Assessing Subjectivity in Sensor Data Post Processing via a Controlled Experiment

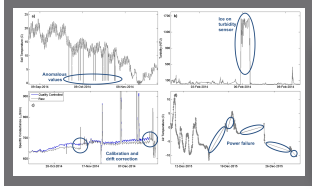
Amber Spackman Jones, (435) 797-7147, amber.jones@usu.edu
Jeffery S. Horsburgh, David P. Eiriksson

IN41C-0050

This project is funded by National Science Foundation grant 1208732. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

The Problem

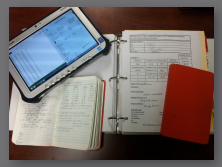
Environmental data collected by *in situ* sensors need quality control that often requires making edits in post processing to generate approved datasets. Technicians with the same level of training and using the same input datasets may produce different results, affecting the overall quality and comparability of finished data products.

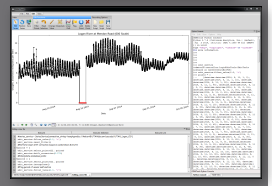


The Experiment

To assess the effects of subjective decision making, we designed an experiment:

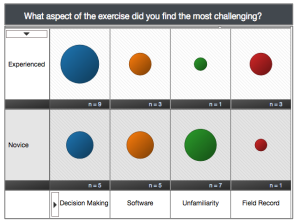
- Participants (n=27) included novices unfamiliar with and technicians experienced in quality control.
- Performed quality control post processing on the same datasets: one year of water temperature, pH, and specific conductance.
- Given consistent set of guidelines, field notes, and tools.
- Used ODMTools: https://github.com/ODM2/ODMToolsPython/





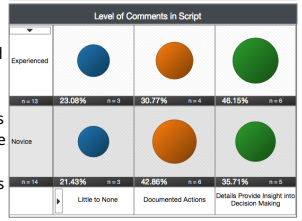
QC Challenges

- Experienced users: **DECISION MAKING** was the most challenging aspect of performing quality control.
- Novice users: **UNFAMILIARITY** with the process and data was the most challenging aspect.

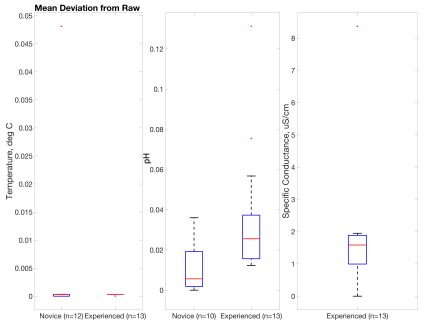


Script Detail

- Comments in scripts generated by participants can provide insight into QC decisions.
- The level of detail in comments was similar between the novice and experienced users.
- Level of comment detail affects the reproducibility of results.



Deviation from Raw Data



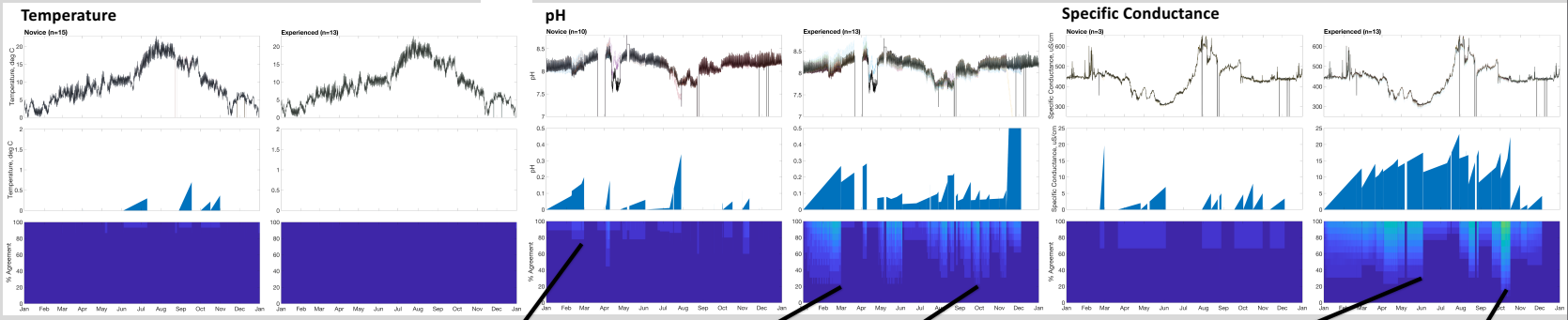
- To what degree did participants alter data?
- Compare to sensor precision:

Temp: ±0.01 degC **pH: ±0.1** **SpCond: ±2 uS/cm**

Variation in Corrections

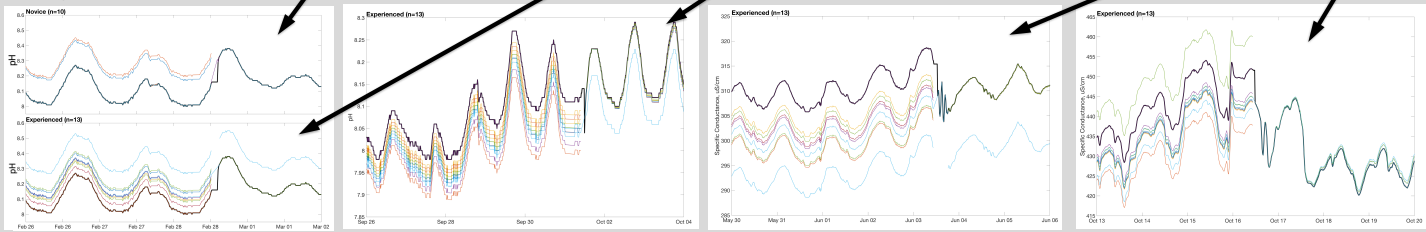
Raw Data — QC Data by Individual Participants

Agreement between participants indicates what percentage of participants agreed on the same result (within ±0.01) for each time stamp



Drift Correction Variability

- Field calibration events require retrospective drift corrections.
- While most participants used the same method of correction, there are discrepancies in the gap values determined by each participant.
- Drift correction was implemented by all experienced participants vs. only a few novice participants.



Ranges of Linear Drift Correction Gap Values:

0.2 (sensor precision ±0.1) | 0.1 (sensor precision ±0.1) | 18 uS/cm (sensor precision ±2 uS/cm) | 23 uS/cm (sensor precision ±2 uS/cm)

Conclusions

- There is a **surprising amount of variation** in resulting datasets
- **Greater degree of changes by the experienced group** vs. the novice group: both in deviation from raw and overall variability.
- **Experienced group more willing to alter data** vs. novice.
- **Degree of agreement higher for variables** that do not undergo calibration and do not need drift correction (e.g., temperature).
- **Consistency at calibration events** when users reference fixed calibration points.
- There is high degree of agreement and the overall deviation from raw **within the range of sensor precision .**
- However, the **greatest differences** between users (for linear drift correction) are **outside the range of sensor precision.**

Variability in different users' QC results may not matter if within the range of sensor precision. If high precision is important, consider a process to support collaborative QC decisions.