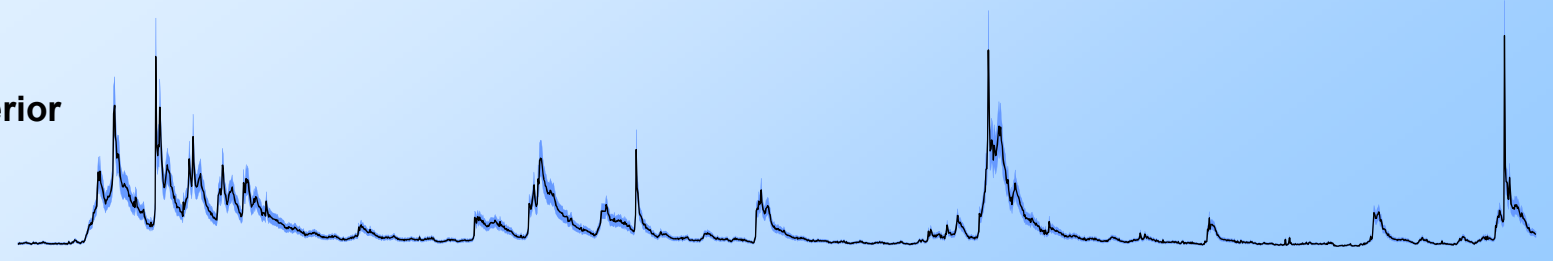
USGS
science for a changing world

Methods and Examples to Quantify Uncertainty Associated with USGS Time-Series Water-Quality Data

Stewart Rounds
USGS
Oregon Water Science Center
Portland, OR

U.S. Department of the Interior
U.S. Geological Survey

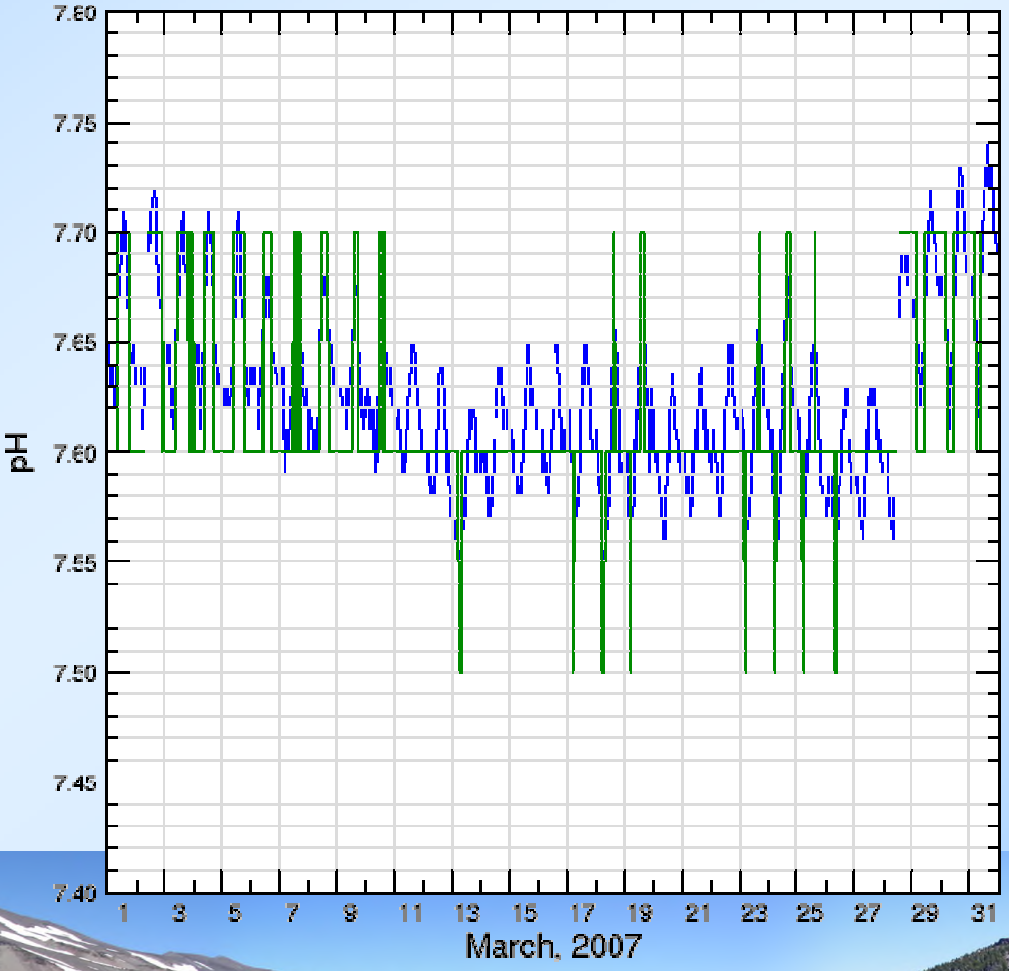


The Vision: Include Uncertainties with Data

Now:

- **Basic time-series graphs from USGS do *not* show uncertainties**
- **Data downloads are:**

  Date, Time, Time Zone, Value, Flag/Remark



Raw data

Rounded data

Rounding is a poor means of expressing uncertainty.

We can do BETTER!

data from USGS

USGS

The Vision: Include Uncertainties with Data

Now:

- **Basic time-series graphs from USGS do *not* show uncertainties**
- **Data downloads are:**

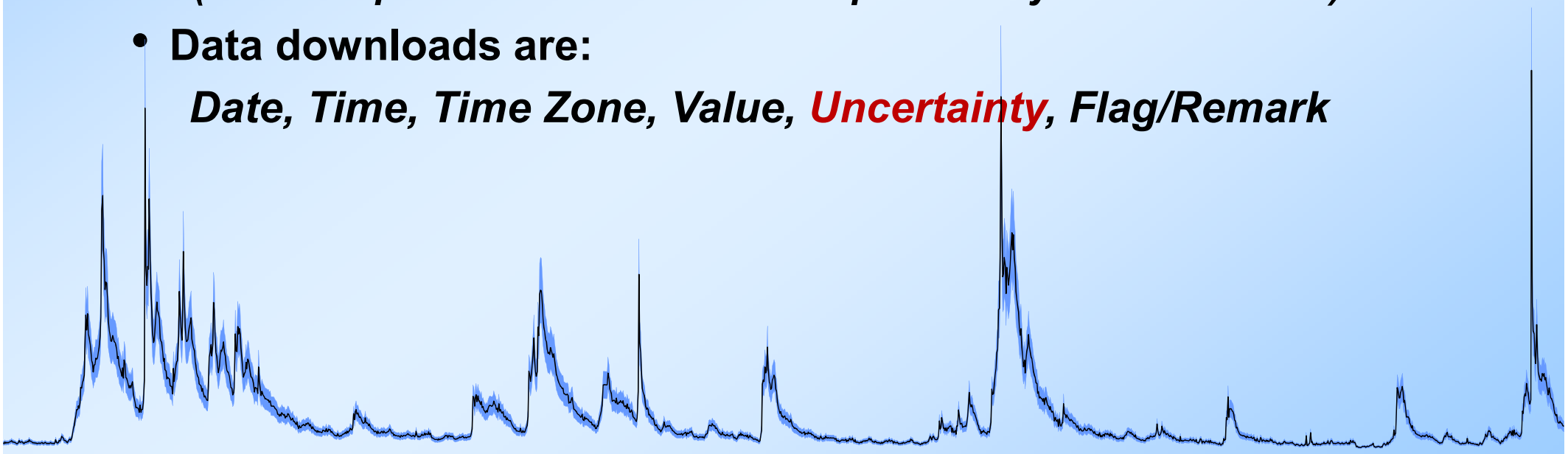
  Date, Time, Time Zone, Value, Flag/Remark

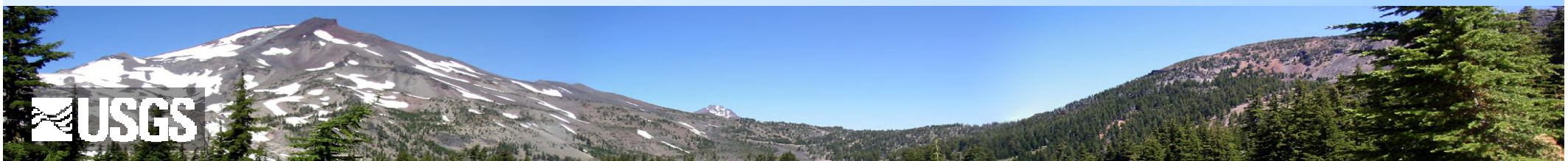
Future:

- **Time-series graphs from USGS show uncertainty bands**

  (with comparison to standards and probability of exceedance)
- **Data downloads are:**

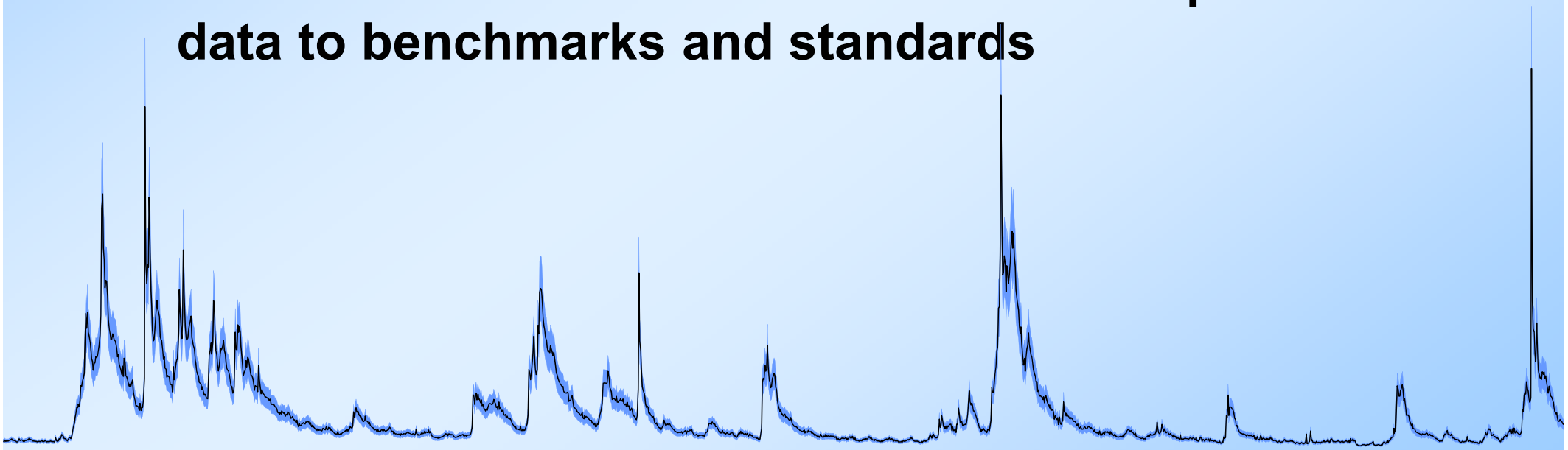
  Date, Time, Time Zone, Value, Uncertainty, Flag/Remark





Why Do We Care About Data Uncertainty?

