Assessment of sea-level variations along the Dutch coastline

J.T. Arens Supervision: R.E.M. Riva
Delft University of Technology - Dept. Geoscience and Remote Sensing

June 2020

Abstract

Long-term sea level change and its spatial and temporal variability measured with tide gauge stations along the Dutch coast have been studied. This study investigates how time series length and modeling choices influence the adoption of a quadratic over a linear sea level model. We apply linear and quadratic models to corrected tide gauges, for differing model start years. Longer models show more consistent results between stations, with less inter-station variability and smaller uncertainties. This improvement of consistency diminishes when using time series longer than 40 years. We find indications of a break-point in trends in the period 1978-1998. Quadratic models result in minor but relevant acceleration for longer time series, but do not perform sufficiently for time series shorter than 20 years. Comparing model quality between linear and quadratic models generally indicate better performance of quadratic models, but results are not conclusive to justify model adoption. A station mean is less conclusive for quadratic models than for linear models and sensitive to choice of stations and model length.
Keywords *mean sea level variability, sea level change, sea level acceleration, tide gauge records, Dutch coast*

1 Introduction

All coastal nations will be affected by climate change through the means of sea level change. In the case of the Netherlands, the struggle against the sea is already encapsulated in it's name, and roughly half the country's flat topography is less than 1 meter above the current sea level. Big infrastructural projects, such as the Delta Works, have been protecting the land from flooding. It is a continuous effort, mainly through the means of sand suppletion and dike reinforcement, to maintain the current Dutch coastline. As these projects are preventative and not reactive, long-term infrastructure planning is necessary. Therefore, it is crucial for policy makers to know what sea level change is to be expected, in order to prepare and adapt the coastal defense strategy.

1.1 Climate scenarios

In order to prepare for future climate, the Dutch meteorological office (KNMI) has provided a climate scenario for the Netherlands in 2014 [1]. This so-called KNMI'2014 scenario can be seen as a translation of the Fifth Assessment Report (AR5) of the IPCC [2] to the regional effect that are to be expected for the Dutch mainland. These include projections of the sea level change and are often used as a basis for policy makers. With the IPCC's Sixth Assessment Report (AR6) being worked on and expected to release 2022, the KNMI is working on new scenarios, with indicative sea level change numbers to be expected in 2021 and a new set of climate scenario's in 2023.

1.2 Zeespiegelmonitor

The Dutch governmental body Rijkswaterstaat, which is tasked with the coastal defense of the country, contracted Deltares to investigate what the current state of the sea level in the Netherlands is and how this compares to the KNMI'14 scenario's. This is done using tide gauge data in the 'Zeespiegelmonitor' (sea level monitor), of which the latest report dates from 2018[3]. In the Zeespiegelmonitor, a mean of six Dutch tide gauge stations' data is analyzed for a time period of 1890 - 2018. Two different models are compared against the linear model. A first model is the so-called broken linear model, which adds an additional linear trend to the model, which starts in 1993. The second model is a quadratic model. Both models have been found to perform a better fit to the mean data than the linear model, but not significant. Therefore, a main conclusion of the Zeespiegelmonitor is that there is currently no acceleration in the sea level change for the Dutch coast. These results contradict the KNMI'2014 scenarios and studies in global sea level change[4]. This study aims to provide a more thorough look into the causes as to why no acceleration is visible in the Dutch tide gauge data and how modeling choices might influence the result.

1.3 Research goals

Tide gauge stations along the Dutch coastline have relatively long record lengths, which allows for extensive analysis. Modeling choices have to be made when investigating the long tide gauge records. Which parts of the record are included and which parts are omitted? For minimizing uncertainties, it is generally better to use long time series. However, if there would be a break-point in the sea level, in which non-linear effects would only start playing a role after a specific year, it might not be possible to find a significant quadratic model which fits the complete time series. Certain corrections, such as for example the wind correction used in the Zeespiegelmonitor, are not available for the same length as tide gauge records. Correction by means of a mean regression value for the missing years can be performed, but might increase uncertainties. Variability between stations might also result in a model applied to a station mean not accurately representing what is actually happening at any of the stations individually.

This project is meant to investigate the tide gauge time-series along the Dutch coast and to explore what uncertainty exists in these time series and derived trends. The goal is to use simple models to investigate what might cause this discrepancy between tide gauge analysis and the KNMI'2014 scenarios. The main research questions we aim to answer are listed below.

Time series length

- How does the usage of differing time series length influence found trends and confidence intervals?
- Is it possible to find a more suitable balance between time series length and trend significance?
- Can possible break points be indicated by analyzing trends found on differing time series lengths?

Inter-station variability

- What variability is there between different stations along the Dutch coast?
- Based on the variability, what can be concluded about the use of a station mean to summarize national effects?

Model adoption

- Can the results of the Zeespiegelmonitor be reproduced to asses influences of choices on model adoption?
- Could shorter time series allow for different sea level change models (accelerating) to become feasible?
- What will acceleration model uncertainties show concerning the adoption of a linear model over a quadratic model?

2 Data and methods

2.1 Datasets

Monthly mean tide gauge data has been retrieved from the Permanent Service for Mean Sea Level[5],[6]. This service collects tide gauge and bottom pressure data from all over the world. The local vertical datum on which tide gauges operate is oftentimes updated or discontinued, which brings trouble for time series analysis due to height jumps when a datum is changed. The PSMSL solves this issue through the mean of a 'Revised Local Reference (RLR) adjustment, which defines a vertical datum 7m below mean sea level, to ensure all monthly mean sea level readings are positive. This also makes it easier to correct for datum shifts. RLR data from PSMSL contain these corrections, resulting in time series with a continuous datum.

All nine PSMSL stations along the Dutch coastline have been included in this study. Furthermore, we included two stations along the Belgian coast and one on the German island of Borkum, all three of which are relatively close to the Netherlands and therefore might also provide some information about parts of the Dutch coastline. The geographical locations are shown in Figure 1. The latest available data from PSMSL is December 2018 for the Dutch stations, December 2017 for the Belgian stations of Oostende and Zeebrugge and December 2016 for the German station of Borkum Fischerbalje. We refer to these latest available months as 'present'. Tide gauge time series are used from January 1948 when available, which corresponds with used wind and sea level pressure models. Exceptions are Zeebrugge (Jan 1961), Roompot-Buiten (Jan 1987) and Borkum Fischerbalje (Jan 1963), due to absence of RLR data before these dates or in the case of Zeebrugge a data gap between 1942 and 1961. For all PSMSL data, the monthly mean is represented by a value in the middle of each month.

