Assimilation of ASCAT near-surface soil moisture into the French SIM hydrological model

C. Draper¹, J.-F. Mahfouf¹, J.-C. Calvet¹, E. Martin¹, and W. Wagner²

¹Météo-France, CNRM/GAME, URA1357, CNRS, Toulouse, France
²Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Wien, Austria

*now at: NASA GSFC/GEST, Greenbelt, USA

Received: 11 May 2011 – Accepted: 16 May 2011 – Published: 1 June 2011

Correspondence to: C. Draper (clara.draper@nasa.gov)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

The impact of assimilating near-surface soil moisture into the SAFRAN-ISBA-MODCOU (SIM) hydrological model over France is examined. Specifically, the root-zone soil moisture in the ISBA land surface model is constrained over three and a half years, by assimilating the ASCAT-derived surface degree of saturation product, using a Simplified Extended Kalman Filter. In this experiment ISBA is forced with the near-real time SAFRAN analysis, which analyses the variables required to force ISBA from relevant observations available before the real time data cut-off. The assimilation results are tested against ISBA forecasts generated with a higher quality delayed cut-off SAFRAN analysis. Ideally, assimilating the ASCAT data will constrain the ISBA surface state to correct for errors in the near-real time SAFRAN forcing, the most significant of which was a substantial dry bias caused by a dry precipitation bias. The assimilation successfully reduced the mean root-zone soil moisture bias, relative to the delayed cut-off forecasts, by close to 50% of the open-loop value. The improved soil moisture in the model then led to significant improvements in the forecast hydrological cycle, reducing the drainage, runoff, and evapotranspiration biases (by 17%, 11%, and 70%, respectively). When coupled to the MODCOU hydrogeological model, the ASCAT assimilation also led to improved streamflow forecasts, increasing the mean discharge ratio, relative to the delayed cut off forecasts, from 0.68 to 0.76. These results demonstrate that assimilating near-surface soil moisture observations can effectively constrain the SIM model hydrology, while also confirming the accuracy of the ASCAT surface degree of saturation product. This latter point highlights how assimilation experiments can contribute towards the difficult issue of validating remotely sensed land surface observations over large spatial scales.
1 Introduction

The last decade has seen considerable interest in the possibility of improving hydrological and meteorological model forecasts by assimilating remotely sensed near-surface soil moisture data (Crow and Wood, 2003; Reichle and Koster, 2005; Balsamo et al., 2007; Drusch, 2007). This interest has motivated recent advances in soil moisture remote sensing, from both purpose designed L-band sensors (Kerr et al., 2001; Entekhabi et al., 2004), and preexisting suboptimal C- and X-band sensors (Wagner et al., 1999; Owe et al., 2001). As a result remotely sensed near-surface soil moisture data are available for the first time with sufficient quality and legacy to test their use in real world modeling problems, and in particular EUMETSAT is now providing the first operationally supported near-surface soil moisture product. This product, derived from Advanced SCATterometer (ASCAT) microwave radiometer observations, is being assimilated into the UK Met Office’s operational NWP model (Dharssi et al., 2011), and will soon be introduced into ECMWF’s operational IFS (De Rosnay et al., 2009).

This study seeks to determine whether an operational hydrological model, specifically Météo-France’s SAFRAN-ISBA-MODCOU (SIM) model, might also benefit from the assimilation of these ASCAT soil moisture observations. SIM is a three-part model, consisting of (i) a low-level atmospheric analysis (the Système d’analyse fournissant des renseignements atmosphériques à la neige; SAFRAN), which provides the forcing for (ii) a land surface model (Interactions between Surface, Biosphere, and Atmosphere; ISBA), which in turn provides surface moisture fluxes to (iii) a hydrogeological model (MODCOU), which provides forecasts of aquifer levels and streamflow. SIM is run operationally at Météo-France in near-real time, with a three hour data cut-off, allowing observations from approximately 1200 automatic weather stations to be used in the SAFRAN atmospheric analysis. The output from SIM is used for water resource monitoring (e.g., see http://climat.meteofrance.com/chgt_climat2/bilans_climatiques), and to provide the initial conditions for an ensemble streamflow prediction system which will be used for operational flood forecasting (Thirel et al.,...
A second stream of SIM is run once a month, using additional observations from 3000 climatological observing stations that report once-monthly.

Evaluation studies based on the delayed cut-off (once-monthly) SIM stream have demonstrated that the SAFRAN analysis provides accurate meteorological variables for forcing the ISBA land surface model (Quintana-Seguí et al., 2008; Vidal et al., 2010), resulting in accurate forecasts of the spatial and temporal variability of observed water fluxes and streamflow (Habets et al., 2008). In contrast, the near-real time (operational) SAFRAN analysis generates precipitation forcing with substantial errors (Quintana-Seguí et al., 2008), due to the underestimation of precipitation by automatic weather station rain gauges (Canellas, 2005), and the tendency for the sparser observation network to detect fewer rain events. To prevent the errors in the near-real time SAFRAN precipitation forcing from accumulating in the near-real time SIM model states, the relevant land surface variables are reset once a month using values from the delayed cut-off stream.

