AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2017WR021521

Key Points:
- The seasonal cycle and interannual variability of water loading in snow-dominated mountains can be estimated from GPS displacement data
- Snow water equivalent estimates are biased low due to mass leakage, which can be reduced by linear scaling or addition of GPS stations
- Accurate estimation of snow water equivalent requires independent estimation of other terrestrial water storage terms

Supporting Information:
- Supporting Information S1

Correspondence to:
T. L. Enzminger,
thomas.enzminger@colorado.edu

Citation:
Enzminger, T. L., Small, E. E., & Borsa, A. A. (2018). Accuracy of snow water equivalent estimated from GPS vertical displacements: A synthetic loading case study for western U.S. mountains. *Water Resources Research*, *54*, 581–599. https://doi.org/10.1002/2017WR021521

Received 13 JUL 2017
Accepted 8 JAN 2018
Accepted article online 18 JAN 2018
Published online 30 JAN 2018

Accuracy of Snow Water Equivalent Estimated From GPS Vertical Displacements: A Synthetic Loading Case Study for Western U.S. Mountains

Thomas L. Enzminger[1], Eric E. Small[1], and Adrian A. Borsa[2]

[1]Department of Geological Sciences, University of Colorado Boulder, Boulder, CO, USA, [2]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA

Abstract GPS monitoring of solid Earth deformation due to surface loading is an independent approach for estimating seasonal changes in terrestrial water storage (TWS). In western United States (WUSA) mountain ranges, snow water equivalent (SWE) is the dominant component of TWS and an essential water resource. While several studies have estimated SWE from GPS-measured vertical displacements, the error associated with this method remains poorly constrained. We examine the accuracy of SWE estimated from synthetic displacements at 1,395 continuous GPS station locations in the WUSA. Displacement at each station is calculated from the predicted elastic response to variations in SWE from SNODAS and soil moisture from the NLDAS-2 Noah model. We invert synthetic displacements for TWS, showing that both seasonal accumulation and melt as well as year-to-year fluctuations in peak SWE can be estimated from data recorded by the existing GPS network. Because we impose a smoothness constraint in the inversion, recovered TWS exhibits mass leakage from mountain ranges to surrounding areas. This leakage bias is removed via linear rescaling in which the magnitude of the gain factor depends on station distribution and TWS anomaly patterns. The synthetic GPS-derived estimates reproduce approximately half of the spatial variability (unbiased root mean square error ~50%) of TWS loading within mountain ranges, a considerable improvement over GRACE. The inclusion of additional simulated GPS stations improves representation of spatial variations. GPS data can be used to estimate mountain-range-scale SWE, but effects of soil moisture and other TWS components must first be subtracted from the GPS-derived load estimates.

1. Introduction

Winter snowpack in mountain headwater regions is a critical water resource in the western United States (WUSA), accounting for as much as 70–80% of regional annual runoff and dominating the hydrologic regimes of most watersheds (Doesken & Judson, 1996). Basin-wide snow water equivalent (SWE)—the amount of water contained in snowpack—is a function of snow extent, depth, and density. SWE is one of the most useful snowpack metrics in water budget forecasting as it improves the skill of streamflow models and facilitates accurate predictions of water yield throughout the year (Koster et al., 2010). It is therefore critical to improve the accuracy and resolution of SWE estimates in snow-dominated watersheds.

Mountain SWE is currently estimated using in situ measurements, remote sensing products, and models. In situ measurements, such as snow telemetry (SNOTEL) and snow course data, are the most accurate at very fine spatial scales, but cannot accurately represent SWE over large areas, as snow accumulation is highly spatially heterogeneous (Serreze et al., 1999). Snow cover and depth can be measured via satellite imaging or airborne scanning lidar; however, SWE cannot be calculated from these data without distributed estimates of snow density (Molotch, 2009; Painter et al., 2016). Modeling snow density introduces uncertainty into estimates of SWE derived from remotely sensed snow depth data (Raleigh & Small, 2017). The GRACE (Gravity Recovery and Climate Experiment) satellites directly measure changes in terrestrial water storage (TWS)—defined as the total water mass stored on or near the land surface as snowpack, soil moisture, surface water, groundwater, and in biomass (Wahr et al., 2004). GRACE data can be used to estimate changes in snow mass if other components of TWS are known. However, uncertainty in the estimation of the numerous other components of TWS, as well as the coarse spatial resolution and 2–6 month latency of the GRACE data product make it unsuitable for real-time SWE monitoring at the mountain-range scale (Famiglietti &

Rodell, 2013). Even newly developed GRACE mascon solutions at 1° × 1° resolution (e.g., Save et al., 2016) are too coarse for mountain SWE monitoring, as the width of most mountain ranges is considerably less than 100 km. Model-based estimates of SWE vary significantly from model to model and must be validated against other SWE estimates, with which they do not always correlate well (Clow et al., 2012; Rutter et al., 2009).

GPS (Global Positioning System) measurements of the solid Earth elastic response to loading have shown promise as a means of quantitatively estimating changes in TWS. Bevis et al. (2005) estimated seasonal fluctuations in surface water mass over the central Amazon basin from a single GPS station. Steckler et al. (2010) used stream gauge, GPS, and GRACE data to estimate both solid Earth elastic properties and TWS change due to monsoonal flooding in Bangladesh. Borsa et al. (2014) estimated total water mass lost during the recent California drought using trends in median displacement from over 700 continuous GPS stations in the WUSA.

The solid Earth undergoes instantaneous, reversible elastic deformation in response to changes in surface or near-surface loading (Farrell, 1972). This deformation is most strongly manifested as vertical displacement of Earth's surface, which can be measured at the millimeter scale by GPS stations anchored in bedrock or deep soil. The land surface depresses in response to hydrologic loading from precipitation and rebounds, when unloading occurs through evaporation or runoff (Dong et al., 2002). The solid Earth response to loading is largest near a load, but extends beyond the boundaries of the load itself (Na & Baek, 2011). Therefore, displacement at each GPS station is an integrated response to loading at all distances. Data from a sufficiently dense network of continuous GPS stations can thus be used to produce gridded estimates of changes in hydrologic load. These estimates represent an improvement on TWS interpolated from in situ point observations, which can suffer from the spatial aliasing of short-wavelength TWS variability.

Earth surface displacements due to nonhydrological forces such as tectonic uplift and glacial isostatic adjustment also appear in GPS station positions, but these are primarily secular signals whose effects can be removed by detrending (e.g., Borsa et al., 2014). In the absence of volcanic or seismic transients, changes in hydrologic load have been shown to be the primary drivers of GPS vertical displacement on seasonal to interannual timescales (Dong et al., 2002).

It has been proposed that seasonal changes in TWS, and therefore in GPS-measured vertical displacement, in snow-dominated hydrologic regimes primarily reflect changes in SWE (Ouellette et al., 2013). Winter snow loading in mountainous regions has been shown to cause downward vertical displacement of the land surface of up to ∼12 mm—a magnitude that is readily resolved by GPS stations (Argus et al., 2014). Ouellette et al. (2013) used vertical position to estimate SWE change. However, their one-dimensional model provides a coarse approximation of true loading and does not consider changes in loading outside the immediate vicinity of the GPS station. Argus et al. (2014) and Fu et al. (2015) estimated seasonal load changes in the snow-dominated Sierra Nevada in California and Cascade Range in the Pacific Northwest, respectively. A large fraction of the multiannual TWS loss found by Borsa et al. (2014) is likely attributable to drastic decreases in mountain SWE.

The magnitude of error associated with GPS-based TWS estimates remains poorly constrained. There are no ground truth TWS metrics of comparable spatial scale and coverage (Rodell & Famiglietti, 2002). Monitoring SWE using GPS presents the added challenge of isolating the change in SWE from other components of TWS affecting the seasonal hydrologic loading signal.

Previous studies addressing SWE estimation from GPS data have estimated error by comparison to in situ measurements, remote sensing products, and hydrologic model estimates. Comparing GPS-inverted TWS loads to SNOTEL station data is problematic (Fu et al., 2015; Ouellette et al., 2013). There is no way to account for spatial variability in SWE beyond the limited SNOTEL capture area, and assumptions must be made about the fraction of TWS change attributable to SWE. Ouellette et al. (2013) assumed 100% of TWS change came from SWE, an assumption unlikely to be valid anywhere with significant soil or surface water. The accuracy of comparisons between GRACE and GPS estimates (e.g., Argus et al., 2014; Fu et al., 2015) at the mountain-range scale is not well known because the GRACE solution is susceptible to errors in grid cells spanning the margins of mountain ranges, which encompass strong gradients in TWS loads. Moreover, as both methods estimate TWS change, SWE must be isolated from the TWS signal in both cases. Comparison to a hydrologic model (e.g., Argus et al., 2014; Fu et al., 2015) entails accounting for unknown or poorly

