

Technical Note

A Note on the Collection and Cleaning of Water Temperature Data

Colin Sowder^{1,*} and E. Ashley Steel²

¹ Statistics Department, University of Washington, Seattle, WA 98115, USA

² PNW Research Station, USDA Forest Service, 400 N 34th Street, Suite 201, Seattle, WA 98103, USA; E-Mail: asteel@fs.fed.us

* Author to whom correspondence should be addressed; E-Mail: csowder@uw.edu; Tel.: +1-206-616-7430; Fax: +1-206-685-7419.

Received: 15 June 2012; in revised form: 13 July 2012 / Accepted: 1 August 2012 /

Published: 20 August 2012

Abstract: Inexpensive remote temperature data loggers have allowed for a dramatic increase of data describing water temperature regimes. This data is used in understanding the ecological functioning of natural riverine systems and in quantifying changes in these systems. However, an increase in the quantity of yearly temperature data necessitates complex data management, efficient summarization, and an effective data-cleaning regimen. This note focuses on identifying events where data loggers failed to record correct temperatures using data from the Sauk River in Northwest Washington State as an example. By augmenting automated checks with visual comparisons against air temperature, related sites, multiple years, and available flow data, dewatering events can be more accurately and efficiently identified.

Keywords: accuracy; precision; errors; quality assurance

1. Introduction

Reliable data describing water temperature regimes is needed to understand ecological functioning of natural streams and rivers and to quantify anthropogenic impacts such as forest management, urbanization, hydropower, climate change, and river restoration. Small, relatively inexpensive water temperature loggers became available in the early 1990s. Since then, these loggers have been installed in tens of thousands of streams and rivers around the world [1].

These modern temperature sensors can quickly and easily be installed using steel cable or epoxy to rocks [2,3]. The data loggers can measure and record water temperature at any time interval, e.g., every minute, every hour, or every day. However, this vast accumulation of water temperature data has brought new challenges. First, large quantities of data require complex data management systems. Second, it is unclear what metrics should be synthesized and reported in order to capture the “thermal regime.” Third, the data loggers may come out of the water as a result of storms, droughts, anchoring failures, logger malfunction, or public interference. Unless the loggers are checked daily (or even hourly in some situations), it is impossible to know for sure that the logger was correctly logging water temperature during deployment. The desire for more records in unregulated streams increases this uncertainty, as unregulated streams are inherently more variable. The effort required to clean and prepare water temperature data for analysis is, in fact, often larger than the effort required for collecting the data in the field. This third challenge is the focus of our research note. Management agencies recognize the difficulty of cleaning raw temperature data, and many of the most complete data cleaning protocols are available from this grey literature (see, for example [3–6]). We combine these protocols with our own visual comparisons to provide a simple set of processing steps for cleaning time series of water temperature data. As these data are collected at an ever increasing number of locations every day, our synthesis of data cleaning procedures can help researchers use stream temperature data efficiently and accurately.

2. Methods: Synthesis and Application

Accuracy and precision can be checked before deployment. Most of the commonly used temperature loggers for remote field use cannot be calibrated. Temperature logger manufacturers suggest that a quick accuracy check can be performed by submerging the logger in melting ice and confirming that the logger records 0 °C [7]. However, the simplest method for checking accuracy and precision is to initialize all loggers several days before deployment and place them in a common environment. One can assume that loggers recording the same temperature are all accurate; any one logger recording a value far from the other loggers is likely inaccurate. Variance across loggers can be used to verify manufacturer-reported precision. In addition, this check can be used to test the battery health of all loggers prior to deployment. A similar check can be repeated following recovery to ensure logger accuracy and precision over the deployment period [4].

