2A.8

EVALUATING THE USE OF THE HIGH-RESOLUTION LAND DATA ASSIMILATION SYSTEM IN AGRICULTURAL DECISION SUPPORT

Wiliam Myers*, Fei Chen, Seth Linden, and Ying Zhang
National Center for Atmospheric Research, Boulder, CO

1. INTRODUCTION

Technological advances in weather forecasting and land surface modeling have led to an increased ability to predict soil temperature and moisture. These variables are critical building blocks to the development of high level agriculture-specific models such as pest models and plant development models. The output of these domain-specific models would be presented to agricultural end users through Agricultural Decision Support Systems (DSSs) targeted to the specific needs of user groups.

The National Center for Atmospheric Research (NCAR) has teamed up with DTN/Telvent, a major US agricultural weather provider, to develop a soil condition forecast system, run downstream agriculture-specific models, and integrate the data into a DSS. NCAR's Dynamic Integrated ForeCast system (DICast) was used in conjunction with variations of the High Resolution Land Data Assimilation System (HRLDAS), a land-surface modeling (LSM) system, to produce soil condition forecasts.

A major area of interest to this NASA-funded project was whether soil temperature and moisture forecasts could be improved through the use of remotely sensed data from the MODIS satellite to help better determine the vegetation coverage. Prior to this, the land surface model had used vegetation indices from a climatological data set with monthly temporal and 15km spatial resolution. The MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation (FPAR) data provide a more current vegetation status at higher spatial resolution and should improve the LSM initial conditions.

2. PROJECT ORGANIZATION

The project was organized into three main areas of research and development. The first was geared toward core development of and enhancements to the HRLDAS model. This work included exploring and improving the thermal transfer within the LSM. It also involved the extension of HRLDAS to incorporate the MODIS vegetation data. This research was carried out using retrospective analyses and observations from 2005-2007. The HRLDAS output in these retrospective runs was compared to soil temperature and moisture observations from the Soil Climate Analysis Network (SCAN), a mesonet operated by the National Resources Conservation Service under the US Department of Agriculture.

The second effort was to develop an operational weather and soil forecast system. This system was designed to be modular so that new weather forecasts or HRLDAS upgrades could be easily integrated. The system ran once per day, at 0900 UTC, in order to have forecasts available to users early in the morning. The raw soil temperature and moisture forecast data were also made available to DTN/Telvent through web-based graphics. The forecast domain was the central and eastern US, and is shown, along with sample soil temperature and moisture forecasts in figures 1 and 2. This is the main area of the DTN/Telvent agricultural user base. It is also a region in which dry-land farming is predominant. That is, in this region, farmers depend on precipitation rather than irrigation. This simplifies the soil forecast process since human effects are minimized. Unfortunately, there are relatively few soil observation sites within this domain. This makes verification more difficult. Due to the agricultural focus of this project, the tuning and verification efforts focused on forecasts at the 5 cm and 10 cm depths made during the growing season which we defined as the months of March through August.

The third effort was by the agricultural forecast advisors at DTN/Telvent. They used the web-based graphics when advising customers and also worked with users to determine the best way to incorporate this data into their DSS, DTN Online.

3. RESULTS

Early in the project, HRLDAS was configured with 4 subsurface nodes, at 5 cm, 25 cm, 75 cm and 150 cm. This matched the depths available in the North American Model (NAM). As the project progressed, it was determined that, with this configuration, the distance between near-surface nodes was too great to effectively model the heat transfer. This led to the addition of two nodes and a restructuring of the node depths. The node depths were set to 1 cm, 5 cm, 20 cm, 50 cm, 100 cm, and 166 cm. The middle four of these six nodes matched the SCAN observational depths. Besides providing a better verification basis, this change also reduced the 5 cm soil temperature forecast errors.

---

* *Corresponding author address:* William Myers, NCAR, 3450 Mitchell Lane, Boulder, CO 80302;
e-mail: myers@ucar.edu

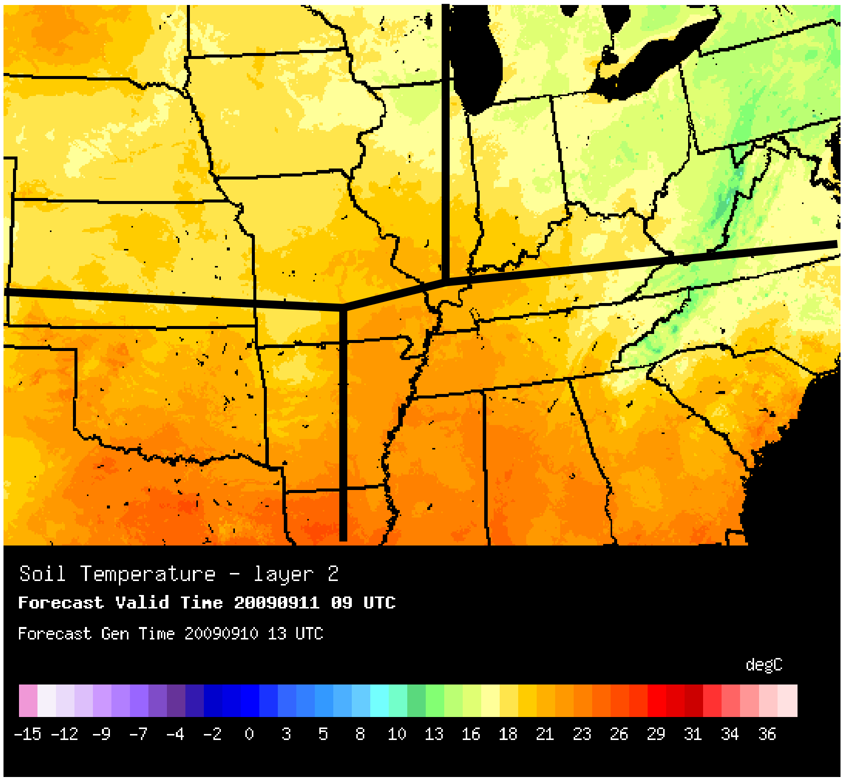

Figure 1: Sample HRLDAS gridded 5 cm soil temperature forecast. The 4 regions described in the results section are roughly obtained by splitting the domain in half in both the vertical and horizontal. The region boundaries are indicated.

