Analyzing the Inundation Pattern of the Poyang Lake Floodplain by Passive Microwave Data

HAOLU SHANG
State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China, and Delft University of Technology, Delft, Netherlands

LI JIA
State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China

MASSIMO MENENTI
Delft University of Technology, Delft, Netherlands

(Manuscript received 24 January 2014, in final form 16 September 2014)

ABSTRACT

The soil wetness condition is a useful indicator of inundation hazard in floodplains, such as the Poyang Lake floodplain. Special Sensor Microwave Imager (SSM/I) passive microwave data were used to monitor water-saturated soil and open water areas of the Poyang Lake floodplain from 2001 to 2008, capturing the inundation patterns of this area in space and time. The polarization difference brightness temperature (PDBT) at 37 GHz is sensitive to the water extension even under dense vegetation. The zero-order radiative transfer model was simplified to retrieve the vertical–horizontal (V–H)-polarized effective emissivity difference from the PDBT at 37 GHz. Vegetation fractional area and vegetation transmission function were derived from NDVI to represent the vegetation attenuation. This effective emissivity difference has a quasi-linear relationship with the fractional area of water-saturated soil and standing water, no matter the frequency. Using the multifrequency-polarization surface emission ($Q_p$) model and the Dobson model of the soil–water mixture, the two segments of this relationship were combined into a quasi-linear model. Comparing the retrieved water-saturated soil and standing water area of Poyang Lake with the lake area obtained from the MODIS and synthetic aperture radar (SAR) image at higher spatial resolution, the calculations show a good fit with the MODIS and SAR data, with $R^2 = 0.7664$ and relative RMSE $= 17.74\%$. The cross-correlation analysis shows that the Poyang Lake extension fluctuates with a 5-day time lag with the upstream land area of water-saturated soil and standing water. Since the closure of the Three Gorges Dam, this relationship is more evident.

1. Introduction

The soil wetness condition is a very important factor to describe the regional soil water-storage capacity, which determines the relative magnitude of surface runoff from precipitation. Thus, for hydrological applications, especially flood simulation and early warning, the soil wetness condition is a very critical variable. However, it is difficult to measure the regional soil wetness condition by field measurements. A passive microwave radiometer at C or L band can provide soil moisture over large regions, but only from the topsoil down to a certain depth, which is determined by the wavelength. Because of the shallow soil layer sensed at 37 GHz (detected soil depth is around 0.8 mm), when this layer is water saturated, additional water will become vertical and horizontal water flow, or standing water after the deep soil layers are saturated. We could then consider the fractional area of water-saturated topsoil and standing water as an indicator of the residual water-storage capacity of the whole region. The
The degree of soil water saturation is the bridge connecting the normalized difference vegetation index (NDVI). We will develop this concept later.

Stable, attenuation by vegetation can be obtained from time series. This suggests that when soil moisture is rather dry, the quasi-linear relationship between the degree of soil water saturation and the PEED was first observed by De Ridder (2000) at 19 GHz by simulating the smooth surface soil emissivity with the Wang–Schmugge model. But, the dielectric constant of a soil–water mixture is frequency dependent. Simulated soil emissivity will also be different between models. Thus, we need to validate the quasi-linear relationship at 37 GHz and the consistency of this relationship across models.

Because of the revisit frequency and swath width of microwave radiometers on polar-orbiting satellites, time series of passive microwave data have gaps of 3–5 days. To make a consistent daily time series, we need some statistical methods, such as a moving window filter, to fill these observation gaps. In addition, the effects of raindrops on microwave observations at 37 GHz are not negligible. Geolocation and resampling introduce errors in the conversion from swath brightness temperature (BT) to the gridded data. Erroneous observations therefore need to be identified and removed. In this case, we applied the Harmonic Analysis of Time Series (HANTS) algorithm (Menenti et al. 1993; Roerink et al. 2000; Verhoef 1996) to fill gaps and remove noisy samples. The fast Fourier transform (FFT) and HANTS have been developed and applied to time series of satellite observations, for example, NDVI and land surface temperature, to study vegetation phenology and land surface climate (Alfieri et al. 2013; Jia et al. 2011; Julien et al. 2006; Menenti et al. 1993, 2010; Moody and Johnson 2001; Roerink et al. 2000, 2003; Verhoef 1996). Different from the FFT by using all observations regardless of quality, the HANTS identifies and removes outliers in data samples. We aimed at evaluating its applicability to satellite microwave radiometric data.

The objective of this paper is to study the inundation pattern in the up- and downstream area of the Poyang Lake floodplain in relation with the WSS and standing water area retrieved from passive microwave observations. Poyang Lake is the largest lake in the Yangtze River basin. As the major natural water reservoir of the Yangtze River basin, the inundation pattern of this floodplain has a significant influence on both local and downstream water resource management.

There are five sections in this paper: introduction, study area and dataset, methods, results, and conclusions. In the methods, the algorithm for retrieving the WSS and standing water area from the PDBT at 37 GHz
is developed. Then, we demonstrate the details of the algorithm from data procession to parameter derivation step by step. In the results, the preprocessed data are compared with the original data to show the improvement after gap filling and noise removal. After solving the parameter-derivation problems, the retrieved fractional area of the WSS and standing water of Poyang Lake area is compared with the lake area obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) and advanced synthetic aperture radar (ASAR) data.

2. Study area and dataset

The study area is the floodplain including Poyang Lake [see the squares in Fig. 1, which are the Equal-Area Scalable Earth Grid (EASE-Grid) with a spatial resolution of 25 km (Brodzik and Knowles 2002)]. It is the largest lake in the Yangtze River basin, located between 28.048° and 29.384°N and 115.444° and 117.007°E. There is a clear dry season from October to March, and the flooding season is from April to September. The low residual water-storage capacity of this floodplain was one of the major causes of the extreme flooding over the Yangtze River basin in 1998. Thus, it is chosen as our study area.

