




# Comparing Methods and Defining Practical Requirements for Extracting Harmonic Tidal Components from Groundwater Level Measurements

Daniel Schweizer<sup>1</sup>  · Vincent Ried<sup>1</sup> ·  
Gabriel C. Rau<sup>1,2</sup> · Jonathan E. Tuck<sup>3</sup> ·  
Petre Stoica<sup>4</sup>

Received: 8 September 2020 / Accepted: 21 December 2020  
© The Author(s) 2021

**Abstract** The groundwater pressure response to the ubiquitous Earth and atmospheric tides provides a largely untapped opportunity to passively characterize and quantify subsurface hydro-geomechanical properties. However, this requires reliable extraction of closely spaced harmonic components with relatively subtle amplitudes but well-known tidal periods from noisy measurements. The minimum requirements for the suitability of existing groundwater records for analysis are unknown. This work systematically tests and compares the ability of two common signal processing methods, the discrete Fourier transform (DFT) and harmonic least squares (HALS), to extract harmonic component properties. First, realistic conditions are simulated by analyzing a large number of synthetic data sets with variable sampling frequencies, record durations, sensor resolutions, noise levels and data gaps. Second, a model of two real-world data sets with different characteristics is validated. The results reveal that HALS outperforms the DFT in all aspects, including the ability to handle data gaps. While there is a clear trade-off between sampling frequency and record duration, sampling rates should not be less than six samples per day and records should not be shorter than 20 days when simultaneously extracting tidal constituents. The accuracy

---

✉ Daniel Schweizer  
daniel.schweizer@kit.edu

<sup>1</sup> Institute of Applied Geosciences (AGW), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

<sup>2</sup> School of Civil and Environmental Engineering, The University of New South Wales, Sydney, Australia

<sup>3</sup> Department of Electrical Engineering, Stanford University, Stanford, USA

<sup>4</sup> Information Technology Department, Division of Systems and Control, Uppsala University, Uppsala, Sweden

of detection is degraded by increasing noise levels and decreasing sensor resolution. However, a resolution of the same magnitude as the expected component amplitude is sufficient in the absence of excessive noise. The results provide a practical framework to determine the suitability of existing groundwater level records and can optimize future groundwater monitoring strategies to improve passive characterization using tidal signatures.

**Keywords** Tidal subsurface analysis · Tidal constituents · Signal analysis · Harmonic least squares · Non-uniform sampling

## 1 Introduction

Reliable detection and extraction of harmonic components embedded in measurements is crucial for a range of different applications in the geosciences. These include, but are not limited to, the prediction of ocean tides (Pawlowicz et al. 2002), investigating the propagation of seismic waves (Tary et al. 2014), identifying oscillations in climate signals (Ghil et al. 2002) and quantifying water flux in near-surface sediments using temperature measurements (Wörman et al. 2012; Rau et al. 2014; Halloran et al. 2016).

One emerging application is the characterization of the subsurface using the groundwater response to Earth and atmospheric tides. The impacts of astronomical tides on groundwater systems have long been observed, and methods have been developed and applied to characterize subsurface systems (Bredehoeft and Papaopulos 1965; Hsieh et al. 1987; Van der Kamp and Gale 1983; Xue et al. 2016). The advantage of such techniques is that they are passive (Allègre et al. 2016) and thus can be widely applied to existing data sets (McMillan et al. 2019). This approach enables the quantification of subsurface hydro-geomechanical properties such as hydraulic conductivity, specific storage or compressibility, and is termed tidal subsurface analysis (TSA) (McMillan et al. 2019). However, the first step towards property quantification is the extraction of amplitudes and phases of distinct harmonic constituents from measurements that contain other signals or noise that is not white.

Many different signal analysis methods have been developed, each tailored to the challenges of specific applications. Generally speaking, the suitability of any particular methodology depends on the requirements for the spectral analysis (Tary et al. 2014). One of the oldest and most popular approaches is the Fourier transform, a theorem stating that any continuous periodic signal can be decomposed into a sum of properly chosen sinusoidal functions (Stein and Shakarchi 2011). For discrete measurements, the reformulated discrete Fourier transform (DFT) was introduced and translated into the fast Fourier transform (FFT), an algorithm designed for computational efficiency (Nussbaumer 1981). The FFT is implemented in all major software platforms designed for scientific computations, such as MATLAB, Scientific Python (*SciPy*) or *R*, which allows for easy implementation in data analysis workflows that require signal processing. The FFT has been used in a broad range of scientific fields over many decades and has therefore been hailed as one of the most important numerical algorithms ever created (Rockmore 2000).

With the DFT, the frequency content of a signal is estimated without a priori knowledge. While this can be a major advantage, it comes at significant costs; for example, estimates of harmonic components are prone to error, especially when their frequencies do not comply with the discretization in the frequency domain which is dictated by the record length in the time domain. As a result, the decomposition of harmonic components is influenced by aliasing and spectral leakage (Havin and Jöricke 1994; Smith et al. 2001; Stoica and Moses 2005). Separating harmonics with nearby frequencies, such as those caused by different constituents (equivalent to components) of the astronomical tides (McMillan et al. 2019), is difficult to achieve. To overcome such frequency separation issues, the minimum record duration is estimated based on the Nyquist theorem (Havin and Jöricke 1994; Acworth et al. 2016).

The presence of data gaps inherent to real-world measurements can complicate the analysis, because DFT requires a uniform sampling rate. Data gaps are usually filled by interpolation which adds to the processing steps and can affect the results in undesired ways (Munteanu et al. 2016). Finally, the magnitude of harmonic components may be smaller than the sensor resolution, and therefore these components may fall below detectability (Rau et al. 2019). Despite these limitations, the DFT remains one of the most popular approaches applied in the geosciences. However, its reliability for analyzing real-world data sets is often neglected, probably because signal processing is considered as a stepping stone rather than the research subject.

Many other signal analysis techniques have been developed to overcome the shortcomings of the DFT, such as methods that apply fitting between a model and data based on minimizing a metric that quantifies the difference; for example, the least squares of the differences between a signal and a sum of harmonic functions at defined frequencies evaluated at the times at which the signal was sampled. The popular *Lomb-Scargle* approach (Lomb 1976; Scargle 1982) falls within this category and was specifically developed for non-uniformly sampled data (VanderPlas 2018). It has received much attention and is a well-established tool applied to data sets from across the scientific disciplines. However, Stoica et al. (2009) illustrated that it does not have any particular advantages for spectral analysis based on least-squares periodograms (LSP). Further, the limitations of least-squares-based approaches when analyzing real-world data sets have not been established.

With signal extraction being a necessary step of TSA, the quality of subsequent natural property quantification is directly proportional to the reliability and accuracy of the signal analysis approach deployed. The aim of this paper is to (i) systematically compare the performance of the DFT with that of a harmonic least-squares (HALS) approach when estimating amplitudes and phases of harmonic constituents with known frequencies, and (ii) define practical limits for record duration, sampling rate, measurement resolution (here termed signal quantization), signal-to-noise ratio and gap fraction for the reliable extraction of harmonic constituents from real-world time-series measurements. The results provide important criteria and guidelines for the types of groundwater data sets that can be analyzed using TSA and for future groundwater monitoring strategies (Rau et al. 2020b).

## 2 Methodology

### 2.1 Estimation of Harmonic Component Properties

#### 2.1.1 Discrete Fourier Transform (DFT)

In the geosciences the DFT is a commonly applied methodology. The use of the DFT does not require any primary information about the signal before processing. If a data set is complete, i.e., regularly sampled in time and without gaps, the DFT can provide a good estimate of the frequency spectrum. Because of this simplicity, the Fourier transform is widely used in many scientific fields to extract frequency information about a process that is not easily accessible in the time domain. The frequency spectrum can be expressed as follows (Havin and Jöricke 1994; Stoica and Moses 2005)

$$\hat{Y}_j = \sum_{n=0}^{N-1} y_n(t_n) e^{-i \frac{2\pi j n}{N}}, \quad j = 0, \dots, N-1, \quad (1)$$

where  $y_n$  are discrete samples at times  $t_n$  with the time index  $n$ ;  $N$  is the number of discrete samples;  $\hat{Y}_j$  is a complex coefficient at a discrete frequency with index  $j$ . The amplitudes and phases for each identified frequency component can be quantified as

$$\hat{A}_j = \frac{2}{N} \text{abs}(\hat{Y}_j) = \frac{2}{N} \sqrt{\Re(\hat{Y}_j)^2 + \Im(\hat{Y}_j)^2}, \quad (2)$$

and

$$\hat{\phi}_j = -\text{atan} \left( \frac{\Re(\hat{Y}_j)}{\Im(\hat{Y}_j)} \right), \quad (3)$$

where  $\Re$  and  $\Im$  denote the real and imaginary parts of a complex number, respectively, and phase values always fall within  $-\pi \leq \phi_k \leq \pi$ .