The next step in the data quality assurance process is initialization of the loggers, including start time, recording unit, and recording interval. Setting all loggers for a particular project to the same start time, recording unit, and recording interval will greatly simplify data management. In the Sauk River, Washington State example we explore here, 48 observations per day were recorded. Many other projects have chosen to record water temperature every hour. The choice will depend on the metrics of most interest. If maximums or minimums are of greatest interest, a more frequent recording interval is more likely to record a value close to the true maximum or minimum. If variance is of greatest interest, a more frequent interval will provide a more precise estimate. A more frequent interval will also provide a more precise estimate of mean temperature but may not be worth the cost of greater data management, storage, and cleaning demands. Note that, as data are processed, metrics will treat observations taken at 10:01:00 in one stream as occurring at the same time as those taken at 10:29:00

in another. This incongruity can be avoided by initializing all loggers to begin at, say, 12:00:00 am. Finally, if the period of the experiment overlaps with a switch from standard time to daylight saving time, many loggers can be set to observe standard time to simplify data processing.

Other researchers have carefully covered deployment protocols to maximize data quality [2–4]. The goal of in-field deployment is to minimize the chance of recording erroneous information. Depending on the river system, protocols suggest cabling the logger to nearby trees, rebar or rock-filled sandbags [3] or epoxying the logger to large rocks or manmade structures [2]. While protection from direct sunlight can be provided by undercut banks and nestling loggers below large boulders, a plastic casing should be employed as a sun shield at any site with the possibility of direct sun exposure, as direct sun exposure has been shown to bias underwater temperature logger readings [8]. However, there is no perfect system. Cabling can lead to dewatering if the cable is too short and the flow is reduced; long cables can lead to dewatering when high flow events pendulum the logger up onto the shore. Epoxy is a good alternative where appropriate structure exists; however, many streams and rivers do not have such conveniences. River systems remain inherently variable—streams can dry up in the summer, floods can bury loggers in silt, and smaller side channels can move to entirely new locations in the course of a single storm. Despite efforts to protect the logger in the field, some field loggers are likely to come out of the water naturally [5]. Finally, human interference can lead to erroneous water temperature recordings for a short period if the curious public pulls the logger out of the water and fiddles with it for a while before replacing, and for long periods, if the logger is permanently removed from the water.

Post-recovery data cleaning begins by removing observations recorded before and after the logger is correctly positioned in the stream channel. Observations outside the deployment period can be automatically cleaned using field notes indicating the exact times of deployment and recovery. During recovery, field notes should also indicate any anomalies such as whether the logger was out of the water or buried in sand. When graphing data, these times can be visually verified if the ambient temperature is dramatically different than the stream temperature. This convenient check can also be forced by keeping the loggers warmed with inexpensive hand warmers or cooled with ice packs in the hours before deployment and after recovery.

Once the data has been downloaded and observations outside of the deployment period have been removed, a series of automatable checks can be used to flag individual files for potential errors. Automated flags can include a thermal maximum relevant to the river system being studied, a thermal minimum (recommended $-1\text{ }^{\circ}\text{C}$), and rapid thermal variation (hourly and/or daily) [9]. When cleaning our Sauk River dataset, we flagged data for probable errors when they exceeded a thermal maximum of $25\text{ }^{\circ}\text{C}$, a thermal minimum of $-1\text{ }^{\circ}\text{C}$, or a daily change of $10\text{ }^{\circ}\text{C}$.

