First simulations with a whole atmosphere data assimilation and forecast system: The January 2009 major sudden stratospheric warming

H. Wang,1,2 T. J. Fuller-Rowell,1,2 R. A. Akmaev,2 M. Hu,1 D. T. Kleist,3 and M. D. Iredell1

Received 16 August 2011; revised 20 November 2011; accepted 21 November 2011; published 22 December 2011.

[1] A Whole atmosphere Data Assimilation System (WDAS) is used to simulate the January 2009 sudden stratospheric warming (SSW). WDAS consists of the Whole Atmosphere Model (WAM) and the 3-dimensional variational (3DVar) analysis system GSI (Grid point Statistical Interpolation), modified to be compatible with the WAM model. An incremental analysis update (IAU) scheme was implemented in the data assimilation cycle to overcome the problem of excessive damping by digital filter in WAM of the important tidal waves in the upper atmosphere. IAU updates analysis incrementally into the model, thus avoids the initialization procedure (i.e., digital filter) during the WAM forecast stage. The WDAS simulation of the January 2009 SSW shows a significant increase in TW3 (terdiurnal, westward propagating, zonal wave number 3) and a decrease in SW2 (semidiurnal, westward propagating, zonal wave number 2) wave amplitudes in the E region during the warming, which can be attributed likely to the nonlinear wave-wave interactions between SW2, TW3 and DW1 (diurnal, westward propagating, zonal wave number 1). There is a delayed increase in SW2 in the E region after the warming, indicating a modulation by the changing large-scale planetary waves in the lower atmosphere during the SSW. These tidal wave responses during SSW appeared to be global in scale. An extended WAM forecast initialized from WDAS analysis shows remarkably consistent tidal wave responses to SSW, indicating a potential forecasting capability of several days in advance of the effects of the large-scale tropospheric and stratospheric dynamics on the thermospheric and ionospheric variability.


1. Introduction

[2] Quite dramatic changes in thermospheric wind and temperature and ionospheric electrodynamics in response to sudden stratospheric warming (SSW) have been shown in both modeling studies and observations. Liu and Roble [2002] studied the effects of an internally generated SSW in an upper atmosphere general circulation model (GCM) coupled to a lower atmosphere GCM. While midwinter warmings in the polar stratosphere have long been known to be accompanied by mesospheric coolings [e.g., Labitzke, 1972], changes in the global circulation were predicted to result in a heating in the lower thermosphere above 100 km [Liu and Roble, 2002]. An increase in daytime ion temperature in the lower thermosphere and a slight cooling above were found by Goncharenko and Zhang [2008] in midlatitude radar observations during the January 2008 SSW. Chau et al. [2009] and Goncharenko et al. [2010] analyzed the ionospheric response at equatorial and low latitudes to the January 2008 and 2009 SSWs, respectively. Estimates of total electron content (TEC) in the Western Hemisphere showed a substantial shift of the daytime maximum plasma density from the afternoon to the morning hours during the warmings. The drastic transformation of the daily variation of plasma density has been attributed to changes in the diurnal pattern of the vertical plasma drifts determined from radar observations at Jicamarca near the geomagnetic equator in the Peruvian sector. These new results have triggered an explosion of studies of mechanisms and types of possible connections between terrestrial and space weather during SSW and other large-scale perturbations in the lower atmosphere [e.g., Chau et al., 2011].

[3] Fuller-Rowell et al. [2010] studied the effects on neutral dynamics and electrodynamics in the upper atmosphere...
of a moderate SSW internally generated in a free annual run of the Whole Atmosphere Model (WAM) [Akmaev et al., 2008; Akmaev, 2011]. Consistent with the observations and previous simulations, the warming in the polar stratosphere by about 30 K was accompanied with the cooling in the mesosphere. A substantial increase of the maximum amplitude of the migrating tidal mode TW3 was perhaps the single most pronounced dynamical effect. The effects of the warming on electrodynamics were investigated by feeding the wind fields from WAM input into the Coupled Thermosphere-Ionosphere-Plasmasphere-electrodynamics model (CTIPe) before and during the warming. The simulation showed a tendency for the peak of the daytime upward drift to increase and to occur at earlier times during the peak of the warming, and for the afternoon minimum to be lower, in general agreement with observations [e.g., Chau et al., 2009, 2011; Goncharenko et al., 2010].

