Development of a high resolution land surface dataset for the South Asian monsoon region

C. K. Unnikrishnan[1], M. Rajeevan[2,]*, S. Vijaya Bhaskara Rao[3] and Manoj Kumar[4]

[1]National Atmospheric Research Laboratory, Gadanki 517 112, India
[2]Earth System Science Organization/Ministry of Earth Sciences, Prithvi Bhavan, New Delhi 110 003, India
[3]Sri Venkateswara University, Tirupati 517 502, India
[4]Centre of Excellence in Climatology, Birla Institute of Technology, Mesra, Ranchi 835 215, India

In this study, we report the development of a high resolution land surface dataset for the South Asian monsoon region for studies on land surface processes, and land and atmosphere coupling. The high resolution land data assimilation system was used to develop the land surface dataset utilizing TRMM rainfall and ECMWF atmospheric variables as forcing parameters. The dataset was developed at a spatial resolution of 0.5° and temporal resolution of 1 h and spans a period of 6 years, i.e. 1 January 2005 to 31 December 2010. The major highlights in the development of the present dataset are higher spatial and temporal resolution of land surface parameters, use of sub-daily forcing parameters including rainfall, use of MODIS land-use data in lieu of USGS land-use data and weekly varying vegetation fraction instead of monthly vegetation climatology. A comparison of soil moisture and soil temperature with limited surface observations of the IMD suggests reasonable reliability of the land surface data. The model sensible heat flux data are compared with *in situ* measurements at Ranchi and MEERA reanalysis data. The sensitivity analysis shows that the land surface data are sensitive to rainfall and green vegetation cover data used as the forcing parameters. The dataset has been used to discuss the variations of land surface processes associated with active and break spells and a severe heat wave observed in 2009. The present dataset will be useful for many applications, including initializing numerical models for weather prediction. This high resolution land surface dataset is available for research on request.

Keywords: Land–atmosphere coupling, land surface dataset, monsoon spells, soil moisture and temperature.

THE South Asian monsoon is characterized by a wide spectrum of spatial and temporal variability due to interaction of land, atmosphere and oceans. It is well known that the variability of the South Asian monsoon is linked to tropical sea-surface temperature anomalies. However, land surface conditions also make significant contribution to the monsoon variability. Variation of the characteristics of land surface leads to a variation in the fluxes of heat and water vapour to the atmosphere, which in turn can lead to changes in the circulation and precipitation. Unfortunately, there are only few studies addressing the important role of land surface processes on the South Asian monsoon circulation and rainfall. Lack of good observations of land surface parameters like soil moisture and soil temperature over the monsoon region remained a major constraint over the years.

Previous studies suggested the links of land surface processes and monsoon variability using global and regional climate models. Charney[1] proposed a positive feedback mechanism between decreasing vegetation cover and the increase in drought conditions across the Sahel region of western Africa. However, the sensitivity of the atmosphere to variations in the response to land surface conditions may vary in space and time[2]. The impact of soil moisture variations is found to be significant in relatively dry monsoon regions, such as the South Asian monsoon region, but such a response is not evident in the humid monsoon regions, such as Southeast Asia. The Global Land and Atmosphere Coupling Experiment (GLACE) in which 12 atmospheric general circulation models participated[3,4] identified the Indian monsoon region as one of the hot spots where soil moisture variations have a significant impact on the precipitation at synoptic timescales. Rajendran *et al.*[5] using a global climate model showed that interactive soil moisture improves simulations of active and weak spells and the northward propagation of convective rain bands compared to the simulations when a fixed hydrology was assumed. The land surface and atmosphere coupling studies mainly focused on the synoptic scale, sub-seasonal scale, seasonal and interannual scale[3,4,6–9]. The role of land surface process on monsoon circulation has been discussed in the literature[10–19].

In India, observational and modelling studies of processes such as the atmosphere–hydrosphere feedbacks (in which land surface processes will be important) are envisaged and planned under the Continental Tropical Convergence Zone (CTCZ) experiment (http://www.imd.

*For correspondence. (e-mail: mn.rajeevan@nic.in)

gov.in/SciencePlanofFDPs/CTCZ%20Science%20Plan.pdf) being supported by the Ministry of Earth Sciences (MoES), Government of India. One of the focus studies of CTCZ will be to examine the role of land surface processes on the dynamics of active and break events and northward propagation of convective systems over the monsoon region.

Land surface modelling initiatives and scope

Availability of reliable data of land surface parameters over the Indian summer monsoon region to initialize and validate land surface processes in the dynamical models is still a challenge. Historical records of land surface data are also not available over most of the continents, including India. Land surface fields derived from atmospheric and climate models such as ECMWF or NCEP reanalysis data are not suitable for climate studies as they are affected by large errors in model simulated precipitation[20]. Recently, efforts have been made to simulate past land surface conditions using comprehensive land surface models forcing with realistic forcing. The past initiatives for developing the datasets are Global Soil Wetness Project[21], the North America Land Data Assimilation System (NLDAS)[22], and the Global Land Data Assimilation System (GLDAS)[23]. These studies have focused on producing realistic soil moisture and other land surface fields for the recent periods for improving weather and seasonal climate forecasts. These studies generally used coarser global reanalysis data for forcing the land surface model. However, it has been shown that land surface model-simulated fields are sensitive to the large biases in the reanalysis forcing data. A central difficulty for performing global land simulations is a lack of realistic atmospheric forcing data for driving land surface models. The forcing data often require sub-daily data of precipitation, surface air temperature, specific humidity, wind speed and downward solar radiation, which are not readily available from observations. Gottschalck *et al.*[24] examined the sensitivity of various precipitation datasets on the quality of land surface data over the US and found very high sensitivity on the precipitation data. Recently, Boone *et al.*[25] performed an analysis of comparison of different land surface datasets over the African monsoon region under the African Monsoon Multidisciplinary Analysis (AMMA). Under this project, different land surface models were used to create high resolution (0.50°) land surface datasets over the African monsoon region. Contrary to the previous studies, they have used sub-daily forcing data (in place of daily forcing data) as well as TRMM rainfall data (in place of reanalysis precipitation data), which have improved the model simulations over the region.

In the present study, we report the development of a high resolution land surface process dataset over the South Asian monsoon region using an off-line land surface model. For developing the land surface dataset, the land surface model, High Resolution Land Data Assimilation System (HRLDAS)[26] was used with the forcing derived from observed TRMM rainfall and the ECMWF model atmospheric parameters at 0.5° resolution. As in the case of Boone *et al.*[25], we have also used sub-daily forcing data at every 1 h for simulating land surface processes over the South Asian monsoon region. For this exercise, an advanced version of the land surface model (version 3.3) was used, while for the GLDAS dataset, an earlier version model (version 2.7) was used. Another improvement was made using the MODIS land-use data in place of USGS land-use data. The major improvements in the development of the present data are higher spatial and temporal resolution of the results (at 0.5° spatial resolution and hourly time interval), use of sub-daily (hourly) forcing parameters including rainfall, use of MODIS land-use data in lieu of USGS land-use data and weekly varying vegetation fraction instead of monthly vegetation climatology.

