TEMPORAL ERROR VARIABILITY OF COARSE SCALE SOIL MOISTURE PRODUCTS - CASE STUDY IN CENTRAL SPAIN

Simon Zwieback
Institute of Environmental Engineering
ETH Zurich
CH-8093 Zurich
Switzerland

Wouter Dorigo, Wolfgang Wagner
Institute of Photogrammetry and Remote Sensing
Vienna University of Technology
A-1040 Wien
Austria

ABSTRACT
The triple collocation technique, which retrieves the error variances of three sets of measurements of the same parameter, is applied to soil moisture records in central Spain: ASCAT remote sensing observations, REMEDHUS in-situ probes, and the ERA Interim model. The objective is the estimation of the temporal variability of the error of ASCAT. The three data sets have to be calibrated with respect to each other as they show different mean values and dynamic ranges. The time-variant estimation of both the error and the calibration parameters is shown to be very sensitive to the extents of the temporal windows used and the calibration procedure. Due to the temporal fluctuations of the calibration constants, artefacts such as seasonal variations and extreme values are introduced. This case study shows that the temporal analysis of the errors using the collocation technique can lead to spurious results when the data sets have to be referenced with respect to one another.

Index Terms— Soil moisture, Spaceborne radar, Hydrology

1. INTRODUCTION
Soil moisture is considered one of the 'Essential Climate Variables' as it links the water and energy balance of both atmosphere and the Earth surface. Consequently coarse scale soil moisture products, as those derived from ASCAT, AMSR-E, and SMOS, are applied in a broad range of applications and scientific disciplines [1, 2]. An accurate characterization of the errors is of vital importance in such studies; many different procedures for the estimation of the errors have thus been suggested. The triple collocation technique has been frequently applied to estimate the error variance of global soil moisture products [3]. It requires three independent data sets, which are treated on an equal footing: estimates of the errors of all three data sources are provided. Additionally, it permits the calibration of the data sets with respect to one another; this is generally deemed necessary in the analysis of soil moisture products due to the significant differences in the nature of the various data sets. In the majority of scientific studies, the error characteristics are assumed to be time-invariant (e.g. [4, 5]). However, error sources such as vegetation cover and atmospheric influences change over time. The spatial scaling properties of soil moisture are subject to temporal variations as well and were e.g. studied by [1, 6]. It is thus of great interest to study the temporal variance of the retrieval errors.

In this study the calibration parameters and the errors are retrieved simultaneously for the ASCAT product, measurements from an in-situ probe and the ERA Interim model. The results are very sensitive to the choice of temporal windows for the estimation of either error or calibration constants. Similarly, the method of referencing the data sources with respect to one another has a considerable impact. This study thus shows that the application of the collocation technique to the estimation of time-variable errors can potentially yield spurious results when the data sets are not adequately calibrated.

2. MATERIALS AND METHODS
2.1. Study area
The study area is located in the Duero basin in central Spain (41.5°N, 5.5°W). The climate is semi-arid with an average precipitation of 385 mm and the land cover dominated by agricultural crops (largely cereals) [7]. The years 2007-2009 are considered.

2.2. Data sets
The following three data sets were used:
- in-situ frequency domain probe K10 of the REMEDHUS soil-moisture network [7, 8] (sampling depth: 0-5 cm, hourly measurements)
- ASCAT remotely sensed soil moisture [2] (C-band radar, resolution: 50 km, rescaled to range [0, 1], GP:

TU Wien acknowledges the support of the ESA through the Climate Chance Initiative (CC) Soil Moisture project
2.3. Collocation technique

