AMSR-E Surface Temperature Retrievals: The Potential for LSM Data Assimilation

Gouweleeuw, B.T.[1], M. Owe[2] and T. Holmes[3]

[1] CSIRO Land and Water, Canberra ACT 2601, Australia.
[2] NASA/GSFC, Greenbelt, 20771 MD, USA
[3] Vrije Universiteit, Amsterdam 1081 HV, The Netherlands
Email: bingo@hsb.gsfc.nasa.gov

Keywords: *Surface temperature, passive microwave imagery, Land Surface Model, data assimilation.*

EXTENDED ABSTRACT

Improved accuracy in defining initial conditions for fully-coupled Numerical Weather Prediction models (NWPs) along with continuous internal bias corrections for baseline data generated by uncoupled Land Surface Models (LSMs) is expected to lead to improved short-term to long-range weather forecasting capability. Because land surface parameters are highly integrated states, errors in land surface forcing, model physics and parameterization tend to accumulate in the land surface stores of these models, such as soil moisture and surface temperature. This has a direct effect on the models' water and energy balance calculations, and may eventually result in inaccurate weather predictions.

For the Oklahoma Mesonet data base surface temperature estimates obtained with a recently improved retrieval algorithm from the Advanced Microwave Scanning Radiometer (AMSR) on board NASA's Earth Observing System (EOS) Aqua satellite are evaluated against different combinations of model output of the Community Noah Land Surface Model and Community Land Model (CLM2) operated within NASA/GSFC's Land Information System (LIS) and atmospheric forcing data of a variety of sources, i.e. the NCEP Global Data Assimilation System (GDAS), the European Centre for Medium-Range Weather Forecast (ECMWF) and the North American Data Assimilation System (NLDAS), based on Eta data and supplemented with observation-based precipitation and radiation data. The surface temperature retrievals and LSM output are further evaluated against station measurements from the Mesonet observational grid in Oklahoma.

Preliminary analysis presented here shows the satellite derived surface temperature estimates - uncorrected for bias - are not necessarily superior to the LSM simulations, evaluated against the Mesonet observational grid benchmark. In general, data assimilation systems take into account observational errors and are able, despite errors in the observations, to obtain improvement of LSM results, as long as the temporal trends are well represented. Further, most assimilation systems use a bias removal prior to actual assimilation. Here, a simple (linear) correction decreases the AMSR T_s error beyond the error of simulated T_s. Therefore, it will be interesting to see how the satellite-derived surface temperature will behave in an assimilation scheme in a follow-up study.

1. INTRODUCTION

Surface temperature is a key parameter in many energy balance applications such as evaporation modeling, climate models, and radiative transfer modeling. Ground observations are generally useful for local applications, however, they are highly intensive in man-power and equipment costs. Furthermore, ground observations of surface temperature are point measurements and since variability can be high, especially in regions with discontinuous vegetation, scaling up to spatial averages is often difficult.

The most common remote sensing method for surface temperature observation is thermal infrared (TIR). The MODIS Terra satellite provides daily global cover of land surface temperature on a 1km resolution. However, TIR is affected by aerosols, particulates and other contaminants, usually requiring some sort of atmospheric correction. The existence of cloud cover will usually render TIR observations unusable.

Higher frequency microwave emissions at vertical polarization possess a strong physical relationship with the thermodynamic temperature of the emitting surface. Microwave sensors are less affected by atmospheric conditions. Therefore, they have potential to provide reliable estimates of averaged surface temperature with a near-all-weather capability on a scale and coverage compatible with NWPs.

A recently developed and further improved theoretically-based land surface parameter retrieval model (Owe et al., 2001; De Jeu et al., 2003; Owe et al., 2005) has demonstrated significant potential for providing independent measurements of land surface parameters, e.g. surface soil moisture and surface temperature. It has enabled the construction of a continuous historical global database of satellite derived land surface parameters from 1978 through to the present, developed from Nimbus-SMMR, DSMP-SSM/I, TRMM-TMI, and AQUA-AMSR microwave brightness temperature measurements. Satellite retrievals of these parameters from this database may be combined with modeled and observational data in a data assimilation scheme in order to generate the best possible data fields. These data may then serve for initialization and continuous bias correction for NWP models.