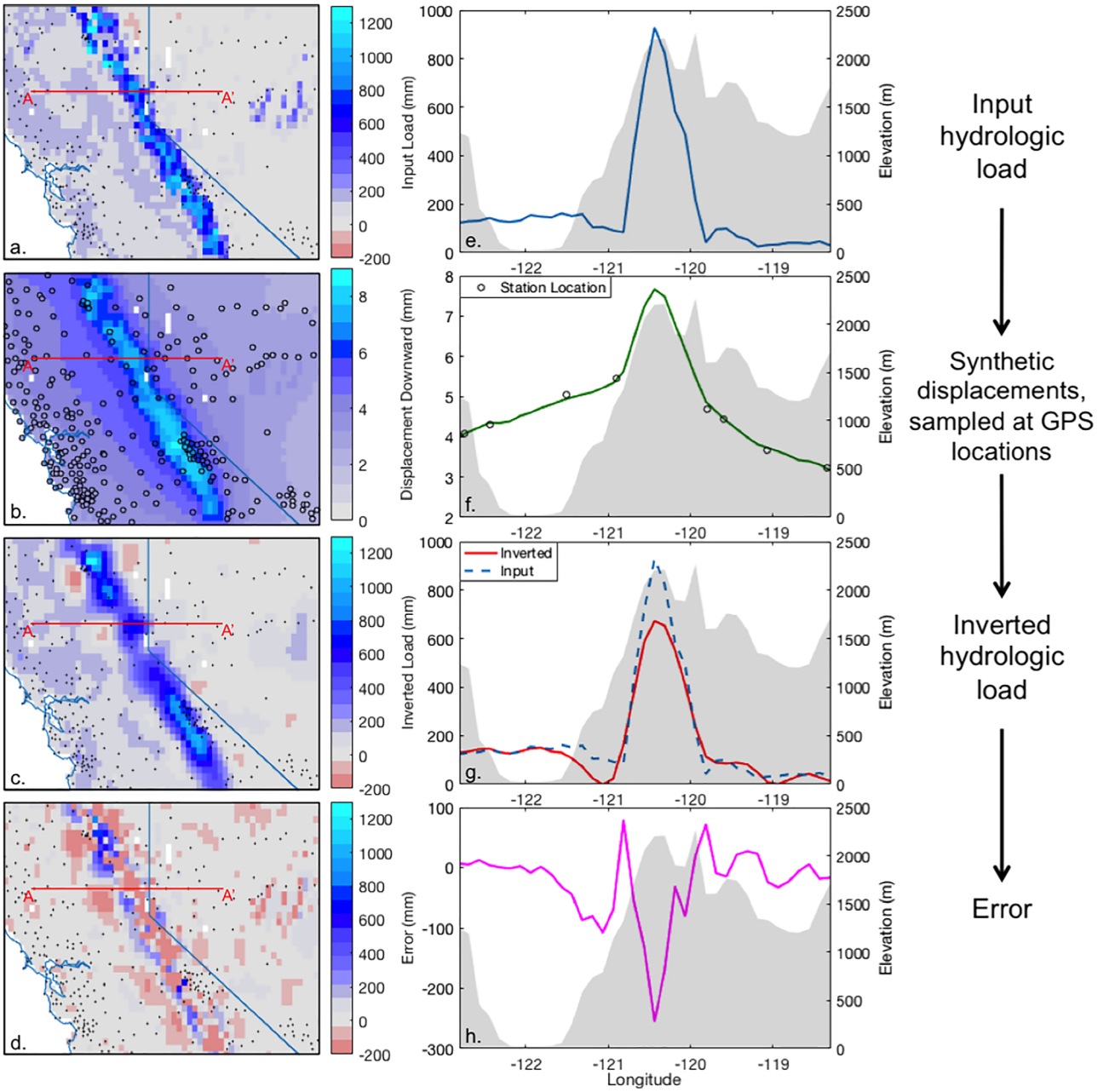

Figure 1. Schematic of experimental design. (a) Gridded anomalies from an input hydrologic load for April 2010 drive (b) synthetic vertical displacements, sampled at real GPS station locations (open circles). (c) Inverted hydrologic load anomalies are recovered from the GPS station displacements. (d) Error is quantified by differencing input and inverted loads. Transects along A-A′ give profile views of (e) input load, (f) synthetic displacement, (g) inverted load, and (h) error across the Sierra Nevada for latitude 39.44°N, with surface elevation shown for reference (light grey).

constrained errors within the model. Hydrologic models may also exclude components of TWS, thereby potentially misattributing change in one component of TWS (e.g., groundwater) to another component (e.g., SWE).

Here, we expand upon the approach of Fu et al. (2015) to quantify the accuracy of mountain SWE estimated from GPS vertical position data. We use a realistic hydrologic (TWS) load distribution to both generate a synthetic earth surface displacement field and to validate the inverted TWS result. A similar approach was taken by Landerer and Swenson (2012) to quantify errors in GRACE. We first calculate synthetic displacements at 1,395 real GPS station locations from a series of gridded monthly hydrologic load anomalies (Figures 1 and 2).

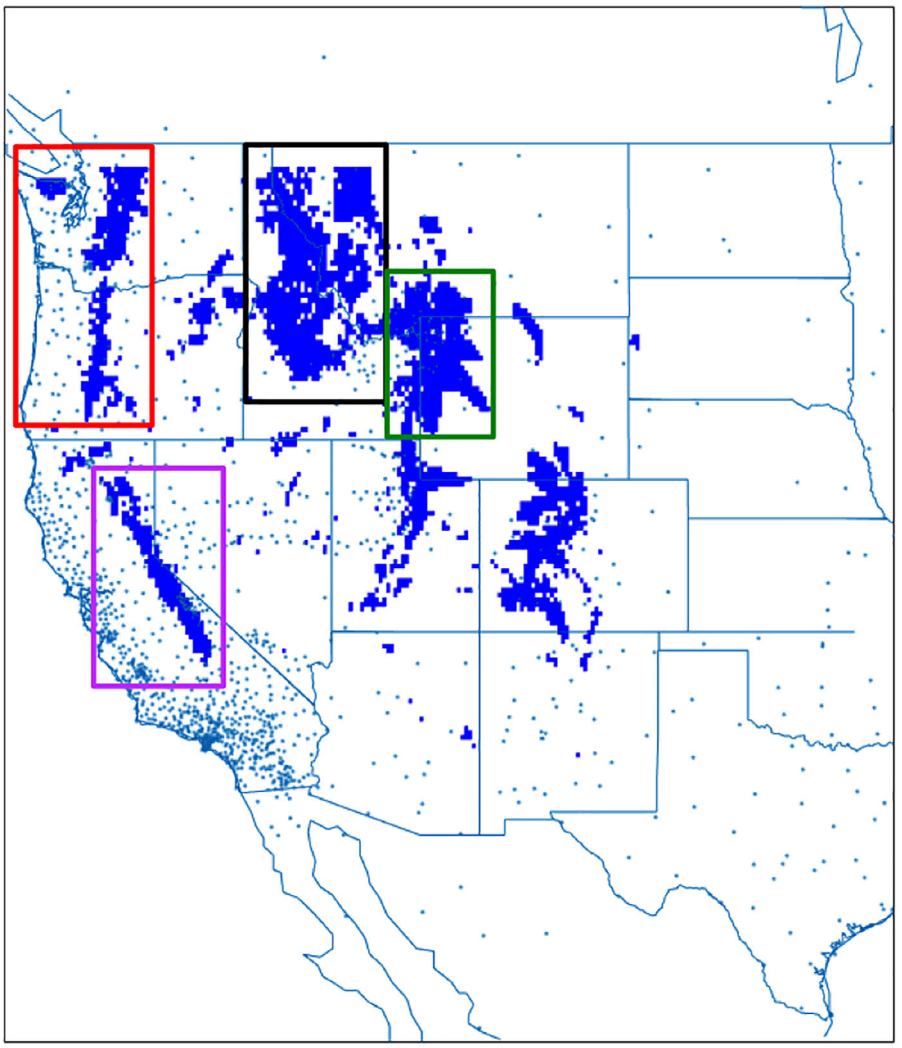

Figure 2. Spatial mask based on SNODAS March average SWE (blue grid cells), and distribution of GPS station locations (points) used in this study, shown relative to the full study domain (25∼53°N, 95∼125°W). Colored boxes denote the individual subregions: Sierra Nevada (purple), Pacific Northwest (red), northern Rockies (black), and Yellowstone area (green). Only the blue grid cells within the boxes comprise the subregions as we have defined them.

We then invert the synthetic displacement field to recover the monthly load anomaly at every grid point. We finally compare the inverted and input loads in order to investigate how error varies with station density and spatial loading patterns. The input loads are not intended to be a completely accurate representation of real hydrologic loading; their purpose is to provide a known input with a realistic spatial distribution of water mass. Thus, for this study, we take the input TWS anomaly to be an error-free load distribution from which we can quantify error associated with the inversion and with the current distribution of GPS stations. Quantifying these errors is crucial to developing this method into a viable product for water resources monitoring.

In this study, we consider the accuracy of mountain-range-scale SWE estimated from synthetic GPS vertical position data to address the question of whether continuous GPS data can be used to monitor seasonal SWE accumulation and melt and fluctuations in peak SWE from year to year. We quantify error magnitude in estimated SWE and its relationship to GPS station density in four mountainous subregions of the WUSA: the Sierra Nevada, Pacific Northwest, Yellowstone area, and northern Rockies. We additionally investigate the partitioning of the hydrologic load into SWE and soil moisture over the course of the snow accumulation season. Last, we evaluate two potential improvements to the accuracy of SWE estimates made using this method: (1) simulating increased station density to reduce gaps in the station network and (2) applying

gain factors to the inverted result similar to those applied to the GRACE solution (Landerer & Swenson, 2012; Long et al., 2015).

2. Methods

2.1. Input Hydrologic Loads

We use input hydrologic (TWS) loads following Fu et al. (2015) consisting of SWE from the Snow Data Assimilation System (SNODAS) and total column soil moisture from the North America Land Data Assimilation System (NLDAS-2) Noah model. SNODAS is the most advanced operational system for monitoring SWE at relatively high spatial resolution, assimilating satellite, airborne, and ground measurements with output from a snow mass and energy balance model (Carroll et al., 2001). NLDAS-2 provides a high-resolution gridded estimate of total soil moisture storage (Xia et al., 2012), which is necessary to realistically model displacement magnitudes. Ground truth soil moisture data from in situ probe networks are only available for certain locations, and thus cannot be used as an input. Fu et al. (2015) found good correspondence between the hydrologic loads predicted by these inputs and calculated from real GPS displacement data in Oregon and Washington.

It would be possible to use an input load that only includes SWE. However, the inclusion of soil moisture is likely to better represent true hydrologic loading, where soil moisture is a significant component of TWS. In addition, inverting for an estimate of total TWS from vertical displacement data, then estimating, and removing the soil moisture component, more closely approximates how GPS-based SWE estimation would be carried out in practice.

We generated 151 monthly distributions of TWS from October 2003 to April 2016 on a $1/8^\circ$ grid spanning 25°N∼53°N and 95°W∼125°W. This domain represents the western half of the native NLDAS-2 grid; the higher resolution (1 km × 1 km) SNODAS data are reinterpolated onto this grid. Loads at grid points corresponding to surface water bodies or the ocean were set to zero—we assume no temporal variations in loading at these points. The TWS load is thus a known combination of SWE and soil moisture, with both varying independently through space.

We converted all monthly loads to anomalies relative to the long-term average TWS load over the entire 151 month study interval. These anomalies, which are used to calculate synthetic displacements and validate the inverted load estimates, are calculated independently for each grid cell from

$$TWSA(x,y,t) = SWE(x,y,t) + sm(x,y,t) - \left(\overline{SWE}(x,y) + \overline{sm}(x,y)\right), \quad (1)$$

where *TWSA* is terrestrial water storage anomaly, *SWE* is snow water equivalent, *sm* is soil moisture, $\overline{SWE}$ is long-term average SWE, $\overline{sm}$ is long-term average soil moisture, *x* and *y* are the spatial coordinates, *t* is time in months, and all components are reported in mm of water equivalent. The anomalies relative to the long-term average used here differ from monthly average anomalies, in which a given month (e.g., April 2011) is compared to the long-term average for that month (e.g., all Aprils; Figure 3). Anomalies relative to the long-term average have a mean of zero but preserve the shape of the annual water cycle, in which a loading maximum is typically reached during the spring snow accumulation season, with subsequent hydrologic unloading as snow melts and soil dries over the course of the summer and early fall. Generally, therefore, the anomalies are positive in the spring (Figure 1a) and negative in the fall, although this is not necessarily the case in times of extreme drought.

