Journal of Hydrology 578 (2019) 124018

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Improvements in the forecasts of near-surface variables in the Global Forecast System (GFS) via assimilating ASCAT soil moisture retrievals

Jifu Yin[a,b,*], Christopher R. Hain[c], Xiwu Zhan[b], Jiarui Dong[d], Michael Ek[e]

[a] ESSIC/CICS, University of Maryland College Park, College Park, MD 20740, USA
[b] NOAA NESDIS Center for Satellite Applications and Research, College Park, MD 20740, USA
[c] ASA Marshall Space Flight Center, Earth Science Branch, Huntsville, AL 35801, USA
[d] NOAA/NCEP Environmental Modeling Center, College Park, MD 20740, USA
[e] NCAR Research Applications Laboratory, 3090 Center Green Drive, Boulder, CO 80301, USA

ARTICLE INFO

This manuscript was handled by Marco Borga, Editor-in-Chief

Keywords:
Data assimilation
Soil moisture
Numerical weather prediction
ASCAT
Global forecast system

ABSTRACT

Recent research has shown that assimilating satellite soil moisture (SM) retrievals into the land surface models (LSMs) improves simulations of land-atmosphere water and energy exchanges. With satellite SM retrievals becoming widely and continuously available, it is desirable to examine the impact of assimilating them into numerical weather prediction models in order to improve numerical weather forecast skills. Based on the development of the coupled system of National Centers for Environmental Prediction (NCEP)-Global Forecast System (GFS) and National Aeronautics and Space Administration (NASA)-Land Information System (LIS) in this paper, we designed an experiment to demonstrate the impacts of assimilating the Advanced Scatterometer (ASCAT) SM data products on the weather forecasts of GFS. With respect to the surface air temperature analysis product of National Oceanic and Atmospheric Administration (NOAA)-Climate Prediction Center (CPC) and CPC's morphing method-based precipitation data, improvement from the ASCAT SM assimilation for probabilities of high quality forecasts can reach up to 1.7% for GFS precipitation, 3.1% for 2-meter minimum temperature, and 3.1% for 2-meter diurnal temperature range predictions, respectively. These results suggest that satellite SM data assimilation could be beneficial for GFS numerical weather forecasts of NOAA NCEP.

1. Introduction

Soil moisture (SM) is an important variable for regional and global numerical weather prediction models that impacts the energy and water exchanges between land surface and the atmosphere (Koster et al., 2004; Seneviratne et al., 2010; Zhan et al., 2012; Yin et al., 2015c, 2016). Currently, a number of microwave satellite SM products are operationally generated with certain accuracy (Naeimi et al., 2009; Kerr et al., 2010; Jones et al., 2010; Entekhabi et al., 2010; Li et al., 2010; Parinussa et al., 2014; Yin et al., 2014; Yin et al., 2015a,b). It is thus desirable to examine the impacts of assimilating satellite SM retrievals into weather prediction models and in turn improve numerical weather forecast skills.

Several global weather forecast centers are currently assimilating or testing the assimilation of remotely-sensed SM observations. Ongoing research at the European Centre for Medium-Range Weather Forecasts has shown positive impacts from the assimilation of remote-sensing SM data (Drusch and Viterbo, 2007; Scipal et al., 2008; de Rosnay et al., 2013). Building on the Extended Kalman Filter and the offline Joint UK Land Environment Simulator land surface model (LSM), a new land data assimilation system was developed by the UK Met Office aiming to correctly propagate surface information to deeper soil layers using remote sensing observations (Dharssi et al., 2010). However, research related to data assimilation of satellite SM at the National Centers for Environmental Prediction (NCEP) of National Oceanic and Atmospheric Administration (NOAA) in the US has only recently begun.

The Global Forecast System (GFS) of NCEP is an important numerical weather prediction model that provides medium-range weather forecasts. For each GFS run, a set of initial values of system state variables including SM is required. For the SM state variable, the current GFS version uses estimates that are provided by previous GFS forecasts. Because of uncertainties associated with precipitation estimates and other meteorological forcing data for the Noah LSM, initial values used for the GFS runs may not represent the true SM information, which may contribute to downstream errors in GFS forecasts. It is thus desirable to maximize the impacts of satellite SM products in an

* Corresponding author at: Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, 5825 University Research Court suite 4001, College Park, MD 20740, USA.
E-mail address: jyin@umd.edu (J. Yin).

https://doi.org/10.1016/j.jhydrol.2019.124018
Received 27 April 2018; Received in revised form 1 August 2019; Accepted 3 August 2019
Available online 06 August 2019

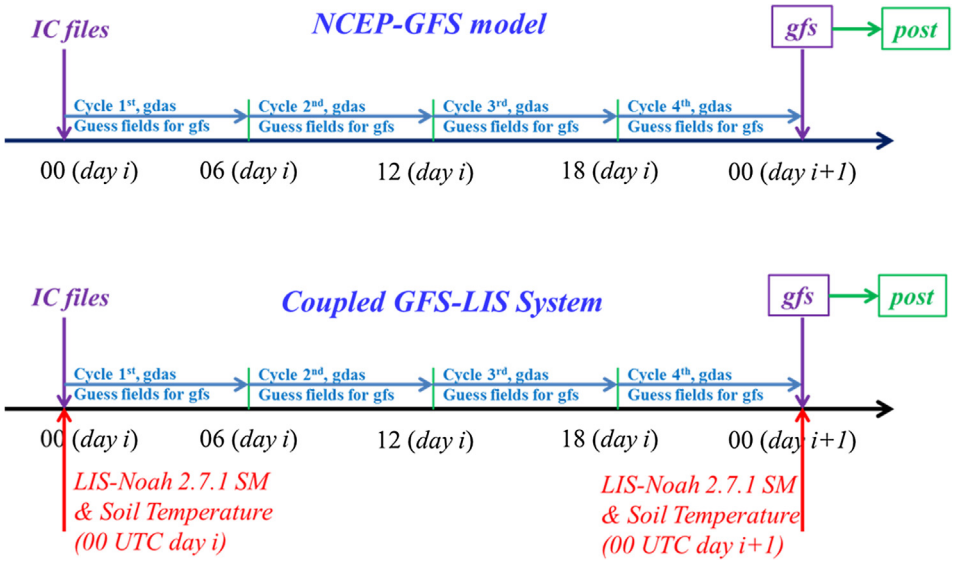

Fig. 1. Steps of the GFS and the coupled GFS-LIS system. The acronym indicates initial conditions.

Table 1
Perturbation parameters for meteorological forcing inputs and for state variables. The acronyms P, SW, LW and SM indicate precipitation, downward shortwave, long-wave radiation and soil moisture, respectively.

Perturbation type	Std dev	Cross correlation for forcing variable perturbations			
		P	SW	LW	
P	0.5 (mm)	1.0	−0.8	0.5	
SW	0.3 (W m^{-2})	−0.8	1.0	−0.5	
LW	50 (W m^{-2})	0.5	−0.5	1.0	
Perturbation type	Std dev	Cross correlation for forcing variable perturbations			
		SM1	SM1	SM1	SM4
SM1 (0–10 cm)	6.0×10^{-3} m^3 m^{-3}	1.0	0.6	0.4	0.2
SM2 (10–30 cm)	1.1×10^{-4} m^3 m^{-3}	0.6	1.0	0.6	0.4
SM3 (30–60 cm)	6.0×10^{-5} m^3 m^{-3}	0.4	0.6	1.0	0.6
SM4 (60–100 cm)	4.0×10^{-5} m^3 m^{-3}	0.2	0.4	0.6	1.0

advanced data assimilation system. Meanwhile, the NASA-LIS (Land Information System) offers a framework to develop a generic sequential data assimilation subsystem (Kumar et al., 2006, 2008). In particular, the LIS includes effective tools such as the ensemble Kalman Filter (EnKF) that is widely used for sequential assimilation of hydrologic variables (Evensen, 1994; Burgers et al., 1998).

Several remote sensing SM products from passive and active microwave satellite sensors have been developed to address issues in the in situ SM observations resulting from sparse spatial distributions. Recent studies indicate that SM retrievals from Advanced Scatterometer (ASCAT) on board the Meteorological Operation (MetOp) satellite program (Bartalis et al., 2007) are comparable to passive SM retrievals (Liu et al., 2011; Zeng et al., 2015). The ASCAT provides continuous and reliable long-term SM service with retrievals from MetOp-A launched in October 2006, MetOp-B launched in September 2012 (Wagner et al., 2013) and MetOp-C launched in November 2018. However, there are typically large uncertainties in SM retrievals from densely vegetated areas due to the robust impacts of vegetation structure and water content. In particular, backscatter for active C-band sensors (such as ASCAT) is impacted by diurnal variations in vegetation water content more strongly than previously thought (Van Emmerik et al., 2015).

