CrossMark

ORIGINAL ARTICLE

An algorithm for soil moisture estimation using GPS-interferometric reflectometry for bare and vegetated soil

Clara Chew[1] · Eric E. Small[1] · Kristine M. Larson[2]

Abstract Ground-reflected global positioning system signals measured by a geodetic-quality GPS system can be used to infer temporal changes in near-surface soil moisture for the area surrounding the antenna. This technique, known as GPS-interferometric reflectometry, analyzes changes in the interference pattern of the direct and reflected signals, which are recorded in signal-to-noise ratio (SNR) data, as interferograms. Temporal fluctuations in the phase of the interferogram are indicative of changes in near-surface volumetric soil moisture content. However, SNR phase is also highly sensitive to changes in overlying vegetation, and thus, the effects of seasonal vegetation changes on the ground-reflected signal must be considered. Here a method is described for determining whether SNR data are significantly corrupted by vegetation and for correcting these effects. Absolute soil moisture content must be determined for each site using ancillary data for the residual moisture content. Accounting for vegetation effects significantly improves the agreement between GPS-derived soil moisture and in situ measurements.

Keywords Hydrogeodesy · Soil moisture · Multipath · Global positioning systems

✉ Clara Chew
    claracchew@gmail.com; clara.chew@colorado.edu

[1] Department of Geological Sciences, University of Colorado, 2200 Colorado Ave. 399 UCB, #285, Boulder, CO 80309-0399, USA

[2] Department of Aerospace Engineering Sciences, University of Colorado, 1111 Engineering Dr. 429 UCB, Boulder, CO 80309-0429, USA

Introduction

In recent years, GPS multipath signals have been used opportunistically to infer land surface conditions, such as snow depth, soil moisture, and changes in vegetation conditions. The use of ground-reflected navigation signals in remote sensing was first proposed by Hall and Cordey (1988) to determine surface wind velocity over the ocean. Subsequent studies primarily evaluated the potential of GPS ground reflections to estimate sea surface wind speed (Komjathy et al. 2000b; Zavorotny and Voronovich 2000; Cardellach et al. 2003), sea level (Martin-Neira 1993; Lowe et al. 2002) or to estimate sea ice thickness (Komjathy et al. 2000a). More recent studies have explored the ability of GNSS ground reflections to estimate land surface conditions, including snow depth (Rodriguez-Alvarez et al. 2011a; Cardellach et al. 2012), soil moisture (Masters 2004), and vegetation state (Ferrazzoli et al. 2011; Rodriguez-Alvarez et al. 2011b, 2012).

The above studies focused on using specially designed GPS receivers to infer changes in environmental conditions. A parallel effort has shown that geodetic-quality GPS antennas/receivers can also estimate changes in soil moisture (Larson et al. 2008a, b, 2010), snow depth (Larson and Nievinski 2013; McCreight et al. 2014), and vegetation water content (Wan et al. 2014). This technique, known as GPS-interferometric reflectometry (GPS-IR), relates temporal changes in signal-to-noise ratio (SNR) data (hereafter SNR interferograms) to changes in environmental conditions around a GPS antenna for an area which scales with the monument height (Katzberg and Garrison 1996).

In principle, GPS antennas and receivers are well suited for land surface remote sensing because the GPS satellite transmit frequency is L-band (microwave frequency) with wavelengths of ∼19 (L1) or 24.4 (L2) cm. Microwave

remote sensing of land surface properties has been studied for over 30 years (Ulaby et al. 1982; Tsang et al. 1985; Wang et al. 1986). For microwave frequencies, the permittivity or dielectric constant of the ground is highly dependent on its water content (Hallikainen et al. 1985). For example, the dielectric constant of water is around 80, compared to a dielectric constant of 3.5 for dry soils. The permittivity of a material determines the extent to which an electromagnetic wave will reflect off of the material, with higher permittivity materials resulting in a greater reflection (Fuks and Voronovich 2000).

Because of this effect, a ground-reflected GPS signal will be altered by the amount of water contained near the ground surface. The GPS system can record temporal changes in the ground-reflected signal, which correlate with changes in properties such as soil moisture, snow depth, and vegetation water content. However, the reflection is also affected by surface roughness and surface temperature, though these effects are negligible for the sites described in this study.

A modeling study described a retrieval algorithm for soil moisture for the top 5 cm of the soil column under bare soil conditions (Chew et al. 2014). A subsequent modeling study showed the effects of changing vegetation canopies on SNR data and how these effects could obfuscate soil moisture estimations (Chew et al. 2015). This was also demonstrated in a field study (Wan et al. 2014). Here, we present an improved algorithm for estimating soil moisture in the top 5 cm of the soil column using GPS-IR with considerations made for vegetated environments. We will present the algorithm and use observations from two different GPS stations to depict data processing steps and the resulting soil moisture time series. These data serve as illustrations of how the algorithm produces soil moisture estimations—a complete validation dataset is provided elsewhere.

Soil moisture measurement: standard methodologies

Volumetric soil moisture, or the ratio of the volume of water to the volume of bulk soil, is currently estimated using both in situ and remote sensing techniques. In situ soil moisture monitoring is performed either by physically drying a volume of soil to calculate the amount of water or through the use of electromagnetic probes (Rajkai and Ryden 1992; Robock et al. 2000). These methods provide information about soil moisture content for a volume <1 m^3. Soil moisture typically varies over relatively small spatial scales, meters to hundreds of meters (Famiglietti et al. 1998; Gomez-Plaza et al. 2001; Brocca et al. 2007; Baroni et al. 2013), so it is often desirable to have a larger,

areal-averaged estimate. This can be accomplished either by using a large number of in situ probes that are combined to yield a larger-scale average or via remote sensing methods with larger footprints.

Soil moisture remote sensing is usually done using a monostatic radar or radiometer. Both monostatic radars and radiometers may be mounted on towers or truck booms (Schwank et al. 2005), flown on airplanes (Jackson et al. 1995; Jackson and Le Vine 1996; Mladenova et al. 2011), or be mounted on satellites or spacecraft (Kerr et al. 2001; Entekhabi et al. 2010). Those that are mounted on towers or truck booms can give soil moisture estimates on the field scale, or tens to hundreds of square meters, while those on airplanes and satellites have successively larger sensing areas. These measurements tend to be conducted during field campaigns and do not tend to yield long-term records. The exception to this would be instruments on satellites, such as the Soil Moisture and Ocean Salinity or upcoming Soil Moisture Active–Passive satellites, which give estimates of soil moisture approximately every 3 days (Entekhabi et al. 2010). The sensing area of spaceborne radars or radiometers is between 10 and 1600 km^2.

GPS-interferometric reflectometry and SNR data

GPS-interferometric reflectometry estimates changes in soil moisture and other land surface properties by recording temporal changes in characteristics of SNR data recorded by a GPS receiver. The interference between the direct and reflected, or multipath, signals produces an oscillatory pattern superimposed on the direct signal for low satellite elevation angles (Fig. 1a). The SNR data for each rising or setting satellite track are first converted from dB-Hz to a linear scale (volts/volts). A low-order polynomial is then fit to the data to retain only the interference pattern (Fig. 1b, c).

Initial studies characterized the SNR interferogram using the following equation (Larson et al. 2008a, 2010):

$$\text{SNR} = A \cos \left(\frac{4\pi H_0}{\lambda} \sin E + \phi \right) \qquad (1)$$

where H_0 is the a priori reflector height. For bare, flat ground H_0 is within a few centimeters of the height of the antenna phase center above the top of the soil. E is the elevation angle of the satellite, A is an amplitude term, λ is the GPS wavelength, and ϕ is a phase shift. Phase and amplitude are calculated using least-squares estimation.

