A GIS framework for surface-layer soil moisture estimation combining satellite radar measurements and land surface modeling with soil physical property estimation

M. Tischler a, M. Garcia b, C. Peters-Lidard c, M.S. Moran d,*, S. Miller e, D. Thoma d, S. Kumar b, J. Geiger c

a U.S. Army Corps of Engineers, Engineering Research and Development Center, Topographic Engineering Center, 7701 Telegraph Rd, Alexandria, VA 22312, USA
b University of Maryland — Baltimore County, NASA Goddard Space Flight Center, Code 614.3, Greenbelt, MD 20771, USA
c NASA, NASA Goddard Space Flight Center, Code 614.3, Greenbelt, MD 20771, USA
d USDA Agricultural Research Service, Walnut Gulch Experimental Watershed, Southwest Watershed Research Center, 2000 East Allen Rd., Tucson, AZ 85719, USA
e Department of Renewable Resources, University of Wyoming, Box 3354, 1000 E. University Dr, Laramie, WY 82071, USA

Received 2 November 2005; received in revised form 9 May 2006; accepted 26 May 2006
Available online 1 August 2006

Abstract

A GIS framework, the Army Remote Moisture System (ARMS), has been developed to link the Land Information System (LIS), a high performance land surface modeling and data assimilation system, with remotely sensed measurements of soil moisture to provide a high resolution estimation of soil moisture in the near surface. ARMS uses available soil (soil texture, porosity, \( K_{\text{sat}} \)), land cover (vegetation type, LAI, Fraction of Greenness), and atmospheric data (Albedo) in standardized vector and raster GIS data formats at multiple scales, in addition to climatological forcing data and precipitation. PEST (Parameter EStimatation Tool) was integrated into the process to optimize soil porosity and saturated hydraulic conductivity (\( K_{\text{sat}} \)), using the remotely sensed measurements, in order to provide a more accurate estimate of the soil moisture. The modeling process is controlled by the user through a graphical interface developed as part of the ArcMap component of ESRI ArcGIS. Published by Elsevier Ltd.

Keywords: GIS; ARMS; Model integration; Soil moisture; Land Information System; Parameter estimation

1. Introduction

The Army Remote Moisture System (ARMS) was conceived out of the need for a better understanding and estimation of profile soil moisture in environmentally diverse, potentially hostile, data poor, or physically inaccessible areas. Such knowledge is particularly useful to the Army with respect to traffic-ability modeling and construction engineering, in addition to a variety of other applications. The project was intended to devise a method by which profile soil moisture over watershed-size areas (1000–25 000 km²) could be characterized and predicted down to the tactical scale (<100 m). Furthermore, the system is intended be applicable over any area of interest, regardless of the amount of existing data or resolution.

The solution was to develop a combination system relying heavily on land-surface modeling, but additionally using remotely sensed estimates of surface soil moisture to provide more accurate results by employing parameter estimation. The final requirements were that the system reside within a Geographic Information System, specifically ArcGIS, and be usable by persons with little or no hydrology, soil science,
or remote sensing background. Therefore, the individual components of ARMS were linked using ArcObjects, Visual Basic for Applications (VBA), and external executables and were all controlled within the ArcGIS framework.

The user friendly nature, spatial data analysis tools, and visualization capabilities of a Geographic Information System (GIS) provide a good framework to which models relying on or predicting spatial data can be appended. The most popular GIS is the Arc series of software tools developed by ESRI, including ArcInfo, ArcView, and most recently, ArcGIS. Alternatively, the GRASS GIS (Neteler and Mitasaovva, 2005) exists as an open source choice, particularly when running on UNIX platforms. The popularity of integrating hydrologic models into a GIS in fact prompted the creation of the ArcHydro data model and tools by ESRI for ArcGIS (Morehouse, 2002). Catchment-Sim, an open source standalone GIS package was developed specifically for hydrologic model integration (Ryan and Boyd, 2003). Similar to ARMS, Hydrological Simulation Program — FORTRAN (HSPF) was created using a windows interface with integrated GIS tools to model and visualize point source and non-point source pollution (Shen et al., 2005).

