Re-evaluating the Triple-Colocation Analysis for Estimating Aquarius Satellite Salinity Error

Gary Lagerloef, Hsun-Ying Kao, David Carey
Earth and Space Research, Seattle, WA

2019 Salinity Continuity Workshop
29-30 April 2019, Santa Rosa, CA
Outline

• Triple-point Analysis (TPA): An *error variance calculation* based on a set of simultaneous co-located measurements from three different systems (e.g. satellite, in situ and numerical analysis). [Stoffelen, 1998]

• For Aquarius, we have used a variation of Stoffelen’s approach.

• Each approach derives a unique error variance of each of the three measurements.

• Each also assumes that the errors of each measurement system are *un-correlated* with one-another.

• Here, we seek to verify if this assumption is true; and consider the consequences if not.
Statement of the Problem

Aquarius L2 data co-located data:

1) Aquarius (AQ)
2) In Situ (IS)
3) Hycom (HY)

$S_T$ is true SSS

$\varepsilon_n$ is the measurement error for n

\[
\begin{align*}
S_1 &= ST + \varepsilon_1 \\
S_2 &= ST + \varepsilon_2 \\
S_3 &= ST + \varepsilon_3
\end{align*}
\]

(1)

\[
\begin{align*}
S_1 - S_2 &= \varepsilon_1 - \varepsilon_2 \\
S_2 - S_3 &= \varepsilon_2 - \varepsilon_3 \\
S_1 - S_3 &= \varepsilon_1 - \varepsilon_3
\end{align*}
\]

(2)

\[
\begin{align*}
\Delta S_{12} &= \Delta \varepsilon_{12} \\
\Delta S_{23} &= \Delta \varepsilon_{23} \\
\Delta S_{13} &= \Delta \varepsilon_{13}
\end{align*}
\]

\[
\langle \Delta S_{ij}^2 \rangle = \langle \varepsilon_i^2 \rangle + \langle \varepsilon_j^2 \rangle
\]

\[
\langle \varepsilon_{ij}^2 \rangle = 0 \quad \text{for } i \neq j
\]

Assume error cross-correlations are zero
\[ <\Delta S_{12}^2> = <\varepsilon_1^2> + <\varepsilon_2^2> \]
\[ <\Delta S_{23}^2> = <\varepsilon_2^2> + <\varepsilon_3^2> \]
\[ <\Delta S_{13}^2> = <\varepsilon_1^2> + <\varepsilon_3^2> \]

Closed set of 3 equations with 3 unknowns \(<\varepsilon_1^2>, <\varepsilon_2^2>, <\varepsilon_3^2>\)

\[ <\varepsilon_{12}>= <\varepsilon_{23}>= <\varepsilon_{13}>= 0 \]

3 equations to compute \(<\varepsilon_i^2>\) from the observed \(<\Delta S_{ij}^2>\).

These are the equations that we have used for *Aquarius* triple-point error estimates, Assuming error cross-correlations are negligible.
Revised Triple-Point analysis equations:

- Without correlated errors (as we have been using).
- Additional error correlation terms

\[
<\varepsilon_1^2> = \left( <\Delta S_{12}^2> + <\Delta S_{13}^2> - <\Delta S_{23}^2> \right)/2 \\
+ <\varepsilon_{12}> + <\varepsilon_{13}> - <\varepsilon_{23}>
\]

\[
<\varepsilon_2^2> = \left( <\Delta S_{12}^2> + <\Delta S_{23}^2> - <\Delta S_{13}^2> \right)/2 \\
+ <\varepsilon_{12}> + <\varepsilon_{23}> - <\varepsilon_{13}>
\]

\[
<\varepsilon_3^2> = \left( <\Delta S_{13}^2> + <\Delta S_{23}^2> - <\Delta S_{12}^2> \right)/2 \\
+ <\varepsilon_{13}> + <\varepsilon_{23}> - <\varepsilon_{12}>
\]
Regression Analysis

Data Input

- **DI**
  - AQ
  - IS
  - HY

**Match-up Sequence**
1) Aquarius (AQ)
2) In Situ (IS)
3) Hycom (HY)

**Difference**

\[ S_T \text{ is true SSS} \]

\[ S_1 = ST + \varepsilon_1 \]
\[ S_2 = ST + \varepsilon_2 \]
\[ S_3 = ST + \varepsilon_3 \]

**Data Difference**

- **DD**
  - AQ-IS
  - IS-HY

**Match-up Sequence**

\[ S_1 - S_2 = \varepsilon_1 - \varepsilon_2 \]
\[ S_2 - S_3 = \varepsilon_2 - \varepsilon_3 \]

**Matrix DD lacks any information of \( S_T \)**

\( S_T \) is the “common signal” to be resolved by regression analysis (next slide).

