GOES DATA ASSIMILATION IN MM5: APPLICATION FOR TEXAS AIR QUALITY STUDY 2000

Arastoo Pour Biazar[1*], Richard McNider[1], Kevin Doty[1], Stephanie L. Haines[1], William Lapenta[2], Ron Suggs[2], and Gary Jedlovec[2]

Introduction:

Accurate predictions of boundary layer characteristics are crucial to air quality forecasts. While the ambient temperature and the amount of sunlight reaching the surface greatly impacts the reaction rates in the atmosphere, the boundary layer height directly affects the tracer concentrations within the boundary layer. Therefore reducing the uncertainties in the meteorological input to a photochemical model can greatly enhance the air quality predictions.

Surface moisture availability and surface heat capacity are critical to accurate predictions of temperature and boundary layer characteristics. However, in the absence of direct routine measurements of these parameters, they are a source of uncertainty in the mesoscale meteorological models [McNider et al., 1994]. For example, in models such as MM5 (Pennsylvania State University/National Center for Atmospheric Research (PSU/NCAR) Mesoscale Model Version 5) surface moisture availability has been specified based on land use classification and climatological conditions or through land surface hydrology models. Use of the first technique leads to errors due to inaccuracies in relating land use types to moisture and due to deviations from climatological norms. The second technique - the use of a surface hydrological model - while perhaps an improvement, still requires specifications of difficult to quantify parameters such as root zone moisture, plant physiological characteristics such as root uptake, stomatal resistance, soil hydrologic conductivity, and antecedent precipitation.

Heat capacity of the surface in models such as MM5 has also been specified based on land use classification. While specification of heat capacity is relatively straight forward for single composition objects such as water, stone, concrete etc. The practical specification of heat capacity on a 4 km or 12 km grid where the surface is made up of everything from buildings, to streets, to grass, to trees to standing water is extremely difficult. This is especially true in the highly heterogeneous urban and suburban environment.

There have been many advances in assimilating the observational data into the numerical models [cardinali et al., 2003; Seo et al., 2003; Xie et al., 2002]. The utilization of point measurements from National Weather Service stations into the numerical weather prediction models has proven to be valuable [Zapotocny, et. al., 2002; NWS, 1999, 2000]. However, the satellite data offers few advantages over the surface monitors.

* Corresponding author address: Arastoo Pour Biazar, Earth System Science Center, University of Alabama, Huntsville, AL 35899
[1] Earth System Science Center, University of Alabama in Huntsville, Huntsville, Alabama
[2] NASA Marshall Space Flight Center, Huntsville, Alabama

First, the extensive spatial coverage that the satellite data offers extends over the data scarce regions where the surface stations are sparse. Second, satellite pixel size (1 to 4-km for GOES) is comparable to the size of model grid spacing and satellite pixel radiative measurements provide a natural averaging process needed for the grid average values in the model.

In the current study we have utilized skin temperature, surface albedo, and insolation retrievals from GOES Imager to infer surface moisture availability and heat capacity compatible with the model parameterization. The assimilation technique, based on McNider, 1994, uses the GOES skin temperature tendencies during the mid-morning time frame to improve specification of surface moisture. The satellite assimilation technique for the recovery of moisture availability has been tested in both case study modes (McNider et al. 1994, 1995) and operational modes (Lapenta et al 2000) and has been shown to improve model performance. GOES skin temperature tendencies in the evening period are used to adjust model heat capacity. MM5 version 3.4 was modified for the use of satellite tendencies and was applied for the following simulations.

MM5 simulations span over the period of August 23, to September 2, 2000 (period of Texas Air Quality Study 2000 modeling activities) and are performed over four domains from 108- (continental U.S.) to 4-km resolution (east Texas).

