International Journal of Remote Sensing
Vol. 32, No. 23, 10 December 2011, 8033–8051

Taylor & Francis
Taylor & Francis Group

Soil hydraulic parameters estimated from satellite information through data assimilation

SUJITTRA CHAROENHIRUNYINGYOS*†, KIYOSHI HONDA†, DAROONWAN KAMTHONKIAT‡ and AMOR V. M. INES§

†School of Engineering and Technology, Remote Sensing and Geographic Information Systems, Asian Institute of Technology, Bangkok, Thailand

‡Department of Geography, Faculty of Liberal Arts, Thammasat University, Bangkok, Thailand

§International Research Institute for Climate and Society, The Earth Institute at Columbia University, Palisades, New York, USA

(*Received 12 March 2009; in final form 27 August 2010*)

Leaf area index (LAI) and actual evapotranspiration (ET_a) from satellite observations were used to estimate simultaneously the soil hydraulic parameters of four soil layers down to 60 cm depth using the combined soil water atmosphere plant and genetic algorithm (SWAP–GA) model. This inverse model assimilates the remotely sensed LAI and/or ET_a by searching for the most appropriate sets of soil hydraulic parameters that could minimize the difference between the observed and simulated LAI (LAI_{sim}) or simulated ET_a (ET_{asim}). The simulated soil moisture estimates derived from soil hydraulic parameters were validated using values obtained from soil moisture sensors installed in the field. Results showed that the soil hydraulic parameters derived from LAI alone yielded good estimations of soil moisture at 3 cm depth; LAI and ET_a in combination at 12 cm depth, and ET_a alone at 28 cm depth. There appeared to be no match with measurement at 60 cm depth. Additional information would therefore be needed to better estimate soil hydraulic parameters at greater depths. Despite this inability of satellite data alone to provide reliable estimates of soil moisture at the lowest depth, derivation of soil hydraulic parameters using remote sensing methods remains a promising area for research with significant application potential. This is especially the case in areas of water management for agriculture and in forecasting of floods or drought on the regional scale.

1. Introduction

Soil moisture is an important state variable in the critical zone (Brantley *et al.* 2006) that provides linkage to the lithosphere, hydrosphere and atmosphere. It controls the partitioning of energy and mass in the vadose zone hence also determines the rate of plant transpiration, soil evaporation, runoff and recharge (Georgakakos 1996). Therefore, an understanding of the spatio-temporal properties of soil moisture and its physical control is essential for a wide range of applications, for example, in the areas of meteorology, agriculture and hydrology, especially for early warning

*Corresponding author. Email: sujittra_geo@hotmail.com

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2011 Taylor & Francis
http://www.tandf.co.uk/journals
http://dx.doi.org/10.1080/01431161.2010.532170

systems (e.g. flood or drought) as well as in forecasting agriculture production. Spatio-temporal attributes of soil moisture depend greatly on variations in precipitation, soil properties, topographic features and vegetation characteristics (Das and Mohanty 2006). Soil hydraulic properties are important physical characteristics of soil that determine its ability to transport and retain water. Hence, quantification of these properties is crucial for modelling the surface/subsurface hydrological processes.

Remote sensing systems using both active and passive sensors are able to observe soil moisture at a regional scale but could not provide data at greater depths below the soil surface (Jackson *et al.* 1995). As a result, it has not been possible to provide soil moisture information in the root zone, knowledge of which is needed for agricultural and water management applications. Agro-hydrological and hydrological models, on the other hand, can simulate soil moisture in the root zone, including spatio-temporal variations, but these models require site-specific parameters (physical and empirical) describing the soil and vegetation properties in order to simulate soil moisture. Likewise, model parameters (e.g. soil hydraulic parameters) are difficult to measure in the field, restricting their application at the regional level. If the necessary parameters could be obtained from remote sensing data, it would be an important contribution to the simulation of soil moisture for numerous large scale applications in agriculture and water management.

Data assimilation is a technique of data integration between a model and external sources (e.g. remote sensing data) for estimating parameters and/or improving the state of the model variables used for simulations (Swinbank *et al.* 2003). Quantification of soil hydraulic properties can be accomplished by integrating easily observable variables into a simulation model using data assimilation techniques (Vrugt *et al.* 2003, Mertens *et al.* 2006, Ines and Mohanty 2008). The advantage with data assimilation is that it reduces the problems that arise from incorporating measurement data that are inadequate, insufficient or unclear.

Several studies have attempted to optimize the land surface model's soil parameters through data assimilation by using different reference data sources to adjust the relevant model parameters. Many researchers (e.g. Enthekabi *et al.* 1994, Georgakakos 1996, Gupta *et al.* 1999, Heathman *et al.* 2003, Moran *et al.* 2004, Hogue *et al.* 2005, Liu *et al.* 2005, Ines and Mohanty 2008) have used near-surface soil moisture or soil temperature from remote sensing data and/or direct measurements from the field to optimize soil parameters. Ines and Droogers (2002) found that the soil moisture profile is a more stable indicator for quantifying the effective soil hydraulic parameters of the root zone because the two are directly related. However, the estimated soil hydraulic parameters inferred from such field observations may be difficult to apply in large scale situations, while remotely sensed vegetation indices or fluxes could be more effective.

This study aims to quantify effective soil hydraulic parameters in the soil profile for soil moisture simulation using the time series of remote sensing data, namely the actual evapotranspiration (ET_a) and leaf area index (LAI), based on the inversion of a physically based soil water atmosphere plant (SWAP) model (Van Dam 2000) with a genetic algorithm (GA, Goldberg 1989). For the inverse estimation, the information contained in the remote sensing data was tested in order to derive the subsurface soil hydraulic properties using, independently or in combination, the satellite-derived data. This method was applied in a rice cropping area in Ubon Ratchathani Province of Thailand.

2. Study area

Ubon Ratchathani Province covers an area of 16 113 square kilometers and is located in the lower part of the north-east region of Thailand (see figure 1). Rice is the main crop in this area. The study site is located in the upper part of Ubon Ratchathani Province in Trakan Phutphon District (15° 42′ 22″ N, 105° 00′ 21″ E). Most of the farmers in this area can grow rice in only one cropping per year because their fields are mostly rainfed. The cropping cycle starts with growing of the seedlings at the beginning of the rainy season from late April to late May. Transplanting of rice seedlings into the flooded fields occurs one month later, typically in late June. Rice is usually harvested around late October to mid November, after the fields have been drained.

Figure 1. Study site in Trakan Phutphon District, Ubon Ratchathani Province, Thailand.