This study is motivated by the possibility that the near-real time SIM land surface states could be more effectively constrained by assimilating the ASCAT near-surface soil moisture observations. This possibility is tested by constraining the root-zone soil moisture in the near-real time SIM model by assimilating three and a half years of the ASCAT observations with a Simplified Extended Kalman Filter (SEKF). The results of this experiment are examined (i) to establish the impact of the assimilation on the modeled hydrological cycle and surface moisture storage; and (ii) to determine whether the assimilation corrected the SIM forecasts for the errors in the near-real time SAFRAN forcing. The results of the assimilation are assessed against SIM forecasts generated with the delayed cut-off SAFRAN atmospheric analysis. In addition to indicating that the near-real time SIM model could benefit from the ASCAT data, a positive result from this experiment would also confirm that the ASCAT soil moisture observations accurately detect near-surface soil moisture, and that observations of near-surface soil moisture contain enough information to effectively constrain a hydrological model.
2 Data and methods

2.1 The SIM hydrological model

SIM (Habets et al., 2008) is run at approximately 0.07° resolution over France. The SAFRAN (Quintana-Seguí et al., 2008) analyses of the low-level atmosphere are performed every 6 h, however the 6-hourly analyses are interpolated to hourly timesteps before being used to force ISBA. ISBA (Noilhan and Planton, 1989; Noilhan and Mahfouf, 1996) then outputs hourly estimates of the land surface states, and exchanges of heat and moisture between the low-level atmosphere, vegetation, and soil. The three layer version (Boone et al., 1999) of ISBA is used in SIM. Finally, MODCOU (Ledoux et al., 1989) is run once daily, to compute the daily evolution of aquifer storages and three-hourly streamflow forecasts.

2.2 ASCAT remotely sensed soil moisture

ASCAT is a real aperture backscatter radar observing at 5.255 GHz (C-band), with approximately 25 km resolution. It orbits on EUMETSAT’s Meteorological Operational (MetOp) satellite, which was launched in 2007 to replace the ageing European Remote Sensing (ERS) satellites. MetOp is in a sun-synchronous orbit, with equator crossing times of approximately 09:30 LT (LT = local time, descending overpass) and 21:30 LT (ascending overpass). ASCAT provides good spatial coverage, and observes approximately 80% of the globe each day.

Soil moisture observations are derived from ASCAT radar backscatter coefficients using the empirical change detection approach developed at the Vienna University of Technology (TU-Wien) by Wagner et al. (1999). This approach is based on the assumption that over a long data record, the highest observed reflectivity can be equated to the maximum soil moisture, while the lowest reflectivity can be equated to the minimum soil moisture, and a linear relationship can be used to interpolate the values in between. For full details refer to Wagner et al. (1999) and Naeimi et al. (2009).
The output from the change detection method is an observation loosely referred to as the “surface degree of saturation” (SDS), and defined by:

$$\text{SDS} = \frac{(w_{sfc} - w_{\text{min}})}{(w_{\text{max}} - w_{\text{min}})}$$

where $w_{sfc}$ is the moisture in the near-surface soil layer, and $w_{\text{min}}$ and $w_{\text{max}}$ are the minimum and maximum $w_{sfc}$ occurring at that location. C-band microwave observations are sensitive to soil moisture in a thin surface layer, of up to 1 cm depth, hence the SDS relates only to this thin surface layer. The SDS is reported exclusively in percentage units here, to avoid confusion with volumetric ($m^3 m^{-3}$) measures of soil moisture. Note that the SDS is localised, in that equivalent values at different locations do not necessarily indicate equivalent soil moisture, due to spatial differences in the soil moisture bounds.

In response to differences in the ERS and ASCAT observing behaviour, the change detection model parameters used in the ASCAT retrieval algorithm were recently updated (Wagner et al., 2010) to use parameters derived from the ASCAT data record, rather than the ERS values that were initially adopted for ASCAT (Naeimi et al., 2009). This update has further improved the soil moisture observations retrieved from ASCAT, resulting in excellent agreement with other soil moisture estimates. For example, Brocca et al. (2010b) found correlations and anomaly correlations of 0.67 and 0.58 against in situ data in Italy (one site, 13 months) and of 0.75 and 0.60 against modeled estimates (six sites, 13 months), while Brocca et al. (2010a) calculated a mean correlation of 0.65 against in situ data throughout Europe (14 sites, 2 years at most locations) and 0.76 against modeled estimates (20 sites, 2 years). Additionally, Draper et al. (2010) found a mean correlation and anomaly correlation of 0.70 and 0.62 against in situ data in France (12 sites, 3.5 years).

The ASCAT level 2 surface degree of saturation (SM OBS1) product, supplied by TU-Wien at 0.125° resolution, has been used here. This product includes the aforementioned update to the change detection model parameters. Since scatterometer observations taken in the evening have been shown to produce less accurate soil moisture fields than early morning observations (Albergel et al., 2009), and since there is
a spurious relationship in ISBA between the near-surface and root-zone soil moisture during the daytime (Draper et al., 2011), only the descending overpass ASCAT observations have been assimilated.

Observations of densely vegetated regions have been removed, based on an ASCAT Estimated Soil Moisture Error (provided with the ASCAT data) threshold of 20%. Additionally, observations with an urban fraction greater than 15% in the ECOCLIMAP database (Masson et al., 2003) have been removed, as have observations with a topographic complexity flag (provided with the ASCAT data) greater than 15%, and/or a wetland fraction (provided with the ASCAT data) greater than 5%.

The remaining data were projected from the 0.125° Discrete Global Grid used by TU-Wien to the ~0.07° SIM grid using a nearest neighbour approach. Observations of frozen surface conditions, temporary surface water, or snow-cover were initially identified based on the probabilistic surface state flag provided with the ASCAT data. However an initial investigation revealed that this probabilistic method did not reliably remove the occurrence of surface freezing. Frozen surface conditions manifest in the data as anomalously low soil moisture observations, which can have a significant detrimental impact on the assimilation. Consequently, an additional screening for frozen surface conditions has been applied, by excluding the ASCAT data whenever SIM forecasts nonzero frozen near-surface soil moisture.