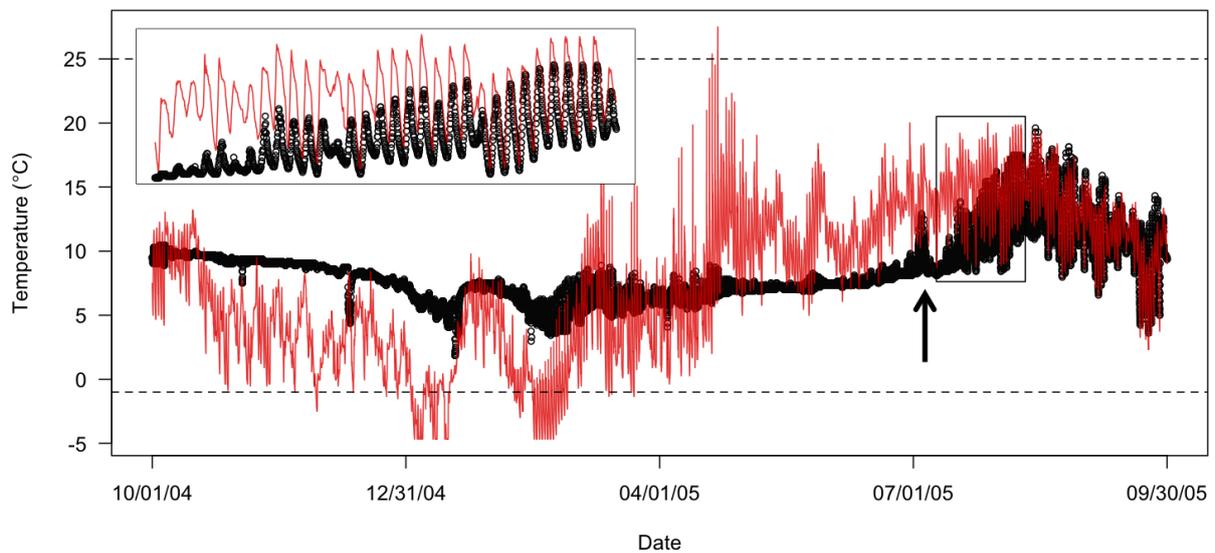
Though the recommended automated checks are a good baseline, they are insufficient for identifying all errors or specifying the exact time and duration of an anomalous event. Human judgment can improve quality assurance procedures. The human eye is a powerful tool for identifying complex and familiar patterns. For example, computer algorithms alone remain stymied by the CAPTCHA system, the online system that uses images of words to verify a human user. Visual inspection of each time series of data will be required for catching anomalies that are not anticipated by the automated checks (such as a dewatering event in the spring when air temperatures do not exceed the thresholds of the automated checks) and for specifying the time and duration of errors in the raw

data files. While it is tempting to remove all observations flagged by automatic checks, it is also possible that an extreme or unexplainable event is accurate and particularly important to study. Comparisons between the data record being cleaned and air temperature data, data from other years at the same site, data from nearby loggers, and flow data improve one's ability to detect potential anomalies, posit their causes, and fine-tune estimates of start and end times of anomalies.

2.1. Stream to Air Temperature Comparisons

If an air temperature logger was deployed in conjunction with the stream loggers, a plot of site data compared against air temperature data can identify periods in which the stream logger closely matches air temperature suggesting erroneous observations [4,6]. A close correspondence between water and air temperature is a strong indication that the stream logger was out of the water. In temperate regions in summer, the air is often much warmer than the stream and, in winter, the air can be colder. In all seasons, water temperatures are generally less variable day to day than ambient air temperatures. Figure 1 presents an example from a side channel of the Sauk River in which a comparison between in-stream logger recordings and air temperature can be used to identify a dewatering event that was missed by automated checks.

Figure 1. Identifying a dewatering event by comparing water temperature data to reference air temperature data.