3. Materials

3.1 *The soil water atmosphere plant–genetic algorithm*

The SWAP model is a physically based agro-hydrological model for simulating the relationships among soil, water, atmosphere and plants (Van Dam 2000). The model uses the Richards equation to simulate the vertical soil water movement. Soil hydraulic functions in the model are defined by the Mualem–Van Genuchten (MVG) equations (Mualem 1976, Van Genuchten 1980), that describe the capacity of the soil to store, release and transmit water under different environmental and boundary conditions (see equations (1) and (2)):

$$S_e = \frac{\theta(h) - \theta_{res}}{\theta_{sat} - \theta_{res}} = \left[\frac{1}{1 + |\alpha h|^n}\right]^m, \text{and} \tag{1}$$

$$K(h) = K_{sat} S_e^{\lambda} \left[1 - \left(1 - S_e^{1/m}\right)^m\right]^2, \tag{2}$$

where S_e is the relative saturation (–); θ_{res} and θ_{sat} are the residual and saturated soil water contents ($cm^3\ cm^{-3}$); α (cm^{-1}), n (–), m (–) and λ (–) are the shape parameters of the retention and the conductivity functions; K_{sat} is the saturated hydraulic conductivity (cm day^{-1}); $m = 1 - n^{-1}$. The values of these parameters are distinct among the soil (textural) types and have to be defined as inputs for the simulation model. They are considered as unknown parameters for this study. Potential evapotranspiration (ET_p) in the SWAP model is calculated using the Penman–Monteith equation. Considering actual environmental conditions, the actual evapotranspriation (ET_a) is calculated by applying the root water uptake reduction factor due to water and/or salinity stress to the ET_p, if present. To simulate crop growth, SWAP uses the world food studies (WOFOST) model. Leaf area index (LAI), calculated on a daily basis, is determined by the fraction of carbon assimilates that are partitioned into the leaves, as a result of photosynthetic activity (Van Dam 2000).

A genetic algorithm (GA) is a mathematical model based on natural genetics, where the mechanics of nature have been abstracted to be used in search and optimization problems. GA techniques are different from other optimization techniques in a number of significant ways. A GA consists of three basic operators, selection, crossover and mutation, which are iteratively processed over many generations to produce the best individual(s) that can represent the optimal solution to a problem (Goldberg 1989). The genetic algorithm used in this study is the modified microGA, which is considered to be powerful in a wide range of situations for solving optimization problems (Ines and Droogers 2002, Ines and Mohanty 2008).

SWAP–GA is a combination of soil water atmosphere plant (SWAP) and genetic algorithm (GA) models taking advantage of the strength of the genetic algorithm to determine the unknown model parameters by using the information from satellite data to optimize an objective function. Usually, the difference between a simulated and observed variable is considered as a function of the unknown parameters. Difficult input parameters in the SWAP, such as soil hydraulic parameters, are estimated by assimilating the remote sensing data (LAI and ET_a) into the model. In turn, SWAP was used to provide continuous estimates of soil moisture at a high temporal resolution and validated using field observations. The SWAP–GA model was then run until

the best fit between the model simulations (ET_a and LAI, independently or combination) and observations (Satellite ET_a and LAI, independently or combination) was obtained. Details of the field observation system where SWAP–GA was applied are given below.

3.2 *Instruments and devices*

A near real-time weather station and flux observation system were installed in a paddy field at Trakan Phutphon District in Ubon Ratchathani Province with financial support from the Thailand Research Fund (TRF). The instruments used at the meteorological stations in the study area are shown in figure 2. Different sensors were set up to measure the general weather, energy flux and soil moisture. A General Weather Station–Davis Vantage Pro (ExploraTrack.com, Cannon Beach, OR, USA) was installed at 2 m, along with two Bowen thermocouple instruments to measure the Bowen ratio at the heights of 2 and 10 m, in order to determine the actual evapotranspiration (ET_a). Data from the different sensors were stored in a database in the control box, which could be remotely accessed via an IPStar satellite link. A sensor control box for the whole system, built on a Linux server, was also installed. A data acquisition system (DAQs) from National Instrument (NI) was set up to record the data from the sensors. A net radiometer was set up at a height of 2.5 m for recording the net radiation. In addition to the Davis automatic weather station, another wind speed sensor and rain gauge were also installed at a height of 10 m. Underground soil moisture sensors were installed at 3, 12, 28 and 60 cm soil depths.

3.3 *Meteorological data*

Meteorological data, comprising the daily average minimum and maximum temperatures, humidity, solar radiation, wind speed and rainfall, were required for the SWAP model. The meteorological data for the study area were obtained from the weather

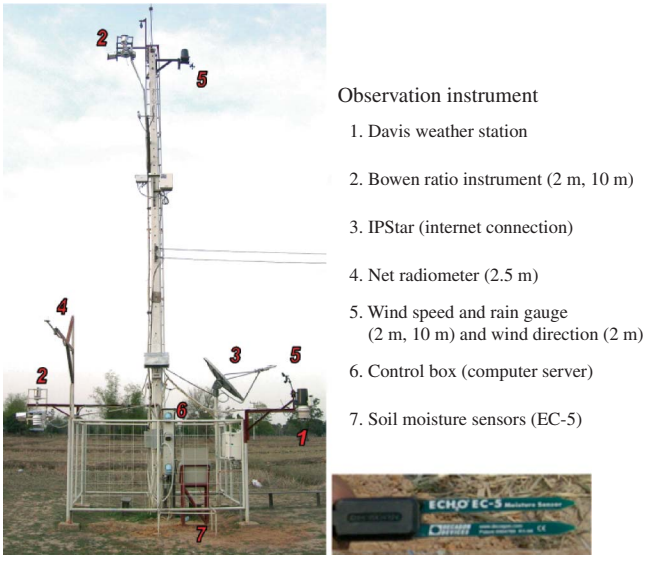

Figure 2. Observation system and configurations.

station installed in the paddy field. However, there were certain periods when the need for system maintenance meant that data were unavailable from these systems. At those times, meteorological data from the Thai Meteorological Department (TMD) Agro-meteorological station in Ubon Ratchathani Province were used.

3.4 *Soil data and soil moisture measurement*

Site soil texture data were required as input for the model, but the information needed was site specific. Global soil data are available from the Food and Agriculture Organization (FAO) map (FAO 1991), and at the provincial level from the Thailand soil map provided by the Land Development Department. In this study, soil texture was analysed in the laboratory, the results of which are shown in table 1.

Observed site soil moisture data were used for the validation purposes. The soil moisture profile was observed using EC-5 soil moisture sensors (Decagon Devices Incorporated, Pullman, WA, USA) at four vertical locations covering the rooting zone of crops (3, 12, 28 and 60 cm). These devices measure the dielectric constant or permittivity of a material by finding the rate of change of voltage in the sensor that is embedded in the porous medium. The equations to convert the voltage into volumetric water content at the four different soil depths, which were determined from the soil samples at each level, were based on Singh-Gill (2007). The equations obtained for the different soil depths are as follows:

$$\text{at } 3\,\text{cm} \quad y = 1.6444x - 0.6292, \text{with } R^2 = 0.99, \tag{3}$$

$$\text{at } 12\,\text{cm} \quad y = 1.4306x - 0.559, \text{with } R^2 = 0.98, \tag{4}$$

$$\text{at } 28\,\text{cm} \quad y = 1.3462x - 0.4967, \text{with } R^2 = 0.97 \tag{5}$$

$$\text{and at } 60\,\text{cm} \quad y = 1.666x - 0.6573, \text{with } R^2 = 0.98, \tag{6}$$

where x is the raw voltage (mV) and y is the volumetric water content ($\text{cm}^3\ \text{cm}^{-3}$).

