Validating WaPOR Data

Draft v1.0

SOLOMON SEYOUM
IHE DELFT

Validating WaPOR datasets

WaPOR is a FAO's portal to monitor Water Productivity through Open-access of Remotely sensed derived data. The WaPOR database is a comprehensive database that provides information on gross Biomass and Gross Biomass Water Productivity, actual Evapotranspiration and Interception, Transpiration, Net Primary Production, Total Biomass Production, Land Cover Classification, Precipitation, and more for Africa and the Near East in near real time covering the period 1 January 2009 to date. The database is in continuous development, currently at version 2 (WaPOR 2.0) and soon will release version 3 (WaPOR 3.0?).

Validation of WaPOR data has been done through the Quality Asesssemnt report (refxxx). This document will discuss and propose methods how to validate WaPOR data specially how to compare WaPOR data with locally available in-situ data.

Remote sensing data can be compared with in-situ observation qualitatively and/or quantitatively. Qualitative comparison includes visually comparing the RS data with the in-situ data, observing variation of the data at different locations based on known factors affecting the data such as elevation, land cover, slope etc. For example, evapotranspiration is known to be higher in forested areas than bare land, irrigated crops transpire much water than rainfed crops. Analyzing evapotranspiration per land cover can provide useful information about the quality of evapotranspiration data. Precipitation is expected to be higher in mountainous areas compared to precipitation at lowlands. Comparison also can be made using modelling results, for example how model results change when using RS data versus in-situ observation are used.

Comparison of observed in-situ data with RS data

Comparison of gridded remote sensed water balance components (precipitation, evapotranspiration etc) with gauge-based in situ observations is normally relies on two approaches: point-to-pixel and pixel-to-pixel comparisons (Macharia et al. 2022, Cerón et al. 2020; Dembélé and Zwart 2016; Dinku et al. 2018.

Point-to-pixel comparisons relates observed values to values of the corresponding satellite pixel where the point measurement is located where as for pixel-to-pixel comparisons spatial interpolation of point-based measurements to the same spatial resolution of the remotely sensed data is needed (Dinku et al. 2018).

The strengths and weaknesses of these approaches has been discussed in Macharia et al. 2022, Stampoulis and Anagnostou 2012; Nerini et al. 2015. The interpolation of point-based observations to spatial data has been said to smoothen the values and therefore decrease the accuracy of extreme values. Different interpolation methods produce different results bring large errors and uncertainties to the outcome when using data from sparse and highly unevenly distributed stations, which is a most common situation in large parts of the world.

Comparison Metrics

Different metrics have been used to compare both point to pixel and pixel to pixel data comparison.

Well-studied and common validation metrics used include Pearson's correlation coefficient (CC), bias, root mean square error, Nash-Sutcliff efficiency, probability of detection, false alar ratio, mean absolute error, multiplicative bias, critical success index, Spatial similarity index, principal component analysis etc (Macharia et al. 2022, Cerón et al. 2020).

Some of these metrics are applicable only for point measurements and high-resolution temporal data (upto daily) and other are applicable for spatial data comparison. The Probability of detection, False alarm ratio and Critical success index are used to validate remotely sensed precipitation data with in-situ measurement.

Table 1 - Description of the comparison metrics

Evaluation metrics	Formula	Unit	Values range	Best value
Pearson's correlation coefficient (CC)	$CC = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$	-	[− 1, 1]	1
Coefficient of determination (R^2)	$R^2 = (CC)^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 (y_i - \bar{y})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}$	-	[0, 1]	1
Nash-Sutcliffe Efficiency (NSE)	$NSE = 1 - \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$	-	(-∞, 1]	1
Kling-Gupta Efficiency (KGE)	$KGE = 1 - \sqrt{(\rho - 1)^2 + (\frac{\sigma_y}{\sigma_x} - 1)^2 + (\frac{\bar{y}}{\bar{x}} - 1)^2}$ Where ρ is Pearson's correlation coefficient, σ_y *and* σ_x standard deviations of y and x serieses	-		1
Bias	$BIAS = \sum_{i=1}^{n} \frac{(x_i - y_i)}{n}$	-	[0, + ∞)	0
Root mean square error (RMSE)	$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(x_i - y_i)^2}{n}}$	mm	[0, + ∞)	0
Mean Absolute error (MAE)	$MAE = \frac{1}{n} \sum_{i=1}^{n} \lvert x_i - y_i \rvert$	mm	[0, + ∞)	0
Probability of detection	$POD = \frac{H}{H + M}$	-	[0, 1]	1
False alarm ratio	$FAR = \frac{F}{H + F}$	-	[0, 1]	1
Critical success index	$CSI = \frac{F}{H + F + M}$	-	[0, 1]	1

Where x_i is the value of the remotely sensed product during the i th period; y_i is the observed point value (such as precipitation at a at rain gauge station) during the i th period; $\overline{x}$ and $\overline{y}$ are the average of the values in the corresponding periods (for the remotely sensed value and point observed values, respectively; n is total number of observations and H, F and M represent hit, false alarm and misses respectively of the remotely sensed data.