The tidal signal in the tide gauges is mostly filtered out of the analysis automatically, due to the harmonic nature of the tides. The use of a monthly mean sea level time series results in a negated effect from higher-frequency tidal effects such as semidiurnal (high- and low) tides, daily inequalities and spring-neap tidal cycles. Comparing only differences between trends based on full-year model length differences also filters out any semi-annual tidal effects. This means that the only significant tidal effect that will affect the analysis will be the nodal cycle, which has a period of roughly 18.6 years. The nodal cycle is caused by the nodal precession of the moon's orbit around the Earth. The moon's orbit is declined by an average of 23.5°, but this declination varies with ±5° in a harmonic fashion, with a period of 18.6 years. The effect of this variation in declination is called the nodal cycle and its amplitude depends on the location of the tide gauge. We assume the amplitude of the nodal cycle to be the self-consistent equilibrium tide, following the assumption given by Woodworth, 2012[7], that this is the best way to deal with the nodal tide. The equilibrium tide amplitudes are provided by Thomas Frederikse[8]. This dataset refers to the amplitude during a nodal minimum in 1922.

Tide gauge data has been corrected for wind stress and sea level pressure using the Reanalysis data products from the United States' National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR), which are provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, from their website. In this study, the near-surface (.995 sigma level) zonal (U-wind) and meridional (V-wind) monthly mean winds from the NCEP/NCAR Reanalysis 1 have been used, together with the sea level pressure. The NCEP/NCAR Reanalysis 1 ingests atmospheric data from a wide range of sources, all over the globe, using an unchanged data assimilation system[9]. This results in a consistent climate analysis, as all data is processed using the same assimilation system. Currently, the temporal coverage of the reanalysis is from January 1948 to present, on a 2.5°lat by 2.5°lon global grid. Different reanalysis products which allow for a wider time window, for example the NOAA-CIRES-DOE Twentieth Century Reanalysis (V3)[10], which dates back to 1836, before the availability of any RLR tide gauge data. However, this product is available only until 2015 and includes a change in sea surface temperature boundary conditions in 1980-1981. Combining reanalysis products to create a continuous parametrization of wind speed and sea level pressure is possible, but using the NCAR/NCEP Reanalysis is sufficient for our purpose, as preliminary time series analysis showed that found linear or quadratic trends by including data before 1948 do not vary a lot from trends computed from 1948 to present.

2.2 Tide gauge correction

We correct the tide gauge data for influences of the nodal cycle, wind stresses and sea level pressure, as these factors all introduce variability to the monthly mean sea level signal. By correcting the tide gauge data for these factors, more accurate trend modeling can be performed on the relative sea level. An example of this reconstruction is shown in Figure 2 for the station of Vlissingen. First of all, the tide gauge signal is corrected for the equilibrium tide of the nodal cycle. As the data refers to the amplitude of the nodal cycle during a nodal minimum in 1922, the height of the nodal tide at any given time $\eta_{nodal}(t)$ is computed using the amplitude A in mm, t in decimal years and a sinusoidal function. The nodal tide height is then simply subtracted from the tide gauge height η_{tg} from the time series.

$$\eta_{nodal}(t) = -A * cos(\frac{2\pi}{18.612958}(t - 1922.7)) \quad (1)$$

$$\eta_1(t) = \eta_{tg}(t) - \eta_{nodal}(t) \quad (2)$$

Further correction occurs through linear regression with the resulting, nodal-corrected, MSL and the zonal (τ_u) & meridional (τ_v) wind stresses and sea level pressure. In order to quantify how much each component of the correction adds to a reduction of variance of the observed tide gauge height, a sequential regression & correction method has been employed, in which the most significant components are removed from the signal before performing

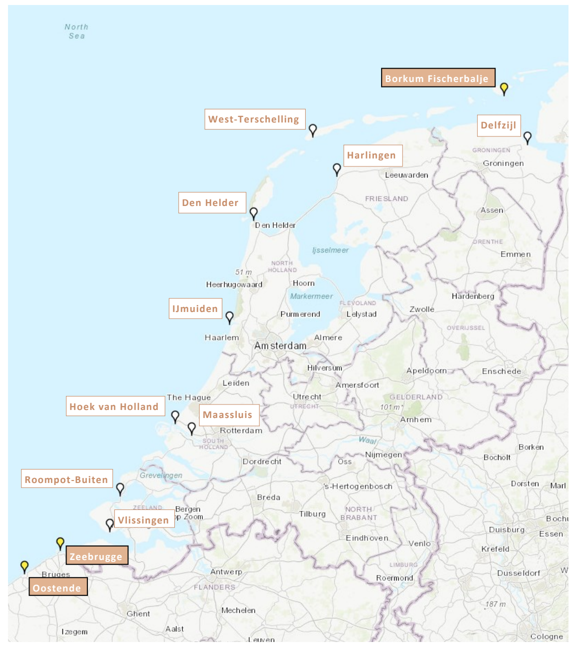


Figure 1: Locations of the investigated tide gauge stations, as retrieved from the PSMSL, April 28 2020. White pins with white boxes indicate the Dutch stations, which have been used to compute the tide gauge mean. Yellow pins with brown boxes indicate the Belgian and German stations close to the Dutch border. From South to North along the coastline: Oostende (BE), Zeebrugge (BE), Vlissingen, Roompot-Buiten, Maassluis, Hoek van Holland, IJmuiden, Den Helder, West-Terschelling, Harlingen, Delfzijl & Borkum Fischerbalje (DE). Background map and location pins retrieved from PSMSL Catalogue Viewer.

regression on the next component. This methods results in a correction for zonal wind stress first and foremost. The difference between meridional wind stress and sea level pressure influences are minimal, but on average meridional wind stresses have a bigger impact for most stations and therefore we correct for meridional wind stress afterward and sea level pressure last.