3.5 *Crop data*

Crop data collected in the paddy field covered the whole period of the cropping season (July–November), from land preparation, transplanting and flowering, up to the ripening stage of the rice. LAI was measured in the field once or twice per week using the LAI meter (LAI-2000) from LI-COR Company (Lincoln, NE, USA). An LAI meter is a tool for measuring the extent of the plant canopy, as it directly affects the interception and absorption of light for plant growth. LAI is computed using a

Table 1. Soil texture analysis of Trakan Phutphon rice field.

Soil layer	Sand (%)	Silt (%)	Clay (%)	Textural class	Organic matter (g cm^{-3})
3 cm	39	49	12	Loam	1.14
12 cm	32	44	24	Loam	0.93
28 cm	32	36	32	Clay loam	0.66
60 cm	16	31	53	Clay	0.31

model of radiative transfer in vegetative canopies. Measurements of LAI were collected within the paddy field from 20 to 30 locations around the weather station, and then averaged to represent the LAI on the day of measurement. Crop information, such as crop height (measured in the field every week), planting and harvesting dates, and yields, was also recorded.

3.6 *Satellite data*

Satellite remote sensing is widely accepted as one of the most promising sources of data for agriculture research and land surface studies. Satellite imaging allows the land surface to be monitored (temporally), and information derived on a regional scale at much lower cost than manually collecting field measurements. The moderate resolution imaging spectroradiometer (MODIS) was chosen for this study because it provides high temporal resolution, 36 multispectral bands and can deliver a range of derived products (leaf area index, surface reflectance, surface temperature and surface emissivity) suitable for data assimilation in this study. In addition, MODIS is available free of charge through the internet.

3.6.1 Leaf area index. The leaf area index (LAI) and fraction of photosynthetically active radiation (FPAR) MODIS 8 day composite (MOD15A2) with spatial resolution of 1 km was used in this study. The satellite LAI is retrieved by multiplying the MOD15A2 product by the factor 0.1. Equation (7) is the general equation that was used to convert MOD15A2 into LAI values:

$$\mathrm{LAI}_{(x,y)} = \mathrm{MOD15A2}_{(x,y)} \times 0.1, \tag{7}$$

where $\mathrm{LAI}_{(x,y)}$ is the leaf area index, $\mathrm{MOD15A2}_{(x,y)}$ is the original digital number, and 0.1the multiplication factor for LAI.

A time series of LAI covering the whole period of the cropping season was retrieved based on the pixel at the same location as the weather station in the field. The LAI derived from MODIS showed slightly lower values than the observed LAI (see figure 3), but a comparison between the two sets of values showed a similar increasing trend. At the beginning of the growing period, LAI from the satellite and ground measurements were nearly the same. After mid July 2007, LAI from ground measurements showed higher values than the satellite-derived LAI. This mismatch in data could be due to the difference in spatial resolution (field vs pixel) of the LAI measurements. Nevertheless, the satellite and field measurements showed strong linear correspondence with a coefficient of determination $R^2 = 0.89$ and with root mean squared error (RMSE) = 0.19. Since the LAI output of SWAP was calibrated at the field level, the conversion of the MODIS LAI product to this level was also necessary for the assimilation process. Without this conversion, the resulting soil hydraulic parameters may not be representative of the actual field data. LAI from Satellite ($\mathrm{LAI}_{\mathrm{sat}}$) was calibrated with field observations using a hand-held LAI meter. The conversion equation of LAI from MODIS ($\mathrm{LAI}_{\mathrm{sat}}$) to LAI from Site ($\mathrm{LAI}_{\mathrm{site}}$) is given in equation (8):

$$\mathrm{LAI}_{\mathrm{site}} = 1.2384(\mathrm{LAI}_{\mathrm{sat}}) \times 0.1859. \tag{8}$$

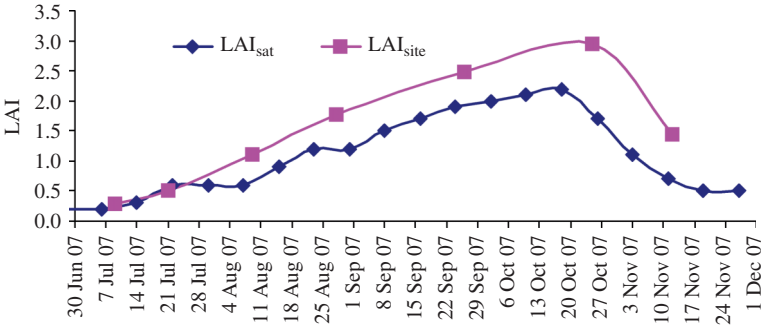

Figure 3. Comparison between LAI_{sat} and LAI_{site}.

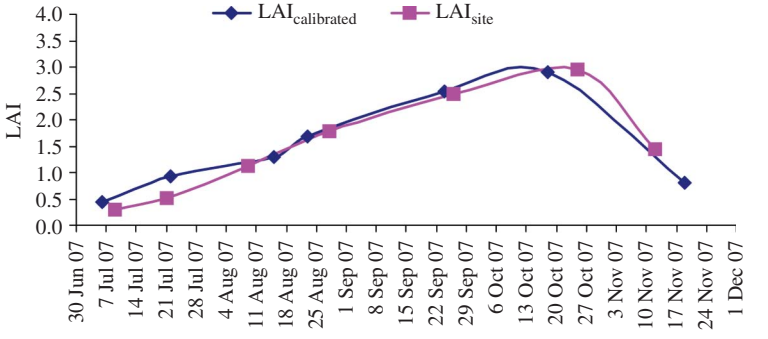

Figure 4. Comparison between LAI_{site} and $LAI_{calibrated}$.

After applying equation (8), the LAI values from the satellite after calibration ($LAI_{calibrated}$) were found to be a close match with the LAI values measured from the site (see figure 4). The time series of LAI values were then used as conditioning data in the assimilation process.

3.6.2 Actual evapotranspiration (ET_a) calculated by the surface energy balance algorithm for land (SEBAL). The surface energy balance algorithm for land (SEBAL) (Bastiaanssen *et al.* 1998a, b) is a remote sensing model for estimating daily actual evapotranspiration (ET_a) of land surfaces. SEBAL calculates the instantaneous and 24 h surface heat flux. SEBAL is based on the well-known surface energy balance equations: $R_n = K_\downarrow - K^\uparrow + L_\downarrow - L^\uparrow$ and $R_n = H + G + \lambda_{ET}$, where R_n is net radiation, $K_\downarrow$ and $K^\uparrow$ are incoming and outgoing shortwave radiation, $L_\downarrow$ and $L^\uparrow$ are incoming and outgoing longwave radiation, H is sensible heat flux, G is soil heat flux, and λ_{ET} is latent heat flux; all units in Wm^{-2}.

