Estimating hydraulic aquifer parameters from tide-induced groundwater fluctuations

A case study in Schouwen-Duiveland

F.H.B. (Floris) van 't Klooster

Delft University of Technology

TU Delft

Estimating hydraulic aquifer parameters from tide-induced groundwater fluctuations

A case study in Schouwen-Duiveland

by

F.H.B. (Floris) van ’t Klooster

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday April 29, 2021 at 10:00 AM.

Student number:	4742737	
Thesis committee:	Prof. Dr. ir. M. Bakker,	TU Delft
	Dr. ir. W. J. Zaadnoordijk,	TNO, TU Delft
	Dr. ir. E. Abraham,	TU Delft
	Dr. ir. G. Schoups,	TU Delft

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

Cover photo: View over part of Schouwen-Duiveland, in the bottom of the picture a part of the study site and on the left in the middle the Oosterschelde. Photographer: Siebe Swart

TU Delft

Preface

With this thesis, I made the final step in concluding my MSc studies in Water Management at the faculty of Civil Engineering at Delft University. Although it took me quite some time to complete this thesis (1.5 years), quite remarkably I enjoyed writing it (almost) continuously. Partly because of my interest and the challenges faced in performing the hydraulic parameter estimation, but also because of the support of friends, family and of course my committee.

I would first like to thank Willem Jan Zaadnoordijk, for the received feedback but mostly for the many meetings we had. Your feedback and insights were very helpful, not only about the subject but also more broadly on how to tackle the different aspects of such a big project. I really appreciate all these meetings and enjoyed them as well. I would also like to thank Mark Bakker, I greatly appreciate all the detailed feedback I received and although it took some time to implement all the feedback, I greatly appreciate the effort and the amount of time you have put into it, which I know is not that self-evident at the TU. Furthermore, I want to thank Edo Abraham and Gerrit Schoups for chairing the committee and for providing valuable feedback.

Furthermore, a big thanks to my friends. Especially my roommates and study-mates for their support and the great time I had during my thesis. Lastly, I want to thank my family and my parents specifically, for their support during my study.

F.H.B. (Floris) van 't Klooster
Delft, April 2021

Summary

In this report, the tidal method is used to estimate hydraulic aquifer parameters. The principle of the tidal method is to use head fluctuations in observation wells, caused by tidal motion in a sea or river, to determine regional hydrological characteristics of an aquifer system. As the wave propagates into the aquifer, the amplitude of the signal decreases and the phase increasingly lags behind the tide at sea. The extent of this is determined by hydrological aquifer characteristics, hence aquifer parameters can be estimated, by calibrating a model that reproduces the measured tidal propagation. Knowledge of hydraulic aquifer parameters is important since these are needed in groundwater models.

The case study was performed on Schouwen-Duiveland where use was made of hydraulic head fluctuations from three groundwater observation wells and water-level observations in the Oosterschelde. First, these time series were analyzed to estimate the amplitudes and phases of the constituents present in the data. Noise in the data was reduced with the use of Pastas (a model to analyze hydrological time series; Collenteur et al., 2019) and a Butterworth filter, after which both methods were compared. Pastas was used to decompose the fluctuations observed in the groundwater to different contributions of hydrological stresses (e.g. rain and evaporation) and Butterworth was used to flatten the frequency response for frequencies that are not of interest. The use of the Butterworth filter is preferred, partly because it produces the smallest standard deviations for the amplitude and phase estimates and because it is easier to use. Moreover, especially for the wells with a low signal to noise ratio, the Butterworth filter is better at extracting the tidal signal.

In addition, a graphical determination of the amplitude and phase was performed to check if this relatively quick and easy analysis gives accurate estimates of the M2 amplitude and phase as well. It was concluded that with a high signal to noise ratio and a dominantly present constituent (M2 in this case), the amplitude can be reliably estimated. Determining the phase with this method did not give satisfactory results.

In the inverse modeling part, the hydraulic aquifer parameters are determined, which is done with a least-squares minimization of the observed and modeled amplitudes and phases. For the optimization, both a one-aquifer and a two-aquifer model were used, both based on the one-dimensional, multi-layer solution presented by Bakker (2019). The optimization was performed with a global optimizer.

The obtained fit to the observations was somewhat poor, moreover, unrealistic optimal parameter estimates indicate that the model, to some extent, incorrectly represents the real system. This also implies that some of the model simplifications do significantly impact the results. Simplifications presumed to mostly influence the model results include homogeneity, one-dimensional flow and the use of a straight shoreline. The presence of some model error signifies that the resulting parameter estimates should be treated with care.

The hydraulic conductivity and the resistance of the one-layer model were consistently estimated at their upper boundary and the storage in the leaky layer was estimated to be negligible. Therefore, only one parameter group (i.e. the diffusivity) could be estimated with the optimization. The fit of the model was not perfect, a compromise must be sought based on what residual the model minimizes. As a result, all diffusivity estimates within and around the range [1.07E6, 1.31E6] m^2/day are all considered to be reasonable estimates.

For the resistance of the aquitard (CU) and the storage in both the aquifer (S_sU) and the aquitard (σU) it was analyzed which parameter values gave unreasonable results. This resulted in a lower bound for CU (CU=663 days) and an upper bound for σU (σU=1.22E-4 m^{-1}). For S_sU reasonable results are obtained between 2.28E-6 m^{-1} and 1.87E-5 m^{-1}. The lower bound for the CU estimate seems to comply with estimates based on two models (REGIS-II and GeoTOP). Finally, the two-aquifer model was considered to be useless for parameter estimation. Presumably, this is predominantly caused by a small amount of amplitude and phase data compared to the number of parameters to be optimized.

Contents

1

Introduction

Groundwater is the most important source of Dutch drinking water and it is intensively used by the (farming) industry (Dufour, 1998). The recent reoccurring droughts in the Netherlands have stressed the importance of the sustainable use of this increasingly scarce source of freshwater (Polman et al., 2019). Of course, sustainable management requires knowledge of the groundwater system.

A specific case close to this case study, which requires sustainable management, is the 'Getij Grevelingen' project. The region and the state want to bring back some tidal movement to improve the quality of the water in the Grevelingenmeer, located on the north side of Schouwen (zwdelta, 2020). This will change the water system significantly and the effect of this on the groundwater needs to be studied. Often this is easier said than done. Groundwater systems are usually complex and there is often a limited amount of data available (Hill et al., 2000). Still, we need accurate models to make informed decisions about water resources. These models need parameters to describe the system. And these parameters need to be estimated.

The most common ways to determine the hydraulic properties of an aquifer system are pumping & slug tests (e.g. Trefry and Bekele, 2004; Istok and Dawson, 2014; Merritt, 2004). By applying a known stress to the aquifer in conjunction with pressure measurements, aquifer parameters can be determined. Of these two types of tests, only the pump test can be used to estimate hydraulic properties on a regional instead of local scale (OhioEPA, 2006). However, pumping tests are resource expensive and can be undesirable when the groundwater contains contaminants (e.g. industrial waste, salt) (Trefry and Bekele, 2004; Merritt, 2004). Moreover, the results of a pumping test are hard to analyze when the aquifer system is also subjected to other stresses. Therefore the use of methods that rely on natural stresses can be convenient. One of these is the tidal method.

The principle of the tidal method is to use head fluctuations in observation wells, caused by the tidal motion in a sea or river, to determine regional hydrological characteristics (e.g. hydraulic conductivities and storage coefficients) of an aquifer system (Carr and Van Der Kamp, 1969; Rotzoll et al., 2013). These head fluctuations in the aquifer are caused by the tidal motion of overlying or adjacent surface-water. Depending on site characteristics (e.g. the presence of a confining aquifer at the bottom of the sea), these fluctuations in the aquifer are either caused by a weight change upon the aquifer and/or by a hydraulic connection with the surface-water (Jacob, 1950; Liu et al., 2008). In the aquifer, the fluctuation then propagates land-inwards. As it does, the amplitude decreases, and the phase lag behind the tide at sea increases (e.g. Merritt, 2004). The amount of damping and the increase in phase lag is determined by hydrological aquifer characteristics (e.g. hydraulic conductivity & specific storage of the aquifer) (Carr and Van Der Kamp, 1969). Hence aquifer parameters can be estimated, by calibrating a model that reproduces the measured tidal propagation.

For this case study, a linear aquifer system is assumed, which is valid for laminar flow and thus valid for most groundwater systems, except for example when the medium is very porous or in close vicinity to a large pumping well (Bear and Verruijt, 1987). If a system can be governed by linear equations, the principle of superposition can be used. This allows for the separate consideration of the various sinusoidal components of the tide. The most dominant constituent is the M2 tide but a more elaborate list of constituents can be found in Appendix B.1.

Being able to consider these constituents separately is convenient, not in the least because the amount of damping and the increase in phase lag is also dependent on the frequency. But also because many of the proposed (analytic) solutions require single constituents as input.

Many such models or analytical solutions have been proposed, of which some were used in case-studies as well. The simple case solution proposed by (Jacob, 1950), has been used many times in subsequent studies (e.g. Ferris, 1952; Dong et al., 2015). Jacob's solution considered fluctuations in a confined aquifer, taking the compressibility of the water and the soil into account. Van der Kamp (1973) presented a solution from which the confining top extended to infinity below the sea. Later Li and Jiao (2001) presented a solution which includes leakage. See Appendix C.1 for a schematic view of the solutions from Jacob (1950), Van der Kamp (1973) and Li and Jiao (2001). The above-described solutions are all one-dimensional and specific cases of the multi-layer solution provided by Bakker (2019), which is used in this case study. A more complete summary of analytical studies on the tidal method can be found in Li and Jiao (2003).

More complex systems/solutions have also been subject of study, such as a sloping beach (e.g. Su et al., 2003; Vandenbohede and Lebbe, 2007) and the effects of wave runup (Nielsen, 1999). Numerical studies have also increasingly been published (e.g. Merritt, 2004; Vandenbohede and Lebbe, 2007; Zhou et al., 2014). A review study performed by Li and Jiao (2003) already pointed out that the solution presented by Jacob (1950), which is used in many subsequent studies, is based on assumptions that are essentially always violated. These assumptions include a one-layer aquifer with no horizontal deformations that stops at the shoreline, a straight coastline and a non-sloping beach. Consequently, many analytical studies found inconsistent estimates of the assessed parameters, depending on whether the tidal lag, attenuation, or both were used in the analysis (e.g. Trefry and Bekele, 2004; Zhou, 2008; Guarracino et al., 2012). Although increasingly complex models have been used in field studies, it seems that the statement made by Van der Kamp (1973) is still valid today. Van der Kamp stated that in general it can be said that although for some studies the modeled results are in compliance with the theory, results are not so good in other studies, and on the whole, no satisfactory explanations for the discrepancies have been presented.

1.1. Objective and approach

In this study, a simple model is considered. Advantages of using a simpler model are that often fewer parameters have to be used and computation times are shorter. The latter is convenient, especially if many model runs must be performed (e.g. when doing Monte Carlo Simulations). Moreover, using a simpler model often results in a better understanding of the dynamics of the model (Hill, 2006). See for example Hill (2006) or Hill and Tiedeman (2006) for a more elaborate discussion on the use of simplicity in developing models.

To attain a simple model, the multi-layer analytical solution presented by Bakker (2019) is used. This model allows for a subsurface schematization with multiple aquifers and aquitard layers, which is present in the Southwestern part of the Netherlands (Dufour, 1998), besides, this model is suitable for the inclusion or exclusion of hydrological parameters. Moreover, because it is an analytical model its computation times are small. Multiple layers are not considered in former analytical case studies known by the author (e.g. Smith and Hick, 2001; Fakir and Razack, 2003; Trefry and Bekele, 2004; Dong et al., 2015). However, Trefry and Bekele (2004) did some numerical tests and found that adding a layer with a high conductivity underneath a less conducting layer, could resolve some inconsistency. Therefore the modeling of multiple layers might decrease inconsistent parameter estimates encountered in some of the cited studies. Hence it is sensible to use a model that can have multiple layers if needed.

The purpose of this study is twofold: (1) building a simple model that is sufficiently able to reproduce the observed tidal fluctuations in the groundwater and in the process gain more insight into the dynamics of the system; (2) estimating hydraulic aquifer parameters. Issues of concern are mitigating the inconsistency between estimates based on attenuation and tidal lag. This was often encountered in former case studies and results in contradictory hydraulic parameter estimates. The main research question is:

Can the hydraulic aquifer parameters, on Schouwen-Duiveland, be reliably estimated from tide-induced groundwater fluctuations?

The following sub-questions are considered to determine an efficient and effective execution of the data preparation part.

- *How can the tidal signal be extracted from the measured hydraulic heads?*
- *How accurate is graphical* determination of the amplitude and phase from the M2 tide?*

*Graphical determination is used as an abbreviation for a method described in Section 3.5. But essentially, the amplitude and phase are determined merely by using a plot of the hydraulic head against the time.

As mentioned, the estimation of the hydraulic aquifer parameters will be done using the tidal method. The tidal method has the advantage that the frequency of the head response caused by the astronomical tide is exactly known. Therefore, other stresses or processes, causing hydraulic head fluctuations, do not need to be identified to determine the effect of the astronomical tide on the hydraulic head. Only time series of the hydraulic head and possibly the surface water are required, which makes this technique relatively inexpensive and suitable for data-scarce areas (Trefry and Bekele, 2004). Moreover, this method can be performed with just a short time series. An alternative would be to work with a groundwater model that includes the effects of other stresses. This would require a more complex model, but the extra information might lead to more accurate parameter estimates. Nevertheless, such a model was not considered as this paper focuses on the tidal method.

The tidal method only deals with the influence of the tide, so all fluctuations considered are sinusoidal and therefore, in addition to working in the time domain, it is also possible to work in the frequency domain (i.e. working with amplitudes and phases). Working in the time domain would be more suitable if other stresses (e.g. rain and evapotranspiration) would be included in the model. However, in the frequency domain, patterns like the above described exponential decaying amplitude and linear increasing phase lag can be detected more easily and the results are easier to interpret. Therefore, most results are presented in the frequency domain in this report. Note that 'extracting the tidal signal' in the first sub-question means, conveying the signal to the amplitude and phase of the different tidal constituents.

Lastly, determining the phase and amplitude, of the M2 tide, graphically could potentially serve as a quick first-order approximation, which is why the second sub-question is considered. It is worth noting that the M2 constituent is a semi-diurnal lunar tide, which is dominant in most locations around the world and in the Oosterschelde as well (Battjes and Labeur, 2017).

1.2. Most important approximations

To make the analysis feasible, some approximations have to be introduced. A more elaborate list of the assumptions and approximations is given in section 5.2, but some of the more notable adopted approximations in this research are the following: hydrogeological units are homogeneous and isotropic. Furthermore, it is assumed that resistance to vertical flow in aquifer layers is negligible (Dupuit-Forchheimer approximations) and that water in confining layers flows vertically (De Glee approximations).

1.3. Report structure

The report is divided into two parts: Data preparation & Inverse modeling. In Figure 1.1 the general approach used this study is clarified with a block diagram. In the data preparation part, the data from the time series of the groundwater and sea-level fluctuations is mapped into an amplitude and a phase, for each constituent that will be used for the inverse modeling. This procedure is called a 'harmonic analysis'. In addition, the groundwater fluctuations observed in the observation wells are first filtered from disturbances not caused by the astronomical tide in the Oosterschelde. The latter was done to improve the precision** of the amplitude and phase estimates (more elaborately discussed in Chapter 3).

In the inverse modeling part, the observed and modeled amplitudes and phases are used to determine hydraulic aquifer parameters, which is done by minimizing a least-squares objective function.

** To avoid confusion, the terms precision and accuracy, which are often used in this report, are described here. Estimates are precise if they have a small variance, regardless of how accurate the results are. An estimate is accurate if they are distributed evenly around the true value.

In Chapter 2 the study area and the sources of the data are discussed. Followed by the methodology and results, Chapter 3 & 4 respectively, for the data preparation part of the report. In Chapters 5 & 6 the methodology and results of the inverse modeling part are discussed. A discussion is presented in Chapter 7, followed by a conclusion and some recommendations in Chapter 8.

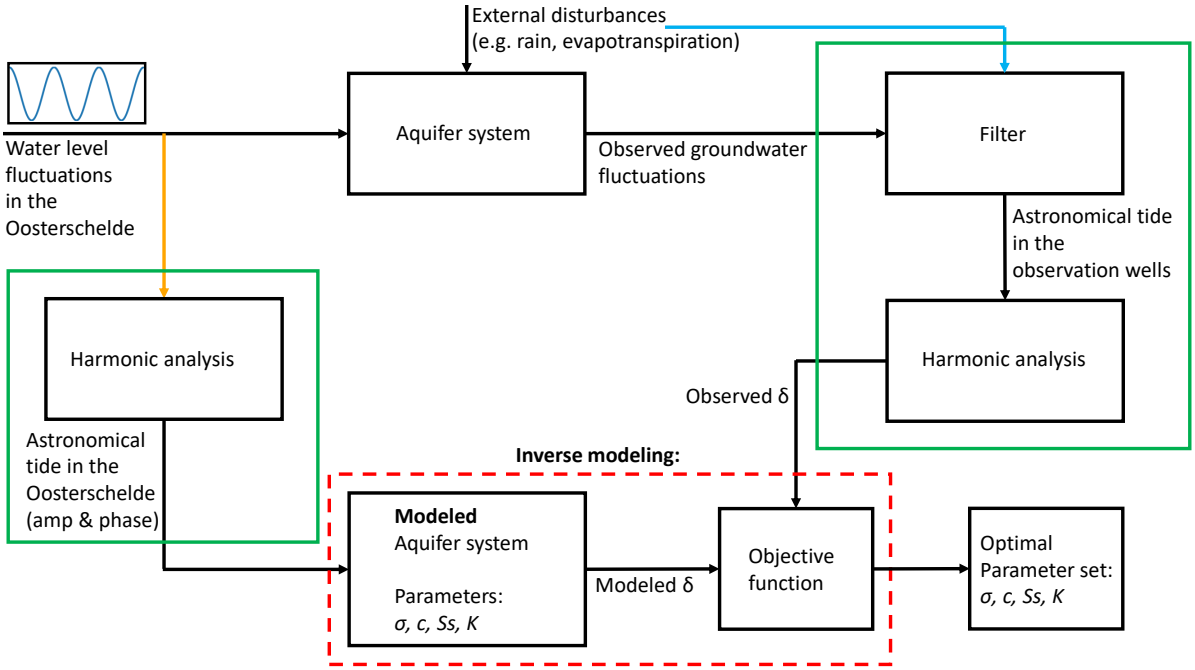


Figure 1.1: Block diagram of the general approach that is used in this study. Where δ is the astronomical tide in the observation wells expressed in terms of the amplitudes and phases of the different constituents. The steps in the green boxes are executed in the data preparation part of the report (Chapters 3 & 4) and the inverse modeling part (red dotted box) is executed in Chapters 5 & 6. Note that two arrows were given a different color. One arrow is colored orange because the data was modified a bit (i.e. translated near the study site and mapped to the astronomical tide; see Section 2.2.2). Another input arrow is blue, which is colored because only one of the two filters, considered in this study, needs these external disturbances as input (see Section 3.2).

2

Study Area & Data sources

The study area of this research is located on the south side of Schouwen (former island, now part of Schouwen-Duiveland) located in Zeeland (exonym: Zealand) (Figure 2.1). This site is suitable because of its vicinity to the Oosterschelde which has a considerable tidal range (±2.8 meter). And because a string of observation wells across the island have been equipped with dataloggers that register the head every 15 minutes. Such a high temporal resolution is convenient if the tidal method is performed because it allows for the determination of small period tidal constituents.

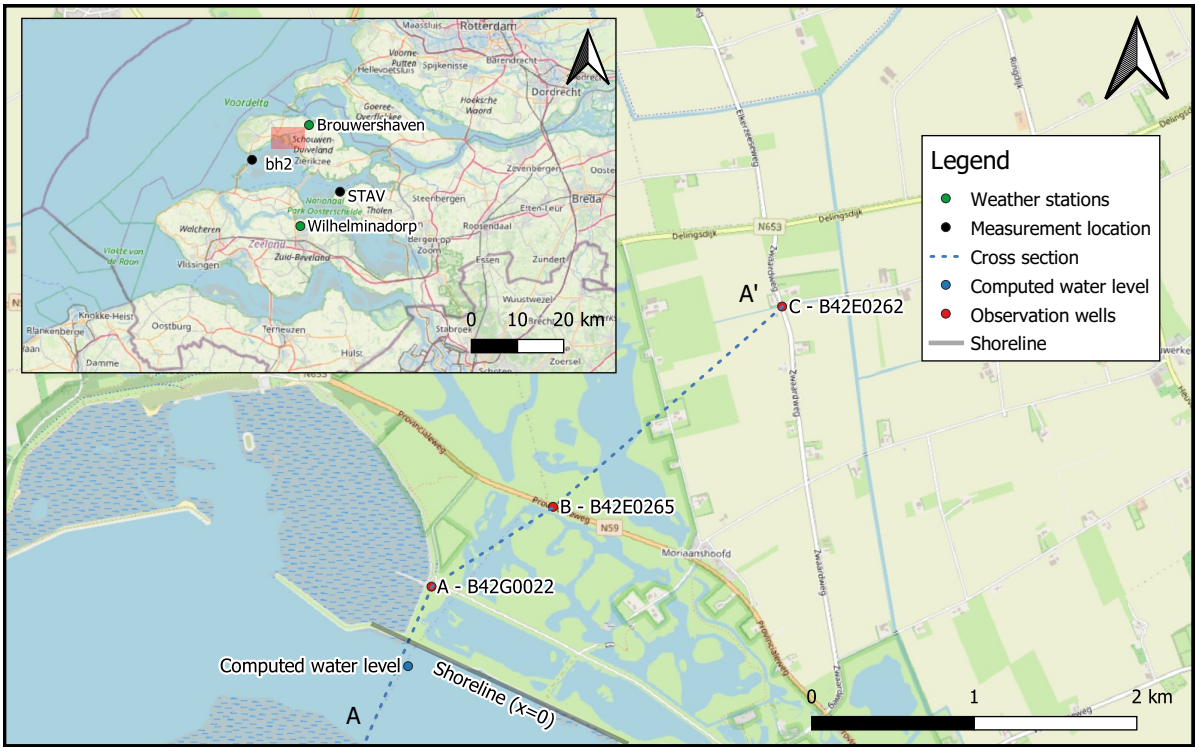


Figure 2.1: Map of the study area (marked transparent red in the small overview map), with the measuring points of the astronomical tide (black dot), weather stations, with location from cross Section of Figure 2.2 & 7.1

2.1. Geohydrological data

At the site of the case study, the geohydrology consists of multiple aquifers (see Figure 2.2). The location of this vertical cross Section is drawn in Figure 2.1. The colored layers in Figure 2.2 are different hydrogeological units from the REgional Geohydrological Information System (REGIS-II; Vernes et al., 2005; Vernes et al., 2010; `www.dinoloket.nl`). The top dark green layer is a Holocene layer, which is a complex layer consisting of

sand, clay and peat. Note that the deeper part of the Oosterschelde cuts through the Holocene confining layer, this can also be observed in Figure 2.3. Below the Holocene layer, several sandy formations are indicated in yellow, light-green and light-blue. The two somewhat darker layers below the yellow layer are two aquitards.

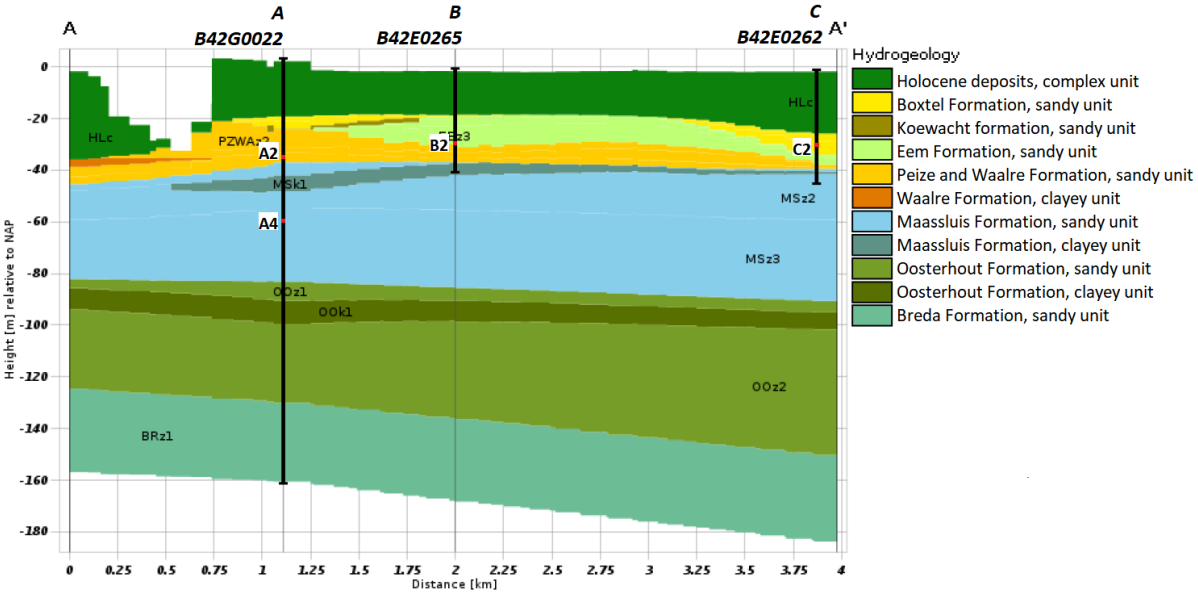


Figure 2.2: Hydrogeological cross Section of the study site, with the depth and location of the observation wells (black line) and an indication of the height from the piezometer screens (red marks on the black line). See Figure 2.1 for the location of the cross Section. This data is publicly available (at: `www.dinoloket.nl`)

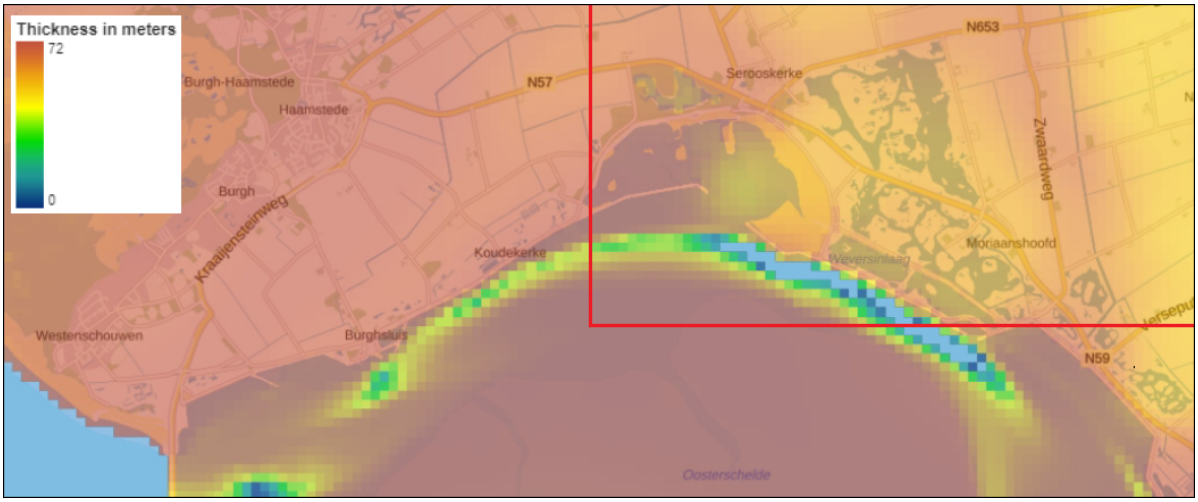


Figure 2.3: Thickness of the Holocene deposits. With the study area from Figure 2.1 outlined in red.

2.1.1. Maps of the hydraulic conductivity

For each of the hydrogeological units representative values of the hydraulic conductivity have been derived (Vernes et al., 2005). These estimates can be used to constrict the optimization, but more on that in Section 5.5.3. Below a description of how maps of the hydraulic conductivity have been derived.

In the Regis-II model, the hydraulic conductivities are interpolated from borehole logs, for which the hydraulic conductivities are derived from the relation between lithology, depositional environment and hydraulic characteristics (Hummelman et al., 2019). These hydraulic conductivities are scaled up to representative values for each hydrogeological unit (Hummelman et al., 2019). It has turned out that the hydraulic

conductivities for each hydrogeological unit are best described by a log-normal distribution (Hummelman et al., 2019). In Regis-II the statistics for each hydrogeological unit have been derived with the use of a Monte Carlo approach. By drawing randomly from a log-normal distribution of each lithoclass within a hydrogeological unit, repeating it 10'000 times and subsequently calculating the mean and standard deviation for each hydrogeological unit (Hummelman et al., 2019).

The 'upper aquifer' is located between the complex (Holocene) layer and a thin clay lens (i.e. a clayey unit from the Maassluis formation; see Figure 2.2). According to the Regis-II model the upper aquifer at the study site consists of 10 sand units, all with different grids for the mean and standard deviation of the hydraulic conductivity. In this report a similar Monte Carlo approach, as described above, was used to calculate the statistics on the hydraulic conductivity of the combined 10 sand units. First, a random value is drawn from the log-normal distribution of each sand unit. Subsequently, these values are mapped to a hydraulic conductivity for the whole upper aquifer. For an aquifer with N (10 in this case) sand units with a thickness d_i, i,..,N with hydraulic conductivities k_i, i,..,N the hydraulic conductivity for the whole aquifer layer (K_h) is given by

$$K_h = \frac{\sum_{i=1}^{N} d_i k_i}{\sum_{i=1}^{N} d_i}. \tag{2.1}$$

The above steps were repeated 2E5 times. After which the 2.5, 50 and 97.5 percentiles of the hydraulic conductivity were calculated. The resulting map for the 50 percentile is presented in Figure 2.4. Maps for the 2.5 and 97.5 percentiles for the hydraulic conductivity of both the upper and the lower aquifer are presented in Appendix C.2, along with the 2.5 and 97.5 percentiles for the resistance of the clay lens (i.e. the clayey unit from the Maassluis formation). The maps in Appendix C.2 are derived with an approach similar to the above-described method. Lastly, note that the lower aquifer is located between the clayey unit from the Maassluis formation and a clayey unit from the Oosterhout formation (see Figure 2.2).

The model of Bakker (2019), which is used for the inverse modeling, uses a constant thickness for each aquifer/aquitard layer. A constant thickness of 20.3 m was used for the upper aquifer. The maps for the hydraulic conductivity (2.4, C.2 & C.3) are therefore transformed by dividing by the true aquifers thickness (see Figure C.4) and subsequently multiplying with 20.3 m. The same approach was used for the lower aquifer.

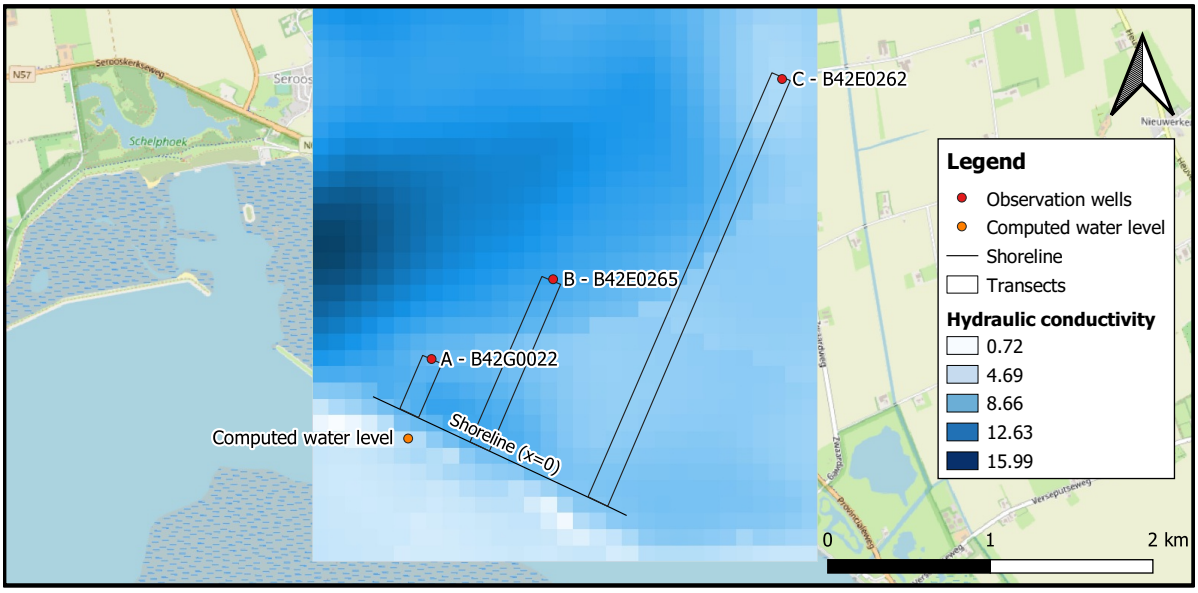


Figure 2.4: A map of the median (50th percentile) hydraulic conductivity (m/day) at the study site, the map is based on the statistics from the Regis-II model (Vernes et al., 2005). The 3 polygons, called transects in the legend, were used to calculate some basic statistics for each transect, but more on that in Section 5.5.3. Lastly note that the thickness of the aquifer associated with the presented hydraulic conductivity is constant (20.3 m) instead of varying (see Figure C.4).

2.2. Hydrological Data

2.2.1. Groundwater data

Data of the head (used as shorthand for hydraulic head in the rest of the report) in the groundwater from several observation wells across Schouwen is provided by TNO (The Geological Survey of the Netherlands). The temporal resolution of this data is 15 min. This sampling rate is high enough to avoid aliasing, which is an effect that causes different signals to become indistinguishable when sampled (Godin, 1972). To avoid it, a sampling rate is needed of at least twice the highest frequency that is desired to be extracted from the signal (Witte, 1993). This 'criteria' is also called the Nyquist condition (Deltares, 2014). So for this case, with 3MS8 (which has a period of 3hour and 5min) as the highest considered frequency, a 1.5-hour temporal resolution would suffice.

The head that is registered by the loggers deviates somewhat from hand measurements that were taken. However, this error is presumed to be constant over time, so that calculations for which the reference level is not important are still correct (e.g. for determining the phase and amplitude). Hence, the data is presumed to be of sufficient quality for this analysis.

Some general information about the observation wells is given in Table 2.1. The identification of these wells, used by TNO, are the 'Well identification' names, to increase the readability of the report the 'well numbers' will be used to name the wells throughout this thesis. In Figure 2.5 the hydraulic head fluctuations of the different wells are plotted. No corrections have been made to the density of the water in the observation wells. Although temperature and especially salinity will influence the density of the water, it is not expected to have a big influence. TNO used a density of 1000 kg/m^3 for the calculations of the head.

The influence of the tide is best seen in A2, which is the well closest to the sea. The signal quickly damps out, this is also visible in the signal to noise ratio, which is a lot smaller in B2 and is even worse in C2. A4 has its screen in a deeper aquifer (see Figure 2.2). The time series of A4 is ranging from 2-aug-2018 till 14-dec-2018 and does not have the jump in the data apparent in the other time series, especially visible in the time series of B2 and C2 (see Figure 2.5). These jumps, which were induced by the replacement of pressure loggers, will be further discussed in Section 3.2.2.

Table 2.1: Characteristics of the observation wells, the levels are reported relative to the NAP (Normaal Amsterdams Pijl, a national datum level used in the Netherlands)

Well number	**Well identification**	**Coordinates (RD)**	**Screen depth (top / bottom)**	**Land-surface level (m)**	**Amplitude M2 (cm)**
A2	B42G0022	47029, 412045	Tube 002 (-34.1m / -35.1m)	2.98	37.4
A4	B42G0022	47029, 412045	Tube 004 (-59.3m / -60.3m)	2.98	37.4
B2	B42E0265	47777, 412536	Tube 002 (-29.6m / -30.6m)	-1.1	1.5
C2	B42E0262	49186, 413768	Tube 002 (-30.4m / -31.4m)	-1.5	0.3

With both the Land-surface level and screen depths relative to NAP.

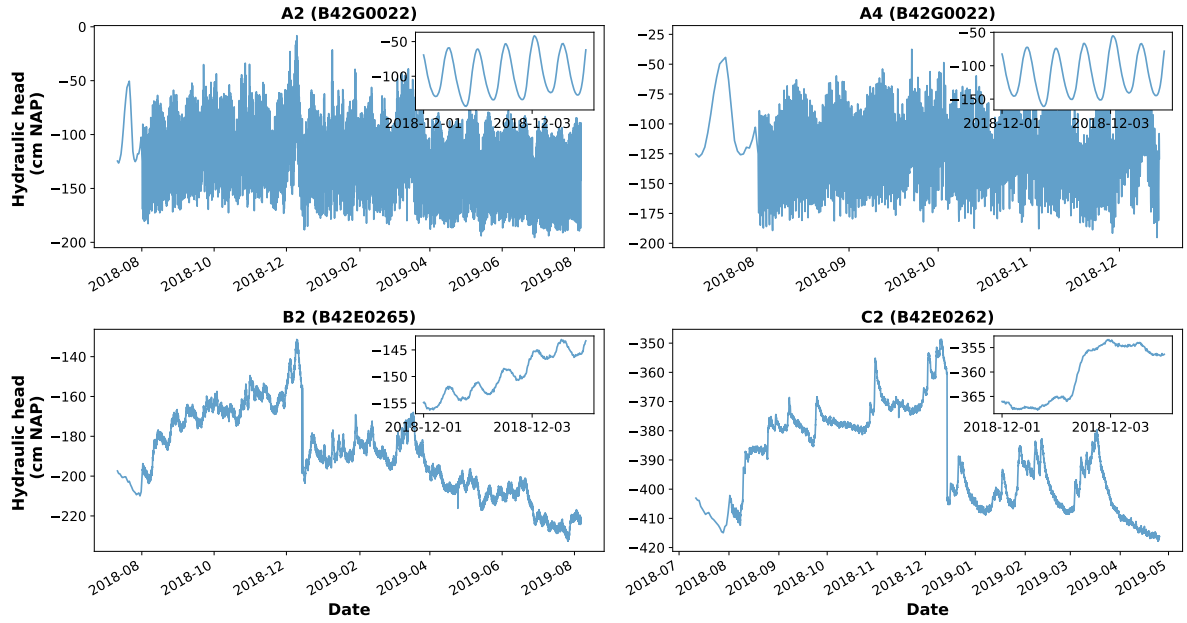


Figure 2.5: Plots of the hydraulic head in the first aquifer layer for four of the observation wells.

2.2.2. Water level in the Oosterschelde

Information on the water level in national waters is collected by Rijkswaterstaat (RWS), the executive agency of the Dutch Ministry of Infrastructure and Water and is, for most measurement locations, available at `www.rijkswaterstaat.nl`. Collection of this data is done at several places in the Netherlands, among which are the measurement locations shown in the small overview map from Figure 2.1. The water level data can be used to compute the astronomical tide, which is the tide that results from gravitational effects, without atmospheric influences. This is computed by RWS by performing a harmonic analysis, using 94 constituents, on part of the water level data (Rijkswaterstaat, 2013). However, data on water level fluctuations caused by (astronomical) tides are not provided for a location near Schouwen (Figure 2.1). Conveniently Piet Lievense (RWS) was kind enough to compute the astronomical water level near Schouwen.

Lievense used the following method for this. For both the M2-amplitude and the M2-phase a (digitized) map of their spatial distribution in the Oosterschelde is available (See Figures C.10 & C.11). The origin of the maps is not clear but probably it was made by analyzing time series of water levels combined with model calculations. The water-level data near the site is then obtained by shifting the data, from a reference location, with the M2-phase difference and subsequently multiplied with the ratio of the M2-amplitude. Similarly, the water level series can be obtained by using two reference locations. The water level series is then determined using distance-based interpolation. With the distance in terms of phase difference and not in terms of geographical distance. In this manner, two water level series were obtained. One from a measuring point at Neeltje Jans (bh2) and one from a weighted average based on bh2 and a measuring point at Stavenisse (STAV), see Figure 2.1 for their locations.

The quality of this transposed data was tested by analyzing the presence, or lack of, logical patterns (e.g. is the amplitude of the computed tidal component in-between the amplitude observed in bh2 and STAV). But more about this in Section 4.1.1.

2.3. Meteorological Data

Meteorological data is gathered by the Royal Netherlands Meteorological Institute (KNMI). On Schouwen there are a few precipitation stations from the KNMI. However, the prevailing geostrophic wind direction is southwest in the Netherlands (Daniels, 2016). Meaning that most rainstorms come from that direction. For this reason the rain data from 'Brouwershaven' (see Figure 2.1), a precipitation station from the KNMI (Royal Netherlands Meteorological Institute), was used since it is located north-east (downwind) from the observation wells.

Wilhelminadorp is the nearest weather station where evapotranspiration is (indirectly) measured (see Figure 2.1). The reference evapotranspiration is calculated with Makkink's equations. The Makkink's reference evapotranspiration is defined as the evapotranspiration of well-watered short grass (De Bruin, 1987).

3

Methodology: Data preparation

In this Chapter the methodology that is used to prepare the data, for the inverse modeling, is described. First, an harmonic analysis, which is a method to decompose the different components of the tide, is presented. In Section 3.2 the unwanted disturbances in the observations are discussed. Subsequently, two methods to filter the data are presented, the 'Pastas' model in Section 3.3 and a Butterworth filter in Section 3.4. Lastly, in Section 3.5 a quick (graphical) method to determine the M2 amplitude and phase is presented.

3.1. Harmonic analysis

The tide consists of several 'harmonics', these are sinusoidal functions with a distinct frequency, amplitude and phase. The tidal signal, in the Oosterschelde, consists of a few major and many minor harmonic components (often also called constituents). For the tidal method, the amplitude and phase of the different constituents need to be deduced from the tidal signal. This procedure is called a 'harmonic analysis' and to perform it, the frequencies of the constituents of interest need to be known. Obtaining the frequencies of interest can, for example, be done by using a Fourier transform technique in which the frequencies of interest are determined from analyzing the power spectrum (Godin, 1972). However, in the Netherlands, the most important tidal frequencies are precisely known, so there is little to gain from a Fourier transform technique.

To determine the amplitudes and phases of the different constituents, the exact frequencies of the constituents are used in a least squares method, described by e.g. Godin (1972) and Foreman and Neufeld (1991). The astronomical tide can be described by the following equation:

$$y_k = A_0 + \sum_{j=1}^{M} A_j \cos(\omega_j t_k - \Phi_j), \tag{3.1}$$

where A_0 is the average waterlevel; A_j, ω_j and Φ_j are the constituents amplitude, frequency and phase respectively, for constituent j. y_k and t_k are the observation and time respectively at time k. This formula is nonlinear for Φ, which complicates the algebra needed to find the solution for this least squares problem. Moreover computing the unknowns for a nonlinear problem is computationally more demanding compared to a linear problem. Fortunately, this formula can be made linear by solving for the unknowns C_j and S_j as follows:

$$A_j \cos(\omega_j t_k - \Phi_j) = C_j \cos(\omega_j t_k) + S_j \sin(\omega_j t_k). \tag{3.2}$$

Subsequently A_j and Φ_j can be calculated with:

$$A_j = \sqrt{(C_j^2 + S_j^2)} \quad \text{and} \quad \Phi_j = \arctan(S_j / C_j). \tag{3.3}$$

Note: Depending on the programming language, atan2 or a similar function should be used for arctan, so that $-\pi < \Phi \leq \pi$ instead of $-\frac{1}{2}\pi \leq \Phi \leq \frac{1}{2}\pi$.

This system of equations is overdetermined as there are more knowns than unknowns, the least squares method is used to minimize the sum of the squared residuals. Or in matrix-vector form, minimize:

$$||\mathbf{B}\vec{x} - \vec{y}||^2, \tag{3.4}$$

where **B** is a matrix containing equations 3.1 substituted with equations 3.2, $\vec{x}$ is a vector containing the unknowns (A_0, C_j and S_j) and $\vec{y}$ is a vector containing the water height data (y_k). This multiple linear regression algorithm can be used to analyze both the Oosterschelde water level and the groundwater head data.

Lastly, note that many papers analyzing the tide (i.e. Godin, 1972; Godin, 1986; Pawlowicz et al., 2002) consider corrections for long period modulations and perturbations in solar, moon and earth orbits. However, these changes are gradual so that they can be considered as effectively constant for time series up to a year (Pawlowicz et al., 2002). So for the purpose of this paper it was considered unnecessary to go through the trouble of working with such modulations.

3.1.1. The constituents considered

The tidal frequencies used in this analysis, are the 31 tidal components identified as most important along the Dutch coast by Rijkswaterstaat (Rijkswaterstaat, 2013). All considered constituents are listed in Appendix B.1 along with their frequency.

The minor harmonic constituents are often hard to distinguish from each other as their frequency is close and time series are often short and/or containing too much error or other influences (Merritt, 2004). The Rayleigh's resolution limit (see equation 3.5) indicates if two frequencies can both be resolved with the data at hand. This is a rule and not a strict law but in general, violation leads to unreliable results (Godin, 1972). If a smaller constituents frequency is less than a Rayleigh's resolution apart from a more important constituent, then these constituents may be unresolvable (Pawlowicz et al., 2002). Below the Rayleigh resolution limit after Deltares (2014):

$$\Delta\omega = \frac{2\pi}{T}, \tag{3.5}$$

where $\Delta\omega$ is the Rayleighs resolution limit in radians per second, and T the length of the time series in seconds. Note that $\Delta\omega$ is also the smallest frequency that can be resolved for a given time series (Deltares, 2014). To use all the 31 constituents regarded as most important along the Dutch coast, a time series of at least 366 days is required. This condition is only satisfied for the time series in wells A2 and B2 who are both 369 days long. The time series in well C2 is too short to both estimate M2 and MPS2 accurately. So MPS2 is not considered in well C2. Although, dropping MPS2 from the calculation is not expected to contaminate the solution because the amplitude of MPS2 is approximated to be negligibly small (4.5e-3 cm) in well C2, based on the calculated attenuation (in %) from M2 and the calculated amplitude of MPS2 in the Oosterschelde. Also note that, although time series are often not long enough to resolve all constituents accurately, these constituents are still present and their hidden presence is more troublesome than calculating them directly if the time series is of sufficient duration (Godin, 1972). And thus dropping MPS2 from the harmonic analysis for all observation wells would not be sensible.

The time series in well A4 has a much shorter record (134 days), which is too short to satisfy the Rayleighs resolution limit for all of the 31 constituents. Simply including all constituents will lead to unreliable results (Deltares, 2014). On the other hand, their hidden presence can contaminate the rest of the solution. The best solution for this problem, according to Pawlowicz et al. (2002), is to apply inference (also called astronomical coupling). To apply this method an amplitude ratio and relative difference in phase between the two unresolved constituents is needed from nearby data (Pawlowicz et al., 2002). In this case, the data from well A2 was used. To make the astronomical coupling an integral part of the least-squares solution, an extra term must be added to the right-hand side of Equation 3.2 for the jth constituent from which another constituent is inferred, see Appendix A.1 for a full derivation.

3.1.2. Standard deviation

The standard deviation of parameters, estimated with a least-squares multiple linear regression model, can be calculated by means of matrix algebra. Assuming that the estimator$\vec{x}$ is unbiased (Kutner et al., 2005), or in equation form:

$$E\{\vec{x}\} = \beta, \tag{3.6}$$

where E is the expectation and β is a parameter. The estimated variance-covariance matrix $\mathbf{cov_{xx}}$ is given by:

$$\mathbf{cov_{xx}} = MSE(\mathbf{B}^T\mathbf{B})^{-1}, \tag{3.7}$$

where $\mathbf{B}$ is a matrix of constants (same as in equation 3.4) and with the mean squared error (MSE) given by:

$$MSE = \frac{1}{n-p}\sum_{i=1}^{n}(y_k - y'_k)^2, \tag{3.8}$$

where y_k and y'_k are the observed and modeled water height respectively at time k. The MSE has $n-p$ degrees of freedom, with n the number of observations and p the number of parameters to be estimated.

The obtained variance-covariance matrix ($\mathbf{cov_{xx}}$), of the unknowns C_j and S_j, are used to construct cumulative distribution functions for the amplitude and phase estimates of the different constituents (see Section 3.1.3).

3.1.3. Monte Carlo simulation

Monte Carlo simulation is used to estimate the average and standard deviation of the amplitude and phase for each constituent. An additional benefit is that probability distributions can be created, which can be used to easily analyze the data visually.

It is assumed that the error present in the data of TNO and Rijkswaterstaat is random and has a normal probability distribution. This assumption is justified in Chapter 4 for the Oosterschelde data and data in well A2 (for one of the filters). Because both show a normally distributed error which implies independent and normal observations (Kutner et al., 2005). For the other observation wells this assumption is presumed to be correct as well.

The probability distribution of the output of a model can be calculated using numerical integration, with a Monte Carlo simulation and in some cases exactly with analytical formulas (Jonkman et al., 2017). Numerical integration can be suitable if the number of parameters is limited (preferably ≤ 2), but can become quite computationally expensive, as the computation time increases exponentially with the number of stochastic variables. Analytical formulas are only suitable for some simple cases.

Furthermore, there are approximations where the joint probability function is simplified. This is computationally more efficient, but the distribution can deviate from the true probabilistic distribution if requirements like small variances of the data are not met. However, this requires repeatedly checking the statistical output. Therefore, Monte Carlo simulation is the selected method to calculate the probability distributions in this report.

For the Monte Carlo Simulation (MCS), random samples have to be generated as input variables. This can for example be done by using the cumulative distribution function (CDF) of the data. In this report, a multivariate normal distribution is obtained by using the estimated unknowns (C_j and S_j) and the associated variance-covariance matrix (see Section 3.1.2). To create m random variables from a multivariate normal distribution, the Cholesky decomposition method can be used (see e.g., Thomopoulos, 2012).

$$\mathbf{X} = C\mathbf{Z} + \vec{\mu}, \tag{3.9}$$

where $\mathbf{X}$ is an $m \times n$ matrix (with the variables as rows (m) and the number of samples (n) drawn from the normal distribution). Matrix $\mathbf{Z}$ is an $m \times n$ matrix, generated using the standard normal distribution, $z \sim$ N(0,1). m random variates are generated (z_1,...,z_m) and inserted into the rows of matrix $\mathbf{Z}$. $\vec{\mu}$ is a vector with the optimized parameters and $\mathbf{C}$ is the Cholesky decomposition of the covariance matrix defined as

$$\mathbf{cov_{xx}} = \mathbf{C}\mathbf{C}^T. \tag{3.10}$$

where $\mathbf{cov_{xx}}$ is the variance-covariance matrix. Note that $\vec{u}$ is the same as $\vec{x}$, from Equation 3.4, but without the estimated average water level.

