A ten-year comparison of water levels measured with a geodetic GPS receiver versus a conventional tide gauge

Kristine M. Larson*

Dept. Aerospace Engineering Sci., University of Colorado, Boulder, Colorado

Richard D. Ray

NASA Goddard Space Flight Center, Greenbelt, Maryland

Simon D. P. Williams

National Oceanography Centre, Proudman Building, Liverpool UK

**Corresponding author address:* Dept. Aerospace Engineering Sciences, University of Colorado, Boulder, CO 80309, USA.

E-mail: kristinem.larson@gmail.com

ABSTRACT

A standard geodetic GPS receiver and a conventional Aquatrak tide gauge, collocated at Friday Harbor, Washington, are used to assess the quality of ten years of water levels estimated from GPS sea-surface reflections. The GPS results are improved by accounting for (tidal) motion of the reflecting sea surface and for signal propagation delay by the troposphere. The RMS error of individual GPS water level estimates is about 12 cm. Lower water levels are measured slightly more accurately than higher water levels. Forming daily mean sea levels reduces the RMS difference with the tide gauge data to approximately 2 cm. For monthly means, the RMS difference is 1.3 cm. The GPS elevations, of course, can automatically be placed into a well-defined terrestrial reference frame. Ocean tide coefficients, determined from both the GPS and tide-gauge data, are in good agreement, with absolute differences below 1 cm for all constituents save K_1 and S_1. The latter constituent is especially anomalous, probably owing to daily temperature-induced errors in the Aquatrak tide gauge.

1. Introduction

A number of published reports have now documented how a standard geodetic-quality GPS receiver, situated at the coast with an unobstructed view of the sea, can act as an "accidental tide gauge" (Larson et al. 2013a,b; Löfgren et al. 2014a). GPS reflections off the sea surface, normally a source of error and noise for geodetic positioning, generate characteristic oscillations in received signal strength, and these can be analyzed to determine the height of the receiving antenna above the reflecting surface. With this approach, the receiver requires no special modifications, neither a second antenna facing down towards the sea nor a tilted antenna (Löfgren et al. 2011). Here we examine ten years of such sea-level measurements from a GPS receiver and compare them with simultaneous measurements from a collocated conventional tide gauge.

As previous studies have indicated, and as we document further here, analysis of GPS reflections is capable of supplying useful sea-level data for any number of applications. The technique cannot, however, completely replace conventional tide gauges. The precision of an individual water level estimate from a single GPS satellite overflight is far worse than the precision of a single tide-gauge reading. Moreover, the sampling rate is fundamentally limited by the number of satellite overflights. This is particularly an issue when GPS data are used from a site where no effort was made to improve the accuracy of water reflections. For this reason, the study of short-period phenomena (e.g., seiches or tsunamis) is likely to prove challenging when using the GPS technique. However, as suggested previously (Larson et al. 2013b), and discussed in greater detail below, averaging GPS measurements over periods of a day or longer yields mean sea levels that are nearly comparable to those obtained with conventional systems. Also as shown below, tides can be determined with comparable accuracies, even though their periods are obviously subdaily.

In fact, for certain problematic tidal constituents like S_1, determinations from GPS may be even more accurate than those from conventional gauges.

One great advantage of GPS-based measurements, in addition to the serendipitous recovery of sea level from a system not designed for it, is that resulting sea levels can be immediately placed into a well-defined terrestrial reference frame, with any vertical land motion precisely determined from the primary geodetic measurements of the GPS system. Inadequately known land motion is a problem that routinely plagues studies of long-term trends in mean sea level (e.g. Wöppelmann and Marcos 2016), and addressing that problem is automatically an integral part of the system. In addition, GPS reflections require none of the traditional infrastructure like stilling wells that are susceptible to storm damage and biological fouling and that need regular maintenance. (Modern tide gauges based on microwave radar sensors also dispense with stilling wells; e.g., see Figure 2 of Woodworth and Smith (2003).)

The data analyzed here were collected at Friday Harbor, 48.546°N, 123.013°W, located about 130 km northwest of Seattle, Washington, on San Juan Island, which sits between the Strait of Juan de Fuca and the Strait of Georgia. The GPS instrument sits about 345 m east of the tide gauge, at a point with better lines-of-sight for reflections. In the next section we discuss past work with GPS reflections, followed by a description of the GPS site, the conventional tide gauge, and how we analyzed both data sets. The main comparison results are given in Section 5.

2. Past Work

There are several different methods and experimental setups for ground-based GPS reflection sea level measurements. Here, however, the focus is on measuring water levels using the Signal to Noise Ratio (SNR) data from commercial off the shelf (COTS) geodetic receivers. SNR data are distinct from typical GPS ranging data in that they tell you nothing about the distance between the transmitting satellite and receiving antenna. However, they have the advantage that fluctuations in SNR levels caused by reflected signals can be easily observed for natural surfaces such as soil, water, snow, and ice (Larson et al. 2008, 2009). Unlike ranging data, where sophisticated models are needed for orbits, satellite and receiver clocks, relativity, and atmospheric delays, the background model for SNR data is a low-order polynomial. This smooth behavior is primarily controlled by the gain pattern of the geodetic antenna, which reduces direct signal power at lower elevation angles. Geodetic-quality COTS receivers always calculate SNR signals from both GPS frequencies (L1 and L2). For a variety of reasons, the quality of SNR data varies by receiver manufacturer, model, and frequency. This issue will be discussed further in the next section.

COTS GPS receivers were first used in two water-reflection experiments in 1998 at San Diego, California and Wallops Island, Virginia (Anderson 2000). However, the antennas were tilted 20° from zenith towards the ocean to improve reception. Benton and Mitchell (2011) estimated water-level reflections at two cliff sites ($\sim$30 m) overlooking the North Sea. Their retrieved heights were only accurate to a few meters. The first GPS tidal reflection study using SNR data from an upright COTS unit was presented by Larson et al. (2013a). Comparisons were made for three months of data from Onsala (Sweden) and Friday Harbor (USA). Validation of the methodology at the Onsala site was limited because there was no collocated tide gauge. At Friday Harbor, there was a collocated tide gauge, but because the authors restricted their study to GPS satellites that transmit the new L2 signal (5 at the time of that study), there were an insufficient number of observations to determine meaningful tidal coefficients. They found a correlation of 0.98 with respect to the NOAA tide gauge and an RMS residual of $\sim$ 10 cm.

Subsequently Larson et al. (2013b) evaluated one year of SNR data for a site in Kachemak Bay (Alaska). This site had enough satellite tracks that tidal coefficients could be estimated. These showed excellent agreement with a NOAA tide gauge operating in Seldovia with the largest tidal components, M_2 and S_2, agreeing to better than 2%. Much of this difference could be attributed to variations in the tide since the GPS is $\sim$30 km from the tide gauge. Larson et al. (2013b) also emphasized the need for correcting for a non-stationary reflecting surface during the measurements which results in a biased spectral peak, particularly at sites with large tidal range.

Löfgren et al. (2014a) extended these preliminary results by analyzing GPS data from five sites, including Friday Harbor and Onsala. The latter had recently had a tide gauge installed so that a more direct comparison could be made between the traditional tide gauge and the GPS results. The three new sites were located at O'Higgins (Antarctica), Burnie (Australia), and Brest (France). They found correlation coefficients of between 0.89–0.99 with respect to the co-located tide gauges, RMS differences from 6.2 cm to 43 cm and agreements of between 2.4% and 10% of the tidal range. Unlike the previous studies which excluded data below 5 degrees, **?** used very low elevation angle measurements (down to 0.5°at one site).

Santamaría-Gómez et al. (2015) examined both L1 and L2 SNR data at co-located GPS and tide gauge sites to estimate a leveling tie between the instruments and hence produce the ellipsoidal height of the tide gauge. They used data from 8 sites, primarily in France, including the sites at Brest and Burnie previously used by Löfgren et al. (2014a). Since they were estimating a static height they used the tide gauge data in the processing as a way to improve the estimate. They found agreements with in-situ leveling results typically at the 3 cm or smaller level. However, they also found biases in the results when using satellite elevations lower than 12° and between the L1 and L2 signals that were larger than 15 cm at two sites.

3. Friday Harbor Instrumentation

The tide gauge at Friday Harbor is one of the continuously operating CO-OPS (Center for Operational Oceanographic Products and Services) gauges maintained by the U.S. National Oceanic and Atmospheric Administration (NOAA). Digital hourly data are available from the site since 1934 with occasional gaps; data at 6-minute sampling, which we employ here, are available since 1996. Each 6-minute measurement represents an average of 181 one-second measurements, with additional filtering imposed by the gauge's protective well. The gauge now operating at the site is a standard acoustic Aquatrak gauge, a design in widespread use in the NOAA network for over two decades. While this type of tide gauge is more than adequate for our purposes, for the discussion below it is pertinent to note that Aquatrak gauges can be prone to errors from temperature-induced variations in the speed of sound within the enclosed sounding tube (Porter and Shih 1996; Hunter 2003). NOAA is slowly replacing their acoustic systems with microwave radar systems (Park et al. 2014).