Let \( y'_i \) denote the value of data set \( i \) at time \( t \). If the errors are additive, uncorrelated and their expectation zero, the triple collocation technique permits the retrieval of the error variance of each of the data sources. When the observations are not referenced with respect to each other, it is possible to introduce the calibration constants \( \alpha_i \) and \( \beta_i \) [3]:

\[
y'_i = \beta_i x^t + \alpha_i + e_i
\]

where \( x^t \) is the (unknown) soil moisture at time \( t \). The method of moments estimator for \( \alpha_i \) and \( \beta_i \) are given by (assuming no autocorrelation of the errors) [3]:

\[
\hat{\beta}_i = \frac{\sum_{t=1}^{N} y'^t_i y'^t_j}{\sum_{t=1}^{N} y'^t_j y'^t_j}
\]

\[
\hat{\alpha}_i = \frac{1}{N} \sum_{n=1}^{N} y^t_i - \hat{\beta}_i y'^t_i
\]

where \( N \) is the number of measurements, \( y'^t_i = y'_i - \frac{1}{N} \sum_{t=1}^{N} y'^t_i \), and the index 1 in Eq. (2) and (3) refers to the data sets to which the others are referenced. In this study, this will be taken to be the in-situ probe. The remaining two data sources can then be expressed in terms of the in-situ measurements by inverting Eq. (1)

\[
y'_i = \frac{(y'_i - \hat{\alpha}_i)}{\beta_i}
\]

The actual triple collocation technique consists of the following set of estimators \((i \neq j \neq k)\):

\[
\hat{\sigma}_i^2 = \frac{1}{N} \sum_{t=1}^{N} \left( \hat{y}'_i - \hat{y}'_j \right) \left( \hat{y}'_i - \hat{y}'_k \right)
\]

Note that the properties of this estimator are poorly understood when the multiplicative bias \( \beta_i \) has to be estimated from the data [3].

If either the calibration constants or the error statistics are expected to vary with time, they can be estimated by windowing: at time \( t \), only the measurements at \( t' \) with \( |t-t'| < w \) are considered in Eq. (3)-(5). The window size for the calibration constants is denoted by \( w_c \), the one for the errors by \( w_e \) – both expressed in months. When the calibration parameters are not constants, there are two ways to perform the rescaling: i) individually \( I \) (in the estimation of the errors centred around \( t \), each measurement \( y'_i \) is scaled according to the estimated calibration constants determined using a window of size \( w_c \) around time \( t' \)); ii) non-individually \( N \) (all the measurements in the window are converted using the calibration constants determined at time \( t \)).

3. RESULTS

The time series of the measured/modelled soil moisture are shown in Fig. 1. Note the differences in the mean and the spread, which necessitate the calibration of the data sets as described in Sec. 2.3. The results of the collocation technique are shown in Fig. 2. The different subplots display the estimated RMS errors and calibration constants for varying sizes of \( w_c \) and \( w_e \). The emphasis is put on possible seasonal fluctuations, with window sizes of 3 and 4 months; note that an infinite window means that all available data were used in the estimation.

3.1. Estimation: \( w_c = w_e \)

Figures 2a) and b) show the estimated RMS error of the ASCAT soil moisture for individual \( I \) and non-individual \( N \) rescaling, respectively – the window sizes \( w_c \) and \( w_e \) are taken to be the same. These errors are scaled to the in-situ probe. Note the large values encountered during spring 2009 for a window size of 4 months: the temporal extent of this outlier is several months for \( I \), but only a few days for \( N \). This exception notwithstanding, the curves for a window size of 3 and 4 months are similar: there is an annual variation.

3.2. Estimation: \( w_c \neq w_e \)

The effect of varying \( w_c \) for \( w_e = 3 \) is demonstrated in Fig. 2 c) (1) and d) (1)). Besides the outlier in spring 2009, which is still evident, the most striking feature is the difference encountered in the phase of the seasonal variation: the blue line \((w_c = \infty)\) reaches its maximum around three months before
the black and red lines, corresponding to \( w_c = 3 \) and \( w_c = 4 \), respectively.

3.3. Estimation: calibration constants

The estimated calibration constants \( \hat{\alpha} \) and \( \hat{\beta} \) for the ASCAT sensor (with respect to the in-situ probe) are shown in Fig. 2 e) and f), respectively. Note that the time series exhibit large fluctuations for window sizes of both three and four months.

3.4. Estimation: referenced to ASCAT

Figures 2g) and h) illustrate the same results as a) and b), but referenced with respect to ASCAT by multiplying by \( \hat{\beta} \) (obtained with the appropriate window \( w_c \)). The temporal variability of the calibration constants results in qualitatively different features. The outlier in April 2009 is still evident for \( w_c = 4 \) in case of 4 rescaling, but not for \( N \).