This analysis assumes that the SNR data have a constant frequency and amplitude as a function of sine of E, which is a simplification. Despite this, strong correlations are observed between ϕ and in situ measurements of soil moisture (Larson et al. 2008b), which were later

Fig. 1 SNR interferograms are observations on DOY 120 in 2012 at the Oklahoma GPS station (ok13). **a** SNR interferogram from PRN 7. **b** Same SNR data as in **a**, but converted to a linear scale and detrended with a low-order polynomial. **c** Detrended SNR interferograms from four different satellite tracks for DOY 120 in 2012 at GPS station ok13. Interferograms are vertically offset for clarity. Useable and non-useable tracks for GPS-IR are labeled "good" and "bad," respectively. **d** Lomb–Scargle periodograms for the interferograms in **c**
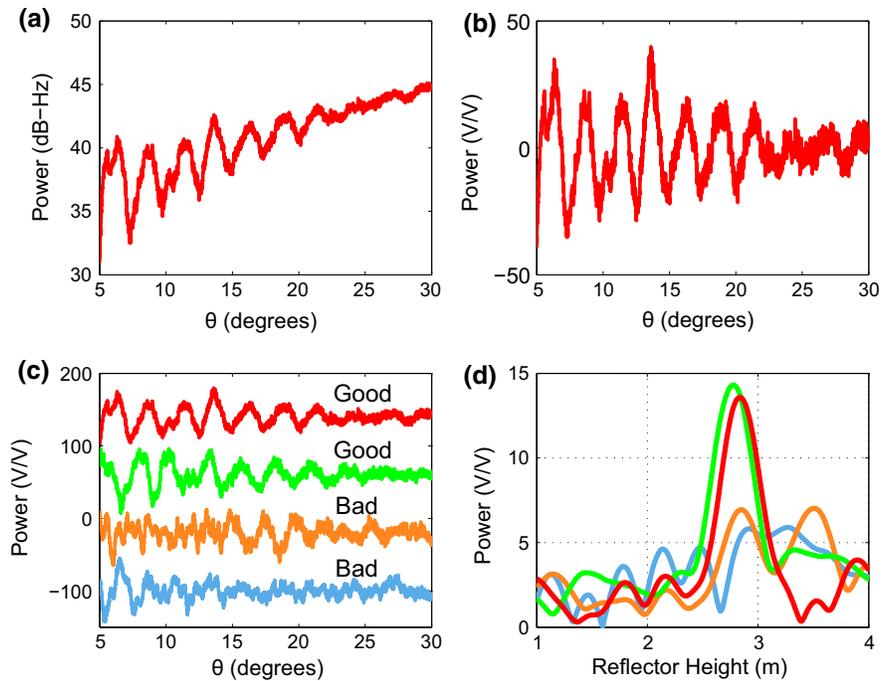
corroborated using a model of ground-reflected signals (Chew et al. 2014).

Alternatively, a Lomb–Scargle periodogram (LSP) can be used to estimate temporal changes in the dominant frequency of the SNR interferogram. This is the method used in GPS-IR estimations of snow depth (Larson and Nievinski 2013). An LSP is like a fast Fourier transform, except that it can be used for unevenly sampled data. Examples of LSPs are shown in Fig. 1d. The dominant frequency may be converted to an effective reflector height (H_{eff}) using

$$H_{\mathrm{eff(t)}} = \frac{1}{2} f_{m,t} \lambda \qquad (2)$$

where f_m is the peak frequency of the LSP and t is the value for a single day. The change in effective reflector height (ΔH_{eff}) throughout the year is defined by:

$$\Delta H_{\mathrm{eff(t)}} = H_0 - H_{\mathrm{eff(t)}} \qquad (3)$$

The power of the peak frequency of the LSP, or the LSP amplitude (A_{LSP}), is also affected by soil moisture, vegetation, and other site conditions such as topography.

Metrics derived from SNR data (ΔH_{eff}, A, A_{LSP}, and ϕ) have been found to be affected to various extents by changes in soil moisture (Larson et al. 2008a; Zavorotny et al. 2010; Chew et al. 2014), vegetation (Wan et al. 2014), and snow depth (Larson and Nievinski 2013). Two modeling studies found phase to be the best indicator of changes in soil moisture (Chew et al. 2014) and A to be the metric most related to changes in vegetation permittivity and canopy height (Chew et al. 2015). Phase, however, was

also found to be significantly affected by changes in vegetation (Chew et al. 2015). Thus, depending on the extent of vegetation at a GPS site, the effect of vegetation must first be quantified and removed before soil moisture can be estimated from GPS SNR data.

PBO H₂O network

The soil moisture algorithm to be described in the next section is currently implemented at stations in the EarthScope PBO H₂O network with the greatest variations in vegetation (Fig. 2; Larson and Small 2013). All sites in the PBO H₂O network have choke-ring antennas and nearly all use Trimble NetRS receivers. They are located in areas with limited topographic variations and are located in ecosystems characterized as grass or shrub lands. Some of the sites record data at 1-s intervals, and the rest record data at 15-s intervals.

Signal-to-noise ratio data from the new L2C GPS signals are used by PBO H₂O because the quality of the data are higher than either the legacy L1 C/A or L2P from non-code tracking receivers (Dunn 2010). The frequency of the L2C signal corresponds to a maximum penetration depth of 5 cm (Njoku and Entekhabi 1996). Since GPS-IR only analyzes reflections from low/grazing satellite elevation angles, the real sensing depth is likely to be 2.5 cm or less (Chew et al. 2014). PBO antennas tend to be between 1.5 and 2.1 m tall, which corresponds to a sensing area of approximately 120 m² per satellite track (Larson and
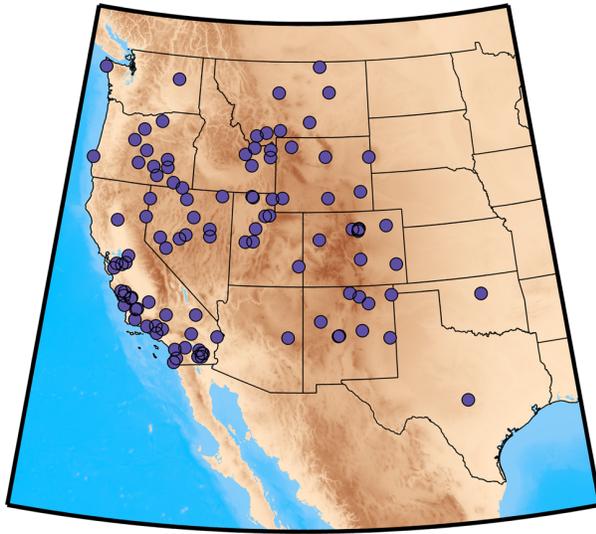
Fig. 2 Map showing locations of PBO H$_2$O soil moisture stations in the continental USA

Nievinski 2013). Multiple tracks are then combined so that the aggregate sensing area is approximately 1000 m^2.

GPS soil moisture estimation algorithm

Here, we describe the algorithm that is currently used to produce the soil moisture estimations for stations in the PBO H$_2$O network with the greatest variations in vegetation (Fig. 3). Included in the algorithm is a method for determining whether or not a phase time series is corrupted by vegetation effects. The method for removing these effects is presented in the subsequent section.