Where the above data processing resulted in less than 100 observations for a model grid cell (less than 10% coverage over the 3.5 year study period) the remaining data have not been used. Finally, the ASCAT observations have been normalised to match the CDF of the SIM_NRT near-surface soil moisture, as described in Sect. 3.1.

2.3 The Simplified Extended Kalman Filter (SEKF)

The SEKF was initially formulated by Balsamo et al. (2004) and Mahfouf et al. (2009). The Extended Kalman Filter equations for the $i$-th model state forecast and update at time, $t_j$, are:
\[
x^b(t_i) = M_{i-1} \left[ x^a(t_{i-1}) \right] \tag{2}
\]

and

\[
x^a(t_i) = x^b(t_i) + K_i \left( y^o_i - \mathcal{H}_i \left[ x^b(t_i) \right] \right) \tag{3}
\]

where \( x \) indicates the model state, and \( y \) is the observation vector. The superscripts \( a, b, \) and \( o \) indicate the analysis, background, and observations, respectively. \( M \) is the nonlinear state forecast model, and, \( \mathcal{H} \) is the nonlinear observation operator. \( K \) is the Kalman gain, given by:

\[
K_i = P H^T \left( H_i P H^T_i + R \right)^{-1} \tag{4}
\]

\( H \) is the linearisation of \( \mathcal{H} \), and \( P \) and \( R \) are the covariance matrices of the model background and observation errors, respectively. The traditional EKF also evolves the background error covariance matrix through a forecast and analysis cycle, while for the Simplified EKF the same \( P \) matrix is used at the start of each assimilation cycle. Draper et al. (2009) found that for the assimilation of near-surface soil moisture into ISBA, the analysed soil moisture generated by the EKF and the SEKF are not substantially different, and hence the simplified EKF was used here.

Since ISBA does not model horizontal exchanges, the assimilation is performed as an individual 1-D assimilation at each model grid. For each model grid cell, ISBA partitions soil moisture into three variables: the near-surface soil moisture (\( w_1 \); defined over the depth of bare soil evaporation), the root-zone soil moisture (\( w_2 \); defined over the depth of transpiration), and the deep-layer soil moisture (\( w_3 \); representing long term surface moisture storage). In these experiments the state update vector included \( w_1 \) and \( w_2 \), and the ASCAT observations were assumed to be the observation-equivalent of the model \( w_1 \). An integration of the forecast model was used as the observation operator, and \( \mathcal{H} \) has been linearised by finite differences. The impact of \( w_2 \) on \( w_1 \) increases with time, and a 24-h forecast length was chosen for the observation operator, as a compromise between a long enough forecast length that \( w_2 \) has a

5434
reasonable impact on $w_1$, and short enough that it can be linearised without significant loss of accuracy.

**The assimilation experiment**

The assimilation experiment consisted of three simulations of SIM, each summarised in Table 1. For the assimilation of the ASCAT data, referred to as SIM_ASCAT, ISBA was forced with the near-real time (NRT) SAFRAN analysis. The performance of the assimilation was measured by comparing the resultant ISBA forecasts to an open-loop ISBA simulation generated with the more accurate delayed cut-off (DEL) SAFRAN analysis, referred to as SIM_DEL. The improvements generated by SIM_ASCAT were benchmarked against the performance of an ISBA open-loop forced with the NRT SAFRAN analysis, and referred to as SIM_NRT.

Each simulation was conducted for the 3.5 years from January 2007. For the SIM_ASCAT and SIM_NRT experiments, ISBA was initialised and forced with archived fields from Météo-France’s near-real time SIM chain, while the SIM_DEL ISBA output was extracted directly from Météo-France’s archives.

For the SIM_ASCAT simulation, the observation error covariances were based on the error estimates provided with the ASCAT SDS data. This is the first study to make use of these error estimates, and an initial investigation confirmed that they have some skill in detecting errors in the ASCAT soil moisture (Draper et al., 2010). The ASCAT error estimates are provided in the SDS units and are relative to the ASCAT soil moisture climatology. Consequently, the errors were linearly rescaled to be consistent with the model soil moisture climatology, by preserving the ratio between the original SDS and ASCAT error estimates at each grid cell. The original ASCAT SDS error estimates ranged between 3.5 and 20% (since observations with an error greater than 20% were screened out), with a median value of 9.0%. The rescaled error estimates ranged between 0.02 and 0.20 m$^3$ m$^{-3}$, with a median value of 0.05 m$^3$ m$^{-3}$. This median value is consistent with errors typically expected for remotely sensed soil moisture, and is
slightly higher than the root mean square error of 0.04 m$^3$ m$^{-3}$ estimated for ASCAT at the French SMOSMANIA monitoring sites by Draper et al. (2010).

The background error covariance matrix was based on that used by Draper et al. (2011) to assimilate AMSR-E near-surface soil moisture observations into a two-layer version of ISBA: $P$ was assumed diagonal, and the $w_1$ and $w_2$ error standard deviations were set at $0.5 \times (w_{fc} - w_{wilt})$ and $0.2 \times (w_{fc} - w_{wilt})$, where $w_{fc}$ and $w_{wilt}$ are the soil moisture at field capacity and wilting, respectively. These values generate mean error standard deviations close to 0.04 and 0.02 m$^3$ m$^{-3}$, for $w_1$ and $w_2$ respectively.

### 3 Results and discussion

Before presenting the results of assimilating the ASCAT data (Sect. 3.3), the ASCAT SDS are directly compared to the near-surface soil moisture from SIM\_NRT (Sect. 3.1). Additionally, the SIM\_NRT and SIM\_DEL soil moisture fields are compared (Sect. 3.2), to establish the impact of the NRT SAFRAN errors on the modeled soil moisture.