The Friday Harbor GPS site known as SC02 was originally installed in 2001 for tectonic studies by the PANGA group (http://pangea.cwu.edu). At that time it operated a Trimble 4700 receiver, a geodetic-quality dual-frequency carrier phase receiver. It sampled measurements every 30 seconds until it was adopted into the EarthScope Plate Boundary Observatory (http://earthscope.org), a geodetic network installed in the western United States by the National Science Foundation in June 2006. At that time the Trimble 4700 was replaced with a newer Trimble model, the NetRS receiver, and the sampling rate was increased to 15 seconds. These sampling rates refer to how often observations are generated for geodetic users. They are not averages over 15 s (or 30 s), but instead are over much shorter intervals ($<$0.1 seconds) at the stated sampling interval (15 or 30 s).

There has been only one major equipment change since 2006 (April 29, 2015), when both the receiver and antenna were changed. The new receiver (Trimble NetR9) and antenna can track multiple constellation signals (GPS, GALILEO, GLONASS) and the third GPS frequency, known as L5. The new antenna should have the equivalent phase center for both L1 and L2 signals. However, based on offsets seen in positioning time series when similar adjustments have been made in GPS networks, we cannot discount a small (mm level) offset in the GPS water level signals at that time.

As shown in the photograph (Figure 1), the SC02 antenna is on a tripod monument. The legs of this monument were drilled $\sim$10 m into bedrock so that the positions estimated from the GPS data would be "anchored." The GPS antenna is $\sim$2 m above soil and covered by a radome. A mapview of the Friday Harbor station location is shown in Figure 2.

Although the remainder of this paper is concerned with *relative* sea levels—i.e., sea levels relative to instruments affixed to land—it is worth noting the absolute vertical land motion as determined by the GPS geodetic measurements at SC02. The station is included in the recent compilation by Blewitt et al. (2011), who report a vertical rate of $+0.25 \pm 0.68$ mm/y in the IGS08 terrestrial reference frame.

4. Data Analysis

a. GPS Reflection Sensing Zone

The sensing zone of a GPS reflection measurement is dependent on H, e, and Az, which are the height of the antenna above the reflecting surface, the angle of the satellite with respect to the horizon, and the satellite azimuth, respectively. These sensing zones are very long and thin ellipses that are offset from the GPS antenna. As the elevation angle increases, the sensing zone becomes smaller and closer to the antenna. Each rising and setting satellite arc will thus have a different sensing zone. Before proceeding to estimate reflection parameters, it is necessary to define an azimuth and elevation angle mask. The azimuths and elevation angles are chosen so that the sensing zones are on water. Fig. 2 shows the reflection mask corresponding to this study. Three ellipses are shown for each satellite track. The longest one is computed for an elevation angle of

5°, the second one 9°, and the last one 13°. Above 13° the Fresnel zone starts to include the shore. Azimuthally, the location of this particular GPS site allows data from only $\sim 50°$–$240°$. An additional region—shown in yellow—was masked because it produced significantly more outliers than the other regions.

b. Estimation of Reflector Height

The primary observation used in GPS water level studies is the SNR reflector height H. To estimate H, the SNR data are translated from units of dB-Hz to a linear scale, and the direct signal effect is removed using a low-order polynomial. These SNR residuals δ are modeled as:

$$\delta = A \cos\left(\frac{4\pi H}{\lambda}\sin e + \phi\right) \tag{1}$$

where λ is the GPS wavelength (19 cm for the L1 frequency). The angle e is calculated using the GPS navigation message, which is sufficiently accurate for these applications. Because the data are not evenly sampled, a Lomb-Scargle periodogram was used to extract H. The Lomb-Scargle periodograms were calculated using an oversampling interval which resulted in a precision of 3 mm. Reflector height estimates were retained only if their amplitudes (A) were greater than 7 volts/volts.

Although the new L2 signal is more precise for GPS reflection studies than the original L1 recorded by this receiver type (Larson et al. 2010), here we have opted to use only L1 signals. Our reasoning is that it is preferable to have access to signals from the entire GPS constellation (as is the case for L1) than to access an inhomogeneous L2 dataset. That inhomogeneity for the new L2 signal is caused by annual satellite launches between 2006 and 2013, followed by 3 launches per year since 2013. We have assessed some of these L2 SNR data and find they have a reflector height precision of 8 cm rather than the 12 cm observed for L1.

Here we will consider only relative sea level measurements. For a discussion of absolute sea level measurements with GPS the reader is directed to Santamaría-Gómez and Watson (2016).

c. Corrections to Reflector Height

If the reflecting surface is non-stationary during the measurements, then Larson et al. (2013b) showed that the spectral peak will be biased by an amount equal to

$$\dot{H}\,\frac{\tan e}{\dot{e}} \tag{2}$$

where $\dot{H}$ and $\dot{e}$ are the time derivatives of H and e. Of course, if we are estimating H then we do not know $\dot{H}$. Larson et al. (2013b) used an iterative solution for the two unknowns; first determining a biased estimate of H, computing $\dot{H}$ from this initial time series to provide a height correction, then applying that correction to produce the final solution. Löfgren et al. (2014a) found that the initial time series were too noisy to produce a reasonable correction, so instead they exploited the fact that the largest changes, except on days with strong meteorological forcing. are generally caused by diurnal and semi-diurnal tides. They proceeded to fit a daily sinusoidal fit using mean frequencies of the dominant tides in the diurnal and semi-diurnal bands and derived the height rates from the fit. We take a similar approach but directly solve for the height rate effect during the tidal analysis. Traditionally, in a least-squares tidal analysis one would solve for the sine and cosine coefficients (S_i, C_i) of N independent tidal frequencies ω_i (N depending on the length of the series) with known nodal amplitude factors, f_i, and equilibrium phases (including nodal corrections), ϑ_i, as given by

$$H = \sum_{i=1}^{N} C_i f_i \cos(\omega_i t + \vartheta_i) + S_i f_i \sin(\omega_i t + \vartheta_i) \tag{3}$$

Instead, we modify the analysis to account for the height rate term to give

$$\begin{aligned} H + \dot{H}\frac{\tan(e)}{\dot{e}} &= \sum_{i=1}^{N} C_i f_i \left[\cos(\omega_i t + \vartheta_i) - \omega_i \sin(\omega_i t + \vartheta_i)\frac{\tan e}{\dot{e}}\right] \\ &\quad + S_i f_i \left[\sin(\omega_i t + \vartheta_i) + \omega_i \cos(\omega_i t + \vartheta_i)\frac{\tan e}{\dot{e}}\right] \end{aligned}$$

This assumes there is no contribution to $\dot{H}$ from other influences such as meteorological forcing. This is a reasonable assumption over long periods of times required to estimate tidal constituents and where the tidal range is large; however, on any individual day where there may be an event such as a large storm surge, the residual H after tides are removed may still have a height-rate bias. Yet even when the tidal range is small this is still probably an acceptable method. For instance, at Tregde (Norway) where the tidal range is 60 cm and the full range (total water level envelope between 2012 and 2015) was 138 cm, the height-rate bias calculated from the tidal analysis has an 86% correlation with the bias calculated from the tide gauge data and the variance of the tidally-induced height-rate bias accounts for 73% of the total height-rate bias. In comparison, Brest (France), which has a maximum tidal range on the order of 7 m, has a correlation of 98% and the tidally-induced height-rate bias accounts for around 81% of the total height-rate bias.

We also apply a tropospheric correction to our data to remove a height bias at low elevation angles. We used a combination of the Vienna Mapping Function (Böhm et al. 2006) and the global pressure and temperature wet-tropospheric delay model, GPT2w (Böhm et al. 2015). We note that, to first order and for a fixed elevation range, the delay δ is a linear function of the reflector height. That is,

$$\delta = \alpha H \tag{4}$$

For Friday Harbor, α is -0.0137 m/m for a fixed elevation elevation angle range of 5° to 13°. The details of this correction and an analysis of tropospheric delay in GNSS-MR sea level studies is the subject of a paper in preparation. Santamaría-Gómez et al. (2015) speculated that tropospheric delay could have some role in the bias found at low elevation angles in their results, but they concluded it could not be the only reason. Roussel et al. (2014) used ray tracing to calculate an elevation angle correction due to geometric bending of the signal in the neutral atmosphere but they applied it only to the specular reflection point position. Santamaría-Gómez and Watson (2016) also corrected their SNR data for the geometric bending due to tropospheric delay and found a reduction in height bias.

5. Comparison of Collocated Measurements

The main results of our comparison of the two sea-level systems at Friday Harbor are discussed in this section. The topics are ordered by frequency: first an analysis of the raw GPS estimates, including extremes, then tides, then mean sea levels with averaging periods of daily and then monthly.

a. Individual sea-level estimates

Over the course of the examined ten-year period (2006–2015), the sampling rate of individual sea-level estimates from the GPS reflections is summarized by Figure 3, which displays histograms of the number of water-level estimates obtained each day and the time in minutes between successive estimates. Almost all days during the period yielded between 20 and 40 estimates with a median of 30. This number will always be necessarily limited by the number of satellite overflights.

