

# Error Estimation of Pathfinder version 5.3 SST Level 3C using Extended Triple Collocation approach

Korak Saha<sup>1,4</sup>, Prasanjit Dash<sup>2,3</sup>, Xuepeng Zhao<sup>4</sup> and Huai-min Zhang<sup>4</sup>

1. CISSS-Maryland, University of Maryland, College Park, MD 20740, USA; NCEI, NOAA/NESDIS, Silver Spring, MD 20910, USA
2. Cooperative Institute for Research in Atmosphere (CIARA), Colorado State University, Fort Collins, CO 80523, USA
3. Center for Satellite Applications and Research (STAR), NOAA/NESDIS, College Park, MD 20740
4. National Centers for Environmental Information (NCEI), NOAA/NESDIS, Silver Spring, MD 20910, USA
5. National Centers for Environmental Information (NCEI), NOAA/NESDIS, Asheville, NC 28801, USA



## Introduction:

Generally, for validation purposes, satellite-derived SST products are compared against the in-situ SSTs which have inaccuracies due to spatial/temporal inhomogeneity between in-situ and satellite measurements. A standard deviation in their difference fields usually have contributions from both satellites as well as the in-situ measurements. A real validation of any geophysical variable must require the knowledge of the "true" value of this variable. Therefore, a one-to-one comparison of satellite-based SST with in-situ data does not truly provide us the real error in the satellite SST and there will be ambiguity due to errors in the in-situ measurements and their collocation differences. A Triple collocation (TC) or three-way error analysis using three mutually independent error-prone measurements, is used to estimate root-mean square error (RMSE) associated with each of the measurements with a high level of accuracy without treating any one system a perfectly-observed "truth". In this study, we are estimating the absolute random errors associated with Pathfinder Version 5.3 Level-3C (PF53) SST Climate Data Record. Along with the in-situ SST data, the third source of the dataset used for this analysis is the AATSR reprocessing of climate (ARC) dataset for the corresponding period. All three SST observations are collocated, and statistics of the difference between each pair are estimated. Instead of using a traditional TC analysis we have implemented the **Extended Triple Collocation (ETC)** approach to estimate the correlation coefficient of each measurement system w.r.t. the unknown target variable along with their RMSEs.

## Background:

**Triple Collocation Error Analysis**-TCM is a technique for estimating the unknown root-mean square Errors (RMSEs) of three mutually independent measurement systems, without treating any one system a perfectly observed "TRUTH" [Stoffelen, 1998].

Assumptions: 1.) A linear error model; 2.) Errors are uncorrelated with each other and the target variable (e.g. SST); 3.) Errors from independent sources have "Zero" mean.

1.  $X_i = \alpha_i + \beta_i t + \epsilon_i$
2.  $Cov(\epsilon_i, \epsilon_j) = 0, i \neq j$  and  $Cov(t, \epsilon_i) = 0$
3.  $E(\epsilon_i) = 0$

**Extended Triple Collocation**-ETC used the covariance between difference measurement systems to calculate the RMSEs ( $\sigma$ ) and the corr. coefficient ( $\rho$ ) of the measurement system w.r.t. the "Unknown TRUTH" [McCole et al. 2014].

$$\sigma_\epsilon = \begin{bmatrix} \sqrt{Q_{11} - \frac{Q_{12}Q_{13}}{Q_{23}}} \\ \sqrt{Q_{22} - \frac{Q_{12}Q_{23}}{Q_{13}}} \\ \sqrt{Q_{33} - \frac{Q_{13}Q_{23}}{Q_{12}}} \end{bmatrix} \quad \rho_{t,x} = \pm \begin{bmatrix} \sqrt{\frac{Q_{12}Q_{13}}{Q_{11}Q_{23}}} \\ \text{sign}(Q_{13}Q_{23}) \sqrt{\frac{Q_{12}Q_{23}}{Q_{22}Q_{13}}} \\ \text{sign}(Q_{12}Q_{23}) \sqrt{\frac{Q_{13}Q_{23}}{Q_{33}Q_{12}}} \end{bmatrix}$$

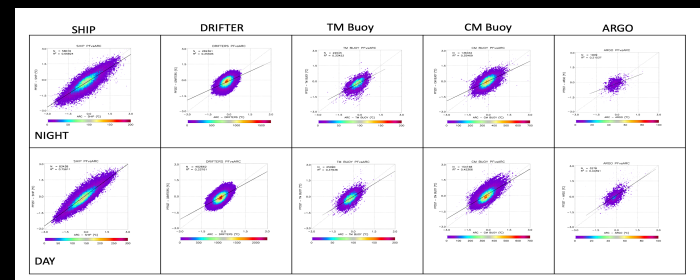
ETC also provides the unbiased SNR as a function of corr. coefficient, which provides information on sensitivity and confidence in the dataset errors.

## DATA Used:

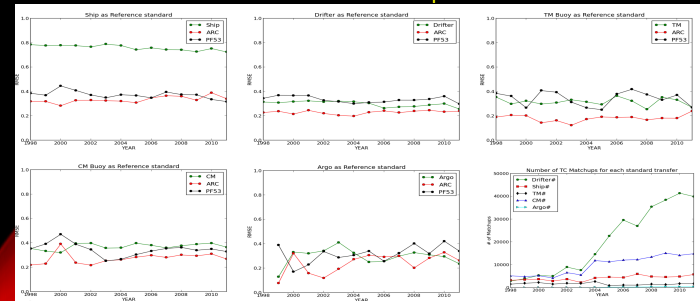
1. Pathfinder SST version 5.3 (PF53) L3C SST - <https://doi.org/10.7289/v52i68xx>
2. AATSR Reprocessing for Climate (ARC) dataset version 1.1.1 - <http://catalogue.ceda.ac.uk/uuid/a44cd6735b7046e13da2ca0bec33c7a9>
3. Iquam v2 in-situ SSTs - <https://www.star.nesdis.noaa.gov/so/d/sst/iquam/v2/>

## Results:

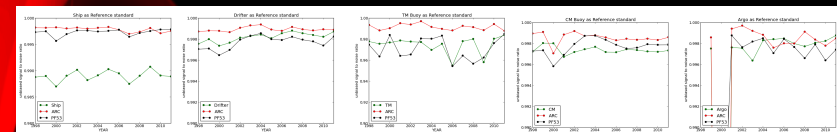
### 1. Assumption of Data Independence



### 2. RMSEs and number of Matchups



### 3. Correlation Coefficient



**Source:** Saha, K.; Dash, P.; Zhao, X.; Zhang, H.-M. Error Estimation of Pathfinder Version 5.3 Level-3C SST Using Extended Triple Collocation Analysis. Remote Sens. 2020, 12, 590.

## Conclusion:

1. RMSEs associated with PF53 is estimated using ETC on three sets of SST data
2. The RMSE ranged from **0.31 to 0.37 K for PF53**, and 0.18 to 0.33 K for the ARC data.
3. These values are reasonable, as is evident from corresponding very high (~0.98) unbiased signal-to-noise ratio (SNR) values.