GOES Satellite Retrievals:
The assimilation technique uses the following remotely-sensed data products, all of which are based on Geostationary Operational Environmental Satellite (GOES) visible and infrared channels (McNider et al., 1994; Suggs et al., 1998):
(1) Clear Sky Composite – an intermediate product used for estimation of both albedo and insolation. It consists of the minimum albedo observed at a particular time over a twenty-day period in the GOES visible band.
(2) Infrared Cloud Mask – a cloud mask computed from the GOES 3.7 and 10.7 micron channels, using thresholding and spatial coherence to identify clouds using cloud-land differences in solar reflectance (daytime) and emissivity (nighttime).
(3) Albedo and Insolation – computed following Gautier et al., (1980) and Diak and Gautier (1983) from GOES visible channel observations. Albedo is determined using the Clear Sky Composite, and Insolation is estimated as total insolation from direct and diffuse sources, including cloud attenuation.
(4) Skin Temperature and Total Precipitable Water – retrieved simultaneously using a physical split window technique (Jedlovec 1987; Suggs et al. 1998) with at least two longwave window GOES channels. The technique is based on perturbation of the radiative transfer equation, using first-guess profiles of temperature and moisture from the mesoscale model forecasts and an assumed emissivity of 0.98. Retrievals can be made at model resolution, with pixel averaging of the observed radiances, or at pixel resolution.

Satellite Assimilation Techniques:
The surface energy budget in MM5 based on a "force-restore" method developed by Blackadar (Blackadar 1979; Zhang and Anthes 1982) can be represented as,

$$C_g\left(\frac{dT_G}{dt}\right) = \left(R_N - H - G\right) - E \qquad (1)$$

Where dT_G/dt is the surface temperature tendency, C_g the surface heat capacity (per unit area), R_N is the net radiation (including incident shortwave, incoming atmospheric longwave, and outgoing longwave), H is the sensible heat flux, G is the soil heat flux, and E is the latent heat flux.

Following McNider et al., 94, we invoke the critical assumption that in the morning all of the terms in the model's surface energy budget are the same as for the actual energy budget observed by the satellite except for the latent energy term E. This is based on the idea that we know the least about evapotranspiration and that in the mid-morning the energy budget is most sensitive to moisture availability (Wetzel et al. 1984). With this assumption we take the difference of the surface energy budgets for the model and satellite to obtain

$$E_s = C_g\left[\left(\frac{dT_G}{dt}\right)_m - \left(\frac{dT_G}{dt}\right)_s\right] + E_m \qquad (2)$$

Where $\left(dT_G/dt\right)_s$ is calculated from hourly GOES-derived surface temperature products retrieved at model grid points. Hereafter, we represent the model quantities by the subscript *m* and satellite quantities by *s*. Allowing $h = \frac{C_g}{E_m}\left[\left(\frac{dT_G}{dt}\right)_m - \left(\frac{dT_G}{dt}\right)_s\right]$ to represent the adjustment needed to bring the model moisture flux in agreement with the satellite inferred flux; the recovered latent heat flux can be expressed as:

$$E_s = (1 + h)\ E_m \qquad (3)$$

The way in which the moisture flux is adjusted within the model is dependent upon the flux formulation used. In MM5, surface specific humidity is not a prognostic variable. Therefore, we adjust what is called the moisture availability parameter (M) which represents the *fraction* of possible evaporation for a saturated surface (equal to 1 over open water and 0 over a non-evaporating surface). Using the definition of the latent heat flux in the Blackadar boundary layer scheme in MM5 and using the satellite inferred latent heat flux from above, the satellite inferred moisture availability is given by

$$M_s = E_s \frac{\ln\left(\frac{ku_* z_a}{k_a} + \frac{z_a}{z_l}\right) - \varphi_h}{\rho k u_* \left(q_{sfc}\left(T_g\right) - q_a\right)} = \left(1 + h\right) M_m \qquad (4)$$

Where $q_{sfc}\left(T_g\right)$ is the saturation mixing ratio of the surface, q_a is the mixing ratio of the air immediately above the surface, k is the von Karman constant, u_* is the frictional velocity, z_a the height of the lowest model layer, z_l the depth of the molecular layer, k_a

is a background molecular diffusivity, and φ_h is a non-dimensional stability parameter for heat and water vapor.

In addition to the recovery of moisture availability, as mentioned above, the GOES derived surface insolation is also assimilated into the MM5 surface energy budget via direct insertion. This insures that we meet our assumptions and that the retrieved skin temperature is consistent with the insolation used in the surface energy budget.

