Improved Quantitative Precipitation Forecasts by MHS Radiance Data Assimilation with a Newly Added Cloud Detection Algorithm

XIAOLEI ZOU
Department of Earth, Ocean and Atmospheric Sciences, The Florida State University, Tallahassee, Florida

ZHENGKUN QIN
Center of Data Assimilation for Research and Application, Nanjing University of Information Science and Technology, Nanjing, China, and Department of Earth, Ocean and Atmospheric Sciences, The Florida State University, Tallahassee, Florida

FUZHONG WENG
National Environmental Satellite, Data, and Information Service, National Oceanic and Atmospheric Administration, Washington, D.C.

(Manuscript received 4 January 2013, in final form 1 March 2013)

ABSTRACT
Satellite microwave humidity sounding data are assimilated through the gridpoint statistical interpolation (GSI) analysis system into the Advanced Research core of the Weather Research and Forecasting (WRF) model (ARW) for a coastal precipitation event. A detailed analysis shows that uses of Microwave Humidity Sounder (MHS) data from both NOAA-18 and MetOp-A results in GSI degraded precipitation threat scores in a 24-h model forecast. The root cause for this degradation is related to the MHS quality control algorithm, which is supposed to remove cloudy radiances. Currently, the GSI cloud detection is based on the brightness temperature differences between observations and the model background state at two MHS window channels. It is found that the GSI quality control algorithm fails to identify some MHS cloudy radiances in cloud edges where the ARW model has no cloud and the water vapor amount is low. A new MHS cloud detection algorithm is developed based on a statistical relationship between three MHS channels and the Geostationary Operational Environmental Satellite (GOES) imager channel at 10.7 μm. The 24-h quantitative precipitation forecast is improved rather than degraded by MHS radiance data assimilation when the new cloud detection algorithm is added to the GSI MHS quality control process. The temporal evolution of 3-h accumulative rainfall distributions compared favorably with that of multisensor NCEP observations and GOES-12 imager observations. The precipitation threat scores are increased by more than 50% after 3–6 h of model forecasts for 3-h rainfall thresholds exceeding 1.0 mm.

1. Introduction
Assimilation of the imager channel radiances from Geostationary Operational Environmental Satellites-11 and -12 (GOES-11 and GOES-12) resulted in a large positive impact on coastal quantitative precipitation forecasts (QPFs) near the northern Gulf of Mexico (Zou et al. 2011; Qin et al. 2013). In our earlier experiments, the National Centers for Environmental Prediction (NCEP) gridpoint statistical interpolation (GSI) scheme is used to produce an analysis field for the Advanced Research core of the Weather Research and Forecasting (WRF) model (ARW). Imager radiances from GOES-11/12, temperature sounding data from the Advance Microwave Sounding Unit-A (AMSU-A), moisture sounding data from the Microwave Humidity Sounder (MHS), as well as infrared radiances from the Atmospheric Infrared Sounder (AIRS) and High Resolution Infrared Sounder (HIRS), were individually added to the conventional data in GSI and all show some positive impacts. However, it was found that an all-data assimilation experiment did not produce better QPF scores than any single type of satellite data assimilation. The
degradation in QPF skill in the all-data assimilation experiment was mainly caused by adding MHS data into the assimilation. The present study further investigates the quality control of the MHS data in the GSI system, develops a new algorithm to detect the MHS cloudy radiances, and assesses the impacts of all satellite data assimilations on QPFs.

Satellite data assimilation is an active area of research. Vukicevic et al. (2006) carried out four-dimensional variational data assimilation (4DVAR) experiments at cloud-resolving scales and found beneficial impacts of GOES imager radiance observations at 10.7 and 12.0 μm to the spatial distribution of the ice cloud mass predicted by the Regional Atmospheric Modeling System (RAMS). Zupanski et al. (2011) investigated the impact of the synthetic advanced baseline imager (ABI) channel at 10.35 μm, which is sensitive to the hydrometeors (e.g., cloud ice and snow) at the cloud top. The assimilation of cloud- and precipitation-affected observations into weather forecasting systems started at the European Centre for Medium-Range Weather Forecasts (ECMWF; Bauer et al. 2006a,b). The simulated radiance is a weighted combination of cloudy- and clear-sky radiances. The cloudy radiance simulation involves only a single cloudy calculation (including any precipitation). Improvements in the forecasts of tropical moisture and wind fields by assimilation of Special Sensor Microwave Imager (SSM/I) data were shown in Kelly et al. (2008). Studies on the assimilation of cloud-affected or water-vapor-sensitive radiances can also be found in Singh et al. (2010, 2011), Stengel et al. (2010), and Otkin (2010, 2012a,b). Rain- and cloud-affected microwave radiances continue to be challenging. Large biases between simulated and observed brightness temperatures in cloudy and rainy areas are also major challenges (Geer et al. 2008; 2010).

