collection designs are being planned or existing ones augmented, we advise users limited to manual, low-frequency sampling with handheld instruments to coordinate their sampling with Landsat satellite flyovers. The Landsat flyover calendar can be viewed using the USGS Landsat Acquisition Tool (<https://landsat.usgs.gov/landsat_acq>). While it is possible to widen the Landsat-*in situ* match window beyond the narrow window used in this study (Sharaf et al. 2019 used a 3 day time window, for example) within the lakeCoSTR tool to increase the likelihood of matches with low-resolution data, this broadening will add more error to the Landsat-*in situ* comparison process. Importantly, the usefulness in temporal proximity of *in situ* measurements to Landsat flyover is not limited to water temperature. Future satellite data tool and algorithm development, especially for parameters like chlorophyll-a that are more variable than temperature both temporally and spatially, will rely on well-matched Landsat-*in situ* data.

Additionally, for both handheld measurements and sensors placed *in situ* for extended periods of time, our analyses suggest that both the vertical location in the water column relative to the surface and the horizontal location within the lake matter (Appendix S1: Figure S1). Vertically, measurements taken as near to the surface as possible should be most representative of the satellite measurements. At Lake Sunapee, there was less error in the regression models developed from sensors located close to the surface (0.1 m) versus lower in the water column (1.5 m) (Appendix S1: Figure S1h); we hypothesize that depth may also influence prediction error in other systems. Spatially, if the number of sites is limited, having *in situ* data from a central, deep-water location generated better Landsat-*in situ* regressions than relying solely on nearshore, littoral locations. Where resources permit, having multiple locations that experience different temperature regimes through the year (e.g., Woolway and Merchant 2018) may help to generate a realistic idea of the horizontal variability in temperature across the lake surface, which can aid the user’s inspection and interpretation of the lakeCoSTR histograms and scene summary.

# Literature Citations

Deming, W. E. 1943. Statistical Adjustment of Data. Statistical Adjustment of Data. Oxford, England: Wiley.

Sharaf, Najwa, Ali Fadel, Mariano Bresciani, Claudia Giardino, Bruno J. Lemaire, Kamal Slim, Ghaleb Faour, and Brigitte Vinçon-Leite. 2019. “Lake Surface Temperature Retrieval from Landsat-8 and Retrospective Analysis in Karaoun Reservoir, Lebanon.” Journal of Applied Remote Sensing 13 (4): 044505. <https://doi.org/10.1117/1.JRS.13.044505>.

Woolway, R. Iestyn, and Christopher J. Merchant. 2018. “Intralake Heterogeneity of Thermal Responses to Climate Change: A Study of Large Northern Hemisphere Lakes.” Journal of Geophysical Research: Atmospheres 123 (6): 3087–98. <https://doi.org/10.1002/2017JD027661>.