Equation 1 is extremely sensitive to moisture availability changes and can become unstable. This high sensitivity/instability causes oscillation in the model LST tendency and produces large disagreements between the model and the satellite tendencies. At times the disagreement is so large that the adjustment term (*h*) causes the adjusted moisture availability to become negative or to exceed its upper limit, 1. In order to minimize the error and to insure that the model tendency converges to the satellite tendency, we have modified our method to constrain *h* by nudging the model moisture availability to its satellite inferred value. The nudging factor δ_M is expressed as:

$$\delta_M = G(\Delta t_m, \Delta t_s)\, H(\Delta x_m, \Delta x_s)\, R(M)\, S(h) \approx \frac{.06\, h}{h^2 + 25.} \qquad (5)$$

Where *G* is the time step nudging factor that accounts for the fact that satellite tendencies are hourly while the model tendencies are calculated at each time step ($G \sim \Delta t_m / \Delta t_s$). *H* accounts for higher spatial variation of GOES skin temperature retrievals (~ 4km grid) vs. the model grid resolution ($G \sim \Delta x_s / \Delta x_m$). *R* is the response factor, adjusted based on the duration of assimilation. And *S* is defined as a function of *h* to insure that the adjustment to moisture availability is only performed when our assumptions are valid and the difference between satellite observed tendencies and that of the model are reasonable and can be ascribed only to the latent heat flux term.

$$S(h) = \frac{10\, h}{h^2 + 25.} \qquad (6)$$

The function varies between 0 and 1 and insures that the largest adjustment only will take place when *h* is within a reasonable range. High sensitivity of the surface energy budget to radiation forcing in the morning and also to the changes in moisture availability can cause large variations in *h*. The function *S* also insures that there is minimal adjustment to *M* at lower values of *h*. The technique suffers from the fact that the surface heat capacity is not adjusted in conjunction with the adjustment in *M*. Using the new adjustment term, the moisture availability is adjusted accordingly.

$$M_s = (1 + \delta) M_m \qquad (7)$$

Carlson (1986) showed that while moisture availability was the most sensitive variable in the mid-morning time frame in the surface energy budget, heat capacity was the most sensitive parameter in the early evening. We outline below a process for using GOES LST tendencies to adjust the bulk heat capacity.

We now consider that the model C_{bm} may be different than the satellite C_{bs} within the model and satellite surface energy budgets. If we assume the moisture availability has been correctly specified (or surface evaporation is negligible late in the day when stomata have closed), and there is negligible difference in net radiation, sensible, latent, and soil heat fluxes, we can subtract the model and satellite energy budget equations and solve for C_{bs} to get:

$$C_{bs} = C_{bm}\left(\frac{dT_G}{dt}\right)_m \bigg/ \left(\frac{dT_G}{dt}\right)_s \qquad (8)$$

The initial value for the model C_{bm} would be determined in the normal fashion via a simple lookup table. The new C_{bs} would be subsequently used as the model value. Again, as for the moisture availability some averaging and constraints on C_{bs} is necessary. Our initial attempts at recovering the heat capacity for the Texas-2000 study period has shown improvements in the nighttime temperature predictions.

MM5 Simulations:
For MM5 simulations, we performed a set of control simulations on 108/36/12/4 km domains that have been used in Texas AQS2000 modeling studies. The model configuration for these simulations were similar to those used by John Nielsen-Gammon (personal communications) for his base case studies. That is, using FDDA gridded input, no observational nudging, simple ice moisture scheme (Dudhia), using look-up table for moist physics, Grell scheme for cumulus parameterization, MRF PBL scheme, RRTM longwave scheme, 5-layer soil model, and shallow convection scheme. The simulations for 108 and 36-km domains were performed simultaneously with two-way nesting, while for the 12 and 4-km domains we used one way nesting.