Data assimilation is the process of finding the model representation which is most consistent with the observations (Lorenc, 1995). In essence, data assimilation merges a range of diverse data fields with a model prediction to provide that model with the best estimate of the current state of the natural environment so that it can then make more accurate predictions. A number of options for data assimilation are currently being implemented and tested within the Land Information System (LIS) developed at NASA Goddard Space Flight Center (see below), and will soon be made available.

Here, as a prequel to such more sophisticated data assimilation efforts and to tentatively asses its feasibility, the LSM surface temperature output is simply compared to AMSR-E retrievals and station data.

2. DATA SETS

2.1. Model Simulations

The Land Information System (LIS) developed at NASA Goddard Space Flight Center is an interoperable platform capable of integrating the use of land surface models, data management techniques and high performance computing (Kumar et al., 2006).
The community Noah land surface model (Ek et al., 2003) and the Community Land Model, version 2:0 (CLM2) (Dai et al, 2002; Zeng et al., 2002), are two of the LSMs currently supported by LIS. Both are stand- alone, 1-D models, which are freely available: Noah from the National Centers for Environmental Prediction (NCEP) and CLM2 from The National Center for Atmospheric Research (NCAR). The LSMs can be executed in either coupled or uncoupled mode. In uncoupled mode, as applied in the present study, near-surface atmospheric forcing data is required as input. Here, forcing data from the NCEP Global Data Assimilation System (GDAS), the European Centre for Medium-Range Weather Forecasts (ECMWF) and the North American Data Assimilation System (NLDAS), based on Eta data and supplemented with observation-based precipitation and radiation data (Cosgrove et al, 2003), are used for the year under consideration, i.e. 2003, thus overlapping the AMSR-E lifetime (2002-present). The forcing data are: large scale precipitation, convective precipitation, specific humidity, surface pressure, downward solar radiation, downward thermal radiation, air temperature, and wind velocity. The temporal resolution of the NLDAS forcing time is an hour, while GDAS and ECMWF have a 3 hr time step.

The LSMs simulate a range of water- and energy balance variables, of which surface (skin) temperature is of most interest for the present analysis. The models apply finite difference spatial discretization methods and (semi-)implicit time-integration schemes to numerically integrate the governing equations of the physical processes of the soil-vegetation-snow pack medium, including the surface energy balance equation, the Richards' equation (1931) for soil hydraulics, the diffusion equation for soil heat transfer, the energy-mass balance equation for the snow pack, and equations for the conductance of canopy transpiration.

2.2. Observed Data

A data set of near-surface temperature (2-3 mm) derived from the 37 GHz microwave signal from the AMSR instrument on board the EOS Aqua satellite for the full year of 2003 is made available by Owe et al. (2005). The 37 GHz AMSR-E footprint is an oval of 10 km square (sampling interval: 10 km), where the derived surface temperature fields are resampled in a 0.25 degree grid. A subset covering the state of Oklahoma, USA is cut from the global dataset. The choice for this location is motivated by the presence of sets of observational data for the corresponding period of time (i.e. the year 2003), made available by the Oklahoma Mesonet (Brock et al., 1995).

Observed data is collected from 11 stations located within three 0.5 degree grids (Fig. 1), (spatially) representative of different types of land cover. The 5 cm profile temperature measurements are made available every half hour and are extrapolated to 2 mm depth temperature estimates using a soil heat transfer algorithm developed by Owe et al., 2005. With regard to the AMSR-E antenna, there are two types of measurements, a night-time (local time ~ 01:30, GMT ~ 08:30h) and a day-time (local time ~ 13:30, GMT ~ 20:30h).