3.2. Noise in the groundwater data

The water level data from the Oosterschelde is already astronomical, so noise from atmospheric influences or influence from river discharge has been removed already (Rijkswaterstaat, 2013). However, observations of the groundwater head contain head fluctuations caused by other influences (e.g. rain and evapotranspiration) than the astronomical tide in the Oosterschelde (see Figure 2.5). For the purpose pursued in this report head fluctuations not caused by the astronomical tide are considered as noise. Two types of noise can be distinguished, noise in the form of physical factors from which their attribution is not yet determined and white noise. Two methods are analyzed for removing noise: by means of flattening the frequency response for frequencies that are of no interest (by using a Butterworth filter; see Section 3.4) and by determining the response to known influences and removing these. The latter is done with the use of a software program called 'Pastas' (see Section 3.3). Both methods are compared to investigate which method is more effective and efficient in reducing the noise of the data.

The time series of the groundwater fluctuations are more than 9 months long. Moreover, the tidal components have distinct frequencies while recharge (rain minus evaporation) forcing does not. So removing noise is not expected to increase the accuracy of the amplitude and phase estimates. However, reducing the error variance reduces the mean square error and thus the standard deviation of the estimated parameters (see Equations 3.7 & 3.8). In other words, reducing the noise, present in the data, increases the precision of the estimates.

3.2.1. Atmospheric & earth tides

Not only oceanic tides can be encountered, atmospheric and earth tides can also be measured. These forcings are potentially an extra source of noise. Because a lot of the cited papers write about earth tides and have occasionally included atmospheric tides into their analysis, it is appropriate to briefly address both topics.

Atmospheric tides (sometimes called barometric tides) were analyzed in some studies, that also studied ocean-tides, for example by Merritt (2004) to estimate the loading efficiency, which is the ratio between the instantaneous head response in the groundwater below a waterbody and the water-level changes in the waterbody (Bakker, 2016). Merritt found decent results for some data sets and inaccurate or meaningless results for other data sets. According to Godin (1972) the apparent oscillation of atmospheric tides is about 1 millibar. However, the bulk of this is due to non-tidal forcings such as heating and cooling of the land. The true atmospheric tide is in the order of 0.03 millibar (Godin, 1972). Such a small variation is easily concealed by stronger (daily) perturbations in the atmosphere. Shih et al. (2000) did a spectral analysis on water level fluctuations in coastal aquifers, coupled with atmospheric pressure fluctuations and found that the latter do not significantly affect the semi-diurnal and diurnal tide. Dong et al. (2015) mostly found the same and further state that barometric influences mainly act in a low-frequency domain (i.e. periods $\geq$ 22 hours), which is characteristic for longer period climatic patterns (Spane, 2002). Therefore, the atmospheric tide is not considered in this paper.

Earth tides are also mentioned in several papers that use the tidal method (e.g. Carr, 1971; Merritt, 2004; Dong et al., 2015). Expansion and contraction of the soil matrix causes pressure in the soil matrix to rise and fall. This effect is especially felt in deep and confined layers (Merritt, 2004). Water level fluctuations caused by earth tides are of the order of 1 cm in amplitude (Van Der Kamp and Gale, 1983; Bredehoeft, 1967). Despite the amplitude being significant, especially for wells located further inland, no paper on the tidal method known to the author, mentions earth tides as a possible source for error. This is somewhat curious, although according to Gregg (1966) and Meyer (1974) earth tides are exactly out of phase with oceanic tides. Theorizing that as the moon or sun passes over an area, the bulge in the earth's crust would cause lateral stretching, and an increased pore size, which causes the hydraulic pressure to decrease (Merritt, 2004). Moreover, unlike most studies performing the tidal method (e.g. Trefry and Bekele, 2004; Merritt, 2004), amplitude and phase data from shallow water constituents were obtained. Fluctuations, in the groundwater, with frequencies of shallow water constituents cannot be caused by earth tides, because they result from non-linear interactions (i.e. advection and bottom friction) occurring in shallow waters (i.e. very small depth compared to typical length dimensions) and not directly from gravitational pull (Deltares, 2014; Battjes and Labeur, 2017). Both the latter and the fact that earth tides are out of phase with oceanic tides can help to make sure that fluctuations are purely caused by the propagation of oceanic tides (see Section 4.5.2). Lastly, it is worth noting that analyzing earth tides can also yield information about formation parameters, see for example Van Der Kamp and Gale (1983).

3.2.2. Step in the data

Due to the replacement of a pressure logger, a step in the water level data is present, as is clearly visible in Figure 2.5 for wells B2 and C2. Thereafter, it was also found that this step is accompanied by a temporal error. Which became apparent when the modeled and observed fluctuations were observed. The extent of the time shift was determined as follows: The M2 and seven other important constituents were fitted through a 3 months time series before and after the step. A phase shift of 61, 58 and 48 minutes were found for A2, B2 and C2 respectively. It is presumed that a shift of 60 minutes had occurred in all observation wells. Discrepancies of the phase shifts from these 60 minutes are assumed to be caused by uncertainty in the phase estimated, due to a low signal to noise ratio.

It remains uncertain if the time series before or after the step was shifted. But it was assumed that the shift had occurred during the replacement of the divers because there where two (three in well C2) double measurements and the replacement of the diver (i.e. reading the data, resetting the logger and waiting for the time to be a quarter of an hour) also takes about half an hour. So the most plausible is that the time series in front of the step is the correct one.

Some hand measurements have been done as well. Unfortunately, the hand measurements and measurements from the data logger do not correspond, for C2 this discrepancy was even more than 20 cm, and could therefore not be used to determine the correct time of the time series.

The data was further adjusted by removing the data with two head observations at the same date. After which, the step in the data was removed. In Figure 3.1 the data before and after these adjustments is plotted. The height of the step was estimated with 'Pastas' (see Section pastas method), and is estimated to be: 23.2, 41.3 and 41.4 cm for well A2, B2 and C2 respectively. Alternatively, the step can be estimated by calculating the difference between the double data points. This gives comparable results for wells B2 and C2. For A2 this method estimates the step to be 9cm bigger, so that the head data corrected for the step (see Figure 3.1, left plot) would be 9cm higher after the step. This would give a less smooth transition of the time series before and after the step. Therefore, the height of the step estimated with 'Pastas' was used to correct for the step of the unfiltered data. Lastly, the Butterworth was found to be less suitable for estimating the step. Therefore, the step estimated with 'Pastas' was also subtracted from the signal, before applying the Butterworth filter. But more about this in Section 4.2.

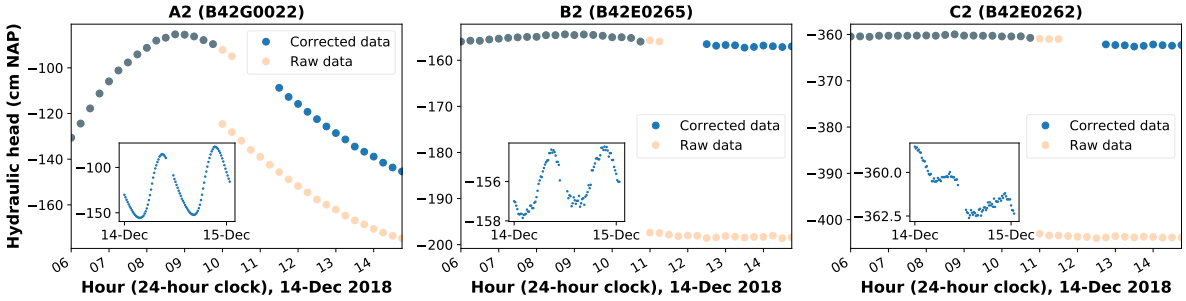


Figure 3.1: Plots of the hydraulic head in the first aquifer layer for three of the observation wells. Note that overlapping orange and blue dots appear as gray dots.

3.3. Removing noise with 'Pastas'

Transfer Function Noise (TFN) modeling can be used to transform input series to an output series using a statistical model (Collenteur et al., 2019). Such a model can be used to decompose the fluctuations observed in the groundwater head to different contributions of hydrological stresses (e.g. rain, evaporation and surface water level fluctuations). TFN modeling is implemented in a software program called 'Pastas', which is used in this part of the research. Pastas is an open source (Python) package for simulating and analyzing groundwater time series analysis (Collenteur et al., 2019).

The Pastas model allows for multiple stresses. The head response to a different stress is simulated through convolution of the stress (e.g rain) with a response function. Pastas includes several response functions (e.g. Gamma, Exponential and Hantush). The scaled gamma function is often used for rain and evaporative stresses but can be used for other stresses as well because it is versatile in its shape (Collenteur et al., 2019).

The exponential response function can be used for a stress with an almost immediate effect on the head. In this research, the scaled gamma response function (see Equation 3.11) was used to simulate the effect, on the hydraulic head, of the astronomic and non-astronomic fluctuations in the Oosterschelde, precipitation and evaporation. The precipitation and evaporation are modeled with gamma response functions, only differing by a factor. The scaled gamma response function is:

$$\theta(t) = A\frac{t^{n-1}}{a^n\Gamma(n)}e^{-t/a} \quad \text{with} \quad t \geq 0, \tag{3.11}$$

where A is a scaling factor, a & n are shape factors and Γ is the Gamma function. The step in the data is modeled using an exponential response function (see Equation 3.12).

$$\theta(t) = A(1 - e^{-ta}) \quad \text{with} \quad t \geq 0, \tag{3.12}$$

where A is a scaling factor and a is a shape factor (both formulas after Collenteur et al., 2019). In the optimization a was fixed to be 1e-20 so that it effectively simulates an immediate drop in the hydraulic head.

Pastas also includes a noise model. A noise model is intended to make all noise white, which is needed to ensure that the statistics and thus also the confidence intervals, around the estimated parameters are reasonable. Unfortunately, the noise model could not be used for the times series of well C2. In general noise models are known to have more troubles with high-frequency data. So for well C2 autocorrelation present in the data could not be removed.

Subsequently, the parameters of the response functions are estimated with a least-squares approach. In Pastas the minimization algorithms are local optimization schemes. However, because of the presence of a lot of local minima, it was hard to find the global optimum, even with reasonable initial value estimates. Therefore, the differential evolution algorithm (global optimizer) from LMFIT was added to the code (Newville et al., 2014).

Below an artificial example to demonstrate how Pastas works. In Figure 3.2 (a,b,c) three hydrological stresses that affect the head response are plotted. Figure 3.2 (d) shows the head response itself. The response functions are simulated with the scaled gamma function (see Equation 3.11). The response function of the evaporation and precipitation is the same only differing by a factor (called 'Recharge_f' in the table of model information in Figure 3.3). The upper plot of Figure 3.3 shows the modeled and observed head, with the residual plotted below it. The plots in the third row show the estimated head response to the sea level, with the associated optimized response function plotted on the right. The lowest plot shows the estimated head response to the recharge (combined effect of the rain and evaporation) with the associated optimized response function plotted next to it. The estimated parameters and their estimated standard deviation are shown in the table.

As demonstrated here, Pastas is able to quantify the effect of different stresses on the hydraulic head. Subsequently, head fluctuations caused by other stresses than the astronomical tide, in this case the recharge, can be removed from the hydraulic head data.

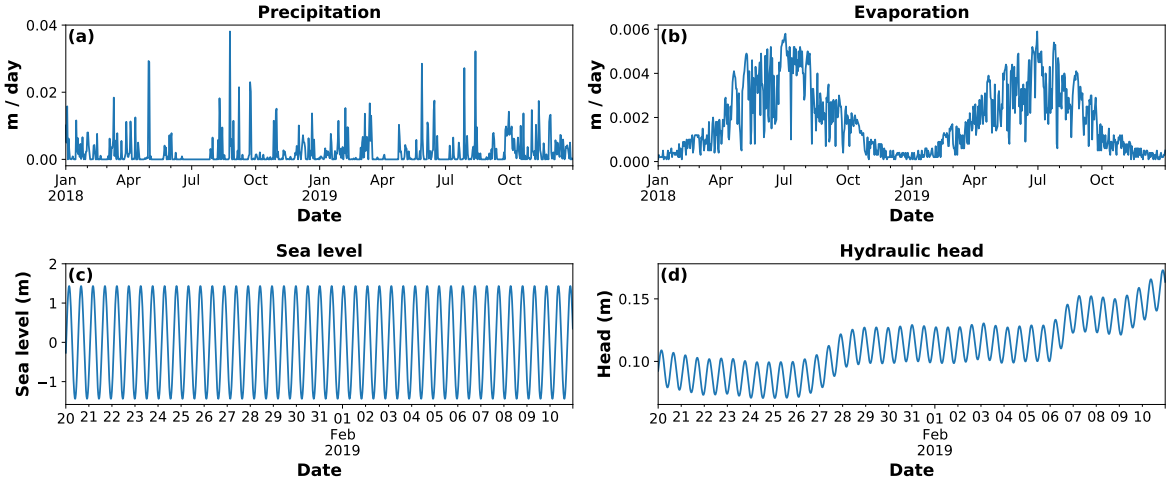


Figure 3.2: Three hydrological stresses (plot a,b,c) and the associated hydraulic head (d).

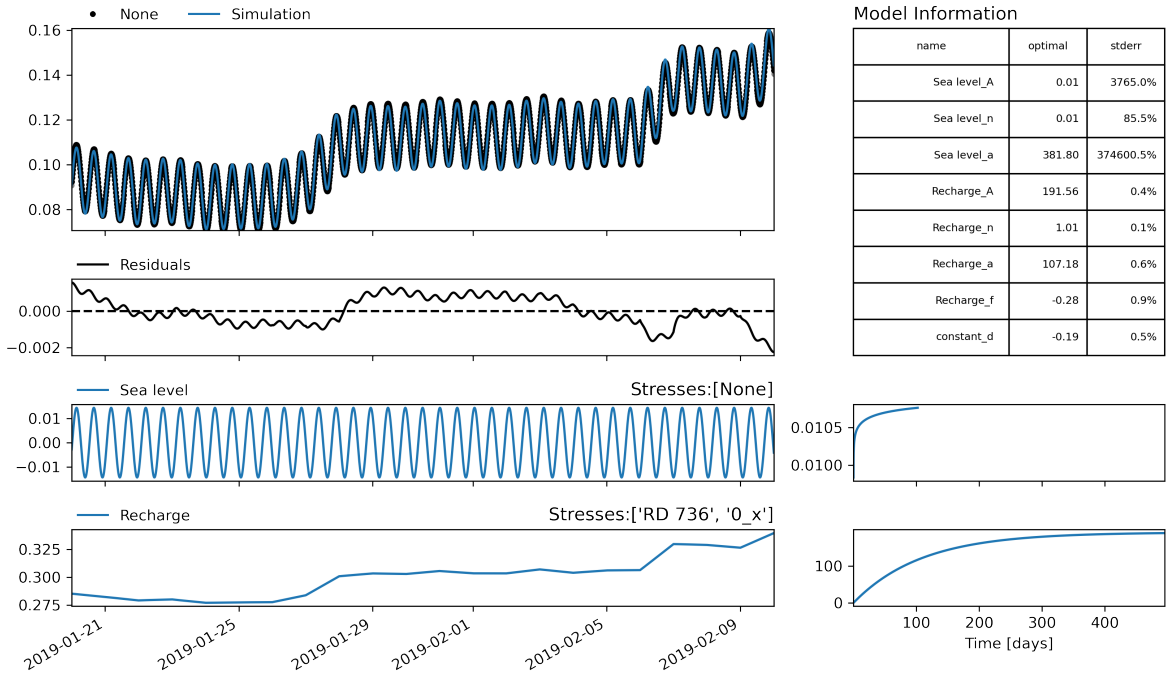


Figure 3.3: Decomposition of the groundwater fluctuations with Pastas for an artificial times series, results are in meter.

3.4. Removing noise with a Butterworth filter

We can remove fluctuations caused by known physical factors, which is what is done with the Pastas model. It is also possible to acknowledge that all frequencies below a threshold are of no interest so that these can be removed. Trefry and Bekele (2004) also used such a 'filter approach', they focused on high frequency forcing and ignored (low-frequency) seasonal and climatic effects. To remove the irrelevant frequencies, a low-pass Butterworth filter is used. An example of a more commonly known low-pass filter is a moving average. The Butterworth filter, well known in electrical engineering, is maximally flat in the passband and has virtually no distortions of the low-frequency components (Roberts and Roberts, 1978). The Butterworth filter is defined by the square of its transfer function, which represents the relation between the input signal and output signal. The Butterworth filter is given by the expression:

$$|H_B(i\omega)|^2 = \frac{1}{1+(\omega/\omega_c)^{2n}}, \tag{3.13}$$

where i is the imaginary unit, ω_c is the cutoff frequency and n is the order of the filter, which is halve the order of the Butterworth filter. In Figure 3.4 the magnitude of the frequency response (gain), of the used Butterworth filter, is plotted against the frequency. Note that the cutoff frequency from equation 3.13, which was set at 4e-6 rad/s, is where the gain equals the square root of 0.5. The passband, transition band and stopband are also indicated in the figure. The passband is the range of frequencies that can pass through the filter, the stopband is the range of frequencies for which the filter does not allow the signal to pass through and lastly the transition band is the range of frequencies in which the transition between the startband and stopband takes place.

For digital filters in general, the transfer function can be expressed in terms of coefficients b and a as:

$$H(z) = \frac{b_0 + b_1 z^{-1} + ... + b_n^{-n}}{1 + a_1 z^{-1} + ... + a_n z^{-n}}, \tag{3.14}$$

with z being (in general) a complex variable coming from the Z-transform of the finite data sequence :

$$Z[x_k] = \sum_{k=1}^{N-1} x_k z^{-k}, \tag{3.15}$$

where x_k ($k = 0, 1, ..., N-1$) is the (unfiltered) finite data sequence.

The filter coefficients (b and a) are determined with the signal sub-package from the SciPy library of Python (Virtanen et al., 2020). A 6th order transfer function, with a cutoff frequency of 4e-6 rad/s resulted in a satisfactory transfer function. Subsequently, SciPy was also used to apply the transfer function forward and backward, resulting in a 12th order Butterworth filter. Note that a higher-order filter ensures a steeper decline of the transfer function. For more details see i.e. Roberts and Roberts (1978).

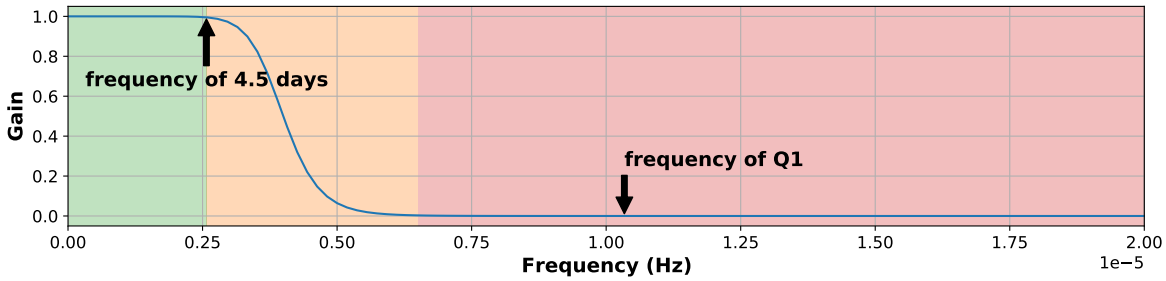


Figure 3.4: The used Butterworth filter, which is its transfer function squared. The passband, transition band and stopband are colored in green, orange and red respectively. Note that the start and end of these bands are not strictly defined, so the colored rectangles are merely indicative.

To illustrate the effect of the Butterworth filter a small example was made. A cosine function with a long period (the SM constituent; i.e. 14.8 days) was added to a cosine with a small period (the M2 constituent; i.e. 1.4e-4 rad/s). Subsequently, the signal was filtered with the Butterworth filter, from which the gain against the frequency is plotted in Figure 3.4. Figure 3.4 shows that frequencies with a period around 4.5 days and higher can pass through the filter, while signals with a frequency with a small period such as M2 are not allowed to pass. Note that from the constituents that we are trying to resolve in the groundwater observation wells, Q1 is the constituent with the biggest period.

The obtained filter is plotted in Figure 3.5 (a). Subsequently, the obtained filter is subtracted from the raw data so that a filtered data-set is obtained. In Figure 3.5 (b) both the raw and filtered data are plotted. Lastly, note that because the low-pass filter was subtracted from the data, technically a high-pass filter is applied to the data.

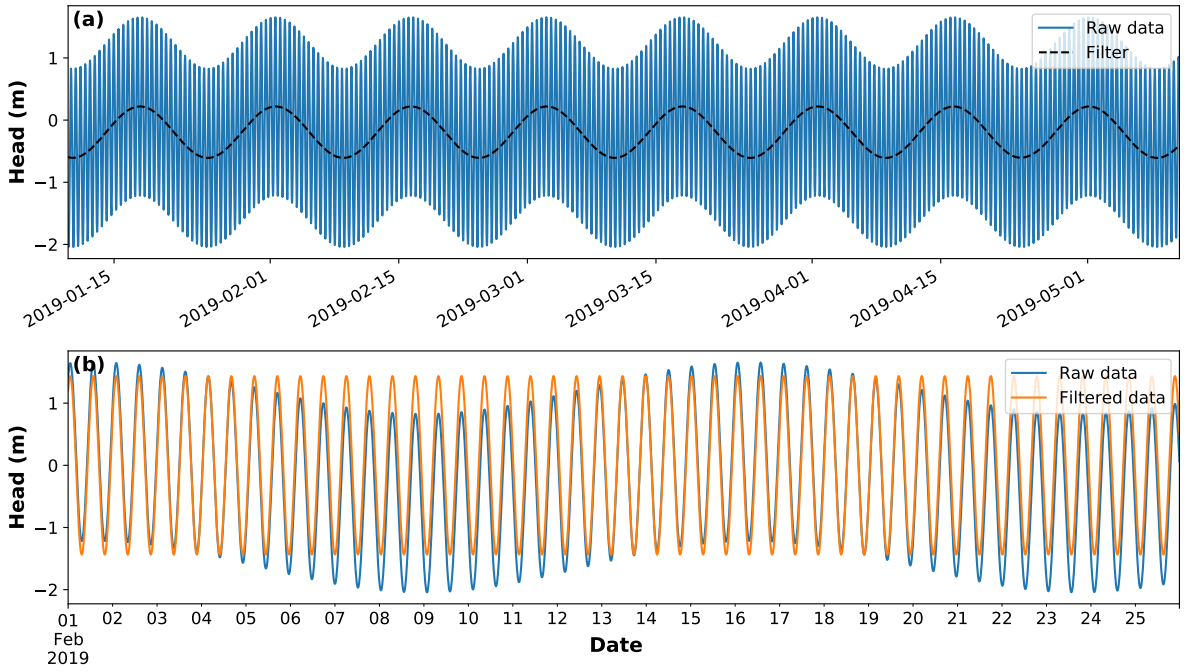


Figure 3.5: Plots of the raw data and the obtained filter (plot a) and a plot with both raw data and the filtered data (plot b).

3.5. Graphical determination of the Amplitude & phase

Obtaining the phases and amplitudes of the constituents can be a time-consuming process. Since the M2 tide is so dominantly present in the Oosterschelde the question arises if good estimates of the phase and amplitude from the M2 constituent can be obtained by graphically inspecting the time series, instead of using an harmonic analysis. To estimate the amplitude graphically, the difference between a few maxima and the subsequent minima, of the observed groundwater fluctuations, is divided by two. The time interval of a few maxima and minima between the reference date (02-aug-2018 00:00) is used to estimate the phase. This method was for example used by Wesseling (1960). In section 4.4 the amplitude and phase of the M2 tide obtained both graphically and with the harmonic analysis are presented.

4

Results: Data preparation

In this chapter, the results from the data preparation are presented. The first section contains the results from the harmonic analysis of the Oosterschelde water level data. In Section 4.2, the results for the groundwater levels from the harmonic analysis with the use of different 'filters' are compared. In Section 4.3 the obtained phase and amplitude estimates are analyzed. Subsequently, in Section 4.4, the amplitude and phase of the M2 constituent estimated with the harmonic analysis are compared with a graphically determined amplitude and phase. The chapter is concluded with a discussion of the results. Lastly, the main purpose of this chapter is to determine what data can be used for the inverse modeling part of this report.

4.1. Harmonic analysis of the Oosterschelde water level data

4.1.1. The translated water-level data in the Oosterschelde

As described in Section 2.2.2 the water level series at the site are obtained both from reference location 'bh2' and from both reference location 'bh2' and 'STAV' (see Figure 2.1 for their locations). In this section, the results of both translations are analyzed, with the purpose of deciding which data-set to use for the rest of the report and to give an indication of the error accompanied by this translation.

In Figure 4.1 the probability distribution of the amplitude and phase of the M2 constituent is plotted for the reference and calculated locations, the graphs of the other constituents can be found in Appendix C.4. Some of the observed patterns are discussed below. Note that the harmonic analysis (Section 3.1) and a Monte Carlo simulation (Section 3.1.3) were used to create these probability density plots.

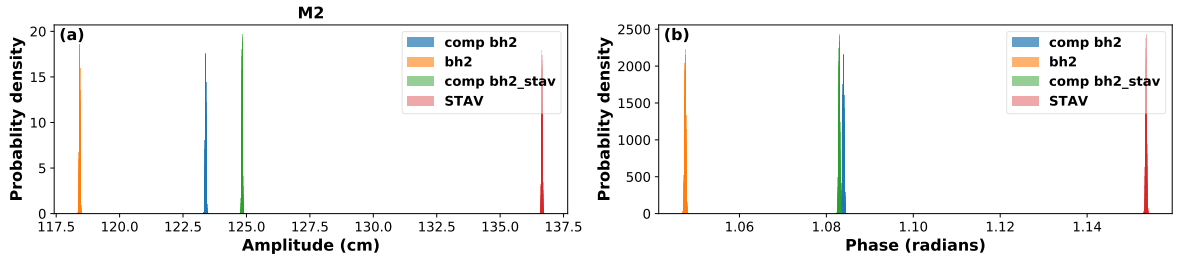


Figure 4.1: Probability density graphs for the amplitude (plot a) and phase (plot b) of the M2 constituent in the Oosterschelde for the study site (comp) and the measuring points at Neeltje Jans (bh2) and Stavenisse (STAV). The probability density functions are calculated with a harmonic analysis of the astronomical tide. In orange and red the decomposed tide based on the astronomical tide at measuring points bh2 and Stavenisse (STAV) respectively (see Figure 2.1). In blue the computed water level (see Figure 2.1) translated from bh2 and in green the computed water level calculated from an interpolation based on the translated time series from bh2 and STAV (see Section 2.2.2).

The phase increases from bh2 to STAV for all constituents except Q1, 3MS2 and μ2. For all constituents (except μ2) the computed tide based on both STAV and bh2 is approximately 2 times closer to bh2 than to STAV, for both the amplitude and phase estimates. For some of the constituents (2MN6, M6 & 2MS6) the

probability distribution of the amplitudes calculated at the study site is not between those for bh2 and STAV for both computed time series at the study site (see Appendix C.4). Using these constituents, even after translation, would involve too much uncertainty, therefore the results for these constituents are considered to be unreliable.

The computed tide based solely on bh2 is not between those for bh2 and STAV for multiple constituents both for the amplitude and/or phase. Moreover, the amplitude and phase are frequently estimated to be much closer to bh2 compared to STAV. While, based on Figures C.10 & C.11, the computed amplitude and phase are expected to be approximately 2 times closer to bh2 than to STAV. Because of the above considerations, the computed time series based on both bh2 and STAV is used to calculate the amplitude and phase at the study site.

From the probability plots in Figure 4.1 and Appendix C.4 a rough estimate of the error made by translation can be acquired. For that, the results were assessed based on two considerations. First, both methods (i.e. obtaining the water level at the site from just bh2 or from both bh2 and STAV) should be reasonably able to estimate the water-level near the site and therefore, a large difference between both estimates indicates a high level of uncertainty. Second, as mentioned above, almost all constituents have an amplitude and phase estimate roughly 2 times closer to bh2 than STAV, which is also what would be expected based on Figures C.10 & C.11. Departures from this expectation indicate a higher level of uncertainty.

Based on this rough analysis, the amplitude and phase of the constituents with a frequency around 2.8e-4 (3MS4, MN4, M4, MS4, Mk4) are regarded as unreliable. The relative error of this group of constituents is approximately 8% for the amplitude and roughly has an absolute error of 0.15 radians for the phase. Whereas for the other constituents the relative error of the amplitude is on average estimated smaller than 2%, while the absolute error of the phase is estimated around 0.02 radians. Consequently, data from this group of constituents with a frequency around 2.8e-4 rad/s will not be used for the inverse modeling part of this report. Lastly, note that (as mentioned above), the constituents with a frequency around 4.2e-4 rad/s (2MN6, M6 & 2MS6) are also considered as unreliable.

4.1.2. Results of the harmonic analysis from the Oosterschelde

A harmonic analysis was performed on a 2 year-long (2018-2019) time series, calculated from the measuring locations bh2 and STAV was used (see Section 4.1.1). In Table 4.1 results for some of the main constituents are presented. The standard deviations of the estimated parameters are quite small, especially for the constituents with a large amplitude, such as M2. Both the modeled and observed water level fluctuations are plotted in Figure 4.2(a). The model is able to produce quite a good fit for the observations. A complete list of the constituents used in the model can be found in Appendix B.1.

For the linear least-squares model to be appropriate the residual needs to be reasonably normally distributed and independent (Kutner et al., 2005). The former is plotted in Figure 4.2(c) and this seems to be the case. Autocorrelation of the residual can be detected by plotting the residual against time, which is an effective though subjective method (Kutner et al., 2005). This is done in Figure 4.2(b). The residual clearly shows a non-random pattern, which is confirmed as well with an ACF (autocorrelation function) plot (d).

Table 4.1: Amplitude and phases of the main constituents in the Oosterschelde

	O1	**μ2**	**N2**	**M2**	**2MN2**	**S2**	**K2**
ω (rad/s)	6.76e-05	1.356e-04	1.379e-04	1.405e-04	1.432e-04	1.454e-04	1.458e-04
Period	25h 49min	12h 52min	12h 39min	12h 25min	12h 11min	12h	11h 58min
Amplitude (cm)	9.4	12.71	19.5	124.83	11.1	30.19	7.94
SD Amplitude	0.22%	0.16%	0.11%	0.02%	0.19%	0.07%	0.26%
Phase (rad)	-3.14	1.42	1.49	-2.53	2.72	2.64	-1.63
SD Phase	0.07%	0.11%	0.07%	0.01%	0.07%	0.03%	0.16%

SD is the standard deviation. The phase is calculated back to the phase at 02-aug-2018 00:00.

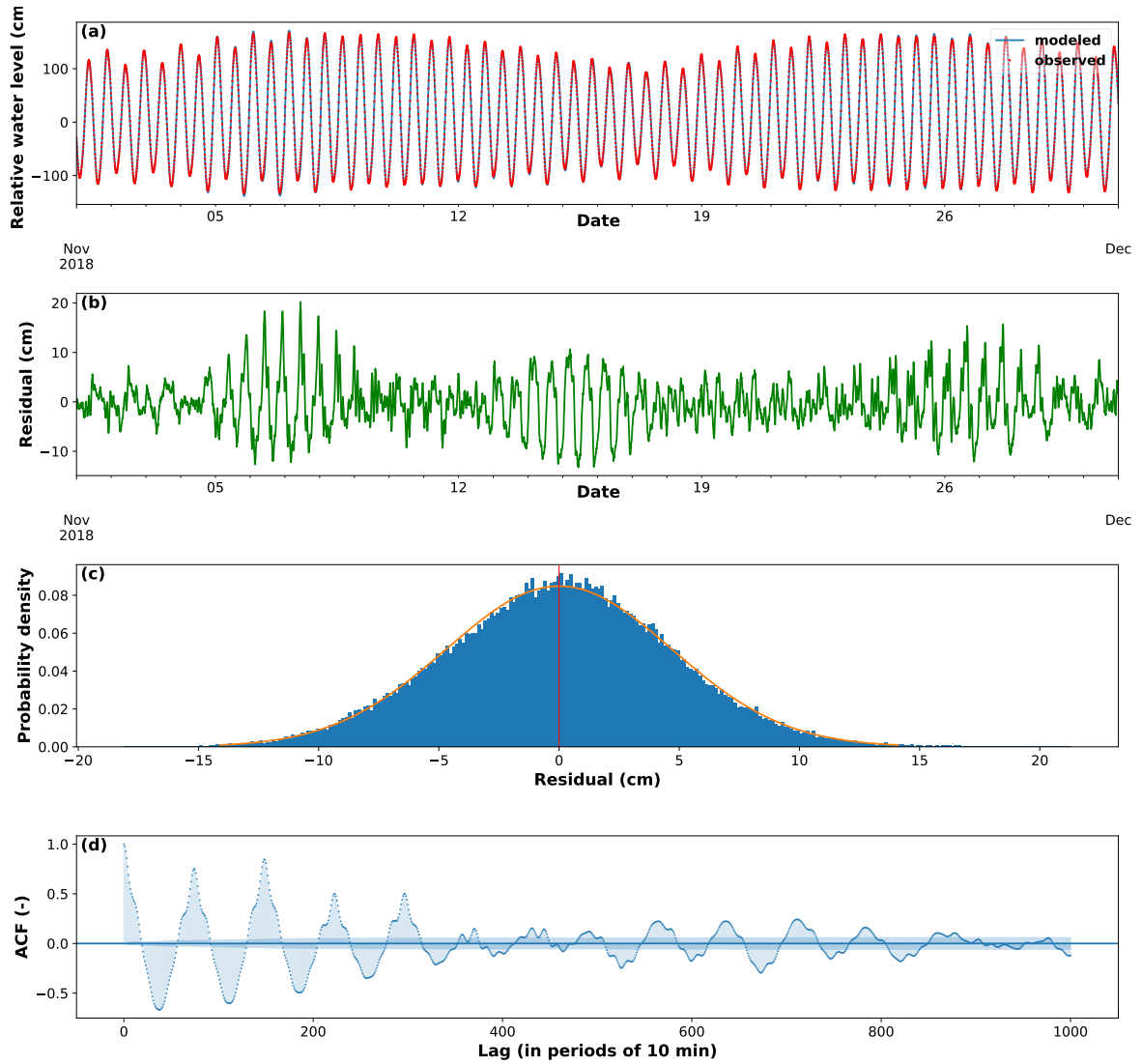


Figure 4.2: Plots of the modeled and observed water fluctuations in the Oosterschelde (a) and the associated residual (b) are plotted for November 2018. With the probability density of the residual for the whole time series in Figure (c) and an ACF plot in Figure (d).

4.2. 'Pastas' vs Butterworth filter residual removal

In this section, the results of the harmonic analysis for the different observation wells are presented. For the analysis different filtering techniques are used to preprocess the observations. The results of applying the filter to the data are presented in Section 4.2.1. In section 4.2.2 the results obtained without using a filter ('Raw data') are compared with results obtained by applying a Butterworth and a Pastas filter. The purpose of this section is to compare the different filtering techniques and simultaneously checking the quality of the decomposition fit.

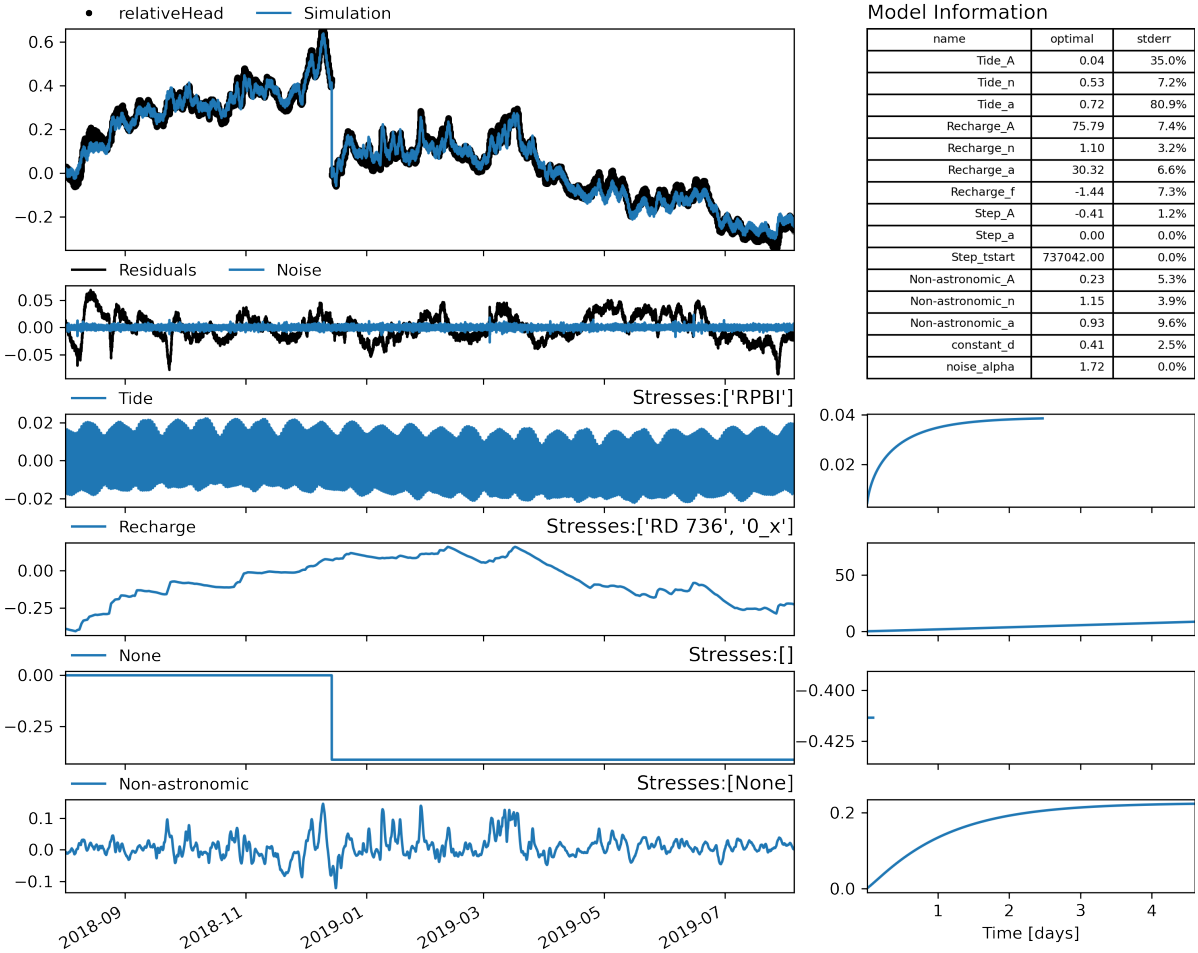
4.2.1. Filtering the data

In Figure 4.3 the result of the decomposition with Pastas for well B2 is presented. Results for wells A2 and C2 can be found in Appendix C.5. The decomposed stresses that cause the head response in the aquifer are the astronomical tide in the Oosterschelde (named 'Tide' in Figure 4.3), the recharge which is the combined effect of rain and evaporation and the non-astronomic water level fluctuations in the Oosterschelde.

Especially for well A2 and B2 the model fits the observed head quite well, with explained variances of 97.6 and 98.8 respectively. The decomposed tide and non-astronomic tide look similar for the three observation

wells. Although, obviously the response time (graphs in the right column) is faster when the well is closer to the Oosterschelde. Also note that the recharge response is quite slow, while the response to the tide and the non-astronomic tide is relatively faster. Lastly, note that (as mentioned in Section 3.3), no noise model is included for well C2.

Since the response time of the non-astronomical stress is quite fast, subtracting it from the hydraulic head may affect the amplitude and phase estimates. Therefore, two time series were created with Pastas. One where both the head response to recharge and non-astronomical fluctuations in the Oosterschelde are subtracted from the groundwater heads (Pastas 1). And one where just the response to recharge was subtracted from the groundwater heads (Pastas 2).



name	optimal	stderr
Tide_A	0.04	35.0%
Tide_n	0.53	7.2%
Tide_a	0.72	80.9%
Recharge_A	75.79	7.4%
Recharge_n	1.10	3.2%
Recharge_a	30.32	6.6%
Recharge_f	-1.44	7.3%
Step_A	-0.41	1.2%
Step_a	0.00	0.0%
Step_tstart	737042.00	0.0%
Non-astronomic_A	0.23	5.3%
Non-astronomic_n	1.15	3.9%
Non-astronomic_a	0.93	9.6%
constant_d	0.41	2.5%
noise_alpha	1.72	0.0%

Figure 4.3: Decomposition of the groundwater fluctuations with Pastas for the times series from well B2, results are in meter. The explained variance of the model is 98.80

In Figure 4.4 the result of the noise removal with the Butterworth filter are presented for well C2. Results for wells B2, C2 (and A4 as well) can be found in Appendix C.6.

Note that the step has already been removed from the time series. The height of the removed step was estimated with 'Pastas'. Alternatively one could use the step estimated by means of the double data, or by trial and error, as removing it is mostly just about avoiding statistical problems and is not significantly changing the estimated parameters.

Initially, the step was left to be removed by the Butterworth filter. Both by applying the filter on the whole time series. And by applying the filter on the data before and after the step separately and subsequently estimate the step by adding an average water level, for the part of the time series after the step, to the least-squares estimation (described in Section 3.1). However, both these methods gave unsatisfactory result, especially for

wells B2 and C2 this led to a big over- and undershoot, in front and after the step respectively. This significantly flattened the probability density function of the residual and also resulted in a more persistent and severe autocorrelation around the step.

Further note that the variance in the data is significantly reduced compared to the raw data (see Figure 4.4). Besides, the high-frequency fluctuations (see the main plot in Figure 4.4(b)) seem to have no visible phase and/or amplitude distortions.

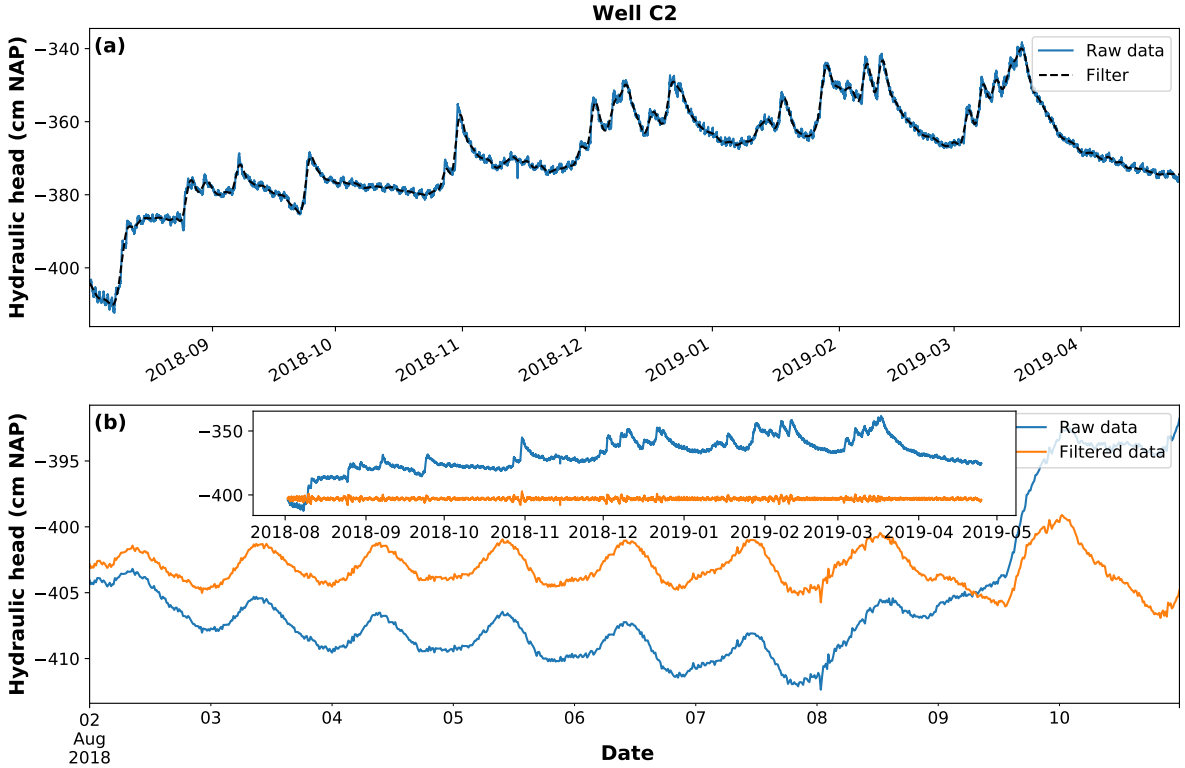


Figure 4.4: Plots of the hydraulic head (raw data) and the obtained filter (plot a) and a plot with both the hydraulic head (raw data) and the filtered data (b), with 9 days in August in the main plot and the whole time series in the subplot.

4.2.2. Comparison of the filters

Probability density graphs

In Figure 4.5 the probability density of the estimated amplitude and phase is plotted for the different methods, for the M2 constituent found in wells A2, B2 and C2. Note that the probability density graphs are normalized, so that the area under each histogram sums up to 1.

In well A2 the amplitude and phase estimated with Pastas1 is in disagreement with the results of the other two filters and the raw data. The latter are largely in agreement, with an increasingly smaller confidence interval from raw data, to Pastas 2 to Butterworth. This increasingly smaller confidence interval is also observable in wells B2 and C2. For these wells, the probability density graph of the raw data is almost flat compared to the filtered data. Results for the other constituents in wells A2, B2 and C2 can be found in Appendix C.8.

Note that no plots have been made for the SA and SM constituents in the groundwater wells. For the Butterworth filter these constituents, which have a period larger than 14 days, are filtered out (see Figure 3.4). For the raw data and the data filtered with Pastas, the results were substantially different from the other constituents (e.g. the amplitude did not decrease from the Oosterschelde to well A2). It is anticipated that this is due to the large periods of both constituents combined with undesired fluctuations still present in the signals. Therefore, proper estimation of the amplitude and phase from these constituents in a least-square sense is not possible. Moreover, the period of SA is almost a year, so according to Rayleigh's resolution limit (see Section 3.1.1) the time series of A2 and B2 are only just long enough and the time series of C2 is too short for reliable results.

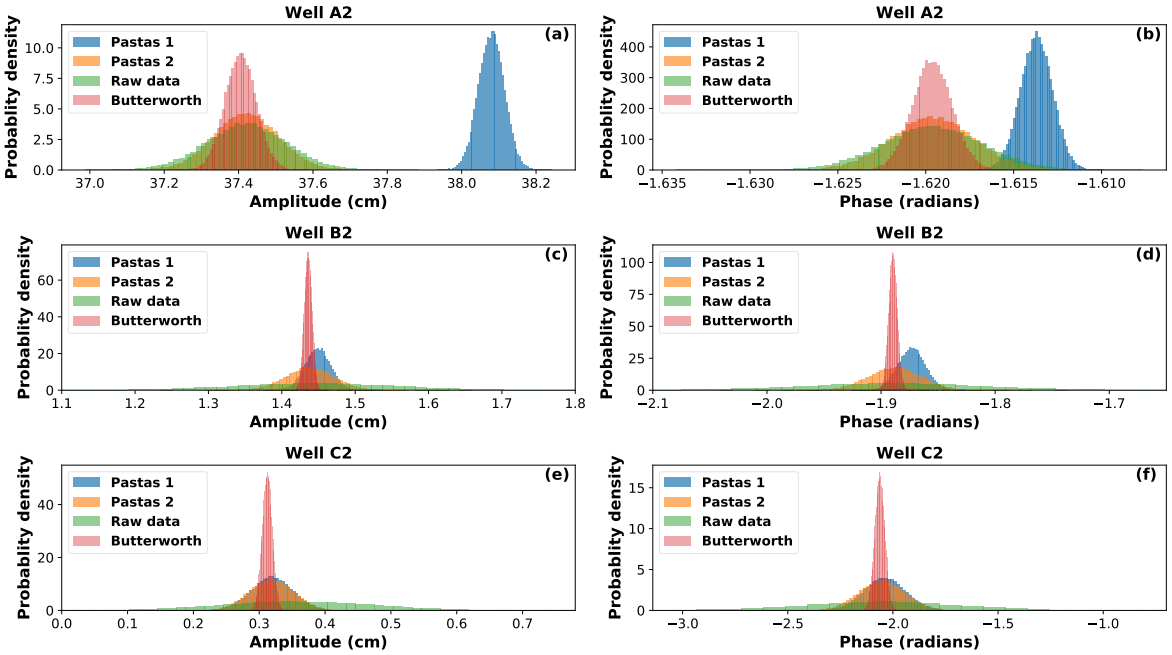


Figure 4.5: Probability density graphs for the amplitude and phase of the M2 constituent observed in wells A2, B2 and C2 plotted in the upper, middle and lower row respectively.

Analysis of the residual

In order to have accurate parameter and uncertainty estimates, the error should be random and independent. To check this, probability density graphs of the residual, graphs of the residual against the time and ACF graphs were plotted for all filters and all observation wells.

In Figure 4.6 the distribution of the residuals, modeled minus observed groundwater level, of the different time series is plotted. In general, the distribution of the well closest to the sea (A2) is the closest to a normal distribution. Figure 4.7 contains plots of the residual against time and Figure 4.8 contains ACF plots. When no noise is removed there is a lot of autocorrelation, this seems to be slightly better for Pastas 2. Pastas 1 performs better, compared to the raw data and Pastas 2, especially for well A2, but is still rather poor for the other wells. The Butterworth filter performs best considering wells B2 and C2, for which it has the smallest residuals and the least autocorrelation. For well A2 the error is less autocorrelated compared to Pastas1, but the residual for Pastas 1 is more normally distributed. Well A4 is not used in this comparison, however, plots of the modeled and observed head fluctuations and associated residual along with a probability density graph of the residual can be found in the Appendix C.9.

For the rest of the report, the results obtained with the Butterworth filter are used. A discussion about this can be found in Section 4.5. However, it is important to note that the present autocorrelation causes the calculated confidence intervals to be inaccurate. The estimated confidence intervals will be smaller than the true confidence intervals (Halcoussis, 2005). Nevertheless, the estimated confidence intervals are still plotted and used because a large confidence interval still shows that the estimates are uncertain.