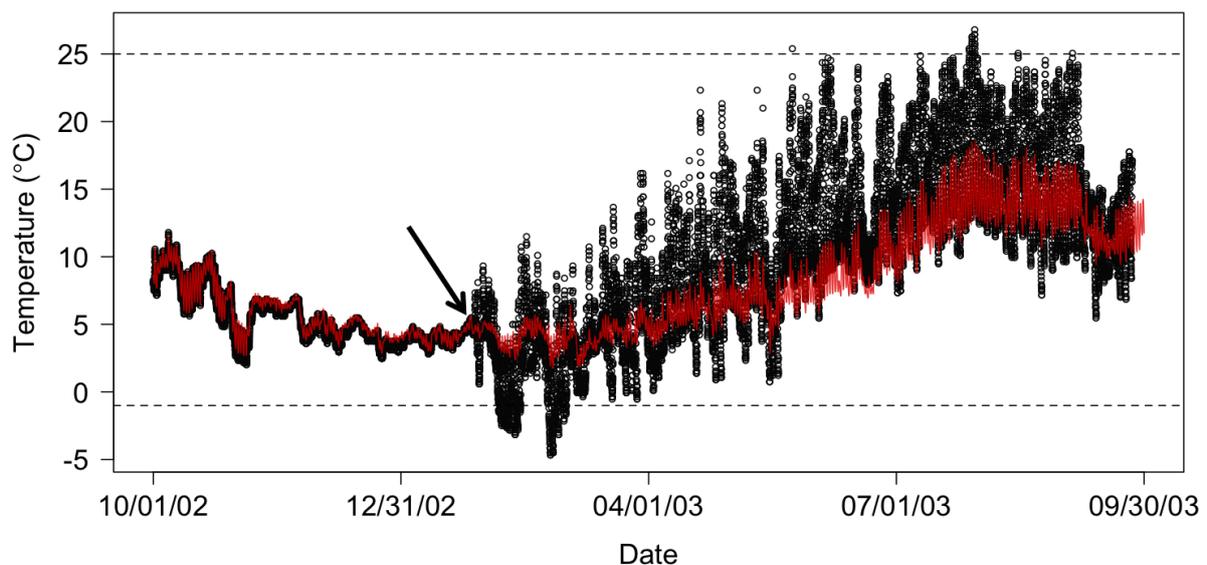
In Figure 1, water temperature data (open black circles) from a side channel on the Sauk River, near Darrington WA, and reference air temperatures (red line) are used to identify a dewatering event. Prior to coming out of the water, the in-stream logger recorded less variable and cooler temperatures than the air temperature logger. Without an air temperature comparison, a change in the logger's readings can be seen in late July 2005. However, it is unclear whether the logger was recording erroneously before or after the change, or if the logger was correctly recording temperatures throughout the time period. The air temperature comparison shows that in late July 2005, the logger begins to follow the air temperature, even exceeding it at times, suggesting that the logger came out of the water. The highlighted region from 10 July to 10 August shows the transition to recording air temperature.

Conservatively, all data from a spike on 4 July 2005, until the end of the hydrologic year were removed. Note that the erroneous water temperature data never exceeded the thermal thresholds of the automated checks (dashed lines).

2.2. Nearby Site Comparison

If air temperatures are unavailable or inconclusive, temperature measurements from a similar, nearby site can be used for comparison. In Figure 2, water temperature data from two sites on the mainstem Sauk River are used to identify a dewatering event in early 2003. Automated checks flagged the upstream site as having potentially erroneous data and field notes confirm that this logger was out of the water at the time of recovery in late September 2003. Prior to coming out of the water, there was a strong correspondence between the upstream logger (open black circles) and the downstream logger (red line). After 27 January 2003, the upstream logger becomes more variable than the downstream logger. The change suggests that the data are erroneous and should be removed. It should be noted that events such as flooding could cause multiple loggers to become dewatered simultaneously, so it is possible that temperature measurements recorded across loggers are highly correlated yet are in error.

Figure 2. Identifying a dewatering event by comparing nearby sites.