1. Need to quantify uncertainty in order to manage it
2. Need to communicate the accuracy of datasets
   → Improves acceptance and appropriateness of use
3. Would improve and extend the value and applicability of data
4. Would create a better framework for comparisons of data to benchmarks and standards



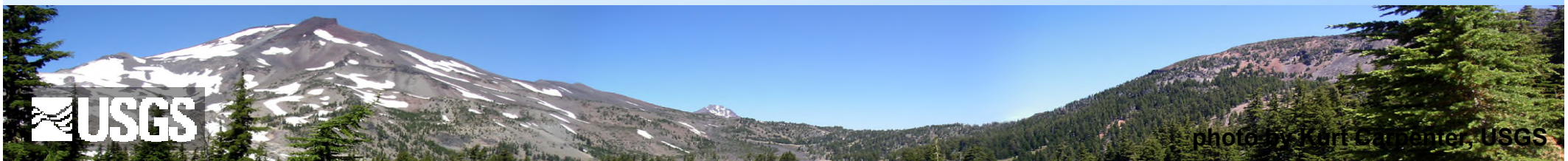


photo by Kurt Carpenter, USGS

Analysis Must Consider All Sources of Error

Measurement Error

- The difference between a measured value and a known reference or true value
- Composed of random and predictable components
  - Predictable components are repeatable and correctable
- Many sources of error
  - Avoidable Errors

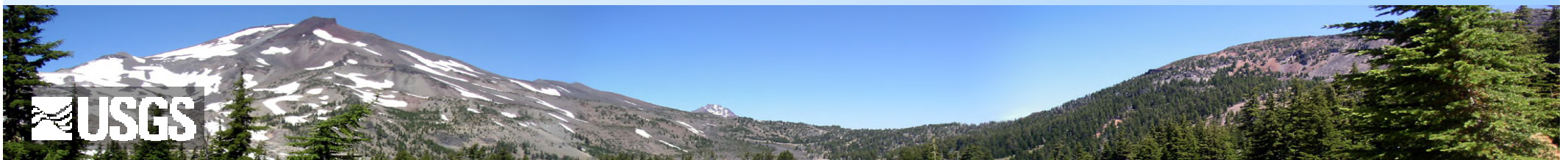
    User error, site-selection errors, site-installation errors
  - Unavoidable Errors (at some level)

    Signal noise/variability, site variability, interferences

    Probe fouling, probe failure, calibration drift

    Deviation from lab verification samples or other references

    Surrogate model error



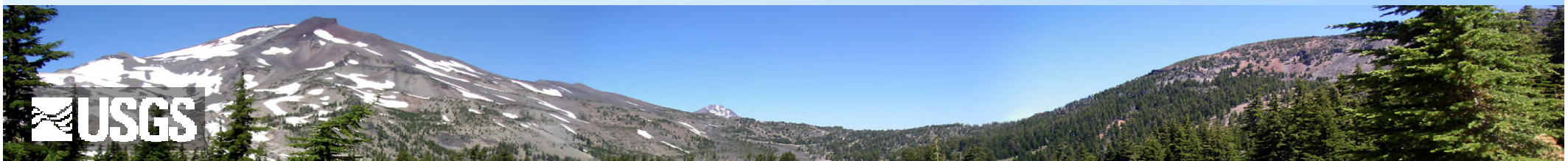
Uncertainty is Derived from Error Analysis

Measurement Error

- The difference between a measured value and a known reference or true value
- Composed of random and predictable components
  - Predictable components are repeatable and correctable
- Many sources of error

Uncertainty

- A description of the degree of accuracy of the final corrected data
  - Can be expressed in many ways, including statistical representations (std. dev.) or a simple half-width of a probable interval



Methods to Quantify Uncertainty

Rigorous statistical approach

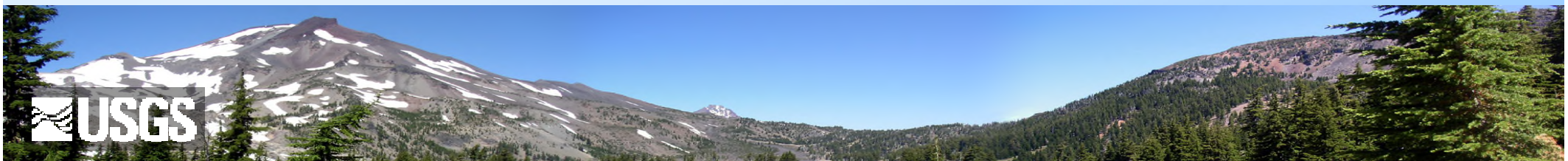
- The GUM— "Guide to the Expression of Uncertainty in Measurement" by the Joint Committee for Guides in Metrology
- Strict combination of variance approach w/ correlation coefficients and dependent cross terms

Root Mean Square Error approach

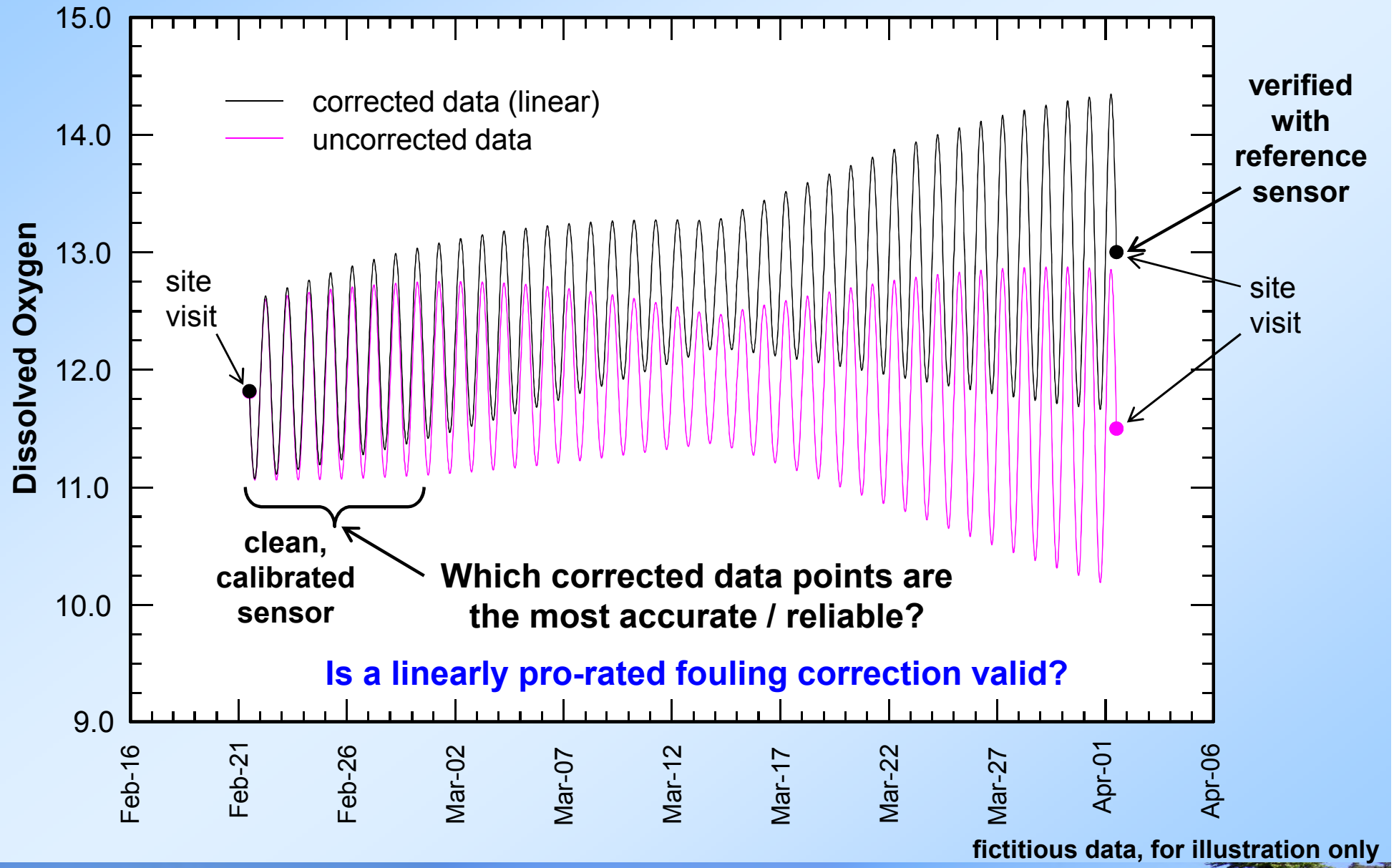
- Simpler method of accounting for all sources of error
- Estimates a most probable value of the cumulative error
- Assumes independence of error sources
  (note the lack of correlation coefficients and cross-terms)

$$E_P = \sqrt{E_1^2 + E_2^2 + \cdots + E_n^2} = \sqrt{\sum_i E_i^2}$$

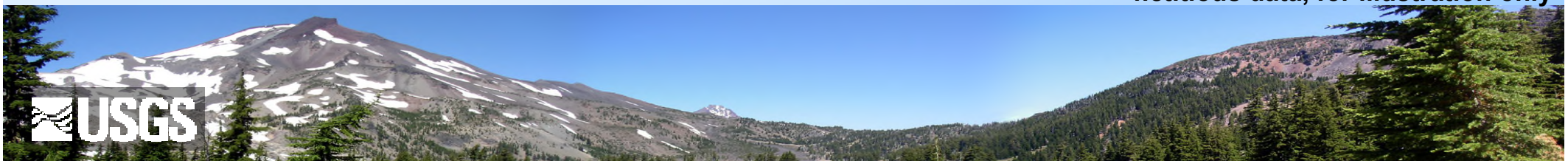
E_P is the combined probable error; E_i are the component errors



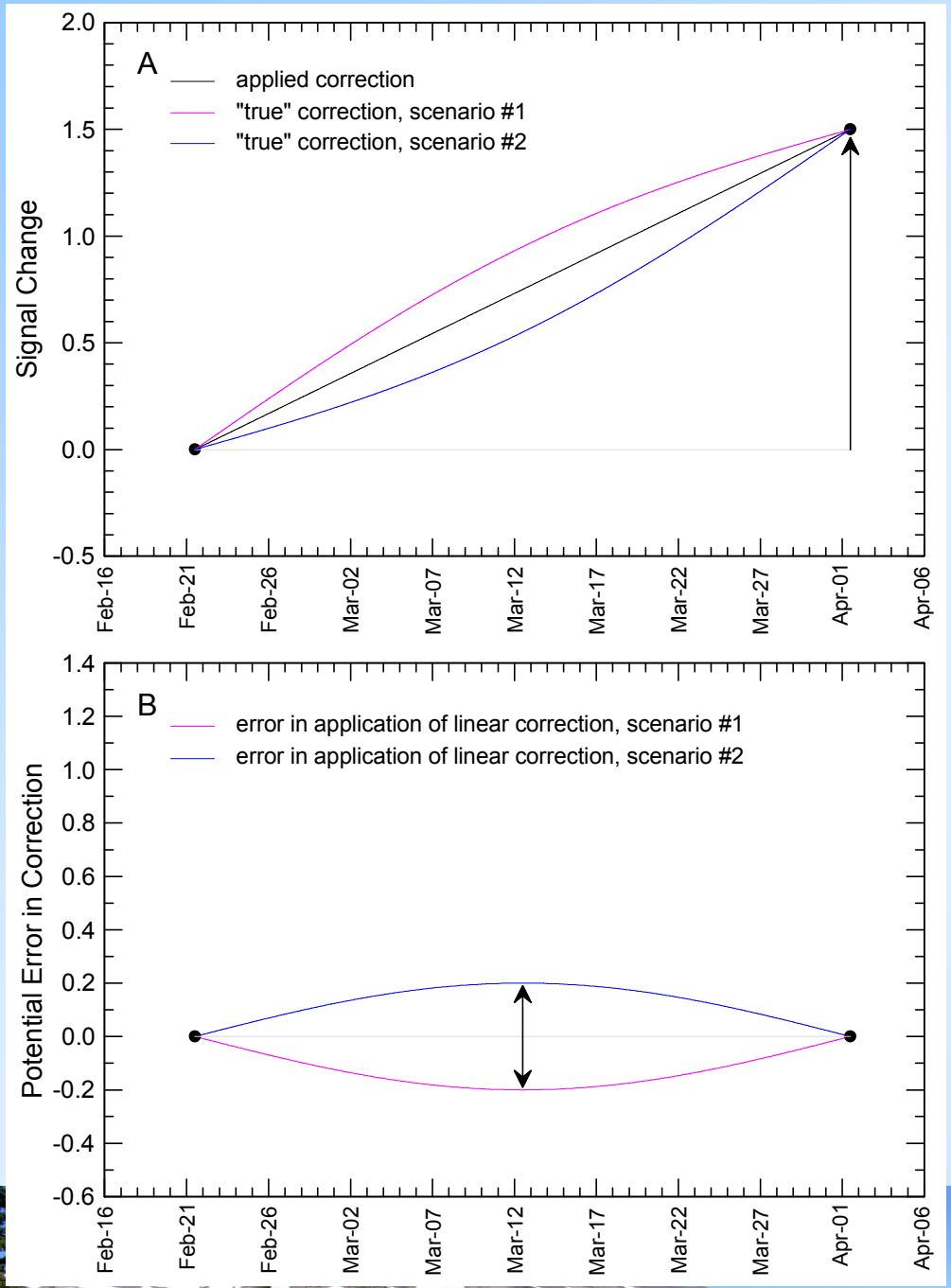
Focus on Assumptions of Fouling Corrections