\( \varepsilon_n = S_n - S_T \) is the error of the \( n^{th} \) measurement system.
Test Case and Summary

One day co-location data set; 1 Sep 2011
V5.0 Level 2 data

Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>$\varepsilon_1$</th>
<th>$\varepsilon_2$</th>
<th>$\varepsilon_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_1$</td>
<td>1.0000</td>
<td>0.2394</td>
<td>0.2710</td>
</tr>
<tr>
<td>$\varepsilon_2$</td>
<td>0.2394</td>
<td>1.0000</td>
<td>0.9995</td>
</tr>
<tr>
<td>$\varepsilon_3$</td>
<td>0.2710</td>
<td>0.9995</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

1) Aquarius (AQ)
2) In Situ (IS)
3) Hycom (HY)

- Yes, the cross correlations are definitely non-zero.
- Future work:
  - Evaluate the effect on the derived error variances
  - More extensive study of Level 2 data
  - Biases in the three co-located measurements were removed and the effect needs to be studied more extensively
  - Apply the analysis to the monthly co-located analysis used to derive the monthly errors in the V5.0 Validation analysis document, and revise results as needed.
Final Comments

- Although the Aquarius mission is over, it has provided a legacy data set from which there is still much to learn.
- Aquarius data lives!

V5.0 Mean
Backup Material
Regression Analysis

1. Regression: \( \mathbf{R} = \mathbf{DD} \backslash \mathbf{DI} \)
2. Inverse: \( \mathbf{DI}_r = \mathbf{DD}^* \mathbf{R} \) is expected to contain measurement error (\( \mathbf{\varepsilon}_n \)) but not \( S_T \)
3. \( \mathbf{DI} - \mathbf{DI}_r = \mathbf{S}_T \) Solves for \( S_T \)
4. \( \mathbf{\varepsilon}_n = \mathbf{DI}_r \)
How to estimate uncertainty of Aquarius and validation data

The satellite salinity measurement $S_S$ and the in situ validation measurement $S_V$ are defined by:

$$S_S = S \pm \varepsilon_S$$
$$S_V = S \pm \varepsilon_V$$

where $S$ is the true surface salinity averaged over the Aquarius footprint area and microwave optical depth in sea water (~ 1 cm). $\varepsilon_S$ and $\varepsilon_V$ are the respective satellite and in situ measurement errors relative to $S$. The mean square of the difference $\Delta S$ between $S_S$ and $S_V$ is given by:

$$<\Delta S_{SV}^2> = <\varepsilon_S^2> + <\varepsilon_V^2>$$  \hspace{1cm} (1)

where $<>$ denotes the average over a given set of paired satellite and in situ measurements, and $<\varepsilon_S \varepsilon_V> = 0$.

Likewise, define HyCOM salinity interpolated to the satellite footprint as $S_H = S \pm \varepsilon_H$, and mean square differences

$$<\Delta S_{HV}^2> = <\varepsilon_H^2> + <\varepsilon_V^2>$$  \hspace{1cm} (2)  

HyCOM vs in situ validation data

$$<\Delta S_{SH}^2> = <\varepsilon_S^2> + <\varepsilon_H^2>$$  \hspace{1cm} (3)  

Satellite vs HyCOM

Equations (1)-(3) comprise three equations with three variables given by:

$$<\varepsilon_S^2> = \frac{<\Delta S_{SV}^2> + <\Delta S_{SH}^2> - <\Delta S_{HV}^2>}{2}$$  \hspace{1cm} (4)  

satellite measurement error

$$<\varepsilon_H^2> = \frac{<\Delta S_{SH}^2> + <\Delta S_{HV}^2> - <\Delta S_{SV}^2>}{2}$$  \hspace{1cm} (5)  

HyCOM measurement error

$$<\varepsilon_V^2> = \frac{<\Delta S_{SV}^2> + <\Delta S_{HV}^2> - <\Delta S_{SH}^2>}{2}$$  \hspace{1cm} (6)  

In situ validation measurement error
Towards the true near-surface wind speed: Error modeling and calibration using triple collocation

Ad Stoffelen
Royal Netherlands Meteorological Institute, de Bilt, Netherlands
<table>
<thead>
<tr>
<th>Year</th>
<th>SMOS</th>
<th>Aquarius</th>
<th>SMAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Situ Ocean Surface Data