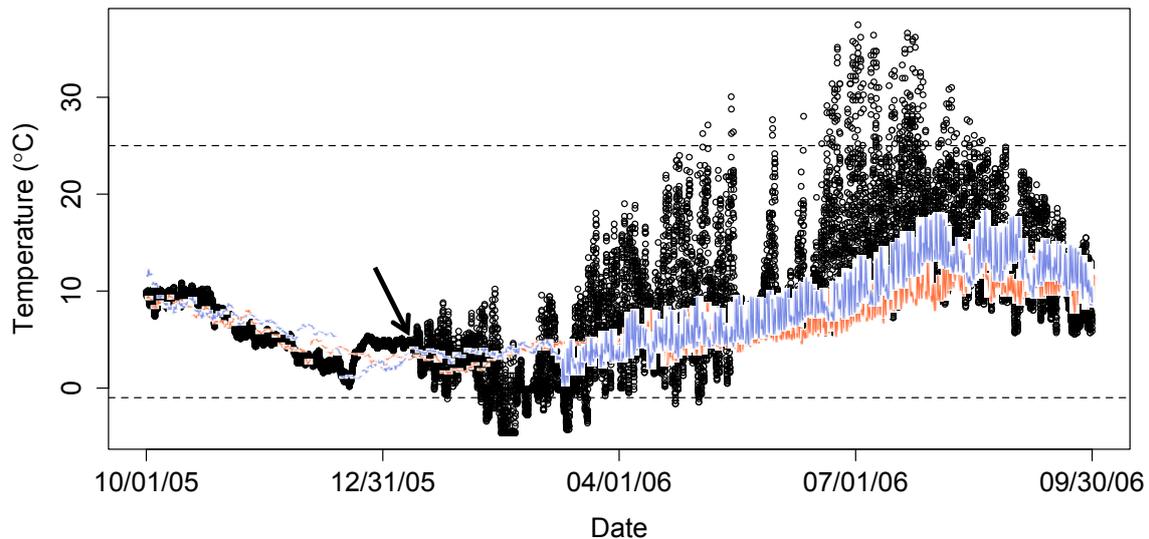
2.3. Comparisons Across Years

An additional comparison can be made for studies conducted over numerous years. Though correspondence will not be as strong between years as in either the air comparison or the similar site comparison, comparing observations from a single site over multiple years of data can also be demonstrative in identifying errors. Often multiple years of data will suggest a window of reasonable values and expected daily variation across seasons. When data from one year are dramatically different, there may be data errors.

In Figure 3, water temperature data from one side channel of the Sauk River are compared across multiple years. Across the winter months, despite differing weather each year, data from 2005 (open black circles), 2007 (orange line), and 2008 (blue line) showed similar patterns. Beginning in January,

the data from 2005 show a dramatically different pattern than data from the other years. The change suggests that the data from January to September 2006 are erroneous and should be removed. Note that this record was flagged for errors during automated checked because the observations exceeded 25 °C and fell below −1 °C (dashed lines).