Therefore, we developed a coupled GFS-LIS system in this paper and then evaluated the impacts of ASCAT SM assimilation on the numerical weather forecasts from GFS over sparsely vegetated areas. The descriptions of the GFS, LIS-Noah 2.7.1, and the structure of coupled GFS-LIS system are briefly introduced in the next section. The EnKF-based numerical experiment focused on assimilating the ASCAT SM product into the developed GFS-LIS coupled system is documented in Section 3. Descriptions of input data sources and methodology are provided in Section 4. Based on comparisons with 2-meter minimum temperature (Tmin), 2-meter maximum temperature (Tmax), and 2-meter diurnal temperature range (DTR) from the NOAA-CPC (Climate Prediction Center) surface air temperature analysis product and precipitation retrieved from the CPC's morphing method (CMORPH), the coupled verification of GFS forecast is shown in Section 5. Discussion and a brief summary are then provided in Sections 6 and 7, respectively.

2. Model descriptions

2.1. NCEP Global Forecast System (GFS)

The GFS operationally running at NOAA-NCEP is a three-dimensional hydrostatic global spectral model. It uses the Global Data Assimilation System (GDAS) to provide guess fields for the full forecasts. The GDAS runs for each cycle (00, 06, 12, and 18 UTC); however, to save time and space in experiments, the GFS is initially set up to run only for the 00 UTC cycle. In this paper, the GFS model version T670 is run at T254 spatial resolution that is about 35 km in Gaussian projection (grid dimensions: 1344 at longitude by 672 at latitude). The GFS produced 3-hour forecasts 168 h in advance during the May 16–30, 2014 time period. The vertical was divided into 47 layers with the first 37 layers ranging from 1000 mb to 100 mb with 25 mb intervals and the following 10 layers set as 70 mb, 50 mb, 30 mb, 20 mb, 10 mb, 7 mb, 5 mb, 3 mb, 2 mb, and 1 mb.

2.2. LIS-Noah land surface model

LIS integrates the use of ground and satellite observations with well-documented LSMs and advanced tools to accurately characterize land surface states and fluxes (Kumar et al., 2006, 2008). To match the LSM used in GFS T670-254, the LIS-Noah 2.7.1 model was used in this paper.

The Noah LSM is a one-dimensional soil–vegetation–atmosphere transfer model that has four soil layers with thicknesses of 10 cm, 30 cm, 60 cm and 100 cm from the surface (Ek et al., 2003). The first three layers are for estimations in non-forested regions while the last soil layer is for simulations in forested areas (Ek et al., 2003). Based on

Table 2
Rainfall forecast skill scores according to 3-hourly rainfall intensity-based classification.

CPC 25 km precipitation products		Predicted Rainfall				
		Non-precipitation 0 mm	Little rain 0–3 mm	Moderate Rain 3.1–12 mm	Heavy Rain 12.1–48 mm	Extreme Rain > 48 mm
Non-precipitation	0 mm	5	3	2	1	0
Little rain	0–3 mm	3	5	3	2	1
Moderate Rain	3.1–12 mm	2	3	5	3	2
Heavy Rain	12.1–48 mm	1	3	3	5	3
Extreme Rain	> 48 mm	0	1	2	3	5

layer-based heat diffusion, standard diffusion, and gravity drainage equations, the Noah LSM is able to model dynamics of SM and soil temperature (Ek et al., 2003; Reichle et al., 2010). In this paper, the LIS-Noah 2.7.1 model was spun up by cycling 3 times through the period from 1 January 2008 to 30 April 2014. Simulations for each case were conducted during the May 1–31, 2014 period with half-hour time-step inputs and daily outputs at 25 km resolution. Specifically, the soil temperature and SM from LIS-Noah 2.7.1 outputs for the May 1–31, 2014 period were placed into the GFS initial condition files.

2.3. Coupled GFS-LIS system

The LIS-Noah 2.7.1 outputs with and without benefits from SM assimilation were remapped from the lat/lon projection at 25 km spatial resolution to a Gaussian projection with respect to the spatial resolution (~35 km) of GFS model. LIS was set up to run in parallel with each GFS simulation, providing updated initial land surface conditions for each 00Z GFS run. There the soil temperature and SM fields from LIS-Noah were placed into the GFS initial condition files (Fig. 1). The resulting analysis and forecast fields were finally converted to 3-hour outputs at 25 km spatial resolution in lat/lon projection by the "post" processor in the GFS model (Fig. 1).

3. Methodology of the experiment

3.1. EnKF-based SM data assimilation

The EnKF has been widely used in sequential SM data assimilation (Evensen, 1994; Burgers et al., 1998). Based on sequential Bayesian filtering and Monte Carlo approximation, ensemble forecasts and state variables are alternately updated in EnKF (Reichle et al., 2002). Forward propagation of ensemble model states is conducted using the LSM when EnKF runs in the ensemble forecast step. In the EnKF, error variances for forecast and measurement estimates vary in time and depend on the dynamics of the previous updates. The Kalman gain relates to forecast error covariance. As in the LIS examples (Kumar et al., 2009), a 3% constant value was used to define error variances.

3.2. Experiment design

Based on the GFS-LIS coupled system, we designed a set of numerical experiments to investigate the impacts of satellite SM assimilation on numerical weather prediction. The ASCAT SM is expected to perform reasonably over the sparsely vegetated areas because dense forests and shrubs are generally opaque to C-band radar (Wagner et al., 2013). Specifically, lower backscatter variations (< 2 dB) may result in great SM retrieval uncertainties over tropical forests and densely vegetated regions (Wagner et al., 2013), where average green vegetation fraction (GVF) climatologies are generally larger than 0.5 (Jiang et al., 2010). LIS provides monthly Advanced Very High Resolution Radiometer (AVHRR)-Normalized Difference Vegetation Index (NDVI)-based GVF maps. This study is focused on the areas where the GVF values are less than 0.5 (Jiang et al., 2010) over the global domain during the May 16 to 30, 2014 period.

The basic structure of the experiments is as follows: 1) The first simulation, the original GFS (ORG) run, is used to emphasize GFS T670-254 original performance without benefits of data assimilation and modifications on land surface parameters (land cover, GVF, soil texture, and slope). 2) The second simulation, the open-loop (OLP) run, uses a single realization with replacement of land surface parameters including land cover, GVF, soil texture, and slope maps consistent with those used in LIS-Noah 2.7.1 LSM. This simulation is chosen to assess the influences of different land surface parameters on GFS forecasts, and in turn to highlight the impact of satellite SM data assimilation. 3) The third simulation, the data assimilation (DA) run, is focused on assimilating the ASCAT SM product into the GFS-LIS coupled system using the EnKF. Because the land surface parameters in GFS are replaced as stated in the OLP case, the differences of the model results between DA and OLP cases are based only on the impact of the ASCAT SM data assimilation in the DA case.

The EnKF is a Monte Carlo method that requires an ensemble, and thus perturbations are added to make ensemble members differ from each other (Kumar et al., 2009). The input forcing data (precipitation, downward long- and short-wave radiation) of the Noah model are perturbed as described in Table 1 (Kumar et al., 2009). The implicit assumption is that a systematic bias in model output should not be caused by adding zero-mean Gaussian noise in the EnKF analysis step (Ryu et al., 2008). The variance values in the NASA LIS were manually adjusted to make the normalized innovations satisfy the requirement of normal distribution for optimal simulations. Ensemble size of EnKF was set as, which is the optimal ensemble size in a sequential SM assimilation system (Yin et al., 2015a). The benefits of assimilating the ASCAT SM product are expected to be highlighted by the suboptimal forcing and initialization used in the DA case.

4. Data and methodology

4.1. ASCAT SM product

ASCAT is onboard the polar orbiting MetOp-A, MetOp-B, and MetOp-C satellites. It is a real aperture radar using vertically polarized antennas. The nominal spatial resolution of the C-band (5.255 GHz) ASCAT is 50 km, and its experimental resolution is 25 km. The morning overpass time of the ASCAT occurs at 09:30 local time, and the evening overpass occurs at 21:30 local time. The ASCAT SM product used in this paper is developed using a time series-based change detection method at the Vienna University of Technology (Wagner et al., 1999; Naeimi et al., 2009). It yields the surface degree of saturation from 0 to 100%, representing the driest and wettest observations, respectively (Bartalis et al., 2007).