There are several parameters described in this section, and altering the parameters will affect the estimated soil moisture time series to greater or lesser degrees. A list of the parameters can be found in Table 1. They have been ordered in terms of their relative importance in the algorithm, with more important parameters being higher in the table. The parameters used in the processing of observations or smoothing the time series were chosen after observing their effects at several PBO stations.

Selection of useful satellite tracks

Not all satellite tracks can be used for GPS-IR. Tracks should have consistent reflections between satellite elevation angles of 5°–25° or 5°–30°; oscillations in the interferogram for higher angles are obscured by the antenna gain pattern. Tracks should also not be obstructed by trees or buildings or reflect from man-made surfaces like roads. A track should have a relatively stable singular dominant frequency for periods of the year when vegetation water content or height is nearly constant. In general, the power (A_{LSP}) of the dominant frequency (H_{eff}) should be at least twice as high as the power of the noise or second most powerful frequency in the periodogram.

Examples of SNR data from two useful satellite tracks (red and green) and two tracks with significant noise corruption (blue and orange) are shown in Fig. 1c. The noise-corrupted satellite tracks in Fig. 1c lack any dominant frequency, which could introduce significant errors in subsequent phase and amplitude estimation (Fig. 1d). Part of the reason for these two tracks lacking a dominant frequency is due to the fact that the two ground tracks pass over a region of sloped topography (see Fig. 4 for a digital elevation model and ground paths of the satellite tracks). Satellite tracks that encompass areas with large topographic changes within 50 m of the antenna should be avoided. There is no simple rule for determining when topographic changes are too extreme, though we have found that in general retrievals are the best when the topographic gradient does not exceed 4 %. See also Larson and Nievinski 2013 for a more detailed discussion.

In general, the more the satellite tracks that can be used at a site, the more reliable the final soil moisture time series tend to be. PBO H$_2$O currently requires at least five useful satellite tracks per day at each site. Figure 5 (top) shows a histogram of the number of tracks used at all PBO H$_2$O stations in the fall of 2014. Figure 5 (bottom) shows for one sample station how the number of useful tracks has increased over the years as more satellites have been launched. Temporary decreases in the number of tracks occur due to abnormal noise that might be present in one or more interferograms.

Estimation of each track's a priori reflector height, H_0

The a priori reflector height (H_0) is not known perfectly and has to be estimated from the SNR data. PBO H$_2$O uses the median of H_{eff} data that are free of snow events and significant vegetation for the value of H_0 for each satellite track.

Calculate SNR metrics: ϕ, A, H_{eff}, and A_{LSP}

SNR metrics for each day and each satellite track are calculated using the procedure described in the previous section. Examples of A, A_{LSP}, and H_{eff} time series from one satellite track at a PBO H$_2$O GPS site located in Oklahoma (okl3) are shown in Fig. 6. There are clear seasonal changes in the A and A_{LSP} time series; seasonality is less apparent in the H_{eff} time series. The absolute magnitudes of A and A_{LSP} are unrelated to soil moisture or vegetation changes and are instead primarily related to the transmit power of the satellite and the gain of the antenna. Secondarily, there is a

Fig. 3 Flowchart depicting steps in the algorithm to estimate soil moisture for a GPS site
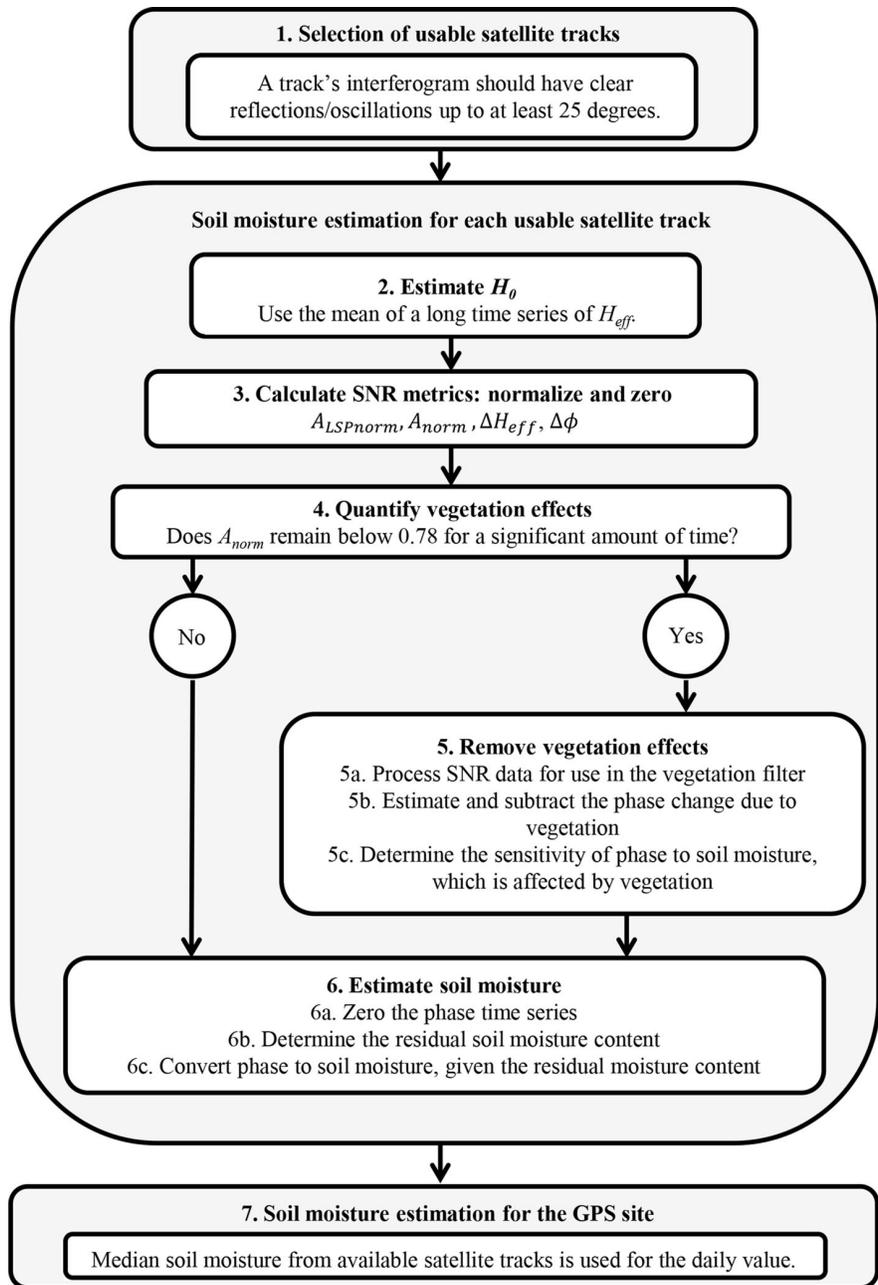
dependence on the temperature of the receiver. Both A and A_{LSP} on each day are normalized as follows:

$$A_{LSPnorm(t)} = \frac{A_{LSP(t)}}{A_{LSP_{20\%}}} \qquad (4)$$

$$A_{norm(t)} = \frac{A_{(t)}}{A_{20\%}} \qquad (5)$$

where the 20 % subscript denotes the mean of the top 20 % of A or A_{LSP} values in the time series for an individual

satellite track; 20 % is a parameter that may be changed, depending on the amount of noise present in the time series or length of the total time series. Any normalized values that are above 1.0 are set to 1.0. Examples are shown in Fig. 6. Both $A_{LSPnorm}$ and A_{norm} for these satellite tracks are very similar to one another, though there are differences during times with the most vegetation.