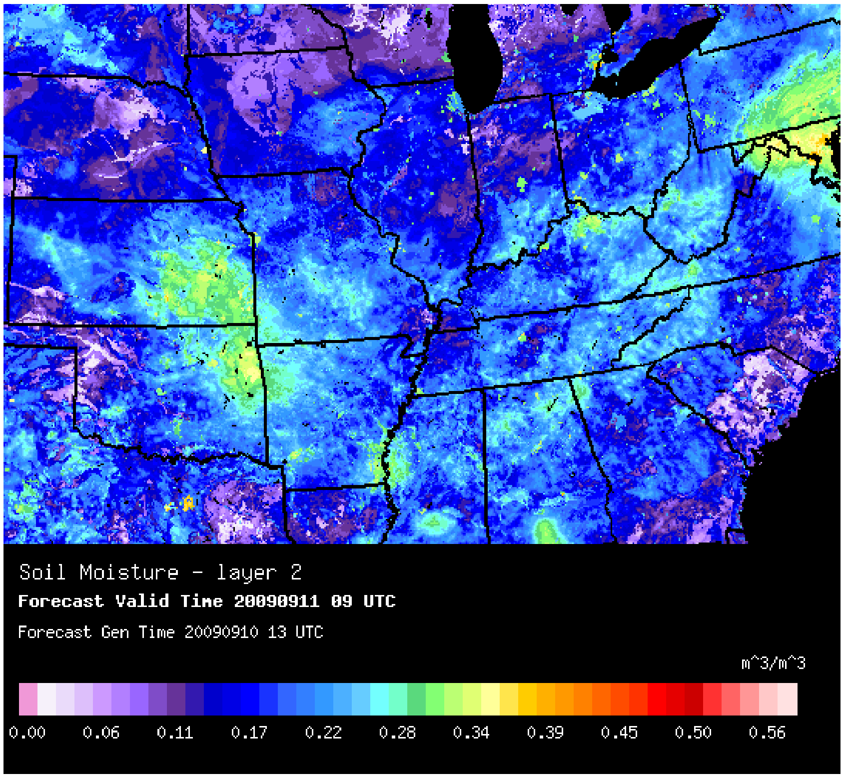

Figure 2: Sample HRLDAS gridded 5 cm soil moisture forecast.

The parameter czil, a surface heat exchange coefficient in HRLDAS, affects the turbulence near the surface and heat transfer at the surface. Originally, this parameter was constant across the whole domain. Research showed that varying czil would improve the soil temperature forecasts for one vegetation type while worsening them another. It was determined that by making czil depend on vegetation type some of this variability could be eliminated. Again, the results varied regionally. The improvement due to the incorporation of this algorithm can be seen in figures 3-7.

The MODIS data, available once every 8 days from polar orbiting satellites, was used to initialize the LSM vegetation fraction. In regions and seasons where the vegetation was near normal, little difference in the soil forecasts was expected. However, in areas where vegetation was above or below normal levels, the influence of using the MODIS data was expected to be noticeable. The improvement the MODIS data had on the czil algorithm can also be seen in figures 3-7. Some of the issues associated with using the MODIS data are addressed in the Discussion section below.

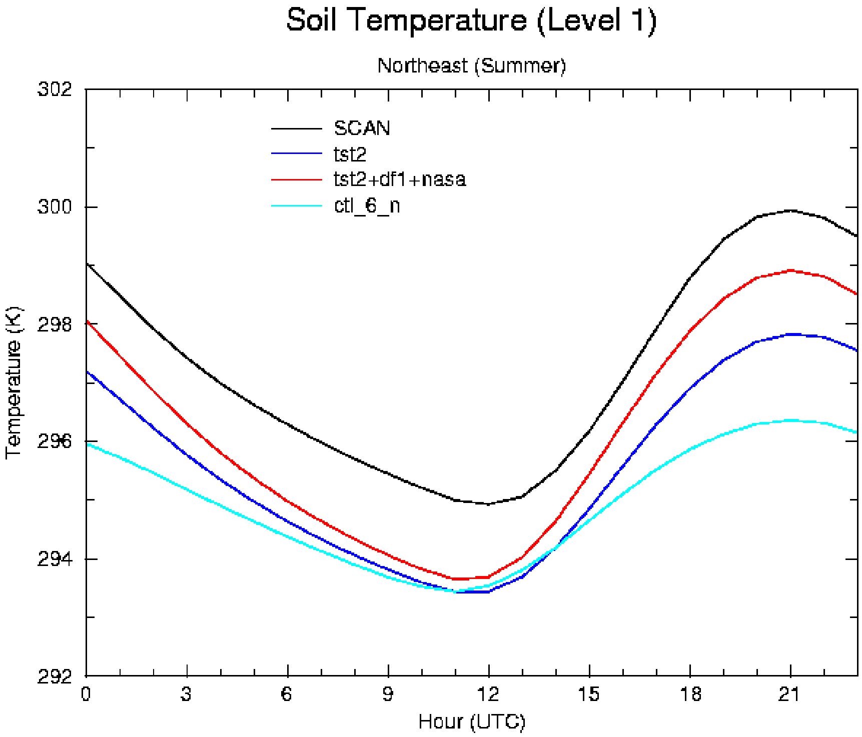

Figure 3: Summer forecasts in the NE show iterative improvement in 5 cm soil temperature times series. The observations (black) are better matched in successive runs starting from baseline 6-layer HRLDAS (ctl_6_n) to CZIL algorithm (tst2), and finally to runs with the MODIS data (tst2+df1+nasa). Despite the successive improvements, there is still a bias of more than 1°K.

