Improvement of daytime land surface skin temperature over arid regions in the NCEP GFS model and its impact on satellite data assimilation

Weizhong Zheng,¹,² Helin Wei,¹,² Zhuo Wang,³ Xubin Zeng,³ Jesse Meng,¹,² Michael Ek,¹ Ken Mitchell,¹ and John Derber¹

Received 3 March 2011; revised 16 February 2012; accepted 17 February 2012; published 31 March 2012.

Comparison of the land surface skin temperature (LST) from the National Centers for Environmental Prediction (NCEP) operational Global Forecast System (GFS) against satellite and in situ data in summer 2007 indicates that the GFS has a large and cold bias in LST over the arid western continental United States (CONUS) during daytime. This LST bias contributes to large errors in simulated satellite brightness temperatures over land by the Community Radiative Transfer Model (CRTM) and hence the rejection of satellite data in the NCEP Gridpoint Statistical Interpolation (GSI) system, especially for surface-sensitive satellite channels. The new vegetation-dependent formulations of momentum and thermal roughness lengths are tested in the GFS. They substantially reduce the large cold bias of daytime LST over the arid western CONUS in the warm season. This, in turn, significantly reduces the large biases of calculated satellite brightness temperatures found for infrared and microwave sensors in window or near-window channels, so that many more satellite data can be assimilated in the GSI system. In the arid western CONUS, the calculation of surface emissivity for microwave sensors in the CRTM can be further improved, and the new microwave land emissivity model together with increased LST via changes in surface roughness length formulations reduces biases and root-mean-square errors in the calculated brightness temperature.


1. Introduction

In recent years, satellite measurements are being increasingly used in weather and climate prediction systems [Uppala et al., 2005; Marshall et al., 2007; Sakamoto and Christy, 2009]. Satellite data differ from many conventional data in that the measurements are often indirect observations of meteorological parameters. For example, satellite radiances are the observations which measure upwelling radiation at the top of atmosphere. At the National Centers for Environmental Prediction (NCEP), most satellite radiance measurements in various spectral channels are assimilated as radiances into the operational Global Data Assimilation System (GDAS) through the Gridpoint Statistical Interpolation (GSI) and the Joint Center for Satellite Data Assimilation (JCSDA) Community Radiative Transfer Model (CRTM) [Han et al., 2005]. A new three-dimensional variational data assimilation (3DVAR) analysis system, GSI, was implemented into the GDAS on 1 May 2007, replacing the Spectral Statistical Interpolation (SSI) 3DVAR system, which had been operational since 1991 [Parrish and Derber, 1992; Derber and Wu, 1998; Wu et al., 2002; Kleist et al., 2009]. The atmospheric analysis is generated every 6 h by the GSI with the NCEP operational Global Forecast System (GFS) previous forecast as the background. This analysis is then used as the initial conditions for GFS subsequent forecasts, and the cycle continues. Radiance data from many satellite sensors have been assimilated in the current operational GDAS such as the High Resolution Infrared Radiation Sounder (HIRS), Microwave Humidity Sounder (MHS), Advanced Microwave Sounding Unit-A (AMSU-A) and -B (AMSU-B), the Atmospheric Infrared Sounder (AIRS) on the EOS-Aqua, Geostationary Operational Environmental Satellites (GOES) sounder and the Meteorological Operational satellite (METOP) Infrared Atmospheric Sounding Interferometer (IASI).

[3] Currently, satellite measurements over the ocean have been successfully utilized to improve numerical weather prediction (NWP) [Marshall et al., 2007]. However, it is noticed that the utilization of satellite data assimilated over land in the GSI is far less than over ocean, because of the difficulty in simulating land surface emissivity and temperature. Figure 1

---

¹NOAA/NCEP/EMC, Camp Springs, Maryland, USA.
²IMSG, Kensington, Maryland, USA.
³Institute of Atmospheric Physics, University of Arizona, Tucson, Arizona, USA.

Copyright 2012 by the American Geophysical Union.
shows a case on the spatial distribution of assimilated NOAA-18 AMSU-A channel 15 data from the NCEP radiance assimilation monitoring for the operational GDAS. Over the western continental United States (CONUS), less data is assimilated at 18:00 UTC than at 12:00 UTC. One of the chief reasons is that there is a much larger bias in the current NCEP operational GFS predicted land surface skin temperature (LST) over desert and arid regions during daytime in the warm season than in the early morning. In fact, LST predicted by the GFS is a critical factor in determining the simulated top-of-atmosphere brightness temperature ($T_b$) for satellite surface-sensitive channels in the GSI. With large bias in LST, the CRTM simulates unreasonable $T_b$, thus large amounts of satellite data are rejected in the GSI/CRTM analysis step, especially for surface-sensitive satellite channels.