### 3.1 Comparison of SIM and ASCAT $w_1$

Previous studies evaluating the ASCAT near-surface soil moisture within the SIM France domain (Albergel et al., 2009; Brocca et al., 2010a; Draper et al., 2010) have focused on a handful of locations where in situ data are available. Consequently, the quality of the ASCAT SDS data across France has been checked here by direct comparison to SIM\_NRT near-surface soil moisture fields. The ASCAT and SIM $w_1$ show the expected strong association in their temporal behaviour. Figure 1 shows maps of the correlation ($r_{abs}$) and anomaly correlation ($r_{anm}$; defined relative to the 31 day moving average) between the SIM\_NRT $w_1$ and ASCAT SDS time series at each grid cell. Both statistics were consistently very high. For $r_{abs}$, the mean value across France was 0.68, and
81% of the SIM model grids had a value greater than 0.60, while for $r_{\text{anm}}$, the mean was 0.62, and 68% of grids had a value greater than 0.60.

Both maps in Fig. 1 have similar spatial patterns, in terms of the regions of high and low values, including several locations with low correlations (<0.3) in regions of mountainous terrain. In each case these were adjacent to locations where the ASCAT data were screened out due to complex terrain and/or vegetation cover, suggesting that the low correlations were associated with ASCAT errors, and the parameters used to screen-out the ASCAT data were insufficiently rigorous. Since the SIM and ASCAT soil moisture are derived using totally independent methods, the strong temporal agreement between them provides very strong evidence that both are accurately detecting the true near-surface soil moisture dynamics.

To compare the absolute values from SIM and ASCAT, the ASCAT SDS have been converted to volumetric soil moisture using Eq. (1), with the soil moisture bounds from SIM.NRT. There were substantial differences between the absolute values of the resulting $\omega_1$ from ASCAT and SIM. For the 3.5 year study period, the mean across the domain of the absolute difference at each grid cell was 0.016 m$^3$ m$^{-3}$. On average, ASCAT was drier than SIM.NRT, and the mean $\omega_1$ values were 0.229 m$^3$ m$^{-3}$ for ASCAT and 0.236 m$^3$ m$^{-3}$ for SIM.NRT, while the mean temporal standard deviations were 0.070 and 0.073 m$^3$ m$^{-3}$, respectively.

While the presence of systematic differences between modeled and observed soil moisture is well established (Reichle et al., 2004), it is disappointing to see these differences persist even after the ASCAT SDS has been fairly strongly constrained to the SIM.NRT climatology, by fitting it to the SIM $\omega_1$ range at each grid cell. These persistent differences may reflect fundamental differences in the quantities defined by the modeled and the observed soil moisture (see Koster et al., 2009). Alternatively, they may also point to deficiencies in the change detection soil moisture retrieval method used for ASCAT, errors in the modeled soil moisture, or inconsistencies between the methods used to define the upper and lower bounds for each data set.
In response to the bias between the SIM and ASCAT soil moisture, the ASCAT data have been rescaled to better match the SIM_NRT $w_1$ climatology prior to the assimilation, using the CDF-matching technique of Reichle and Koster (2004). The CDF-matching effectively removed the differences in the mean and standard deviation of the ASCAT and SIM soil moisture, and the resulting values for the rescaled ASCAT $w_1$ are the same as reported above for SIM_NRT.

### 3.2 Impact of NRT forcing errors on soil moisture

Before using SIM_DEL as a benchmark for evaluating the impact of the ASCAT assimilation, the soil moisture from the SIM_NRT and SIM_DEL simulations have been compared to identify the impact of the NRT forcing errors on the modeled soil moisture. Since there is a strong seasonal cycle in the impact of the assimilation in Sect. 3.3, all statistics describing the experiments from this point forward are limited to three complete years, from May 2007–April 2010 (although the time series plots will show the full 3.5 year assimilation period). Calculating the statistics using the full 3.5 year period does not change the qualitative conclusions drawn from the results.

Comparing SIM_NRT and SIM_DEL revealed that they have similar temporal variability. For $w_1$, $r_{abs}$ averaged across the domain for the three years from May 2007 was 0.95, and 98% of the model grid cells had a value above 0.90. Likewise the $r_{anm}$ was also consistently very high, giving a mean of 0.95, with 96% of the grid cells above 0.90. For $w_2$, the mean $r_{abs}$ was 0.95, with 88% of the grid cells above 0.90, while the mean $r_{anm}$ was 0.94, with 87% of the model grid cells above 0.90. The lower mean correlations for $w_2$ were caused by lower values in mountainous regions, where the variable terrain increases the spatial variability in the near-surface atmosphere, emphasising the impact of the enhanced observation density in the DEL SAFRAN analysis.

While their temporal behaviour was similar, there were substantial differences between the absolute soil moisture simulations from SIM_DEL and SIM_NRT, particularly for $w_2$. Figure 2 shows the mean difference in $w_1$ and $w_2$ between the two simulations,
demonstrating that SIM_NRT was consistently drier than SIM_DEL. The mean bias in the SIM_NRT $w_1$, relative to SIM_DEL, was $-0.013 \text{ m}^3 \text{ m}^{-3}$ ($-0.13 \text{ mm}$, compared to a seasonal range in the mean $w_1$ of approximately 2 mm in Sect. 3.3). Likewise, for $w_2$ the mean bias was $-0.008 \text{ m}^3 \text{ m}^{-3}$ (or $-12.1 \text{ mm}$, compared to a seasonal range in the mean $w_2$ of close to 100 mm in Sect. 3.3). This dry bias is associated with the errors in the near-real time precipitation forcing that were mentioned previously. Figure 2c shows that the NRT precipitation is biased dry across the SIM domain, by as much as $-200 \text{ mm yr}^{-1}$ in places. Additionally, there is a reasonably strong spatial correspondence between the precipitation and soil moisture biases in Fig. 2, including the same isolated regions of positive bias.

