Cross-Validation of AMSR-E Soil Moisture Retrievals and the Surface Soil Moisture Simulations from NASA's Global Land Data Assimilation System (GLDAS)

Alok Kumar Sahoo, George Mason University, Fairfax, VA
(asahoo@gmu.edu)
Mentors: **Xiwu Zhan*, Matthew Rodell** & Kristi Arsenault***
(*UMBC-GEST/NASA-GSFC Hydrological Sciences Branch, Code 974.1)
(NASA-GSFC Hydrological Sciences Branch, Code 974.1)**

Abstract: *Soil moisture is a critical hydrosphere state variable that often limits the exchanges of water and energy between the atmosphere and land surface, controls the partitioning of rainfall between evaporation and runoff, and impacts vegetation photosynthetic rate and soil microbiologic respiratory activities. Accurate measurements of this variable are thus required for global water and energy cycle as well as the carbon cycle. A global soil moisture data product is currently steadily generated from the observations of the Advanced Microwave Scanning Radiometer onboard of NASA's Aqua satellite. The accuracy of this data product has not been yet validated globally and assessment of the product quality is required for its applications. This summer student research project was to learn global soil moisture remote sensing technique and research experience in a relatively short time period. The project started with setting up runs of the NASA's Global Land Data Assimilation System and then processing the already in house AMSR-E soil moisture data set with the available software and data processing tools. By directly comparing the soil moisture observation data fields from the AMSR-E data with the corresponding simulation data fields from GLDAS for a certain time period, the statistics of the comparison will be analyzed for understanding their quality and characteristics. We are also planning to use soil moisture field experiment data (SMEX02) to judge the accuracies of these data fields.*

Keywords: AMSR-E, GLDAS, LIS, SMEX02, soil moisture

1. Introduction

Soil moisture is a critical element for both global water and energy budget. Remote sensing is a widely used technique to deal with large-scale spatial and temporal characterizations of soil moisture fields. However, satellite remote sensing data products contain uncertainties due to imperfect instrument calibration and inversion algorithms, geophysical noise, representativeness error, communication breakdowns, and other sources. It is therefore essential that the accuracy and credibility of these remotely-sensed fields be evaluated for their use in critical research and applications[1]. Measuring global soil moisture by any particular technique is very difficult because we have very few observation sites to cover the whole globe; we have very simple models which don't include the complex physics related to soil moisture and satellites can see only few inches below the earth's surface where as soil moisture can be found up to couple of meters below the surface[11]. So, data assimilation is a good technique to incorporate data from all the above three sources to produce an accurate global soil moisture map. In this study, we attempt to cross-compare the global soil moisture data products retrieved from EOS-AQUA AMSR-E sensor with the ground observed soil moisture data as well as the GLDAS/LIS model soil moisture data to validate the satellite product and do the statistical inferences. The following sections of this report explains the motivation for this project work, objective and work plan of this summer project, brief description about the GLDAS/LIS model, AMSR-E instrument and SMEX02 ground observation and their data products, data processing techniques used to extract the final data from the above three sources, results and discussion, conclusion and future work to be done after this project work.

2. Motivation

This summer project was a part of the current priority research focus of the NASA's Land Surface Hydrology Program (LSHP) to validate AMSR-E global soil moisture data product. This was also a part of the long term objective to develop an accurate global soil-moisture dataset. The accurate measurement of soil moisture can help us to solve many global issues because,

- **Soil moisture** is a key state variable in land surface hydrology. It controls the proportion of rainfall that percolates, runs off, or evaporates from land.
- **Soil moisture** plays a major role in surface energy budget. It controls the amount of sensible heat and latent heat energy coming out of the earth's surface.
- **Soil moisture** also enables photosynthesis in plants that use solar energy to convert carbon dioxide and water into the oxygen and food necessary for animal life on Earth. Soil moisture is important for maintaining crop and vegetation health, and its monitoring on a global basis will allow drought prone areas to be monitored for signs of drought.
- Accurate weather forecasts require the rate of transfer of **soil moisture** to the atmosphere, whether by evaporation or plant transpiration.

3. Objective and work plan

The main objective of my summer project is to cross-validate AMSR-E soil moisture data with the GLDAS output soil moisture data for the globe as well as compare both the output data with the SMEX02 ground observation data. This work is a part of the AMSR-E satellite land product validation project by NASA's Land Surface Hydrology Program (LSHP). The schematic diagram of my summer project work plan has been shown in Figure 1. The soil moisture product from the LIS model is 3-hourly, global, binary data. The soil moisture data from the AMSR-E is daily, global data in HDF-EOS format and the same product from the SMEX02 field observation is daily, local, ASCII data. So, I need to collect theses 3 products from three different sources and convert them to the same spatial as well as temporal resolution. For my summer project, I converted all the products into 3-day global data with 25km spatial resolution (except the local observation data which was 3-day local point data). Then I need to display and draw time series of the three datasets using existing softwares like GrADS or Matlab and compare the datasets and draw the conclusion.