[4] The NCEP operational GFS is a spectral model (current operational horizontal resolution is T574) with 64 vertical levels, defined using a hybrid sigma-pressure coordinate, with substantial upgrades in recent years (http://www.emc.ncep.noaa.gov/GFS). In particular, the Noah land surface model (LSM) (Version 2.7.1) replaced the Oregon State University (OSU) LSM, which had been operational in the GFS since the mid-1990s [Chen et al., 1996; Koren et al., 1999; Ek et al., 2003; Mitchell et al., 2005]. LST is derived from the surface energy budget. The Noah LSM has four soil layers (10, 30, 60, 100 cm thick), including updated treatments of frozen soil physics, infiltration and runoff, snowpack, canopy resistance, ground heat flux, soil thermal conductivity, direct surface evaporation and green vegetation cover. In terms of land surface characteristics, 9 soil texture classes [Zobler, 1986, 1999] and 13 vegetation types [Dorman and Sellers, 1989] are used. Green vegetation fraction (GVF) is obtained with the NESDIS 5 year (from April 1985 to March 1991 with year 1988 excluded) Normalized Difference Vegetation Index (NDVI) monthly climatology [Gutman and Ignatov, 1998]. Momentum roughness lengths over land are prescribed for each month based on calculations from the vegetation and land use data set of Dorman and Sellers [1989]. Monthly variation of snow-free surface albedo is derived in reference to Staylor and Wilbur [1990], and for a snow case, its albedo is calculated in the Noah LSM. Longwave emissivity is prescribed to be unity (blackbody emission) for all surfaces.

[5] Our investigation in the GFS testing has revealed that a major cause of the cold daytime LST bias is related to the treatment for roughness lengths, particularly thermal roughness length ($z_{0t}$) in the physics of surface turbulent heat transfer. The computation of LST in land models depends on the treatment of the ratio of the roughness length for momentum $z_{0m}$ over the roughness length for heat $z_{0t}$, the characterization of vegetation, and other model details [Yang et al., 2008]. For instance, an annual maximum GVF along with seasonally variable leaf area index (LAI) is used in the Community Land Model (CLM4) [Oleson et al., 2010], while seasonally varying GVF along with a constant LAI is used in the Noah land model [Mitchell et al., 2004]. Furthermore, in land models with prescribed annually maximum GVF (e.g., CLM4), bare soil and vegetated area are treated separately. In contrast, bare soil and vegetated area are treated together in Noah.

[6] In all these land models, an important physically based constraint should be the convergence of the turbulent fluxes and LST to bare soil values when the aboveground biomass approaches zero in a grid cell (e.g., when the leaf and stem area index becomes zero in CLM3 or when GVF becomes zero in Noah). In the earlier version of CLM, the convergence of the under-canopy turbulent exchange coefficient $C_s$ was not considered, which led to an excessive warm bias of

Figure 1. Spatial distribution of assimilated NOAA-18 AMSU-A channel 15 (89 GHz) data from the NCEP radiance assimilation monitoring for the operational GDAS at (a) 06:00 UTC, (b) 12:00 UTC, (c) 18:00 UTC, 1 July, and (d) 00:00 UTC, 2 July 2007.
around 10 K in monthly mean ground temperature over semiarid regions [Bonan et al., 2002]. This bias can be reduced by the use of a more appropriate GVF data [Barlage and Zeng, 2004]. To substantially reduce this bias, it is necessary to consider the convergence of $C_t$ to the bare soil formulation as aboveground biomass goes to zero [Zeng et al., 2005]. The convergence of canopy roughness length and displacement height to bare soil values as the aboveground biomass becomes zero has also been demonstrated to significantly improve the wintertime simulation of turbulent fluxes in CLM3 [Zeng and Wang, 2007]. In contrast to the 10 K warm bias of LST in the earlier version of CLM, Noah has a cold bias of around 10 K or more in the early afternoon of summer over semiarid regions (demonstrated in section 2). Previous efforts to reduce the LST bias in Noah has focused on the adjustment of the coefficient $C_{st}$ in the computation of the ratio of momentum roughness length to thermal roughness length ($z_{st}/z_{om}$) [e.g., Mitchell et al., 2004], varying from 0.2 in the earlier version of Noah, 0.1 in the version of Noah used in the Weather Research and Forecasting Model (WRF) [Chen and Zhang, 2009], to (implicitly) zero in the operational GFS at NCEP (that is, $z_{st} = z_{om}$). More recently, this issue was addressed from the context of land-atmosphere coupling strength by Chen and Zhang [2009]. The goal of this paper is to revise the computation of $z_{om}$ and $z_{st}/z_{om}$ in Noah, which represents a parallel effort with that of Chen and Zhang [2009], and to address the improvement of LST prediction and its effect on satellite data assimilation.

[7] In this study, the new formulations of momentum and thermal roughness lengths are tested to reduce the GFS warm season midday cold bias in LST. The impact of new roughness length changes on the $T_b$ calculation for satellite infrared and microwave sensors in the GSI is investigated, especially with regards to improving satellite data assimilation. The NOAA-17 HIRS-3 and NOAA-18 AMSU-A sensors are selected as the infrared and microwave sensors to focus on, respectively. These two sensors have been used in the NCEP GSI system for several years and their data quality and performance are well understood.