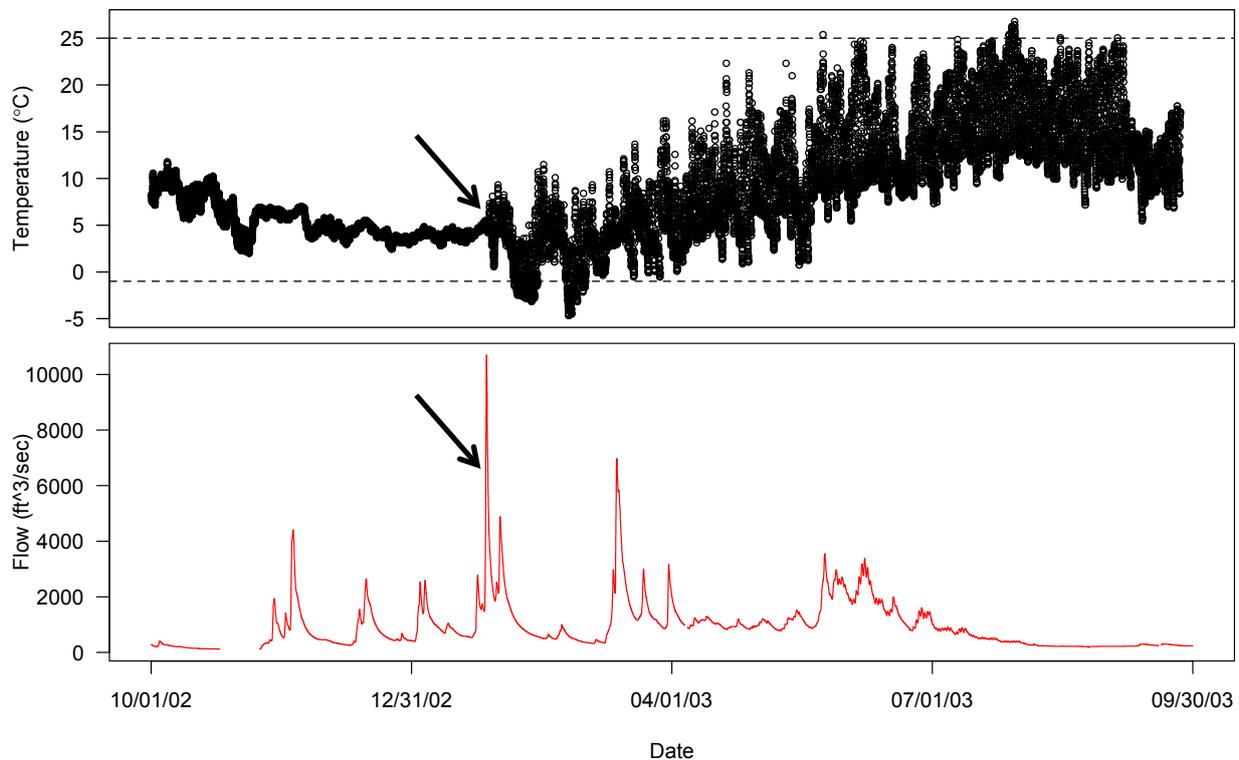
Figure 3. Identifying a dewatering event by comparing a site to multiple reference years at the same site.



2.4. Comparison with Flow Data

Dewatering events generally occur when the river experiences extreme conditions such as droughts or floods. When other comparisons are unavailable or inconclusive, a final check may be to plot stream flow from either the stream in question or a similar site in the same river system on top of the water temperature data if flow data is available. Floods can precede a logger getting deposited on a bank, and low flow near the logger can suggest a stream drying up. Flow data is available from automated flow monitoring systems. For the Sauk River, flow data is publicly available from the U.S. Geological Survey [10]. Figure 4 demonstrates a high flow day preceding a dewatering event. In this figure, water temperature data from the upstream site on the mainstem Sauk River and USGS flow information are used to identify a dewatering event. A high flow (red line) event on 26 January 2003, immediately precedes the dewatering event experienced by the logger (open black circles) identified in Figure 2 on 27 January 2003, providing a precise estimate of exactly when the erroneous data begin, and a clear explanation of the dewatering event.

Figure 4. Identifying a dewatering event by comparing temperature records with publicly available flow information.



3. Discussion and Conclusions

Field notes, automated checks, and graphical comparisons can provide excellent estimates of whether and when errors occurred within a time series of water temperature data. While field notes and automated checks are often sufficient for determining if an error occurred, a combination of the four graphical comparisons can help identify exactly when, and for what duration, the readings were in error. All four comparisons add value, and each has its own strengths and weaknesses. Comparisons to air temperature clearly identify dewatering events, but require the installation of air temperature sensors, which can increase the cost of monitoring systems over large areas. Comparisons between sites do not require additional sensors, but require care in determining what constitutes a similar site as well as noting extreme events such as floods or droughts may cause multiple sensors to read erroneously simultaneously. Comparisons across years are only applicable for studies that have multiple years to compare, and then can only identify significant deviations from past trends. Finally flow data is not available in all cases, and can help identify erroneous readings caused by flooding and by drought. Recommendations for data-processing are summarized in Table 1.