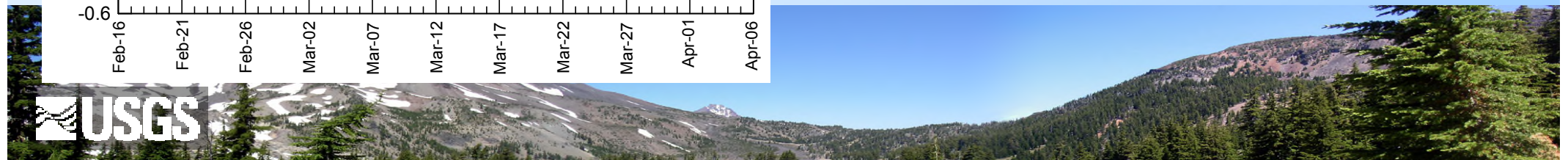
fictitious data, for illustration only



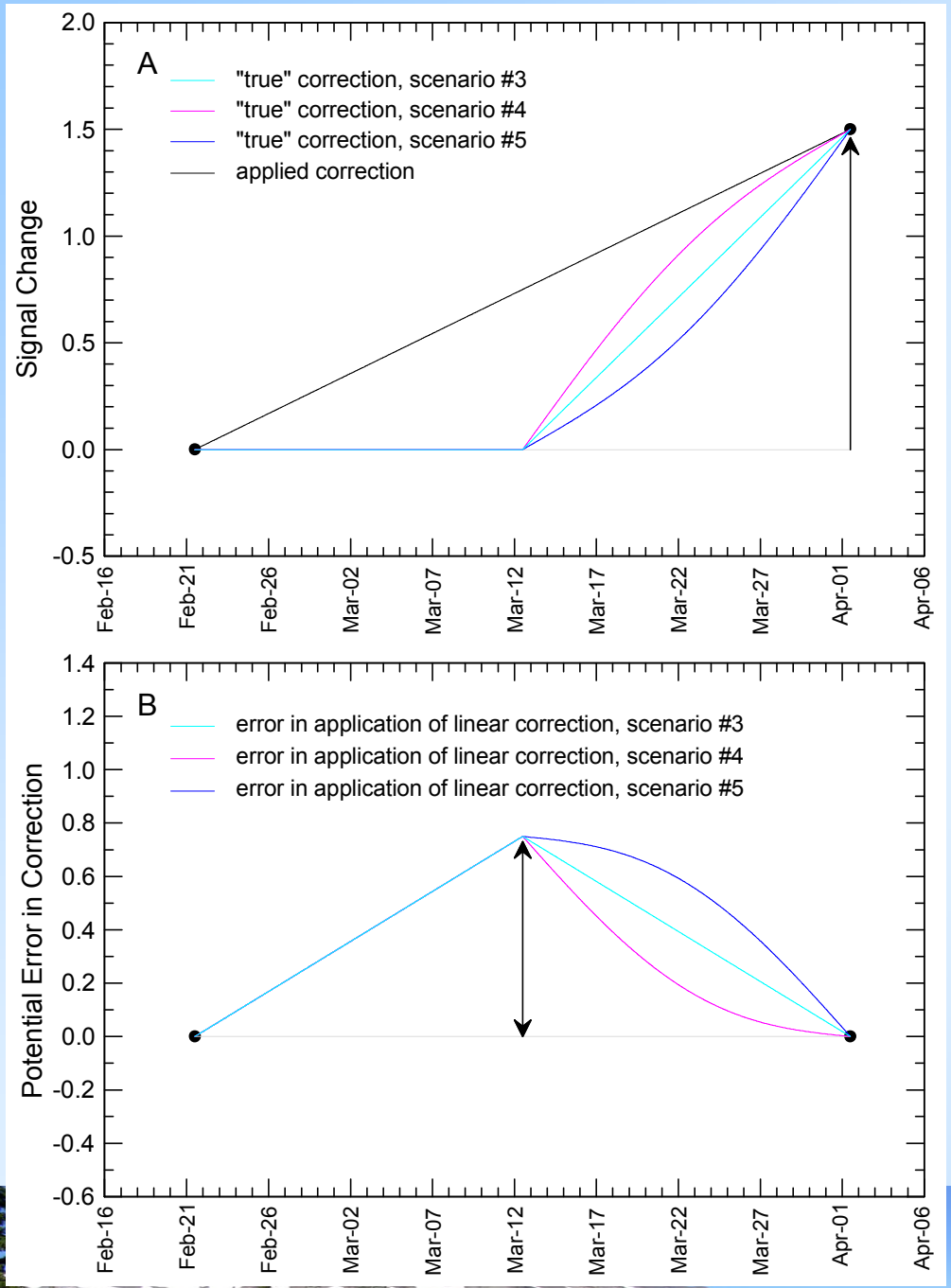
Potential Errors in Application of Corrections



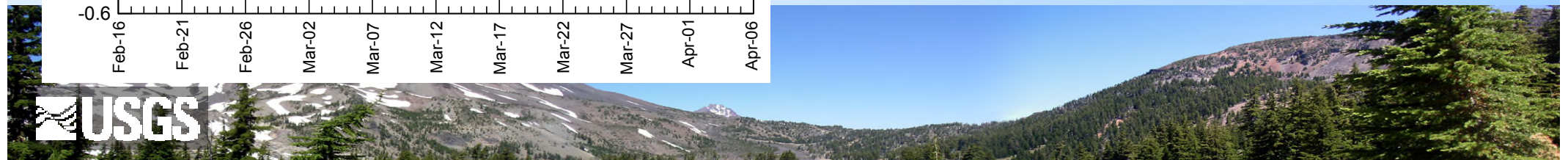
- **Error only exists if the applied correction is different from the "true" correction**
- **Errors may arise from the assumed functional shape of a correction**



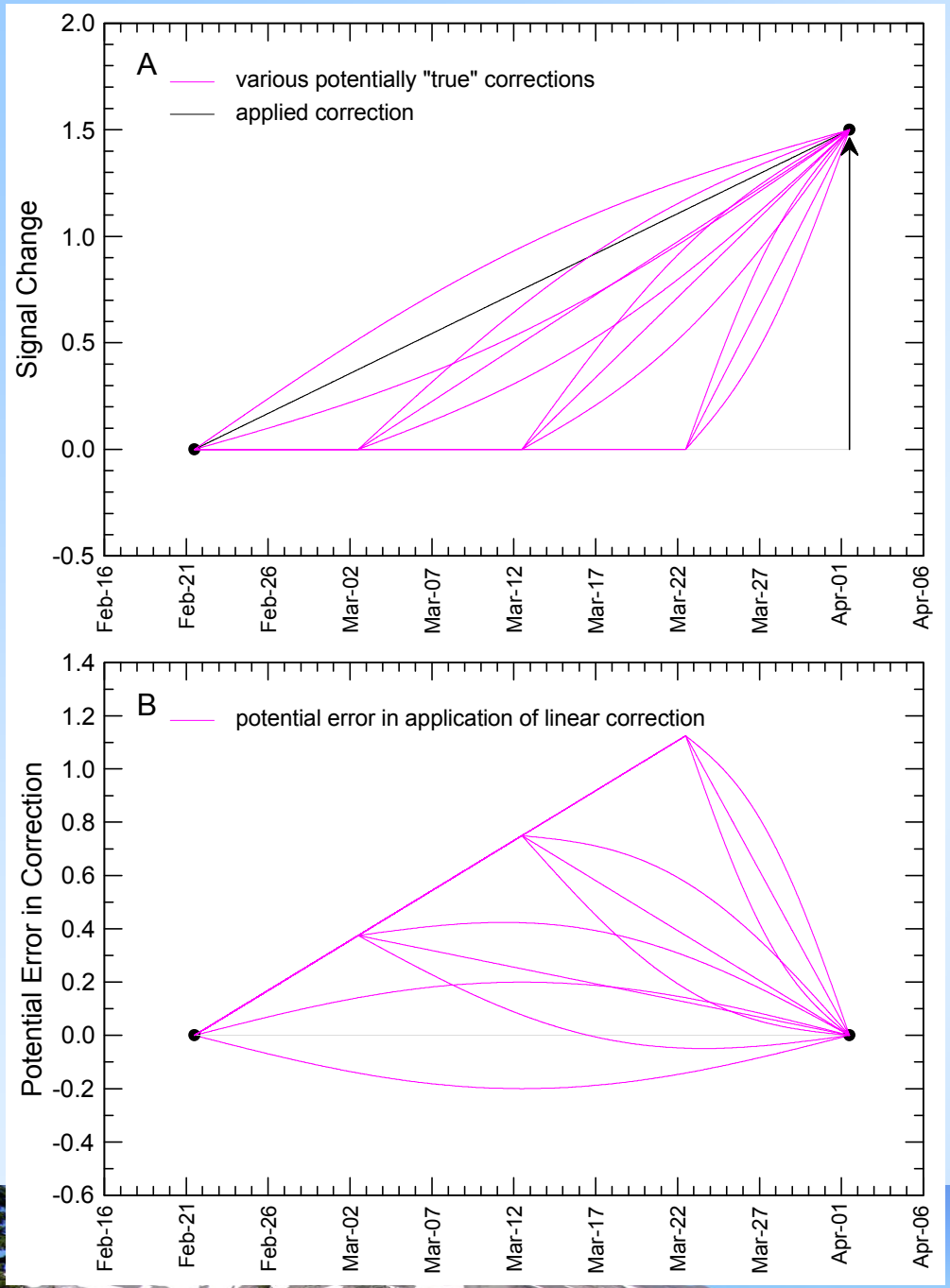
Potential Errors in Application of Corrections