Changes in the H_{eff} time series are calculated using 3. An example of the ΔH_{eff} time series is shown in Fig. 6. In general, ΔH_{eff} does not change by much more than

Table 1 List of parameters used in the algorithm

Description	Symbol, if used in the text	Value
Response of ϕ to SMC, bare soil	S	1.48 cm^3 cm^{-3} deg^{-1}
Residual moisture content	SMC$_{resdi}$	Obtained from STATSGO dataset
Phase zeroing factor	–	Mean of the top (for filtering) or bottom (for residual determination) 15 % of observed values
Amplitude normalization factor	$\overline{A_{\mathrm{LSP}_{20\%}}}$ or $\overline{A_{20\%}}$	Mean of the top 20 % of observed values
Savitzky–Golay filter: frame length	–	99
Savitzky–Golay filter: polynomial order	–	2
Padding to decrease edge effects: number of repetitions	–	30
Padding to decrease edge effects: value	–	Mean of either the first or last 15 observations

Parameters are ordered such that ones that affect the final soil moisture time series more than others are higher in the list
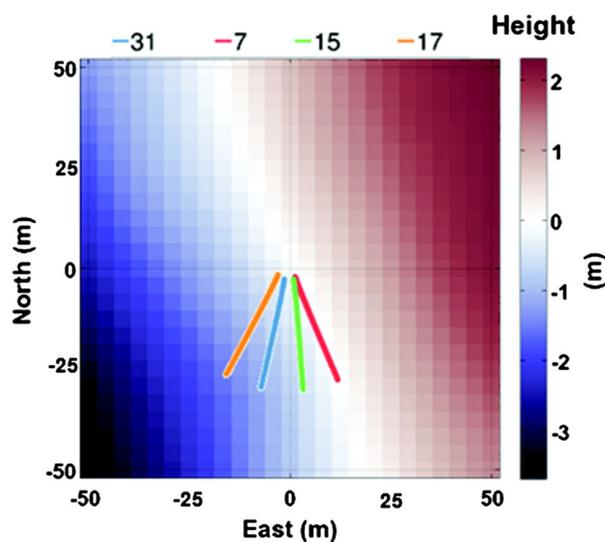

Fig. 4 Digital elevation model (DEM) and the approximate reflection point paths for four satellite tracks at the GPS station okl3 in Oklahoma. The DEM is referenced with respect to (0, 0), which is the location of the antenna. The path traces represent specular reflection points for satellite elevation angles between 5° and 40°. The two western tracks are the "bad" tracks in Fig. 1c; the two eastern tracks are the "good" tracks in Fig. 1c

Fig. 5 Number of satellite tracks varies depending on the GPS site and through time. (*Top*) Histogram of the number of satellite tracks used at PBO H$_2$O stations during the fall of 2014. (*Bottom*) Time series of the number of tracks used at site p038. Step increases occur when additional satellites were launched. Because some satellites pass over a site twice per day, it is possible to increase the number of tracks by four when a new satellite is launched, though oftentimes not all passes are able to be used due to ground-track obstructions

5 cm throughout the year, though significant snow or vegetation water content at the site can produce larger changes.

Quantify vegetation effects

In the absence of field measurements of seasonal vegetation water content change, the A_{norm} time series may be used to estimate whether or not there are significant vegetation effects corrupting the $\Delta\phi$ time series. Chew et al. (2014) showed with an electrodynamic model that an increase in soil moisture produces a decrease in SNR amplitude. The mode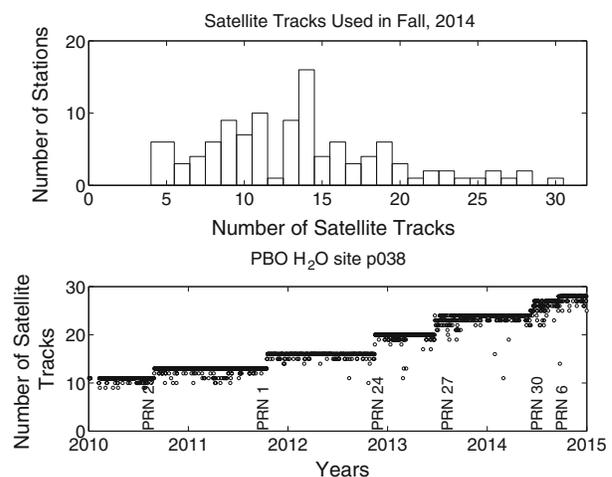l indicates that an increase in moisture content by 0.4 cm^3 cm^{-3} will result in A_{norm} decreasing from 1.0 to 0.78. Since the difference between the residual and saturated moisture contents for a typical soil is about 0.4 cm^3 cm^{-3}, we should not see a decrease in A_{norm} beyond 0.78 in the absence of vegetation growth. This means that a satellite track whose A_{norm} time series stays below 0.78 for an extended period of time is likely to be affected by something other than soil moisture variations, such as changes in vegetation water content. A track whose time series remains above 0.78 may have small vegetation effects, though for these cases it is difficult to determine whether decreases in A_{norm} are from vegetation or from soil moisture changes.

Fig. 6 SNR metrics and field observations from the Oklahoma GPS site okl3. (*Rows 1–4*) SNR metrics for three satellite tracks: PRN 15 (*pink*), 12 (*blue*), and 5 (*green*). The *left column* is raw observed amplitude, LSP amplitude, effective reflector height, and phase. The *right column* contains the same metrics, except they have been either normalized (in the case of amplitude), subtracted from the estimated a priori antenna height, or zeroed with respect to the baseline phase values. (*Bottom row*) Mean and standard deviations of vegetation field measurements made near okl3
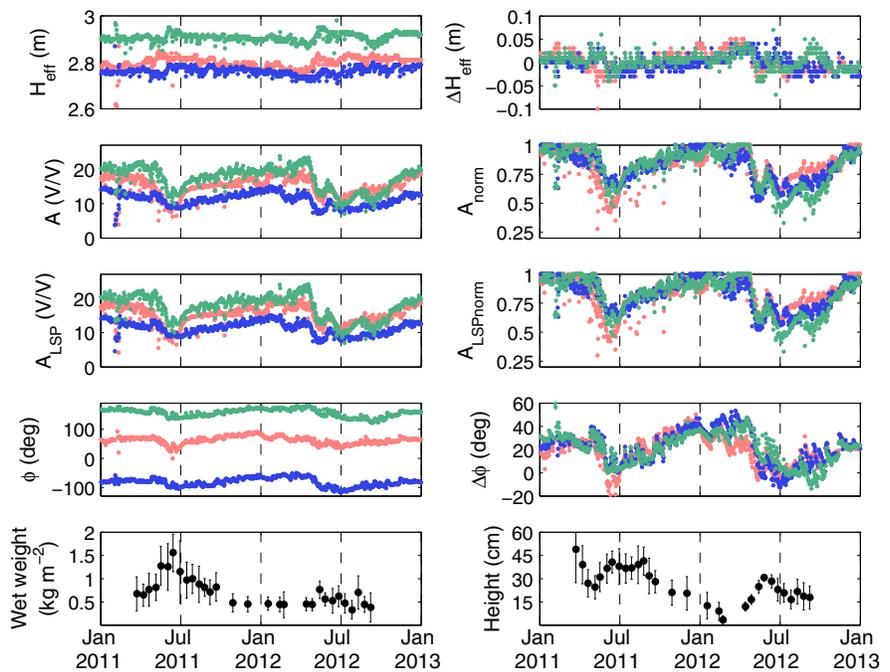
Many PBO H$_2$O sites are located in arid natural environments in the western USA, which typically do not experience significant seasonal changes in vegetation. An example of a site where vegetation effects are negligible is GPS station mfle in Boulder, Colorado. The land cover classification of this site is "desert steppe." We have made in situ measurements of vegetation status at this site (Wan et al. 2014). The highest vegetation water content observed over a several year period is ∼0.2 kg m^{-2}, almost ten times less than observed at the Oklahoma site (Fig. 6). As expected, A_{norm} only rarely drops below 0.78 for some satellite tracks. This indicates that vegetation effects are not large compared to soil moisture variations. Although most of the PBO H$_2$O sites are in arid environments, approximately 20 % of the SNR data appears to be significantly affected by vegetation (Fig. 7), as indicated by the A_{norm} time series dropping below the 0.78 threshold. The rest of the data may have smaller or no effects from vegetation, similar to the case at mfle. In these cases, one may skip to Step 6 in the algorithm below.

