

Abstract

The quality of water level time series data strongly varies with periods of high and low quality sensor data. In this paper we are presenting the processing steps which were used to generate high quality water level data from water pressure measured at the Time Series Station (TSS) Spiekeroog. The TSS is positioned in a tidal inlet between the islands of Spiekeroog and Langeoog in the East Frisian Wadden Sea (southern North Sea). The processing steps will cover sensor drift, outlier identification, interpolation of data gaps and quality control. A central step is the removal of outliers. For this process an absolute threshold of 0.25 m/10 min was selected which still keeps the water level increase and decrease during extreme events as shown during the quality control process. A second important feature of data processing is the interpolation of gappy data which is accomplished with a high certainty of generating trustworthy data. Applying these methods a 10 years dataset of water level information at the TSS was processed and the results were submitted to WDC MARE data base system PANGAEA (<http://doi.pangaea.de/10.1594/PANGAEA.843740>).

1 Introduction

The Time Series Station (TSS) Spiekeroog measures different time series data in the East Frisian Wadden Sea. The station is positioned in the tidal inlet between the East Frisian Islands Spiekeroog and Langeoog (Fig. 1, left) and is measuring hydrographic, atmospheric and biological parameters since 2002 (Reuter et al., 2009) as a part of a long term study of biogeochemical processes of tidal flats (Rullkötter, 2009). The goal of the Time Series Station was to have a platform at a fixed position for measurements during the whole year. Especially autumn and winter months were of interest because of extreme events.

Time series are often analysed for hindcast of events and trends in the past (Visser et al., 1996; Gräwe and Burchard, 2012). Due to the long term aspect of these mea-

ESSDD

8, 345–364, 2015

Processing of water level derived from water pressure data

L. Holinde et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Processing of water level derived from water pressure data

L. Holinde et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



surements data sets are often affected by sensor degradation and maintenance resulting in time series combining periods of high and low quality of data. To assess the quality of the time series data and to correct low quality data sections different processing steps are needed. In this work we describe the processing of water pressure data recorded at the TSS Spiekeroog. The method emphasises the drift of sensor data and the handling of outliers and data gaps. At the end the quality of the data will be verified by a frequency and storm surge analyses.

2 Methods

The data processing method is divided into five steps. For these steps meta-information about the time series, measurement station and nearby measurements is required. These steps are:

1. Subtraction of a trend
2. Removal of outliers
3. Calculation of supporting points
4. Interpolation of missing data
5. Quality control of processed data

The data used for this paper were exported at an interval of 10 min for the time interval from 2003 to 2012. The sensor is mounted in a tube 1.5 m above the seafloor with other sensors to measure temperature and conductivity (Fig. 1, right).

2.1 Subtraction of a trend

At first the time series of the water level data is divided into different sections. The divisions depend on the maintenance of the pressure sensor. Figure 2 shows the water

pressure data (blue) and sensor maintenance (red) for the TSS Spiekeroog. Each section is tested for a trend and, when found, it will be subtracted. These trends can result from short term and long term changes. The short term changes are based on the tidal cycles and wind stress. The long term changes can be a result of sensor drift or sea level changes due to climate change.

For the analysis presented in this work the long term trend is not addressed. Therefore the subtraction of a trend for each section of the time series is possible resulting in a time series with a mean value of zero.

During this step the transition from pressure data to water level data will also be performed for which the Gibbs Sea Water (GSW) Oceanographic Toolbox (V3.01, 11.05.2011) for MATLAB (R2014b, The MathWorks) was used. The toolbox is based on the International Thermodynamic Equation of SeaWater – 2010 (TEOS-10, IOC, SCOR and IAPSO, 2010).

2.2 Removal of outliers

The next step of the data processing is to remove outliers. Outliers occur for example during sensor maintenance. To detect probable outliers the speed of water level change (gradient) between two adjacent data points is calculated. For this, the distribution of the gradients is presented in the histogram of Fig. 3. Most values (99.96 %) have a maximum absolute gradient of 0.25 m for a Δt of 10 min. Consequently, the probability is high that an absolute gradient exceeding 0.25 m/10 min is an outlier and should be removed. In addition, each year is scanned visually in search for outliers missed by the gradient method.