Remotely sensed estimates of surface albedo, surface temperature and surface emissivity are used to compute reflected shortwave and emitted longwave radiation away from the land surface. The soil heat flux is computed as an empirical fraction of the net radiation using surface temperature, surface albedo and the Normalized Difference Vegetation Index (NDVI). The sensible heat flux is computed first for two specific land surfaces: one pixel with high surface temperature where latent heat flux is negligible (dry pixel) and one cold pixel where sensible heat flux is negligible

(wet pixel). The aerodynamic resistance is the transfer coefficient for heat transport and is calculated from the logarithmic wind profile between the blending height, where the wind speed is constant, and the surface roughness length for momentum transfer. Combining the aerodynamic resistance with the extremes of sensible heat flux allows the assessment of the range of near-surface vertical air temperature differences in the specially selected land surfaces. The near-surface temperature is used to interpret the vertical air temperature differences over the region of interest, assuming linearity between the surface temperature and the vertical thermal gradients in the air layer adjacent to the land–atmosphere interface. The resulting evaporative fraction describes the energy partitioning of the surface energy balance as the latent heat flux or net available energy fraction, with the net available energy being defined as the difference in net radiation and soil heat flux. The instantaneous evaporative fraction has been shown in the literature to be similar to the 24 h evaporative fraction (Brutsaert and Chen 1996), thus allows the estimate of the daily latent heat flux. Detailed description of SEBAL can be found in Bastiaanssen *et al.* (1998a).

SEBAL has been widely used and validated in many regions of the world for estimation of daily ET_a from remotely sensed data (Bastiaanssen *et al.* 1998a, b, Bastiaanssen 1999). C code for SEBAL (Aung and Honda 2008) was developed for this study by building upon a flexible and powerful C image handling library (Honda 2008). The MODIS products that were used to estimate ET_a comprised the daily surface reflectance products (MOD09GA) with a spatial resolution of 500 m from bands 1 to 7, solar zenith angle, quality control band, and land surface temperature and emissivity from the MODIS daily land surface temperature product (MOD11A1) with a spatial resolution of 1 km. Julian day, wind speed, height of vegetation, altitude of the target area, and the locations of hot and cold pixels were also used as inputs in SEBAL.

The quality of ET_a computations in SEBAL depends on the careful selection of two anchor points, that is, hot and cold pixels. The NDVI, albedo and temperature image outputs from the preliminary SEBAL processing were used for the manual selection of hot and cold pixels. A pixel that had low NDVI, high temperature and high albedo was considered as a hot pixel (e.g. bare agricultural field or dry land), while a pixel having high NDVI, low temperature and low albedo was denoted as the cold pixel (e.g. well-irrigated crop surface with the ground fully covered by vegetation).

The comparison between the ET_a derived from the satellite (ET_{asat}) through the SEBAL model and ET_a derived from field measurements using Bowen's ratio method (ET_{aBowen}) is shown in figure 5. At the beginning of the cropping season, during the second week of July, ET_{aBowen} already showed high values, with ET_a measurements ranging from 0.6 to 0.7 cm day^{-1}, while the ET_{asat} was between 0.3 and 0.4 cm day^{-1}. The high values of ET_a during this period may have resulted from the condition of the field, which was flooded. After the harvesting period, during in the third week of November, the ET_{aBowen} values were lower, between 0.3 and 0.5 cm day^{-1}, while ET_{asat} was between 0.1 and 0.2 cm day^{-1}. The ET_{aBowen} was generally higher than the ET_{asat} calculated by SEBAL (with 500 m resolution), which may have been affected by the difference in the scale of measurements used.

The growing period for rice in the study area is during the peak of the rainy season. The study area was nearly covered by clouds for most of the duration of the study, resulting in limited satellite data being available to estimate the ET_a. As for Bowen's

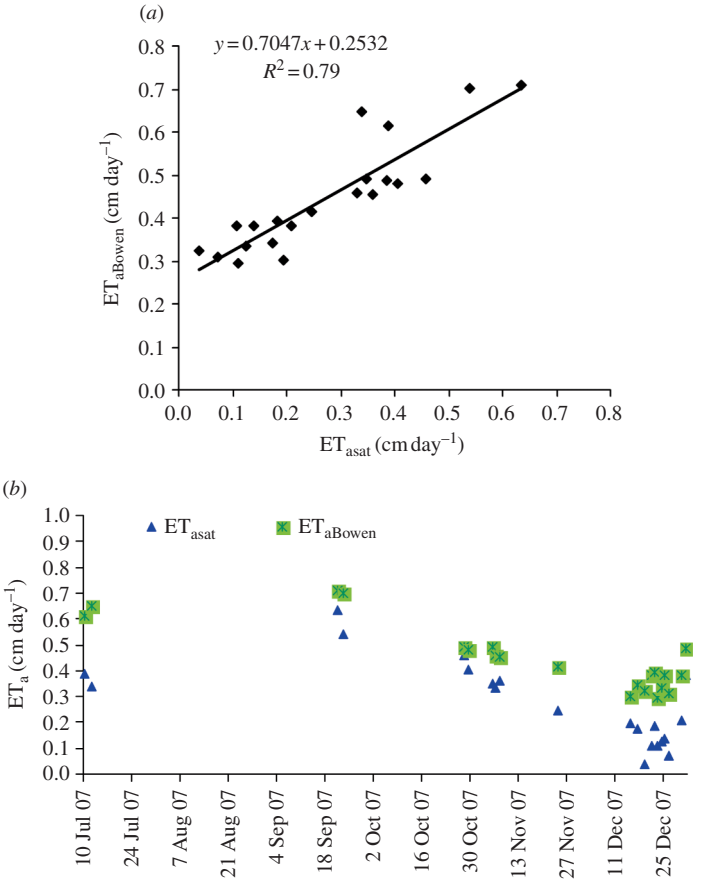

Figure 5. Comparison between ET_{asat} and ET_{aBowen}: (*a*) the scatter plot and (*b*) the time series of ET_a.

method, there were frequent power cuts in the field in the rainy season, thus fewer Bowen data could be obtained during this period as well. Nevertheless, the comparison of available ET_{asat} and ET_{aBowen} showed a good correlation, with $R^2 = 0.79$ and RMSE = 0.19 (see figure 5(*a*)).

4. SWAP input data for assimilation

4.1 *Input data*

The required input for assimilation is weather, crop and soil data. The relevant data, taken from the observation station, were converted into the units supported by the SWAP model. The data comprised humidity (kPa), maximum and minimum temperature (°C), solar radiation (kJ m^{-2}), rainfall (mm) and wind speed (m sec^{-1}). The soil data included soil texture (% sand, silt, clay) and organic matter, which were taken from the field. Crop data comprised the height of rice measured in the field every week, together with planting and harvesting dates recorded from the field.

4.2 *Assimilation of the unknowns*

Soil hydraulic parameters are among the main factors that affect the simulation of soil moisture, but are difficult to measure by field or laboratory methods. Soil hydraulic parameters were therefore set as unknown parameters in the assimilation procedure. In this study, five Mualem–Van Genuchten soil hydraulic parameters were estimated for each soil layer, which were considered simultaneously. For all layers, these unknown parameters were θ_{res}, the residual water content ($cm^3\ cm^{-3}$); θ_{sat}, the saturated water content ($cm^3\ cm^{-3}$); K_{sat}, the saturated hydraulic conductivity ($cm\ day^{-1}$); α, the main drying curve (cm^{-1}) and n, the parameter that accounts for the pore size distribution. The exponent of the hydraulic conductivity that accounts for tortuous soil (λ) was fixed at 0.5 (Mualem 1976).

4.3 *Reference data used for assimilation*

In this study, the LAI and ET_a from the satellite were used as reference data for the assimilation process. Plants affect the soil moisture dynamics owing to water uptake by their roots during the transpiration process. The LAI makes good reference data for assimilation as it directly controls for plant transpiration. ET_a is the loss of water to the atmosphere by the combined processes of soil evaporation and plant transpiration, thus also good reference data for assimilation as it does not only involve soil but plant processes as well. These two variables are therefore expected to indirectly provide information on the properties of the soil surface and its subsurface.