The reader is reminded that the frequency bins are uniformly spaced, with

$$f_j = j \frac{f_d}{N}, \quad j = 0, \dots, N-1, \quad (4)$$

where  $f_d$  is the sampling frequency. Further, the frequency resolution  $\Delta f$  is the minimum interval between each frequency bin of the DFT that can be resolved and is defined as

$$\Delta f = \frac{f_d}{N} = \frac{1}{\tau}, \quad (5)$$

where  $\tau$  is the total duration of the time series.

The inverse relationship between the frequency resolution and the total duration of the time series justifies taking a preferably long data set (i.e., the longer the duration, the better the DFT resolves the components in the spectrum). Energy from frequency components whose frequencies lie between bins is distributed to the neighboring bins. This is referred to as spectral leakage (Havin and Jöricke 1994). Furthermore, the DFT assumes the input data set to be finite (i.e., a continuous spectrum with one complete

period of a periodic signal). Periodic continuation of the discrete signal beyond the considered series duration can lead to discontinuities at the transition (i.e., the two endpoints of the waveform). These affect our ability to distinguish the frequencies of the original spectrum using DFT, as they appear in the spectrum as additional high-frequency components. This well-known effect was minimized by multiplying the time record with a Hanning window (Fig. 1h), which tapers the magnitude at the beginning and at the end of a finite-length record towards zero and therefore prevents discontinuities. While many different window shapes exist, the Hanning window is often used to reduce interference from leakage (Harris 1978). For a comprehensive discussion of the DFT please refer to relevant signal analysis literature, for example Smith (2007).

### 2.1.2 Harmonic Least Squares (HALS)

The harmonic least-squares method for amplitude and phase estimation (HALS) is an optimization approach which aims to minimize the sum of the squared residuals between a model combining harmonics with known frequencies and some discretely measured data points (Stoica and Moses 2005) and is defined as follows

$$\min_{a_k, b_k} \sum_{n=0}^{N-1} \left[ y(t_n) - \sum_{k=1}^K [a_k \cos(\omega_k t_n) + b_k \sin(\omega_k t_n)] \right]^2. \tag{6}$$

Here,  $N$  is the number of discrete samples,  $y(t_n)$  is the value of a sample at time  $t_n$  and  $K$  is the total number of tidal constituents  $k$  with angular frequency  $\omega_k = 2\pi f_k$ . By adding up the harmonics ( $K$ ), multiple tidal constituents, such as those found in groundwater measurements (see Table 1), are taken into account. The sample timings  $y(t_n)$  in Eq. 6 are much more flexible compared to the requirement for equally spaced samples  $y_n(t_n)$  in Eq. 1. The solution to this minimization problem can be simply obtained by solving a system of linear equations for the coefficients  $a_k$  and  $b_k$ . It is worth noting that the solution can be affected by numerical errors that arise from the limited precision of the arithmetic operations performed by standard computers.

The coefficients  $a_k$  and  $b_k$  are converted into amplitude

$$\hat{A}_k = \sqrt{a_k^2 + b_k^2}, \tag{7}$$

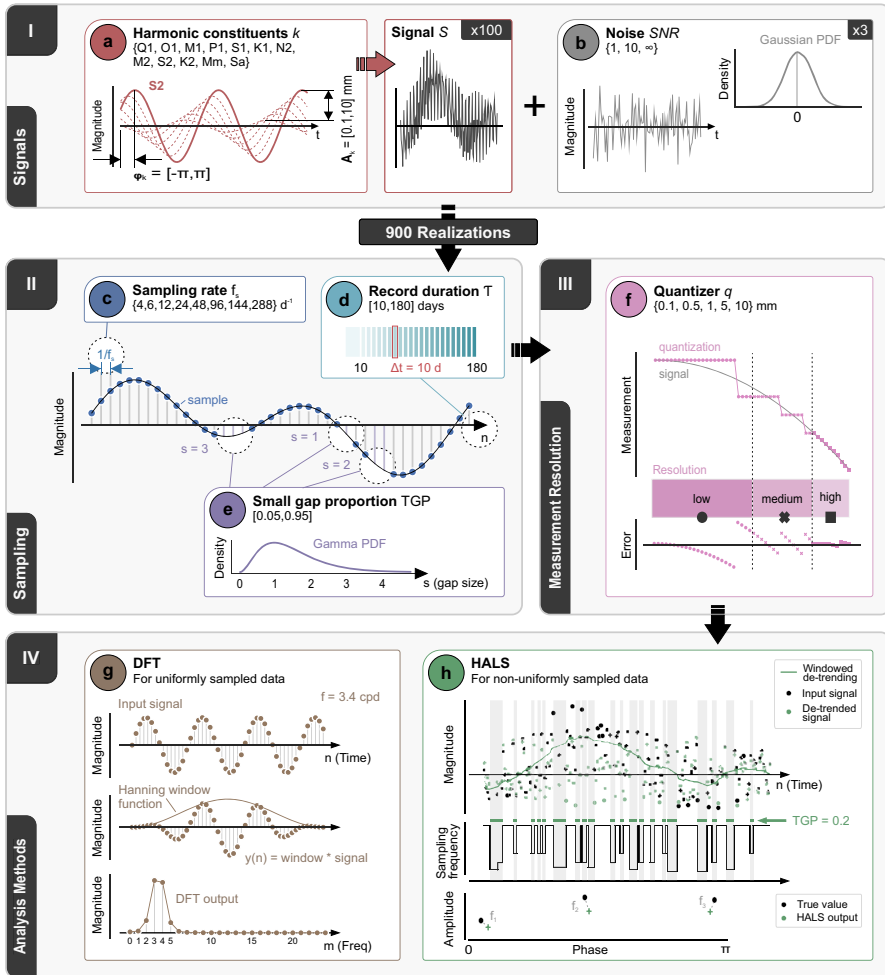
and phase

$$\hat{\phi}_k = \text{atan} \left( \frac{a_k}{b_k} \right), \tag{8}$$

with phase values always within  $-\pi \leq \phi_k \leq \pi$ .

### 2.1.3 De-trending of Records

Both DFT (Eq. (1)) and HALS (Eq. (6)) work best when a signal can be approximated by a summation of harmonic components and white noise. Most real-world ground-



**Fig. 1** A workflow for the systematic generation and analysis of the synthetic signal data set. In sub-figure (h), the windowed de-trending function is applied to a non-uniformly sampled input signal (black dots). The resulting trend (green line) is subtracted from the original signal to get the de-trended signal (green dots). The grey vertical bars illustrate the missing samples for a time series with a TGP of 0.2, while the change in sampling frequency  $f_s$  is shown in the second plot below. Finally, at the bottom, the HALS analysis results for three frequencies and their deviation from the true value (grey broken line) are shown

water head measurements contain components with random but longer periods and higher magnitudes compared to the target constituents (Table 1). A de-trend function should precede application of both methods to improve signal extraction. This was specifically developed for non-uniformly sampled data. The function fits a linear trend to the data using least squares (Oliphant 2006) and then subtracts the resulting trend from the original time series (Fig. 1h). In order to eliminate any frequency components lower than a specific cut-off value, the de-trending was done in segments by moving an averaging window with a predefined time length across the record. By defining the

**Table 1** Overview of the main tidal constituents reported in groundwater head measurements (Merritt 2004; McMillan et al. 2019), their frequency spacing and minimum required record length to resolve harmonic properties. The latter is calculated for the two closest constituents (top to bottom), using Eq. (5), with the frequency spacing being the frequency resolution

Tidal constituent ( $K$ )	Period [d]	Frequency ( $f_k$ ) [cpd]	Spacing ( $\Delta f_k$ ) [cpd]	Min. duration ( $\tau$ ) [d]
Q1	1.1195	0.89324		
O1	1.0758	0.92954	0.0363	27.55
M1	1.0347	0.96645	0.0369	27.09
P1	1.0027	0.99726	0.0308	32.46
S1	1.0000	1.00000	0.0027	364.96
K1	0.9973	1.00274	0.0027	364.96
N2	0.5274	1.89598		
M2	0.5175	1.93227	0.0363	27.56
S2	0.5000	2.00000	0.0677	14.76
K2	0.4986	2.00548	0.0055	182.48
Lunar	27.5542	0.03629		
Solar	365.2425	0.00274		

window size in terms of a time length instead of a number of samples, this approach works well for non-uniformly sampled time series. The approach acts as a high-pass filter, where the cut-off frequency ( $f \geq 0.2$  cpd) was selected to preserve the target tidal frequencies ( $f_k > 0.89$  cpd). Boundary effects were reduced by defining the window step size as a fraction of its own size, resulting in a window overlap. Here, a window size of 5 d with an overlap of 3 was chosen as a compromise between an effective high-pass filter and reasonable computing times. De-trending can be optimized by reducing the window size (i.e., by approximating the cut-off frequency to the smallest  $f_k$ ) and increasing the overlap (i.e., by smoothing the de-trending curve). In the present case, the high-pass filter removes all trends caused by constituents with a period  $d \geq 5$ , whereby all the tidal constituents sought are retained in the signal. For availability of this de-trending function please refer to the acknowledgements.