2.3 Calculation of supporting points

The removed outliers and gaps of the original time series represent missing information making it more difficult to interpret the dataset especially since many methods for the analysis of time series require evenly sampled data (Karl et al., 1982). To fill these gaps

Processing of water level derived from water pressure data

L. Holinde et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



short term rises in the water level data. The German Hydrographic Institute (BSH, “Bundesamt für Schifffahrt und Hydrographie”) has published values for different magnitudes of storm floods at the German North Sea Coast (BSH, 2015). These values classify a weak storm flood with water levels between 1.5 and 2.5 m above mean high water, a severe storm flood with water levels between 2.5 and 3.5 m and a very severe storm flood with water levels more than 3.5 m above mean high water.

3 Results

The different processing steps were applied to the whole time series presented in Fig. 2. During the processing of the first three steps it became obvious that greater amounts of data and/or comparison measurements are missing in the first two years (2003 and 2004) and at the end of 2012. Following these observations these three years were excluded for further data processing in this work.

Exemplary the results of the first four processing steps for data from 2008 are shown in Fig. 4. For the quality control the fully processed data set (2005–2011) will be used.

3.1 Data processing

Figure 4 presents the time series at different stages of the data processing for a time period from 1 January 2008 to 31 December 2008. The top graph shows the measured pressure data before the validation (blue) and pressure sensor maintenance (red). At the second vertical red line of the graph a shift in the water level is apparent and after the third a gap in the time series.

The middle graph in Fig. 4 shows the time series after the subtraction of a linear trend and the removal of outliers. Two important details have to be mentioned about this graph. First, after the removal of the trend the whole time series has a mean value of zero. That means that afore mentioned shifts in the top graph were removed from the time series. Second, some outliers which also occurred near the shift have also

Processing of water level derived from water pressure data

L. Holinde et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



4 Discussion

4.1 Subtraction of a trend

In the first processing step a piecewise linear trend was subtracted. While it is easier to identify outliers by this way, it is not possible to analyse the dataset for long term sea level changes. The manual of the utilized pressure sensors states that the sensor degradation is $\pm 0.1 \text{ \% year}^{-1}$ of a maximum of 30 dBar. In winter 2008/09 the water level date evinced a decrease by 0.8 dBar in 8 months. By contrast, the “Niedersächsische Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz” (NLWKN) has claimed that the water level is increasing (NLWKN, 2006). A possible reason for the decreasing values of the water level is bio fouling which is especially severe in spring and summer.

4.2 Removal of outliers

145 outliers were removed which represent only 0.04 % of the measured 368 064 values. A larger portion of data points (12 253, 3.33 %) is missing due to gaps in the observations. These gaps can happen during power failure or computer problems at the TSS. The comparatively small number of outliers implies that the measurement system is very robust. But the use of a gradient threshold for outlier detection can also lead to the removal of correct data if the water level is indeed increasing very fast. A comparison between Fig. 4 and the results of the storm flood analysis (Fig. 7 and Table 1) shows that no values during extreme events were removed. In addition, Fig. 3 shows that about 99.96 % of all values have an absolute gradient lower than 0.25 m in 10 min. All of these pieces of information indicate that the selection of a gradient threshold of 0.025 m min^{-1} is an acceptable choice for this study.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



4.3 Calculation of supporting points

A comparison between the time series data from Spiekeroog and Neuharlingersiel reveals that both are in good agreement. Constant values below -2 m at Neuharlingersiel are explained by the sensor falling dry during low water. If one of these values would be taken as a support point then this could lead to an overestimation of the interpolated water level data. The cross correlation analysis reveals a 20 min time lag between the Time Series Station Spiekeroog and Neuharlingersiel since the tidal wave which moves through the North Sea reaches the station first. A comparison with the local tidal chart (BSH, 2011) shows a delay between Spiekeroog and Neuharlingersiel between 1 and 5 min. The discrepancy could be a result of the 10 min sampling rate leading to possible errors of $\pm 5\text{ min}$.

4.4 Interpolation of missing data

Sturges (1983) suggests that it is still possible to interpolate data with gaps as long as $1/3$ of the times series and a total number of missing values as high as half the time series. In this time series only 3.37% of the data are missing. Using the supporting points it was possible to interpolate the missing data also for longer gaps ($> 7\text{ h}$) with reasonable results. This can be observed by the identification of no additional extreme events at the TSS and the similarities between the measured and interpolated data (Fig. 5).

4.5 Data quality

The Fourier analysis has shown two pronounced differences between the original and processed data. The first is the high peak at the beginning of the curve with a frequency of 0 representing the mean of the original data. The second difference is the strength of the tidal peaks. For the original time series data these peaks are not as distinct as

Processing of water level derived from water pressure data

L. Holinde et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



for the processed data. These differences result from the previous processing steps and illustrate their usefulness.