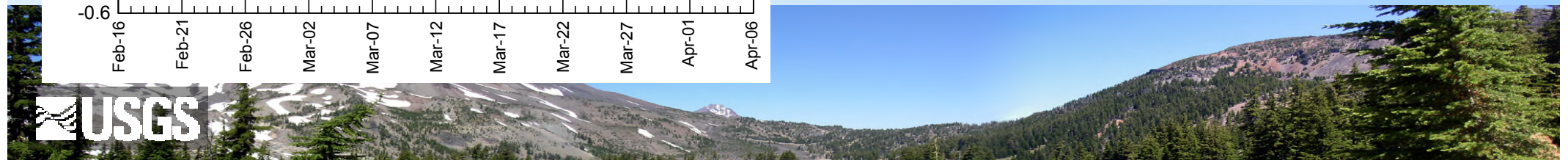
- **Error only exists if the applied correction is different from the "true" correction**
- **Errors may arise from the assumed functional shape of a correction**
- **Errors may arise from the assumed timing of a correction**



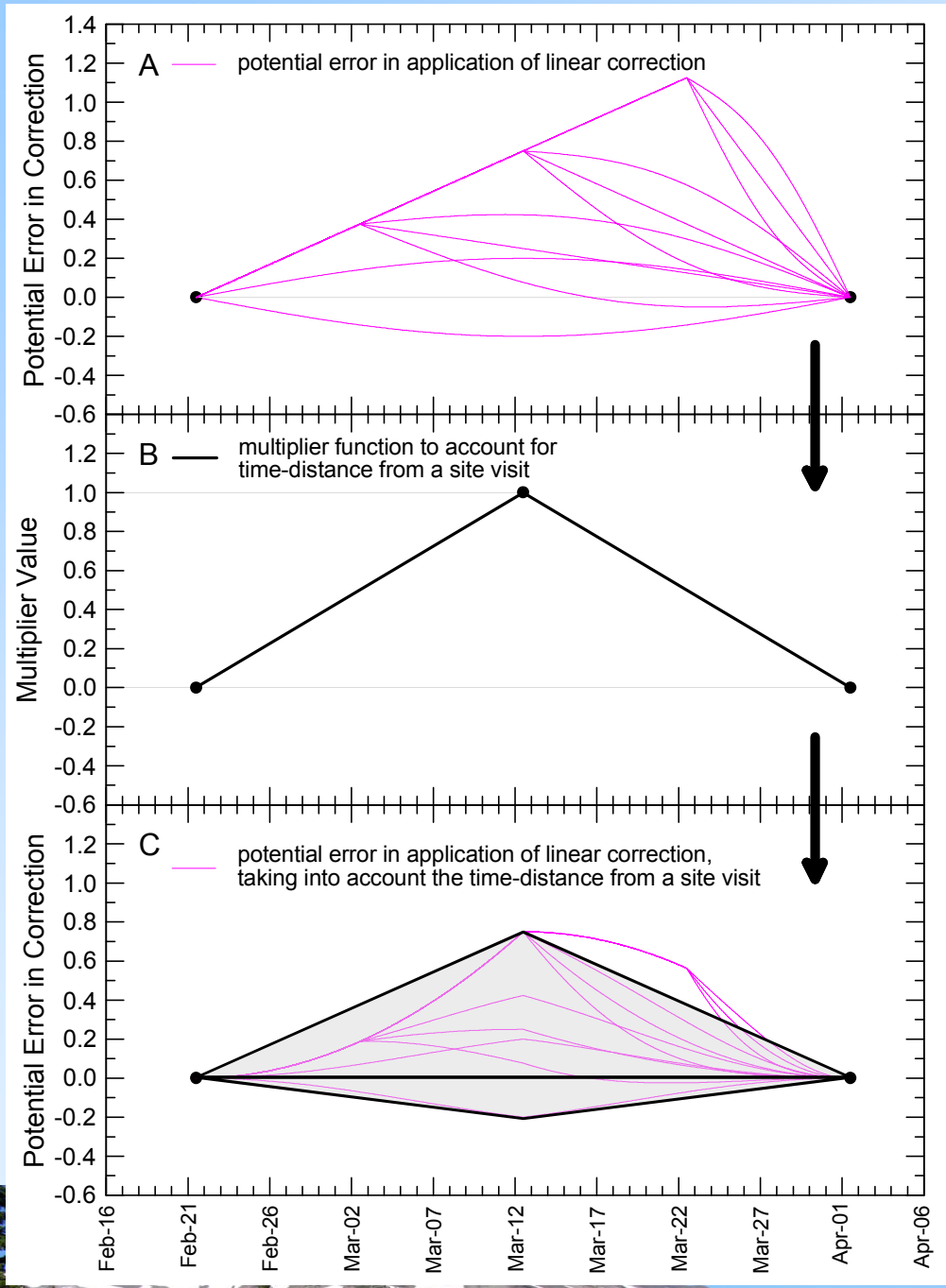
Potential Errors in Application of Corrections



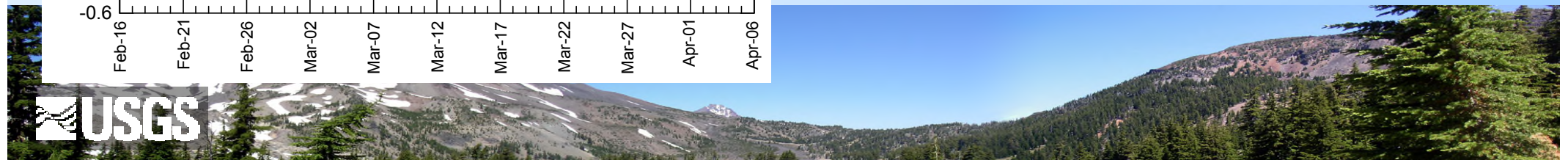
- **A range of potential errors in the timing and functional shape of a correction begins to form a pattern**



Potential Errors in Application of Corrections



- **A range of potential errors in the timing and functional shape of a correction begins to form a pattern**
- **Applying a triangular filter represents the greater likelihood that:**
  1. **corrected data are more accurate if they are closer to a site visit, and**
  2. **the true correction is not as likely to be initiated right before the second site visit**



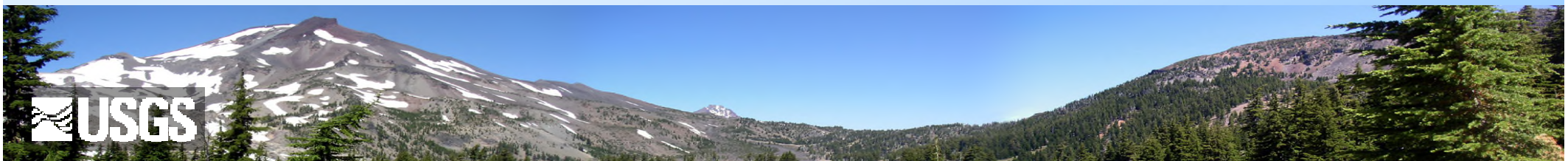
Assessment of Fouling Assumptions

Wiped vs Unwiped SC and DO sensors

- YSI EXO and 6-series sensors
- Fanno Creek near Portland, OR

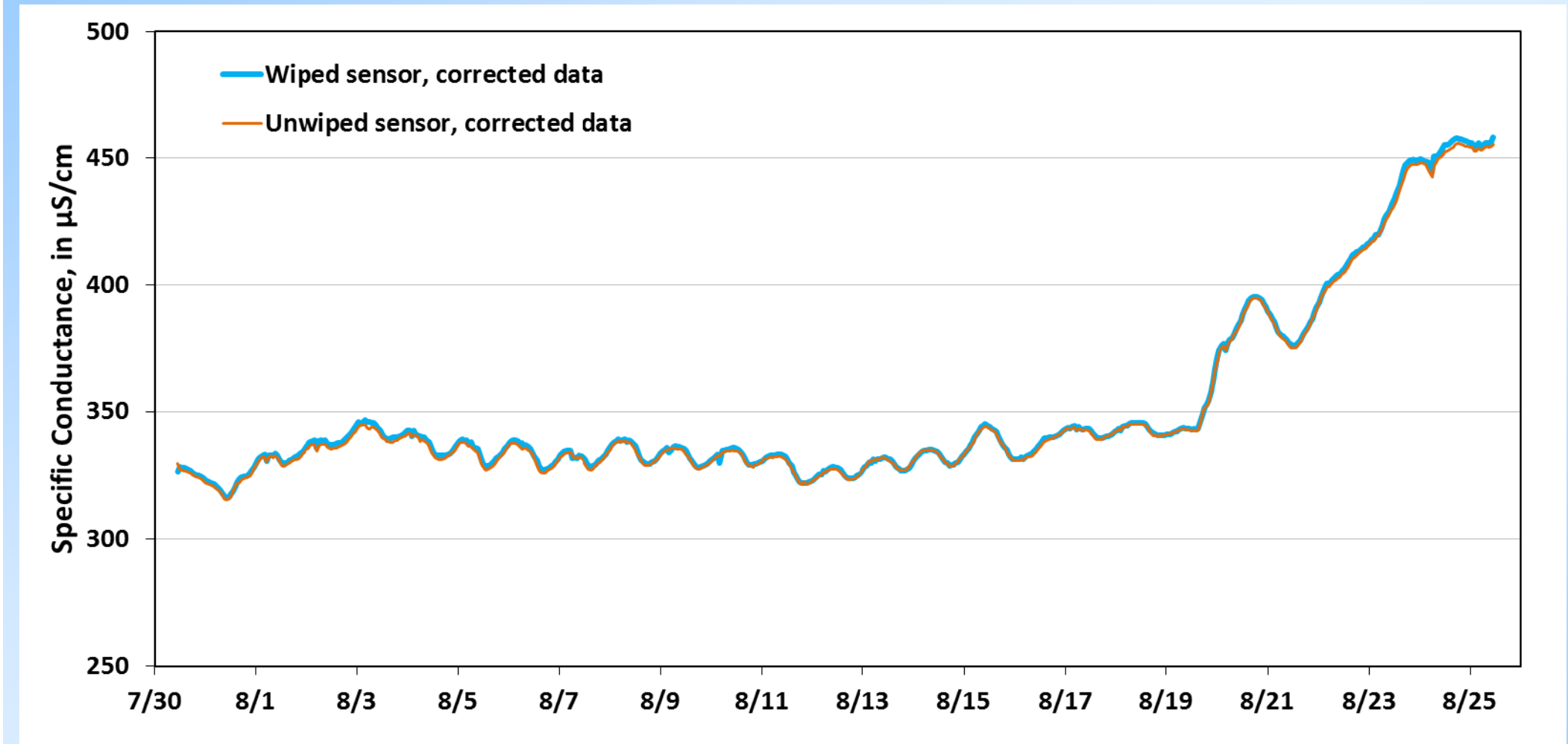
photo from USGS

sensor images from YSI;
not a product endorsement



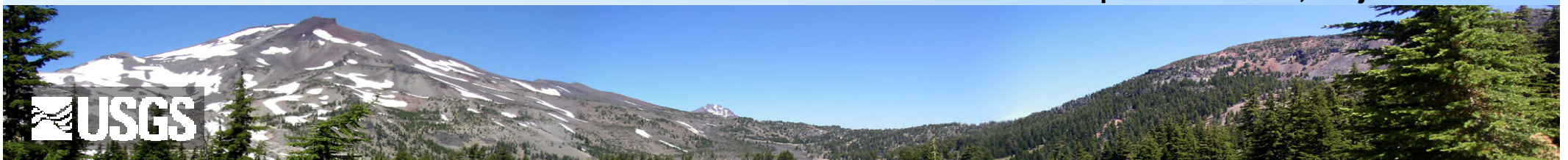
Wiped vs Unwiped SC Data

Fanno Creek at Durham Road (14206950)