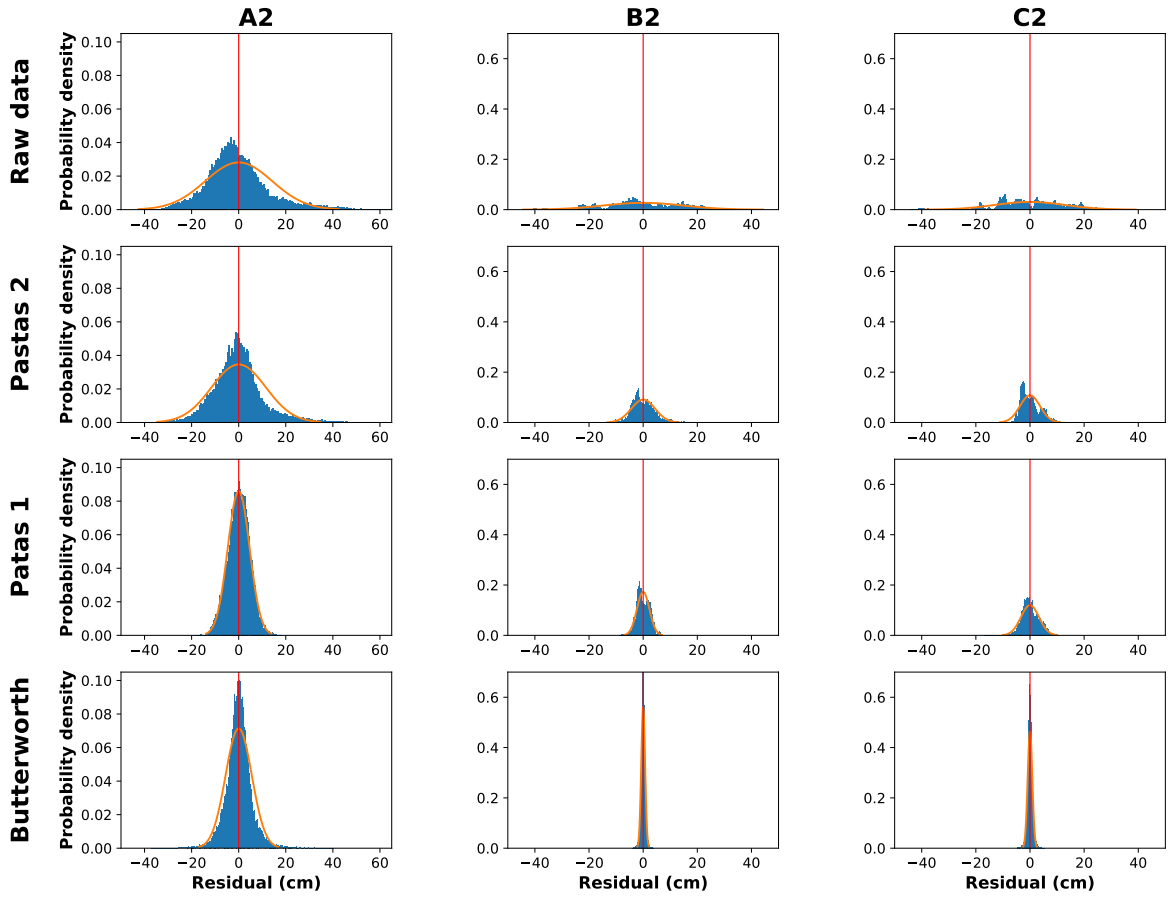


Figure 4.6: Probability density graphs of the residuals (modeled minus observed groundwater level) for the groundwater time series A2, B2, C2 (columns) for various treatments of the time series (rows). The plots also include the mean (vertical red line) and the normal distribution (orange line) based on the mean and standard deviation of the data. Note that the plots have a fixed x-scale and y-scale for every column.

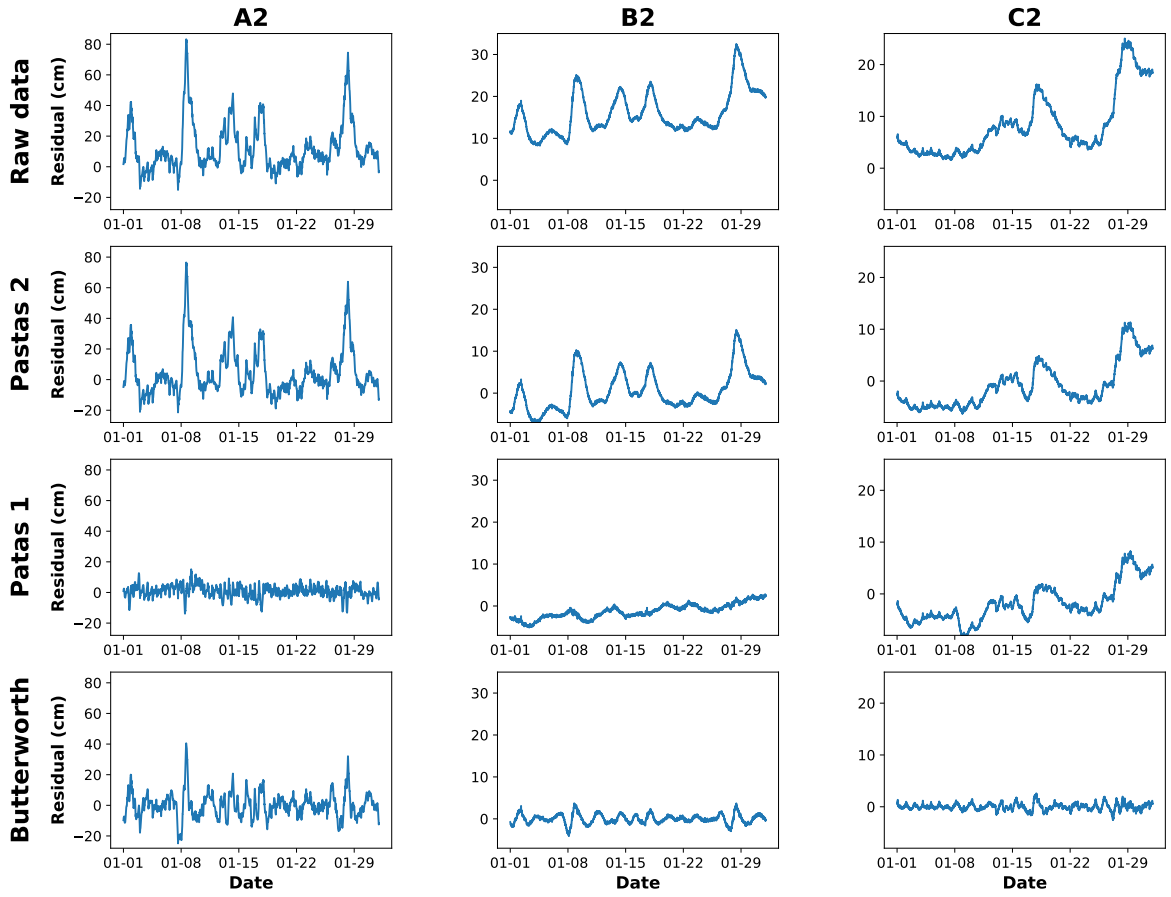


Figure 4.7: Plots of the residuals (modeled minus observed groundwater level) against time for January 2019 for the groundwater time series A2, B2, C2 (columns) for various treatments of the time series (rows). Note that the plots have a fixed x-scale for every column.

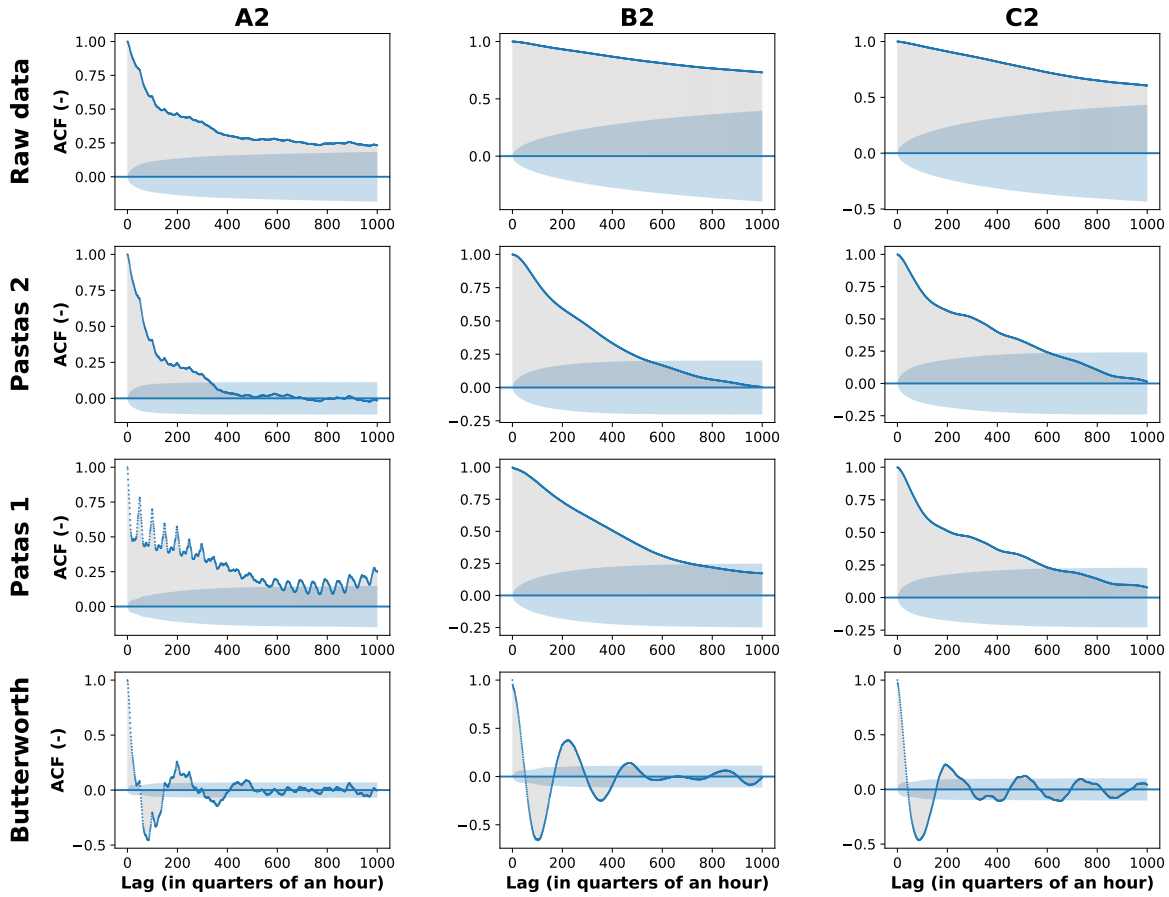


Figure 4.8: Plots of the autocorrelation of the residual for the groundwater time series A2, B2, C2 (columns) for various treatments of the time series (rows). The 95% confidence intervals are drawn as blue cones, values outside the cone are likely to be autocorrelated.

4.3. Analysis of the amplitude & phase data

In this section, several ways to visualize the data are used to analyze it, with the purpose of identifying what data is appropriate to use for the inverse modeling.

4.3.1. Probability density plots

For some constituents, especially in well C2, the probability density plots showed that the estimated parameters are unreliable. In Figure 4.9 the probability density graph of the 3MS8 constituent in well C2 is plotted. The probability distribution function of the amplitude is right-skewed, and the phase is not normally distributed, in fact, if the lowest calculated probability is assigned to all bins, the area underneath the graph would still add up to more than 0.28. Several other constituents appear to have a high probability for every bin in well C2 (see Appendix C.8 for the other Figures). These are not considered further in this analysis (O1, MNS2, NLK2, K2, 2SM2 2MK3, MK3, 3MS4, MN4, MK4, 2MN6, M6, 2MS6, M8, 3MS8).

In Appendix B.1 a table is presented for each observation well, containing the amplitude and phase estimates plus their associated standard errors. For wells B2 and C2 unreliable amplitude and phase estimates are indicated with a mark. Again note that, because the times series is autocorrelated, the standard errors estimated with the least-squares model are smaller than the true values.

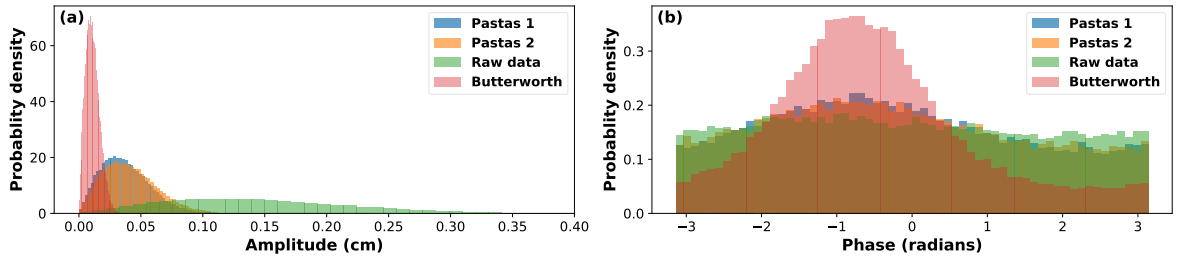


Figure 4.9: Probability density graphs for the amplitude(a) and phase(b) of the 3MS8 constituent observed in well C2

4.3.2. Amplitude and phase plots

In Figure 4.10 the amplitude and phase of the M2 constituent, obtained with the harmonic analysis, are plotted against the (perpendicular) distance to the gap in the Holocene layer (see Figure 5.1). See Appendix C.10 for the other constituents. The plotted trend-line is based on the observed phase/amplitude in the Oosterschelde, A2 and if available C2. Because of the dissimilar results in well B2 (see also Section 4.5.2), these amplitudes and phases were not used to construct the trend-line. The other observations are quite close to the trend-line. Also, note that the observed phase is determined relative to the Oosterschelde and that a multiple of 2π was added to the observations if needed to obtain a logical phase shift.

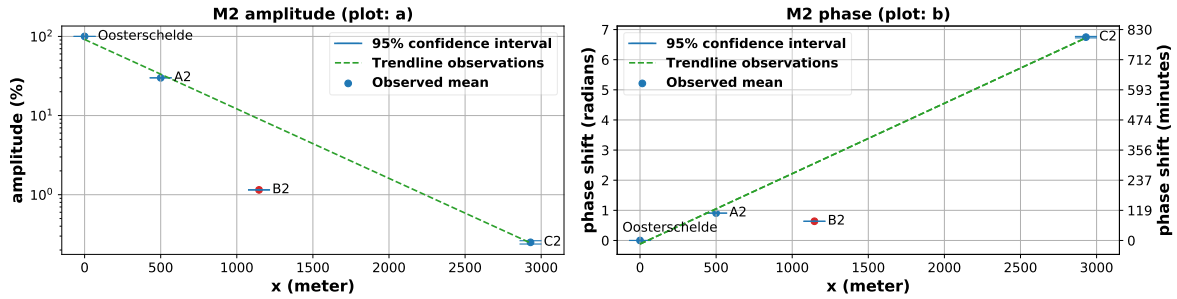


Figure 4.10: Plot of the amplitude (plot: a) and phase shift (plot: b) against the distance to the Oosterschelde, for constituent M2. With the observed mean (blue dot), a trend-line through the observed means (green dash). Note that, although plotted, observations from B2 (red dot) were not used to draw the trend-line. Lastly, note that the observed phase is determined relative to the Oosterschelde and that a multiple of 2π was added to the observations if needed to obtain a logical phase shift.

For the constituents for which the amplitude and phase in C2 are still considered (see Section 4.3.1), the amplitude decreases approximately exponentially and the phase shift increases approximately linearly. Both observations are also expected based on some of the cited case studies (e.g. Trefry and Bekele, 2004; Merritt, 2004; Dong et al., 2015). Although it must be noted that this is not the case for M4, for which the amplitude in A2 is estimated to be bigger than the amplitude in the Oosterschelde. Furthermore, it can be noted that for almost all constituents both the amplitude and phase of B2 are far below the trend-line. Moreover, for some of the constituents, the amplitude in C2 is bigger than the amplitude in B2 (Q1, P1, K1, 3MS2; see Appendix C.8).

Lastly, the constituents with a frequency around 2.8E-4 rad/s (i.e. 3MS4, MN4, M4, MS4 & MK4) had a bigger amplitude in A2 compared to the amplitude at sea. However, there is no configuration possible in this one-dimensional model that can simulate an amplitude increase. So using these constituents in this set-up will not make sense, therefore these constituents are not considered for the inverse modeling. Note that this group of constituents was already classified as unreliable in Section 4.1.1 based on large possible translation errors.

4.3.3. Amplitude and phase against frequency

Playing with the model from Bakker (2019) showed that the amplitude attenuation (in %) and phase shift (in radians) should be approximately the same for constituents with similar frequencies. Moreover, when observing the signal on a fixed location in the aquifer, the phase shift is expected to increase and the amplitude

attenuation is expected to increase when going from a low to a high frequency (e.g. Jacob, 1950). In Figure 4.11 the amplitude, in percentage compared to the amplitude observed in the Oosterschelde, is plotted against the frequencies used in the harmonic analysis, for the different observation wells. In Figure 4.12 the phase shift is plotted against the frequency.

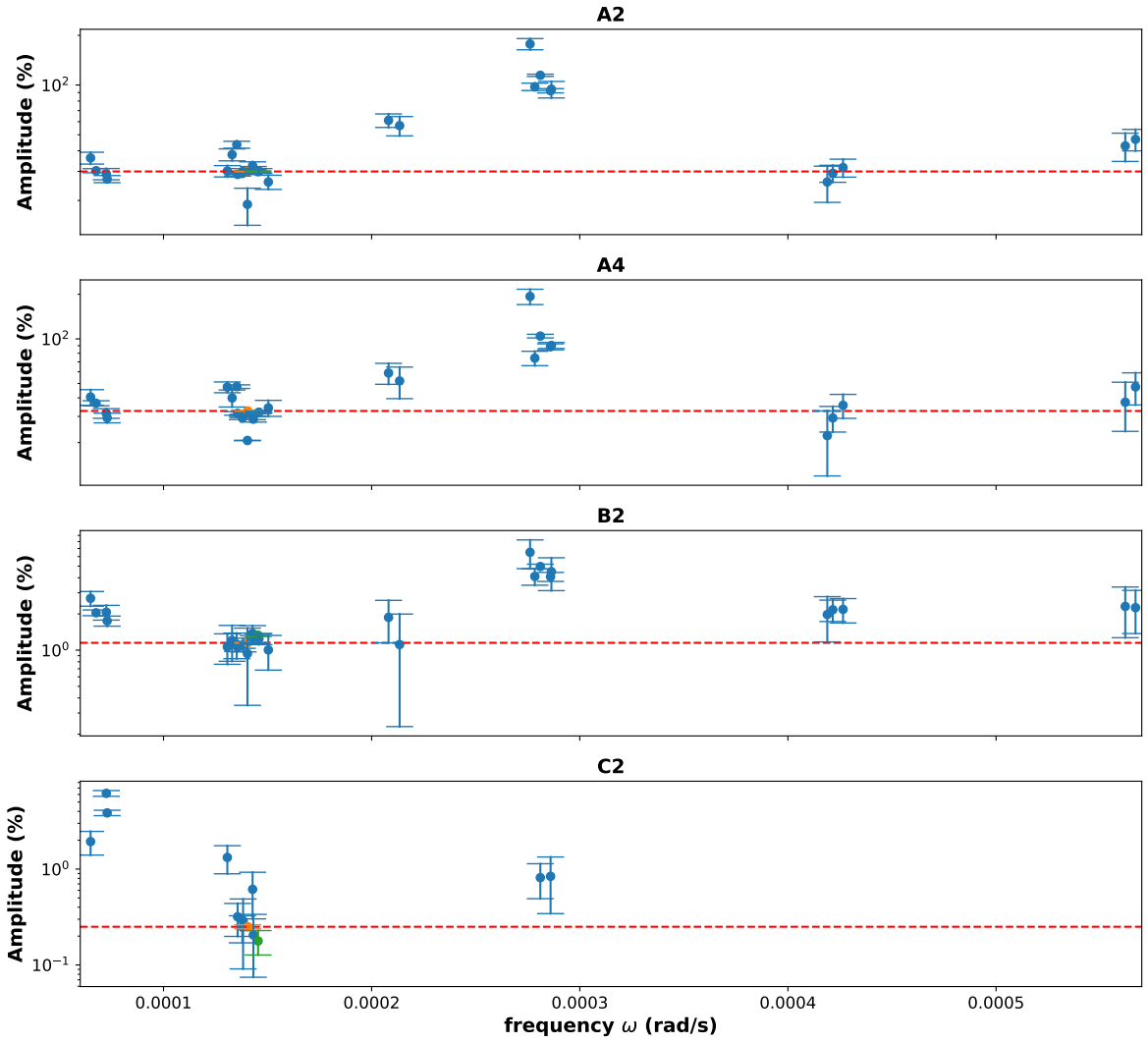


Figure 4.11: Plots of the amplitude(%) against the frequency for the observation wells A2, A4, B2 and C2 (from top to bottom). With in orange the M2 and in green the S2 constituent. At the same height as M2 a red dotted horizontal line is plotted. Note that the 95% error bars are calculated with the formulas from Section 3.1.2 and thus only include the estimated error of the harmonic analysis for that specific well.

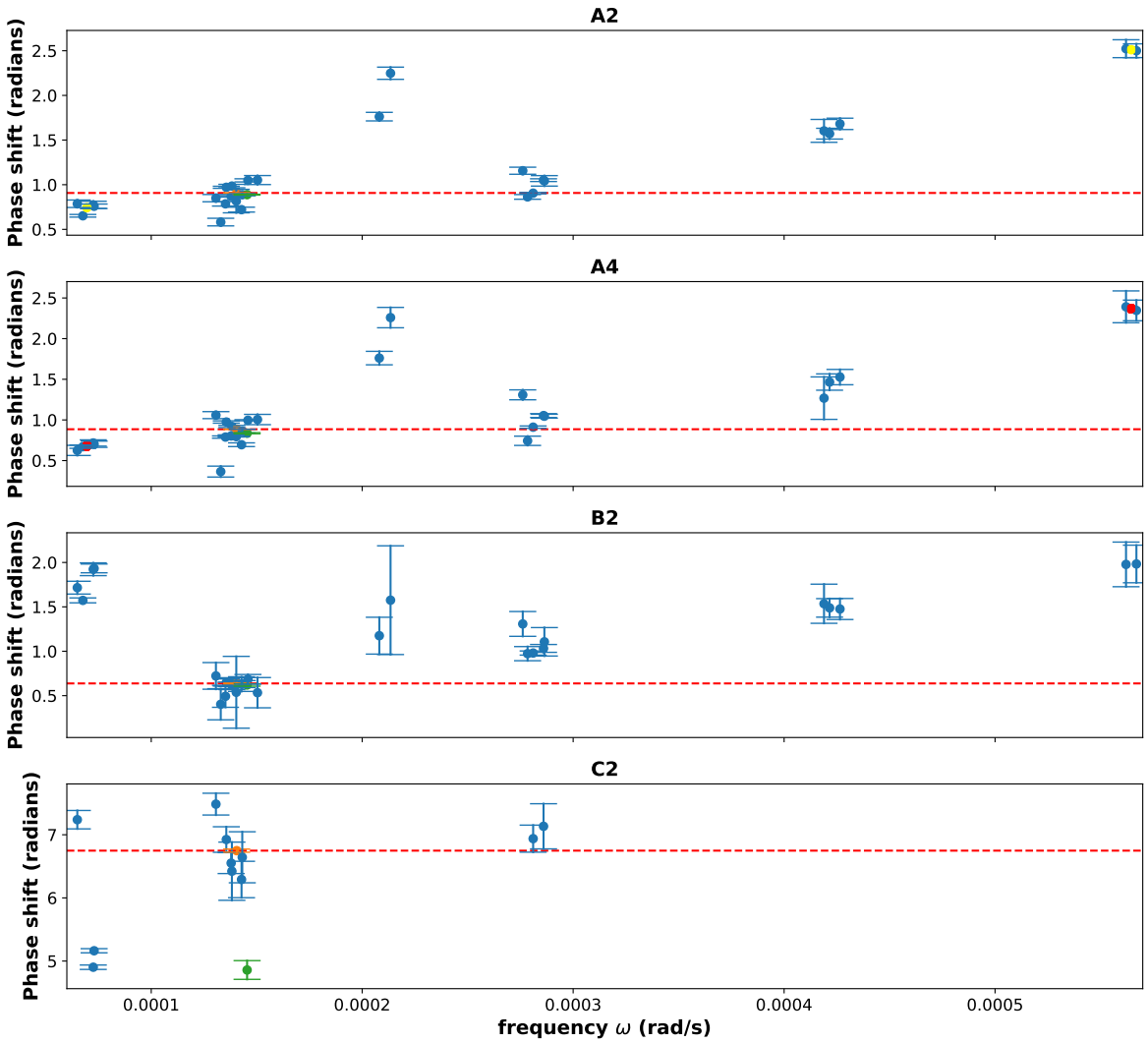


Figure 4.12: Plots of the phase shift (in radians compared to the Oosterschelde) against the frequency for the observation wells A2, A4, B2 and C2 (from top to bottom). With, in orange the M2 and in green the S2 constituent. At the same height as M2 a red dotted horizontal line is plotted. Note that the 95% error bars are calculated with the formulas from Section 3.1.2 and thus only include the estimated error of the harmonic analysis for that specific well.

In Figure 4.11 the expected decrease in amplitude (in %), when the frequency increases, cannot be detected. For the phase shift against the frequency (Figure 4.12) the expected increase is visible, mainly in observation wells A2 and A4. Also note that constituents with a similar frequency roughly have the same percentile decrease in amplitude and a similar phase shift, although in both Figures 4.11 & 4.12 a large spread is observed especially in wells B2 & C2.

Lastly, it is interesting to note that the phase shift in A4 is sometimes bigger than the shift observed in A2 (e.g. for M2 and S2). But this is not the case for all constituents. Moreover, the amplitude of the constituents is sometimes higher in A4 compared to A2 (e.g. for M2 and S2), but not for all constituents.

4.4. Amplitude & phase determination: graphical vs harmonic analysis

In table 4.2 the amplitudes and phases of the M2 tide, for the different observations wells, obtained both graphically (Section 3.5) and with the harmonic analysis (Section 3.1) are presented.

Table 4.2: Comparison of the amplitude and phase obtained graphically vs determination by harmonic analysis with different filters

	Amplitude (cm)				Phase (Radians)			
Time series	**Oosterschelde**	**A2**	**B2**	**C2**	**Oosterschelde**	**A2**	**B2**	**C2**
Graphical method	124.05	37.86	1.71	0.47	3.13	-2.29	3.00	-2.42
No filter used	124.83	37.41	1.45	0.36	-2.53	-1.62	-1.90	-1.91
Butterworth filter		37.41	1.44	0.31		-1.62	-1.89	-2.06
Pastas 1		38.08	1.47	0.33		-1.61	-1.88	-2.06
Pastas 2		37.42	1.45	0.33		-1.62	-1.90	-2.07

The phase is calculated back to the phase at 02-aug-2018 00:00.

If the signal to noise ratio is high, which is the case in the Oosterschelde and in well A2, then the graphical estimates of the amplitude are quite close to the estimates obtained with an harmonic analysis. For wells B2 and C2 the relative difference of the amplitude estimates is quite big. Lastly, the phase is off by at least 5% ([difference/2π]*100%) in all observation wells and 10% off in the Oosterschelde.

4.5. Discussion

Before moving on to the methodology and results of the inverse modeling, it is only logical to discuss the results of the data preparation. The rest of the results are discussed in Chapter 7, including a part about errors in the data, which is not considered to be highly relevant for now.

4.5.1. Pastas vs Butterworth filter

In the results, the filters have been compared with their probability distributions of the residuals, a plot of the residual against time, an ACF plot and probability density functions for all the considered constituents. A discussion of these results is presented below.

Statistics of the residual from the harmonic analysis for the three observation wells

As mentioned, the error should be random and independent for accurate parameter and uncertainty estimates. To check this, probability density graphs of the residual, graphs of the residual against the time and ACF graphs were plotted for all filters. These plots show that the distribution of the residual is far from normal for the time series with no filter considering wells B2 and C2 (see also Appendix C.38 for probability density graphs without fixed axes). For Pastas1 in well A2 the residual seems reasonably normally distributed. And the residuals from the Butterworth filter are symmetric but not quite normally distributed. From the graphs of the residual against the time and the ACF plots, it can be concluded that the residual is autocorrelated. Especially for the raw data and both Pastas filters there is quite some autocorrelation.

For the least-squares model to provide accurate uncertainty predictions, the error should be independent and normal (Hill and Tiedeman, 2006). This condition is not met here and this can result in serious underestimation of the confidence intervals (Kutner et al., 2005). In the rest of the report, the estimated confidence intervals are still used, as large confidence intervals are still indicative for poor estimates. Nevertheless, autocorrelation by itself does not imply that the estimates are biased (Halcoussis, 2005). Biased estimates can result from a relevant variable that is missing from the model, the latter is often causing the autocorrelation. For the harmonic analysis, with or without the use of a filter, it is presumed that no useful information is missing from the model and thus that the parameter estimates are unbiased. Although the estimated amplitudes and phases will vary more from their true value compared to the case without autocorrelation (Halcoussis, 2005).

An attempt was made to validate the amplitude and phase estimates with an harmonic analysis, although this attempt was impeded by a short data-set and the Rayleighs resolution limit. Nevertheless, the amplitudes and phases were calculated for a smaller times series of 250 days (a smaller times series would require more astronomical coupling on top of the M2-MPS2 combination; see Section 3.1.1) instead of 369 days, this resulted in fairly similar results. Which again indicates unbiased phase and amplitude estimates.

Probability density plots of the phases and amplitudes for the three observation wells

From Figure 4.6 it becomes clear that the use of a filter considerably decreases the standard deviation of the

estimates, especially when the signal to noise ratio decreases. The application of a filter also leads to different estimated amplitudes and phases (mostly) in B2 and C2. Also, note that the phase sometimes has an estimated distribution ranging from -π to +π (see Figure 4.10). This results in skewed and/or imprecise amplitude estimates. The consequence is that for the wells with a small signal to noise ratio, especially when no filter is applied, a lot of constituents are not accurately estimated. Furthermore, it is clear that the standard deviations resulting from the calculations with the Butterworth filter are a lot smaller compared to the other filters.

Figure 4.5 and Appendix C.8 show probability density plots for the constituents in the different wells. For most of the probability distribution plots from wells B2 and C2, the amplitude and phase of the different constituents are estimated to be close to each other for the different filters. For A2 this is not the case and the Pastas 1 filter is often a bit off compared to the other filters and the raw data. This is especially striking for the results of the M2 constituent since the M2 constituent is predominantly present in the signal and the time series is sufficiently long for a good estimate of both the amplitude and phase. Pastas 2 for which only the head fluctuations caused by rain and evaporation were subtracted did a more similar job compared to the Butterworth filter and the use of no filter. Therefore, it is concluded that this deviation is caused by the non-astronomic fluctuations in the Oosterschelde. Looking at the response time in Figure 4.3 of the non-astronomic fluctuations it is rather quick and partly within the frequency domain of the tidal constituents. However, more research is necessary to conclude which time series is better and this is out of the scope of this research. The signal of the non-astronomical fluctuations could be approximated with a Fourier series, containing several frequencies, which would all propagate differently through the aquifer. So using only one response function does not accurately simulate what happens in the aquifer. Therefore, the Butterworth and Pastas2 filter are presumed to have better estimates for the phase and amplitude.

Conclusion
The use of the Butterworth filter is recommended. Using the Butterworth filter produces the smallest confidence intervals for the amplitude and phase estimates. Moreover, especially for the wells with a low signal to noise ratio, the Butterworth filter is better at extracting the tidal signal. Besides, the residuals of all wells, using the Butterworth filter, are symmetrical. Unfortunately, the residuals are not normal and independent. The estimated parameters are presumed to be unbiased, but the range of the confidence intervals is underestimated. Overall the Butterworth filter is considered to perform best and is used to produce the amplitude and phase estimates for the inverse modeling part of the report.

4.5.2. Analysis of the amplitude and phase data

Amplitude and phase against distance plots
For the constituents where the amplitude and phase could be estimated in C2, the (log-linear) trend-line approximately follows the observations (see Appendix C.10). Although B2 is excluded from this statement (more details in the next paragraph). As mentioned in Section 4.3, an exponentially decaying amplitude and a linearly increasing phase shift is expected based on some of the cited case studies (e.g. Trefry and Bekele, 2004; Merritt, 2004; Dong et al., 2015) and so the data can be labeled as trustworthy to some extent. Although it must be noted that this does not tell if the amplitude and phase data will produce consistent estimates of the assessed parameters. This question will be addressed in Chapter 6.

In almost every amplitude plot, the amplitude of B2 is not close to the trend-line constructed for the data points of the Oosterschelde, A2 and C2. See for example Figure 4.10. For some constituents, the amplitude in B2 is even lower than the amplitude in C2, which is of course unexpected. A possible explanation are errors in the hydraulic head data or an incorrect interpretation of the data. However, similar magnitudes of the head response to recharge in wells A, B and C, which were calculated with Pastas, indicate that at least the amplitudes are unlikely to be estimated much too small because of for example a measuring error. Therefore, the dissimilar result in well B2 cannot be (fully) attributed to a measurement error. Hence, the head fluctuations in well B are presumed to be correct.

Nevertheless, the head data in well B is troublesome for several reasons. First, if the amplitude attenuation is underneath the plotted trend-line then the phase shift should be above the trend-line. n*2π (with n being an integer) could be added to the phase observation in B2, so that the amplitude and phase are (slightly) more in agreement. Such a fast attenuation and big phase shift can be caused by a layer with a low diffusivity (i.e.

a small hydraulic conductivity and/or high specific storage). However, there is no suggestion of that present in the REGIS-II model, or in other words, there is no obvious explanation (e.g. a more clayey transect) for this deflecting behavior. Moreover, with a one-layer aquifer, it is possible only to simulate an exponentially decaying amplitude and a linearly increasing phase shift, although some deviation from this is possible with a two-layer model (i.e. a confined aquifer with a clay lens and an aquifer below it). However, the results of such an (optimized) two-layer model, which includes the data from well B2, were quite poor. Consequently, amplitude and phase data from well B2 is not used for the inverse modeling part of the report.

That being said, the incompetence of the model to simulate head fluctuations, that are presumed to be correct, indicates that the model is, at least partly, incorrectly representing the real system (e.g. Hill and Tiedeman, 2006; see also Section 7.2.1 for a discussion about the simplifications).

Amplitude and phase against frequency plots
Constituents with similar frequencies should approximately have the same amplitude attenuation and phase shift. So based on Figures 4.11 & 4.12, it can already be concluded that at least not all constituents are calculated as precise, with the harmonic method, as the plotted confidence intervals would imply.

The M2 (orange) and S2 (green) are modeled with a different color in Figures 4.11 & 4.12. Both constituents have a large amplitude in the Oosterschelde 124.8 and 30.2cm respectively and therefore their amplitude and especially their phase should be estimated more precisely. In A2 & A4, in both the phase and amplitude plots, S2 and M2 have almost the same values. Because S2 has a frequency very close to M2 this is to be expected. However, when the frequency increases, a smaller amplitude (in % compared to the Oosterschelde) and a bigger phase shift are expected. These trends are not observed between M2 and S2. Moreover, two constituents close to each other gives some extra information, but this difference is small and the uncertainty is relatively big. Therefore it was decided not to use the results of the S2 constituent. Furthermore, none of the other constituents showed the expected pattern for both the amplitude and phase in both A2, A4 and B2. Therefore, it was decided to only use both the amplitude and phase data of the M2 constituent, in wells A2, A4 and C2.

It is noted that the amplitudes in observation well C2 are quite small, except for constituent M2 which still has an amplitude of 0.33 cm, which is similar to the magnitude of S2 observed in well B2, which appears to be still modeled correctly (i.e. close to the amplitude of M2). Therefore the results of M2 in C2 are trusted to be reasonably accurate, unlike the results in C2 for the other constituents.

Unfortunately, the amplitude against the frequency plot does not show the expected decreasing pattern when going from a lower to a higher frequency. Nevertheless, for the phase against the frequency, the expected increasing trend is visible. The lack and presence of the expected pattern, in Figures 4.11 & 4.12, might be caused by different sensitivities to system parameters.

Since information obtained from the data is limited, the information from the phase against frequency plot might prove to be valuable additional information for the inverse modeling. Therefore the 10 clusters that are present in A2 and A4 are further analyzed.

Because the clusters show quite some spread, centroids of the clusters were calculated to increase the usability of the data and to partly cancel out errors from individual constituents. The calculated centroids are plotted in Figure 4.13. Unfortunately, not all centroids are useful for inverse modeling. Since the 1.4e-4 rad/sec centroid is very close to the phase of M2, this centroid won't be used. The centroid around a frequency of 2.8e-4 rad/sec (i.e. 3MS4, MN4, M4, MS4 & MK4) is regarded as unreliable because of large uncertainty in the translation from the reference locations (bh2 and STAV) to the study site (see Section 4.1.1).

Furthermore, the amplitudes of the constituents with a frequency of approximately 4.2e-4 (2MN6, M6 & 2MS6) are not estimated accurately in the Oosterschelde see Appendix C.4 and Section 4.1.1. Therefore, this centroid is also not considered for inverse modeling. So only the centroids of the diurnal tides and high-frequency tides (M8 and 3MS8) can potentially be used as extra information for the inverse modeling. The data of these centroids can be found in a table in Appendix B.2.

Lastly, note that from Figures 4.12 & 4.13 it appears that the effect of earth tides is insignificant. As mentioned in Section 3.2.1, fluctuations caused by earth tides, whom might disturb the results, are exactly out of phase with oceanic tides. The variance of the clusters plotted in Figure 4.12 is a lot smaller than π, so frequencies are roughly in phase. Therefore, the effect of earth tides is presumed to be negligible.

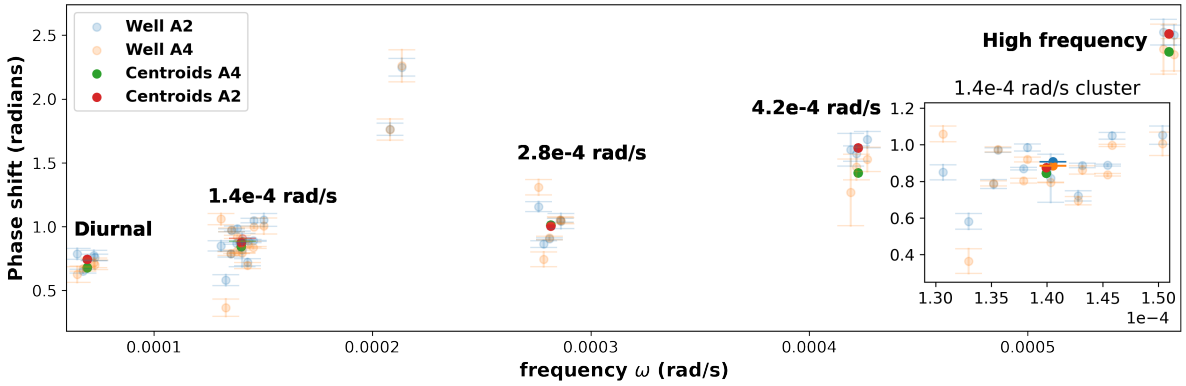


Figure 4.13: A plot of the centroids from the observed 5 clusters. With the A4 centroids in green and the A2 centroids in red. The phase shifts (in radians compared to the Oosterschelde) against the frequency for the observation wells A2 & A4 are also plotted (with a high transparency). Within the subplot, a more detailed plot from the 1.4e-4 rad/s cluster, where the M2 constituent is made opaque. Lastly, note that for the '2.8e-4 rad/s' cluster the A4 centroid has almost the same phase shift as the A2 centroid.

4.5.3. Amplitude & phase determination: graphical vs harmonic analysis

A graphical determination of the amplitude and phase were done to check if this less time-consuming method would give accurate results as well.

If the signal to noise ratio is high, which is obviously the case in the Oosterschelde but also in A2, then the estimates of the amplitude are quite accurate (i.e. close to the estimate based on the harmonic analysis). The difference between the two methods is 0.5% for the amplitude in the Oosterschelde and 1.2% for the amplitude in well A2. For wells B2 and C2 the amplitude was overestimated with 18% and 31% respectively, which is not sufficiently accurate for using these results for the tidal method. The phase in the Oosterschelde and A2 is 0.62 and 0.67 respectively off from what is expected. The phase estimated graphically in B2 is far off from what is estimated by using a harmonic analysis. The result for C2 is closer to what is expected, but because the signal to noise ratio is even worse for this well compared to B2, this is presumed to be a lucky shot.

So at this site, determining the amplitude of M2 graphically is a reliable method if the signal to noise ratio is high. The extent to which this method still produces reliable results can not be determined from these results. Determining the phase graphically does not give good results.

4.5.4. Data selected for the inverse modeling

Because of unreliable observations or in some cases not entirely accurate and of little added value, only a few observations are used for the inverse modeling. A small summary of which data will be used for the inverse modeling is presented in the paragraph below.

First, the amplitude and phase data of well B2 is not used for optimization, because it deviates a lot from the other data and cannot be correctly modeled by the used model. Therefore the 'information' provided by this well is of little benefit and is likely to even obscure the results. The data that is used for the inverse modeling is the amplitude and phase of the M2 tide observed in wells A2, A4 and C2. Moreover, also the average phase shift in A2 and A4 of the diurnal constituents and two high-frequency constituents (M8 & 3MS8) are used for the inverse modeling.

5

Methodology: Inverse modeling

The main goal of this research is to estimate hydraulic parameters from tide-induced groundwater fluctuations. To estimate these parameters a set of observations and a mathematical model that includes the most important system characteristics are needed. This procedure of parameter estimation is often called 'inverse modeling'.

First, the modeling strategy and the used model are discussed in Section 5.1. In Section 5.2 the adopted approximations are discussed. The model-setup is discussed in Section 5.3. Subsequently, the derivation and some insights in both the one- and two-layer solutions, which are implemented in the used model, are presented in Section 5.4. In Section 5.5 the calibration of the model is discussed. In the final section, the measures used to assess the model performance are presented.

5.1. Modeling strategy

Groundwater systems are often complex and data is limited, therefore it is difficult to obtain a tractable model that represents the system accurately enough to produce useful results. Collecting more data can be a solution but is not done in this research. Another solution is to carefully disregard some of the complexity of the system (Hill et al., 1998). This can be done by choosing a lower level of parameterization of the problem (i.e. a simpler model).

In this study both a one-layer and a two-layer model are considered, the former being a more simple model. An advantage of having a simple model is that fewer parameters need to be estimated. Moreover, a simple model is more transparent, meaning that the model dynamics are more understandable (Hill and Tiedeman, 2006). Transparency decreases as the complexity of the model increases. A more complex model, such as a two-layer model can be warranted by knowledge of the hydrology of the system, an inability to reproduce the observations and the complexity allowed by the data (Hill and Tiedeman, 2006). Although a two-layer model might prove to be problematic because of a small ratio of observations over parameters. Nevertheless, both a one and two-layered model are considered also because a model becomes more defensible if more defensible ideas about the system are evaluated (Hill and Tiedeman, 2006).

The solution and code selected for this research is the code presented by Bakker (2019). This code is able to simulate one-dimensional flow in a multi-layer aquifer covered by a leaky layer and allows for the inclusion of aquitards in the model (see Figure 2.2). It is expected that this model is able to include the most relevant processes that play a role in the system. Another advantage of using the code of Bakker (2019) is that it uses a set of analytical equations, which leads to short computation times. This is useful as results can be generated quickly and is also convenient if many model runs must be performed (e.g. when doing Monte Carlo Simulations). The model presented by Bakker (2019) uses the following parameters as input: the transmissivity, storage coefficient and loading efficiency of the aquifer(s) and the resistance to vertical flow, storage coefficient and loading efficiency of the leaky layer(s) (aquitards). It returns the head, amplitude and phase shift for the given tidal component at arbitrary locations. Note that some parameters such as the storage coefficient of the leaky layers can be set to zero if desired.

The complexity of the modeled system combined with a limited number of observations often results in a modeling process that is troubled by nonuniqueness, insensitivity and instability (Hill and Tiedeman, 2006). Nonuniqueness is the problem of having multiple parameter sets with the same goodness of fit. Insensitivity occurs when the observations contain an insufficient amount of information to estimate the parameters. And instability occurs when a small deviation in the model, such as a small parameter value change, results in a profound change of model outcome. The detection of these problems is important to understand the value of the obtained results (Hill and Tiedeman, 2006).

5.2. Approximations

In order to model the system using the code of Bakker (2019), some approximations are needed. The following approximations are adopted:

- Dupuit-Forchheimer approximation: groundwater flows vertically without resistance in an aquifer
- Flow in leaky layers is vertical
- All layers are homogeneous
- Impermeable base
- All parameters are constant in time.
- The loading efficiency is 1 for both the aquifers and aquitards*

* The loading efficiency can be smaller than 1 if part of the load exerted on the aquifer is initially carried by the solid skeleton so that only part of the stress is transmitted to the soil water (Merritt, 2004). However, if the compressibility of the soil is much larger than that of the water then the loading efficiency is near 1. This is the case if the sand and clay layers are uncemented (Bakker, 2016).

5.3. Model set-up

Figure 5.1 gives a schematic overview of the geometrical features of the model. The following is a short discussion of these features. The governing equations are discussed in Section 5.4.

The model of Bakker (2019) requires uniform thickness for the different aquitard and aquifer layers. Judging from Figure 2.2 such a schematization may be a reasonable approximation. Furthermore, both Figure 2.2 & Figure 2.3 show that the Oosterschelde cuts through the Holocene layer. The edge of the Holocene layer approximately coincides with the shore of the Oosterschelde. Therefore, the left boundary (x=0) of the model is taken at the shore (see Figure 2.1). The left boundary condition of the model is the tide observed in the Oosterschelde (see Box 5.1 for more details). The right boundary condition is a head of zero for $x \rightarrow \infty$. The thickness of the upper aquitard, upper aquifer, second aquitard and second aquifer is approximately 16.6, 20.3, 5, and 45.1 m respectively. These are rough estimates based on Regis-II (see Figure 2.2). The (perpendicular) distance to the Oosterschelde (i.e. x=0 line in Figure 2.1) for wells A, B and C is approximately 500, 1145 and 2930 m, respectively.

Box 5.1: The left boundary condition

The left boundary condition of the model is taken equal to the tide observed in the Oosterschelde. This corresponds to assuming that the sea cuts through the aquifer layers at x=0. In addition, this boundary condition also corresponds with assuming that the loading efficiency is 1, along with assuming there is no resistance to vertical flow at x=0. The latter is true for the upper aquifer (i.e. there is no confining layer at x=0), but not entirely for the lower aquifer (see the clayey unit from the Maassluis formation in Figure 2.2; and Figures C.8 & C.9). However, because this clay lens is only a few hundred meters further from shore, the effect of this clay lens on the validity of the boundary condition is assumed to be negligible.

Two set-ups are tested in this paper, a one-aquifer system: consisting of the upper leaky layer and upper aquifer. And a two-aquifer system for which a second leaky layer and second aquifer layer are added to the model. The parameters to optimize are the transmissivity and storage coefficient of the aquifer(s) and the resistance to vertical flow and the storage coefficient of the leaky layer(s), of which some can be fixed or

estimated collectively in the optimization. The parameters and constants of the one-aquifer and two aquifer model are specified in Table 5.1 and Table 5.2 respectively.

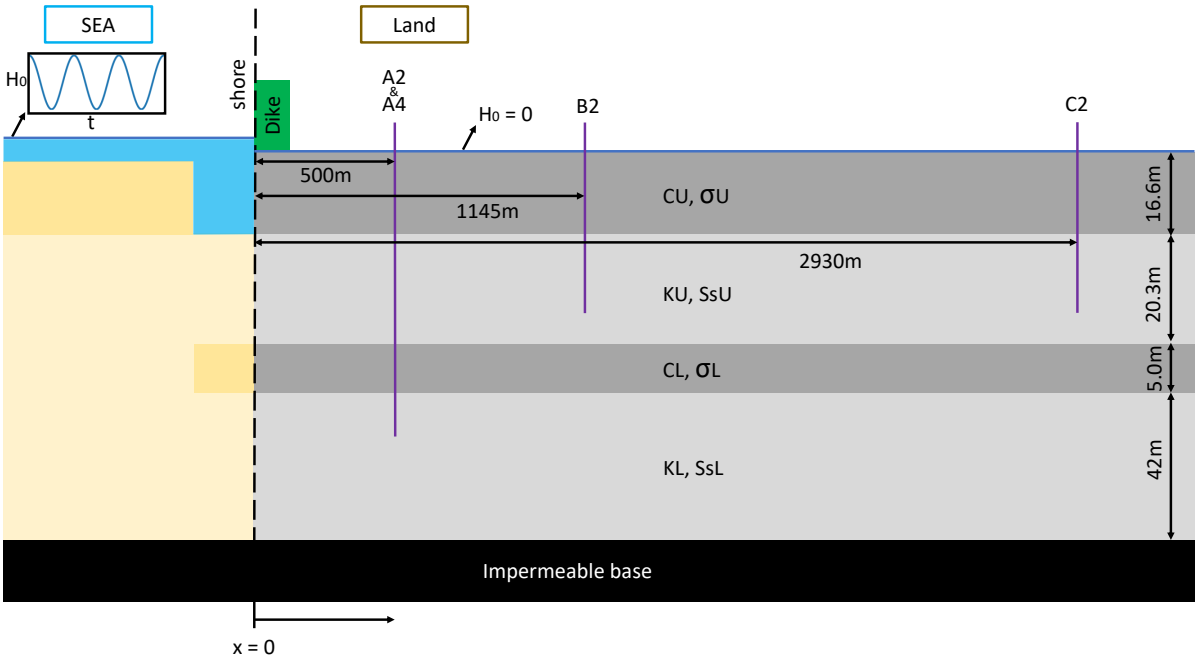


Figure 5.1: A schematization of the model. Note that the left boundary of the model is at $x = 0$. For the one-aquifer system the layers below the upper sand layer are impermeable, which can also be modeled by taking $CL \to \infty$. Note that well B2 is not used for the inverse modeling (see Section 4.5.2). Lastly note that the parameters of the upper sand and clay layer are denoted with 'U' and the parameters from the lower sand and clay layer are denoted with 'L'.

Table 5.1: The parameters and constants of the one-aquifer model.

Parameters		**Unit**	**Description**
KU		m/day	Hydraulic conductivity of aquifer
S_sU		m^{-1}	Specific storage of aquifer
CU		days	Resistance of leaky layer
σU		m^{-1}	Specific storage of leaky layer
Constants	**Value**		
βU	1	-	Loading efficiency of both the aquifer and leaky layer

Note that the U in the name of the parameters emphasizes that the parameter is used for the Upper aquifer system (see also Figure 5.1).

Table 5.2: The parameters and constants of the two-aquifer model.