Discrepancies in the storm flood events between the TSS Spiekeroog and Neuharlingersiel originate from the different positions of the measurement sites. During storm floods the wind blew from northern or western directions forcing the water into the harbour of Neuharlingersiel where it can accumulate and leads to increased water levels. The water level is also increasing at the measurement station but here it is not possible for the water to pile up resulting in lower water levels than at Neuharlingersiel.

5 Conclusions

Processing water level data resulted in a relevant long term dataset. Here it should be emphasised that:

- From 99.96 % of the time series data only a linear trend was subtracted which leads to no difference in the ratio between the values.
- During the removal of outliers no values during extreme events were deleted. This was concluded derived from a comparison of storm floods between the Time Series Station Spiekeroog and Neuharlingersiel.
- The calculation of supporting points has yielded a 20 min time lag between the TSS Spiekeroog and Neuharlingersiel. This is greater than comparable data from the BSH and needs further analysis.
- The interpolated data follows the same trend as the measured data at the TSS Spiekeroog.
- A spectral analysis has shown that all major tidal frequencies can be found. Also during the storm flood analysis six events were found for the Time Series Station and eight for Neuharlingersiel. The difference comes from the different positions of the measurement stations.

Processing of water level derived from water pressure data

L. Holinde et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Processing of water level derived from water pressure data

L. Holinde et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



All of the above shows that the described processing steps achieve good results if one accepts certain prerequisites. The first is that it will not be possible to analyse the water level time series with respect to changing sea levels because of climate changes due the removal of a trend. The second is that although the time series is now without gaps the interpolated data can be wrong. A possibility to circumvent part of these prerequisites would be to install another device for the measurement of the water level. This device should then use a different method for measuring sea level changes e.g. based on radar signals from an above water installation.

6 Data Provenance and Structure

The data used in this paper is available at the WDC MARE data base system PAN-GAEA (<http://doi.pangaea.de/10.1594/PANGAEA.843740>). It was divided into yearly sections. Each section provides the original water pressure data (L1), the water level after the subtraction of a trend and removal of outliers (L2) and the final product of this paper the water level data where gaps were interpolated (L3).

Acknowledgements. Special thanks to Axel Braun and Waldemar Siewert for the maintenance on the station during the whole year and all weather conditions as well as for Helmo Nikolai for servicing the institute boats. The authors are grateful to Rainer Reuter for the concept and construction of the Time Series Station Spiekeroog.

Part of the data was acquired during the project BioGeoChemistry of Tidal Flats (DFG, grant no. FOR 432).

References

Badewien, T. H., Zimmer, E., Bartholoma, A., and Reuter, R.: Towards continuous long-term measurements of suspended particulate matter (SPM) in turbid coastal waters, *Ocean Dynam.*, 59, 227–238, doi:10.1007/s10236-009-0183-8, 2009. 358

Processing of water level derived from water pressure data

L. Holinde et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



BSH: Gezeitenkalender 2011, Hoch- und Niedrigwasserzeiten für die Deutsche Bucht und deren Flussgebiete, Hamburg, Germany, 2011. 353

BSH: Informationen zu Sturmfluten, available at: <http://www.bsh.de/de/Meeresdaten/Vorhersagen/Sturmfluten/> (last access: 20 January 2015), 2015. 350

5 Gräwe, U. and Burchard, H.: Storm surges in the Western Baltic Sea: the present and a possible future, *Clim. Dynam.*, 39, 165–183, 2012. 346

IOC, SCOR and IAPSO: The international thermodynamic equation of seawater – 2010: Calculation and use of thermodynamic properties., Report, Unesco, Paris, France, 2010. 348

10 Karl, T. R., Koscielny, A. J., and Diaz, H. F.: Potential errors in the application of principal component (eigenvector) analysis to geophysical data, *J. Appl. Meteorol.*, 21, 1183–1186, doi:10.1175/1520-0450(1982)021<1183:PEITAO>2.0.CO;2, 1982. 348

NLWKN: Jahresbericht 2005 des NLWKN, Report, Niedersächsischer Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz, Norden, Germany, 2006. 352