In this paper we employed the daily 25 km ASCAT SM data sets during 2008–2014 period over sparse vegetation areas where GVF values are lower than 0.5 and scaled them to the same SM climatology of the Noah 2.7.1 LSM using linear transformation that is represented as (Koster et al., 2009):

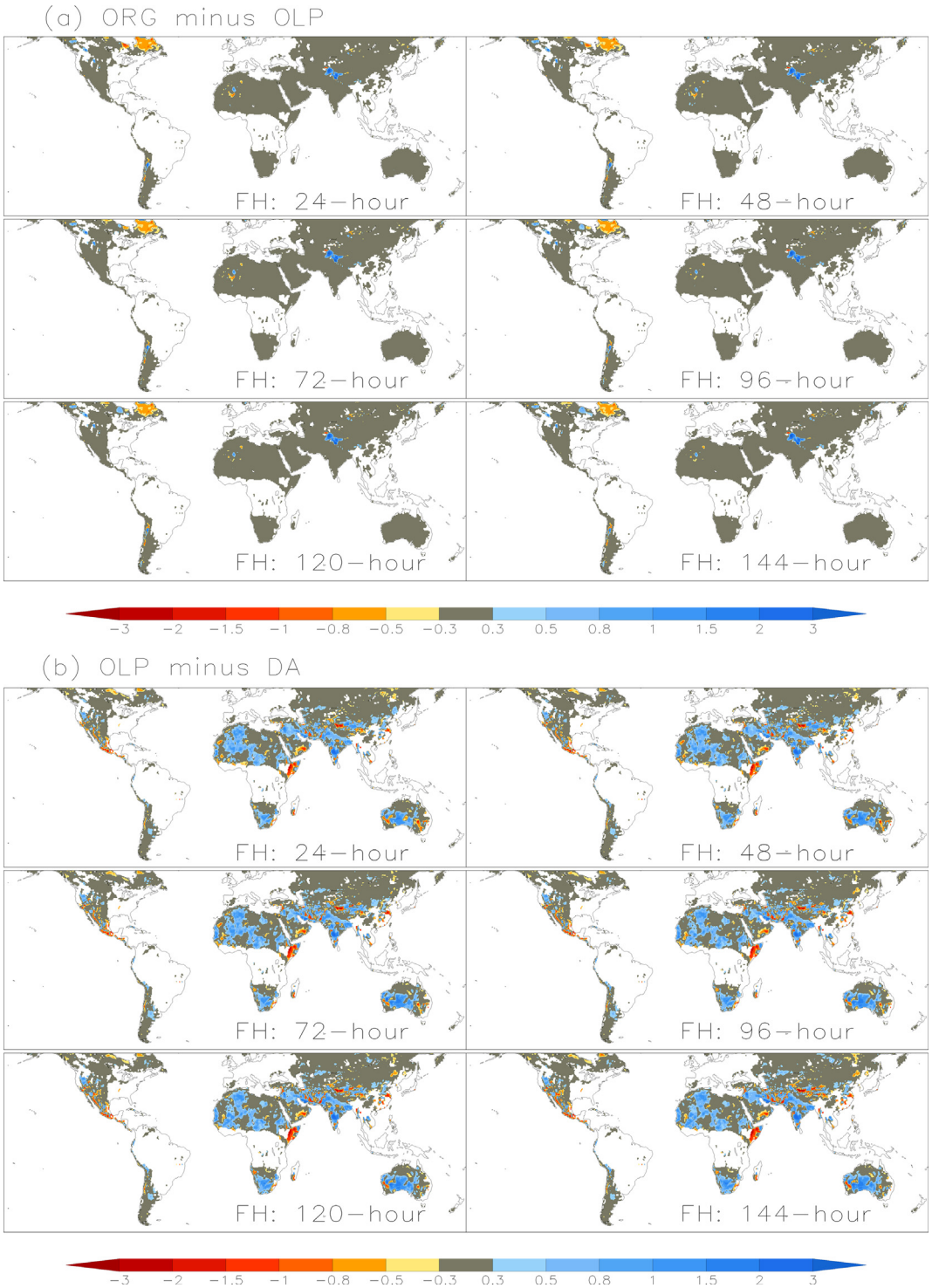

Fig. 2. With respect to the 50 km CPC-observations-based products, RMSE differences (in K) in predicted 2-meter minimum temperature (T_{min}) for 24-, 48-, 72-, 96-, 120-, 144-forecasthour (FH) over the May 16–30, 2014 period: (a) ORG minus OLP; and (b) OLP minus DA. The blue (red) color indicates improvement (degradation), while grey color means insignificant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$\Phi_{ASCAT/Noah} = \mu_{Noah} + (\Phi_{ASCAT} - \mu_{ASCAT})\frac{\sigma_{Noah}}{\sigma_{AsCAT}} \quad (1)$$

where $\Phi_{ASCAT/Noah}$ is the re-scaled ASCAT SM retrieval and Φ_{ASCAT} is the raw ASCAT SM retrieval. μ_{Noah} (σ_{Noah}) and μ_{ASCAT} (σ_{ASCAT}) are the averaged SM (the standard deviation) for the LIS-Noah 2.7.1 model and the ASCAT SM estimations over the 2008–2014 period, respectively.

4.2. NOAA-CPC surface air temperature analysis product

In support of various purposes, including climate monitoring, forecast verification, and studies on climate change and on the development of regional-to-global scale Earth system models, NOAA-CPC has developed an operational global daily land-only 2-meter temperature analysis at 50 km spatial resolution (ftp://ftp.cpc.ncep.noaa.gov/

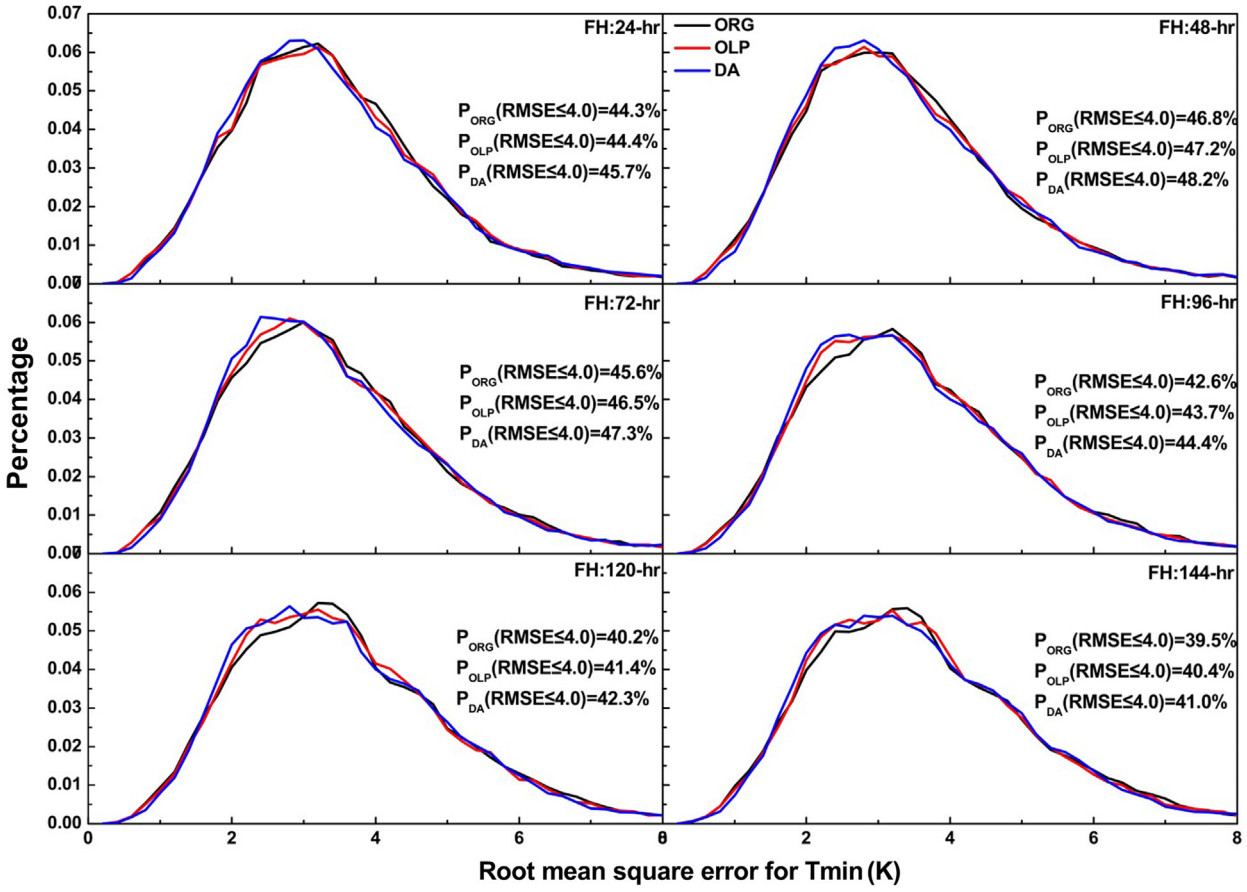

Fig. 3. With respect to the NOAA-CPC surface air temperature analysis product, the study domain-averaged frequency probability as a function of RMSE for the 24-, 48-, 72-, 96-, 120- and 144-FH predicted 2-meter minimum temperature (T_{min}) over the May 16–30, 2014 period.

precip/wd52ws/global_temp). The analysis is based on ~6000 Global Telecommunication System stations around the global and covers the 1979-present period. In this paper, the daily T_{min}, T_{max}, and DTR datasets over the May 15-June 5, 2016 period were used to assess the GFS forecasts. The GFS-LIS coupled system-based T_{min} and T_{max} estimations were resampled from 3-hour 25 km to daily 50 km to match the temporal and spatial resolution of the NOAA-CPC 2-meter temperature analysis.