## 2.2 Characteristics of Tidal Constituents

The frequencies of astronomical tides are measured in cycles per day (cpd) and have been well documented in the literature (Agnew 2018), including their impact on groundwater systems (Merritt 2004; McMillan et al. 2019). Table 1 shows the frequencies of significant constituents that have been reported in the literature (Merritt 2004; McMillan et al. 2019) and will therefore be used in this analysis.

A major challenge for TSA is the fact that tidal constituents are buried in many other signals. Further, the frequencies of tidal constituents embedded in groundwater head measurements can be quite close. Table 1 illustrates the frequency spacing between the two closest components and the theoretically required minimum record duration to resolve the respective peaks using Eq. (5). Acworth et al. (2016) demonstrated

that  $M2$  and  $S2$  have the highest information content, making them the most useful constituents that can be resolved with a uniformly sampled groundwater level record of just 15 days. However, separating other constituents such as  $P1$ ,  $S1$  and  $K1$  would require a record length of 1 year. The uniformly spaced frequency resolution of the DFT presents a particular challenge when identifying and resolving tidal constituents (Table 1), for two reasons: (i) it is impossible to find a record duration which is a multiple of all the estimated frequencies, and (ii) it is impractical to limit the analysis to a defined record duration. Despite these shortcomings, the DFT is commonly used (Acworth et al. 2016; Allègre et al. 2016; Xue et al. 2016; Rau et al. 2018; Qu et al. 2020). It is suggested that HALS would be a better solution since the problem is well-posed because the desired tidal frequencies are known.

### 2.3 Synthetic Data Set for Method Testing

To compare and analyze the performance of DFT and HALS in extracting harmonic tidal constituents, synthetic signals were generated with the following form

$$y(t) = \sum_{k=1}^K A_k \sin(2\pi f_k t + \phi_k). \quad (9)$$

Here,  $t$  is time, and  $A_k$  and  $\phi_k$  are the amplitude and phase of each mode  $k$ , with values drawn at random from uniform distribution  $U_A[0.1, 10]$  and  $U_\phi[-\pi, \pi]$ , respectively. The synthetic signal  $S$  consists of  $K = 12$  superposed sinusoidal modes which represent typical tidal constituents that have been documented in the literature and are summarized in Table 1 (Merritt 2004; McMillan et al. 2019). Moreover, white noise was added to the signal in Eq. (9) such that

$$y(t) = S + C\mu, \quad (10)$$

where

$$C = \sqrt{\frac{P_{signal}}{SNR}}. \quad (11)$$

Here,  $C$  is a scaling factor and  $\mu$  is a Gaussian distributed random variable with zero mean and unit standard deviation  $[\mathcal{N}(0, 1)]$ . The scalable noise term in Eq. (10) defines the noise level of the signal and was calculated based on the total signal power  $P_{signal}$  and the signal-to-noise ratio  $SNR = \{1, 10, \infty\}$ . To express the power ratio between signal and noise ( $SNR$ ) more conveniently, a logarithmic decibel (dB) scale was chosen such that

$$dB = 10 \lg \frac{P_{signal}}{P_{noise}}. \quad (12)$$

Thus, a signal with  $dB = 0$  corresponds to a power ratio of 1.

A set of 900 synthetic signal realizations was generated from Eq. (10) (Fig. 1), based on 100 different signal combinations  $S$  from  $A_k$  and  $\phi_k$  in Eq. (9), while additionally considering three random noise realizations for every given  $SNR$  in Eq. (11). The



**Table 2** Overview of the signal parameters and their set of values used to generate synthetic time series representative of groundwater head measurements

Parameter	Symbol	Unit	Value range
Amplitude	$A$	mm	[0.1, 10]
Phase	$\phi$	rad	$[-\pi, \pi]$
Signal-to-noise ratio	SNR	[-]	{1, 10, $\infty$ }
Record duration	$\tau$	d	{10, 20, 30, ..., 180}
Sampling frequency	$f_s$	$d^{-1}$	{4, 6, 12, 24, 48, 96, 144, 288}
Gap proportion	TGP	[-]	{0, 0.05, 0.1, ..., 0.95}
Total length	$N$	n	[40, 51840]
Quantization	$q$	mm	{0.1, 0.5, 1, 5, 10}

random noise was generated separately for every signal realization. The resulting data set was then systematically sampled at a frequency  $f_s$  over a record duration  $\tau$  with a total gap proportion (TGP) and quantization  $q$  to create 1,860 time series for each element of the set. The complete parameter space is defined in Table 2. In total, the data set consists of 1,674,000 time-series realizations that were generated and analyzed. The different time-series configurations are designed to represent real-world groundwater head measurements and can be broadly classified as:

- i uniformly sampled time series, and
- ii non-uniformly sampled time series that are characterized by missing values at random locations, resulting in small data gaps.

While all time series with uniform sampling were analyzed by both DFT and HALS, the analysis of the non-uniformly sampled time series was limited to HALS. It should be noted that the latter configuration is representative of most time-series measurements in groundwater monitoring. The gaps are caused, for example, by temporarily removing the logger for maintenance purposes or by replacing it due to failure (Rau et al. 2019). For the synthetic time series under consideration, these gaps were created by removing sample points from the uniformly sampled time series of the signal at random locations and of different sizes (Fig. 1e). For this purpose, a similar approach as in Munteanu et al. (2016) was chosen, whereby the gap size distribution was defined by a gamma function, so that

$$f(x, \alpha, \beta) = \frac{x^{\alpha-1} \times \exp\left\{\frac{-x}{\beta}\right\}}{\beta \times \Gamma(\alpha)}, \quad x > 0; \alpha, \beta > 0 \quad (13)$$

where  $\alpha$  and  $\beta$  are the shape and scale parameter of the gamma probability density function (PDF), respectively. The actual probability of each gap size  $P(s)$  was controlled by drawing both mean ( $\bar{x} = \alpha \times \beta$ ) and variance ( $s^2 = \alpha \times \beta^2$ ) of the gamma function at random from a uniform distribution  $U_g[1, 5]$ . Each gap size probability was then scaled by the TGP, which represents the total number of points removed from the time series and is defined as a proportion of the original number of sampling

points  $N$ . Thus, the total number of gaps  $G$  of a certain size  $s$  that are removed from the signal at a random location  $x$  as successive samples ( $n_g = [n_x, n_{x+s})$ ) can be described as follows

$$G(s) = \frac{P(s) \times TGP}{s} \times N, \quad (14)$$

where the resulting real number was rounded to the nearest integer value. The overlap of the randomly created gaps is minimized while

$$N - (G \times 2) - \sum n_g > 0, \quad (15)$$

by only selecting gap locations that are buffered by at least one sample value on either side of the gap. Non-uniform sampling was limited to  $f_d = \{6, 12, 24, 48\}$  and  $\tau = \{60, 120, 180\}$  to minimize the computational load while sampling representative time series at the lower end of  $N$ .

In practice, every measurement device has a limited resolution, which leads to a rounding effect on the digits of the signal magnitude (Fig. 1f). In signal theory this is referred to as quantization, and the approximation leads to a small error. This effect was considered by applying a uniform quantizer  $Q$  to each discrete time series, which can be expressed using a floor function

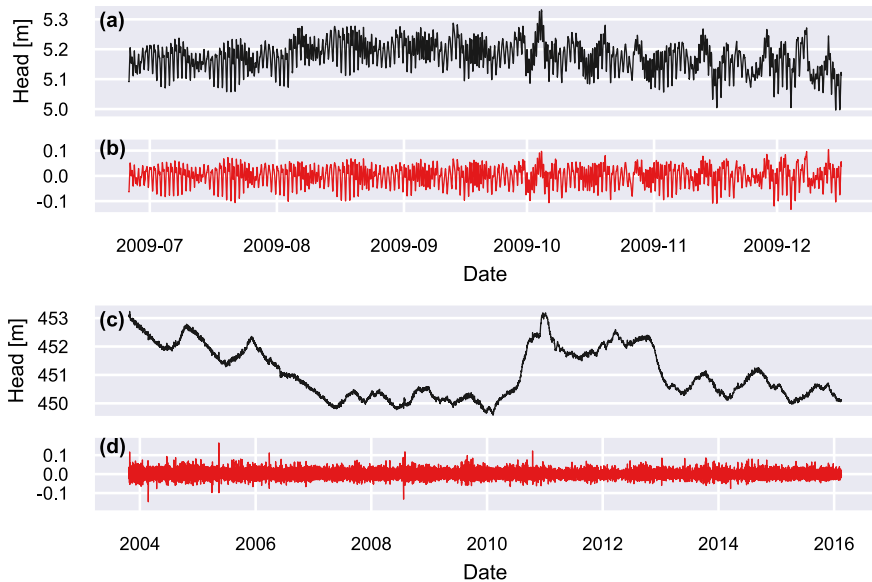
$$Q(x) = q * \left\lfloor \frac{x}{q} + \frac{1}{2} \right\rfloor, \quad (16)$$

where  $x$  is a real number and  $q$  is the quantization step size. The floor function returns the greatest integer that is less than or equal to the enclosed term. Thus, for  $q = 1$ , the quantizer is simply rounding to the nearest integer. The quantization range represents the range of groundwater pressure transducer resolutions that are found on the market (Rau et al. 2019).