4. DISCUSSION AND CONCLUSIONS

The results shown in Fig. 2 clearly indicate temporal fluctuations of both the RMS errors and the calibration constants. When the latter are estimated using a different window size (subfigures c) and d)), the estimated errors exhibit dissimilar annual variations. As the only difference between these two cases is the estimation of the calibration constants, the change over time of these parameters is the cause of the deviation. That the calibration constants do depend on the window can be seen in subfigures e) and f).

Not only are the estimated errors susceptible to artefacts induced by the choice of the window size, but also the estimation of the calibration constants \( \alpha \) and \( \beta \) is not robust. The estimator \( \hat{\alpha} \) of Eq. (3) is essentially an arithmetic mean, which is well known to be sensitive to outliers. The estimation of \( \hat{\beta} \) in Eq. (2) consists basically of the ratio of two covariances. If the covariance in the numerator is small (e.g. due to the corresponding correlation coefficient being close to zero), so will be \( \hat{\beta} \). After the rescaling of Eq. (4), where \( \hat{\beta} \) features in the denominator, the result will be spuriously large. This is exactly what happens for window sizes on the order of 4 months in April 2009. In Fig. 1 it can be seen that the ERA Interim curve decays more slowly than the other two data sets, implying low covariances between the different soil moisture time series for certain window sizes and thus a low \( \hat{\beta} \), cf. Fig. 2f). The corresponding rescaled values \( y' \), which are huge, lead to the outliers described in Sec. 3.

The reasons for the temporal variation of \( \hat{\beta} \) can be gleaned from the soil moisture time series, Fig. 1. During April and June (except for 2009), \( \hat{\beta} \) (ASCAT) is small, as ASCAT uses its dynamic range, and effectively so does the in-situ sensor (0 - 0.18). In summer and early autumn, the retrieval of \( \hat{\beta} \) is very unstable, as neither sensor uses its dynamic range due to dryness. From November to February, \( \hat{\beta} \) is large as ASCAT makes use of its dynamic range, whereas the in-situ probe has a limited range of values. Note that autocorrelation appears to be an issue, in particular in spring 2009: the drying out appears to take place at different rates for the different sensors. This could e.g. be related to different sampling/modelling depth or an inaccurate model representation of percolation and evapotranspiration. In such cases, the estimation of the calibration constants is conceptually difficult, as it is not clear which part of the deviation should be accounted for by the calibration constants and which by the random errors.

The temporal variations of both the calibration constants and the error structure are thus closely connected. The inference drawn from the present example are inherently non-robust to variations in the window sizes \( w_c \) and \( w_c \): subfigures a), b), g) and h) show different temporal behaviour. These effects can easily dominate more subtle phenomena, such as changes in the scaling behaviour of point measurements, the study of which might well be made impossible using the techniques and tools applied above.

The difficulties of the temporal estimation of the RMS error and the calibration constants encountered in this case study might well be an exception. They nevertheless indicate that additional research is necessary before the collocation technique can be applied routinely to such scenarios. The crucial aspects appear to be i) the choice of the windows for estimation of both the calibration parameters and the RMS errors; ii) the rescaling during the estimation (e.g. 4 and \( N \)); iii) the robust estimation of the calibration constants (particularly \( \beta \)); and iv) the treatment of autocorrelated errors. These issues and the inherent non-robustness of the collocation technique require particular care (e.g. sensitivity studies) when analyzing and interpreting real data.

5. REFERENCES


Fig. 2. Results of the triple collocation technique applied to the ASCAT, ERA Interim, and in-situ datasets. a) and b): estimated ASCAT RMS error ($w_c = w_e$) for individual $I$ and non-individual $N$ rescaling, respectively. c) (1) and d) (NI): estimated ASCAT RMS error ($w_c \neq w_e$). e) and f): estimated calibration constants $\alpha$ and $\beta$ for ASCAT with respect to in-situ. g) (1) and h) (NI): estimated ASCAT RMS error (scaled to ASCAT).