5. Methodology

The SWAP–GA data assimilation technique was used to estimate the soil hydraulic parameters at four different soil depths (3, 12, 28 and 60 cm) in the soil profile, which were further used to simulate soil moisture at each of the soil depths. The schematic diagram of this study is shown in figure 6. The leaf area index (LAI) retrieved from the MODIS satellite, calibrated with LAI from field measurement, and the actual evapotranspiration (ET_a) derived from SEBAL, were used as reference inputs to the SWAP–GA model to find the set of soil hydraulic parameters. The SWAP–GA estimated five Mualem–Van Genuchten soil hydraulic parameters for each soil layer simultaneously, thus 20 unknown soil hydraulic parameters were estimated at once.

The initial values of the unknown parameters were generated randomly by the GA between their given minimum and maximum values (Ines and Mohanty 2008). By selection, crossover and mutation, the GA determined a set of parameters that could explain the pattern of the satellite observations (ET_{asat} and LAI_{sat}), and matched the LAI_{sat} with simulated LAI (LAI_{sim}), or ET_a from satellite (ET_{asat}) with simulated ET_a (ET_{asim}), until the difference between the simulated (ET_a or LAI from SWAP) and observed data (ET_{asat} or LAI_{sat}) was minimized using an objective function. This evaluation was repeated until the GA arrived at a set of parameters that gave the best solution. The objective function used in this study could be categorized into three cases: using LAI alone, using ET_a alone, and combined use of ET_a and LAI together (see equations (9)–(11)).

Case one (input is ET_a):

$$\text{objective function} = \frac{1}{N}\sum_{t=1}^{N} |\, \mathrm{ET}_{\mathrm{aobs},t} - \mathrm{ET}_{\mathrm{asim},t} \,|\,, \tag{9}$$

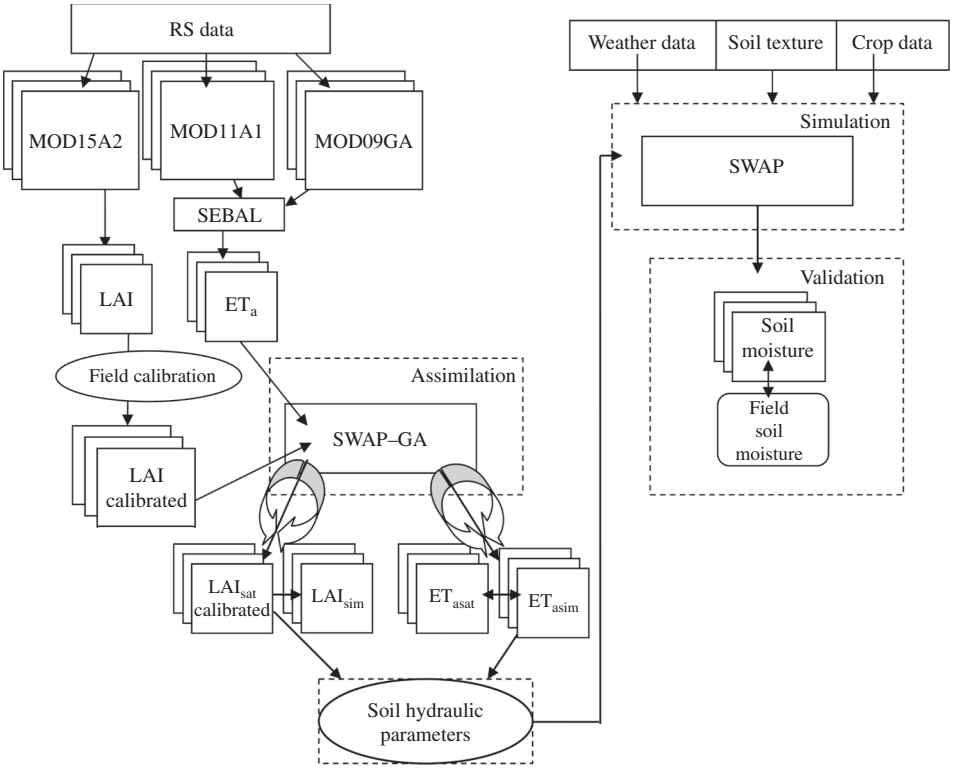

Figure 6. Framework of the study.

case two (input is LAI):

$$\text{objective function} = \frac{1}{N}\sum_{t=1}^{N} |\,\text{LAI}_{\text{obs},t} - \text{LAI}_{\text{sim},t}\,|\,, \tag{10}$$

case three (inputs are ET_a and LAI):

$$\text{objective function} = \frac{1}{N}\sum_{t=1}^{N} |\,0.5\left[\text{LAI}_{\text{obs},t} - \text{LAI}_{\text{sim},t}\right] + 0.5\left[\text{ET}_{\text{aobs},t} - \text{ET}_{\text{asim},t}\right]\,|\,, \tag{11}$$

where N is the time domain, 'sim' signifies simulated data, 'obs' signifies observed data and t = time.

The ET_a and LAI values have been normalized based on the maximum and minimum observed values, allowing them to be added regardless of differences in terms of their units or range of measurements. Values for unknown parameters were found when assimilation was completed or when a matching model was reached for ET_a and LAI values from the satellite (ET_{asat}/LAI_{sat}) and simulated data (ET_{asim}/LAI_{sim}). The resulting parameters could then be applied separately to the SWAP along with the

meteorological data, soil and crop parameters, to simulate the soil moisture in the soil profile.

6. Results

The results of the assimilation of LAI and ET_a into the SWAP model were generally promising. Assimilated LAI from the model showed a very strong correlation with LAI_{sat} (see figure 7), with $R^2 = 0.93$, while simulated ET_a (ET_{asim}) also showed a good match with $ET_{asat,}$ with $R^2 = 0.52$ (see figure 8). Despite the limited availability of satellite data to estimate the ET_{asat} during the rainy season due to cloud contamination, still a certain level of correspondence between the observed (satellite) and simulated was achieved.

Soil hydraulic parameters obtained either from LAI, ET_a, or their combination at four soil depths are shown in table 2. These parameters were applied separately to the SWAP model to simulate soil moisture. The comparison between observed and simulated soil moisture at each soil depth using different reference inputs is shown in figure 9. Results show that soil moisture was highly sensitive to the occurrence of rainfall, especially in the upper layers of the soil, which showed the largest fluctuation. However, during heavy rainfall, the loss of electricity supply in the field meant that no data could be obtained.

Generally, the simulated soil moisture and that observed from the field showed a very good correspondence up to depths of 28 cm. However, soil hydraulic parameters

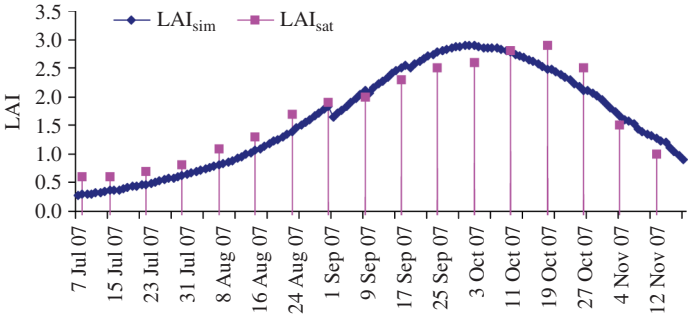

Figure 7. Comparison between LAI_{sat} and LAI_{sim} from the model.

