Article

Applicability of Multi-Frequency Passive Microwave Observations and Data Assimilation Methods for Improving Numerical Weather Forecasting in Niger, Africa

Mohamed Rasmy [1,*], Toshio Koike [1] and Xin Li [2]

[1] Department of Civil Engineering, The University of Tokyo, Tokyo 113-8656, Japan;
  E-Mail: tkoike@hydra.t.u-tokyo.ac.jp

[2] Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of
  Sciences, Lanzhou 730000, China; E-Mail: lixin@lzb.ac.cn

* Author to whom correspondence should be addressed; E-Mail: rasmy@hydra.t.u-tokyo.ac.jp;
  Tel.: +81-358-418-874.

Abstract: The development of satellite-based forecasting systems is one of the few affordable solutions for developing regions (e.g., West Africa) that cannot afford ground-based observation networks. Although low-frequency passive microwave data have been used extensively for land surface monitoring, the use of high-frequency passive microwave data that contain cloud information is very limited over land because of strong heterogeneous land surface emissions. The Coupled Atmosphere and Land Data Assimilation System (CALDAS) was developed by merging soil moisture information estimated from low-frequency data with corresponding high-frequency data to estimate cloud information and, thus, improve weather forecasting over Niger, West Africa. The results showed that the assimilated soil moisture and cloud distributions were reasonably comparable to satellite retrievals of soil moisture and cloud observations. However, assimilating soil moisture alone within a mesoscale model produced only marginal improvements in the forecast, whereas the assimilation of both soil moisture and cloud distributions improved the simulation of temperature and humidity profiles. Rainfall forecasts from CALDAS also correlated well with satellite retrievals. This indicates the potential use of CALDAS as a reliable forecasting tool for developing regions. Further developments of CALDAS and the inclusion of data from several other sensors will be researched in future studies.

Nomenclature

AMSR	Advanced Microwave Scanning Radiometer
AMSR-E	AMSR on Earth Observing System
AMSR-2	AMSR-2nd generation on GCOM-W1 satellite
NCEP	National Centers for Environmental Prediction
LPRM	NASA Land Parameter Retrieval Model
ARPS	The Advanced Regional Prediction System
CALDAS	Coupled Atmosphere and Land Data Assimilation System
EnKF	Ensemble Kalman Filter
CMDAS	Cloud Microphysics Data Assimilation System
LDAS	Land Data Assimilation Sysatem
LDAS-A	LDAS coupled with an Atmospheric model
NWP	Numerical Weather Prediction
IR data	InfraRed data
SiB2	Simple Biosphere model version-2
TRMM	Tropical Rainfall Measuring Mission

1. Introduction

Africa is a continent that is vulnerable to the effects of global warming and climate change/variability, which will increase the incidences of abnormal weather, such as frequent flooding and prolonged drought [1]. Numerical Weather Prediction (NWP) in West Africa is still underdeveloped because of several problems (e.g., lack of data, inadequate forecast system, difficulties in simulating land/ atmospheric processes, and lack of forecast verification exercises). NWP is mainly associated with the initial-boundary-values problem; *i.e.*, given an estimate of the initial state of the atmosphere and land surface, the model forecasts the evolutions of both. A greater number of high quality observations that more fully represent the complete nature of the initial atmosphere and land surface would greatly improve the forecasting capabilities of NWP models.Therefore, the lack of reliable in situ data with which to initialize the NWP model is the most challenging problem regarding forecasting in underdeveloped regions.

The capability of remote sensing to contribute to the prediction of the Earth's weather and climatic systems has been demonstrated in many scenarios. Although many studies have been based on the use of satellite data in West Africa, only a few have considered the use of satellite data for NWP modeling and validation [2]. Accordingly, this research investigates the use of passive microwave satellite observations with a mesoscale model in West Africa. In microwave frequencies, the dielectric property of materials allows the quantitative estimation of moisture quantities such as soil moisture, vegetation water, snow water content, and cloud condensate [3–6]. Thus, the potential for the use of space-borne microwave

observations in NWP has been advanced considerably by the launch of several sophisticated platforms (e.g., Terra, Aqua, and the Global Change Observation Mission).

Furthermore, low-frequency microwave observations provide information on soil moisture that is very useful for enhancing NWP model forecasting through improved initialization [7,8]. However, the few studies that have been performed on the assimilation of soil moisture observations into NWP models have shown several limitations (e.g., retrieval and preprocessing of soil moisture products, Cumulative Distribution Function matching, difficulties in defining an observation operator that changes spatially and temporally, and time constraints for the use of near-real-time applications). Thus, they recommended the application of assimilation schemes that are more advanced [9,10]. To overcome such problems, an on-line system (*i.e.*, Land Data Assimilation System coupled with Atmospheric model (LDAS-A)) was developed to directly assimilate lower frequency microwave radiance from the Advanced Microwave Scanning Radiometer (AMSR) for updating realistic soil moisture content in a mesoscale model and the results were presented in [11]. However, because of limited satellite observations (e.g., AMSR-E and AMSR-2, twice daily maximum), the improved land surface conditions often suffered from significant errors and drift due to biases in the predicted forcing (e.g., rainfall and solar radiation) that misguided the subsequent forecast [11,12]. Thus, this particular problem rendered the on-line land data assimilation ineffective in NWP models.

Clouds directly affect surface forcing (*i.e.*, downward radiation and rainfall), and thus they affect the estimation of the Earth's surface water and energy budgets. Similar to microwave sounding observations (e.g., Advanced Microwave Sounding Unit (AMSU)), passive microwave images (e.g., AMSR-E and AMSR-2) also contain information on cloud fields at higher frequencies (e.g., 89 GHz). It is also noteworthy to mention the Global Precipitation Measurement mission, which is the first coordinated international satellite network to provide near-real-time observations of radar and microwave images across a range of frequencies (10–183 GHz) every 3 h anywhere on the globe. However, the increasing abundance of data arising from several platforms has not been used well within NWP models. In fact, the 89-GHz frequency is more sensitive to cloud information compared with other channels of the AMSR-E, but it also contains information of land surface emission. Obtaining atmospheric information from combined land-atmosphere signals is very challenging because the strong heterogeneous land surface emissions mask the very weak atmospheric signals. When land surface emissions are estimated as accurately as possible, then the weak signal of the cloud fields from the combined signals can be converted into realistic representations of the cloud fields for the NWP models.

Consequently, the Coupled Atmosphere and Land Data Assimilation System (CALDAS) was developed as an improvement of the LDAS-A. CALDAS assimilates lower frequency data of AMSR-E/AMSR-2 to improve the representation of land surface conditions and it merges them with higher frequency data of AMSR-E/AMSR-2 to improve the representation of cloud conditions over land surfaces. CALDAS results from the Tibetan Plateau have shown that simultaneous assimilation of land and cloud conditions improved the biases in land surface states, cloud representation, and forcing to land surface models, which enhanced land-atmosphere interactions [12].

As a continuation from our previous study [12], this work addresses the application of CALDAS as a weather forecasting system, simulating short-term weather conditions over a mesoscale domain

of Niger in West Africa. Model performance is investigated using available ground-based data and satellite observations.

2. Dataset, Models and Method

2.1. Dataset

In this research, data from the National Centers for Environmental Prediction (NCEP) were used for model simulations, and AMSR-E brightness temperature data were used for soil moisture and cloud data assimilation. The other data sets listed under this section were used for the analysis and validation of the model outputs.