[8] The HIRS is an operational atmospheric sounding sensor which measures scene radiance with the 20 spectral channels, i.e., 12 longwave (6.7–15 μm), 7 shortwave infrared (3.7–4.6 μm) and 1 visible (0.69 μm) channel(s). The swath width is 2160 km with a 10 km resolution, from a spacecraft altitude of 837 km. The HIRSs have been carried on the NOAA Polar Operational Environmental Satellites (POES) for nearly 30 years, and their measurements have been extensively used for many weather and climate studies as well as forecasts [e.g., Kalnay et al., 1996; Andersson et al., 2005; Zapotocny et al., 2005, 2008; Uppala et al., 2005; Sakamoto and Christy, 2009].

[9] AMSU-A is a passive microwave sensor with 12 sounding channels and 3 window channels at 23.8, 31.4 and 89 GHz. Microwave observations in the window or near-window channels are strongly affected by surface characteristics and are very difficult to effectively use in NWP models because the assimilation requires both accurate surface temperature and emissivity calculations, which are more challenging over land than over ocean [Weng et al., 2001; Prigent et al., 2005; Karbou et al., 2006, 2007, 2010]. While the microwave land emissivity model [Weng et al., 2001] enables the GSI system to assimilate microwave measurements, there still remain outstanding problems in using satellite measurements over arid or bare soil regions such as western CONUS. Since both surface temperature and emissivity contribute to $T_b$, simulations, the significant errors in GFS surface temperature predictions over arid regions can produce large errors in $T_b$ simulations.

[10] In section 2, we present comparisons of LST over the CONUS from the NCEP operational GFS, satellite and ground observations. The new treatment of momentum and thermal roughness lengths and improvement of LST in the GFS are addressed by section 3. The impact study of surface skin temperature on satellite data assimilation in the GSI system is given in section 4. The summary is given in section 5.

2. Analysis of Land Skin Temperatures From GFS, GDAS, and GLDAS

[11] In this section, we compare LST over CONUS from GFS, GDAS and GLDAS (Global Land Data Assimilation System) to the GOES-derived satellite measurements and SURFRAD in situ data in summer 2007. The GFS LST is obtained from 12 h to 36 h forecasts of the previous day’s cycle at 12:00 UTC. In the GDAS data, we should mention that the GOES brightness temperature is already assimilated with the GSI.

[12] Developed jointly by NASA Goddard Space Flight Center (GSFC) and NOAA NCEP, GLDAS ingests satellite- and ground-based observational data products, using advanced land surface modeling and data assimilation techniques, to generate optimal fields of land surface states and fluxes for weather and climate predictions [Rodell et al., 2004]. The NCEP GLDAS data, covering 1979 to present, was generated on the same horizontal grid, land mask, terrain field, soil and vegetation types, seasonal cycle of green vegetation fraction and surface albedo as in the GFS [Meng et al., 2006]. The NCEP GLDAS is forced with the satellite-gauged merged (CMA) precipitation [Xie and Arkin, 1997] and other near-surface atmospheric variables from GDAS.

[13] The GOES-derived LST is produced with the split window technique through the partnership in GOES land surface retrievals between NESDIS and the University of Maryland [Pinker et al., 2003, 2009]. It provides gridded fields of hourly LST at 0.5° spatial resolution in cloud-free conditions throughout the year. The GOES-derived LST demonstrates a remarkable ability to match the station-observed mean diurnal cycle when verified against spring and summer flux station LST observations over the SGP ARM (Southern Great Plains/Atmospheric Radiation Measurement) network [Mitchell et al., 2004].

[14] Figure 2 shows the average clear-sky LSTs at 18:00 UTC for July 2007, which is midday in local time over the central United States. The GOES LST was retrieved only over clear-sky conditions, and the data were missing for 5 days in July 2007 (16, 18, 19, 26 and 27 July). Therefore, for a fair comparison, the average temperatures from GFS, GDAS and GLDAS include only the time slices when the GOES data are available at each grid cell. The horizontal resolution of GFS, GDAS and GLDAS data is 0.3125° by 0.3125° in longitude and latitude. The GOES-derived data...
have a horizontal resolution of 0.5° by 0.5° in longitude and latitude, and are bilinearly interpolated to the same resolution as other data sets. In the GDAS data assimilation system, skin temperature is the only surface field to be updated for a diagnostic purpose and the forecast model does not use it. We can see from Figure 2 that over the eastern CONUS, LSTs in GFS, GDAS and GLDAS are pretty close to the GOES-derived data. However, over the arid western CONUS, GFS, GDAS and GLDAS have a similar problem, showing much cooler LSTs compared to the GOES-derived data. The cold bias can reach up to 12°C in large areas.