The above findings have several implications for assimilating the ASCAT data. It is extremely unlikely that data assimilation could improve the preexisting very high ($\sim 0.95$) correlations between SIM_NRT and SIM_DEL (this would require the observations to be extremely accurate in correcting for SAFRAN forcing errors, while not making significant corrections for other errors in the SIM soil moisture), and nor is it necessary to improve such high values. However, there is a substantial low bias in the SIM_NRT soil moisture relative to SIM_DEL, associated with the low-biased NRT precipitation forcing. The assessment of the impact of assimilating the ASCAT data will then be focused on determining whether this low bias can be corrected. Since the assimilated ASCAT soil moisture data were rescaled to the SIM_NRT $w_1$ climatology, the assimilated data will have the same low bias as the SIM_NRT $w_1$ (over the full data record). Hence, for the assimilation to correct the soil moisture in response to the biased precipitation forcing will require that the rescaled ASCAT data still contain a signal of the low-biased response to individual precipitation events.
3.3 Impact of the assimilation on soil moisture

3.3.1 Comparison to SIM_DEL

The assimilation had a strong tendency to add moisture to the surface, with a mean of 38.5 mm yr\(^{-1}\) added in the last three years. The assimilation added net moisture at nearly all model grid cells, with only a handful of isolated occurrences of net moisture removal. These locations do not correspond to the locations of positive precipitation and soil moisture biases in Fig. 2. Nor do the regions of strongest moisture addition correspond to the regions of strongest precipitation and soil moisture biases, although this could be due to the nonlinearity of the relationship between soil moisture increments and changes in soil moisture storage, or other complicating factors such as the frequency of the assimilated observations.

Figure 3 shows maps of the soil moisture bias between SIM_ASCAT and SIM_DEL. Comparing this to Fig. 2 demonstrates that the positive soil moisture increments generated by the ASCAT assimilation reduced the negative SIM_NRT soil moisture biases. For \(w_1\), there were very small reductions in the net bias at most grid cells (at 78\% of cells across the domain, and at 94\% of the cells which have ASCAT observations), with slightly larger reductions of approximately 0.02 mm in the north of France. Overall, the mean bias for the three years from May 2007 was slightly reduced to \(-0.11\) mm (from \(-0.13\) mm for SIM_NRT). The impact of the assimilation on the \(w_2\) biases was more substantial, and while the magnitude of the (negative) \(w_2\) bias was reduced across most of France, a small positive bias was introduced in the northeast and southwest (corresponding to very small negative biases in the SIM_ASCAT \(w_1\)). In the northeast the positive bias occurs in the Champagne-Picardie region, which has unusual soil properties, characterised by dark soils and a large soil moisture holding capacity, which are not well represented by soil parameters used in SIM. Overall, the assimilation reduced the mean bias for \(w_2\) to \(-5.6\) mm (from \(-12.1\) mm), while the absolute bias was reduced at 73\% of the grid cells (and at 89\% of the cells with ASCAT data).
To check that the assimilation also improved the absolute fit to the SIM_DEL time series, Fig. 4 shows maps of the reduction in the RMSE, relative to SIM_DEL, generated by assimilating the ASCAT soil moisture. For both soil layers the assimilation decreased the RMSE across most of France, although it was increased in the northeast and southwest, corresponding to the small positive biases in the SIM_ASCAT $w_2$. For $w_1$, the assimilation reduced the RMSE at 59% of the grid cells (and at 71% of cells with ASCAT data), although the reductions were very small, and the mean RMSE was unchanged from 0.28 mm for both SIM_NRT and SIM_ASCAT. Consistent with the larger bias correction for $w_2$, the impact on the $w_2$ RMSE was also greater, and the RMSE was reduced at 57% of the model grid cells (and at 69% of cells with ASCAT data), although there were relatively large increases in the RMSE in the northeast. Overall, the mean $w_2$ RMSE was reduced from 16.6 to 15.8 mm.

Figure 5 shows time series of the average volume of moisture added to the surface by SIM_ASCAT each day. Very little moisture was added or subtracted during the winter months, due to the widespread occurrence of frozen surface conditions, as well as the reduced vertical soil moisture coupling in ISBA during winter. During the nonwinter months there was a tendency towards the addition of surface moisture, with net moisture added more often (approximately 60% of the mean daily increments were positive), and with the positive increments tending to be larger than the negative increments (by a factor of 1.5, on average).

Figure 6 shows the temporal evolution on the impact of the assimilation on the model $w_2$. Prior to the assimilation, the SIM_NRT $w_2$ was persistently biased low, by −10 to −20 mm. The assimilation generally reduced the magnitude of the negative $w_2$ biases in Fig. 6, with the greatest reductions (of around 10 mm) occurring through the summer, and persisting into early winter, before being gradually lost in late winter. The assimilation also reduced the spatial RMSE between the simulated $w_2$ and SIM_DEL (by up to 5 mm) on most days.
3.3.2 Comparison to SMOSMANIA in situ observations

Since comparison to SIM_DEL cannot be used to test whether the assimilation has improved the temporal behaviour of the modeled soil moisture, the assimilation results have also been compared to in situ data from the SMOSMANIA network. The SMOSMANIA network (Calvet et al., 2007; Albergel et al., 2008) consists of 12 soil monitoring stations spanning from the Mediterranean to Atlantic coasts in southwest France, each spaced approximately 45 km apart. The deepest soil moisture sensors observe at 30 cm, much shallower than the root-zone soil moisture depths used in ISBA (approximately 1 m). Consequently, the time scales of soil moisture variability in the 30 cm SMOSMANIA data and the SIM root-zone soil moisture time series are qualitatively very different, preventing a meaningful comparison between them. Hence, only the near-surface (5 cm) SMOSMANIA observations have been used, and these have been compared to the (1 cm) near-surface soil moisture from SIM.