4.3. CMORPH precipitation product

The Climate Prediction Center morphing method (CMORPH) incorporates half-hourly interval, geostationary satellite-based high quality precipitation estimates derived from passive microwave data. The current CMORPH product uses motion vectors derived from the DMSP 13, 14 & 15 (SSM/I), the NOAA-15, 16, 17 & 18 (AMSU-B), and AMSR-E and TMI aboard NASA's Aqua and TRMM spacecraft (Joyce et al., 2004). To generate spatially and temporally complete microwave-derived precipitation analyses, the precipitation intensity and shape are modified between microwave sensor scans using a time-weighted linear interpolation (Joyce et al., 2004). The 3-hourly 25 km CMORPH datasets were retrieved from May 16 to June 5, 2016 to assess GFS precipitation predictions in this paper.

4.4. Performance measures

4.4.1. Rainfall forecast skill scores

Many skill scores have been used to evaluate the quality of rainfall forecast, but a directive is an objective method of converting the predicted probability distribution into a discrete number or category (Gringorten, 1951; Vislocky and Young, 1988). In this study, rainfall events were classified as non-precipitation, little rain, moderate rain, heavy rain or extreme rain on the basis of the thresholds of 3-hourly intensity in Table 2. And the rainfall forecast skill scores (*S*) were then ranged from 0 to 5 with indicating from the worst to the best prediction performance (Vislocky and Young, 1988) with respect to the CMORPH precipitation product. Thus, estimations on the forecasting accuracy (*FA*) of the predicted non-precipitation and precipitation for *i, j* grid location are then expressed as

$$\mathrm{FA}(i,j) = \frac{\mathrm{S}(i,j)}{5.0} \times 100\% \quad (2)$$

4.4.2. Root mean square error

Root mean square error (RMSE) is a widely applied measure of the differences between the model and observed variables. Temperature and precipitation forecasts are very important to weather warnings and human activity. The predicted precipitation, T_{min}, T_{max}, and DTR are thus verified using the observations-based products in this work. Specifically, the 3-hourly model 25 km 2-meter temperature simulations at Coordinated Universal (UTC) time are resampled as daily 50 km at local time. Then the RMSE is used to evaluate the predicted precipitation with respect to the CMORPH product and verify the GFS T_{min}, T_{max} and DTR forecasts with respect to the NOAA-CPC 2-meter temperature analysis.

5. GFS forecast verification

With respect to the 50 km NOAA-CPC surface air temperature analysis product, Fig. 2 shows RMSE differences (in Unit: K) in predicted

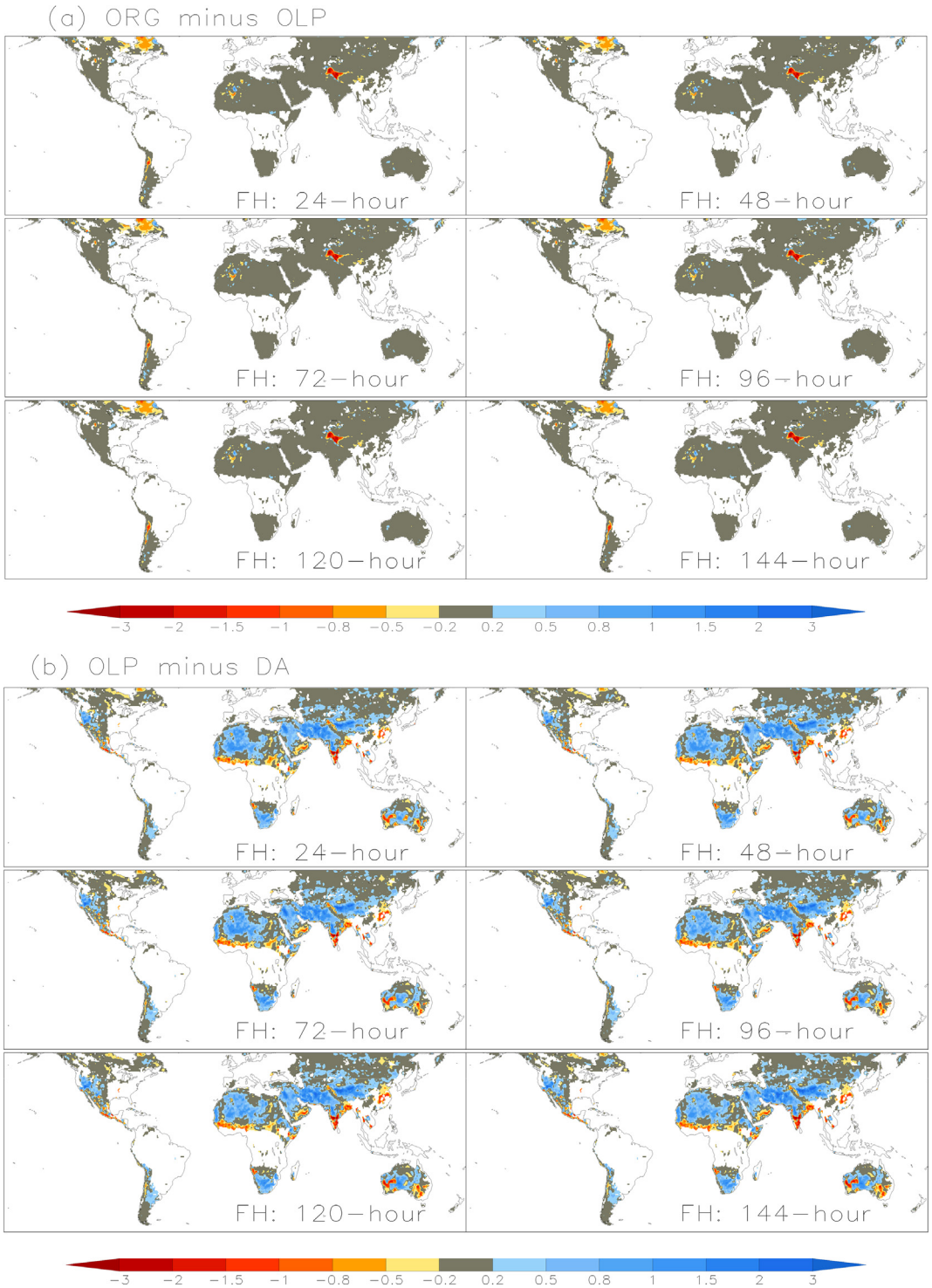

Fig. 4. With respect to the 50 km NOAA-CPC surface air temperature analysis product, RMSE differences (in Unit: K) in predicted 2-meter diurnal temperature range (DTR) for 24-, 48-, 72-, 96-, 120-, 144-FH over the May 16–30, 2014 period: (a) ORG minus OLP; and (b) OLP minus DA. The blue (red) color indicates improvement (degradation), while grey color indicates insignificant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

T_{min} for 24-, 48-, 72-, 96-, 120- and 144-forecast hour (FH) over 16–30 May 2014 period. The blue color indicates improvement while the red color means degradation. The forecasting results for the 6 FHs present similar patterns. The slight differences in the predicted T_{min} between ORG and OLP cases can be found in Fig. 2a with improvements over the small areas located in the northwestern China and degradations distributed in the northeastern Canada. However, this situation can be clearly improved by DA case in Fig. 2b. Relative to the OLP run, the GFS T_{min} predictions exhibited overwhelmingly positive responses to DA run on almost the entire study domain with some improvements larger than 2 K in Central Australia, southern India, and South Africa (Fig. 2b), whereas the slight degradations are mainly located over the Horn of Africa.