2.1.1. Initial and Boundary Conditions

The initial and lateral boundary conditions were derived from the NCEP data. These NCEP-FNL (Final) operational global analysis data are available as a $1° \times 1°$ grid at 6-hourly intervals. To obtain the corresponding initial and boundary conditions required to run the mesoscale model, analysis data available at 26 pressure levels were used. Variables (e.g., pressure, geopotential height, temperature, relative humidity, and u and v winds) required for the mesoscale model simulations were interpolated from the pressure levels.

2.1.2. AMSR-E Brightness Temperature Data

AMSR-E is a total power passive microwave radiometer that measures horizontally and vertically polarized brightness temperatures at separate frequencies of 6.925, 10.65, 18.7, 23.8, 36.5, and 89.0 GHz. Its individual measurements have spatial resolutions varying from ∼0.05° at 89.0 GHz to ∼0.5° at 6.925 GHz. Because of spatial resolution differences, we used the nearest neighbor interpolation method to resample the data to fit the horizontal resolution of our model (∼0.05°). We used vertical polarizations of 6.925 and 10.65 GHz to retrieve soil moisture information, because the atmosphere can be considered transparent at these frequencies. The 89-GHz band of the AMSR-E shows high sensitivity to cloud water content and precipitation compared with other channels. Consequently, the 89-GHz vertical polarization was used to assimilate cloud information over the land surface. The motivation for using vertical polarization is that it is less sensitive than horizontal polarization to land surface roughness and vegetation. The calibrated brightness temperature (Level 1B) data were obtained from the Japan Aerospace Exploration Agency (JAXA).

2.1.3. *In Situ* Data

In situ data were obtained from the African Monsoon Multidisciplinary Analysis database. Atmospheric profile observations were retrieved from radiosondes (Vaisala radiosonde RS92) launched four times daily from the Niamey (2.17°E, 13.48°N) station in June 2006. The associated soil conditions were either missing or unsuitable for validation during the simulation period. Consequently, comparisons were performed using satellite-derived soil moisture products.

2.1.4. Satellite-Derived Soil Moisture Products

JAXA's soil moisture products are based on AMSR-E data by applying the 10–36 GHz algorithm for simultaneous retrieval of the soil moisture and vegetation water content from two indices: the polarization index and index of soil wetness [13]. These global moisture products are available twice daily at 0.1° spatial resolution.

National Aeronautics and Space Administration (NASA) soil moisture products were derived from the AMSR-E data and Land Parameter Retrieval Model (LPRM). The LPRM uses dual polarized channels (6.925 or 10.65 GHz) for the retrieval of both surface soil moisture and vegetation water content [14]. These global moisture products are also available twice daily at 0.25° spatial resolution.

2.1.5. Satellite Cloud and Rainfall Products

The global full-resolution (∼0.04°) infrared (IR) data, merged from several geostationary satellites (∼11 micron channels), and available from NASA at 30-min intervals, were used as cloud observations. Lower brightness temperatures in the IR images indicate sufficiently thick clouds whose cloud tops radiate at the atmospheric temperature at higher altitudes.

Rainfall data derived from the Tropical Rainfall Measuring Mission (TRMM) were used for model validation. These gridded estimates (3B42) are available with 3-hourly temporal resolution and $0.25° \times 0.25°$ spatial resolution in a global belt extending across latitudes from 50°S–50°N.

2.2. Models

CALDAS (Figure 1) has three subsystems: (1) a land-atmosphere coupled mesoscale model (ARPS-SiB2); (2) a Land Data Assimilation System (LDAS); and (3) a Cloud Microphysics Data Assimilation System (CMDAS). Here, the combinations of the ARPS-SiB2-LDAS and ARPS-SiB2-LDAS-CMDAS models are named LDAS-A and CALDAS, respectively. All three models (*i.e.*, ARPS-SiB2, LDAS-A, and CALDAS) shared the same atmosphere-land couple model (*i.e.*, ARPS-SiB2) and same settings of physics and parameterizations. ARPS-SiB2 model simulation was a simple dynamical downscaling of GCM analysis data. The only difference between ARPS-SiB2 and LDAS-A simulations is that LDAS-A used additional satellite observed soil moisture contents compared to ARPS-SiB2. Similarly, the only difference between LDAS-A and CALDAS simulations is that CALDAS used additional satellite observed cloud data. Detailed information and formulations on LDAS-A and CALDAS are given in [11,12], respectively; however, the following sections provide a brief explanation of their formulation.

The Advanced Regional Prediction System (ARPS) is a comprehensive regional- to storm-scale prediction model, and its atmospheric prediction component is a three-dimensional nonhydrostatic compressible model [15]. The LDAS consists of the Simple Biosphere model version 2 (SiB2) [16], which functions as a model operator for LDAS as well as the land surface scheme for ARPS, together with a physics-based radiative transfer model as the observation operator, and an Ensemble Kalman Filter (EnKF) [17] as the assimilation algorithm.

Figure 1. Schematic diagram of the Coupled Atmosphere and Land Data Assimilation System (CALDAS).
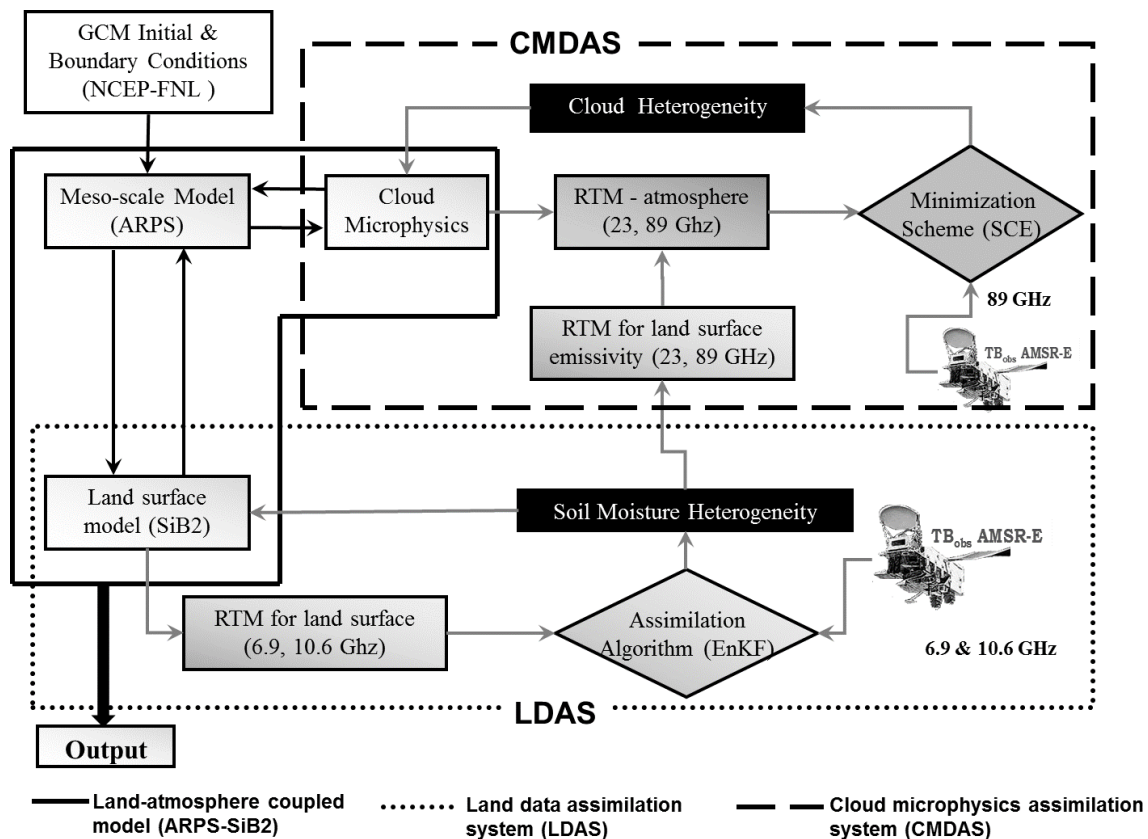
A land surface radiative transfer algorithm has been developed based on the large contrast between the dielectric constants of dry soil (∼4) and water (∼80) for lower microwave frequencies (*i.e.*, 1–15 GHz). By neglecting atmospheric and rainfall effects and by assuming that the reflection at the surface is much less than the radiation from the surface and vegetation layers at lower frequencies, the brightness temperature T_b at the satellite level is given by [18]:

$$T_b = (1 - R_p) * T_s * \exp(-\tau_c) + (1 - \omega_c) * T_c * (1 - \exp(-\tau_c)) \tag{1}$$

here, R_p is the surface reflectivity and T_s is the surface physical temperature (K), ω_c is the single-scattering albedo of the canopy, τ_c is the vegetation optical thickness and T_c is the canopy temperature (K).

In the field of hydrology, the EnKF has been applied to soil moisture estimation and it has been found to perform well against the variational assimilation method [19,20]. An overview of the EnKF for soil moisture assimilation is given below.

Consider X=$[w_1, \ w_2, \ w_3]^T$ as a state variable and the first estimate, for which w_1, w_2 and w_3 are the soil moisture contents of the surface, root and deep soil layers, respectively. The first estimate is used to create an ensemble of size (N) by adding pseudo-random noise with known statistics. By dropping the time notation, each member of state variable X_i is given by

$$X_i = \bar{X} + e_i \qquad e_i (i = 1, 2 .... N) \sim N(0, P) \tag{2}$$

where e_i is the random error vector of each member obtained from a multivariate Gaussian distribution with zero mean and error covariance matrix P, and $\bar{X}$ is the expectation of the first estimate X. In the forecast step, the forecast state member X_i^f is determined from the nearest analysis state member X_i^a according to

$$X_i^f = M(X_i^a) + u_i \qquad u_i \sim N(0, Q) \tag{3}$$

where M is the model operator and u_i is the model error vector of each member, obtained from a multivariate Gaussian distribution with zero mean and error covariance matrix Q.

In the analysis step, the AMSR-E observation data are perturbed by adding a random observation error and each member of the analyzed state variable X_i^a is updated as

$$X_i^a = X_i^f + K((Y_o + v_i) - H(X_i^f)) \qquad v_i \sim N(0, R) \tag{4}$$

where K is a Kalman gain matrix, H is the observation operator, Y_o is the observation, R is the observation error covariance and v_i is a random error vector of the observation with zero mean and covariance matrix R.

CMDAS was developed for use over only sea surfaces to improve the atmospheric moisture fields by assimilating higher-frequency AMSR-E observations [21]. To apply CMDAS over a land surface, land surface emissivity derived from the assimilated soil moisture was used as a boundary condition. To estimate the effects of atmospheric absorption, emission, and scattering on the upwelling radiation at 89 GHz, the 4-stream fast model [22] was used, and the Shuffled Complex Evolution technique [23] was adopted as a minimization scheme.

Compared with land data assimilation (50 members), cloud data assimilation is computationally very expensive because of the several hundred iterations of cloud parameters that are required to minimize the cost function. Nevertheless, the CALDAS model is enabled on parallel computing technology to satisfy the increasingly high-performance computing requirements of operational NWP models.

2.3. Method

1. As shown in Figure 1, the land-atmosphere mesoscale model (ARPS-SiB2) was established using initial and boundary conditions from NCEP-FNL data.
2. The ARPS model was integrated for a predefined period (10 min) and the calculated atmospheric forcing data transferred to the SiB2 model.
3. At the beginning of the SiB2 integration, the ensemble (50 members) of soil moisture profiles was generated. SiB2 was executed independently for each ensemble member of the soil moisture profile, retaining the same model parameters and atmospheric forcing. At the end of the SiB2 calculation, the mean values of the updated soil state and fluxes were computed and fed back to the ARPS model as the lower boundary conditions of the atmospheric model. Then, the ARPS-SiB2 model was integrated forward in time.
4. At times, when AMSR-E observations were available, the brightness temperatures at 6.9 and 10.65 GHz were perturbed to produce an ensemble of observations with prescribed statistics. The SiB2-driven ensemble of soil moisture profiles, surface temperature, and canopy temperature

were used to obtain the simulated brightness temperatures using the forward microwave radiative transfer model. The EnKF calculated the assimilated soil moisture profiles using simulated and observed brightness temperatures, as shown in Equation (4). In the case of soil moisture assimilation (no cloud assimilation), the updated soil state and fluxes were fed back to the ARPS model and the ARPS-SiB2 model was integrated forward in time.

5. In the case of cloud data assimilation, CMDAS was activated as soon as the LDAS completed the soil moisture assimilations. The control variables (profiles of temperature, specific humidity, pressure, air density, mixing ratio of cloud water, rain water, hail, snow, and cloud ice) were obtained from ARPS as an initial state to run the model operator (Lin's ice microphysics [24]). The 4-stream fast model calculated the modeled brightness temperatures for 89 GHz at the satellite level by considering the land surface as the lowest boundary. Land surface emissivity was calculated using assimilated soil moisture content. The Shuffle Complex Evolution scheme was used to estimate the assimilated cloud parameters (*i.e.*, cloud liquid water, snow, and rain) by minimizing the cost function calculated between the modeled and observed brightness temperatures. Then, the updated soil state and fluxes were fed back to the ARPS-SiB2 model.

6. Finally, with the reinitialized land surface and atmospheric conditions, the ARPS-SiB2 model was integrated forward in time to predict the land and atmospheric evolution until the next AMSR-E observations were available. The results from the ARPS-Sib2 model were recorded at 30-min intervals.

3. Experiment Descriptions

The performance of CALDAS was assessed based on three simulations: (1) the ARPS-SiB2 (hereafter ARPS) run, where a one-way nesting procedure employed the land-atmosphere mesoscale model without any assimilation; (2) the LDAS-A run, in which an ARPS run was accompanied by sequential land data (soil moisture) assimilation; and (3) the CALDAS run, in which the LDAS-A run was accompanied by the cloud microphysics data assimilation.

Study Domain and Model Configuration

Two model domains, domain-1 ($\sim$0.23° horizontal resolution) and domain-2 ($\sim$0.05° horizontal resolution) were established (Figure 2). Domain-1 was used to downscale the initial and lateral boundary conditions derived from NCEP-FNL by ARPS to domain-2. Then, ARPS, LDAS-A , and CALDAS were applied to domain-2 independently. The physical parameterization options were configured with a 1.5-order turbulent kinetic-energy-based closure scheme for sub-grid-scale turbulent mixing, latitude-dependent Coriolis parameters, Kain and Fritsch cumulus parameterization [25] (only for domain-1), and Lin ice microphysics.