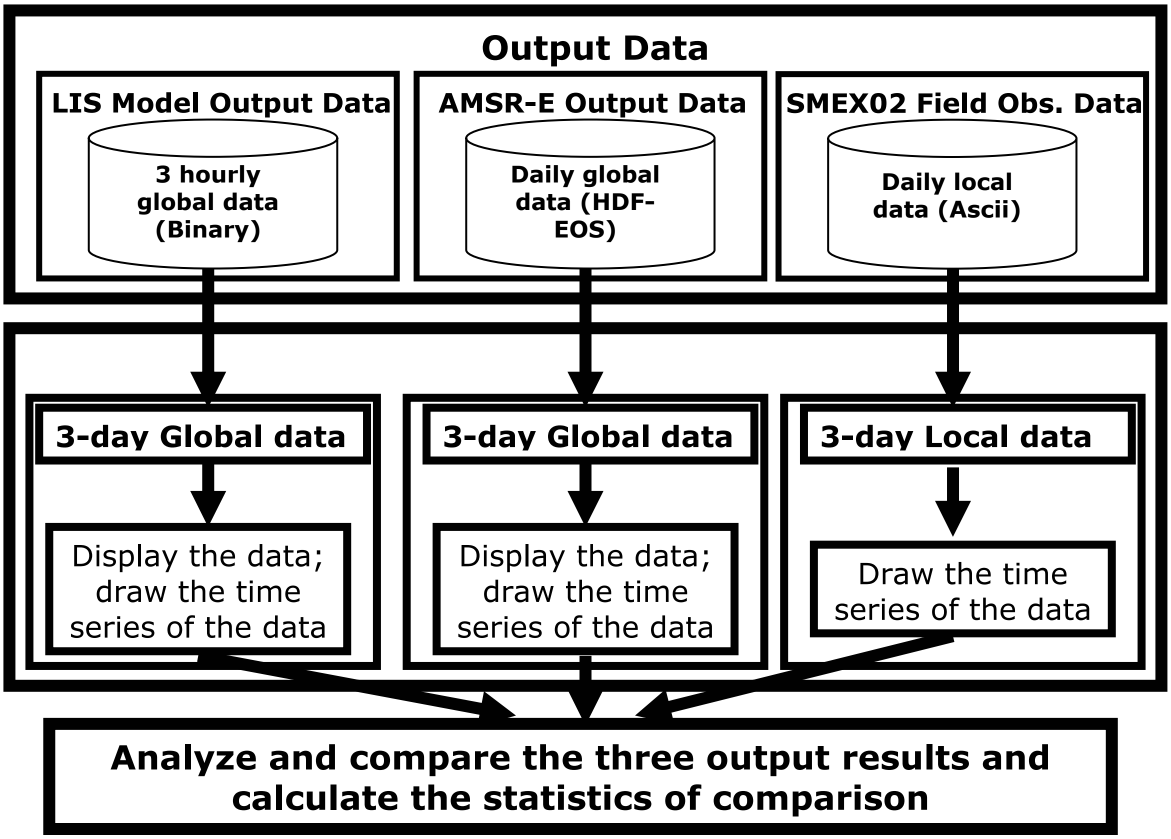

Figure 1 Outline chart of the summer project

4. GLDAS/LIS model and output data products

A Global Land Data Assimilation System (GLDAS) has been developed jointly by scientists at the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC) and the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Prediction (NCEP) in order to provide forecast simulations that will lead to more accurate reanalysis and simulations by numerical weather prediction (NWP) models. GLDAS follows the same concept as that of NLDAS (North American Land Date Assimilation System), but it computes all the parameters for the globe instead of only the North-America. GLDAS makes use of the new generation of ground and space-based observation systems, which provide data to constrain the modeled land surface states. Constraints are applied in two ways. First, by forcing the land surface models (LSMs) with observationbased meteorological fields, biases in atmospheric model-based forcing can be avoided. Second, by employing data assimilation techniques, observations of land surface states can be used to curb unrealistic model states.

Through innovation and an ever-improving conceptualization of the physics underlying earth system processes, LSMs have continued to evolve and to display an improved ability to simulate complex phenomena. Concurrently, increases in computing power and affordability are allowing global simulations to be run more routinely and with less processing time, at spatial resolutions that could only be simulated using supercomputers five years ago. GLDAS harnesses this low-cost computing power to integrate observation based data products from multiple sources within a sophisticated, global, high-resolution land surface modeling framework. GLDAS is very unique as it is a global, high-resolution, offline (uncoupled to the atmosphere) terrestrial modeling system that incorporates satellite- and ground-based observations in order to produce optimal fields of land surface states and fluxes in near–real time[3, 10].

The Land Information System (LIS) is a high performance land surface modeling and data assimilation system, based on GSFC's Land Data Assimilation Systems. Figure 2 shows the schematic diagram of LIS/GLDAS model. There are 3 parts in the LIS model, namely input data, driver routines and land surface models (LSMs). The input data can be broadly divided into 2 parts, i.e. *parameter data* which represent the properties of the land surface that change on time steps of a day or longer, e.g., soil, land cover, topography etc. and *forcing data* which include the atmospheric inputs to the land surface models, including precipitation, radiation, and surface winds, temperature, pressure and humidity. User interface (called as LIS card file) file is very user friendly. The user has to specify which parameter data and forcing data and the name of the LSM he/she wants to use in the user interface file. Apart from that, he/she has also to specify the initial and final date of the model run, the time step for the model to reinitialize the model, the time step for the model to write the output files etc. The driver routines read this LIS card file and accordingly they collect the specified input data, interpolate all the data to the same grid size as required for the output and provide that data to the specified LSM. LSM is the core part of the LIS model. All the physical equations have been applied through the LSM to calculate the output parameters. Again the driver routines take these output data from the LSM and write into output data files. The user then can use softwares like GrADs or Matlab to postprocess that output data and analyze it. The user doesn’t need to worry about the LIS model except the user interface file unless he/she wants to customize the model. The most promising feature of the LIS model is that it can be expanded in many ways. The user can use his forcing and parameter data as well as the LSM in the LIS model through the available plug-ins. So, LIS model is an expandable system[4].

