This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Quadruple Collocation Analysis for Soil Moisture Product Assessment

Nazzareno Pierdicca, *Senior Member, IEEE*, Fabio Fascetti, Luca Pulvirenti, *Member, IEEE*, Raffaele Crapolicchio, and Joaquín Muñoz-Sabater

***Abstract*—For validating remotely sensed products, the triple collocation (TC) is often adopted, which is able to retrieve the independent error variances of three systems observing the same target parameter. In this letter, three years of soil moisture data derived from the Advanced SCATterometer (ASCAT) aboard the MetOp satellite and the Soil Moisture and Ocean Salinity (SMOS) radiometer are analyzed and compared with the ERA Interim/Land model outputs and the ground measurements available from the International Soil Moisture Network. As we have four sources, a novel quadruple collocation (QC) approach is developed, which is more precise than TC since it uses the sources jointly. The results of QC show that the ERA model has the lowest error variance, while ground measurements are likely to be affected by the difficulty to represent a mean soil moisture within the satellite field of view by a limited number of stations. Moreover, the ASCAT retrievals outperform the SMOS ones if only anomalies with respect to the seasonal trend are considered, while the opposite occurs when the whole dynamic of soil moisture variation is considered.**

***Index Terms*—Advanced SCATterometer (ASCAT), quadruple collocation (QC), Soil Moisture and Ocean Salinity (SMOS), soil moisture, triple collocation (TC).**

I. Introduction

VALIDATION of remotely sensed products is generally performed through a comparison with an independent set of observations from conventional *in situ* sensors, which are assumed to provide the true (or reference) value. The comparison is quantified by computing statistical scores, such as bias, root-mean-square error, and correlation coefficient. To deal with the fact that even the reference can be affected by errors, the triple collocation (TC) approach is often adopted in the literature [1], [2]. Although initially applied to scatterometer wind retrievals over ocean, it has been extensively applied to soil moisture products as well [2]–[4]. The TC estimates the standard deviation of the random errors affecting each data source related to the same target parameter. It requires setting up a database of collocated (in space and time) observations from three systems, in order to solve a set of equations derived from the variances and covariances among observations, and assuming that the errors are statistically independent. The error standard deviation magnitudes are then referred to the dynamic range of one of the systems.

Within the framework of the EUMETSAT Satellite Application Facility (SAF) on Support to Operational Hydrology and Water Management (H-SAF), some studies on the comparison among different soil moisture products were previously performed [5], [6]. They involved two remotely sensed soil moisture data sets, i.e., the European Space Agency (ESA) Soil Moisture and Ocean Salinity (SMOS) [7] and the H-SAF product derived from the Advanced SCATterometer (ASCAT) onboard the MetOp satellite [8]. The aforementioned papers used the ERA-Interim/Land model simulations produced by European Centre of Medium Range Weather Forecasts (ECMWF), hereafter denoted as ERA-LAND, and the ground measurements available in the frame of the International Soil Moisture Network (ISMN) [9] as references. In particular, in [6], a series of applications of the TC was performed by considering the four possible combinations of collocated data of three systems out of four. From this approach, different error variances can be retrieved for each system according to the considered combination; however, the reference cannot be the same in all combinations.

In this letter, a novel quadruple collocation (QC) approach, based on the same assumptions of the TC, is developed. It solves the system of equations relating observations, variances, and covariances through a root-mean-square minimization, because there are fewer unknown variables (i.e., the four error standard deviations and standard deviation of the target parameter) than equations. This solution is more precise with respect to the TC, where the number of unknowns is equal to that of the equations, and copes with the problem of having a common reference for the error magnitude. A study about the QC was accomplished in [10], where, however, three systems were used to retrieve the error variance, and eventually, the fourth one was exploited to estimate an additional unknown, i.e., the cross-correlation among system errors. Conversely, we introduce a new approach solving a redundant/overconstrained system of equations to reduce the sampling errors.

Manuscript received October 9, 2014; revised December 29, 2014; accepted February 25, 2015.

N. Pierdicca and F. Fascetti are with the Department of Information Engineering, Electronics and Telecommunications, Sapienza University of Rome, Rome 00185, Italy (e-mail: nazzareno.pierdicca@uniroma1.it; f.fascetti@diet.uniroma1.it).