Because integrated values of moisture fields were assimilated by CALDAS, the assimilated values were distributed with predefined profiles. Cloud Liquid Water Content (CLWC) was assumed to have a parabolic distribution (single-layer cloud) with zero values both above the top and below the bottom of the cloud layers. The lower and upper bounds of integrated CLWC were set to 0 and 1.5 kg·m^{-2}, respectively.

Figure 2. Mesoscale model domains.
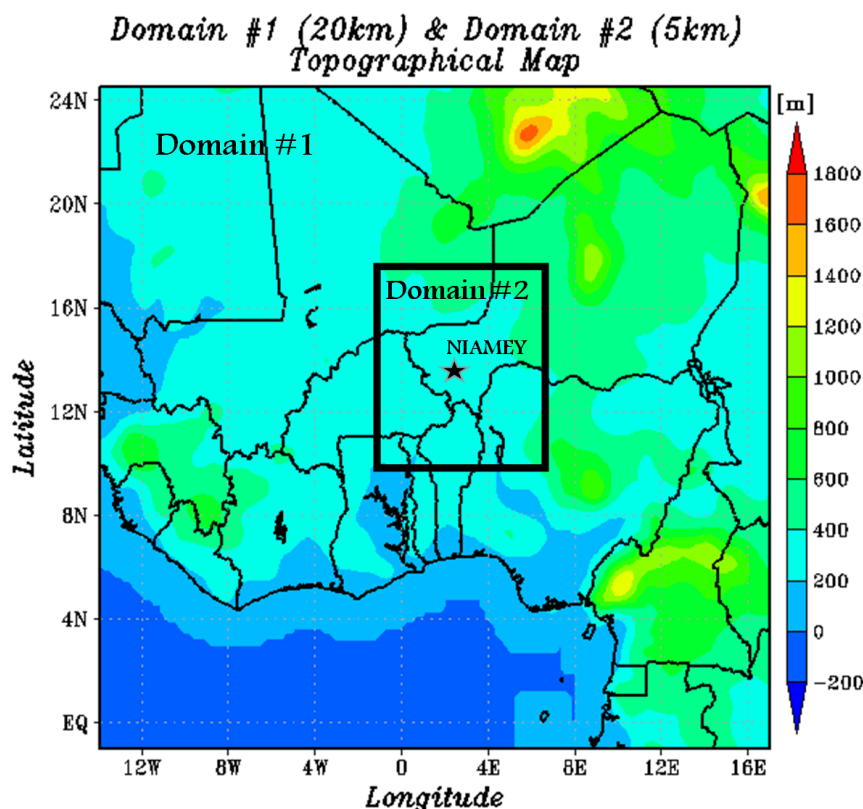
In this study, the assimilation of the 23-GHz channel for water vapor was not performed because such assimilation was found to cause errors (*i.e.*, underestimation) in the assimilated water vapor profile. This underestimation in turn caused assimilated clouds to evaporate quickly to compensate for the water vapor deficiency (not shown). To avoid this particular problem, air was assumed saturated only within the cloud layer. However, more accurate information on the water vapor profiles would improve the development of convective systems in the model and thus, enhance rainfall prediction. Therefore, we are investigating the use of water vapor profile information within the CALDAS system, obtained from the AMSU and Atmospheric Infrared Sounder (AIRS).

The distributions of rain and snow followed a skewed profile. They began to form at the cloud tops, grew to their maximum at the cloud bottoms, and subsequently decreased because of evaporation and the breakup of the raindrops or snowflakes. Further information on the model and the setup can be referenced from [12].

4. Results and Discussions

This study investigated the applicability of multi-frequency passive microwave observations for improving weather forecasting. Simulations were performed during a cloudy period starting from 6 June 2006 to 7 June 2006. The AMSR-E observed full coverage of domain-2 at 0140 UTC on 6 June 2006, which captured the cloud activities over the domain. The results from domain-2 are discussed in the following sections.

4.1. Distribution of Surface Soil Moisture

To examine the reliability of assimilated land surface emission, model simulations of soil moisture content were compared with AMSR-E brightness temperatures at 6.9 GHz and satellite-derived soil moisture products immediately after land data assimilation. Although a linear relationship cannot be assumed between the surface soil moisture content and brightness temperatures at 6.9 GHz, visual classification can be performed based on lower brightness temperatures being related to higher soil moisture content. Figure 3a,b represent the spatial distributions of simulated surface soil moisture for both the ARPS and the CALDAS models, and Figure 3c represents the distribution of brightness temperatures at 6.9 GHz. The magnitude of the soil moisture content simulated by ARPS (derived from NCEP-FNL) was higher in most of the model grids (particularly in the south) and produced completely different spatial distributions when compared with the observed brightness temperatures. Conversely, the distribution of CALDAS soil moisture content was reasonably comparable with that of brightness temperature at 6.9 GHz, and the clusters of dry and wet regions observed at 6.9 GHz were well defined in the CALDAS simulation.