## 2.4 Description and Analysis of Field Measurements

The two signal processing methods were further applied to real groundwater head measurements from two boreholes at different locations, which are representative of frequently occurring subsurface and measurement conditions:

- Borehole **BLM-1** is located in Inyo County (California, USA) at the latitude  $36.408130^\circ$ , longitude  $-116.471360^\circ$  and height of 688 m (WGS84). The well is screened in a carbonate rock aquifer at a depth of 884 m below land surface with a screen length of 106 m. The groundwater heads were measured between 25 June 2009 22:00:00 [UTC] and 12 December 2009 16:30:00 [UTC], using a high-quality gauge transducer manufactured by In-Situ Inc. (USA). The data set comprises 16,683 data points sampled every 15 minutes (96 samples per day), with no missing values. This data set shows an unusually strong Earth and atmospheric tide influence and thus is an ideal case for investigating the impact of Earth and atmospheric tides on groundwater systems (Rau et al. 2020a).



**Fig. 2** Groundwater head measurement time series from **a** borehole BLM-1 (Inyo County, CA, USA) and **c** BH3 (Baldry, Australia) and the de-trended time series in **b** and **d**, respectively

– Borehole **BH3** is located in Baldry (New South Wales, Australia) at the longitude  $-32.868088^\circ$ , latitude  $148.536771^\circ$  and height of 450 m (WGS84). The well is screened in a fractured rock aquifer below a confining layer of silt at 20 m depth with a screen length of 1 m (Acworth et al. 2016). The groundwater heads were measured between 24 October 2003 01:00:00 [UTC] and 15 February 2016 20:00:00 [UTC], using a high-quality gauge transducer manufactured by In-Situ Inc. (USA). The data set comprises 107,948 measurements sampled hourly (24 samples per day) over 4,497.8 days, with no missing values. This unusually long continuous data set provides an ideal test bed for achieving high-frequency resolution (here  $10^{-4}$  cpd) when the DFT is used, resulting in sharp peaks in the frequency spectrum and reliable identification of closely spaced frequency components. The DFT spectrum for this data set was shown in McMillan et al. (2019).

Figure 2 panels (a) and (c) show the groundwater head records measured over different time periods at each location. The amplitude and phase of the estimated tidal constituents are considered stationary for both borehole data sets and thus represent the integrated average of the function of the true values over the interval given by the length of the data sets.

The accuracy of DFT and HALS in quantifying harmonic constituents was analyzed and compared in the following way:

- (a) The model of both methods was trained on the first 75% of each data set. Thus, the model was trained on 12,512 and 80,961 samples for BLM-1 and BH3, respectively. To evaluate the performance of each method, the error variance of the residual between model and measurements was calculated.

- (b) For the DFT, the modeled signals were calculated using only the frequency bins closest to the tidal frequencies. The other bins in the spectrum were discarded to exclude artifacts originating from the de-trending or from noise and also to increase practicability.
- (c) Both models are validated on the remaining 25% of each data set by calculating the variance of the residual between each model prediction and the corresponding measurements. Thus, the validation was based on 4,171 and 26,987 samples for BLM-1 and BH3, respectively.

The error variance  $\sigma$  is determined by

$$\sigma = \frac{1}{N} |\mathbf{M} - \mathbf{D}|^2, \quad (17)$$

where the vectors  $\mathbf{M}$  and  $\mathbf{D}$  contain the modeled and de-trended signal, respectively. Comparing the error variance of the residual allows for a quantitative assessment of model performance, whereby a smaller variance indicates better performance.

### 3 Results and Discussion

#### 3.1 Comparison of DFT and HALS

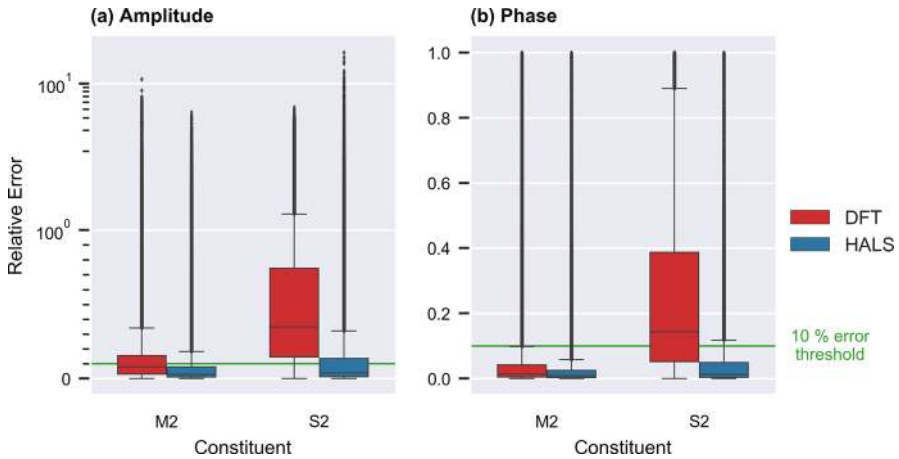
When applying TSA, an accurate assessment of amplitudes and phases of the constituents  $M2$  and  $S2$  is particularly important in order to derive reliable subsurface properties. The efficacy of both the DFT and HALS methods was evaluated as a function of the detection accuracy, with results shown in Fig. 3. The direct comparison was limited to time series with  $\tau \geq 30$  days. Since the true parameter value is known, the accuracy was defined in terms of the relative error  $RE$  as

$$RE_A = \frac{|\hat{A} - A_{true}|}{A_{true}}, \quad (18)$$

and

$$RE_\phi = \frac{|\pi - ||\hat{\phi} - \phi_{true}| - \pi||}{|\phi_{true}|}, \quad (19)$$

with the nominator in Eq. (19) taking into account that the maximum distance between two phases is  $\pi$ . A target value (TV) of 10% relative error was defined as the maximum acceptable relative error (i.e., results with a  $RE$  less than 10% are regarded as sufficiently accurate). While this is an arbitrary threshold, it is a small value compared with the compound uncertainties typical of active hydraulic investigations that are standard in hydrogeology (Raghavan 2004). It should be noted that the  $RE$  distribution of both methods is strongly right-skewed. In Fig. 3 this is evident from the off-centered median value and interquartile range. In order to take account of the skewness in the data and to give less weight to outliers, the mean relative error for each grid cell in Figs. 4, 5, 6 and 7 was calculated using the log-transformed  $RE$  values. The result was inverted again for better readability.



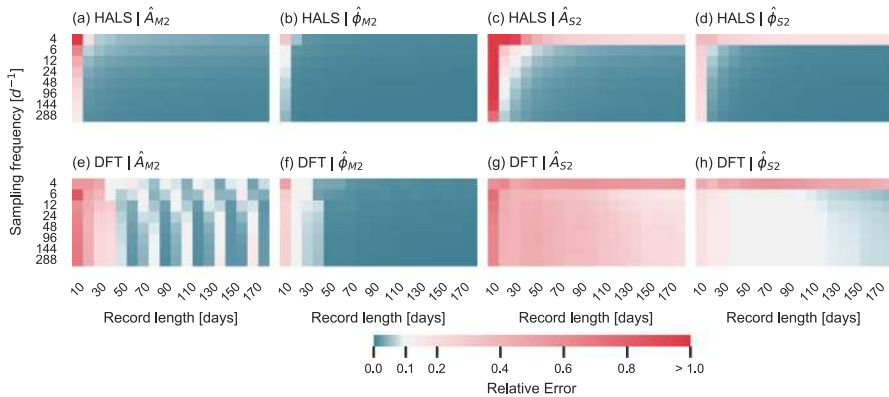
**Fig. 3** Comparison of the two signal analysis methods HALS and DFT for  $\tau \geq 30$  days in estimating the **a** amplitude and **b** phase of the target constituent  $M2$  and  $S2$ , respectively. The boxes are the interquartile range (IQR), with the black horizontal line denoting the median of all calculated  $RE$  values. The whiskers extend to  $1.5 * IQR$ , and the black dots are outliers. Note that **a** is plotted on a symmetric log scale with base 100 and linearly increasing values between 0 and 100

Overall, 86% and 62% of  $\hat{A}_{M2}$ , and 95% and 90% of  $\hat{\phi}_{M2}$  were below the TV for HALS and DFT, respectively, whereas only 76% and 18% of  $\hat{A}_{S2}$ , and 86% and 40% of  $\hat{\phi}_{S2}$  were below the TV for HALS and DFT, respectively. Thus, both methods are consistently more accurate in estimating the  $M2$  properties compared to  $S2$ .