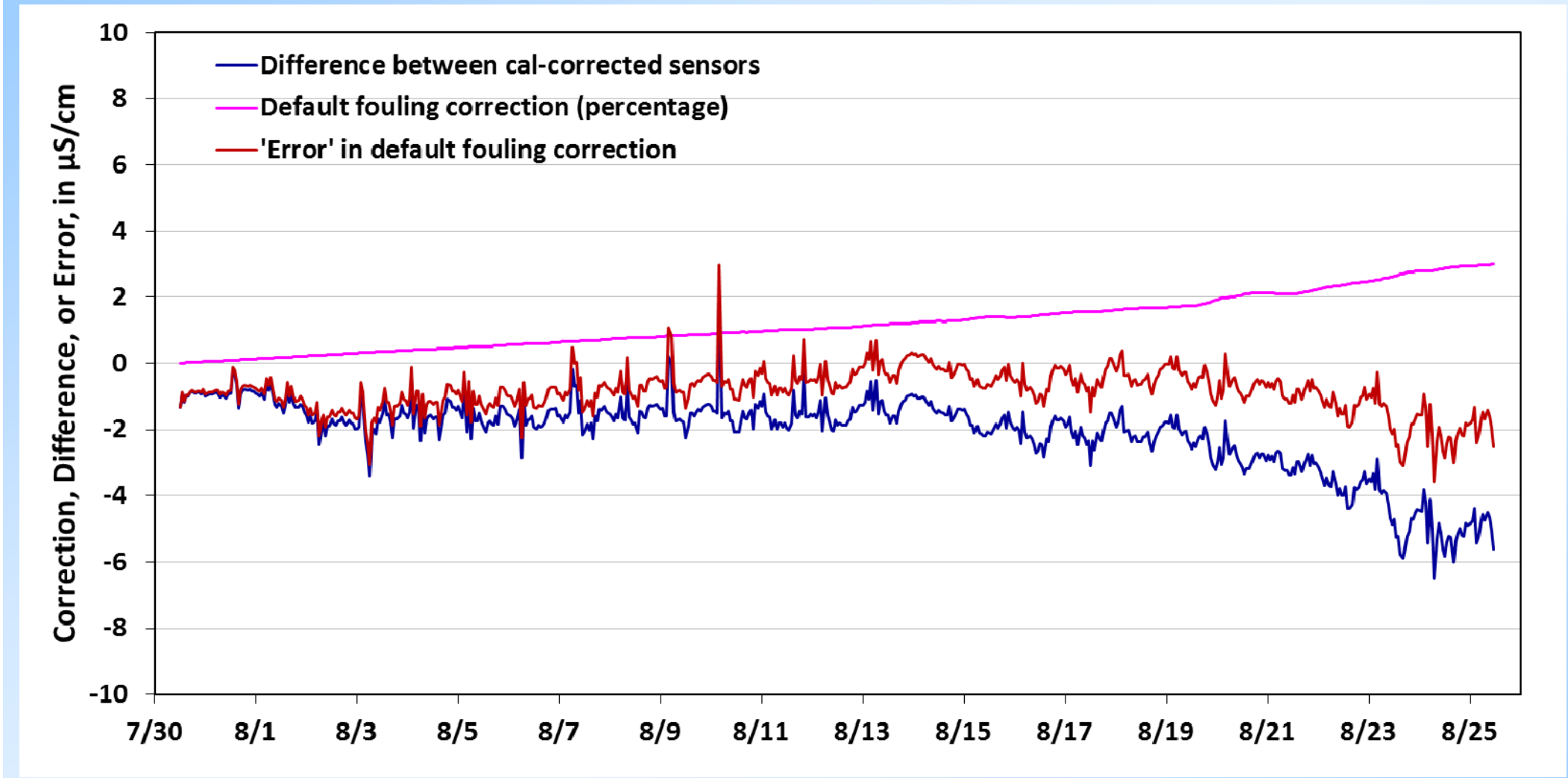
Date in 2015

provisional data, subject to revision



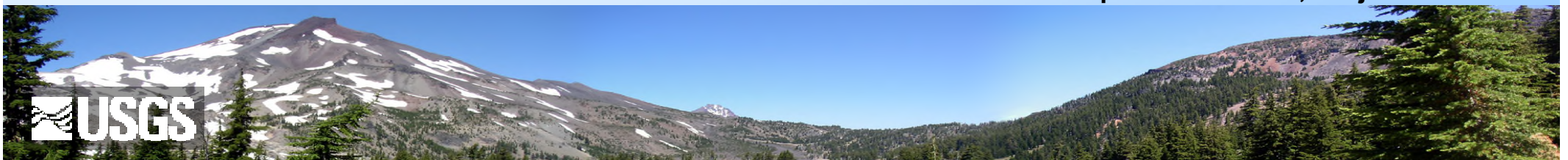
Wiped vs Unwiped SC Data

Fanno Creek at Durham Road (14206950)