For each station, the corresponding latitude and longitude grid cells are identified and the variables retrieved. Since the NCEP/NCAR Reanalysis 1 provides zonal (u) and meridional (v) wind velocity, these are used to compute the wind stress, using similar methods as Frederikse & Gerkema (2018)[11], including the parameterization of the drag coefficient C_D following Pugh & Woodworth (2014)[12].

$$\frac{\tau_u}{\rho_{air}} = C_D u\sqrt{u^2 + v^2} \tag{3}$$

$$\frac{\tau_v}{\rho_{air}} = C_D v\sqrt{u^2 + v^2} \tag{4}$$

$$C_D = 0.8 + 0.065\sqrt{u^2 + v^2} \tag{5}$$

Zonal and meridional wind stress regression is then performed in the aforementioned sequential fashion. We do not specify the air density ρ_{air}, as its effect is accounted for in the regression. The regression model is expressed as an ordinary linear least squares problem. Firstly, the mean height after nodal correction is subtracted from the time series, simplifying the model by not including an intercept. Adding an option for an intercept to the design matrix is equally valid and yields the same results. All ordinary least square solutions have been computed using the Statsmodels Python module[13].

$$\eta_2(t) = \eta_1(t) - mean\{\eta_1(t)\} \tag{6}$$

The regression model is defined below, in which y is the regressand, in our case τ_u/ρ_{air}. β contains the regression coefficient. The design matrix X then follows as the entries of the nodal-corrected, mean-removed, MSL $\eta_2(t)$.

$$y = X\beta \tag{7}$$

$$X = \begin{bmatrix} \eta_2(t_0) \\ \eta_2(t_1) \\ \vdots \\ \eta_2(t_m) \end{bmatrix} \tag{8}$$

The solution to this model is the commonly known Least-Squares solution. This solution is then used to predict the height based on the regression of the zonal wind stress, and this prediction is then removed from the signal.

$$\hat{\beta} = (X'X)^{-1}X'y \tag{9}$$

$$\eta_{\tau_u}(t) = \eta_2(t) * X\hat{\beta} \tag{10}$$

$$\eta_3(t) = \eta_2(t) - \eta_{\tau_u} \tag{11}$$

In similar fashion, the regression is sequentially performed on the meridional wind stress, with $y = \frac{\tau_v}{\rho_{air}}$ and $\eta_3(t)$ and on sea level pressure, with $y = slp$ and $\eta_4(t)$.

$$\eta_4(t) = \eta_3(t) - \eta_{\tau_v} \tag{12}$$

$$\eta_{cor}(t) = \eta_4(t) - \eta_{slp} \tag{13}$$

The resulting corrected sea level record, $\eta_{cor}(t)$, is used for all subsequent trend modeling.

2.3 Trend modeling

Trend modeling is performed to investigate the influence of record length on found sea level change. In general, a model is computed every 10 years of record length, with the exception of the first 20 years, in which a 5 year gap was used in order to better visualize any intricacies that might occur when using these relative short time series in modeling. A total of 9 dates have been picked for investigation, being 1949, 1959, 1969, 1979, 1989, 1999, 2004, 2009 and 2014. These years all denote a record length from that particular year till present (i.e. the most recent available data). In the case of the stations with shorter time series, only the years for which there was near-continuous RLR data available are included.

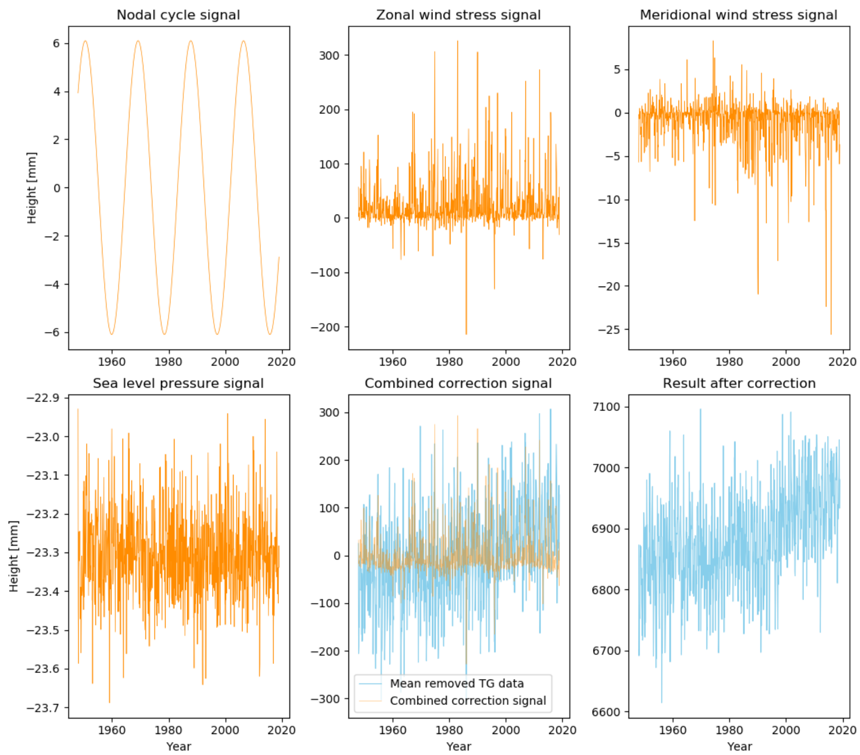


Figure 2: Correction terms for the MSL, based on the tide gauge station in Vlissingen. Left top: the nodal cycle equilibrium tide signal. Middle top: effects of the zonal wind stress. Right top: effects of the meridional wind stress. Left bottom: effects of the sea level pressure. Middle bottom: a comparison between the original signal and the cumulative correction term. Right bottom: the resulting corrected time series.

$$\eta_{cor}(t) = a * t + \eta_0 + \epsilon \tag{14}$$

$$X = \begin{bmatrix} 1 & t_0 \\ 1 & t_1 - t_0 \\ 1 & t_2 - t_0 \\ \vdots & \vdots \\ 1 & t_m - t_0 \end{bmatrix} \tag{15}$$

After picking the preferred dates, the data point, i.e. the month closest to the chosen dates, has to be found. This actually results in the starting month of all models to become December the year prior. Next, ordinary least squares models are constructed for trend modeling, very similar to the regression models for the tide gauge correction in Equations 8 and 9. The linear model is shown in Equation 14, in which the linear term is the regression coefficient a. In the ordinary least squares sense, this trend is computed using the following design matrix X, with t_0 the chosen start year (and t_m the most recent available date). The quadratic model, Equation 16, enhances the linear model by adding a quadratic term b to the equation. The used sea level acceleration is defined as twice the quadratic term.

$$\eta_{cor}(t) = b * t^2 + a * t + \eta_0 + \epsilon \tag{16}$$

$$X = \begin{bmatrix} 1 & t_0 & t_0^2 \\ 1 & t_1 - t_0 & (t_1 - t_0)^2 \\ 1 & t_2 - t_0 & (t_2 - t_0)^2 \\ \vdots & \vdots & \vdots \\ 1 & t_m - t_0 & (t_1 - t_0)^2 \end{bmatrix} \quad (17)$$

2.4 Uncertainty analysis

The model coefficients' standard error is computed using the heteroskedasticity-consistent covariance matrix estimator (HC0) from White (1980)[14], in which r_i are the residuals. In least squares sense, this is expressed as in Equation 18. These are used for both the linear and acceleration in Figure 3 and Figure 4 to indicate error margins.

$$\sigma = \sqrt{diag(X'X)^{-1}X' * diag(r_i^2)X(X'X)^{-1}} \quad (18)$$

However, for analysis whether a linear model or quadratic model fits the data better, different techniques have to be used. A simple method is to compare the errors of the different models by computing the root mean squared error (RMSE), based on the residuals r_i between the model and η_{cor}. A smaller RMSE indicates a better fit to the available data.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{i=n} r_i^2} \quad (19)$$

3 Results

3.1 Time series length

Tide gauge records along, and in close proximity of, the Dutch coast have been analyzed for a period of 1948 - 2018. They are corrected for nodal cycle & wind stresses. Unlike the Zeespiegelmonitor, the choice has been made to also regress for sea level pressure to correct for the inverse barometer effect, as Wunsch & Stammer (1997)[15] have concluded that inverse barometric can indeed decrease sea level variance. Different model start moments have been used to register how both a linear and quadratic model differ over the length of the time series chosen. The found linear trend and in the case of the quadratic model, the acceleration, are shown in Figures 3 and 4. A comparison between stations can be found in the appendix.