Over the whole ten years we obtained 107688 individual GPS water level estimates. We matched each GPS water-level estimate $\eta_{\mathrm{GPS}} = -H$ with a corresponding tide-gauge value η_{TG} by linearly interpolating the 6-minute gauge data in time. Both time series were demeaned and then used to form a time series of differences

$$\Delta\eta = \eta_{\text{GPS}} - \eta_{\text{TG}}. \tag{5}$$

The standard deviation of $\Delta\eta$ was found to be 11.6 cm. The $\Delta\eta$ differences form a distribution having slightly positive skewness and kurtosis, implying somewhat more large positive differences than negative differences. A standard deviation of 11.6 cm is of course much larger than might be obtained when comparing two conventional tide gauges, which today aim for sub-cm differences (e.g. Martin Miguez et al. 2012). In practice, collocated conventional gauges routinely yield values around 1–3 cm. For example, Woodworth and Smith (2003) found a standard deviation of 1.4 cm when hourly measurements from two different gauges at Liverpool; Pérez et al. (2014) quote values between 1.3 and 3.3 cm for 5-minute data from 17 pairs of gauges located along the Spanish coast.

Figure 4 displays a variation on the so-called Van de Casteele diagram (Martin Miguez et al. 2008), which has been found useful in comparison tests of tide gauges, since it can indicate scale problems, timekeeping errors, and other problems (Pérez et al. 2014). Usually a time series of a few days, possibly longer, is plotted as a continuous curve with abscissa $\Delta\eta$ and ordinate η; here we have computed a two-dimensional density of the corresponding pairs $(\Delta\eta, \eta_{\text{TG}})$ for the entire ten-year period. The skewed distribution of $\Delta\eta$ in Figure 4a is evident, mostly for $\eta > 0$ for which the spread in $\Delta\eta$ is skewed toward positive values. The central axis of the distribution, however, appears very close to the $\Delta\eta$ zero line, so unlike some skewed distributions, a scale problem may or may not be indicated. We therefore computed a least-squares fit to the relation

$$\eta_{\text{GPS}} = \beta\,\eta_{\text{TG}} + c$$

and found $\beta = 1.0084 \pm 0.0010$. A factor of 1.0084 (i.e., a possible scale error of 8.4‰) is much smaller than corresponding errors found by Pérez et al. (2014); among their 17 pairs of gauges they found errors between $-79‰$ and $+26‰$, the largest (at Ibiza) was attributed to a pressure gauge affected by seasonal variations in seawater density. Woodworth and Smith (2003) obtained $6‰$ when comparing a radar and pressure tide gauge, which they also attributed primarily to errors in seawater density required for the latter.

Figure 4b shows the variance of $\Delta\eta$ as a function of the elevation η. Over the elevation range $\eta \in (-150, 120)$ cm, where most of the data lie, there is a clear tendency for the variance of $\Delta\eta$ to rise with increasing water level. For example, the variance of $\Delta\eta$ near $\eta = 100$ cm is about 145 cm^2 whereas the variance near $\eta = -150$ cm is about 118 cm^2. Assuming the tendency is unrelated to the Aquatrak gauge, we conclude that the GPS water-level estimates are more accurate for lower water levels. The effect is undoubtedly real, because more reflection cycles are present in low-tide data (large H) than high-tide data (small H).

Although not especially germane to the topic at hand, readers may notice that the distribution of η in Figure 4a is highly asymmetric, with the peak occurring about 60 cm above mean sea level. The coefficient of skewness for η is -0.625, one of the most negative coefficients from a tide gauge that we are aware of. The cause stems from the fairly unusual tides at Friday Harbor, where the three largest constituents are K_1, O_1, and M_2, with K_1 the largest (see next section). These three constituents happen to be phase-locked, with frequencies satisfying the relationship $\omega_K + \omega_O = \omega_M$, and their phases are such that whenever K_1 and O_1 combine to form either high water or low water, M_2 always acts to lower the sum (see discussion by Woodworth et al. 2005, Section 3).

b. Extremes

Sea level extremes as measured by tide gauges are of the greatest practical importance (Pugh 1987). It is therefore of interest to understand how the statistics of extremes as seen in the GPS data compare with those from the tide gauge. Probabilities for the rare and largest flood events, needed for civil planning purposes, generally require more than ten years of data (Arns et al. 2013), but comparison of some 10-year statistics is nevertheless still enlightening.

Figure 5 compares GPS and tide-gauge daily maximum and daily minimum sea levels as observed over the whole ten years. Relative to the tide-gauge data, the GPS maxima are seen to be biased high, with a median offset of $+8.4$ cm. The GPS minima are biased low, with a median offset of -5.3 cm. Neither of these results is surprising. Because the GPS measurement errors are much larger than the tide gauge errors, extracting the extreme values from each day will nearly always result in a GPS maximum biased high and a minumum biased low. The RMS differences in the extremes are 13.2 cm and 10.9 cm, respectively, which are comparable to the high and low-water variances shown in Figure 4b.

Sea level extremes are often characterized by annual percentiles of measured water levels, typically in the interval 99% to 99.9%. Woodworth and Blackman (2004), in their search for systematic changes in extreme high waters, prefer the 99th percentile level; the 99.9th level can be impacted by a small number of incorrect measurements, although the 99th level might significantly underestimate the true extreme. We show both in Figure 6.

Unlike the GPS daily maxima in Figure 5, the GPS percentiles in Figure 6 are generally lower than the tide-gauge values, at least for the 99.9th percentile. The 99th percentiles are more comparable, although with slightly less year-to-year variability in the GPS. We have determined that the differences in the 99.9th case are caused mostly by the non-uniform sampling in the GPS time series, and less so by its inherently higher measurement noise. By resampling the tide-gauge data at the times of the GPS measurements, we can obtain a percentile time series comparable to the GPS series of Figure 6. Thus, the occasional coarser sampling in the GPS time series, as documented in Figure 3b—in contrast to the uniform 6-minute sampling of the tide gauge—evidently leads to missing the peak values of some high-water extremes.

Note that as more geodetic receivers are deployed which track all GNSS systems, not just GPS, this sampling problem will be considerably reduced. The errors in the sea level extremes will then be solely a function of the measurement errors in the systems.

c. Tide estimates

In our analysis of the ocean-tide signals extracted from GPS-based water-level measurements at Kachemak Bay, Alaska (Larson et al. 2013b), a site of extraordinarily large tides, we obtained results that appeared nominally accurate. However, the closest tide gauge was 30 km away, and in light of the complicated macrotidal environment, it was unclear whether observed discrepancies were due to instrumentation or to real changes in the tide over 30 km. For the collocated instruments at Friday Harbor there is no such uncertainty.

As noted above, the tidal regime at Friday Harbor is somewhat unusual since it is predominantly diurnal. The largest constituent is K_1, at frequency 1 cycle per sidereal day. Owing to the shallow-water location there is also a large number of nonlinear compound tides. Tidal analysis of the tide gauge data, followed up by a spectral analysis of the residuals, reveals 102 tidal constituents with amplitudes above 1 mm. There are pronounced tidal lines up through species 10 (i.e., ten cycles per day), but above species 12 the lines become insignificant. In our analysis of the full ten years of data, we accounted for 131 tidal constituents. In our analyses of annual data, we reduced this to 112 constituents, since some of the constituents included in the full set cannot be separated except in multi-year time series.

Before discussing the main tide results, it is worth noting the effect on estimated GPS tides of the additional enhancements to the GPS processing discussed above in Section 4c. Figure 7 compares GPS and tide-gauge estimates with and without correcting for motion $\dot{H}$ of the reflecting surface and with and without the wet-troposphere correction. With only a few exceptions the estimates with the corrections improves the tide-gauge comparisons. The tropospheric correction is especially useful for K_1. For the remainder of this paper, we describe only results from the fully corrected data.

A selected set of final estimated tidal constituents, computed from the whole ten years of tide-gauge and GPS data, is tabulated in Table 1. The rightmost column of the table gives the absolute value of the complex difference between the tide-gauge and GPS coefficients; that is, it tabulates $|A_1 e^{-iG_1} - A_2 e^{-iG_2}|$, for amplitudes A_i and Greenwich phase lags G_i. For the most part, the agreement between the tidal coefficients is sub-cm, with the largest differences occurring for the largest constituents. There are no evident systematic differences in terms of the GPS being either consistently higher or lower than the tide gauge. The discrepancy at K_1 is fairly large, but again this may simply reflect its large amplitude; the GPS and tide-gauge amplitudes for K_1 are nearly identical, and the discrepancy arises from a 1° difference in phase. The reasonably good agreement at K_1 is noteworthy for a GPS-based system, since that tidal frequency is essentially identical to the orbital frequency of the GPS satellite constellation (Agnew and Larson 2007). In the same way that GPS positioning can sometimes be prone to K_1 errors (e.g. King et al. 2008) because the satellite geometry (and thus K_1) is correlated with (say) multipath error, GPS reflection data could be similarly correlated with geometrical errors. It is thus reassuring that leakage of such effects into K_1 appears here to be small.

The robustness of the GPS tidal solutions may be further assessed by examining the year-to-year consistency of annual estimates. These are displayed in Figure 8 for the three largest constituents. Except for the K_1 constituent, the scatter in the GPS estimates is quite comparable to the scatter in the tide-gauge estimates. For the K_1, O_1 and M_2 constituents, the standard deviations of the ten tide-gauge estimates are: 0.41, 0.31, 0.37 cm, respectively, while the standard deviations of the ten GPS estimates are: 0.72, 0.40, 0.32 cm.