Figure 3. Spatial distribution of volumetric surface soil moisture (m^3/m^3) at 0200 UTC on 6 June 2006; (**a**) ARPS; (**b**) CALDAS; (**c**) dvanced Microwave Scanning Radiometer (AMSR)-E brightness temperature (K) observed at 6.9 GHz; (**d**) Japan Aerospace Exploration Agency (JAXA) product; and (**e**) National Aeronautics and Space Administration (NASA) product, respectively.
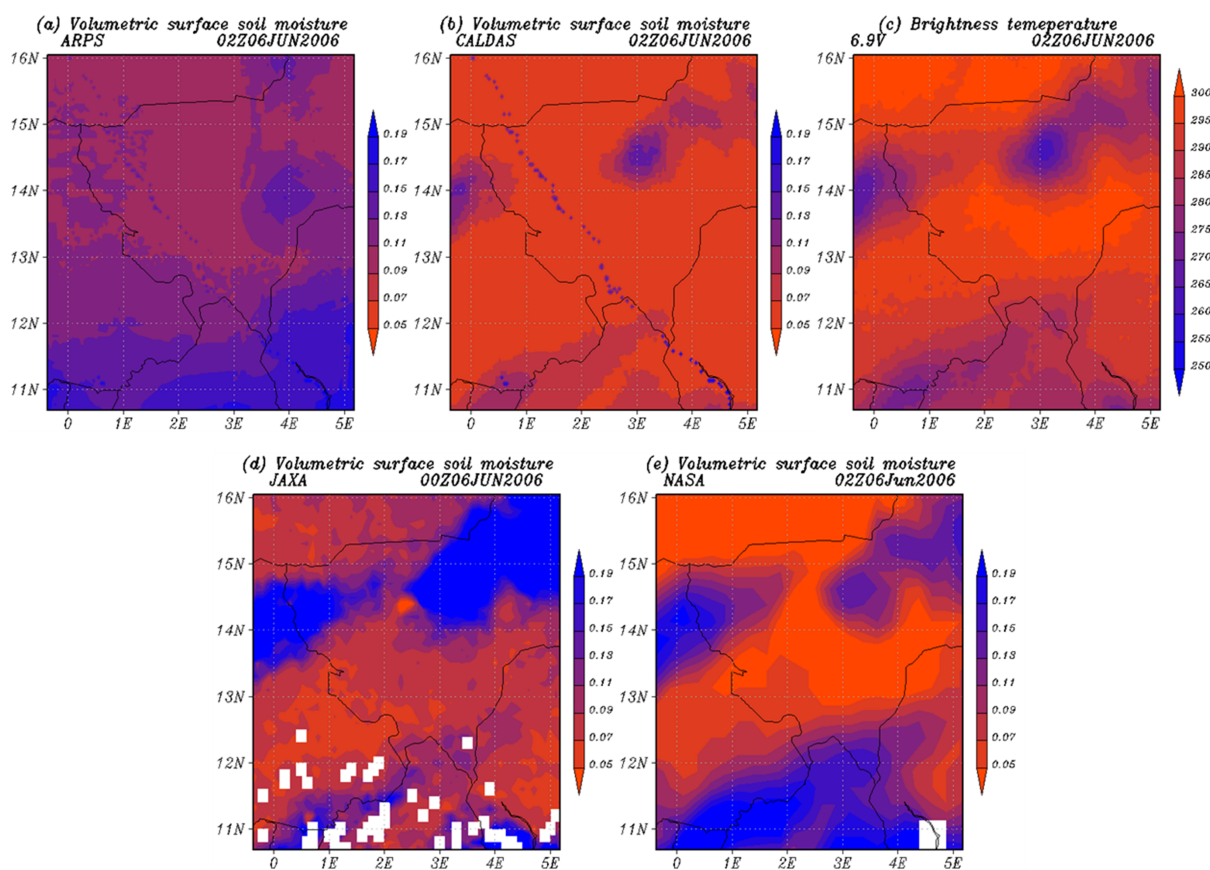
Quantitative validation of assimilated soil moisture content was not possible owing to lack of *in situ* soil moisture data during this period. Therefore, JAXA and NASA (c-band) soil moisture products were obtained for independent verification of the assimilated soil moisture information. As shown in Figure 3d,e, the distribution of CALDAS assimilated soil moisture was comparable and correlated reasonably well with the products of JAXA (Pearson product-moment correlation coefficient R = 0.6) and NASA (R = 0.5). It is also worth noting that the absolute values of soil moisture obtained from the JAXA and NASA products differed greatly, even though they both used the same satellite information. A previous study [26] on the validation of the JAXA and NASA satellite-derived soil moisture products reported significant differences between them and that neither provided reliable estimates for all the conditions represented by the four watershed sites in the United States. The significant differences between both products in our domain also indicated that further coordinated validation studies are necessary to understand and resolve the problems associated with the retrieval algorithms for African regions. However, as CALDAS performed better than ARPS and spatially, its soil moisture correlated well with the satellite-derived soil moisture products, the improved soil moisture content from CALDAS (using the lower frequencies) was used to calculate land surface emission at higher frequencies to facilitate the integration of cloud microphysics data by assimilating AMSR-E higher-frequency observations over land surfaces.

4.2. Comparisons of Cloud Condensate With Satellite IR Observations

The quantitative validation of individual or integrated atmospheric moisture variables (*i.e.*, cloud liquid water, ice, snow, and hail) is still difficult, even at point locations, because of the unavailability of the necessary data sets. Therefore, to investigate the retrieval capability of the land and atmosphere coupled data assimilation processes and to evaluate the reliability of the assimilated cloud related parameters, qualitative comparisons were performed of the spatial distributions between vertically integrated condensate (summation of liquid and solid phases) and IR satellite data. The results were investigated ∼1 h after the assimilation (*i.e.*, from 0300 UTC on 6 June 2006).

As shown in the first column of Figure 4, the IR observation reflected very active cloud activity over the model domain, which was successfully introduced by the assimilation of passive microwave observations in CALDAS. The assimilated cloud distributions from CALDAS were coherent and compared well with the observed IR cloud cell distributions. Conversely, the ARPS and LDAS-A models predicted almost no cloud activity at this time. In addition, hourly evolutions of cloud activity (each column) showed that the seeding of cloud information using the 89-GHz observation within the model enhanced the forecasting capability of the model. Thus, the clouds predicted by CALDAS correlated well with IR observations 6 h after the assimilation, whereas ARPS and LDAS-A showed much lower cloud activity during this period. The identical results of cloud condensate obtained from ARPS and LDAS-A during the model forecasting showed that the assimilation of soil moisture did not improve the cloud simulation, and therefore, the reliable initialization of cloud information is shown to be crucial for improving weather forecasting during this particular period.

Figure 4. Hourly variation of integrated condensate (kg/m^2) and IR data (K); first row: ARPS, second row: Land Data Assimilation System coupled with Atmospheric model (LDAS-A), third row: CALDAS, and fourth row: infrared (IR) brightness temperature.
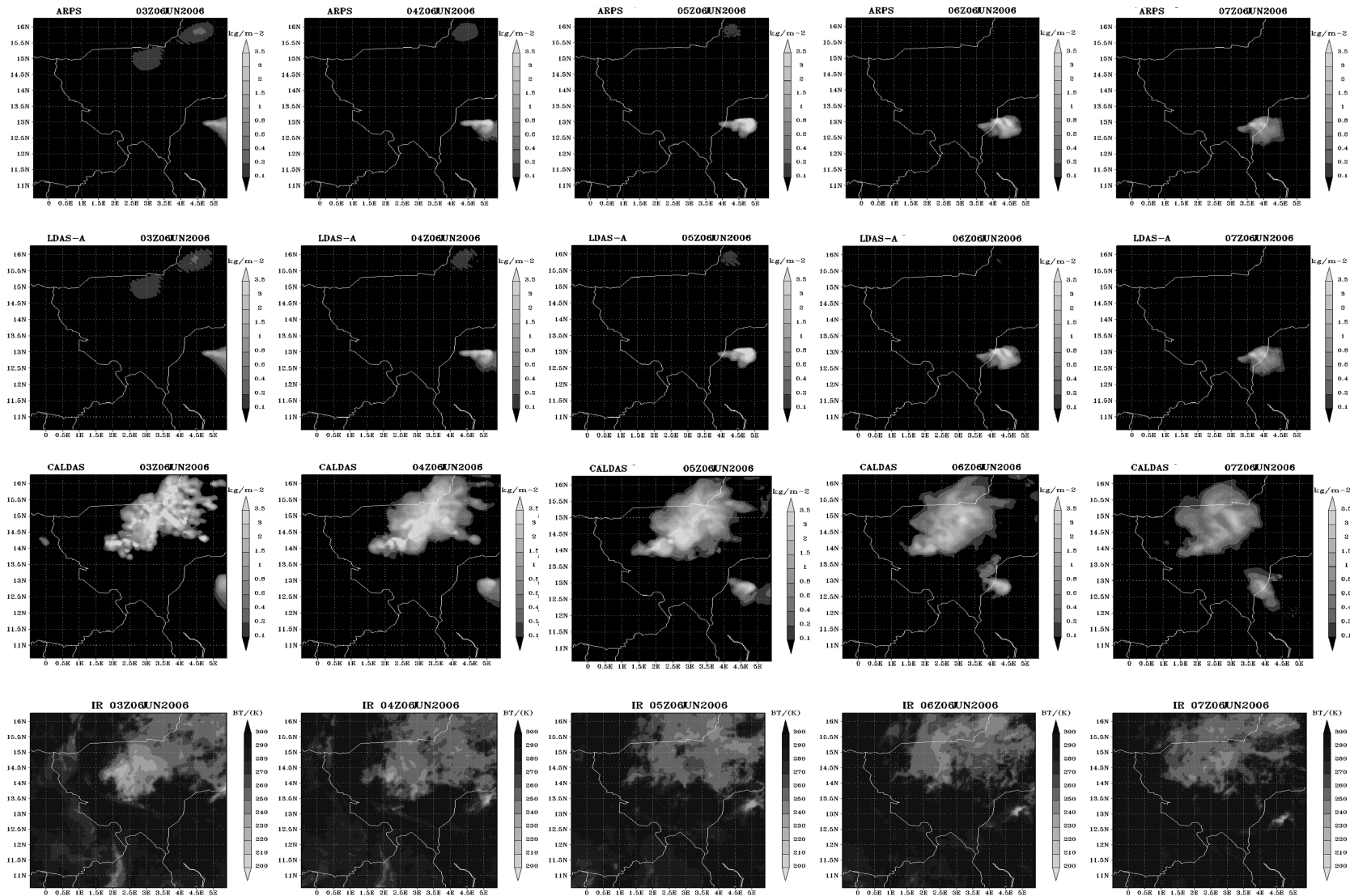
Figure 5 depicts the diurnal variation in spatial correlations calculated between cloud condensate simulated by the models and IR observations for 6 June 2006. To eliminate the shift between the AMSR-E and IR observations, the IR data and the model's cloud condensate were scaled up to 0.25°. Negative values in the spatial correlation indicate mismatches between the IR data and model simulations of cloud positions, whereas positive values indicate coherence between the IR observations and model simulations. As shown, the spatial correlations derived from ARPS and LDAS-A were very similar and showed mostly negative or very low spatial correlations with the IR observations. However, the CALDAS results showed a rapid increase (+0.49) in the spatial correlation at 0300 UTC. Two hours after the assimilation, the correlation reached a maximum of +0.52. This result indicates that CALDAS was reinitialized with improved cloud distributions (which coincided well with the IR data) and that the assimilation of cloud condensate enhanced the simulated cloud activities over the model domain. Moreover, the correlation of CALDAS was larger than ARPS and LDAS-A for about 6 h from the assimilation before decreasing. Therefore, CALDAS, as a multi-frequency assimilation system, has the potential to improve cloud information over land surfaces. However, quantitative information on each atmospheric moisture variable and microphysical property must be investigated in the future, using reliable data sets, to assess the full capabilities of the CALDAS model.