Figure 3 shows that for longer models, linear trends between stations seem to vary quite a bit already, ranging from 1.19 ± 0.28 mm/yr for West-Terschelling to 2.55 ± 0.24 mm/yr for Hoek van Holland (2σ interval) when starting in 1948, with a Dutch station mean of 1.84 ± 0.24 mm/yr. Individual trends slightly vary when choosing a model with a time span of 40 years (starting in 1978), but follow a slightly upward change, ranging from 1.69 ± 0.50 mm/yr for Maassluis to 3.00 ± 0.78 mm/yr for Borkum-Fischerbalje and the Dutch station mean being 2.36 ± 0.54 mm/yr. Uncertainties clearly become bigger when using the 40 year model compared to the 70 year model, as expected. A prestudy has shown that usage of longer time series than currently shown, would result in similar effects on uncertainty and inter-station variability. Due to climate effects being less profound in the first half of the twentieth century, we assume that the numbers for a model starting in 1890 could drive the linear trend down, similarly to the results shown when comparing 1948 with 1978 as start year. As the trends are based relative sea level (RSL), subsidence signals in the found trends will naturally not be affected in this assumption.

When using even shorter models, with start dates after 1988, trends found change, for some stations quite rapid. As found trends on average increase, a logical conclusion is that the sea level change is increasing. However, inter-station and intra-station variability increases significantly and so do the associated uncertainties. Local variability becomes more prominent in the tide gauge signal when using shorter models. These local variabilities are not well explained through the corrections applied and seem to affect different stations in different ways. For some stations, we know where to look for explanations. For example, it is known that subsidence effects due to gas extraction greatly affects the station of Delfzijl and to lesser extent Hoek van Holland [16]. In the

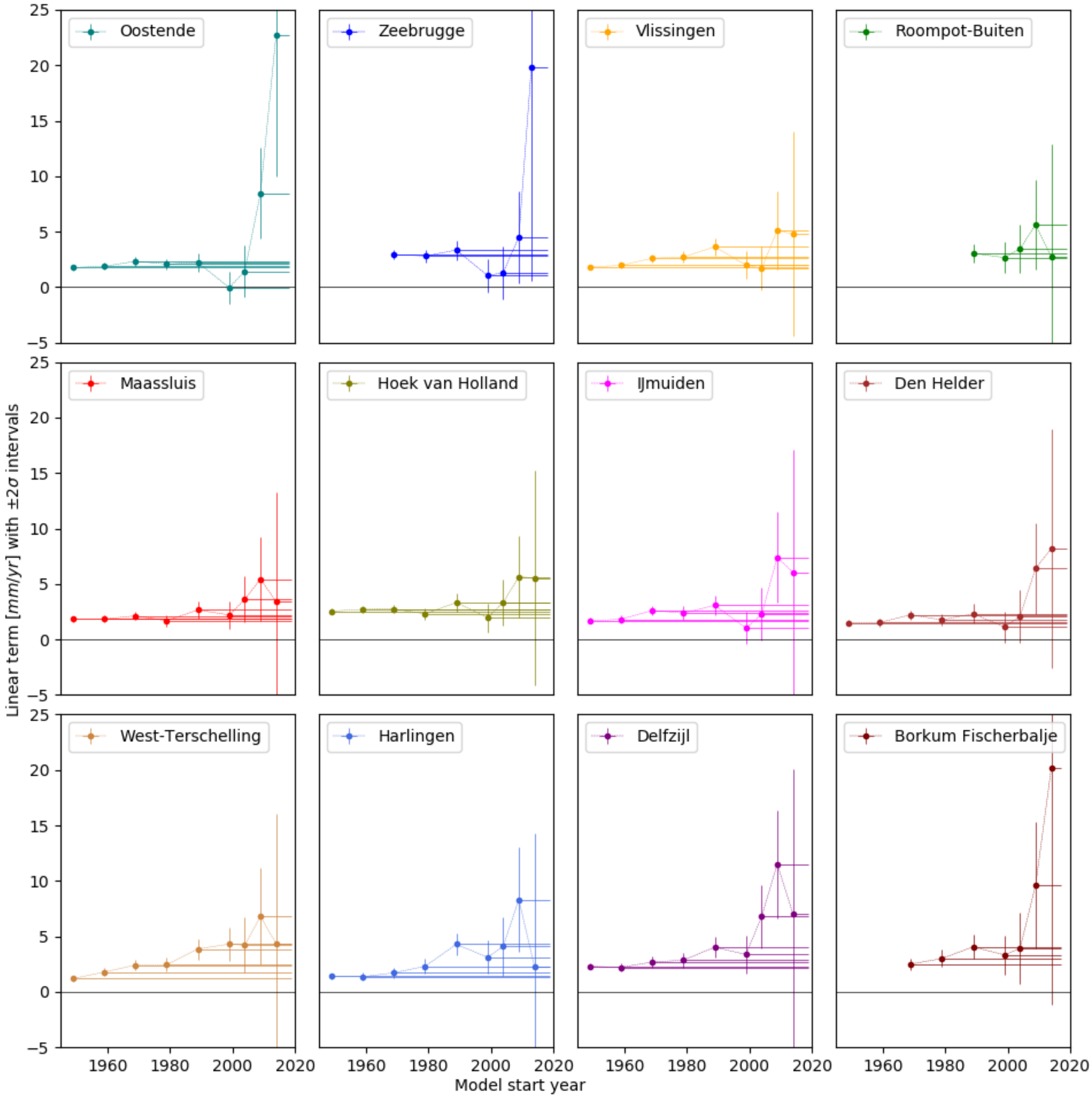


Figure 3: Found sea level change rate, based on a linear model, for different model start years. Each dot indicates the start year, with corresponding horizontal line showing the complete length of the time series used for the model value. Vertical error bars indicate the 2σ confidence intervals.

case of Delfzijl, this is related to the Groningen gas field and might explain relatively high trends found recently compared to other stations.