The most discrepant result in Table 1 is actually for the very small constituent S_1 where the difference of 1.37 cm is nearly as large as the GPS-based amplitude of 1.6 cm. The frequency of S_1 is 1 cycle per mean solar day, coincident with the mean daily heating and cooling cycle, so it is unsurprising to find that measurements of S_1 can be plagued by systematic errors (Ray and Egbert 2004). In the present case, it seems possible, if not likely, that the disagreement in our S_1 estimates stems mostly from errors in the Aquatrak tide gauge. Inadequately compensated changes in temperature within the acoustic sound channel are a known error source in these gauges (Porter and Shih 1996; Hunter 2003). In contrast, thermal effects in the GPS instrumentation are likely to be small (e.g. Munekane 2013). Moreover, since the daily heating/cooling cycle is likely to have a significant seasonal dependence, similar errors could arise at the P_1 and K_1 frequencies, which are both 1 cpy from the S_1 frequency. The discrepancy at K_1 has already been noted, and the discrepancy at P_1 does appear slightly inflated; for example, it is larger than that of M_2 even though the M_2 amplitude is more than twice the P_1 amplitude. Finally, the difference indicated in Table 1 for the annual cycle Sa is also somewhat pronounced, and this could similarly arise from temperature problems.

In summary, it appears that tidal analysis of the (unequally spaced) GPS time series is capable of yielding results comparable to analysis of standard hourly tide gauge data. For particular tidal constituents prone to instrumental thermal problems, the GPS may possibly be superior.

d. Daily mean sea levels

Once tidal coefficients are determined, they can be used to remove short-period tidal variability from the GPS time series. The removal will not be quite as effective as de-tiding with a standard "tide killing" filter applied to an equally spaced time series (e.g. Pugh 1987), and the removal will be even less perfect if only a short time series is available for the tidal analysis (although iteration can be done as the time series lengthens). Nonetheless, the procedure should be adequate for subsequently forming mean sea levels. For this de-tiding step, only the short-period tides (i.e., of periods diurnal and shorter) are removed, since long-period tides are traditionally retained in a time series of daily mean sea levels.

Daily sea levels may be determined from the de-tided GPS reflection data by the straightforward method of computing simple averages of all water levels obtained during each day. This method has the advantage of simplicity, but it can be improved upon. We here used an approach similar to that used for the Kachemak Bay work (Larson et al. 2013b). Nominally hourly sampling is formed by averaging within a running window of size 6 h (a calculation equivalent to applying an order-0 Savitzky-Golay filter; experiments with an order-1 filter did not appear to yield greater accuracy); the output is passed through a low-pass filter with a half-power cutoff at 60 h, from which daily means are formed. The last step is identical to the procedure employed by the University of Hawaii Sea Level Center to form daily means from hourly tide-gauge data.

For the ten-year period of analysis, the rms difference between the GPS and tide-gauge daily means is 2.07 cm. Figure 9 shows two full-year comparisons, year 2008 having the best rms agreement and year 2010 having the worst. It is probably no coincidence that 2010 also had the most gaps in the GPS time series.

Figure 10 shows spectra of the GPS and tide-gauge daily sea levels as well as their spectral coherence. Close examination of the spectra (notably in the zoomed inset) shows a tendency for slightly larger variance in the GPS data. The coherence remains close to 1 at all periods longer than about 10 days, dropping off only at the shortest periods.

e. Monthly mean sea level

Time series of the tide-gauge and GPS-based monthly mean sea levels are shown in Figure 11. The rms difference between the two time series is 1.28 cm. By way of comparison, Pérez et al. (2014), analyzing their 17 pairs of tide gauges, found rms differences in monthly means between 0.25 cm and 1.99 cm, with most values less than 1 cm. We conclude that the accuracy of the GPS-based monthly means at Friday Harbor is nearly comparable to what can be achieved with standard operating tide gauges.

6. Conclusions

From our analysis of ten years of L1 SNR GPS data at Friday Harbor we find that individual water-level estimates have an rms error of about 12 cm. The errors are slightly reduced at lower water levels, slightly raised at higher levels. Forming daily mean sea levels significantly reduces the error, so that the rms difference with the Aquatrak tide gauge was 2.1 cm, and some part of this difference must owe to errors in the Aquatrak system. Forming monthly means further reduces the rms differences to 1.3 cm.

Thus, it is clear that a standard geodetic-quality GPS receiver, properly sited with a sufficiently open view of the sea, can act as a serendipitous tide gauge, supplying useful sea-level information for a number of applications. It is worth emphasizing that no part of the instrumentation sits in the water, so the kind of regular maintenance needed for most tide gauges is eliminated. Moreover, the difficult task of tying the sea-level measurements into a well-defined terrestrial reference frame becomes automatic. Indeed, for studies of global mean sea level, the problem of vertical land motion at tide gauges is a critical one (e.g. Wöppelmann and Marcos 2016). This has motivated an international campaign to deploy GPS receivers (or similar geodetic instrumentation) at a large global network of tide gauges (Schöne et al. 2009). Were such geodetic GPS stations properly sited near the shore, they could also provide important redundancy for the primary tide gauges.

It would be unrealistic to conclude from our study that GPS reflection technology can completely replace conventional tide gauges. For example, the Global Sea Level Observing System (GLOSS) requirements for tide gauges call for 1 cm precision in individual sea-level readings and a sampling rate of 1 hour or better (IOC, 2006, Appendix 1). The GPS reflection measurements described here cannot meet these requirements. The precision of individual water-level estimates is much worse than 1 cm. And although the sampling rate is often much better than 1 hour (see Figure 3b), it is necessarily limited by the number of satellite overflights, precision of the L1 SNR data, and by the geometry of the site. The simplest way to increase the number of overflights is to use signals from non-GPS satellite constellations (GLONASS, GALILEO, BEIDOU) and more frequencies, such as L2C and L5 (Löfgren and Haas 2014b; Strandberg et al. 2016). More advanced SNR analysis techniques have also recently been proposed. These methods have been tested at two sites (in Sweden and Australia) and are significantly more precise than using the Lomb-Scargle periodogram alone (Strandberg et al. 2016). Another limitation for the GPS reflection method is the roughness of the surface. One metric we can use to evaluate how well the method works for rough surfaces is wind speed (Löfgren and Haas 2014b). In that study, reflection measurements were successful up to wind speeds of 17.5 m/sec. However, this is not an upper limit, as no GPS data were collected in conditions with larger wind speeds.

If considerations of water reflections are taken into account, it is straightforward to improve the precision of a GPS tide gauge by either raising the height of the antenna and/or moving the antenna closer to the shore. On many of the Great Lakes, for example, the GPS antenna has been deployed on the end of a pier, significantly improving the reflection zone (Michael Craymer, personal communication, August 30, 2015). Moreover, for our work reported here, the emphasis has been on GPS data because the instrument we used tracked only GPS satellites until mid-2015. The Friday Harbor site currently tracks all GNSS signals. In coming years that would mean perhaps as many as 120 satellite signals. While this might not improve the precision of an individual reflector height measurement, it would certainly provide better temporal sampling and more accurate mean sea levels. A GNSS tide gauge might then be useful for the study of short-period phenomena like seiches or tsunamis.

Acknowledgments. The tide gauge data at Friday Harbor were obtained from NOAA, http://tidesandcurrents.noaa.gov/waterlevels.html?id=9449880. Monthly tide gauge data, used for further comparisons, were obtained from the Permanent Service for Mean Sea Level. GPS data from SC02 were provided by the EarthScope Plate Boundary Observatory via UNAVCO (http://pbo.unavco.org). We thank UNAVCO staff for maintaining SC02. KL's work on reflections has been supported by the National Science Foundation (AGS 1449554). RR's work is supported by the Sea Level Change program of the National Aeronautics and Space Administration. SW's work is supported by NERC National Capability funding to the NOC Marine Physics and Ocean climate directorate. Permanent Service for Mean Sea Level data were retrieved from http://www.psmsl.org/data/obtaining.

References

Agnew, D. C., and K. M. Larson, 2007: Finding the repeat times of the GPS constellation. *GPS Solutions*, **11 (1)**, 1–8.

Anderson, K. D., 2000: Determination of water level and tides using interferometric observations of GPS signals. *J. Atmos. Oceanic Tech.*, **17 (8)**, 1118–1127.

Arns, A., T. Wahl, I. D. Haigh, J. Jensen, and C. Pattiaratchi, 2013: Estimating extreme water level probabilities: A comparison of the direct methods and recommendations for best practise. *Coastal Eng.*, **81**, 51–66.

Benton, C. J., and C. N. Mitchell, 2011: Isolating the multipath component in GNSS signal-to-noise data and locating reflecting objects. *Radio Science*, **46 (6)**, RS6002, doi:10.1029/2011RS004767.

Blewitt, G., C. Kreemer, W. C. Hammond, and J. Gazeaux, 2016: MIDAS robust estimator for accurate GPS station velocities with step detection. *J. Geophys. Res.: Solid Earth*, **121**, 2054–2068.

Böhm, J., G. Möller, M. Schindelegger, G. Pain, and R. Weber, 2015: Development of an improved empirical model for slant delays in the troposphere (GPT2w). *GPS Solutions*, **19 (3)**, 433–441.

Böhm, J., B. Werl, and H. Schuh, 2006: Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-range Weather Forecasts operational analysis data. *J. Geophys. Res.*, **111 (B2)**, B02 406.

