NWS-CPC's Monitoring and Prediction of US Soil Moisture & Associated Land Surface Variables: Land Data Reanalysis

Y. Fan, H. M. van den Dool, K. Mitchell and D. Lohmann

CPC/EMC/NCEP/NOAA

I. Introduction

Land surface variables, such as soil moisture, are among the most important components of memory for the climate system. To quantify the memory of land surface variables, we need multi-decadal land data records. Therefore more accurate and long time series of land surface data is very important for understanding of land-surface-atmosphere interaction and for improving our ability to predict weather and climate. The Retrospective U.S. LDAS Project is part of an effort to upgrade NWS-CPC's system for monitoring and prediction of US soil moisture and associated land surface variables by replacing the current leaky bucket model (Huang et al, 1996) with the more advanced Noah land surface model. Thus the main effort of this work is to rerun the Noah land surface model retroactively as far back as possible: 1948-present, which is essentially a 'Land Reanalysis'.

II. Model & Forcing Data

Since the late 1990's one of the several LDAS (Land Data Assimilation System) land surface models, the Noah land surface model (LSM), has been run at EMC/NCEP in real time forward mode. This Noah LSM (Mitchell et al, 2002) is physically much more complete and interesting than the leaky bucket model – a simple water budget model, which currently operates at CPC/NCEP on monthly temperature (T) and precipitation (P) input data at 344 US Climate Divisions. Additionally, the spatial and temporal resolution of the Noah LSM is also very much higher than that of the leaky bucket model. During the past several years, the Noah LSM underwent substantial upgrades, including an increase of the number of model layers, modifications to the canopy conductance formulation, bare soil evaporation, vegetation phenology, surface runoff and infiltration, thermal roughness length treatment in the surface layer exchange coefficients, cold season processes etc. These model enhancements significantly improve the performance of the Noah LSM (Ek et al 2003).

The Noah LSM used for the Retrospective U.S. LDAS Run was configured on the North America LDAS grid, which is a 464 X 224 1/8th-degree grid bounded by 25 N, 53N, 67W, and 125W. On this grid, NASA/GSFC uses high resolution source datasets (typically 1-km) to derive common 1) land/sea mask, 2) terrain heights, 3) dominant and subdominant vegetation classes, and 4) vegetation parameters for use by all LDAS LSMs. Similarly on this grid, NWS/OH uses high resolution soil databases to derive soil characteristics (such as texture) and companion soil physical parameters. The parameters of the Noah LSM and fixed fields are derived from the above existing high resolution vegetation, soil coverage and orography. Meanwhile, the parameters of leaky bucket model were estimated using Oklahoma data alone and applied to the entire model domain.

As the first step of the above effort, a 51 year (1948-1998) 1/8th degree grid and hourly retroactive LDAS forcing input data set was generated, namely air temperature (T_a), air humidity (q), surface pressure (p_*), wind (u,v), surface downward shortwave radiation ($S\downarrow$), and surface downward longwave radiation ($IR\downarrow$) from NCEP/NCAR Global Reanalysis while the new and improved precipitation from CPC (Higgins & Shi). This new CPC daily precipitation over 1948-1998 is different from the one we processed before in (a) using new PRISM precipitation climatology vs no PRISM climatology, and (b) least squares distance weighting as the objective analysis scheme vs Cressman analysis scheme and (c) on 1/8th degree US LDAS grid vs 1/4th degree grid.

In order to make a homogeneous forcing data set at the required spatial and temporal resolution and at the same time use the available observed data as much as possible, some unique procedures and techniques were used. The details are as follows:

The procedures for creating atmospheric forcing from NCEP-NCAR Global Reanalysis are:

(1) extract 7 (T_a, q, p_*, u, v, $S\downarrow$, $IR\downarrow$) variables from NCEP/NCAR global Reanalysis data (which is 4 times daily, 1948-1998);
(2) convert data from Gaussian grid to LDAS grid using linear spatial interpolation;
(3) interpolate 6-hour data to hourly data using linear temporal interpolation, except for T (see item 7);
(4) perform a pre-processing for downward shortwave radiation by averaging two 6-hour periods;
(5) perform solar zenith angle correction to downward solar radiation;
(6) perform elevation correction to temperature, humidity, surface pressure, downward longwave radiation.
(7) determine the hourly temperature from (a) 6 hourly temperature (4 times daily) and (b) the Reanalysis 6-hourly maximum and minimum temperature for each 6 hour interval (allowing for an error in the data that was stored).