Currently, there are two sensors providing dual polarization measurements around 37 GHz: the Special Sensor Microwave Imager (SSM/I) and the Advanced Microwave Scanning Radiometer for EOS (AMSR-E). In this case, the SSM/I on board the Defense Meteorological Satellite Program (DMSP) F13 satellite is used. It is a seven-channel, four-frequency radiometric system that has provided global observations at a constant angle of incidence of 51.2° since July 1987 (Hollinger et al. 1990). No data were acquired from 2 December 1987 to 12 January 1988 because of overheating of the sensor. The period of our study is from 2001 to 2008, to ensure that the observations by MODIS overlap with the enclosure of the Three Gorges Dam (in November 2002). The footprint of the 37-GHz observations is about 28 km × 37 km, with geolocation uncertainty of about 8 km. We used the data acquired at 0558:00 local time (LT). The AMSR-E on board the Aqua satellite measures radiation since May 2001 at six frequencies in the range of 6.9–89 GHz, all at dual polarization, with a constant angle of incidence of 55°. The 36.5-GHz data have a footprint of 14 km × 8 km and are acquired at 0130:00 LT. The SSM/I brightness temperature has been intercalibrated with that of AMSR-E, and both are registered into the EASE-Grid system. Thus, for the same pixel area, the time series of PDBT at similar frequencies (37 GHz for SSM/I and 36.5 GHz for AMSR-E) should have similar fluctuations. The PDBT time series of these two sensors are compared to validate the consistency, observation bias, and possible noise. The daily SSM/I and AMSR-E data gridded into the EASE-Grid (Brodzik and Knowles 2002) were
downloaded from the National Snow and Ice Data Center (NSIDC) (Armstrong et al. 1998). The Poyang Lake area covers 10 pixels of EASE-Grid data, as indicated in Fig. 1 by the white-numbered squares. The upstream area of the Poyang Lake covers the four major tributary river systems coming from the west, southwest, southeast, and east, and indicated in Fig. 1 by the gray-shaded squares with black numbers.

This study uses two MODIS data products: the MOD09A1 8-day composite surface reflectance data of MODIS bands 1–7, which are used in the calculation of vegetation fractional area and vegetation optical thickness; and the MOD13A2 16-day composite NDVI data, which are used to derive the relationship between NDVI and vegetation optical thickness. The spatial resolution of the reflectance dataset is 500 m, while the NDVI dataset is 1 km. Because of the nonlinear dependence of NDVI on red and near-infrared reflectance, these two bands will be averaged first over each EASE-Grid and then used to calculate its NDVI. MOD13A2 is used in the parameter derivation, mainly because of its high quality data. MOD09A1 is chosen from the best quality atmosphere condition during each 8-day measurement. MOD13A2 corrects the viewing angle difference based on MOD09A1: thus, large off-nadir and forward-scattering view angles and large solar zenith angles are avoided in MOD13A2.

3. Methods

The remote sensing data, that is, the MOD09A1 and SSM/I 37-GHz brightness temperature data, are used to retrieve the water-saturated soil and open water area. These data are first preprocessed to remove noise and daily gaps. The zero-order radiative transfer model is simplified, in order to use these processed remote sensing data to retrieve the surface PEED. A linear model is then developed to calculate the WSS and open water area from the PEED. The overview of the theory of our method is explained in section 3a, and the detailed explanation of each step is given in sections 3b–d.

a. Theory and flowchart of retrieving fractional area of WSS and standing water from passive microwave measurements

1) SIMPLIFIED RADIATIVE TRANSFER MODEL

Satellite microwave radiometers measure the radiance emitted by the land surface, such as soil, standing water, and snow. This radiance is determined by the surface temperature, surface emissivity, and surface roughness, and it is attenuated by the vegetation canopy and atmosphere. The zero-order radiative transfer model is often used to describe the energy transfer from the land surface to the radiometer. We simplified this model to retrieve the PEED from the PDBT $\Delta T$ at 37 GHz as follows:

$$\Delta T = T_s \Delta \varepsilon [(1 - f_{\text{veg}}) + f_{\text{veg}} \exp(-2\tau_0/\mu)],$$

where $T_s$ is the surface brightness temperature; $\Delta \varepsilon$ is PEED; $f_{\text{veg}}$ is the vegetation fractional area; $\mu = \cos(\theta)$, where $\theta$ is zenith view angle; and $\tau_0$ is the optical thickness of vegetation.

The single scatter albedo of the vegetation canopy at two polarizations is assumed to be the same. The apparent brightness temperature of the sky is negligible compared with the land surface temperature. Under the above two assumptions, we get Eq. (1) from the zero-order radiative transfer model.

2) LINEAR MODEL TO RETRIEVE FRACTIONAL AREA OF WSS AND STANDING WATER

Since the PEED is obtained from Eq. (1), the fractional area of WSS and standing water $f_{\text{ws}}$ can be derived from the following linear equation:

$$f_{\text{ws}} = (\Delta \varepsilon - \Delta \varepsilon_{\text{dry}}) / (\Delta \varepsilon_{\text{sat}} - \Delta \varepsilon_{\text{dry}}),$$

where $\Delta \varepsilon_{\text{dry}}$ and $\Delta \varepsilon_{\text{sat}}$ are the PEED of completely dry soil and water-saturated soil, respectively. Their derivation method will be discussed in section 3d. The variable $\Delta \varepsilon_{\text{sat}}$ has the same value with that of water; thus, the fractional area in Eq. (2) represents the sum of the WSS and standing water.