With respect to the NOAA-CPC surface air temperature analysis

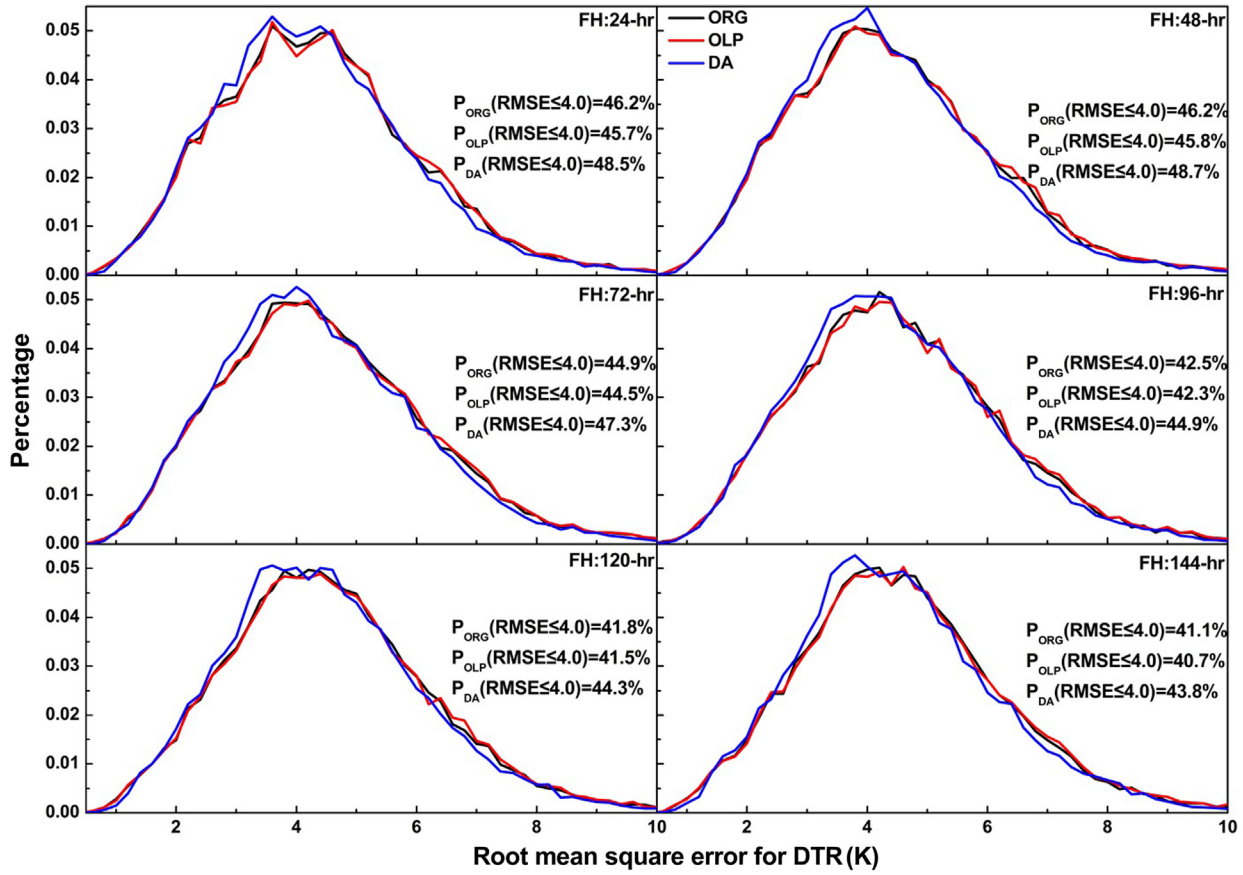

Fig. 5. With respect to the NOAA-CPC surface air temperature analysis product, the domain-averaged frequency probability as a function of RMSE for the predicted 2-meter diurnal temperature range (DTR) for 24-, 48-, 72-, 96-, 120- and 144-FH from May 16 to 30, 2014.

product, Fig. 3 shows statistical results for the RMSE values of the 24-, 48-, 72-, 96-, 120- and 144-FH T_{min} forecasts over the May 16–30, 2014 period. Curves shifting toward the left indicate improvements in reducing the probability of large RMSE while curves shifting towards the right indicate degradations. Compared to the ORG case, the accumulative probabilities of less than 3 K RMSE values for 24-, 48-, 72-, 96-, 120- and 144-FH T_{min} predictions are slightly improved by 0.1%, 0.4%, 0.9%, 1.1%, 1.2%, and 0.9% by OLP case, respectively. Relative to OLP run, DA case presents a left shifting tendency with increased probability of providing lower RMSE values. With benefits of ASCAT SM assimilation, the accumulative probabilities of less than 3 K RMSE values for the OLP run are increased by 1.3%, 1.0%, 0.8%, 0.7%, 0.9%, and 0.6% for 24-, 48-, 72-, 96-, 120- and 144-FH T_{min} predictions, respectively.

Similar to Fig. 2, with respect to the 50 km NOAA-CPC surface air temperature analysis product, Fig. 4 denotes RMSE differences in predicted DTR for 24-, 48-, 72-, 96-, 120- and 144-FH from May 16 to 30, 2014. The blue color shading shows improvement, yet the red color shading indicates degradation. The large degradations for the OLP case are observed in Central Asia, northeastern Canada, and western South America (Fig. 4a). The DA case yields significant improvement on DTR forecasts across the study domain in comparison to the OLP case by exhibiting larger than 3 K improvements in west U.S., central Australia, South Africa, and central Asia (Fig. 4b). The different patterns in Fig. 4b indicate that the DA case successfully forecasts DTR in comparison to OLP run and shows significant improvements over sparsely vegetated areas.

With respect to the NOAA-CPC surface air temperature analysis product, Fig. 5 shows the study domain-averaged frequency probability as a function of RMSE values for DTR during the May 16–30, 2014 period. The statistical density function of frequency probability shifting toward the left indicates improvement while shifts toward the right indicate degradation. Relative to the ORG case, the OLP case exhibits marginal influences on DTR forecasts from 24- to 144-FH, whereas the DA case performs better with significantly shifting toward the left. The OLP run shows slight degradations over the ORG case and reductions in accumulative probabilities of RMSE less than 4 K for 24-, 48-, 72-, 96-, 120- and 144-FH DTR forecasts by −0.5%, −0.4%, −0.4%, −0.2%, −0.3% and −0.4%, respectively. Compared to the OLP run, the accumulative probabilities of RMSE less than 4 K for 24-, 48-, 72-, 96-, 120- and 144-FH DTR forecasts are enhanced by 2.8%, 2.9%, 2.8%, 2.6%, 2.8%, 3.1% with benefits of ASCAT SM assimilation.

Based on the rainfall forecast skill scores in Table 1 and forecast accuracy computed according to Eq. (2), differences in predicted rainfall accuracy (in %) for 24-, 48-, 72-, 96-, 120- and 144-FH over the 16–30 May 2014 period are shown in Fig. 6. The blue color shading means improvement; the red color shading means degradation. Compared to the ORG case, the OLP run presents a modest performance in northwestern Sahara, Middle East, Canada, and northwestern China, but shows marginal improvements in northeast China and western U.S. (Fig. 6a). The pattern differences between OLP and DA cases demonstrate that the DA case is successful in the northern China and northwestern Sahara with showing improvements larger than 6%, whereas the degradations can be seen in the north India, east Sahara, and South Africa (Fig. 6b).

With respect to the 25 km CMORPH precipitation product, Fig. 7 illustrates study domain-averaged frequency probability as a function of forecasting accuracy for the 24-, 48-, 72-, 96-, 120- and 144-hour precipitation and non-precipitation predictions from May 16 to 30, 2014. Contrary to Figs. 3 and 5, the statistical density function of frequency probability shifting towards the left indicates degraded and shifting towards the right indiciates degradation. For the 24- and 48-hour precipitation forecasts, the OLP case demonstrates great improvements