[4] The purpose of this paper is to present the first whole-atmosphere simulations of a real event, the major January 2009 SSW. In order to follow real weather events, data assimilation (DA) and forecasts were performed using the Grid point Statistical Interpolation (GSI) as the data analysis subsystem and WAM as the forecast model. GSI is used in the National Centers for Environmental Prediction (NCEP) with lower atmosphere numerical weather prediction (NWP) models to produce operational DA and forecasts. Using the winds from the DA system to drive CTIPe, Fuller-Rowell et al. [2011] presented the response of equatorial ionospheric electrodynamics to the SSW, found in very good agreement with the observations. Because large-scale weather processes can be predicted several days in advance, this opens a tantalizing opportunity of developing a true forecast capability for space weather applications. The whole atmospheric data assimilation and forecast system is briefly described in the next section, followed by sections on data assimilation and forecast for the January 2009 SSW, and summary and conclusions.

2. The Whole Atmosphere Data Assimilation and Forecast System

[5] The whole atmosphere data assimilation and forecast system consists of an analysis system and a whole atmospheric model. The analysis system uses NCEP’s GSI system, modified to be compatible with WAM. The 3-dimensional variational (3DVar) analysis technique is used in GSI [Wu et al., 2002]. GSI is used in the NCEP’s operational regional and global data assimilation systems. WAM is an extension of the NCEP’s Global Forecast System (GFS), extended from 64 model levels (with model top at about 60 km) to 150 model levels (with model top at about 600 km) [Akmaev et al., 2008; Akmaev, 2011]. It covers the regions of important ionospheric processes and their variability. WAM includes basic ionospheric effects on neutral atmosphere, i.e., ion drag and Joule heating. The horizontal resolution of WAM used in this study is T62 (about 200 km).

2.1. Modifications to the GSI Analysis System

[6] To use GSI with WAM, several modifications were made to make GSI compatible with WAM. First, the background error (BE) statistics were redefined. The BE statistics (i.e., forecast error variance/covariance) was derived from the 24-hour difference in a 1-year WAM ‘free’ run (i.e., a continuous run without observational data injection). The method is similar in spirit to the so-called NMC-method (NMC stands for National Meteorological Center, i.e., current NCEP) [Parrish and Derber, 1992], but usually the BE statistics are derived from difference in a sample of short-term forecast of different lengths, e.g., 12 hour and 36 hour forecasts, valid at the same time [e.g., Belo Pereira and Berre, 2006]. Considering the existence of the large diurnal variability in the upper atmosphere, the 24-hour difference in the forecast were used. Second, the top-of-model (TOM) pressure is set to 0.005 hPa (about 80 km) in the Community Radiative Transfer Model (CRTM) used for radiance assimilation in GSI. Although the TOM pressure of WAM is $1.5 \times 10^{-9}$ hPa (about 600 km), the calculation of CRTM in WAM-GSI is restricted to 0.005 hPa. The WAM-GSI analysis appeared not affected by this restriction, since in this study WAM-GSI only assimilates observational data up to about 0.1 hPa (about 60 km) from the NCEP’s operational Global Data Assimilation System (GDAS). Third, WAM treats O, O2, and N2 separately, in addition to H2O and O3. The input and output in GSI were modified to handle these additional trace gas components.

2.2. Implementation of the Incremental Analysis Update (IAU) Scheme

[7] Experiments with WAM-GSI show that, when the conventional intermittent 6-hour cycle data assimilation (DA) strategy was used in WAM-GSI, important tides that were well simulated in the WAM ‘free’ runs [Akmaev et al., 2008; Fuller-Rowell et al., 2010] were severely damped by the digital filter required by WAM in the DA cycle. This pathological effect can not be remedied simply by tuning the digital filter used in the model. One way to remedy this damping of tidal waves in the upper atmosphere was to update analysis incrementally into the model, thus avoided the need for the initialization procedure (digital filter in this case) during the forecast stage. We implemented the so-called Incremental Analysis Update (IAU) scheme [Bloom et al., 1996] into the WAM-GSI DA cycle. Impacts of digital filtering and IAU on the middle atmosphere data assimilation is discussed by Sankey et al. [2007]. In work by Eckermann et al. [2009] the nonlinear normal mode initialization scheme was turned off with no apparent deleterious effects in their 3DVar DA system extending to the base of the thermosphere.

[8] Using IAU in WAM-GSI consists of the following steps.

[9] 1. After GSI analysis, analysis increments, i.e., GSI analysis minus first-guess (FG, i.e., WAM forecast), are created on WAM physics grids.

[10] 2. Run WAM for a 6-hour ‘forced’ forecast, restarted from 3 hour earlier before the analysis time, applying increments at each time step as a constant forcing spread over a 6-hour window centered around the analysis time.

[11] 3. Run WAM for a 6-hour ‘free’ forecast, restarted at the end of ‘forced’ forecast in step 2, to provide FG for GSI, and DA cycle back to step 1 again. WAM forecast can continue and branch off from step 3 for a medium-range (16 days) or extended (21 days) forecast. The WAM-GSI-IAU system, or briefly WDAS (Whole atmosphere Data Assimilation System), significantly improved simulation of
tidal waves in the upper atmosphere, as can be seen from the January 2009 SSW simulation in the following section.