We further use the Surface Radiation Budget Network (SURFRAD) data [Augustine et al., 2000] for LST verification. SURFRAD was established in 1993 through the support of NOAA’s Office of Global Programs (now Climate Program Office), in order to support climate research with accurate, continuous, long-term measurements pertaining to the surface radiation budget over the United States. Data are downloaded, quality controlled, and processed into daily files that are distributed in near real time (at http://www.srrb.noaa.gov). Currently seven SURFRAD stations are operating in climatologically diverse regions. There are three stations located in the western CONUS: Desert Rock, Nevada (with a bare soil surface type); Boulder, Colorado (grassland); and Fort Peck, Montana (also grassland). Figure 3 shows the averaged diurnal cycles of LSTs for July 2007. All data sets have hourly LST except for the GFS data set, which has only 18:00 UTC here, and the LST from all gridded data sets are bilinearly interpolated to SURFRAD sites.

We first assess GOES LST against the in situ LST observations. The GOES LST matches the station-observed mean diurnal cycle very well with biases of −2.0K, −3.6K, −0.4K and root-mean-square errors (RMSE) of 2.3K, 3.9K, 1.8K for Fort Peck, Boulder, and Desert Rock, respectively. Boulder and Fort Peck show an overall cool bias while Desert Rock shows a warm bias of 1–3°C that occurs before or after local noon. This kind of smaller cool biases could be related to undetected clouds or less prevalent cloud cover such as subpixel cumulus [Mitchell et al., 2004]. At the Desert Rock station, the GFS, GDAS and GLDAS show much larger cold bias of LST during daytime and still have about 5°C cold bias during nighttime. The largest cold biases appear at 20:00 UTC and can reach up to 15°C. Similarly, daytime LSTs at Boulder have relatively cold biases but not as large as those at Desert Rock, and at Fort Peck the cold biases are even smaller.

The other four SURFRAD stations are used for LST verification over the eastern part (not shown). The LSTs of GFS, GDAS and GLDAS are much closer to the observations in this region.

Figure 4 shows the actual time series of LSTs of GDAS, GOES and SURFRAD in July 2007 for three SURFRAD stations over the western CONUS and one (Penn State, PA) over the eastern CONUS. The GDAS results agree well with the observations at the Penn State station during the whole month, but show cold biases during daytime at the other three stations. In particular, the substantial daytime cold biases consistently exist during the whole month at the Desert Rock station. Nighttime cold biases also exist but are much smaller in magnitude than the daytime biases.

In summary, the LST from the NCEP operational GFS has a smaller bias over eastern CONUS but a substantial cold bias over western CONUS during daytime, especially in arid or bare soil regions. The notably strong cold bias of LST is also prevalent in other regions of the world, yet the impact of our new formulations on this issue was not directly quantified in this study. Figure 5 depicts the vegetation category and GVF in the GFS model on 1 July. Broadleaf or needleleaf evergreen/deciduous trees or cultivations with high GVF values dominate the eastern CONUS, while the western part is covered by grassland, broadleaf shrubs or even bare soil with small GVF values. Comparison between Figures 2 and 5b shows that the cold bias is the
largest over arid regions. The cold bias is not due to vegetation itself but strongly related to GVF, especially for bare soil as GVF approaches zero.

3. New Treatment of Momentum and Thermal Roughness Lengths and Improvement of LST in GFS

In the current NCEP GFS thermal roughness length \( z_{0t} \) is taken as the same value as momentum roughness length \( z_{0m} \). This is usually not true, especially over arid or semiarid regions, where GVF is low (as shown in Figure 5b) and \( z_{0t} \) is typically much smaller than \( z_{0m} \). Thus, the current large \( z_{0t} \) in the GFS contributes to large aerodynamic conductance \( C_h \) and improper excess surface heat flux between the land and the atmosphere for a given surface-air temperature difference [Mitchell et al., 2004]. For a given surface net radiation flux over semiarid regions, it is primarily balanced by the sensible heat flux, and the increased \( C_h \) requests the decrease of LST. We propose a new formulation on \( \ln(z'_{0m}/z_{0t}) \) as follows:

\[
\ln(z'_{0m}/z_{0t}) = (1 - \text{GVF})^2 C_{zil} (u^* z_{0g}/\nu)^{0.5}
\]

where \( z_{0m} \) is the effective momentum roughness length computed in (2) for each grid, \( z_{0t} \) is the roughness length for heat, \( C_{zil} \) is a coefficient (taken as 0.8), \( \nu \) is the Von Karman constant (0.4), \( \nu = 1.5 \times 10^{-5} \text{ m}^2 \text{ s}^{-1} \) is the molecular viscosity, \( u^* \) is the friction velocity, and \( z_{0g} \) is the bare soil roughness length for momentum (taken as 0.01 m).