Figure 5. Hourly variation of spatial correlations calculated from model simulated cloud top temperatures and IR cloud top temperatures for 6 June 2006.
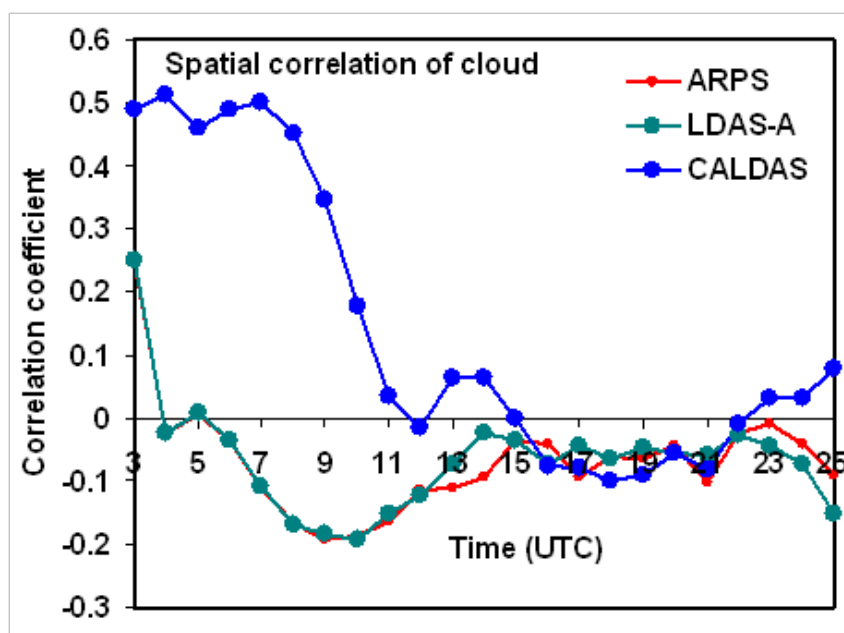
4.3. Evaluation of Land–Atmosphere Interactions

Improving the representations of land surface conditions and cloud will improve the simulations of surface (energy and water) budgets and turbulent heat fluxes, which will eventually improve the representation of land-atmosphere interactions in the model. Radiosonde observations of temperature and humidity in the planetary boundary layer, which reflect average turbulent fluxes over the observation domain, are commonly used to determine the degree to which the land affects the atmosphere, and vice

versa. Radiosonde soundings obtained at Niamey station were used to investigate the model performance in simulating the land-atmosphere feedback mechanism over the model domain.

Figure 6a,b compare the observed potential temperature and specific humidity soundings with the simulated soundings from ARPS, LDAS-A, and CALDAS at 1030 UTC (1130 Local Time (LT)) on 6 June 2006. During this time, the land was heated by the sun and turbulent mixing was created close to the land surface (from surface to ∼875 hPa), as shown by the observed temperature soundings. ARPS underestimated the potential temperature sounding (Mean Bias Error (MBE) = −0.52 K, Root Mean Square Error (RMSE) = 2.00 K) and overestimated the humidity sounding (MBE = 1.85 g/kg, RMSE = 2.92 g/kg) from the surface to ∼700 hPa. In the case of LDAS-A, the assimilation of surface soil moisture resulted in marginal improvements in both the temperature (MBE = −0.48 K, RMSE = 1.94 K) and the humidity (MBE = 1.82 g/kg, RMSE = 2.88 g/kg) profiles, whereas improvements were seen in CALDAS for the humidity profile (MBE = 1.32 g/kg, RMSE = 2.01 g/kg).

Figure 6. Comparison of observed and model soundings at 1030 UTC on 6 June 2006; (**a**) potential temperature (K) and (**b**) specific humidity (g/kg).
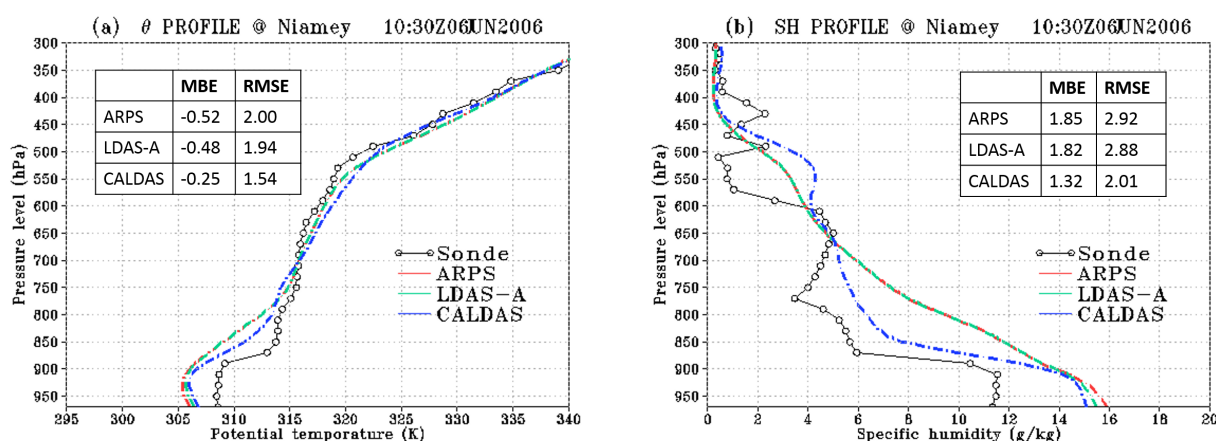