Parameters		**Unit**	**Description**
KU		m/day	Hydraulic conductivity of the upper aquifer
S_sU		m^{-1}	Specific storage of the upper aquifer
CU		days	Resistance of the upper leaky layer
σU		m^{-1}	Specific storage of the upper leaky layer
KL		m/day	Hydraulic conductivity of the lower aquifer
S_sL		m^{-1}	Specific storage of the lower aquifer
CL		days	Resistance of the lower leaky layer
σL		m^{-1}	Specific storage of the lower leaky layer
Constants	**Value**		
βU	1	-	Loading efficiency of the upper aquifer and leaky layer
βL	1	-	Loading efficiency of the upper aquifer and leaky layer

Note that the U and L in the name of the parameters emphasizes that the parameter is used for the Upper and Lower aquifer system respectively (see also Figure 5.1).

5.4. Analytical solutions

In this section, the differential equations, boundary conditions and solutions as described in Bakker (2019) are presented. Besides, the solution is simplified and rewritten to expose present parameter couplings. Furthermore, conditions are derived for when the storage in the leaky layer can be neglected.

5.4.1. One-layer solution

For the one-layer solution, the head in the aquifer, below the land, is governed by the following equation:

$$T\frac{\partial^2 h}{\partial x^2} = S\frac{\partial h}{\partial t} + q_t \qquad x \geq 0, \tag{5.1}$$

where h is the head change caused by the tidal fluctuations, q_t is the vertical specific discharge through the upper confining layer (positive for upward flow), S is the storage coefficient (which is the specific storage S_s multiplied with the aquifer thickness), and T is the transmissivity, which is related to the hydraulic conductivity as follows:

$$T = KH, \tag{5.2}$$

where K is the hydraulic conductivity and H is the aquifer thickness. The complex leakage q_t is a function of the head above the leaky layer (h_0) and within the aquifer, the storage coefficient σ and resistance c of the leaky layer. The head above the aquifer is set to zero, which simplifies the equation for the leakage to (Bakker, 2019; Eq. 12)

$$q_t = gh, \tag{5.3}$$

where g is defined as

$$g = \frac{\sqrt{i\omega_j \sigma c}}{c\ \tanh(\sqrt{i\omega_j \sigma c})}, \tag{5.4}$$

where ω_j is the frequency of constituent j. See Bakker (2019) for the full derivation of the flow through a leaky layer. It is assumed that the sea cuts through the aquifer system (see Box 5.1), therefore the head in the aquifer at the shore is equal to the water level in the sea. The head in the aquifer far from the shore, below the land, is equal to zero.

$$\begin{aligned} h &= h_{sea} = A\cos(\omega_j t) && x = 0 \\ h &= 0 && x \to \infty. \end{aligned} \tag{5.5}$$

Parameter groups
Now that the governing differential equation (DEQ) and the boundary conditions are introduced, it can be demonstrated that some of the parameters in the governing differential equation 5.1 can be grouped.

First, the simpler case without storage in the leaky layer is considered. Note that when the storage in the leaky layer is neglected ($\sigma = 0$), g (Eq. 5.4) is equal to $1/c$. Using the latter and dividing the governing DEQ (Eq. 5.1) by T, we get

$$\frac{\partial^2 h}{\partial x^2} = \frac{S_s}{K}\frac{\partial h}{\partial t} + \frac{h}{CT}. \tag{5.6}$$

The diffusivity (D) and the leakage factor (λ) are introduced as

$$D = \frac{K}{S_s} \qquad \lambda = \sqrt{cT}, \tag{5.7}$$

Equation 5.6 may be written as

$$\frac{\partial^2 h}{\partial x^2} = \frac{1}{D}\frac{\partial h}{\partial t} + \frac{h}{\lambda^2}. \tag{5.8}$$

Hereby demonstrating that the governing DEQ contains parameter groups, which implies that the individual parameters (i.e. K, c & S_s) cannot be determined uniquely.

If the storage in the leaky layer is not neglected ($\sigma \neq 0$), Equation 5.1 can be rewritten using Equation 5.3 substituted with Equation 5.4 and by dividing the whole equation by T

$$\frac{\partial^2 h}{\partial x^2} = \frac{S_s}{K}\frac{\partial h}{\partial t} + \frac{\sqrt{i\omega_j \sigma c}}{c\,\tanh(\sqrt{i\omega_j \sigma c})} h. \tag{5.9}$$

By using the diffusivity (D), the leakage factor (λ) and η

$$D = \frac{K}{S_s} \qquad \lambda = \sqrt{cT} \qquad \eta = \sqrt{\sigma c}, \tag{5.10}$$

Equation 5.9 may be written as

$$\frac{\partial^2 h}{\partial x^2} = \frac{1}{D}\frac{\partial h}{\partial t} + \frac{\eta\sqrt{i\omega_j}}{\tanh(\eta\sqrt{i\omega_j})}\frac{h}{\lambda^2}. \tag{5.11}$$

Again the governing DEQ contains parameter groups, implying that the individual parameters (i.e. K, c, S_s & σ) cannot be determined uniquely. In fact, these parameter groups are also coupled, since all parameter groups contain at least one hydraulic parameter from another parameter group. Implying that if we would randomly change one parameter value of an (optimized) parameter set, all other parameters can be changed so that all parameter groups (Eqs. 5.10) are again identical to the optimized parameter groups, which subsequently gives exactly the same tidal propagation as for the optimized parameter groups. Or in other words, only the parameter groups can be estimated uniquely. Note that this implication also holds for the case without storage in the leaky layer ($\sigma = 0$).

When can the storage in the leaky layer be neglected?
In several studies that presented an analytical one-dimensional solution for a coastal aquifer the storage in the leaky layer is neglected (e.g. Li and Jiao, 2001; Chuang and Yeh, 2007). Some comprehension about for what conditions this is reasonable can be gained from the simplified and rewritten solution to the governing DEQ (Eq. 5.1) and the applied boundary conditions (Eq. 5.5).

Application of the boundary conditions to the governing differential equation yields the following (complex) solution:

$$h = h_{sea} e^{-x\sqrt{\alpha}} e^{i\omega_j t}, \tag{5.12}$$

where i is the imaginary unit and α is defined as

$$\alpha = \frac{g + i\omega_j S}{T}, \tag{5.13}$$

See Bakker (2019) for a complete derivation. The head in the aquifer can be expressed in terms of an amplitude and a phase shift, for this purpose, Equation 5.12 is written as

$$h = (u + iv)e^{i\omega_j t} = \rho e^{i\omega_j (t-\theta)}, \tag{5.14}$$

where ρ is the amplitude and θ is the phase of the tidal fluctuation at x

$$\rho = \sqrt{u^2 + v^2} \qquad \theta = -\frac{1}{\omega_j} \arctan(v/u). \tag{5.15}$$

ρ may be written as (see Appendix A.2; Eq. A.23)

$$\rho = \sqrt{h_{sea}^2 \exp\left\{ -2x \left[\underbrace{\left(\frac{\Re g}{T}\right)^2}_{\text{term 1}} + \underbrace{\left(\frac{\omega_j S + \Im g}{T}\right)^2}_{\text{term 2}} \right]^{1/4} \cos\left[\frac{1}{2} \arctan\left(\underbrace{\frac{\omega_j S + \Im g}{\Re g}}_{\text{term 3}} \right) \right] \right\}}, \tag{5.16}$$

where $\Re g$ and $\Im g$ are the real and imaginary part respectively of g (Eq. 5.4). Furthermore, θ may be written as (see Appendix A.2; Eq. A.24)

$$\theta = \frac{x}{\omega_j} \left[\underbrace{\left(\frac{\Re g}{T}\right)^2}_{\text{term 1}} + \underbrace{\left(\frac{\omega_j S + \Im g}{T}\right)^2}_{\text{term 2}} \right]^{1/4} \sin\left[\frac{1}{2} \arctan\left(\underbrace{\frac{\omega_j S + \Im g}{\Re g}}_{\text{term 3}} \right) \right]. \tag{5.17}$$

For a full derivation see appendix A.2. Note that, terms 1,2 and 3 in Equation 5.16 contain the four model parameters, moreover these terms are present in Equation 5.17 as well.

If σ > 1e-6 then $\Re g$ and $\Im g$ are both functions of σ and c. However, if $\sigma \leq$ 1e-6 than g (Eq. 5.4) can be approximated accurately by its first order Taylor expansion at $\sigma = 0$. $\Re g$ and $\Im g$ are then defined as

$$\Re g = \frac{1}{c} \qquad \Im g = \frac{\omega_j \sigma}{3}, \tag{5.18}$$

If the magnitude of the numerator of 'term 2' from equations 5.16 & 5.17 is predominantly determined by $\omega_j S$ (i.e. S should be at least one order of magnitude bigger than σ) than the storage in the leaky layer can safely be neglected. Moreover, this result gives some information about the sensitivity of the model regarding σ. If for example σ = 2e-6 and S_s is a few orders bigger, then the sensitivity of the model, regarding σ, will likely be small and the resulting estimation of σ should be treated with care.

5.4.2. Two-layer solution

In this section, the governing differential equation and solution for a two-layer system, as described in Bakker (2019), are presented. Note that, Bakker (2019) describes a more general multi-layer solution.

Flow in the aquifer layer n (1 or 2), below the land, is governed by the following equation:

$$T_n \frac{\partial^2 h_n}{\partial x^2} = S_n \frac{\partial h}{\partial t} + q_{n,b} - q_{n+1,t} \qquad x \geq 0, \tag{5.19}$$

Applying the boundary conditions (5.5) to this differential equation yields the following solution:

$$h = \vec{h}_{sea} e^{-x\sqrt{\mathbf{A}}} e^{i\omega_j t}, \tag{5.20}$$

where $\mathbf{A}$ is defined as

$$\mathbf{A} = \mathbf{T}^{-1}(\mathbf{F} + \mathbf{i}\omega_\mathbf{j}\mathbf{S}), \tag{5.21}$$

where **T** and **S** and are diagonal matrices and **F**, for a two-layer solution, is defined as

$$\mathbf{F} = \begin{bmatrix} g_1 + g_2 & -g_2 \\ -f_2 & g_2 \end{bmatrix} \tag{5.22}$$

where f_n and g_n are defined as

$$f_n = \frac{\lambda_n}{c_n \sinh(\lambda_n)} \qquad g_n = \frac{\lambda_n}{c_n \tanh(\lambda_n)}, \tag{5.23}$$

where λ_n is defined as

$$\lambda_n = \sqrt{i\omega_j \sigma_n c_n}. \tag{5.24}$$

See Bakker (2019) for a complete derivation. Further note that, when $\sigma_n = 0$ (i.e. no storage in leaky layer n), $\lambda_n = 0$ and f_n and g_n reduce to

$$\lim_{\lambda \to 0} f_n = \lim_{\lambda \to 0} g_n = \frac{1}{c_n}. \tag{5.25}$$

This slightly reduces the complexity of solution 5.20. For this system with two aquitards solving Equation 5.20 involves taking the square root of matrix A (2×2). The square root of a 2 by 2 matrix can be solved without recursion (e.g. Levinger, 1980). However, further rewriting the solution was not done in this report. Instead, the possible existence of coupled parameters in the solution will be exposed in the optimization process when high correlations (close to ±1) are encountered.

5.5. Model calibration

The methods used to calibrate the hydraulic parameters based on the amplitude and phase data are described in this Section.

5.5.1. Weighting factors

The model is calibrated with both amplitude and phase data, obviously these have different units and therefore weighting must be performed. Weighting should achieve 2 objectives (Hill and Tiedeman, 2006). First, weighting should produce weighted residuals with the same dimensions. Second, weighting factors should reflect the errors in the input data. The influence of data that is less accurate should be weighted less relative to accurate data. Both objectives are achieved by making the weights proportional with the inverse of the variance of the observations, see Hill and Tiedeman (2006) for more details.

Unfortunately using the inverse variances as weights for the objective function resulted in the outcome being predominantly determined by the M2 amplitude. Moreover, the model error is a lot bigger than the observation errors, which makes a focus on a small vs really small standard deviation practically irrelevant. Therefore using the inverse variance as weights was considered to be inappropriate. Nevertheless, it is mentioned here because it might be appropriate for other situations and ideally the data error should be formally included in the model (Hill and Tiedeman, 2006).

Therefore an alternative method of determining weights is used. With this method, weights are chosen such that the data is normalized (i.e. all variables are brought to the same range). This is done by taking the inverse of the difference in phase lag and amplitude attenuation between well C and the Oosterschelde as the weighting factor. However, for the amplitude, this would distort the results because of an approximately exponentially decaying amplitude. This evidently causes the amplitude to differ considerably in magnitude for the different wells. Therefore we take the logarithm, with base 10, of the amplitude in the well over the amplitude in the Oosterschelde in percentage. Subsequently, this logarithmic observation is used for the optimization. Note that these logarithmic amplitudes (in %) are also plotted in Figure 4.10. In Figure 4.10 this effectively means that both the attenuation and phase shift plot have an upper and lower y-axis limit differing by exactly 1 (i.e. a phase shift ranging from 0 to 1 and an amplitude attenuation ranging from -0.23 to 0.77). This weighting method ensures that deviations from the trend-line such as vissible in Figure 4.10 are basically equal to how they would be determined graphically. This method of determining weights may seem

somewhat subjective but in practice that is always the case (Hill et al., 1998).

5.5.2. Objective function & Least-squares minimization

For the analysis of the data, Python is used. We define a function that has parameters as arguments that need to be optimized. Inside the function, we put the model of Bakker (2019) and minimize the weighted sum of squared residuals (WSSE), which is calculated as follows:

$$WWSE = \sum_{k=1}^{M} \left([y_k - y'_k(\vec{b})] * w_k \right)^2, \tag{5.26}$$

with y_k the observed value, y'_k the simulated value, w_k the associated weight, and $\vec{b}$ a vector containing the parameters. Equation 5.26 is the objective function that is minimized. The objective function is minimized using a differential evolution algorithm (global optimizer). The optimization is done with LMFIT, which is a Python package made by Newville et al. (2014) for non-linear least-squares problems. An extra advantage of this package is that it also calculates correlation coefficients and confidence intervals.

As stated in Section 5.3, both a one- and a two aquifer system are considered. For the one-aquifer model configuration, one model is optimized with and another without the centroid data (see Table 5.3), so that a total of 3 models is considered. Two one-layer models are considered to see the effect of the A2 centroid data. The terms 'A2 centroid data' and 'A4 centroid data' will be used in the rest of the report to signify both the 'diurnal' and 'high frequency' centroids of wells A2 and A4 respectively.

Lastly, comparing the WWSE of the different models only makes sense if the residual is calculated for the same observations. Therefore, three WSSE are defined $WSSE_4$, $WSSE_6$ and $WSSE_{10}$ similar to the objective function of the one-aquifer model without centroid data, with centroid data and the two-aquifer model respectively (see Table 5.3).

Table 5.3: The observations and corresponding weights for the different models. Note that the observations are presented in the form in which they are optimized.

				Model names		
Observations	**Value**	**Units**	**Weights**	**With centroid data (one-aquifer model)**	**Without centroid data (one-aquifer model)**	**Two-aquifer model**
$\log_{10}\left(\frac{\text{M2 amp in well A2}}{\text{M2 amp in Oostsch}} * 100\%\right)$	1.48	-	0.38	✓	✓	✓
M2 phase in well A2	0.91	rad	0.15	✓	✓	✓
$\log_{10}\left(\frac{\text{M2 amp in well C2}}{\text{M2 amp in Oostsch}} * 100\%\right)$	-0.60	-	0.38	✓	✓	✓
M2 phase in well C2	6.75	rad	0.15	✓	✓	✓
$\log_{10}\left(\frac{\text{M2 amp in well A4}}{\text{M2 amp in Oostsch}} * 100\%\right)$	1.51	-	0.38			✓
M2 phase in well A4	0.89	rad	0.15			✓
A2 Diurnal centroid	0.74	rad	1	✓		✓
A2 High frequency centroid	2.51	rad	1	✓		✓
A4 Diurnal centroid	0.68	rad	1			✓
A4 High frequency centroid	2.37	rad	1			✓

With amp and Oostsch as abbreviations for the amplitude and Oosterschelde respectively.

5.5.3. Parameter constraints

In the optimization, parameters can be constrained i.e., impose an upper and lower limit on the parameters. In this section, it will be identified what parameter estimates are reasonable. The estimated ranges of realistic parameter estimates will be used as default optimization constraints. Although it must be noted that allowing the model to give unrealistic parameter estimates can be beneficial, as unrealistic parameter estimates can

indicate problems with the model (Hill et al., 1998).

In Appendix C.2.1 maps of the 2.5th and 97.5th percentiles of the hydraulic conductivity are presented. In section 2.1.1 it is explained how these maps are derived. The small rectangular polygons, drawn in these figures, were used to calculate statistics of the transects between the three wells and the Oosterschelde (see e.g. Figure C.4). Conservative estimates of the lower and upper limits are made. So that for the lower limit the minimum value present in the transects of well A and B was chosen from the 2.5 percentile map and for the upper limit, the maximum value present in the transects of well A and B was chosen from the 97.5 percentile map of the hydraulic conductivity. The upper and lower limits of CL and KL were determined similarly. The resulting limits are presented in Table 5.4.

TNO has no raster data publicly available for the resistance of the upper leaky layer CU (i.e. the Holocene layer). However, in a report from Stafleu and Gunnink (2016) maps showing the resistance of this layer are presented. The lowest estimate of CU at the study site is 200 days and the highest 5000 days (see Section 7.3.2 for more details about these estimates). However, no data about the uncertainty of these estimates was provided. Therefore, the lower limit was essentially ignored (CU = 1) and the upper limit was taken ample (CU = 10000).

TNO has no publicly available data on the specific storage of the clay or sand layers. Estimates on specific storage for different materials are for example provided by Jumikis (1962). However, data on for example the amount of consolidation and the porosity of the layers is missing and so it is unknown how comparable the layers are. Therefore the limits for the specific storage are taken ample (see Table 5.4).

Table 5.4: The default parameter constraints used for the optimization.

	Optimization constraints			
Parameters	**Lower limit**	**Upper limit**	**Unit**	**Meaning**
KU	3.42	14.7	m/day	Hydraulic conductivity of the upper aquifer
S_sU	1E-7	5E-3	m^{-1}	Specific storage of the upper aquifer
CU	1	10000	days	Resistance of the upper leaky layer
σU	1E-7	5E-2	m^{-1}	Specific storage of the upper leaky layer
KL	2.71	14.53	m/day	Hydraulic conductivity of the lower aquifer
S_sL	1E-7	5E-3	m^{-1}	Specific storage of the lower aquifer
CL	99	3053	days	Resistance of the lower leaky layer
σL	1E-7	5E-2	m^{-1}	Specific storage of the lower leaky layer

5.6. Model performance

In general, a model is considered better if its fit statistics (e.g. weighted residuals) are smaller, the parameter estimates are more realistic and the weighted residual is normally distributed (Hill and Tiedeman, 2006). Analyzing the latter is not possible in this case because of a small number of observations. Consequently, in this report, the model performance will be evaluated with the weighted sum of squared residuals, besides, the model fit will be evaluated with plots of the amplitude attenuation and phase shift. Furthermore, the extent to which the optimized parameters are realistic will be analyzed, although this analysis is slightly impeded because the parameters are coupled (at least for the one-layer model; see Section 5.4).

6

Results: Inverse modeling

In this chapter, the results of the inverse modeling are presented. The results are discussed in Section 6.4.

Both a one and a two-aquifer model have been optimized using the LMFIT Python package (Newville et al., 2014). First, a one-aquifer model is considered. To see what the effect is of the A2 centroid data, an optimization was performed without (Section 6.1) and with (Section 6.2) this data. In Section 6.3 a two-aquifer system is considered. Table 5.3 presents the observations that were used for the different models. Different sets of observations were used for the different models. Moreover, different weights were assigned to the input parameters (see Section 5.5.1 and Table 5.3).

6.1. Without centroid data (One-aquifer model)

First, a one-layer aquifer schematization is considered, so that only the upper sand and clay layers are parameterized (CU, σU, KU and SsU; see Figure 5.1). The layers below the upper sand layer are regarded as impermeable.

In Table 6.1, the optimal parameter estimates obtained by minimizing the 'without centroid data' model are presented, along with the parameter constraints.

Table 6.1: Presenting the optimal parameter estimates obtained with the: 'Without centroid data' model. Besides, the (default) parameter constraints used for the optimization are presented.

		Optimization constraints			
Parameters	**Optima**	**Lower limit**	**Upper limit**	**Unit**	**Meaning**
KU	14.7	3.42	14.7	m/day	Hydraulic conductivity of aquifer
S_sU	1.12E-5	1E-7	5E-3	m^{-1}	Specific storage of aquifer
CU	9999	1	10000	days	Resistance of leaky layer
σU	1E-7	1E-7	5E-2	m^{-1}	Specific storage of leaky layer

As discussed in Section 5.4.1, only the following parameter groups (Eqs. 6.1) can be estimated uniquely.

$$D = \frac{K}{S_s} \qquad \lambda = \sqrt{cT} \qquad \eta = \sqrt{\sigma c}, \tag{6.1}$$

However, in the optimization KU & CU and thus also λ, are estimated at their upper limit, this remains the case if the upper limit is increased. Therefore the leakage factor (λ) cannot be estimated. σU is estimated at its lower limit, which was kept at 1e-7. So σ < 1e-6 and S_sU is estimated two orders higher, therefore the storage in the leaky layer can safely be neglected (see Section 5.4.1). In other words, fixing σU at 0 gives the same result for the other parameters. The diffusivity (KU/S_sU = 1.31E6 m^2/day) on the other hand remains essentially constant, even when the upper limits of KU & CU are adjusted. Hence, only the diffusivity could

be estimated uniquely in this optimization.

In Figure 6.1 the amplitude attenuation and phase shift (associated with the optimized parameter set) against the distance to the Oosterschelde are plotted. The amplitude and the phase are modeled similarly well, both the phase and amplitude in well A2 are overestimated and both the phase and amplitude are underestimated in well C2. In Figure 6.2 the phase shift in well A2 is plotted against the frequency. The model was not optimized with these two data points, but the phase shift for the Diurnal centroid is modeled quite well. The phase shift for the high-frequency centroid is underestimated quite a bit. Lastly, the $WSSE_4$ and $WSSE_6$ of the optimal parameter set are 1.27E-2 and 1.43E-1 respectively.

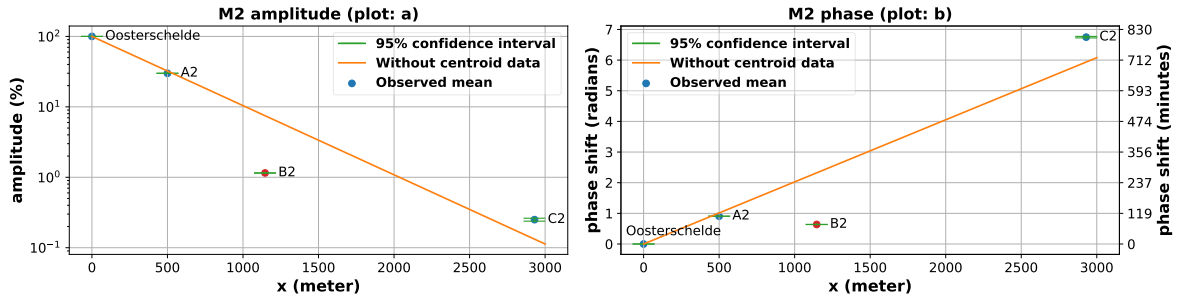


Figure 6.1: Plot of the amplitude (plot: a) and phase shift (plot: b) against the distance to the Oosterschelde, for constituent M2. With the observed mean (blue dot), and the modeled amplitude attenuation and phase shift for the optimized model without centroid data (orange line). Note that, although plotted, observations from B2 (red dot) were not used in the optimization.

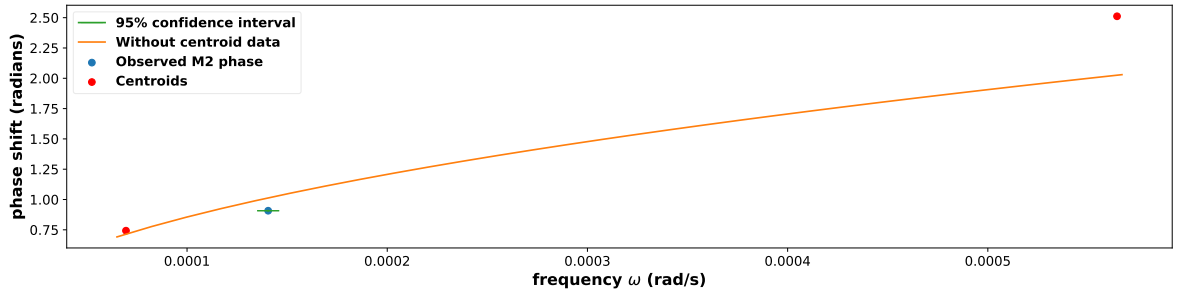


Figure 6.2: Plots of the phase shift (radians) in well A2 (compared to the Oosterschelde) against the frequency for the model without centroid data (orange line). With the M2 constituent in blue and the centroid data in red. Note that the centroid data has not been used to optimize this model.

6.2. With centroid data (One-aquifer model)

In this section, the one-aquifer model is optimized with the centroid data. In table 6.2, the optimal parameter estimates, obtained by minimizing the 'without centroid data' model, along with the parameter constraints (same as in Table 6.1), are presented.

Table 6.2: Presenting the optimal parameter estimates obtained with the: 'With centroid data' model. Besides, the (default) parameter constraints used for the optimization are presented.

		Optimization constraints			
Parameters	**Optima**	**Lower limit**	**Upper limit**	**Unit**	**Meaning**
KU	14.7	3.42	14.7	m/day	Hydraulic conductivity of aquifer
S_sU	1.38E-5	1E-7	5E-3	m^{-1}	Specific storage of aquifer
CU	10000	1	10000	days	Resistance of leaky layer
σU	4.85E-7	1E-7	5E-2	m^{-1}	Specific storage of leaky layer

Similar to the model without centroid data, KU & CU and thus also λ are estimated at their upper limit and again this remains the case if the upper limit is increased. σU is not estimated at its lower limit, but because σ < 1e-6 and S_sU is estimated two orders higher, the storage in the leaky layer can safely be neglected. The diffusivity (KU/S_sU = 1.07E6 m^2/day) on the other hand remains essentially constant, even when the upper limits of KU & CU are adjusted. Hence, similar to the model without centroids, only the diffusivity could be estimated uniquely in the optimization.

Figure 6.3 & 6.4 show that the different models produce quite different results. For the model with centroid data, the amplitude in C2 is modeled poorly but the phase in C2 is modeled well and the phase of the high-frequency centroid is also modeled quite well. For the model without the centroid data, the amplitude in C2 is modeled better compared to the model with centroid data, but the phase in C2 is modeled worse. Also, the phase of the M2 tide in A2 is modeled better compared to the model with centroid data. So depending on which observations are emphasized to reduce the error in the optimization, the attenuation and phase shift can be quite different. Lastly, the $WSSE_4$ and $WSSE_6$ of the optimal parameter set are 3.63E-2 and 5.77E-2 respectively. As expected $WSSE_4$ is lower for the model 'without centroid data' and $WSSE_6$ is lower for the model 'with centroid data'.

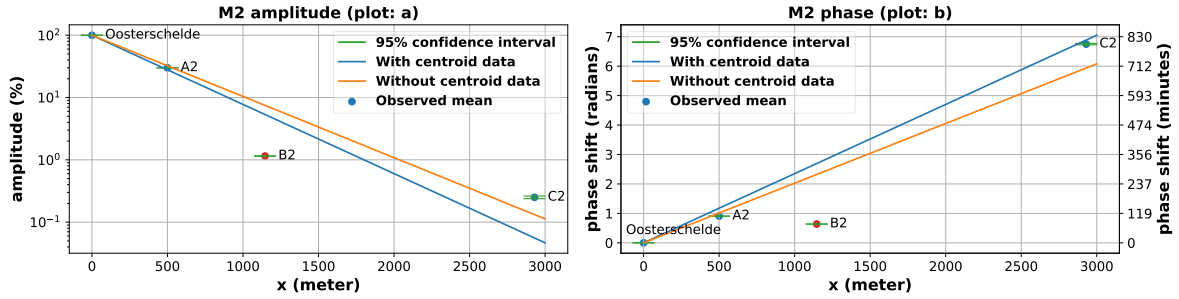


Figure 6.3: Plot of the amplitude (plot: a) and phase shift (plot: b) against the distance to the Oosterschelde, for constituent M2. With the observed mean (blue dot), and the modeled amplitude attenuation and phase shift when centroid data (blue line) and no centroid data (orange line) were used in the optimization. Note that, although plotted, observations from B2 (red dot) were not used in the optimization.

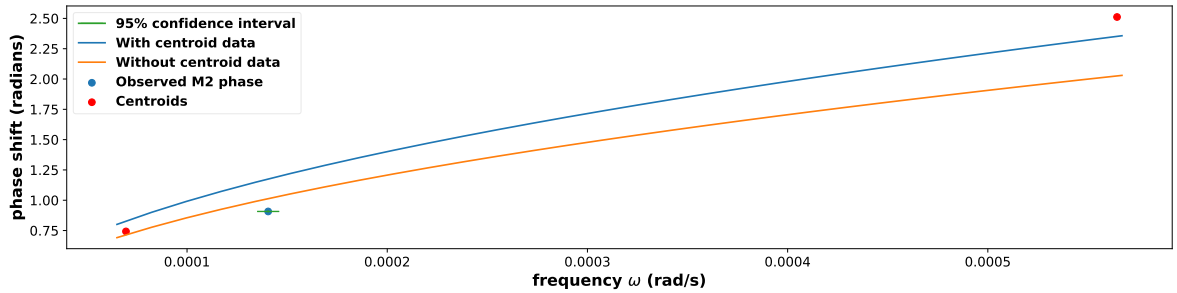


Figure 6.4: Plots of the phase shift (radians) in well A2 (compared to the Oosterschelde) against the frequency. With the M2 constituent in blue and the centroid data in red. And with the modeled amplitude attenuation and phase shift when centroid data (blue line) and no centroid data (orange line) were used in the optimization.

To still make some useful statements about CU, σU and S_sU (besides the diffusivity estimate) a different approach was considered. Instead of seeking an optimal parameter-set, it was analyzed for what parameter values the results are poor. This was done by fixing one of the parameters, though varying its constraints. For the other parameters the default constraints (see Table 6.1 or 6.2) are used. Note that KU is not considered here, because, for KU ranging from 3.42 m/day to 14.7 m/day, good results were obtained.

First, CU is considered. The results of this analysis are that $CU \to \infty$ still gives good results. However, if CU is given a low resistance the model gives a poor fit. In Figure 6.5 the model is plotted with CU = 663 days and the other parameters optimized within the default optimization constraints. The $WSSE_6$ of this result is 0.2869. For σU the above procedure was performed as well, σU is 0 m^{-1} gives good results. However, if σU is too large, this gives a poor fit. In Figure 6.5 the model is plotted with σU = 1.22e-4 m^{-1} and the other parameters optimized, the $WSSE_6$ of this result is also 0.2869. Lastly S_sU is considered, if 2.28E-6 < S_sU < 1.87E-5 m^{-1} is outside of this range, a poor fit is obtained. The $WSSE_6$ associated with this upper and lower limit of S_sU is 0.2869. Note that determining when the fit is poor combined with the associated $WSSE_6$ is a somewhat arbitrary choice.

Note that only the model with the centroid data was used for these calculations. The model without centroid data produced diffusivity (=KU/S_sU) estimates that were a lot higher compared to the optimal estimates from both one-aquifer models. Moreover, the discrepancy between the modeled and observed A2 centroid data was very bad.

The plots in Figure 6.5 are almost similar for CU and σU. The phase shift is simulated quite well, however, the amplitude is far off for both well A and C, although for well A this is hardly noticeable on a log-scale. Note that such a big inconsistency is not so much visible in Figure 6.3. The results for the phase shift are similar compared to the results in Figure 6.4, but especially the estimated amplitude in well A and C are a lot worse. Lastly, both the lower and upper limit of S_sU have a poor fit in the amplitude attenuation plot and an even worse fit in the phase shift plot.

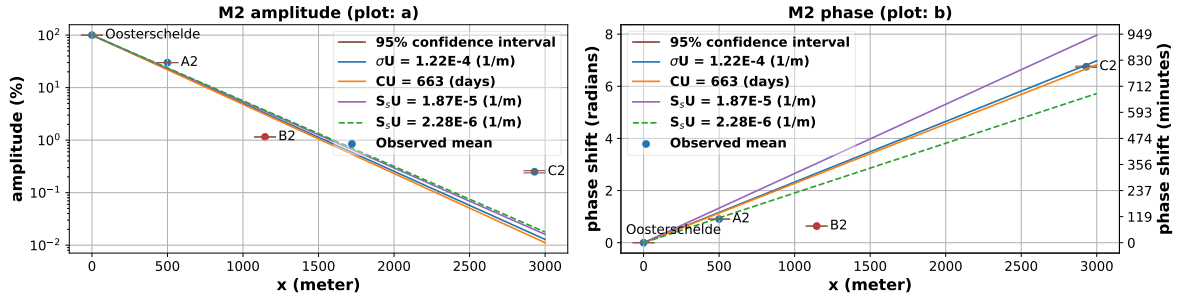


Figure 6.5: Plot of the amplitude (plot: a) and phase shift (plot: b) against the distance to the Oosterschelde, for constituent M2. With the observed mean (blue dot), and the modeled amplitude attenuation and phase shift with σU = 1.22e-4 m^{-1} (blue line), CU = 663 days (orange line), S_sU = 1.87E-5 m^{-1} and S_sU = 2.28E-6 m^{-1} (green dashed line) as fixed values in the optimization. Note that, although plotted, observations from B2 (red dot) were not used in the optimization.

In Table 6.3, the results obtained with the one-aquifer model are summarized. Estimates of the diffusivity were obtained by optimization with both the model 'with centroid data' & 'without centroid data'. For CU, σU and S_sU a range of values is determined which contains likely parameter estimates (based on a limited $WSSE_6$ value). Values outside these ranges are classified as not very likely.

Table 6.3: Summary of the results obtained with the one-aquifer model.

	Estimates			
Parameters	**With centroid data**	**Without centroid data**	**With**	**Meaning**
Diffusivity (KU/S_sU)	1.07E6	1.31E6	m^2/day	Diffusivity of aquifer
	Estimated ranges			
CU	[663, $\longrightarrow >$		days	Resistance of leaky layer
σU	[0, 1.22E-4]		m^{-1}	Specific storage of leaky layer
S_sU	[2.28E-6, 1.87E-5]		m^{-1}	Specific storage of aquifer

6.3. Two-aquifer model

A two-aquifer schematization was considered as well. So that 8 parameters need to be optimized (CU, CL, σU, σL SsU, KL, SsU and SsL; see Figure 5.1). The layers below this model are regarded as impermeable.

For this optimization, the data observed in well A4 can be used as well, so that there are 10 observations to determine 8 parameters (see Table 5.3). The optimization constraints are presented in Table 5.4.

There is little data to constrain the parameter values. This problem became apparent in the optimization process as well. Unlikely parameter estimates were obtained when only the default minimum and maximum parameter bounds were imposed (e.g. KL was estimated to be more than 2 times bigger than KU). By adding mathematical expressions to the optimization this could be prevented (e.g. $0.5 \geq KL/KU \leq 2$). Another complication was that CU was consistently estimated to be smaller than CL. However, according to borehole logs and the REGIS-II model Vernes et al. (2005) this is very unlikely (see Figure 2.2). Although this could also be prevented with a mathematical expression, the resulting optimization estimated CL to be consistently at the lower constraint limit.

The results of the latter model are shown in Figures 6.6 & 6.7. The obtained fit looks similarly good as for the one-aquifer model. The parameter estimates are considered to be poor and are therefore not shown here, in section 6.4 these results are further discussed.

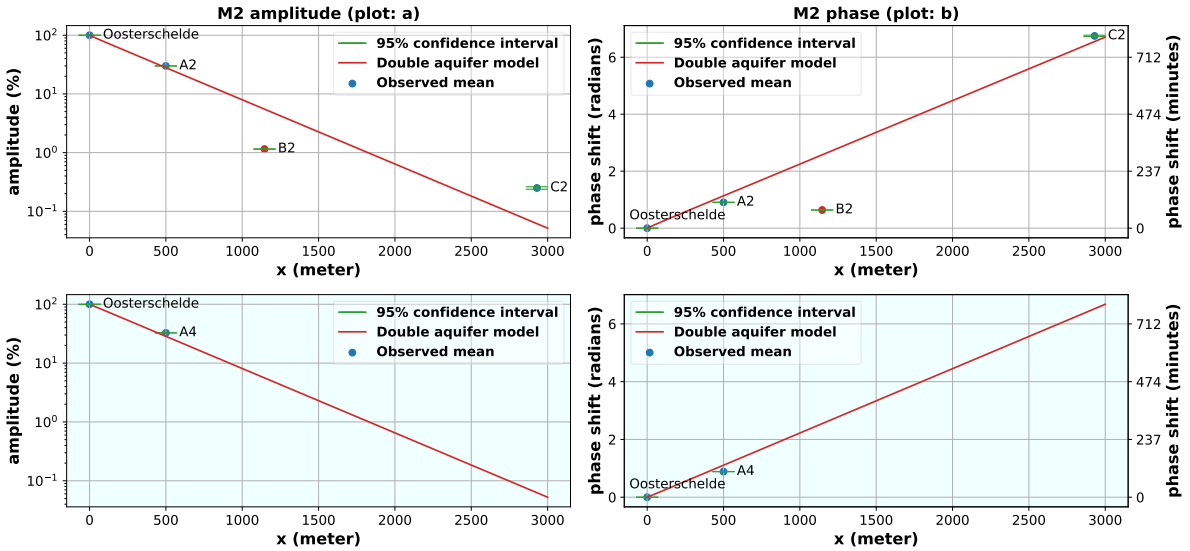


Figure 6.6: Plot of the amplitude (plot: a) and phase shift (plot: b) against the distance to the Oosterschelde, for constituent M2. With the observed mean (blue dot), and the modeled amplitude attenuation and phase shift when centroid data (blue line) and no centroid data (orange line) were used in the optimization. The plots with the white background are from the upper aquifer and the plots with the azure background are from the lower aquifer. Note that, although plotted, observations from B2 (red dot) were not used in the optimization.

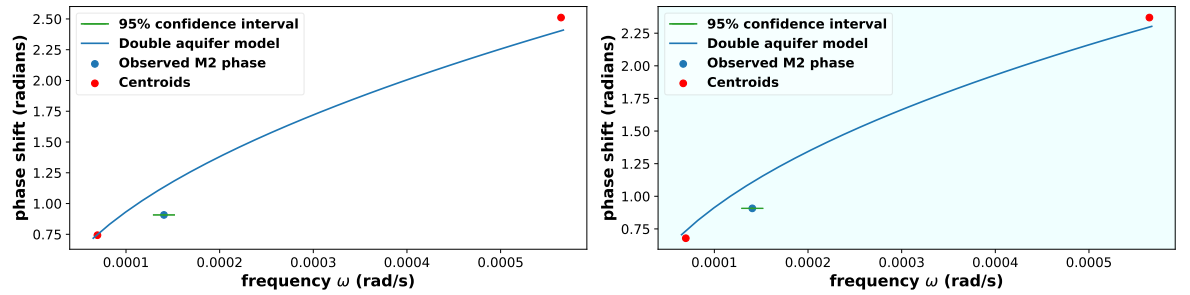


Figure 6.7: Plots of the phase shift (in radians compared to the Oosterschelde) against the frequency. With the M2 constituent in blue and the centroid data in red. And with the modeled amplitude attenuation and phase shift when centroid data (blue line) and no centroid data (orange line) were used in the optimization. The plot with the white background is from the upper aquifer and the plot with the azure background is from the lower aquifer.

6.4. Discussion

In this section, the results of the inverse modeling will be discussed.

6.4.1. One-aquifer model

Results for this schematization were obtained by modeling with and without centroid data. This was done to see the effect of the centroid data on the optimization result. An additional benefit is that, in this case, the two models emphasize on minimizing the residual of different observations. As a result, it can be determined if the latter significantly changes the parameter estimates (i.e. the diffusivity). The diffusivity estimates for the model with and without centroid data are 1.07e6 and 1.31e6 m^2/day respectively. This indicates that changing the emphasize on the different observation does significantly affect the model outcome, although the ratio of the two diffusivity is merely 1.22. A similar conclusion can be drawn for changing the weights (see Table 5.3) as weights also determine which error is minimized most by the model.

In the attenuation and phase shift plots (Figures 6.3 & 6.4) small inconsistencies are observed between the phase and amplitude plots, especially visible when looking at the amplitude and phase in C2. This dis-

crepancy is not presumed to be due to a large 'observation' error. However, it might be explained by a small deviation from the theoretical exponentially decaying amplitude and linearly increasing phase shift. The latter will result in the biggest discrepancy for the well furthest from the sea and thus in C2. Furthermore, because both models reduce the error for some of the observations and less for others, it can not be decided which model is better.

For both models, the fit is not that well indicating that, to some extent, the model is incorrectly representing the system (Hill and Tiedeman, 2006; see also Section 5.6). Another indication of an oversimplified model are unrealistic parameter estimates. The leakage factor (i.e. $\lambda = \sqrt{cT}$), which is one of the parameter groups (see Equations 6.1 & 5.10), was constantly estimated at its upper limit, even when the latter was increased. This also indicates oversimplification of the system and therefore, the resulting parameters should be treated with care. More about the used simplifications in Section 7.2.

On the other hand, the diffusivity estimates of both models are within plausible ranges for confined aquifers (e.g. Domenico and Mifflin, 1965; Malama et al., 2009). The diffusivity (D) was estimated to be 1.07e6 and 1.31e6 m^2/day for the model with and without centroids respectively. Since σU was estimated to be smaller than 1E-6 and 2 orders lower than S_sU, this result is essentially obtained with a model without storage in the leaky layer (see Section 5.4.1). If the model is forced to have significant storage in the leaky layer, the residual increases and the diffusivity estimate slightly increases. Note that no form of sensitivity analysis was done for the diffusivity so the likelihood of D falling inside the range [1.07e6, 1.31e6] m^2/day is unknown. As mentioned, the fit of the model was not perfect for both the model with and without the centroid data. A compromise must be sought based on what residual the model minimizes. As a result, the diffusivity estimates within and around the range [1.07E6, 1.31E6] m^2/day are all considered to be reasonable estimates.

Lastly, it was analyzed for what parameter values the results are unrealistic. For CU a lower limit was estimated, for σU an upper limit and for S_sU both an upper and lower limit (see Table 6.3). These boundaries are of course a bit arbitrary. However, these limits were estimated with KU estimated at its upper limit (14.7 m/day) or at its lower limit (for S_sU = 2.28E-6 m^{-1}). While according to REGIS-II (Vernes et al., 2005) an estimate of 14.7 m/day and 3.42 m/day is at the 97.5 and 2.5 percentile of the estimated hydraulic conductivity respectively. Moreover, the other (not fixed) parameters were not limited with narrow constraints. The above two arguments suggest that boundaries are conservative estimates. It is (again) noted that the optimized model does not produce an optimal fit. Therefore a more conservative estimate is reasonable. Taking these facts into consideration these boundaries (see Table 6.3) can be reasonably trusted.

6.4.2. Two-aquifer model

For the two-aquifer model, 8 parameters had to be estimated, based on 10 observations. Such a low amount of information to constrain the optimization proved to be problematic. Presumably because of high parameter correlations as well. To resolve this, extra information was provided to the model using extra constraints in the form of assumptions (e.g. $CU > CL$). Unfortunately, this resulted in CL being consistently estimated at the imposed lower boundary. Moreover, as mentioned, unreasonable optimized parameter estimates indicate problems with the data or the model (Hill et al., 1998). Furthermore, because of long computation times and a high number of parameters, extensive analysis of this configuration is very time-consuming. In summary, the two-aquifer model seems to have too many parameters to tune and too little data to constrain the results and therefore the two-aquifer model was not used for further evaluation of the hydraulic parameters.

7

Discussion

In this chapter, the main uncertainties and limitations are discussed. It starts with a general section about errors in the data (Section 7.1), followed by a discussion about the simplifications (in Section 7.2) or more specifically: the assumptions, the schematization and some unconsidered physical phenomena. In Section 7.3 the results are compared with case studies that also performed the tidal method. Moreover, in this section, the resistance of the confining layer (CU) is compared with the estimates of two other studies. Lastly, some comments on the wider application of the results of this research are provided in Section 7.4.

7.1. Data errors

Some errors were found in the groundwater level data. The data from the loggers was not correctly validated with hand measurements. So the measurements by hand were not the same as the measurements registered with the logger. However, this error is presumed to be constant over time, so that calculations for which the reference level is not important are still correct (e.g. for determining the phase and amplitude). Another known problem, often occurring with dataloggers, is a small drift in the data (Ritzema et al., 2012). This can slightly affect the estimated parameters, though the effects are presumably negligible when the time series is large enough, which is the case in this study.

The time series of the groundwater level also contained a sudden drop in the water level, which happened during the replacement of the divers. This is not a big issue as the absolute water level is not important when determining the phase and amplitude. Moreover, the data was corrected for the step. However, the time shift, that also occurred when the divers were replaced, is more inconvenient, since it is harder to determine this shift, and because a shift is important as the phase of the components needs to be determined. How the time shift was handled can be found in Chapter 3, Section 3.2.2.

Some errors are inherent to GW observations, e.g. the time lag between the hydrostatic pressure in the aquifer, the pressure in the piezometer and within the measuring device. There are theoretical and experimental methods to determine the time lag (e.g. Hvorslev, 1951). However, if the hydraulic conductivity of the aquifer is high and the diameter of the well is small, no correction is needed. The latter is assumed in this research.

Although these errors can potentially change the results somewhat, they are not expected to alter the conclusions of this report. Lastly, it is again noted that the amplitudes and phases estimated in well B2 are dissimilar to the results obtained from the Oosterschelde and wells A2 & B2 (see Figure 4.10 and Appendix C.10). However, the magnitude of the head response to recharge in wells A, B & C (estimated with Pastas) are of similar magnitude, indicating that the amplitude is measured, at least approximately, correct. Therefore, the dissimilar result in well B2 cannot be (fully) attributed to a measurement error and thus it is presumed to be caused by relevant processes which are not included in the model and not by large data errors.

7.2. Simplifications

First, some of the assumptions and approximations are discussed (Section 7.2.1). Followed by a discussion of both the schematization of the area (Section 7.2.2) and some unconsidered physical phenomena (Section

7.2.3). Concluding with a general conclusion about the whole section (Section 7.2.4).

Like all models, the model used in this study is a simplification of the real system. The validity of the used assumptions must be checked as much as possible and if the model does a poor job the simplifications used must be reevaluated. As discussed in Section 6.4.1, the fit of both the model with and without the centroid data is somewhat poor and the leakage factor (i.e. $\lambda = \sqrt{cT}$), which is one of the estimated parameter groups, has an unrealistic optimal estimate, moreover, the model is unable to also simulate the signal observed in well B2 adequately (see Section 4.5.2). A poor fit to the observations and unrealistic optimal parameter estimates indicate that the model incorrectly represents the real system, often referred to as model error (Hill and Tiedeman, 2006). In this section, some of the used simplifications are discussed.

7.2.1. Assumptions and approximations

In this section, some of the assumptions and approximations are discussed.

Homogeneity & lumped parameters

Homogeneity is assumed. This is obviously a simplification of the often complex groundwater system. Nonetheless, due to a limited amount of data, this assumption is often needed and is adequate for many modeling purposes. A limited amount of data is also the reason that lumped parameters are used. Note that, lumped parameters are spatially averaged, in this case hydrological, parameters.

The schematizations used are based on the hydrogeological model REGIS-II (Vernes et al., 2005). In Figure 2.2, a cross-section of this model at the observation site is shown. The logic of starting with a one-dimensional model with layers of constant thickness can be understood from this Figure. The thickness of the sand and clay layers is also estimated based on the hydrogeological model REGIS-II (Vernes et al., 2005). REGIS-II serves as a good (base) for a geohydrological model schematization, but obviously, REGIS-II has its uncertainties. The lithology, on which REGIS-II is based, and the conductivities are interpolated from boreholes, of which the latter is derived from the relation between the lithology, hydraulic properties and deposition environment. Uncertainties belonging to these interpolation techniques are calculated as well.

Uncertainties are also calculated for the lithoclass in the subsurface voxel model GeoTOP (Stafleu et al., 2019). A cross-section with the most likely lithoclass is presented in Figure 7.1 for the cross-section shown in Figure 2.1. From this figure, it is clear that the observation site is more heterogeneous than Figure 2.2 shows. The schematization of REGIS-II is more coarse and looks a lot different from Figure 7.1, for several reasons, of which one is that a balanced geographical distribution in relation to the geological complexity was intended for the REGIS-II model (Vernes et al., 2005).

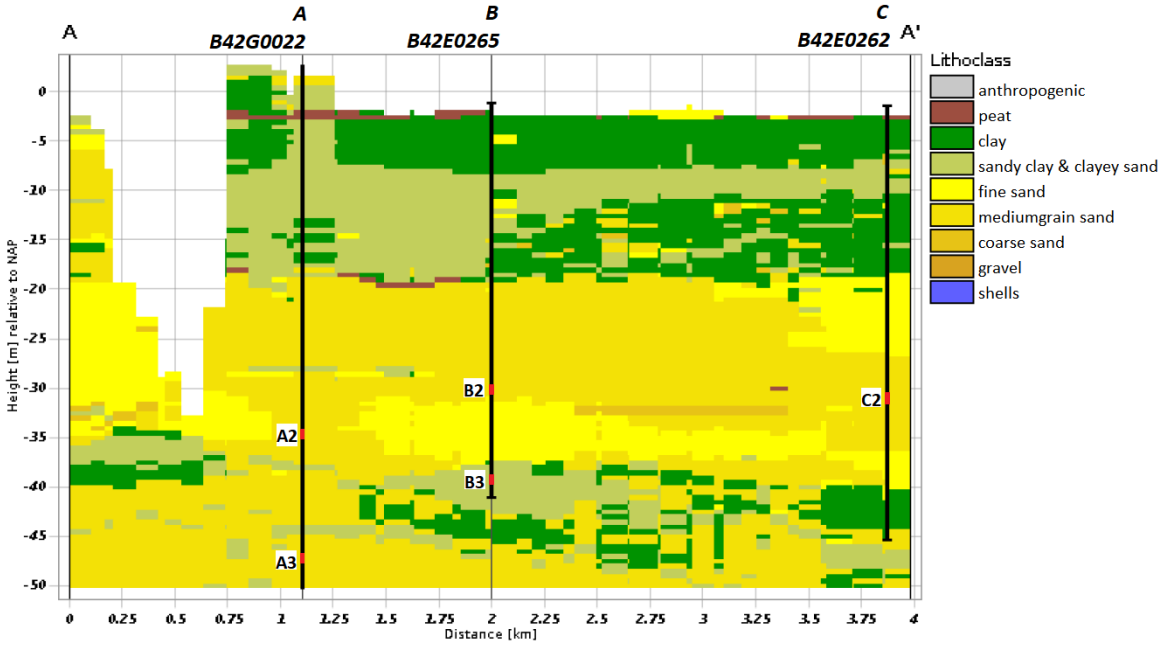


Figure 7.1: Cross Section ofthe GeoTOP model with the most probable lithoclass, with the depth and location of the observation wells (black line) and an indication of the height from the piezometer screens (red marks on the black line). See Figure 2.1 for the location of the cross Section. Publicly available (at: `www.dinoloket.nl`)

Of course, (some) heterogeneity is present and possibly also some preferential pathways. However, not enough information is available to model this. Therefore lumped parameters are a necessity and are thus also used in this report. Consequently, the parameters are estimated using a constant thickness for the different layers, and because the model is one-dimensional, the parameters are also spatially averaged in the lateral direction. Moreover, the thickness of the layers was not differed for the different wells (A & C). A distinction could have been made by optimizing two different models simultaneously. However, the differences in thickness for the used layers were small, moreover, the composition in terms of formations, of the different transects are slightly different. Therefore, too little substantiation was found to make any distinctions. Unfortunately, lumped parameter estimates make it harder to use the estimated parameters for other sites. However as mentioned, assumptions like these are hard to overcome in groundwater modeling, where data is often scarce, which is also the case in this study.