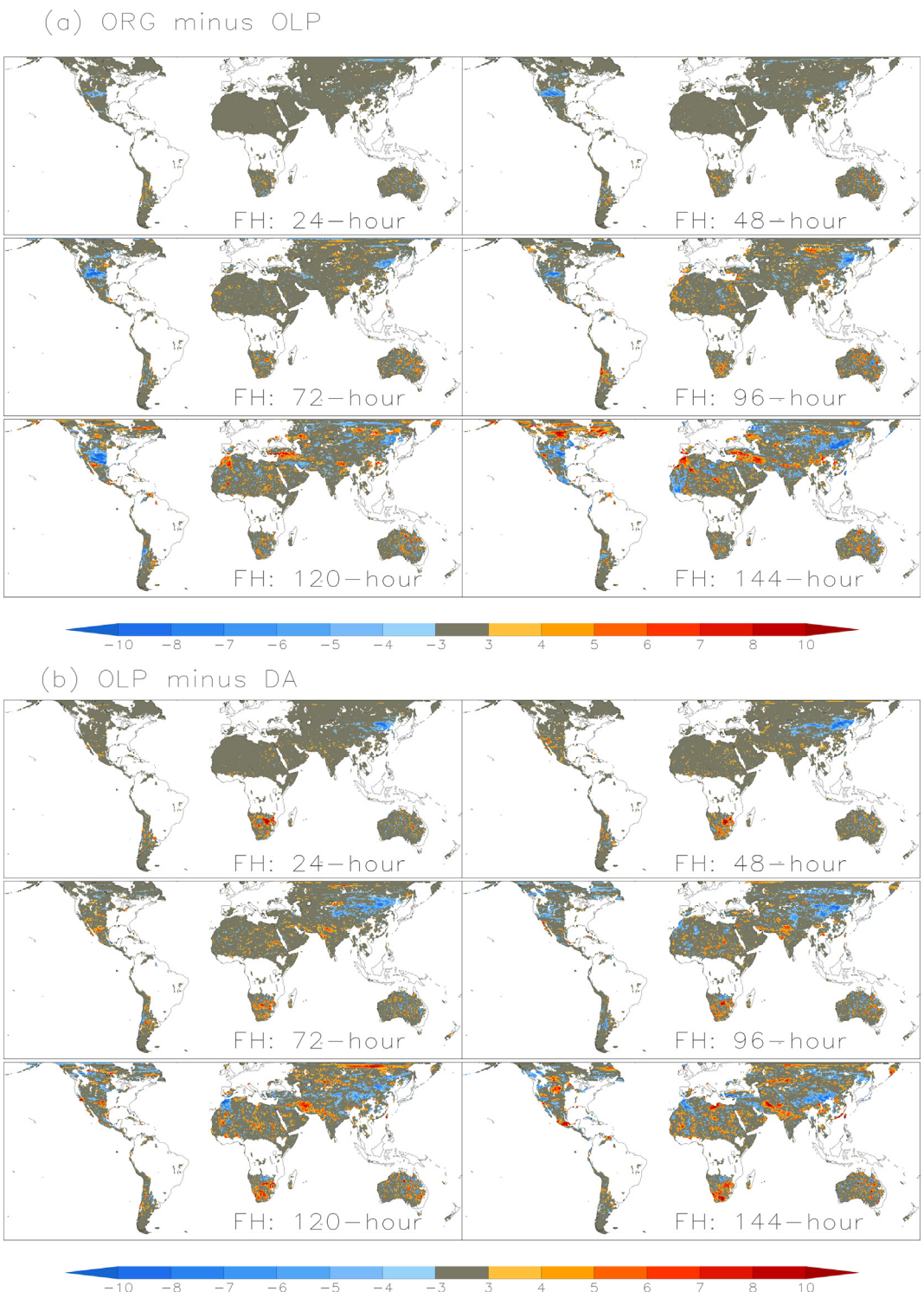

Fig. 6. With respect to the 25 km CMORPH precipitation product, differences in predicted rainfall accuracy (in Unit: %) for 24-, 48-, 72-, 96-, 120-, 144-FH over the May 16–30, 2014 period: (a) ORG minus OLP; and (b) OLP minus DA. The blue (red) color indicates improvement (degradation), while grey color indicates insignificant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

over the ORG case, while there are slight differences between OLP and DA cases. However, the DA case yields greater improvements for 72-, 96-, 120-, and 144-hour forecasts in comparison to the OLP case. Specifically, over the ORG case, the OLP case can enhance the accumulative probabilities of high quality precipitation forecasts, which are defined as the forecasts that have larger than 80% accuracy, by 1.4%, 2.2%, 2.7%, 1.6%, 1.8% and 1.8% for 24-, 48-, 72-, 96-, 120- and 144-FH, respectively. With benefits of assimilating ASCAT SM retrievals, the improvements for 24-, 48-, 72-, 96-, 120- and 144-FH are −0.4%, −0.4%, 0.4%, 3.1%, 0.8% and 2.9% in comparison with the OLP run, respectively.

6. Discussion

This work examined the impact of EnKF-based ASCAT SM assimilation on the forecasts of the NOAA NCEP GFS model over sparsely

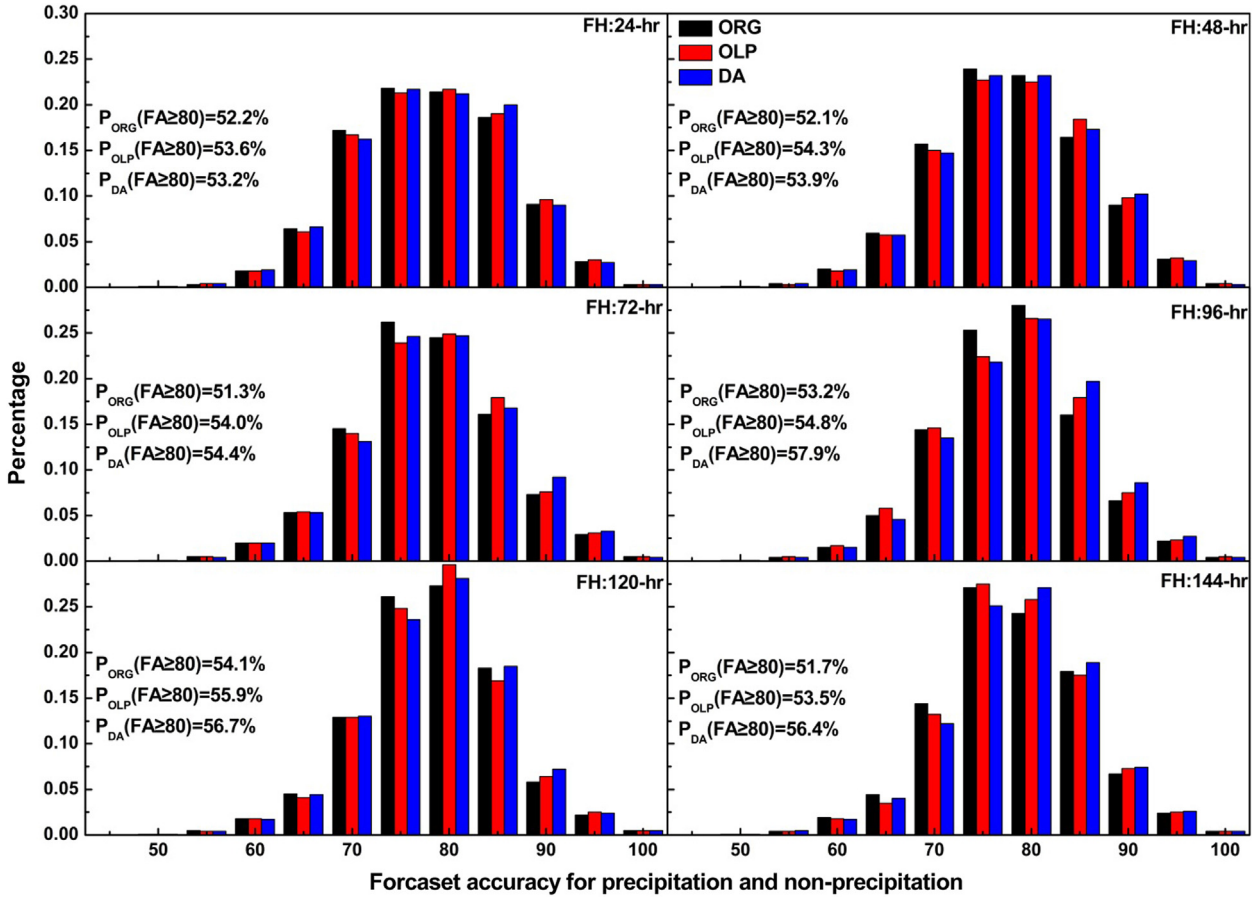

Fig. 7. With respect to 25 km CMORPH precipitation product, the study domain-averaged frequency probability as a function of forecasting accuracy for the predicted precipitation/non-precipitation for 24-, 48-, 72-, 96-, 120- and 144-FH from May 16 to 30, 2014. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