Equation (2) is established based on the definition transform from the water saturation condition to the fractional area of WSS and standing water as follows:

$$f_{\text{ws}} = \sum_{i=1}^{n} (f_i W_i / W_{\text{sat},i}) = W_{r,s} / W_{\text{sat}},$$

where $f_i$ is the fractional area of the $i$th subpixel area with the volumetric soil moisture $W_i$ (cm$^3$ cm$^{-3}$) and the field capacity of this subpixel area $W_{\text{sat},i}$ (cm$^3$ cm$^{-3}$); $n$ is the number of the subpixel areas; $W_{r,s}$ (cm$^3$ cm$^{-3}$) is the volumetric soil moisture of the whole pixel region; and $W_{\text{sat}}$ (cm$^3$ cm$^{-3}$) is the regional field capacity.

In Eq. (3), the spatial heterogeneity of the soil moisture in one pixel region is formulated by the field capacity of the subpixel area, which is determined by its soil texture. The regional water saturation condition is represented by the sum of its subregion conditions. The fractional area $f_i$ and the water saturation condition of each subpixel $W_i/W_{\text{sat},i}$ in Eq. (3) are transformable to each other, because both can be taken as the weight.
factor of the other. Thus, in Eq. (3), the fractional area of WSS and standing water has the same definition as the regional water saturation condition $W_{r,s}/W_{sat}$. According to De Ridder (2000), the regional water saturation condition could be derived from the PEED by a linear relationship, which we will discuss in section 3d.

3) OVERVIEW OF THE RETRIEVAL METHOD

The processing steps to retrieve the fractional area of the WSS and standing water are as follows:

1) Gap filling and noise removal. Preprocess the passive microwave data by moving filter and HANTS to obtain a daily noise-free estimation of the PDBT and vertical BT data at 37 GHz. Aggregate the MOD09A1 and preprocess the NDVI by HANTS to obtain daily noise-free NDVI data (see section 3b).

2) Retrieve the PEED. The parameters of the simplified radiative transfer model in Eq. (1) need to be derived from the satellite observations in order to retrieve the PEED. The land surface temperature can be derived from the vertical BT at 37 GHz, and the vegetation fractional area can be obtained from NDVI data. The derivation of the vegetation optical thickness from NDVI is discussed in detail in section 3c.

3) Retrieve the fractional area of the WSS and standing water. The parameters of the linear model in Eq. (2) need to be determined through the model simulation of soil–water mixture and surface roughness, because there are no field measurements for this area. Then, the fractional area of the WSS and standing water can be retrieved using Eq. (2) (see section 3d).

Besides the steps listed above, the FFT method is used to determine the major parameters used in HANTS (see section 3b). To validate the frequency and model dependence of the parameters’ derivation of Eq. (2), we used two soil–water mixture models, the Wang–Schmugge model and the Dobson model, to simulate the PEED at frequencies of 1.4, 19, and 37 GHz (see section 3c).

b. Gap filling and noise removal

To create a consistent time series of microwave data, erroneous observations need to be identified and removed and the gaps in the time series need to be filled. First, we used a 10-day moving window filter to preprocess the time series of PDBT and vertically polarized BT. In each window, the maximum and minimum non-zero values were removed, and the average of the remaining values was the estimation of the central day. The window size was chosen according to the gap size in the time series (the gap size caused by the revisiting schedule is 3–5 days) and the need to retain a sufficient number of samples. The upper and lower envelopes of the time series are potential erroneous observations (larger scatter of the upper and lower envelope samples is shown in Fig. 2). Extremely low values are probably caused by the raindrops and wind. Besides those, the integration method used in the SSM/I EASE-Grid product introduces extremely high and extremely low values. That is because a weighted sum of the brightness temperature in each $4 \times 4$ array is used to represent the brightness temperature at its central location in the EASE-Grid products (Poe 1990). These weight factors are the description of the antenna pattern. Thus, nearby pixels will significantly influence the gridded data, with possible under- and overestimations of PDBT, so we removed the maximum and minimum value in the moving window.
However, when durative measurements are taken under precipitation, for example, during the rainy season, the filtered time series will be underestimated. To remove such noise in filtered time series, we applied the HANTS algorithm (Menenti et al. 1993; Roerink et al. 2000; Verhoef 1996). The main user-selectable parameters of the HANTS algorithm are the number of harmonics and their period, high/low outlier direction, valid data range, fit error tolerance, and degree of overdeterminedness [see Jia et al. (2011) and Roerink et al. (2000) for detailed information].

To determine the number and periods of the frequencies used in the HANTS algorithm, we have applied the FFT algorithm to 1-yr time series of both 37-GHz PDBT and vertical BT data. The components with the first three largest amplitudes were chosen for the three main land use types in the Poyang Lake floodplain—wetland, open water, and cropland. These components are enough to regress the upper envelope of the original time series (called a reconstruction). The precipitation-contaminated observations have lower values and thus are removed as outliers during the reconstruction. The HANTS algorithm was also applied to NDVI time series to obtain cloud-free daily NDVI. The number and periods of the dominant frequencies are determined in a similar way as above.

c. Parameter derivation of the simplified radiative transfer model

To calculate the fractional area of WSS and standing water using Eq. (1), land surface temperature, fractional vegetation area, and vegetation optical thickness need to be derived from satellite observations. The variable $T_s$ can be obtained from the vertical BT $T_{B,37V}$ at 37 GHz (Holmes et al. 2009):

$$T_s = 1.11(T_{B,37V}) - 15.2. \quad (4)$$

The green vegetation fractional area $f_{veg}$ can be estimated from NDVI as

$$f_{veg} = \frac{(NDVI - NDVI_{soil})}{(NDVI_{veg} - NDVI_{soil})}, \quad (5)$$

where NDVI_{soil} is that of bare soil and NDVI_{veg} is that at full vegetation cover. In this case, NDVI_{soil} is set as 0 because of water-saturated soil. The NDVI_{veg} of rice is set as 0.60, the same as that of shrubland.