In this study, a cloud detection algorithm is first developed and added as an additional step for the MHS quality control (QC) procedure to the GSI system. MHS data from the National Oceanic and Atmospheric Administration-18 (NOAA-18) and MetOp-A satellites with and without adding this new cloud detection algorithm are then assimilated using the NCEP GSI analysis code (Wu et al. 2002; Purser et al. 2003a,b). The impacts of MHS data assimilation on 24-h alongshore and offshore precipitation forecasts are finally assessed. Since the MHS data are thinned in space, the high spatial resolution information from microwave humidity sensors is probably underutilized in the present study. The assimilation cycle is now limited to no more than 2–3 times, with a 6-h cycling interval, to avoid problems arising from lateral boundary conditions for regional forecasts.

This paper is organized as follows: A brief overview of MHS channel characteristics, the NCEP GSI system, and the data assimilation experiment setup is provided in the following section. Section 3 describes a quality control procedure for MHS data used in the GSI system, compares the sensitivity of MHS brightness temperature differences between observations and model simulations (e.g., $O - B$) with that of GOES-12-observed and Global Forecast System (GFS)-modeled clouds, and provides a detailed description of a new cloud detection algorithm. The impacts of MHS data on QPFs with and without incorporating a new cloud detection algorithm are compared in section 4.

2. Data and model

a. MHS data

MHS is currently on board the NOAA-18, NOAA-19, and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) MetOp-A satellites. The MHS from all of the current operational satellites provided observations 4 times or more each day for moisture sounding in our precipitation case study (Qin et al. 2013). This instrument has a spatial resolution of 15 km near nadir and allows for an effective monitoring of atmospheric water vapor having high spatial variability. In the past decade, MHS data have long been incorporated into both operational and research data assimilation systems. For example, ECMWF assimilates satellite data from a wide range of instruments using a 4DVAR analysis system, including satellite humidity sounding data from microwave and infrared sensors (Theijs, 2003). Compared with infrared humidity sensors, microwave radiation can penetrate through nonprecipitating clouds and carries atmospheric humidity information within the clouds. However, observations from both precipitating and nonprecipitating clouds are not assimilated in the current GSI system.

NOAA-18 was launched on 20 May 2005 into an afternoon-configured (1400 LT) orbit with an altitude of 854 km above Earth. Since the launch of NOAA-18, MHS has become part of the Advanced Television and Infrared Observation Satellite (TIROS) Operational Sounder (ATOVs) package. MHS is also being flown on board the first European polar-orbiting satellite, MetOp-A, which was launched on 19 October 2006. MetOp-A is a morning-configured satellite (0930 LT). MetOp-A initialized a new era of international cooperation between NOAA and EUMETSAT.

MHS is a self-calibrating, cross-track-scanning, five-channel microwave radiometer, operating in a spectral
region from 89 to 190 GHz to provide information on atmospheric humidity at various altitudes, as well as data on atmospheric ice and temperature. The antenna beamwidth is a constant $1.11^\circ$. Ninety contiguous scene resolution cells are sampled along a single scan line in $\frac{\pi}{3}$ s. These scan patterns and their geometric resolution translate to a 16-km diameter cell at nadir. The MHS center frequencies for channels 1–5 are 89.0, 157.0, 183.31, 183.31, and 190.00 GHz, respectively. Figure 1 shows the weighting functions for the five MHS channels calculated by the Community Radiative Transfer Model (CRTM), which was developed by the Joint Center for Satellite Data Assimilation (JCSDA; Han et al. 2006; Weng 2007; Han et al. 2007; Chen et al. 2008, 2010). It is seen that MHS channels 3–5 profile the atmospheric water vapor in the troposphere. MHS channels 1 and 2 are near the atmospheric absorption window and are affected by the radiation from both the earth’s surface and emission and scattering from ice-phase clouds. MHS channel 5 has the highest central frequency and is most sensitive to scattering from thin clouds.

b. Data assimilation system and radiative transfer model

The GSI analysis system developed at NCEP is employed for this study. It is a three-dimensional variational data assimilation (3DVAR) system with recursive filters appropriately built into the analysis system (Wu et al. 2002; Purser et al. 2003a,b). The background error covariance matrix is in a gridpoint space and is anisotropic and nonhomogeneous.