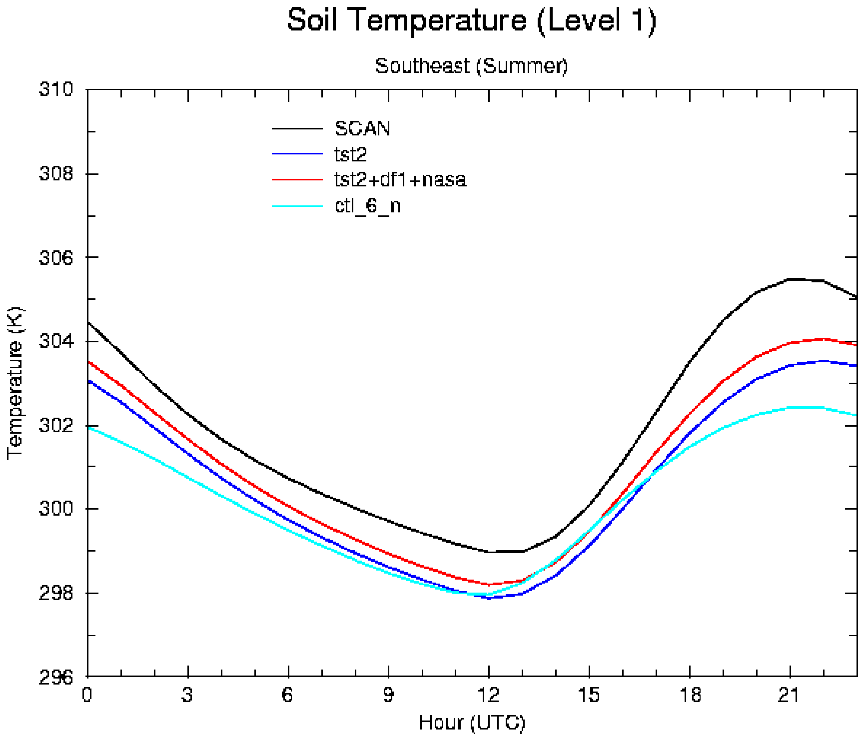

Figure 4: Summer 5 cm soil temperature forecasts for the SE region. Again, an iterative improvement is seen after each upgrade. The bias is less than in the NE region.

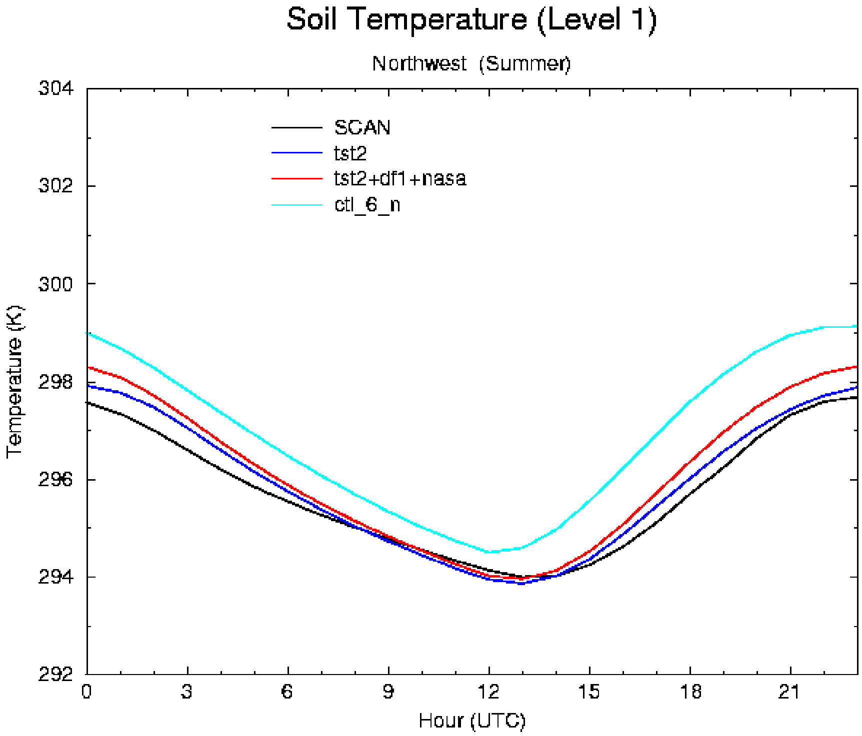

Figure 5: Summer 5 cm soil temperature forecasts for the NW region. The czil upgrade leads to an extremely good match with the observations. The inclusion of the NASA-MODIS data only slightly worsens the forecasts.

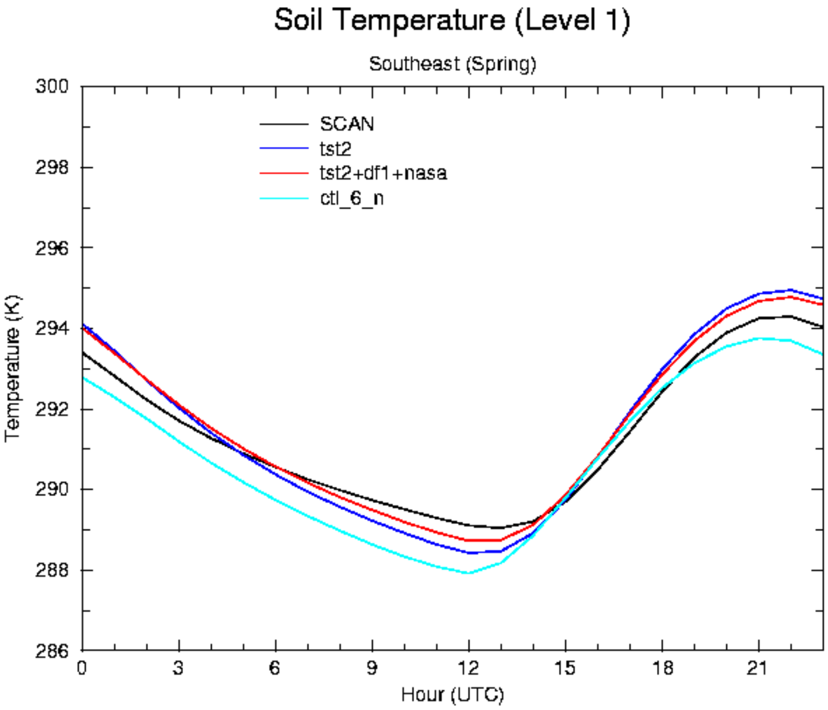

Figure 6: Spring 5 cm soil temperature forecasts for the SE region. The bias is removed with the czil algorithm. The MODIS data provides a slightly better match to the range of the diurnal cycle.