Looking at both Figures 2.2 & 7.1 lumped estimates seem to be quite a good approximation. Although these two Figures show some contradiction about the thickness and continuity of the clayey unit from the Maassluis formation. This clayey unit from Figure 2.2 is rather smooth and without gaps. However, from Figure 7.1 it is clear that the less permeable layers (i.e. sandy clay & clayey clay, clay, peat and sometimes also fine sand) are not as continuous and with the same thickness, for the Maassluis formation, as is the case in REGIS-II, especially around A3. Furthermore, for the two-aquifer model, parameter optimization was problematic, partly because the resistance of the lower clay layer (CL) was consistently estimated to be at the specified minimum (see Section 6.3). The varying thickness of the second clay layer might be the cause of this inconvenience.

Hydraulic head measurements

In the upper aquifer layer two more head measurements have been done (A3 and B3; see Figure 7.1). For a couple of reasons data from these wells has not been used in the optimization. One reason is that it was assumed that the influence of vertical resistance on the groundwater flow is negligible. Which is equivalent to the hydrostatic assumption $\Delta p/\Delta z = -\rho g$ (Bear and Verruijt, 1987). This assumption can be checked by comparing the head in A2 and B2 to the head in A3 and B3 respectively.

Note that in Figure 7.1 it seems that the screen of well B3 is located in a clayey layer. However, the bore log description of this well, which should give a better description of the local situation, suggests the presence of very fine sand. Which makes more sense, because it is unlikely that a piezometer screen has been installed in a clay layer.

In Figure 7.2 the hydraulic heads for wells A and B are plotted for a small time window. The hydraulic head for well B2 is consistently a little more than 12 cm above the head in B3. Likewise, in well A2 the head is consistently (on average 16 cm) above the head in A3.

This indicates that in between well A2 & A3 and B2 & B3 some degree of layering of the bottom profile or a less permeable layer is present, which causes resistance to vertical flow. For both wells this was to be expected based on both the lithology and geohydrology (REGIS-II). However, although the head measurements indicate some amount of layering, differences in density preclude quantifying vertical flow. Furthermore, quantification is also impeded because, as noted before, hand measurements did not match the values measured with the data loggers. Therefore, it can (only) be concluded that there is some layering and that besides this, no peculiar aspects are observed in Figure 7.2.

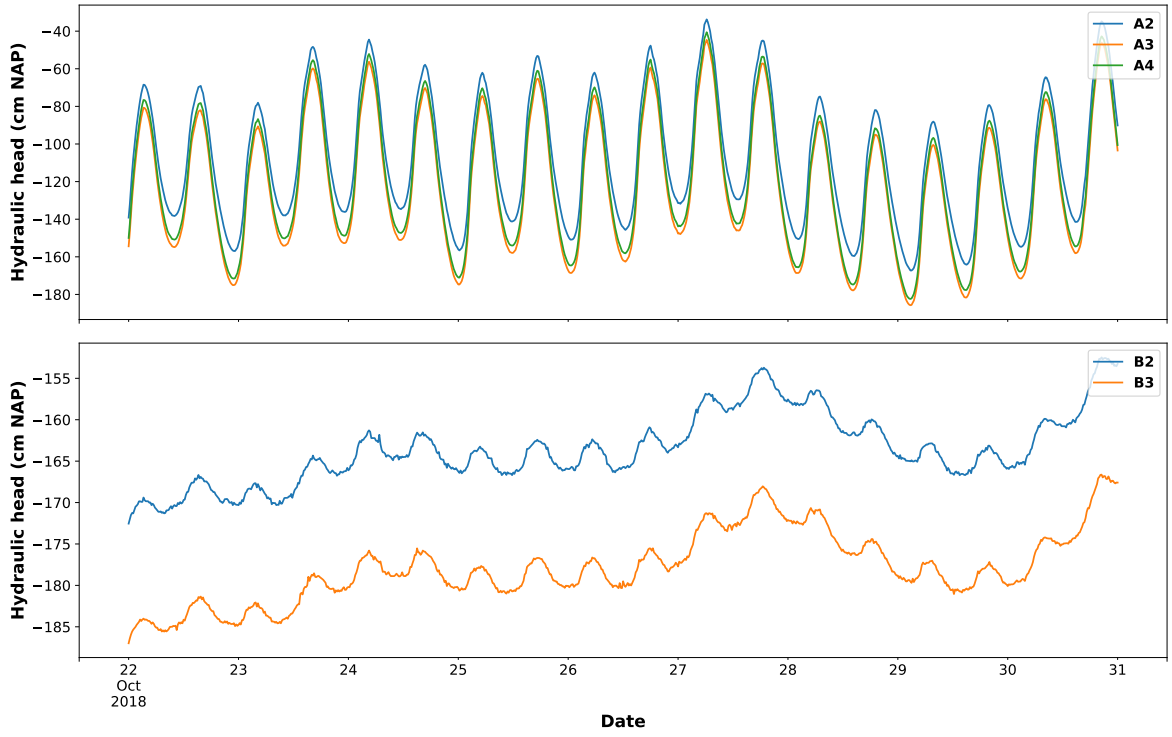


Figure 7.2: Fluctuations of the hydraulic head (cm NAP) for three screens in well A and 2 screens in well B. See Figure 7.1 for their approximate location in the aquifer.

Dupuit-Forchheimer assumption

The validity of the use of the assumptions of Dupuit-Forchheimer and vertical flow in the semi-permeable layer has been investigated by Neuman and Witherspoon (1969) using the finite element method. They found that when the permeability of the aquitard is two or more orders lower than the permeability of the aquifer, this results in errors of usually less than 5% being introduced by this assumption. This condition is easily met for this system.

Although, there is a region around the gap in the Holocene layer (see Figure 2.2 around x=0.5 km), where the vertical flow component in the aquifer might not be negligible. Near the gap in the Holocene layer there might be some significant discharge and thus streamline contraction, this results in head loss (Obergfell et al., 2016). For this research, it is assumed that the decrease in amplitude of the constituents is not affected significantly by this effect. The extent of this effect is unknown but could be estimated with a two or three-dimensional model, the latter is not done in this research.

Temperature

The hydraulic conductivity varies with temperature (± 1 m^2/day per degree Celsius). However, assuming a

constant hydraulic conductivity is useful as it greatly simplifies the modeling process. However, it should be noted that a temperature range of 1.0, 5.5 and 7.0 degrees Celsius was measured (by TNO) for wells A, B and C respectively. Although it can be questioned how representative these temperatures are for the aquifer, especially considering the large differentiation between the temperature range of the different wells.

An attempt was made to roughly check if a temporal variation of the amplitude and phase estimates caused by, for example, temperature could be ruled out. This was done by comparing amplitude and phase estimates of a 3-month time series in warmer and colder periods of the groundwater temperature (one beginning at 2-Augustus and one beginning at 25-January). The amplitude and phase estimates of the time series beginning at 2-Augustus were quite similar to the estimates obtained with the whole time series of 369 days. However, the estimates obtained with the time series beginning at 25-January were quite a bit different, although they were not consistently higher or lower compared to the estimates based on the whole time series. Therefore these differences cannot fully, but possibly partly, be explained by some kind of temporal effect such as a temperature decrease. Consequently, a significant effect of the groundwater temperature on the estimates could not be ruled out with this analysis. Because the effect of temperate on the hydraulic conductivity can potentially be quite considerable, it is recommended to include or analyze the effect of the temperature if the data is at hand.

Impermeable base
The model of Bakker (2019) approximates the base as impermeable. This approximation implies that no water is released from storage when the hydraulic head declines (storativity of zero) (Maas, 1998). Because of consolidation, the storativity of the impermeable layer will often be negligible compared to the storativity of the upper aquitard (Maas, 1998).

7.2.2. Schematization of the area

One-dimensional model
The model of Bakker (2019) provides a solution for one-dimensional flow for tidal propagation, considering an arbitrary number of layers. The thickness of each layer is fixed. For the upper sand and clay layer and the lower sand layer, a fixed thickness does not seem to be a big problem. However, for the second clay layer, this is more troublesome as the thickness of this layer is varying a lot when moving land inwards (see Figure 2.2). Moreover, although not indicated in the REGIS-II model, a cross Section of the lithology (see Figure 7.1) shows that there might be some gaps in this clayey layer. This variation might be troublesome, especially for the two-aquifer model.

Determination of the shoreline (x=0)
The point where x=0, which is the left boundary of the model, is chosen, is subject to discussion as the shoreline and the edge of the Holocene layer do not exactly intersect (see Figure 2.3). Moreover, the edge of the Holocene layer is not observed but extrapolated from the information of nearby boreholes and is therefore not highly accurate. The potential error of an incorrect estimate of the distance between the shore and the observation wells is estimated to be around 50 m. Although this error can alter the results somewhat, the error is not anticipated to change the conclusions of this report.

Non-straight shore line
In Figure 2.1 the straight gray line indicates the shoreline. However, a straight shoreline is obviously a simplification because a large cove is also present next to the observation site. Carr (1971) found that results from piezometers close to tidal creeks gave inconsistent results and stresses the importance of selecting a site with an approximately straight shoreline. The error introduced by a non-straight shoreline and the present cove is unknown.

Furthermore, there are intertidal flats present in this cove (indicated by the light-blue striped surface in Figure 2.1). Both the different geometry and the intertidal flats introduce some uncertainty since the effect of these on the amplitude attenuation and phase shift, in the aquifer, is not quantified. Of course if one of these phenomena has a significant effect, then the current one-dimensional flow schematization would be less appropriate.

7.2.3. Unconsidered physical phenomena

Some 'physical phenomena' such as salt concentrations and wave runup were not considered in this report. The effects of these phenomena and the implications on the results of this report are discussed in this paragraph.

Salt concentrations

Wang et al. (2012) examined the density effect on tidal head fluctuations with a numerical model and concluded that ignoring density variation resulted in an error less than 2.5%. Although the set-up used by Wang et al. is different from this study: they used an unconfined aquifer above a confined aquifer with an impermeable layer in between, it is not expected that the error caused by density effects is much larger for this study.

Non-linear filtering effects

Non-linear filtering effects have been shown for both wave runup and an overheight of the water-table above mean sea level (Vandenbohede and Lebbe, 2007). An overheight of the mean sea level is not the case at this site, but a sloping beach is present at the site. It has been shown that a beach slope has some kind of filtering effect on the tidal signal, resulting in a steeper rise and slower descend of the groundwater level (Vandenbohede and Lebbe, 2007). To the author's knowledge, this effect was only described for unconfined aquifers (e.g. Nielsen, 1999; Robinson and Gallagher, 1999; Teo et al., 2003; Su et al., 2003). However, the steeper rise and slower descend can also be detected in the groundwater observations at this site. Unfortunately, including the beach slope makes the model non-linear, which complicates the analysis.

Sandbank

The sandbank in the cove of 'Schelphoek' falls dry during low tide. The effect of this has not been studied in this report, but it should be considered to include it in a three-dimensional model study.

7.2.4. Conclusion

Although quite a few simplifications have been made, most of them are not expected to alter the results significantly. However, a poor fit to the observations and unrealistic optimal parameter estimates indicate that the model incorrectly represents the real system, implying that some of the simplifications do significantly impact the results. The same conclusion can be drawn for the incompetence of the model to adequately simulate the M2 amplitude & phase in well B2 (see Section 4.5.2).

Especially assuming homogeneity and using lumped parameter estimates is expected to cause some model error. Since heterogeneity is known to be present (see Figure 7.1 and the maps in Appendix C.2), it is likely for the transects, between the observation wells and the Oosterschelde, to have significant differences in their hydraulic properties. Therefore, including some heterogeneity, or at least make some distinctions between the different transects, is likely to decrease part of the model error. This could be done with a 3D numerical model. With such a model, other simplifications, that are expected to have the most impact on the model results (e.g. non-straight shoreline, the cove at Schelphoek and streamline contraction near the gap in the Holocene layer), can be addressed as well. However, the use of such a complex model also has several drawbacks, such as long computation times, a more cumbersome model construction and the need for extra data. Specifically, discharge data would be required because it constraints the model much more than head data, which tends to lead to more accurate models (Hill and Tiedeman, 2006). Besides, adding discharge to the model reduces uncertainty caused by correlated parameters (Hill et al., 1998).

In short, simplifications are needed, not in the least because of a limited amount of data. Consequently, constructing a more complex model requires extra data. For more reliable parameter estimates a pumping test, possibly combined with a numerical model, is needed. However, if fewer resources are available and/or rough estimates are required, the tidal method can be considered.

7.3. Comparison of the results

In this section, first a small comparison with previous case studies that used the tidal method is presented. Second, the results from this study are compared with the resistance of the Holocene layer obtained from both the REGIS-II and GeoTOP model (Vernes et al., 2005; (Stafleu et al., 2019)).

Again note that TNO has no publicly available data on the storativity of the clay or sand layers. Moreover, the hydraulic conductivity, for which TNO has a vast amount of information, was not uniquely determined in this study (only in combination with other parameters). Therefore, only the results for the resistance of the confining layer CU can be compared with the minima of CU estimated in this study.

7.3.1. Tidal method

Previous studies often reported inconsistent parameter estimates depending on if the diffusivity (KU/S_sU) estimate was based on amplitude or phase data (Li and Jiao, 2003). Structural inconsistent estimates would suggest that the model is not able to accurately simulate the hydrological system. Since inconsistency is a recurring issue in tidal method case studies and to see how this study compares to former case studies, it will shortly be addressed here.

Trefry and Bekele (2004) reported the ratio of the diffusivity estimated with amplitude data (D_{amp}) over the diffusivity estimated with phase data (D_{phase}) for several studies (Erskine, 1991; Smith, 1999; Smith and Hick, 2001: Fakir and Razack, 2003; including their own work). They reported a D_{amp} over D_{phase} ratio ranging from 0.096 to 0.41.

The diffusivity estimates from Sections 6.1 & 6.2 were based on both the amplitude attenuation and phase lag. However, the one-aquifer model with centroid data approximately models the phase lag correctly (see Figure 6.3) while the one-aquifer model without centroid data focuses more on minimizing the residual of the amplitude data. Using this data, a ratio of 1.22 is obtained. This is a lot closer to unity, which suggests that more relevant processes are taken into account in this study compared to the above-mentioned. For a more extensive statement, a more in-depth analysis would be required, which is out of the scope of this research.

7.3.2. Resistance of the Holocene layer

TNO (The Geological Survey of the Netherlands) has hydraulically parameterized the Holocene layer in Zeeland with the GeoTOP voxel model and REGIS-II ((Stafleu et al., 2019); Vernes et al., 2005). They characterized hydrogeological units in terms of their hydraulic conductivity and for confining layers, their vertical hydraulic conductivity and thus the resistance of the confining layer (CU). These results can be compared with the minima of CU estimated in this study.

Hydraulic conductivities are often derived from pump tests, laboratory measurements or empirical relations between the conductivity and the observed grain sizes (Stafleu and Gunnink, 2016). In Stafleu and Gunnink (2016) the resistance derived from pumptests and resistances obtained with the REGIS-II and GeoTOP models are compared. Unfortunately, pumptests analyzed in this paper were not located near the study site of this report.

Both GeoTOP and REGIS-II are hydraulically parameterized for the Holocene layer, which is the dark-green layer in Figure 2.2. Some of the differences between both models are that both are based on different geological models. Another big difference is that REGIS-II is parameterized per hydrogeological unit while GeoTOP is parameterized using voxels of 100x100x0.5m (Stafleu and Gunnink, 2016).

For REGIS-II the resistance of the Holocene clay and peat layers is determined to be within the range of 200-1000 days at the study site (Vernes et al., 2010). For GeoTOP the range is roughly between 1000-2000 days (with small parts of the study site that are within 2000-5000 days). The difference between the models is mainly caused by a different definition of the confining layer. In REGIS, sandy tidal trenches, within the Holocene layer, that are often in contact with the aquifer are considered to be part of the aquifer (Stafleu and Gunnink, 2016). Besides, there are differences in the underlying geological models. If the geological model of GeoTOP is combined with the hydraulic conductivities from Regis, the resistance ranges from 2000-5000 days at the study site.

The minimum plausible value for the resistance of the Holocene layer (CU) was estimated to be 663 days. The scale associated with this estimate is approximately 3000 meters (the distance of well C2 to the Oosterschelde is 2930 meters). For both models (REGIS-II and GeoTOP) a CU bigger than 663 days is probable, although for REGIS-II part of the 68% confidence interval falls below this limit. Nonetheless, the estimated minimum of 663 days complies with other estimates in this study area.

7.4. Wider application

In principle, the tidal method can be used in aquifers adjacent to water bodies, subject to sinusoidal movement. This might be along the coast or in estuaries where there is still tidal movement. But more inland near a river might be possible as well. Wesseling and Colenbrander (1961) determined aquifer parameters in the Tielerwaard based on fluctuations in the river stage. Timmers already did analysis on an aquifer near the Yssel (IJssel) in 1955 (Timmers, 1955). In general, the tidal method can be used anywhere, if the groundwater is subject to sinusoidal movement caused by sinusoidal forcing in a nearby water body. However, to make an informed decision on the layout of the model, geohydrological information is needed as well. Detailed information on the latter is available in the Netherlands but is not common everywhere. Moreover, frequent observations of the hydraulic head are needed as well.

In a cross-section from the Netherlands, it can be remarked that the Southwestern part consists of multiple aquifers (see Figure 7.3). So in this part of the Netherlands, a one-dimensional (multi-layer) solution might be a useful model set-up for some locations. Especially if only rough estimates or ranges of the resistance of the upper clay layer and/or storage of the sand and clay layer are required.

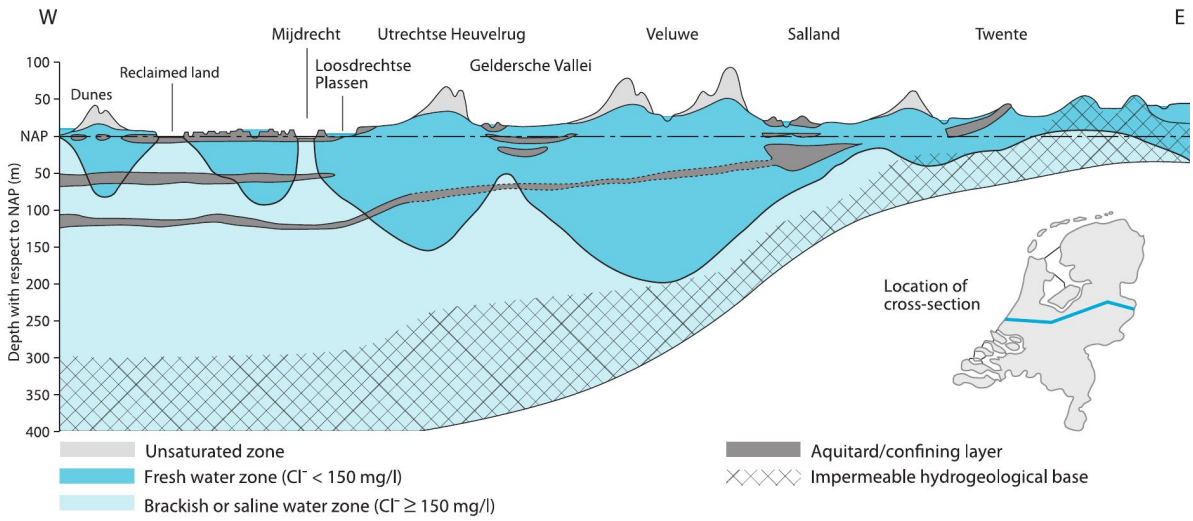


Figure 7.3: Simplified geohydrological cross-section of the Netherlands (after Van de Ven, 1986; source Wong et al., 2007).

8

Conclusion

8.1. Conclusion

The goal of this study was to estimate hydraulic aquifer parameters of the aquifer system below Schouwen-Duiveland, from tide-induced groundwater fluctuations. First, the time series of the water fluctuations were analyzed, to estimate the amplitudes and phases of the constituents present in the data. Subsequently, hydraulic parameters were estimated by optimizing a one-aquifer and a two-aquifer model, both based on the solution presented by Bakker (2019).

Noise in the data was reduced with the use of Pastas (a Python package from Collenteur et al. 2019) and a Butterworth filter. Pastas was used to decompose the fluctuations observed in the groundwater to different contributions of hydrological stresses (e.g. rain and evaporation). The use of the Butterworth filter is preferred because it produces the smallest standard deviations for the amplitude and phase estimates. Moreover, the produced residuals are more symmetrical for well B and C and have less autocorrelation. Besides, pre-processing the data with a filter is easier than decomposing the observed fluctuations for different contributions of the hydrological stresses. In addition, especially for the wells with a low signal to noise ratio, the Butterworth filter is better at extracting the tidal signal.

In addition, a graphical determination of the amplitude and phase was performed to check if this relatively quick and easy analysis gives accurate estimates as well. It was concluded that with a high signal to noise ratio and a dominantly present constituent (M2 in this case), the amplitude can be reliably estimated. Determining the phase with this method did not give satisfactory results.

A somewhat poor fit to the observations and unrealistic optimal parameter estimates indicate that the model, to some extent, incorrectly represents the real system. This also implies that some of the model simplifications do significantly impact the results. Simplifications presumed to mostly influence the model results include homogeneity, one-dimensional flow and the use of a straight shoreline. The presence of some model error also signifies that the resulting parameter estimates should be treated with care.

All hydraulic parameters (CU, σU, KU and SsU) occur coupled in the governing differential equation. However, because KU and CU were consistently estimated at their upper boundary and the storage in the leaky layer was estimated to be negligible, only one parameter group (i.e. the diffusivity) could be estimated with the optimization. The diffusivity was estimated to be 1.07E6 and 1.31E6 m^2/day for the model with and without centroid data respectively. The fit of the model was not perfect for both the model with and without the centroid data. A compromise must be sought based on what residual the model minimizes. As a result, all diffusivity estimates within and around the range [1.07E6, 1.31E6] m^2/day are all considered to be reasonable estimates.

In order to make some useful statements about CU, σU and S_sU (besides the diffusivity estimates), it was analyzed which parameter values gave unreasonable results. This resulted in a lower bound for CU (CU=663 days) and an upper bound for σU (σU=1.22E-4 m^{-1}). For S_sU reasonable results are obtained between 2.28E-6 m^{-1} and 1.87E-5 m^{-1}. For KU all values considered (3.42-14.7 m/day) resulted in a good fit.

The lower bound for the CU estimate seems to comply with estimates based on two models (REGIS-II and GeoTOP). However, the question remains how much the lower aquifer influences the above results. Finally,

the two-aquifer model was considered to be useless for parameter estimation. Presumably, this is predominantly caused by a small amount of amplitude and phase data compared to the number of parameters to be optimized.

8.2. Recommendations

8.2.1. Wells more close to the sea & on the same transect

If new groundwater wells would be constructed, it would be worth considering having (some) wells more close to the sea. Wells closer to the sea have a higher signal to noise ratio, this might allow for the analysis of more than just the dominant constituent and gives more reliable phase and amplitude estimates for the dominant constituent. Besides, having at least two wells on the same transect perpendicular to the sea should also be considered. Since results will be less disturbed by an in-homogeneous aquifer.

8.2.2. Three-dimensional groundwater model including additional stresses

As mentioned in the introduction, alternatively to just considering the tide and the associated hydraulic head response, a groundwater model that includes the effects of other stresses (e.g. rain and evapotranspiration) could be considered as well. This can be done with a 3D numerical model. An additional benefit to this is that this would allow adding more complexity to the model (e.g. add some heterogeneity), which might reduce some model error. However, the use of such a complex model also has some drawbacks, such as long computation times, a more cumbersome model construction and the need for extra data. Specifically, discharge data would be required because it constraints the model much more than head data, which in turn tends to lead to more accurate models (Hill and Tiedeman, 2006). Besides, adding discharge to the model reduces uncertainty caused by correlated parameters (Hill et al., 1998).

However, by adding recharge and discharge data to the model, the parameters can be estimated based on more data. Besides, more relevant processes are included in the model which will presumably lead to a better model and more accurate parameter estimates.

Bibliography

Bakker, M. (2016). The effect of loading efficiency on the groundwater response to water level changes in shallow lakes and streams. *Water Resources Research*, 52(3):1705–1715.

Bakker, M. (2019). Analytic solutions for tidal propagation in multilayer coastal aquifers. *Water Resources Research*, 55(4):3452–3464.

Battjes, J. A. and Labeur, R. J. (2017). *Unsteady flow in open channels*. Cambridge University Press.

Bear, J. and Verruijt, A. (1987). Theory and applications of transport in porous media. *Modeling of groundwater flow and pollution, Dordrecht: Reidel*.

Bredehoeft, J. D. (1967). Response of well-aquifer systems to earth tides. *Journal of Geophysical Research*, 72(12):3075–3087.

Carr, P. (1971). Use of harmonic analysis to study tidal fluctuations in aquifers near the sea. *Water Resources Research*, 7(3):632–643.

Carr, P. A. and Van Der Kamp, G. S. (1969). Determining aquifer characteristics by the tidal method. *Water Resources Research*, 5(5):1023–1031.

Chuang, M.-H. and Yeh, H.-D. (2007). An analytical solution for the head distribution in a tidal leaky confined aquifer extending an infinite distance under the sea. *Advances in water resources*, 30(3):439–445.

Collenteur, R. A., Bakker, M., Caljé, R., Klop, S. A., and Schaars, F. (2019). Pastas: open source software for the analysis of groundwater time series. *Groundwater*, 57(6):877–885.

Daniels, E. E. (2016). *Land surface impacts on precipitation in the Netherlands*. PhD thesis, Wageningen University.

De Bruin, H. (1987). From penman to makkink. In *Evaporation and Weather: Technical Meeting 44, Ede, The Netherlands 25 March 1987. The Hague, Netherlands. 1987. p 5-31. 1 fig, 4 tab, 34 ref.*

Deltares (2014). Delft3D-TIDE, user manual. *Deltares Delft, The Netherlands. 26 May, Version: 5.00.32824*.

Domenico, P. and Mifflin, M. (1965). Water from low-permeability sediments and land subsidence. *Water Resources Research*, 1(4):563–576.

Dong, L., Shimada, J., Kagabu, M., and Yang, H. (2015). Barometric and tidal-induced aquifer water level fluctuation near the Ariake Sea. *Environmental monitoring and assessment*, 187(1):4187.

Dufour, F. (1998). Grondwater in Nederland. *TNO, Delft, Netherlands*.

Erskine, A. (1991). The effect of tidal fluctuation on a coastal aquifer in the UK. *Groundwater*, 29(4):556–562.

Fakir, Y. and Razack, M. (2003). Hydrodynamic characterization of a Sahelian coastal aquifer using the ocean tide effect (Dridrate Aquifer, Morocco). *Hydrological sciences journal*, 48(3):441–454.

Ferris, J. G. (1952). Cyclic fluctuations of water level as a basis for determining aquifer transmissibility. Technical report, US Geological Survey.

Foreman, M. G. and Neufeld, E. (1991). Harmonic tidal analysis of long time series. *The International Hydrographic Review*, 68(1).

Forrester, W. D. (1986). Direct inference of tidal constituents, research note. *The International Hydrographic Review*.

Godin, G. (1972). *The analysis of tides*. Toronto: University of Toronto Press.

Godin, G. (1986). The use of nodal corrections in the calculation of harmonic constants. *The International Hydrographic Review*.

Gregg, D. (1966). An analysis of ground-water fluctuations caused by ocean tides in Glynn County, Georgia a. *Groundwater*, 4(3):24–32.

Guarracino, L., Carrera, J., and Vázquez-Suñé, E. (2012). Analytical study of hydraulic and mechanical effects on tide-induced head fluctuation in a coastal aquifer system that extends under the sea. *Journal of Hydrology*, 450:150–158.

Halcoussis, D. (2005). *Understanding econometrics*. South-Western.

Hill, M. C. (2006). The practical use of simplicity in developing ground water models. *Groundwater*, 44(6):775–781.

Hill, M. C., Cooley, R. L., and Pollock, D. W. (1998). A controlled experiment in ground water flow model calibration. *Groundwater*, 36(3):520–535.

Hill, M. C., D Agnese, F. A., Faunt, C. C., et al. (2000). Guidelines for model calibration and application to flow simulation in the death valley regional groundwater system. *IAHS PUBLICATION*, pages 195–204.

Hill, M. C. and Tiedeman, C. R. (2006). *Effective groundwater model calibration: with analysis of data, sensitivities, predictions, and uncertainty*. John Wiley & Sons.

Hummelman, J., Maljers, D., Menkovic, A., Reindersma, R., Vernes, R., and Stafleu, J. (2019). *Totstandkomingsrapport Hydrogeologisch model (REGIS II)*. TNO.

Hvorslev, M. J. (1951). Time lag and soil permeability in ground-water observations. *US Army Corps of Engineers, Waterways Experiment Station, Vicksburg, Mississippi, Bulletin*, 36.

Istok, J. D. and Dawson, K. J. (2014). *Aquifer testing: design and analysis of pumping and slug tests*. CRC Press.

Jacob, C. (1950). Flow of groundwater. *Engineering hydraulics*, pages 321–386.

Jonkman, S. N., Steenbergen, R. D. J. M., Morales-Nápoles, O., Vrouwenvelder, A. C. W. M., and Vrijling, J. K. (2017). Probabilistic design: Risk and reliability analysis in civil engineering. *Lecture Notes CIE4130. Delft University of Technology*.

Jumikis, A. R. (1962). *Soil mechanics*. D. Van Nostrand Book Company, New York.

Kutner, M. H., Nachtsheim, C. J., Neter, J., Li, W., et al. (2005). *Applied linear statistical models*, volume 5. McGraw-Hill Irwin New York.

Levinger, B. W. (1980). The square root of a 2×2 matrix. *Mathematics Magazine*, 53(4):222–224.

Li, H. and Jiao, J. J. (2001). Tide-induced groundwater fluctuation in a coastal leaky confined aquifer system extending under the sea. *Water Resources Research*, 37(5):1165–1171.

Li, H. and Jiao, J. J. (2003). Review of analytical studies of tidal groundwater flow in coastal aquifer systems. In *Proceedings of the International Symposium on Water Resources and the Urban Environment. Wuhan, PR China, Nov*, pages 9–10. Citeseer.

Liu, S., Li, H., Boufadel, M. C., and Li, G. (2008). Numerical simulation of the effect of the sloping submarine outlet-capping on tidal groundwater head fluctuation in confined coastal aquifers. *Journal of hydrology*, 361(3-4):339–348.

Maas, K. (1998). Over grondwatergetijden (dutch for: On groundwater tides). *Stromingen*, 4(4):33–46.

Malama, B., Revil, A., and Kuhlman, K. L. (2009). A semi-analytical solution for transient streaming potentials associated with confined aquifer pumping tests. *Geophysical Journal International*, 176(3):1007–1016.

Merritt, M. L. (2004). *Estimating hydraulic properties of the Floridan aquifer system by analysis of earth-tide, ocean-tide, and barometric effects, Collier and Hendry Counties, Florida*. US Department of the Interior, US Geological Survey.

Meyer, F. W. (1974). Evaluation of hydraulic characteristics of a deep artesian aquifer from natural water-level fluctuations, Miami, Florida. Technical report, Florida Geological Survey.

Neuman, S. P. and Witherspoon, P. A. (1969). Theory of flow in a confined two aquifer system. *Water Resources Research*, 5(4):803–816.

Newville, M., Stensitzki, T., Allen, D. B., and Ingargiola, A. (2014). LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python.

Nielsen, P. (1999). Groundwater dynamics and salinity in coastal barriers. *Journal of Coastal Research*, pages 732–740.

Obergfell, C., Bakker, M., and Maas, K. (2016). A time-series analysis framework for the flood-wave method to estimate groundwater model parameters. *Hydrogeology Journal*, 24(7):1807–1819.

OhioEPA (2006). *Technical guidance manual for ground water investigations. Chapter 4: Pumping and Slug test, Revision 1*. Ohio Environmental Protection Agency Division of Drinking and Ground Waters.

Pawlowicz, R., Beardsley, B., and Lentz, S. (2002). Classical tidal harmonic analysis including error estimates in matlab using t_tide. *Computers & Geosciences*, 28(8):929–937.

Polman, N., Peerlings, J., and van der Vat, M. (2019). Economische effecten van droogte voor landbouw in nederland: samenvatting. Technical report, Wageningen Economic Research.

Rijkswaterstaat (2013). *Getijtafels voor Nederland 2014*. Sdu Uitgevers.

Ritzema, H., Heuvelink, G., Heinen, M., Bogaart, P., van der Bolt, F., Hack-ten Broeke, M., Hoogland, T., Knotters, M., Massop, H., and Vroon, H. (2012). Meten en interpreteren van grondwaterstanden: analyse van methodieken en nauwkeurigheid. Technical report, Alterra.

Roberts, J. and Roberts, T. D. (1978). Use of the butterworth low-pass filter for oceanographic data. *Journal of Geophysical Research: Oceans*, 83(C11):5510–5514.

Robinson, M. A. and Gallagher, D. L. (1999). A model of ground water discharge from an unconfined coastal aquifer. *Ground Water*, 37(1):80.

Rotzoll, K., Gingerich, S. B., Jenson, J. W., and El-Kadi, A. I. (2013). Estimating hydraulic properties from tidal attenuation in the northern Guam lens aquifer, territory of Guam, USA. *Hydrogeology Journal*, 21(3):643–654.

Shih, D. C.-F., Lee, C.-D., Chiou, K.-F., and Tsai, S.-M. (2000). Spectral analysis of tidal fluctuations in ground water level 1. *Journal of the American Water Resources Association*, 36(5):1087–1099.

Smith, A. J. (1999). Application of a tidal method for estimating aquifer diffusivity: Swan river, western australia.

Smith, A. J. and Hick, W. P. (2001). *Hydrogeology and aquifer tidal propagation in Cockburn Sound, Western Australia*. CSIRO Land and Water Australia.

Spane, F. A. (2002). Considering barometric pressure in groundwater flow investigations. *Water resources research*, 38(6):14–1.

Stafleu, J. and Gunnink, J. (2016). *Hydraulische parameterisering van GeoTOP Zeeland*. TNO.

Stafleu, J., Maljers, D., Busschers, F., Gunnink, J., Schokker, J., and Hummelman, J. (2019). *Totstandkomingsrapport GeoTOP*. TNO.

Su, N., Liu, F., and Anh, V. (2003). Tides as phase-modulated waves inducing periodic groundwater flow in coastal aquifers overlaying a sloping impervious base. *Environmental Modelling & Software*, 18(10):937–942.

Teo, H., Jeng, D., Seymour, B., Barry, D., and Li, L. (2003). A new analytical solution for water table fluctuations in coastal aquifers with sloping beaches. *Advances in Water Resources*, 26(12):1239–1247.

Thomopoulos, N. T. (2012). *Essentials of Monte Carlo simulation: Statistical methods for building simulation models*. Springer Science & Business Media.

Timmers, H. (1955). Determination of soil permeability in situ. *NJAS wageningen journal of life sciences*, 3(2):119–126.

Trefry, M. and Bekele, E. (2004). Structural characterization of an island aquifer via tidal methods. *Water Resources Research*, 40(1).

Van de Ven, G.P., e. a. (1986). *Water, Atlas van Nederland, deel 15*. Staatsuitgeverij, 's-Gravenhage.

Van der Kamp, G. (1972). Tidal fluctuations in a confined aquifer extending under the sea. In *International Geological Congress*, volume 24, pages 101–106. Quebec Montreal.

Van Der Kamp, G. and Gale, J. E. (1983). Theory of earth tide and barometric effects in porous formations with compressible grains. *Water Resources Research*, 19(2):538–544.

Van der Kamp, G. S. (1973). *Periodic flow of groundwater*. Editions Rodopi NV Amsterdam.

Vandenbohede, A. and Lebbe, L. (2007). Effects of tides on a sloping shore: groundwater dynamics and propagation of the tidal wave. *Hydrogeology Journal*, 15(4):645–658.

Vernes, R. W., Hummelman, H. J., and Menkovic, A. (2010). *REGIS Zeeland, Deelrapport B: Hydrogeologische opbouw en hydraulische eigenschappen van Holocene afzettingen*. TNO.

Vernes, R. W., Van Doorn, T. H. M., Bierkens, M. F. P., Van Gessel, S. F., and De Heer, E. (2005). *Van gidslaag naar hydrogeologische eenheid: toelichting op de totstandkoming van de dataset REGIS II*. TNO.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., Vand erPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.

Wang, X., Li, H., Wan, L., Liu, F., and Jiang, X. (2012). Loading effect of water table variation and density effect on tidal head fluctuations in a coastal aquifer system. *Water Resources Research*, 48(9).

Wesseling, J. (1960). Enkele resultaten van de bepaling van hydrologische constanten in het Prunje gebied (Schouwen Duiveland). *Geologie en Mijnbouw*, 39(11):624–630.

Wesseling, J. and Colenbrander, H. J. (1961). De bepaling van de hydrologische bodemconstanten uit de voortplanting van de getijbeweging; in: De waterbehoefte van de Tielerwaard-West. *Commissie ter bestudering van de waterbehoefte van de Gelderse Landbouwgronden*, pages 60–65.

Witte, R. A. (1993). Anti-alias filtering apparatus for frequency domain measurements. US Patent 5,233,546.

Wong, T. E., Batjes, D. A., and de Jager, J. (2007). *Geology of the Netherlands*. Edita Publishing House of the Royal.

Zhou, P., Li, G., Lu, Y., and Li, M. (2014). Numerical modeling of the effects of beach slope on water-table fluctuation in the unconfined aquifer of Donghai Island, China. *Hydrogeology Journal*, 22(2):383–396.

Zhou, X. (2008). Determination of aquifer parameters based on measurements of tidal effects on a coastal aquifer near Beihai, China. *Hydrological Processes*, 22(16):3176–3180.

zwdelta (2020). Het getij terugbrengen in de grevelingen. `https://www.zwdelta.nl/projecten/getij-grevelingen`. Accessed: 2020-06-08.

A

Derivations

A.1. Astronomical coupling

The times series in well A4 is too short (134 days) to satisfy the Rayleighs resolution limit for all of the 31 constituents used (see also Section 3.1.1). The best solution for this problem, according to Pawlowicz et al. (2002), is to apply inference (also called astronomical coupling). This method involves prescribing a relation between the known and inferred constituent for the amplitude and phase. This allows for a 'lumped' estimation of both constituents amplitude and phase (Deltares, 2014). The couplings that need to be made for for a record with the length of A4 are: (K1,P1), (MU2,NLK2), (N2,NU2), (M2,MPS2), (2MN2,λ2), (S2,K2), (MS4,MK4).

Below a solution to make the astronomical coupling an integral part of the least squares solution, partly after the derivation in Forrester (1986). Starting with the same equation (Equation 3.1) as in Section 3.1:

$$y_i = A_0 + \sum_{j=1}^{M} A_j \cos(\omega_j t_i - \Phi_j), \tag{A.1}$$

where A_0 is the average waterlevel; A_j, $\omega_j and\ \Phi_j$ are the constituents frequency, amplitude and phase respectively for constituent j. y_i and t_i are the observation and time respectively at time i.

This formula can be transformed with some trigonometry to:

$$y_i = A_0 + \sum_{j=1}^{M} A_j \cos(\omega_j t_i)\ \cos(\Phi_j) + A_j \sin(\omega_j t_i)\ \sin(\Phi_j). \tag{A.2}$$

The kth constituent can be inferred from the pth constituent with the following relationships:

$$R_{pk} = A_k / A_p \quad \text{and} \quad G_{pk} = \Phi_p - \Phi_k. \tag{A.3}$$

Subsequently, the combined solution of the dependent and independent component can be rewritten as:

$$A_p \cos(\Phi_p)\ [\cos(\omega_p t) + R_{pk}\ \cos(\omega_k t + G_{pk})] \quad + \quad A_p \sin(\Phi_p)\ [\sin(\omega_p t + R_{pk}\ \sin(\omega_k t + G_{pk})]. \tag{A.4}$$

If we than substitute part of Equation A.4 with:

$$F1 = R_{pk}\ \cos(\omega_k t + G_{pk}) \quad \text{and} \quad F2 = R_{pk}\ \sin(\omega_k t + G_{pk}). \tag{A.5}$$

,

we can rewrite Equation A.4 to

$$C_p\ [\cos(\omega_p t)\ +\ F1]\ +\ S_p\ [\sin(\omega_p t)\ +\ F2]. \tag{A.6}$$

Where C_p and S_p are defined as

$$C_p = A_p \cos(\Phi_p) \quad \text{and} \quad S_p = A_p, \sin(\Phi_p) \tag{A.7}$$

and are both to be solved with the least squared optimization. Note that, apart from the subscripts and variables F1 & F2, Equation A.6 is similar to the right-hand side of Equation 3.2 from Section 3.1. Which is, for convenience, also shown below:

$$A_j \cos(\omega_j t_i - \Phi_j) = C_j \cos(\omega_j t_i) + S_j \sin(\omega_j t_i). \tag{A.8}$$

Lastly note that the inferred constituent must be smaller than the 'main' component.

A.2. Rewriting the (complex) solution derived in Section 5.4.1

The objective of this appendix section is to rewrite the one-layer solution (Eq. 5.12; which is also stated below, see Equation A.9), derived in Section 5.4.1, in terms of the hydraulic parameters. This is done to gain more insight into how the hydraulic parameters determine the amplitude attenuation and phase shift in the aquifer. The hydraulic parameters in the solution are the transmissivity (T) and storage coefficient (S) of the aquifer; and the resistance to vertical flow (c) and the storage coefficient (σ) of the leaky layer.

The boundary conditions (Eq. 5.5) imposed on the governing equation (Eq. 5.1) yielded the following (complex) solution:

$$h = h_{sea} e^{-x\sqrt{\alpha}} e^{i\omega_j t}, \tag{A.9}$$

where i is the imaginary unit and α is defined as

$$\alpha = \frac{g + i\omega_j S}{T}, \tag{A.10}$$

where ω_j is the frequency of constituent j and where g is defined as

$$g = \frac{\sqrt{i\omega_j \sigma c}}{c \tanh(\sqrt{i\omega_j \sigma c})}. \tag{A.11}$$

Note that, if the storage is neglected, the coupled parameters can already be distinguished by rewriting Equation A.10, see Box A.1.

The square root of the complex number α in Equation A.9, may be written as

$$\sqrt{\alpha} = \sqrt{r}\left[\cos\left(\frac{\phi}{2}\right) + i \sin\left(\frac{\phi}{2}\right)\right], \tag{A.12}$$

where r and ϕ are the polar coordinates of α, where ϕ is the angle and r is the modulus or absolute value of the complex number α. In equation form:

$$|\alpha| = \text{mod}\ \alpha = r = \sqrt{c^2 + d^2}, \tag{A.13}$$

$$\phi = \arctan(d/c) \tag{A.14}$$

where c is the real part of complex number α and d the imaginary part. e^{α} can be written as

$$e^{\alpha} = e^{c+id} = e^c e^{id} = e^c \left[\cos(d) + i \sin(d)\right]. \tag{A.15}$$

Consequently, by substituting Eq. A.12 into Eq. A.15, Eq. A.9 can be rewritten as

$$h = h_{sea} e^{-x\sqrt{r}\cos(\phi/2)} \left\{\cos\left[-x\sqrt{r}\sin\left(\frac{\phi}{2}\right)\right] \sin\left[-x\sqrt{r}\sin\left(\frac{\phi}{2}\right)\right]\right\} e^{i\omega_j t}. \tag{A.16}$$

The head in the aquifer can be expressed in terms of an amplitude and a phase shift. For this purpose, Equation A.16 is written as

$$h = (u + iv) e^{i\omega_j t} = \rho e^{i\omega_j (t-\theta)}, \tag{A.17}$$

where ρ is the amplitude and θ is the phase of the tidal fluctuation at x

$$\rho = \sqrt{u^2 + v^2}, \tag{A.18}$$

$$\theta = -\frac{1}{\omega_j} \arctan(v/u). \tag{A.19}$$

and where u & v are given as

$$\begin{aligned} u &= h_{sea} e^{-x\sqrt{r}\cos(\phi/2)} \cos\left[-x\sqrt{r}\sin\left(\frac{\phi}{2}\right)\right], \\ v &= h_{sea} e^{-x\sqrt{r}\cos(\phi/2)} \sin\left[-x\sqrt{r}\sin\left(\frac{\phi}{2}\right)\right]. \end{aligned} \tag{A.20}$$

By using the 'Pythagorean' identity: $\sin^2(x) + \cos^2(x) = 1$; the amplitude ρ (Eq. A.18) may be written as

$$\rho = \sqrt{h_{sea}^2 \exp\left[-2x(c^2+d^2)^{1/4}\cos\left(\frac{1}{2}\arctan\frac{d}{c}\right)\right]}. \tag{A.21}$$

Subsequently, by using simple trigonometry (i.e. $\tan(\phi) = \sin(\phi) \,/\, \cos(\phi)$) and by substitution of both Equation A.13 an Equation A.14, θ (Eq. A.19) may now be written as

$$\theta = -\frac{x}{\omega_j}(c^2+d^2)^{1/4}\sin\left[\frac{1}{2}\arctan\left(\frac{d}{c}\right)\right]. \tag{A.22}$$

c and d are the real and imaginary part of α (Eq. A.10), hence ρ can also be written as

$$\rho = \sqrt{h_{sea}^2 \exp\left\{-2x\left[\underbrace{\left(\frac{\Re g}{T}\right)^2}_{\text{term 1}} + \underbrace{\left(\frac{\omega_j S + \Im g}{T}\right)^2}_{\text{term 2}}\right]^{1/4} \cos\left[\frac{1}{2}\arctan\left(\underbrace{\frac{\omega_j S + \Im g}{\Re g}}_{\text{term 3}}\right)\right]\right\}}, \tag{A.23}$$

where $\Re g$ and $\Im g$ are the real and imaginary part respectively of Equation A.11. θ may be written as

$$\theta = \frac{x}{\omega_j}\left[\underbrace{\left(\frac{\Re g}{T}\right)^2}_{\text{term 1}} + \underbrace{\left(\frac{\omega_j S + \Im g}{T}\right)^2}_{\text{term 2}}\right]^{1/4} \sin\left[\frac{1}{2}\arctan\left(\underbrace{\frac{\omega_j S + \Im g}{\Re g}}_{\text{term 3}}\right)\right]. \tag{A.24}$$

Note that, terms 1,2 and 3 in Equations A.23 & A.24 are written in terms of the four model parameters (i.e. T, S, c and σ), where g (Eq. A.11) is a function of the resistance (c) and the storage coefficient (σ) of the leaky layer. The former was the objective of this appendix section.

Box A.1: Coupled parameters

Note that when the storage in the leaky layer is neglected ($\sigma = 0$), g (Eq. A.11) is equal to $1/c$. Subsequently, α may be written as

$$\alpha = \frac{1}{cT} + \frac{i\omega_j S}{T}. \tag{A.25}$$

Notice that parameter $c \times T$ determines the real part of α and S/T the imaginary part. Thus both S and c are linked with T, therefore when storage in the leaky layer is neglected, all parameters (S, c & T) are linked. Therefore, the parameters cannot be estimated uniquely when this configurations is used. The same conclusion was derived in Section 5.4.1, by rewriting the governing differential equation.

B

Tables

B.1. Amplitudes and phases of the main constituents

B.1.1. Oosterschelde

Table B.1: Estimated amplitudes and phases of the main constituents in the Oosterschelde

	ω (rad/s)	Period (days)	Amplitude (cm)	Standard Error Amplitude	Phase (rad)	Standard Error Phase	Unreliable amplitude & phase estimate
SA	1.988e-07	365.853	7.67	0.27%	1.44	0.19%	
SM	4.926e-06	14.764	4.99	0.41%	-1.51	0.27%	
Q1	6.496e-05	1.12	2.78	0.74%	0.34	2.19%	
O1	6.76e-05	1.076	9.4	0.22%	-3.14	0.07%	
P1	7.252e-05	1.003	3.61	0.57%	-2.29	0.25%	
K1	7.292e-05	0.997	6.19	0.33%	2.98	0.11%	
3MS2	1.307e-04	0.557	3.45	0.60%	-0.59	1.01%	
MNS2	1.33e-04	0.547	2.61	0.79%	-0.79	1.00%	
NLK2	1.352e-04	0.538	3.92	0.52%	-0.58	0.91%	
μ2	1.356e-04	0.536	12.71	0.16%	1.42	0.11%	
N2	1.379e-04	0.527	19.5	0.11%	1.49	0.07%	
NU2	1.382e-04	0.526	7.32	0.28%	0.9	0.32%	
MPS2	1.403e-04	0.518	1.72	1.20%	-0.68	1.77%	
M2	1.405e-04	0.518	124.83	0.02%	-2.53	0.01%	
λ2	1.428e-04	0.509	4.78	0.43%	-3.1	0.14%	
2MN2	1.432e-04	0.508	11.1	0.19%	2.72	0.07%	
S2	1.454e-04	0.5	30.19	0.07%	2.64	0.03%	
K2	1.458e-04	0.499	7.94	0.26%	-1.63	0.16%	
2SM2	1.504e-04	0.484	3.23	0.64%	-1.86	0.34%	
2MK3	2.081e-04	0.349	1.44	1.43%	-0.85	1.67%	
MK3	2.134e-04	0.341	1.08	1.88%	-1.44	1.32%	
3MS4	2.761e-04	0.263	0.6	3.42%	-1.36	2.52%	
MN4	2.784e-04	0.261	1.65	1.24%	-1.08	1.16%	✗
M4	2.810e-04	0.259	4.59	0.45%	1.21	0.37%	✗
MS4	2.86e-04	0.254	2.98	0.69%	-0.06	11.72%	✗
MK4	2.864e-04	0.254	0.76	2.70%	2.01	1.35%	✗
2MN6	4.189e-04	0.174	1.76	1.17%	0.51	3.12%	✗
M6	4.216e-04	0.173	3.24	0.63%	2.9	0.30%	✗
2MS6	4.265e-04	0.171	2.78	0.74%	1.51	0.66%	✗
M8	5.621e-04	0.129	0.99	2.09%	-1.18	1.75%	
3MS8	5.67e-04	0.128	1.17	1.75%	-2.53	0.69%	

The phase is calculated back to the phase at 02-aug-2018 00:00.
All 31 components considered to be most important by Rijkswaterstaat (2013) are considered.
It is important to note that, because the times series is autocorrelated, the standard deviations estimated with the least squares model are smaller than the true values. Lastly, ✗ marks the rows with unreliable data (see Section 4.1.1).