Use of TWS anomalies rather than absolute loads better approximates the inversion procedure that would be used to estimate SWE from real displacement data. Raw GPS data are typically detrended to correct for secular tectonic movement and glacial isostatic adjustment, yielding a displacement residual with a mean of zero with fluctuations on timescales of months to several years (e.g., Chew & Small, 2014). The seasonal cycle ordinarily dominates these fluctuations. When the displacement residuals are inverted for load, the load time series will also be mean zero with fluctuations reflecting seasonal hydrologic loading and unloading.

2.2. Forward Model

We calculated the expected load-induced vertical displacement at the location of each of the 1,395 GPS stations currently in operation in the study domain from the load anomaly distribution for each month (Figures

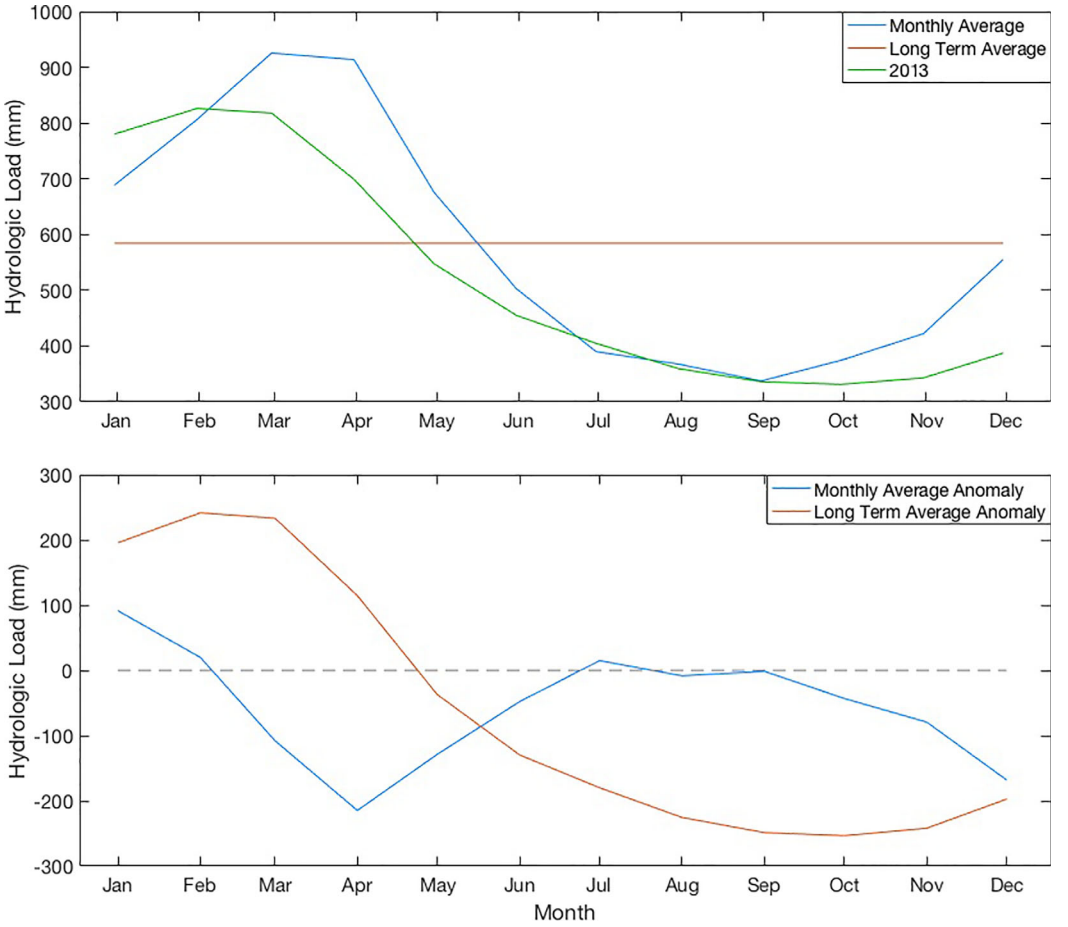

Figure 3. (top) Sierra Nevada monthly average hydrologic load (blue) and long-term average load (red) over October 2003 to April 2016, compared to monthly load over 2013 (green) from the input hydrologic load. (bottom) 2013 load anomaly relative to the monthly average (blue) and the long-term average (red). This study utilizes anomalies relative to the long-term average. Although April 2013 was dry relative to other Aprils (negative monthly average anomaly), it was wet relative to the average for all months over the period of record (positive long-term average anomaly).

1b and 2). In practice, some stations—such as those exhibiting signs of poroelastic effects or transient volcanic deformation—would be omitted due to signal interference (e.g., Borsa et al., 2014). As this study used only synthetic data and was concerned with quantifying error based on station density and loading characteristics, we did not remove any stations from our analysis. Only ∼10% of stations are omitted in practice (e.g., Borsa et al., 2014), thus the station distribution used in this synthetic experiment is reasonable. Expected displacement at each GPS station location was calculated using the SPOTL (Some Programs for Ocean Tide Loading) geophysical modeling program in nontidal mode. SPOTL takes a uniform grid of mass anomalies and generates the corresponding elastic surface displacement at any point on the planet, assuming a spherically symmetric Earth with depth-dependent elastic properties specified in a one-dimensional rheological model (Agnew, 2012). We used multiple calls to SPOTL to generate synthetic displacements at each station location for each of the 151 monthly load anomalies. Although we used SPOTL with a specific choice of rheological model (Gutenberg-Bullen; Agnew, 2012), the vertical displacement response generated by most one-dimensional models converges at load distances greater than ∼10 km (Na & Baek, 2011). Thus, the choice of rheological model is not critical to our analysis.

2.3. Inverse Model

Following the procedure of Borsa et al. (2014), we performed an inversion of the forward-modeled synthetic GPS displacements to estimate the associated load on every grid cell of the $1/8^\circ$ input grid. We regularized our solutions using minimum curvature smoothing, incorporated as a set of additional constraints within

Table 1
Summary of Area and Station Density for WUSA Mountainous Subregions

Subregion	Area[a] (km^2)	Number of stations[b]	Station density ($\frac{stations}{10{,}000\ km^2}$)
Sierra Nevada	31,568	36	11.4
Pacific Northwest	44,157	18	4.08
Yellowstone	67,322	41	6.09
Northern Rockies	11,6740	15	1.28

[a]Total area of grid cells included in each subregion (Figure 2). [b]Number of stations within the subregion boundaries as previously defined.

the design matrix. Specifically, for each time step in our analysis, we minimized the damped least squares problem

$$\|\mathbf{Gm}-\mathbf{d}\|_2^2 + \lambda^2\|\mathbf{Lm}\|_2^2, \quad (2)$$

where **d** is a vector of observed vertical displacements at each GPS station for the time step, **m** is the (unknown) equivalent water thickness in all cells of a uniform model grid, **G** is a matrix containing the displacement response at each station for a unit load at each cell of the model grid (calculated using SPOTL), **L** is a smoothing matrix of spatial second-differences, and λ is a regularization parameter, with higher values of λ resulting in smoother solutions (Aster et al., 2005). Additional details are provided in the supporting information Text S1.

We chose a model grid resolution of $1/8^\circ$ and a λ of 0.1 in order to minimize error in the inverted result (supporting information Text S2). This high spatial resolution (relative to previous studies) and minimal smoothing (low λ) allow for better representation of the short-wavelength hydrologic load variability characteristic of topographically complex mountain environments. Although we estimated the load for all nonocean, nonsurface water grid cells within the domain, we limited our analysis of error to SWE-dominated mountainous areas. To do so, we created a spatial mask that excludes all grid cells with a March average SWE load less than 100 mm based on SNODAS (Figure 2). This mask shows good correspondence with high elevation areas in the WUSA. We truncate the spatial mask north of 48.5°N to omit an area with very low station density.

Fu et al. (2015) found that edge effects introduced error into the GPS solution to a distance of up to ~3° between the area of interest and model domain boundary. Truncating the spatial mask at 48.5°N establishes a buffer ≥4.5° between the model domain boundary and any grid cell considered in this error analysis. As this distance exceeds the threshold found by Fu et al. (2015), we ignore edge effects in this study.

2.4. Subregions

We selected four subregions of interest to test the accuracy of our estimated hydrologic loads for a range of GPS station densities, and physiographic province sizes and shapes (Table 1). Each subregion is a subset of the SNODAS-based spatial mask (Figure 2). The Sierra Nevada and Pacific Northwest subregions are long, narrow mountain belts with comparatively high station density both within the boundaries of the subregion and in the surrounding areas. Yellowstone is a small, roughly square area with relatively high station density within the subregion but lower station density outside the subregion boundary. The northern Rockies subregion is larger and more complex in shape with relatively few stations, most of which are concentrated at the southern margin.

2.5. Soil Moisture Fraction of TWS

In order to quantitatively estimate SWE from GPS data, it is necessary to partition the inverted estimate of TWS change into its constituent parts. Although TWS includes all terrestrial water including surface and groundwater, in this study we take TWS in snow-dominated hydrologic regimes to be a combination of snow and soil moisture. We calculate the soil moisture fraction of the known input load in order to better understand the relationship between SWE, soil moisture, and the time variability of surface displacements. This metric gives the fraction of loading over the snow accumulation season that can be attributed to changes in soil moisture rather than SWE.

A soil moisture fraction for each subregion is calculated as

$$F_{sm} = \frac{\Delta sm}{\Delta TWS}, \quad (3)$$

where Δsm is the change in NLDAS-2 soil moisture and ΔTWS is the change in TWS between the first snowfall of the accumulation season (i.e., the first month of the water year in which SNODAS SWE is nonzero) and peak SWE. A separate soil moisture fraction is calculated for every year in the study interval. As solid Earth elastic deformation is a linear process, the soil moisture fraction can be removed from the inverted TWS estimate in order to compare input SWE and inverted SWE directly.

2.6. Additional Stations

We investigated error reduction in the Sierra Nevada subregion by simulating two scenarios where GPS stations were added to the network within the subregion boundary. In the first, 10 stations were added at locations chosen to minimize interstation distance, thereby improving station coverage. This scenario is analogous to a deployment strategy that could be adopted if additional stations were installed for the purpose of monitoring changes in hydrologic loading. In the second, 100 additional stations were simulated at random locations within the subregion (Figure 4). For both scenarios, the complete time series of synthetic displacements was recalculated from the original load anomaly maps with the additional stations included, then inverted to recover load.