Land surface model and methodology

For developing the land surface dataset, the model HRLDAS version 3.3 was used[26]. The model is based on the Noah land surface model[27,28]. An overview of the latest changes in the model is given by Ek *et al.*[28] and the detailed description of this model is documented by Mitchell[29] and available at http://www.emc.ncep.noaa.gov/mmb/gcp/noahlsm/. HRLDAS is an uncoupled (off-line) 2D-Noah land surface model, which is forced from an initial field of soil moisture and soil temperature by surface air temperature, specific humidity, surface pressure, *u* wind, *v* wind, shortwave and longwave radiation, and precipitation to update latest land surface conditions. Other inputs include time-varying vegetation characteristics (like green vegetation fraction) and static surface fields (land-use and soil texture data). HRLDAS uses the

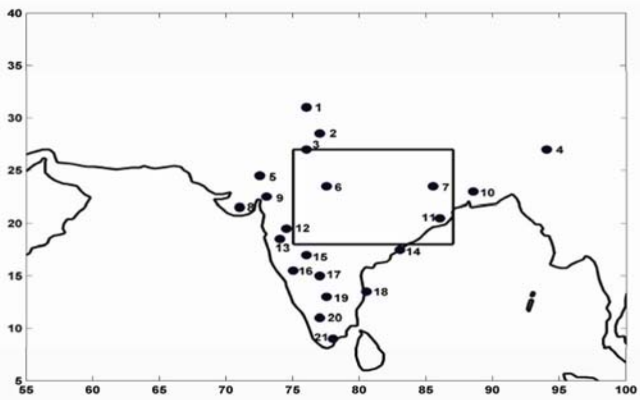

Figure 1. The model domain and observational sites used for validation of soil moisture and soil temperature.

Table 1. Forcing dataset used for the land surface data development

Data	Source
Atmospheric forcing	ECMWF model data at 0.5° resolution (*u* wind, *v* wind, specific humidity, surface pressure, solar radiation and downward longwave radiation)
Rainfall	TRMM 3B42 at 0.25° resolution
Green vegetation fraction	NOAA STAR weekly green vegetation fraction (GVF), NCEP monthly mean climatology
Land use	MODIS 30-s, USGS 24 categories
Soil texture	WRF input based on hybrid 30-s State Soil Geographic Database (STATSGO, now referred to as the US General Soil Map; for CONUS) and 5-min Food and Agriculture Organization (outside CONUS) 16-category soil texture
Terrain height	WRF input based on USGS-derived 30 s topographical height data

Table 2. The output variables from HRLDAS

Variable	Unit
Skin temperature	K
Soil temperature (four levels)	K
Soil moisture (four levels)	m^3m^{-3}
Surface run-off	mm
Ground run-off	mm
Interflow run-off	mm
Surface evaporation	mm
Evapotranspiration	mm
Canopy evaporation	mm
Direct soil evaporation	mm
Plant transpiration	mm
Water equivalent snow depth	m
Canopy drip	mm
Dew fall	mm
Surface sensible heat flux	Wm^{-2}
Surface latent heat flux	Wm^{-2}
Ground heat flux at surface	Wm^{-2}
Residual of surface energy balance	Wm^{-2}

weather research and forecasting (WRF) defined grids and static land surface parameters by default and can handle nested domains without any coupling between the domains. For the present study, HRLDAS was implemented using 0.5° horizontal grid (lat. × long.) and the model was integrated for the period from 1 January 2003 through 31 December 2010. The model domain covers the latitudes from 5°N to 40°N and longitudes from 55°E to 100°E (Figure 1). Four land surface levels at 7 cm, 28 cm, 100 cm and 2.55 m were considered for the computations. The model was forced at 1 h time interval and a time step of 15 min was used for the model integration. Table 1 shows the different forcing data used in the study. For rainfall, we used the TRMM 3B42 hourly rain rate[30], which combines calibrated microwave and infrared precipitation estimates. For atmospheric parameters, as recommended by Boone *et al*.[25], we used the ECMWF analysis and forecast data of *u* and *v* winds, surface specific humidity, surface pressure, solar radiation and downward longwave radiation. We preferred this dataset for forcing as the data are available at higher resolution (at 0.5° spatial resolution and at three hourly temporal resolution). For vegetation fraction, the weekly varying green vegetation fraction (GVF) from NOAA STAR[31] was used. This vegetation fraction is estimated from the normalized difference vegetation index (NDVI) observed by the advanced very high resolution radiometer (AVHRR). Land use and land cover was used from the MODIS 30 s data and soil texture from WRF STATSGO[32]. Terrain height was taken from WRF input USGS-derived 30 s data. Vegetation and soil parameters are based on look-up tables[27].

HRLDAS was initialized using the ECMWF atmospheric data from 1 January 2003 and integrated up to 31 December 2010. The first two years of data were not included in the analysis to avoid the model spin-up. To examine the sensitivity of simulations on land-use data and vegetation fraction, we have carried out additional model simulations. In the first experiment, we used the ECMWF predicted rainfall and TRMM observed rainfall data to examine the sensitivity of rainfall forcing on the model simulations. In the second experiment, we used weekly varying vegetation fraction instead of monthly vegetation climatology. For validation of soil moisture and soil temperature simulations, the India Meteorological Department (IMD) surface observations for the period 2005–2008 were used. These weekly data are available for a few stations spread over the Indian land region as shown in Figure 1. Soil moisture observations are taken using gravimetric methods. We converted from gravimetric per cent by mass to volumetric soil moisture by multiplying the ratio of soil density to water density. Table 2 shows the list of output variables from the HRLDAS available for the period 2005–2010.

Validation of results

To examine the quality of the HRLDAS data over the Indian summer monsoon region, we evaluated the land surface data during the period 2005–2008 using IMD surface observations of the soil moisture and soil temperature. Figure 1 shows the surface observation sites of IMD used for the validation purpose. For the validation, we used 21 surface observatories for the period June 2005 to December 2008. IMD observations are weekly averages of daily observations at local time 0700 IST (+01:30