In order to consider the convergence of \( z_{0m} \) between fully vegetated and bare soil, the effective \( z_{0m} \) is computed:

\[
\ln(z_{0m}) = (1 - \text{GVF})^2 \ln(z_{0g}) + [1 - (1 - \text{GVF})^2] \ln(z_{0w})
\]

To test the new roughness length formulations, the NCEP operational GFS model at the resolution of T382L64 is used for these experiments. A 3 day window (1–3 July 2007) is
chosen for a case study in Figure 4, as most part of the western CONUS was clear during this period. Furthermore, the GFS LST cold biases over arid regions are persistent from day to day (e.g., as shown in Figure 4 at the Desert Rock station). Two 72 h forecasts, a control and sensitivity run, are made from the 00:00 UTC, 1 July 2007 analysis. For the control run, \( z_0^{t} \) is taken as the same value as \( z_0^{m} \), while in the sensitivity run, \( z_0^{t} \) and \( z_0^{m} \) are computed from equations (1) and (2).

[22] Figures 6 and 7 show the 3 day average land surface skin temperature predicted by the GFS and comparison with the GOES-derived satellite measurements and SURFRAD in situ data. At 18:00 UTC (midday in local time over the central United States) in the control run (Figure 6a), a substantial cold bias can be found over the west half of CONUS (i.e., arid or semiarid regions). The new roughness length formulations significantly reduce the cold bias in the western CONUS, while the LST in the eastern CONUS, where the bias is small in the control run, is not much affected. On the other hand, the new formulations cause some positive biases in the southern United States, which might be related to the NESDIS 5 year monthly GVF climatology [Gutman and Ignatov, 1998] used in the GFS. The GVF climatology cannot capture real-time vegetation status [Jiang et al., 2010]. The new real-time weekly GVF products developed by Jiang et al. [2010] are more suitable for use in operational numerical models including the GFS and could further improve the GFS LST in the future. The Desert Rock station (36.63°N, 116.02°W), located in the Nevada desert, is one of the observation stations within the SURFRAD network. The GOES-derived LST is very similar to the SURFRAD observations (Figure 7a). However, the GFS control run produces very low LST, compared to the observations, especially during daytime. During midday, the cold bias can reach up to \(-15^\circ\text{C}\). The sensitivity run produces a reasonable LST diurnal cycle. The daytime LST increases substantially and is very close to the observations. In Figure 7b, the aerodynamic conductance in the control run is too large during daytime, and this is significantly reduced using the new roughness length formulations in the sensitivity run. Note

![Figure 4.](image-url)
that the nighttime LST cold bias is only slightly improved by our new formulations (Figure 7a), as the aerodynamic conductance is not significantly affected at night (Figure 7b). Further work is still needed to resolve this issue, particularly from the perspective of turbulent mixing under stable conditions and soil heat transfer at night.

Figure 5. (a) Vegetation types and (b) green vegetation fraction on 1 July in the GFS model. Vegetation types 1 to 13 in Figure 5a denote broadleaf evergreen trees, broad deciduous trees, broadleaf and needleleaf trees, needleleaf evergreen trees, needleleaf deciduous trees, broadleaf trees with ground cover, groundcover only, broadleaf shrubs with groundcover, broadleaf shrubs with bare soil, dwarf trees and shrubs with groundcover, bare soil, cultivations, and glacial.

Figure 6. Comparison of LST simulated in GFS and verification with the observations. (a) Difference between GFS and GOES in the control run at 18:00 UTC; (b) difference between GFS and GOES in the sensitivity run at 18:00 UTC, averaged from 1 to 3 July 2007.

Figure 7. Average diurnal cycle on 1–3 July 2007 for (a) verification of LST with GOES (black) and SURFRAD (blue) at Desert Rock, Nevada; red and green lines are for the control and sensitivity runs, respectively; (b) aerodynamic conductance \(C_h\) (no \(C_h\) observed or inferred from observations); and (c) sensible heat flux (SH), latent heat flux (LH), net radiation flux (Rnet) and ground heat flux (G).
4. LST Impacts on Satellite Data Assimilation

[24] The NCEP GFS forecast needs fields from the land surface such as upward longwave radiation flux and surface heat fluxes, which largely depend on LST predicted by the model. Moreover, the GFS needs to provide atmospheric profiles and surface fields to the GSI data assimilation system. The GSI uses these GFS forecast fields and uses the CRTM to simulate brightness temperatures for both infrared and microwave sensors. The forecast fields, also referred to as the first guess, are adjusted based on the analysis increment which is derived from an error weighted difference of observed and simulated brightness temperatures [Wu et al., 2002; Derber and Wu, 1998]. A bias correction is applied to the simulated $T_b$.

[25] The calculation of $T_b$ not only requires accurate model input including atmospheric profiles and surface skin temperatures, but also an accurate specification of the surface emissivity. For infrared frequencies, the CRTM computes surface emissivity via the IRSS model for ocean [Van Delst, 2003; Wu and Smith, 1997], and over land, a look-up table of 24 surface types. The 13 GFS vegetation types are matched to these 24 surface types in the CRTM. The database of surface emissivity and reflectivity at infrared frequencies over land are based on surface measured reflectance for these surface types [Vogel et al., 2011].

