Journal of Hydrology 539 (2016) 392–405

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Soil moisture deficit estimation using satellite multi-angle brightness temperature

Lu Zhuo [a,*], Dawei Han [a], Qiang Dai [b]

[a] WEMRC, Department of Civil Engineering, University of Bristol, Bristol, UK
[b] Key Laboratory of Virtual Geographic Environment of Ministry of Education, School of Geography Science, Nanjing Normal University, Nanjing, China

ARTICLE INFO

Article history:
Received 6 February 2016
Received in revised form 5 May 2016
Accepted 23 May 2016
Available online 28 May 2016
This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Magdeline Laba, Associate Editor

Keywords:
SMOS brightness temperature
Soil moisture
Local linear regression (LLR)
Artificial neural networks (ANNs)
Soil moisture deficit (SMD)
Hydrological modelling

SUMMARY

Accurate soil moisture information is critically important for hydrological modelling. Although remote sensing soil moisture measurement has become an important data source, it cannot be used directly in hydrological modelling. A novel study based on nonlinear techniques (a local linear regression (LLR) and two feedforward artificial neural networks (ANNs)) is carried out to estimate soil moisture deficit (SMD), using the Soil Moisture and Ocean Salinity (SMOS) multi-angle brightness temperatures (T_bs) with both horizontal (H) and vertical (V) polarisations. The gamma test is used for the first time to determine the optimum number of T_bs required to construct a reliable smooth model for SMD estimation, and the relationship between model input and output is achieved through error variance estimation. The simulated SMD time series in the study area is from the Xinanjiang hydrological model. The results have shown that LLR model is better at capturing the interrelations between SMD and T_bs than ANNs, with outstanding statistical performances obtained during both training (*NSE* = 0.88, *r* = 0.94, *RMSE* = 0.008 m) and testing phases (*NSE* = 0.85, *r* = 0.93, *RMSE* = 0.009 m). Nevertheless, both ANN training algorithms (radial BFGS and conjugate gradient) have performed well in estimating the SMD data and showed excellent performances compared with those derived directly from the SMOS soil moisture products. This study has also demonstrated the informative capability of the gamma test in the input data selection for model development. These results provide interesting perspectives for data-assimilation in flood-forecasting.

1. Introduction

Although soil moisture comprises only 0.01% of the total amount of water on our planet, its existence plays an important role in influencing the water and energy exchanges at the land surface/atmosphere interface. There is abundant evidence that hydrological processes are significantly conditioned by a river catchment's antecedent wetness state (Massari et al., 2014; Tramblay et al., 2012). In particular the surface soil wetness is an important variable in hydrological modelling because it controls key processes such as runoff and evapotranspiration, and is a vital parameter for flood modelling (Draper et al., 2011; Han et al., 2012)

The Earth thermal emission at microwave bands depends essentially on the soil temperature and the soil water content (Al-Yaari et al., 2014; Rodríguez-Fernández et al., 2015). Recent research activities indicate rising interest in the operational monitoring of the global soil moisture remote sensing. In particular, the data acquired by lower microwave frequencies (e.g., L-band at 1.20–1.41 GHz), both active and passive, have been utilised to provide detailed surface soil moisture fluctuations in recent years (Calvet et al., 2011). The launch of the Soil Moisture and Ocean Salinity (SMOS; (Kerr et al., 2001)) mission in November 2009 and the Soil Moisture Active/Passive mission (SMAP; (Entekhabi et al., 2010)) in January 2015 clearly demonstrates the significance and determination of an advanced global surface soil moisture monitoring system. SMOS is the first mission dedicated to monitoring direct surface soil moisture and sea surface salinity on a global scale (Kerr et al., 2010), and has a longer period of data record since its launch in 2009. Therefore, SMOS is chosen in this study.

The SMOS soil moisture operational algorithm utilises a direct or forward model and an optimal estimation method: a radiative transfer model (e.g., LMEB model is used in the SMOS algorithm (Wigneron et al., 2007)) is applied to estimate L-band brightness temperatures (hereafter T_bs) for a set of physical parameters, soil composition, and moisture content and vegetation opacity (Rodríguez-Fernández et al., 2015). In order to estimate soil moisture, the simulated T_bs are compared with those measured by SMOS

* Corresponding author.
E-mail address: lu.zhuo@bristol.ac.uk (L. Zhuo).

http://dx.doi.org/10.1016/j.jhydrol.2016.05.052
0022-1694/

using an iterative process to minimise the difference between them. This approach then requires in-situ observation data for soil moisture evaluation (Al-Yaari et al., 2014; Al Bitar et al., 2012). However most areas do not have in-situ sensors because they are expensive to set up and impractical to maintain; and they are too sparse for catchment-scale studies (Al-Shrafany et al., 2013; Srivastava et al., 2013b, 2013c; Walker et al., 2004; Wang and Qu, 2009). Since the presence of vegetation can reduce the brightness temperature sensitivity to soil moisture, in the aforementioned method decoupling the effects of soil and vegetation on brightness temperature can pose a major challenge for useful application under such circumstances.

In order to retrieve accurate soil wetness information that can be directly used in a hydrological model and avoid aforementioned shortcomings, a data-driven model is desirable, which could effectively link the inputs to the desired output and is not computationally intensive. This can be achieved by building an inverse model that provides soil moisture information (i.e., soil moisture deficit (SMD) in this study, which is a key soil moisture variable in hydrological models (Zhuo et al., 2015a)) directly from a given set of satellite measured T_bs. Among the data-driven models, nonlinear regression models such as Local Linear Regression (LLR) and Artificial Neural Networks (ANNs) are widely recognised and used as efficient inverse models. Therefore both LLR and ANNs are used in this study.

