SMOS soil moisture data validation in the Aurajoki watershed, Finland
Juho Jakkila, Tiia Vento, Tapani Rousi and Bertel Vehviläinen

ABSTRACT

Soil Moisture and Ocean Salinity (SMOS) level 2 soil moisture observations were validated by the soil moisture simulations of the Watershed Simulation and Forecasting System (WSFS) in the Aurajoki watershed located in Southwest Finland. The validation period included summer seasons 2010–2012. A new 2-layer soil moisture model was developed to simulate soil moisture in two layers, 10 cm thick surface layer and 80 cm thick sub-surface layer, due to low penetration depth of the satellite observations. Both layers of the model were divided into six classes of soil types to characterize the soil properties of the study catchment. The soil parameters for different soil types were estimated using the soil moisture observation network and other model parameters were calibrated against discharge observations. The WSFS with the 2-layer soil moisture model was found appropriate in reproducing observed discharges in the test site. Despite slight dry bias, the SMOS observations were consistent with the simulated surface soil moisture content, giving Nash–Sutcliffe efficiency coefficient of 0.77 and root-mean-square deviation of 0.076. The analysis of the SMOS level 2 data presents promising results for the satellite observations to be used as an indicator of the filling of the soil moisture storage in the WSFS.

Key words | hydrological modeling, satellite observation, soil moisture, soil properties, watershed

INTRODUCTION

The European Space Agency (ESA) Soil Moisture and Ocean Salinity (SMOS) satellite mission aims to provide observations of the global surface soil moisture variation to improve meteorological forecasting, hydrological modeling, monitoring photosynthesis and plant growth and estimation of the terrestrial carbon cycle. The calibration and validation (Cal/Val) campaigns of SMOS have been active since the satellite was launched in 2009. The Cal/Val campaigns focus on monitoring of detailed temporal and spatial variation of soil moisture in SMOS validation sites by collecting in situ soil moisture measurements, land surface and hydrological model simulations and ancillary remote sensing data of satellite and air-borne radiometers (Juglea et al. 2010; dall’Amico et al. 2012; Wanders et al. 2012).

The large spatial variation of soil properties brings the challenge for both observing and modeling of soil moisture content (Lindström et al. 1997; Dingman 2002). The spatial and temporal variation of soil moisture can be monitored at a field scale for instance by time or frequency domain reflectometers (TDR or FDR) (Zehe et al. 2010; Heathman et al. 2012), ground penetrating radars (Huisman et al. 2002; Weihermüller et al. 2007) and ground-based radiometers (Rautiainen et al. 2012; Schlenz et al. 2012). However, the data assimilation of watershed models and validation of soil moisture simulations of the general circulation models of the atmosphere require larger scale soil moisture observations, which can only be provided by satellite observations.

The objective of the SMOS satellite mission is to respond to the challenge with a new technology, which enables the receiving of emitted microwave radiation around the frequency of 1.4 GHz (L-band). The L-band radiometer increases the penetration depth and has lower
sensitivity to surface roughness, vegetation and atmosphere than the higher frequency (C-band) passive microwave instruments (Wigneron et al. 2005; Kerr et al. 2008). The SMOS level 2 soil moisture data are expected to provide an accuracy of 4% volumetric soil moisture content in 35–50 km spatial resolution, 1–3 days revisit time and a couple of centimeters penetration depth (Kerr et al. 2010).

In hydrological modeling the soil spatial variation can be described in detail only at field scale by three dimensional models (Bronstert & Plate 1997; Laine-Kaulio 2011; Warsta et al. 2013). In watershed scale, the conceptual Hydrologiska Byrâns Vatternbalansavdelning (HBV) models have been widely used, especially in operational flood forecasting (Bergström 1976, 1995). The spatial variation of the soil moisture dynamics in the HBV models was improved in the 1990s, when distributed approaches to small sub-basins were introduced (Lindström et al. 1991; Bergström & Graham 1998). However, the lack of the spatial information of the soil properties disables the direct comparison between the soil moisture simulations of the HBV model and the observations of soil moisture content.

The main objective of this study was to develop the HBV-based soil moisture model of Watershed Simulation and Forecasting System (WSFS) for validation of the SMOS satellite soil moisture observations. The new 2-layer soil moisture model was developed to simulate soil moisture content in 10 cm thick surface layer, which is comparable to penetration depth of the SMOS observations. The TDR- or FDR-devices of the automatic soil moisture stations were used to fix the new soil moisture model parameters for different soil types. The 2-layer model also applies the information of the Finnish soil database (Lilja et al. 2009) in order to improve the spatial variation of the soil moisture simulations in the WSFS.

**TEST SITE**

The Hypöistenkoski catchment, in the upper part of the Aurajoki watershed (Figure 1) located in Southwest Finland was chosen as a test site for SMOS level 2 data validation due to good coverage of daily data in the catchment area. The Hypöistenkoski observation station provides discharge records at the outlet of the catchment, which enables the hydrological model calibration of the WSFS. The catchment is mainly covered by cultivated fields and without lakes and large wetland areas runoff is fast responding to precipitation events and snowmelt. Table 1 provides hydrological characteristics of the Hypöistenkoski catchment including discharge information. Runoff is usually at largest after snowmelt in spring, but autumn and winter rains also bring frequently high discharges. During the 1961–2010 period, 58% of the annual maximum discharges occurred in spring (Mar–May), 28% in winter (Dec–Feb), 12% in autumn (Sep–Nov) and 2% in summer (Jun–Aug). The mean high-water discharges MHQ for the seasons are 39.8, 16.0, 27.0 and 11.9 m³/s, respectively. The annual mean value for precipitation is 630 mm and for evapotranspiration 470 mm.