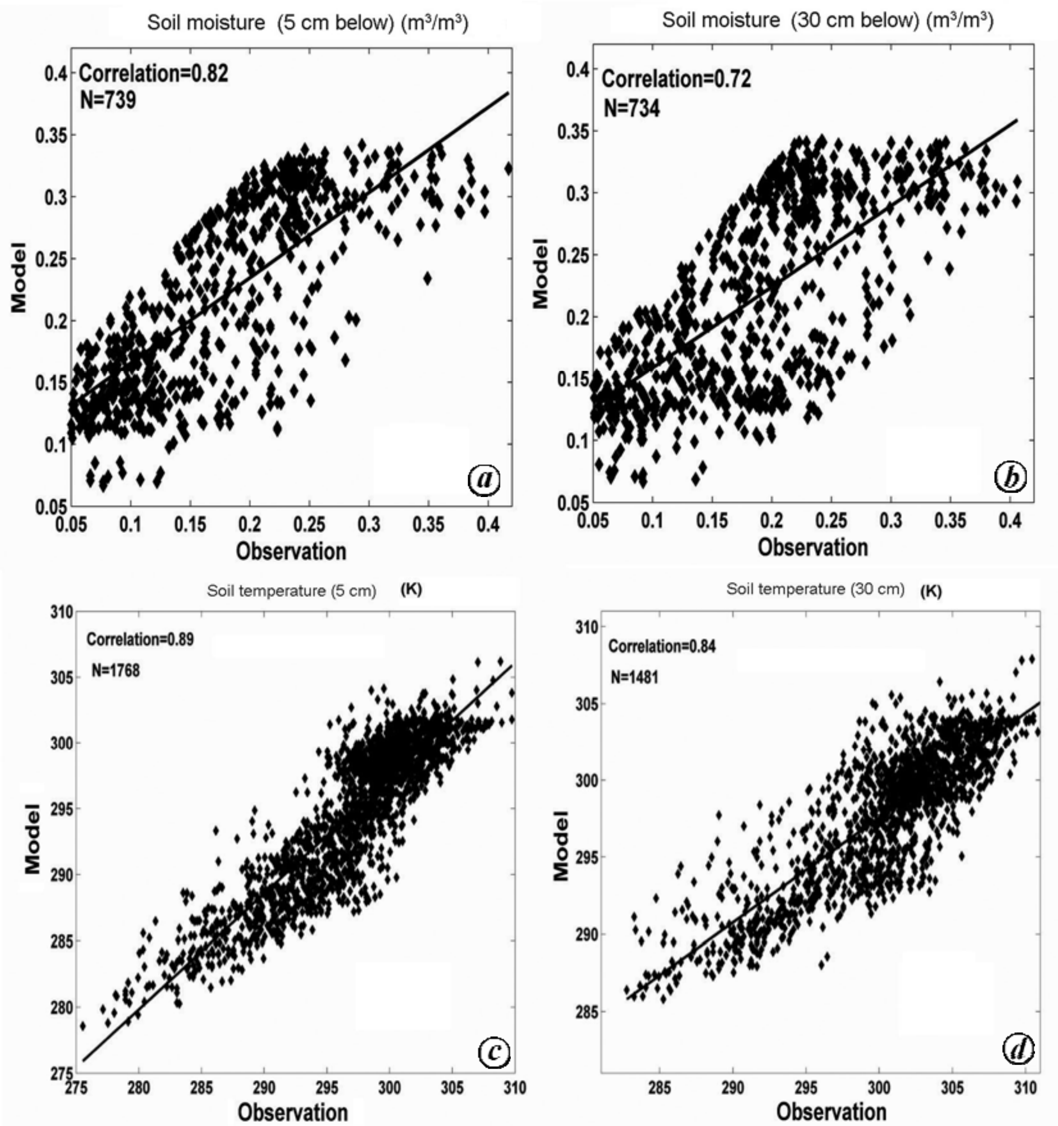

Figure 2. Scatter plot of observed and model soil moisture and soil temperature (annual). ***a***, ***b***, Soil moisture at (***a***) level 1, and (***b***) at level 2. ***c***, ***d***, Soil temperature at (***c***) level 1 and (***d***) at level 2.

UTC). The IMD soil moisture and soil temperature observations are available for 5 and 30 cm depths. Model simulation levels, however, are at 7 cm (level 1) and 28 cm (level 2). Therefore, the model data were interpolated to compare with the observations and hourly model data are averaged to weekly data. The soil moisture and soil temperature simulations at two levels (level 1 and level 2) of HRLDAS are compared against the IMD observations. Figure 2 *a* and *b* shows the scattering plots of annual soil moisture at level 1 and level 2 between the model simulations and observations. Correlations of model soil moisture data with observations are statistically significant at 99% level. Annually averaged soil moisture simulated by the model at level 1 (level 2) shows a significant correlation of 0.82 (0.72) with the IMD soil moisture observations. However, soil moisture at level 1 (level 2) during the monsoon (JJAS) season shows a weaker correlation of 0.67 (0.57) (Figure 3 *a* and *b*). The simulated soil moisture mainly depends on the forcing data and land surface model parameterization. Decrease in correlation during the monsoon season may be due to higher uncertainties in forcing data during the monsoon season, especially the TRMM rainfall data. The scatter plots also reveal systematic differences in the soil moisture simulations. Some differences could arise since model data are averaged over the particular grid, while the IMD observations are point observations. Even over a distance of 0.5° (55 km), soil moisture can vary substantially. Vertical interpolation of data can also contribute to some errors.

Model-simulated soil temperature shows far better correlation with observations compared to soil moisture.

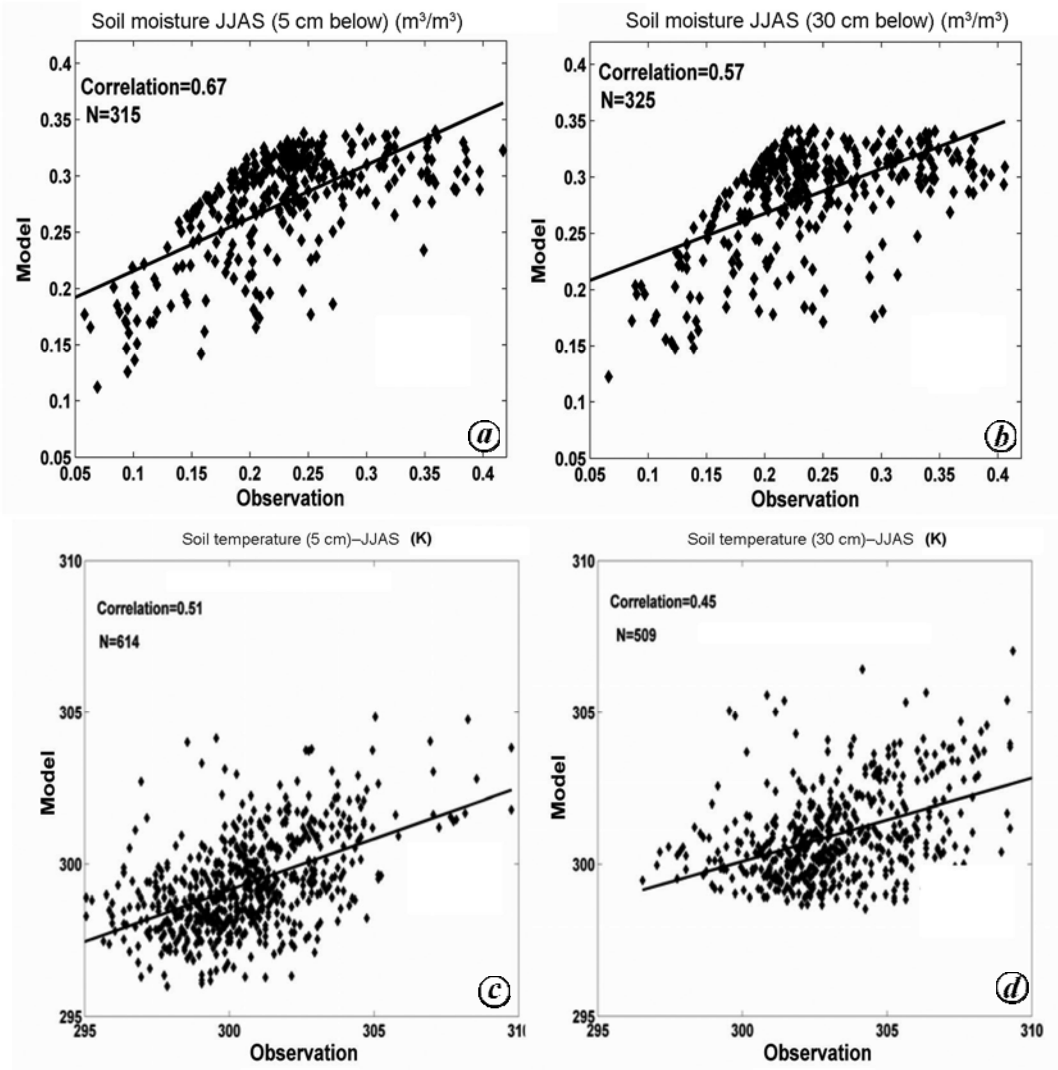

Figure 3. Scatter plot of observed and model soil moisture and soil temperature (June to September). ***a***, ***b***, Soil moisture at (***a***) level 1 and (***b***) at level 2. ***c***, ***d***, Soil temperature at (***c***) level 1 and (***d***) at level 2.