For more than 10 years several different types of watershed models have been integrated with a GIS (Ogden et al., 2001). The complexity of such model integration varies with the project, however. Xu et al. (2001) used a fairly loose integration by relying on the GIS to provide input data for the PDTank model. Miller et al. (2002) used a tighter coupling by scripting GIS functionality to perform hydrologic functions and provide input to the KINEROS and SWAT models in their AGWA ArcView extension. Storck et al. (1998) also used a tightly coupled approach in their analysis of stream flow in the Pacific Northwest using DHSVM (Wigmasta et al., 1994). Frankenberger et al. (1999) present the most extreme case of GIS coupling, where the SMR model is written into the GRASS code rather than being compiled and called externally.

However, in most cases the GIS is used to process and display inputs and outputs to an external model that can be started automatically via scripting, or manually. This allows flexibility in the modeling language, and does not require a model to be rewritten, only simply modified to process the model inputs and outputs. ARMS uses this approach by scripting ArcGIS to provide a Graphical User Interface (GUI), process and convert input data, communicate between external modules, write metadata, and display output data. Similarly, Jeong and Liang (2005) developed a data retrieval, analysis, and visualization system that use GIS tools to accomplish particular tasks, and requires little scientific knowledge of the user.

2. Background

2.1. Remote sensing of soil moisture

Airborne microwave radiometers have been used for over three decades to measure surface soil moisture (Jackson and Schmugge, 1989). Due to the large contrast in the dielectric constant that naturally exists between water and non-saturated soil, this technology has been successful for measuring soil moisture. Water, with a dielectric constant of 80, is vastly different than dry soil, which typically has a dielectric constant <5. In the field, a water-dry soil mixture exists, yielding a dielectric constant somewhere between the two values that can be determined by measuring the soils’ emissivity at microwave frequencies (Schmugge et al., 2002).

Given soil texture and vegetation information, proven and accurate methods exist to convert the measured emissivity to volumetric soil moisture (Jackson and Schmugge, 1991). The actual sampling depth over which the soil moisture can be determined is variable, depending on the soil moisture condition and radar properties. Most studies agree that the penetration depth for microwave sensing is between 0.1 to 0.3 times the wavelength, where the longest wavelengths (L-band) are about 21 cm (Schmugge et al., 2002), which equates to an effective sampling depth of approximately 2–6 cm. For example, several large-scale field experiments have been conducted that used microwave remote sensing to map soil moisture on a watershed scale, including Monsoon ’90 (Kustas and Goodrich, 1994), Southern Great Plains 1997 (Jackson et al., 2002), Southern Great Plains 1999 (Jackson and Hsu, 2001), Soil Moisture Experiment (SMEX) 2002 (Jackson et al., 2003), SMEX 2003 (Jackson et al., 2004), SMEX 2004, and SMEX 2005.

However, active microwave sensors such as Synthetic Aperture Radar (SAR) currently represent the best approach for obtaining spatially distributed surface soil moisture at scales of 10–100 m for watersheds ranging from 1000 to 25 000 km². The magnitude of the SAR backscatter coefficient (σ°) is related to volumetric surface soil moisture (mₛ) through the contrast of the dielectric constants of dry bare soil and water. The perturbing factors affecting the accuracy of mₛ estimation are soil surface roughness and vegetation biomass.

Studies, particularly in the past decade, have generated a multitude of methods, algorithms, and models relating satellite-based images of SAR backscatter to surface soil moisture (Table 1). However, no operational algorithm exists using SAR data acquired by existing spaceborne sensors (Borgeaud and Saich, 1999). A significant limitation of SAR for watershed scale applications is that the sun synchronous satellites can provide only weekly repeat coverage, and even longer for the same orbital path (generally around 35 days). Moran et al. (2004) identified a number of priorities in research, validation and development to improve the accuracy of SAR-derived mₛ estimations, including further studies to interpret the effects of surface roughness and vegetation on the SAR signal, investing in in situ soil moisture measurement networks, launching new sensors, and decreasing the price of SAR imagery.