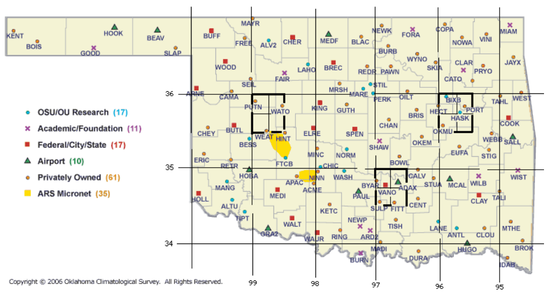

Figure 1. The Oklahoma Mesonet. Observed data are taken from stations located within the three boxed 0.5 grids.

3. ANALYSIS

3.1. Satellite derived surface temperature

Fig. 2 shows an example of the Oklahoma AMSR-E surface temperature (T_s) retrievals at night time (a) and day time (b), together with the three 0.5 degree grids within which observational data are sampled.

Fig. 3 shows scatter plots of observed and satellite retrieved T_s for two of the selected observational sites of the Oklahoma Mesonet**.** The outliers tagged with a blue date are identified as images containing precipitating clouds, resulting in an underestimated satellite derived T_s of over 15 (K). At higher surface temperatures, cloud contamination is less detectable with smaller T_s differences (Fig. 3, right panel). Most of the cloud contaminating conditions occur at day time in summer, when convection is strong. While this eliminates about 1-3% of the data set, frozen soil conditions in winter take out the bulk of the data (over 30%). These phenomena seem to put some emphasis on 'near' in the assessment of the passive microwave retrieval of T_s as a 'near-all-weather' technique. Consequently, this also applies to the passive microwave soil moisture algorithm retrieval, in case the passive microwave T_s estimate is used.

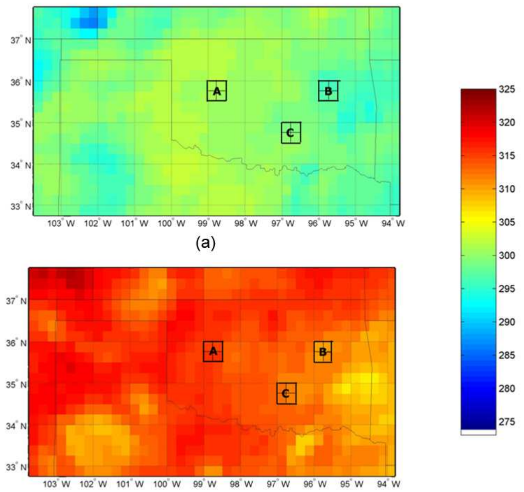

Figure 2. AMSR-E surface temperature retrievals on (a) August 22, 2003, 8:20 AM UCT (b) August 25, 2003, 20:03 UCT. Key, A = Grassland, B = Cropland, C = Wooded Grassland.

Table 1 shows variation of RMSE between the three 0.5 degree grid cells and also from station to station. Further, satellite derived T_s does not necessarily compare better to T_s 2 mm, which was modeled from the observed T_s 5 cm using a soil heat transfer algorithm (Owe et al., 2005). Despite the fact the observed point data only give us an estimate of the 'true' integrated grid T_s at best, it may indicate the heat transfer algorithm needs some further fine tuning. In all, the average difference between observed and satellite derived T_s for our (limited size) data set is over 3 (K). This value exceeds the 2 degree soil temperature threshold, which was set for microwave space based missions in order to achieve a 4 vol. %