The foremost objective of this study is therefore to build an inverse model for the first time that can simulate the relevant hydrological SMD data directly from the SMOS brightness temperatures using various nonlinear modelling techniques. In this study, the SMD is estimated instead of the normal soil moisture because in hydrological modelling the excess runoff is closely linked with SMD, but not directly with the normal soil moisture (i.e., the volumetric soil moisture). The SMD refers to the amount of water needed to bring the soil moisture back to field capacity. Since SMD is directly relevant to hydrology, it is the main purpose of this study. SMOS is the first radiometer in space with full-polarisation and multangular capabilities (Rodríguez-Fernández et al., 2015). Hence, a dedicated retrieval scheme has to be studied. An LLR model and two ANN models are trained and tested for their valuation in SMD retrieval. The modelled SMD values using different techniques are then compared against the Xinanjiang simulated SMD as the target. Furthermore, a well-proven and widely applied computing algorithm called the gamma test (GT) is employed to find the optimal combination of data inputs for SMD calculation. Noori et al. (2011) and Remesan et al. (2008) applied the GT data selection method in hydrological studies, for daily solar radiation estimation and monthly streamflow prediction, and both reported positive performances. In contrast to the conventional allocation method of the training and the testing data, the *M*-test is adopted to find the optimal training dataset which has sufficient information for training any regression models. This will avoid wasting time and effort in allocating excessive training data or using inadequate training data. Therefore, no predefined training and testing data will be specified at the early stage of the study. Finally, the SMD estimates from the aforementioned nonlinear methods are compared with those directly derived from the SMOS soil moisture products (i.e., two different SMOS products are used: one is from the SMOS Barcelona Expert Centre (SMOS-BEC) (SMOS-BEC, 2015) and the other is from the Centre Aval de Traitement des Données SMOS (SMOS-CATDS) (Jacquette et al., 2010)).

2. Study area and data

Pontiac is a medium-sized catchment (1500 km^2) in the Vermilion River, located in the central Illinois area of the U.S. The catchment's topography is flat and mainly used for cultivation purpose as illustrated in Fig. 1b (Bartholomé and Belward, 2005; Hansen et al., 1998). Based on the Global Soil Regions map (USDA, 2005), its soil is predominately Mollisols. The catchment is dominated mainly by hot summer continental climate (Peel et al., 2007). The layout of the Pontiac catchment is shown in Fig. 1a along with the location of its flow gauge, river network, and the North American Land Data Assimilation Systems Phase 2 (NLDAS-2) grid points (i.e., the marked grid points are located at the central of each 0.125° × 0.125° NLDAS-2 grids). The spatial variations of an extracted SMOS T_b dataset (H polarisation) at an incidence angle of 32.5° is shown in Fig. 1c (it has been transformed into NLDAS-2 grid spacing at 0.125° for easier analysis). It can be seen from this retrieved image, the central catchment area has lower T_b values (i.e., relatively wetter soil), while the western upper and lower parts show slightly higher T_b values (i.e., relatively drier soil). This could partially be explained by the location of the river network as indicated in Fig. 1a: the majority of the water concentrates at the central area (i.e., the mainstream) and then flows to the catchment outlet (so the soil can be replenished with water more easily); whereas the soil around the small substream areas has less water availability and tends to be drier. It should be noted that soil moisture does not solely correlate with the variation of brightness temperature but also with other factors such as vegetation cover, local soil properties, and surface roughness.

The Xinanjiang (XAJ) model's hydrological forcing is obtained from the NLDAS-2 (Mitchell et al., 2004). The datasets comprise precipitation (Daly et al., 1994) and potential evapotranspiration at the 0.125° spatial resolution and daily temporal resolution (converted from hourly resolution). Both datasets have been transformed into the catchment-scale using the weighted average method to operate the lumped XAJ model. Readers are referred to Xia et al. (2012) and Zhuo et al. (2015c) for a full description of the NLDAS-2 data products. The observed daily flow data for this study is provided by the U.S. Geological Survey. The observations cover a total period of 24-months from January 2010 to December 2011. The reason for using these two-year data is due to the discontinuity of flow observations in the selected catchment.

2.1. SMOS data

SMOS retrieves the thermal emission from the Earth at the frequency of 1.4 GHz in both polarisations and for incidence angles from 0° to 60°. It is dedicated to providing global surface soil moisture information at an accuracy of 0.04 m^3/m^3 (Kerr et al., 2012). SMOS has a Y-shaped antenna structure, which comprises 69 small antennas (a diameter of 16.5 cm) and 4.5-m long arms to perform interferometry and synthesise an aperture of ~7.5 m (McMullan et al., 2008; Rodríguez-Fernández et al., 2015). The projection of the synthesised beam on the Earth surface is generally presented as an ellipse whose axis ratio and orientation depend on the observed point position (Rodríguez-Fernández et al., 2015). The retrieved observations have a spatial resolution of 35–50 km (Kerr et al., 2010). SMOS follows a sun-synchronous polar orbit with a global coverage at the equator crossing the times of 6:00 A.M. at the local solar time (LST) (ascending) and 6:00 P.M. (LST, descending).

In order to estimate SMD from SMOS T_bs, the Level-3 brightness temperature data from the CATDS is used (Jacquette et al., 2010). This daily global brightness temperature data contains SMOS T_bs in the reference frame of 0.25° EASE grid (Brodzik and Knowles, 2002) on the Earth surface. It provides T_bs measurements acquired at all incidence angles in a given day (averaged in 5° – width angle bins) which have been transformed into the ground polarisation reference frame (i.e., H, and V polarisations). Hence, the quantity