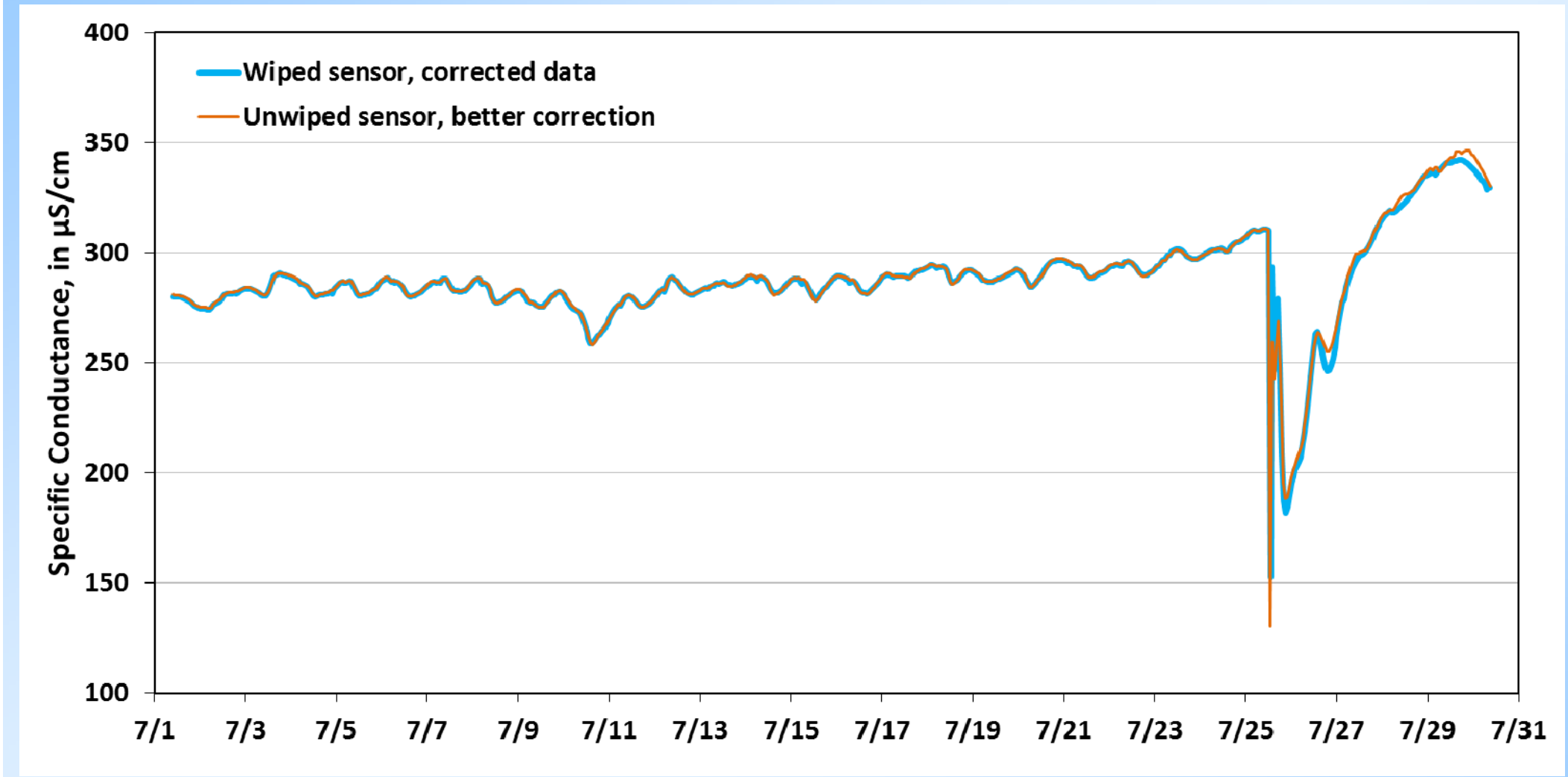


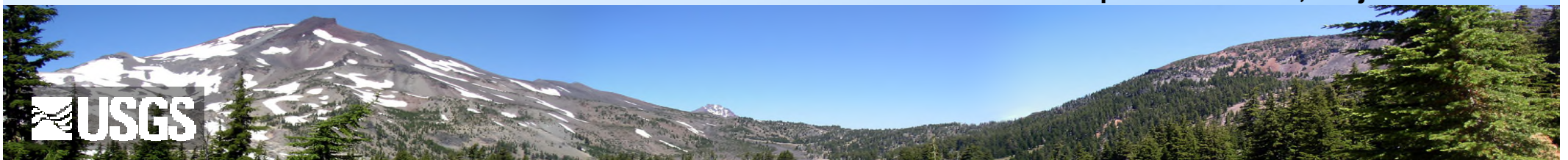
Date in 2015

provisional data, subject to revision

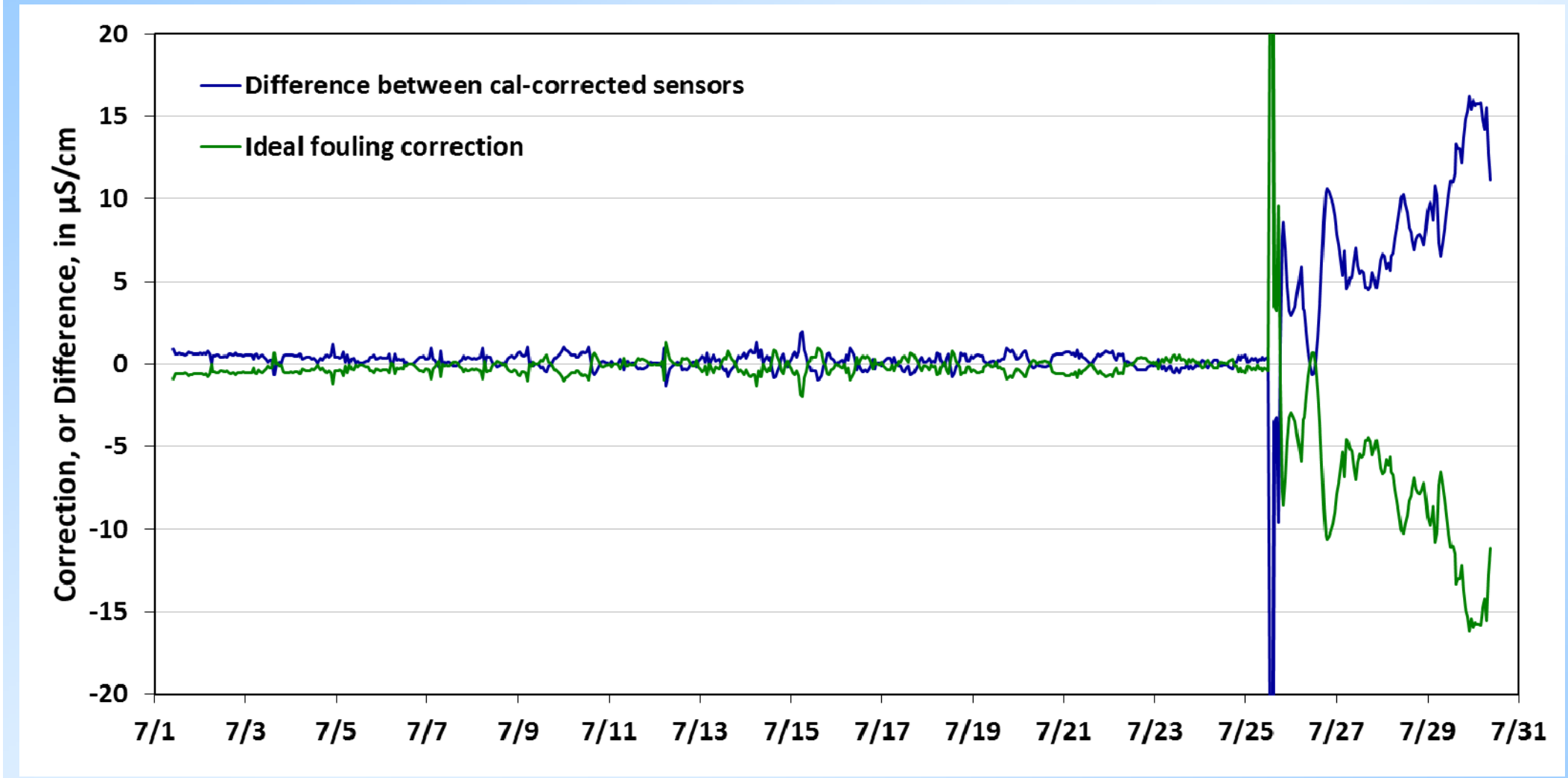


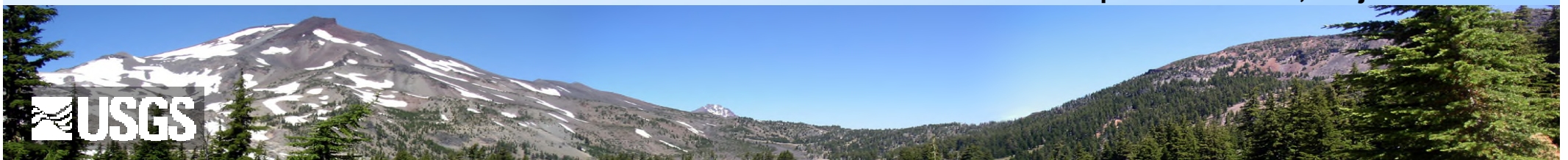
Example 2: Wiped vs Unwiped SC Data





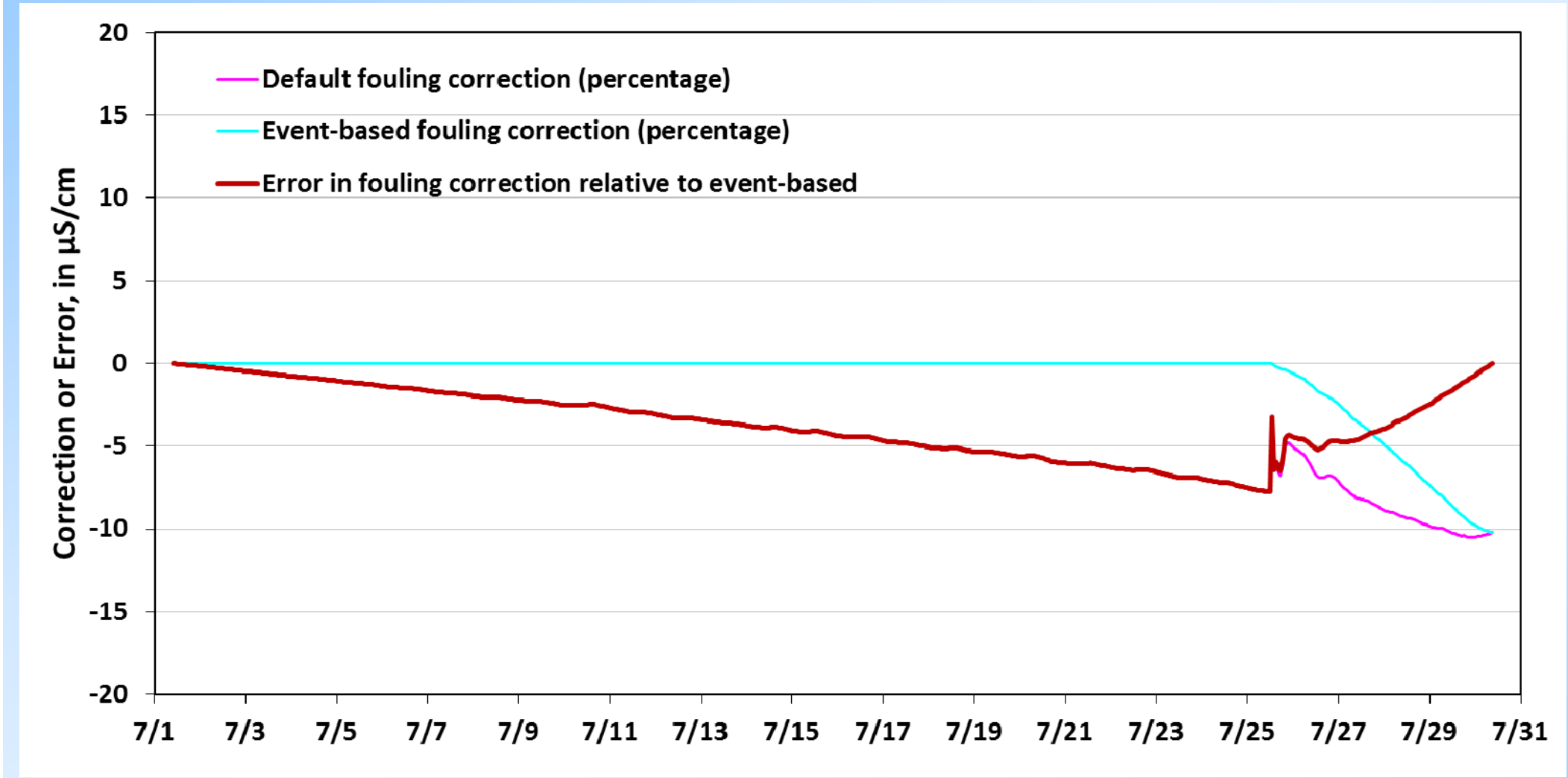
Example 2: Wiped vs Unwiped SC Data





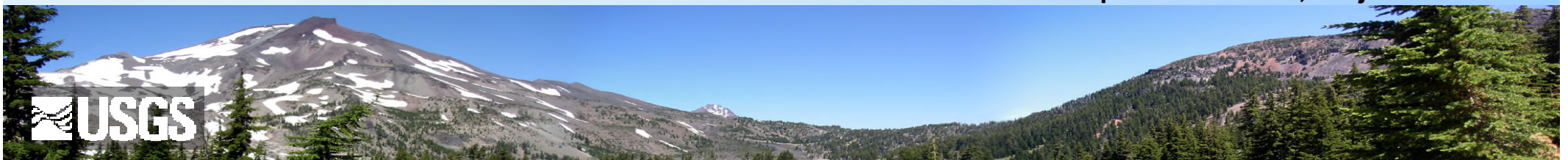
Example 2: Wiped vs Unwiped SC Data

Fanno Creek at Durham Road (14206950)