<table>
<thead>
<tr>
<th>Area (km²)</th>
<th>Lakes (%)</th>
<th>MQ (m³/s)</th>
<th>MHQ (m³/s)</th>
<th>MNQ (m³/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>351.02</td>
<td>0.0</td>
<td>3.3</td>
<td>50.0</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Soil information for the test site was acquired from the Finnish soil database (Lilja et al. 2009) and detailed soil data...
for agricultural fields were provided by a Finnish agricultural service ‘Viljavuuspalvelu’ (www.viljavuuspalvelu.fi). The soil classification is based on FAO/Unesco, World Reference Base for Soil Resources and Finnish soil type classifications (Aaltonen et al. 1994). The Hypöistenkoski catchment is mostly covered by agricultural fields with clayey and sandy soil classes (Cambisols and Regosols, Figure 1). The uncultivated areas in the catchment are typically covered by till (Podzols), peatlands (Histosols) and rocky soil classes ( leptosols). In Figure 2, the soil classes are classified by the soil types of the surface and the sub-surface layers. Seventy-two percent of the upper soil layer in Hypöistenkoski catchment consists of sandy, rocky and clayey soil types. Other areas are covered by organic (15%), till (9%) and silty (4%) soil types. The sub-surface soil layers mainly consist of clayey (49%) or rocky soil types (25%). Organic and till soil types are both found with around 10% coverage. Sandy and silty soil types are scarce in the sub-surface layer.

SMOS level 2 data

The SMOS level 2 soil moisture product is generated from series of dual-polarized multi-angular measurements of brightness temperatures using a radiative transfer model L-MEB (L-band Microwave Emission of the Biosphere), which applies global land use, soil texture and vegetation information (Wigneron et al. 2007; Kerr et al. 2010). Different parametrizations of the algorithms – e.g. radiative transfer models and dielectric mixing models – in the level 2 soil moisture processors have been tested during the Cal/Val campaigns (Kerr et al. 2012). The reprocessed level 2 data of the latest processor version v5.01 for years 2010 and 2011 were obtained for this study. The operational data of the processor version v5.51 were used for year 2012. The main difference between the processor versions v5.01 and v5.51 is the change of the dielectric mixing model from Dobson to Mironov formulation, which is expected to improve the soil moisture estimates (de Rosnay et al. 2009; Bircher et al. 2012; Kerr et al. 2012). The whole data series was filtered by missing values of the SMOS level 2 variable, data quality index (DQX SM), to discard the failed retrieval of the soil moisture data. Also, radio frequency interferences (RFI) were filtered out with the formula:

\[
\frac{N_{RFIx} + N_{RFIy}}{M_{AVA0}} < 0.1
\]

where \( N_{RFIx} \) and \( N_{RFIy} \) are the numbers of the deleted views due to RFI in X and Y polarization and \( M_{AVA0} \) is the number of views initially available for the given SMOS discrete global grid (DGG) node.

Figure 3 shows the filtered level 2 data coverage in Finland, including transboundary watersheds, calculated for the 2011 summer season. Since the area of Finland is largely covered by forests (72%) and lakes (10%) and the retrieval of the SMOS level 2 data over the forest areas is still poor in processor versions v5.01 and v5.51, continuous time
series for soil moisture were found only in a few areas in southwestern and western parts of Finland, where agricultural fields are the dominant land use class. Despite typically poor retrieval of SMOS level 2 data in forested areas, relatively high numbers of retrieved data were also found from the eastern border of Finland, which is mainly covered by wetlands and coniferous forests and from the most northern part of Finland, where the landscape is typically covered by sparsely populated mountain birch forests and wetlands.

The Aurajoki watershed in Southwest Finland was selected as the test site because the Hypöistenkoski catchment is dominated by agricultural fields without significantly large water bodies resulting in good quality SMOS level 2 data series. Figure 4 shows the SMOS DGG nodes 25 km in distance from the Hypöistenkoski catchment and the number of filtered SMOS level 2 data in each DGG node during the summer seasons of 2010–2012. The agricultural land use is more common on the eastern side of the catchment where the landscape is typically covered by sparse mountain birch forests and wetlands.

The Aurajoki watershed in Southwest Finland was selected as the test site because the Hypöistenkoski catchment is dominated by agricultural fields without significantly large water bodies resulting in good quality SMOS level 2 data series. Figure 4 shows the SMOS DGG nodes 25 km in distance from the Hypöistenkoski catchment and the number of filtered SMOS level 2 data in each DGG node during the summer seasons of 2010–2012. The agricultural land use is more common on the eastern side of the catchment where the highest number of the SMOS level 2 data were received. The root-mean-square-deviation (RMSD) of level 2 data between the three closest nodes with more than 400 soil moisture retrievals were 0.05 m$^3$/m$^3$ and increased to 0.04–0.06 m$^3$/m$^3$ with more distant nodes.

The filtered daily average SMOS level 2 soil moisture data within 25 km distance from each sub-basin of the Hypöistenkoski catchment were selected and soil moisture (SM) for each sub-basin was calculated by:

$$SM(k) = \frac{\sum_{i=1}^{n} \left[ \frac{SM_{smos}(i)}{d(i, k)^2} \right]}{\sum_{i=1}^{n} \left[ \frac{1}{d(i, k)^2} \right]}$$

where $SM_{smos}$ is the SMOS soil moisture observation and $d$ is the distance between the center of sub-basin $k$ and SMOS node $i$ and $n$ is the number of the SMOS nodes within 25 km distance from the sub-basin $k$. The 25 km distance was selected to collect all DGG nodes, which footprint covers at least part of the sub-basin area and inverse distance weighting was used to put most of the weight on the closest DGG nodes.