The outliers for  $RE_A$  estimates of the HALS method compared to a narrow and low interquartile range in Fig. 3 indicates that the accuracy of the HALS estimation deteriorates under certain boundary conditions, while it stays robust for most of the investigated signal and sampling parameter range. While HALS clearly outperformed the DFT in simultaneously estimating amplitudes and phases of the tidal constituents, the estimation routine requires proper constraints (see Sect. 3.6). In contrast to the DFT method, in which the maximum of  $\hat{A}$  is limited by the signal power, HALS is an optimization problem that minimizes the residuals and thus has no reasonable inherent constraint on  $\hat{A}$ . Instead, the goodness of fit depends on the polynomial order (i.e., the degrees of freedom) given by the number of estimated constituents, but also on the record duration and the frequency separation of the estimated constituents. The poor performance of DFT in estimating the  $S2$  properties, on the other hand, is likely due to spectral leakage from other nearby frequencies, for example  $S1$ ,  $M2$  and  $K2$  (Table 1). The method’s ability to discern nearby frequencies is strongly limited by the frequency binning, which depends on sample frequency and record duration, which is further analyzed below.

### 3.2 Practical Data Set Requirements for Detecting Harmonic Properties

When assessing whether a data set of measurements is suitable for TSA, record duration  $\tau$  or sampling frequency  $f_s$  are used as indicators in practice because they are



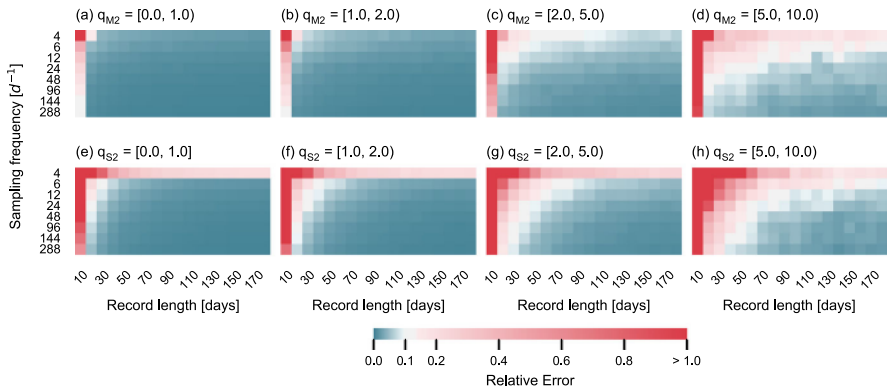
**Fig. 4** Effect of sampling frequency and record duration on the mean accuracy of both HALS (top) and DFT (bottom) in estimating the amplitude (a, c, e, g) and phase (b, d, f, h) of the two target constituents  $M2$  and  $S2$

easy to determine and related to criteria such as the Nyquist-Shannon sampling theorem. Figure 4 shows the interrelated effect of  $f_s$  and  $\tau$  on both  $RE_A$  and  $RE_\phi$  for HALS and DFT, respectively. HALS can reliably extract the constituents phase and amplitude at  $f_s \geq 6$  n/d, while not exceeding the TV of 10% relative error. Only for  $\tau < 50$  d, the sampling frequency increasingly becomes an issue in determining  $\hat{A}_{S2}$  (Fig. 4c) and should be taken into consideration. In essence,  $\tau$  and  $f_s$  are inversely correlated for  $\tau < 50$  d, and  $f_s$  needs to be increased if  $M2$  and  $S2$  are to be reliably detected. Figure 4 shows that even if  $f_s = 288$  d<sup>-1</sup>, the absolute minimum duration for records analyzed with the HALS method is 20 days.

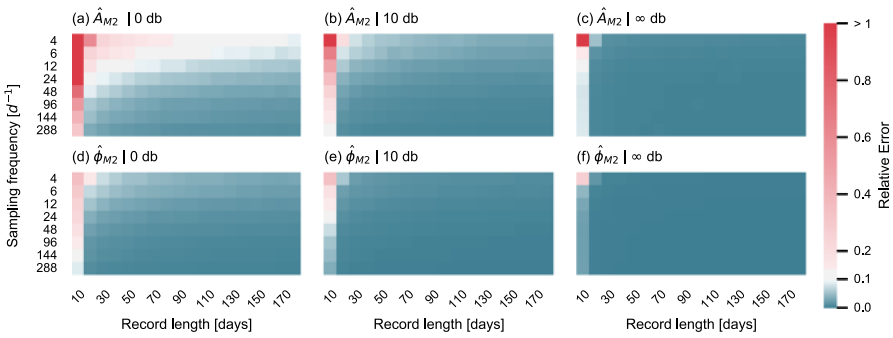
A detection trade-off between  $\tau$  and  $f_s$  was not discernible with the DFT method. This is in accord with the well-known fact that the DFT frequency resolution depends on record duration. Instead, the detection accuracy increased with the length of the record only. Further, the DFT method appears much less robust compared to HALS, and it is not possible to determine a minimum criterion for  $RE_A$  within the realistic constraints given by our analysis. Acworth et al. (2016) suggested that records as short as 16 days could be used to resolve the properties of  $S2$  and  $M2$ , based simply on considering the Nyquist-Shannon theorem (Table 1). However, our analysis clearly demonstrates that under realistic measurement conditions, the presence of additional tidal constituents renders this limit impossible.

### 3.3 Effect of Measurement Resolution and Noise Level

The influence of measurement resolution and noise level on the accuracy of HALS is shown in Figs. 5, 6 and 7. In order to facilitate a comparison of the effect of the quantization step size  $q$  on the accuracy of the amplitude estimate  $RE_A$  across all time series,  $q$  was normalized by the amplitude  $A_k$  of the constituent of the signal under consideration, namely  $M2$  and  $S2$ , such that



**Fig. 5** Effect of sampling frequency and record duration on the mean accuracy of HALS in estimating the amplitude of  $M2$  (a–c) and  $S2$  (d–f) for different ranges of normalized quantization

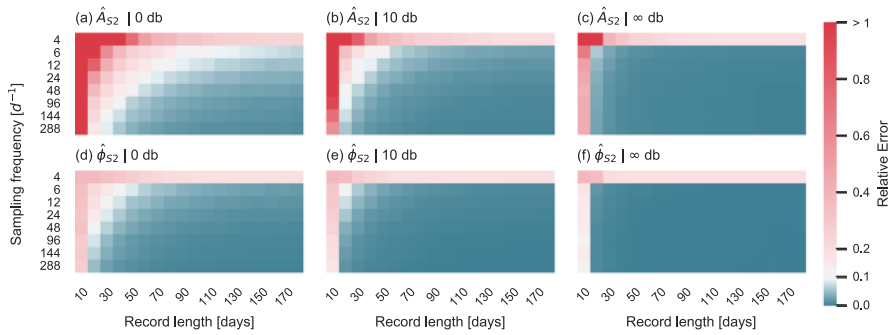


**Fig. 6** Effect of sampling frequency and record duration on the mean accuracy of HALS in estimating  $M2$  amplitude (a–c) and phase (d–f) for three different noise levels

$$q_k = \frac{q}{A_k} \quad k \in \{M2, S2\}. \tag{20}$$

The resulting quantization ratio  $q_k$  thus reflects the relative influence of the rounding error introduced by  $q$  on the identifiability of the constituent  $k$ , which depends on the magnitude of  $k$  itself.

From Fig. 5 it becomes apparent that the effect of the sample parameters  $\tau$  and  $f_s$  on the detection accuracy changes with the  $q_k$ . In fact, as  $q_k$  increases, there is a clear dependency between the accuracy and both sample parameters. For  $q_k < 2.0$ , the minimum sample criteria for HALS are identical to the previously determined limits (Fig. 4). In other words, the quantization step size should be less than twice the amplitude of the constituent under consideration, without impairing the accuracy of the method. Conversely, at higher values of  $q_k$ , the minimum sample criteria need to be adapted for a reliable estimate, even more so for  $\hat{A}_{S2}$ . An increase in the sampling rate seems to be particularly effective here. In practice, the actual measurement resolution that can be achieved is of course limited, which is why this relationship is not transferred to arbitrarily small amplitudes  $A_k \ll 0.1$  mm.