L. Pulvirenti is with CIMA Research Foundation, Savona 17100, Italy, and also with the Department of Information Engineering, Electronics and Telecommunications, Sapienza University of Rome, Rome 00185, Italy (e-mail: luca.pulvirenti@cimafoundation.org).

R. Crapolicchio is with SERCO S.p.a., Frascati 00044, Italy, and also with ESA-ESRIN, Frascati 00044, Italy (e-mail: Raffaele.Crapolicchio@esa.int).

J. Muñoz-Sabater is with the European Centre for Medium-Range Weather Forecasts, Reading RG2 9AX, U.K. (e-mail: Joaquin.Munoz@ecmwf.int).

Digital Object Identifier 10.1109/LGRS.2015.2414654

The QC solution is described in detail in this letter, and the major outcomes of its application to the comparison between ASCAT/HSAF, SMOS, ERA-LAND, and ISMN soil moisture data are discussed and compared to the TC ones.

II. Data Set and the Preprocessing Steps

The large-scale surface soil moisture products (SM-OBS-1), available through the H-SAF project over Europe and North Africa, are produced from the C-band (5.255 GHz), vertically polarized ASCAT data by means of the TU-Wien algorithm [11]. The spatial resolution is around 50 km, and the data are sampled with a spacing of 25 km. Each pixel represents a relative value (between 0% and 100%) of moisture with respect to the driest and wettest conditions, which is the degree of saturation *SD* (i.e., the soil moisture content expressed in percent of porosity). As for SMOS [7], it is an interferometric radiometer that measures the antenna brightness temperature at 1.427 GHz (L-band) at different angles, with a horizontal spatial resolution between 35 and 50 km. The reprocessed level 2 (L2) products, obtained from the version 551 of the processor, provide the volumetric soil moisture content (*SMC*) in cubic meters per cubic meter or in percentage units, sampled over the ISEA4h9 grid, whose spacing is on the order of 15 km.

Satellite data are compared to the ERA-Interim/Land modeled soil moisture and to *in situ* data available from the ISMN. ERA-LAND, produced by the ECMWF, is a global atmospheric reanalysis combined with an ocean and a land surface model available until 2012. Soil moisture is provided at four different layers and at synoptic hours each day over a grid with a space sampling of $0.125 \times 0.125^{\circ}$. ISMN is an international cooperation coordinated by the Global Energy and Water Exchanges Project, in cooperation with the Group of Earth Observation and the Committee on Earth Observation Satellites, with the task of maintaining a global *in situ* soil moisture database. For the comparison, we used the data available from several European networks with a probe depth of 0–5 cm [6] in order to minimize data mismatch due to different probing depths.

All data have been collocated in time and space over the SMOS (ISEA4h9) grid in the time frame from 2010 to 2012. While satellite and modeled data have been collocated in space by a nearest neighbor approach, the *in situ* data have been upscaled to the satellite resolution, through averaging of the station measurements within the satellite field of view (an intermediate value between ASCAT and SMOS). More details on the data processing can be found in [5] and [6]. Since it is necessary to compare quantities with the same units, the H-SAF *SD* retrievals have been converted into *SMC* through a linear transformation using a soil porosity map available from the Global Land Data Assimilation System (GLDAS) website (http://ldas.gsfc.nasa.gov/gldas/).

In the combined database, we have retained data fulfilling these conditions: 1) SMOS retrievals with data quality index (DQX) [5] less than 0.045; 2) ASCAT/HSAF retrievals with less than four bad quality flags up; 3) SMOS and ASCAT/HSAF SMC values below 0.7 m^3/m^3; and 4) ASCAT *SD* values between 0 and 100%.

As QC and TC require a constant mean of the considered quantity (i.e., stationarity), the mean spatial pattern has been estimated by averaging the SMC maps over time and removed. As far as the seasonal variability of the soil moisture is concerned, two approaches have been adopted, whose consequences will be discussed in Section IV. The first one considers it as a temporal drift to be removed, using a suitable harmonic fitting; in this case, we look just at the anomalies of soil moisture. The second approach considers that the period of analysis covers almost three years of data, and so, it assumes the seasonal variability as part of the random variability that is retained in the data to which TC and QC are applied.