Land surface models predict terrestrial water, energy and biogeophysical processes critical for applications in weather and climate prediction, agricultural forecasting, water resources management, hazard mitigation and mobility assessment.

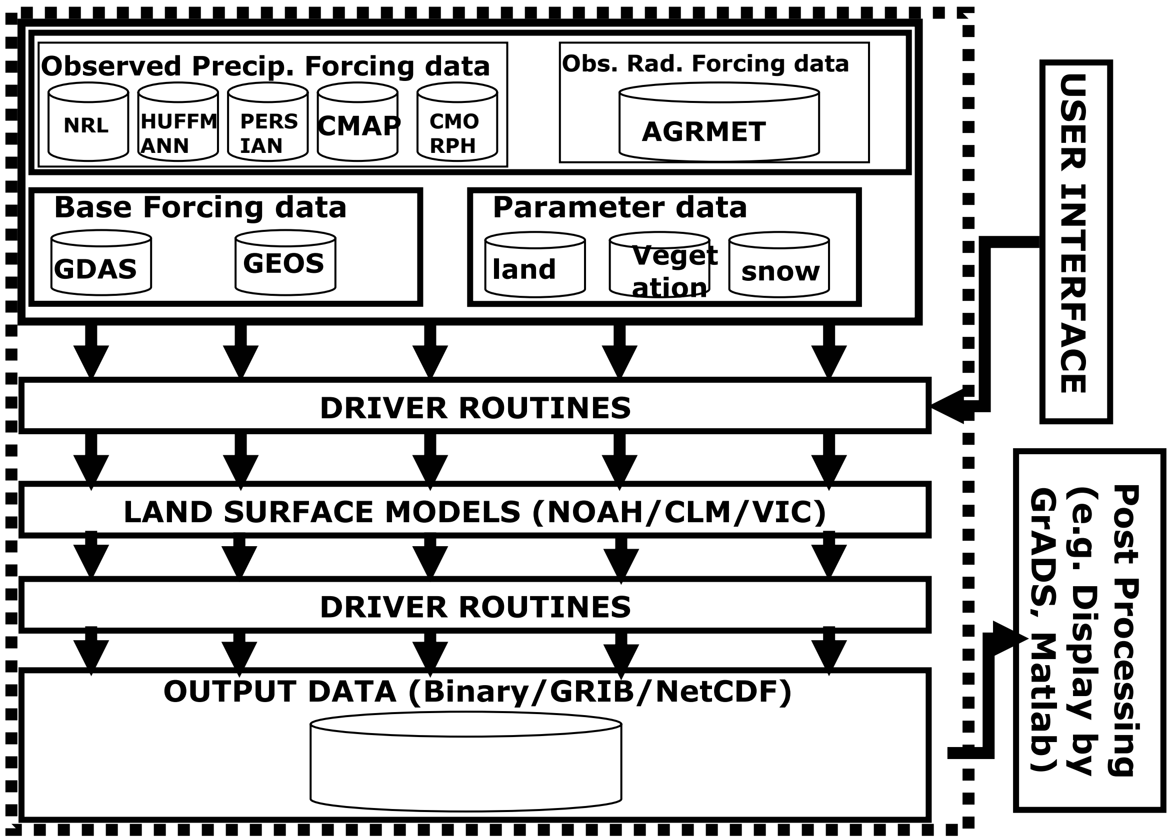

Figure 2 Schematic diagram of GLDAS/LIS model

5. AMSR-E instrument and soil moisture data product

The Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) is a passive microwave radiometer, launched aboard NASA's Aqua Satellite on 4 May 2002. AMSR-E is modified from the Advanced Earth Observing Satellite-II (ADEOS-II) AMSR, designed and provided by NASDA[6]. AMSR-E is one of its kind which is a more focused instrument and is specifically meant to measure the water content in the land surface of the earth. This conically scanning instrument senses microwave radiation (brightness temperatures) at 6 frequencies ranging from 6.9 to 89.0 GHz with both horizontal and vertical polarizations (total 12 channels). The detailed AMSR-E performance characteristics have been shown in Table1. Daily Level-2B and Level-3 products are now available from the National Snow and Ice Data Center (NSIDC), beginning with dates from February 18, 2004. These derived geophysical AMSR-E products include measurements of rainfall, snow, sea ice and many other land and ocean geophysical variables[5, 11].

The AMSR-E C- (6.9GHz) and X-band (10.7GHz) channels are strongly related to land surface soil moisture and are used to generate the global land data products[8, 9]. Soil moisture is the principal retrievable Aqua AMSR land surface parameter. Surface temperature and vegetation water content are also retrieved by the algorithm, and are output with the level 2 soil moisture product as diagnostic parameters. The level 3 product is derived by compositing the level 2 parameters daily into global maps, separating ascending and descending passes so that diurnal effects can be evaluated. Soil moisture is not retrievable where significant fractions of snow cover, frozen ground, dense vegetation, precipitation, open water, or mountainous terrain occur within the sensor footprint (as determined by a classification algorithm and ancillary information). The algorithm products are generated on an earth-fixed grid with ~25- km nominal grid spacing. The spatial resolution of the products is ~60 km (determined by the 6.9 GHz footprint resolution)[2]. Table 2 gives a brief explanation of all the AMSR-E land products.