**Fig. 7** Effect of sampling frequency and record duration on the mean accuracy of HALS in estimating  $S_2$  amplitude (a–c) and phase (d–f) for three different noise levels

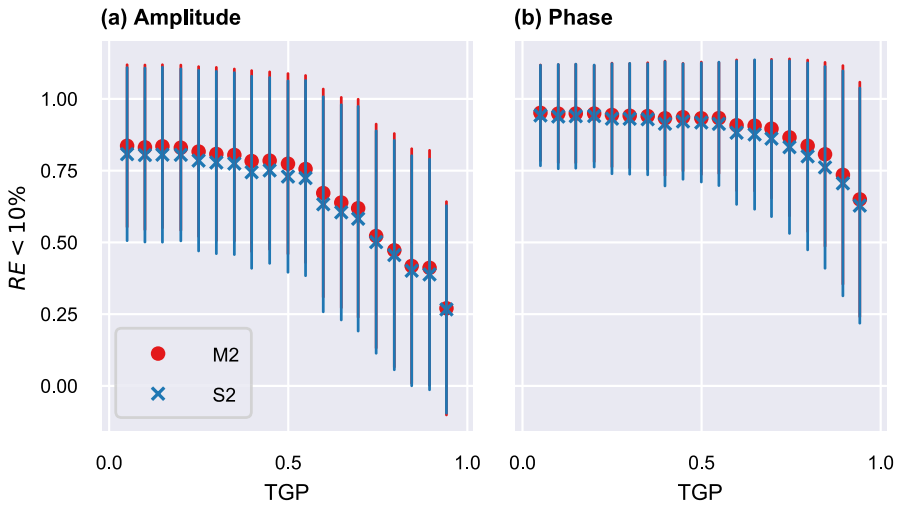
A similar overall tendency in the effect of the sample parameters on detection accuracy was observed for a change in noise level. With an increasing level of noise, the sampling criteria require adjustment. In contrast to  $q_k$ , however, accurate estimates for the properties of both target constituents are feasible even at the highest investigated noise levels of 0 dB. Furthermore, for a noise level approaching zero ( $db = \infty$ ), HALS is effectively limited by the accuracy of the  $S_2$  estimate requiring  $f_s \geq 6 d^{-1}$  and  $\tau \geq 20$  (Fig. 7).

### 3.4 Effect of Non-uniform Sampling on HALS Performance

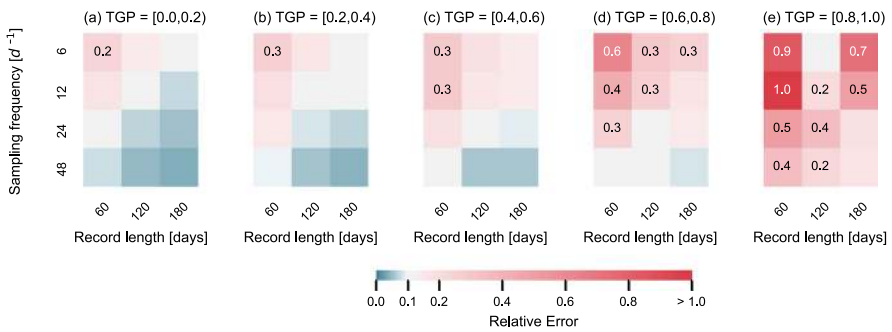
One of the main advantage of the HALS method is that non-uniformly sampled data records can be analyzed directly without further preprocessing steps. The performance of the HALS method in estimating the target constituent properties for time series with increasing TGP is presented in Fig. 8, as the fraction of estimates with a relative error below the aforementioned TV of 10% ( $RE < 10\%$ ). On average, 72% and 70% of all  $\hat{A}$  and 86% and 85% of all  $\hat{\phi}$  met the TV criteria for  $M_2$  and  $S_2$ , respectively. Thus, the HALS method performs well in extracting the target constituent properties for the majority of the investigated time series, even with gaps. However, the fraction of estimates that meet the TV criteria are highly dependent on the TGP, which is apparent from a sharp decrease in accuracy at  $TGP \approx 0.5$ . This is especially the case for the amplitude estimates (Fig 8a). For  $TGP < 0.5$ , on the other hand, the estimates are highly reliable, with about  $79 \pm 25\%$  and  $94 \pm 15\%$  of all  $RE_A$  and  $RE_\phi$  meeting the TV criteria, respectively. Thus, the general dependency of the estimation accuracy on the TGP is much less pronounced for the phase (Fig. 8b). However, the standard deviation indicates that the accuracy of the estimate at any given gap proportion can deviate, depending on the underlying signal realization.

Furthermore, the effect of the gap proportion on the accuracy of the estimate is again dependent on the sampling of the time series (Fig. 9). Data gaps decrease the effective sampling rate, so that even for a  $TGP < 0.2$ , the minimum sampling criteria are slightly more strict than for a time series without gaps. It should also be noted that while the average gap size is controlled by the gamma function (Eq. (13)), once Eq. (15)





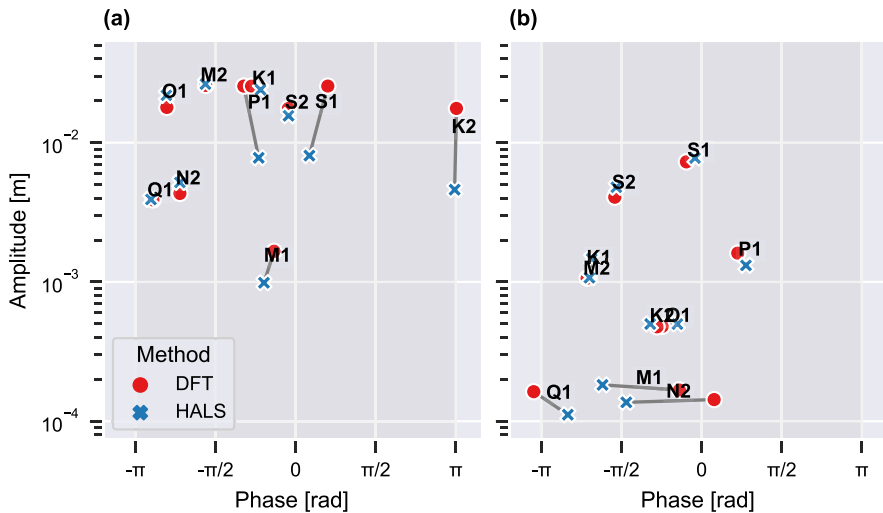
**Fig. 8** Effect of small gap proportion on the performance of HALS in estimating amplitude (a) and phase (b) of the two target constituents *M2* and *S2*. The performance is evaluated in terms of the fraction of time series where the relative error is below the TV ( $RE < 0.1$ ). The vertical bars indicate the standard deviation for the different signal realizations



**Fig. 9** Effect of sampling frequency and record duration on the mean accuracy of HALS in estimating the amplitude of *M2* and *S2* combined, for different gap proportions. The mean accuracy was explicitly indicated for values of  $RE > 0.2$

is violated, the average gap size increases dramatically due to the aggregation of smaller gaps. Thus, the observed steep decrease in estimation accuracy above  $TGP \approx 0.5$  can be partly attributed to the emergence of ever larger gaps.

Groundwater head measurements are generally made at regular intervals, but often contain small, randomly distributed gaps, with the TGP remaining well below the critical 50% (Rau et al. 2019). While these gaps are usually interpolated as a preprocessing step for the DFT method, this not only adds an extra step to the data analysis, but also increases the uncertainty of the parameter estimation. The previous analysis shows that HALS, on the other hand, makes interpolation superfluous and works consistently well for records with small gaps, providing another advantage for HALS as a standard approach to extracting harmonic constituent properties.