The vegetation optical thickness is approximately a linear function of NDVI (van de Griend and Owe 1993). The opposite fluctuations of the time series of NDVI and PDBT in arid areas suggest that when the surface temperature and water saturation are constant, the $\Delta T$ changes because of vegetation transmission. To derive the vegetation optical thickness, we modified Eq. (1) as

$$\Delta T = \Delta T_s (1 - f_{veg}) + f_{veg} \exp[-A(NDVI)], \quad (6)$$

where $\Delta T_s = T_s(\delta)$, which is the PDBT emitted from the land surface. In Eq. (6), the vegetation optical thickness is linearly determined by NDVI with a coefficient $A$. The angle of incidence is also contained in this coefficient. To simplify our process, we defined this exponential function in Eq. (6) as the vegetation transmission function (denoted as $\delta$):

$$\delta = \exp[-A(NDVI)]. \quad (7)$$

So, in fact, we derive the vegetation transmission function of Eq. (7), instead of the vegetation optical thickness, from NDVI.

To determine $A$ using Eq. (6), a set of observations where $\Delta T_s$ is constant and $\Delta T$ only changes with NDVI is necessary. Because $f_{veg}$ can be obtained from NDVI using Eq. (5), the only undetermined parameters in Eq. (6) are two constants: $\Delta T_s$ and $A$. Thus, by regressing NDVI with its dependent variable $\Delta T$ using Eq. (6), we can derive $A$ (see the equation in Fig. 6, described in greater detail below).

Flooded paddy fields fit this assumption very well. Paddy fields are flooded in specific periods of the growing season (as shown in Table 1). Normally, between two adjacent flooding periods, there is a long dry period. Thus, the irrigated paddy fields can be taken as flooded or water-saturated area and the unirrigated paddy fields can be taken as dry soil. The fractional area of paddy fields does not change year by year according to local statistical records, and rice growth demands a strict irrigation plan. This guarantees consistent PEED during the flooding periods. The land surface temperature is also constant in the early morning of the following season.
flooding periods (SSM/I measures in the early morning) because of the relative high thermal capacity of water. Thus, \( \Delta T_f \) (i.e., underneath the rice canopy) of the paddy fields can be taken as a constant during the flooding periods. Then, we can apply Eq. (6) on the coupled \( \Delta T \) and NDVI observation to regress \( A \).

d. Parameter derivation of the linear model

The linear model of Eq. (2) comes from the hypothesis that the regional water saturation condition is linearly related with the retrieved PEED as follows:

\[
W_{r,L}/W_{sat} = (\Delta \varepsilon - \Delta \varepsilon_{dry})/(\Delta \varepsilon_{sat} - \Delta \varepsilon_{dry}).
\]

This has been proven at 19 GHz by De Ridder (2000) using the Wang–Schmugge model (Wang and Schmugge 1980) to simulate the dielectric constant of soil–water mixture for a smooth surface. However, the dielectric constant of a soil–water mixture \( \varepsilon_w \) is frequency \( \nu \) dependent, and there are some other models available, such as the Dobson model (Dobson et al. 1985). In this case, both the Wang–Schmugge and the Dobson models are used to estimate the dielectric constant of soil–water mixture at frequencies of 1.4, 19, and 37 GHz. Then, the PEED of rough soil surface is calculated by the multifrequency-polarization surface emission \( (Q_p) \) model (Shi et al. 2005) as follows:

\[
\Delta \varepsilon = [1 - (Q_{H} + Q_{V})][\Gamma_{s,s}(H, \theta, \nu) - \Gamma_{s,s}(V, \theta, \nu)],
\]

where \( Q_{H} \) and \( Q_{V} \) are the surface roughness parameters for horizontal H and vertical V polarizations, respectively, at the angle of incidence \( \theta \). They can be obtained from the roughness property of the surface, named \( s/l \), which is the ratio of RMS height \( s \) to correlation length \( l \). Variables \( \Gamma_{s,s}(H, \theta, \nu) \) and \( \Gamma_{s,s}(V, \theta, \nu) \) are the reflectivity of a smooth surface for horizontal and vertical polarization, respectively, which is determined by the Fresnel equations:

\[
\Gamma_{s,s}(H, \theta, \nu) = \left( \frac{\cos \theta - \sqrt{\varepsilon_w - \sin^2 \theta}}{\cos \theta + \sqrt{\varepsilon_w - \sin^2 \theta}} \right)^2
\]

(10)

and

\[
\Gamma_{s,s}(V, \theta, \nu) = \left( \frac{\varepsilon_w \cos \theta - \sqrt{\varepsilon_w - \sin^2 \theta}}{\varepsilon_w \cos \theta + \sqrt{\varepsilon_w - \sin^2 \theta}} \right)^2.
\]

(11)

The Dobson (Dobson et al. 1985) and the Wang–Schmugge (Wang and Schmugge 1980) models are used to calculate the dielectric constant of soil–water mixture at frequencies of 1.4, 19, and 37 GHz. The dominant parameter of these two models is soil texture. A silt-loam soil was used in this numerical analysis with 13.5% clay and 30.6% sand, since it is the major soil type in the Poyang Lake area. Because of the calibration difficulty, the parameters of the Dobson model were not recalibrated while the parameters of the Wang–Schmugge model at 37 GHz were recalibrated by Calvet et al. (1995).