III. QC Approach

To evaluate the error associated to the soil moisture estimates provided by the different systems, the formalism introduced in [1] is adopted. Supposing that four systems X, Y, Z, and W provide the x, y, z, and w measures, taking the first one as unit of reference, the following error model has been adopted as in [12], but adding the fourth system

$$\begin{aligned} x &= \theta + \delta_X \\ y &= s_Y(\theta + \delta_Y) \\ z &= s_Z(\theta + \delta_Z) \\ w &= s_W(\theta + \delta_W) \end{aligned} \tag{1}$$

where s_Y, s_Z, and s_W are the gains of the systems with respect to the reference (i.e., system X with a unitary gain being the reference), θ represents the true volumetric soil moisture, and δ_X, δ_Y, δ_Z, and δ_W are the random errors affecting the three systems. The common assumptions reported in TC literature are considered here (see, for instance, [10]). The errors are assumed with zero mean, and variance $\varepsilon_X^2 = \langle \delta_X^2 \rangle$, $\varepsilon_Y^2 = \langle \delta_Y^2 \rangle$, $\varepsilon_Z^2 = \langle \delta_Z^2 \rangle$, and $\varepsilon_W^2 = \langle \delta_W^2 \rangle$ ($\langle \cdot \rangle$ stands for ensemble average); the correlation between different errors is 0 (i.e., $\langle \delta_i \delta_j \rangle = 0$ for each $i, j = X, Y, Z, W, i \neq j$), as well as the correlation between the errors and the true random variable θ (i.e., $\langle \theta \delta_i \rangle = 0$ for each $i = X, Y, Z, W$), although the latter can also be considered deterministic [10]. In case of TC, once the measurements of the first three systems are collocated in time and space, expressing the variances $\sigma_X^2, \sigma_Y^2, \sigma_Z^2$ and the covariances C_{XY}, C_{XZ}, C_{YZ} between all individual pairs of observations, one can write a system of six equations from which the six unknowns can be derived, which are the gains s_Y and s_Z, the true variable variance $\sigma^2 = \langle \theta^2 \rangle$, and the error standard deviations of each system. Recently, in [13], it has been also derived the correlation coefficient between observations and the true parameter. It is further required that the true random variable θ has a constant mean. For simplicity, a zero mean can be also considered, so that the covariances are simply defined as $C_{XY} = \langle xy \rangle, C_{XZ} = \langle xz \rangle$, and $C_{YZ} = \langle yz \rangle$. In case the true variable is actually a random function $\theta(\underline{\mathbf{r}}, t)$ of horizontal surface coordinates $\underline{\mathbf{r}}$ and time t, as in the case of the soil moisture field, the nonstationary component of the fields (which in geostatistics is referred as a "drift") should be removed, as detailed in [6].

When four systems are available, considering the four variances and six covariances, it turns out that the number of equations ($= 10$) becomes larger than the number of unknowns

$(=8)$, since one more gain s_W and error variance ε_W^2 are added. The overconstrained system has to be solved using a minimum root-mean-square criterion. We have solved this algebraic problem and found the following formulas for the four unknown error standard deviations, the three gains, and the variance of the true variable

$$\begin{aligned}\varepsilon_X^2 &= \langle x^2\rangle - \sigma^2; \quad \varepsilon_Y^2 = s_Y^{-2}\langle y^2\rangle - \sigma^2\\ \varepsilon_Z^2 &= s_Z^{-2}\langle z^2\rangle - \sigma^2; \quad \varepsilon_W^2 = s_W^{-2}\langle w^2\rangle - \sigma^2\end{aligned} \tag{2}$$

$$\begin{aligned}s_Y &= \frac{\langle xw\rangle\langle yz\rangle^2\langle yw\rangle + \langle xz\rangle\langle yw\rangle^2\langle yz\rangle}{\langle xw\rangle^2\langle yz\rangle^2 + \langle xz\rangle^2\langle yw\rangle^2}\\ s_Z &= \frac{\langle xy\rangle\langle zw\rangle^2\langle yz\rangle + \langle xw\rangle\langle yz\rangle^2\langle zw\rangle}{\langle xy\rangle^2\langle zw\rangle^2 + \langle xw\rangle^2\langle yz\rangle^2}\\ s_W &= \frac{\langle xy\rangle\langle zw\rangle^2\langle yw\rangle + \langle xz\rangle\langle yw\rangle^2\langle zw\rangle}{\langle xy\rangle^2\langle zw\rangle^2 + \langle xz\rangle^2\langle yw\rangle^2}\end{aligned} \tag{3}$$