CENTER FREQUENCIES (GHz)	6.925	10.65	18.7	23.8	36.5	89.0
BANDWIDTH (MHz)	350	100	200	400	1000	3000
SENSITIVITY (K)	0.3	0.6	0.6	0.6	0.6	1.1
MEAN SPATIAL RESOLUTION (km)	56	38	21	24	12	5.4
IFOV (km x km)	74 x 43	51 x 30	27 x 16	31 x 18	14 x 8	6 x 4
SAMPLING RATE (km x km)	10 x 10	10 x 10	10 x 10	10 x 10	10 x 10	5 x 5
INTEGRATION TIME (MSEC)	2.6	2.6	2.6	2.6	2.6	1.3
MAIN BEAM EFFICIENCY (%)	95.3	95.0	96.3	96.4	95.3	96.0
BEAMWIDTH (degrees)	2.2	1.4	0.8	0.9	0.4	0.18

Table 1 AMSR-E performance Characteristics[6]

Product Level	Parameter	Spatial Resolution	Grid Spacing	Temporal Resolution
2	Soil Moisture	56 km	25 km	swath
2	Vegetation Water Content	56 km	25 km	swath
2	Surface Temperature	56 km	25 km	swath
3	Soil Moisture	56 km	25 km	daily
3	Vegetation Water Content	56 km	25 km	daily
3	Surface Temperature	56 km	25 km	daily
3	Gridded Brightness Temps.	12, 56 km	25 km	daily

Table 2 Description of AMSR-E Land data Products[2]

6. SMEX02 field observation and soil moisture data

Field experiments have been very successful at addressing a broad range of science questions in terrestrial hydrology. The data have been used in studies that went well beyond the algorithm research.

For 2002, soil moisture and water cycle field experiments was conducted that supported the Aqua Advanced Microwave Scanning Radiometer (AMSR), NASA's Global Water and Energy Cycle Program, and will support future satellite missions for Terrestrial Hydrology. Main elements of the experiment were validation of AMSR brightness temperature and soil moisture retrievals, extension of instrument observations and algorithms to more challenging vegetation conditions, integration of land surface and boundary layer measurements, and the evaluation of new instrument technologies for soil moisture remote sensing.

Figure 3 shows all the SMEX sites in USA where experiments have already been done and where it is going to be performed in future. SMEX02 was conducted in Iowa, USA during June 24 to July 15, 2002. The experiments included gravimetric and volumetric soil moisture, meteorological, and soil temperature measurements.

The main elements of the experiment were:

1. To study land-atmosphere interactions and extend remote sensing observations and algorithms to more challenging vegetation conditions than previously studied
2. To validate AMSR-E satellite brightness temperature and soil moisture measurements
3. To evaluate new instrument technologies for soil moisture remote sensing.

Figure 3 Locations of target field experiment sites[2]

7. Data processing

7. 1. GLADAS/LIS model data: The important features of the specifications I used for my model run is given in Table 3. I ran this model for the whole globe from June 1, 2001 to September 30, 2002. I used a restart file given by Dr. Matthew Rodell which was generated after 15 years of spin up. I also spun up my model for one year from June 1, 2001 to May 31, 2002 and I used the model output data after that for the comparison study. I was planning to produce the 2 cm top layer soil moisture data instead of NOAH model default 10-cm top layer soil moisture output to make it more realistic to compare with the AMSR-E satellite data (which has soil moisture from top 2 cm only). But, it required very small time step (2 minutes instead of default 30 minutes) and it was computationally intensive. It was not possible to run the model for a longer time in this short summer program. So, I used the default 10 cm top layer soil moisture output. I accessed all these 3-hourly binary output soil moisture data through GrADS control file and then merged the datasets to create the 3-day global soil moisture dataset. I also used the GrADS script to extract the data and display it and then imported that data into excel sheets to draw the time series of the dataset.

Land Surface Model (LSM)	NOAH
Base Forcing	GDAS
Precipitation Forcing	CMAP
Radiation Forcing	AGRMET
Leaf Area Index (LAI)	AVHRR based LAI
Maximum number of tiles per grid	13
Time step of the run	30 minutes
Latitude Range	60^0 S to 90^0 N
Longitude range	180^0 W to 180^0 E
Output Data Resolution	1/4 degree
Output interval to write the output files	3 hours
Output data format	Binary

Table 3 Inputs I used for my GLDAS/LIS model run

7. 2. AMSR-E satellite data: AMSR-E level 3 (named as AE_Land3 product) soil moisture data was from top 2 cm layer of the earth for the period Jun 19, 2002 to October 3, 2002. It was daily global data in HDF-EOS format. Figure 4 shows the steps used for the AMSR-E data processing. For this AMSR-E data processing part, I used all the software tools provided by Dr. Xiwu Zhan. First, I converted the HDF-EOS data files to binary 25 km resolution dataset. Then I changed the datasets original ease-grid map projection to the lat-long map projection system. Finally to get the global coverage, I merged the daily global product to produce 3-day global soil moisture dataset. Here also I used the GrADS script to extract the data and display the global soil moisture map and then used excel sheets to draw the time series of the AMSR-E soil moisture dataset.

7. 3. SMEX02 observation data: I used the SCAN (Soil Climate Analysis Network) site data which was hourly ASCII data for a point location (42.1^0 N, 93.85^0 W). This data was from June 1, 2002 to August 31, 2002. I just imported these ASCII files to excel sheet to draw the time series of this dataset[7].