4. Results

a. Gap filling and noise removal on time series of PDBT

Our FFT analysis of the PDBT time series showed that there were eight dominant frequencies and that their corresponding periods were 365, 184, 123, 91, 73, 61, 46, and 30 days. The vertically polarized BT time series had exactly the same frequencies, according to their own FFT analysis. We set the other parameters required to apply the HANTS algorithm to the PDBT and vertical BT time series as follows:

- the direction of outliers was lower,
- the fit error tolerance was 1.5 K,
- the degree of overdeterminedness was 80, and
- the valid data ranges were 3–100 K for PDBT and 200–400 K for vertical BT.

The FFT analysis of NDVI time series showed that there were six frequencies and their corresponding periods were 365, 184, 123, 91, 74, and 61 days. The other parameters were set as follows:

- the direction of outliers was lower,
- the fit error tolerance was 0.05,
- the degree of overdeterminedness was 20, and
- the valid data range was 0–1.

The moving window–filtered brightness temperature data are the daily estimation. They follow the overall trend of the raw data in Fig. 3. The raw data time series in Fig. 3 show much scattering, while the moving window–filtered data give a much clearer trend of the time series.

The sharp sags of raw and filtered PDBT during the continuous rainfall is probably caused by the attenuation of raindrops as they occur at the same duration as the precipitation in Fig. 3. The HANTS algorithm throws out these relatively lower-valued samples as outliers. The upper envelope of the filtered time series is followed by the reconstruction time series. HANTS can work because the spatial heterogeneity slows the response of the regional soil moisture to precipitation. The time series of daily region-averaged soil moisture observed in Tibet is much smoother than that at hourly resolution, according to the field observations by Zhao et al. (2013).
b. Parameterization of the vegetation transmission function

To compare with the PDBT data, MOD13A2 NDVI data were aggregated to a 25 km × 25 km EASE-Grid, because the aggregation from NDVI and reflectance (MOD09A1) does not show large differences for the paddy fields (shown in Fig. 4), and the data quality of MOD13A2 is much better than MOD09A1 without sudden low values (cloud-contaminated samples) in the aggregated NDVI time series of MOD13A2. This consistency is only found for cropland area. For other land types with large spatial heterogeneity, NDVI data are calculated from the spatially aggregated reflectance data (MOD09A1).

During each rice cropping season in 2002 (one from April to July and the other from July to October, according to the NDVI time series of Fig. 5), there is a significant peak occurring in PDBT time series at its start time. This indicates the sowing time and the start of irrigation of the paddy fields, since the PDBT increases with WSS and standing water area. The flood duration is about 50 days, as shown in Table 1. It covers several stages of rice phenology and includes three or four observations of the 16-day composite NDVI data. Since the irrigation schedule is similar among fields, the surface emittance and PDBT can be taken as constants during these 50 days at the observation time (early morning). Thus, the radiometer-measured PDBT is only affected by the vegetation fractional area and optical thickness. The vegetation fractional area can be calculated from NDVI using Eq. (5). The vegetation optical thickness, which describes the scattering and absorption of the surface emittance by vegetation, can be parameterized by NDVI using the vegetation transmission function of Eq. (7). The simplified radiative transfer model appears in Eq. (6) to clarify this.

To get sufficient observations to estimate the coefficients in Eq. (6), we chose NDVI data and corresponding average
PDBT data of one EASE-Grid pixel area (25 km x 25 km; centrally located at 28.3807°N, 115.8351°E) during the rice flooding periods from 2001 to 2003 (as shown in Fig. 6), since this area is almost entirely covered by paddy fields and the rice area in this region shows little variability during these 3 years according to local statistical records. By fitting the data points with Eq. (6), we get the regression equation as shown in Fig. 6: $\Delta T_s$ in Eq. (6) is 26.9 K and $A$ is 1.23179. Thus, the vegetation transmission function of Eq. (7) is

$$\delta = \exp[-1.23179(NDVI)].$$  \hfill (12)

In this regression, we did not use Eq. (4) to calculate the land surface temperature, because the model uncertainty of Eq. (4) might be introduced into the derivation of the vegetation transmission function. The uncertainty of this regression is highly dependent on the vegetation fractional area. We compared the linear and quadratic model for deriving vegetation fractional area from NDVI. When the same setting of $NDVI_{\text{veg}}$ and $NDVI_{\text{soil}}$ as described in section 3b is used, the RMSE of our regression with the quadratic model is 3.2 K, higher than that in Fig. 6.

The parameter definition of vegetation fractional area model in Eq. (5) has little influence on this regression. We evaluated different settings of $NDVI_{\text{soil}}$ (−0.05, 0, and 0.05), which is the major model uncertainty in Eq. (5). The RMSE of our regression only varies from 2.7 to 2.8 K. Because there is no field measurement of NDVI for flooded soil, which could have negative values close to water, we simply set $NDVI_{\text{soil}} = 0$.

c. Evaluation of the relationship between the emissivity difference and soil wetness