The Community Radiative Transfer Model (CRTM) developed by JCSDA is used in the GSI system for rapid calculations of satellite radiances and gradients of radiance with respect to input variables. CRTM supports all sensors whose data are assimilated into GSI, covering microwave, infrared, and visible frequency regions (Weng 2007; Han et al. 2007).

c. Experiment setup

A convective precipitation case is investigated in this study. Two data assimilation experiments (EXP1 and EXP2) are carried out to assess the impact of MHS data assimilation with the modified MHS QC process when MHS data were assimilated together with conventional and AMSU-A observations. AMSU-A observations are included to provide much needed temperature sounding information over oceans for the selected convective case. Other satellite observations are excluded to keep the simplicity of this study. Differences between EXP1 and EXP2 were identified only in the amount of MHS data assimilated. Both experiments employ the same data assimilation system (GSI). The 6-h forecasts from NCEP Final Analysis (FNL) are used as the background fields $x_b$ at 1200 UTC 22 May 2008, observations within a 12-h time window are assimilated during two 6-h cycles of EXP1 and EXP2. The ARW 6-h forecasts are used as the background fields of the GSI 3DVAR at 1800 UTC 22 May and 0000 UTC 23 May 2008. The conventional observations are composed of a global set of surface and upper-air reports operationally collected by NCEP, including land surface, marine surface, radiosonde, and aircraft reports from the Global Telecommunications System (GTS), as well as profiler and radar-derived winds, SSM/I oceanic winds and atmospheric total column water (TCW) retrievals, and satellite wind data.

ARW, version V3.0, is selected as the forecast model. The horizontal resolution is 10 km. There are 27 vertical levels from the earth’s surface to the model top specified at 50 hPa. The grid size of the model domain is $250 \times 200 \times 27$. The WRF single-moment three-class microphysics scheme (Hong and Lim 2006), the Kain–Fritsch cumulus parameterization scheme (Kain and Fritsch 1990, 1993; Kain 2004), and the Yonsei planetary boundary layer scheme (Hong and Dudhia 2003) are selected for the ARW runs carried out in this study. The model domain is as large as that shown later (see Fig. 4a).
3. QC and cloud detection

a. GSI QC for MHS data

Although microwave data such as MHS can provide the moisture information in nonprecipitating cloudy conditions, cloudy radiances are currently rejected in the GSI system due to insufficient cloud fields from NWP models required by CRTM, the reduced accuracy of CRTM for simulating cloudy radiances, and a non-constant bias in the presence of clouds. Cloudy radiance assimilation is one of the most challenging problems in NWP and has been actively investigated in recent years (Bauer et al. 2006a,b; Geer et al. 2008; Kelly et al. 2008; Geer et al. 2010). This study focuses on clear-sky MHS radiance assimilation.

The quality control procedure thus aims at not only removing those MHS data that are in precipitation conditions with large cloud particles but also those data that could not be well simulated due to limitations in either CRTM or the ARW model or an inconsistency among the cloud information required by CRTM and provided by the NWP model. In the GSI system, QC for MHS data is based on both observed and modeled MHS data. Quality control including cloud detection for MHS data is done empirically in the GSI system, with a liquid water path index (LWPindex) first calculated over all surface types. Specifically, over ocean, ice, and snow surfaces, LWPindex is calculated as

\[
LWP_{\text{ocean index}} = \begin{cases} 
0.13 \times \left\{ (T_{b,1}^o - T_{b,1}^m) - 33.58 \times \left( \frac{T_{b,2}^o - T_{b,2}^m}{300 - T_{b,2}^o} \right) \right\}, & \text{if } T_{b,2}^o \leq 300 \\
9, & \text{otherwise}
\end{cases}
\]

where \(T_{b,1}^o\) and \(T_{b,2}^o\) are observations of the first and second MHS channels, respectively, and \(T_{b,1}^m\) and \(T_{b,2}^m\) are model simulations of the first and second MHS channels, respectively.

For other surface types, the index LWPindex is calculated as

\[
LWP_{\text{land index}} = 0.85 \times (T_{b,1}^o - T_{b,1}^m) - (T_{b,2}^o - T_{b,2}^m). \tag{2}
\]

If LWPindex is less than 0.0, it is set to 0.