Development of more sophisticated automated algorithms is ongoing for remote sensor data, but these novel techniques have not yet been widely applied to ecological data [11]. Such algorithms will likely be built by quantifying the types of visual assessments we propose here through change point analyses, signal decomposition, and comparisons across multiple different types of records. Validation of algorithms could be accomplished by feeding them datasets from a wide variety of streams with known or induced anomalies and analyzing the algorithms' ability to catch these test anomalies as well as their propensity to return false positives. While the best quality assurance methods will likely

continue to incorporate some human judgment [11], improvements in automated algorithms are needed to reduce data cleaning time.

Table 1. Steps for maximizing quality of water temperature data collected using remote field loggers.

Research or Monitoring Stage	Recommended Procedure	Purpose of Recommended Procedure
Logger initialization	Set all loggers to the same settings: <ul style="list-style-type: none"> - observations per day; - standard Time; - units (C or F); - time of first observation (recommend 0:00:00). 	Facilitates combining data files across locations, years, and projects.
	Check logger accuracy by placing in common environment for a short period. Can be repeated after logger recovery.	Can indicate if one logger is inaccurate pre-deployment and provides an estimate of precision across loggers.
Logger Deployment and Recovery	Take detailed field notes: <ul style="list-style-type: none"> - deployment date and time; - recovery date and time; - indicate if hobo was dewatered (either in the air or buried) at recovery. 	Accurate times of deployment and recovery allow for automated removal of observations recorded before and after logger was in the water. Recovery notes are the first indication of data errors.
	Install air temperature logger outside of direct sunlight in deployment region. Redundant loggers can be deployed in the same channel.	Air temperature data enables visual comparison with stream temperature data. In highly variable channels, redundant loggers increase the likelihood of having at least one logger in the water at all times.
Post-recovery	Automated Checks: <ul style="list-style-type: none"> - flag files with field notes indicating dewatering; - flag observations outside of thermal limits based on river system; - flag days with rapid change in temperature. 	Flagged files are highly likely to contain errors. However, unflagged files may still contain errors and should be visually checked with the methods below.
	Graphical Checks: <ul style="list-style-type: none"> - plot each file individually to look for abnormalities; - graphically compare each logger file to air temperature data if it exists; - graphically compare data across sites; - graphically compare data across years; - if available, graphically compare temperature data with flow data 	Exact times of anomalies can be hard to detect with only the individual file. Graphical comparisons can confirm or identify periods of poor data quality and make identification of the start and end of anomalous records easier.

Once errors have been identified, the question becomes how much data is reliable. To ensure that no dewatered observations are used, the entire data file could be discarded. To ensure that no valid observations are lost, the data file could be used without cleaning. Clearly, a balance is required

between these two types of errors. Study objectives and types of metrics to be calculated will determine which data-cleaning error, retaining erroneous data or discarding good data, is most problematic. Often, using a combination of automated and graphical checks, it will be possible to pinpoint probable errors and, therefore, simultaneously minimize both types of data cleaning error. Despite careful protocols, the data cleaning process is inherently subjective. Therefore, we recommend maintaining both the original and the cleaned data files.

Remote water temperature loggers allow researchers to investigate and quantify the thermal regimes of vastly more streams and rivers than could possibly be accomplished through direct monitoring by field personnel. Research and management continue to demand and collect water temperature data, particularly from unregulated rivers [1]. As more and more loggers are deployed, and deployed for longer periods of time, care must be taken pre- and post- deployment to ensure that the data accurately reflect water temperatures. A series of initialization, deployment, and data cleaning steps (Table 1) can improve data quality, reduce the time and effort of data storage and processing, and increase collaboration and comparison across projects, streams, and regions.

Acknowledgments

This project was supported by NOAA's Northwest Fisheries Science Center and the USDA Forest Service, Pacific Northwest Research Station. We appreciate the field assistance of Karrie Hanson and insights into the original project design from Tim Beechie, both at NOAA's Northwest Fisheries Science Center. We thank Daniel Isaak at the USDA Forest Service, Boise Aquatic Sciences Laboratory and two anonymous reviewers for their comments and suggestions on earlier versions of this manuscript.