We modeled the smooth surface emissivity difference at 1.4, 19, and 37 GHz with both the Wang–Schmugge and the Dobson models, as shown in Fig. 7. Both model

$$\Delta T = 26.8917[-f_v + f_s \times \exp(-1.23179 \times NDVI)]$$

FIG. 6. Scatterplots of PDBT (i.e., $\Delta T$) data and corresponding NDVI observations and the result of curve fitting.
simulations show (Fig. 7) that a quasi-linear relationship exists at all frequencies, but with different slopes and interceptions. The Wang–Schmugge model gives a very similar slope at different frequencies: the emissivity difference increases from 0.17 with volumetric soil moisture, reaching a plateau at $\Delta \varepsilon = 0.3$, when the volumetric soil moisture is close to $0.255 \text{ cm}^3 \text{ cm}^{-3}$ (Fig. 7a), which is the field capacity of the silt-loam soil (De Ridder 2000). The Dobson model simulation shows that the slope of the quasi-linear relationship is dependent on frequencies, while the plateau of $\Delta \varepsilon$ has very similar values at all three frequencies (Fig. 7b). At 1.4 GHz, $\Delta \varepsilon$ increases from 0.16 and reaches the plateau at $\Delta \varepsilon = 0.3$ for a soil volumetric moisture close to $0.14 \text{ cm}^3 \text{ cm}^{-3}$. At 19 GHz, $\Delta \varepsilon$ starts to increase at 0.15 and reaches the same plateau when the soil gets wetter than for 1.4 GHz, at a soil volumetric moisture around $0.26 \text{ cm}^3 \text{ cm}^{-3}$. At 37 GHz, it increases from an even lower $\Delta \varepsilon$ around 0.14 and reaches the plateau when the soil moisture is close to $0.5 \text{ cm}^3 \text{ cm}^{-3}$.

The Dobson model suggests that at the higher frequency the relationship is much closer to linear, while the Wang–Schmugge model gives similar quasi-linear relationships at all frequencies. The interesting thing is that linear or quasi-linear relationships occurred at all frequencies in both model simulations and the plateau of $\Delta \varepsilon$ is almost constant for different frequencies. The Dobson model gives a frequency-dependent slope and interception of $\Delta \varepsilon$.

The difference between the Wang–Schmugge and the Dobson models lies in their theoretical basis. The Wang–Schmugge model is a linear mixture of the soil components’ dielectric constants. It used the wetting point to describe maximum bound water fraction, while the Dobson model used the double-layer model (Dobson et al. 1985) to describe the Debye-type relaxation of soil–water mixture. The Dobson model, however, is a linear mixture of the components’ dielectric constants to the power $\alpha$. The power $\alpha$ is obtained from field measurements. From a theoretical point of view, the Dobson model better describes the dielectric constant at higher frequencies, for example, 37 GHz, because it accounts for the imaginary loss of microwave.

We used the Dobson model to calculate the dielectric constant at 37 GHz in Eqs. (10) and (11). Then, the PEED was simulated by the $Q_p$ roughness model in Eq. (9). The linear relationships between PEED and soil moisture still occur, as shown in Fig. 8, though the interception and the plateau of $\Delta \varepsilon$ are roughness dependent. The $s/l$ is the RMS height divided by the correlation length, which represents the surface roughness: a smaller $s/l$ value means a smoother surface. The slope of the relationships calculated for three different roughnesses slightly decreases with increasing roughness. This suggests that the fractional area of WSS and standing water as a linear function of the PEED [Eq. (2)] should depend on surface roughness. Because the terrain around the lake area is relatively smooth, we set the same roughness condition for all pixels (the white-numbered squares in Fig. 1).

The wind’s influence on surface roughness is not significant, since 58% of the wind speed measured at the Boyang weather station near Poyang Lake is less than $2.0 \text{ m s}^{-1}$ and 97% is less than $4.0 \text{ m s}^{-1}$. That is because Poyang Lake and its floodplain are surrounded by mountains. The shape of the Poyang Lake floodplain looks like a quadrangle, with its northern edge facing to the northwest and its southern edge facing to the southeast. The eastern and western edges are almost parallel. The mean north–south distance is about 140 km, and the mean east–west distance is about 120 km. The lake has an irregular shape; it is located in the central floodplain and its eastern boundary is elongated in the same way as the eastern edge of the floodplain, with a north–south distance of about 100 km and an east–west distance of about 80 km.
lake connects to the Yangtze River to the north through a funnel-shaped area, with a wide mouth to the south and narrow head to the north. Thus, in the lake area, a north–south wind blows on most days. Only in July and August does the monsoon from the Indian Ocean force a south–north wind. The connecting funnel area yields a higher wind speed in the narrow head to the north and the wind speed decreases from north to south over the floodplain, after passing through this connecting area.

Thus, the largest mean wind speed occurs north of Poyang Lake, with an annual mean wind speed between 3 and 4 m s\(^{-1}\) from 1998 to 2007. In the center of the lake area, the mean wind speed is reduced to 2–3 m s\(^{-1}\), while at the Nanchang meteorology station south of the lake, the mean wind speed is only 1.7 m s\(^{-1}\).

To derive the model expressed by Eq. (2), the PEED of dry and water-saturated soil needs to be determined. According to Choudhury (1989), the PEED of dry soil is 0.091. Its surface roughness was calculated for an open water area by Choudhury with a corresponding \(s/l = 0.175\). The PEED of the paddy fields in Fig. 6, however, is close to 0.1 (the land surface temperature is assumed to be 270 K). This suggests that the land surface is much rougher than open water; thus, \(s/l\) is set as 0.25, with a corresponding \(\Delta e_{\text{dry}} = 0.068\). For completely water-saturated soil, we can assume that its PEED is the same as standing water, 0.21, given by Choudhury (1989). Equation (2) is then expressed as

\[
 f_{\text{ws}} = (\Delta e - 0.068)/(0.21 - 0.068). \tag{13}
\]

### d. Retrieved WSS and Standing Water Area of Poyang Lake

Since we derive the vegetation transmission function using Eq. (12), PEED can be retrieved from noise-free PDBT using Eq. (1), where \(T_s\) is obtained from noise-free vertical brightness temperature at 37 GHz [Eq. (4)] and the vegetation fractional area from noise-free NDVI [Eq. (5)]. The fractional area of WSS and standing water can then be calculated from the PEED using Eq. (13).