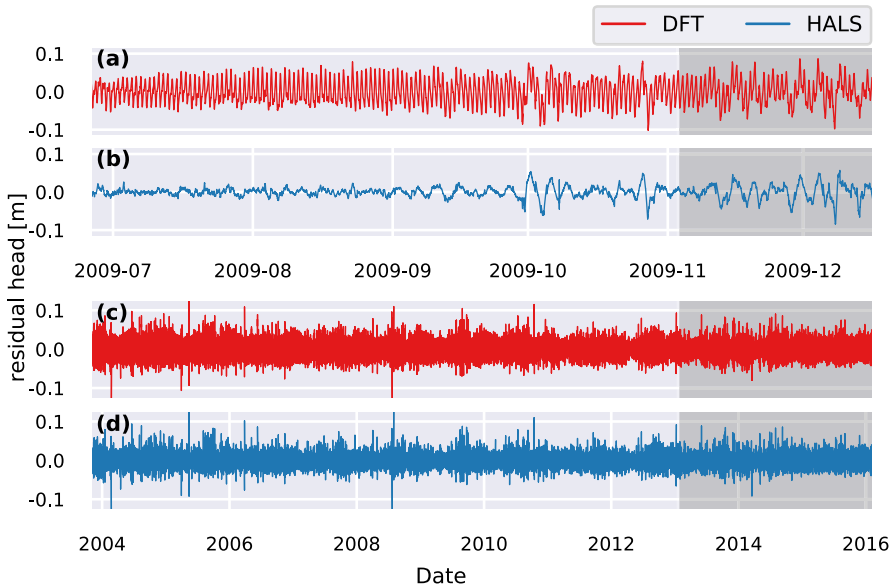


**Fig. 10** Comparison of the amplitudes and phases determined by the DFT and HALS for **a** the BLM-1 data set from Inyo County (California, USA) and for **b** the BH3 data set at Baldry (New South Wales, Australia). The grey lines represent the differences between the DFT and HALS estimates. Note that the amplitude scale is logarithmic for improved visibility

### 3.5 Comparing the Performance of DFT and HALS Using Groundwater Level Measurements

Both methods were tested on the real-world groundwater head measurements described in Sect. 3.2. Figure 10 shows the amplitude and phase estimates obtained for all tidal constituents and both data sets. A comparison of the two data sets clearly shows that the record length in particular has a major impact on the estimation accuracy. Indeed, the two methods result in rather different  $\hat{A}$  estimates for the short data set (Fig. 10a), while they give almost identical  $\hat{A}$  estimates for the longer data set (Fig. 10b). However, it is also interesting to note that the results for some of the constituents are consistently very close, while others differ significantly. For example, the determined  $\hat{A}_{S2}$  and  $\hat{\phi}_{S2}$  are in compliance with each other for both DFT and HALS and in both data sets. On the other hand,  $\hat{A}_{K2}$  and  $\hat{A}_{S1}$  in the BLM-1 data set are rather different for the two methods. These differences are mainly due to spectral leakage inherent to the DFT. This means that the energy of tidal constituents with nearby frequencies, for example,  $K1$  and  $S1$ , are contained within the same frequency bin, especially for records with shorter duration (e.g., BLM-1), leading to an inaccurate frequency resolution. In this regard, the ability of HALS to model specific frequencies generally provides a better performance.

Figure 10 qualitatively illustrated that there can be significant differences in the amplitudes and phases detected using DFT or HALS. When used to characterize groundwater systems based on TSA, this can lead to erroneous interpretations. For example, Hsieh et al. (1987) used the DFT to extract amplitudes but then calculated phases using a HALS-like approach. Figure 10 and the synthetic data analysis indicate



**Fig. 11** Residuals calculated by applying the DFT (a, c) and HALS (b, d) to the real-world groundwater head measurements from boreholes BLM-1 (a, b) and BH3 (c, d). The model for both methods was trained on the first 75% of the record and validated using the remaining 25% (grey-shaded). Error variances were calculated for the training and the validation periods and are listed in Table 3

**Table 3** Summary of the training and validation of the DFT and HALS models for both real-world groundwater head measurements. The values represent the error variance calculated for each subset of the data depicted in Fig. 11

	BLM-1		BH3	
	Training	Validation	Training	Validation
DFT	$9.99 \cdot 10^{-4}$	$1.09 \cdot 10^{-3}$	$3.93 \cdot 10^{-4}$	$3.55 \cdot 10^{-4}$
HALS	$1.56 \cdot 10^{-4}$	$4.62 \cdot 10^{-4}$	$2.87 \cdot 10^{-4}$	$2.97 \cdot 10^{-4}$

that this could be inconsistent. Further, the synthetic data analysis clearly shows that HALS provides more accurate results for real-world groundwater records.

Figure 10 begs the question of which method is more accurate in resolving harmonic component properties. To answer this question, the prediction accuracy of each method was analyzed and compared on the basis of the error variance of the model residuals. The results are summarized in Fig. 11 and Table 3. Overall, the variance of the residuals between model and measurements is small for HALS compared to DFT. Of course, this should come as no surprise, as minimizing the residuals is inherent to the least-squared method, and it is clearly superior in this regard. Nevertheless, it also reflects the earlier results that were obtained by testing the methods with synthetic data and demonstrates that HALS performs better in quantifying the properties of tidal constituents. Furthermore, the model residuals of the validation data (Fig. 11), and thus

the error variance (Table 3), are higher for the DFT than for the HALS method. This is particularly pronounced for the short data set, where the DFT residuals also contain prominent oscillations (Fig. 11a). These oscillations are a result of the spectral leakage stemming from the discrete and fixed frequency resolution inherent to the DFT. Figure 11 panels (c) and (d) show that the residuals of both methods become more similar as the length of the data set increases. Thus, as the frequency resolution of the DFT increases, the spectral leakage is reduced, which leads to an improved ability to detect harmonic components. In contrast, the HALS model residuals are consistently low, indicating that it is more accurate in resolving the harmonic components contained in both records.

### 3.6 General Considerations for the Use of HALS

The analysis presented in this work demonstrates that HALS is superior to the DFT when extracting tidal constituents from realistic groundwater head measurements. While the minimum record duration required to distinguish components appears lower compared to the DFT, it is still the most significant constraint for HALS. It is important to note that Eq. 6 is an optimization problem where the solution can be affected by the limited precision of arithmetic computations, leading to compounded numerical errors. The risk of this occurring increases with decreasing record duration, where a solution can become ill-conditioned. For example, the combination of closely spaced frequency components and short duration records will lead to similar values in the columns of the linear system of equations that is to be solved. This can degrade the accuracy of the solution and lead to overall biased estimates.

Similar to the DFT, there is a trade-off between record duration and minimum frequency spacing for the components determined by HALS (i.e., the smaller the frequency spacing, the longer the record duration must be). To determine whether the solution is ill-conditioned, the conditioning number can be evaluated. This is the ratio of the maximum and minimum singular values of the matrix in the least-squares problem (Eq. 6) and should not exceed a large number, such as  $10^8$  (i.e., depending on the computing environment). If a system becomes ill-conditioned, a solution would be to reduce the number of desired tidal constituents. For example, tidal constituents  $P1$ ,  $S1$  and  $K1$  (Table 1) could first be unified into one harmonic component, as they are closely spaced. Constituents  $S2$  and  $K2$  are the next closest components that could be merged if the previous merger does not improve the conditioning. For TSA, however, the two components  $M2$  and  $S2$  are of primary interest, which is why an analysis with HALS only makes sense as long as the properties of both can be reliably extracted.

## 4 Conclusions

The systematic analysis and comparison of the methods revealed that the harmonic least-squares (HALS) approach is very robust and outperforms the discrete Fourier transform (DFT) in extracting the amplitude and phase of harmonic constituents of known and closely spaced frequencies from time-series measurements. Only for rather short data sets can the variance of the least-squares amplitude estimate increase signif-

icantly. The analysis focused on distinguishing the closely spaced frequencies of the Earth and atmospheric tidal constituents  $M2$  (1.93227 cpd) and  $S2$  (2.0 cpd). Both are generally present as dominant constituents in barometric and groundwater pressure measurements and can be used to quantify subsurface properties. As a benchmark of accuracy, an average relative error of 10% was considered to be acceptable when estimating tidal constituent amplitudes and phases.

It is well known that the DFT frequency domain resolution increases with longer duration of a time series and remains independent of the sampling frequency  $f_s$ . For HALS on the other hand, an accurate estimate of the  $S2$  amplitude requires  $f_s$  to increase as  $\tau$  decreases. This effect was slightly less pronounced for the amplitude estimate of  $M2$ . Furthermore, the results of this study suggested that an absolute minimum record duration of at least 20 days is needed for resolving the properties of both constituents  $M2$  and  $S2$  at more than 24 n/d (for a normalized measurement resolution below 2). However,  $f_s$  should not be lower than 6 n/d, unless the noise level approaches zero ( $SNR = \infty$ ). In the case of  $SNR = \infty$ , HALS can exceed these sampling limitations, in particular when determining the properties of  $M2$ . However, this scenario is unlikely in practice, as environmental measurements are generally noisy.

The synthetic data analysis further included noise and signal resolution, factors that represent the characteristics of measurement hardware used in environmental sensing such as groundwater pressure. As expected, increasing noise levels degrade the detection of harmonic properties, but more so for amplitudes than phases. Even a noise level of 0 dB, which is considered very high for standard sensors, resulted in excellent HALS performance. Furthermore, sensor resolution (here simulated as signal quantization) also degraded the performance. However, results illustrated that a quantization ratio below 2.0 (i.e., quantization normalized by the amplitude of the target constituent) is recommended.