Using the satellite retrievals of skin temperature, surface insolation and surface albedo, we performed the simulations using our satellite assimilation techniques. The first set of simulations did not include the adjustments to the heat capacity. The results were evaluated against both the National Weather Service (NWS) surface observations and also the available boundary layer heights from the TEXAQS2000 campaign. The results from MM5 with satellite assimilation showed improvement in the 2-m temperature predictions over the Texas domain. However, for the 36-km domain that covered most of the southeast, the assimilation run exhibited warm bias for the eastern and northeastern part of the domain.

Results and discussion:
The inferred moisture availability exhibits the overall drying of the surface during the period of study for the Texas-2000 study period. Compared to the control simulations the assimilation runs improve the model predictions of temperature where the control exhibits cool bias, but exacerbate the warm bias in the control run. The assimilation runs

also show improvement over the control runs as the model resolution increases (since the impact of land-use inhomogeneity become more pronounced and comparable to the satellite pixel area).

Figure 1 shows the skin temperatures as observed by satellite for August 30, 2000, at 19:45 GMT for the 36-km domain. The figure also shows the skin temperatures from control simulation and the simulation with satellite assimilation at 20:00 GMT. It is obvious that the assimilation technique is making notable improvement in the skin temperatures compared to the control run. The temperatures in the western part of the domain, while showing an improvement over the control simulation are still cooler than the satellite observations. This is due to the fact that the western part of the domain is initially dry and therefore our method of adjusting the moisture availability is not as effective as it is in the center of the domain. This feature also points out the necessity for recovering surface heat capacity. Panel (d) in figure 1 shows the corresponding adjusted moisture availability. As it can be seen, most of the domain, with the exception of northern-northeastern part of the domain, dries out for this period. This is in agreement with the surface analysis chart for this period.

Figures 2 and 3 exhibit the scatter plots for all four domains in the study (108-, 36-, 12-, and 4-km domains). The results indicate that as MM5 grid spacing decreases in high-resolution simulations, MM5 performance in capturing the spatial/temporal variation in skin temperature degrades. In the 12- and 4-km resolution domains, there is no skill in the control MM5 simulation. On the other hand, the satellite assimilation seems to be improving the model performance in explaining some of this variation. However, the scatter in the data is still substantial. There are several factors that could be contributing to this scatter. These factors include the errors in wind fields, the non-uniformity of bias in the retrievals, surface heat capacity, and/or the limitation of our technique (where the soil is extremely dry and the model is still unable to reproduce the satellite observed tendency).

A more detailed discussion of the results will be presented at the meeting.

References:

Blackadar, A. K., 1979: High resolution models of the planetary boundary layer. *Adv. Environ. Sci. Eng.,* **1**, 50-85.

Cardinali, Carla, Isaksen, Lars, Andersson, Erik. 2003: Use and Impact of Automated Aircraft Data in a Global 4DVAR Data Assimilation System. Monthly Weather Review: Vol. 131, No. 8, pp. 1865–1877.

Diak, G. R. and C. Gautier, 1983: Improvements to a simple physical model for estimating insolation from GOES data. *J. Appl. Meteor.*, **22**, 505-508.

Diak, G. and M. Whipple, 1995: Note on estimating surface sensible heat fluxes using surface temperatures measured from geostationary satellite. J. Geophys. Res., **100**, 25453-25461

Gautier, C., G. R. Diak, and S. Mass, 1980: A simple physical model for estimating incident solar radiation at the surface from GOES satellite data. *J. Appl. Meteor.* , **19**, 1005-101.

Jedlovec, G.J. , 1987: Determination of atmospheric moisture structure from high resolution MAMS radiance data., Ph.D. dissertation. Ph.D. Degree.The University of Wisconsin-Madison, University Microfilm International, Ann Arbor, MI. 187 pgs.

McNider, R. T., J. A. Song, and S. Q. Kidder, 1995: Assimilation of GOES-derived solar insolation into a mesoscale model for studies of cloud shading effects. *Int. J. Remote Sens., 16*, 2207-2231.

McNider, R.T., A.J. Song, D.M. Casey, P.J. Wetzel, W.L. Crosson, and R.M. Rabin, Toward a dynamic -t thermodynamic assimilation of satellite surface temperature in numerical atmospheric models, *Mon. Wea. Rev.*, *122*, 2784-2803, 1994.