The procedure for creating precipitation forcing is:

(1) fill in missing values in CPC hourly precipitation on the 2.5x2.0 degree grid using linear temporal interpolation with the available data at its two sides;
(2) convert the hourly precipitation from 2.5x2.0 to 0.125 degree LDAS grid using linear interpolation;
(3) calculate the hourly weights with the CPC hourly precipitation on 0.125 degree LDAS grid;
(4) apply the hourly weights to daily precipitation to obtain the hourly precipitation on 0.125 degree LDAS grid.

After the hourly precipitation and atmospheric forcing are created, they are merged together and organized in an hourly file in GRIB format. The size of the tarred and compressed data files is around 6 GB per year.

III. Products of the Retrospective US LDAS Run - Data format and availability

Before starting the Retrospective US LDAS Run, proper initialization of the Noah LSM is also important. In order to save some spin-up time, the initial conditions (July 1, 1998) were taken from an operational run from EMC/NCEP, which started from late 1996. We set the time stamp back to 1948 and repeatedly run the Noah LSM for 3 and half years with 1948 forcing. Then we started the Retrospective US LDAS Run from Jan 1, 1948.

The output dataset from the Retrospective US LDAS Run is generated at 1/8th degree LDAS grid and all results from the Retrospective US LDAS Run are stored in GRIB format or binary format. The details of the full dataset are organized as follows:

a). The initial conditions (23Z) from the Retrospective US LDAS Run will be in binary format and stored once per day (9 variables).

b). The products of the Retrospective US LDAS Run are in GRIB format and will be divided into two groups:

(i) The first group mainly includes 8 flux variables (energy balance components) - Net surface shortwave and longwave radiations, latent heat flux, sensible heat flux, ground heat flux, snow phase change heat flux, downward surface shortwave and longwave radiations. The flux terms are averaged over the 3 previous hours and stored 8 times (2Z, 5Z, 8Z, 11Z, 14Z, 17Z, 20Z, 23Z) per day.

(ii) The second group mainly consists of 15 variables (water balance components, surface and subsurface state variables, cold season processes) – i.e. soil temperature, soil moisture (liquid and frozen) and liquid soil moisture for 0-10cm, 11-40cm, 411-100cm and 101-200cm respectively, snowfall, rainfall, total evaporation, surface runoff, subsurface runoff, snowmelt, snowpack water equivalent, plant canopy surface water storage, snow evaporation, potential evaporation, snow depth and snow cover. The water accumulations will be daily accumulations and other variables are averaged daily. These 15 variables are stored once per day.

All above forcing data and the full land dataset from the Retrospective US LDAS Run are archived on NCEP IBM SP HSM storage system. The details of accessing these data will be released late.

IV. Summary

The preliminary results show that the hourly forcing data set is reasonably good, compared with the observations. At the time of this writing, a few test runs have been done and the Retroactive LDAS Run has been run for 45 years, starting from the first day of 1948. The outputs provide an improved soil moisture and associated land surface variable dataset, such as soil temperature (example see Fig.1), snowpack, surface fluxes etc which we never had before. Many interesting data validation, data analysis and model comparisons are underway. The retroactive run will also provide superior model consistent initial conditions for numerical predictions. These studies are important for understanding land memory processes, evaluating and developing new land surface models, and eventually improving our understanding of weather prediction and climate change.