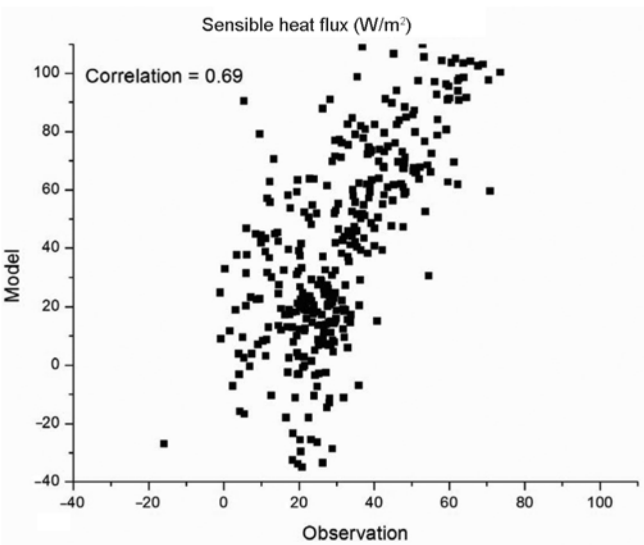

Figure 4. Comparison of sensible heat flux (Wm^{-2}) from the model and *in situ* observations at Ranchi.

Figure 2 *c* and *d* shows the comparison of soil temperature for all the seasons together. The correlation between the model simulated soil temperature and observed soil temperature at level 1 (level 2) is 0.89 (0.84), which is statistically significant at 99% level. During the monsoon season, soil temperature correlation (Figure 3 *c* and *d*) is 0.51 and 0.45 for level 1 and level 2 respectively. This evaluation shows that soil moisture and soil temperature data developed by HRLDAS are reasonably comparable with the observations and can be used for research studies. The sensible heat flux data are validated with *in situ* measurements taken at the Birla Institute of Technology, Ranchi (Figure 4) for the year 2009. The data for sensible heat flux were taken from the flux tower of the Land Surface Atmosphere and Micrometeorological Observational System (LATAMOS) (http://www.bitmesra.ac.in/). This comparison shows a correlation of 0.69, which is significant at 99% level. The sensible heat flux is also validated with the MERRA reanalysis (http://gmao.gsfc.nasa.gov/merra/).

Comparison of sensible heat flux of MERRA reanalysis over the central India region (box shown in Figure 1) shows a correlation of 0.90 (Figure 5). Since we do not have reliable long-term observations of other parameters and surface fluxes (sensible heat and latent heat), the validation of other parameters and surface fluxes could not be carried out. In the next section, the sensitivity of land surface data to the forcing data is discussed.

Sensitivity to forcing data

The simulations based on the land surface model depend on the quality of forcing data[33]. In the HRLDAS model, hourly precipitation and downward solar radiation play primary roles in driving the land modelling system and determining long-term evolution of soil moisture and temperature. In this section, we focus on the sensitivity of soil moisture and soil temperature to changes in the forcing of rainfall and GVF datasets. In the first experiment we made the sensitivity model simulations to rainfall forcing. Also, the ECMWF predicted rainfall (24-hour forecasts) was used for the simulations, which was replaced with the TRMM rainfall dataset. The results show that the use of TRMM rainfall for forcing the land surface model significantly improved the level 1 and level 2 soil moisture simulations. However, soil temperature simulations showed only a marginal improvement. The correlation of level 1 soil moisture with TRMM data (during the JJAS season) increased from 0.48 to 0.67 and for the level 2, it increased from 0.39 to 0.45.

In the second experiment, we examined the sensitivity of weekly varying GVF on land surface processes. Vegetation fraction is an important factor in the partition of surface energy. During the annual cycle, vegetation fraction at a particular location varies with time. It is necessary to examine the sensitivity of time evolution of GVF in land surface processes. Gutman and Ignatov[34] showed that satellite-derived GVF improve the surface flux in numerical weather prediction models. The sensitivity of land surface model to MODIS-based GVF was studied by Miller *et al*.[35], who found an appreciable impact of the MODIS GVF data on the surface energy and water balance during the summer season. To examine the sensitivity due to GVF, two simulations were performed. The first used the monthly mean climatology of GVF and the second used weekly varying satellite-derived GVF derived from satellite data. The results for the JJAS season showed that soil moisture simulations at both the levels have improved. The correlation of soil moisture at level 1 with weekly varying GVF improved from 0.52 to 0.67 and the level 2 soil moisture improved from 0.43 to 0.57 (Table 3). At the same time, soil temperature correlation at level 1 improved from 0.49 to 0.51 and at level 2 from 0.43 to 0.45.

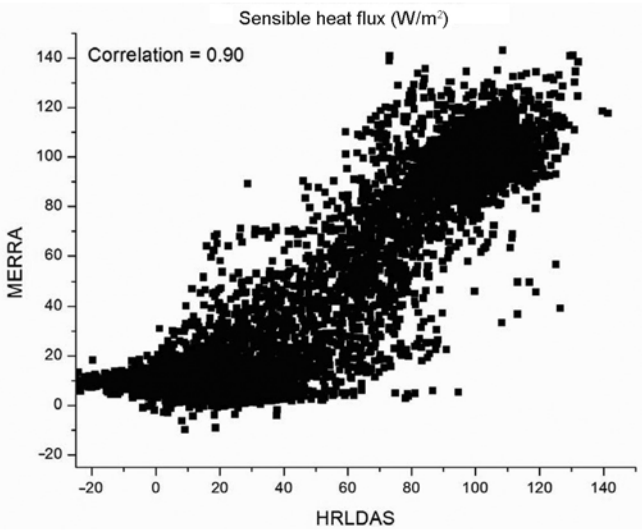

Figure 5. Comparison of sensible heat flux (Wm^{-2}) from the MERRA reanalysis and HRLDAS over central India.

Further, we examined the coupling of GVF in the simulations of soil moisture and soil temperature by calculating percentage of variance as in Zhang *et al*.[36]. The percentage of variance (PV) of monthly mean of variable (x) by vegetation coupling is calculated using the following equation

$$\mathrm{PV}_x = \frac{\sigma_x^2(\mathrm{WGVF}) - \sigma_x^2(\mathrm{CGVF})}{\sigma_x^2(\mathrm{WGVF})},$$

where x is soil temperature, soil moisture, latent heat flux and sensible heat flux and σ_x^2 is the monthly mean variance of x. WGVF refers to experiment without coupling and CGVF refers to the coupling experiment. This percentage of variance change calculated is related to vegetation coupling in the land surface processes. Percentage of variance of latent heat flux (LH), sensible heat flux (SH), top level soil moisture (SM) and soil temperature (ST) due to coupling of vegetation fraction is shown in Figure 6. Among the parameters, latent heat flux has the largest sensitivity to GVF. For soil moisture, the largest coupling is observed over the eastern parts of India, explaining more than 60% of variance. The largest coupling (40–50%) in soil temperature is observed over central and northwest India. The largest coupling in latent heat flux and sensible heat flux is observed over the northwest and north-central India. These results suggest a strong coupling of land surface processes with GVF over the Indian monsoon region. Therefore, for reliable estimation of land surface parameters, it is important to use varying vegetation fraction instead of monthly climatology.

Application of land surface dataset

In this section, we discuss some examples of application of the land surface dataset thus developed for the period 2005–2010.