[26] Microwave emissivity can be obtained according to surface types such as land, snow, sea ice and ocean. A microwave land emissivity model [Weng et al., 2001] is employed using a two-stream radiative transfer approximation that quantifies the land emissivity over various surface conditions such as snow cover, deserts and vegetation. The emissivity model uses a three-layer medium and includes volumetric scattering, emission and interface reflection. It derives the reflection and transmission at the surface-air interface and lower boundary by modifying the Fresnel equations to account for cross-polarization and surface roughness effects. At spectra frequencies between 4.9 and 94 GHz, the emissivity simulated by this model has shown a good agreement with the ground-based radiometer measurements for bare soil, grass land, and snow conditions. However, large discrepancies were found over high latitudes and desert regions [Weng et al., 2001].

[27] In order to examine the impact of LST improvements in the GFS on data assimilation, we choose two sensors, the NOAA-17 HIRS-3 and the NOAA-18 AMSU-A. With these two sensors, we show how the LST improvement increases utilization of infrared (IR) and microwave (MW) satellite measurements in the GSI data assimilation system, respectively.

4.1. NOAA-17 HIRS-3

[28] NOAA-17 HIRS-3 channel 8 (centered at 11.11 microns) is considered a window channel that is very sensitive to surface characteristics. From the NCEP operational radiance monitoring statistics, such window channels over arid or semiarid regions are difficult to be used in the current operational GSI data assimilation system because of large biases in calculated $T_b$. Therefore, in this study we specifically focus on HIRS-3 channel 8. Two GSI analyses were generated valid at 18:00 UTC, 1 July 2007. The control analysis used forecast fields from the GFS control run. The second analysis used forecast fields from the GFS using the new roughness lengths. All satellite measurements and other input data in the GSI are the same in both experiments.

[29] Quality control in the GSI rejects the satellite pixels with relatively large brightness temperature biases and cloud conditions. Figures 8a and 8b show the spatial distribution of satellite pixels used in the GSI. In the control run, most of the satellite data are excluded over the western CONUS according to the rejection criteria. With the improvement in LST, more satellite data are included in the data assimilation system in this region (Figure 8b).

[30] To further illustrate the improvement in $T_b$ simulation, we focus on the western CONUS area within the red box. We calculate $T_b$ biases, RMSE and normalized frequency distribution in 1K bins for all clear-sky pixels within the red box. Figures 8c and 8d give these histograms for land and various land surface categories in typically arid to semiarid regions, like scrub soil (broadleaf shrubs with bare soil in the GFS), scrub (groundcover or broadleaf shrubs with perennial groundcover in the GFS) and bare soil. The frequency distributions for these surface types are strongly skewed to the left (negative bias) in the control run. This is especially true for scrub soil and scrub which contribute large errors to the entire land area since these two categories dominate the western CONUS. These large errors are reduced in the sensitivity experiment (Figure 8d) for all categories. The bias over the land is reduced to $-1.8$K from $-6.0$K and the RMSE is reduced to $3.9$K from $7.7$K with respect to the control run.

[31] In order to quantitatively evaluate the impact of LST on data assimilation, Table 1 lists the percentage of usable data within the red box. After passing the quality control and excluding cloud conditions, only very small percentage of HIRS-3 data are actually utilized in the GSI over the entire west CONUS land. Compared to the control run, the sensitivity run more than doubles the usable data percentage. Over soil or soil regions, the improvement is even more dramatic.

4.2. NOAA-18 AMSU-A

[32] Using NOAA-18 AMSU-A, the brightness temperature at the top of atmosphere for a microwave channel can be expressed as

$$T_b = LST \cdot e^{-\tau(0,H)\mu} + T_{\text{atm}}^1(1 - e) \cdot e^{-\tau(0,H)\mu} + T_{\text{atm}}^0,$$

where $e$ is the surface emissivity, $\tau$ is the atmospheric optical thickness, $\mu$ is the cosine of incident zenith angle, $H$ is the top-of-atmosphere height, and $T_{\text{atm}}^1$ and $T_{\text{atm}}^0$ are the atmospheric downwelling and upwelling brightness temperatures, respectively. For surface-sensitive channels (also...
called window channels), atmospheric absorption is weak, and the second and the third terms on the right side of equation (3) are small so that the $T_b$ calculation largely depends on the surface term (the first term on the right side of equation (3)). Thus, LST and surface emissivity are key factors in determining microwave $T_b$ for window channels. English [1999, 2008] used the simplified error analysis to estimate the contribution of typical LST and emissivity errors to $T_b$ calculation. For the AMSU-A sensor, at a low surface-to-space transmittance ($t$), brightness temperature error shows a significant sensitivity to LST but little sensitivity to emissivity. When transmission is high, both LST and emissivity become important.