NWS, cited 1999: EDAS. NWS Tech. Procedures Bull. [Available online at http://www.nws.noaa.gov/om/tpb/3d-eta.htm.]

——, cited 2000: EDAS. [Available online at http://www.emc. ncep.noaa.gov:/mmb/gcp/logs/eta.log.code.]

Seo, Dong-Jun, Koren, Victor, Cajina, Neftali. 2003: Real-Time Variational Assimilation of Hydrologic and Hydrometeorological Data into Operational Hydrologic Forecasting. Journal of Hydrometeorology: Vol. 4, No. 3, pp. 627–641.

Suggs, Ronnie J., Jedlovec, Gary J., Guillory, Anthony R. 1998: Retrieval of Geophysical Parameters from GOES: Evaluation of a Split-Window Technique. Journal of Applied Meteorology: Vol. 37, No. 10, pp. 1205–1227.

Wetzel, P. J., 1984: Determining soil moisture from geosynchronous satellite infrared data: A feasibility study. *J. Climate Appl. Meteor*., **23,** 375-391.

Xie, Yuanfu, Lu, Chungu, Browning, Gerald L. 2002: Impact of Formulation of Cost Function and Constraints on Three-Dimensional Variational Data Assimilation. Monthly Weather Review: Vol. 130, No. 10, pp. 2433–2447.

Zapotocny, Tom H., Menzel, W. Paul, Nelson, James P., Jung, James A. 2002: An Impact Study of Five Remotely Sensed and Five In Situ Data Types in the Eta Data Assimilation System. Weather and Forecasting: Vol. 17, No. 2, pp. 263–285.

Zhang, D., and R. A. Anthes, 1982: A high-resolution model of the planetary boundary layer: Sensitivity tests and comparisons with SESAME-79 data. *J. Appl. Meteor.,* **21**, 1594-1609.