Table 3. Correlation of simulated soil moisture and temperature with observations during the monsoon season (JJAS) with different forcing datasets

	ECMWF predicted rainfall	TRMM rainfall	GVF climatology	Weekly varying GVF
Correlation with observations (June–September)				
Soil moisture (level 1)	0.48	0.67	0.52	0.67
Soil moisture (level 2)	0.39	0.57	0.43	0.57
Soil temperature (level 1)	0.50	0.51	0.49	0.51
Soil temperature (level 2)	0.45	0.45	0.44	0.45

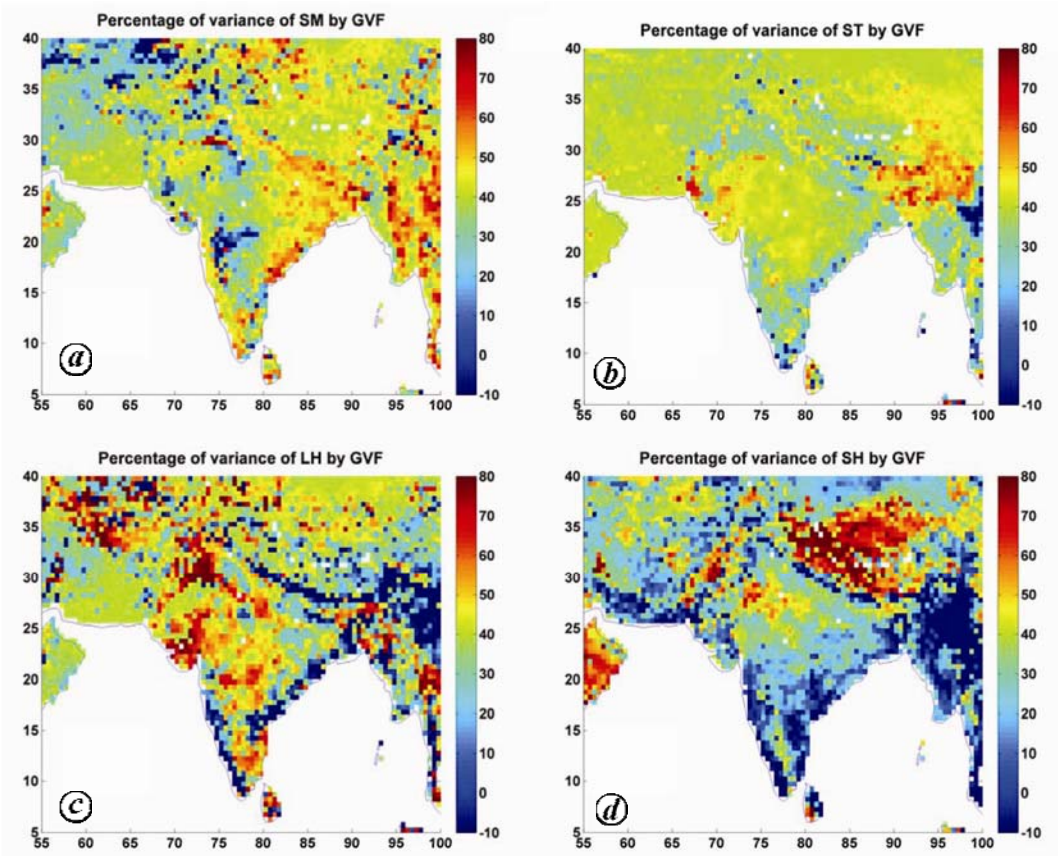

Figure 6. Percentage of variance of (***a***) soil moisture, (***b***) soil temperature, (***c***) latent heat flux and (***d***) sensible heat flux during monsoon season (JJAS) by weekly varying vegetation.

Spatial variations of mean surface fluxes

The spatial variation of soil temperature, soil moisture, latent heat flux and sensible heat flux, during the monsoon season (JJAS) is shown in Figure 7. Maximum soil moisture is observed over the southwest coast, central parts of India and NE India, which are the areas with heavy rainfall. Low soil moisture is observed over southeast India and northwest India ($0.15\ m^3\ m^{-3}$ or less), where rainfall is low during the monsoon season. Soil temperature (level 1) shows a maximum over extreme northwest India and adjoining Pakistan (305–310 K), where the seasonal heat low is observed. Another secondary maximum is observed over southeast India (304–305 K) where seasonal mean rainfall is very low. Minimum soil temperature values are observed over the west coast, NE and central India (296–300 K) where soil moisture is higher.

During the monsoon season, latent heat flux shows maximum values over the monsoon trough region (45–80 Wm^{-2}), NE India (80–90 Wm^{-2}) and southwest coast (40–75 Wm^{-2}). Minimum values are observed over northwest India (10–35 Wm^{-2}) and southeast India (30–60 Wm^{-2}). However, maximum sensible heat flux is observed over northwest India (80–120 Wm^{-2}) and southeast India (60–110 Wm^{-2}) where soil temperatures are also higher. Minimum values are observed over the southwest coast of India, eastern parts of central India (30–40 Wm^{-2}) and northeast India (20–30 Wm^{-2}). Surface flux values from this dataset are found comparable with those quoted in the previous studies[37,38].

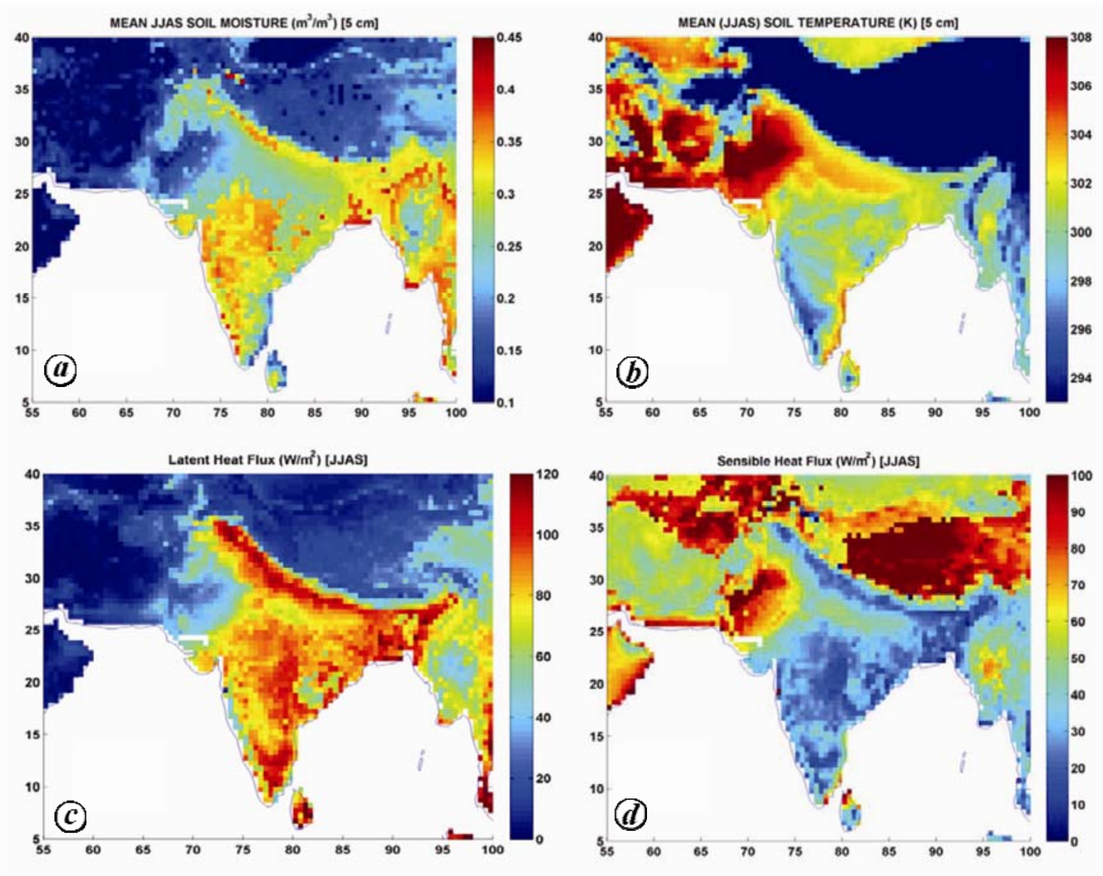

Figure 7. Mean (***a***) soil moisture, (***b***) soil temperature, (***c***) latent heat flux and (***d***) sensible heat flux during the monsoon season (JJAS) for the period 2005–2010.