Figure 9 plots data for NOAA AMSU-A channel 15. The impact of the new roughness length formulations in the GFS on utilization of AMSU-A data in the GSI can be seen from Figures 9a and 9b. Over the western CONUS, many more satellite pixels are used for data assimilation system with a corresponding LST improvement. However, in some regions such as southern California and southern Nevada, as well as the western part of Mexico where it is mostly covered by bare soil or broadleaf shrubs with bare soil, the new roughness length scheme results in more satellite data rejected in the GSI. Further investigation of the surface

<table>
<thead>
<tr>
<th>Surface Category</th>
<th>HIRS-3</th>
<th>AMSU-A</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTL (%)</td>
<td>EXP (%)</td>
<td>CTL (%)</td>
</tr>
<tr>
<td>Land 4.96</td>
<td>11.15</td>
<td>26.36</td>
</tr>
<tr>
<td>Scrub 3.87</td>
<td>10.55</td>
<td>18.27</td>
</tr>
<tr>
<td>Scrub soil 4.30</td>
<td>14.84</td>
<td>10.32</td>
</tr>
<tr>
<td>Bare soil 23.35</td>
<td>25.75</td>
<td>72.59</td>
</tr>
</tbody>
</table>

*Control run (CTL), sensitivity run with the new roughness lengths (EXP) and the second sensitivity run with both the new roughness lengths and the updated land emissivity model (EXP + ε).
The microwave land emissivity model was recently improved over bare soil and desert regions using the empirical emissivity algorithm by Yan and Weng [2009, 2011]. Over bare soil and desert regions, they first generated a desert microwave emissivity training database at window frequencies [34].

The emissivity calculation [Yan and Weng, 2009, 2011] indicates that the land emissivity model in the CRTM produces unreasonable high emissivity over these regions, particularly in window channels, and the mean error of desert emissivity is typically larger than 0.04 and is 0.05 at the AMSU-A frequencies.

**Figure 9.** Spatial distribution of satellite pixels used in GSI: brightness temperature bias of channel 15, NOAA-18 AMSU-A from (a) control run, (b) sensitivity run and (c) sensitivity run with new land emissivity model.

**Figure 10.** Frequency distribution of $T_b$ bias (channel 15) in 1K bins for all clear-sky pixels within the western CONUS (red boxes in Figure 9) from (a) control run, (b) new roughness length sensitivity run, and (c) new emissivity with new roughness length run. $T_b$ biases (first values in parentheses) and RMSE (second values in parentheses) for various land surface categories (as in Figure 8) are also given.
channels using the CRTM under clear-sky conditions. Then they derived fitting coefficients for the emissivity estimate at window channels from the training data set (with the improved GFS LST). At other frequencies, emissivity was interpolated according to a series of mean emissivity spectra along subdesert type. Finally, emissivity polarization was calculated using the existing physical model [Weng et al., 2001] as needed. With this updated land emissivity model, we performed the new roughness length sensitivity experiment again. It can be seen from Figure 9c that even more satellite pixels are accepted in the GSI over the western CONUS.

The statistical analyses indicate a substantial improvement in $T_b$ simulation with the new roughness lengths (Figure 10). In the control run, the frequency distribution for the scrub category is broad and skewed to the left, showing large cold bias (−3.8K) and large RMSE (6.0K). In the new roughness length experiment, the bias was changed from negative to positive, and both the bias and RMSE were reduced (2.4K and 5.6K, respectively) and were further improved with the new land emissivity model (1.6K and 4.7K, respectively) as shown in Figure 10c. Another category, scrub soil, also shows that bias was changed from negative (−1.4K) in the control run to positive but became larger (2.4K) in the new roughness length experiment. Over bare soil regions, however, the simulated $T_b$ shows warm bias (2.3K) in the control run, and even warmer bias and larger RMSE (4.4K and 6.9K, respectively) in the new roughness length experiment. The bias and RMSE got reduced with the new land emissivity model (1.8K and 3.7K, respectively).

Table 1 compares the percentages of assimilated AMSU-A data for these experiments. Over scrub regions, the improvement in the sensitivity run is substantial, but is minimal over scrub soil regions or even worse over bare soil regions. After the emissivity calculation is improved with the updated land emissivity model [Yan and Weng, 2009], more data are assimilated, particularly over scrub soil regions (EXP + $\varepsilon$ in Table 1), again indicating that both surface emissivity and LST need to be improved in order to realistically simulate brightness temperatures over bare soil and desert regions.

To test the robustness of our results, Figure 11 presents another example for improving the satellite data assimilation (at 23.8 and 89 GHz, respectively) with both the improved GFS LST and the updated land emissivity model over North Africa, where the satellite data are rarely assimilated in the current operational GSI. As expected, though...
the GFS LST has a cold bias, the original land surface emissivity model \cite{Weng et al., 2001} produces higher desert emissivity and results in large warm brightness temperature biases, and consequently causes a high volume of rejection of AMSU-A data over North Africa (Figures 11a and 11c). After the reduction of GFS’s cold bias, the warm brightness temperature biases become even larger (not shown). Along with the new emissivity model, the reduction of warm biases in brightness temperature and increase of assimilated satellite data can be seen from Figures 11b and 11d (compared with Figures 11a and 11c). This result again indicates that the improvement of both LST and surface emissivity is important in simulating brightness temperature and then in improving satellite data utilization and assimilation, especially over desert regions.