The WSS and standing water area enclosing Poyang Lake (white-numbered squares in Fig. 1) is calculated from 2001 to 2003 and compared with the Poyang Lake area from MODIS (with a spatial resolution of 250 m) and ASAR (with a spatial resolution of 30 m) data (Yésou et al. 2011, 2013), as shown in Table 2 (Yésou et al. 2011). The trend of our retrieved WSS and standing water area is close to the Poyang Lake area, as shown in Fig. 9. The RMSE of the retrieved WSS and standing water area is 361.22 km\(^2\) or 17.74% of mean area during the same period of time. A wet and dry case of the pixel region centered at 29.0485°N, 116.0954°E is shown in Fig. 10. In the wet case of 4 July

<table>
<thead>
<tr>
<th>Date</th>
<th>Water extension from Yésou et al. (2011) (km(^2))</th>
<th>Retrieved WSS and standing water area from SSM/I (km(^2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>17 Jan 2001</td>
<td>1071.13</td>
<td>1308.71</td>
</tr>
<tr>
<td>12 Apr 2001</td>
<td>2186.69</td>
<td>1617.73</td>
</tr>
<tr>
<td>3 Jul 2001</td>
<td>2771.81</td>
<td>3072.36</td>
</tr>
<tr>
<td>16 Nov 2001</td>
<td>1669.44</td>
<td>947.683</td>
</tr>
<tr>
<td>9 Jan 2002</td>
<td>1174.31</td>
<td>1188.72</td>
</tr>
<tr>
<td>12 Apr 2002</td>
<td>1495</td>
<td>2173.18</td>
</tr>
<tr>
<td>21 Aug 2002</td>
<td>3212.81</td>
<td>3761.77</td>
</tr>
<tr>
<td>10 Nov 2002</td>
<td>2007</td>
<td>1287.61</td>
</tr>
<tr>
<td>25 Jan 2003</td>
<td>1751.09</td>
<td>1978.5</td>
</tr>
<tr>
<td>20 Jun 2003</td>
<td>2285.62</td>
<td>2319.19</td>
</tr>
<tr>
<td>25 Jul 2003</td>
<td>3305.69</td>
<td>3307.12</td>
</tr>
<tr>
<td>7 Nov 2003</td>
<td>1507.89</td>
<td>698.439</td>
</tr>
</tbody>
</table>
2002, the fractional area of WSS and standing water, which is calculated from SSM/I (65.28%), is close to the fractional area of open water (58.81%) observed from MODIS NDVI images. However, in the dry case of 24 October 2002, the fractional area of WSS and standing water area (46.24%) is much larger than that of open water (21.85%) from NDVI. That is because the degraded lake area becomes wet land and is occupied by aquatic vegetation, mainly Phragmites. An optical sensor, like MODIS, cannot observe the standing water beneath the vegetation canopy, while a microwave radiometer can penetrate through the vegetation.

Fig. 9. Scatterplot of lake area estimated with MODIS and ASAR data by Yésou et al. (2011) and WSS and standing water area calculated from SSM/I 37 GHz.

<table>
<thead>
<tr>
<th>DOY</th>
<th>July 4th 2002</th>
<th>October 24th 2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image</td>
<td></td>
<td></td>
</tr>
<tr>
<td>label</td>
<td>Green vegetation</td>
<td>Vegetation soil mixture</td>
</tr>
</tbody>
</table>

Fig. 10. The (left) wet and (right) dry case of the five numbered pixels at Poyang Lake with corresponding NDVI images. The white rectangle inside the NDVI images is the area of the five numbered pixels of SSM/I 25-km EASE-Grid in the Poyang Lake area (Fig. 1) on 4 Jul and 24 Oct 2002. The upper part of the white rectangle represents the percentage (the numbers outside the parentheses) of the WSS and standing water area, which is calculated from SSM/I. The NDVI image is obtained from MOD09A1 with a spatial resolution of 500 m. The fractional open water area calculated from NDVI is shown in the parentheses.
The difference between our retrieved WSS and standing water area and the lake area from MODIS and ASAR data is mainly caused by three factors:

1) The study area is different. Yésou et al. (2011) focus on the Poyang Lake area, while our study area is 6250 km² and covers not only Poyang Lake, but also the wetland and cropland adjacent to Poyang Lake.

2) The observed features are different. The MODIS and ASAR capture the open water, while we observed the WSS and standing water. That means we cannot discriminate open water, flooded vegetation, and water-saturated soil using our model.

3) The temporal resolution is different. SSM/I data supply daily brightness temperature data, while the MODIS data used by Yésou et al. (2011) are 8-day composites and ASAR has a 35-day repeat orbit. Thus, near-daily monitoring is one of the advantages of using 37-GHz microwave radiometer data.

Another very important factor is the different impacts of vegetation on the three types of observation. The MODIS observations cannot penetrate under vegetation. In principle, ASAR can capture the backscatter from the water surface underneath the vegetation, but the scattering by vegetation is very strong. Our model takes the emittance of the WSS and standing water under vegetation cover into account.