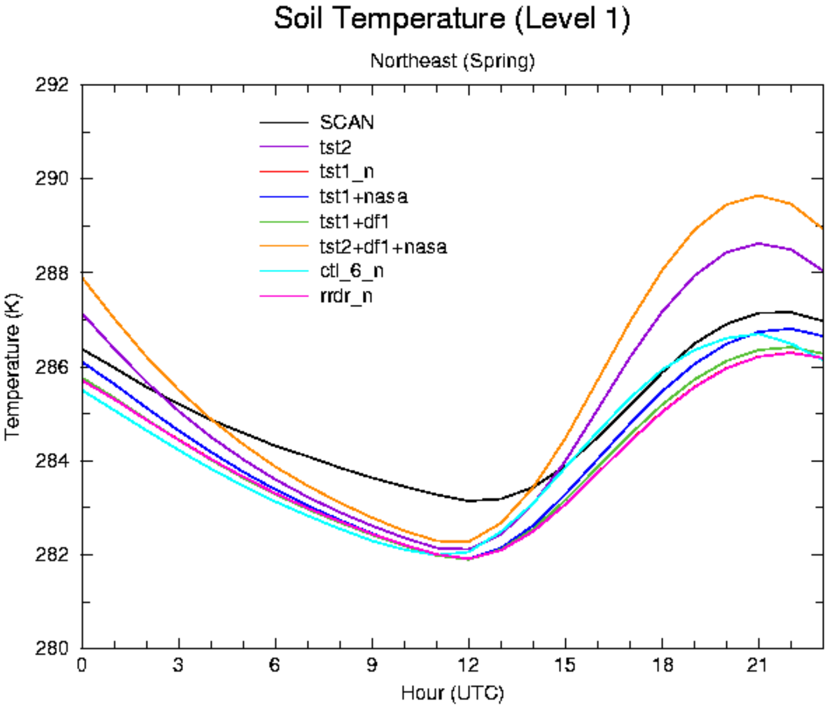

Figure 7: Spring 5 cm soil temperature forecasts for the NE region. Extra time series lines here should be ignored. They represent other variations in the use of MODIS data and czil algorithms. The lines for the czil upgrade (tst2) and MODIS data inclusion (tst2+df1+nasa) are different colors in this plot. Here, the diurnal range is better matched by the control 6-layer run. The czil algorithm and the MODIS data worsen the forecasts. The small observed range in the diurnal cycle may be due to snow coverage for part of the period.

To subjectively summarize the improvements by region and season, the Table 1 attempts to describe the improvements (relative to the 6-layer baseline) due to the czil algorithm and MODIS data incorporation. An estimate of the quality of the forecasts in each region and season is also provided. The subjective quality is based upon error statistics and how well the diurnal cycle was matched. The ++ symbol is intended to represent a much more significant improvement than a + symbol.

It should be noted that snow cover and frozen ground issues affect the NE, NW, and to some extent the SW (KS, OK) in spring. Otherwise, the forecasts are generally good or better.

To further address the question of the value of the MODIS data to HRLDAS soil forecasting, the sites were broken out not by region but rather by vegetation type. Interestingly, this ended up also being a rough regional breakdown. The sites were chosen to have similar land use and soil types. The first of the three areas was the low vegetation region west of the Mississippi River.). These plains sites were generally less vegetated than the Middle Region (E of the Mississippi River and N of Tennessee). The Southern Region (south of the Middle Region) contained the most vegetated sites. The 6-layer with climatological vegetation was used as the control run, and the MODIS data was then added to modify the initial state of the vegetation.

As the growing season progressed, as expected, the RMSEs grew larger as the insolation increased. The average errors aggregated over each regions' sites varied from 1.85°C in April to 2.46°C in June. The errors in the Western Region were by far the largest. This result has been seen before and is attributed to heat transfer issues in HRLDAS in less vegetated areas.

Figures 8-10 compare regional forecasts generated with and without the MODIS data for April, May, and June. During April, the MODIS data does not improve the forecast in any region. In fact, with the MODIS data, the forecasts are significantly worsened in the less vegetated regions. However, as the growing season evolves and more vegetation develops, the situation reverses and the addition of the MODIS data dramatically improves the forecast in the more vegetated regions. In the Western Region where there is still not as much vegetation even in June, the MODIS data does not improve the forecasts. In fact, they are still slightly worsened. However, even in this region, there seems to be a clear relationship between the amount of vegetation and the improvement provided by the MODIS data.

It seems fairly clear that the improvement provided by MODIS data for soil temperature forecasting is related to the amount (or perhaps to the height) of the vegetation present in the forecast area. That is, the MODIS data improves the soil forecasts in regions that are more heavily vegetated but does not improve forecasts in region of relatively sparse vegetation.

The differences using the MODIS data for soil moisture forecasting were less significant. The differences in errors were always less than 0.5%.

	Spring			Summer		
	Czil	NASA	Quality	Czil	NASA	Quality
NE	-	-	Poor	++	++	Good
NW	++	=	Poor	++	-	Excellent
SW	++	-	Fair	++	=	Good
SE	+	+	Excellent	+	+	Good

Table 1: Subjective description of the changes in the 5 cm soil temperature forecast skill due to the HRLDAS czil upgrade and the inclusion of NASA MODIS data are shown.