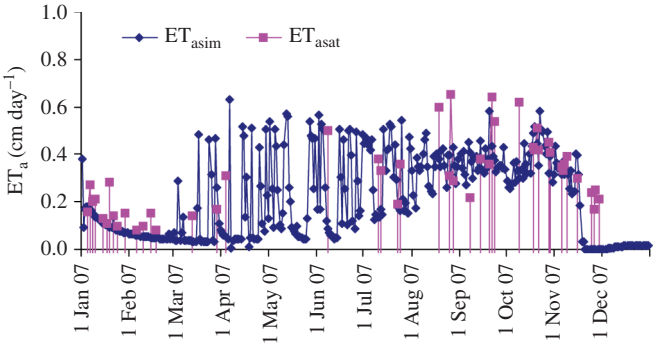

Figure 8. Comparison between ET_{asat} and ET_{asim} from the model.

Table 2. Soil hydraulic parameters estimated from various reference data sources.

Soil layer	Variable		
3 cm	ET_a	LAI	$ET_a + LAI$
θ_{res} (cm^3 cm^{-3})	0.08	0.08	0.06
θ_{sat} (cm^3 cm^{-3})	0.57	0.39	0.47
K_{sat} (cm day^{-1})	45.22	53.75	21.69
α (cm^{-1})	0.03	0.02	0.03
n (unitless)	1.48	1.56	1.54
12 cm	ET_a	LAI	$ET_a + LAI$
θ_{res} (cm^3 cm^{-3})	0.15	0.10	0.08
θ_{sat} (cm^3 cm^{-3})	0.45	0.52	0.51
K_{sat} (cm day^{-1})	38.54	10.84	35.48
α (cm^{-1})	0.01	0.02	0.01
n (unitless)	1.25	1.40	1.23
28 cm	ET_a	LAI	$ET_a + LAI$
θ_{res} (cm^3 cm^{-3})	0.08	0.10	0.15
θ_{sat} (cm^3 cm^{-3})	0.41	0.38	0.39
K_{sat} (cm day^{-1})	42.54	4.26	3.10
α (cm^{-1})	0.01	0.02	0.01
n (unitless)	1.32	1.20	1.34
60 cm	ET_a	LAI	$ET_a + LAI$
θ_{res} (cm^3 cm^{-3})	0.10	0.12	0.10
θ_{sat} (cm^3 cm^{-3})	0.54	0.47	0.52
K_{sat} (cm day^{-1})	20.48	8.00	26.48
α (cm^{-1})	0.01	0.03	0.03
n (unitless)	1.45	1.42	1.53

retrieved from satellite data could not represent the actual value of soil moisture at the lowest depth (60 cm), although the general pattern of soil moisture changes or movements was exhibited.

7. Model evaluation

The soil moisture simulated at each soil depth was evaluated with the soil moisture observations from the field, to measure the accuracy and reliability of the model. Although the simulation period was one year, the model evaluation was done between 23 March and 15 May 2007, except at the 12 cm depth where it was done between 23 March and 20 April 2007. Based on the observations, a sudden change of the soil moisture was found on 21 April 2007, when the value jumped suddenly and then remained static until harvesting time in November. Since it was suspected that the sensor was broken during that period, the data after 20 April 2007 were excluded from the evaluation.

One of the most common statistical methods used to evaluate the accuracy of a model is the root mean square error (RMSE) (cm^3 cm^{-3}), as it evaluates the model performance and measures the total error of the model's results. A lower value of the RMSE indicates better performance of the model (see equation (12)).

In addition, the Pearson correlation coefficient (r) at 0.01 level of significance was used to measure the random error and the correspondence between the simulated and

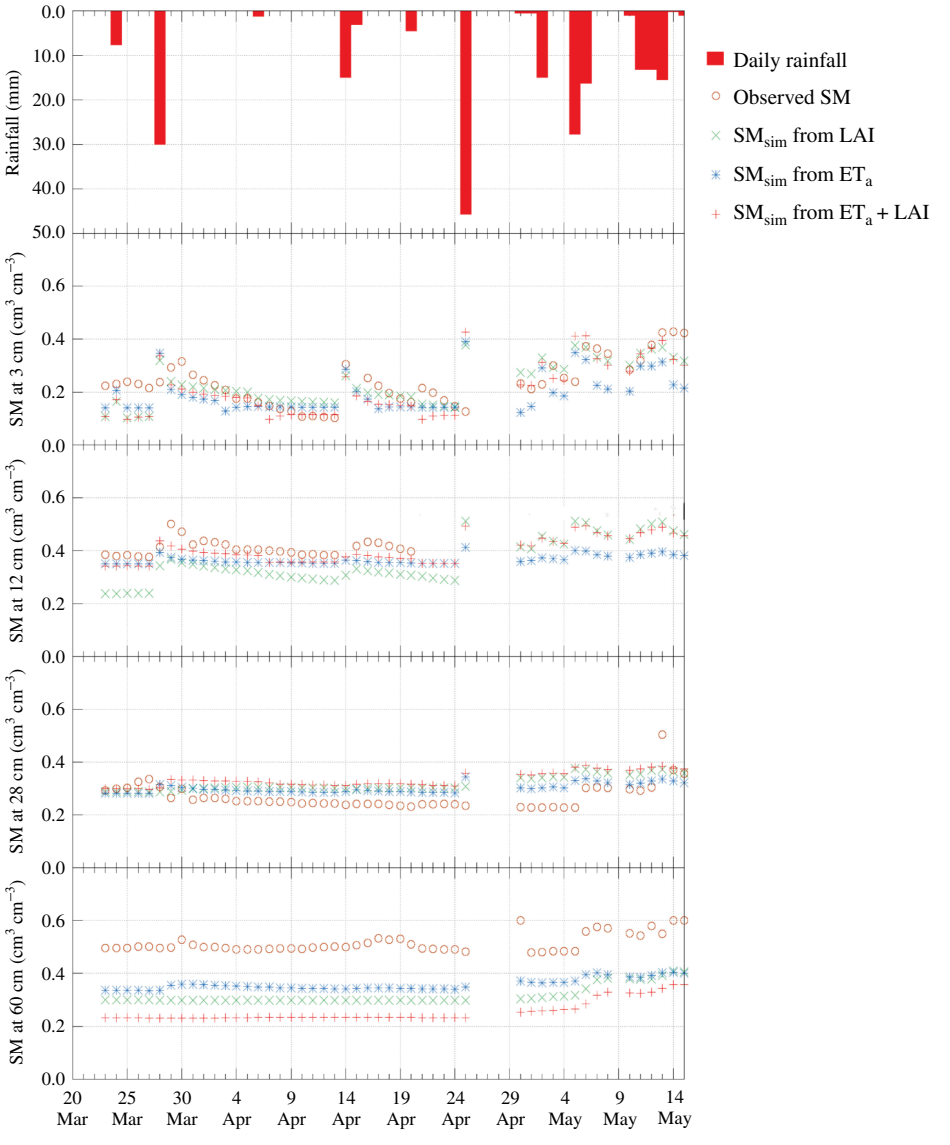

Figure 9. Comparison between soil moisture (SM) from different reference inputs and soil moisture from field observations.

observed soil moisture. A higher value of the correlation coefficient denotes higher efficiency (see equation (13)).