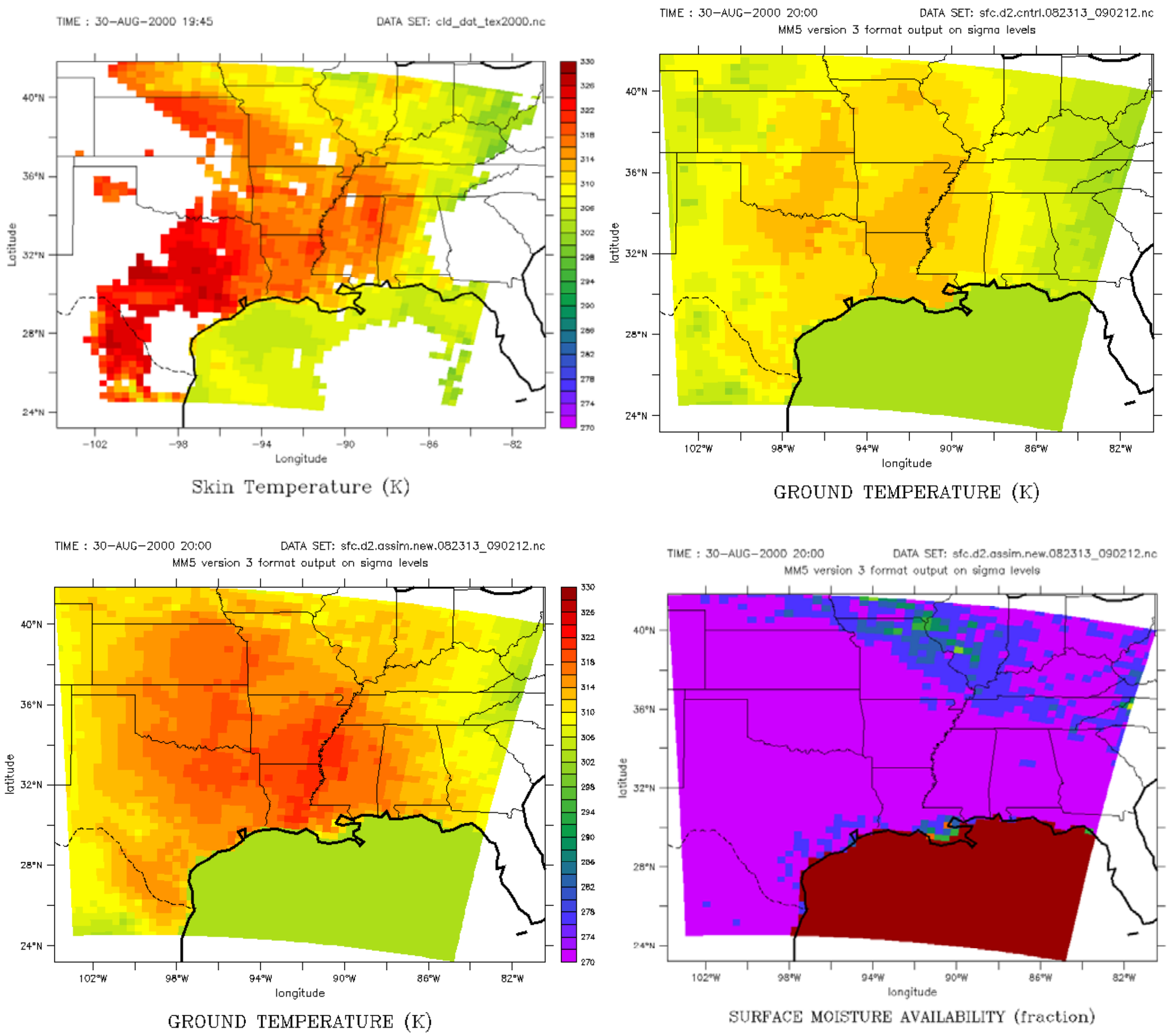
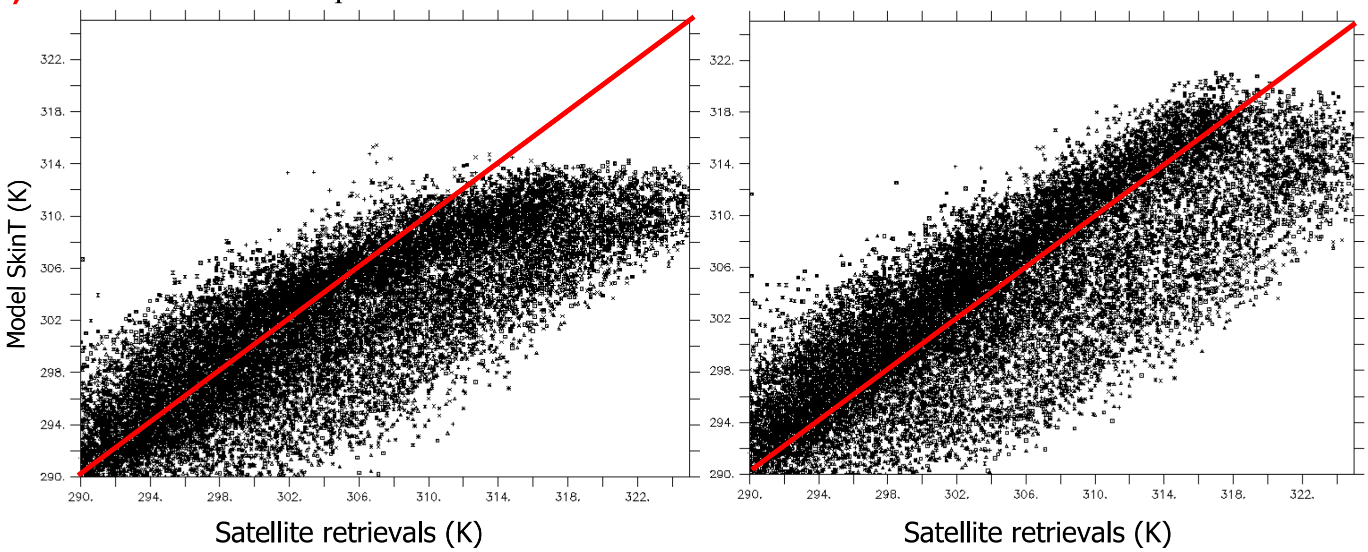


Figure 1. Skin temperature **a)** observed by satellite; **b)** from MM5 control simulation; and **c)** from MM5 simulation with satellite assimilation. Also **d)** recovered surface moisture availability for the assimilation run.

a) 108-km domain scatter plots



b) 36-km domain scatter plots.