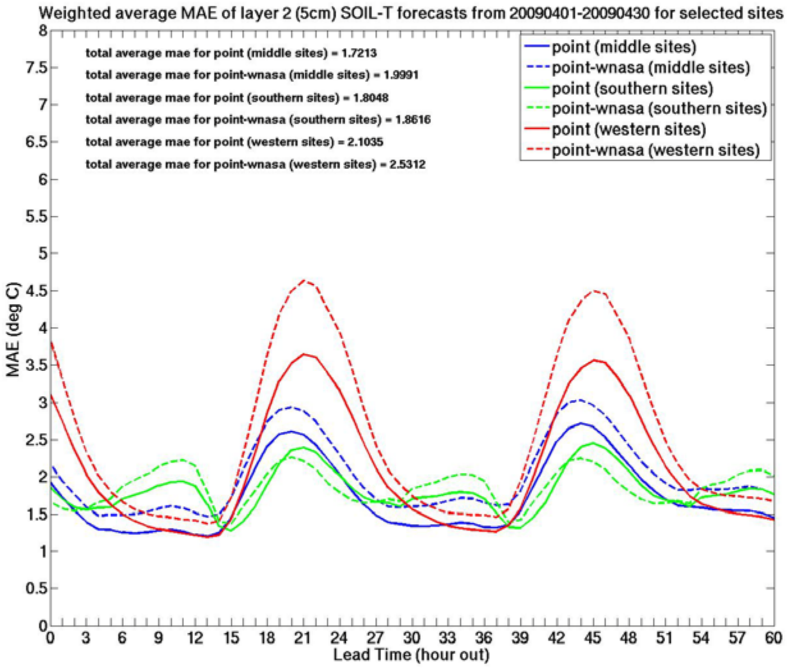

Figure 8: April 5 cm soil temperature forecast errors with and without the MODIS data are shown. A different colored line is used for each region. Dashed lines are used for the forecasts made with MODIS data. It can be seen that the MODIS data increases the errors significantly for the middle and especially the western sites. The errors are largest at night. There is also an interesting smaller peak in the errors near midday.

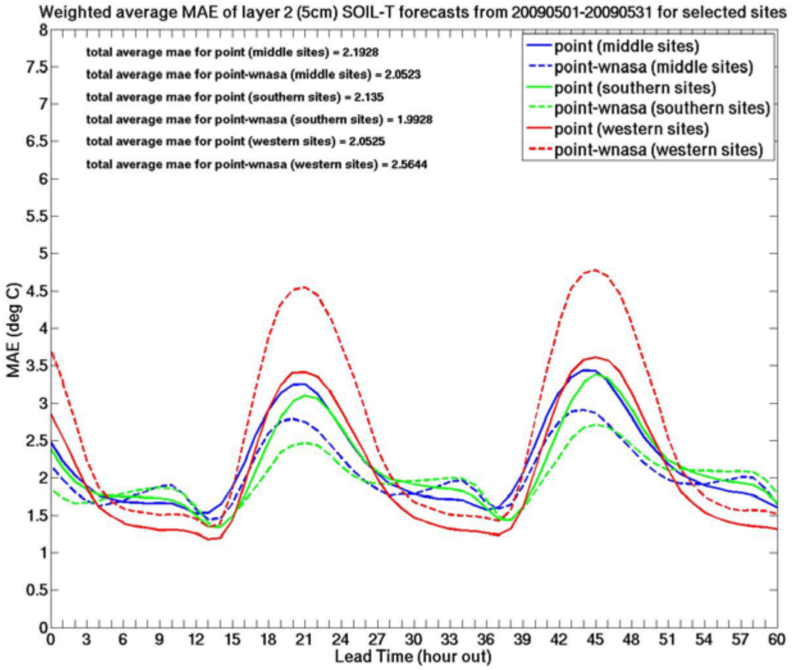

Figure 9: May 5 cm soil temperature forecast errors with and without the MODIS data are shown. With rapidly increasing vegetation in the more forested middle and southern sites, the MODIS data significantly reduce the peak errors. The less vegetated western sites still show larger errors with the MODIS data than without.

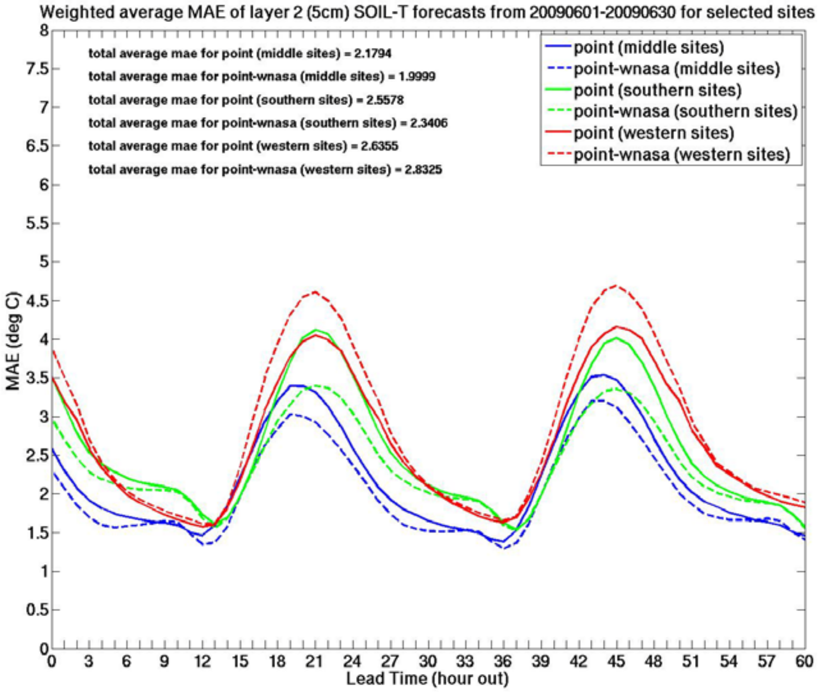

Figure 30: June 5 cm soil temperature forecast errors with and without the MODIS data are shown. The improvement with MODIS data in the increasingly vegetated middle and southern sites is further increased when compared to May. Now the improvement can be seen across the full diurnal cycle. The western sites still have larger errors with the MODIS data than without. However, with increased vegetation, the difference in errors is reduced.

4. DISCUSSION

While the results were encouraging, it became apparent that the forecast skill varied significantly between observational sites. This begs the question of how well soil forecasting can be done given the current parameter sets. For example, there are only 10 soil types and 24 land use types in the current HRLDAS system. Clearly, there is an infinite variety of each of these. It is unlikely that the parameterizations associated with any of these will exactly match the soil chemistry and plants growing there.