Figure 7 compares the observed soundings with the simulated soundings at 2230 UTC (2330 LT) on the same day. The observed atmospheric conditions at this (night) time, as shown by the temperature profile, were characterized by statically stable air with neutrally stratified layers. In the case of the ARPS and LDAS-A models, profiles with higher humidity and lower temperature were simulated. Further investigation of the model results indicated that these two models produced higher soil moisture during this period (not shown). However, the observed trend was captured well by CALDAS and it compared well with the observed soundings of potential temperature (MBE = 0.26 K, RMSE = 1.02 K) and specific humidity (MBE = 1.06 g/kg, RMSE = 1.31 g/kg). The improvement extended from the surface to the middle atmospheric layers (∼650 hPa) in the CALDAS simulation. These results show that the synchronized land and atmosphere re-initialization resulted in better prediction of the atmospheric sounding (*i.e.*, land-atmosphere interactions) during the forecast period of the model.

Figure 7. Same as Figure 6 but for 2230 UTC on 6 June 2006; (**a**) potential temperature (K) and (**b**) specific humidity (g/kg).
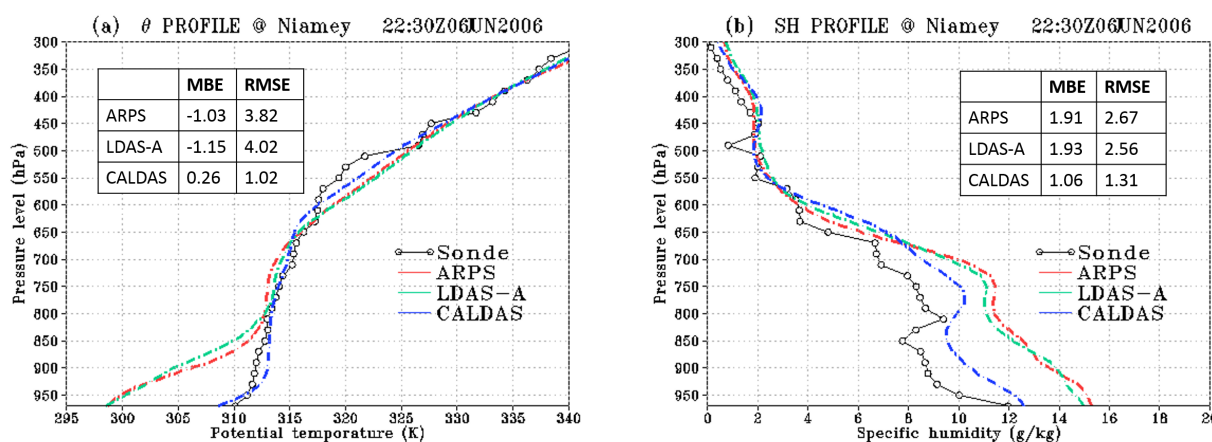

4.4. Evaluation oF Rainfall Forecast

Forecasting precipitation remains very challenging, and the skill and accuracy of quantitative forecasts of precipitation are generally low, even over well-instrumented regions. Figure 8 compares the 6-h (from 03 UTC–09 UTC) accumulated rainfall forecast by ARPS and CALDAS, and estimated by the TRMM satellite on 6 June 2006. As shown in Figure 8a, the location of the rainfall from ARPS was completely different from the location where rain was actually observed by TRMM (Figure 8c). In contrast, the assimilation of the 89-GHz information for cloud data produced rainfall at the same locations as observed by TRMM. Although the rainfall amounts predicted by CALDAS did not correspond completely with the TRMM retrievals, the locations of the heavy rainfall were better predicted by CALDAS. In the case of ARPS, the spatial correlation of rainfall with TRMM was very low ($R = 0.01$), whereas in the case of CALDAS, the spatial correlation of rainfall with TRMM was improved significantly ($R = 0.47$).

The quantification of model accuracy, errors, and uncertainties is crucial in determining the reliability of model forecasts, but it is a very challenging task without accurate observations in this study region. In this research, we utilized satellite retrievals (*i.e.*, soil moisture, cloud distribution, and rainfall) because of the unavailability of ground-based observations (except radiosonde data). However, these products are subjected to significant uncertainties due to bias of retrieval algorithms (e.g., soil moisture measurements obtained from JAXA and NASA differed greatly, even though data from the same sensor were used for the retrievals), few calibration and verification practices, low spatial resolutions, and the indirect nature of satellite measurements, which included uncertainties from various factors (e.g., estimation through cloud top reflectance, thermal radiance, and penetration depth of microwave signals). Consequently, our present discussions on validating model outputs with satellite retrievals are based mainly on statistical correlations, because the retrievals capture spatial distribution well, even though the absolute values differ among the similar products used. Conversely, further coordinated efforts are required to organize more intensive and comprehensive field observations to calibrate and assess the uncertainties of the retrievals, as well as to validate model outputs quantitatively.

Figure 8. Comparison of 6 hours (from 03UTC to 09UTC) accumulated rainfall (mm) obtained from (**a**) ARPS; (**b**) CALDAS; and (**c**) TRMM, respectively.
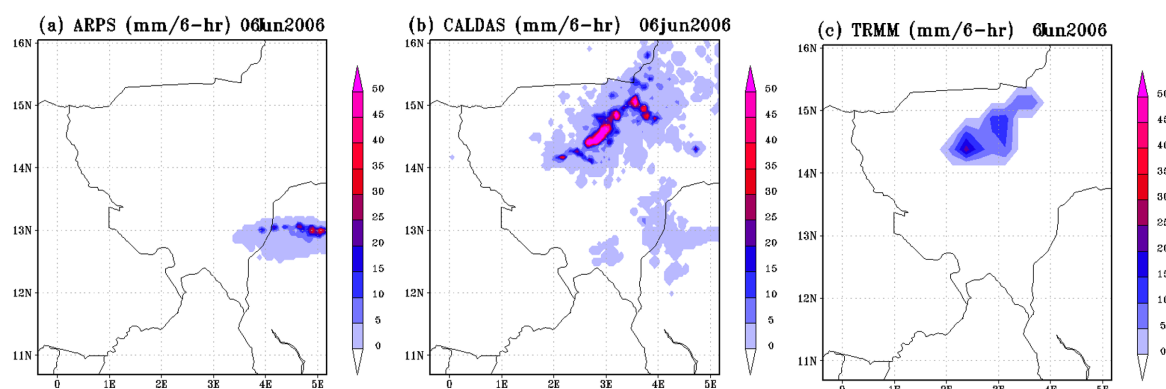
5. Conclusions

This research investigated the applicability of space-borne multi-frequency passive microwave observations for improving weather forecasting by retrieving soil moisture and cloud information using two data assimilation systems (*i.e.*, LDAS-A and CALDAS) over a mesoscale domain of Niger, Africa.

The results showed that the assimilated soil moisture was reasonably comparable with satellite retrievals obtained from NASA and JAXA, whereas the no-assimilation case produced higher soil moisture contents and completely different spatial distributions compared with the assimilated case. However, assimilating soil moisture alone produced only marginal improvements in the model forecasts of atmospheric profiles (*i.e.*, potential temperature and specific humidity).