Coefficient of correlation is used to quantify the correspondence of RS product values and observed time series. BAIS indicates whether the RS product values underestimates (negative) or overestimates (positive) the observed values (Nogueira et al. 2018). RMSE measures the differences between two variables and is always a positive value (Beck et al. 2017). The values of BIAS and RMSE close to zero indicates good agreement between the RS product values and the observed time series (Nogueira et al. 2018).

The probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI) were used to evaluate the accuracy of satellite products in detecting rainfall occurrence at the daily time scale.

Principal Component Analysis (PCA) can be used to evaluate the large-scale spatiotemporal coherence among the observed and remotely sensed data (see Cerón et al. 2020). The details of this method are out of the context of this review. To read more about the method and how it can be used to compare two datasets please refer to Cerón et al. and 2020, Björnsson and Venegas 1997.

Steps to calculate the validation metrics

For point to pixel comparison;

1. Get the time series of the location under consideration (call it the *y* series)
2. Get the location (lat and long) of the point and extract a time series of the corresponding pixel from the remotely sensed product for the same temporal resolution (call it the *x* series)
3. For POD, FAR and CSI, count the number of Hits, False alarms and Misses
4. Use the formula in Table 1 to compute the metrics

Pixel-to-pixel comparison

Pixel-to-pixel data comparison is only possible for gridded datasets

The second approach of observing in-situ (observed) data with RS spatial data is pixel-to-pixel comparisons. This implies that the observed data be spatial. However most of the times observed data is in the form of pint data and therefore spatial interpolation of point-based measurements to the same spatial resolution of the remotely sensed data is needed. There are several ways of producing spatial maps from point observations. For the interpolated spatial map to be representative of the spatial variation of the variable estimated or measured, the point data should be spatially distributed to capture the variation in the underlying process of the variable estimated.

Some of the widely used spatial interpolation methods include Theissen Polygons, Inverse Distance Weighted (IDW), Spline, Triangular Interpolation Network (TIN), Trend Surface Analysis and Kriging. These methods have their own advantage and flaws. It is not the intention of this document to discuss different interpolation methods. You can read more about these methods in Simpson and Wu, (2014), Robinson and Metternicht, (2006), Naoum and Tsanis (2004) and Laslett et al. (1987).

After spatial maps are produced using one of the interpolation methods, pixel to pixel comparison of the interpolated map to the remotely sensed map can be made if the two maps (raster) are in the same projection and spatial resolution.

The maps can be compared in terms of spatial patterns. Statistical testing has been employed to compare spatial patterns. Levine et al., (2009) computed the differences of pixel to pixel and assuming mean pixel differences have a normal distribution, they employed a two-sample t-test to evaluate the null hypothesis that pixel scores are the same. Jones et al. (2016) showed the Structural Similarity (SSIM) index, a method they adapted from determining quality of image compression that can be used. An advantage of SSIM index is that different aspects of spatial comparison such as similarities in local means, variances and covariances can be computed and summarizes them in one metric which can further be summarized by calculating mean on the index to provide a single value to express similarity between two maps.

References

Cerón, W. L., Molina-Carpio, J., Ayes Rivera, I., Andreoli, R. V., Kayano, M. T., & Canchala, T. (2020). A principal component analysis approach to assess CHIRPS precipitation dataset for the study of climate variability of the La Plata Basin, Southern South America. Natural Hazards, 103(1), 767-783.

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H., & Ceccato, P. (2018). Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Quarterly Journal of the Royal Meteorological Society, 144, 292-312.

Jones, E. L., Rendell, L., Pirotta, E., & Long, J. A. (2016). Novel application of a quantitative spatial comparison tool to species distribution data. Ecological Indicators, 70, 67-76.

Laslett, G. M., McBrantney, A. B., Pahl, P. J., & Hutchinson, M. F. Comparison of several spatial prediction methods for soil pH. European Journal of Soil Science. 1987; 38(2): 325–341.

Levine, R.S., Yorita, K.L., Walsh, M.C. et al. A method for statistically comparing spatial distribution maps. Int J Health Geogr 8, 7 (2009). https://doi.org/10.1186/1476-072X-8-7

Macharia, D., Fankhauser, K., Selker, J. S., Neff, J. C., & Thomas, E. A. (2022). Validation and intercomparison of satellite-based rainfall products over Africa with TAHMO in-situ rainfall observations. Journal of Hydrometeorology.

Naoum, S. & Tsanis, I.K. 2004, 'Ranking Spatial Interpolation Techniques Using a GIS-Based DSS', Global Nest: the Int. J., vol. 6, no.1, pp. 1-20.

Robinson, T. P., & Metternicht, G. Testing the performance of spatial interpolation techniques for mapping soil properties. Computers and Electronics in Agriculture. 2006; 50(2): 97–108.

Simpson, G, & Wu, Y. -H. Accuracy and effort of interpolation and sampling: Can GIS help lower field costs? ISPRS International Journal of Geo-Information. 2014; 3: 1317–1333.