**METHODS**

**SMOS data validation**

The resolution of the SMOS data is 35–50 km and the footprint of the soil moisture data covers vast areas with variable soil and land use classes. Because soil moisture content
depends remarkably on the soil properties and due to the lack of multi-station soil moisture observation in the study area, the SMOS level 2 data were validated against the 2-layer soil moisture model. The areal mean simulated surface layer soil moisture was calculated over 10 different soil classes in the four sub-basins of the Hypöistenkoski catchment. The areal weighted mean soil moisture over the sub-basins was compared to daily average SMOS observations covering the catchment. Three different efficiency coefficients presented in this section have been calculated between the SMOS observations and the mean simulated surface layer soil moisture.

The 2-layer model surface soil moisture simulations were compared to SMOS level 2 observations in the summer seasons of 2010–2012. Goodness of fit between the SMOS level 2 data and simulated surface soil moisture was estimated with root-mean-square deviation (RMSD), correlation coefficient (r) and Nash–Sutcliffe efficiency coefficient (R², Nash & Sutcliffe 1970), which are calculated as follows:

\[
\text{RMSD}(x, y) = \sqrt{\frac{\sum_{i=1}^{n} (x_i - y_i)^2}{n}}
\]

\[
r = \frac{1}{n-1} \sum_{i=1}^{n} \left( \frac{x_i - \bar{x}}{s_x} \right) \left( \frac{y_i - \bar{y}}{s_y} \right)
\]

\[
R^2 = 1 - \frac{\sum_{i=1}^{n} (x_i - y_i)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}
\]

where \(n\) is the number of observations, \(x\) is the observed and \(y\) the modeled daily value, \(\bar{x}\) and \(\bar{y}\) are the observed and the modeled means and \(s_x\) and \(s_y\) the observed and the modeled standard deviations. These efficiency criteria were calculated for the whole period excluding winter seasons and separately for each summer season. The summer seasons were defined as April 15–September 15. The simulated soil moisture content before and after the summer seasons was normally close to the saturation level. The influence of freezing and thawing and snow accumulation on SMOS observations described by Rautiainen et al. (2012) was the reason for discarding the observations outside the summer season from this analysis. \(R^2\) were also calculated between simulated and observed discharges in observation stations of Hypöistenkoski and Halinen. \(R^2\) for discharges produced by WSFS with the 2-layer model and with the operational soil moisture model were calculated without restriction to the summer seasons.

**Two-layer soil moisture model**

The WSFS is the main tool used in flood forecasting in Finland. The hydrological model has also been used in research purposes, especially in climate change studies (Vehviläinen & Huttunen 1997; Veijalainen et al. 2010, 2012). The model is based on the HBV-model developed in Sweden (Berghström 1976) and it has been further developed in the Finnish Environment Institute (Vehviläinen 1992; Vehviläinen et al. 2003). The hydrological model of the WSFS is a semi-distributed, conceptual model which includes sub-models for areal precipitation, snow cover, soil moisture, groundwater, river routing and lakes. The input variables, areal temperature and precipitation are calculated for each sub-basin by inverse distance weighting from three closest stations. The areal temperature and potential evaporation from class A-pan observations. The average area of the sub-basins in the Hypöistenkoski catchment is 90 km² and time-step of the model is 1 day.

The new 2-layer soil moisture model formulation is based on the operational soil moisture model of the WSFS described by Vehviläinen (1992). The operational model was modified such that the soil moisture content is simulated in two layers. The surface layer of the 2-layer model was selected to be 10 cm thick, which is comparable to the soil penetration depth of the SMOS observations, corresponding to the surface soil layer depth of Finnish soil database and is appropriate for the model to simulate soil moisture variation in 1 day time-step. The 80 cm thick sub-surface layer represents the lower part of the hydrologically active soil moisture storage. The surface layer of the model responds fast to precipitation events and changes in potential evaporation. Water infiltrates from the surface layer to the sub-surface layer as long as the soil moisture content exceeds the field capacity. Runoff is produced from the surface layer \((QR_{sl})\) only when the ground is frozen. The sub-surface layer produces the main part of the runoff \((QR_{sub})\) to the river routing and lake models as well as percolation to the groundwater storage (Figure 5). The new fixed soil
parameters – soil porosity and field capacity – were introduced into the model. The soil porosity determines the maximum soil moisture content $MVAK$ (Equations (9) and (15)), which is a calibrated parameter in the operational soil moisture model of the WSFS. The field capacity determines the volumetric water content, which limits infiltration, runoff and percolation in Equations (7), (12) and (14).

The water balance of the surface layer soil moisture content $MVS_{sl,i}$ is calculated as follows:

$$\frac{dMVS_{sl,i}(t)}{dt} = YIELD(t) - E_{sl,i}(t) - INF_i(t) - Q_{R,sl}(t) + E_{sub,i}(t)$$

(6)

$$INF_i(t) = YIELD(t) \left[ \frac{MVS_{sl,i}(t) - LIM_{sl,i}}{MVAK_{sl,i} - LIM_{sl,i}} \right]$$

(7)

$$E_{sl,i}(t) = HP(t) \frac{MVS_{sl,i}(t)}{LP}$$

(8)

$$MVAK_{sl,i} = \phi_{sl,i} \cdot d_{sl,i}$$

(9)

$$LIM_{sl,i} = FC_{sl,i} \cdot d_{sl,i}$$

(10)

where $YIELD$ is the sum of snowmelt and rain entering into the soil moisture model, $E_{sl}$ is the actual evaporation from the surface layer, $INF$ is the infiltration rate to the sub-surface layer, $E_{sub}$ is the upward flux of water vapour from the sub-surface layer to the surface layer, $HP$ is the potential evaporation, $LP$ is the level of the soil moisture content after which actual evaporation equals potential evaporation, $MVAK$ is the maximum surface soil moisture content, $\phi$ is the soil porosity, $FC$ is the field capacity and $d$ is the layer depth. Subscript $sl$ refers to the surface layer, $sub$ to the sub-surface layer and $i$ to the soil types.