The assimilated cloud distributions from CALDAS were coherent and they compared well with the observed IR cloud cell distributions. The seeding of cloud information from high-frequency passive microwave radiance enhanced the self-development of simulated cloud activities over the model domain. Consequently, simulated cloud activities exhibited positive correlations for a period of about 12 h following the assimilation, whereas the no-assimilation case demonstrated mostly negative correlations with the IR observations during this period.

The simulated soundings of both potential temperature and specific humidity from CALDAS showed significant improvements from the surface to the middle atmospheric layers (~650), which resulted from improvements in the representations of land and cloud conditions in the model. Furthermore, the rainfall predicted by CALDAS correlated better with the TRMM retrievals than the no-cloud assimilation cases. In particular, the locations of the heavy rainfall were better predicted by CALDAS, whereas the no-assimilation case produced rainfall in completely different locations compared with the TRMM rainfall retrievals.

These results are encouraging in terms of producing reliable land and atmospheric states as well as regional forecasts, especially in poorly gauged or un-gauged regions, because CALDAS requires only satellite data as input. The consideration of using other satellite information such as the Global Precipitation Measurement mission, AMSU, and AIRS within the system will be researched in future studies to improve the model's robustness and performance further.

Acknowledgments

Author Contributions

Mohamed Rasmy coupled the satellite data assimilation systems within a mesoscale model, performed numerical simulations, analyzed satellite retrievals and model results, and wrote the manuscript. Xin Li contributed significantly for developing the data assimilation components and provided constructive comments and suggestions for coupling the components within a mesoscale model. Toshio Koike supervised the research work, outlined the design of data assimilation systems, and assisted with the system developments as well as the manuscript writing.

Conflicts of Interest

References

1. Intergovernmental Panel on Climate Change. *Fourth Assessment Report: Climate Change 2007: The AR4 Synthesis Report*; Intergovernmental Panel on Climate Change: Geneva, Switherland, 2007.
2. Sohne, N.; Chaboureau, J.; Guichard, F. Verification of cloud cover forecast with satellite observation over West Africa. *Mon. Weath. Rev.* **2008**, *136*, 4421–4434.
3. Jackson, T.J.; Schmugge, T. Vegetation effects on the microwave emission of soils. *Remote Sens. Environ.* **1991**, *36*, 203–212.
4. Jackson, T.; le Vine, D.M.; Hsu, A.Y.; Oldak, A.; Starks, P.J.; Swift, C.T.; Isham, J.D.; Haken, M. Soil moisture mapping at regional scales using microwave radiometry: The Southern Great Plains Hydrology Experiment. *IEEE Trans. Geosci. Remote Sens.* **1997**, *37*, 2136–2151.
5. Paloscia, S.; Pampaloni, P. Microwave polarization index for monitoring vegetation growth. *IEEE Trans. Geosci. Remote Sens.* **1988**, *26*, 617–621.
6. Wentz, F.J.; Meissner, T. *Supplement 1 Algorithm Theoretical Basis Document: AMSR-E Ocean Algorithms*; NASA: Santa Rosa, CA, USA, 2007.
7. Koster, R.; Suarez, M.; Higgins, R.; van den Dool, H. Observational evidence that soil moisture variations affect precipitation. *Geophys. Res. Lett.* **2003**, *30*, doi:10.1029/2002GL016571.
8. Koster, R.; Suarez, M.; Liu, P.; Jambor, U.; Berg, A.; Kistler, M.; Reichle, R.; Rodell, M.; Famiglietti, J. Realistic initialization of land surface states: impacts on subseasonal forecast skill. *J. Hydrometeor.* **2004**, *5*, 1049–1063.

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

10. Draper, C.S.; Mahfouf J.-F.; Walker, J. P. An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme. *J. Geophys. Res.* **2009**, *114*, doi:10.1029/2008JD011650.

11. Rasmy, M.; Koike, T.; Boussetta, S.; Lu, H.; Li, X. Development of a satellite land data assimilation system coupled with a mesoscale model in the Tibetan Plateau. *IEEE Trans. Geosci. Remote Sensing* **2011**, *49*, 2847–2862.

12. Rasmy, M.; Koike, T.; Kuria, D.; Mirza, C.; Li, X.; Yang, K. Development of the coupled atmosphere and land data assimilation system (CALDAS) and its application over the Tibetan Plateau. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4227–4242.

13. Fujii, H.; Koike, T.; Imaoka, K. Improvement of the AMSR-E Algorithm for soil moisture estimation by introducing a fractional vegetation coverage dataset derived from MODIS data. *J. Remote Sens. Soc. Jpn.* **2009**, *29*, 282–292.

14. Jeu, R.; Wagner, W.; Holmes, T.; Dolman, A.; Giesen, N.; Friesen, J. Global soil moisture patterns observed by space borne microwave radiometers and scatterometers. *Surv. Geophys.* **2008**, *29*, 399–420.

15. Xue, M.; Droegemeier, K.K.; Wong, V.; Shapiro, A.; Brewster, K.; Carr, F.; Weber, D.; Liu, Y.; Wang, D. The Advanced Regional Prediction System (ARPS)— A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteorol. Atmos. Phys.* **2001**, *76*, 143–166.

16. Sellers, P.J.; Los, S.; Tucker, C.; Justice, C.; Dazlich, D.; Collatz, G. A revised land surface parameterization (SiB2) for atmospheric GCMs, Part II: The generation of global fields of terrestrial biophysical parameters from satellite data. *J. Clim.* **1996**, *9*, 706–737.

17. Evensen, G. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.* **2003**, *53*, 343–367.

18. Fujii, H.; Koike, T. Development of a TRMM/TMI algorithm for precipitation in the Tibetan Plateau by considering effects of land surface emissivity. *J. Meteorol. Soc. Jpn.* **2001**, *79*, 475–483.

19. Reichle, R.H.; McLaughlin, D.B. Hydrologic data assimilation with the ensemble Kalman filter. *Mon. Weath. Rev.* **2002**, *130*, 103–114.

20. Huang, C.; Li, X.; Lu, L. Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter. *Remote Sens. Environ.* **2008**, *112*, 889–900.

21. Mirza, C.R.; Koike, T.; Yang, K.; Graf, T. Retrieval of Atmospheric integrated water vapor and cloud liquid water content over the ocean from satellite data using the 1-D-Var ice cloud microphysics data assimilation system (IMDAS). *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 119–129.

22. Liu, G. A fast and accurate model for microwave radiance calculations. *J. Meteorol. Soc. Jpn.* **1998**, *76*, 335–343.

23. Duan, Q.; Sorooshian, S.; Gupta, V. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* **1992**, *28*, 1015–1031.

24. Lin, Y.; Farley, R.; Orville, H. Bulk parameterization of the snow field in a cloud model. *J. Clim. Appl. Meteorol.* **1983**, *22*, 1065–1089.

25. Kain, J.; Fritsch, J. Convective Parameterization for mesoscale models: The kain-fritsch scheme. *Meteoro. Monogr. Amer. Meteor. Soc.* **1993**, *24*, 165–170.

26. Jackson, T.J.; Cosh, M.H.; Bindlish, R.; Starks, P.J.; Bosch, D.D.; Seyfried, M.; Goodrich, D.C.; Moran, M.S.; Du, J. Validation of advanced microwave scanning radiometer soil moisture products. *IEEE Tran. Geosci. Remote Sens.* **2010**, *48*, 4256–4272.