Another index, called total-column precipitable water (TPWindex), is then defined as

\[
TPW_{\text{index}} = \left( (T_{b,1}^o - T_{b,1}^m) - 7.5 \times LWP_{\text{index}} \right)^2 + LWP_{\text{index}}^2. \tag{3}
\]

In the first QC step, MHS data from all five channels for the same field of view (FOV) is rejected from data assimilation in the GSI system if

\[
TPW_{\text{index}} > 1. \tag{4}
\]

The dependence of TPWindex on LWPindex and the difference of MHS channel 1 between observations \(O\) and background simulations \(B\) (see Fig. 2) shows an elliptic shape with its size defined by TPWindex. The TPWindex increases as LWPindex and \((O - B)_{ch1}\) increase. The first QC step is to eliminate the MHS data with TPWindex greater than one [see Eq. (4)]. Observations whose absolute value of \((O - B)_{ch1}\) is greater than 10 and whose LWPindex value is greater than 1 are rejected.

The second QC step removes those \(O - B\) outliers satisfying the following condition:

\[
|T_{b,1}^o - T_{b,1}^m| > 3\epsilon_i \quad \text{or} \quad |T_{b,2}^o - T_{b,2}^m| > 6 \text{ K}, \tag{5}
\]

where the subscript \(i\) represents the channel number from 1 to 5 and \(\epsilon_i\) the accuracy of observation errors adjusted as

\[
\epsilon_i = \epsilon_c \times (1 - TPW_{\text{index}}^2) \times f_H \times \tau_{i,\text{top}}, \tag{6}
\]
where $e_i$ ($i = 1, 2, \ldots, 5$) represents the input value of the accuracy of observation error for the $i$th channel (which is set to 2.5 K for channels 1–3 and 2.0 K for channels 4 and 5), $\tau_{\text{top}}^i$ is the transmittance at the model top for the $i$th channel, and $f_H = 2000/H$ if the terrain height $H$ is higher than 2000 m and $f_H = 1$ if $H$ is lower than 2000 m.

The third QC step rejects data from all five channels if the data from any of channels 2–5 is rejected in the second QC step.

Figure 3 provides an example showing observations eliminated by the above three QC steps at 1800 UTC 22 May 2008 within the model domain. Outliers removed by the first, second, and third QC steps are indicated by red, black, and green dots, respectively. Observations that passed the GSI QC test are indicated by blue dots.

where $e_i$ ($i = 1, 2, \ldots, 5$) represents the input value of the accuracy of observation error for the $i$th channel (which is set to 2.5 K for channels 1–3 and 2.0 K for channels 4 and 5), $\tau_{\text{top}}^i$ is the transmittance at the model top for the $i$th channel, and $f_H = 2000/H$ if the terrain height $H$ is higher than 2000 m and $f_H = 1$ if $H$ is lower than 2000 m.

The third QC step rejects data from all five channels if the data from any of channels 2–5 is rejected in the second QC step.

Figure 3 provides an example showing observations eliminated by the above three QC steps at 1800 UTC 22 May 2008. For clarity, only MHS data from NOAA-18 are shown. It is seen that outliers are removed quite effectively by the first, second, and third QC steps in the GSI system. Observations that passed the GSI QC test are characterized by small TPW$_\text{index}$ values (less than 1) and small model deviations from observations.

### b. A diagnosis of MHS QC results

Qin et al. (2013) found that when clear-sky radiance data from seven different infrared and microwave sensors on board the Polar Operational Environmental Satellite (POES) and GOES platforms were separately added to conventional data and assimilated into the GSI system, each instrument provided improvements to the coastal QPFs near the northern Gulf of Mexico. However, when all seven different types of satellite instruments were assimilated together, the rainfall forecast skill became significantly lower than any single-type satellite data assimilation. A simple test was performed in Qin et al. (2013) to show that an elimination of MHS data over areas where GOES imager channels detected clouds significantly improved the QPF scores from MHS data assimilation. Qin et al. also pointed to a possibility that not all cloudy radiances are removed in the GSI QC procedure for MHS data.