B.1.2. Observation well A2

Table B.2: Amplitude and phases of the main constituents in the Oosterschelde

	ω (rad/s)	Period (days)	Amplitude (cm)	Standard deviation Amplitude	Phase (rad)	Standard deviation Phase
Q1	6.496e-05	1.12	1.01	4.18%	1.13	3.72%
O1	6.76e-05	1.076	2.85	1.47%	-2.49	0.59%
P1	7.252e-05	1.003	1.04	4.02%	-1.52	2.67%
K1	7.292e-05	0.997	1.67	2.52%	-2.54	0.99%
3MS2	1.307e-04	0.557	1.04	4.04%	0.26	15.41%
MNS2	1.33e-04	0.547	0.99	4.24%	-0.21	20.69%
NLK2	1.352e-04	0.538	1.71	2.48%	0.21	11.77%
μ2	1.356e-04	0.536	3.65	1.15%	2.39	0.48%
N2	1.379e-04	0.527	5.68	0.74%	2.35	0.32%
NU2	1.382e-04	0.526	2.2	1.92%	1.88	1.03%
MPS2	1.403e-04	0.518	0.33	12.82%	0.14	92.89%
M2	1.405e-04	0.518	37.41	0.11%	-1.62	0.07%
λ2	1.428e-04	0.509	1.55	2.72%	-2.38	1.15%
2MN2	1.432e-04	0.508	3.48	1.21%	-2.68	0.45%
S2	1.454e-04	0.5	9.01	0.47%	-2.75	0.17%
K2	1.458e-04	0.499	2.39	1.76%	-0.58	3.03%
2SM2	1.504e-04	0.484	0.84	5.03%	-0.8	6.26%
2MK3	2.081e-04	0.349	0.88	4.77%	0.91	5.31%
MK3	2.134e-04	0.341	0.62	6.83%	0.81	8.43%
3MS4	2.761e-04	0.263	1.06	3.97%	-0.21	19.04%
MN4	2.784e-04	0.261	1.61	2.62%	-0.21	12.28%
M4	2.810e-04	0.259	5.26	0.80%	2.12	0.38%
MS4	2.86e-04	0.254	2.76	1.53%	0.99	1.53%
MK4	2.864e-04	0.254	0.72	5.84%	3.05	1.92%
2MN6	4.189e-04	0.174	0.33	12.59%	2.12	6.09%
M6	4.216e-04	0.173	0.7	6.02%	-1.81	3.33%
2MS6	4.265e-04	0.171	0.66	6.35%	-3.1	2.07%
M8	5.621e-04	0.129	0.42	9.89%	1.34	7.50%
3MS8	5.67e-04	0.128	0.55	7.59%	-0.03	240.66%

The phase is calculated back to the phase at 02-aug-2018 00:00.
29 components considered to be most important by Rijkswaterstaat (2013) are considered. with the exception of SA & SM.
It is important to note that. because the times series is autocorrelated. the standard deviations estimated with the least squares model are smaller than the true values.

B.1.3. Observation well B2

Table B.3: Amplitude and phases of the main constituents in the Oosterschelde

	ω (rad/s)	Period (days)	Amplitude (cm)	Standard deviation Amplitude	Phase (rad)	Standard deviation Phase	Unreliable amplitude & phase estimate
Q1	6.496e-05	1.12	0.07	7.14%	2.05	3.49%	
O1	6.76e-05	1.076	0.19	2.77%	-1.57	1.78%	
P1	7.252e-05	1.003	0.07	7.16%	-0.37	19.73%	
K1	7.292e-05	0.997	0.11	4.93%	-1.37	3.62%	
3MS2	1.307e-04	0.557	0.04	14.52%	0.14	110.00%	
MNS2	1.33e-04	0.547	0.03	16.85%	-0.39	45.32%	
NLK2	1.352e-04	0.538	0.04	12.21%	-0.08	149.80%	
μ2	1.356e-04	0.536	0.13	3.98%	2.07	1.92%	
N2	1.379e-04	0.527	0.21	2.53%	2.11	1.20%	
NU2	1.382e-04	0.526	0.08	6.62%	1.53	4.36%	
MPS2	1.403e-04	0.518	0.02	31.91%	-0.14	291.83%	
M2	1.405e-04	0.518	1.44	0.37%	-1.89	0.20%	
λ2	1.428e-04	0.509	0.07	8.18%	-2.47	3.35%	
2MN2	1.432e-04	0.508	0.14	3.94%	-2.94	1.34%	
S2	1.454e-04	0.5	0.4	1.33%	-3.02	0.44%	
K2	1.458e-04	0.499	0.1	5.41%	-0.95	5.76%	
2SM2	1.504e-04	0.484	0.03	16.35%	-1.32	12.85%	
2MK3	2.081e-04	0.349	0.03	19.61%	0.32	65.00%	
MK3	2.134e-04	0.341	0.01	40.51%	0.14	446.90%	✗
3MS4	2.761e-04	0.263	0.04	13.70%	-0.06	248.84%	
MN4	2.784e-04	0.261	0.07	7.91%	-0.1	77.14%	
M4	2.810e-04	0.259	0.23	2.36%	2.19	1.07%	
MS4	2.86e-04	0.254	0.12	4.40%	0.97	4.51%	
MK4	2.864e-04	0.254	0.03	15.50%	3.12	5.14%	
2MN6	4.189e-04	0.174	0.03	20.77%	2.05	10.84%	
M6	4.216e-04	0.173	0.05	10.31%	-1.89	5.50%	
2MS6	4.265e-04	0.171	0.05	11.67%	2.98	3.99%	
M8	5.621e-04	0.129	0.02	22.95%	0.79	31.56%	
3MS8	5.67e-04	0.128	0.03	19.90%	-0.55	38.52%	

The phase is calculated back to the phase at 02-aug-2018 00:00.
29 components considered to be most important by Rijkswaterstaat (2013) are considered. with the exception of SA & SM.
It is important to note that. because the times series is autocorrelated. the standard deviations estimated with the least squares model are smaller than the true values. Lastly, ✗ marks the rows with the unreliable data.

B.1.4. Observation well C2

Table B.4: Amplitude and phases of the main constituents in the Oosterschelde

	ω (rad/s)	Period (days)	Amplitude (cm)	Standard deviation Amplitude	Phase (rad)	Standard deviation Phase	Unreliable amplitude & phase estimate
Q1	6.496e-05	1.12	0.05	14.10%	1.29	11.22%	
O1	6.76e-05	1.076	0.02	41.86%	1.66	56.63%	✗
P1	7.252e-05	1.003	0.22	3.52%	2.61	1.34%	
K1	7.292e-05	0.997	0.24	3.26%	1.86	1.76%	
3MS2	1.307e-04	0.557	0.05	16.49%	0.62	27.92%	
MNS2	1.33e-04	0.547	0.02	41.73%	-0.44	150.99%	✗
NLK2	1.352e-04	0.538	0.01	45.64%	-0.21	386.21%	✗
μ2	1.356e-04	0.536	0.04	19.19%	2.07	9.79%	
N2	1.379e-04	0.527	0.05	15.99%	1.75	9.47%	
NU2	1.382e-04	0.526	0.02	35.01%	1.04	44.00%	
MPS2	1.403e-04	0.518	-	-	-	-	
M2	1.405e-04	0.518	0.31	2.44%	-2.06	1.19%	
λ2	1.428e-04	0.509	0.03	25.95%	-3.09	9.39%	
2MN2	1.432e-04	0.508	0.02	32.48%	3.07	13.17%	
S2	1.454e-04	0.5	0.05	14.45%	1.22	12.24%	
K2	1.458e-04	0.499	0.02	40.74%	-1.51	48.82%	✗
2SM2	1.504e-04	0.484	0.02	40.29%	-0.42	148.28%	✗
2MK3	2.081e-04	0.349	0.01	50.36%	0.94	138.52%	✗
MK3	2.134e-04	0.341	0.01	49.86%	-0.44	538.19%	✗
3MS4	2.761e-04	0.263	0.01	51.12%	0.55	390.17%	✗
MN4	2.784e-04	0.261	0.02	39.88%	-0.5	121.27%	✗
M4	2.810e-04	0.259	0.04	20.24%	1.87	11.45%	
MS4	2.86e-04	0.254	0.03	30.19%	0.79	44.69%	
MK4	2.864e-04	0.254	0.01	51.50%	0.27	810.20%	✗
2MN6	4.189e-04	0.174	0.01	51.97%	0.47	404.84%	✗
M6	4.216e-04	0.173	0.01	48.52%	-1.28	137.87%	✗
2MS6	4.265e-04	0.171	0.01	52.02%	0.27	757.35%	✗
M8	5.621e-04	0.129	0.01	52.28%	0.01	18628.24%	✗
3MS8	5.67e-04	0.128	0.01	51.03%	-0.52	250.04%	✗

The phase is calculated back to the phase at 02-aug-2018 00:00.
29 components considered to be most important by Rijkswaterstaat (2013) are considered. with the exception of SA & SM.
It is important to note that. because the times series is autocorrelated. the standard deviations estimated with the least squares model are smaller than the true values. Lastly, ✗ marks the rows with the unreliable data.

B.1.5. Observation well A4

Table B.5: Amplitude and phases of the main constituents in the Oosterschelde

	ω (rad/s)	Period (days)	Amplitude (cm)	Standard deviation Amplitude	Phase (rad)	Standard deviation Phase	Coupled Constituents
Q1	6.496e-05	1.12	1.13	6.20%	0.97	6.46%	
O1	6.76e-05	1.076	3.46	2.00%	-2.47	0.82%	
P1	7.252e-05	1.003	1.14	3.84%	-1.57	2.43%	✓
K1	7.292e-05	0.997	1.82	3.84%	-2.6	1.47%	✓
3MS2	1.307e-04	0.557	1.63	4.28%	0.47	9.11%	
MNS2	1.33e-04	0.547	1.04	6.72%	-0.42	15.99%	
NLK2	1.352e-04	0.538	1.87	1.43%	0.21	6.69%	✗
μ**2**	1.356e-04	0.536	4	1.43%	2.4	0.59%	✗
N2	1.379e-04	0.527	5.72	1.31%	2.29	0.57%	✓
NU2	1.382e-04	0.526	2.21	1.31%	1.81	0.72%	✓
MPS2	1.403e-04	0.518	0.36	0.17%	0.12	1.47%	✗
M2	1.405e-04	0.518	40.76	0.17%	-1.64	0.11%	✗
λ**2**	1.428e-04	0.509	1.43	2.27%	-2.4	0.94%	✓
2MN2	1.432e-04	0.508	3.19	2.27%	-2.71	0.84%	✓
S2	1.454e-04	0.5	9.58	0.66%	-2.8	0.23%	✗
K2	1.458e-04	0.499	2.54	0.66%	-0.63	1.03%	✗
2SM2	1.504e-04	0.484	1.11	6.28%	-0.85	7.45%	
2MK3	2.081e-04	0.349	0.85	8.20%	0.91	9.16%	
MK3	2.134e-04	0.341	0.56	12.31%	0.82	15.19%	
3MS4	2.761e-04	0.263	1.15	6.03%	-0.06	107.61%	
MN4	2.784e-04	0.261	1.22	5.70%	-0.33	17.21%	
M4	2.810e-04	0.259	4.8	1.46%	2.12	0.69%	
MS4	2.86e-04	0.254	2.64	2.36%	0.99	2.39%	✓
MK4	2.864e-04	0.254	0.69	2.36%	3.06	0.78%	✓
2MN6	4.189e-04	0.174	0.29	23.99%	1.78	14.68%	
M6	4.216e-04	0.173	0.7	9.94%	-1.92	5.27%	
2MS6	4.265e-04	0.171	0.74	9.34%	3.03	3.13%	
M8	5.621e-04	0.129	0.37	18.64%	1.21	16.26%	
3MS8	5.67e-04	0.128	0.56	12.56%	-0.18	69.46%	

The phase is calculated back to the phase at 02-aug-2018 00:00.
29 components considered to be most important by Rijkswaterstaat (2013) are considered. with the exception of SA & SM.
It is important to note that. because the times series is autocorrelated. the standard deviations estimated with the least squares model are smaller than the true values. Lastly, coupled constituents are indicated alternately with a ✓ and a ✗.

B.2. Table with the centroid data

See Section 4.5.2 for more information.

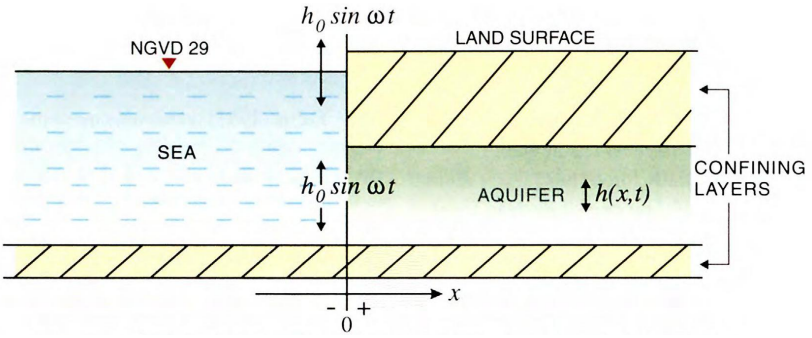
Table B.6

Centroid names	Average frequency ω (rad/s)	A2 Average phase shift (rad)	A4 Average phase shift (rad)	Used for the inverse modeling
Diurnal	6.950e-05	0.743	0.679	✓
1.4e-4	1.399e-04	0.874	0.845	
2.8e-4	2.816e-04	1.004	1.012	
4.2e-4	4.223e-04	1.619	1.421	
High frequency	5.645e-04	2.511	2.369	✓

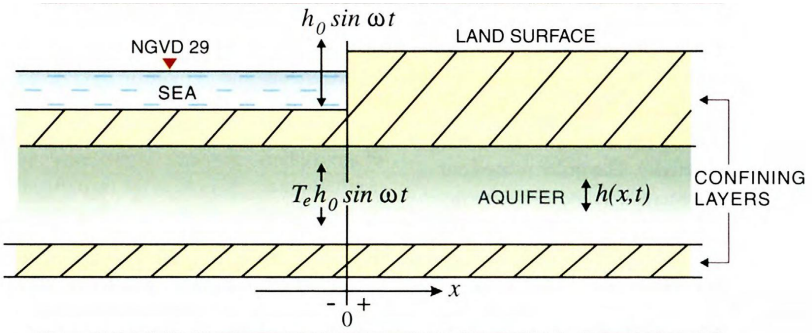
C

Figures

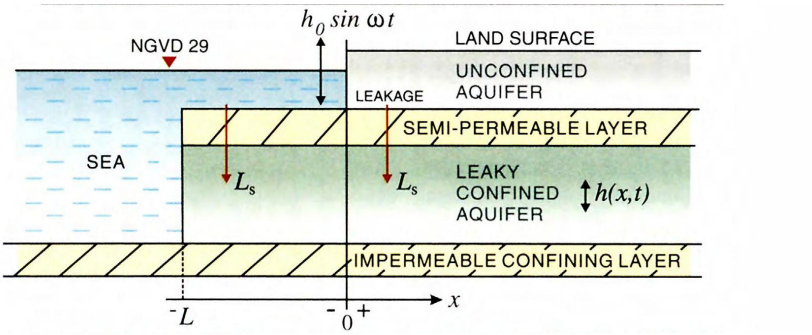
C.1. Schematic drawings of sea-aquifer models from previous studies



(a) Model of Jacob (1950) and Ferris (1952) - aquifer and overlying confining layer crop out at the coastline.



(b) Model of Van der Kamp (1972) - a completely confined aquifer extending under the sea.



(c) Model of Li and Jiao (2001) - a leaky contained aquifer system extending under the sea for a distance L.

Figure C.1: Schematic views of different sea-aquifer models (copied from: Merritt, 2004)

C.2. Maps of the geohydrology and aquifer thickness at the study site

C.2.1. Upper aquifer

This Appendix-section contains maps of the 2.5th and 97.5th percentiles (Figures C.2 & C.3) of the hydraulic conductivity, from the upper aquifer, at the study site. In addition, a map of the aquifer thickness is presented (Figure C.4). Note that the upper aquifer is located between the complex (Holocene) layer and a thin clay lens (i.e. a clayey unit from the Maassluis formation; see Figure 2.2).

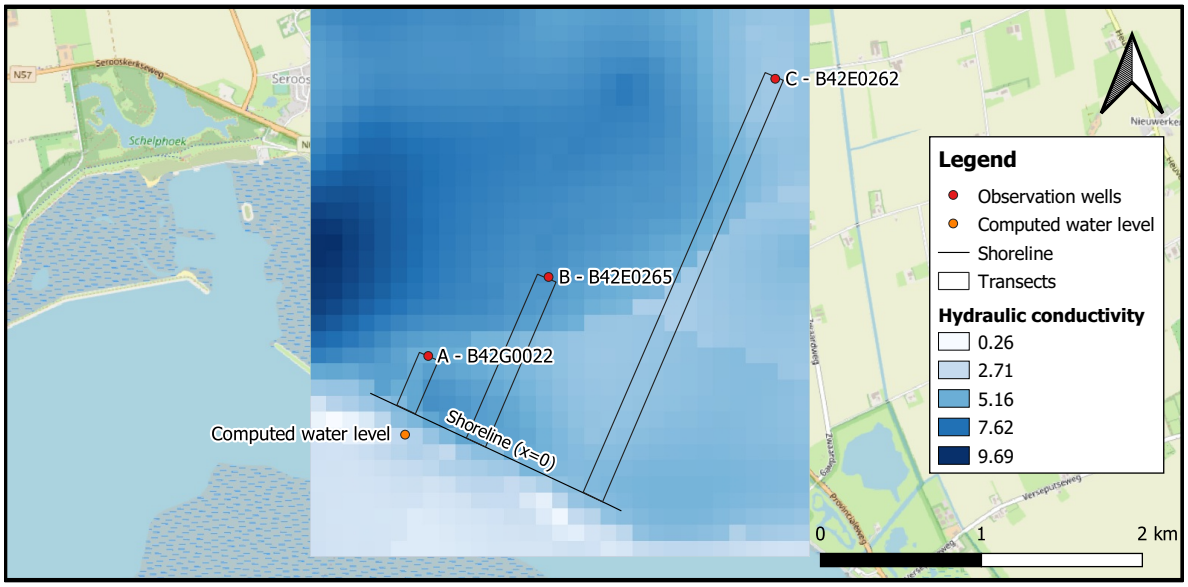


Figure C.2: A map of the 2.5th percentile of the hydraulic conductivity (m/day), from the upper aquifer, at the study site. The map is based on the statistics from the Regis-II model (Vernes et al., 2005). The 3 polygons, called transects in the legend, were used to calculate some basic statistics for each transect, but more on that in Section 5.5.3. Lastly note that the thickness of the aquifer associated with the presented hydraulic conductivity is constant (20.3 m) instead of varying (see Figure C.4).

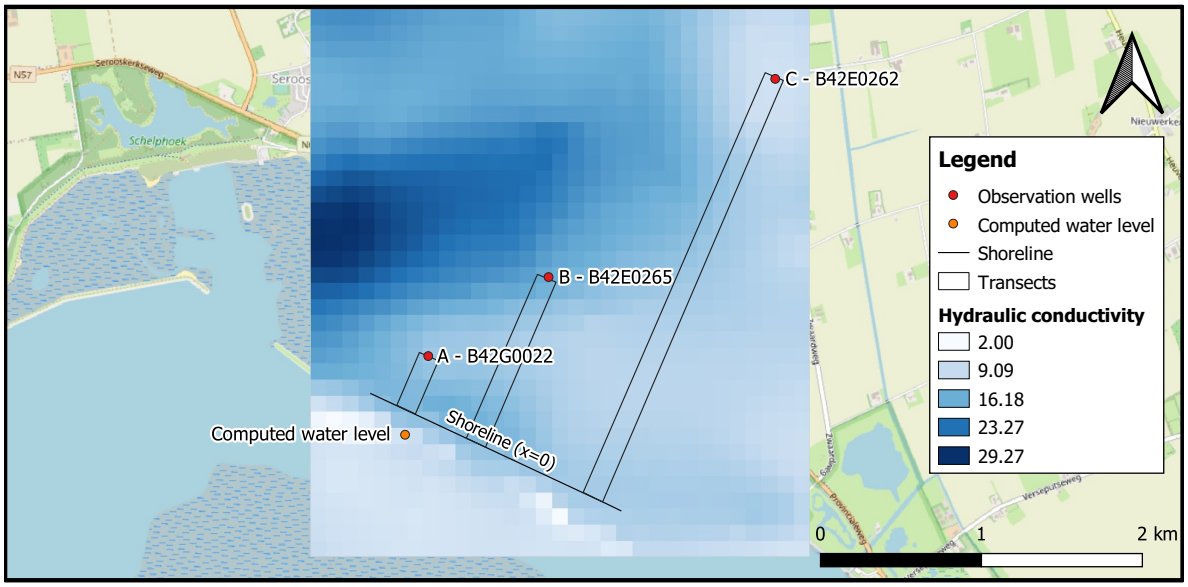


Figure C.3: A map of the 97.5th percentile of the hydraulic conductivity (m/day), from the upper aquifer, at the study site. The map is based on the statistics from the Regis-II model (Vernes et al., 2005). The 3 polygons, called transects in the legend, were used to calculate some basic statistics for each transect, but more on that in Section 5.5.3. Lastly note that the thickness of the aquifer associated with the presented hydraulic conductivity is constant (20.3 m) instead of varying (see Figure C.4).

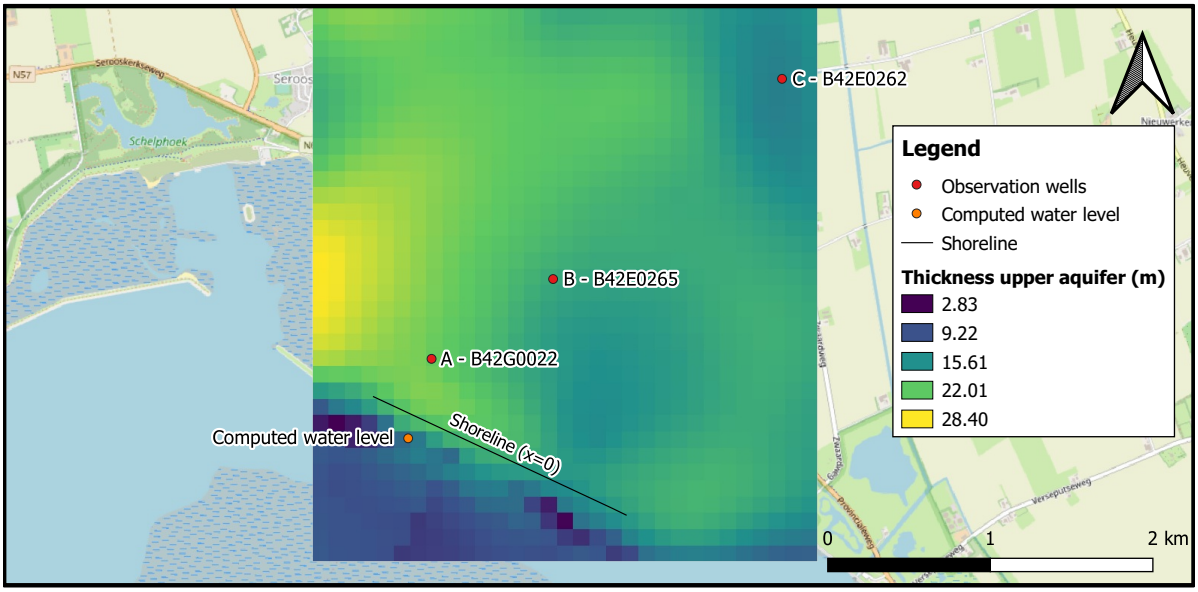


Figure C.4: The thickness of the upper aquifer at the study site. The map is based on estimates from the Regis-II model (Vernes et al., 2005).

C.2.2. Lower aquifer

This Appendix-section contains maps of the 2.5th and 97.5th percentiles (Figures C.5 & C.6) of the hydraulic conductivity, from the lower aquifer, at the study site. In addition, a map of the aquifer thickness is presented (Figure C.7). Note that the lower aquifer is located between the clayey unit from the Maassluis formation and a clayey unit from the Oosterhout formation (see Figure 2.2).

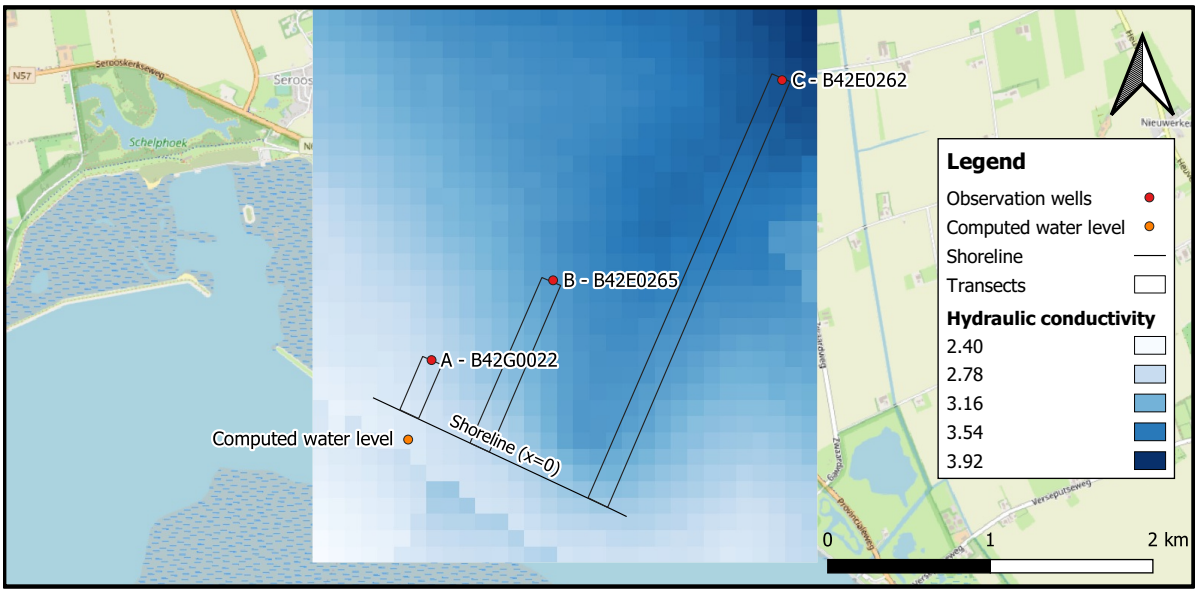


Figure C.5: A map of the 2.5th percentile of the hydraulic conductivity (m/day), from the lower aquifer, at the study site. The map is based on the statistics from the Regis-II model (Vernes et al., 2005). The 3 polygons, called transects in the legend, were used to calculate some basic statistics for each transect, but more on that in Section 5.5.3. Lastly note that the thickness of the aquifer associated with the presented hydraulic conductivity is constant (42 m) instead of varying (see Figure C.4).

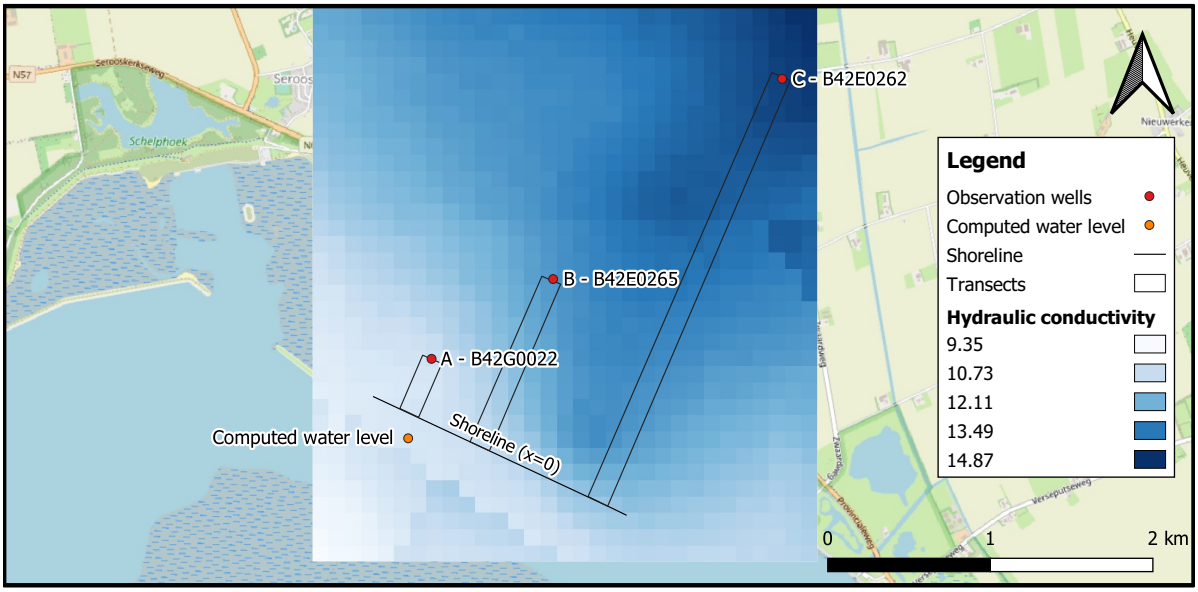


Figure C.6: A map of the 97.5th percentile of the hydraulic conductivity (m/day), from the lower aquifer, at the study site. The map is based on the statistics from the Regis-II model (Vernes et al., 2005). The 3 polygons, called transects in the legend, were used to calculate some basic statistics for each transect, but more on that in Section 5.5.3. Lastly note that the thickness of the aquifer associated with the presented hydraulic conductivity is constant (42 m) instead of varying (see Figure C.4).

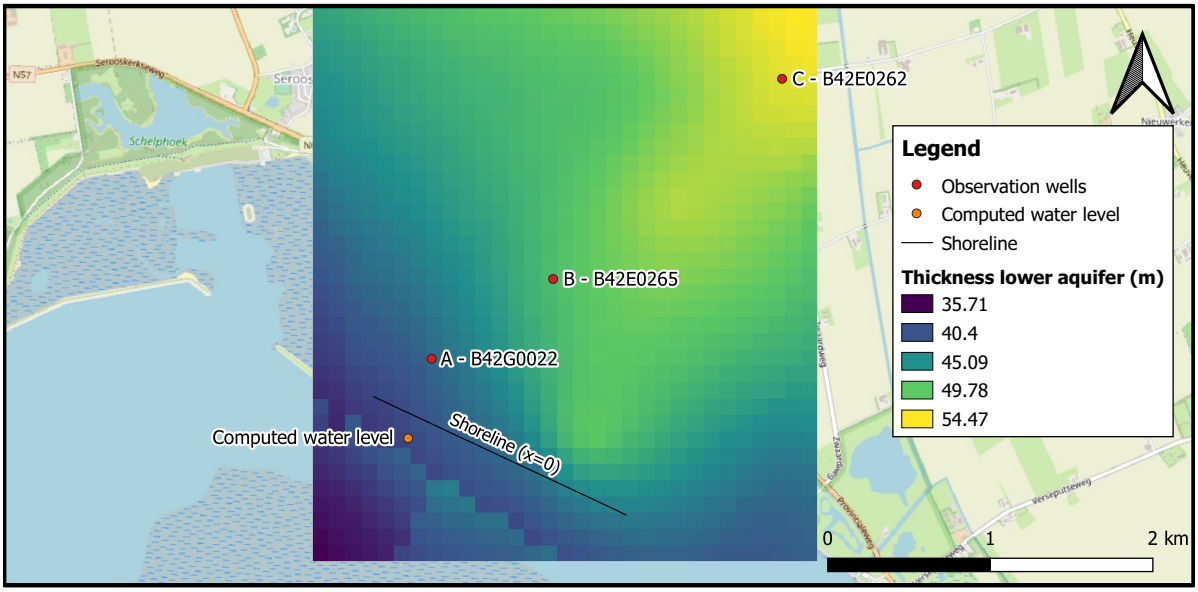


Figure C.7: The thickness of the lower aquifer at the study site. The map is based on estimates from the Regis-II model (Vernes et al., 2005).

C.2.3. Clay lens

This Appendix-section contains maps of the 2.5th and 97.5th percentiles (Figures C.8 & C.9) of the resistance from the clay lens between the upper and lower aquifer (i.e. the clayey unit from the Maassluis formation), at the study site.

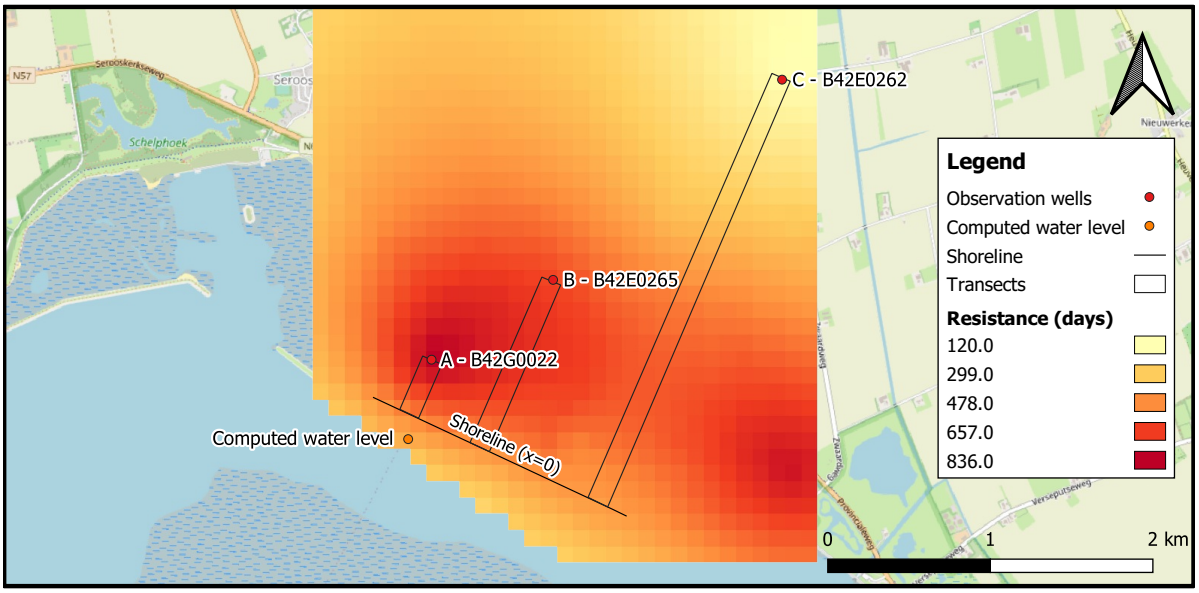


Figure C.8: A map of the 2.5th percentile of the resistance (days) from the clay lens between the upper and lower aquifer (i.e. the clayey unit from the Maassluis formation), at the study site, The map is based on the statistics from the Regis-II model (Vernes et al., 2005). The 3 polygons, called transects in the legend, were used to calculate some basic statistics for each transect, but more on that in Section 5.5.3.

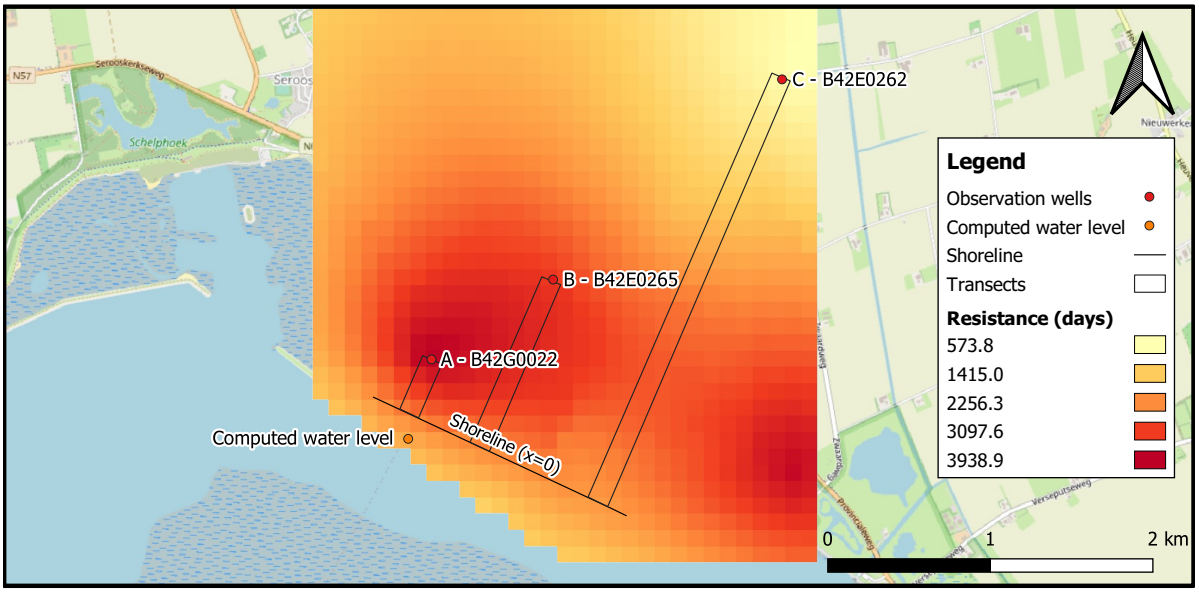


Figure C.9: A map of the 97.5th percentile of the resistance (days) from the clay lens between the upper and lower aquifer (i.e. the clayey unit from the Maassluis formation), at the study site, The map is based on the statistics from the Regis-II model (Vernes et al., 2005). The 3 polygons, called transects in the legend, were used to calculate some basic statistics for each transect, but more on that in Section 5.5.3.

C.3. Spatial maps of the M2 amplitude and phase in the Oosterschelde

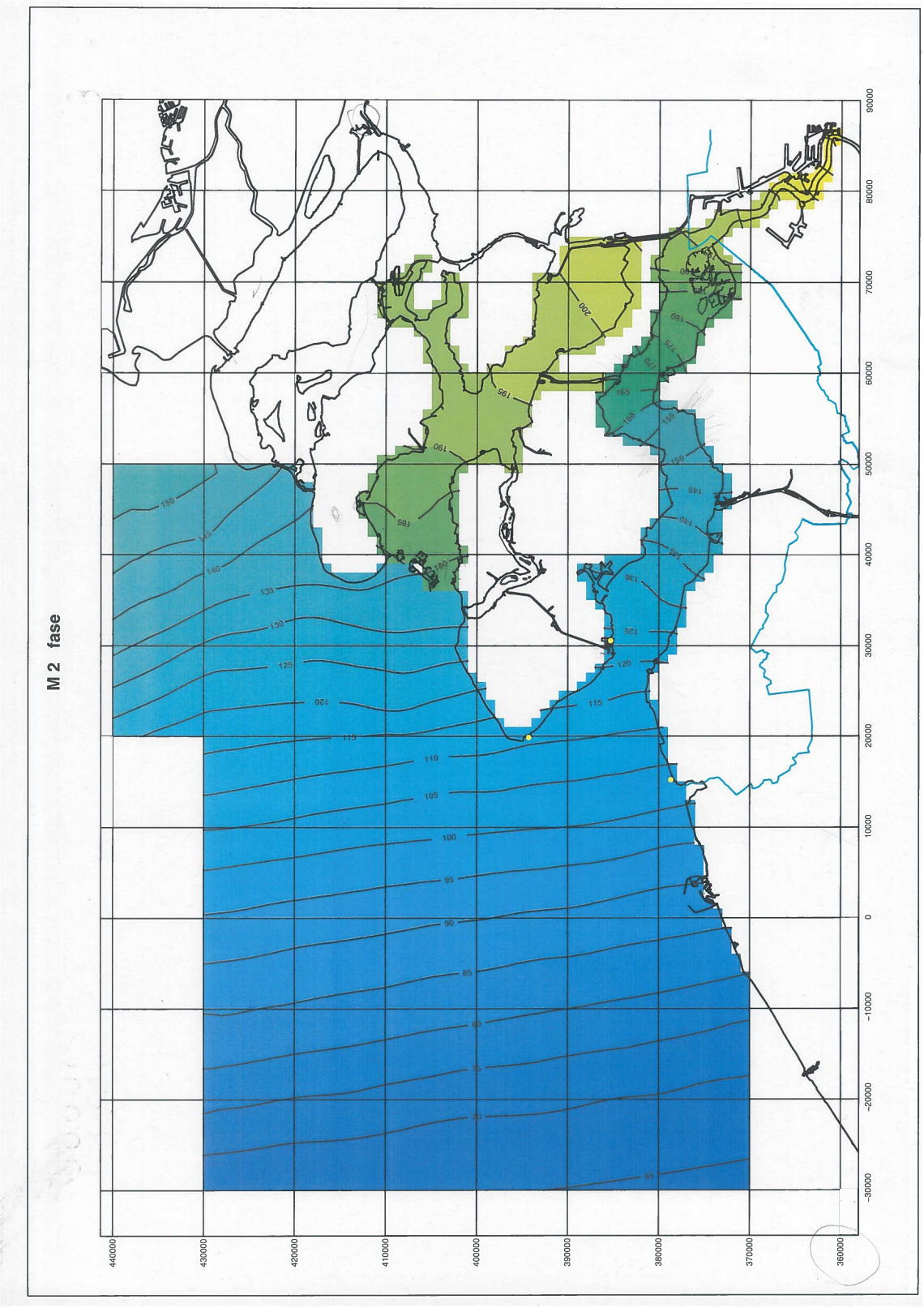


Figure C.10: Spatial map of the M2 phase

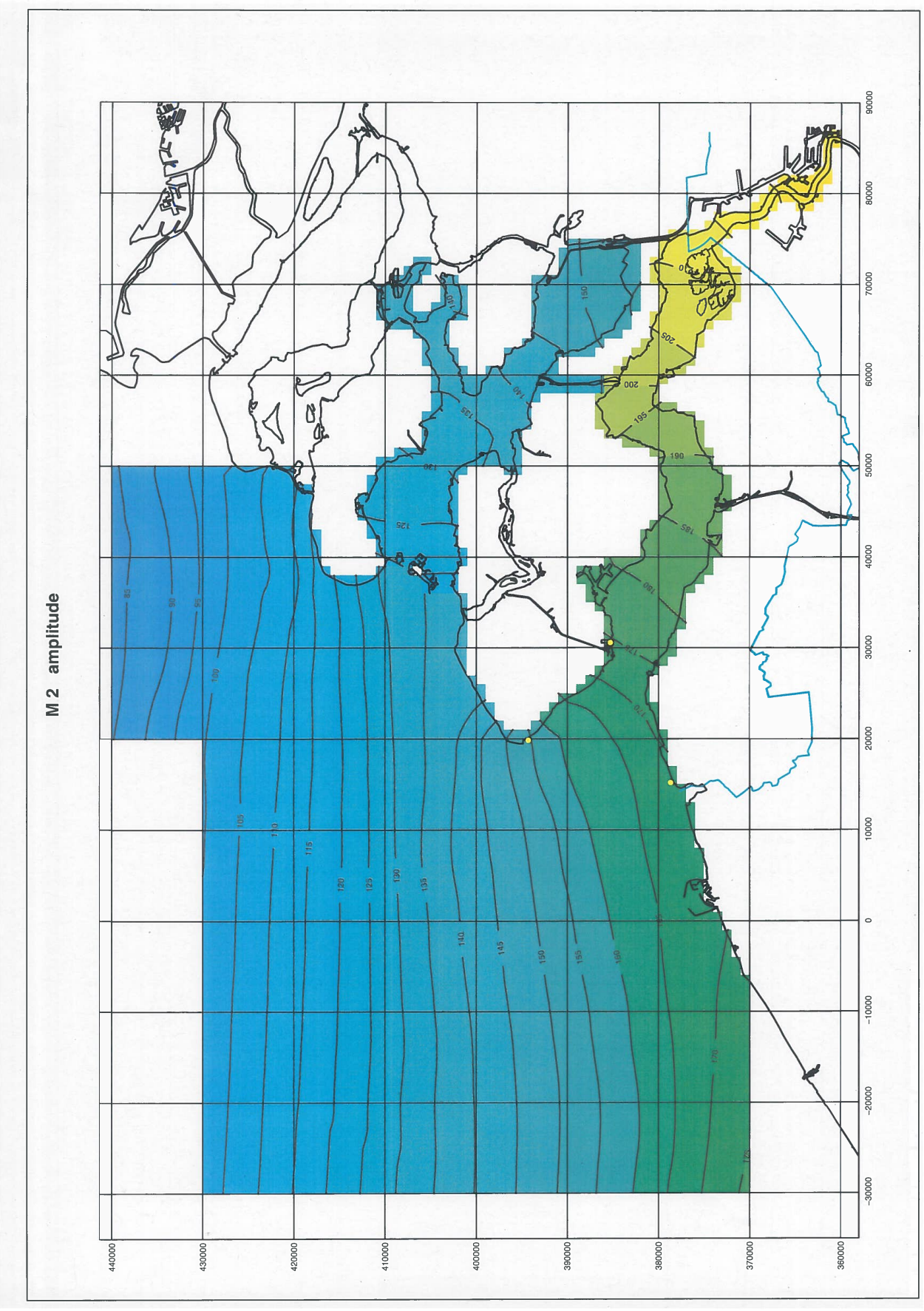


Figure C.11: Spatial map of the M2 amplitude

C.4. Tidal propagation in the Oosterschelde graphs

Note that the constituents are plotted from low to high frequency.

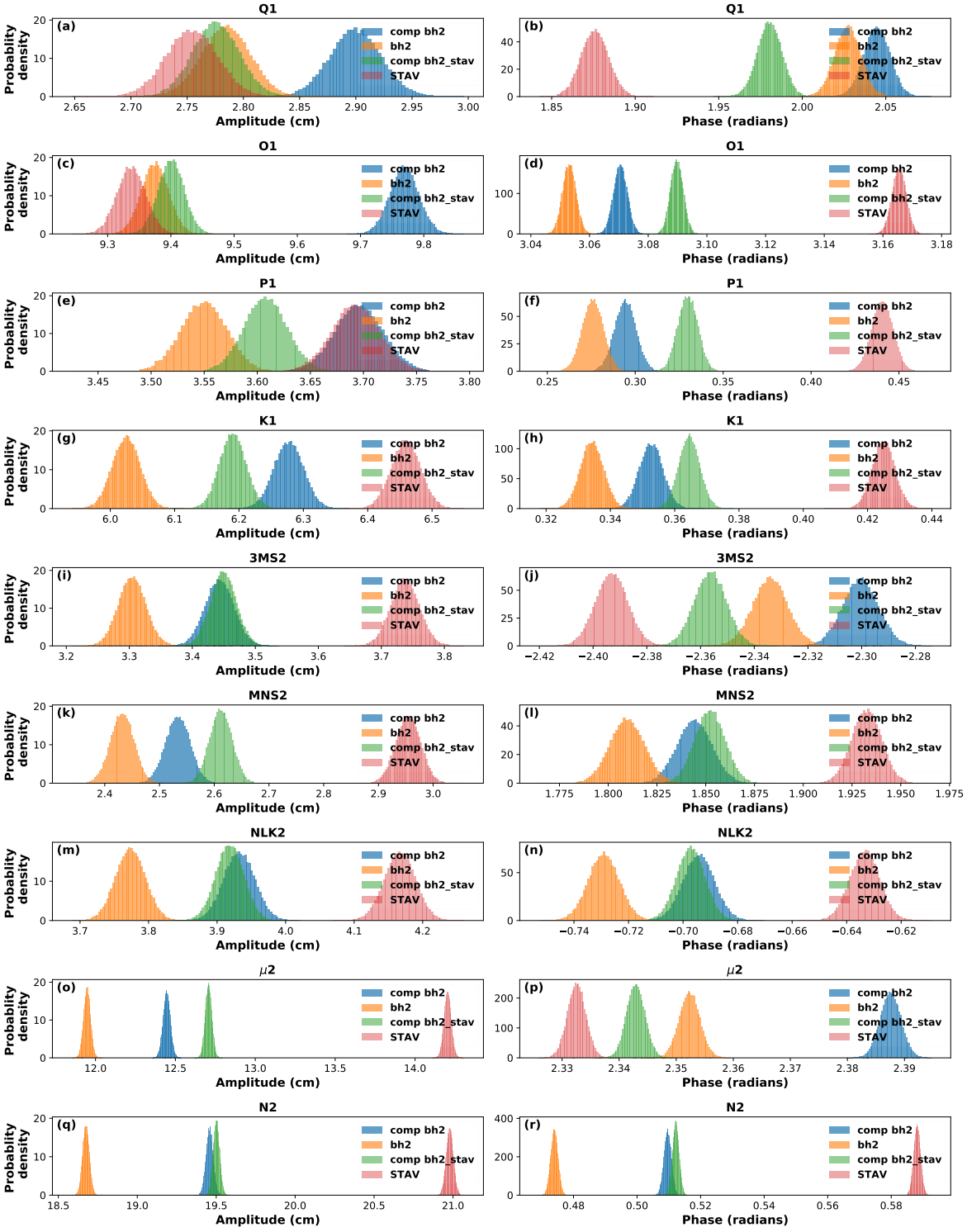


Figure C.12: Probability density graphs for a part (1/3) of the different constituents observed in the Oosterschelde for the study site (comp) and the measuring points at Neeltje Jans (bh2) and Stavenisse (STAV). The probability density functions are calculated with a harmonic analysis of the astronomical tide. In orange and red the decomposed tide based on the astronomical tide at measuring points bh2 and Stavenisse (STAV) respectively (see Figure 2.1). In blue the computed water level (see Figure 2.1) translated from bh2 and in green the computed water level calculated from an interpolation based on the translated time series from bh2 and STAV.