Table 2 presents the correlation statistics between each of the SIM_NRT, SIM_ASCAT, and SIM_DEL near-surface soil moisture, and the SMOSMANIA in situ observations. Two interesting features are revealed by these statistics. First, SIM_DEL consistently had higher correlations with the SMOSMANIA time series than SIM_NRT did, giving higher mean $r_{abs}$ ($r_{anm}$) for SIM_DEL of 0.70 (0.59), compared to 0.66 (0.53) for SIM_NRT. This supports the assumption in this work that the SIM_DEL soil moisture is more accurate than SIM_NRT, while also suggesting that the SMOSMANIA observations are sufficiently accurate to detect the difference in accuracy between SIM_NRT and SIM_DEL. Second, while assimilating the ASCAT data did improve the SIM_NRT correlations at most sites (only $r_{abs}$ at CDM is degraded), the improvements were marginal, and so generated only a slight improvement in the mean $r_{abs}$ ($r_{anm}$) to 0.67 (0.54) for SIM_ASCAT. While this is encouraging, particularly given the consistency of the higher correlations for SIM_ASCAT, these improvements are not statistically (or practically) significant.
3.4 Impact of the assimilation on the water balance

Figure 7 shows time series of the monthly mean surface water balance terms, from SIM_DEL, SIM_NRT, and SIM_ASCAT, while Fig. 8 shows time series of the monthly mean difference from SIM_DEL of each term, for each of SIM_NRT and SIM_ASCAT. Precipitation is imposed by the forcing, while all other variables are forecast by ISBA. The precipitation in SIM_NRT (and SIM_ASCAT) was persistently biased low, with a tendency for larger biases in winter, generating a large mean monthly bias of −16.8 mm month\(^{-1}\) for the three years from May 2007. The monthly mean analysis increments from SIM_ASCAT are included in both figures. As discussed previously the analysis increments tended to be positive, giving a mean monthly increment of +2.9 mm month\(^{-1}\), although the largest increments (of approximately 10 mm month\(^{-1}\)) were similar in size to the largest differences in the other water balance terms.

Close to half of the SIM_NRT precipitation bias was transferred into a bias in the drainage forecasts. The SIM_NRT drainage was persistently biased low, with the largest errors occurring in winter, coinciding with the drainage maxima. Over the three year period, the mean monthly drainage bias for SIM_NRT was −7.3 mm month\(^{-1}\). Since runoff and drainage are both triggered when soil moisture exceeds saturation, the seasonal cycle and pattern of biases in the runoff was similar. However, the volume of runoff, and hence the errors in the runoff, was much smaller, and the mean monthly runoff bias was just −2.0 mm month\(^{-1}\). For both discharge and runoff the SIM_NRT bias was close to 30% of the SIM_DEL forecasts, compared to the NRT precipitation bias of 20% of the SIM_DEL forcing. The wetter soil moisture in SIM_ASCAT reduced the magnitude of the negative biases in drainage and runoff, to −6 mm month\(^{-1}\) and −1.8 mm month\(^{-1}\), a reduction of 17% and 11% of the SIM_NRT original biases, respectively.

Evapotranspiration has the opposite seasonal cycle to drainage and runoff, with maxima in summer. In each year the SIM_NRT evapotranspiration was biased low in late-summer, when surface drying causes transpiration to become moisture limited. This
negative evapotranspiration bias during summer (larger than \(-5 \text{ mm month}^{-1}\)) was offset by a small positive bias during the wet months (of approximately 1 mm month\(^{-1}\)), generating a mean bias of just \(-1.2 \text{ mm month}^{-1}\). The evapotranspiration bias is relatively small, being just 7\% of the SIM\_DEL forecast precipitation over the three years, even though the evapotranspiration is one of the largest terms in the water balance. Assimilating the ASCAT data effectively decreased the evapotranspiration bias during the latter half of summer, while not affecting the positive biases during winter (which are not easily attributed to soil moisture errors). This reduced the three year bias to \(+0.4 \text{ mm month}^{-1}\), or 70\% of the magnitude of the original bias, with a reversal of sign.

The seasonal behaviour of the monthly change in surface moisture storage is less consistent than the other terms. Despite the consistent negative precipitation bias, the monthly mean error in the change in storage is often positive, due to the underestimation of the aforementioned fluxes. Periods of positive and negative errors offset each other, giving a small three year mean of just 0.1 mm month\(^{-1}\). The impact of the assimilation on the monthly change in moisture storage is also less consistent than for the other water balance terms. There were instances of the assimilation both increasing and decreasing the monthly errors, resulting in no net change in the monthly mean bias.