In the 2-layer model, the sub-surface layer produces runoff $QR_{sub}$ to the rivers and lakes and percolation $PO$ to the groundwater storage. Water flows upwards from the sub-surface layer to the surface layer; if evaporation decreases the surface layer soil moisture content below the sub-surface layer soil moisture content. The water balance of the sub-surface layer is calculated as follows:

$$\frac{dMVS_{sub,i}(t)}{dt} = INF_i(t) - Q_{R,sub}(t) - PO_i(t) - E_{sub,i}(t)$$

(11)

$$QR_{sub}(t) = INF_i(t) \left[ \frac{MVS_{sub,i}(t) - LIM_{sub,i}}{MVAK_{sub,i} - LIM_{sub,i}} \right]$$

(12)

$$E_{sub,i}(t) = E_{sl,i}(t), \quad \text{if} \quad \frac{MVS_{sub,i}(t)}{d_{sub,i}} > \frac{MVS_{sl,i}(t)}{d_{sl,i}}$$

(13)

$$PO_i(t) = PC \left[ MVS_{sub,i}(t) - LIM_{sub,i} \right]$$

(14)

$$MVAK_{sub,i} = \phi_{sub,i} \cdot d_{sub,i}$$

(15)

$$LIM_{sub,i} = FC_{sub,i} \cdot d_{sub,i}$$

(16)

The calibrated parameters of the 2-layer model in Equations (6)–(16) are the surface layer infiltration parameter $EX_{sl}$, the sub-surface layer runoff parameter $EX_{sub}$, the evaporation parameter $LP$ and the percolation parameter $PC$. These parameters, like all of the calibrated parameters of the WSFS, were calibrated against the discharge observations of the Hypöistenkoski observation station.

The 2-layer model utilizes new soil information from the Finnish soil database (Lilja et al. 2009). In addition, more detailed field scale soil information of the surface layer soil properties from Finnish agricultural service was applied
for cultivated areas classified as Vertic Cambisols in Finnish soil database. The sub-basins of the WSFS were divided in both layers into six classes of soil types: sand, silt, clay, organic material, till and rock to characterize the soil properties of the catchment. The soil classes, with areal coverage over 2%, are simulated separately in each sub-basin improving the spatial resolution of the soil moisture model to approximately 5 km². The combination of two layers of soil types, with areal coverage exceeding 2%, form altogether 10 soil classes in Hypöistenkoski catchment. The percentage coverage of each soil class in the test site is presented in Table 2.

### Table 2: Soil classification based on Finnish soil database (Lilja et al. 2009) and percentage in the Hypöistenkoski catchment

<table>
<thead>
<tr>
<th>Soil classification</th>
<th>Surface layer</th>
<th>Sub-surface layer</th>
<th>Percentage of area in the test site (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertic Cambisol 1</td>
<td>Clay</td>
<td>Clay</td>
<td>20</td>
</tr>
<tr>
<td>Vertic Cambisol 2</td>
<td>Sand</td>
<td>Clay</td>
<td>23</td>
</tr>
<tr>
<td>Vertic Cambisol 3</td>
<td>Silt</td>
<td>Clay</td>
<td>2</td>
</tr>
<tr>
<td>Vertic Cambisol 4</td>
<td>Organic</td>
<td>Clay</td>
<td>2</td>
</tr>
<tr>
<td>Dystric/Lithic Leptosol</td>
<td>Rock</td>
<td>Rock</td>
<td>25</td>
</tr>
<tr>
<td>Fibric/Terric Histosol 1</td>
<td>Organic</td>
<td>Organic</td>
<td>8</td>
</tr>
<tr>
<td>Haplic Podzol 1</td>
<td>Till</td>
<td>Till</td>
<td>9</td>
</tr>
<tr>
<td>Haplic Podzol 2</td>
<td>Sand</td>
<td>Sand</td>
<td>5</td>
</tr>
<tr>
<td>Eutric Regosol</td>
<td>Silt</td>
<td>Silt</td>
<td>3</td>
</tr>
<tr>
<td>Umbric Gleysol 1</td>
<td>Organic</td>
<td>Clay</td>
<td>2</td>
</tr>
<tr>
<td>Gleyic Podzol 1</td>
<td>Organic</td>
<td>Till</td>
<td>1</td>
</tr>
<tr>
<td>Fibric/Terric Histosol 2</td>
<td>Organic</td>
<td>Silt</td>
<td>1</td>
</tr>
<tr>
<td>Gleyic Podzol 2</td>
<td>Organic</td>
<td>Sand</td>
<td>0</td>
</tr>
<tr>
<td>Eutric Cambisol 1</td>
<td>Sand</td>
<td>Clay</td>
<td>0</td>
</tr>
<tr>
<td>Fibric/Terric Histosol 3</td>
<td>Organic</td>
<td>Rock</td>
<td>0</td>
</tr>
<tr>
<td>Water body</td>
<td>–</td>
<td>–</td>
<td>0</td>
</tr>
</tbody>
</table>

*Vertic Cambisols are classified as fields for which more detailed surface layer soil information is available. Fields in the Hypöistenkoski catchment have a surface layer soil composition of 48% sand, 43% clay, 5% organic and 4% silt soil types. The sub-surface layer of fields is classified as clay.

