**Error Characterization of Global Land Evapotranspiration Products: A Collocation approach**

Changming Li1, Hanbo Yang1\*, Wencong Yang1, Ziwei Liu1, Yao Jia1, Sien Li2, Dawen Yang1

1State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University. Beijing 100084, China

2Center for Agricultural Water Research in China, China Agricultural University, Beijing 100083, China

\*Correspondence: Hanbo Yang ([yanghanbo@tsinghua.edu.cn](mailto:yanghanbo@tsinghua.edu.cn) )

## S2: Deviation of Collocation Analysis Methods

### Triple Collocation method (TC)

In triple collocation analysis, three independent products of a geophysical variable are typically assumed to be linearly related to the true signal (Mccoll et al., 2016). This linear model can be expressed as:

Where is the product, is the true signal; and are the ordinary least squares intercept and slope; is zero-mean random error. This model is referred as additive error model, while a multiplicative error model is also available in conjunction with log transformation.

The main assumptions adopted in TC method contain: (1) error orthogonality, assuming that the random error is independent with the true signal, which can be expressed as: ; (2) zero error cross-correlation, requiring the independence of each two products, which can be expressed as: ; (3) the random error of each products is zero-mean, which means . Based on these assumptions, with the variance of the truth signal , the covariances between the products are expressed as:

Thus, the variances and the Pearson correlation () of each product against the true signal can be derived as:

### Single Instrumental Variable based algorithm (IVS)

By introducing an instrumental variable (), (Su et al., 2014) proposed the instrumental- variable-based algorithm (IVS) to calculate geophysical data uncertainty requiring only two independent products. Given that the random errors of the two independent products are both white noises, (Su et al., 2014) directly taken the lag-1 time series of one product ( for example) as the third input , i.e., . Where is the true signal; and are the ordinary least squares intercept and slope; is zero-mean random error.

In this way, addition to the original triple collocation scheme, the white-noise random error assumption is required, which can be expressed as: .

The covariance between the original products and the instrumental variable can be expressed as:

Where is the lag-1 auto-covariance of the true signal:

Here, we define the IVS-estimated scaling ratio of the two products as:

Therefore, the variances of each product against the true signal can be derived as:

The error variances of and and their correlation with truth can be solved for as:

### Double instrumental variable algorithm (IVD)

Based on IVS method, (Dong et al., 2019) recommended a more robust solution using the lag-1 variables from both geophysical products, denoted as the double instrumental variable algorithm, or IVD.

The estimate of is modified by including one additional instrumental variable, thus, serially lag-1 geophysical observations from both products are used as instrumental variables.

The covariance of and (i.e., ) is expressed as:

The ratio of and can be solved for as:

Where is the scaling ratio estimated by IVD.

Thus, the error variances and Pearson correlation can be solved for as:

This modification reduces the impact of random sampling errors on scaling ratio estimates, which leads to reduced uncertainty in estimates of and (Zhou et al., 2021).

### Quadruple collocation method (QC)

(Gruber et al., 2016) extended The TC algorithm to include a fourth dataset (i.e., quadruple collocation, or QC) and demonstrated that the error variances can be estimated with a least squares’ solution following the same TC assumptions. Furthermore, the zero ECC assumption can be relaxed. On the condition that only one pair within the four datasets have non-zero ECC, estimates of ECC can be obtained from the least-squares solution (Zwieback et al., 2019; Sun et al., 2021). All the errors are mutually independent, except for two specific products (taken product and for example), which can be expressed as:

Therefore, the covariance can be derived as:

The collocation system of equations can be shown as:

And the final equations for variance, Pearson correlation and ECC of specific products can be solved for as:

### Extended double instrumental algorithm (EIVD)

Combining the benefits of both instrumental variable algorithm and quadruple collocation method, (Dong et al., 2020) proposed an extended double instrumental variable algorithm, or EIVD. The extended system adopts following assumptions:

(1) All the errors are assumed orthogonal to the true signal

(2) All the errors are mutually independent, except for and

(3) The random errors of all products are serially white:

(4) The random error is zero-mean

Therefore, the covariance and lag-1 auto-covariance can be derived as:

Thus, the variances of the truth and the errors can be estimated in a linear system:

And the final equations for variance, Pearson correlation and ECC of specific products can be solved for as:

## References

Dong, J., Crow, W. T., Duan, Z., Wei, L., and Lu, Y.: A double instrumental variable method for geophysical product error estimation, Remote Sensing of Environment, 225, 217-228, 10.1016/j.rse.2019.03.003, 2019.

Dong, J., Wei, L., Chen, X., Duan, Z., and Lu, Y.: An instrument variable based algorithm for estimating cross-correlated hydrological remote sensing errors, Journal of Hydrology, 581, 124413, 10.1016/j.jhydrol.2019.124413, 2020.

Gruber, A., Su, C. H., Zwieback, S., Crow, W., Dorigo, W., and Wagner, W.: Recent advances in (soil moisture) triple collocation analysis, International Journal of Applied Earth Observation and Geoinformation, 45, 200-211, 10.1016/j.jag.2015.09.002, 2016.

McColl, K. A., Roy, A., Derksen, C., Konings, A. G., Alemohammed, S. H., and Entekhabi, D.: Triple collocation for binary and categorical variables: Application to validating landscape freeze/thaw retrievals, 176, 31-42, 10.1016/j.rse.2016.01.010, 2016.

Su, C.-H., Ryu, D., Crow, W. T., and Western, A. W.: Beyond triple collocation: Applications to soil moisture monitoring, Journal of Geophysical Research: Atmospheres, 119, 6419-6439, 10.1002/2013jd021043, 2014.

Sun, J., McColl, K. A., Wang, Y., Rigden, A. J., Lu, H., Yang, K., Li, Y., and Santanello, J. A.: Global evaluation of terrestrial near-surface air temperature and specific humidity retrievals from the Atmospheric Infrared Sounder (AIRS), Remote Sensing of Environment, 252, 10.1016/j.rse.2020.112146, 2021.

Zhou, J., Crow, W. T., Wu, Z., Dong, J., He, H., and Feng, H.: A triple collocation-based 2D soil moisture merging methodology considering spatial and temporal non-stationary errors, Remote Sensing of Environment, 263, 10.1016/j.rse.2021.112509, 2021.

Zwieback, S., Bosch, D. D., Cosh, M. H., Starks, P. J., and Berg, A.: Vegetation–soil moisture coupling metrics from dual-polarization microwave radiometry using regularization, Remote Sensing of Environment, 231, 10.1016/j.rse.2019.111257, 2019.