Annual cycle of surface fluxes

Figure 8 shows the annual variation of soil moisture, soil temperature, latent heat flux and sensible heat flux averaged over central India (box shown in Figure 1). Soil moisture peaks in the JJAS season coinciding with the main rainy season over the region. However, large year-to-year variations are observed. The year 2009 was a major drought year with significant reduction of monsoon seasonal rainfall (23% deficiency). The effect of the major drought can be seen in the soil moisture variations (with lower values) during the 2009 monsoon season. The year 2010 was a good monsoon year and marked with higher soil moisture in post-monsoon season. Soil temperature peaks during the pre-monsoon season. Once the monsoon sets in, soil temperature falls appreciably. Soil temperature also shows year-to-year variations. Latent heat flux also peaks during the JJAS monsoon season like soil moisture, but with large year-to-year variations. Sensible heat flux, on the other hand, peaks up during the pre-monsoon season coinciding with the peak of soil temperature. Along with the onset of monsoon rains over the region, sensible heat flux reduces sharply due to reduction of soil temperature.

Surface fluxes during active and break monsoon spells

During the monsoon season, there are spells in which rainfall activity is stronger with excess rainfall, called the active spells. Similarly, break spells are observed with suppressed rainfall. We have considered active and break spells of the Indian monsoon in 2009 using the objective criteria proposed by Rajeevan *et al*.[39]. Following the criteria an active period was identified between 19 and 23 July and break period between 29 July and 10 August during the 2009 monsoon season. Figure 9 shows the time-series anomalies of rainfall, latent heat flux, sensible heat flux, soil moisture and soil temperature averaged over central India (box shown in Figure 1) during July and August 2009. Active period (i.e. 19–23 July) is marked with an increase of soil moisture but small decrease of soil temperature, sensible heat flux and latent heat flux. But, during the break period (29 July–10 August), there was a sharp decrease in soil moisture and latent heat flux. At the same time, sensible heat increased by 20 Wm^{-2} and soil temperature increased by 2 K due to less cloudy days and suppressed rainfall.

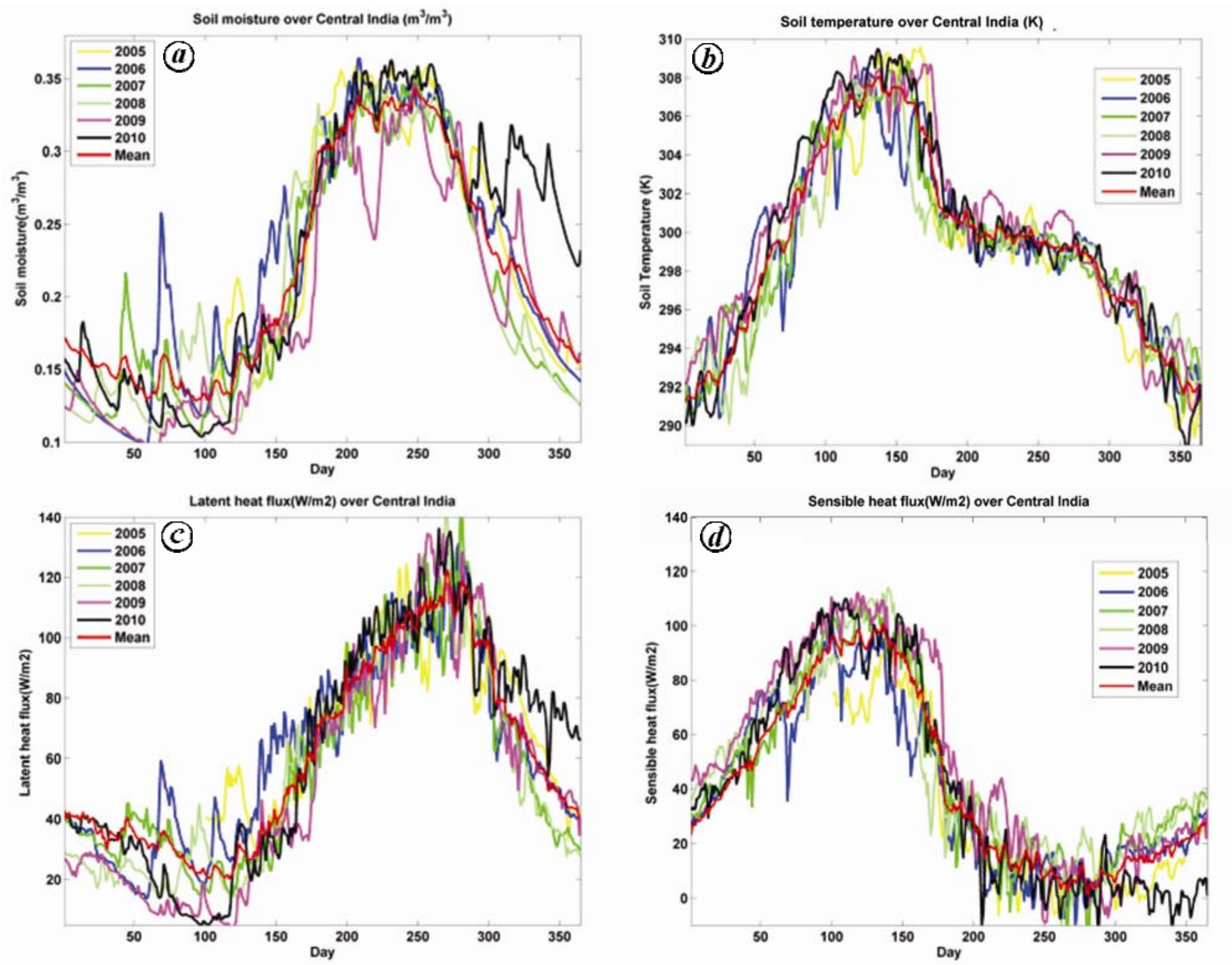

Figure 8. Annual cycle of (***a***) soil moisture, (***b***) soil temperature, (***c***) latent heat flux and (***d***) sensible heat flux averaged over central India during different years for the period 2005–2010.

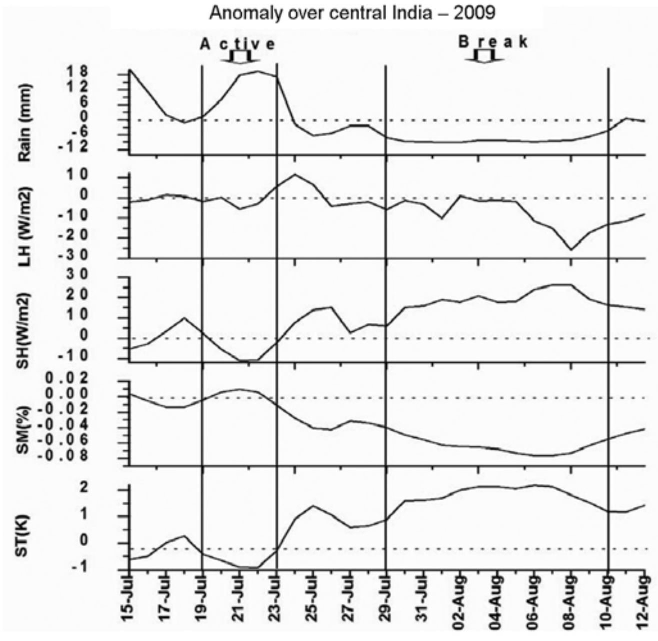

Figure 9. Time series of rainfall, latent heat flux, sensible heat flux, soil moisture and soil temperature averaged over central India during the active and break periods of 2009 monsoon season.