The inter-station variability starts to increase in the period 1978-1998, which could indicate that this is a period where a possible break point might be found. The used methods and unexplained shorter time-scale variabilities make it difficult to actually quantify such a possible break point, but do provide an insight where to look.

A possible break point can also be identified when looking at the acceleration found in the quadratic model in Figure 4. With the exception of Harlingen, with an acceleration of 0.28 ± 0.12 mm^2/yr, and West-Terschelling, with 0.21 ± 0.12 mm^2/yr, (2σ interval) in 1978, all found accelerations before 1988 are smaller than 0.2 mm^2/yr. More clear variations can be seen to emerge when using models starting in the period of 1978-1998, similarly to the found break point period in the linear models. After 1998, both acceleration magnitude and their uncertainties in acceleration become very big. Borkum-Fischerbalje and Zeebrugge show big accelerations, while other stations show deceleration. It must be noted that linear components of these quadratic models often counteract the effect of the quadratic components. For example, when looking at the models starting in 2008, Zeebrugge provides an acceleration of 5.02 ± 4.27 mm^2/yr, paired with a linear term of -18.38 ± 21.88 mm/yr. Conversely, stations with a negative acceleration show similar signals, with linear terms often negating the effect of strong deceleration. We conclude that application of quadratic models time series shorter than 20 years is deemed unstable, given the used corrections. Due to this instability, the Dutch station mean is not very meaningful for these shorter models. For models that run longer than 20 years, most individual acceleration numbers found do seem to fall within the 2σ-confidence interval of the station mean.

3.2 Acceleration uncertainty

To determine whether found acceleration can be deemed significant within the quadratic models, four model lengths have been chosen to investigate the uncertainty. In Figure 5, the magnitude of acceleration terms have been plotted against the corresponding uncertainty. Please note the different scales for the 10 year model length plot. Points that fall in between the uncertainty lines are up for debate whether the found acceleration will be considered significant or not. This feeds into a discussion whether a trend with a uncertainty of equal magnitude should be discarded as too uncertain or does still hold some value.

What can be concluded from these plots is that, first and foremost, longer models result in more accurate acceleration terms. For the 10-year model, none of the found accelerations have a magnitude bigger than their 2σ uncertainties, whereas a 70-year model shows that all stations but Hoek van Holland and Maassluis (which are very close geographically) can be considered significant. The Dutch station mean has an estimated acceleration of $0.044 \pm 0,013(1\sigma)$ mm^2/yr over the period 1948-2018, which means there is less than 0.05% chance that the actual acceleration is smaller than or equal to zero. Shorter models show that even though most stations and the Dutch station mean have bigger uncertainties than actual estimated acceleration, some individual station's acceleration is still very certain. For example, take a look at Zeebrugge in the 30 year length model. The found acceleration for Zeebrugge is $-0.34 \pm 0.11(1\sigma)$ mm^2/yr, which means the signal is bigger than its 3σ uncertainty.

3.3 Model selection

For model selection, we have used the RMSE of both the linear and quadratic model and compared these in Figure 6. Most notable are that the differences in RMSE between linear and quadratic models are minimal, which indicates that, in general, both models perform equally well. When looking into absolute numbers, the quadratic model performs slightly

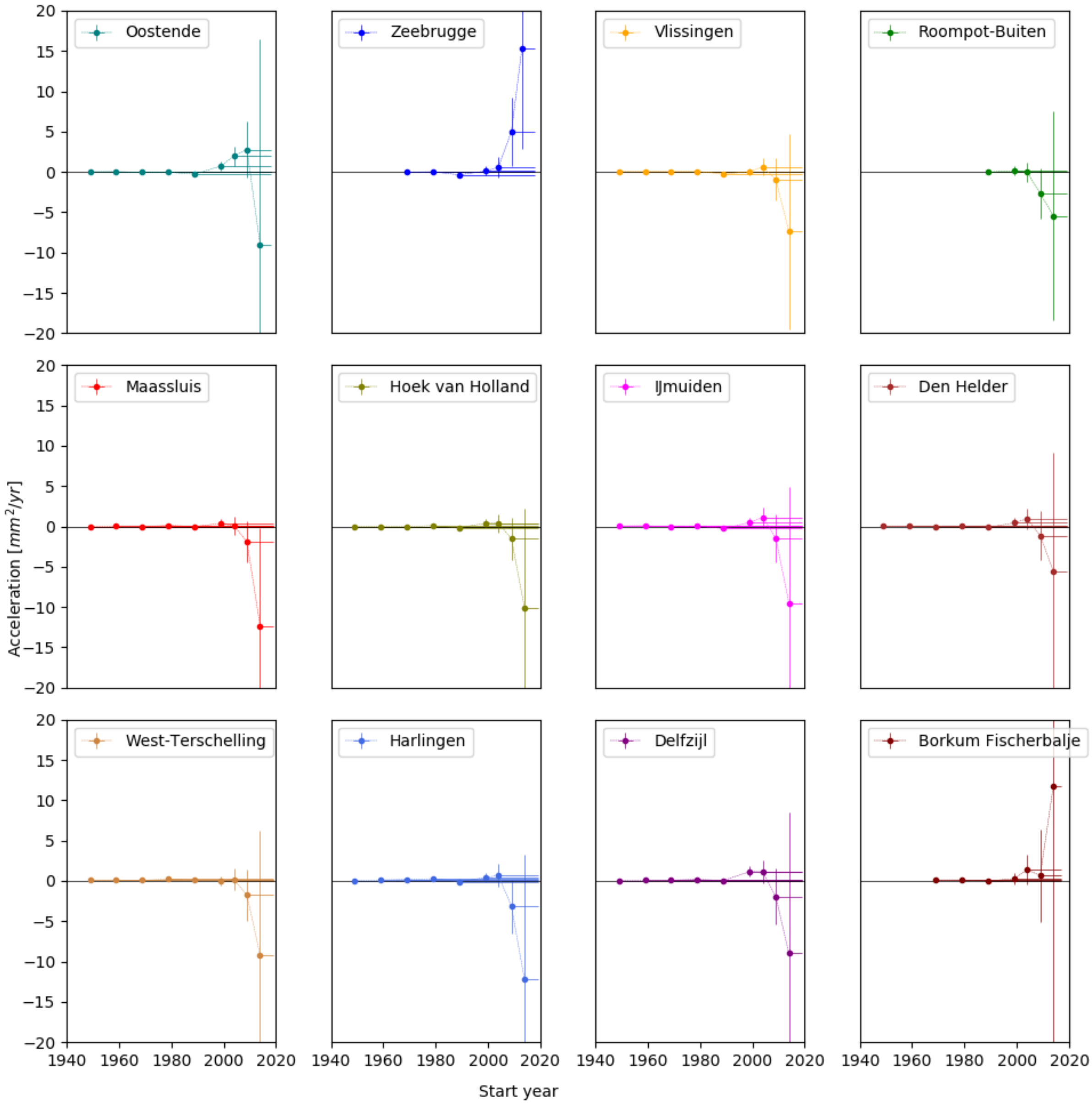
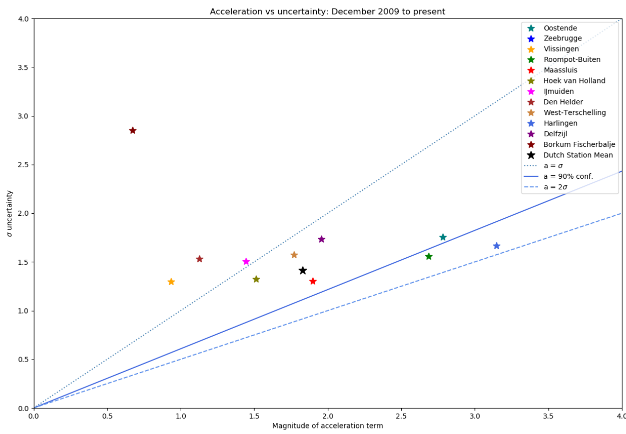
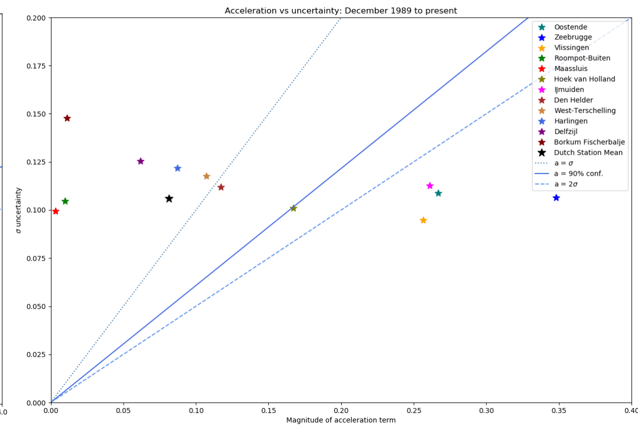