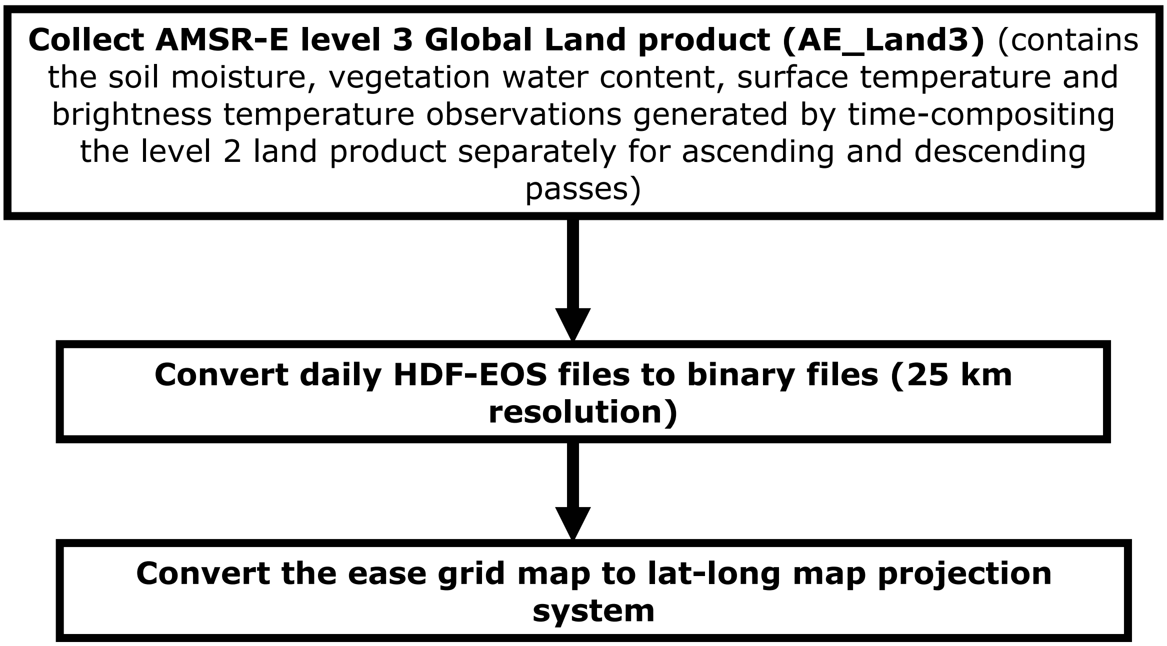

Figure 4 AMSR-E soil moisture data processing steps

8. Results and discussion

Figure 5a shows the global soil wetness map generated from the LIS/GLDAS model for the period July 25-27, 2002. All the values are in volumetric percentage and it ranges from 0 to 50 %. As expected, all the deserts namely Kalahari and Sahara deserts in Africa, Gobi and Saudi Arabian deserts in Asia, Western Australia, Western USA and some parts of South America show very low soil moisture percentage values ranging from 0 to 15 %. Some of the regions show higher values ranging from 35 to 45 % which might be because of the heavy precipitation on the previous days, e.g. higher values in the south-east Asian countries are because of the south-east and south-west monsoon rainfall. Soil wetness values more than 45% are mainly attributed to the snow cover regions like Greenland here in the map. Figure 5b shows the global soil moisture map generated from the AMSR-E brightness temperature data for the same time period as that of the Figure 1a. Unlike LIS model, AMSR-E soil moisture map doesn't include the snow cover of Greenland. Here also all the values are in volumetric percentage. The main notable point on this map is that the soil moisture values for the whole globe ranges from 0 to 20 % only. So in general, this soil wetness map is very dry as compared to the wetness map generated from LIS model. This map follows the same pattern as that of the LIS model throughout the globe except over India and north of Canada where it is again very dry. Figure 6a shows the soil wetness map from the LIS model for the SMEX02 region (39.875 to 45.875 N and 95.875 to 89.875 W) for June 19-21 and July 25-27, 2002 periods. We can see the systematic change of soil moisture in each image. The very high soil moisture values for the period July 25-27 might be because of the precipitation. Figure 6b shows the same maps as that of Figure 6a, but generated from AMSR-E data. Here, we miss the systematic pattern of soil moisture and the values are very low again. To compare the model and satellite soil moisture data with the SMEX02 ground observation data, I have plotted the time series graph of the soil moisture values for a point location (42.1^0 N, 93.85^0 W) which is shown in Figure 7. SMEX02 data is from June 1 to August 31, 2002; LIS model data is from June 1 to September 30, 2002 and AMSR-E data is from June19 to October 3, 2002 (with some of the days are missing). Each point in this graph is a 3-day averaged data point. The figure clearly indicates that LIS model soil moisture values are almost within the range of the soil moisture values of the SMEX02 observation though they don't follow the exact pattern. So, these two datasets are comparable. But at the same time, AMSR-E soil moisture values are very low as compared to the other two datasets and there is a little variation in the AMSR-E data throughout the summer.