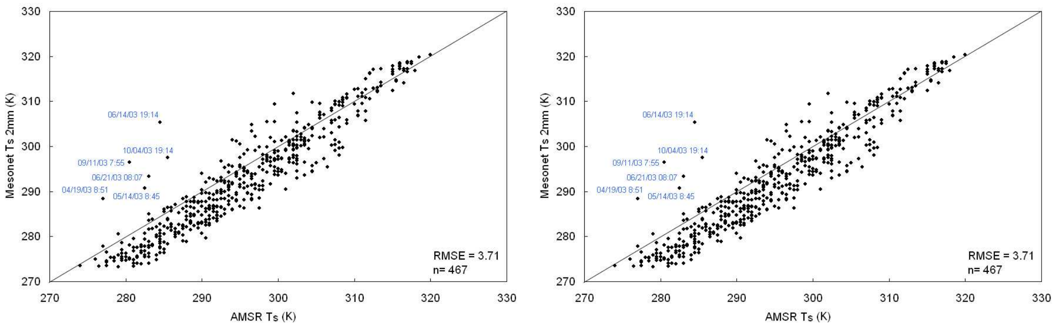

Figure 3. Observed T_s vs. AMSR-E T_s for stations WATO (left) and BYAR (right), uncorrected for bias.

precision (or less) in soil moisture retrieval (Entekhabi et al., 2004). In an attempt to remove systematic bias in the data sets, a simple linear correction was carried out. AMSR generally overestimates the Mesonet observation, mainly in the lower half temperature range (see Fig. 3). Bias, however, is not constant, neither over the temperature range (i.e. time dependent), nor for the individual stations (i.e. place dependent). Although the correction removes close to 1 (K) error on the average, RMSE remains above the 2 degree threshold.

Table 1. RMSE of observed and satellite derived T_s for selected Mesonet stations.

Grid	Station	RMSE Ts 5cm (K) bias correction No	RMSE Ts 5cm (K) bias correction Yes	n	RMSE Ts 2mm (K) bias correction No	RMSE Ts 2mm (K) bias correction Yes	n
A	PUTN	3.43	2.29	439	4.67	3.50	404
	WATO	3.76	2.49	493	3.71	2.62	467
	WEAT	3.73	2.30	467	4.16	3.12	430
	HINT	4.12	2.53	472	3.82	2.51	457
B	BIXB	2.67	2.62	467	4.94	3.73	434
	HECT	2.57	2.52	480	2.87	2.60	471
	PORT	2.60	2.51	486	2.53	2.79	473
	OKMU	2.58	2.46	482	3.69	3.08	463
	HASK	3.24	3.20	499	2.75	2.69	492
C	BYAR	2.91	2.61	489	3.25	2.80	467
	VANO	4.82	2.46	452	5.20	3.00	422
Mean		3.31	2.54		3.87	2.95	

3.2. Simulated surface temperature

Fig. 4 shows the surface temperature at AMSR-E night overpass time simulated by the different LSM and atmospheric forcing combinations, six in total. Fig. 5 is identical, but at day time. At first glance, the simulations look similar, indicating corresponding cooler and warmer surface areas, both for the day and night time overpass. The cooler areas, dynamic in space and time, likely relate to (convective) clouds, again indicating the 37 GHz antenna's weather dependence. The cooler areas are less prevalent or absent in the NLDAS forced simulations at day time, underscoring its deviating temporal and spatial properties. Conversely, the similarity of the GDAS and ECMWF forced simulation underlines the close correspondence of these atmospheric forcing data sets.

Fig. 6 shows the scatter plots of observed and simulated T_s for the identical stations as in Fig. 3 for the LSM-atmospheric forcing combination with the smallest RSME. For these two cases this is the CLM2 model with NLDAS forcing. Table 2 contains all 11 observational stations and shows the CLM2-NLDAS combination on the average compares best with an RMSE of just above 3 (K), uncorrected for bias. Some (linear) bias is present and RMSE shows improvement after correction, although less than in the AMSR data. Further, the number of eliminated data (sub zero) is limited and constant compared to the AMSR T_s data set, as the sample size (n) in Fig. 6 indicates. In all, the RMSE of the simulated and satellite derived data set, as evaluated against the observed T_s data set benchmark, is comparable. However, the satellite derived T_s does not necessarily perform better than the simulated T_s. In fact, for this particular analysis before bias correction, it compares slightly worse to the observed data. In general, data assimilation systems take into account observational errors and are able, despite errors in the observations, to obtain improvement of LSM results, as long as the temporal trends are well represented. Further, most assimilation systems use a bias removal prior to actual assimilation. A simple (linear) correction carried out here decreases the AMSR T_s RMSE beyond that of the simulated T_s. Therefore, it will be interesting to see how the satellite-derived surface temperature will behave in an assimilation scheme in a follow-up study.