$$\begin{aligned}\sigma^2 = {} & (s_Y\langle xy\rangle + s_Z\langle xz\rangle + s_W\langle xw\rangle + s_Y s_Z\langle yz\rangle\\ & + s_Y s_W\langle yw\rangle + s_Z s_W\langle zw\rangle) \,/\, \big(s_Y^2 + s_Z^2 + s_W^2 + s_Y^2 s_Z^2\\ & + s_Y^2 s_W^2 + s_Z^2 s_W^2\big).\end{aligned} \tag{4}$$

The mathematical proof is reported in the Appendix.

The reader can refer to [14] for a critical review of TC (and QC) assumptions and the impact of their possible failure. Generally speaking, a correlation among errors, or of errors with the target parameter, or a lack of error stationarity (in space and time) can introduce systematic factors affecting in the same way the different systems. For instance, seasonal effects could introduce correlations between errors of different satellites, even in case they are working at different frequency bands and operating in passive and active modes (as in the case of SMOS and ASCAT).

IV. Analysis of the Results

A. TC

As the first step, the TC technique has been applied to the four data sets (SMOS, ISMN, ASCAT/H-SAF, and ERA-LAND) by considering any possible combination of triplets. The results of the different combinations are reported in Table I, with *SMC* expressed in percentage unit. For each cell of Table I, the first line represents the results obtained by estimating and removing the seasonal variability (i.e., looking at the temporal anomalies), while the second line reports the results achieved by retaining the seasonal variability in the data. For the TC analysis, one system must be chosen as reference; thus, the error standard deviations are expressed in its observation space. When the ISMN is considered, it is chosen as reference, while in the last column of Table I, ERA-LAND is the reference.

Looking at the first line of each cell of columns 2–4 (i.e., ISMN as reference and seasonal variability removed from the data), it is shown that the error standard deviation is on the order of 5% for SMOS and ISMN, slightly less for ASCAT/H-SAF, and surprisingly lower for ERA (around 3%). If the seasonal variability is retained in the data (second line of each cell),

TABLE I
Results of the TC Evaluating All the Possible System Combinations. In Each Cell, the Upper Figure Refers to the Temporal Anomalies (Seasonal Variability Is Removed), and the Lower One Refers to the Collocation of Data Including the Seasonal Variability. σ Is the True Variable Standard Deviation, Whereas the Subscript of Gain S and Error Standard Deviation ε Refers to the Different Systems

	ISMN H-SAF SMOS	ISMN H-SAF ERA	ISMN SMOS ERA	ERA H-SAF SMOS
σ [%]	3.88 5.23	3.70 5.74	4.02 5.58	6.00 6.04
$s_{\text{H-SAF}}$	1.94 1.46	2.13 1.21	-	1.36 1.22
s_{SMOS}	1.05 0.94	-	0.98 0.83	0.63 0.84
$s_{\text{ERA/LAND}}$	-	1.68 1.12	1.42 1.19	1 1
$\varepsilon_{\text{ISMN}}$ [%]	4.97 5.32	5.11 4.77	4.86 4.95	-
$\varepsilon_{\text{H-SAF}}$ [%]	4.61 5.94	4.04 7.60	-	6.14 7.21
$\varepsilon_{\text{SMOS}}$ [%]	4.94 5.36	-	5.41 6.44	8.70 5.85
ε_{ERA} [%]	-	2.59 3.07	3.50 2.59	4.65 4.11

SMOS slightly outperforms ASCAT/H-SAF, the ISMN error standard deviation is slightly less than $\varepsilon_{\text{SMOS}}$ and $\varepsilon_{\text{H-SAF}}$, and ERA still exhibits the lowest error standard deviation. In general, the trend of each data set is kept for all of the TC combinations; for example, when considering the anomalies, if either ASCAT/H-SAF or ERA is introduced in the TC analysis, the SMOS gain with respect to the reference ISMN is around 1 (the same applies for the ASCAT/H-SAF gain). It is interesting to note that, removing only the spatial pattern, SMOS outperforms ASCAT/H-SAF, whereas the opposite is observed when the temporal anomalies are considered. A possible explanation is that the ASCAT/H-SAF retrieval algorithm is based on a change detection approach which relates directly the variations in radar backscatter to soil moisture changes, whatever the temporal scale is. Conversely, the retrieval algorithm of SMOS is based on a forward model of surface emissivity which reveals itself more suitable to account for other environmental variables involved in the seasonal cycle and to sense the whole dynamic range of moisture.