### 3.5 Impact of the assimilation on river flow

Finally, the drainage and runoff forecasts from SIM\_NRT and SIM\_ASCAT have been routed through the surface river network with the MODCOU model, for comparison against SIM\_DEL streamflow forecasts. An example hydrograph is shown in Fig. 9. Consistent with the previous results for soil moisture, SIM\_NRT simulated the timing of flood events from SIM\_DEL very well, while consistently underestimating the magnitude of the peak flows. Also, assimilating the ASCAT data reduced this underestimation of the peak flows. In this example the discharge ratio, relative to SIM\_DEL, was increased.
from 0.76 to 0.78 by the assimilation, while the Nash-Sutcliffe Efficiency ($E$), increased from 0.74 to 0.77.

Similar results, indicating a generally positive impact of the assimilation on streamflow, were obtained across France. For the discharge ratio, Fig. 10 shows that SIM_ASCAT reduced the absolute error in the discharge ratio (i.e., difference from unity) at most (88%) of the gauging stations, increasing the mean discharge ratio for the 907 stations modeled by MODCOU from 0.68 for SIM_NRT to 0.76 for SIM_ASCAT. Additionally, the Nash-Sutcliffe Efficiency, which is more sensitive to the accuracy of the (timing and magnitude) of the peak flows, was also improved by the assimilation. Figure 11 shows that assimilating the ASCAT data increased $E$ at most (82%) stations, increasing the mean from 0.62 for SIM_NRT to 0.68 for SIM_DEL.

### 4 Summary and conclusions

The experiments presented here have demonstrated that the root-zone soil moisture in land surface models can be improved by the assimilation of ASCAT-derived near-surface soil moisture. Specifically, assimilating the ASCAT data into the ISBA model forced with the near-real time (NRT) SAFRAN low-level atmospheric analysis improved the model hydrology, relative to the ISBA model forced with the higher quality delayed cut-off (DEL) SAFRAN analysis. The temporal agreement between the SIM_NRT and SIM_DEL soil moisture simulations was very high (the mean correlation and anomaly correlation were 0.95 and 0.94, respectively), making it extremely difficult and also unnecessary for the assimilation to correct the small errors in the temporal behaviour between SIM_NRT and SIM_DEL. However, there was a substantial dry bias in SIM_NRT soil moisture relative to SIM_DEL, associated with a dry bias in the NRT SAFRAN precipitation. Assimilating the ASCAT near-surface soil moisture data effectively reduced the dry bias in SIM_NRT, and the root-zone soil moisture bias (RMSE) relative to SIM_DEL was reduced from $-12.1$ mm ($16.6$ mm) for the SIM_NRT open-loop, to $-5.6$ mm ($15.8$ mm) by the assimilation of ASCAT data (SIM_ASCAT).
Additionally, assimilating the near-surface soil moisture observations also had a strong effect on the model hydrology throughout the seasonal cycle, including the subsequent moisture flux and streamflow forecasts. For the root-zone soil moisture, the assimilation had the greatest impact during the nonwinter months (reducing the magnitude of the negative bias by as much as 10 mm), when the near-surface soil moisture observations were not masked by a frozen surface. This in turn reduced the magnitude (by 5–10 mm month$^{-1}$) of the (already small) negative evapotranspiration biases that developed towards the end of summer and in autumn, when transpiration is moisture limited. Then, during the wetter months, the increased root-zone soil moisture also resulted in reductions in the negative drainage biases (by up to 5 mm month$^{-1}$), as well as the much smaller biases for runoff (also triggered when soil moisture exceeds saturation). Since the analysis increments were minimal during the winter months, the reductions in the root-zone soil moisture bias (i.e., increases in $w_2$) accumulated during the previous year were largely lost by the end of each winter, due to the reduced drainage bias (i.e., enhanced drainage forecasts). Finally, the improved discharge and runoff forecasts resulted in improved streamflow forecasts from the 907 gauging stations modeled by MODCOU, relative to SIM_DEL, in terms of both the net flow (the mean discharge ratio was increased from 0.68 to 0.76), and the fit to hydrograph (the mean Nash-Sutcliffe Efficiency was increased from 0.62 to 0.68).

The results of this study are extremely encouraging in terms of the quality of the ASCAT soil moisture data set. First, the consistently high correlation (mean 0.68) and anomaly correlation (mean 0.62) between SIM and ASCAT near-surface soil moisture across the France domain indicates that both are detecting temporal changes in soil moisture very well. Second, the improved root-zone soil moisture relative to the SIM_DEL generated by the assimilation suggests that the ASCAT observations are sufficiently accurate to detect soil moisture errors introduced by the NRT forcing.

It was not expected that a bias-blind assimilation of near-surface soil moisture could correct SIM_NRT $w_2$ biases. The assimilated $w_1$ observations were by design (CDF-matching) unbiased relative to SIM_NRT, and consequently the assimilation had little
influence on the $w_1$ biases, and yet it reduced the $w_2$ biases. The mechanism by which this occurred is unclear, however it is thought to be related to the longer memory of $w_2$ (compared to $w_1$), combined with nonlinear aspects of the model-response to the analysis updates. This issue will be investigated further in the future. Despite this uncertainty, the tendency for the assimilation to improve the $w_2$ RMSE (by reducing the timeseries of the biases), as well as the small improvements in the correlations with the in situ near-surface soil moisture observations from SMOSMANIA, support the interpretation that the results obtained here reflect positively on the accuracy of the ASCAT data.

Finally, this work highlights the potential of using model output and assimilation studies as tools to evaluate remotely sensed near-surface soil moisture. This is an extremely important issue, given the difficulty of directly verifying soil moisture at the scales and spatial coverage of remotely sensed data sets. Additionally, in this study the drainage and hence streamflow forecasts contained a strong signal of the biased root-zone soil moisture. Since streamflow integrates soil moisture conditions over space and time, and it is regularly observed at a large number of locations globally, streamflow forecast skill presents an ideal metric for assessing long term (seasonal scale or longer) soil moisture errors in a model. Unfortunately the SIM model does not account for human management of river flows, making a meaningful evaluation of SIM streamflow forecasts against observed streamflow difficult.