This is a US Government work and is in the public domain in the United States of America.

References

1. Isaak, D.J. *New Studies Describe Historic & Future Warming in Northwest US Streams*; Air, Water and Aquatic Environments Sciences Program: Fort Collins, CO, USA, 2012; Climate-Aquatics Blog #23. Available online: http://www.fs.fed.us/rm/boise/AWAE/projects/stream_temp/stream_temperature_climate_aquatics_blog.html (accessed on 1 August 2012).
2. Isaak, D.J.; Wollrab, S.; Horan, D.; Chandler, G. Climate change effects on stream and river temperatures across the northwest U.S. from 1980–2009 and implications for salmonid fishes. *Clim. Chang.* **2012**, *113*, 499–524.
3. Mauger, S. *Water Temperature Data Logger Protocol for Cook Inlet Salmon Streams*; Cook Inletkeeper: Homer, AK, USA, 2008; pp. 1–10. Available online <http://inletkeeper.org/resources/contents/water-temperature-data-logger-protocol> (accessed on 1 August 2012).
4. Bilhimer, D.; Stohr, A. *Standard Operating Procedures for Continuous Temperature Monitoring of Freshwater Rivers and Streams Conducted in a Total Maximum Daily Load (TMDL) Project for Stream Temperature, Version 2.2*; Washington State Department of Ecology: Olympia, WA, USA, 2008; SOP Number EAP044. Available online: http://www.ecy.wa.gov/programs/eap/qa/docs/ECY_EAP_SOP_Cont_Temp_Monit_TMDL_v2_3EAP044.pdf (accessed on 1 August 2012).

5. Dunham, J.; Chandler, G.; Rieman, B.; Martin, D. *Measuring Stream Temperature with Digital Data Loggers: A User's Guide*; Department of Agriculture, U.S. Rocky Mountain Research Station, Forest Service: Fort Collins, CO, USA, 2005; General Technical Report RMRS-GTR-150WWW; pp. 1–15. Available online: http://www.fs.fed.us/rm/pubs/rmrs_gtr150.pdf (accessed on 1 August 2012).
6. Kashiwagi, M.T.; Prochaska, A. *Quality Assurance Document for Temperature Monitoring*; Draft Version 1.3; Maryland Department of Natural Resources, Monitoring and Non-tidal Assessment Division: Annapolis, MD, USA, 2010. Available online: http://www.dnr.state.md.us/streams/pdfs/QA_TemperatureMonitoring.pdf (accessed on 1 August 2012).
7. *How do I do a Quick Accuracy Check*; Onset Corporation: Pocasset, MA, USA, 2012. Available online: <http://www.onsetcomp.com/support/faq/quick-accuracy-check> (accessed on 14 August 2012).
8. Isaak, D.J.; Horan, D.L. An assessment of underwater epoxies for permanently installing temperature sensors in mountain streams. *North Am. J. Fish. Manag.* **2011**, *31*, 134–137.
9. Rieman, B.E.; Chandler, G.L. *Empirical Evaluation of Temperature Effects on Bull Trout Distribution in the Pacific Northwest*; U.S. Environmental Protection Agency: Boise, ID, USA, 1999. Available online: http://www.fs.fed.us/rm/boise/publications/fisheries/rmrs_1999_riemanb001.pdf (accessed on 1 August 2012).
10. *USGS Current Water Data for the Nation*; U.S. Geological Survey: Reston, VA, USA, 2012. Available online: <http://waterdata.usgs.gov/nwis/rt> (accessed on 8 February 2012).
11. Porter, H.J.; Hanson, P.C.; Lin C.C. Staying afloat in the sensor data deluge. *Trends Ecol. Evol.* **2012**, *27*, 121–129.

Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).