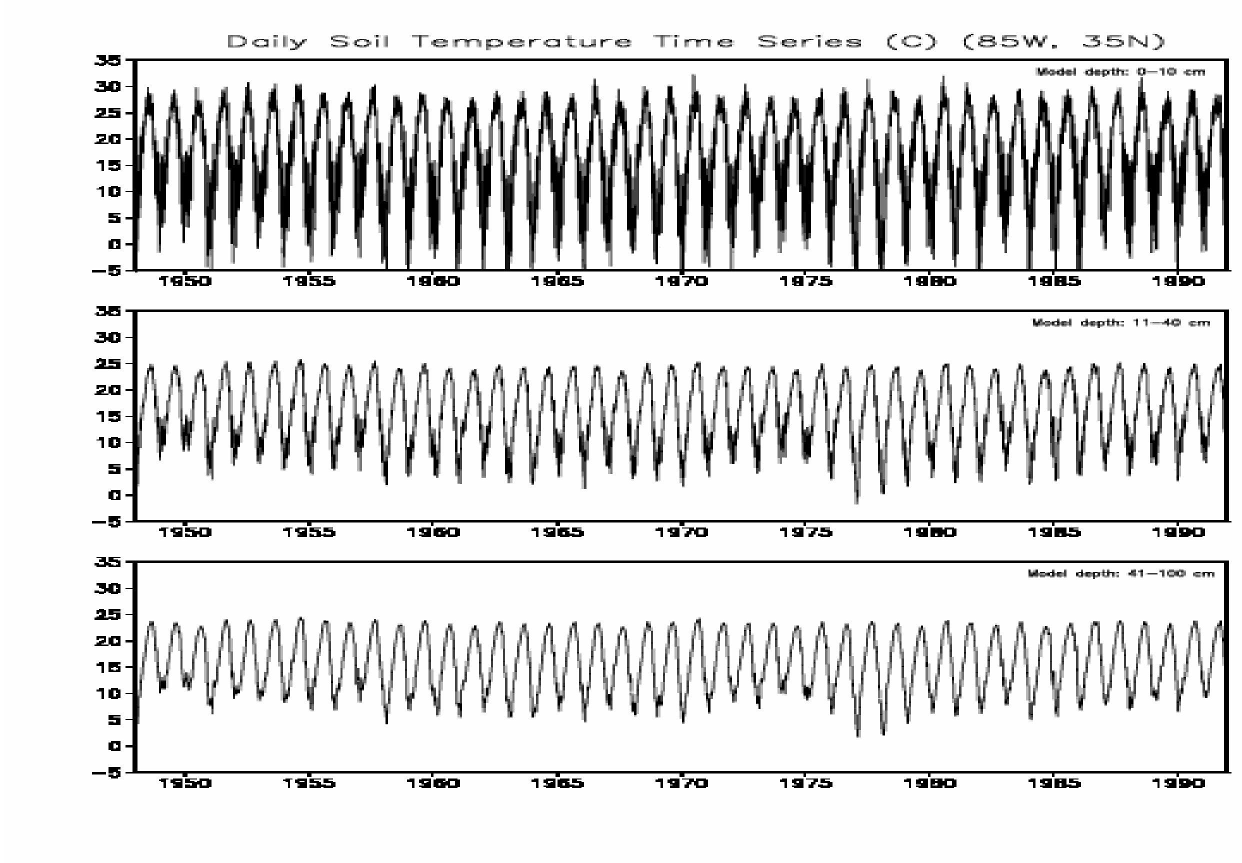

Fig.1 The soil temperature time series for the top 3 model layers

References

Huang, J., H.M. Van den Dool, and K.P. Georgakakos, 1996: Analysis of model-calculated soil moisture over the United States (1931-1993) and applications to long-range temperature forecasts. J.Climate, 9. 1350-1362.

Ek, M.B., K.E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J.D. Tarplay, 2003: Implementation of Noah land-surface model advances in the NCEP operational mesoscale Eta model (submitted to JGR)

Maurer, E.P. et al, 2003: A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States. J.Climate, 15, 3237-3251.

Mitchell, K.E. et al, 2002: The Community Noah Land Surface Model (LSM) – User's Guide (v2.5.2)