Holgate, S. J., and Coauthors, 2013: New data systems and products at the Permanent Service for Mean Sea Level. *J. Coastal Res.*, **29**, 493–504.

Hunter, J. R., 2003: On the temperature correction of the Aquatrak acoustic tide gauge. *J. Atmos. Oceanic Tech.*, **20 (8)**, 1230–1235.

King, M. A., C. S. Watson, N. T. Penna, and P. J. Clarke, 2008: Subdaily signals in GPS observations and their effect at semiannual and annual periods. *Geophys. Res. Lett.*, **35**, L03 302, doi:10.1029/2007GL032252.

Larson, K. M., J. J. Braun, E. E. Small, V. U. Zavorotny, E. D. Gutmann, and A. L. Bilich, 2010: GPS multipath and its relation to near-surface soil moisture content. *IEEE JSTARS*, **3 (1)**, 91–99.

Larson, K. M., E. D. Gutmann, V. U. Zavorotny, J. J. Braun, M. W. Williams, and F. G. Nievinski, 2009: Can we measure snow depth with GPS receivers? *Geophys. Res. Lett.*, **36 (17)**.

Larson, K. M., J. S. Löfgren, and R. Haas, 2013a: Coastal sea level measurements using a single geodetic GPS receiver. *Adv. Space Res.*, **51**, 1301–1310.

Larson, K. M., R. D. Ray, F. G. Nievinski, and J. T. Freymueller, 2013b: The accidental tide gauge: a GPS reflection case study from Kachemak Bay, Alaska. *IEEE Geosci. Remote Sens. Lett.*, **10 (5)**, 1200–1204.

Larson, K. M., E. E. Small, E. D. Gutmann, A. L. Bilich, J. J. Braun, and V. U. Zavorotny, 2008: Use of GPS receivers as a soil moisture network for water cycle studies. *Geophys. Res. Lett.*, **35 (24)**.

Löfgren, J. S., R. Haas, H.-G. Scherneck, and M. S. Bos, 2013: Three months of local sea level derived from reflected GNSS signals. *Rad. Sci.*, **46**.

Löfgren, J. S., R. Haas, and H.-G. Scherneck, 2014: Sea level time series and ocean tide analysis from multipath signals at five GPS sites in different parts of the world. *J. Geodynam.*, **80**, 66–80.

Löfgren, J. S., and R. Haas, 2014: Sea level measurements using multi-frequency GPS and GLONASS observations. *EURASIP J. Adv. Signal Proc.*, **2014**, 1–13, doi:10.1186/1687-6180-2014-50.

Martin Miguez, B., L. Testut, and G. Wöppelmann, 2008: The Van de Casteele test revisited: an efficient approach to tide gauge error characterization. *J. Atmos. Oceanic Tech.*, **25 (7)**, 1238–1244.

Martin Miguez, B., L. Testut, and G. Wöppelmann, 2012: Performance of modern tide gauges: towards mm-level accuracy. *Sci. Marina*, **76 (S1)**, 221–228.

Munekane, H., 2013: Sub-daily noise in horizontal GPS kinematic time series due to thermal tilt of GPS monuments. *J. Geodesy*, **87**, 393–401.

Park, J., R. Heitsenrether, and W. Sweet, 2014: Water level and wave height estimates at NOAA tide stations from acoustic and microwave sensors. *J. Atmos. Oceanic Tech.*, **31 (10)**, 2294–2308.

Pérez, B., A. Payo, D. López, P. L. Woodworth, and E. Alvarez Fanjul, 2014: Overlapping sea level time series measured using different technologies: an example from the REDMAR Spanish network. *Nat. Hazards Earth Syst. Sci.*, **14**, 589–610.

Porter, D. L., and H. H. Shih, 1996: Investigations of temperature effects on NOAA's next generation water level measurement systems. *J. Atmos. Oceanic Tech.*, **13**, 714–725.

Pugh, D. T., 1987: *Tides, Surges and Mean Sea-Level*. Wiley & Sons, Chichester.

Ray, R. D., and G. D. Egbert, 2004: The global S_1 tide. *J. Phys. Oceanogr.*, **34**, 1922–1935.

Roussel, N., F. Frappart, G. Ramillien, J. Darrozes, C. Desjardins, P. Gegout, F. Pérosanz, and R. Biancale, 2014: Simulations of direct and reflected wave trajectories for ground-based GNSS-R experiments. *Geosci. Model Dev.*, **7**, 2261–2279.

Santamaría-Gómez, A., and C. Watson, 2016: Remote leveling of tide gauges using GNSS reflectometry: case study at Spring Bay, Australia. *GPS Solutions*, in press, doi:10.1007/s10291-016-0537-x.

Santamaría-Gómez, A., C. Watson, M. Gravelle, M. King, and G. Wöppelmann, 2015: Levelling co-located GNSS and tide gauge stations using GNSS reflectometry. *J. Geodesy*, **89 (3)**, 241–258.

Schöne, T., N. Schön, and D. Thaller, 2009: IGS tide gauge benchmark monitoring pilot project (TIGA): scientific benefits. *J. Geodesy*, **83**, 249–261.

Strandberg, J., T. Hobiger, and R. Haas, 2016: Improving GNSS-R sea level determination through inverse modeling of SNR data. *Rad. Sci.*, **51**, 1286-1296.

Woodworth, P. L., and D. E. Smith, 2003: A one year comparison of radar and bubbler tide gauges at Liverpool. *Int. Hydrogr. Rev.*, **4 n.s. (3)**, 1–8.

Woodworth, P. L., and D. L. Blackman, 2004: Evidence for systematic changes in extreme high waters since the mid-1970s. *J. Climate*, **17**, 1190–1197.

Woodworth, P. L., D. L. Blackman, D. T. Pugh, and J. M. Vassie, 2005: On the role of diurnal tides in contributing to asymmetries in tidal probability distribution functions in areas of predominantly semidiurnal tide. *Estuar. Coastal Shelf Sci.*, **64**, 235–240.

Wöppelmann, G., and M. Marcos, 2016: Vertical land motion as a key to understanding sea level change and variability. *Rev. Geophys.*, **54**, 64–92, doi:10.1002/2015RG000502.

LIST OF TABLES

TABLE 1. Estimated amplitudes A and phase lags G of selected tidal constituents, based on data collected during 2006–2015.

	Acoustic gauge		GPS		
Tide	A (cm)	G	A (cm)	G	Diff (cm)
Sa	6.1	274.8°	5.8	277.6°	0.37
Ssa	1.5	227.7°	1.6	220.1°	0.21
Mf	2.0	168.2°	2.0	162.4°	0.20
Q_1	7.4	250.0°	7.5	249.9°	0.13
O_1	43.4	258.1°	44.0	258.6°	0.78
P_1	23.6	278.7°	23.1	278.0°	0.54
S_1	2.6	31.2°	1.6	59.2°	1.37
K_1	76.0	280.0°	76.0	279.0°	1.33
J_1	4.0	311.6°	4.0	310.5°	0.08
N_2	12.1	342.4°	12.0	343.1°	0.15
M_2	56.0	10.5°	56.4	10.2°	0.50
S_2	13.3	36.0°	13.2	34.9°	0.25
MK_3	1.2	26.8°	1.2	33.9°	0.16
M_4	1.7	121.2°	1.5	121.1°	0.17
MS_4	1.0	131.4°	0.8	131.4°	0.17
M_6	0.5	236.0°	0.4	255.1°	0.18

LIST OF FIGURES

FIG. 1. Friday Harbor GPS station SC02.