Figure 6 shows that there are large seasonal changes in A_{norm} at the GPS station in Oklahoma (okl3) that are attributed to changes in vegetation. The mean and standard deviation of vegetation measurements made at the field site during the growing seasons of 2011 and 2012 are shown in Fig. 6. The corresponding phase time series in Fig. 6 also shows large decreases in phase during times of vegetation growth, though higher frequency variations are attributed to soil moisture fluctuations.
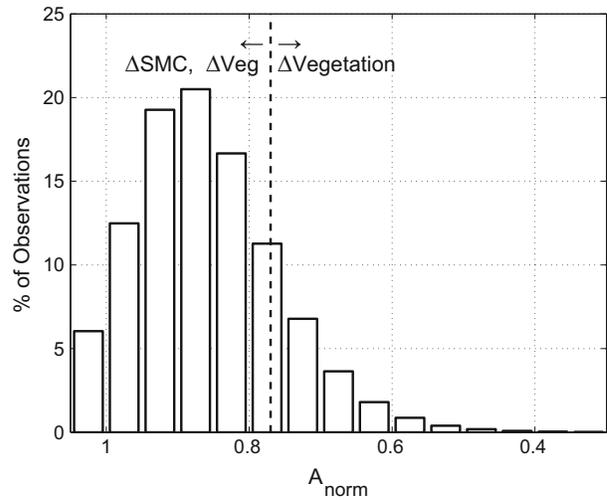
Fig. 7 Observed normalized amplitude data from the PBO H$_2$O sites, shown as the percentage of total observations that fall into each 0.05 normalized amplitude bin. The *dashed line* indicates when data are likely affected by vegetation. Data to the *right* of the line are likely affected. Data to the *left* of the line may be affected to a lesser extent by vegetation, or by soil moisture

If the GPS site does have a significant vegetation effect as determined by the A_{norm} time series, then one should proceed to the section "Algorithm for removing vegetation effects" before estimating soil moisture in Step 6. If there is a significant vegetation effect but the effect is of a short-enough duration, then the phase data when A_{norm} is below

0.78 may be removed, without using the vegetation filter. However, simply removing the phase data without implementing a vegetation correction will result in error for periods surrounding the affected interval, which are still affected by vegetation growth. It is up to the user whether or not greater error is acceptable for the trade-off in simplicity of data analysis.

Soil moisture estimation

At GPS sites with limited seasonal change in vegetation, as indicated by the A_{norm} time series, soil moisture estimation is relatively straightforward. We use data from the site mfle, which was described above, as an example.

Zero the phase time series

For each year of data, the phase time series for each satellite track is zeroed, i.e., the lowest values are set to near zero. An example of this zeroing is shown in Fig. 6. Sites in the PBO H$_2$O network are zeroed by first calculating the mean of the lowest 15 % of observed phase data for each year and each track. The mean is then subtracted from the phase time series; 15 % is a parameter that may be decreased or increased, depending on the amount of noise present in the data.

Phase time series are zeroed yearly; this assumes that soil moisture at the site reaches its residual value at some point during the year. This assumption is generally valid for the PBO H$_2$O sites. Yearly zeroing is done so that any period of anomalously low phase data, either due to non-removal of vegetation effects or random noise, will not affect the entire time series.

Determine the residual soil moisture content

Because temporal phase changes are relative and the minimum moisture content of a soil is never zero, the baseline phase value needs to be associated with the soil's residual moisture content. Residual moisture content, SMC_{resid}, is correlated with soil texture, and these data can be found in publicly available data sets such as the US Geological Survey's STATSGO data set (Schwarz and Alexander 1995). However, optimally the residual moisture content of the soil at the site would be measured through the use of gravimetry.

Estimate the soil moisture time series for each satellite track

Modeled relationships between phase and soil moisture are used to compute the estimated soil moisture time series,

$$SMC_t = S\Delta\phi_t + SMC_{resid} \qquad (6)$$

where S is the expected slope between phase and soil moisture. For time series with no significant vegetation affect, $S = 1.48$ cm^3 cm^{-3} deg^{-1} (Chew et al. 2014).

Estimate the soil moisture time series for the GPS site as a whole

To estimate site-averaged soil moisture on the daily time-scale, the median soil moisture value of all tracks for each day is used. Uncertainties are computed as the standard deviation of the soil moisture estimates from all satellite tracks. For sites with a large number of satellite tracks, the mean soil moisture value of all tracks could be used.

Figure 9a shows in situ soil moisture data from Campbell Scientific 616 probes that have been buried at 2.5 cm depth within 250 m of the desert steppe site (mfle). Also shown are soil moisture estimates using GPS-IR, using data from the GPS antenna. There is a good agreement between in situ measurements and GPS soil moisture estimations at this site, even without a vegetation filtering algorithm, as expected given the A_{norm} time series.

Algorithm for removing vegetation effects

In this section, we describe how to correct the phase time series to remove the effects from variations in vegetation state around a GPS site. The vegetation adjustment algorithm corresponds to the box "5. Remove vegetation effects" in Fig. 3. We use data from the Oklahoma GPS site okl3 to demonstrate the importance of adjusting phase time series for vegetation. Figure 6 (bottom) shows that the Oklahoma GPS station is surrounded by vegetation that varies in water content and height, both seasonally and from year to year. Figure 10a shows how following the simple soil moisture estimation procedure described above would lead to poor agreement with in situ data. Because an increase in vegetation causes a decrease in phase, and the residual moisture content is set to the lowest observed phase value, vegetation growth causes the resulting soil moisture estimations to be too high in winter. Clearly, an adjustment for vegetation is needed at this site. This could be determined solely from the A_{norm} time series, which shows values as low as 0.5 during the growing season.