To see how effective the GSI QC process is in removing observations that are located within clouds, we show a spatial distribution of all observations that are removed (Fig. 4a) and retained (Fig. 4b) by the QC process at 1800 UTC 22 May 2008. The $O - B$ values of MHS channel 3 are indicated in Figs. 4a and 4b. Distributions of these observations are overlapped on the cloud-sensitive GOES channel 4 (10.7 $\mu$m) brightness temperature fields. It is seen that while some cloudy points are effectively removed (Fig. 4a) by the GSI QC procedure described in section 3a, there are still some cloudy points remaining, mostly around cloud edges after QC (Fig. 4b). Cloud effects on MHS data could be assessed by comparing them with collocated GOES channel 4 $O - B$ values since the latter is more sensitive to all types of clouds. Figures 4c and 4d present scatterplots of $O - B$ of MHS channel 3 against $O - B$ of GOES-12 imager channel 4 for those observations shown in Figs. 4a and 4b, respectively. Note that the MHS $O - B$ values for cloudy observations are not always more negative than those of points that passed the GSI QC test. In Fig. 4e, the geographical locations of observations exceeding two standard deviations of the GOES-12 imager channel 4 $O - B$ values from the mean are indicated in the same colors as in Fig. 4d. It is seen that those observations with differences from the mean greater than two standard deviations on the negative side (red dots) are located at and near the cloud edges, and observations with differences from the mean greater than two standard deviations are located over land. The former is probably related to differences in water vapor emission and cloud-scattering effects between the observations and model simulations. The latter is related to surface emissivity and terrain.

MHS radiances are sensitive to water vapor and clouds. To diagnose some residual cloudy pixels after GSI QC testing, we show in Fig. 5 the LWP and relative humidity at 300 hPa at 1800 UTC 22 May 2008 from

![Fig. 3. Scatterplots for TPW$_\text{index}$ and $O - B$ for MHS channels (top) 1 and (bottom) 2 from NOAA-18 at 1800 UTC 22 May 2008 within the model domain. Outliers removed by the first, second, and third QC steps are indicated by red, black, and green dots, respectively. Observations that passed the GSI QC test are indicated by blue dots.](image-url)
FIG. 4. Spatial distribution of all observations that (a) did not and (b) did pass the GSI QC test at 1800 UTC 22 May 2008 with $O - B$ values of MHS channel 3 indicated in color. (c) Scatterplots of $O - B$ of MHS channel 3 ($y$ axis) and $O - B$ of GOES-12 imager channel 4 ($x$ axis) for those data in (a). The $O - B$ values in (c) are indicated in the same color as in (a). (d) Scatterplots of $O - B$ of GOES-12 imager channel 4 for those data in (b). (e) Spatial distribution of all the observations shown in (d). Observations with their differences from the mean (solid curve) greater than two standard deviations on the negative side are indicated in red and observations with their differences from the mean greater than two standard deviations on the positive side are indicated in green. GOES channel 4 provides a background reference for cloud in (a), (b), and (e).
EXP1. The MHS cloudy pixels that are detected by the GSI QC procedure seem to be a combined result of the following two phenomena: (i) the modeled cloud coverage is smaller than the observations (Fig. 5a) and (ii) the sharpest gradient of the background specific humidity is located within the observed clouds instead at the cloud edges. It is possible that the modeled calculated water vapor emission is small (i.e., smaller B) due to reduced water vapor at the cloud edges, making the cloud detection based on the values of $O - B$ of MHS channels 1 and 2 (see section 3.1) insensitive to clouds.

c. A new cloud detection algorithm

The differences between the observations and the model simulation of GOES imager channel 4 at 10.7 μm are very sensitive to clouds, as shown in Fig. 4. On the other hand, the two MHS window channels, channel 1 at 89-GHz frequency and channel 2 at 150-GHz frequency, as well as the lowest MHS sounding channel 5 at 190.31-GHz frequency, are sensitive to ice clouds and precipitation, with channel 1 being most sensitive to total-column ice water and channel 5 being most sensitive to ice scattering from thin clouds. These three MHS channels are less sensitive to water vapor than are the other two MHS sounding channels (i.e., channel 3 at 183 ± 1 GHz and channel 4 at 183 ± 3 GHz). In addition, the two MHS window channels are affected greatly by the surface emissivity, which is much larger over land than over ocean. Therefore, two linear regression equations are first established between GOES channel 4 and the three MHS channels (e.g., channels 1, 2, and 5) over land and ocean separately. For all MHS data from NOAA-18

Fig. 5. (a) LWP (mm) and (b) 300-hPa relative humidity (contour, %) and brightness temperature observations from GOES channel 4 (shaded). (c) Cross section of the GOES channel 4 brightness temperature observations (red curve, also shaded), LWP (blue curve), and relative humidity (green curve) from EXP1 along 82°W [see the green line in (b)] at 1800 UTC 22 May 2008.
and MetOp-A during a 5-day period from 0000 UTC 17 May to 1800 UTC 21 May 2008, we obtain the following regression equations:

\[
(O - B)_{\text{GOES,land}}^\text{regression} = 0.009 \times T_{b,\text{MHS,ch1}}^{\text{obs}} + 0.085 \times T_{b,\text{MHS,ch2}}^{\text{obs}} + 0.877 \times T_{b,\text{MHS,ch5}}^{\text{obs}} - 274.255
\]  