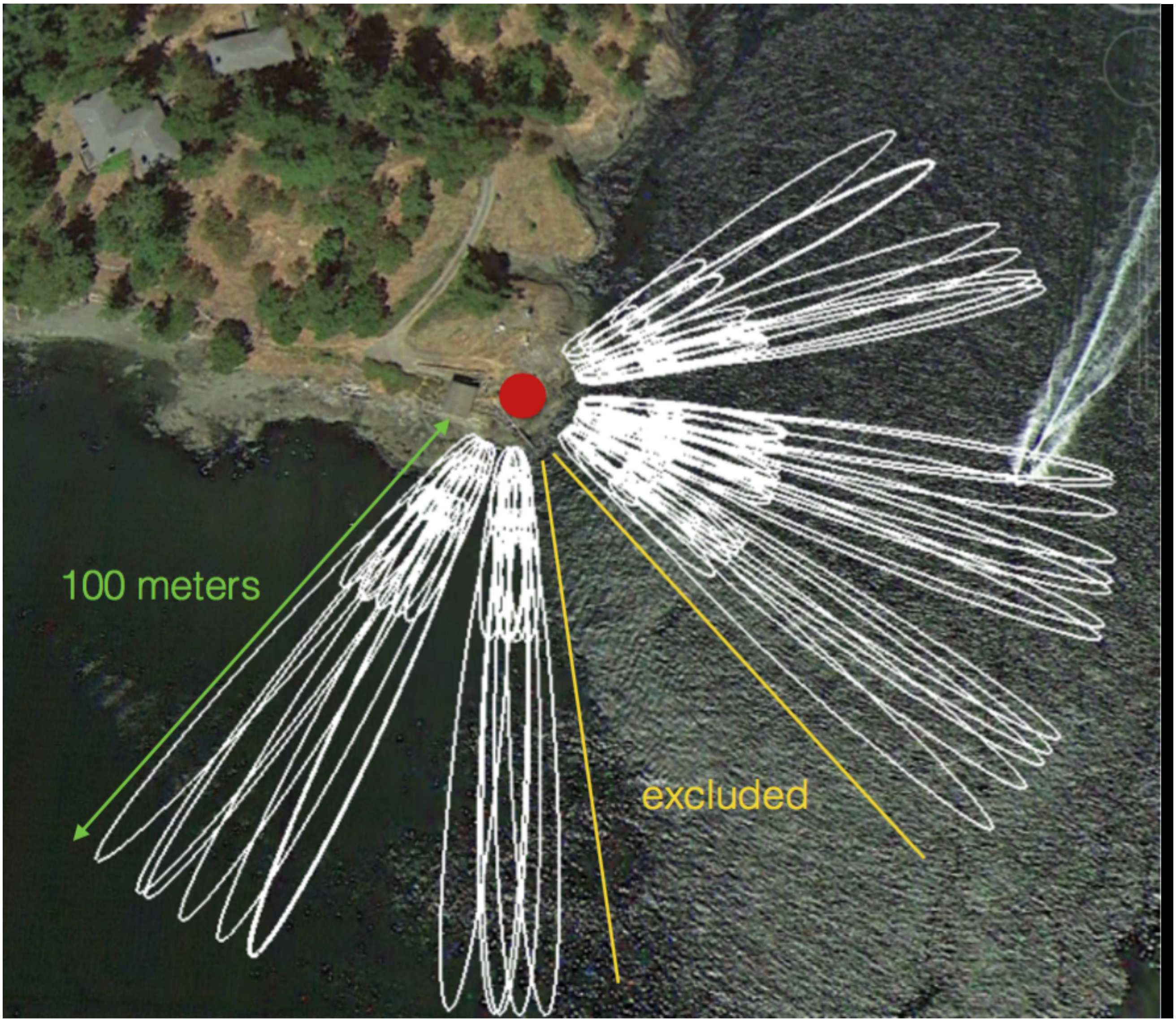


FIG. 2. Location of the SC02 GPS site at Friday Harbor, seen as red circle. The Fresnel zones for reflector height of 5 meters and elevation angles 5°, 9°, and 13° are shown in white for the satellite tracks used in this study. At far right is the wake of a small boat. The tide gauge sits about 345 m to the west, at the end of a long pier. Image obtained from Google Maps.

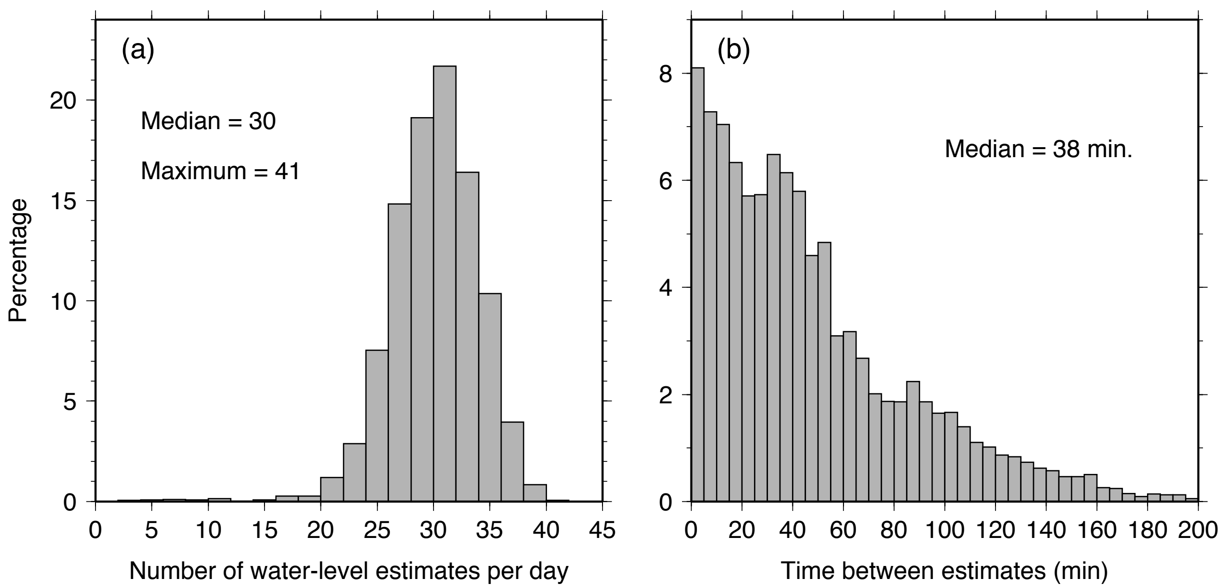


FIG. 3. (a) Histogram of number of GPS-based sea-level estimates each day, obtained over the ten-year period 2006–2015. (b) Histogram of time (in minutes) between successive GPS sea-level estimates.

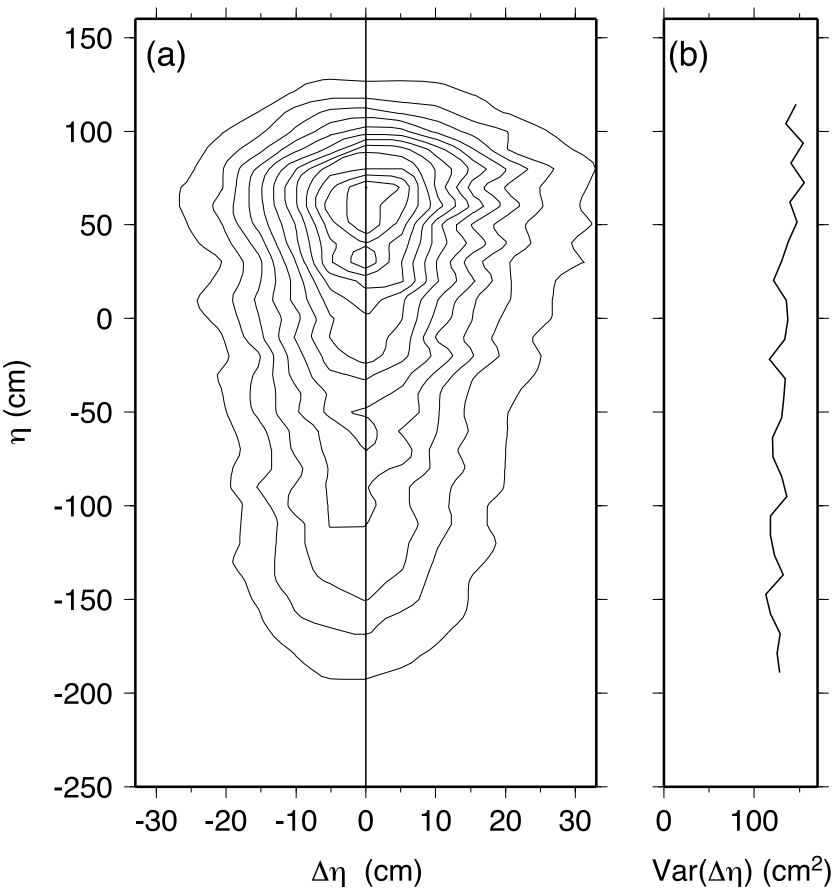


FIG. 4. (a) Van de Casteele diagram as a two-dimensional density of the differences $\Delta\eta$ between the GPS and Aquatrak water level measurements as a function of the water level η. Contour levels are linear, in arbitrary units. The mean of η is set to zero. (b) Variance of $\Delta\eta$ as a function of water level. There is a slight tendency for reduced variance with lower water levels, suggesting that the GPS estimates are likely more accurate for lower water levels.

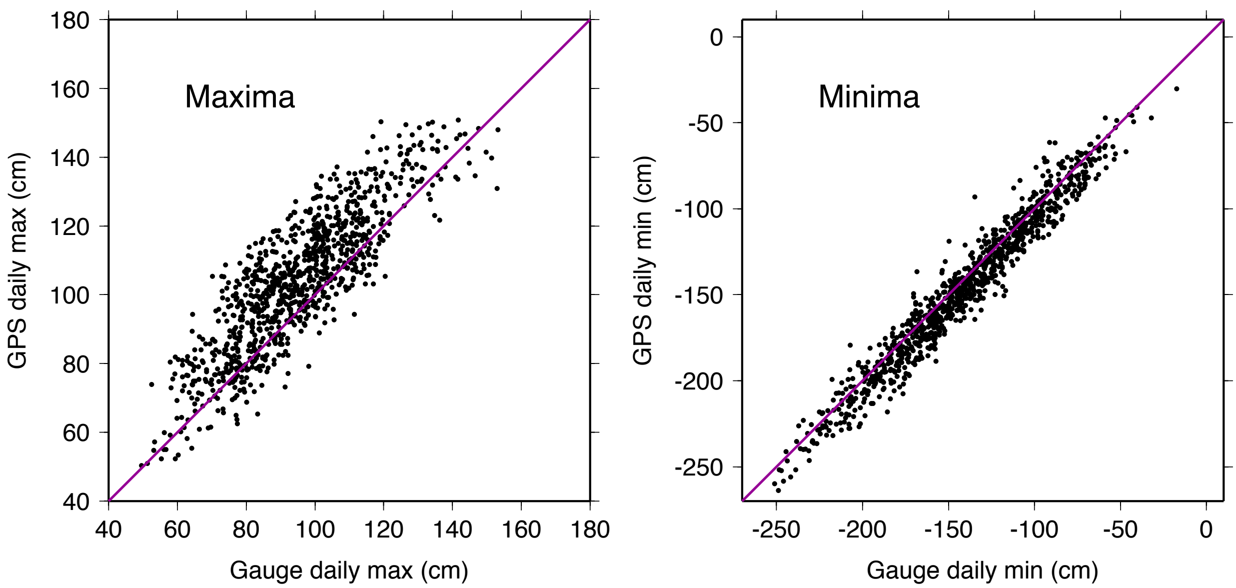


FIG. 5. Comparison of daily sea level extremes as measured by the GPS and tide gauge. Owing to its random measurement noise of $\sim$12 cm, the GPS daily maxima are biased high and the daily minima are biased (slightly) low. Note the scale difference between the two panels.