At sites affected by vegetation, one can use the following vegetation filtering algorithm to mitigate the effect of vegetation. This filter was developed using the soil vegetation model described and validated in Chew et al. (2015). Over 15,000 modeled SNR interferograms and associated SNR metrics were simulated using random combinations of the vegetation parameters required by the

model and an underlying soil moisture content of
0.15 cm³ cm⁻³ as input to the soil vegetation model.
Changes in SNR metrics with respect to their bare soil
values were calculated and saved in a database. The data-
base can be queried to estimate the change in phase corre-
sponding to any observed combination of A_{norm}, $A_{LSPnorm}$,
and ΔH_{eff} time series. The same random combinations of
vegetation parameters were then used with underlying soil
moisture contents of both 0.05 and 0.45 cm³ cm⁻³, and the
soil vegetation model was run again. Another database was
created from these simulations, which shows how SNR
metrics are expected to respond to changes in underlying
soil moisture given a particular vegetation state.

Now, we describe how to use the model database to
remove vegetation effects from observed SNR phase data.
This is done on a track-by-track basis, as the input data
(SNR metrics) are all track specific.

Process data for use in the vegetation filter

The vegetation filter assumes that the phase time series has
been zeroed such that data most affected by vegetation go
below zero. In order to conform to this convention, the
phase time series should be zeroed with respect to the
median of the highest values in the time series. We have
found the median of the highest 15 % of observed phase
data works well.

A low-pass filter is used to remove high-frequency noise
associated with soil moisture fluctuations from the A_{norm},
$A_{LSPnorm}$, and ΔH_{eff} time series. The ends of the A_{norm},
$A_{LSPnorm}$, and ΔH_{eff} time series are first padded to decrease
edge effects. Thirty repetitions of the mean of the first and
last 15 days of A_{norm}, $A_{LSPnorm}$, and ΔH_{eff} values are ap-
pended to the beginning and end of each time series.

A_{norm}, $A_{LSPnorm}$, and ΔH_{eff} are smoothed using a Sav-
itzky–Golay filter, a least-squares smoothing method, with
a polynomial order of two and frame size of 99 (Table 1).
We have found the S–G filter to be more successful than a
moving average or median filter when the change in
vegetation is either rapid and significant or short lived. This
is because the S–G filter is more successful at retaining the
maximum or minimum extents of a time series, though this
comes at the expense of potentially also retaining noise
unrelated to vegetation changes. For many cases, however,
a moving average filter is sufficient, though care must be
taken to not over-smooth the time series.

**Estimate and subtract the phase change
from vegetation from the observed time series**

A linear nearest neighbor search algorithm is used to find
the estimated phase caused by vegetation fluctuations,
given observed A_{norm}, $A_{LSPnorm}$, and ΔH_{eff} and the modeled

database described above (Fig. 8, top). Thus, one can es-
timate the effects of vegetation on phase using time series
of the other three SNR metrics. No field observations of
vegetation amount are needed. Estimated phase changes
from the database are smoothed through time using the
same Savitzky–Golay or moving average filter, and the
padded ends are removed.

The expected phase changes due to vegetation are sub-
tracted from the observed, unsmoothed phase time series,
resulting in phase time series that indicate soil moisture
variations only.

$$\phi_{SMC,t} = \Delta\phi_t - \phi_{veg,t} \tag{7}$$

where $\phi_{smc,t}$ is the expected phase change due to soil
moisture at time t, $\Delta\phi_t$ is the original observed time series at
time t, and ϕ_{veg} is the predicted phase change due to
vegetation at time t. This relationship is a simplification, as
it is based on the assumption that the total phase change is a
linear combination of the phase change due to soil moisture
and the phase change due to vegetation. Chew et al. (2015)
showed that it is only an approximation, though we have
found it does not significantly affect our results.

**Determine the sensitivity change between phase
and soil moisture due to vegetation**

Expected sensitivity changes between phase and soil mois-
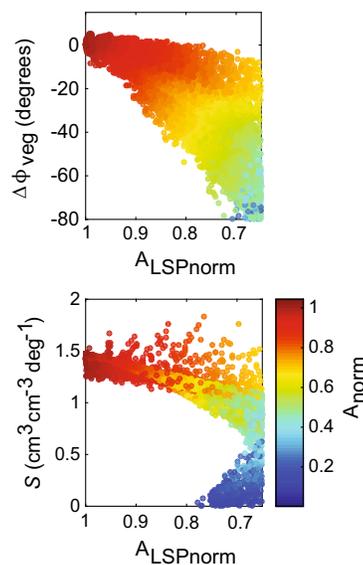ture are estimated, using the same procedure and linear

Fig. 8 Depiction of the model simulations used in the vegetation
filter, though here changes in effective reflector height are not shown.
(*Top*) Modeled relationships between $A_{LSPnorm}$ and $\Delta\phi$, due to
changes in vegetation only. Points are *colored* by changes in A_{norm}.
(*Bottom*) Modeled relationships between $A_{LSPnorm}$ and the sensitivity
of $\Delta\phi$ to soil moisture changes (S), given only changes in the
vegetation canopy. Points are colored by changes in A_{norm}

Fig. 9 *Blue lines* are average soil moisture estimates from in situ probes installed at 2.5 cm depth within 250 m the GPS antenna mfle in Colorado. *Yellow points* are the mean soil moisture estimates resulting from theta probe surveys. **a** *Black points* and *gray error bars* are the median and standard deviation of soil moisture estimates using GPS-IR, if no vegetation corrections are made. **b** *Magenta points* and *gray error bars* are the median and standard deviation of soil moisture estimates using GPS-IR, if the vegetation filtering algorithm is used to correct the phase data for vegetation effects
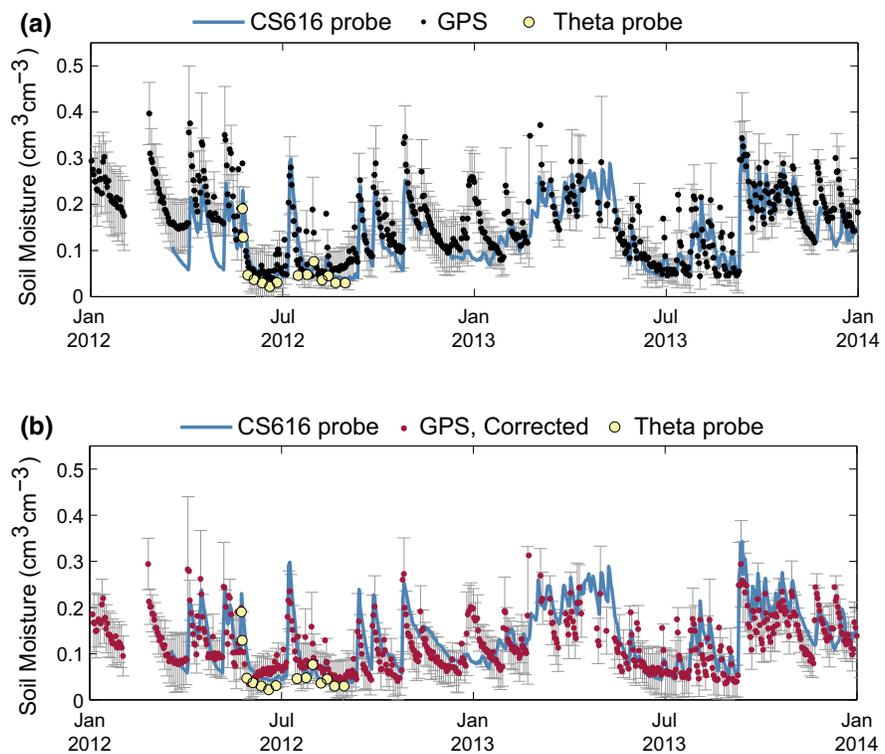
search algorithm described above (Fig. 8, bottom). The sensitivity of phase to soil moisture will change depending on the extent of overlying vegetation (Chew et al. 2015), and thus, the slope of the relationship between phase and soil moisture, S, will change throughout the year. As shown in Fig. 8 (bottom), changes in S are large and cannot be ignored. Equation 6 used in Step 6 thus does not have a constant value of S for data processed with the vegetation filter, but rather a value that varies depending upon the vegetation state.