B. QC

Considering that we have a collocated database of four systems, we can overcome the ambiguity inherent to the accomplishment of individual TC analyses. These would assign different errors according to how the systems are combined, whereas, by applying the QC approach, unambiguous estimates of the error in the same reference scale are derived. It is also more robust since it is derived from a larger set of data through a minimum root-mean-square error estimate. It is not significantly more demanding from the computation point of view, as it took 0.01 s to process 2585 records of QCs on an Intel i3 based PC. The most demanding step is the detrending of

TABLE II
RESULTS OF THE QC. THE LEFT COLUMN REFERS TO THE TEMPORAL ANOMALIES, AND THE RIGHT COLUMN REFERS TO THE DATA INCLUDING THE SEASONAL VARIABILITY

ISMN ERA/LAND H-SAF SMOS		
σ [%]	3.90	5.54
$s_{H\text{-}SAF}$ [%]	2.03	1.31
s_{SMOS} [%]	1.02	0.88
s_{ERA} [%]	1.53	1.15
ε_{ISMN} [%]	4.96	5.00
$\varepsilon_{H\text{-}SAF}$ [%]	4.25	6.83
ε_{SMOS} [%]	5.23	5.82
ε_{ERA} [%]	3.06	3.07

spatial and temporal drifts. The results are reported in Table II, where all the estimates are expressed in the scale of ISMN and the second and third columns refer to data without and with temporal drift included, respectively. It can be noticed that the results of the QC are in agreement with those obtained by the four TCs, as expected. Indeed, a consistent and more precise solution is provided.

The good performances of the ERA/LAND are confirmed, with error standard deviation of about 3%, considering either the anomaly or the entire dynamic range of the parameter. This could be surprising, but is a direct consequence of the good correlation exhibited by the ERA/LAND data with respect to the other systems. Despite the reasons why a reanalysis can be inaccurate, the result suggests to take into account the many sources of inaccuracies in satellite products. ERA-LAND has proved to be a substantial improvement of land variables, as soil moisture [15], and at the end, over a long period, a relatively good quality reanalysis could be more reliable than satellite products. In addition, regarding the main input for soil moisture, ERA-LAND does not strictly assimilate precipitation; however, the precipitation forcing comes from ERA-Interim precipitation rescaled with the Global Precipitation Climatology Project v2.1, therefore of reasonable quality.

The *in situ* data, upscaled at the satellite resolution, do not offer very good performances, which again remain stable in the two cases (i.e., overall dynamic range or anomalies). This is not really unexpected, as we are actually measuring the capability of a pointwise measurement to represent the average soil moisture within an area that is equal to the field of view of the satellite sensors or the resolution of the model (order of 30–40 km). Moreover, the *in situ* network is quite sparse, and in some areas, it may not provide a suitable coverage. Ground probes remain the only reliable source of point measurements.

Regarding the two satellite products, ASCAT/H-SAF has a considerably lower error standard deviation when looking at the temporal anomalies, but the opposite occurs when considering the whole variability of the soil moisture, including the seasonal one.

V. CONCLUSION

Within the framework of the H-SAF product validation activity, an extensive comparison between SMOS, ASCAT/HSAF, ISMN, and ERA Interim/Land soil moisture products has been performed. The comparison has been carried out over both Europe and Northern Africa territory, using the data acquired from 2010 to 2012. To this aim, we have developed a novel QC approach to estimate the error variance of the four systems, which is based on the same hypotheses of the TC but offers better robustness and the capability to refer to a consistent reference system.

The QC has surprisingly indicated that ERA Interim/LAND yields the best performances (SMC error standard deviation on the order of 3%); SMOS (5.8 % error standard deviation) has a slightly better behavior than ASCAT/HSAF (6.8% error standard deviation) when considering data including the seasonal variability, while the opposite (i.e., 5.2% and 4.3%, respectively) is observed when looking at the temporal anomalies. The *in situ* data (around 5% error standard deviation) suffer from the problem of representing a large field of view by pointwise sparse data.