The automatic soil moisture measurement network of the Finnish Environment Institute (SYKE) consists of eight TDR- or FDR-stations located in different parts of Finland. Both instrument types determine the dielectric constant of soil which correlates strongly with volumetric water content. The dielectric constant of ice is low compared to liquid water and close to the value of soil particles, which causes the decline in TDR- and FDR-measurements during the ground frost formation. Thus the observations in winter-time are unreliable. Water content measurements of unfrozen soil are available at each station for approximately 7 months a year.

The measurement network was used for estimating the new fixed soil parameters – soil porosity and field capacity – for different soil types by comparing the observations to the 2-layer model soil moisture simulations. The parameters provide soil characteristics such as water holding capacity of different soil types for the 2-layer model. In order to determine these values, the 2-layer model soil moisture simulations were fitted to the measurements by adjusting porosity and field capacity values. Figures 6 and 7 show the soil moisture measurements and fitted 2-layer model simulations at Tarvasjoki and Rautavaara soil moisture station. Tarvasjoki is located in the Aurajoki watershed close to the Hypöistenkoski catchment and Rautavaara is located in eastern Finland. At Tarvasjoki station the soil type is clay and the TDR-device measures soil moisture at 5, 10, 30 and 50 cm depths. The Rautavaara station is equipped with FDR-devices, which are installed in silt soil type at 5, 15, 30, 50, 70 and 100 cm depths. The surface layer observations are from the topmost measurements (at 5 cm depth) and the sub-surface observations are calculated from the deeper measurements as a mean value by integrating over the depth ranges between each sensor depths. At these stations the 2-layer model is fixed to simulate soil moisture at corresponding depths: surface soil layer at 7.5 cm thick layer at Tarvasjoki and 10 cm at Rautavaara station and sub-surface layers at 50 and 100 cm thick layers, respectively.

In spring, soil is saturated due to snowmelt in both layers (Figures 6 and 7). The rapid increase of soil moisture in the beginning of April is fabricated by the TDR- and FDR-devices during ground frost melting. The instrument gradually reaches the real volumetric water content of the unfrozen soil. In the summer, the surface layer which is directly affected by rainfall and evaporation is rather fast responding and shows larger variation in volumetric water content.
content than the sub-surface layer. The increase of the sub-surface layer water content depends on the infiltration rate from the surface layer and begins after the surface layer reaches the field capacity.

The individual parameter values for sandy, silty and clayey soil types were found separately for both layers in each soil moisture station, because the soil properties may vary between the soil horizons. For instance, in the Rautavaara station the surface layer soil porosity is clearly around 0.25 m$^3$/m$^3$ whereas in the sub-surface layer the maximum soil moisture values indicate a porosity value closer to 0.20 m$^3$/m$^3$. In Tarvasjoki, the best fit with observed soil moisture was obtained with surface and sub-surface layer porosity values of 0.40 and 0.70 m$^3$/m$^3$, respectively. Due to lack of observations in the organic, till and rocky soil types, the parameter values were obtained from the literature (Andersson & Wiklert 1972; Laine-Kaulio 2011). Some of the fitted parameter values for soil porosity and field capacity shown in Table 3 were remarkably below the reference values. This is partly due to the fact that one value is unable to represent the heterogeneity of the soil properties but also implies that the parameters still require some freedom of variation due to conceptual formulation of the 2-layer model.

**RESULTS**

**Soil moisture simulations in various soil types**

The large spatial variation in soil properties and consequently in surface soil moisture content within a SMOS
pixel requires that the validation of the SMOS data should be based on multiple *in situ* soil moisture observation stations and/or a distributed soil moisture modeling approach. This section presents the simulations of the areal mean soil moisture of the most dominant soil types in the Hypöistenkoski catchment.

The 2-layer model mean soil moisture simulations in the Hypöistenkoski catchment for the most common soil

### Table 3

<table>
<thead>
<tr>
<th>Soil type</th>
<th>Soil porosity (m³/m³)</th>
<th>Field capacity (m³/m³)</th>
<th>Layer depth (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Surface layer</td>
<td>Sub-surface layer</td>
<td>Surface layer</td>
</tr>
<tr>
<td>Sand</td>
<td>0.14 (0.47)</td>
<td>0.10 (0.47)</td>
<td>0.06 (0.03)</td>
</tr>
<tr>
<td>Silt</td>
<td>0.25 (0.46)</td>
<td>0.21 (0.46)</td>
<td>0.18 (0.18)</td>
</tr>
<tr>
<td>Clay</td>
<td>0.40 (0.62)</td>
<td>0.70 (0.62)</td>
<td>0.20 (0.41)</td>
</tr>
<tr>
<td>Organic</td>
<td>0.60 (0.95)</td>
<td>0.80 (0.95)</td>
<td>0.20 (0.33)</td>
</tr>
<tr>
<td>Till</td>
<td>0.45 (0.45)</td>
<td>0.30 (0.34)</td>
<td>0.20 (0.20)</td>
</tr>
<tr>
<td>Rock</td>
<td>0.25</td>
<td>0.21</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Figure 7 | Soil moisture observations from Rautavaara station in eastern Finland and fitted soil moisture simulations with the 2-layer model from surface (5 cm depth) and sub-surface (down to 100 cm depth) layers.
types in both layers are shown in Figure 8. Fine particle soil types (clay, silt) have a large surface area and they have a larger water holding capacity than coarse particle soil types. Organic soil types can hold a large amount of water because of the high porosity (0.60 m$^3$/m$^3$ in the surface layer, Table 3) in comparison to the field capacity (0.20 m$^3$/m$^3$). Therefore, organic and fine particle soil types cause longer delays to the runoff compared to other soil types. Coarse particle soil types such as sandy soil types have poor water retaining qualities and they dry up easily. Rocky soil types are shallow, consisting of 2 and 16 cm thick surface and sub-surface layers and they are easily drained, producing fast runoff. Till as a heterogeneous soil type containing a mixture of fine and coarse particle soil types has a relatively high water holding capacity. The porosity and field capacity values of till are close to the values of clay in the surface layer, and to the values of silt in the sub-surface layer.