15 Reuter, R., Badewien, T. H., Bartholomä, A., Braun, A., Lübben, A., and Rullkötter, J.: A hydrographic time series station in the Wadden Sea (southern North Sea), *Ocean Dynam.*, 59, 195–211, doi:10.1007/s10236-009-0196-3, 2009. 346

Rullkötter, J.: The back-barrier tidal flats in the southern North Sea a multidisciplinary approach to reveal the main driving forces shaping the system, *Ocean Dynam.*, 59, 157–165, doi:10.1007/s10236-009-0197-2, 2009. 346

20 Stanev, E. V., Al-Nadhairi, R., Staneva, J., Schulz-Stellenfleth, J., and Valle-Levinson, A.: Tidal wave transformations in the German Bight, *Ocean Dynam.*, 64, 951–968, doi:10.1007/s10236-014-0733-6, 2014. 351

Sturges, W.: On interpolating gappy records for time-series analysis, *J. Geophys. Res.-Oceans*, 88, 9736–9740, doi:10.1029/Jc088ic14p09736, 1983. 353

25 Visser, M., Batten, S., Becker, G., Bot, P., Colijn, F., Damm, P., Danielssen, D., Eynde, D. V. A. D., Føyen, L., Frohse, A., Groeneveld, G., Laane, R., Raaphorst, W. V., Badach, G., Schultz, H., and Sundermann, J.: Time series analysis of monthly mean data of temperature, salinity, nutrients, suspended matter, phyto- and zooplankton at eight locations on the Northwest European Shelf, *Deutsche Hydrografische Zeitschrift*, 48, 299–323, 1996. 346

Processing of water level derived from water pressure data

L. Holinde et al.

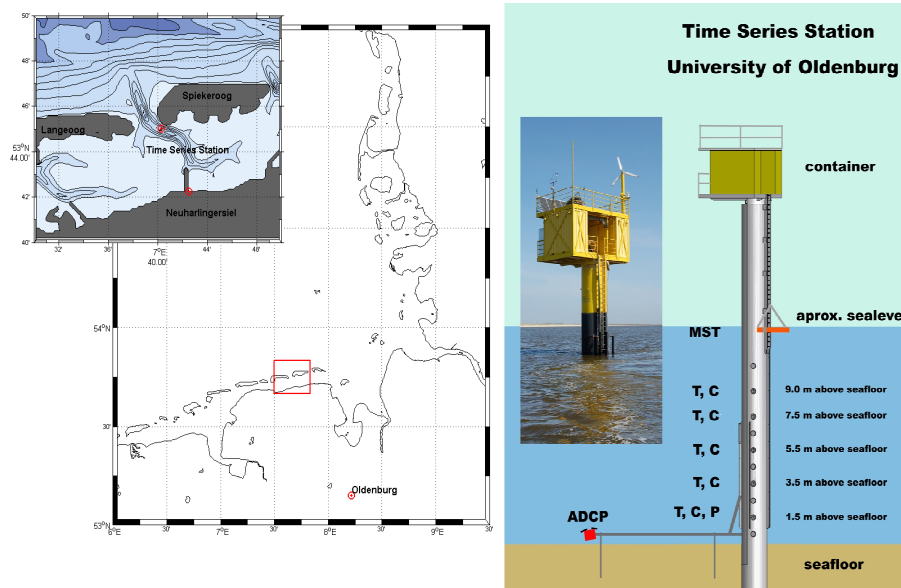


Figure 1. Left: Time Series Station Spiekeroog in the tidal inlet between the East Frisian Islands Spiekeroog and Langeoog. Right: Schematic of the Time Series Station Spiekeroog with attached sensors, T: temperature sensor, C: conductivity sensor, P: pressure sensor, ADCP: acoustic doppler current profiler, MST: multispectral transmissiometer (Badewien et al., 2009).