The QC technique represents a powerful evolution of TC, although it can suffer from the difficulty to gather four observation systems which obey the fundamental hypothesis of independent errors. A further step of this research will be the consideration of systems having different spatial resolutions. This introduces a representativeness error, which is a correlated component of the random error due to the small-scale variability of the target parameter.

APPENDIX
QC ANALYTICAL FORMULATION

Our starting point is (1). For the sake of simplicity, we assume the expected values of the true quantity θ and that of the errors to be zero. Then, by deriving the observation variances from (1), under the hypothesis already done for TC (i.e., $\langle\delta_i\delta_j\rangle = \langle\theta\delta_i\rangle = 0$ for each $i, j = X, Y, Z, W$), the system of equations in (2) can be easily obtained. Furthermore, six additional nonlinear equations can be considered by deriving the six possible covariances combining the four observations modeled as in (1) and applying again the aforementioned hypothesis

$$\begin{aligned}\langle xy\rangle &= s_Y\sigma^2;\ \langle xz\rangle = s_Z\sigma^2\\ \langle xw\rangle &= s_W\sigma^2;\ \langle yw\rangle = s_Y s_W\sigma^2\\ \langle yz\rangle &= s_Y s_Z\sigma^2;\ \langle wz\rangle = s_Z s_W\sigma^2\end{aligned} \tag{A1}$$

where it was assumed that $s_X = 1$ for the reference system. Equation (A1) is nonlinear, as it contains products among unknown quantities (i.e., gains and variance σ^2). To overcome this difficulty, we have considered the following ratios:

$$\begin{aligned}a &= \langle xw\rangle/\langle yw\rangle;\ b = \langle xz\rangle/\langle yz\rangle\\ c &= \langle xy\rangle/\langle yz\rangle;\ d = \langle xw\rangle/\langle zw\rangle\\ e &= \langle xy\rangle/\langle yw\rangle;\ f = \langle xz\rangle/\langle zw\rangle.\end{aligned} \tag{A2}$$

which can be rearranged in matrix form obtaining the following linear system:

$$\mathbf{B} = \begin{bmatrix}1\\1\\1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}a & 0 & 0\\ b & 0 & 0\\ 0 & c & 0\\ 0 & d & 0\\ 0 & 0 & e\\ 0 & 0 & f\end{bmatrix}\begin{bmatrix}s_Y\\ s_Z\\ s_W\end{bmatrix} = \mathbf{AX} \tag{A3}$$

where $\mathbf{X}$ is the vector of the unknown gains (dimension 1×3), $\mathbf{A}$ is the matrix of the known coefficients (dimension 3×6), and $\mathbf{B}$ is a unity vector (dimension 1×6). The terms of the matrix $\mathbf{A}$ are the ratios between covariances estimated from the collocated QC database. Equation (A3) can be solved in the minimum root-mean-square sense using the so-called pseudoinverse matrix [16] according to

$$\mathbf{X} = (\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{B}. \tag{A4}$$

After straightforward algebraic manipulations, the solution is the following:

$$\mathbf{X}=\begin{bmatrix} s_Y \\ s_Z \\ s_W \end{bmatrix}=\begin{bmatrix} (a^2+b^2)^{-1} & 0 & 0 \\ 0 & (c^2+d^2)^{-1} & 0 \\ 0 & 0 & (e^2+f^2)^{-1} \end{bmatrix}\begin{bmatrix} a+b \\ c+d \\ e+f \end{bmatrix} \tag{A5}$$

which, once we substitute (A2) in (A5), after simple algebraic manipulations, provides the formula of the gains in (3). In order to derive the error standard deviation in (2), the true variable variance σ^2 still needs to be derived. To do this, we use (A1) and solve them again in the minimum root-square sense by minimizing the following objective function with respect to σ^2

$$\begin{aligned}\mathbf{F}(\sigma^2) &= \left[\langle xy\rangle - s_Y\sigma^2\right]^2 + \left[\langle xz\rangle - s_Z\sigma^2\right]^2 \\ &\quad + \left[\langle xw\rangle - s_W\sigma^2\right]^2 + \left[\langle yz\rangle - s_Y s_Z\sigma^2\right]^2 \\ &\quad + \left[\langle yw\rangle - s_Y s_W\sigma^2\right]^2 + \left[\langle zw\rangle - s_Z s_W\sigma^2\right]^2. \end{aligned} \tag{A6}$$

Computing the derivative with respect to σ^2 and solving for the derivative equal to zero, the solution reported in (4) is obtained in terms of the covariances computed from the QC database.