Table 1

<table>
<thead>
<tr>
<th>Promising approaches using SAR and optical sensors for mₛ estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach</td>
</tr>
<tr>
<td>Semi-empirical SAR algorithm</td>
</tr>
<tr>
<td>SAR for mₛ change detection</td>
</tr>
<tr>
<td>SAR data fusion — passive</td>
</tr>
<tr>
<td>and active microwave</td>
</tr>
<tr>
<td>SAR data fusion — microwave and optical</td>
</tr>
<tr>
<td>SAR plus microwave scattering model</td>
</tr>
</tbody>
</table>
2.2. Land-surface models and modeling soil moisture

Land-surface models (LSM’s) are computer simulations that attempt to describe a host of dynamic environmental processes including evaporation, transpiration, soil water infiltration, diurnal soil and air temperature fluctuations, and radiative flux. The methods used to estimate such processes differ widely between LSM’s, as do their necessary input and intended application. This research used the NOAH soil-vegetation-atmosphere-transfer type LSM to model soil moisture throughout the profile. NOAH is a community model, and the result of collaboration between several organizations for well over a decade.

There are several variables in three categories that are required to run NOAH. The initial condition data reflect the environmental variables, such as soil temperature and soil moisture, at the beginning of the modeling period. Static parameter data describe the unchanging physical characteristics of the soil and vegetation, such as porosity and fraction of greenness. Finally, forcing data are the atmospheric and climatic data, such as precipitation and solar radiation, needed for computation at each timestep during the model run. With initial conditions, parameter data, and forcing data, NOAH can then be successfully run to model a number of hydrologic variables.

Currently, the Land Information System (LIS) (Kumar et al., 2004) designed at the National Aeronautics and Space Administration Goddard Space Flight Center (NASA GSFC), supports the NOAH model and its parameterization to run over large areas in a gridded fashion, rather than a single point. It is the LIS implementation of NOAH that is used within ARMS. The specific algorithms used by NOAH to simulate soil hydrological and thermodynamic processes are described in Sridhar et al. (2002).

The NOAH community model is currently used by several institutions in an operational capacity to model environmental variables at several different scales. For example, the Global Land Data Assimilation System (Rodell et al., 2004) at NASA GSFC uses NOAH as a land surface model and produces output at 0.25 degree globally and the National Center for Environmental Prediction uses NOAH as part of its global mesoscaleEta model at 12 km resolution (Ek et al., 2003). In addition, CAPS, the predecessor to NOAH, has been used in several Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) experiments to evaluate several LSM’s over identical space and time.

2.3. Modeling with parameter estimation

To run a land-surface model, variables describing the vegetation, soil, and other physical characteristics of the area of interest must be properly initialized. Usually, look-up tables based on soil texture and vegetation classification taken from published studies are used to ascertain appropriate parameters. In this method, all areas of similar soil type (texture) have the same soil properties, and all areas of similar vegetation type (landcover classification) have the same vegetative properties. For example, NOAH uses the Cosby et al. (1984) classification system for soil texture to retrieve soil hydraulic parameters such as saturated hydraulic conductivity ($K_{sat}$) and porosity, and the Dorman and Sellers (1989) classification system to classify and parameterize vegetation related parameters such as the rooting distribution and stomatal resistance. Alternatively, several pedotransfer functions (Brakensiek et al., 1984; Cosby et al., 1984; Saxton et al., 1986) and neural network programs (Schapp et al., 2001) exist that will equate standard, easily measured soil parameters, such as particle size and bulk density, to parameters much more difficult to measure, such as $K_{sat}$, porosity, and air entry pressure. However, such methods are frequently empirical in nature and work well for mesoscale and larger resolution applications, but they do not approximate local physical characteristics for smaller resolution studies. Consequently, they prove to be inadequate for use in modeling (Sobieraj et al., 2001).