2.7. Gain Factors

We further investigated error reduction through linear scaling of the inverted loads to better match the known input loads. We considered two time-varying rescaling strategies—subregion gain factors and grid-cell gain factors—designed to adjust the slope of the relationship between input and inverted load onto the 1:1 line (representing perfect correlation between the two).

To estimate the subregion gain factors, we consider the average inverted load relative to the average input load over each of the four mountainous subregions, performing a separate analysis for each month of the year. For a given month (e.g., March), we calculate a regression coefficient and constant (i.e., slope and intercept) between all subregion-averaged input and inverted loads from 2003 to 2016 (Figure 5). For the Sierra Nevada subregion, for example, the resulting March regression constant was 7 mm, indicating an expected average inverted load of 7 mm for an average input load of 0 mm. The regression coefficient was 0.81, indicating that the subregion-averaged inverted load was ~81% of the subregion-averaged input load for any given March in the study interval. We rescaled the inverted load in each cell of the subregion by subtracting 7 mm then dividing by 0.81, thereby adjusting the relationship onto the 1:1 line.

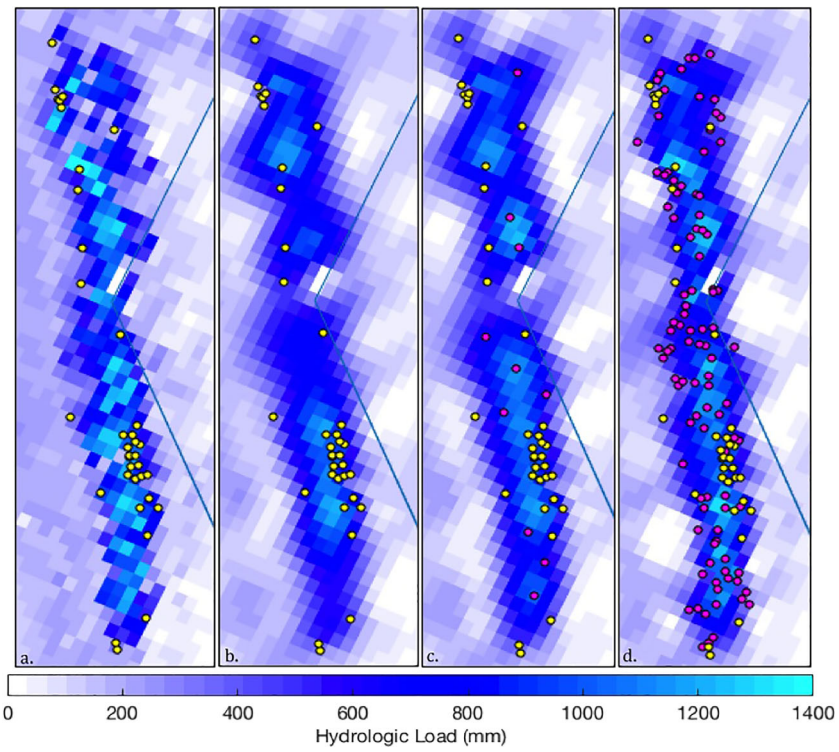

Figure 4. March 2011 (a) input load, (b) inverted load from actual GPS station distribution, (c) inverted load with 10 simulated additional stations, and (d) inverted load with 100 simulated additional stations over the Sierra Nevada subregion. Real station locations are yellow points; simulated stations are red points. The solid blue line is the California-Nevada border. Each grid cell is $1/8°$.

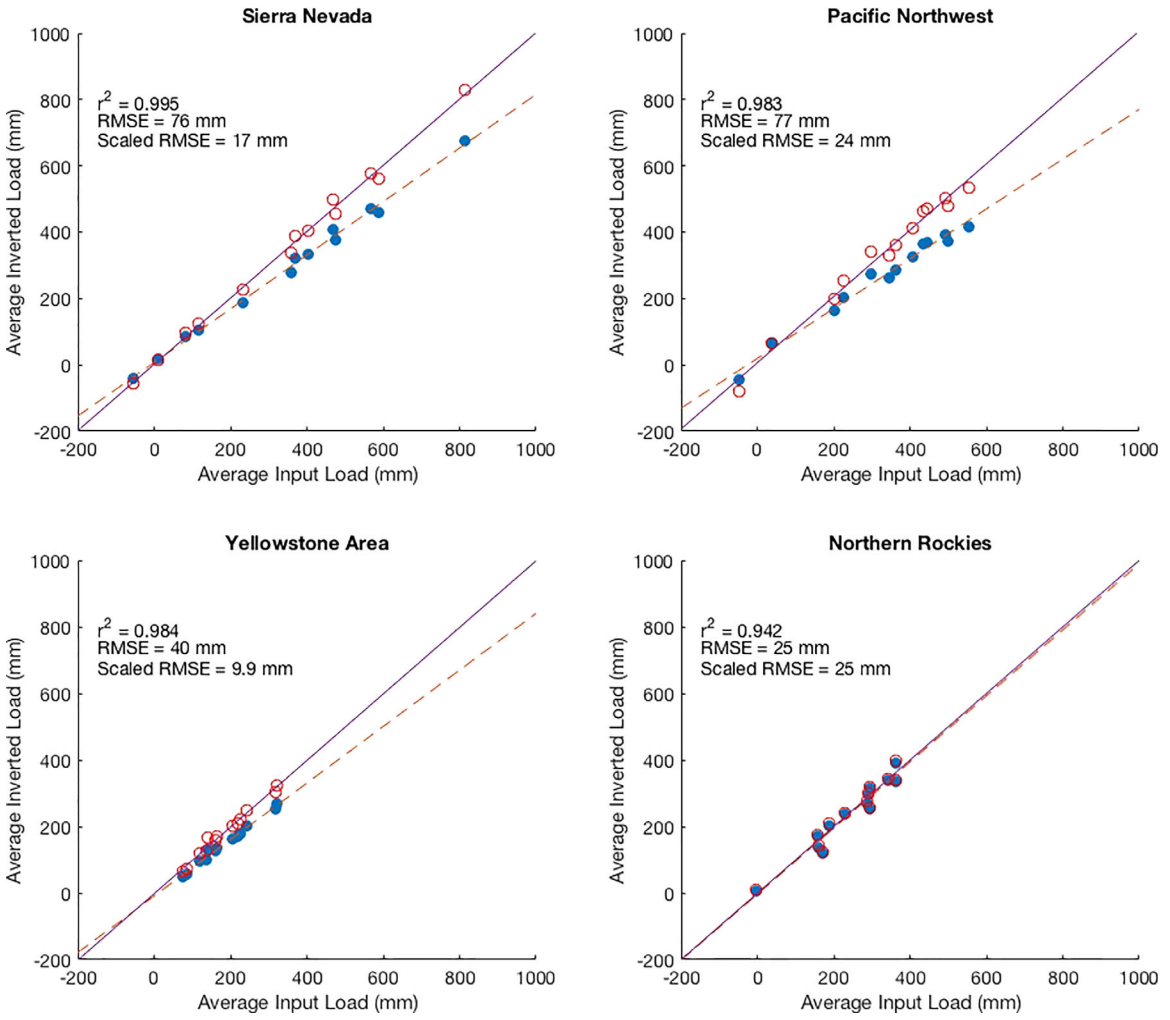

Figure 5. Subregion-averaged input load versus inverted load for all Marches in the time period analyzed. Blue points represent the unscaled relationships; red circles represent the relationships scaled onto the 1:1 line. The dashed (unscaled) and solid (scaled) lines are best-fit lines for each data set. RMSE relative to the 1:1 line is reduced by scaling for all subregions except the northern Rockies.

For the grid-cell gain factors, we use the same procedure described above, but calculate the 12 monthly gain factors on a grid-cell-by-grid-cell basis. Thus, for a given month (e.g., March), each grid cell in the subregion is scaled by a different value.

In both cases we use different gain factors for each month of the year to account for seasonal variations in hydrologic loading patterns. Use of a time-invariant gain factor fails to address important differences in loading characteristics between, for example, the annual maximum and minimum snow loads. At peak SWE, a large load gradient follows the topographic gradient between snow-dominated mountains and the surrounding lowlands. By contrast, the early fall is characterized by much more uniform loading between mountains and lowlands due to the lack of mountain SWE. As we discuss below, the presence or absence of a load gradient profoundly impacts error magnitude. Thus, the magnitude of the gain factor applied to the inverted result must be allowed to vary seasonally. For each subregion, the magnitude of the gain factor reaches a maximum at the time of peak SWE—when load gradients between mountains and surrounding areas are the strongest—and declines to a minimum in the early fall (Figure 6). We chose to apply the grid-cell gain factors for similar reasons. Load gradients vary significantly at short-wavelengths in mountain environments, and thus individual rescaling of grid cells may be preferable for error reduction at fine spatial scales (e.g., supporting information Figure S3).

The gain factors we calculate are optimal for the subregion, station distribution, inversion resolution, and λ used in our analysis. For a different distribution of GPS stations, a differently defined subregion, or a

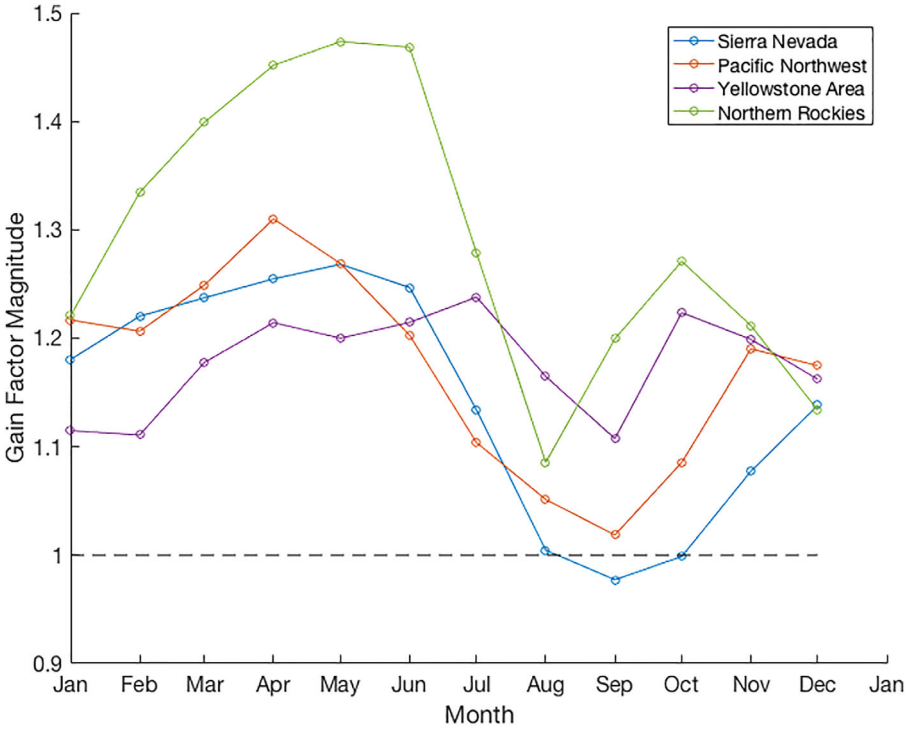

Figure 6. Subregion gain factor magnitude by month, calculated as the inverse of the regression coefficient for the relationship between subregion-averaged input and inverted loads. Higher magnitudes indicate more leakage and correspond to peak annual SWE, when load gradients are expected to be highest.

different λ, our gain factors would not be optimal and thus would need to be recalculated. Specifically, higher λ requires greater scaling (discussed below) as high magnitude input loads are redistributed over larger areas in the inversion (supporting information Figure S4).