Furthermore, the resolution of the data sets used to initialize HRLDAS is seemingly unsatisfactory. It was recognized that the land use and soil type characteristics for each verification site were potentially incorrect. They had been taken from the HRLDAS (4.5 km) grid cell in which the site was located. The grid cells values had been populated from a 1-km USGS data set developed in the early 1990s. However, within each 4.5 km grid cell there ended up being a fair amount of variability. After extracting the sites' land use and soil type directly from the USGS grid, the land use and occasionally the soil type at a verification site often did not match the grid center's land use and soil type.

This led us to believe that if we corrected the soil-type and/or land-use at the mischaracterized sites that the modeled soil temperatures would improve. Running with climatological vegetation data (15km resolution), the overall verification showed little improvement in soil temperature error. This was somewhat disappointing. However, several sites did show significant improvement while a handful of sites were significantly worsened. We began by investigating the sites which performed worst (even when configured with "correct" land use and soil type data). One site, Rock Springs, PA was investigated in depth. This investigation brought up several important issues.

Using the HRLDAS grid cell containing the observation site, the Rock Springs site's land use was characterized as Deciduous Broadleaf Forest. The SCAN web pages contain a photo of this site. It is clearly not in the woods (Figure 11). Instead, it is in an open field with dense woods approximately 300m to the south. Google Earth was used to verify the lat-longs and that the photo was reasonable (see Figure

12). However, after setting the land use at the observing site as Dryland Cropland and Pasture, the modeled soil temperatures at the site were even worse with the correct land use type! We could only attribute this to a potential mismatch between the land use, and the climatological LAI and FPAR values used in this run. The LAI and FPAR values for this site came from a 15km climatological data set and were consistent with a heavily wooded area. Apparently, the model felt that the crops were incredibly high and dense. We decided that this mismatch would probably be corrected by using the MODIS 1-km LAI and FPAR.

The run using the Aqua and Terra MODIS data gave slightly better results overall (for all the mischaracterized sites combined) than either of the previous runs. At Rock Springs however, the results fell between the original run and the run with the "correct" land use type. Upon closer inspection, the MODIS LAI and FPAR data seemed to be dominated by the nearby woods. That is, the LAI/FPAR values were similar to the climatological values, and very different than the LAI/FPAR values for adjacent non-forested grid cells (see Figure 12).

The high LAI and FPAR values seemed to extend well beyond the forested region. This led us to question the calibration of the MODIS data. It should also be noted that the Terra and Aqua LAI and FPAR values differed somewhat over the same region (although this was not the case for our experiment shown in Figure 12). Eventually, by "moving" the observation site about 1km to the north and 1km west, we were able to run HRLDAS with LAI and FPAR values consistent with cropland. The error dropped significantly and was well below the error obtained from any other configuration. This would indicate that having the correct land-use type along with the correct LAI and FPAR data does improve the model's performance.

Afterwards, we examined other observation sites. Several were classified as Cropland/Woodland Mosaic. The SCAN web site photos confirmed this. Although all the sites were in the cropland, trees could be seen from approximately 20m to 100m from the measurement stations. This indicates that the 1-km MODIS data may not have sufficient resolution to accurately describe the vegetation at a specific point, especially in areas of rapidly changing land use. Higher resolution remotely-sensed data would help greatly with this issue. 100m resolution would probably be satisfactory.

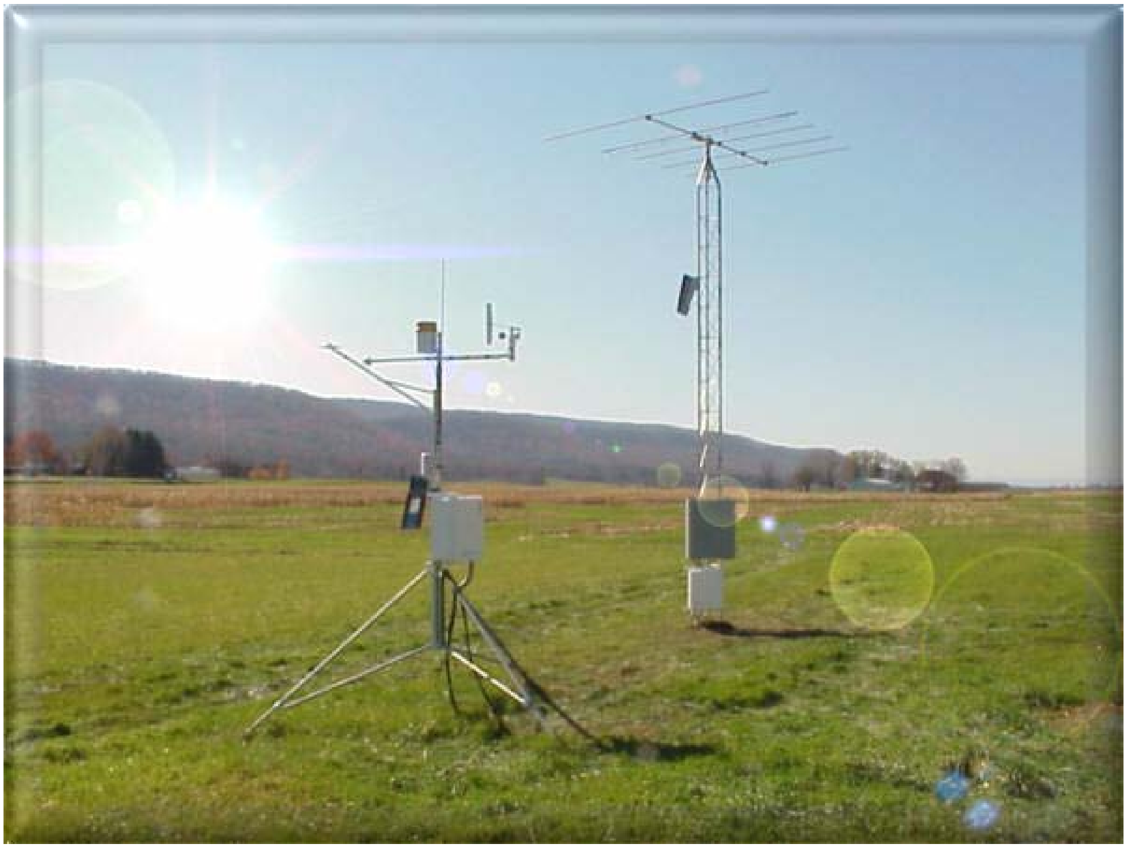

Figure 11: The Rock Springs observing site clearly lies in cropland rather than woodland.