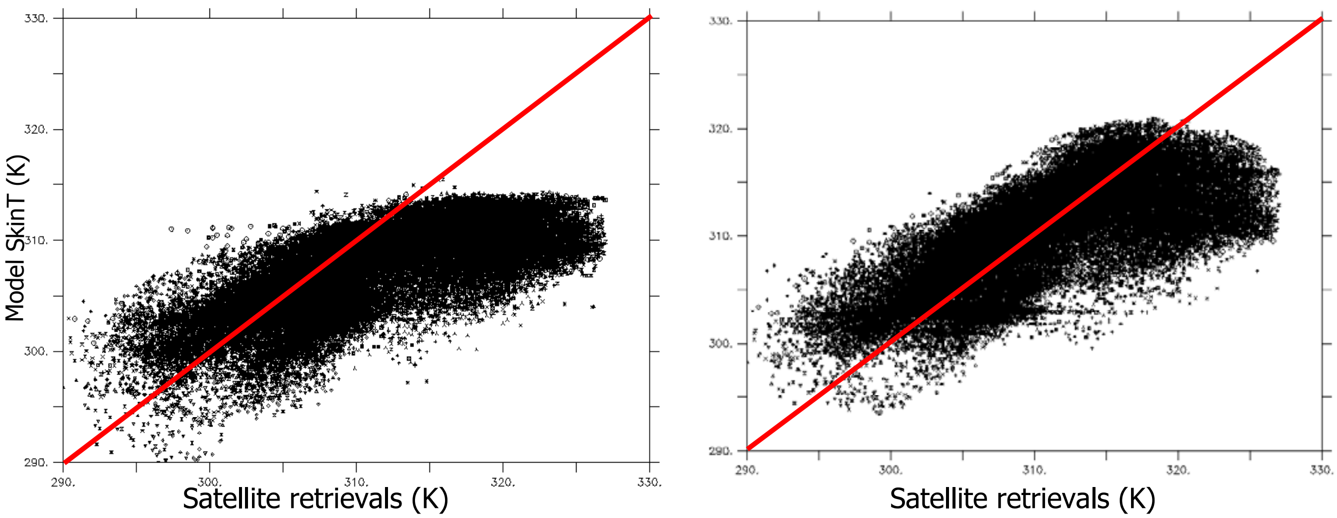
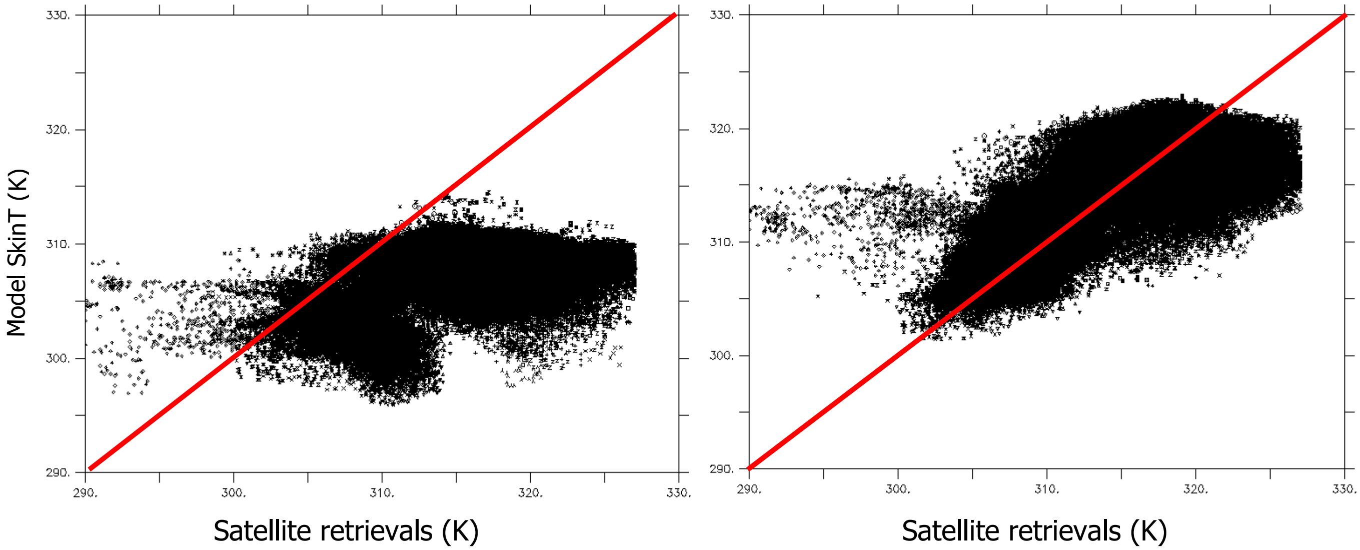