In the spring of 2010, the surface layer water content of organic soil types varied largely as a response to dry and
wet periods. This delayed the infiltration to the sub-surface layer resulting in a smaller increase in water content of organic soil types than clayey soil types in the sub-surface layer. Sub-surface organic soil types are overlaid by surface layer organic soil types whereas sub-surface clayey soil types are covered by various soil types, sandy, clayey, silty and organic soil types (Table 2). During the dry summer period water content in the coarse soil types decreased to a very low level, but fine particle soil types and organic soil types hold still around 0.1 m³/m³ water content in the surface layer.

Because of the shallow surface and subsurface layers of the rocky soil types, a major part of the 30 mm precipitation event at the beginning of August infiltrated through the surface layer to the sub-surface layer producing runoff and percolation to the groundwater storage. Other surface layer soil types had a larger soil moisture deficit as a result of the preceding dry period. Thus the water content only slightly exceeded the field capacity values keeping the infiltration rate low. Thus the increase in the sub-surface layer water content of clayey and organic soil types remained too small to produce runoff and increased soil moisture content after the event almost totally evaporated from the study area.

In September, when rainfall amounts were much larger and evapotranspiration was smaller, more water infiltrated to the sub-surface layer. Over 100 mm accumulated precipitation in less than 2 weeks increased the water content of the surface layer close to the saturation level and infiltration became substantial. The sub-surface water content exceeded field capacity and runoff was produced.

**SMOS level 2 data validation**

This section presents the results of the SMOS level 2 data validation against the surface soil moisture simulations of the 2-layer model. The areal weighted average surface and sub-surface layer simulations of soil moisture content in the Hypöistenkoski catchment with filtered SMOS level 2 data and discharge simulations at the Hypöistenkoski observation station are shown in Figures 9–11. The unfiltered, RFI contaminated SMOS data are also shown in gray. The discharge simulations with 2-layer soil moisture model are compared to observations and simulations with the operational soil moisture model of the WSFS.

The summer of 2010 was warm and dry, which caused lower mean soil moisture values than in the following years. The drying up of the soil began in late June and the 2-layer model surface soil moisture content and SMOS observations reached the first minimum at the end of July and the second at the beginning of September (Figure 9).

In 2011, the first month after the snowmelt was exceptionally dry, which explains the lower than average soil moisture content compared to the summer of 2012. A total of 30 days’ precipitation between April 15 and May 15 only amounted to 12 mm. The minimum surface soil moisture was reached in mid-June. The minimum value in the 2-layer model was 0.10 m³/m³ and in the SMOS observations was 0.04 m³/m³ (Figure 10).

In 2012, the minimum surface soil moisture content (0.13 m³/m³) in the 2-layer model occurred in late August, when the SMOS observations matched perfectly with the model simulations. However, the minimum value of SMOS observations (0.09 m³/m³) was already recorded in mid-July, when the 2-layer model surface soil moisture was 0.19 m³/m³ (Figure 11).

The response of the surface layer soil moisture content to the precipitation events in the study period were most remarkable, when the 5 day precipitation totals exceeded 20 mm after a dry period. These events occurred on August 3–7, 2010, September 10–14, 2010, June 13–17, 2011, June 1–5, 2012 and August 18–22, 2012. The 2-layer model surface soil moisture simulations and SMOS observations show similar increases as a response to the precipitation events (Figures 9–11). In most of the events, the surface soil moisture increased temporarily close to the areal mean maximum soil moisture content (0.31 m³/m³), but the sub-surface layer moisture content remained too low to produce significant runoff. The most remarkable runoff peaks in the summer seasons occurred in April–May after snowmelt and in August–September due to lower evapotranspiration.

The RFI contamination was detected on only a few occasions. However, in August–September 2012, it affected the SMOS data remarkably and similar filling of the surface
soil moisture storage as in year 2010 was thus not realized by the SMOS data.

Table 4 shows the areal precipitation and evapotranspiration sums and the 2-layer model mean surface soil moisture simulation in the Hypöistenkoski catchment with corresponding mean SMOS observations for the whole study period and separately for each season. $R^2$ between the SMOS level 2 data and simulated surface soil moisture for the whole study period was 0.77 and RMSD was 0.76 (Table 5). RMSD was remarkably higher and $R^2$ correspondingly lower in 2011 due to more remarkably dry bias of the SMOS observations (Figure 10). The correlation coefficient was strongest in the relatively dry summer of 2010, when only a few precipitation events occurred. The variation in the efficiency criteria in different summers was evidently caused by weather conditions, but also the different processor versions of the SMOS observations between 2010–2011 and 2012 affected the results between these years.

Figure 9 | Precipitation, 2-layer model mean surface and sub-surface layer soil moisture simulations with SMOS level 2 data in the Hypöistenkoski catchment. Discharge simulations with the 2-layer and operational soil moisture models in WSFS and observations in the Hypöistenkoski station in the summer of 2010.
Discharge

The goodness of the 2-layer model was assessed by comparing the discharge simulations with the operational soil moisture model in the WSFS and discharge observations in the calibration period 2005–2012. The $R^2$ for both model versions were calculated for the discharge observation stations at Hypöistenkoski and downstream site Halinen, located in the city of Turku (Table 5). The $R^2$ values with the 2-layer model were slightly deteriorated compared to the operational model version at both stations, from 0.80 to 0.76 in Hypöistenkoski and from 0.88 to 0.81 in Halinen.