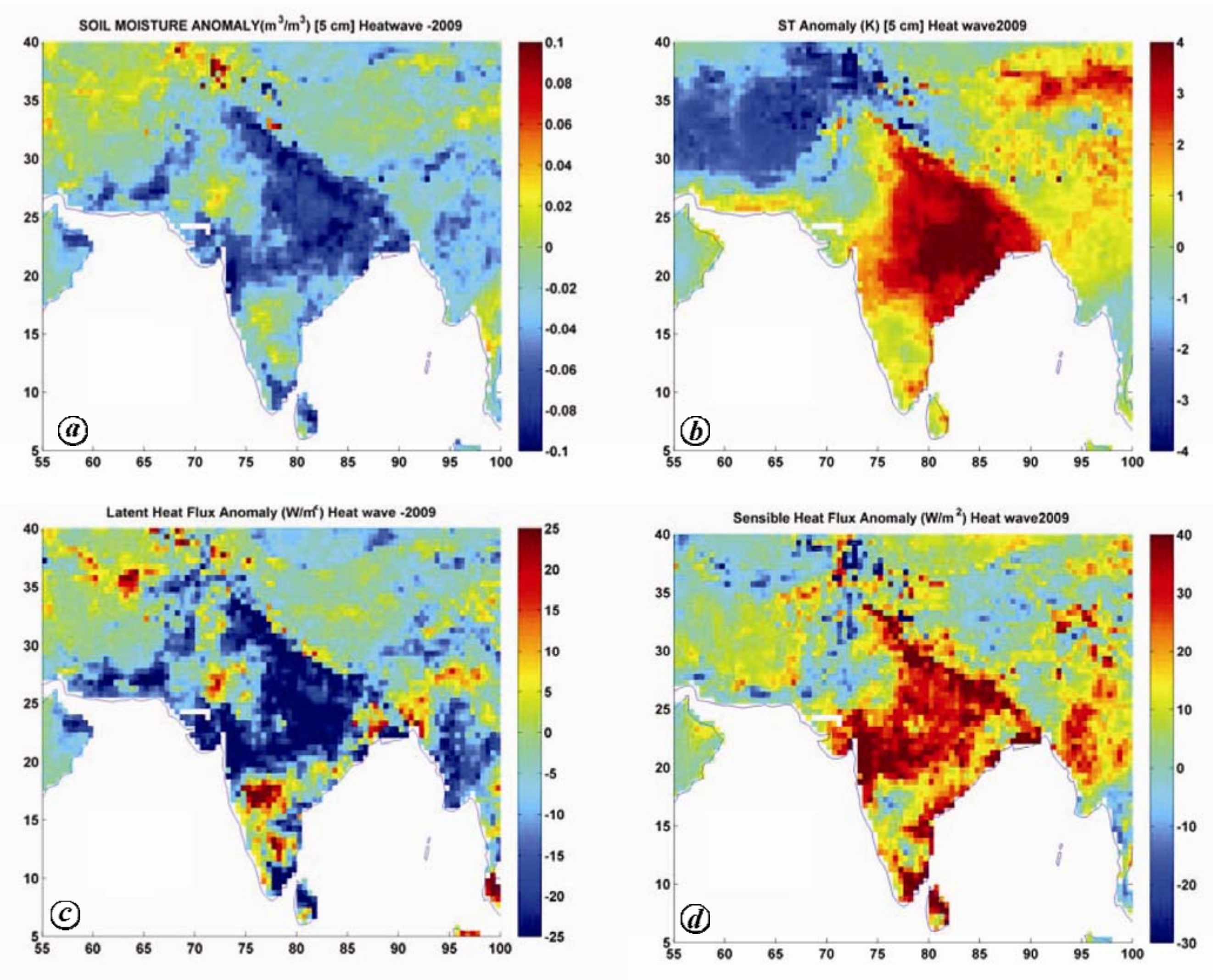

Figure 10. Spatial anomalies of (***a***) soil moisture, (***b***) soil temperature, (***c***) latent heat flux and (***d***) sensible heat flux during the June 2009 heat wave.

Surface fluxes during a heat wave

During the pre-monsoon season, heat waves with abnormal daytime temperatures are frequent over the northern parts of India. During June 2009, there was a heat wave (11–26 June), which affected different parts of central India with abnormally high temperatures. This abnormal heat wave occurred due to late monsoon onset over the region and the presence of an anti-cyclonic flow in the middle troposphere (not shown here). Figure 10 shows the spatial variation of anomalies of soil moisture, soil temperature, sensible heat flux and latent heat flux during the 2009 heat wave period. It is seen that during the heat wave, the soil moisture latent heat flux anomalies are negative over central India due to large scale drying of the atmosphere. At the same time, soil temperature and sensible heat flux were higher indicating the severity of the heat wave.

Discussion and conclusion

In this article, we have discussed the development of a high resolution land surface dataset for the South Asian monsoon region using an off-line land surface model (HRLDAS). The land surface dataset was developed for a 6-year period (2005–2010) at a 0.5° resolution and hourly time resolution. The HRLDAS model was forced by hourly TRMM rainfall, 3-hourly atmospheric parameters derived from ECMWF data, MODIS land-use data and weekly varying GVF. The dataset thus developed was compared with surface observations of soil moisture, soil temperature and sensible heat flux. The results suggest that the model simulations of soil moisture and temperature are comparable with the observations. The high resolution dataset can be effectively used for many applications. As an example, the dataset has been used to demonstrate the variations of soil moisture, soil temperature and surface flux associated with the monsoon active break spells and heat-wave conditions. One useful application of this dataset could be initialization of land surface conditions in numerical models to assess the impact of land surface processes on weather and climate predictions. Another application could be to understand the land surface–atmosphere coupling over the monsoon region and soil moisture feedback in the dynamics of active break phases of the Indian summer monsoon. The present dataset will be useful in the on-going Continental Tropical Convergence Zone experiment.

The present study has some caveats on the validation of the results. Due to want of adequate land surface data, a comparison was possible only for soil moisture and soil temperature. The surface sensible heat flux is compared with limited observations at Ranchi. There is no source of observations of surface fluxes (latent heat and sensible heat) from the region for validating them. While the model data are the average of whole grid (0.5°), IMD observations are point observations. These constraints highlight the need for development of land surface observations, including meso networks over the South Asian monsoon region. As the quality of the data products is sensitive to the forcing data, there is further scope to improve the quality of the data using better forcing data. The TRMM data may have some biases over land regions. The better option could be to make use of ground-based hourly rain-gauge data to force the land surface model. Since IMD has installed many automatic weather stations and rain gauges recently, it may be possible to develop a high resolution hourly rainfall dataset using ground-based observations. The Ministry of Earth Sciences, Government of India has taken an initiative to develop such a high resolution dataset, which is underway. We also have plans to extend the land surface dataset backwards so that long time-series data are developed and thus can be used to assess the impact of land surface conditions on seasonal prediction of monsoon rainfall. The present dataset is available for academic research on request.