\cite{38} It is also noted that in this paper we limit the impact study of GFS LST in the operational data assimilation system GSI. For example, CRTM/GSI uses the microwave land emissivity model from \cite{Weng et al., 2001} to estimate the surface emissivity for the various surface conditions encountered over the global continents, except over desert regions where the newly developed land emissivity model \cite{Yan and Weng, 2009, 2011} is applied to avoid large emissivity errors from the former model \cite{Weng et al., 2001}. Alternative approaches are also available to improve surface emissivity estimates such as using global land surface emissivity maps \citep[e.g.,][]{Prigent et al., 2005; Karbou et al., 2006; Aires et al., 2010}. This should be worthy of future study on the impacts of various emissivity calculations together with the improved GFS LST.

\cite{39} The new roughness length scheme not only increases utilization of satellite data and reduces errors in brightness temperature simulation at window channels, but also improves the simulation of sounding channels. Figure 12 presents AMSU-A channel 4 (52.8 GHz) within the western CONUS as an example. Channel 4 is sensitive to the average air temperature in a deep layer from the surface to about 7 km of the troposphere. In the GFS, the new roughness length formulations reduce the substantial cold bias in LST over arid regions during daytime. At the same time, due to the correction in land-atmosphere turbulent heat exchange with the new roughness length formulations, biases in atmospheric profiles are also reduced, particularly in the lower troposphere. This issue will be further addressed in a future paper.

5. Summary

\cite{40} Satellite radiance data assimilation in various spectral channels suffers from large biases in LST predicted by the NCEP GFS. This paper compared LST over the CONUS from the NCEP GFS, GDAS, GLDAS and the GOES-derived satellite measurements and SURFRAD in situ data in summer 2007. It was found that the GFS, GDAS and GLDAS all have large cold biases in LST over the arid western CONUS in warm season during daytime, when compared against the GOES-derived or the surface observations from SURFRAD stations. The differences can reach up to 12K or more at 18:00 UTC (approximately midday in western CONUS). The large cold bias of LST results in large errors in the CRTM simulated satellite brightness temperatures over land and rejection of satellite data in the GSI for surface-sensitive satellite channels.

\cite{41} The new vegetation-dependent formulations of momentum and thermal roughness lengths were tested in the NCEP GFS model to reduce the substantial cold bias in LST over arid and semiarid regions during daytime in the warm seasons. This case study has shown that with an increased daytime LST over western CONUS, there is an obvious reduction in the large cold biases of calculated brightness temperatures found for infrared and microwave satellite sensors in window or near-window channels, so that many more satellite measurements can be used in the GSI data assimilation system.

\cite{42} Brightness temperature calculation is affected by both LST and land surface emissivity in the CRTM. Over desert or bare soil regions, unreasonable surface emissivity for microwave sensors in the CRTM was corrected, and the new emissivity model together with increased LST via changes in surface roughness length formulations gave smaller biases and RMSE in the calculated brightness temperature. With the results in this paper and additional extensive tests, the new formulations (1) and (2) were implemented in the operational GFS in May 2011.

\cite{43} Note that the GFS forecast model and GSI are the two main components within the complex NCEP global operational forecast and data assimilation system. In this system, a new analysis is generated every 6 h with the GSI to initialize a new global model GFS forecast. This forecast, in turn, provides the background for the next GSI analysis \citep{Kleist et al., 2009}. In this paper, we only generated static GSI analyses by GSI with no feedback of the analyses onto the subsequent GFS forecast. Results from fully cycled GFS/GSI experiments will be reported in our separate paper.

\cite{44} Acknowledgments. The authors would like to thank our many collaborators at EMCCJSDA and NASA Hydrological Science Branch for their useful comments and suggestions, in particular, Christa D. Peters-Lidard for the land surface model (Noah); Russ Treadon and George Gayno for the GSI tests and for detailed comments and constructive suggestions on the manuscript; and Paul van Delst, Fuzhong Weng, Ronald Vogel, Banghua Yan, Yong Han and Mark Liu for the JSCDA CRTM. The work of X. Zeng and Z. Wang was funded by NOAA grant NA07NES4400002. Several anonymous reviewers are thanked for helpful comments.


J. Derber, M. Ek, J. Meng, K. Mitchell, H. Wei, and W. Zheng, NOAA/NCEP/EMC, WWBG, Rm. 207, 5200 Auth Rd., Camp Springs, MD 20746, USA. (Weizhong.Zheng@noaa.gov)

Z. Wang and X. Zeng, Institute of Atmospheric Physics, University of Arizona, 1118 E. 4th St., Tucson, AZ 85721-0081, USA.