Figure 2. Skin temperature scatter plots, MM5 vs. satellite observed, for 14:00-21:00 GMT, August 25-30, 2000, over land. Left panel is from control MM5 simulation, and the right panel from MM5 with satellite assimilation. a) 108-km domain (domain 1); b) 36-km domain (domain 2).

c) 12-km domain scatter plots



d) 4-km domain scatter plots.

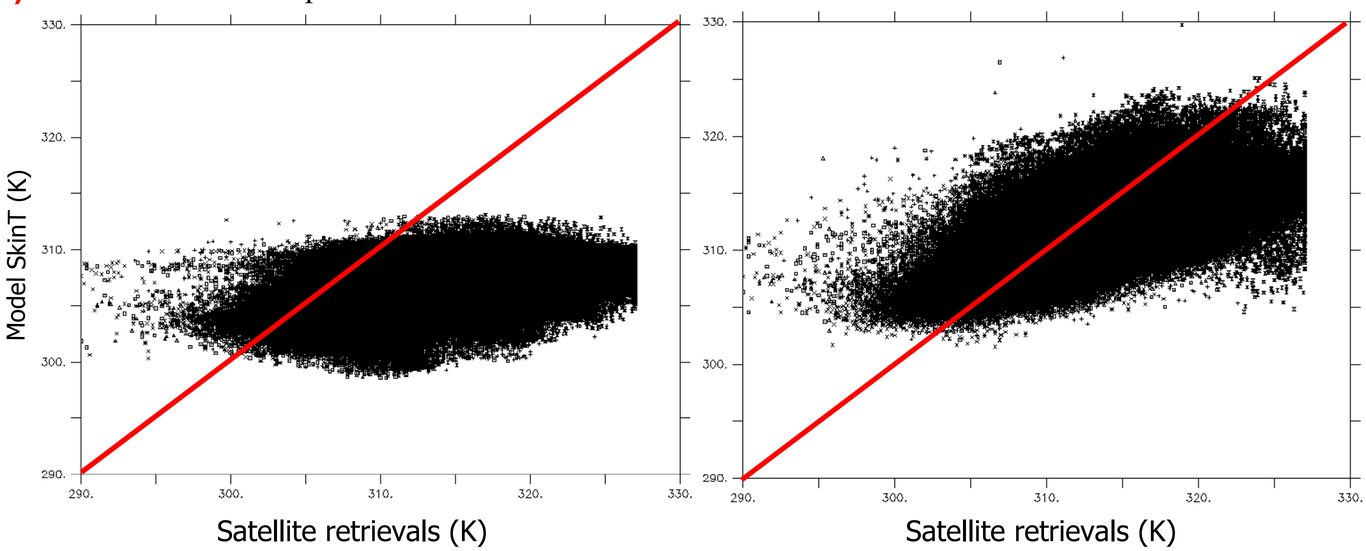


Figure 3. Skin temperature scatter plots, MM5 vs. satellite observed, for 14:00-21:00 GMT, August 25-30, 2000, over land. Left panel is from control MM5 simulation, and the right panel from MM5 with satellite assimilation. c) 12-km domain (domain 3); d) 4-km domain (domain 4).