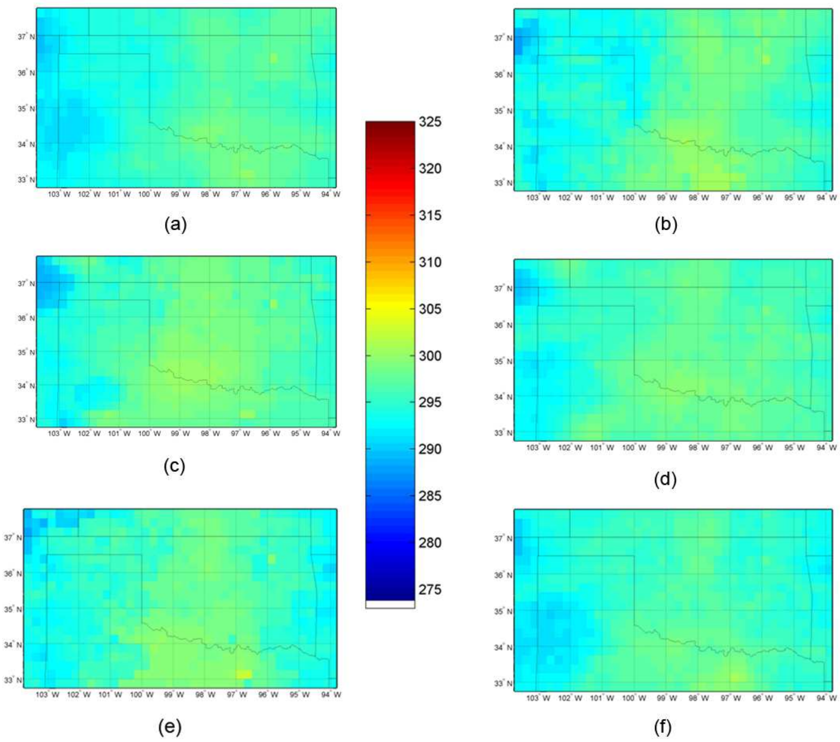

Figure 4. Simulated T_s on August 22, 2003 at the AMSR-E night time overpass for the Oklahoma domain by six combinations of LSM and atmospheric forcing (a) CLM2-ECMWF (b) Noah-ECMWF (c) CLM2-GDAS (d) Noah-GDAS (e) CLM2-NLDAS (f) Noah-NLDAS.