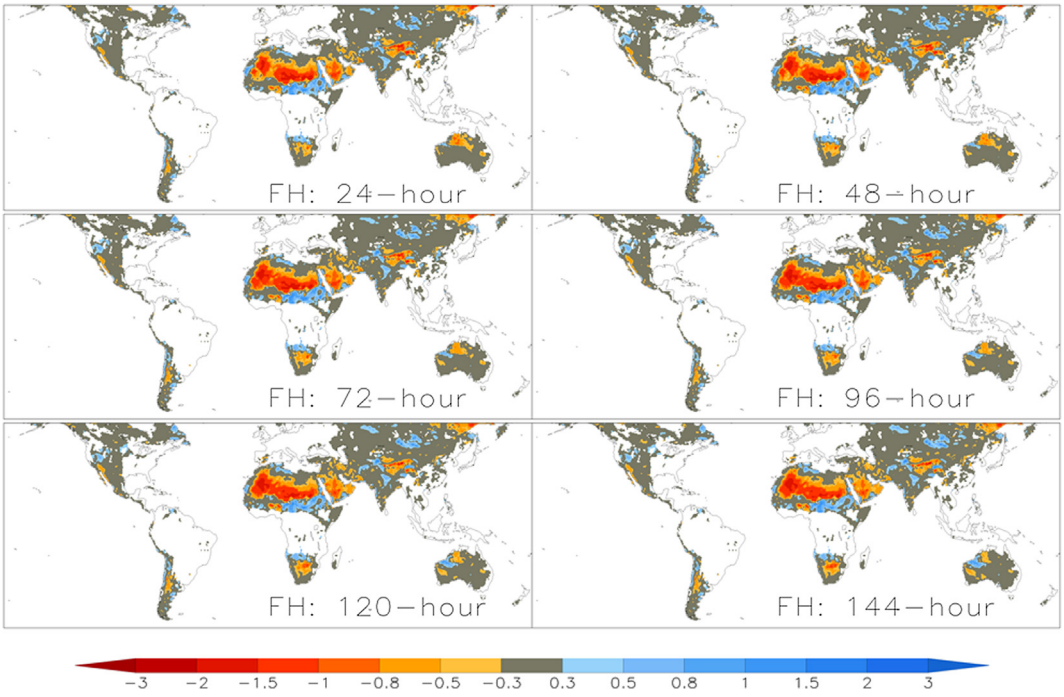

Fig. 8. With respect to 50 km NOAA-CPC surface air temperature analysis product, RMSE differences (OLP minus DA, in Unit: K) in predicted 2-meter maximum temperature (T_{max}) for 24-, 48-, 72-, 96-, 120-, 144-FH over the May 16–30, 2014 period. The blue (red) color indicates improvement (degradation), while grey color indicates insignificant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

vegetated areas (GVF less than 0.5). The results presented in Section 5 indicate that the GFS model performs better with benefits of assimilating ASCAT SM data into the coupled GFS-LIS system. The GFS-LIS coupled system produced 3-hour forecasts for 168-hr in advance for each day over a biweekly period, thus making the sample size 112 (8 three-hourly forecasts/day × 14-day) in Figs. 2, 4, 6 and 8. The statistical results shown in Figs. 3, 5 and 7 have huge sample sizes (roughly 10^8) with representation of the study domain-averaged RMSE patterns. The strong consistency of results implies that the results in this paper are qualitatively stable and likely represent a longer analysis period. However, several limitations should be considered in interpreting this result when applying global remote sensing SM data products in numerical weather prediction models. These limitations are discussed below.

6.1. Physical considerations

In this paper, clear improvements of ASCAT SM assimilation on GFS T_{min} and DTR forecasts over Sahara areas can be found in Figs. 2 and 4. However, the obvious degradations of assimilating ASCAT SM product on T_{max} predictions in the Sahara areas are exhibited in Fig. 8. On one hand, the results show that the ASCAT retrievals should include useful information in their seasonal cycle and anomaly signals over sand areas, which can utilize the model SM increments to correct T_{min} and DTR simulations; on the other hand, however, further improvements on ASCAT SM product over sand areas are expected to address noise reduction.

Previous studies indicate that the weather forecast model generally performs better at nighttime than daytime due to the benefits of ASCAT SM assimilation (Dharssi et al., 2011; Schneider et al., 2014); the results in this paper demonstrated this very clearly. The non-hydraulic equilibrium of the soil may lead to decoupling occurrence during the daytime (Leroux et al., 2014) while the remotely sensed SM retrievals are generally closer to the ground measurements during the nighttime (Albergel et al., 2009; Leroux et al., 2014). Thus ASCAT SM data of descending orbits (a.m) perform better than that of ascending orbits (p.m) (Wagner et al., 1999; Albergel et al., 2009).

Additionally, the daily ASCAT SM product was assimilated into the LIS-Noah model at 00Z of each day, and then LIS was set up to run in parallel with each GFS simulation, providing updated initial land surface condition for each 00Z GFS run (Fig. 1). There, the soil temperature and SM fields from LIS/Noah were placed into the GFS initial condition files. Benefits of assimilating the ASCAT SM product were thus directly used for the first 6-hour GDAS guess cycle and the T_{min} generally showed during the 00Z–06Z time. The T_{max} generally showed during the 12Z–18Z time during the third GDAS cycle runs. Positive information with benefits of data assimilation might be reduced when propagating from the first GDAS cycle to the third GDAS cycle. Meanwhile, model uncertainties generally tend to increase in the long term. However, the significant improvements on T_{min} and DTR forecasts over the global domain potentially overshadow the degradations on T_{max} estimations in the Sahara region.

6.2. Impacts of land surface parameters

In Noah LSM, vegetation information from land cover and GVF directly affects energy and water balances (Chen and Dudhia, 2001; Yin et al., 2015c, 2016). And the hydraulic conductivity in Noah model is associated with soil variables including soil texture and slop (Chen and Dudhia, 2001; Ek et al., 2003). Due to the differences of the land surface parameters including land cover, GVF, soil texture, and slope, there are distinct differences in forecasts between the ORG and OLP cases. The OLP and DA cases are thus driven by the same land surface parameters to ensure the model output differences only come from turning on/off the ASCAT SM data assimilation.

6.3. Additional future work

The current operational GFS version T1534 at NCEP is a global numerical weather prediction system at a base horizontal resolution of ~12 km (grid dimensions: 3072 at longitude by 1536 at latitude in Gaussian projection), which is higher than the T670-254 used in this paper, but their dynamical cores and mathematical processes of producing forecast outputs are consistent. The operational GFS T1534 performance is thus expected to be improved with the benefits of ASCAT SM data assimilation. However, the revisit time of ASCAT is 2–3 days. Compared to the ASCAT, Soil Moisture Operational Products System (SMOPS) developed at NOAA/NESDIS provides a near real time operational global blend of all available microwave soil moisture retrievals on a daily basis (Yin et al., 2015a, 2019). Building on the reasonable quality control of the blended SM retrievals (Yin et al., 2014), it can be expected to extend the improvements on GFS model performance from sparse vegetation areas to the entire global domain with benefits from the assimilation of the near real time SMOPS blended SM data in near future.

7. Conclusions

With the development of the coupled GFS-LIS system in this paper, three numerical experiments were carried out to examine the impacts of assimilating satellite soil moisture data into numerical weather prediction models. The 2-meter minimum temperature and 2-meter diurnal temperature range forecasts with benefits of ASCAT soil moisture assimilation are closer to the observations with high quality forecast probability, showing improvements of 1.7% and 3.1%, respectively. Positive impacts of assimilating ASCAT soil moisture product on precipitation predictions are also found with the study domain-averaged probability of high quality forecasts improved as much as 3.1%. Based on the results, assimilating remotely-sensed soil moisture observations into GFS model is suggested for NCEP GFS forecast operations.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was jointly supported by NOAA's Climate Program Office's Modeling, Analysis, Predictions, and Projections program and NOAA NGGPS-R2O program (NOAA-NWS-NWSPO-2015-2004117). We would like to thank Dr. Krishna Kumar from NOAA/NESDIS/ Joint Center for Satellite Data Assimilation (JCSDA) for helping build the coupled GFS-LIS system. Dr. Fanglin Yang and Dr. Helin Wei from NOAA-NCEP-EMC are thanked for their suggestions on analysis of the results from GFS model outputs. We are also grateful to Dr. Wei Shi and Dr. Pingping Xie from NOAA-CPC for providing the observations-based temperature and precipitation products.

References

Albergel, C., Rüdiger, C., Carrer, D., Calvet, J., Fritz, N., Naeimi, V., Bartalis, Z., Hasenauer, S., 2009. An evaluation of ASCAT surface soil moisture products with in-situ observations in Southwestern France. Hydrol. Earth Syst. Sci. 13 (2), 115–124.

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., Anderson, C., 2007. Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). Geophys. Res. Lett. 34, L20401. https://doi.org/10.1029/2007GL031088.

Burgers, G., Leeuwen, P.J.V., Evensen, G., 1998. Analysis scheme in the ensemble Kalman filter. Mon. Weather Rev. 126, 1719–1724.

Chen, F., Dudhia, J., 2011. Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. Month. Weather Rev. 129, 215–231.

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., Isaksen, L., 2013. A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. Q. J. R. Meteorol. Soc. 139 (674), 1199–1213.