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



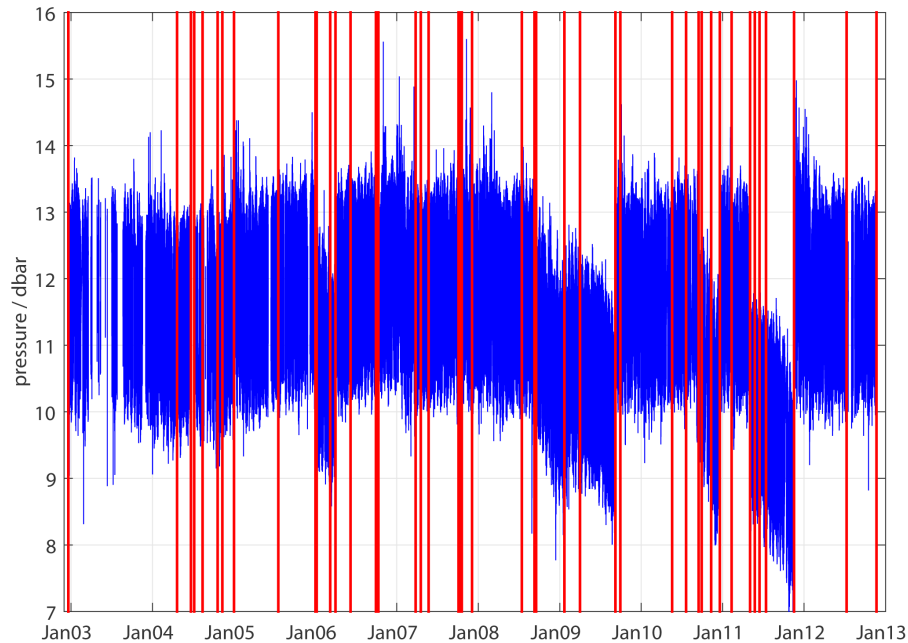


Figure 2. Measured water pressure data (blue) at the Time Series Station Spiekerroog before the validation and times of sensor maintenance (red vertical lines).

Processing of water level derived from water pressure data

L. Holinde et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Processing of water level derived from water pressure data

L. Holinde et al.

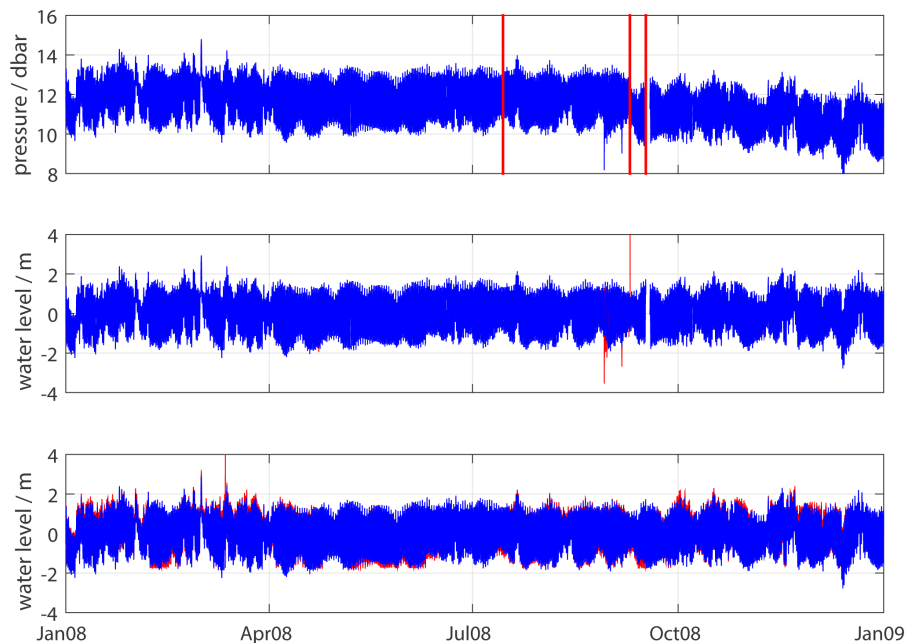


Figure 4. The subfigures show the validation process for the year 2008. Top: The time series before the validation (blue) and sensor maintenance (red vertical lines) are shown. Middle: The time series (blue) and outliers (red) after the first two steps of the validation. Bottom: The time series after the interpolation (blue) and the comparison data of Neuharlingersiel (red) are shown.

[Title Page](#)[Abstract](#)[Instruments](#)[Data Provenance & Structure](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