(7)

and

\[
(O - B)_{\text{GOES,ocean}}^\text{regression} = -0.536 \times T_{b,\text{MHS,ch1}}^{\text{obs}} + 1.132 \times T_{b,\text{MHS,ch2}}^{\text{obs}} + 0.537 \times T_{b,\text{MHS,ch5}}^{\text{obs}} - 321.318,
\]  

(8)

where \(T_{b,\text{MHS,ch1}}^{\text{obs}}, T_{b,\text{MHS,ch2}}^{\text{obs}}, \) and \(T_{b,\text{MHS,ch5}}^{\text{obs}}\) represent brightness temperature observations for MHS channels 1, 2, and 5, respectively. In addition, \((O - B)_{\text{GOES,ocean}}^\text{regression}\) represent the differences between the observations and the model simulation of GOES imager channel 4.

Thresholds for cloud detection are determined empirically. This is illustrated in Fig. 6. First, differences between the observations and model simulations of GOES imager channel 4, \((O - B)^{\text{obs}}_{\text{GOES,land}}\) and \((O - B)^{\text{obs}}_{\text{GOES,ocean}}\), are arranged in decreasing order over land (Fig. 6a) and ocean (Fig. 6b). When the data are arranged in this order, the \(O - B\) differences in the brightness temperatures for GOES channel 4 over land (Fig. 6a) experience an initial decrease, then remain nearly constant, and finally decrease quite rapidly. Over the ocean, the \(O - B\) differences in brightness temperatures for GOES channel 4 (Fig. 6b) are nearly constant at the beginning of the order data series, and then decrease rapidly. The rapid decrease in the \(O - B\) differences of brightness temperatures in the latter half of the ordered data series on each day reflects the presence of clouds when the background state is clear. Although data from five different days are used, the monotonic curves of \((O - B)^{\text{obs}}_{\text{GOES,land}}\) and \((O - B)^{\text{obs}}_{\text{GOES,ocean}}\) have similar variations with their rapid decreasing lined along two nearly straight lines, defined by \(R_{\text{threshold,land}}^\text{land} = -4\)K and \(R_{\text{threshold,ocean}}^\text{ocean} = -2\)K. The thresholds for cloud detection, \(R_{\text{threshold,land}}^\text{land}\) and \(R_{\text{threshold,ocean}}^\text{ocean}\), are thus determined as these values. The similarity of the variations of \((O - B)^{\text{obs}}_{\text{GOES,land}}\) and \((O - B)^{\text{obs}}_{\text{GOES,ocean}}\) with data counts for five different days suggests the robustness of the proposed method for determining the threshold for cloud detection of the atmospheric conditions this case represents in the studying region. A longer time period and a larger range of atmospheric conditions with different cloud properties are needed to examine the validity of these thresholds as well as the proposed cloud detection algorithm.

Finally, cloudy points are determined by the following inequalities:

\[
(O - B)^{\text{regression}}_{\text{GOES,land}} \leq R_{\text{threshold,land}}^\text{land}\]  

and

\[
(O - B)^{\text{regression}}_{\text{GOES,ocean}} \leq R_{\text{threshold,ocean}}^\text{ocean}.
\]

(9)

It is emphasized that the above proposed cloud detection algorithm [Eqs. (7)–(9)] is based on the MHS data themselves, instead of using direct measurements of clouds from GOES imager sensors.

To see how effectively the above-proposed algorithm identifies cloudy points, the procedure is applied to GOES-12 channel 4. Figure 7 provides spatial distributions of all the MHS observations from NOAA-18.
that pass the GSI QC test at 1800 UTC 22 May and 1800 UTC 23 May. Cloudy radiance data over both land and ocean are identified by the cloud detection defined by Eq. (9) with its right-hand-side terms calculated by the regression Eqs. (7) and (8). It is noted that some clear-sky or thin cloudy radiance data are removed over the ocean.

Frequency distributions of MHS channel 1–5 brightness temperature observations over land and ocean with data with and without passing the new cloud detection during 1800 UTC 22 May and 1800 UTC 23 May are shown in Fig. 8. It is seen that the observations removed by the cloud detection algorithm are in general colder than those retained for all five MHS channels over land (Fig. 8). The mean brightness temperature for MHS channel 1 over ocean is more than 40 K colder than that over land, which is partially due to a much smaller surface emissivity over ocean than land near 89 GHz. Brightness temperatures removed by the cloud detection algorithm over land are also systematically colder than those retained for MHS channels 3–5 over ocean. The land–ocean contrasts in brightness temperature are not significantly different for tropospheric humidity sounding channels 2–5.