The mean bias error (MBE) ($cm^3\ cm^{-3}$) was also computed as it provides information of the bias of linear correlation, by allowing a comparison of the actual deviation between the simulated and observed values, term by term. A negative MBE means an underestimation while a positive MBE means an overestimation. The ideal value of MBE should be zero, and the unit is the same as the source data, which in this case is $cm^3\ cm^{-3}$ (see equation (14)).

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \left(\theta_{\mathrm{sim},i} - \bar{\theta}_{\mathrm{obs},i}\right)^2}{n}}, \tag{12}$$

$$r = \frac{\sum_{i=1}^{n} \left(\theta_{\mathrm{sim},i} - \bar{\theta}_{\mathrm{sim}}\right)\left(\theta_{\mathrm{obs},i} - \bar{\theta}_{\mathrm{obs}}\right)}{\sqrt{\sum_{i=1}^{n} \left(\theta_{\mathrm{sim},i} - \bar{\theta}_{sim}\right)^2 \sum_{i=1}^{n} \left(\theta_{\mathrm{obs},i} - \bar{\theta}_{\mathrm{obs}}\right)^2}}, \tag{13}$$

$$\mathrm{MBE} = \frac{1}{n} \sum_{i=1}^{n} \left(\theta_{\mathrm{sim},i} - \theta_{\mathrm{obs},i}\right), \tag{14}$$

where θ_{sim}= simulated data, θ_{obs}= observed data, $\bar{\theta}_{\mathrm{sim}}$= average of simulated data, $\bar{\theta}_{\mathrm{obs}}$= average of observed data, n = the number of data and i = the data index.

As shown in table 3, the model performance using ET_a and LAI as conditioning data was good at soil depths of 3 to 28 cm. At 3 cm, LAI showed a good match between the observed and simulated soil moisture values with RMSE = 0.06, r = 0.73 and MBE = –0.01. At 12 cm, the combination of ET_a and LAI showed a good performance with RMSE = 0.04, r = 0.81 and MBE = –0.04. At 28 cm, using ET_a gave RMSE = 0.05, r = 0.43 and MBE = 0.03. At this soil depth, r was quite low, which could have been because the soil was mixed with gravel and stone, which the SWAP model could not handle. On the other hand, although the model performance at 60 cm showed the lowest accuracy with RMSE = 0.15 and MBE = –0.15, the linear correlation between the observed and simulated soil moisture appeared promising, with r = 0.78.

Table 3. Model evaluation results.

Soil layer	Variable		
3 cm	ET_a	LAI	ET_a+ LAI
RMSE ($cm^3\ cm^{-3}$)	0.08	0.06	0.07
r (unitless)	0.64	0.73	0.78
MBE ($cm^3\ cm^{-3}$)	−0.05	−0.01	−0.03
12 cm	ET_a	LAI	ET_a+ LAI
RMSE (unitless)	0.06	0.11	0.04
r (unitless)	0.70	0.87	0.81
MBE ($cm^3\ cm^{-3}$)	−0.05	−0.10	−0.04
28 cm	ET_a	LAI	ET_a+ LAI
RMSE ($cm^3\ cm^{-3}$)	0.05	0.07	0.08
r (unitless)	0.43	0.35	0.34
MBE ($cm^3\ cm^{-3}$)	0.03	0.05	0.06
60 cm	ET_a	LAI	ET_a+ LAI
RMSE ($cm^3\ cm^{-3}$)	0.15	0.20	0.26
r (unitless)	0.75	0.78	0.76
MBE ($cm^3\ cm^{-3}$)	−0.15	−0.20	−0.26

8. Discussion

The MODIS satellite imagery, although free of charge, provides the advantages of high temporal resolution and access to several derived information products, but has limitations in tropical areas. When data were acquired on cloudy and smoggy days, the optical sensor provided poor quality of images. The estimation of ET_a using MODIS through the SEBAL model was also affected by cloud cover during the rainy season, imposing certain limits on both availability and accuracy. Satellite images needed to cover dry land where the latent heat flux is assumed to be zero and wet land where the sensible heat flux is assumed to be zero. Unfortunately, the study site is a rainfed rice area located in the tropical region, and it was impossible to get cloud-free images every day throughout the cropping season.

The selection of the hot and cold pixels is the most critical part of the SEBAL model because such selections could result in over- or underestimation of the ET_a. If too high a temperature of hot pixel is selected, it could result in an overestimation of the ET_a, and if too low a temperature of cold pixel is selected it will result in underestimation of ET_a. A complication also arises from the fact that the hot and cold pixels are liable to change between images acquired throughout the growing season. In the dry season, when harvesting is completed in an area, such area could be chosen as a hot pixel since the land is bare. Whereas during the rainy season, the same area may have crops or plants and may be waterlogged, and thus could be chosen as a cold pixel. Therefore, a good area to select a cold pixel should be irrigated fields, which have crops or plants all year round. The selection of the hot pixels is even trickier in the rainy season, as it is difficult to find a pixel which is satisfactory. When clouds cover bare land, temperatures could be low, with the result that a hot pixel may not be selected. Also, bare land can sometimes be covered with grass, and selection of a pixel using the multiple criteria of temperature, NDVI and albedo might become difficult. In such cases, more priority is given to temperature than to albedo or NDVI, as the selection of the hot pixel determines the upper limit of the energy balance.

In the case of model assimilation, LAI derived from the MODIS LAI product cannot be used directly because it underestimates the results. In this study, it was therefore necessary to calibrate the LAI from ground measurements, to adjust the satellite LAI to reflect the field conditions, before the data could be used in the assimilation process.

This research work has effectively shown that soil moisture data up to the depth of 28 cm can be inferred from satellite information. Furthermore, a new concept of coupling ET_a with LAI from satellite data in the assimilation process, to obtain soil hydraulic parameters for further soil moisture simulation, has been suggested. However, in this research, equal weights were given to LAI and ET_a in the objective function. Giving different weights according to the quality of the satellite data at each crop growth stage may further improve the results.