Dharssi, I., Bovis, K.J., Macpherson, B., Jones, C.P., 2011. Operational assimilation of ASCAT surface soil wetness at the Met Office. Hydrol. Earth Syst. Sci. 15, 2729–2746.

Dharssi, I., Bovis, K., Macpherson, B., Jones, C., 2010. Assimilation of ASCAT surface soil wetness. UK Met Office Forecasting R&D Technical Report No. 548. http://research.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/reports/548.pdf.

Drusch, M., Viterbo, P., 2007. Assimilation of screen-level variables in ECMWF's integrated forecast system: a study on the impact on the forecast quality and analyzed soil moisture. Month. Weather Rev. 135, 300–314.

Ek, M.B., Mitchell, K.E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., Tarpley, J.D., 2003. implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. J. Geophys. Res. 108 (D22), 8851. https://doi.org/10.1029/2002JD003296.

Entekhabi, D., Njoku, E.G., O'Neill, P.E., et al., 2010. The soil moisture active passive (SMAP) mission. Proc. IEEE 98 (5), 704–716.

Evensen, G., 1994. Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99 (C5), 10143–10162.

Gringorten, I.I., 1951. The verification and scoring of weather forecasts. JASA 46 (255), 279–296.

Jiang, L., Kogan, F.N., Guo, W., Tarpley, J.D., Mitchell, K.E., Ek, M.B., Tian, Y., Zheng, W., Zou, C.Z., Ramsay, B.H., 2010. Real-time weekly global green vegetation fraction derived from advanced very high resolution radiometer-based NOAA operational global vegetation index (GVI) system. J. Geophys. Res. 115, D11114. https://doi.org/10.1029/2009JD013204.

Jones, L.A., Ferguson, C.R., Kimball, J.S., Zhang, K., Chan, S.T.K., McDonald, K.C., Njoku, E.G., Wood, E.F., 2010. Satellite microwave remote sensing of daily land surface air temperature minima and maxima from. AMSR-E IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 3, 111–123.

Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydromet. 5, 487–503.

Kerr, Y.H., Waldteufel, P., Wigneron, J.-P., et al., 2010. The SMOS mission: new tool for monitoring key elements of the globalwater cycle. Proc. IEEE 98 (5), 666–687.

Koster, R.D., et al., 2004. Regions of strong coupling between soil moisture and precipitation. Science 305, 1138–1140.

Koster, R.D., Guo, Z., Yang, R., Dirmeyer, P.A., Mitchell, K., Puma, M.J., 2009. On the nature of soil moisture in land surface models. J. Clim. 22, 4322–4325.

Kumar, S.V., et al., 2006. Land information system: an interoperable framework for high resolution land surface modeling. Environ. Modell. Software 21, 1402–1415.

Kumar, S.V., Reichle, R.H., Peters-Lidard, C.D., Koster, R.D., Zhan, X., Crow, W.T., Eylander, J.B., Houser, P.R., 2008. A land surface data assimilation framework using the Land Information System: description and application. Adv. Water Resour. 31, 1419–1432.

Kumar, S.V., Reichle, R.H., Koster, R.D., Crow, W.T., Peters-Lidard, C.D., 2009. Role of subsurface physics in the assimilation of surface soil moisture observations. J. Hydrometeor. 10, 1534–1547.

Leroux, D.J., Kerr, Y.H., Al Bitar, A., Bindlish, R., Jackson, T.J., Berthelot, B., et al., 2014. Comparison between SMOS, VUA, ASCAT, and ECMWF soil moisture products over four watersheds in US. IEEE Trans. Geosci. Remote Sens. 52 (3), 1562–1571.

Li, L., Gaiser, P.W., Gao, B.-C., Bevilacqua, R.M., Jackson, T.J., Njoku, E.G., Rudiger, C., Calvet, J.-C., Bindlish, R., 2010. WindSat global soil moisture retrieval validation. IEEE Trans. Geosci. Remote Sens. 48 (5), 2224–2241.

Liu, Y.Y., Parinussa, R.M., Dorigo, W.A., et al., 2011. Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrol. Earth Syst. Sci. 15, 425–436.

Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., Wagner, W., 2009. An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations. IEEE Trans. Geosci. Remote Sens. 47 (7), 1999–2013.

Parinussa, R.M., Wang, G., Holmes, T.R.H., Liu, Y.Y., Dolman, A.J., de Jeu, R.A.M., Jiang, T., Zhang, P., Shi, J., 2014. Global surface soil moisture from the Microwave Radiation Imager onboard the Fengyun-3B satellite. Int. J. Remote Sens. 35 (19), 7007–7029.

Reichle, R., Kumar, S., Mahanama, S., Koster, R., 2010. Assimilation of satellite-derived skin temperature observations into land surface models. J. Hydrometeor. 11, 1103–1122.

Reichle, R.H., Mclaughlin, D.B., Entekhabi, D., 2002. Hydrologic data assimilation with the ensemble Kalman filter. Month. Weather Rev. 130, 103–114.

Ryu, D., Crow, W.T., Zhan, X., Jackson, T.J., 2008. Correcting unintended perturbation biases in hydrologic data assimilation. J. Hydrometeor. 10, 734–750.

Schneider, S., Wang, Y., Wagner, W., Mahfouf, J.-F., 2014. Impact of ASCAT soil moisture assimilation on regional precipitation forecasts: a case study for Austria. Month. Weather Rev. 142, 1525–1540.

Scipal, K., Drusch, M., Wagner, W., 2008. Assimilation of a ERS scatterometer derived soil moisture index in the ECMWF numerical weather prediction system. Adv. Water Resour. 31, 1101–1112.

Seneviratne, S.I., Corti, T., Davin, L.E., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., Teuling, A.J., 2010. Investigating soil moisture–climate interactions in a changing climate: a review. Earth Sci. Rev. 99, 125–161.

Van Emmerik, T., Steele-Dunne, S.C., Judge, J., van de Giesen, N., 2015. Impact of diurnal variation in vegetation water content on radar backscatter from maize during water stress. IEEE Trans. Geosci. Remote Sens. 53 (7), 3855–3869.

Vislocky, R.L., Young, G.S., 1988. Improving your weather forecasts through a better knowledge of skill scores. Forecast Skill Scores 13 (3), 15–17.

Wagner, W., et al., 2013. The ASCAT soil moisture product: A review of its specifications, validation results, and merging applications. Meteorologische Zeitschrift 22, 5–33.

Wagner, W., Lemoine, G., Rott, H., 1999. A method for estimating soil moisture from ERS scatterometer and soil data. Remote Sens. Environ. 70 (2), 191–207.

Yin, J., Zhan, X., Zheng, Y., Liu, J., Hain, C.R., Fang, L., 2014. Impact of quality control of satellite soil moisture data on their assimilation into land surface model. Geophys. Res. Lett. 41, 7159–7166. https://doi.org/10.1002/2014GL060659.

Yin, J., Zhan, X., Zheng, Y., Hain, C., Liu, J., Fang, L., 2015a. Optimal ensemble size of Ensemble Kalman Filter in sequential soil moisture data assimilation of land surface model. Geophys. Res. Lett. 16 (28), 6710–6715. https://doi.org/10.1002/2015GL063366.

Yin, J., Zhan, X., Zheng, Y., Liu, J., Fang, L., Hain, C.R., 2015b. Enhancing model skill by assimilating SMOPS blended soil moisture product into Noah land surface model. J. Hydrometeorol. 16, 917–931.

Yin, J., Zheng, Y., Zhan, X., Hain, C.R., Zhai, Q., Duan, C., Wu, R., Liu, J., Fang, L., 2015c. An assessment of impacts of land-cover changes on root-zone soil moisture. Int. J. Remote Sens. 36 (24), 6116–6134.

Yin, J., Zhan, X., Zheng, Y., Hain, C.R., Ek, M., Wen, J., Fang, L., Liu, J., 2016. Improving noah land surface model performance using near real time surface albedo and green vegetation fraction. Agric. For. Meteorol. 218–219, 171–183.

Yin, J., Zhan, X., Liu, J., Shull, M., 2019. An inter-comparison of noah model skills with benefits of assimilating SMOPS blended and individual soil moisture retrievals. Water Resour. Res. https://doi.org/10.1029/2018WR024326.

Zeng, J., Li, Z., Chen, Q., Bi, H., Qiu, J., Zou, P., 2015. Evaluation of remotely sensed and reanalysis soilmoisture products over the Tibetan Plateau using in-situ observations. Remote Sens. Environ. 163 (15), 91–110.

Zhan, X., Zheng, W., Meng, J., Dong, J., Ek, M., 2012. Impact of SMOS soil moisture data assimilation on NCEP-GFS forecasts. Geophys. Res. Abstr. 14 EGU2012–12724-1.