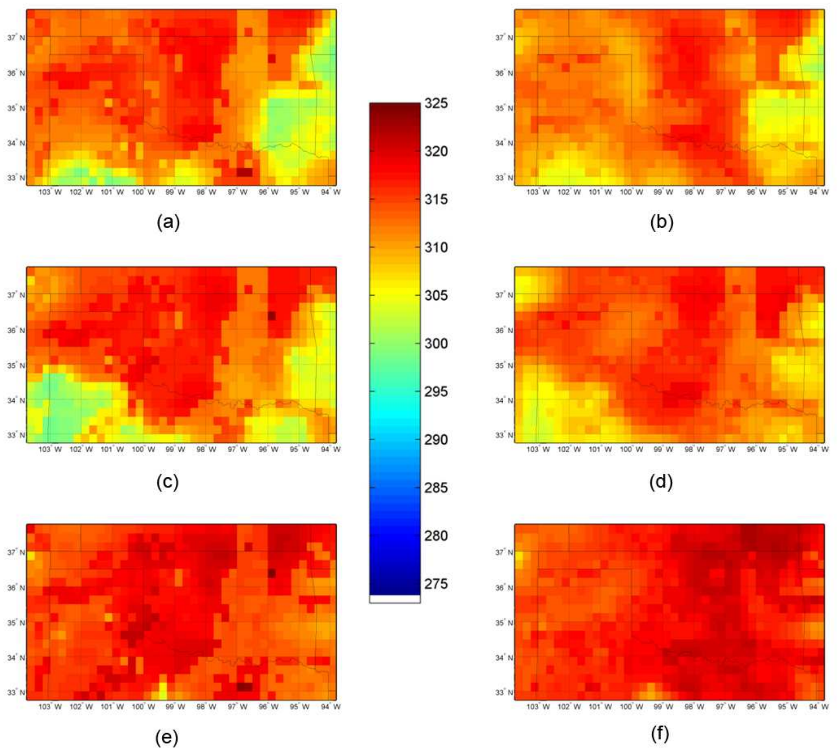

Figure 5. Simulated T_s on August 25, 2003 at the AMSR-E day time overpass for the Oklahoma domain by six combinations of LSM and atmospheric forcing (a) CLM2-ECMWF (b) Noah-ECMWF (c) CLM2-GDAS (d) Noah-GDAS (e) CLM2-NLDAS (f) Noah-NLDAS.

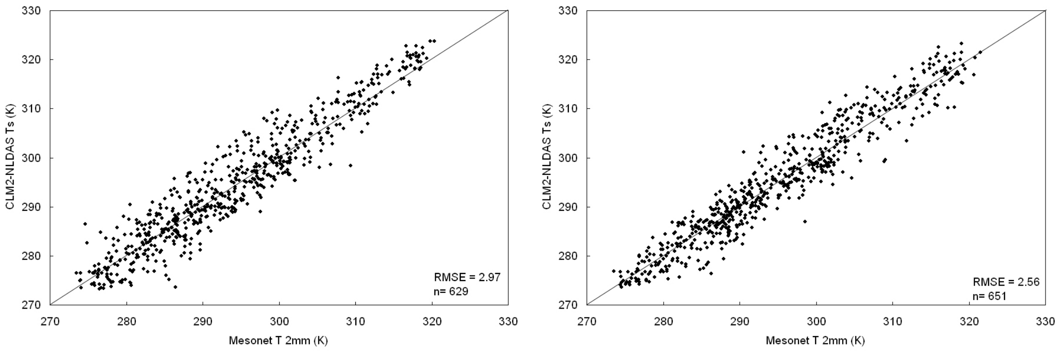

Figure 6 Observed T_s vs. simulated T_s for stations WATO (left) and BYAR (right), uncorrected for bias.

Table 2 RMSE of observed T_s 2mm and simulated T_s for selected Mesonet stations.

Grid	Station	RMSE (K) CLM2 ECMWF		NOAH ECMWF		CLM2 GDAS		NOAH GDAS		CLM2 NLDAS		NOAH NLDAS	
		bias correction											
		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
A	PUTN	3.96	3.50	3.09	3.09	3.78	3.77	3.22	3.20	2.83	2.80	2.96	2.80
	WATO	3.33	3.31	3.41	3.14	3.34	3.33	3.26	3.15	2.97	2.73	3.65	2.84
	WEAT	3.87	3.43	3.28	3.15	3.67	3.64	3.38	3.34	3.13	2.89	3.48	2.93
	HINT	3.53	3.39	3.66	3.22	3.41	3.39	3.43	3.21	3.10	2.70	3.86	2.86
B	BIXB	3.92	3.85	3.31	3.31	4.08	3.92	3.44	3.44	3.52	3.49	3.20	3.17
	HECT	3.12	3.02	3.21	2.86	2.74	2.71	3.14	2.75	3.08	2.71	3.87	2.84
	PORT	3.16	3.09	2.94	2.85	2.89	2.85	2.86	2.73	3.01	2.79	3.39	2.82
	OKMU	3.46	3.36	3.23	2.96	3.41	3.31	3.23	3.07	3.04	2.95	2.73	2.72
	HASK	3.45	3.35	3.36	3.19	3.11	3.07	3.36	3.12	3.46	3.22	4.57	3.41
C	BYAR	2.90	2.87	2.84	2.75	2.87	2.87	2.78	2.74	2.56	2.47	2.60	2.52
	VANO	4.08	3.34	3.29	2.67	4.53	3.50	3.29	2.84	3.32	2.91	3.56	2.58
Mean		3.53	3.32	3.24	3.02	3.44	3.40	3.22	3.05	3.09	2.88	3.44	2.86