After intersensor calibration, the retrieved sea surface temperature from these two datasets has a difference of less than 0.5 K. Though the brightness temperature data of the two radiometers have both been registered into the EASE-Grid system, the instantaneous field of view (IFOV) of original data is rather different: 37 km × 28 km for SSM/I and 14 km × 8 km for AMSR-E. Thus, AMSR-E observes detailed changes in the lake area, especially during the dry season. The PDBT of SSM/I seems to be saturated earlier than AMSR-E, as shown by, for example, the peaks during May–October, since the AMSR-E observes a higher fractional WSS and standing water area in its IFOV than SSM/I. Taking the peaks in March and November in Fig. 11 as an example, the PDBT of SSM/I does not increase simultaneously with the AMSR-E, but 10 days later. In the period of 4–24 November 2003 (rectangle in Fig. 11), the SSM/I PDBT is about 5 K lower than AMSR-E, with the retrieved WSS and standing water area 345 km² lower than that from AMSR-E. The time series trend should be almost the same since two sensors both measure before sunrise. Over all comparison samples, our retrieved WSS and standing water area from SSM/I is closer to the Poyang Lake area from MODIS and SAR than that from AMSR-E, except for the samples in November 2003. Thus, there is a need for future work on the analysis of these two datasets.

To study the relationship between the soil water saturation conditions in the upstream catchment and the inundation pattern downstream, we retrieved the WSS and standing water area in the upper catchment. The upstream area of the Poyang Lake is indicated in Fig. 1 by the black-numbered EASE-Grid cells, and it includes the major tributary river systems from the western, southwestern, southeastern, and eastern directions. The downstream area is the lake area indicated by the white-numbered EASE-Grid cells.

The time series of the Poyang Lake area from 2001 to 2008 have similar fluctuations with the upstream WSS and standing water area (Fig. 12). For most periods, the Poyang Lake area increases with the upstream WSS and
standing water area. This is proven by the cross-correlation analysis between these two time series (Fig. 13): the maximum cross-correlation values from 2001 to 2008 are all higher than 0.5 and the time lag is no more than 5 days. The time lag indicates that the two time series are better correlated if shifted by this time lag relative to the other. In this case, it means that the increase of WSS and standing water area upstream will come earlier than that of the Poyang Lake area.

There are some peaks of the upstream WSS and standing water area that do not lead to the simultaneous increasing of lake area in Fig. 12. The reason is that the water source changes for the lake area. According to the long-term hydrological data in Poyang Lake, before July of each year, the water from upstream of the Poyang Lake floodplain (it is the subbasin of the Yangtze River basin) is the major source of lake. During this time, the lake area fluctuation is closely related with the WSS and standing water upstream in most years. There are some spikes in the upstream not leading to the same response in Poyang Lake, for example, at the end of April 2002 and 2003. That is because the surface runoff and streamflow is consumed by irrigation and other human activities before reaching the lake. From August of each year, the Poyang Lake area is supplied from both the upstream Poyang Lake floodplain and the Yangtze River. Thus, the asynchronism between upstream WSS and standing water area and the lake area becomes larger. For example, the spikes in the lake area in September 2001 and 2008 are probably caused by the water from the Yangtze River.

It is also interesting to observe that since 2003, the maximum cross-correlation value increased significantly. The major reason why the correlation coefficient changed after 2002 is that the water supplied by the Yangtze River to Poyang Lake is limited to the period from August to September, that is, the flooding season of the Yangtze River. Thus, the water source of Poyang Lake is mainly from the upstream catchment, and the correlation coefficients increase from 2003 to 2006. Since 2007, the dam stopped storing water and released the streamflow as usual. An obvious decrease then occurs in the coefficients of Fig. 13. This again proves that the water supplement from the Yangtze River will significantly influence the lake changes since August.

5. Conclusions

We used a simplified radiative transfer model and linear model to retrieve the fractional area of water-saturated soil (WSS) and standing water from the polarization difference brightness temperature (PDBT) at 37 GHz. The moving filter is first used to fill the observation gaps of the PDBT time series. Noise-free daily PDBT, vertical brightness temperature (BT), and NDVI

![Figure 12](image-url)  
**FIG. 12.** Comparison between the upstream WSS and standing water area and that of the Poyang Lake area.

![Figure 13](image-url)  
**FIG. 13.** Max cross-correlation values between time series of upstream WSS and standing water area and that of the Poyang Lake area from 2001 to 2008 and the days that max time lag occurs.
are derived from the Harmonic Analysis of Time Series (HANTS) algorithm of their raw data time series. The vegetation transmission function is obtained from the regression between NDVI and its dependent variable PDBT for flooded paddy fields, under the assumption that the land surface temperature and PEED of the flooded rice is constant during its growing season. The quasi-linear relationship between the PEED and fractional area of WSS and standing water exists no matter what the frequencies are. This indicates that the regional water-storage capacity can be derived from PDBT observations of other frequencies. The retrieved WSS and standing water area in the Poyang Lake region shows a good fit with the lake area from the MODIS and SAR data. Our method seems to provide satisfactory estimates of the water-storage capacity of the upper catchment of the Poyang Lake, and as such, it is a useful early indicator of flooding events in the lake area.

Acknowledgments. This work has been supported by the CAS/SAFEA International Partnership Program for Creative Research Teams (Grant KZKD-EW-TZ-09) and the EU-FP7 CEOP-AEGIS project (Grant 212921). This investigation is a cooperative effort of the Delft University of Technology, Delft, Netherlands, and the State Key Laboratory of Remote Sensing Sciences, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China. We thank Dr. Hervé Yéou and Dr. Claire Huber of SERTIT—University of Strasbourg for sharing their data on the Poyang Lake area with us. Many thanks also to Prof. Jiancheng Shi, the chair of the State Key Laboratory of Remote Sensing Sciences, for his constructive suggestions on our work.

REFERENCES