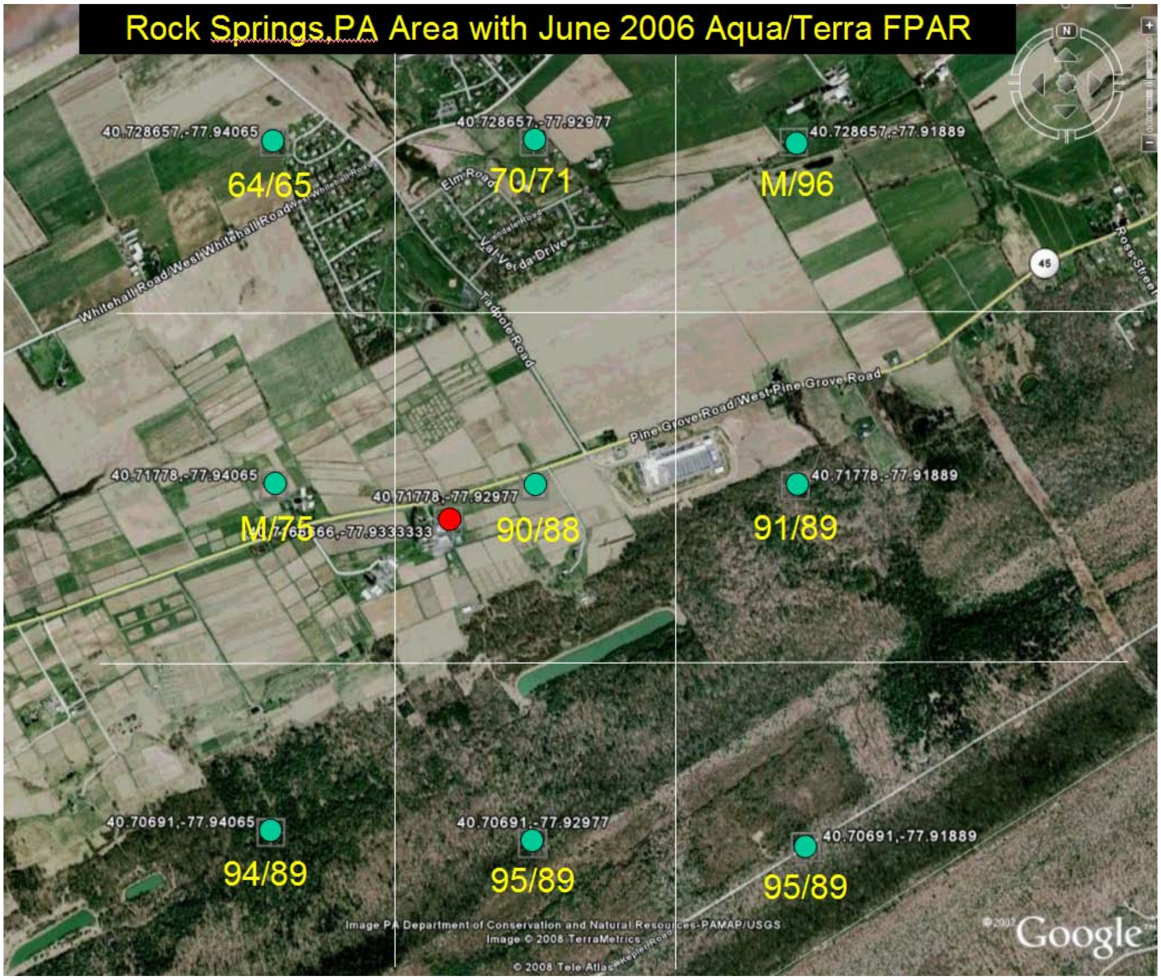

Figure 12: MODIS FPAR values from mid-June 2006 near the Rock Springs, PA soil observation site (red dot). The MODIS FPAR values from Aqua and Terra are provided for each 1-km grid cell (green dots at cell center). The grid cell in which the observation site lies has FPAR values similar to the more wooded areas to south. However, the observation site lies in cropland which should have a FPAR value more like the grid cell to the northwest. Is this a registration or resolution issue?

While this is a little discouraging, it should be noted that many sites' modeled soil temperatures were significantly improved by using the higher resolution USGS Land Use and Soil Type data as well as the MODIS data. Without any manipulation, i.e. "moving" the sites, the overall errors (for all sites) were decreased. Ultimately, a definitive answer as to whether or not improvements using higher resolution data leads to improved modeled soil temperatures will have to be based on statistics over a large numbers of sites. Although there is much variability between sites, the overall errors are decreased using the high resolution data despite the aforementioned issues.

5. CONNECTING WITH THE END USER

Towards the end of the project NCAR worked with DTN to determine how the soil forecast data could be used in a DSS to present meaningful information to the end users. It became clear that deterministic soil temperature and moisture values alone were not very valuable to the end user. Based on feedback, it was determined that the trend in soil temperature and moisture was most important in making decisions such as when to plant and when to harvest. NCAR prototyped some software to come up with forecast, observed, and climatological trends. Figures 13 and 14 are examples of how this trend data could be displayed. Note that in these two examples the departure from climatology in HRLDAS predictions generally agrees with that in observations.