A solution to empirically derived parameters in localized modeling is to combine the land-surface model with a parameter calibration scheme. Using a parameter calibration or estimation procedure allows mathematically refined parameters representing local phenomenology to supersede the default values for a particular variable. This method has been shown to improve modeling results (Wood et al., 1998). Several documented methods exist for optimizing input parameters, differing mostly in type and number of criteria. A simple calibration scheme could employ the Root Mean Squared Error (RMSE) between the modeled output and measured value for a variable, such as streamflow (Boyle et al., 2000) or microwave brightness temperature (Burke et al., 1997), as the single criterion to estimate input parameters. The input parameter to be optimized is iterated up and down in parameter space between feasible values, continuously minimizing the RMSE. More advanced and complex examples of parameterization schemes include the Multiobjective Generalized Sensitivity Analysis algorithm to estimate better hydraulic and thermal soil properties for the SiSPAT-RS model (Demarty et al., 2004), and the Multiobjective Complex Evolution global optimization method to parameterize 13 variables of the Sacramento Soil Moisture Accounting Model (Yapo, 1998).

ARMS uses the model independent Parameter ESTimation tool (PEST) (Watermark Numerical Computing, 2004) to refine the soil hydraulic parameters which have been shown to strongly influence the movement of water through the soil profile. PEST uses the Gauss-Marquardt-Levenberg algorithm and is flexible in how it interacts with and is controlled by the other components of LIS. Detailed description of the mathematical model, inputs, and limitation of PEST can be found in the user manual accompanying the model.

3. ARMS framework

3.1. GIS Interface

A graphical user interface was created using VBA to allow a user to identify the GIS layers and tables that are to be used in the modeling procedure. A wizard-type series of forms prompts a user to enter input data describing the area of interest (AOI), initial conditions, and forcing data. ARMS can use
either spatially explicit characterizations of the AOI, as in a landcover classification raster layer, or the user can set a parameter to be uniform across the AOI. The same method is used for setting required initial conditions. For example, if a raster layer representing the initial soil moisture of the surface layer exists, a user can choose that layer. Alternatively, a user can specify that the initial soil moisture is constant (e.g., 25% soil moisture) across the domain, as in Fig. 1. Soil classification, landcover classification, saturated hydraulic conductivity, porosity, fraction of greenness, initial soil moisture for the profile, initial soil temperature for the profile, skin temperature, and albedo may all be designated in one of the above methods. At a minimum, the user must have a spatially explicit representation of soil type and landcover type. It is assumed that a user would at least have these data for their AOI. If no other characteristics of the domain are known but soil and landcover, necessary parameters can be calculated from just these two inputs through internal look-up tables derived from pedotransfer functions. At a maximum, all of the above mentioned parameters may be represented by GIS layers.

Control parameters for LIS and NOAH are also set using the GUI. These parameters include start time, end time, timestep, domain boundary, cell size, number of soil layers, and soil layer thicknesses. The domain boundary can be set by choosing the extent of an existing layer, or by manually entering in the lower left coordinates of the AOI, cell size, and number of columns and rows at that cell size within the AOI (Fig. 2).

The forcing data exist in a .dbf format table added to the ArcMap session. The table has a predefined format so that each column represents a different forcing variable, and each row represents a different timestep. The precipitation data are also defined with the GUI, and is the most complex variable to configure for use in ARMS. Currently, two methods are available for entering precipitation. If spatially explicit precipitation files exist for every timestep as raster layers, an index file can be created which ARMS will recognize. The index file is a simple file that for each timestep points to the appropriate precipitation file. While straightforward, the amount of precipitation data can be tremendous, and rarely available. To compensate, a multiquadric-biharmonic rainfall interpolator was developed for use in ARMS that will interpolate breakpoint rain gauge data temporally and spatially across the domain (Hardy, 1990). This interpolator runs as a preprocessing procedure before the LIS model is called.

The parameter estimation parameters are the last set of variables that are set with the ARMS GUI. These include identifying the remotely sensed soil moisture and the time at which the images were collected. The imagery is added as raster layers to the ArcMap session.