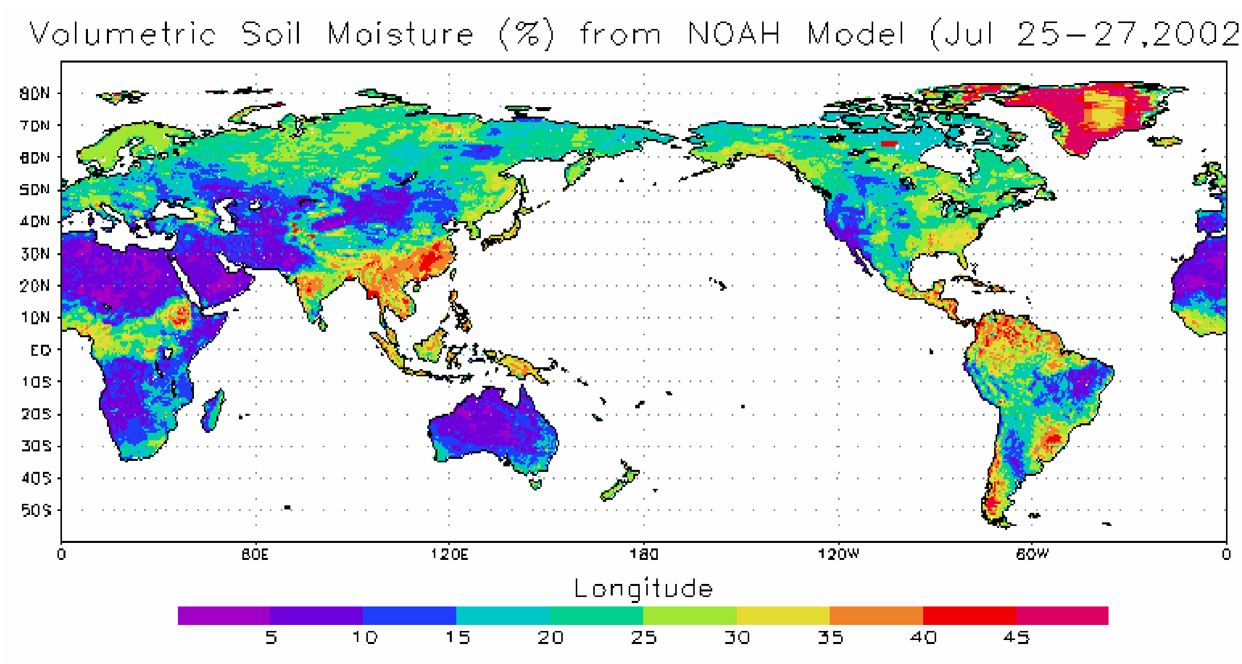

Figure 5a The Global soil wetness map generated from the LIS/GLDAS model

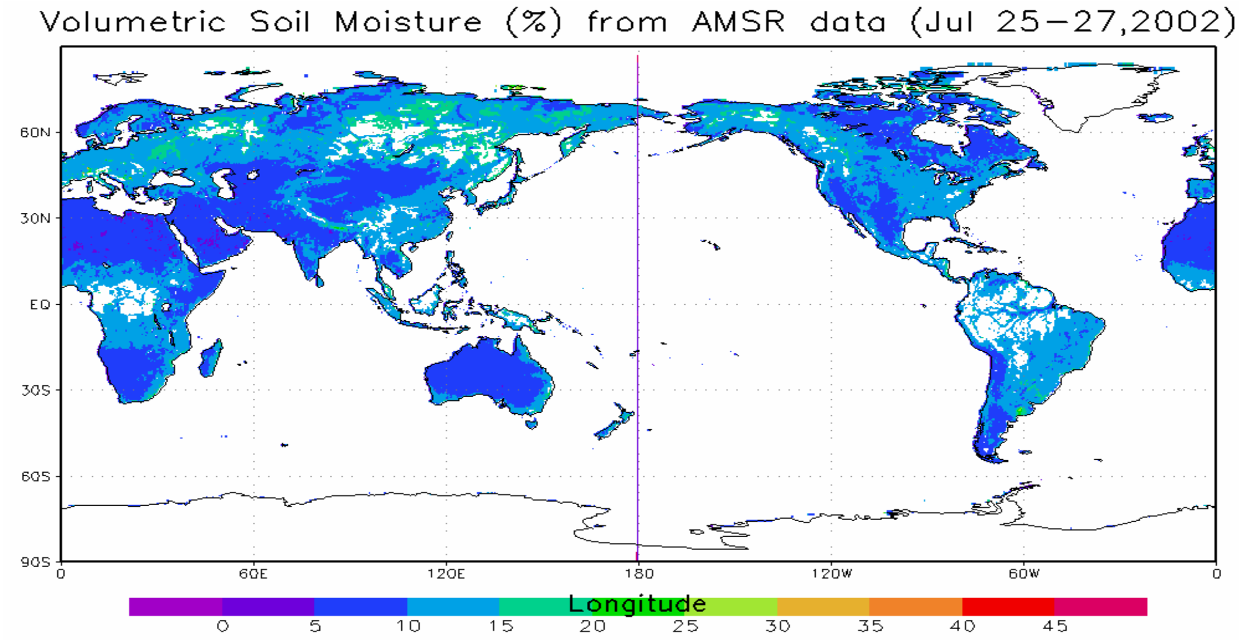

Figure 5b The Global soil wetness map generated from the AMSR-E data

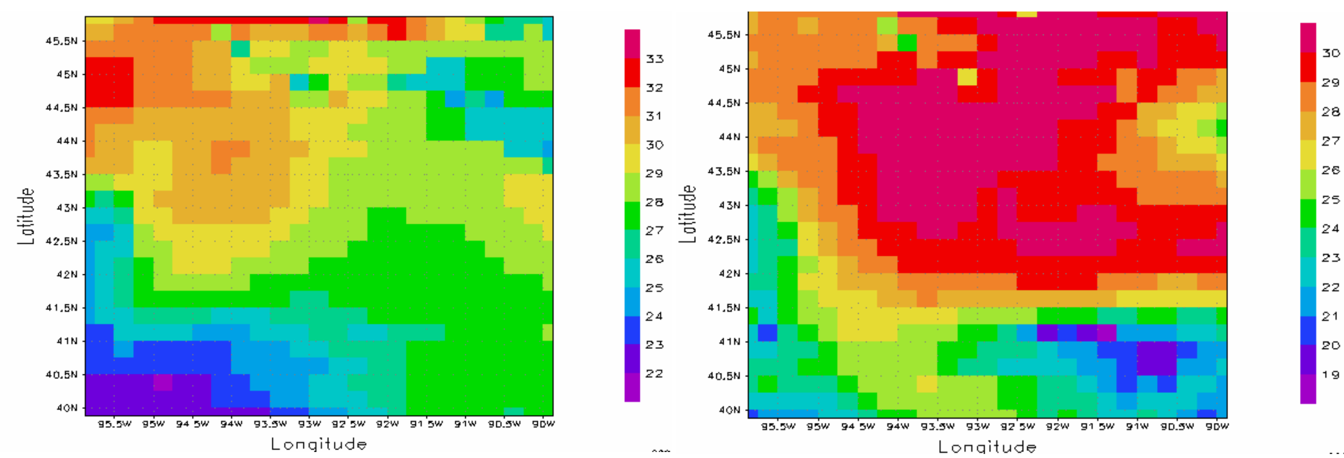