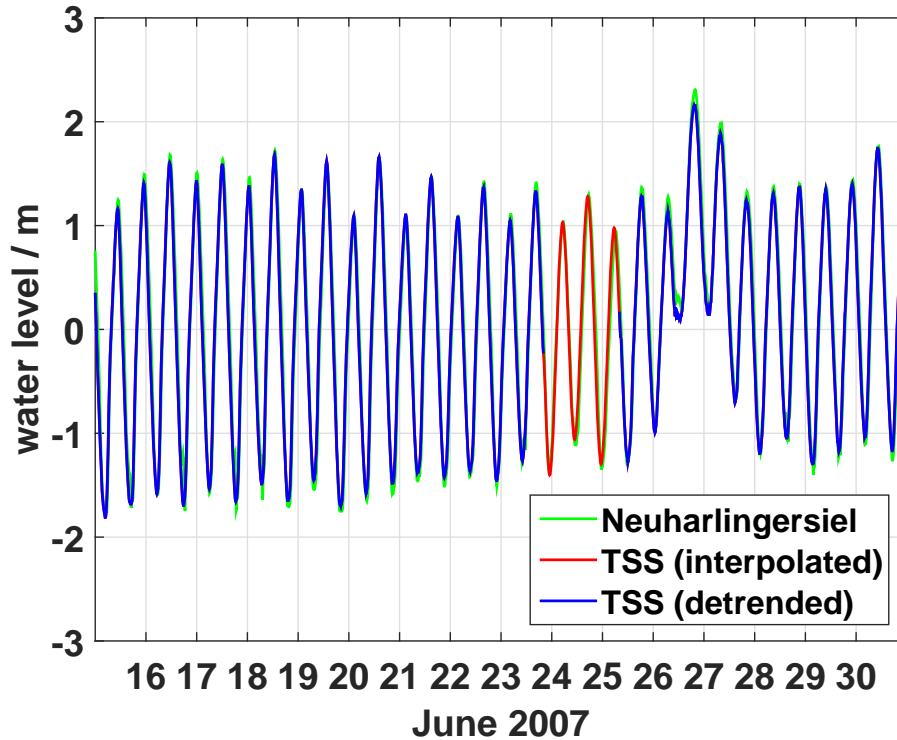


Figure 5. All three curves show data for two weeks in June 2007. The blue line shows the data after the removal of a trend and outliers and the red curve the interpolated data at the TSS Spiekeroog. The green curve shows data from Neuharlingersiel.

Processing of water level derived from water pressure data

L. Holinde et al.

Title Page

Abstract Instruments

Data Provenance & Structure

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Processing of water level derived from water pressure data

L. Holinde et al.

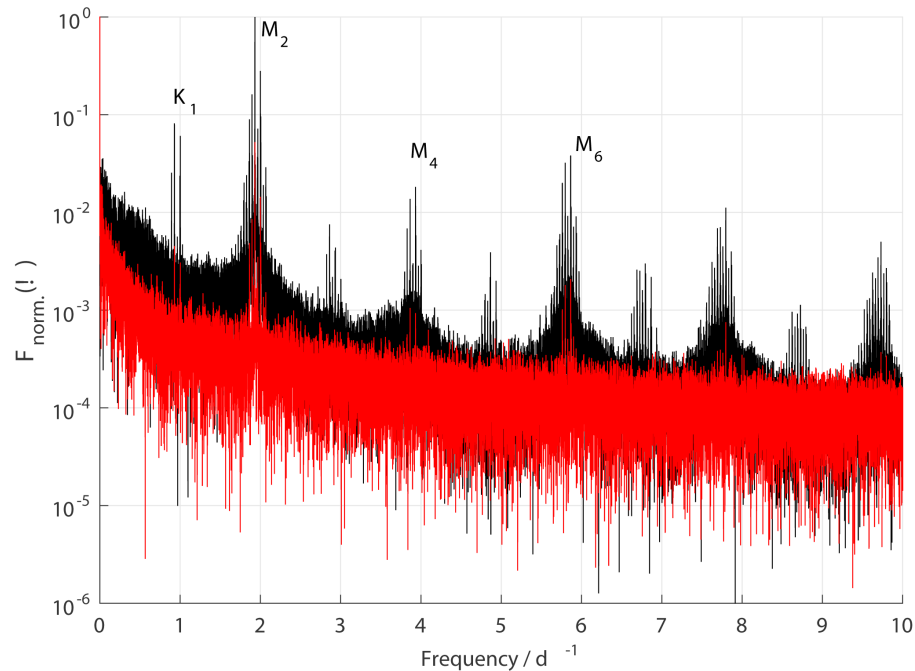


Figure 6. Fast Fourier Transformation (FFT) of the original (red) and processed (black) water level at Time Series Station Spiekeroog. K_1 : Lunar diurnal constituent; M_2 : principal lunar semidiurnal constituent; M_4 and M_6 : shallow water overtide of principal lunar semidiurnal constituent.

[Title Page](#)[Abstract](#)[Instruments](#)[Data Provenance & Structure](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