As the gain factors were calculated from relationships between input and inverted loads, it was necessary to test their utility beyond the loading derived from NLDAS and SNODAS. To validate the gain factors, we produced two additional load anomaly time series in which the original hydrologic load was nonlinearly perturbed (supporting information Text S3). We performed a full forward model and inversion for each perturbed time series, then scaled the inverted results with both empirically derived gain factors from the original time series. We finally compared the scaled inverted results back to the perturbed input time series.

It is well established that spatial patterns of snow accumulation are consistent from year to year, except in cases of extreme anomalies (Schirmer et al., 2011). This suggests that stable relationships should exist between input and inverted loads for any given grid cell. However, as we discuss below, load redistribution in the inverted result due to smoothing and the irregular distribution of GPS stations introduces noise into this relationship at many grid cells. Thus, nonlinear perturbation is a way of testing the ability of the gain factors to reduce error when the load distribution deviates from its normal spatial pattern. Error reduction in the perturbed time series by the empirically derived gain factors supports the idea that model-derived gain factors can be applied to results inverted from real GPS data (supporting information Tables S1 and S2).

3. Results

3.1. Accuracy of Inverted Hydrologic Loads

The inversion accurately reproduces the seasonal cycle of hydrologic loading averaged over each subregion, regardless of the load's spatial variance or interannual variability (Figure 7). Both the Sierra Nevada and Pacific Northwest have snow loads with high spatial and interannual variance, while Yellowstone and the northern Rockies have lower-variance loads with a more consistent seasonal cycle. In all cases, the

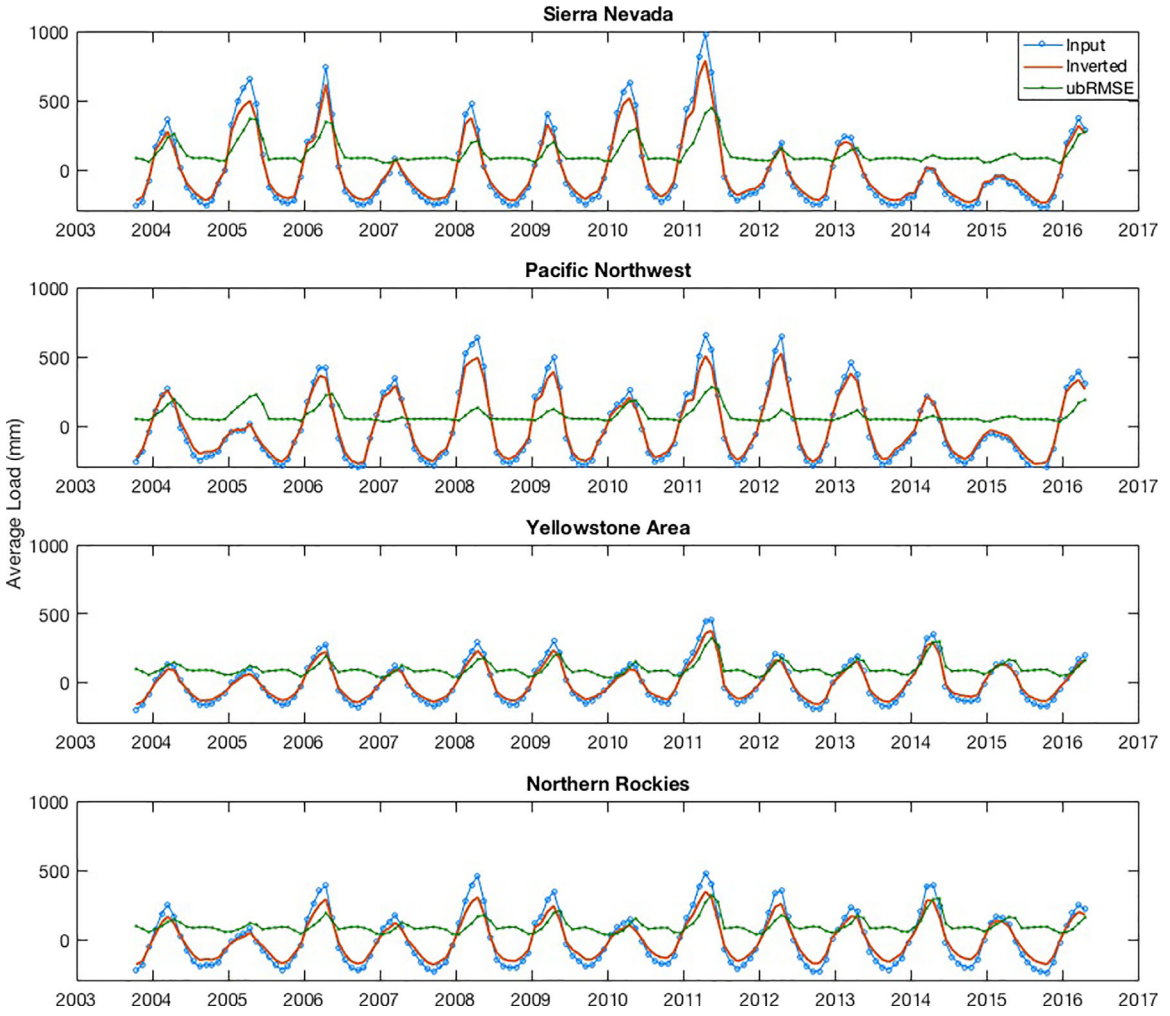

Figure 7. TWS anomaly averaged over each subregion. A low bias is evident in the magnitude of the inverted load (red) relative to the magnitude of the input load (blue). Time series of ubRMSE (green) indicate that maximum error at the grid cell scale occurs at or shortly after peak annual loading.

timing of seasonal loading and unloading in the input model is accurately reproduced by inverting the synthetic GPS station displacement field.

While the shape of the seasonal loading time series is well captured, the inversion systematically underestimates the magnitude of peak positive and negative loading (Figure 7). For both positive and negative loads, the magnitude of the error increases nearly linearly with the magnitude of the load (Figure 8). All four subregions exhibit similar load-error relationships, in which average error becomes more negative with increasing average load. This relationship is manifested in the subregion time series as a reduced amplitude of seasonal loading in the inverted result compared to the input load (Figure 7).

At the mountain-range scale, year-to-year variations in peak hydrologic load are well represented. March average input and inverted loads are highly correlated in all subregions (Figure 5). Three of the four subregions have $r^2 > 0.98$ for this linear relationship, while the northern Rockies subregion has a slightly lower r^2 (0.94). The inverted load is biased low relative to the input load, evident as a slope <1 in the best-fit lines (Figure 5). The inverted load can be rescaled to correct for this bias, yielding accurate representation of the year-to-year changes in TWSA (discussed below).

Average March and September error metrics for each subregion are summarized in Table 2. March and September are the months of maximum and minimum loading, respectively. They represent the months where absolute load is most underestimated. We define bias as the difference between the average inverted load and average input load over all grid cells within a region of interest. For each subregion, we also calculate

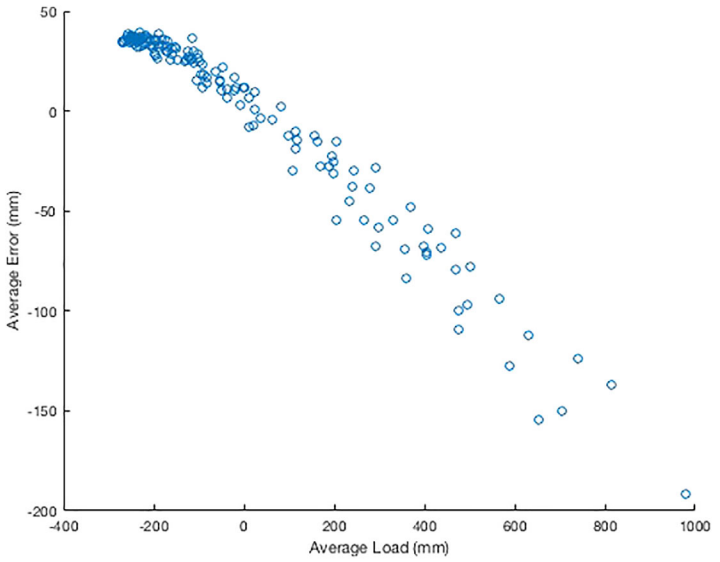

Figure 8. Average load versus average error for the Sierra Nevada subregion. Each point represents the average values for an individual month. Average error grows more negative with increasing average load. This trend is evident in all subregions.

an unbiased root mean square error (ubRMSE) after bias correction of the inverted load. We normalize bias and ubRMSE by the subregion-averaged input load for that month, giving error as a percentage of the average load over the subregion. March normalized bias for the Sierra Nevada, Pacific Northwest, and Yellowstone ranges between 15% and 20%. The northern Rockies subregion, which has considerably lower station density, has a normalized bias of 29%. March normalized ubRMSE ranges between 50% (Sierra Nevada) and 65% (Yellowstone) for the four subregions. The corresponding error metrics for September are similar in magnitude. These results demonstrate that the subregion in which the average input and inverted loads are the closest, is not necessarily the subregion in which the inversion best reproduces the spatial pattern of the input load on a cell-by-cell basis.

Because of smoothing, the displacement generated by a high-magnitude load over a small region of the input grid is inverted to yield a lower magnitude load over a larger area. This inversion "leakage" of some of the load in a given grid cell into surrounding grid cells is responsible for both the largest magnitude errors at the grid-cell scale and the systematic underestimation of average loading at the subregion scale. Within the Sierra Nevada subregion, for example, the input load map often contains single cells with high-magnitude loads surrounded by cells with much smaller loads (Figure 4a). This pattern is particularly common at the peak of the snow accumulation season (March–April). When the displacement fields from these months are inverted, the high-magnitude loads are distributed to surrounding grid cells, resulting in a dramatic load underestimation in the high-magnitude grid cell, and a slight load overestimation in the grid cells around it (Figures 1h and 4b). At the regional scale, relatively high-magnitude mountain snow loading is distributed into surrounding areas that have much smaller loads (Figure 9). The magnitude of average loading is underestimated by the inversion within the Sierra Nevada subregion but overestimated in a band $\sim$4 grid cells wide outside the subregion boundary. Thus, the slope of the relationship between input versus inverted load is <1 within the subregion and >1 in the surrounding area (Figure 10).