Figure 6a The soil wetness map from the LIS model for the SMEX02 region for June 19-21, 2002 (left) and July 25-27, 2002 (right) period

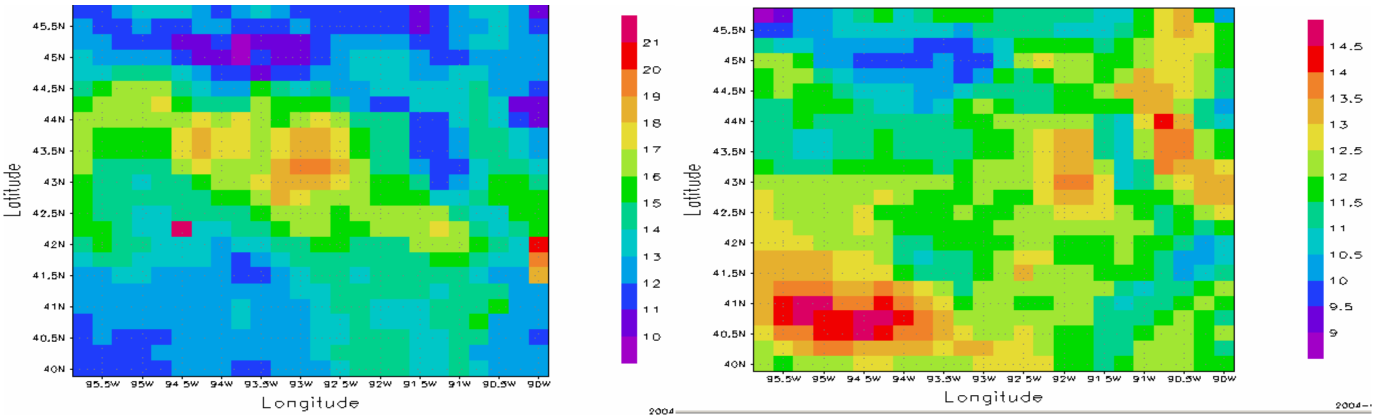

Figure 6b The soil wetness map from AMSR-E for the SMEX02 region for June 19-21, 2002 (left) and July 25-27, 2002 (right) period