3. Data Assimilation and Forecast Experiments for the January 2009 SSW

[12] The WDAS system described above is used to simulate the January 2009 SSW. The January 2009 SSW was a vortex split type of SSW [Fuller-Rowell et al., 2011]. It was a significant event with a singular rapid polar cap temperature increase and especially the observed thermospheric and ionospheric responses [e.g., Chau et al., 2011]. All conventional and satellite observational data from NCEP operational GDAS system are injected into WDAS. In this section the polar-cap temperature in the stratosphere at 10 hPa (about 30 km altitude) from the WDAS analysis and forecast will be shown. Space-time spectral analysis is also performed to show the responses to SSW of major tidal waves in the E region, the so-called dynamo region in the 90–160 km altitudes.

3.1. The Polar Cap Temperature Change at 10 hPa

[13] The polar cap temperature change during the January 2009 SSW is well captured by WDAS system. Figure 1 shows the two-month time series of the stratospheric polar-cap temperature from WDAS and GDAS analyses. In both analyses the temperature at 10 hPa averaged over the 60°N polar cap increased over 50°K sharply within about a week, peaked on 23 January 2009. Note that in the stratosphere the WDAS analysis followed closely the NCEP operational GDAS analysis. In addition, shown in Figure 1 is also the average polar-cap temperature at 10 hPa from the extended (21-day) WAM forecast initialized at 00 UTC on 15 January 2009 from the WDAS analysis. The overall temperature changes followed the analysis closely, especially in the first 7 days.

3.2. Tidal Wave Responses in the E Region

[14] The standard space-time spectral analysis [e.g., Wheeler and Kiladis, 1999] was applied to the hourly model output of WDAS analysis and forecast. The data were grouped in 3 days and linear trends were removed. We are specifically interested in the change of major tidal waves during the January 2009 SSW.

[15] The migrating SW2 (semi-diurnal, westward propagating, zonal wave number 2) wave is a major driver of the E region dynamo winds [e.g., Millward et al., 2001]. The migrating TW3 (ter-diurnal, westward propagating, zonal wave number 3) wave was also found under significant change in a WAM model-generated moderate SSW [Fuller-Rowell et al., 2010]. Figure 2 shows the time series of the averaged E region SW2, TW3 and DW1 (diurnal, westward propagating, zonal wave number 1) wave amplitudes of the zonal wind from the WDAS analysis and an extended
forecast. Several features of these tidal waves can be noted from Figure 2.

1. SW2 is a dominant component of the E region tidal waves in the NH. Its change could have important impacts on the overall dynamo winds in this region.

2. There is a significant increase in TW3 and a rather large decrease in SW2 during the warming.

3. SW2 in both NH and SH attained a peak amplitude about 10 days after the SSW peak, followed quite closely the stratospheric warming trend earlier, indicating a modulation of SW2 by the changing large-scale planetary waves (PWs) in the lower atmosphere during SSW.

4. Both SW2 and TW3 were also under significant change in the SH, indicating a global response of SW2 and TW3 to SSW.

5. The tidal waves from the extended forecast are remarkably consistent with the WDAS analysis, indicating a great forecasting capability of the WDAS system, several days, even weeks in advance.

3.3. Nonlinear Wave-Wave Interaction

The large increase in TW3 and decrease in SW2 at the time of warming shown in Figure 2 prompt a closer analysis on the relation between the two. Figure 3 is a latitudinal-vertical view of the SW2 and TW3 wave amplitudes of the zonal wind from WDAS analysis two days apart: (a) SW2 and (b) TW3 on 21 January 2009; (c) SW2 and (d) TW3 on 23 January 2009. Drawn on the same color scale. Model levels shown on the left y-axis are on constant pressure levels, with their approximate altitudes shown on the right y-axis.
among the members of the triad may lead to rapid rate of change of wave amplitudes averaged in the upper E region where TW3 maxima were located.

zonal wind from WDAS analysis two days apart. On 21 January 2009 (2 days before the SSW peak), SW2 has two noticeable local maxima with a comparable amplitude of about 30 ms$^{-1}$: one around 120 km altitude and 40°N and one around 100 km altitude and 55°N (Figure 3a). On 23 January 2009 (on the day of the SSW peak), the upper local maximum decreased about 5 ms$^{-1}$, leaving only one local maximum around 100 km. On contrast, TW3 has one local maximum around 150 km and 50°N; and its peak amplitude increased about 10 ms$^{-1}$ in two days, from about 20 ms$^{-1}$ (Figure 3b) to 30 ms$^{-1}$ (Figure 3d). These variations suggest a possible connection between the two tidal waves.