Acknowledgements. The EUMETSAT H-SAF Associated Scientist Program is acknowledged for funding the work presented here.

The publication of this article is financed by CNRS-INSU.
References


5448

Canellas, C.: Intercomparison de pluviomètres, Metéo-France Note de la Direction de la Production, 2005. 5430


**Table 1.** Details of each SIM simulation.

<table>
<thead>
<tr>
<th>SAFRAN forcing</th>
<th>Assimilated data</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIM_DEL</td>
<td>DEL</td>
</tr>
<tr>
<td>SIM_NRT</td>
<td>NRT</td>
</tr>
<tr>
<td>SIM_ASCAT</td>
<td>NRT</td>
</tr>
</tbody>
</table>
Table 2. Correlation statistics between the in situ observations from SMOSMANIA, and each \( w_1 \) from each of SIM_NRT, SIM_ASCAT, and SIM_DEL, from May 2007 to April 2010. All correlations are calculated using only days on which all data sets are available, and all are significant at 1%.

<table>
<thead>
<tr>
<th></th>
<th>SIM_NRT</th>
<th>SIM_ASCAT</th>
<th>SIM_DEL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( r_{ab} )</td>
<td>( r_{anm} )</td>
<td>( r_{ab} )</td>
</tr>
<tr>
<td>SBR</td>
<td>0.77</td>
<td>0.65</td>
<td>0.78</td>
</tr>
<tr>
<td>URG</td>
<td>0.64</td>
<td>0.66</td>
<td>0.67</td>
</tr>
<tr>
<td>CRD</td>
<td>0.70</td>
<td>0.56</td>
<td>0.73</td>
</tr>
<tr>
<td>PRG</td>
<td>0.68</td>
<td>0.46</td>
<td>0.70</td>
</tr>
<tr>
<td>CDM</td>
<td>0.76</td>
<td>0.55</td>
<td>0.72</td>
</tr>
<tr>
<td>LHS</td>
<td>0.65</td>
<td>0.45</td>
<td>0.65</td>
</tr>
<tr>
<td>SVN</td>
<td>0.63</td>
<td>0.53</td>
<td>0.64</td>
</tr>
<tr>
<td>MNT</td>
<td>0.55</td>
<td>0.52</td>
<td>0.56</td>
</tr>
<tr>
<td>SFL</td>
<td>0.67</td>
<td>0.45</td>
<td>0.67</td>
</tr>
<tr>
<td>MTM</td>
<td>0.50</td>
<td>0.41</td>
<td>0.55</td>
</tr>
<tr>
<td>LZC</td>
<td>0.71</td>
<td>0.62</td>
<td>0.72</td>
</tr>
<tr>
<td>NBN</td>
<td>0.67</td>
<td>0.49</td>
<td>0.67</td>
</tr>
</tbody>
</table>
Fig. 1. Maps of (a) absolute correlation and (b) anomaly correlation between SIM $w_1$ and ASCAT SDS from January 2007 to May 2010.
Fig. 2. Net bias from May 2007 to April 2010 between SIM_NRT and SIM_DEL for (a) $w_1$ (mm), (b) $w_2$ (mm), and (c) precipitation forcing (mm yr$^{-1}$).
Fig. 3. Net bias from May 2007 to April 2010 (in mm) between SIM_ASCAT and SIM_DEL for (a) $w_1$, and (b) $w_2$. 
Fig. 4. Improvement in the RMSE (mm) relative to SIM_DEL from the assimilation of ASCAT, for (a) $w_1$ and (b) $w_2$, from May 2007 to April 2010.
Fig. 5. Time series of the spatial mean volume of moisture (mm day$^{-1}$) added to the surface ($w_1 + w_2$) through assimilation of the ASCAT SDS.
Fig. 6. Time series of the spatial mean (a) $w_2$, and (b) $w_2$ bias (dotted lines) and $w_2$ RMSE (solid lines) relative to SIM_DEL. SIM_DEL is plotted in black, SIM_NRT in red, and SIM_ASCAT in blue, and both plots are in mm.
Fig. 7. Monthly water balance in mm month$^{-1}$ for SIM_DEL (black), SIM_NRT (red), and SIM_ASCAT (blue). Each panel shows (a) precipitation, (b) runoff, (c) drainage, (d) evapotranspiration, (e) change in surface moisture storage (all soil layers, liquid plus solid), and (f) the analysis increments (for SIM_ASCAT).
Fig. 8. Mean error relative to SIM_DEL in mm month$^{-1}$ of the monthly water balance terms for SIM_NRT (red) and SIM_ASCAT (blue). Each panel shows (a) precipitation, (b) runoff, (c) drainage, (d) evapotranspiration, (e) change in surface moisture storage (all soil layers, liquid plus solid), and (f) the analysis increments (for SIM_ASCAT).
**Fig. 9.** Discharge (m$^3$ day$^{-1}$) from SIM_DEL (black), SIM_NRT (red), SIM_ASCAT (blue), for the River Seine at Poses.
Fig. 10. Scatterplot of the discharge ratio, assuming SIM_DEL as the truth, for SIM_ASCAT vs. SIM_NRT, for the 907 gauging stations simulated by MODCOU.
Fig. 11. Scatterplot of the Nash-Sutcliffe efficiency, assuming SIM_DEL as the truth, for SIM_ASCAT vs. SIM_NRT, for the 907 gauging stations simulated by MODCOU.