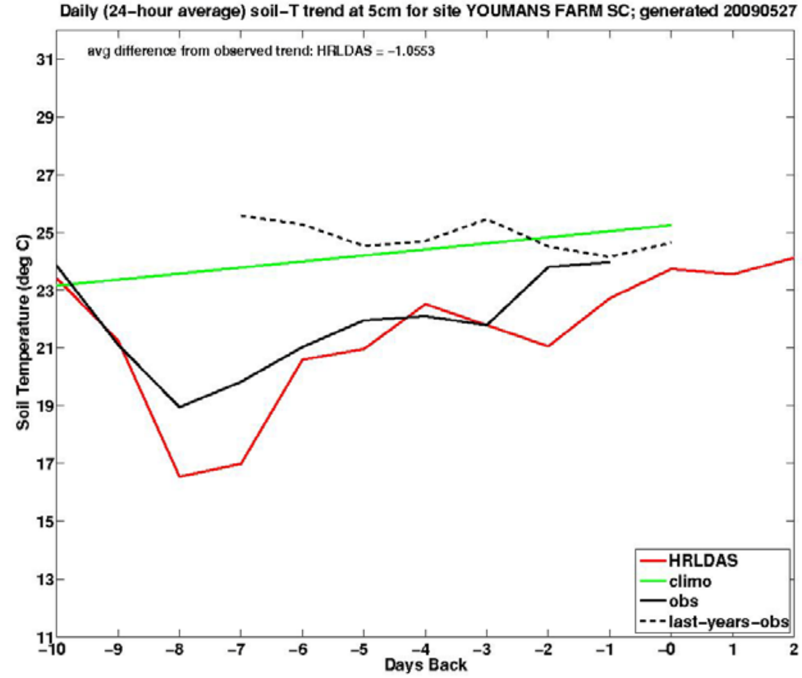

Figure 13: Initial efforts at communicating soil temperature conditions to the end user. The user can see normal conditions (climo) compared to the forecasts and observations. Last year's observations from this time are also shown.

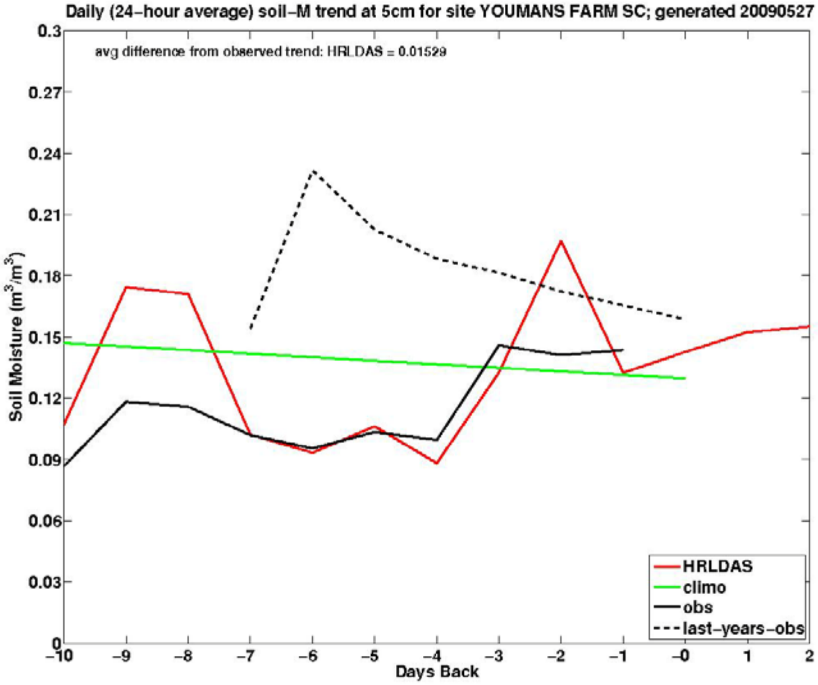

Figure 14: Same as figure 13, but showing soil moisture.

6. CONCLUSIONS

Land Surface Models, forced by weather forecasts, can produce reasonable soil temperature forecasts. Surface heat transfer algorithm upgrades do improve the soil temperature forecast. However, this is an ongoing research effort as there is still room for significant improvement in the heat exchange calculations.

The use of MODIS vegetation data for LSM initialization (instead of climatological LAI) can significantly reduce the soil temperature forecast errors for sites in which the MODIS values are reasonable. The error reductions are highly correlated to the amount of vegetation present at any location and time. The MODIS data reduces forecast errors much more significantly in areas that are heavily vegetated but has little impact in drier areas with sparse vegetation.

The forecast performance varies from site to site depending on how well the land use, soil type, and MODIS data in HRLDAS match specific observation sites. For many of the sites, the resolution of these data sets is too coarse to correctly classify the real characteristics and subsequently the forecasts are not as good for these sites.

While it is not easily verifiable, we believe that even in mixed land use areas the HRLDAS soil forecasts provide a good approximation of the average soil temperature and moisture over the entire grid cell and this data could be used effectively in a real-time decision support system. Much of the difficulties in verification arise because the observation site may not be typical of its surroundings. A much more involved project will be required to investigate whether this hypothesis bears true.

7. ACKNOWLEDGEMENTS

This research is sponsored by the NASA Earth Observing System, in response to NRA #NNH05ZDA001N, Research Opportunities in Space and Earth Sciences (ROSES-2005). The material within is based upon work supported by NASA under award No. NNA06CN03A titled “A Soil Temperature and Moisture Decision Support System for Agriculture”. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Aeronautics and Space Administration.

This project would like to acknowledge the data sets and help provided by the USDA Natural Resource Conservation Service (NRCS) who manage the SCAN data sets, the Oklahoma Mesonet, and various NASA labs who organize the MODIS data.

8. REFERENCES

Chen,F., K. Manning, M. LeMone, S. Trier, J. Alfieri, R. Roberts, M. Tewari, D. Niyogi, T. Horst, S. Oncley, J. Basara, and P. Blanken, 2005: Description and Evaluation of the Characteristics of the NCAR High-Resolution Land Data Assimilation System, Vol.46, pp. 694-713

Trier, S., F. Chen, and K. Manning, 2004: A Study of convection initiation in a mesoscale model using high-resolution land surface initiation conditions. *Monthly Weather Review,* Vol.132, pp. 2954-2976

Schaefer, G., M. Cosh, and T. Jackson, 2007: The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN), Journal of Atmospheric and Oceanic Technology, Vol.24, pp. 2073-2077