ACKNOWLEDGMENT

The authors would like to thank DPC for its support and, in particular, Dr. S. Puca. This work has been accomplished in the frame of the EUMETSAT H-SAF project validation managed by the Italian Department of Civil Protection. SMOS data have been acquired in the frame of the ESA Cat. 1 project (N. 9720).

REFERENCES

[1] A. Stoffelen, "Toward the true near-surface wind speed: Error modeling and calibration using triple collocation," *J. Geophys. Res.*, vol. 103, no. C4, pp. 7755–7766, Apr. 1998.

[2] K. Scipal, T. Holmes, R. de Jeu, V. Naeim, and W. Wagner, "A possible solution for the problem of estimating the error structure of global soil moisture data sets," *Geophys. Res. Lett.*, vol. 35, no. 24, 2008, Art. ID. L24403.

[3] D. G. Miralles, W. T. Crow, and M. H. Cosh, "Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations," *J. Hydrometeorol.*, vol. 11, no. 6, pp. 1423–1427, Dec. 2010.

[4] W. A. Dorigo *et al.*, "Error characterization of global active and passive microwave soil moisture data sets," *Hydrol. Earth Syst. Sci.*, vol. 14, pp. 2605–2616, 2010.

[5] N. Pierdicca, L. Pulvirenti, F. Fascetti, R. Crapolicchio, and M. Talone, "Analysis of two years of ASCAT- and SMOS-derived soil moisture estimates over Europe and North Africa," *Eur. J. Remote Sens.*, vol. 46, pp. 759–773, 2013.

[6] N. Pierdicca, F. Fascetti, L. Pulvirenti, R. Crapolicchio, and J. Muñoz-Sabater, "Analysis of ASCAT, SMOS, *in-situ* and Land Model soil moisture as a regionalized variable over Europe and North Africa," *Remote Sens. Environ.*, submitted for publication.

[7] Y. Kerr *et al.*, "Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 8, pp. 1729–1736, Aug. 2001.

[8] Z. Bartalis *et al.*, "Initial soil moisture retrievals from the MetOp-A Advanced Scatterometer (ASCAT)," *Geophys. Res. Lett.*, 34, no. 20, 2007, Art. ID. L20401.

[9] W. A. Dorigo *et al.*, "The International Soil Moisture Network: A data hosting facility for global *in situ* soil moisture measurements," *Hydrol. Earth Syst. Sci.*, vol. 15, pp. 1675–1698, 2011.

[10] S. Zwieback, K. Scipal, W. Dorigo, and W. Wagner, "Structural and statistical properties of the collocation technique for error characterization," *Nonlin. Process. Gephys.*, vol. 19, pp. 69–80, 2012.

[11] V. Naeimi, K. Scipal, Z. Bartalis, S. Hasenauer, and W. Wagner, "An improved soil moisture retrieval algorithm for ERS and MetOp scatterometer observations," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 1999–2013, Jul. 2009.

[12] D. Cimini *et al.*, "On the accuracy of integrated water vapor observations and the potential for mitigating electromagnetic path delay error in InSAR," *Atmos. Meas. Tech.*, vol. 5, pp. 1015–1030, 2012.

[13] K. A. McColl *et al.*, "Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target," *Geophys. Res. Lett.*, vol. 41, no. 17, pp. 6229–6236, Sep. 2014.

[14] M. T. Yilmaz and W. T. Crow, "Evaluation of assumptions in soil moisture triple collocation analysis," *J. Hydrometeorol.*, vol. 15, no. 3, pp. 1293–1302, Jun. 2014.

[15] G. Balsamo *et al.*, "ERA-Interim/Land: A global land-surface reanalysis data set," *Hydrol. Earth Syst. Sci.*, vol. 19, pp. 389–407, 2015 doi:10.5194/hess-19-389-2015.

[16] C. Radhakrishna Rao, "Analysis of dispersion for multiply classified data with unequal numbers in cells," *Sankhyd, Indian J. Statisti.*, vol. 15, no. 13, pp. 253–280, Jul. 1955.