Once the phase data are processed with the vegetation filter, one can calculate a soil moisture time series for each track (Step 6, Fig. 3). Residual moisture content is selected in the same way as for a bare soil case (Step 6b). When converting phase to soil moisture, S is a function of time, as described in Step 5c. The soil moisture time series for a phase time series filtered for vegetation will thus be a variant of (6):

$$\text{SMC}_t = S_t \Delta\phi_{\text{SMC,t}} + \text{SMC}_{\text{resid}} \tag{8}$$

As described above, the median of the track soil moisture time series is taken to be representative of the soil moisture of the site as a whole. There are additional uncertainties due to assumptions made in the vegetation filtering algorithm, which are not yet incorporated into final soil moisture estimates.

If the vegetation filter is used on the observations from mfle, there are only small changes in the resulting soil moisture time series (Fig. 9b) relative to the uncorrected

time series (Fig. 9a). Although soil moisture estimates using the filter are improved in early 2012, the rest of the time series is largely unchanged.

Conversely, the vegetation filter makes a large difference in soil moisture estimations at the Oklahoma GPS site, okl3. The vegetation-corrected soil moisture time series for okl3 is shown in Fig. 10b, using the mean soil moisture value for all tracks. Agreement between the GPS-derived soil moisture estimations and those from in situ data has been much improved over the non-vegetation-filtered estimations (Fig. 10a). In particular, the large wintertime deviations described in the previous section have been significantly reduced. A more complete validation dataset is presented elsewhere.

Future work

Extending the soil moisture algorithm presented here to include stations located in areas with greater topography or for sites with other antenna types remains a subject of active research. In order to use different antenna types, their gain and phase patterns must be known to simulate their specific relationships between soil moisture and SNR data. Including data from other GNSS constellations in our soil moisture estimations could enhance the spatial and temporal coverage of tracks at some sites, though issues such as the non-repeating ground tracks of other GNSS satellites would

GPS Solut

Fig. 10 *Blue lines* are average soil moisture estimates from in situ probes installed at 2.5 cm depth within 20 m the GPS antenna okl3 in Oklahoma. *Yellow points* are the mean soil moisture estimates resulting from theta probe surveys. **a** *Black points* and *gray error bars* are the median and standard deviation of soil moisture estimates using GPS-IR, if no vegetation corrections are made. **b** *Magenta points* and *gray error bars* are the median and standard deviation of soil moisture estimates using GPS-IR, if the vegetation filtering algorithm is used to correct the phase data for vegetation effects
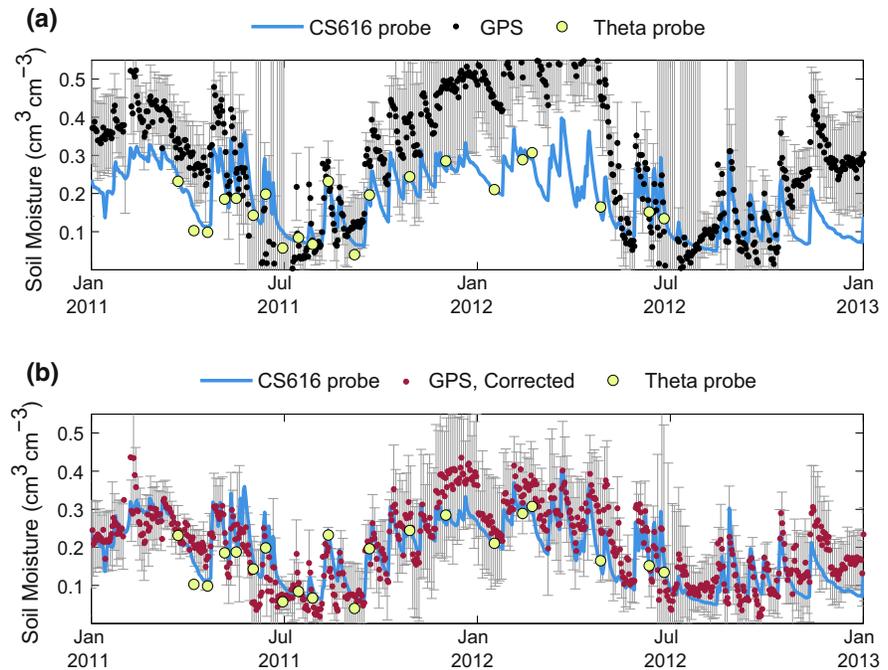
first need to be modeled. Errors due to assumptions made in the vegetation filtering algorithm should also be incorporated into final soil moisture estimations.

Conclusions

An algorithm is presented for estimating near-surface soil moisture fluctuations for the immediate area surrounding a geodetic-quality GPS antenna. This algorithm is being used operationally at ~ 120 GPS sites in the western USA (Larson and Small 2013). The algorithm mitigates the effect of vegetation in estimated soil moisture time series by using a soil vegetation model to predict relationships between SNR metrics due to changes in vegetation state. SNR metrics A_{norm}, $A_{LSPnorm}$, and ΔH_{eff} are used to describe expected changes in ϕ due to vegetation, and these changes are subtracted from the ϕ time series to retain only the effect of soil moisture fluctuations on ϕ, which are then converted to soil moisture estimates.

Two example soil moisture time series, derived from GPS SNR phase data, were presented along with in situ soil moisture data. Results from this study show that GPS-IR is a viable technique to be used at GPS stations situated in environments with a relatively high amount of natural vegetation if appropriate steps are taken to model the vegetation effect.

References

Baroni G, Ortuani B, Facchi A, Gandolfi C (2013) The role of vegetation and soil properties on the spatio-temporal variability of the surface soil moisture in a maize-cropped field. J Hydrol 489:148–159. doi:10.1016/j.jhydrol.2013.03.007

Brocca L, Morbidelli R, Melone F, Moramarco T (2007) Soil moisture spatial variability in experimental areas of central Italy. J Hydrol 333:356–373. doi:10.1016/j.jhydrol.2006.09.004

Cardellach E, Ruffini G, Pino D et al (2003) Mediterranean balloon experiment: ocean wind speed sensing from the stratosphere, using GPS reflections. Remote Sens Environ 88:351–362. doi:10.1016/S0034-4257(03)00176-7

Cardellach E, Fabra F, Rius A et al (2012) Characterization of dry-snow sub-structure using GNSS reflected signals. Remote Sens Environ 124:122–134. doi:10.1016/j.rse.2012.05.012

Chew CC, Small EE, Larson KM, Zavorotny VU (2014) Effects of near-surface soil moisture on GPS SNR data: development of a retrieval algorithm for soil moisture. IEEE Trans Geosci Remote Sens 52:537–543

Chew CC, Small EE, Larson KM, Zavorotny VU (2015) Vegetation sensing using GPS-interferometric reflectometry: theoretical effects of canopy parameters on signal-to-noise ratio data. IEEE Trans Geosci Remote Sens 53:2755–2764

Dunn MJ (2010) Global Positioning Systems Wing (GPSW) systems engineering & integration, interface specification IS-GPS-200, pp 1–226. Retrieved from http://www.gps.gov/technical/icwg/IS-GPS-200E.pdf