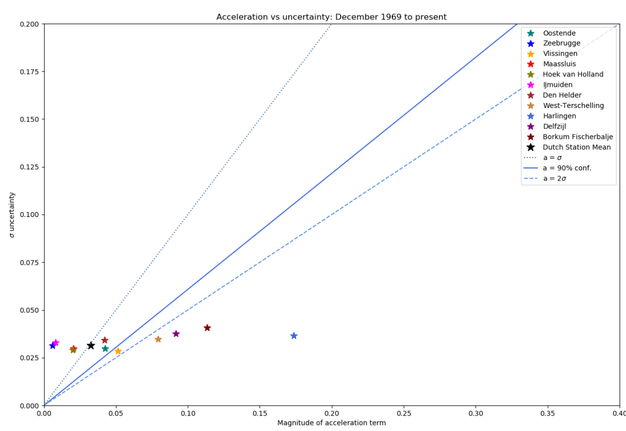
Figure 4: Found sea level acceleration, based on a quadratic model, for different model start years. Each dot indicates the start year, with corresponding horizontal line showing the complete length of the time series used for the model value. Vertical error bars indicate the 2σ confidence intervals.



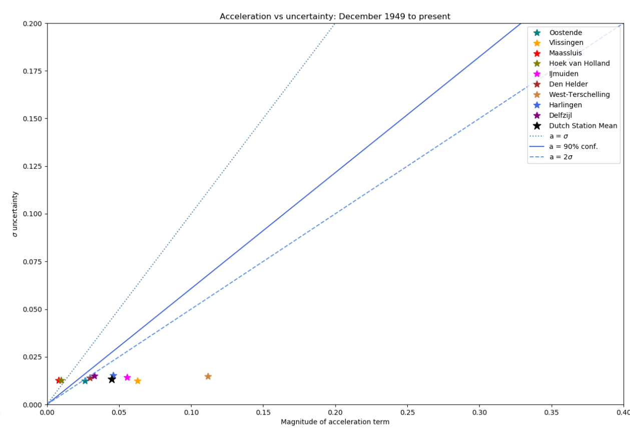
(a) 10 year model length.



(b) 30 year model length.



(c) 50 year model length.



(d) 70 year model length.

Figure 5: Acceleration terms versus their corresponding uncertainty. Points left of/above the three uncertainty lines have a bigger uncertainty than acceleration and could be deemed insignificant. Points to the right of/below these lines have a smaller uncertainty than acceleration and therefore are significant. The three lines indicate three different, commonly used, confidence intervals: 1σ (68% confidence), 90% confidence and 2σ (95% confidence).

better for more start years than the linear model. With the exceptions of Zeebrugge, Roompot-Buiten and Maassluis, the quadratic model is considered a better fit more often than not. The same accounts for the Dutch station mean, where 7 out of 9 model start years result in a slight preference for the quadratic model. There seems to be no clear predictor into which years correspond to a quadratic preference. This leads to our conclusion that the RMSE is not suitable for model selection in this case.

For comparison of the linear model with the quadratic model and deciding which fits best, the Zeespiegelmonitor uses the Akaike information criterion (AIC)[17], which is able to provide a trade-off between goodness of fit and model simplicity. They assume that the model with the smallest AIC is the best model. A short look into the AIC values we found, shows that based on this assumption, quadratic models are generally better for models with length longer than 40 years and worse for the shorter models. It must be noted that there is some variability in these results and we did not thoroughly look into this as we had some concerns with the assumption made that the model with the smallest AIC is the best. More on this in the discussion.

4 Discussion and conclusions

When applying either linear or quadratic models, we think it is sensible to shorten time series length in order to better quantify the current sea level change. This allows for more accurate corrections due to better data availability. A time series start year of 1948 or later allows the inclusion of the wind stress and sea level pressure corrections from NCEP/NCAR Reanalysis 1. In 1978, uncertainties and inter-station variability are still relatively small and so any choice of model start year in between 1948 and 1978 are sufficient to investigate modern-day tide gauge data. Models in which older data is incorporated, but given less weight to compensate for less accurate corrections might also be possible. Break points could possibly be found in the period between 1978 and 1998. If one wishes to further investigate whether there is a break point, a time series which includes a run-up to the possible break point is advised and therefore we recommend to use data starting in 1948 - 1968. Using shorter time series (1988-2018 records) for individual stations is possible, but it must be noted that the quadratic models do not perform well.