1. Charney, J. G., Dynamics of deserts and drought in the Sahel. *Q. J. R. Meteorol. Soc*., 1975, **101**, 193–202.
2. Yasunari, T., Land–atmosphere interaction. In *The Asian Monsoon* (ed. Wang, B.), Springer, Praxis, Berlin, 2006, pp. 459–478.
3. Koster, R. D. *et al*., Regions of strong coupling between soil moisture and precipitation. *Science*, 2004, **305**, 1138–1140.
4. Koster, R. D. *et al*., GLACE: The Global Land–Atmosphere Coupling Experiment. Part I: Overview. *J. Hydrometeorol.*, 2006, **7**, 590–610.
5. Rajendran, K., Nanjundiah, R. and Srinivasan, J., The impact of surface hydrology on the simulation of tropical intraseasonal oscillation in NCAR (CCM2) atmospheric GCM. *J. Meteorol. Soc. Jpn. Ser. II*, 2002, **80**, 1357–1381.
6. Koster, R. D. *et al.*, The second phase of the Global Land–Atmosphere Coupling Experiment: soil moisture contributions to sub seasonal forecast skill. *J. Hydrometeorol*., 2011, **12**, 805–822.
7. Dirmeyer, P. A., Schlosser, C. A. and Brubaker, K. L., Precipitation, recycling and land memory: an integrated analysis. *J. Hydrometeorol*., 2009, **10**, 278–288.
8. Zhang, J., Wu, L. and Dong, W., Land atmosphere coupling and summer climate variability over East Asia. *J. Geophys. Res*., 2011, **116**, D05117. doi:10.1029/2010JD014714.
9. Betts, A. *et al*., The land surface atmosphere interaction: a review based on observational and global modeling perspective. *J. Geophys. Res.*, 1996, **101**, 7209–7225.
10. Shukla, J. and Mintz, Y., The influence of land surface evapotranspiration on the earth's climate. *Science*, 1982, **214**, 1498–1501.
11. Meehl, G. A., Influence of the land surface in the Asian summer monsoon: external conditions versus internal feedbacks. *J. Climate*, 1994, **7**, 1033–1049.
12. Ferranti, L., Slingo, J. M., Palmer, T. N. and Hoskins, B. J., The effect of land-surface feedbacks on the monsoon circulation. *Q. J. R. Meteorol. Soc.*, 1999, **125**, 1527–1550.
13. Webster, P. J., Mechanisms of low-frequency variability: surface hydrological effects. *J. Atmos. Sci.*, 1983, **40**, 2110–2124.
14. Saha, S. K., Halder, S., Kumar, K. K. and Goswami, B. N., Pre-onset land surface processes and 'internal' interannual variabilities of the Indian summer monsoon. *Climate Dyn*., 2011, **36**, 2077–2089.
15. Saha, S. K., Halder, S., Rao, A. S. A. and Goswami, B. N., Modulation of ISOs by land–atmosphere feedback and contribution to the interannual variability of Indian summer monsoon. *J. Geophys. Res.*, 2012, **117**, D13, 16.
16. Fennessy, M. and Shukla, J., Impact of initial soil wetness on seasonal atmospheric prediction. *J. Climate*, 1999, **12**, 3167–3180.
17. Douville, H., Relevance of soil moisture for seasonal atmospheric predictions: is it an initial value problem? *Climate Dyn*., 2004, **22**, 429–446.
18. Douville, H., Relative contribution of soil moisture and snow mass to seasonal climate predictability: a pilot study. *Climate Dyn*., 2010, **34**, 797–818.
19. Douville, H., Conil, S., Tyteca, S. and Voldoire, A., Soil moisture memory and West African monsoon predictability: artifact or reality? *Climate Dyn*., 2007, **28**, 723–742.
20. Qian, T., Dai, A., Trenberth, K. E. and Oleson, K. W., Simulation of global land surface conditions from 1948 to 2004. Part I: forcing data and evaluations. *J. Hydrometeorol.*, 2006, **7**, 953–975.
21. Dirmeyer, P. A., Dolman, A. J. and Sato, N., The pilot phase of the global soil wetness project. *Bull. Am. Meteorol. Soc*., 1999, **80**, 851–878.
22. Mitchell, K. E., The multi-institution North American land data assimilation system (NLDAS): utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.*, 2004, **109**, D07S90.
23. Rodell, M. *et al*., The global land data assimilation system. *Bull. Am. Meteorol. Soc.*, 2004, **85**, 381–394.
24. Gottschalck, J., Meng, J., Rodell, M. and Houser, P., Analysis of multiple precipitation products and preliminary assessment of their impact on global land data assimilation system land surface states. *J. Hydrometeorol.*, 2005, **6**, 573–598.
25. Boone, A. *et al*., The AMMA land surface model intercomparison project (ALMIP). *Bull. Am. Meteorol. Soc.*, 2009, **90**, 1865–1880.
26. Chen, F. *et al*., Description and evaluation of the characteristics of the NCAR high-resolution land data assimilation system. *J. Appl. Meteorol. Climatol*., 2007, **46**, 694–713.
27. Chen, F. and Dudhia, J., Coupling an advanced land surface hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Weather Rev.*, 2001, **129**, 569–585.
28. Ek, M. B. *et al*., Implementation of Noah land surface model advances in the National Centers for environmental prediction operational mesoscale Eta model. *J. Geophys. Res*., 2003, **108**, 8851.

29. Mitchell, K., The Community Noah Land Surface model: User's Guide. Public Release Version 2.7.1, NCEP/EMC, 2005; http://www.ral.ucar.edu/research/land/technology/lsm/noah/Noah_LSM_USERGUIDE_2.7.1.pdf
30. Huffman, G. *et al*., The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, 2007, **8**, 38–55.
31. Jiang, L. *et al.*, Real-time weekly global green vegetation fraction derived from advanced very high resolution radiometer-based NOAA operational global vegetation index (GVI) system. *J. Geophys. Res.*, 2010, **115**, D11114.
32. Miller, D. A. and White, R. A., A conterminous United States multilayer soil characteristics data set for regional climate and hydrology modeling. *Earth Interactions*, 1998, 1–26.
33. Luo, L. *et al*., Validation of the North American Land Data Assimilation System (NLDAS) retrospective forcing over the southern Great Plains, *J. Geophys. Res.*, 2003, **108**, 8843.
34. Gutman, G. and Ignatov, A., Satellite-derived green vegetation fraction for the use in numerical weather prediction models. *Adv. Space Res.*, 1997, **19**, 477–480.
35. Miller, J., Barlage, M., Zeng, X., Wei, H., Mitchell, K. and Tarpley, D., Sensitivity of the NCEP/Noah land surface model to the MODIS green vegetation fraction dataset. *Geophys. Res. Lett.*, 2006, **33**, L13404.
36. Zhang, J., Wang, W. C. and Leung, L. R., Contribution of land atmosphere coupling to summer climate variability over the contiguous United States. *J. Geophys. Res.*, 2008, **113**, D22109.
37. Urankar, G., Prabha, T. V., Pandithurai, G., Pallavi, P., Achuthavarier, D. and Goswami, B. N., Aerosol and cloud feedbacks on surface energy balance over selected regions of the Indian subcontinent. *J. Geophys. Res.*, 2012, **117**, D04210.
38. Bhat, G. S. and Arunchandra, S. C., On the measurement of the surface energy budget over a land surface during the summer monsoon. *J. Earth Syst. Sci*., 2008, **117**, 911–923.
39. Rajeevan, M., Gadgil, S. and Bhate, J., Active and break spells of the Indian summer monsoon. *J. Earth Syst. Sci.*, 2010, **119**, 229–247.

ACKNOWLEDGEMENTS. We thank the National Atmospheric Research Laboratory, Gadanki, India for support to carry out this study; Dr Jimy Dhudia, NCAR, USA for useful discussions on land surface modelling and the India Meteorological Department, Pune for providing soil moisture and soil temperature data. The TRMM rainfall data used in this study were generated with the Giovanni on-line data system, developed and maintained by NASA GES DISC. MERRA data used in this study were provided by Giovanni (GES, NASA). We also thank NOAA NESDIS STAR, NCEP and ECMWF for providing the datasets used in this study, and the anonymous reviewers for their useful suggestions and comments that helped improve the manuscript.

Received 22 May 2013; revised accepted 10 July 2013