Once the domain parameters, static parameters, forcing data, precipitation data, and parameter estimation data are defined, ARMS reformats necessary files, writes several new text files, and modifies existing files which the separate modules will use. For example, nearly all input variables to LIS are read through a card file, a text file that identifies the dozens of variables on which the modeling procedure is dependant. These include model setup parameters, pathnames to parameter files, and domain boundary variables. ARMS modifies the card file depending upon the values entered in the GUI. The forcing data are converted from .dbf format to a specifically

Fig. 1. The GUI allows the user to set initial conditions of the soil layers to an image (raster layer) or constant value.
formatted text file to be read by LIS during runtime. All raster files that are chosen as inputs are also converted to a text format for access by the separate modules. Once all the input data are converted, and the appropriate control files written, parameter estimation and modeling can begin.

3.2. Parameter estimation

The parameter estimation routine in ARMS uses the 1-dimensional version of NOAH and is controlled by the Parameter EStimation Tool (PEST) (Watermark Numerical Computing, 2004). The PEST module jointly optimizes $K_{\text{sat}}$ and porosity for the AOI. Running PEST over the entire domain is not feasible for several reasons, though mainly computational time and power. Instead, stratified samples of cells from across the AOI are chosen for the procedure. The cells are chosen by determining the total number of unique combinations of soil and landcover types in the domain. A file is created identifying the combination to which each cell in the domain belongs. PEST is run for a specified number of cells that represent each unique combination. Currently, PEST is hard coded to run for up to 8 cells in each unique combination of landcover and soil types. If the total number of cells for a particular combination is less than 8, then PEST will run on as many as are present.

For each cell that is to be used in PEST, the initial conditions, static parameters, forcing data, and remotely sensed soil moisture values are extracted from the appropriate parent files, and written to text files that PEST will read. Then, 1-D NOAH is run by PEST until optimized parameters are converged upon. The detailed process by which this happens is described in the PEST User Manual (Watermark Numerical Computing, 2004). Once optimized parameters exist for each chosen cell, those optimized values are averaged within each unique combination. Optimized $K_{\text{sat}}$ and porosity raster layers are created from the average values for each unique combination across the domain by using the reference file that tracks to which combination each cell belongs. These optimized parameters can supersede the original $K_{\text{sat}}$ and porosity estimations for use in LIS.

3.3. LIS and NOAH

The modeling kernel of ARMS is the NOAH LSM accessible through LIS. LIS resides as an external Windows executable file compiled using Compaq Visual Fortran v6.6. The model is a modified version of the standard LIS software package designed to work with the files which ARMS provides. Any spatially explicit LIS inputs the user may have are added to the ArcMap session as raster layers. These can be either data available for the area, or the result of parameter estimation. ARMS reformats the raster layers into text files that LIS can then read and process. The modification of a card file communicates to LIS how the model will be run, and with which parameters. LIS will process all the inputs given to it by ARMS, define a domain, and then rely on NOAH to perform the actual modeling at each timestep for each cell in the domain. The computations performed by NOAH are outlined and explained in Sridhar et al. (2002). After NOAH has completed calculations across the domain at a particular timestep, LIS will write output files summarizing the calculations. In addition, a text file of the output profile soil moisture is written that can easily be imported to ArcGIS for display.

4. Field study

The ARMS system has been tested using data from the Monsoon '90 field experiment (Kustas and Goodrich, 1994) at the Walnut Gulch Experimental Watershed in southern Arizona. During Monsoon '90, daily gravimetric soil moisture data were collected at eight micrometeorological-energy flux...
(Metflux) sites, in addition to standard meteorological variables and surface fluxes. In addition, an airborne L-band Push Broom Microwave Radiometer (PBMR) mounted on a NASA C-130 aircraft was flown at an altitude of 600 m above the ground to yield soil moisture products derived from measured microwave brightness temperature (Tb) (Schmugge et al., 1994). Tb data were collected over an approximately $8 \times 20$ km area with a 40 m horizontal resolution for six days: 212 (Jul. 31), 214 (Aug. 2), 216 (Aug. 4), 217 (Aug. 5), 220 (Aug. 8), and 221 (Aug. 9).