Finally, the influence of small data gaps inherent to real-world data sets was analyzed. The results demonstrated that the accuracy of the HALS method in detecting the harmonic constituents was reliable as long as the TGP remained below 50%. This relatively high threshold is one of the main advantages of the HALS method, which makes further preprocessing steps such as interpolation or resampling superfluous. The superiority of HALS was further underpinned by a comparison of amplitude and phase estimates from two different real-world groundwater head records.

The analysis presented herein suggests that when applying tidal subsurface analysis (TSA) (McMillan et al. 2019), HALS should be used instead of DFT. In combination, these approaches provide a powerful tool for groundwater resource investigations. The results establish practical criteria which can be used to determine the suitability of existing groundwater head records for TSA. In addition, recommendations for future groundwater monitoring strategies can be derived, with which the accuracy of the passive characterization can be maximized (Rau et al. 2020b).

**Acknowledgements** The idea for this work was conceived by GCR. This manuscript is a significant advancement to a BSc thesis written by Vincent Ried submitted to the Geophysical Institute at KIT. A generalized Python function to perform HALS using Eq. (6) and the windowed de-trending function can be accessed on GitHub: <https://github.com/ds138/HALS>, where they are located in the HALS.py script. We thank Paula Cutillo and Shannon Mazzei from the National Park Service (NPS) in California (USA)

for providing the groundwater pressure data set from BLM-1. This data set is also available online: [https://nwis.waterdata.usgs.gov/nwis/gwlevels/?site\\_no=362402116280901](https://nwis.waterdata.usgs.gov/nwis/gwlevels/?site_no=362402116280901). Some of the data used in this paper were collected with equipment provided by the Australian Federal Government-financed National Collaborative Research Infrastructure Strategy (NCRIS). The groundwater data are available through the NCRIS Groundwater Database: <http://groundwater.anu.edu.au>.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement no. 835852. PS was partly supported by the Swedish Research Council (grants VR 2017-04610 and 2016-06079).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Acworth RI, Halloran LJS, Rau GC, Cuthbert MO, Bernardi TL (2016) An objective frequency domain method for quantifying confined aquifer compressible storage using Earth and atmospheric tides. *Geophys Res Lett* 43(22):611–671
- Agnew DC (2018) An improbable observation of the diurnal core resonance. *Pure Appl Geophys* 175(5):1599–1609 ISSN 14209136
- Allègre V, Brodsky EE, Xue L, Nale SM, Parker BL, Cherry JA (2016) Using earth-tide induced water pressure changes to measure in situ permeability: a comparison with long-term pumping tests. *Water Resour Res* 52(4):3113–3126 ISSN 19447973
- Bredehoeft JD, Papaopulos IS (1965) Rates of vertical groundwater movement estimated from the Earth's thermal profile. *Water Resour Res* 1(2):325–328 ISSN 19447973
- Ghil M, Allen MR, Dettinger MD, Ide K, Kondrashov D, Mann ME, Robertson AW, Saunders A, Tian Y, Varadi F, Yiou P (2002) Advanced spectral methods for climatic time series. *Rev Geophys* 40(1):3-1
- Halloran LJ, Rau GC, Andersen MS (2016) Heat as a tracer to quantify processes and properties in the vadose zone: a review. *Earth Sci Rev* 159(June):358–373 ISSN 00128252
- Harris F (1978) On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc IEEE* 66(1):51–83 ISSN 0018-9219
- Havin V, Jörnicke B (1994) The uncertainty principle in harmonic analysis, vol 28. Springer, Berlin Heidelberg, Berlin, Heidelberg. ISBN 978-3-642-78379-1
- Hsieh PA, Bredehoeft JD, Farr JM (1987) Determination of aquifer transmissivity from earth tide analysis. *Water Resour Res* 23(10):1824–1832
- Lomb NR (1976) Least-squares frequency analysis of unequally spaced data. *Astrophys Space Sci* 39(2):447–462 ISSN 0004640X
- McMillan TC, Rau GC, Timms WA, Andersen MS (2019) Utilizing the impact of earth and atmospheric tides on groundwater systems: a review reveals the future potential. *Rev Geophys* 57(2):281–315 ISSN 8755-1209
- Merritt ML (2004) Estimating hydraulic properties of the Floridan Aquifer System by analysis of earth-tide, ocean-tide, and barometric effects, Collier and Hendry counties, Florida. Technical report, USGS, Tallahassee, Florida
- Munteanu C, Negrea C, Echim M, Mursula K (2016) Effect of data gaps: comparison of different spectral analysis methods. *Ann Geophys* 34(4):437–449 ISSN 14320576
- Nussbaumer HJ (1981) The Fast Fourier transform. In: *Fast Fourier transform convolution algorithms*, 2nd edn. Springer, Berlin, Heidelberg, pp 80–111. ISBN 978-3-662-00553-8
- Oliphant TE (2006) *A guide to NumPy*. Trelgol Publishing USA

- Pawlowicz R, Beardsley B, Lentz S (2002) Classical tidal harmonic analysis including error estimates in MATLAB using T\_TIDE. *Comput Geosci* 28(8):929–937 ISSN 00983004
- Qu S, Wang G, Shi Z, Zhou P, Xu Q, Zhu Z (2020) Temporal changes of hydraulic properties of overburden aquifer induced by longwall mining in Ningxia coalfield, northwest China. *J Hydrol* 582:124525 ISSN 00221694
- Reviews of Geophysics (2004) A review of applications to constrain pumping test responses to improve on geological description and uncertainty. 42(4):RG4001. ISSN 8755-1209
- Rau GC, Acworth RI, Halloran LJS, Timms WA, Cuthbert MO (2018) Quantifying compressible groundwater storage by combining cross-hole seismic surveys and head response to atmospheric tides. *J Geophys Res Earth Surf* 123(8):1910–1930 ISSN 21699003
- Rau GC, Andersen MS, McCallum AM, Roshan H, Acworth RI (2014) Heat as a tracer to quantify water flow in near-surface sediments. *Earth Sci Rev* 129:40–58 ISSN 00128252
- Rau GC, Cuthbert MO, Acworth RI, Blum P (2020a) Technical note: disentangling the groundwater response to earth and atmospheric tides to improve subsurface characterisation. *Hydrol Earth Syst Sci Dis* 2020:1–19
- Rau GC, Cuthbert MO, Post VEA, Schweizer D, Acworth RI, Andersen MS, Blum P, Carrara E, Rasmussen TC, Ge S (2020b) Future-proofing hydrogeology by revising groundwater monitoring practice. *Hydrogeol J* 28(8):2963–2969
- Rau GC, Post VEA, Shanfield M, Krekler T, Banks EW, Blum P (2019) Error in hydraulic head and gradient time-series measurements: a quantitative appraisal. *Hydrol Earth Syst Sci* 23(9):3603–3629 ISSN 1607-7938
- Rockmore D (2000) The FFT: an algorithm the whole family can use. *Comput Sci Eng* 2(1):60–64 ISSN 15219615
- Scargle JD (1982) Statistical aspects of spectral analysis of unevenly spaced data. *Astrophys J* 263:835–853 ISSN 0004-637X
- Smith J O (2007) Mathematics of the discrete Fourier transform (DFT): with audio applications. W3K Publishing. ISBN 097456074X
- Smith KE, Green M, Thomas GO, Jones KC (2001) Behavior of sewage sludge-derived PAHs on pasture. *Environ Sci Technol* 35(11):2141–50 ISSN 0013-936X
- Stein EM, Shakarchi R (2011) Fourier analysis: an introduction, vol 1. Princeton University Press. ISBN 978-0691113845
- Stoica P, Li J, He H (2009) Spectral analysis of nonuniformly sampled data: a new approach versus the periodogram. *IEEE Trans Signal Process* 57(3):843–858 ISSN 1053-587X
- Stoica P, Moses RL (2005) Spectral analysis of signals. Upper Saddle River, New Jersey 07458: Pearson Prentice Hall. ISBN 0-13-113956-8
- Tary JB, Herrera RH, Han J, van der Baan M (2014) Spectral estimation-what is new? What is next? *Rev Geophys* 52(4):723–749 ISSN 87551209
- Van der Kamp G, Gale JE (1983) Theory of earth tide and barometric effects in porous formations with compressible grains. *Water Resour Res* 19(2):538–544 ISSN 19447973
- VanderPlas JT (2018) Understanding the Lomb–Scargle periodogram. *Astrophys J Suppl Ser* 236(1):16 ISSN 0067-0049
- Wörman A, Riml J, Schmadel N, Neilson BT, Bottacin-Busolin A, Heavilin JE (2012) Spectral scaling of heat fluxes in streambed sediments. *Geophys Res Lett* 39(23):n/a–n/a. ISSN 00948276
- Xue L, Brodsky EE, Erskine J, Fulton PM, Carter R (2016) A permeability and compliance contrast measured hydrogeologically on the San Andreas fault. *Geochem Geophys Geosyst* 17(3):858–871 ISSN 15252027