3.2. Soil Moisture Fraction of TWS

We present soil moisture partitioning results for the Sierra Nevada; results from other regions are similar. Increases in SWE and soil moisture over the snow accumulation season are highly correlated, with a temporal lag in peak soil moisture of 1 month or less relative to peak SWE (Figure 11). Over the melt season, SWE

Table 2
March and September Error Metrics by Subregion

Month	Subregion	Average input load (mm)	Average inverted load (mm)	Bias (mm)	Norm. Bias	Spatial ubRMSE[a] (mm)	Norm. ubRMSE	Unscaled RMSE[b] (mm)	Scaled RMSE[c] (mm)
March	Pacific Northwest	326	276	50	**0.15**	197	0.6	77	24
	Yellowstone	186	149	37	0.2	121	0.65	40	10
	N. Rockies	259	184	75	0.29	158	0.61	25	25
	Sierra Nevada	341	283	58	0.17	170	**0.5**	76	17
September	Pacific Northwest	−272	−241	−31	**0.11**	123	0.45	32	2
	Yellowstone	−163	−129	−34	0.21	85	0.52	34	4
	N. Rockies	−209	−156	−53	0.25	93	0.44	53	8
	Sierra Nevada	−247	−211	−36	0.15	77	**0.31**	36	1

[a]Spatial ubRMSE is calculated by comparing input and inverted load at all points within the snow mask for each region. [b]Unscaled RMSE is calculated by comparing average input and inverted load (blue points in Figure 5). [c]Scaled RMSE is calculated after temporal scaling (red circles in Figure 5).

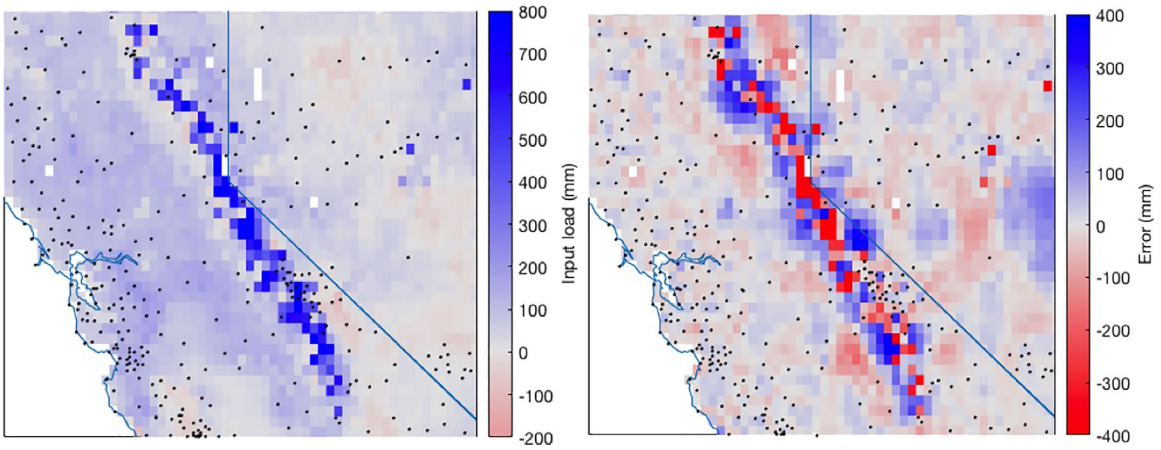

Figure 9. Input load (left) and error (right) for April 2016 in the Sierra Nevada subregion. Grid cells with high-magnitude input loads have large negative error, while adjacent cells—particularly those in areas lacking GPS stations—have large positive errors. This distribution of errors is evidence of leakage from high-load cells to their neighbors due to both smoothing and the distribution of stations.

declines to 0 mm, 1–3 months before soil moisture reaches its annual minimum. The soil moisture component of our input hydrologic load has a consistent seasonal cycle and low interannual variability, with an annual range of 200–250 mm, regardless of snow conditions. The consistency of seasonal soil moisture fluctuations may be a weakness typical of the finite soil depth used in land surface models (e.g., Houborg et al., 2012). Interannual variability in SWE is much higher: the change in SWE over the snow accumulation season was just 58 mm in 2015 compared to 967 mm in 2011. Because the lag between peak SWE and peak soil moisture is small, and the soil moisture increase is similar from year to year, variations in the soil moisture fraction (equation (3)) are largely determined by the amount of SWE. In high (low) SWE years, soil moisture fraction is small (large). It varies from 0.2 in 2011 to 0.75 in 2015—a year in which there was virtually no SWE in the Sierra Nevada. Thus, from 20% to 75% of TWS change over the snow accumulation season in the Sierra Nevada can be attributed to soil wetting, with the remainder from snow. These results indicate seasonal soil moisture changes are considerable. The magnitude of the soil moisture fraction is, however, contingent on the magnitude of the soil moisture input—the accuracy of which is not well known. Inverted

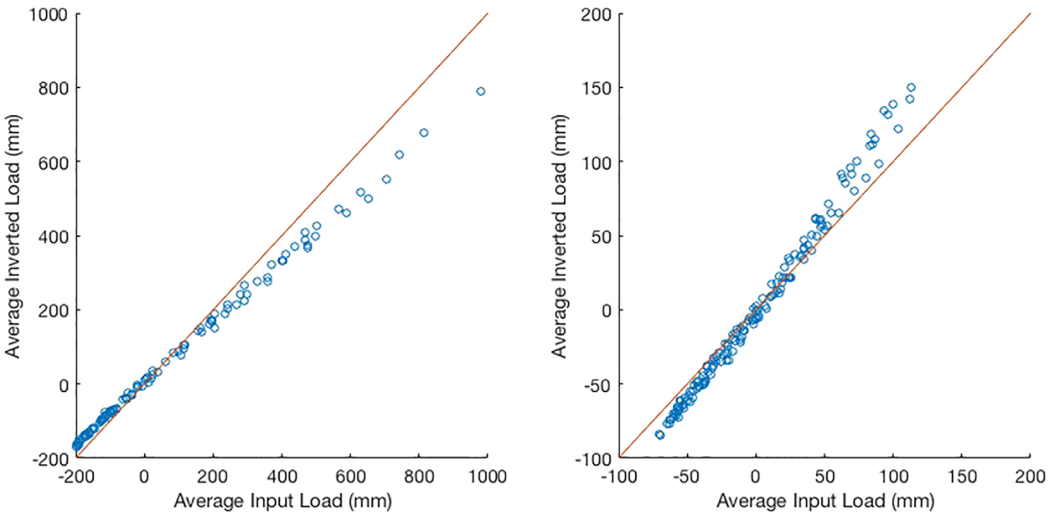

Figure 10. Scatters of average input load versus average inverted load within the Sierra Nevada subregion (left) and in the area surrounding the subregion but excluding the subregion itself (right). The 1:1 line is shown in red. Within the subregion, the inversion underestimates the magnitude of loading. Outside the subregion, the inversion overestimates loading, indicating some of the load from the subregion is distributed to grid cells outside the subregion boundary.

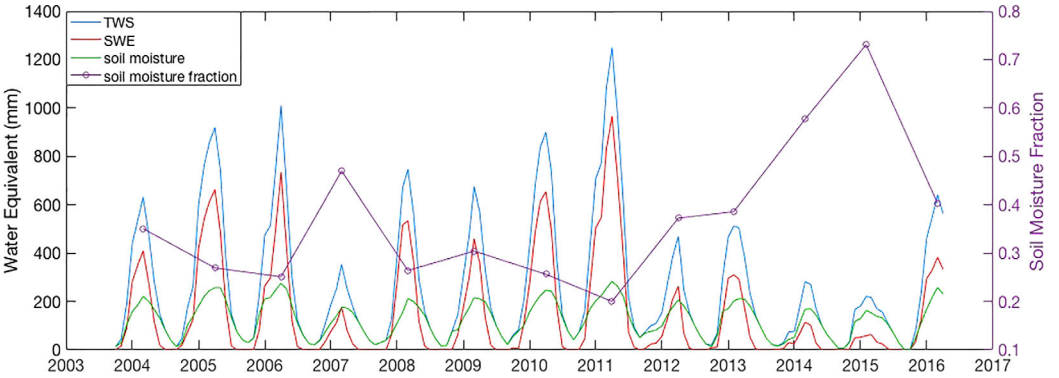

Figure 11. Input load partitioned into its constituent parts (SWE, soil moisture), with soil moisture fraction over the snow accumulation season.

TWS should not be considered an indicator of SWE alone, even in a snow-dominated hydrologic regime such as the Sierra Nevada.

3.3. Additional Stations

There is a clear improvement in the estimation of peak average loading when stations are added to the network (Table 3, supporting information Figure S5). March normalized bias is reduced from 17% to 10% in the 10 station case and to 6% in the 100 station case. March normalized ubRMSE is also reduced from 50% to 48% in the 10 station case and to 34% in the 100 station case. Similar reductions are evident in the September error metrics for the two scenarios. The reduction in ubRMSE indicates an improvement in the ability of the inversion to reproduce spatial patterns of loading. This improvement is evident in the inverted load maps for the base and additional station cases (Figure 4). The inversions with additional stations (plots c and d) more accurately reproduce the input load (plot a) than the inversion without extra stations (plot b).

3.4. Gain Factors

Results from the inversion show a systematic underestimation of peak load magnitude due to smoothing-related leakage. Although adding more stations ameliorates this problem, it is not a viable means of solving it completely. Therefore, we considered the utility of scaling the inversion, similar to the approach taken for

Table 3
March and September Sierra Nevada Error Metrics for Scaled and Additional Station Cases

March	**Average input load: 341 mm** Average inverted load (mm)	Average bias (mm)	Normalized bias	Average ubRMSE (mm)	Normalized ubRMSE
Unscaled	283	58	0.17	170	0.5
Subregion	341	0	**0**	175	0.51
Grid-cell	341	0	**0**	146	0.43
10 stations	307	34	0.1	163	0.48
100 stations	320	21	0.06	117	**0.34**
September	**Average input load: −247 mm** Average inverted load (mm)	Average bias (mm)	Normalized bias	Average ubRMSE (mm)	Normalized ubRMSE
Unscaled	−211	−37	0.15	77	0.31
Subregion	−247	0	**0**	77	0.31
Grid-cell	−247	0	**0**	13	**0.05**
10 stations	−229	−18	0.07	78	0.32
100 stations	−236	−12	0.05	50	0.21

Note. Unscaled error metrics are given for comparison.