The springtime observed maximum discharges in Hypöistenkoski were 38.0 m$^3$/s in 2010, 54.0 m$^3$/s in 2011 and 35.6 m$^3$/s in 2012. The WSFS with the 2-layer model overestimated the snowmelt discharge peaks on average by 16%, whereas with the operational soil moisture model spring flood peaks were overestimated by 7%. The timing of the spring flood was good ($\pm$1 day) in both model versions. The 30 day discharge sums were overestimated on
Figure 11 | Precipitation, 2-layer model mean surface and sub-surface layer soil moisture simulations with SMOS level 2 data in the Hypöistenkoski catchment. Discharge simulations with the 2-layer and operational soil moisture models in WSFS and observations in the Hypöistenkoski station in the summer of 2012.

Table 4 | Summer season precipitation and evapotranspiration sums, mean SMOS level 2 observation and 2-layer model surface and sub-surface layer soil moisture in the Hypöistenkoski catchment in 2010–2012

<table>
<thead>
<tr>
<th></th>
<th>Precipitation total (mm)</th>
<th>Evapotranspiration total (mm)</th>
<th>SMOS mean soil moisture (m$^3$/m$^3$)</th>
<th>2-layer model mean surface soil moisture (m$^3$/m$^3$)</th>
<th>2-layer model mean sub-surface soil moisture (m$^3$/m$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer 2010  (15.4–15.9)</td>
<td>260</td>
<td>288</td>
<td>0.10 (77 d)</td>
<td>0.16</td>
<td>0.14</td>
</tr>
<tr>
<td>Summer 2011  (15.4–15.9)</td>
<td>371</td>
<td>351</td>
<td>0.12 (126 d)</td>
<td>0.19</td>
<td>0.17</td>
</tr>
<tr>
<td>Summer 2012  (15.4–15.9)</td>
<td>330</td>
<td>327</td>
<td>0.17 (114 d)</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td>Summer seasons average 2010–2012 (15.4–15.9)</td>
<td>320</td>
<td>322</td>
<td>0.15 (317 d)</td>
<td>0.19</td>
<td>0.18</td>
</tr>
</tbody>
</table>
average by 8% with the 2-layer model and only by 2% with the operational WSFS model.

The maximum observed discharges in the summer months (June–August) were small (1.6–3.5 m$^3$/s) and usually slightly overestimated with both soil moisture models. In autumn (September–November), the discharges increased due to decreased evapotranspiration and the maximum observed discharges varied between 20.0 and 38.6 m$^3$/s. The WSFS with the 2-layer model underestimated the autumn discharge peaks on average by 20% and with the operational model by 29%. The observed mean discharges in the summer months were low, 0.30–0.70 m$^3$/s, and were overestimated with the 2-layer model by a factor of 1.5–1.7.

The operational soil moisture model in WSFS produced lower mean discharges and smaller bias to the observations. In autumn, the observed mean discharges were 1.4–4.9 m$^3$/s and overestimated by a factor of 1.5–1.9 with the 2-layer model and by 1.1–1.4 with the operational soil moisture model.

### DISCUSSION

The SMOS level 2 data validation against the 2-layer soil moisture model in the Hypöistenkoski catchment showed reasonably good results in 2010 and 2012. In 2011, the dry bias of the SMOS data was more remarkable. The difference in bias between years 2011 and 2012 is most likely caused by the change in the dielectric mixing model of the SMOS level 2 processor. The difference between years 2010 and 2011 could be explained by the different weather conditions. Year 2010 was remarkably drier than year 2011. The penetration depth of the L-band radiometer is dependent on the surface soil moisture content becoming shallower with wet surfaces. The surface layer of the 2-layer soil moisture model is 10 cm and considerably deeper than the SMOS penetration depth (3–5 cm), which may explain the better results in 2010 due to a drier summer than in 2011.

The decline of the SMOS observations during the dry periods and the timing of the minimum values of the surface soil moisture were mostly consistent with the 2-layer surface soil moisture simulations. The larger bias in SMOS data in 2011 compared to years 2010 and 2012 could also be partly explained by earlier snowmelt, more efficient evaporation after the snowmelt and/or lower initial soil moisture content in the spring than produced by the 2-layer model. However, the summer season precipitation total increased to 359 mm during the wetter periods after mid-June and surface soil moisture content was expected to rise to the same average level as in 2012, but the SMOS observations still showed strong bias (0.07 m$^3$/m$^3$) between August 15 and September 15. Therefore the dry bias in the SMOS observations in 2011 and consequently poorer RMSE and $R^2$ values than in 2012 is estimated to depend on the different processor versions of the SMOS level 2 data from 2011 and 2012.

In the new SMOS level 2 processor version v5.51, the Dobson dielectric model was replaced by Mironov formulation, which seems to decrease the bias between SMOS level 2 data and the 2-layer model surface soil moisture simulations between years 2011 and 2012. On the other hand, r decreased in the new processor version and careful studies of the reprocessed data with the Mironov model for 2010 and 2011 should be made before reaching final conclusions.