9. Conclusions

The movement of water in the soil varies in space and time with the complex and dynamic interactions between climate, soil and vegetation. LAI and ET_a derived from satellite data can be used as reference inputs for data assimilation in order to obtain reliable soil hydraulic parameters to simulate soil moisture. The SWAP–GA model could be used to estimate and develop reliable soil hydraulic parameters from satellite information in various input domains. The estimation of soil moisture in the top surface layers (3 cm) was found to be accurate using LAI to obtain the soil hydraulic

parameters. Since the moisture movement in the soil surface is directed to fulfill the transpiration demand at the canopy, LAI can be used as conditioning data. At 12 cm, a combination of LAI and ET_a was found to be a strong predictor of soil hydraulic parameters. At this depth, the LAI still influenced the soil moisture dynamics in the soil profile through plant transpiration via the upper rooting system. When coupled with ET_a, which is the main parameter that accounts for water loss through plant transpiration and evaporation from the soil, a good determination of the soil moisture characteristics at this layer could be observed. At 28 cm, ET_a was also found to be a good predictor of soil hydraulic parameters for simulating soil moisture. The impact of LAI was reduced at greater depths. At 60 cm, the soil hydraulic parameters derived by ET_a compared well with the other variables, however the response was not enough to obtain a match between the observed and simulated soil moisture. This may be due to the fact that the information from satellite LAI and ET_a has limitations inferring the characteristics of soil moisture at deeper layers. Moreover, at 60 cm the soil was saturated, and this may have contributed to some difficulty in simulating the characteristics at such depth. Finally, it is concluded that satellite remote sensing, coupled with SWAP–GA, is a promising tool for determining the effective soil hydraulic parameters for soil moisture simulation, and is amenable to scaling up to regional or large scale applications. This technique could reduce the need for costly and complex field measurements (albeit they are still important in model validation) and also could overcome the problem of data availability in the optical remote sensing for soil moisture simulation. Furthermore, as reliability of soil moisture simulation plays a significant role in agricultural management and yield estimation, early detection of dry soil conditions is not only very useful for agricultural activities but also for reliable flood and drought prediction.

Acknowledgements

Our sincere gratitude goes to the Thailand Research Fund (TRF) for funding this study (Grant Agreement #RDG49O0006). Special thanks also go to all local collaborators, especially the officers of the Agricultural Office and Thai Meteorological Department for their valuable information and suggestions. Part of the work of AVMI was supported by NOAA (Cooperative Grant Agreement #NA05OAR4311004).

References

Aung, K.S. and Honda, K., 2008, *SEBAL C Code* (Bangkok: Asian Institute of Technology).

Bastiaanssen, W.G.M., 1999, SEBAL-based sensible and latent heat fluxes in the irrigated Gediz basin, Turkey. *Journal of Hydrology*, **229**, pp. 87–100.

Bastiaanssen, W.G.M., Meneti, M., Feddes, R.A. and Holtslag, A.A.M., 1998a, A remote sensing surface energy balance algorithm for land (SEBAL). Part 1, formulation. *Journal of Hydrology*, **212–213**, pp. 198–212.

Bastiaanssen, W.G.M., Meneti, M., Feddes, R.A. and Holtslag, A.A.M., 1998b, A remote sensing surface energy balance algorithm for land (SEBAL). Part 2, validation. *Journal of Hydrology*, **212–213**, pp. 213–229.

Brantley, S.L., White, T.S., White, A.F., Sparks, D., Ritcher, D., Pregitzer, K., Derry, L., Chorover, J., Chadwick, O., April, R., Anderson, S. and Amundson, R. 2006, *Frontiers in Exploration of the Critical Zone: Report of a Workshop Sponsored by the National Science Foundation (NSF)*, 24–26 October 2005, Newark, DE. 30 pp.

Brutsaert, W. and Chen, D., 1996, Diurnal variation of surface fluxes during thorough drying (or severe drought) of natural prairie. *Water Resources Research*, **32**, pp. 2013–2019.

Das, N.N. and Mohanty, B.P., 2006, Root zone soil moisture assessment using remote sensing and vadose zone modeling. *Vadose Zone Journal*, **5**, pp. 296–307.

Enthekabi, D., Nakamura, H. and Njoku, E., 1994, Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations. *IEEE Transactions on Geoscience and Remote Sensing*, **32**, pp. 438–448.

FAO, 1991, *The Digitized Soil Map of the World*. World Soil Resources Report 67(1) (Rome: FAO).

Georgakakos, K.P., 1996, Soil moisture theories and observations. *Journal of Hydrology*, **184** (special issue), pp. 131–152.

Goldberg, D.E., 1989, *Genetic Algorithms in Search and Optimization and Machine Learning* (Boston, MA: Addison-Wesley).

Gupta, H.V., Bastidas, L.A., Sorooshian, S., Shuttleworth, W.J. and Yang, Z.L., 1999, Parameter estimation of a land surface scheme using multi-criteria methods. *Journal of Geophysical Research*, **104**, pp. 19491–19504.

Heathman, G.C., Starks, P.J., Ahuja, L.R., Jackson, T.J., 2003, Assimilation of surface soil moisture to estimate profile soil water balance content. *Journal of Hydrology*, **279**, pp. 1–17.

Hogue, T.S., Bastidas, L., Gupta, H., Sorooshian, S., Mitchell, K. and Emmerich, W., 2005, Evaluation and transferability of the Noah land surface model in semi-arid environments. *Journal of Hydrology*, **6**, pp. 68–84.

Honda, K., 2008, *C Image Handling Utility* (Bangkok: Asian Institute of Technology).

Ines, A.V.M. and Droogers, P., 2002, Inverse modeling in estimating soil hydraulic functions: a genetic algorithm approach. *Hydrology and Earth System Sciences*, **6**, pp. 49–65.

Ines, A.V.M. and Mohanty, B.P., 2008, Near-surface soil moisture assimilation for quantifying effective soil hydraulic properties under different hydroclimatic conditions. *Vadose Zone Journal*, **7**, pp. 39–52.

Jackson, T.J., Le Vine, D.M., Schmugge, T.J. and Schiebe, F.R., 1995, Large area mapping of soil moisture using ESTAR passive microwave radiometer in Washita '92. *Remote Sensing of Environment*, **53**, pp. 27–37.

Liu, Y., Gupta, H.V., Sorooshian, S., Bastidas, L.A. and Shuttleworth, W.J., 2005, Constraining land surface and atmospheric parameters of a locally coupled model using observational data. *Journal of Hydrometeorology*, **6**, pp. 156–172.

Mertens, J., Stenger, R., and Barkle, G.F., 2006, Multiobjective inverse modeling for soil parameter estimation and model verification. *Vadose Zone Journal*, **5**, pp. 917–933.

Moran, M.S., Peters-Lidard, C.D., Watts, J.M., Mcelroy, S., 2004, Estimating soil moisture at the watershed scale with satellite-based radar and land surface models. *Canadian Journal of Remote Sensing*, **30**, pp. 805–826.

Mualem, Y., 1976, A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, **12**, pp. 513–522.

Singh-Gill, J.V., 2007, Real time observation and simulation of soil moisture using agro–hydrological model and field server. Masters thesis, Asian Institute of Technology, Bangkok, Thailand.

Swinbank, R., Shutyaev, V., Lahoz, W.A., (Eds.), 2003, *Data Assimilation for the Earth System, NATO Science Series IV, Earth and Environmental Sciences,* vol. 26, p. 338 (Dordrecht: Kluwer Academic Publishers).

Van Dam, J.C., 2000, Field-scale water flow and solute transport. SWAP model concepts, parameter estimation and case studies. Doctoral thesis, Wageningen University, Wageningen, Netherlands.

Van Genuchten, M.T.H., 1980, Closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, **44**, pp. 892–898.

Vrugt, J.A., Gupta, H.V., Bastidas, L.A., Bouten, W. and Soorooshian, S., 2003, Effective and efficient algorithm for multi-objective optimization of hydrological models. *Water Resources Research*, **39**, 1214, 19 pp.