GRACE TWS estimates (Landerer & Swenson, 2012; Long et al., 2015). Subregion and grid-cell gain factors were calculated to bring the relationship between inverted and input loads onto the 1:1 line (e.g., Figure 5).

Subregion scaling is critical to generate accurate estimates of year-to-year changes in SWE at the mountain-range scale. Inverted TWS anomalies during the snow season underestimate the magnitude of positive input anomalies due to smoothing-induced mass leakage away from snow-dominated areas (Figure 5). From May through June, the gain factor for the Sierra Nevada is approximately 1.2 (Figure 6). By contrast, the gain factor is very close to 1.0 between August and October, when the load in the mountains is similar to that in the adjacent valleys. Prior to scaling, the RMSE between input and inverted peak SWE is ∼75 mm in the Sierra Nevada and Pacific Northwest, 40 mm for Yellowstone, and 25 mm for the northern Rockies (Table 2). These errors are approximately 20% of the average annual load. After scaling, the RMSE is ∼20 mm or less in all subregions—the magnitude of peak SWE from the inversion, including fluctuations from year to year, matches the input load almost exactly (supporting information Figure S5).

Like the subregion gain factors, the grid-cell gain factors yielded subregion-averaged inverted loads identical to the average input loads. Moreover, application of the grid-cell gain factors reduced ubRMSE to 43% in March and 5% in September. For both months, the grid-cell gain factor represents the most effective means of reducing both bias and ubRMSE.

4. Discussion

4.1. Area-Averaged Load Estimation

Year-to-year variations in peak loading are well represented at the mountain-range scale. The relationship between subregion-averaged input and inverted loads is linear with a high r^2 across all four subregions (Figure 5). Even the northern Rockies subregion, which has a very low station density and uneven station distribution relative to other subregions, has an r^2 of 0.94. This suggests that the station density and distribution necessary to accurately capture annual variations in peak loading at the mountain-range scale are already in place for most, if not all, of the major snow-dominated environments in the WUSA. Although smoothing introduces a leakage bias—causing the average inverted load to underestimate the average input load—the high correlation between the two loads allows for bias correction via simple linear scaling. At the spatial scales we tested, gain factors were effective at reducing leakage bias to zero or near zero. Thus, empirically derived gain factors allow for accurate TWS estimation via GPS displacement at the mountain-range scale.

Our procedure for deriving empirical gain factors for GPS-estimated loads consists of the following steps:

1. Based on the area and time interval of interest, eliminate any GPS stations that are subject to significant data gaps or are affected by deformation or poroelastic effects.
2. Using a realistic distribution of water mass from models or remote sensing, calculate expected displacement at each GPS station location for a representative epoch. Invert for load using a range of λ values. Choose λ that minimizes error between the input and inverted loads over the area of interest.
3. With the chosen station distribution, λ, and inversion resolution, calculate synthetic displacements and invert for load estimates. The calculation should be repeated over the full-time interval of interest using data from intervals with both normal and anomalous loading. Calculate gain factors based on the relationship between input and inverted loads.
4. Process real GPS data from the chosen station distribution over the time interval of interest to remove effects from atmospheric loading, earthquake-related offsets, etc. (e.g., Argus et al., 2014; Borsa et al., 2014; others).
5. Invert real GPS data, and apply gain factors from synthetic experiments to the inverted load estimates.

Although this study has focused on area-averaged load estimation at the mountain-range scale, estimates at the basin or catchment scale are more likely to be useful in water management decisions. An important next step, therefore, will be testing the accuracy of GPS-based estimates of load over individual catchments and drainage basins. Smaller areas will suffer from a more substantial leakage bias. Below a certain spatial scale or threshold number of grid cells, it is possible that the linear relationship between average input and inverted loads will be lost, particularly in basins with no GPS station nearby. In addition, many basins include both snow-dominated mountain headwater regions and lowlands where snow is absent. They therefore will likely be subject to leakage both out of the snow-dominated zone and into the nonsnow-

dominated zone, the effects of which we have not considered here. In such cases, scaling may not be as useful in terms of increasing the accuracy of area-averaged load estimates.

Fu et al. (2015) modeled GPS vertical displacement and TWS change in the Cascade Range of Oregon and Washington following the same procedure and using the same input load described here. The Cascade Range comprises the majority of our Pacific Northwest subregion, which also includes the Olympic Range. Their study found a maximum difference between the inverted result and input load of ~8% of the change in loading over the snow accumulation season. We calculate an average March normalized bias of 15%, nearly double that presented in Fu et al. (2015). This discrepancy most likely stems from differences in the regions considered. Our spatial mask removes any grid cell with less than 100 mm average March SWE. The Cascade Range as defined in Fu et al. (2015) extends outward to grid cells omitted by our spatial mask. The wider spatial extent of the Cascade Range defined in Fu et al. (2015) captures more of the leakage from snow-dominated areas. Here, some mass is lost to leakage beyond the subregion boundary and is therefore omitted from the inverted result. Extending our subregion boundaries outward to an extent similar to that in Fu et al. (2015) reduced the amount of leakage across the boundary and lowered March normalized bias to ~10%.

4.2. Grid-Cell Scale Load Estimation

At the grid-cell scale ($^1/_8°$), limited GPS station coverage reduces the ability to resolve short-wavelength spatial patterns, even when minimal smoothing is used. Since there is insufficient information in the GPS observations to fully reproduce the input load, short-wavelength features are largely determined by the smoothing constraint in the inversion. The amount of leakage from grid cells with high-magnitude input loads depends both on the magnitude of loading in neighboring grid cells and the local distribution of GPS stations. Spatial patterns in the input loads are relatively consistent from year to year; however, small inconsistencies in the input load are magnified in the inverted load due to gaps in the station network. This variability in leakage introduces scatter into the relationship between input and inverted loads, particularly in grid cells for which there is no GPS station in the immediate vicinity. Calculating gain factors for grid cells without a strong linear relationship between input and inverted loads does not significantly improve loading estimates for these grid cells relative to the unscaled inverted result. Thus, grid-cell scaling cannot perfectly reproduce spatial loading patterns, although it can provide modest error reduction at the grid-cell scale.

While scaling may not be effective, our results suggest grid-cell-scale load estimates can be improved through increases in station density. Improvements to Sierra Nevada ubRMSE were minimal in the 10 station case but substantial in the 100 station case (Figure 4). Reductions to grid-cell-scale error comparable to those found in the 100 station case are likely possible with fewer additional stations if they are chosen to minimize interstation distance rather than placed randomly. Increasing station density also reduces leakage from areas of high loading, thereby improving the accuracy of both the magnitude and spatial distribution of loading in the inverted result. Similar improvements are likely possible in the other subregions.

Although increasing station density is a viable means of grid-cell-scale error reduction in a modeling study, it may not represent a realistic means of improving GPS-based SWE monitoring at very fine spatial scales. We have shown that large increases in the number of stations within a subregion are necessary to accurately reproduce spatial loading patterns of SWE. Installation and maintenance costs may make increases to the density of the GPS station network of the magnitude required infeasible. Moreover, at spatial scales up to ~10s of km, near-surface heterogeneity in rheological properties (e.g., between different rock types) can cause local variations in the solid Earth response to loading (Na & Baek, 2011; Wang et al., 2012). A one-dimensional rheological model such as we use is insufficient to fully describe the Earth elastic response to loading at these fine spatial scales, limiting the benefit from adding stations in regions that are already well instrumented. On the other hand, increasing station density may reduce the uncertainty associated with noise in the daily GPS displacement record, thereby affecting a portion of the total error budget we have not considered in this study.

Increasing the density of the current network should remain a priority; however, future work should also explore incorporating data from in situ monitoring networks such as SNOTEL to constrain the spatial distribution of mountain SWE at high spatial resolution. Gridded estimates of SWE produced from GPS vertical

displacements could also be validated or improved via comparison to remotely sensed snow cover maps from satellite or airborne lidar imaging.

While the spatial loading patterns produced by the inversion do not accurately reproduce the input, they represent a substantial improvement over the GRACE solution. We have shown that ubRMSE increases at coarser spatial resolutions—as short-wavelength variability is incorporated into a single spatially averaged grid cell value (supporting information Figure S1). The GRACE solution is available at approximately 1° grid cells, making each grid cell a spatial average of 64 grid cells at the $^1/_8$° spatial resolution discussed here. In consequence, the "mountain-range scale" for the GRACE solution can be as few as five grid cells. Moreover, a single GRACE grid cell may contain part of a long, narrow mountain belt as well as the lowlands to either side (Argus et al., 2014). The hydrologic load from GRACE for such a grid cell, while an accurate spatial average of loading over the area it contains, gives very little information about loading occurring solely in the mountains. Spatial patterns of loading at even finer scales cannot be represented at all.

4.3. Estimating SWE From TWS

The results presented above document the error in TWS anomaly estimation at the mountain-range scale. However, SWE is the most critical water resource metric in mountainous environments, so estimates of this variable are more useful than TWS anomalies for water resources management. Vertical displacement recorded by GPS measures TWS anomalies. These anomalies must be partitioned into TWS components (SWE, soil moisture, etc.) or assumed to be entirely associated with SWE accumulation and loss.

The assumption that a change in TWS corresponds entirely to a change in SWE is unlikely to be valid except in high alpine regions lacking soil cover or significant surface water bodies. Depending on the year, anywhere between 25% and 80% of TWS change over the snow accumulation season in our input load is attributable to SWE, with the remainder coming from soil moisture. In reality, the percentage of TWS change attributable to SWE will be somewhat lower, as we do not consider changes in components such as groundwater and surface water. The assumption that TWS change consists entirely of SWE would result in an overestimation of peak SWE, with the magnitude of error related to the soil moisture fraction. In 2011, for example, change in Sierra Nevada average TWS over the snow accumulation season is 1,209 mm, while change in SWE is 967 mm. Thus, 20% of the 1,209 mm TWS change comes from soil moisture. If other components of TWS are ignored, SWE will be overestimated by 25% of its true value.

The alternative to assuming that all TWS change comes from SWE is to take a water balance approach, where SWE is the residual after all other TWS components are removed. Numerous studies have adopted this method to estimate water budget components using TWS from GRACE (e.g., Rodell et al., 2004, 2011; Strassberg et al., 2009; others). Accurate estimation of the residual, however, requires accurate estimation of all other components of TWS that change significantly over the snow accumulation season. The most straightforward approach is to use hydrologic models and observational data to provide constraints on non-SWE TWS components. This approach will always introduce errors; however, SWE is typically by far the largest seasonally varying load in these environments. It therefore makes sense to derive TWS from GPS displacements, and then calculate SWE as a residual by estimating the other components of the water budget.