3.3. Satellite derived and simulated surface temperature

In perspective of the tentative assessment of the feasibility of surface temperature data assimilation, the AMSR-E surface temperature retrievals are next compared to LSM output fields over the entire Oklahoma domain, as depicted in Fig 2. The Root Mean Square Difference (RMSD, rather than the Error, since the Mesonet data are used as benchmark) is computed in three ways: (1) without threshold (2) eliminating all subzero data and data with RMSD > 10 (K) (3) eliminating all subzero data and data with RMSD > 15 (K).

Table 3 shows that the CLM2-GDAS combination compares best to the satellite derived T_s over the entire Oklahoma domain with an RMSD of just over 3 (K) when applying the most restrictive filter (2). This seems consistent with the evaluation against observed data, be it that the NLDAS forced simulations compare worse here. This is possibly explained by the averaging temporal and spatial resolution effect of the

Table 3 RMSD of satellite derived and simulated T_s over the Oklahoma domain.

LSM Forcing	RMSD	n	Ts>0(K) RMSD>10	n (%)	Ts>0(K) RMSD>15	n (%)
CLM2 ECMWF	3.81	401919	3.24	88.6	3.39	90.2
NOAH ECMWF	3.98	402161	3.55	88.0	3.70	91.5
CLM2 GDAS	3.58	400490	3.11	90.2	3.26	91.7
NOAH GDAS	3.67	400507	3.25	87.5	3.40	91.2
CLM2 NLDAS	4.11	401149	3.45	82.6	3.63	86.5
NOAH NLDAS	4.47	401936	3.85	84.6	4.15	90.4

GDAS and ECMWF data. Further, the most rigorous filter (2) eliminates about 10-20% of the data, while threshold (3) eliminates about 4% less. This again implies the frozen soil condition, as in the observed data evaluation, cancels the bulk of the data. The total percentage of data eliminated, however, is lower here, indicating the selection of data to be cancelled by hand in the observed data evaluation was somewhat less forgiving than then the automated procedure. As in the observed data evaluation, the minimum difference between simulated and satellite derived T_s is over 3 (K), which again exceeds the 2 degree temperature threshold, which was set for microwave space based missions in order to achieve a 4 vol. % precision (or less) in soil moisture retrieval (Entekhabi et al., 2004). In all, the variation of RMSE between the various LSM-atmospheric forcing combinations is low, apart from the ones forced with NLDAS. This seems to indicate some temporal and spatial constraints on the assimilation of passive microwave T_s into LSMs.

4. CONCLUSION

Evaluation of different data sources of surface temperature indicates that satellite derived passive microwave Ts is not necessarily a superior estimate compared to simulated T_s, if evaluated against a data set of observed point measurements. In general, data assimilation systems take into account observational errors and are able, despite errors in the observations, to obtain improvement of LSM results, as long as the temporal trends are well represented. Further, most assimilation systems use a bias removal prior to actual assimilation. Here, a simple (linear) correction decreases the AMSR T_s error beyond that of the simulated T_s. Therefore, it will be interesting to see how the AMSR derived surface temperature will behave in an assimilation scheme in a follow-up study.

It should be pointed out that the atmospheric forcing data sets used in the present study are mainly reanalysis data making use of in-situ observations. As a consequence, the model simulations over Oklahoma with a dense observation network are relatively accurate. However, no or little observation data are available for most areas and the quality of the forcing data - and the model simulations - decreases. Hence, in these areas more scope is present for remote sensing data to constrain these models.