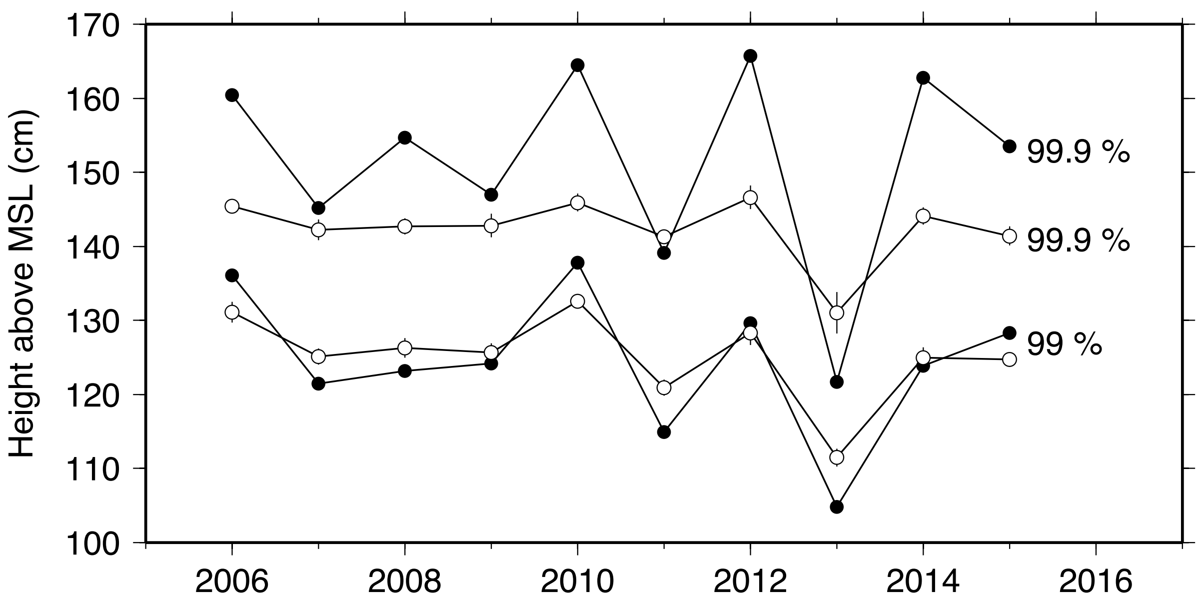


FIG. 6. Annual high-water percentile time series from the Friday Harbor tide gauge (closed circles) and from our GPS estimates (open circles). The differences in the 99.9 percentile series are caused mainly by the non-uniform sampling of the GPS data, which causes some water-level peaks to be missed.

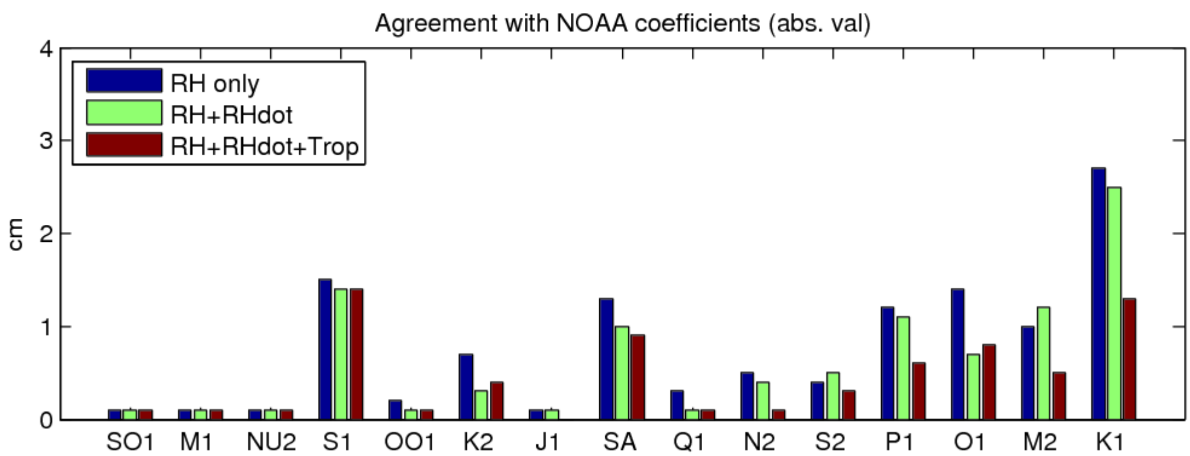


FIG. 7. Absolute differences between GPS and tide-gauge tide estimates, as function of improvements made in GPS processing (see Section 4c). Constituents are displayed in order of increasing amplitude, left to right.

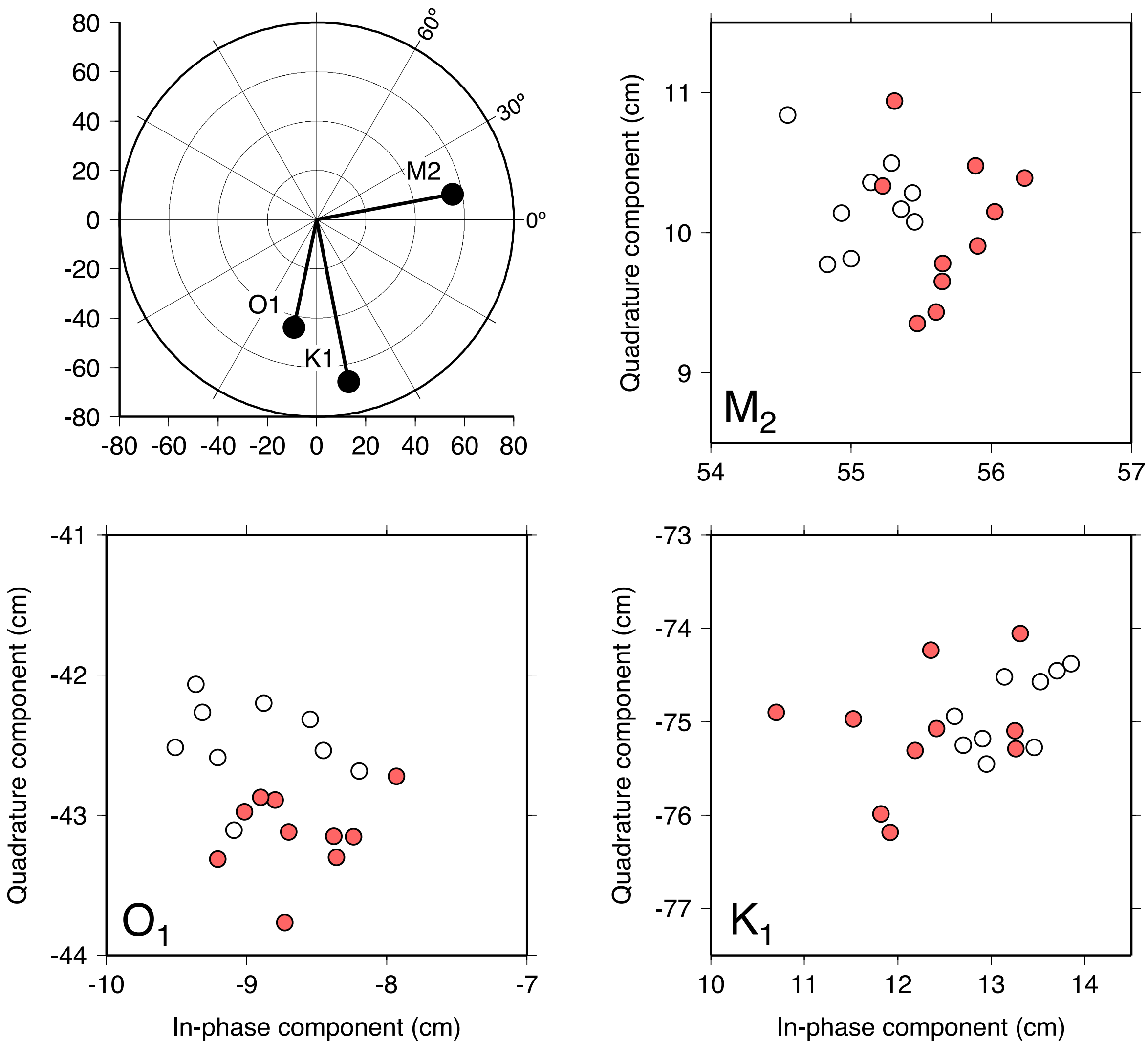


FIG. 8. Upper left panel shows approximate amplitudes and phase lags of the three largest ocean tides at Friday Harbor. Other panels are 'zoom' views showing annual estimates from the tide gauge data (open circles) and from the GPS reflection data (red circles). For the K_1, O_1 and M_2 constituents, the standard deviations of the annual tide gauge estimates are: 0.41, 0.31, 0.37 cm; the standard deviations of the annual GPS estimates are: 0.72, 0.40, 0.32 cm. Thus, the scatter of estimates is comparable except for the GPS estimates of K_1. Mean amplitudes and phases over the whole ten year period are tabulated in Table 1.