The major uncertainty in the SMOS validation in this study is brought by the accuracy of the 2-layer soil moisture model of the WSFS. The approach of using the hydrological model instead of in situ soil moisture measurements was
chosen due to lack of soil moisture observation stations in the test site. The 2-layer model is a conceptual model and requires calibration to the local conditions. In this study the most important soil moisture parameters (soil porosity and field capacity) were fixed by fitting them against the observations at soil moisture stations outside the test site. The calibrated parameters of the 2-layer soil moisture model and the parameters of the other sub-models of the WSFS were calibrated against discharge observations in the Hypöistenkoski station with individual parameter values in each sub-basin. For this reason, the test site does not correspond exactly with the footprint of the SMOS observations, but the data of the closest DGG nodes cover a slightly larger area than the Hypöistenkoski catchment. The RMSD between the three closest SMOS nodes was 0.03 m$^3$/m$^3$. Thus this approach leaves some uncertainty to the validation of the SMOS level 2 data against the surface soil moisture simulations. Also, the difference between the penetration depth of the SMOS product (3–5 cm) and the surface soil moisture layer of the 2-layer model (10 cm) brings uncertainty to SMOS validation as discussed above. The depth of the surface layer of the 2-layer model was chosen to be 10 cm, due to the 1 day time-step of the WSFS. The shallower surface layer would have caused unrealistic behavior in the surface soil moisture variation.

The benefit of this approach is that the SMOS validation against the HBV-based hydrological model brings important information about the applicability of the SMOS level 2 data in hydrological model assimilation and flood forecasting. This has been one of the ultimate goals of the SMOS mission as well as the main emphasis of this study.

**CONCLUSIONS**

The satellite soil moisture observations have recently developed after a new technology has enabled the use of L-band microwave radiometers. Compared to previously used C-band radiometers, the new technology adopted in the SMOS satellite decreases the influence of vegetation, atmosphere and surface roughness on the emitted brightness temperature signal from the Earth surface, which can be related to the surface soil moisture content (Kerr et al. 2010). The real-time monitoring of the large scale spatial and temporal variations of the soil moisture content yields valuable information for the data assimilation of the hydrological models used in flood forecasting.

This study focused on the development of the 2-layer soil moisture model in WSFS for validation of the SMOS level 2 data. The new model utilizes the most detailed national scale information of the soil properties for producing the most reliable spatial and temporal variation of the areal mean soil moisture in the study area. The test site was selected after examining the SMOS level 2 data coverage in Finland. A continuous data series with long-term discharge observations was found from a catchment of Hypöistenkoski in the Aurajoki watershed, located in Southwest Finland.

The results of the SMOS level 2 observations were in good agreement with the surface soil moisture simulations of the 2-layer model. The Nash–Sutcliffe efficiency coefficient for the areal mean surface soil moisture was 0.77. Dry bias of the SMOS data was detected throughout the whole study period, being most significant in 2011. The change of dielectric mixing model in SMOS data processor between 2011 and 2012 seems to decrease the bias. The RMSDs in different years varied between 0.064 and 0.088 and correlation coefficients between 0.42 and 0.76 and were similar in level to several studies in which the SMOS observations have been compared to in situ soil moisture observations in different parts of the world (Albergel et al. 2012; Jackson et al. 2012). The similar response of the SMOS observations and the surface soil moisture simulations after dry periods, when the uncertainties in the soil moisture deficit of the watershed models are at their largest, encourages the further development and utilization of the soil moisture satellite products in hydrological models.

Despite the promising results presented in this study, the extensive use of the satellite soil moisture observations in the WSFS will still require further development of the 2-layer soil moisture model, which is computationally expensive compared to the current operational soil moisture model of the WSFS. The slight deterioration in the hydrological model performance to reproduce the observed discharges implies that the runoff production of the 2-layer model demands improvements before the model can be taken into operational use in Finland. On the other hand, the
2-layer model is expected to improve the soil moisture simulations of the WSFS. The main benefits of the 2-layer model are: (1) by applying the Finnish soil database the model is able to produce more detailed and reasonable spatial variation for soil moisture; and (2) by including the surface soil moisture layer the model simulations produce reliable data for comparison and validation of the satellite soil moisture products.

The national scale validation of the SMOS soil moisture observations and their utilization in the data assimilation of the hydrological models and flood forecasting also faces challenges in the development of the radiative transfer models of the SMOS level 2 soil moisture processors. The SMOS spatial resolution is coarse for capturing the soil moisture content over various land use classes and soil types. Also, the attenuation of the soil emission signal by vegetation cover leads to poor retrieval of the level 2 data from forested areas. In Finland, 72% of the land area is covered by forests and 10% by inland water bodies, including more than 2600 lakes with an area over 1 km². For these reasons, the SMOS level 2 data are still only obtained in some extensively cultivated areas and also in some parts of northern and eastern Finland, where the wetlands and sparse trees are the dominant land use classes. During the Cal/Val campaigns, the different parameterizations of the radiative transfer and dielectric mixing models are considered and tested for the SMOS soil moisture processor to improve the brightness temperature modeling of mixed pixels with vegetated surfaces, bare soils and water bodies (Schlenz et al. 2011; Gherboudj et al. 2012; Kerr et al. 2012).

The development of the parameterization for the boreal coniferous forests could be expected to fill some of the gaps in the level 2 data coverage in Finland. Also, the finer resolution observations to avoid mixed pixel signals in lake areas and information from the sub-surface layer soil moisture content to cover hydrologically important soil moisture variation are hoped to be achieved in future satellite soil moisture missions.

ACKNOWLEDGEMENTS

This study was carried out in the Freshwater Centre of the Finnish Environment Institute (SYKE). The research has been funded by the Finnish Funding Agency for Technology and Innovation (TEKES) and Academy of Finland. The work was part of the projects ‘SMOS Calibration and Validation Activities in Northern Areas (NORA)’ and ‘Advanced methods for global monitoring of soil moisture and ocean salinity’ (GlobSMOS). The authors would like to thank Paula Havu, Antti Taskinen and the reviewers for their detailed and valuable comments, Risto Mäkinen for establishing and maintaining the soil moisture observation stations and Juho-Pekka Vehviläinen for his contribution to the early development work of the 2-layer soil moisture model.

REFERENCES


First received 19 December 2012; accepted in revised form 5 September 2013. Available online 17 October 2013