Figure 9 shows scatterplots of the $O - B$ of MHS channels 1–5 against the $O - B$ of GOES-12 imager channel 4 using all data from 1800 UTC 22 May to 1800 UTC 23 May 2008. It is found that nearly all observations characterized by large negative $(O - B)_{GOES-12}$ are eliminated by the new cloud detection algorithm added to the GSI MHS QC procedure. The $(O - B)_{MHS}$ values of outliers are in general negative and greater in magnitude than those that pass the cloud detection check, especially for MHS channels 1–3. Results in Fig. 9 confirm that the cloud detection algorithm proposed in this study effectively recognizes most cloudy points that the GSI QC fails to identify.

The frequency distribution of $O - B$ for MHS channels 1–5 over land and over ocean for observations that passed the GSI QC test and those that further passed an additional QC step with the new cloud detection are shown in Fig. 10. The frequency distributions of the differences between the model and the observations for those data to be assimilated are nearly Gaussian. A remaining small negative bias is noticed for MHS channel 2 over land.

4. Impacts on QPFs

The impacts of MHS radiance data assimilation with a new QC algorithm on QPFs are assessed by comparing forecasts from two pairs of data assimilation experiments described in section 2c. Scan biases calculated based on the error statistics of the innovations $H(x^b) - T_{obs}$, where $H(x^b)$ represents the CRTM simulations, are shown in Fig. 11a. Global biases are presented in Fig. 11b, in which scan biases are removed. Biases over ocean are positive for MHS channels 3–5. The magnitudes of bias are less than 0.5 K over ocean. Negative biases of about $-1.45$ and $-0.25$ K are found for MHS channels 1 and 2 over land, respectively. The differences in biases between land and ocean are associated with the differences in surface emissivity over land and ocean and their impact on the simulations of the two MHS window channels. For the three MHS sounding channels, the biases over land are comparable to those over ocean. These biases are subtracted in the MHS data assimilation.
FIG. 8. Frequency distributions of MHS channel 1–5 brightness temperature observations over land (solid) and ocean (dashed) that passed (blue) and did not pass (red) the new cloud detection algorithm for all data during 0000 UTC 17 May–1800 UTC 21 May 2008.
The impacts of MHS data assimilation with and without incorporating the proposed QC step are examined. Figure 12 presents the temporal patterns of the evolution of geopotential at 500 hPa from 0000 to 2400 UTC 23 May 2008 from EXP1 and EXP2, as well as the differences between these two forecast experiments. There is a subtropical high to the south and a low pressure trough to the north of the Gulf coast.
FIG. 10. Frequency distributions of $O - B$ for MHS channels 1–5 over land (solid) and over ocean (dashed) for all observations before (red) and after (blue) implementation of an additional QC step with the new cloud detection technique.
which is a typical synoptic weather pattern during the late spring and early summer as the land warms up from daytime heating. The winds near and within the northwest quadrant of the subtropical high bring warm and moist air from the Gulf of Mexico to the coastal regions, which collides with the cool dry air brought by the low pressure system over the land from the north, which leads to the development of thunderstorms. The movement and development of the low pressure trough system toward the southeast determines the subsequent southeastward movement of the clouds and the precipitation system. At 0600 UTC 23 May, a secondary trough developed to the east of the main trough in EXP2, the two combined into one at 0900 UTC, and further intensified until 1800 UTC 23 May 2008. The trough in EXP1 moved into the coastal area at 1200 UTC 23 May, and continued its further deepening while propagating farther to the southeast in EXP2. The trough in EXP1 did not move southeastward as far as that in EXP2 did.

Consistent with the temporal evolution of the upper-level trough, the 3-hourly accumulative precipitation predicted by EXP1 is confined over land (Fig. 13). The movement of the NCEP multisensor observed hourly precipitation (Fig. 14) and the cloud system in GOES imager channel 4 compares favorably with the temporal movement of the 3-hourly accumulative precipitation predicted by EXP2. Recall that the NCEP multisensor precipitation observations have limitations in their spatial coverage into the ocean since they are based on Next Generation Doppler Radar (NEXRAD) located over land. The rainfall patterns over the Gulf of Mexico oceanic regions correspond reasonably well to cloud distributions from the GOES imager channel.