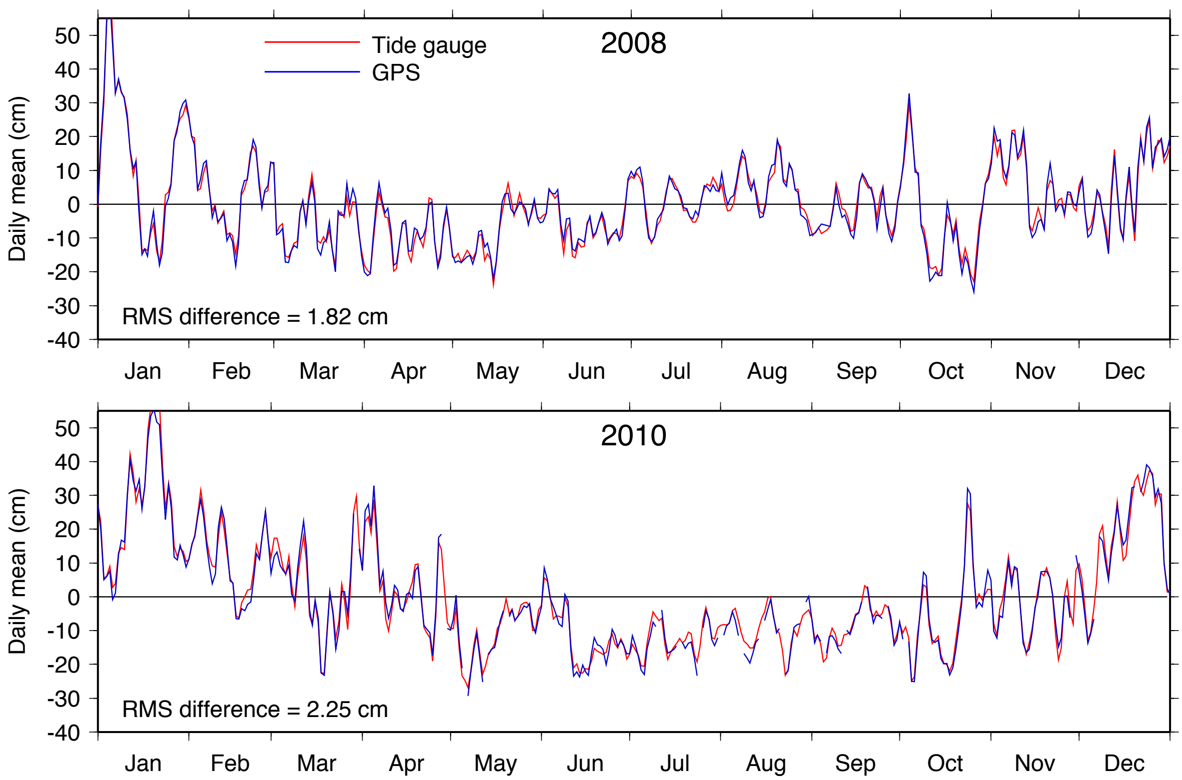
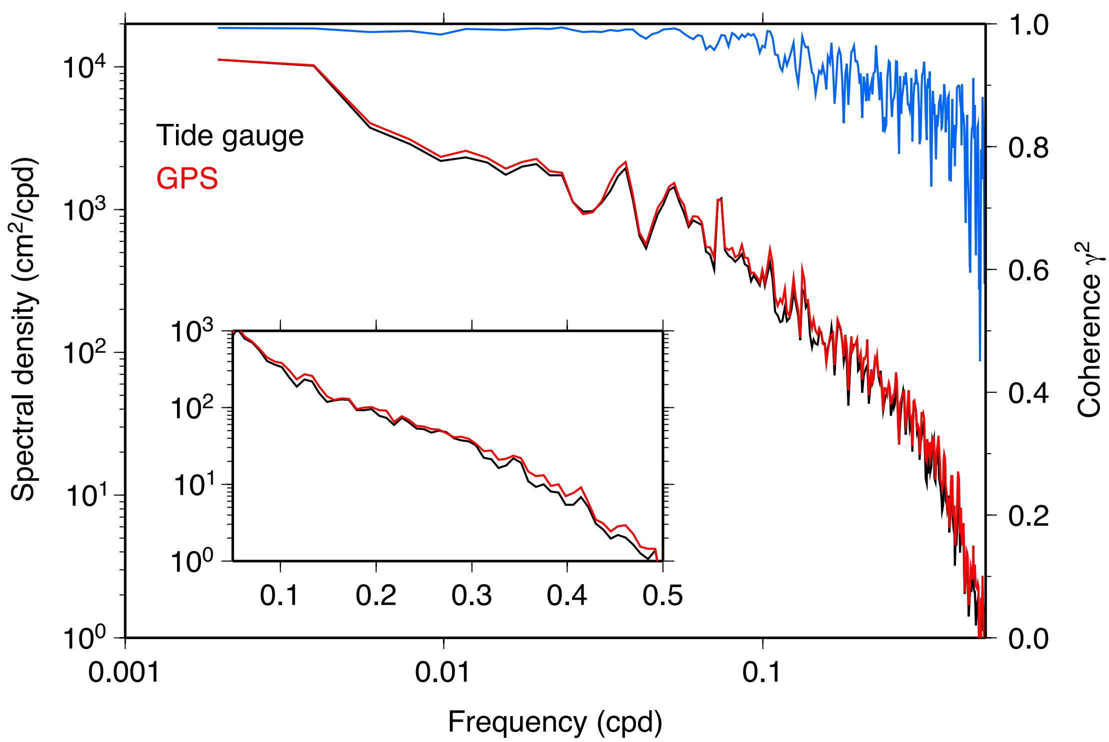


FIG. 9. Daily mean sea levels for 2008 and 2010. Red lines mark daily means deduced from the Friday Harbor tide gauge; blue lines mark means deduced from GPS reflections. Year 2008 had the best and 2010 the worst agreement between the two time series.



623 FIG. 10. Spectra of daily mean sea levels from the Friday Harbor tide gauge and from the GPS analysis. Blue line is the coherence γ^2, with ordinate axis at left.

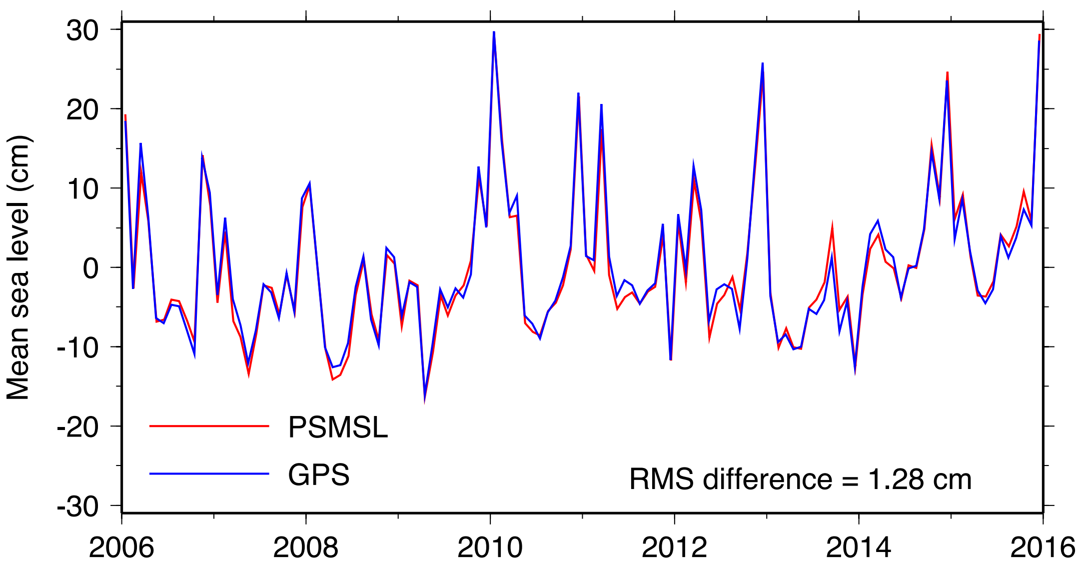


FIG. 11. Monthly mean sea levels for 2006–2015. Red line is from the Permanent Service for Mean Sea Level. (Holgate et al. 2013). Blue line is computed from the GPS reflection data. A least-squares estimate of the difference in trends of these two series is 0.8 ± 0.5 mm/y (1-σ), assuming AR(1) noise.