Due to inter-station variability, using a station mean to make conclusions on a national level is not advised for these shorter time series. For time series starting in 1948-1968, the station mean does represent the individual stations decently well, but some concerns do arise for usage of quadratic models and station means. Individual stations show relatively big variability in acceleration, but low variability in corresponding uncertainty. The result is that some stations will show an acceleration which is not or hardly significant due to uncertainties being higher than found acceleration terms. Other stations show very significant accelerations, several times bigger than corresponding uncertainties. Due to the small amount of stations available, a station mean is very sensitive to the amount of stations of latter or former characteristics. Choosing to include or omit certain stations greatly influences the mean acceleration found and whether the result can be considered significant or not. We advise strongly to not base conclusions on acceleration of the sea level along the Dutch coast solely on a station mean, but to always incorporate individual station differences.

It is quite common in academics to be very careful of making conclusions which are on the edge of significance, which is a good thing. These conclusions are oftentimes the ones communicated to the wider public through press releases and media. But whether a found acceleration is deemed significant is a subjective choice. What criterion should be used? If one chooses to adopt a more stringent criterion of 2σ, conclusions as to whether the sea level is accelerating or not can wildly differ from a more relaxed criterion of 1σ. In Figure 5, we deliberately included three different confidence intervals which are often used in statistical analysis to visualize the consequences of such a choice. If, for example, one would find an acceleration of 1 ± 1 mm^2/yr (2σ), one

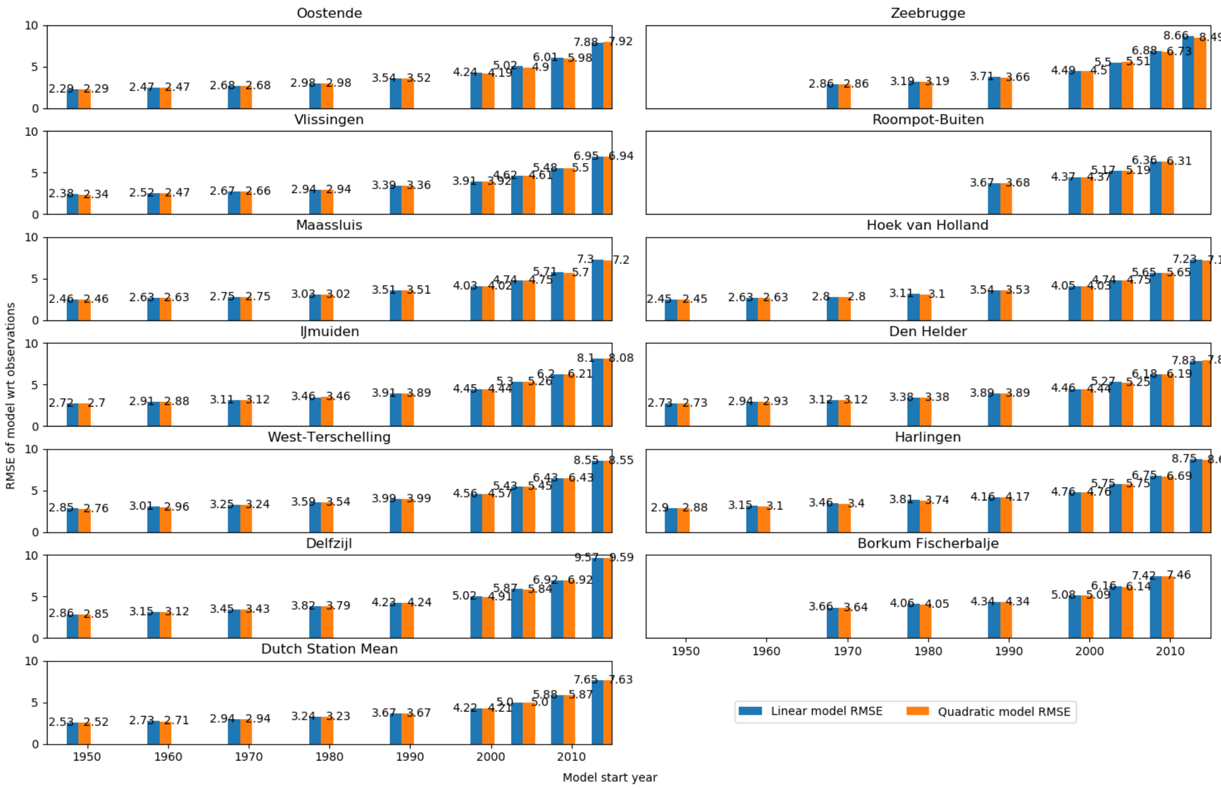


Figure 6: Linear and quadratic model RMSE for each station and model start year. A relatively lower RMSE indicates a better fit of the model to the corrected MSL time series and might support adoption of a linear over a quadratic model, or vice versa.

could say that there is no significant acceleration, implying the acceleration equals 0 $\mathrm{mm}^2/\mathrm{yr}$. If several stations all show such a similar signal, a conclusion of no acceleration might be too careful. In the scope of this study, concerning the current knowledge on global sea level change, it might be actually be a more sensible approach to also be careful of stating that there is no acceleration. One could for example state that current data or methods are not sufficient to increase certainty of the found acceleration. This is the conclusion we draw concerning model adoption.

This discussion also touches on the assumption of the Zeespiegelmonitor that a smaller AIC means a better model by definition. The AIC is a more sophisticated method than that. In principle, a smaller AIC means the information loss compared to the truth is estimated to be smaller. This can be used to quantify the so-called relative likelihood, which is the probability that a model with a bigger AIC actually minimizes the estimated information loss. For two candidate models mdl_1 and mdl_2, where mdl_1 has a smaller AIC, the relative likelihood of mdl_2 is expressed in Equation 20.

$$\exp\{\frac{AIC(mdl_1) - AIC(mdl_2)}{2}\} \tag{20}$$

Based on the found values, one could make choices to omit candidate model mdl_2. If the relative likelihood of candidate model mdl_2 is high, however, this means that it is likely for the model mdl_2 to be closer to the truth. In such a case, omitting might not be the best option. Conclusions could be made that the data is not sufficient to select a model or a weighted average between the models mdl_1 and mdl_2 could be used to compute a new model, which might have a better resulting AIC than both. For the results of the Zeespiegelmonitor, this means that the broken linear model is considered best. The quadratic model then is 0.61 times as probable to minimize information loss as the broken linear model. The regular linear model is 0.14 times as probable to minimize information loss as the broken linear model, or 0.22 as probable when compared to the quadratic model. These are all values for which omission of models is questionable. A more thorough investigation into the differences of found AIC based on time series length might give useful insights.