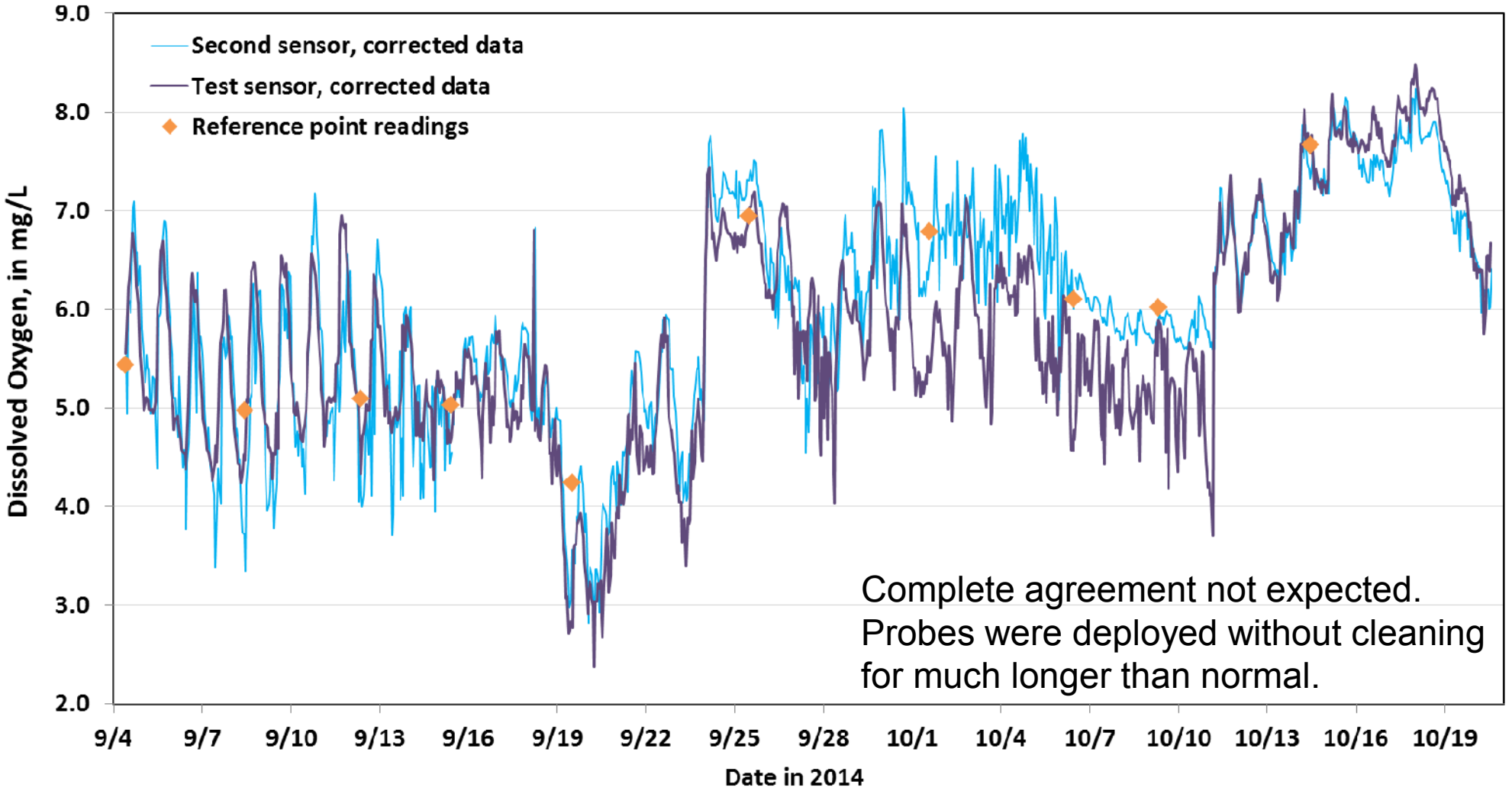
Date in 2015

provisional data, subject to revision

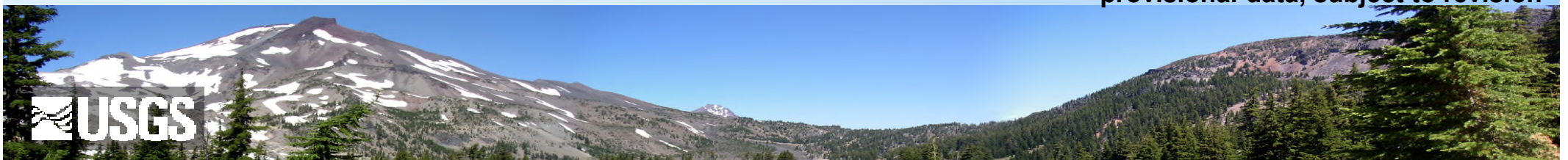


Example 3: Wiped vs Unwiped DO Data

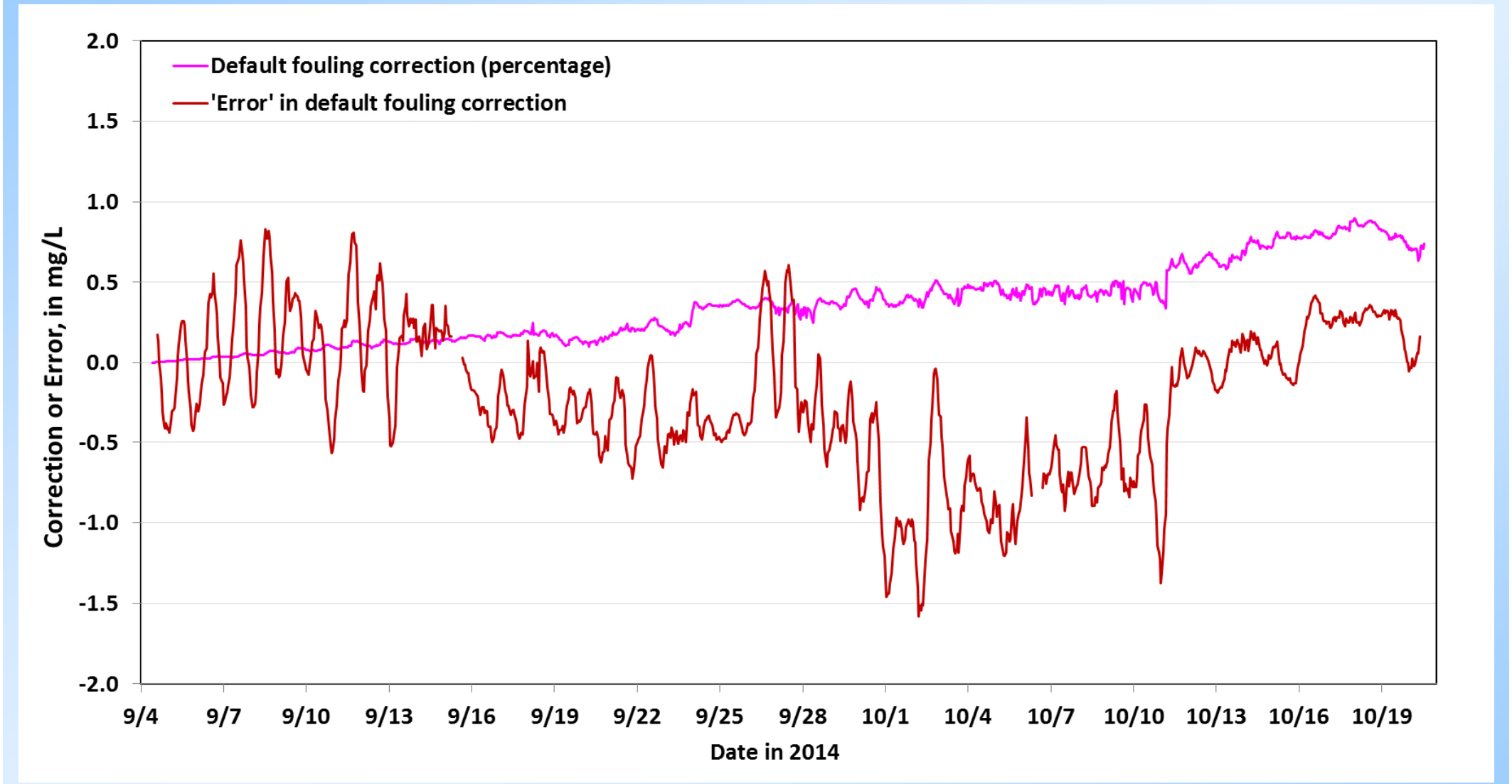
Fanno Creek at Durham Road (14206950)



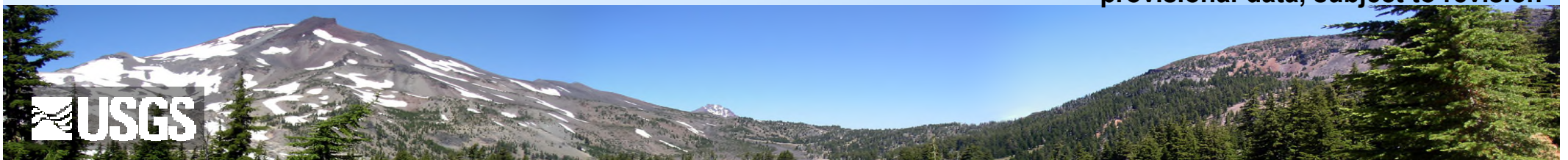
provisional data, subject to revision



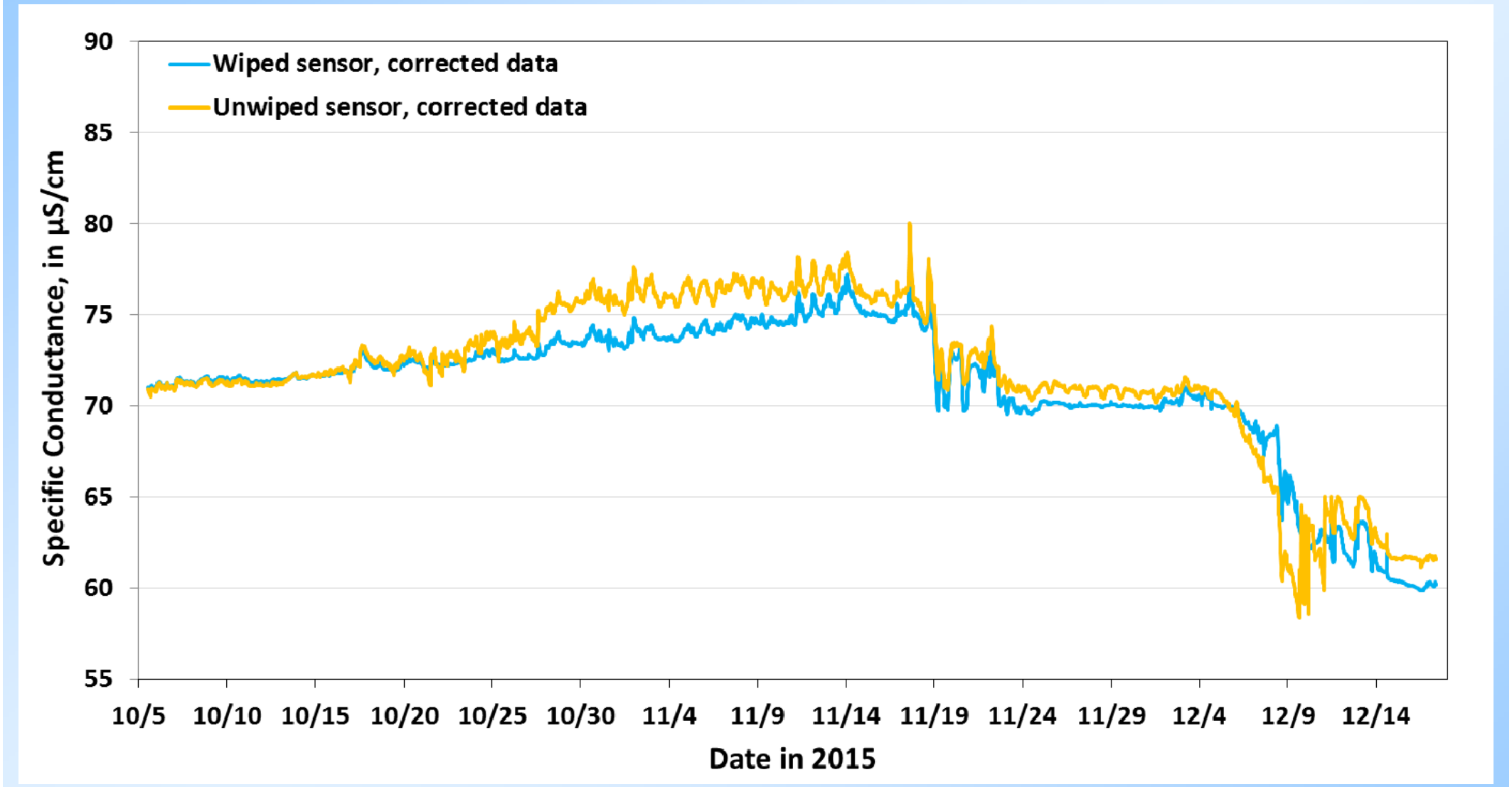
Example 3: Wiped vs Unwiped DO Data



provisional data, subject to revision



Example 4: Wiped vs Unwiped SC Data



provisional data, subject to revision

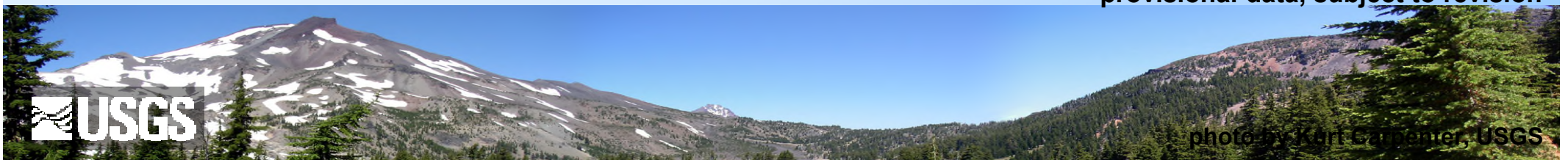
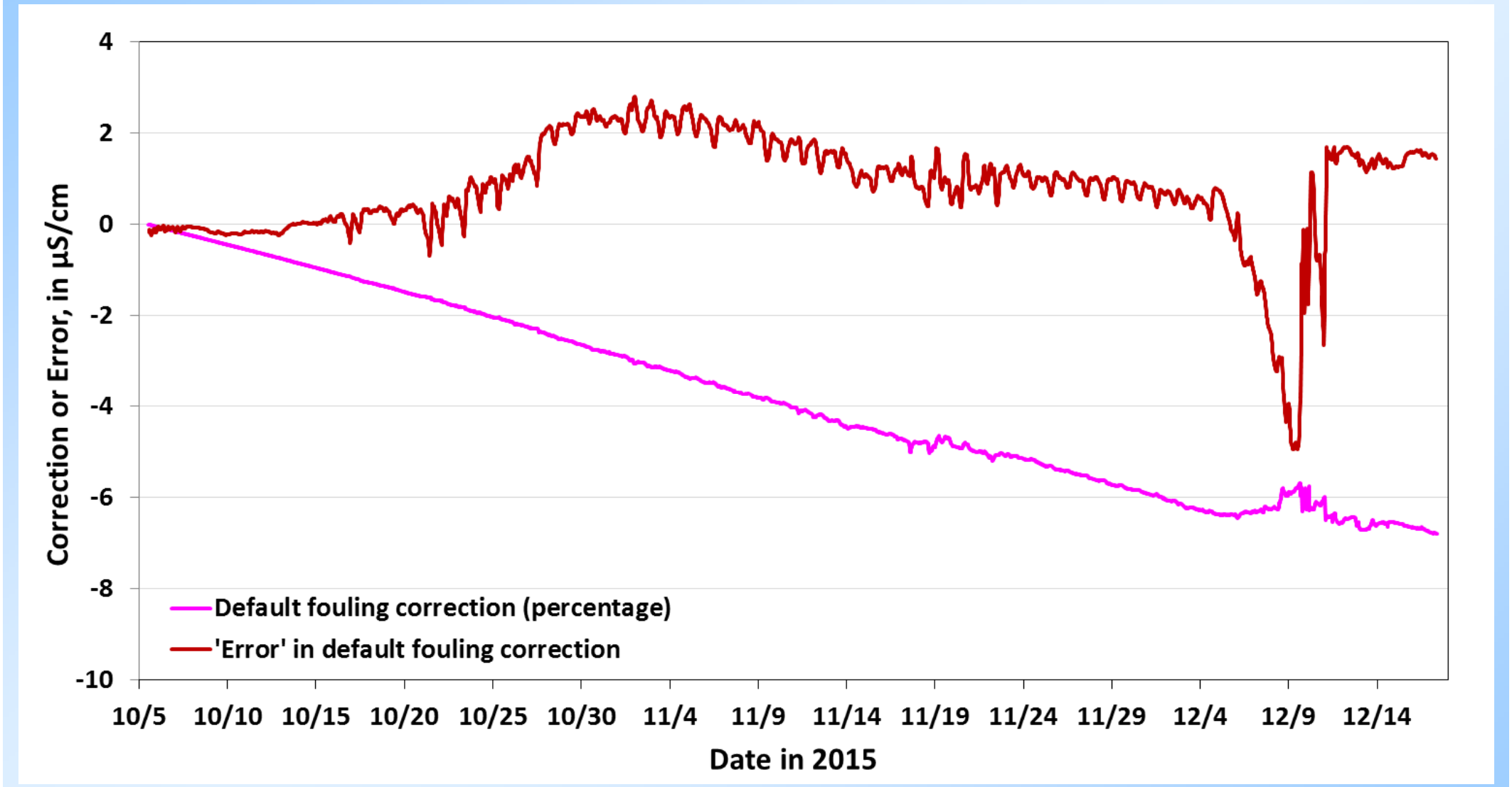