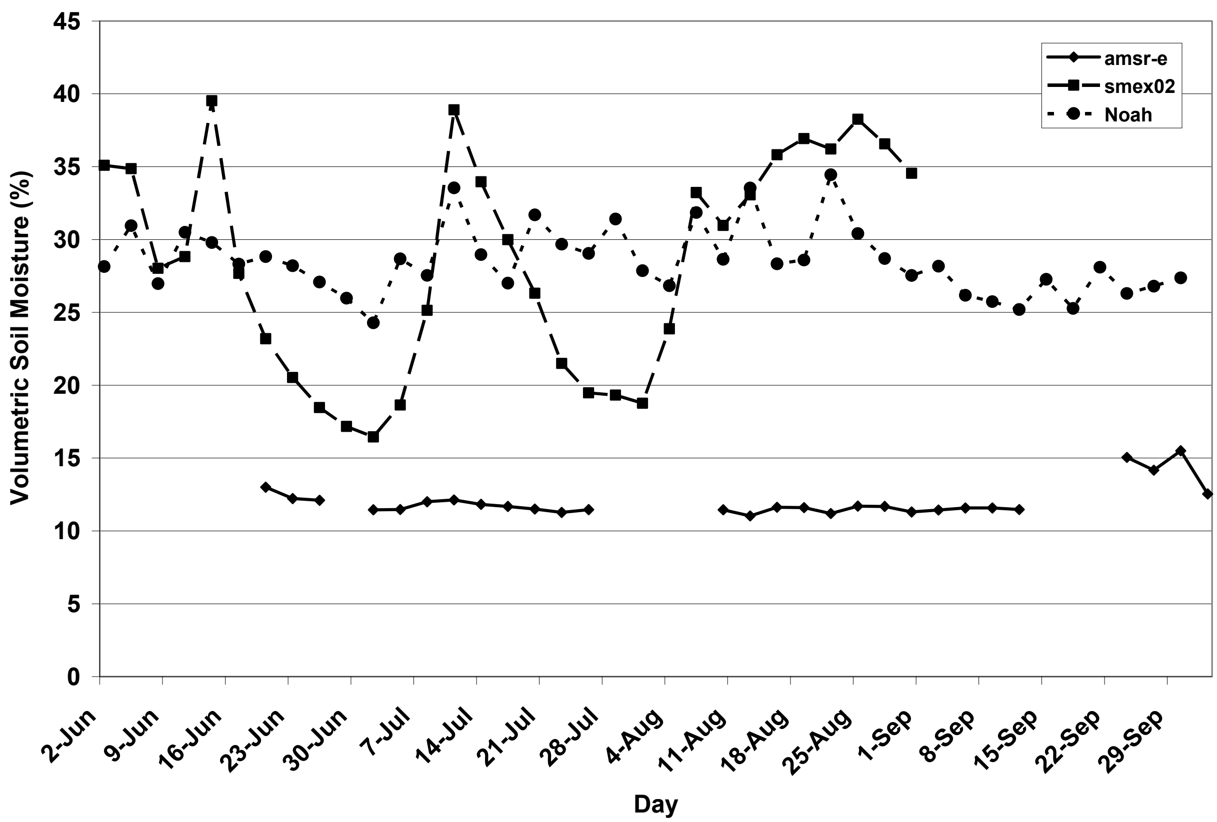

Figure 7 Time series plot of the SMEX02, GLADS/LIS model and AMSR-E soil moisture datasets

9. Conclusion Recent study in this paper indicates that we can't compare the current AMSR-E soil moisture data with any other soil moisture data because AMSR-E soil moisture data is very low from reality. LIS model data looks like a good dataset and we can use that for further scientific research.

10. Future Work The results shown here is a very preliminary work. We need to study deep and do quantitative study and draw statistical inferences. For this comparison study, we used 3-day averaged soil moisture data. Instead, we can use the swath data of the satellite and the ground observed data and the model data closest to the time of the overpass of the satellite over the study area. This comparison will be more accurate than averaging the data over 3-day period. Since AMSR-E soil moisture values are very low, we need to recheck our AMSR-E soil moisture data extraction programs. Also, we need to check the AMSR-E soil moisture algorithm. Finally after the AMSR-E soil moisture validation work, we want to assimilate this soil moisture data with a global model to generate a good global soil moisture data product which is more close to the observation and more realistic.

Acknowledgements: I am very happy for the amount of work we have done in this project during this 10-week period. I thank Dr. Menas Kafatos, my Ph. D. supervisor, for supporting me for this summer internship and allowing me to work at NASA-GSFC. I acknowledge Dr. Xiwu Zhan for providing me the software tools to process the AMSR-E data, Dr. Mathew Rodell for providing the restart file to start my GLDAS/LIS model run and Kristi Arsenault for providing me the required forcing and parameter data to run the GLDAS/LIS model. I am also grateful to Jim Geiger for his technical support on the software issues related to the GLDAS/LIS model. I am thankful to Dr. Christa Peters-Lidard, Dr. Jon Gottschalck and Dr. Jesse Meng for their critical comments on NOAH model soil moisture top layer depth issue, CMAP precipitation data related issues and AGRMET radiation forcing data related issues respectively. The feedbacks taken from Dr. Randal Koster and Dr. Paul Dirmeyer are also highly appreciated. All the members of the HSB group were very supportive for my work during this summer program. At the last, I don't find words to express my gratitude to Marci Delaney and Anathea Brooks for helping me in all the non-technical issues and making my stay at NASA-GSFC very comfortable.

References

[1] Eymard, L., Bernard, R. and Lojou, J. Y. (1993), "Validation of microwave radiometer geophysical parameters using meteorological model analysis", *Int. J. Remote Sens.,* Vol. 14, No. 10, pp. 1945-1963.

[2] http://eospso.gsfc.nasa.gov/ftp_docs/validation/SDVPlanversion2.pdf

[3] http://ldas.gsfc.nasa.gov/

[4] http://lis.gsfc.nasa.gov/

[5] http://nsidc.org/data/amsr/

[6] http://www.ghcc.msfc.nasa.gov/AMSR/

[7] http://www.nsidc.org/data/amsr_validation/soil_moisture/

[8] Njoku, E.G. and Li, L. (1999), "Retrieval of land surface parameters using passive microwave measurements at 6-18 GHz", *IEEE Trans. Geosci. Remote Sens.,* Vol. 37, No. 1, pp. 79-93.

[9] Njoku, E.G., Jackson, T. J., Lakshmi, V., Chan, T. K. and Nghiem, S. V. (2003), "Soil moisture retrieval from AMSR_E", *IEEE Trans. Geosci. Remote Sens.,* Vol. 41, No. 2, pp. 215-229.

[10] Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D. and Toll, D. (2004), "The global land data assimilation system", *Bull. Amer. Met. Soc*., Vol. 85, No. 3, pp. 381-394.

[11] Zhan, X., Houser, P. R., Walker, J. P. and Rodell, M. (2004), "Validation of AMSR-E Soil Moisture Product Using Data Assimilation Techniques", COAA2004, June 28-30, 2004, Beijing.