A further consideration is that the retrieval of passive microwave satellite derived surface temperature is hampered by weather conditions: frozen soil conditions in winter and (precipitating) clouds in summer. These phenomena appear to put some emphasis on 'near' in the assessment of the passive microwave retrieval of T_s as a 'near-all-weather' technique. Consequently, this also applies to the passive microwave soil moisture algorithm retrieval, in case the passive microwave T_s estimate is used.

REFERENCES

Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson and M. D. Eilts. 1995: The Oklahoma Mesonet: A Technical Overview. *Journal of Atmospheric and Oceanic Technology*, **12**, 5-19.

Cosgrove, B. A., Lohmann D., Mitchell, K. E., Houser, P. R., Wood, Schaake, J. C., E. F., Robock, A., Marshall, C., Sheffield, J., Duan, Q., Luo, L., Higgins, R. W., Pinker, R. T. Tarplay, J. D., and Meng, J., 2003: Real-time and retrospective forcing in the North American Land Data Assimilation Sytsem (NLDAS) project, J. of Geoph. Res., **108**, D22, 8842-8854.

Dai, Y., X. Zeng, R.E. Dickinson, I. Baker, G. Bonan, M. Bosilovich, S. Denning, P. Dirmeyer, P. Houser, g. Niu, K. Oleson, A. Schlosser and Z.-L. yang, 2002, The Common Land Model (CLM), Bull. Amer. Meteor. Soc., **84(4)**, 1013-1023.

De Jeu, R.A.M. and Owe. M, 2003: Further validation of a new methodology for surface moisture and vegetation optical depth retrieval, *International Journal of Remote Sensing*, **24** (22), 4559-4578(20).

Ek, M. B., K.E. Mitchell, Y. Lin, E. rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land-surface model advances in the NCEP operational mesoscale mesoscale Eta model, J. Geophys. Res. 108, D22, 8851, doi:10.1029/2002JD003296

Entekhabi, D., E. Njoku, P. Houser, M. Spencer, T. Doiron, J. Smith, R. Girard, S. Belair, W. Crow, T. Jackson, Y. Kerr, J. Kimball, R. Koster, K. McDonald, P. O'Neill, T. Pulz, S. Running, J. Shi, E. Wood, and J. Van Zyl, 2004: The Hydrosphere State (HYDROS) Mission concept: An Earth system pathfinder for global mapping of soil moisture and land freez/thaw, *IEEE Trans. Geosci. Rem. Sens.*, **42**(10), 2184-2195.

Kumar, S. V., C. D. Peters-Lidard, Y. Tian, P. R. Houser, J. Geiger, S. Olden, L. Lighty, J. L. Eastman, B. Doty, P. Dirmeyer, J. Adams, K. Mitchell, E. F. Wood and J. Sheffield, 2006. Land Information System - An Interoperable Framework for High Resolution Land Surface Modeling. *Environmental Modelling & Software,* **21**, 1402-1415.

Lorence, A.C., 1995: Atmospheric Data Assimilation, Meteorological Office Forecasting Research Division Scientific Paper, **34**.

Owe, M., De Jeu, R.A.M., and Walker, J.P., 2001: A methodology of surface soil moisture and vegetation optical depth retrieval using the microwave polarization index, *IEEE Transactions on Geoscience and Remote Sensing*, **39**, 1643-1654.

Owe, M., T. Holmes, R. De Jeu, 2005: A physically based model with remote sensing inputs for improved soil temperature retrievals", In: *Proceedings, International Symposium on Remote Sensing*, **5976**, International Society for Optical Engineering, Bellingham, WA.

Richards, L. A., 1931: Capillary conduction of liquids in porous media, *Physics*, **1**, 318-333.

Zeng, X., m. Shaikh, Y. Dai, R.E. Dickinson and R. Myneni, 202, Coupling of the Common Land Model to the NCAr Community Land Model, J. Climate, **15**, 1832-1854.