[23] Nonlinear interactions of migrating diurnal and semidiurnal tides have long been proposed as a possible mechanism for generation of the semidiurnal tide in the lower thermosphere [e.g., Glass and Felton, 1975], although observational studies may suffer limitations of spatial (e.g., point observations) or time (e.g., satellite observations) resolutions. Here we provide a plausible case of demonstration of the mechanism in the simulation of a fully nonlinear model. According to nonlinear wave-wave interaction theory [e.g., Pedlosky, 1986], the SW2 wave is donating energy to TW3 for several days. Then from 19 January 2009 there is a rapid growth of TW3 and a relative large decrease in SW2. DW1 has on average relative small amplitude compared to SW2 and TW3 in the E region (see Figure 2). Its sign of change shown in Figure 4 is not as consistent as the other two. But considering its small amplitude and the averaging used, as well as the presence of dissipative processes and other waves, this alone can not eliminate the possibility of the wave-wave interaction between the three waves. Discussion and references on semi-diurnal tide, including nonlinear interaction, and its climatology in the extended Canadian Middle Atmospheric Model can be found in work by Du and Ward [2010].

[23] The above analysis shows that it is likely that nonlinear wave-wave interaction between SW2, TW3 and DW1 has contributed to the rapid growth of TW3 during the warming. Other possible mechanisms were mentioned by Fuller-Rowell et al. [2010]. At the SSW peak time, TW3 attained a comparable magnitude to that of SW2. The appearance of large amplitude TW3 wave is likely to have significant impact on the E region dynamo winds and electrodynamic. This is investigated by Fuller-Rowell et al. [2011].

4. Summary and Conclusions

[24] WAM is combined with NCEP GSI data analysis system to follow the whole atmospheric dynamics from 0–600 km for a period in January and February 2009. This time period includes one of the biggest major SSW events in recent history. The forecast model WAM is an extension of the NCEP GFS model. GSI uses the 3DVar technique, and was modified to be compatible with WAM. When the conventional intermittent 6-hour cycle DA strategy was used in WAM-GSI, the tidal waves in the thermosphere were severely damped. These waves significantly influence the ionosphere, thus it is vital to have representative signatures of these waves in the analysis for successful simulation of ionospheric variability. To overcome this problem, an IAU scheme was implemented in WAM-GSI. IAU updates analysis incrementally into the model, thus avoids the initialization procedure (i.e., digital filter) during the WAM forecast stage. The data assimilation and forecast experiments with the WAM-GSI-IAU system, or briefly the Whole atmosphere Data Assimilation System (WDAS), show significant improvement for the simulation of tidal waves in the upper atmosphere.

[25] The simulation of the January 2009 SSW with the WDAS system shows the following.

[26] 1. WDAS analysis is in good agreement with NCEP’s operational GDAS analysis.

[27] 2. At the time of the warming, there is a significant increase in TW3 and a decrease in SW2 wave amplitudes in the E region. These changes are likely to have significant impact on the E region dynamo winds and ionospheric electrodynamic [see Fuller-Rowell et al., 2011].

[28] 3. The rapid growth of TW3 during the warming can be attributed likely to the nonlinear wave-wave interactions between SW2, TW3 and DW1.

[29] 4. The delayed increase in SW2 in the E region after the SSW peak indicates a modulation by the changing
large-scale planetary waves in the lower atmosphere during the SSW. These tidal wave responses during SSW appeared to be global in scale.

[30] 5. The extended forecasts showed fairly consistent tidal wave responses to SSW, indicating a potential forecasting capability of several days in advance of the effect of the large-scale tropospheric and stratospheric dynamics on thermospheric and ionospheric variability.

[31] Acknowledgments. We thank R. Treaton for answering various GSI-related questions. Funding for this work was provided by NASA Heliophysics Theory Program, AFOSR MURI NADIR, and NASA LWS Strategic Capabilities. Computacional resources were provided by NOAA NCEP R&D high performance computing system.

Robert Lysak thanks the reviewer for his or her assistance in evaluating this paper.

References


R. A. Akmaev, Space Weather Prediction Center, NOAA, W/NP9, 325 Broadway, Boulder, CO 80305, USA.

T. J. Fuller-Rowell, M. Hu, and H. Wang, CIRES, University of Colorado at Boulder, 216 UCB, Boulder, CO 80309, USA. (houjun.wang@noaa.gov)

M. D. Iredell and D. T. Kleist, Environmental Modeling Center, NOAA, 5200 Auth Rd., Camp Springs, MD 20746, USA.