The Metflux data in combination with a network of 88 rain gauges provided the forcing data necessary to use ARMS. To test the effectiveness of the system, two model runs were performed. First, the model was run with the best possible soil parameters (Run 1). This included soil texture, porosity, and $K_{sat}$ values derived from the Soil Survey Geographic Database (SSURGO). This dataset is available only for the United States, and provides the most detailed estimates of soil property information. Next, the model was run for the Monsoon ’90 period using soil textures derived from the Food and Agricultural Organization (FAO) Digital Soil Map of the World (Run 2). This dataset provides very coarse, but globally available data coverage. Porosity and $K_{sat}$ values initially used in this model run were derived from the default tables within LIS, which are based on Cosby et al. (1984). For the first model run (Run 1) with SSURGO data, no parameter estimation was used. PEST was used in the second run (Run 2) to optimize porosity and $K_{sat}$, based on the comparison between the model output and PBMR derived soil moisture values collected in the modeling period. Fig. 3 compares the watershed averages of $K_{sat}$ and porosity for the different soil data sources, and pest.

The ARMS system employing PEST using FAO soil data minimized both the bias (Fig. 4) and Root Mean Squared (RMS) error (Fig. 5) across the watershed over the modeling period when compared with a traditional run using the most detailed soil data available.

5. Summary

ARMS provides the framework to combine remotely sensed data, parameter estimation, and land surface modeling for the estimation of soil moisture at the watershed scale. The integration of these models takes place within a GIS, made possible by using the VBA integrated development environment (IDE) contained within ArcMap on a Windows platform. As a result, a graphical user interface exists for a user to easily parameterize the land surface model and to define the modeling domain. Scripts perform data format conversions from propriety GIS file formats to text files that are read and processed by the separate modules of ARMS. Control files, such as the LIS card file, are created within ARMS, allowing a precompiled executable to be used, rather than compiling source code after each modification, as in traditional hydrologic modeling using C or Fortran source code. A parameter estimation module exists within ARMS that optimizes soil hydraulic parameters to which land surface models have been shown to be

![Fig. 3. The $K_{sat}$ and porosity values averaged over the watershed for different data types. The PEST optimization was performed using FAO soil type as input.](image1)

![Fig. 4. The bias over the watershed for the Monsoon ’90 modeling period. The bias was reduced significantly using the parameter optimization routines within ARMS.](image2)

![Fig. 5. The RMS error over the watershed for the Monsoon ’90 modeling period. The RMS error was reduced significantly using the parameter optimization routines within ARMS. The intrinsic error in the PBMR data used to validate the results is 4.5% ± 1.9% (Peters-Lidard et al., 2003).](image3)
sensitive. Modifications to the robust LIS system were made to customize the internal models for use with ARMS.

ARMS was designed and developed with the military user in mind, for use in trafficability assessment, construction engineering, electromagnetic signal propagation and adsorption, and countermine efforts. ARMS allows a user with little formal background in hydrological sciences, remote sensing, computer science, or soil science to determine spatial surface soil moisture using a scientifically sound procedure with robust modeling tools for use in a variety of applications.

Modeling soil moisture with ARMS over Walnut Gulch using the Monsoon '90 dataset yielded promising results, when compared with traditional modeling. A more accurate assessment of the state of the soil moisture over the watershed was made by reducing errors using the tools within the modeling system. In the future, modeling efforts will be conducted in different regions to test and improve ARMS in a variety of terrain and landcover regimes. Ultimately, ARMS will be able to run worldwide with fairly limited inputs and provide a high resolution soil moisture product to feed a number of other tactical decision aids. Currently, the system is not yet operational. In the future, information regarding ARMS and the release of software can be directed to the United States Army Corps of Engineers Topographic Engineering Center.