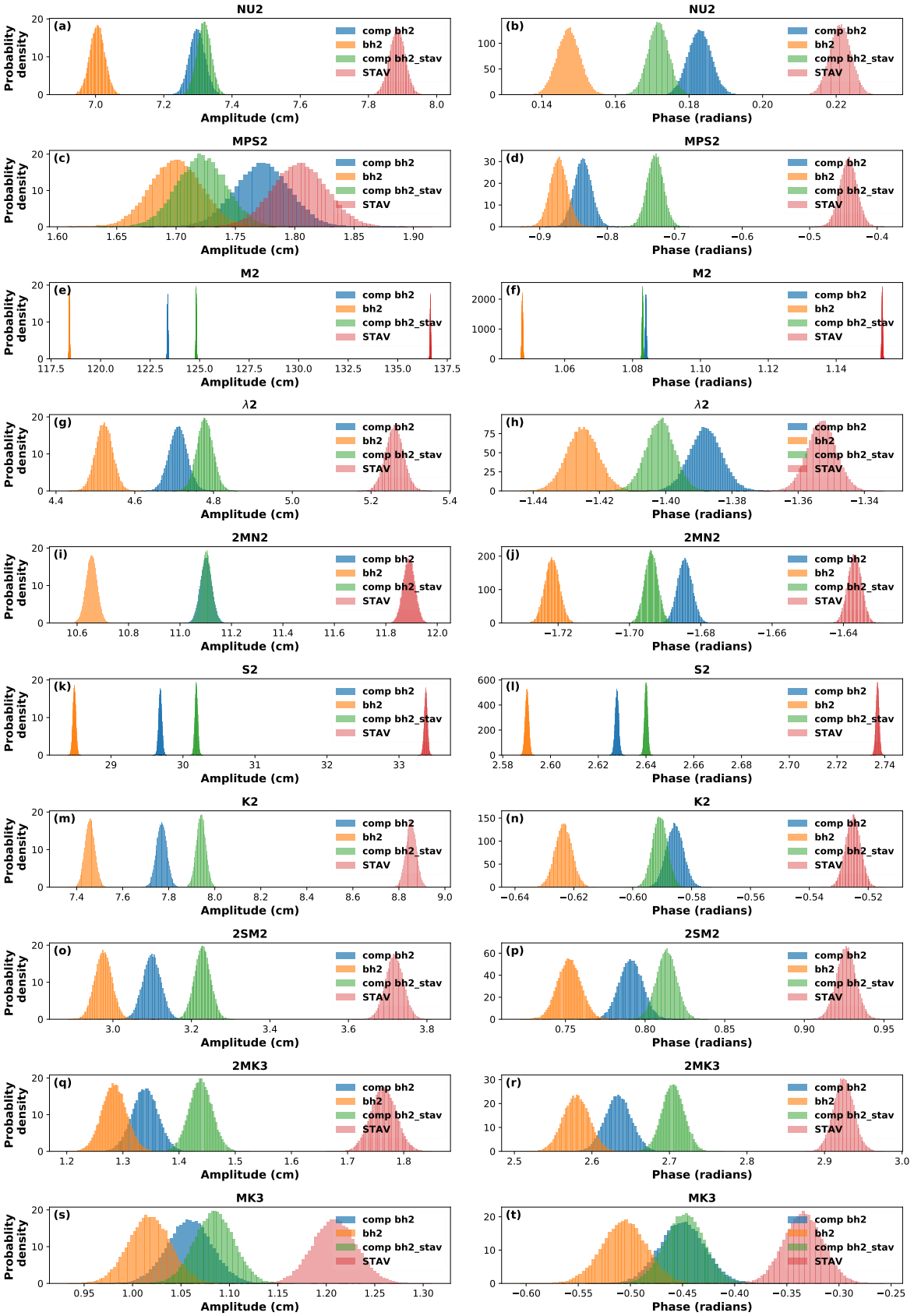


Figure C.13: Probability density graphs for a part (2/3) of the different constituents observed in the Oosterschelde for the study site (comp) and the measuring points at Neeltje Jans (bh2) and Stavenisse (STAV). The probability density functions are calculated with a harmonic analysis of the astronomical tide. In orange and red the decomposed tide based on the astronomical tide at measuring points bh2 and Stavenisse (STAV) respectively (see Figure 2.1). In blue the computed water level (see Figure 2.1) translated from bh2 and in green the computed water level calculated from an interpolation based on the translated time series from bh2 and STAV.

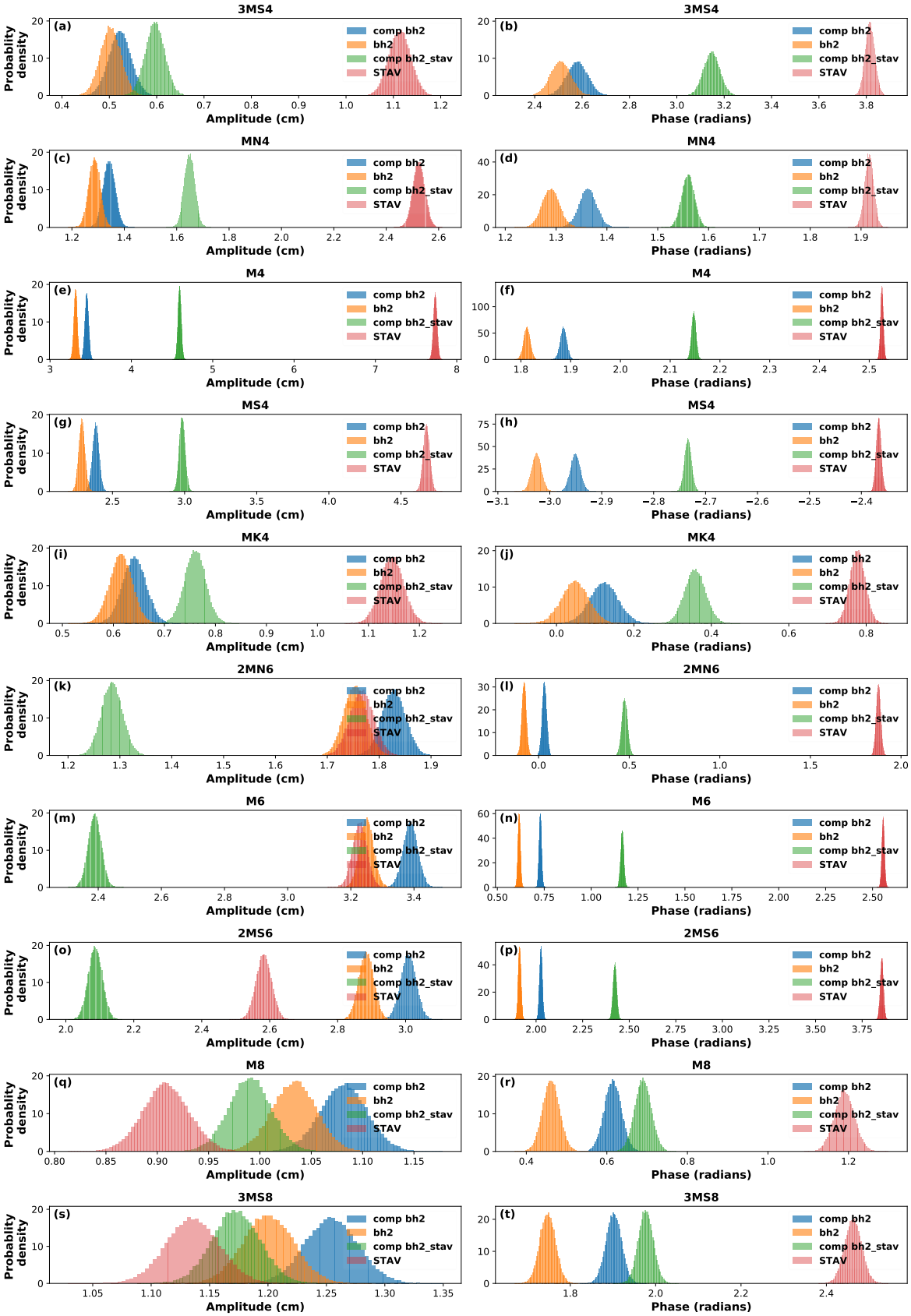
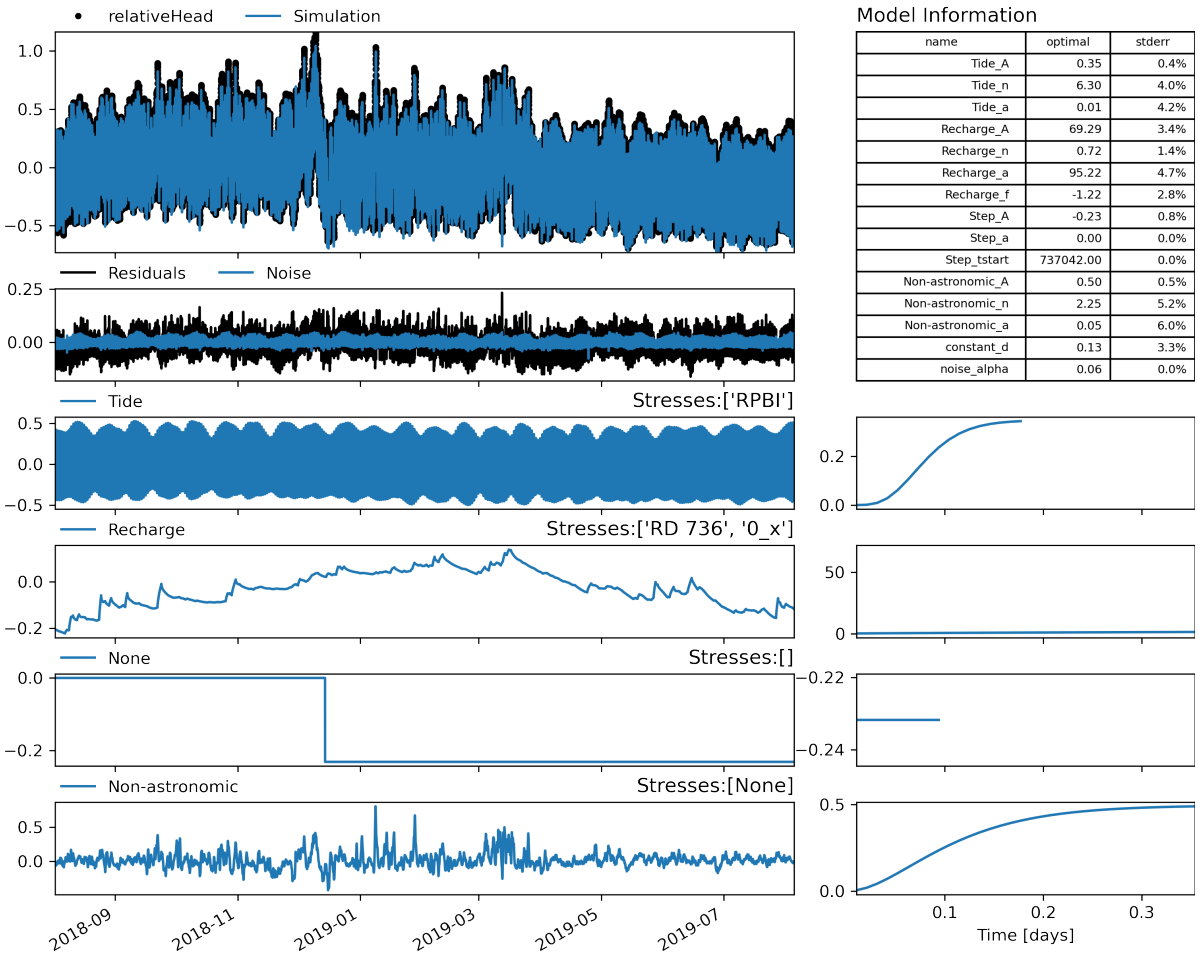


Figure C.14: Probability density graphs for a part (3/3) of the different constituents observed in the Oosterschelde for the study site (comp) and the measuring points at Neeltje Jans (bh2) and Stavenisse (STAV). The probability density functions are calculated with a harmonic analysis of the astronomical tide. In orange and red the decomposed tide based on the astronomical tide at measuring points bh2 and Stavenisse (STAV) respectively (see Figure 2.1). In blue the computed water level (see Figure 2.1) translated from bh2 and in green the computed water level calculated from an interpolation based on the translated time series from bh2 and STAV.

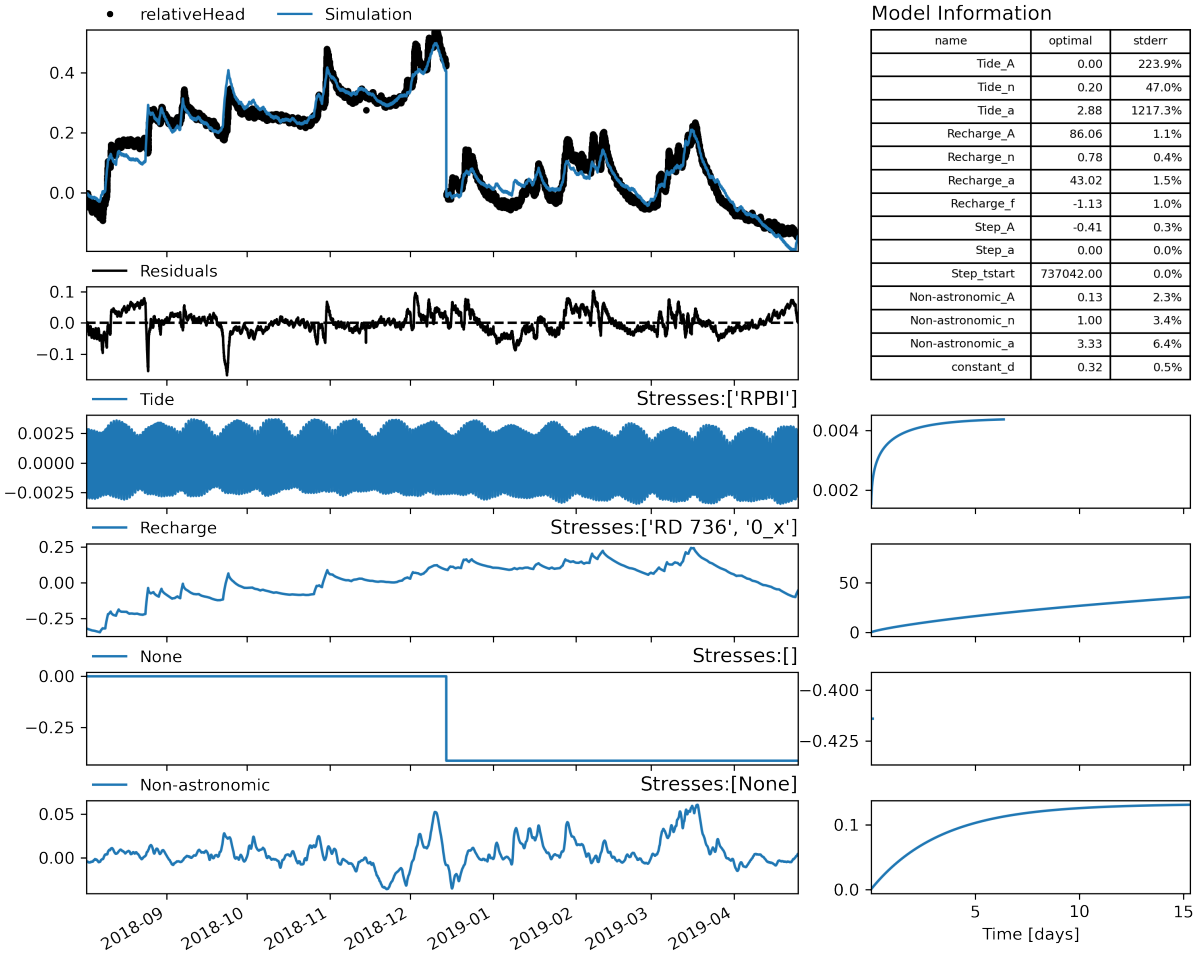
C.5. Removing noise with Pastas, wells A2 and C2



Model Information

name	optimal	stderr
Tide_A	0.35	0.4%
Tide_n	6.30	4.0%
Tide_a	0.01	4.2%
Recharge_A	69.29	3.4%
Recharge_n	0.72	1.4%
Recharge_a	95.22	4.7%
Recharge_f	-1.22	2.8%
Step_A	-0.23	0.8%
Step_a	0.00	0.0%
Step_tstart	737042.00	0.0%
Non-astronomic_A	0.50	0.5%
Non-astronomic_n	2.25	5.2%
Non-astronomic_a	0.05	6.0%
constant_d	0.13	3.3%
noise_alpha	0.06	0.0%

Figure C.15: Decomposition of the groundwater fluctuations with Pastas for the times series from well A2, results are in meter. The decomposed stresses that cause the head response in the aquifer are the astronomical tide in the Oosterschelde (named 'Tide' in the plot), the recharge and the non-astronomic water level fluctuations in the Oosterschelde. The explained variance of the model is 97.60



Model Information

name	optimal	stderr
Tide_A	0.00	223.9%
Tide_n	0.20	47.0%
Tide_a	2.88	1217.3%
Recharge_A	86.06	1.1%
Recharge_n	0.78	0.4%
Recharge_a	43.02	1.5%
Recharge_f	-1.13	1.0%
Step_A	-0.41	0.3%
Step_a	0.00	0.0%
Step_tstart	737042.00	0.0%
Non-astronomic_A	0.13	2.3%
Non-astronomic_n	1.00	3.4%
Non-astronomic_a	3.33	6.4%
constant_d	0.32	0.5%

Figure C.16: Decomposition of the groundwater fluctuations with Pastas for the times series from well C2, results are in meter. The decomposed stresses that cause the head response in the aquifer are the astronomical tide in the Oosterschelde (named 'Tide' in the plot), the recharge and the non-astronomic water level fluctuations in the Oosterschelde. The explained variance of the model is 95.63

C.6. Removing noise with Butterworth, wells A2, B2 and A4

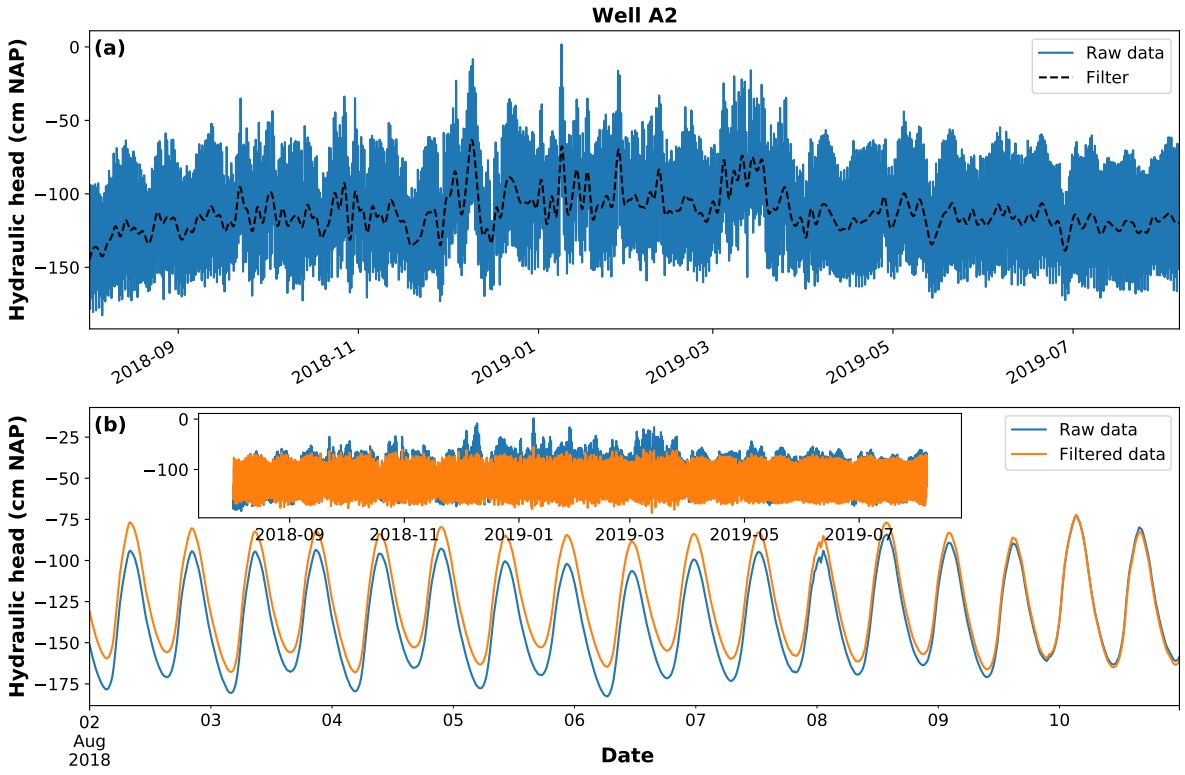


Figure C.17: Plots of the hydraulic head (raw data) and the obtained filter (plot a) and a plot with both the hydraulic head (raw data) and the filtered data (b), with 9 days in August in the main plot and the whole time series in the subplot.

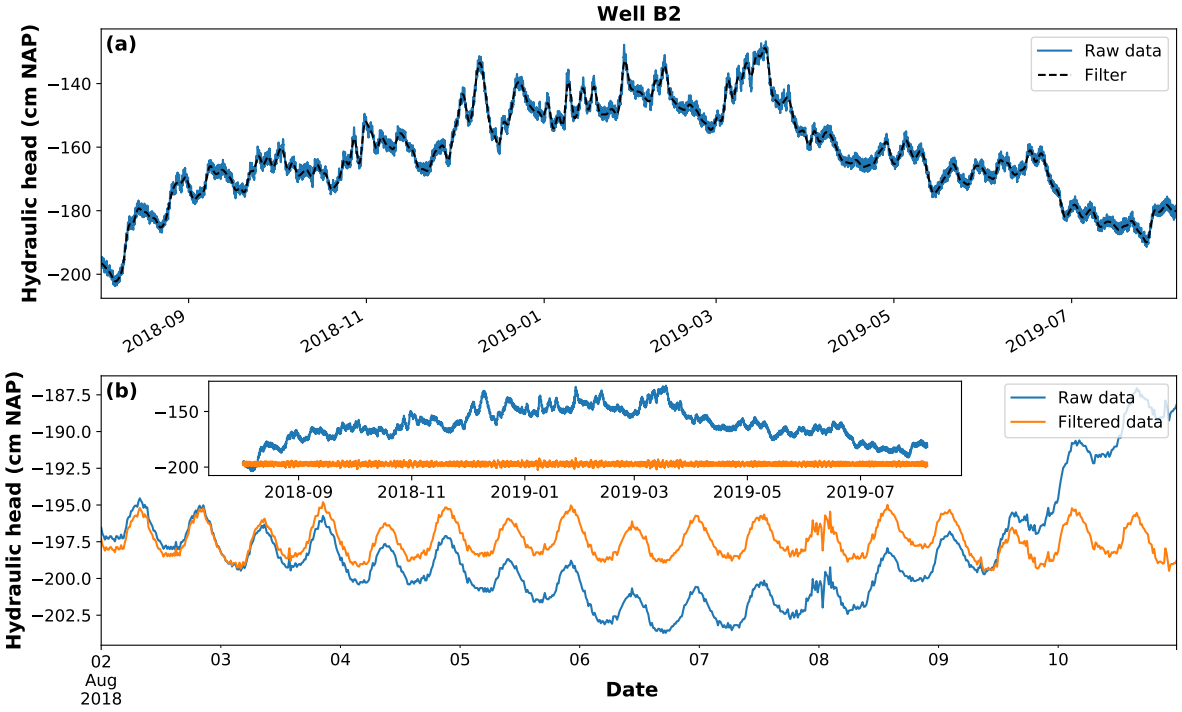


Figure C.18: Plots of the hydraulic head (raw data) and the obtained filter (plot a) and a plot with both the hydraulic head (raw data) and the filtered data (b), with 9 days in August in the main plot and the whole time series in the subplot.

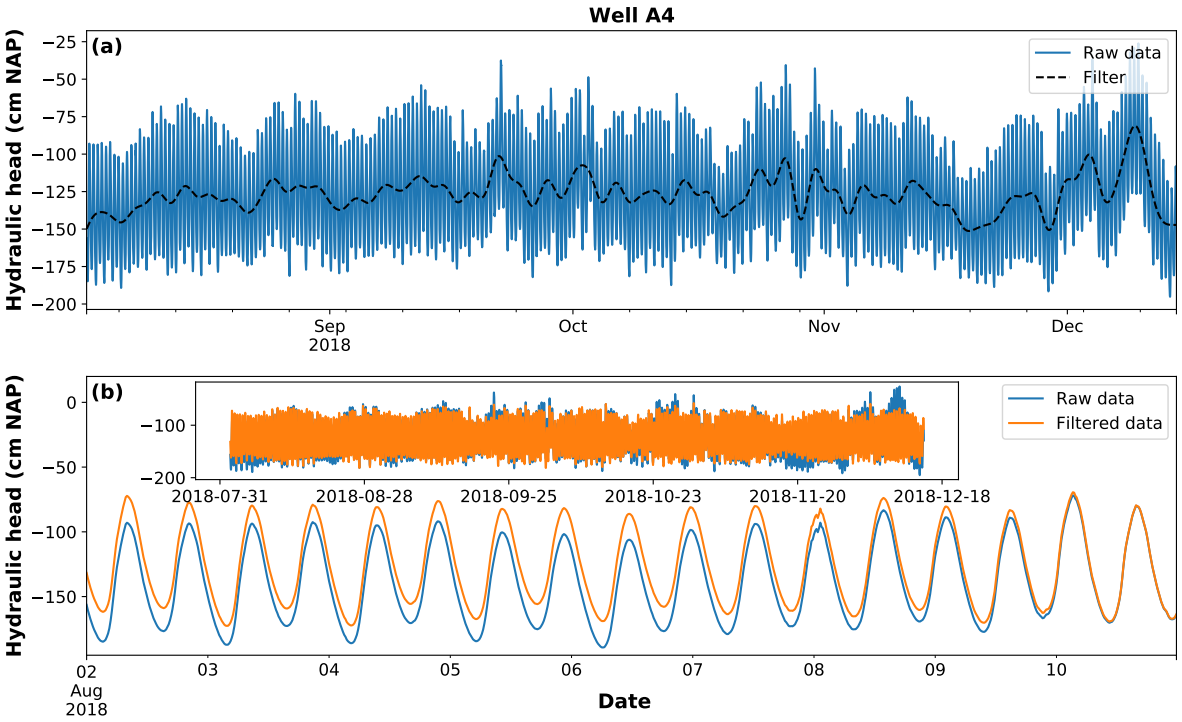


Figure C.19: Plots of the hydraulic head (raw data) and the obtained filter (plot a) and a plot with both the hydraulic head (raw data) and the filtered data (b), with 9 days in August in the main plot and the whole time series in the subplot.

C.7. Modeled and (filtered) observed fluctuations in the different groundwater wells

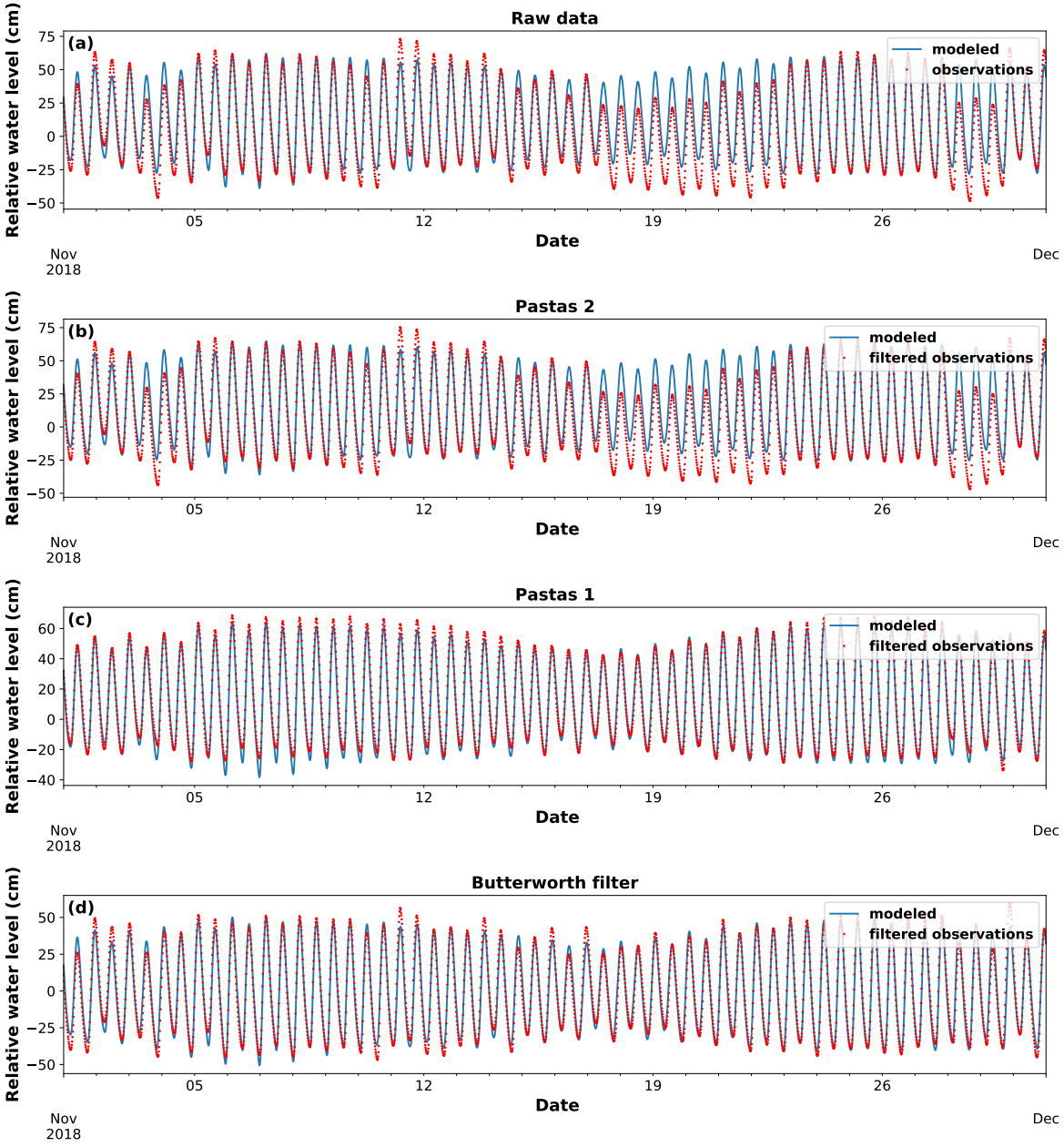


Figure C.20: Modeled groundwater fluctuations (blue line) & (filtered) groundwater observations (red dots) for December 2018, with different filtering techniques. For observation well A2.

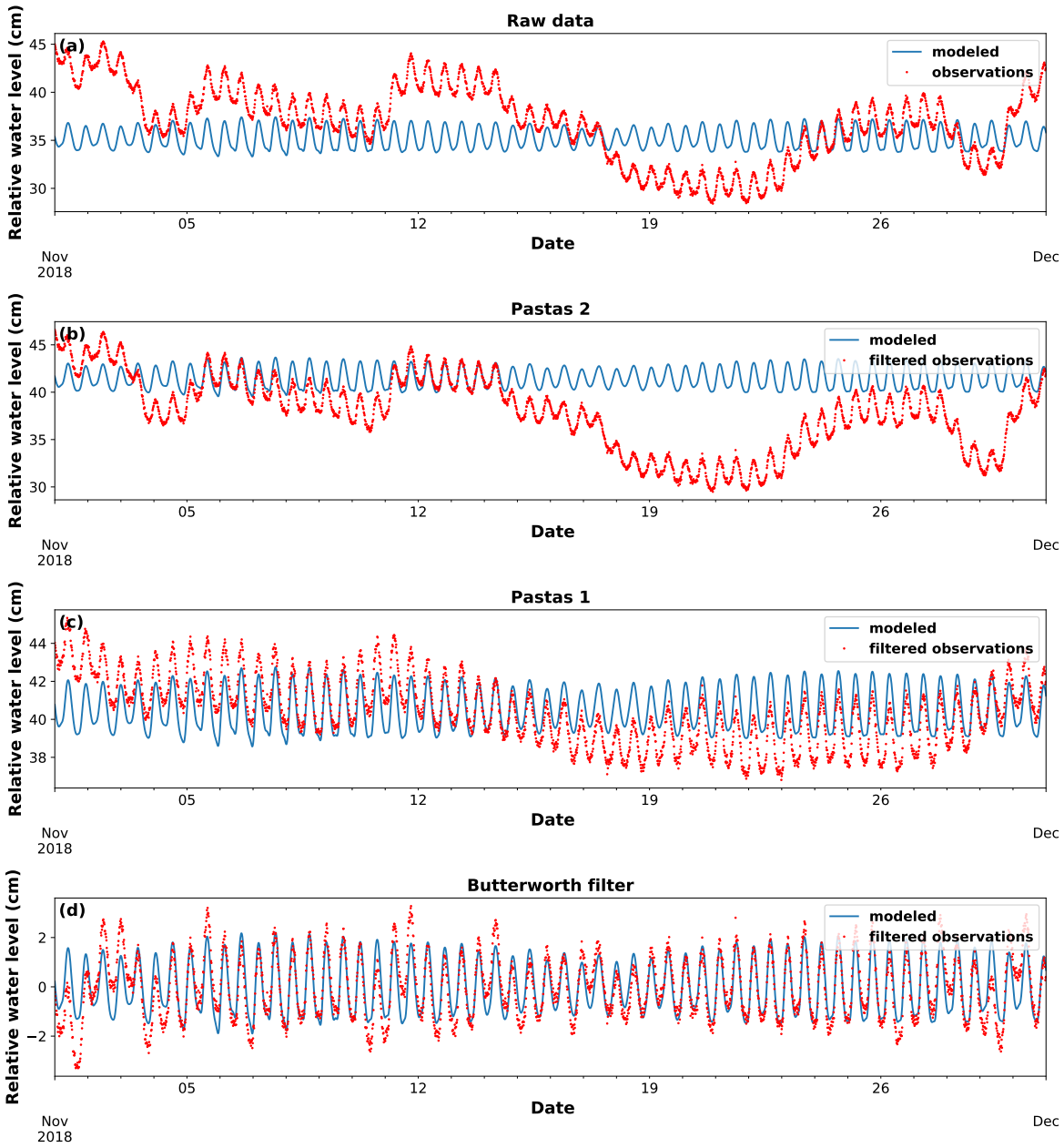


Figure C.21: Modeled groundwater fluctuations (blue line) & (filtered) groundwater observations (red dots) for December 2018, with different filtering techniques. For observation well B2.

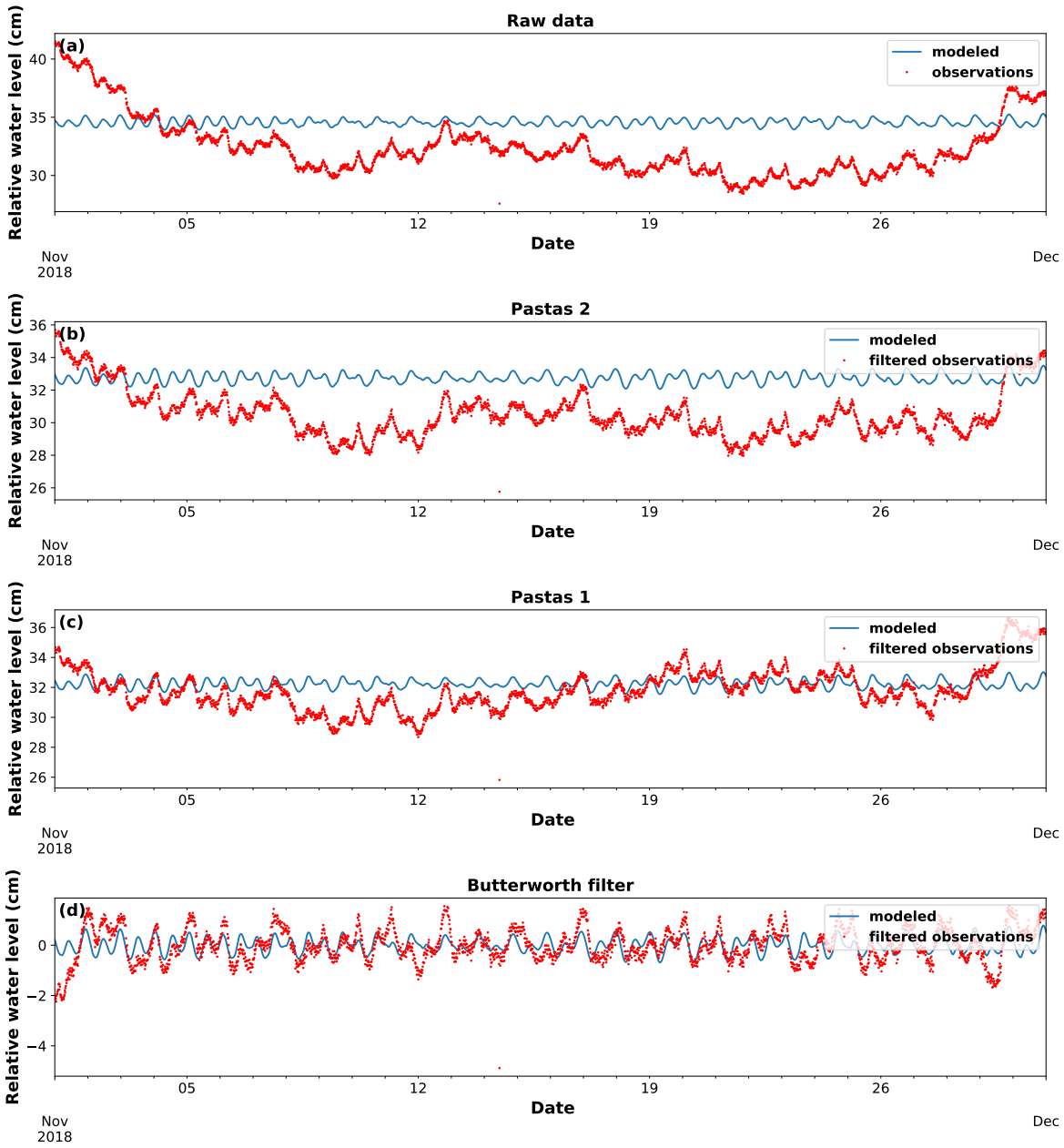


Figure C.22: Modeled groundwater fluctuations (blue line) & (filtered) groundwater observations (red dots) for December 2018, with different filtering techniques. For observation well C2.

C.8. Probability density graphs for the different constituents observed in the observation wells

C.8.1. Observation well A2 (B42G0022)

Note that the constituents are plotted from low to high frequency.

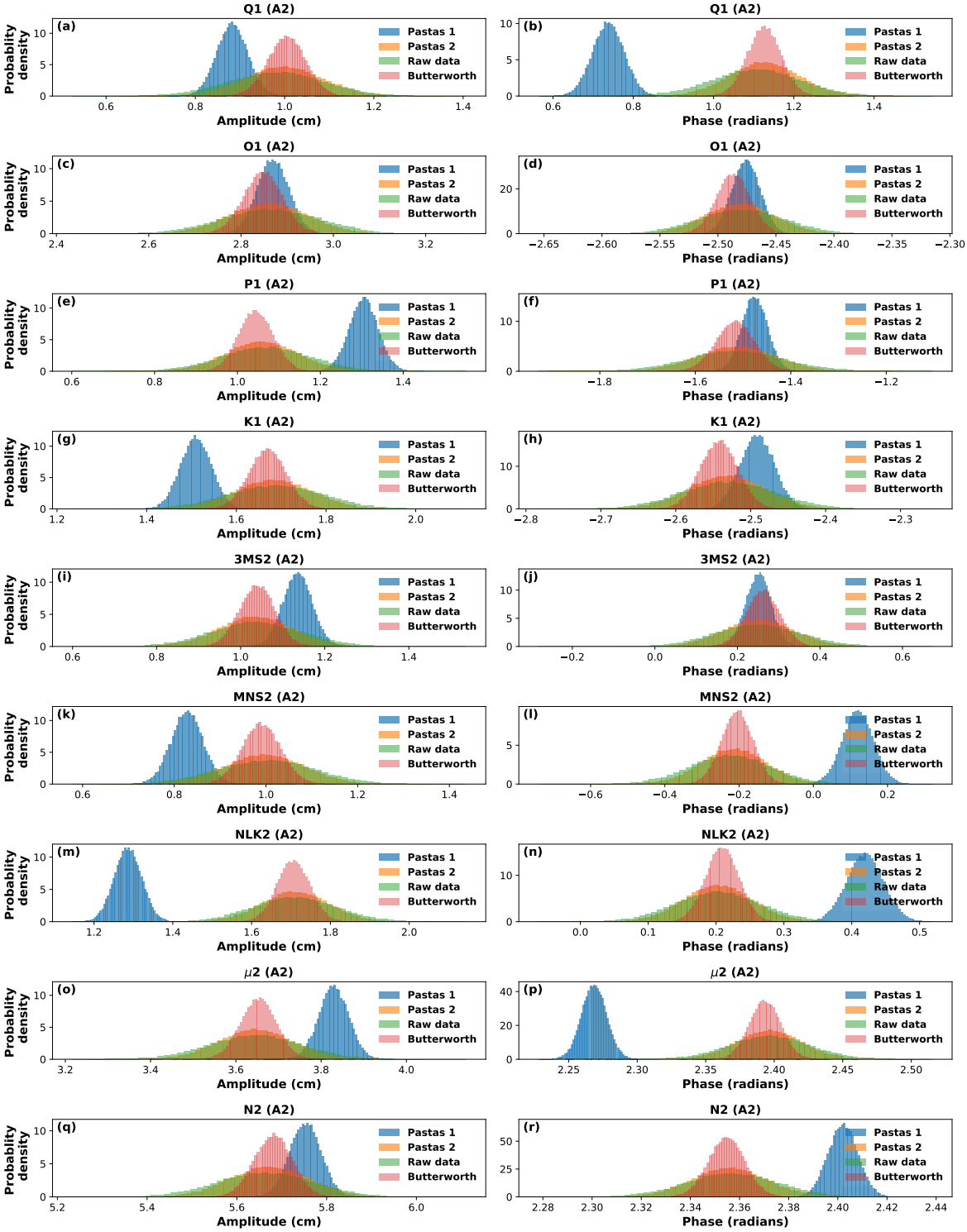


Figure C.23: Probability density graphs for a part (1/3) of the different constituents observed in well A2

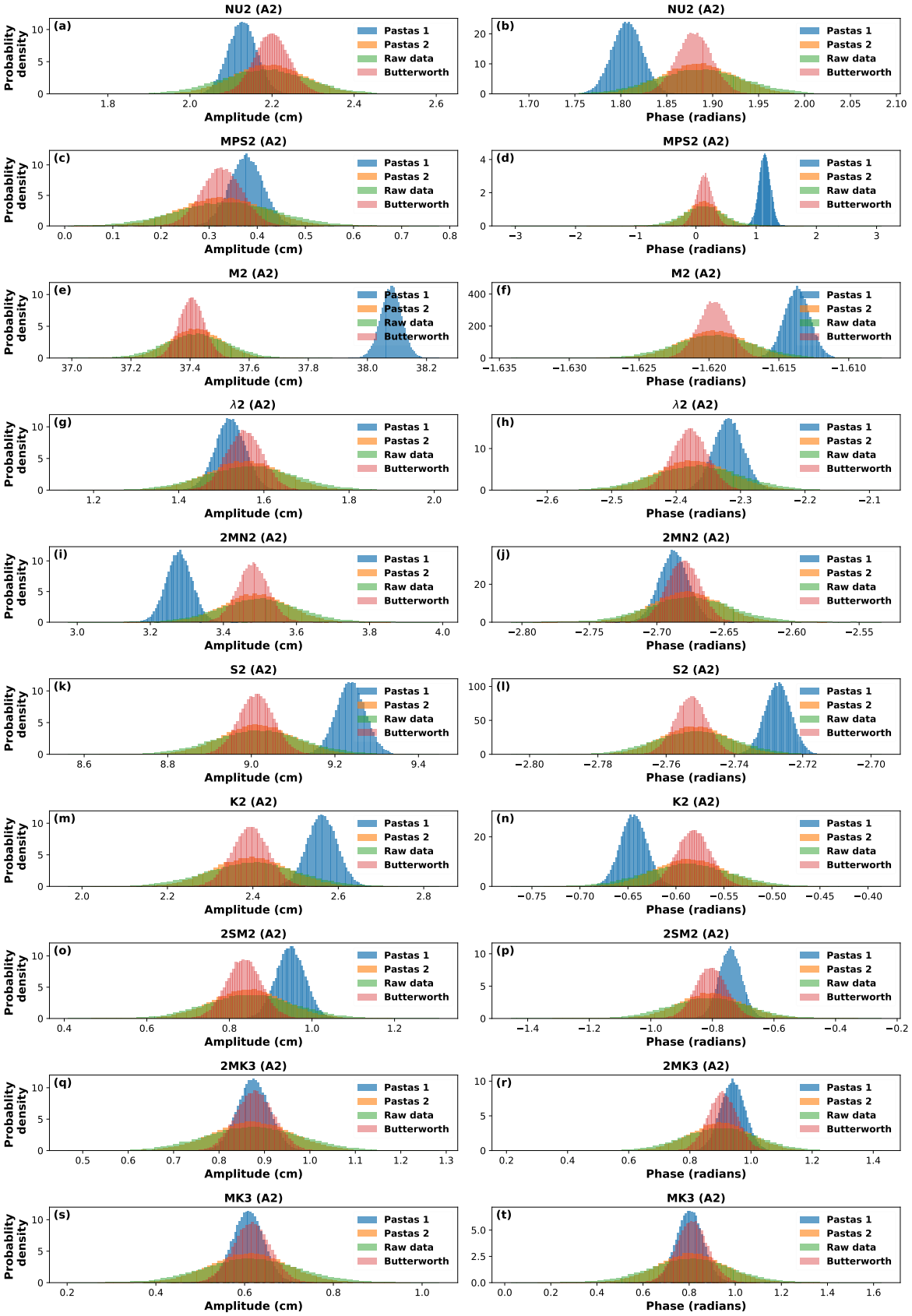


Figure C.24: Probability density graphs for a part (2/3) of the different constituents observed in well A2

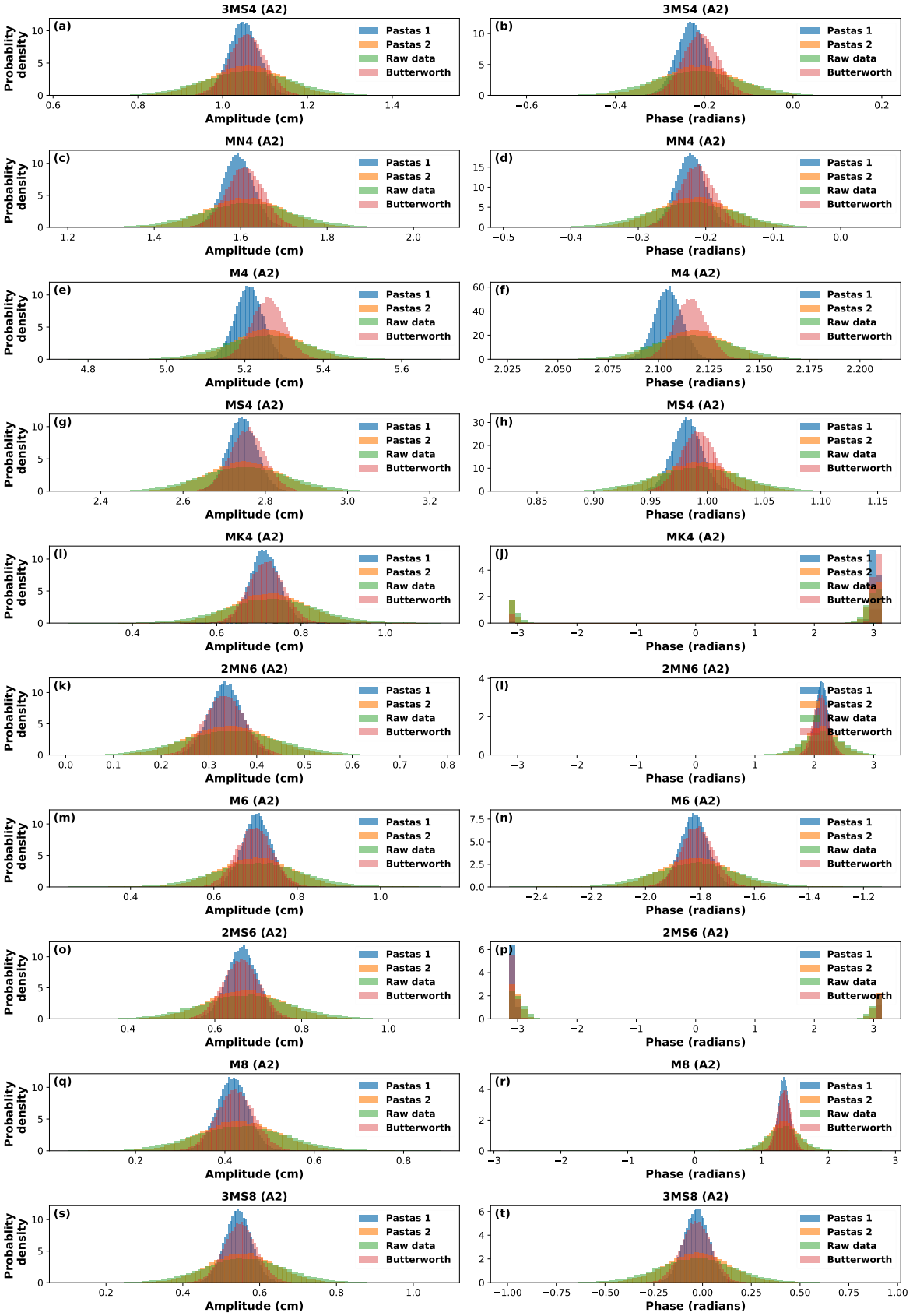


Figure C.25: Probability density graphs for a part (3/3) of the different constituents observed in well A2

C.8.2. Observation well B2 (B42E0265)

Note that the constituents are plotted from low to high frequency.