Entekhabi BD, Njoku EG, Neill PEO et al (2010) The soil moisture active passive (SMAP) mission. Proc IEEE 98:704–716

Famiglietti JS, Rudnicki JW, Rodell M (1998) Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas. J Hydrol 210:259–281. doi:10.1016/S0022-1694(98)00187-5

Ferrazzoli P, Guerriero L, Pierdicca N, Rahmoune R (2011) Forest biomass monitoring with GNSS-R: theoretical simulations. Adv Sp Res 47:1823–1832. doi:10.1016/j.asr.2010.04.025

Fuks IM, Voronovich AG (2000) Wave diffraction by rough interfaces in an arbitrary plane-layered medium. Waves Random Media 10:37–41

Gomez-Plaza A, Martinez-Mena M, Albaladejo J, Castillo VM (2001) Factors regulating spatial distribution of soil water content in small semiarid catchments. J Hydrol 253:211–226

Hall CD, Cordey RA (1988) Multistatic scatterometry. In: Proceedings of IGARSS'88 symposium, Edinburgh, pp 561–562

Hallikainen MT, Ulaby FT, Dobson MC et al (1985) Microwave dielectric behavior of wet soil-part 1: empirical models and experimental observations. IEEE Trans Geosci Remote Sens GE 23:25–34

Jackson TJ, Le Vine DE (1996) Mapping surface soil moisture using an aircraft-based passive microwave instrument: algorithm and example. J Hydrol 184:85–99

Jackson TJ, Le Vine DM, Swifi CT et al (1995) Large area mapping of soil moisture using the ESTAR passive microwave radiometer in washita'92. Remote Sens Environ 53:27–37

Katzberg SJ, Garrison JL (1996) Utilizing GPS to determine ionospheric delay over the ocean. NASA Technical Memorandum 4750

Kerr YH, Waldteufel P, Wigneron J et al (2001) Soil moisture retrieval from space: the soil moisture and ocean salinity (SMOS) mission. IEEE Trans Geosci Remote Sens 39:1729–1735

Komjathy A, Maslanik J, Zavorotny VU, et al. (2000a) Sea ice remote sensing using surface reflected GPS signals. In: Geoscience and remote sensing symposium 2000, Honolulu, pp 2855–2857

Komjathy A, Zavorotny VU, Axelrad P et al (2000b) GPS signal scattering from sea surface: wind speed retrieval using experimental data and theoretical model. Remote Sens Environ 73:162–174

Larson KM, Nievinski FG (2013) GPS snow sensing: results from the Earthscope Plate Boundary Observatory. GPS Solut 17:41–52. doi:10.1007/s10291-012-0259-7

Larson K, Small E (2013) Using GPS to study the terrestrial water cycle. EOS Trans Am Geophys Union 94:505–512. doi:10.1029/2010GL042951.Water

Larson KM, Small EE, Gutmann E et al (2008a) Using GPS multipath to measure soil moisture fluctuations: initial results. GPS Solut 12:173–177. doi:10.1007/s10291-007-0076-6

Larson KM, Small EE, Gutmann ED et al (2008b) Use of GPS receivers as a soil moisture network for water cycle studies. Geophys Res Lett 35:1–5. doi:10.1029/2008GL036013

Larson KM, Braun JJ, Small EE et al (2010) GPS multipath and its relation to near-surface soil moisture content. IEEE J Sel Top Appl Earth Obs Remote Sens 3:91–99

Lowe ST, Zuffada C, Chao Y et al (2002) 5-cm-Precision aircraft ocean altimetry using GPS reflections. Geophys Res Lett 29:10–13

Martin-Neira M (1993) A passive reflectometry and interferometry system (PARIS): application to ocean altimetry. ESA J 17:331–355

Masters DS (2004) Surface remote sensing applications of GNSS bistatic radar: soil moisture and aircraft altimetry. PhD dissertation, University of Colorado, Boulder

McCreight JL, Small EE, Larson KM (2014) Snow depth, density, and SWE estimates derived from GPS reflection data: validation in the western U.S. Water Resour Res 50:6892–6909. doi:10.1002/2014WR015561.Received

Mladenova I, Lakshmi V, Jackson TJ et al (2011) Validation of AMSR-E soil moisture using L-band airborne radiometer data from National Airborne Field Experiment 2006. Remote Sens Environ 115:2096–2103. doi:10.1016/j.rse.2011.04.011

Njoku EG, Entekhabi D (1996) Passive microwave remote sensing of soil moisture. J Hydrol 184:101–129. doi:10.1016/0022-1694(95)02970-2

Rajkai K, Ryden BE (1992) Measuring areal soil moisture distribution with the TDR method. Geoderma 52:73–85

Robock A, Vinnikov KY, Srinivasan G et al (2000) The global soil moisture data bank. Bull Am Meteorol Soc 81:1281–1299

Rodriguez-Alvarez N, Aguasca A, Valencia E, et al. (2011a) Snow monitoring using GNSS-R techniques. In: Geoscience and remote sensing symposium, pp 4375–4378

Rodriguez-Alvarez N, Bosch-Lluis X, Camps A et al (2011b) Review of crop growth and soil moisture monitoring from a ground-based instrument implementing the interference pattern GNSS-R technique. Radio Sci. doi:10.1029/2011RS004680

Rodriguez-Alvarez N, Bosch-lluis X, Camps A et al (2012) Vegetation water content estimation using GNSS measurements. IEEE Geosci Remote Sens Lett 9:282–286

Schwank M, Mätzler C, Member S et al (2005) L-Band radiometer measurements of soil water under growing clover grass. IEEE Trans Geosci Remote Sens 43:2225–2237

Schwarz GE, Alexander RB (1995) Soils data for the conterminous United States derived from the NRCS State Soil Geographic (STATSGO) data base. [Original title: State Soil Geographic (STATSGO) data base for the conterminous United States.]

Tsang L, Kong JA, Shin RT (1985) Theory of microwave remote sensing, 1st edn. Wiley-Interscience, London

Ulaby FT, Moore RK, Fung AK (1982) Microwave remote sensing active and passive-volume II: radar remote sensing and surface scattering and emission theory. Addison-Wesley, Reading

Wan W, Larson KM, Small EE et al (2014) Using geodetic GPS receivers to measure vegetation water content. GPS Solut. doi:10.1007/s10291-014-0383-7

Wang JR, Engman ET, Shiue JC, Rusek M (1986) Observations of microwave dependence on soil moisture, surface roughness, and vegetation covers. IEEE Trans Geosci Remote Sens GE-24:510–516

Zavorotny VU, Voronovich AG (2000) Scattering of GPS signals from the ocean with wind remote sensing application. IEEE Trans Geosci Remote Sens 38:951–964

Zavorotny VU, Larson KM, Braun JJ et al (2010) A physical model for GPS multipath caused by land reflections: toward bare soil moisture retrievals. IEEE J Sel Top Appl Earth Obs Remote Sens 3:1–11

Clara Chew received her Ph.D. in Geological Sciences from the University of Colorado at Boulder in 2015. She received her B.A. in Environmental Studies from Dartmouth College in 2009.

Eric Small is a Professor of Geological Sciences at the University of Colorado at Boulder. His current research focuses on land surface hydrology.

Kristine Larson is a Professor of Aerospace Engineering Sciences at the University of Colorado at Boulder. Her research focuses on GPS reflections.