Figures 15 and 16 display the conventional threat scores (TSs) and equitable threat score (ETS) of 3-h accumulative rainfall from EXP1 and EXP2 at four selected thresholds, respectively. The TS (Fig. 15) and ETS (Fig. 16) values are nearly the same, implying a low bias and a lower number of forecast “hits” due to random chance. EXP2 outperforms EXP1. The new MHS QC algorithm significantly improves QPFs when only conventional data and AMSU-A satellite data are assimilated. The largest improvements occurred for larger thresholds by incorporating the modified QC procedure for MHS radiance data assimilation.

5. Summary and conclusions

This paper examines the impact of adding a cloud detection algorithm to the MHS quality control procedure for data assimilation and forecasts of the limited-area ARW using the NCEP GSI analysis system. A pair of data assimilation experiments (EXP1 and EXP2) was carried out to produce model initial conditions for a coastal convective precipitation forecast case using MHS, AMSU-A, and conventional data 12 h prior to convective initiation. Experiment EXP1 assimilated only conventional data, from AMSU-A and the MHS dataset, while experiment EXP2 is the same as EXP1, but with a cloud detection algorithm added to the MHS QC procedure. Since the cloud detection scheme is used to remove cloudy MHS observations missed by the GSI system, it is important to remember that fewer observations are assimilated in EXP2 than in EXP1.

A careful analysis of the MHS QC procedure in the GSI system is first conducted. It was found that the GSI QC process fails to identify some cloudy radiance data near the cloud edges. The main reasons are that the model clouds and the water vapor distributions do not match in their spatial extent. Since infrared channels are more sensitive to clouds, a new cloud detection algorithm based on a linear regression between MHS and GOES imager channels 4 data is then developed. It was shown that adding this new cloud
FIG. 12. Temporal patterns of evolution of geopotential at 500 hPa from 0000 to 2400 UTC 23 May 2008 from EXP1 (dashed curve) and EXP2 (solid curve) and differences between EXP2 and EXP1 (EXP2 − EXP1, shaded).
FIG. 13. The 3-hourly accumulative precipitation (mm) from 0000 to 2400 UTC 23 May 2008 predicted by EXP1.
FIG. 14. The 3-hourly accumulative precipitation (mm) of (left) NCEP multisensor observations and (middle) EXP2, and (right) GOES imager channel 4 brightness temperature from 0000 to 2400 UTC 23 May 2008.
detection technique to the existing MHS QC process in the GSI system successfully removes most of the cloudy points. Compared to EXP1, the modified MHS data assimilation process made a significant difference in the forecasting of an upstream trough that moved from the northwest to the southeast and modulated the movement of the convective precipitation of the case studied.

The evaluation of the ARW quantitative precipitation forecast accuracy against multisensor 3-hourly rainfall and the 8-km high-resolution GOES imager channel radiance revealed very encouraging results for the case investigated in this study. The threat scores of the EXP2 precipitation forecast for all thresholds, especially for thresholds equal to or greater than 5 mm, significantly increased after 6 h into the model forecasts when the new cloud detection algorithm for the MHS data was incorporated. This study shows that MHS radiances over clear-sky conditions prior to convective initiation, when assimilated with a careful QC method, had a significant positive impact for QPFs compared with the control experiment, EXP1.

A convective precipitation case is investigated in this study. The robustness of the new cloud detection algorithm must be further tested before it could be added to the GSI model. The preliminary results from this study highlight the potential benefit of assimilating MHS radiance observations for improved coastal precipitation forecasts. In the future, continuing effort will be made (i) to assess the impacts of MHS...
brightness temperature measurements on regional NWP for many more cases, (ii) to investigate the applicability of the proposed MHS cloud detection algorithm on global scales, and (iii) to carry out a careful analysis of the QC procedure for other types of satellite data in the regional GSI system for hurricane forecasts similar to what was done in this study for MHS data.

FIG. 15. TSs of 3-h accumulative precipitation (mm) at 1-, 5-, 10-, and 15-mm thresholds over the entire model domain from 0000 to 2400 UTC 23 May 2008 for EXP1 (red) and EXP2 (blue).

FIG. 16. As in Fig. 15, but for ETS.
Acknowledgments. This work was jointly supported by the Chinese Ministry of Science and Technology under 973 Project 2010CB951600, NSF Project AGS-1037936, the NOAA GOES-R Risk Reduction Program, and the project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The views and opinions contained in this paper are those of the authors and should not be construed as an official National Oceanic and Atmospheric Administration or U. S. government position, policy, or decision.

REFERENCES