photo by Kurt Carpenter, USGS

Example 4: Wiped vs Unwiped SC Data



provisional data, subject to revision

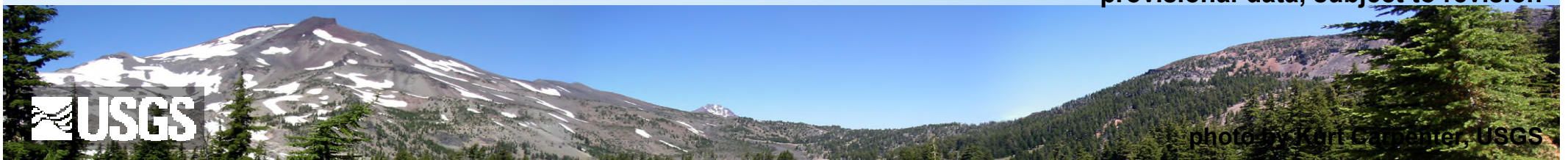
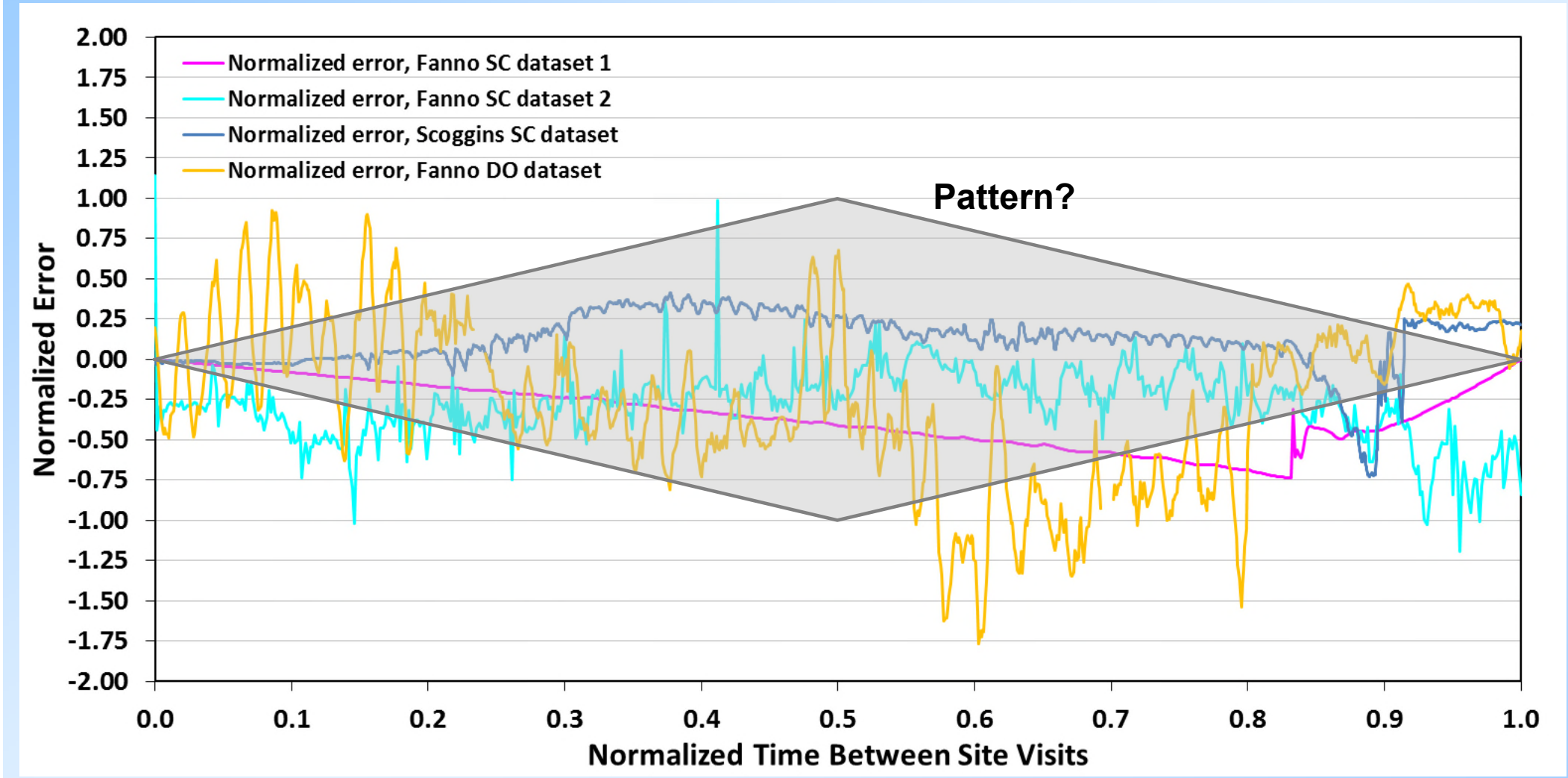
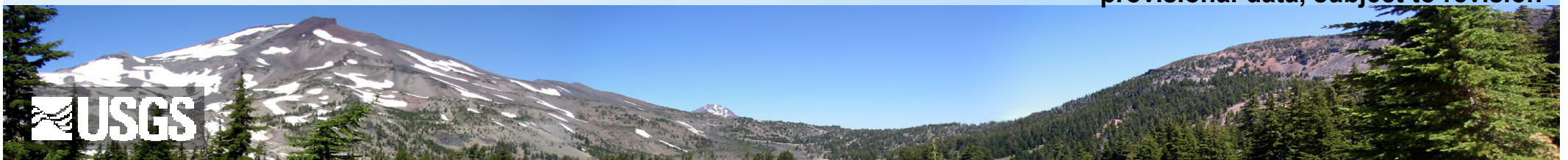


photo by Kurt Carpenter, USGS

Errors Normalized to Maximum Fouling Correction



provisional data, subject to revision



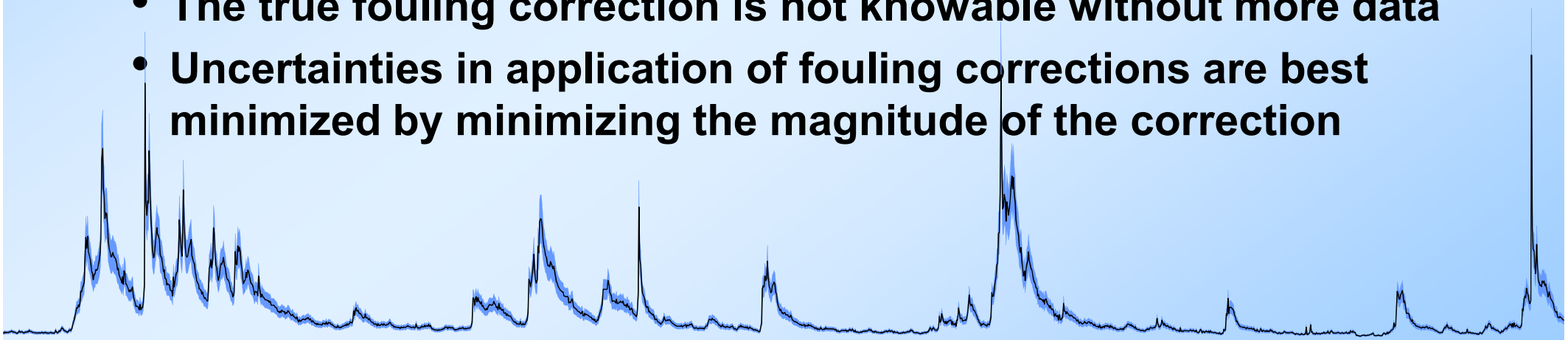
Take-Home Messages

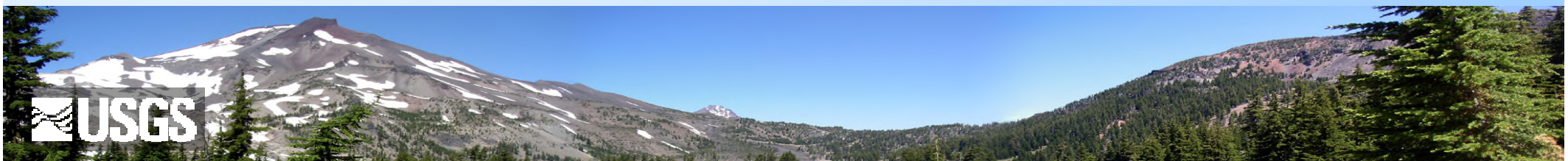
Methods/Guidelines under development:

- Rounding is not a valid method of expressing uncertainty
- Statistical approach needed to combine error estimates
- Avoid measurement errors, if possible
- Some errors unavoidable, and related to our assumptions

Fouling not always linear over time:

- Hydrologic events sometimes dictate initiation of fouling
- Linear pro-ration is a good first-cut assumption
- The true fouling correction is not knowable without more data
- Uncertainties in application of fouling corrections are best minimized by minimizing the magnitude of the correction





Contact, and Thanks

Stewart Rounds

USGS Oregon Water Science Center
sarounds@usgs.gov
503-251-3280

Help and Insights from Data Uncertainty Team:

USGS:	Stacey Archfield, Zach Freed, Janice Fulford, David Holtschlag, Brian Pellerin, Pat Rasmussen, Susan Wherry
YSI / Xylem:	Rob Ellison
Aquatic Informatics:	Stu Hamilton, Brian Gouge