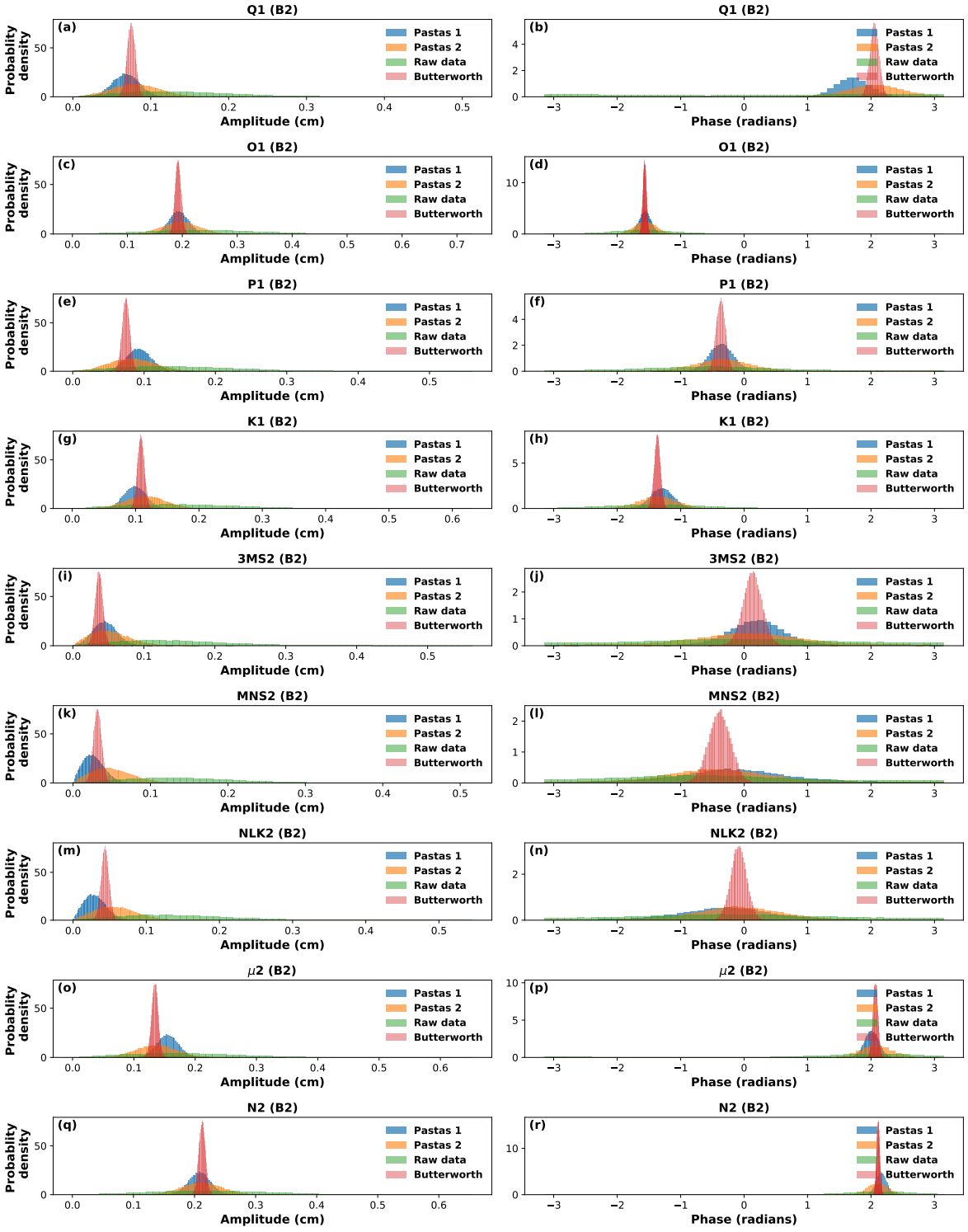


Figure C.26: Probability density graphs for a part (1/3) of the different constituents observed in well B2

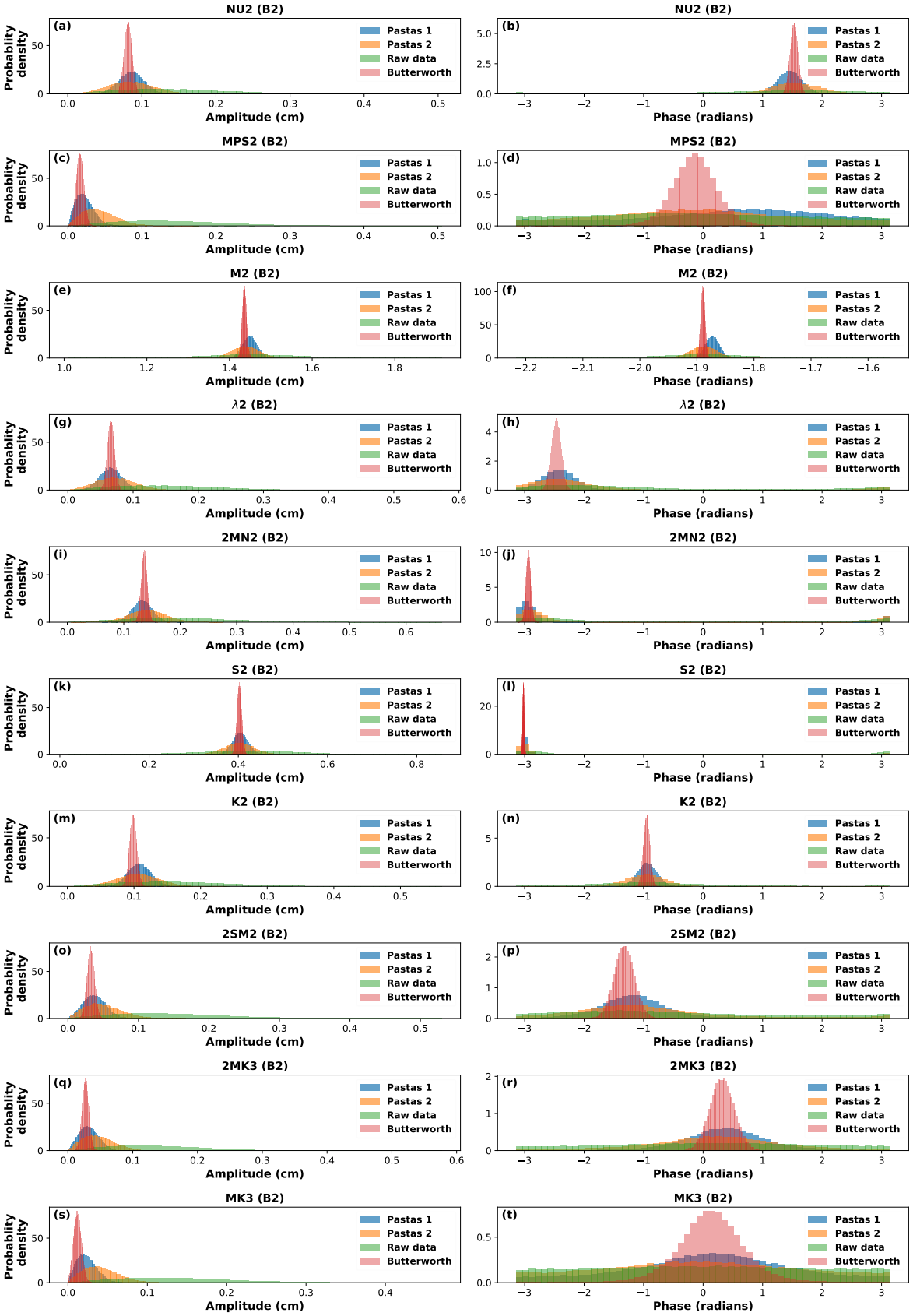


Figure C.27: Probability density graphs for a part (2/3) of the different constituents observed in well B2

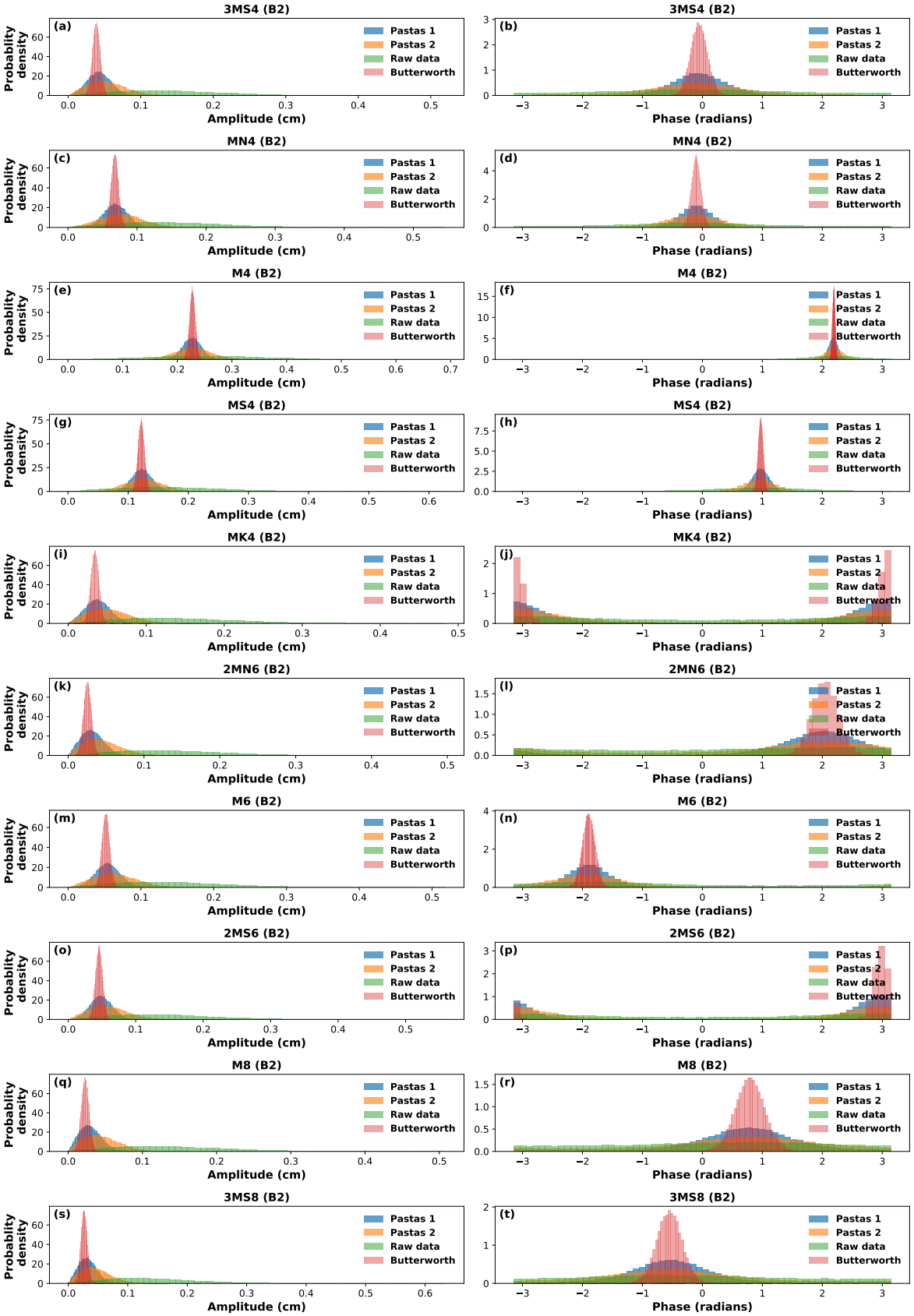


Figure C.28: Probability density graphs for a part (3/3) of the different constituents observed in well B2

C.8.3. Observation well C2 (B42E0262)

Note that the constituents are plotted from low to high frequency.

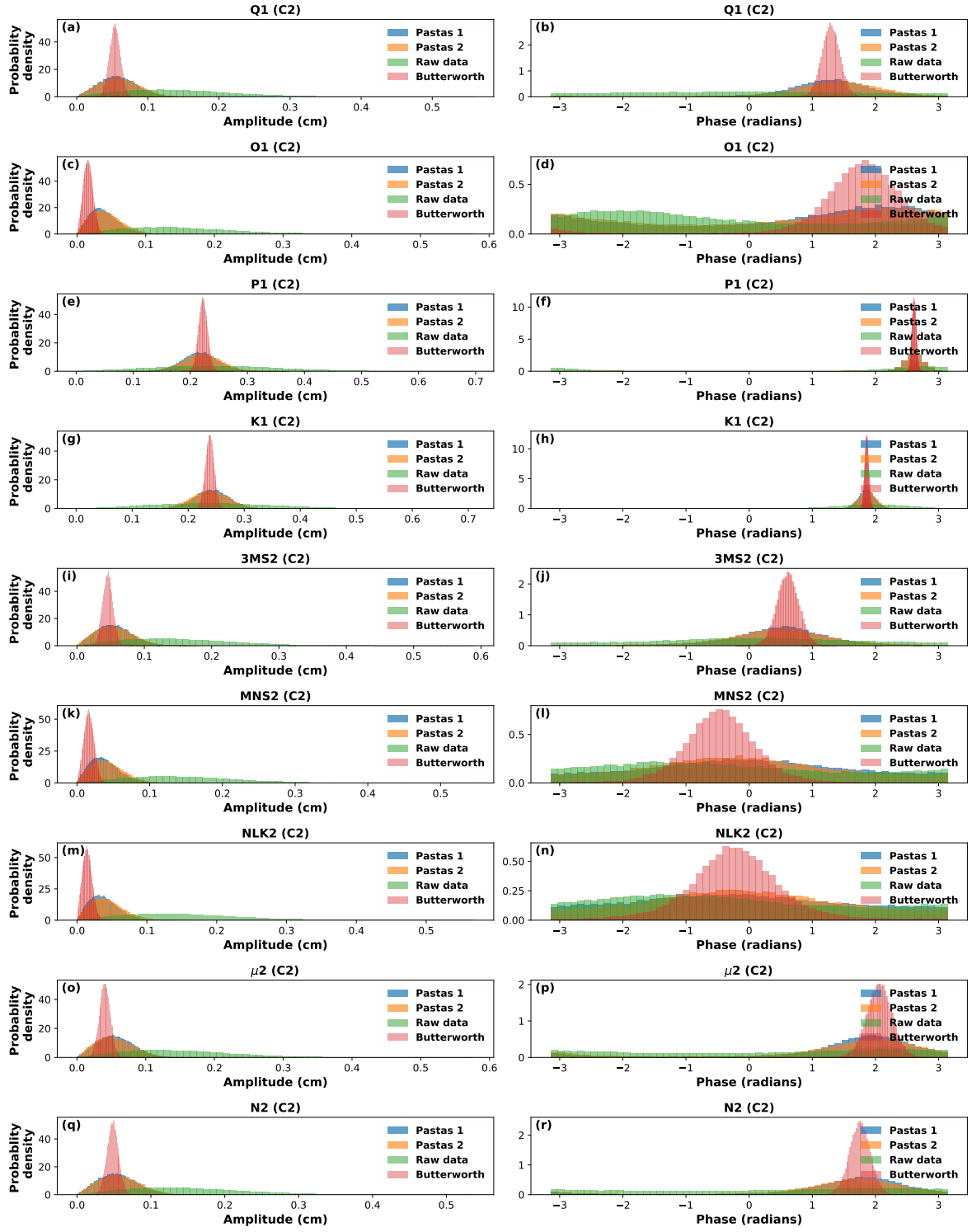


Figure C.29: Probability density graphs for a part (1/3) of the different constituents observed in well C2

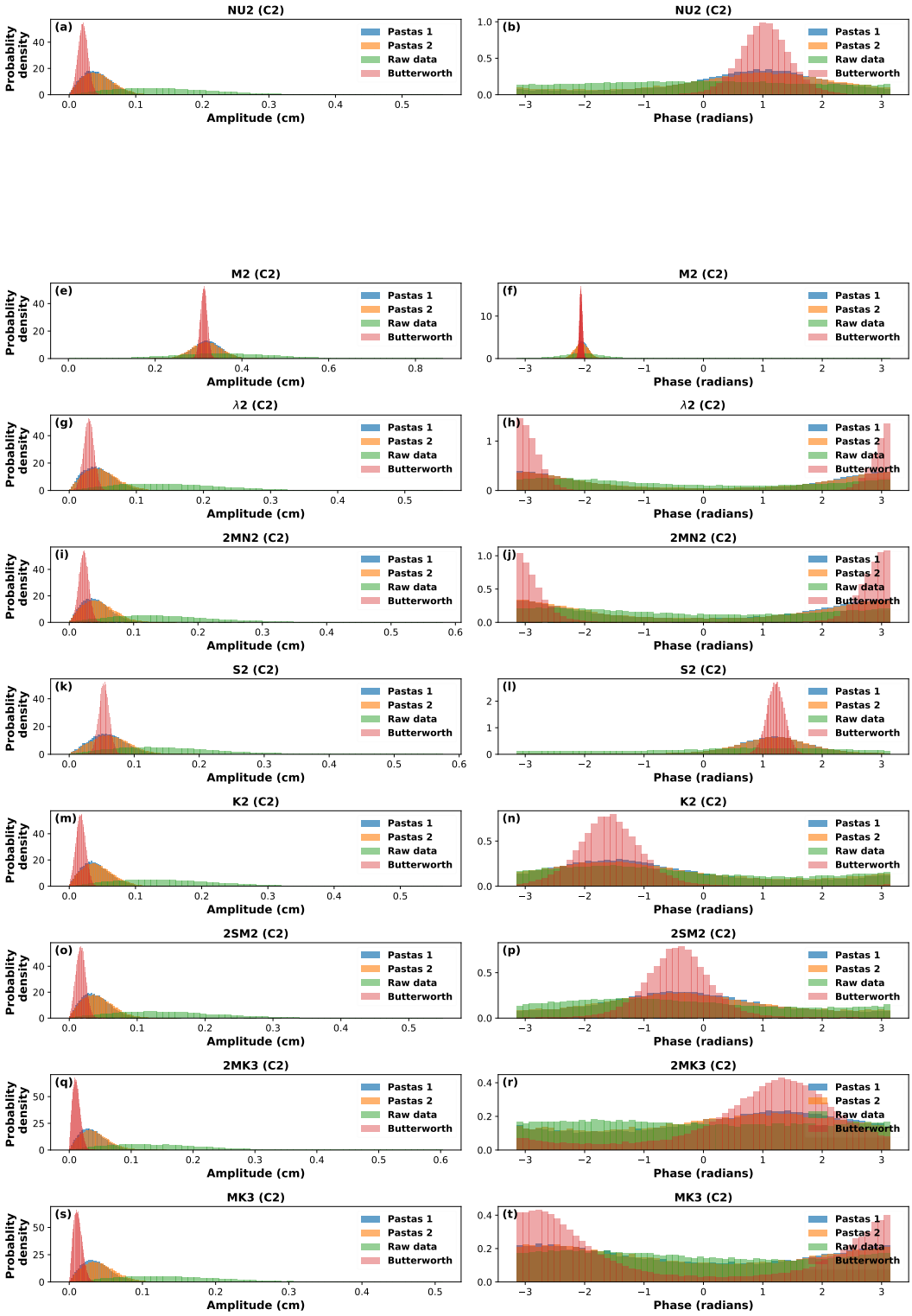


Figure C.30: Probability density graphs for a part (2/3) of the different constituents observed in well C2. Note that because of Rayleighs resolution limit and an expected insignificant contribution to the signal, MPS2 was not included in the calculations (see also Section 3.1.1).

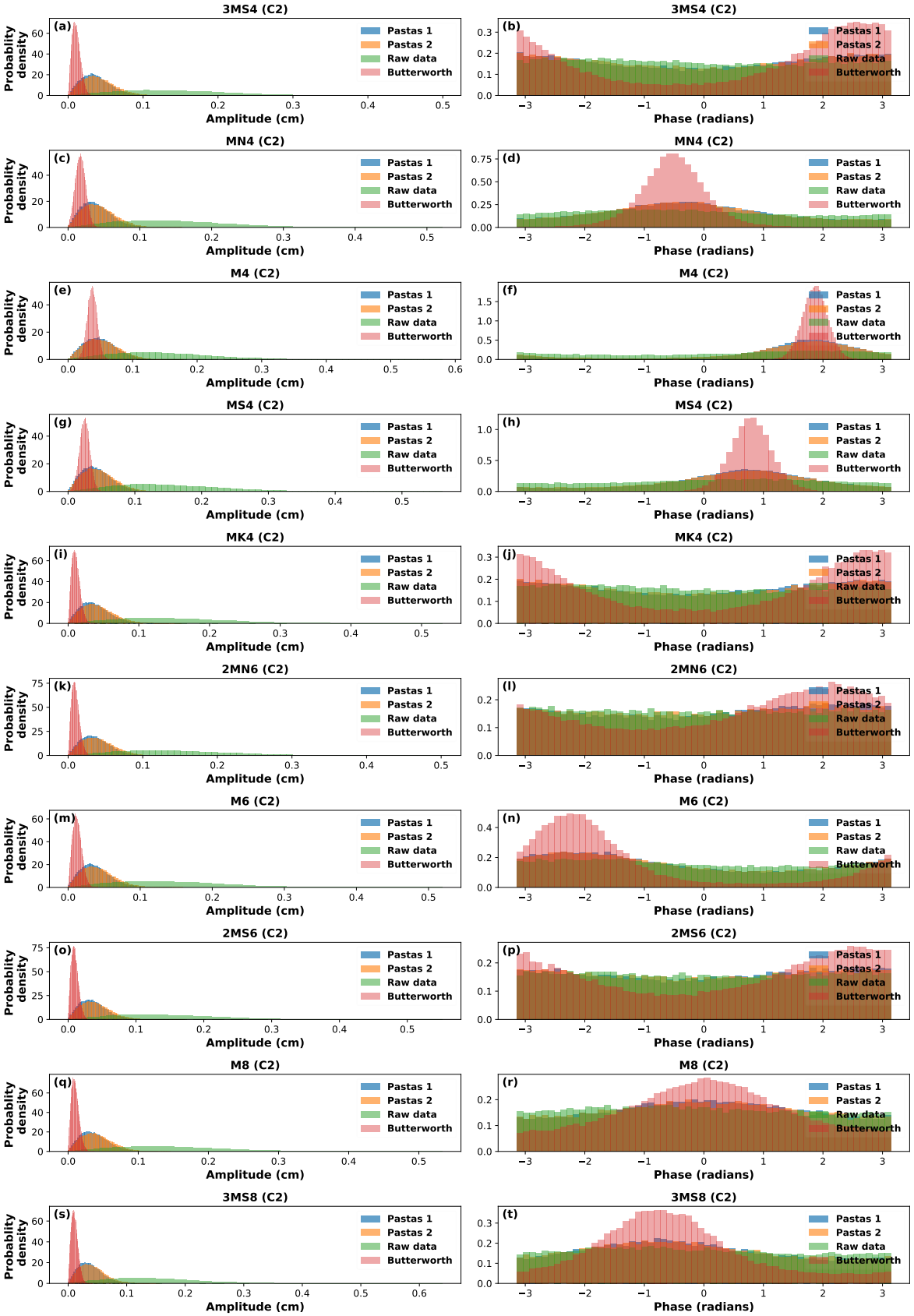


Figure C.31: Probability density graphs for a part (3/3) of the different constituents observed in well C2

C.9. Results obtained with the Butterworth filter for well A4

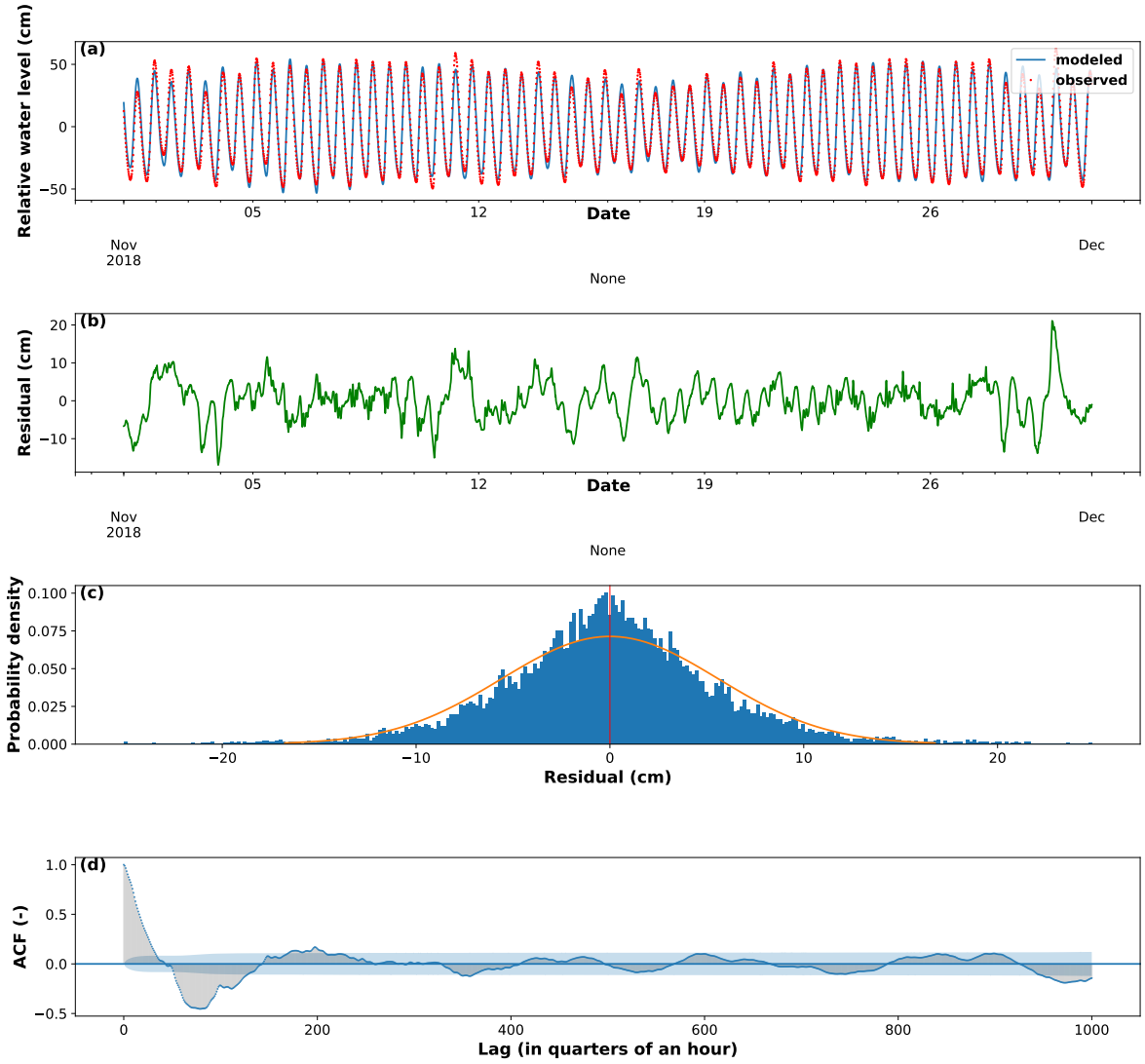


Figure C.32: Plots of the modeled and observed water fluctuations in the Oosterschelde (a) and the associated residual (b) are plotted for November 2018. With the lower Figure (c) showing the probability density of the residual.

C.10. Amplitude & phase plots against the distance to the Oosterschelde

Note that the constituents are plotted from low to high frequency.

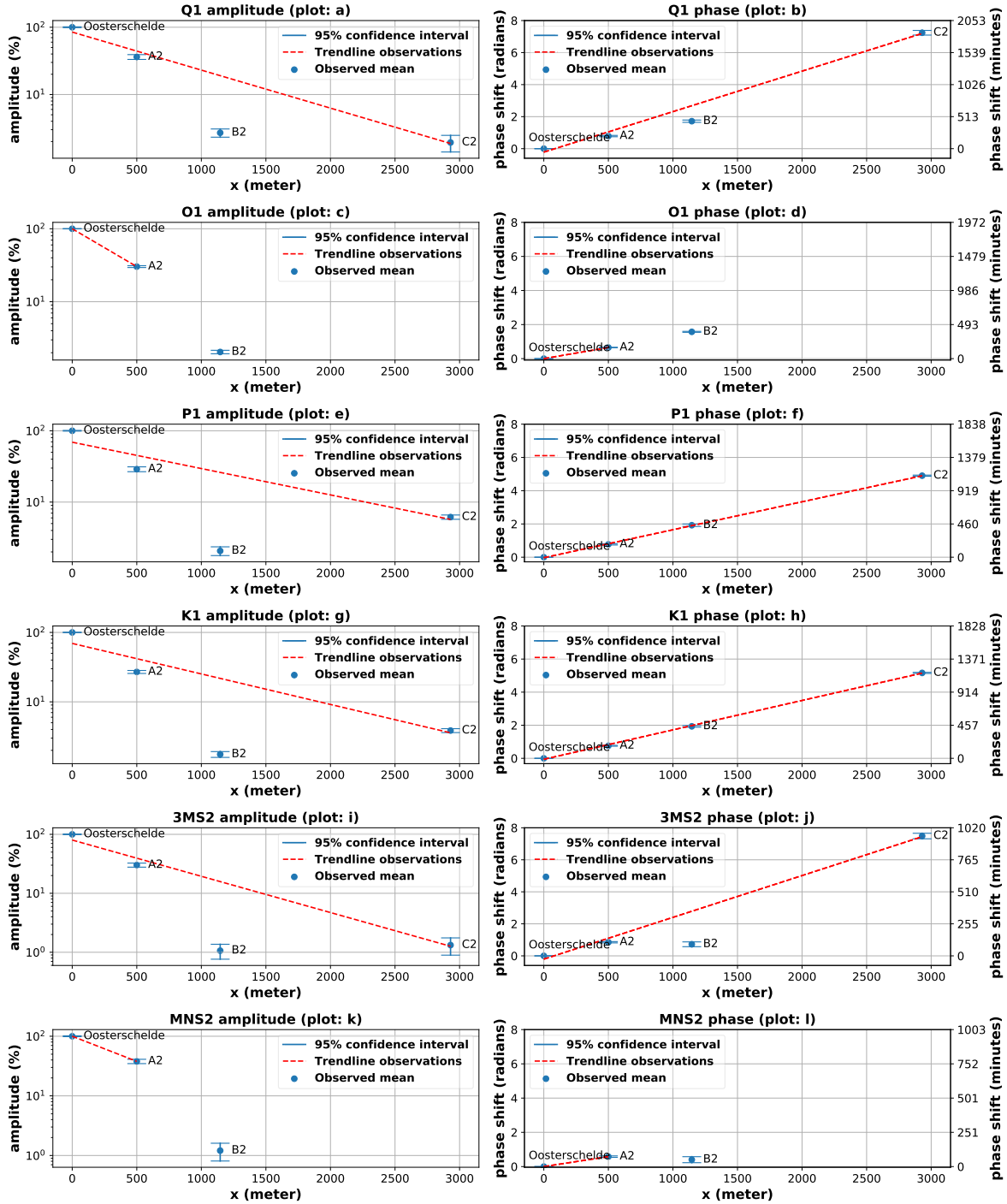


Figure C.33: Plots of the amplitude and phase shift against the distance to the Oosterschelde, for a part (1/5) of the different constituents observed in the observation wells. With the observed mean (blue dot), a trend-line through the observed means excluding B2 (red dash). Note that the observed phase is determined relative to the Oosterschelde and that a multiple of 2π was added to the observations if needed to obtain a logical phase shift.

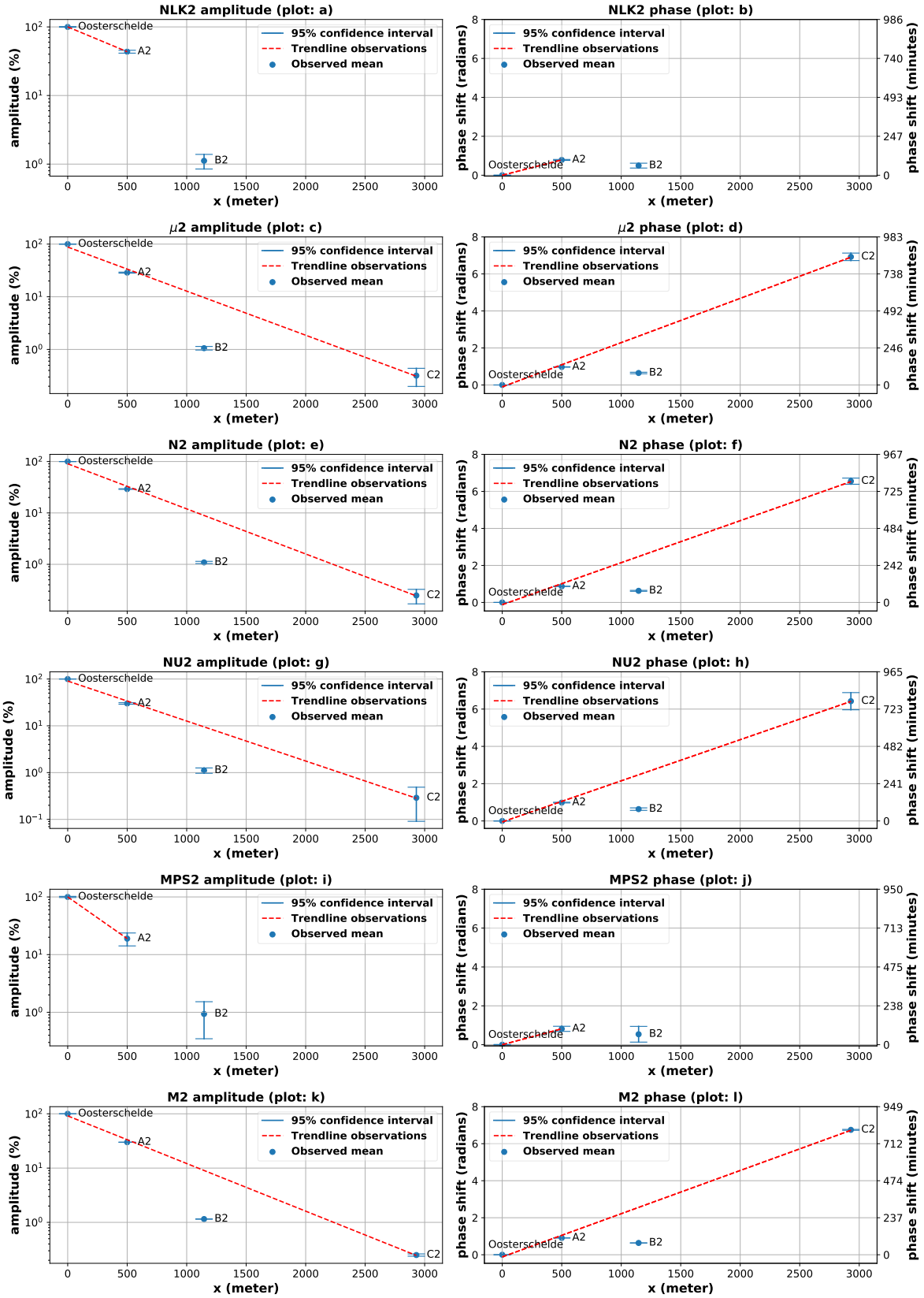


Figure C.34: Plots of the amplitude and phase shift against the distance to the Oosterschelde, for a part (2/5) of the different constituents observed in the observation wells. With the observed mean (blue dot), a trend-line through the observed means excluding B2 (red dash). Note that the observed phase is determined relative to the Oosterschelde and that a multiple of 2π was added to the observations if needed to obtain a logical phase shift.

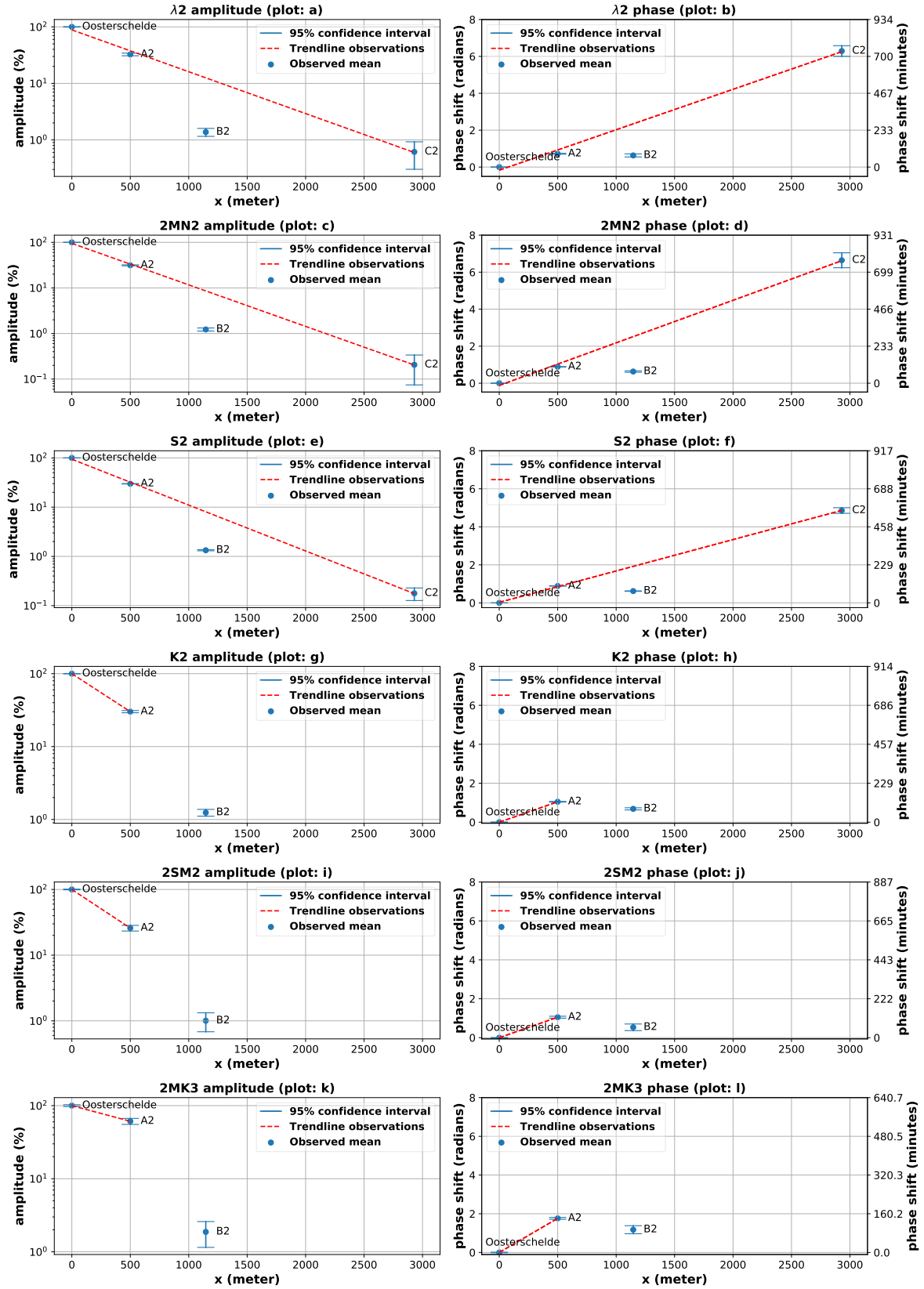


Figure C.35: Plots of the amplitude and phase shift against the distance to the Oosterschelde, for a part (3/5) of the different constituents observed in the observation wells. With the observed mean (blue dot), a trend-line through the observed means excluding B2 (red dash). Note that the observed phase is determined relative to the Oosterschelde and that a multiple of 2π was added to the observations if needed to obtain a logical phase shift.

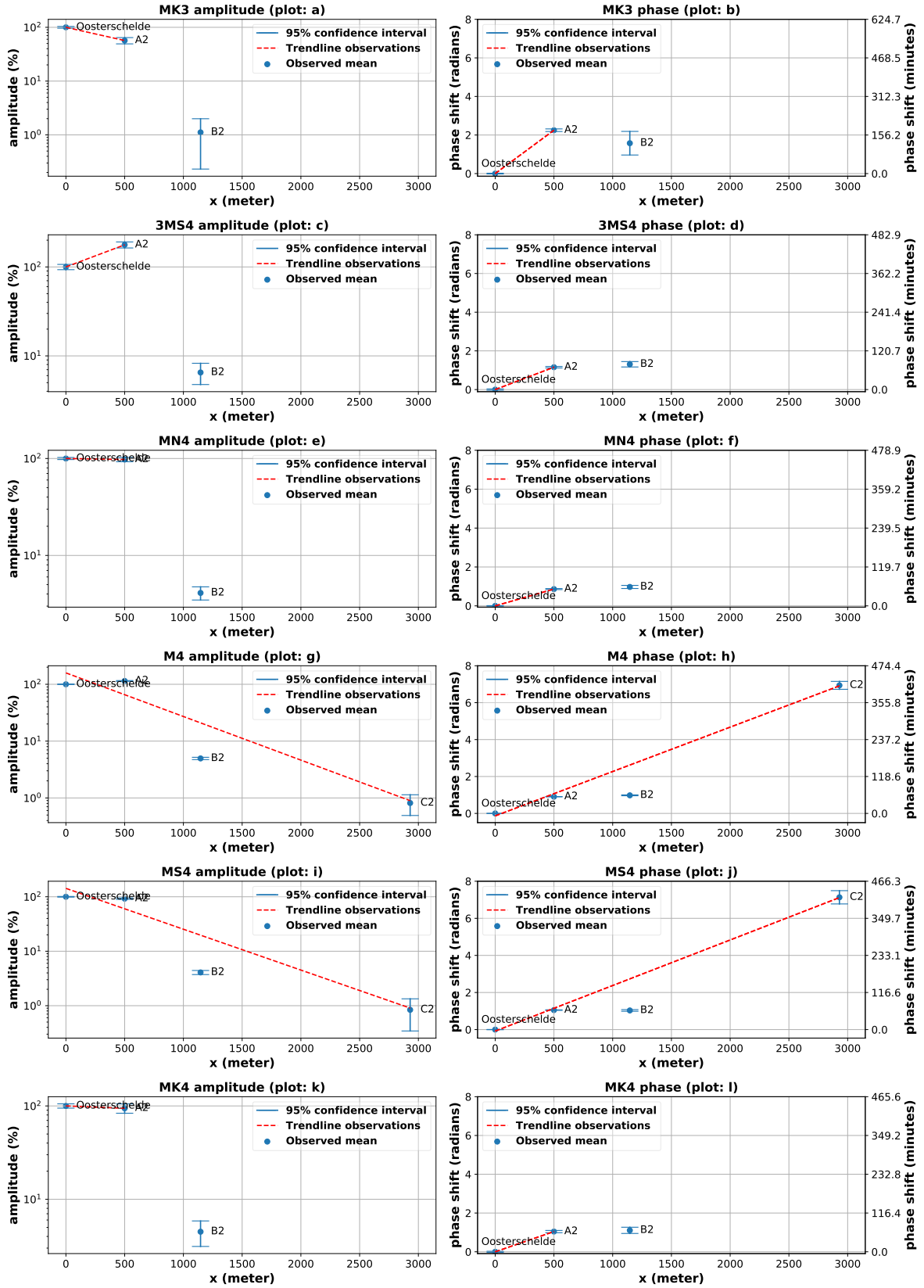


Figure C.36: Plots of the amplitude and phase shift against the distance to the Oosterschelde, for a part (4/5) of the different constituents observed in the observation wells. With the observed mean (blue dot), a trend-line through the observed means excluding B2 (red dash). Note that the observed phase is determined relative to the Oosterschelde and that a multiple of 2π was added to the observations if needed to obtain a logical phase shift.

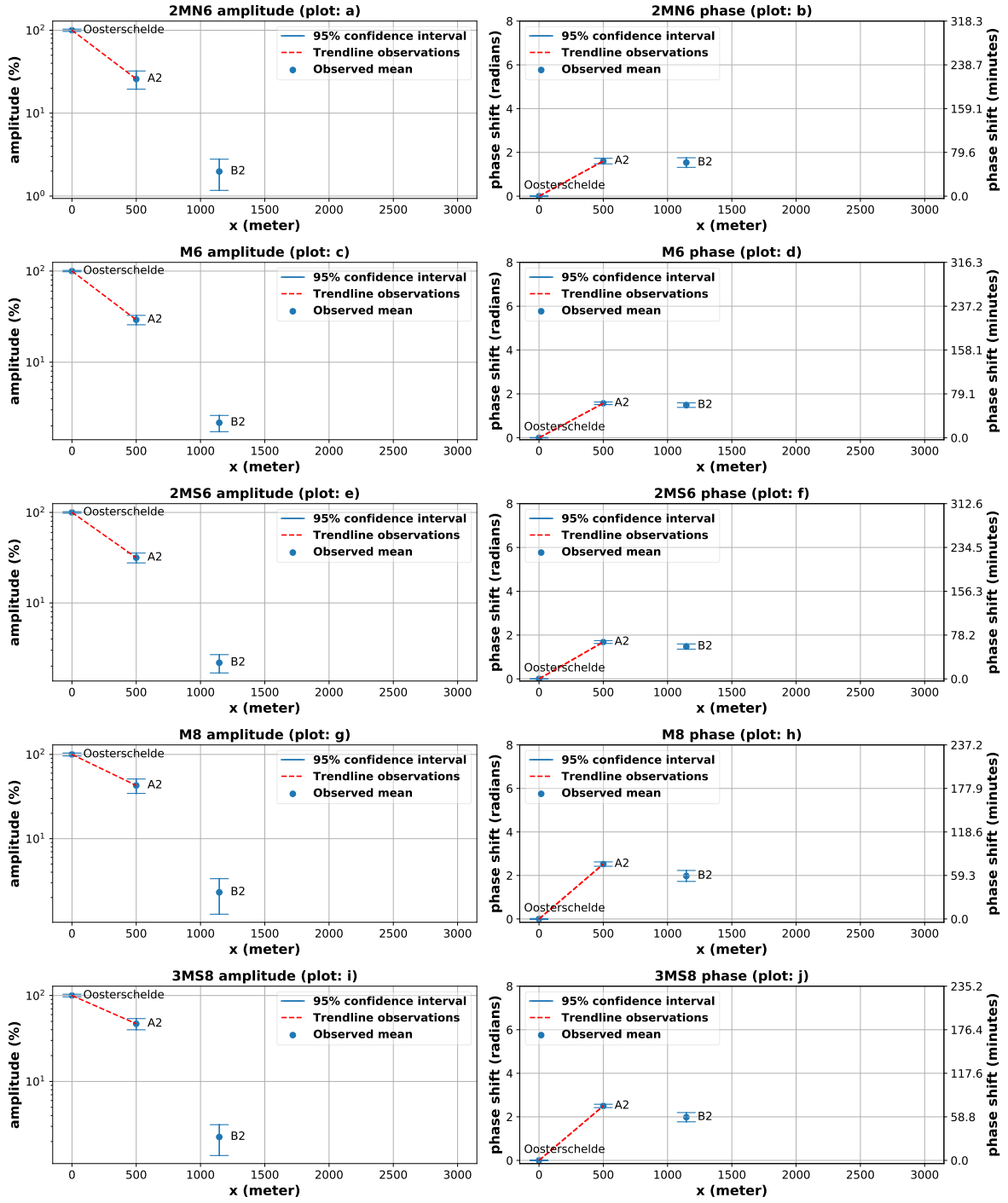


Figure C.37: Plots of the amplitude and phase shift against the distance to the Oosterschelde, for a part (5/5) of the different constituents observed in the observation wells. With the observed mean (blue dot), a trend-line through the observed means excluding B2 (red dash). Note that the observed phase is determined relative to the Oosterschelde and that a multiple of 2π was added to the observations if needed to obtain a logical phase shift.

C.11. Probability density graphs of the residual; without fixed x or y-axes

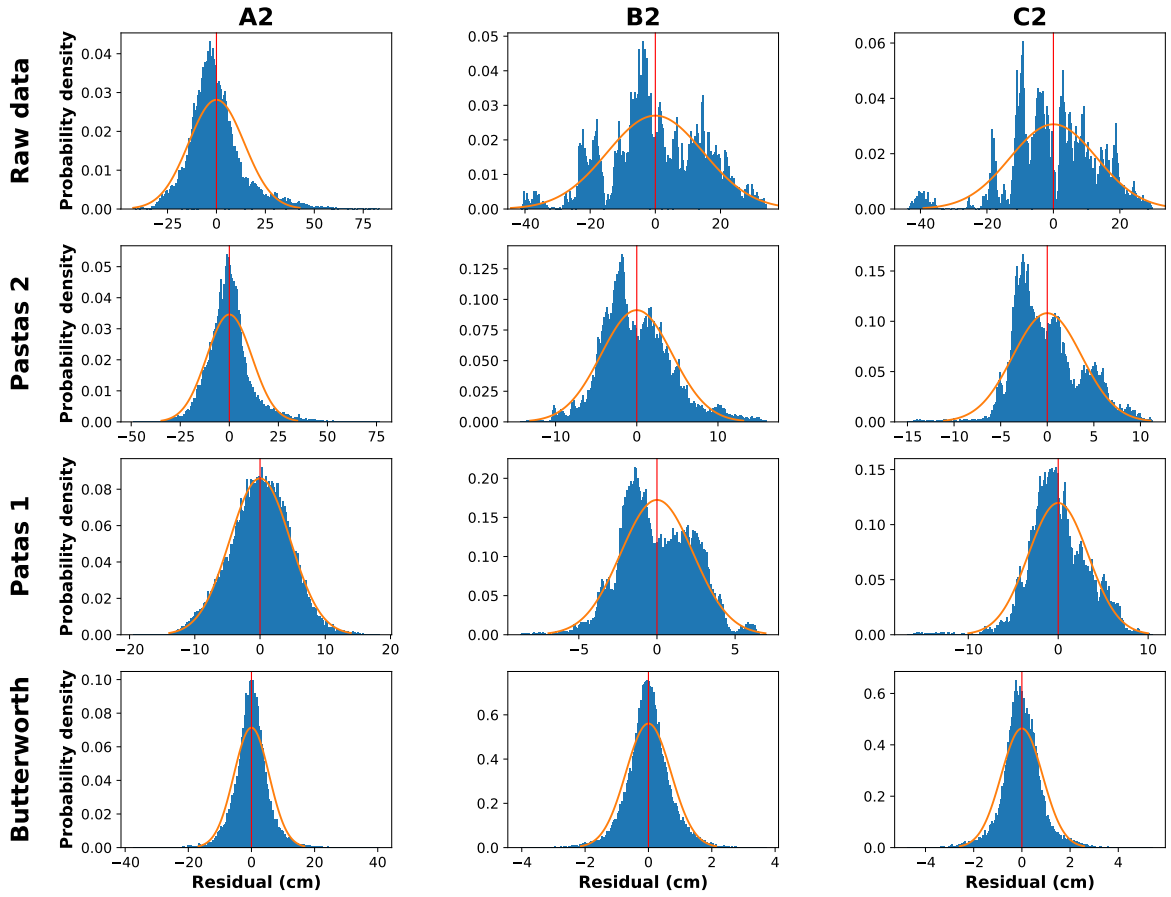


Figure C.38: Probability density graphs of the residuals (modeled minus observed groundwater level) for the groundwater time series A2, B2, C2 (columns) for various treatments of the time series (rows). The plots also include the mean (vertical red line) and the normal distribution (orange line) based on the mean and standard deviation of the data. Note that this plot is similar to Figure 4.6, but does not have fixed x or y-axes.