A final source of error in SWE estimation is the several mm mean formal uncertainty in daily GPS vertical positions (Amos et al., 2014). Previous studies have treated this uncertainty as random noise, removing it by filtering or averaging daily data over monthly intervals (Fu et al., 2015; Ouellette et al., 2013; others). As we have discussed monthly SWE estimates only, we ignore this uncertainty. However, it must be considered in order to estimate SWE at higher temporal resolutions.

5. Conclusions

The GPS station network in the WUSA is dense enough to accurately monitor changes in hydrologic loading at the mountain-range scale. The timing of seasonal loading is well captured in all subregions. Year-to-year variations in area-averaged peak loading suffer from a leakage bias. However, this can be corrected via scaling of the inverted result. At the grid-cell scale, the current GPS network fails to accurately reproduce the input spatial loading patterns. While adding GPS stations would improve spatial loading patterns, the current network already provides a substantial improvement over the GRACE solution in terms of monitoring spatially variable loading in mountainous regions. In addition to establishing a denser GPS network, future

work should make use of in situ SWE and remotely sensed snow cover data to constrain spatial loading patterns in areas where GPS station coverage is lacking.

Accurate GPS-based mountain-range-scale estimation of SWE requires accounting for other components of TWS. The accuracy of model or network-based estimates of soil moisture is critical to deriving SWE estimates from GPS-based TWS anomalies. Spatiotemporal changes in surface and groundwater also have the potential to significantly impact SWE estimates in certain locations. Accurate estimation of other components of TWS will allow GPS-based SWE monitoring at high spatiotemporal resolution in well-instrumented drainage basins—a crucial objective for water resources management in the WUSA.

Acknowledgments
This work was funded by NSF grants EAR-1521474, EAR-1521127, and EAR-1614218. NLDAS data were acquired via the Giovanni system, maintained by the NASA GES DISC. SNODAS data were acquired via the NSIDC Polaris Search and Order interface. Data generated in this study can be accessed via the University of Colorado Boulder Institutional Repository (scholar.colorado.edu). We thank Donald Argus and two anonymous reviewers for their helpful comments.

References

Agnew, D. C. (2012). *SPOTL: Some programs for ocean-tide loading* (Tech. Rep., Prog. Ver. 3.3.0.2). La Jolla, CA: Scripps Institution of Oceanography, University of California.

Amos, C. B., Audet, P., Hammond, W. C., Bürgmann, R., Johanson, I. A., & Blewitt, G. (2014). Uplift and seismicity driven by groundwater depletion in central California. *Nature*, *509*(7501), 483–486. https://doi.org/10.1038/nature13275

Argus, D. F., Fu, Y., & Landerer, F. W. (2014). Seasonal variation in total water storage in California inferred from GPS observations of vertical land motion. *Geophysical Research Letters*, *41*, 1971–1980. https://doi.org/10.1002/2014GL059570

Aster, R. C., Borchers, B., & Thurber, C. (2005). *Parameter estimation and inverse problems* (2nd ed.). Amsterdam, the Netherlands: Elsevier.

Bevis, M., Alsdorf, D., Kendrick, E., Fortes, L. P., Forsberg, B., Smalley, R., et al. (2005). Seasonal fluctuations in the mass of the Amazon River system and Earth's elastic response. *Geophysical Research Letters*, *32*, L16308. https://doi.org/10.1029/2005GL023491

Borsa, A. A., Agnew, D. C., & Cayan, D. R. (2014). Ongoing drought-induced uplift in the western United States. *Science*, *345*(6204), 1587–1590. https://doi.org/10.1126/science.1260279

Carroll, T., Cline, D., Fall, G., Nilsson, A., Li, L., & Rost, A. (2001). *NOHRSC operations and the simulation of snow cover properties for the coterminous U.S.* Paper presented at the 69th Annual Meeting of the Western Snow Conference (pp. 1–14), Sun Valley, Idaho.

Chew, C. C., & Small, E. E. (2014). Terrestrial water storage response to the 2012 drought estimated from GPS vertical position anomalies. *Geophysical Research Letters*, *41*, 6145–6151. https://doi.org/10.1002/2014GL061206

Clow, D. W., Nanus, L., Verdin, K. L., & Schmidt, J. (2012). Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA. *Hydrological Processes*, *26*(17), 2583–2591. https://doi.org/10.1002/hyp.9385

Doesken, N. J., & Judson, A. (1996). *The snow booklet: A guide to the science, climatology, and measurement of snow in the United States* (2nd ed.). Fort Collins, CO: Colorado Climate Center.

Dong, D., Fang, P., Bock, Y., Cheng, M. K., & Miyazaki, S. (2002). Anatomy of apparent seasonal variations from GPS-derived site position time series. *Journal of Geophysical Research*, *107*(B4), 2075. https://doi.org/10.1029/2001JB000573

Famiglietti, J. S., & Rodell, M. (2013). Water in the balance. *Science*, *340*, 1300–1301. https://doi.org/10.1126/science.1236460

Farrell, W. E. (1972). Deformation of the Earth by surface loads. *Reviews of Geophysics*, *10*(3), 761–797. https://doi.org/10.1029/RG010i003p00761

Fu, Y., Argus, D. F., & Landerer, F. W. (2015). GPS as an independent measurement to estimate terrestrial water storage variations in Washington and Oregon. *Journal of Geophysical Research: Solid Earth*, *120*, 552–566. https://doi.org/10.1002/2014JB011415

Houborg, R., Rodell, M., Li, B., Reichle, R., & Zaitchik, B. F. (2012). Drought indicators based on model-assimilated gravity recovery and climate experiment (GRACE) terrestrial water storage observations. *Water Resources Research*, *48*, W07525. https://doi.org/10.1029/2011WR011291

Koster, R. D., Mahanama, S. P. P., Livneh, B., Lettenmaier, D. P., & Reichle, R. H. (2010). Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. *Nature Geoscience*, *3*(9), 613–616. https://doi.org/10.1038/ngeo944

Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, *48*, W04531. https://doi.org/10.1029/2011WR011453

Long, D., Longuevergne, L., & Scanlon, B. R. (2015). Global analysis of approaches for deriving total water storage changes from GRACE satellites. *Water Resources Research*, *51*, 2574–2594. https://doi.org/10.1002/2014WR016853

Molotch, N. P. (2009). Reconstructing snow water equivalent in the Rio Grande headwaters using remotely sensed snow cover data and a spatially distributed snowmelt model. *Hydrological Processes*, *23*, 1076–1089.

Na, S., & Baek, J. (2011). Computation of the load love number and the load Green's function for an elastic and spherically symmetric Earth. *Journal of the Korean Physical Society*, *58*(5), 1195–1205. https://doi.org/10.3938/jkps.58.1195

Ouellette, K. J., De Linage, C., & Famiglietti, J. S. (2013). Estimating snow water equivalent from GPS vertical site-position observations in the western United States. *Water Resources Research*, *49*, 2508–2518. https://doi.org/10.1002/wrcr.20173

Painter, T. H. Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrkee, F., et al. (2016). The airborne snow observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sensing of Environment*, *184*, 139–152. https://doi.org/10.1016/j.rse.2016.06.018

Raleigh, M. S., & Small, E. E. (2017). Snowpack density modeling is the primary source of uncertainty when mapping basin-wide SWE with lidar. *Geophysical Research Letters*, *44*, 3700–3709. https://doi.org/10.1002/2016GL071999

Rodell, M., & Famiglietti, J. S. (2002). The potential for satellite-based monitoring of groundwater storage changes using GRACE: The High Plains aquifer, Central US. *Journal of Hydrology*, *263*(1–4), 245–256. https://doi.org/10.1016/S0022-1694(02)00060-4

Rodell, M., Famiglietti, J. S., Chen, J., Seneviratne, S. I., Viterbo, P., Holl, S., et al. (2004). Basin scale estimates of evapotranspiration using GRACE and other observations. *Geophysical Research Letters*, *31*, L20504. https://doi.org/10.1029/2004GL020873

Rodell, M., Mcwilliams, E. B., Famiglietti, J. S., Beaudoing, H. K., & Nigro, J. (2011). Estimating evapotranspiration using an observation based terrestrial water budget. *Hydrological Processes*, *25*(26), 4082–4092. https://doi.org/10.1002/hyp.8369

Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., et al. (2009). Evaluation of forest snow processes models (SnowMIP2). *Journal of Geophysical Research*, *114*, D06111. https://doi.org/10.1029/2008JD011063

Save, H., Bettadpur, S., & Tapley, B. D. (2016). High-resolution CSR GRACE RL05 mascons. *Journal of Geophysical Research: Solid Earth*, *121*, 7547–7569. https://doi.org/10.1002/2016JB013007

Schirmer, M., Wirz, V., Clifton, A., & Lehning, M. (2011). Persistence in intra-annual snow depth distribution: 1. Measurements and topographic control. *Water Resources Research*, *47*, W09516. https://doi.org/10.1029/2010WR009426

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A., & Pulwarty, R. S. (1999). Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. *Water Resources Research*, *35*(7), 2145–2160. https://doi.org/10.1029/1999WR900090

Steckler, M. S., Nooner, S. L., Akhter, S. H., Chowdhury, S. K., Bettadpur, S., Seeber, L., et al. (2010). Modeling earth deformation from monsoonal flooding in Bangladesh using hydrographic, GPS, and gravity recovery and climate experiment (GRACE) data. *Journal of Geophysical Research*, *115*, B08407. https://doi.org/10.1029/2009JB007018

Strassberg, G., Scanlon, B. R., & Chambers, D. (2009). Evaluation of groundwater storage monitoring with the GRACE satellite: Case study of the High Plains aquifer, central United States. *Water Resources Research*, *45*, W05410. https://doi.org/10.1029/2008WR006892

Wahr, J., Swenson, S., Zlotnicki, V., & Velicogna, I. (2004). Time-variable gravity from GRACE: First results. *Geophysical Research Letters*, *31*, L11501. https://doi.org/10.1029/2004GL019779

Wang, H., Xiang, L., Jia, L., Jiang, L., Wang, Z., Hu, B., et al. (2012). Load Love numbers and Green's functions for elastic Earth models PREM, iasp91, ak135, and modified models with refined crustal structure from Crust 2.0. *Computers & Geosciences*, *49*, 190–199. https://doi.org/10.1016/j.cageo.2012.06.022

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research*, *117*, D03109. https://doi.org/10.1029/2011JD016048