References

[1] Van den Hurk,B., Siegmund, P., Klein Tank, A., Attema, J., Bakker, A., Beersma, J., Bessembinder, J., Boers, R., Brandsma, T., Van den Brink, H., Drijfhout, S., Eskes, H., Haarsma, R., Lenderink, G., Loriaux, J., Van Meijgaard, E., Van Noije, T., Van Oldenborg, G.J., Selten, F., Siebesma, P., Sterl, A., De Vries, H., Van Weele, M., De Winter, R. & Van Zadelhoff, G. (2014), KNMI'14: Climate Change scenarios for the 21st Century – A Netherlands perspective, KNMI, De Bilt, The Netherlands. Scientific Report WR2014-01, www.climatescenarios.nl

[2] Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex & P. M. Midgley (2013), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.https://www.ipcc.ch/report/ar5/wg1/

[3] Baart,F., Rongen, G., Hijma, M., Kooi, H., de Winter & R., Nicolai, R. (2019), Zeespiegelmonitor 2018, Deltares, Delft, The Netherlands. Tech.Rep. 11202193-000-ZKS-0004 https://www.deltares.nl/app/uploads/2019/03/Zeespiegelmonitor-2018-final.pdf

[4] Jevrejeva, S., Moore, J. C., Grinsted, A., Matthews, A. P., & Spada, G. (2014). Trends and acceleration in global and regional sea levels since 1807. Global and Planetary Change, 113, 11–22. https://doi.org/10.1016/j.gloplacha.2013.12.004

[5] Holgate, S., Matthews, A., Woodworth, P., Rickards, L., Tamisea, M., Bradshaw, E., Foden, P., Gordon, K., Jevrejeva, S. & Pugh, J. (2013) New Data Systems and Products at the Permanent Service for Mean Sea Level. Journal of Coastal Research: Volume 29, Issue 3: pp. 493 – 504. https://doi.org/10.2112/jcoastres-d-12-00175.1

[6] Permanent Service for Mean Sea Level (PSMSL), 2020, "Tide Gauge Data", Retrieved 28 Apr 2020 from http://www.psmsl.org/data/obtaining/complete.

[7] Woodworth, P. L. (2012). A Note on the Nodal Tide in Sea Level Records. Journal of Coastal Research, 280, 316–323. https://doi.org/10.2112/jcoastres-d-11a-00023.1

[8] Frederikse, T., Riva, R., Kleinherenbrink, M., Wada, Y., van den Broeke, M., & Marzeion, B. (2016). Closing the sea level budget on a regional scale: Trends and variability on the Northwestern European continental shelf. Geophysical Research Letters, 43(20), 10,864-10,872. https://doi.org/10.1002/2016gl070750

[9] Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynods, R., Jenne, R. & Joseph, D. (1996). The NCEP/NCAR 40-Year Reanalysis Project. Bulletin of the American Meteorological Society, 77(3), 437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2

[10] Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., Allan, R., Yin, X., Vose, R., Titchner, H., Kennedy, J., Spencer, L. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Crouthamel, R., Moore, G.W.K., Przybylak, R., Pudmenzky, C., Reason, C., Slonosky, V., Smith, C., Tinz, B., Trewin, B., Antónia Valente, M., Wang, X., Wilkinson, C., Wood, K. & Wyszyński, P. (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. Quarterly Journal of the Royal Meteorological Society, 145(724), 2876–2908. https://doi.org/10.1002/qj.3598

[11] Frederikse, T., and Gerkema, T. (2018). Multi-decadal variability in seasonal mean sea level along the North Sea coast. Ocean Science, 14(6), 1491–1501. https://doi.org/10.5194/os-14-1491-2018

[12] Pugh, D. and Woodworth, P. (2014): Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes. Cambridge University Press, Cambridge, New York. https://doi.org/10.1063/PT.3.2753

[13] Seabold, S. and Perktold, J. (2012)Statsmodels: Econometric and statistical modeling with python. Proceedings of the 9th Python in Science Conference 92-96.

[14] White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica, 48(4), 817. https://doi.org/10.2307/1912934

[15] Wunsch, C., & Stammer, D. (1997). Atmospheric loading and the oceanic 'inverted barometer' effect. Reviews of Geophysics, 35(1), 79–107. https://doi.org/10.1029/96rg03037

[16] Hijma, M. & Kooij, H. (2017) Bodemdaling in het kustfundament en de getijdenbekkens, Deltares. *Part I*:Tech. Rep. 11200538-00. *Part II*: Tech. Rep. 11202190-001-ZKS-0001

[17] Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716–723. https://doi.org/10.1109/tac.1974.1100705

5 Appendix

Composite images have been created to visualize how linear and quadratic models differ in between stations and can be found in Figure 7 and Figure 8.

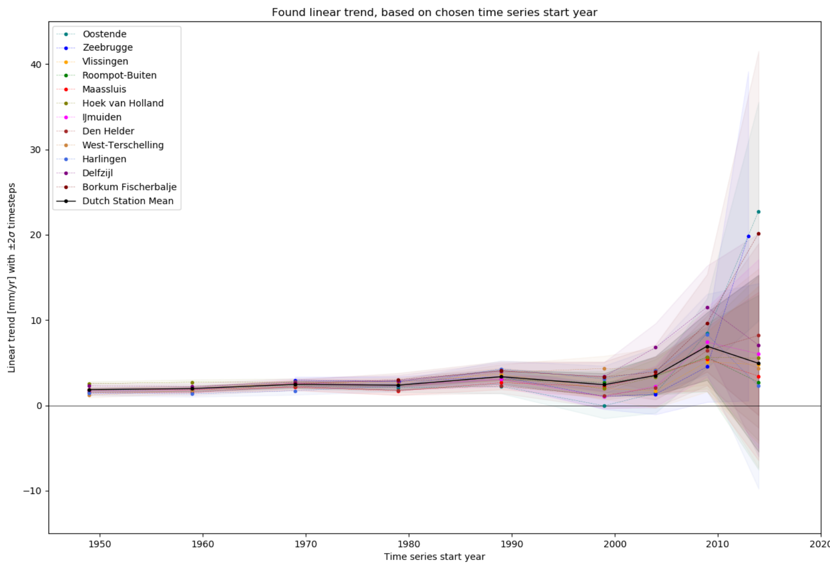


Figure 7: Composite of found sea level change rate, based on a linear model, for different model start years. Colored areas indicate the 2σ uncertainty associated with the trend found for each given start year. A Dutch station mean has been included in black.

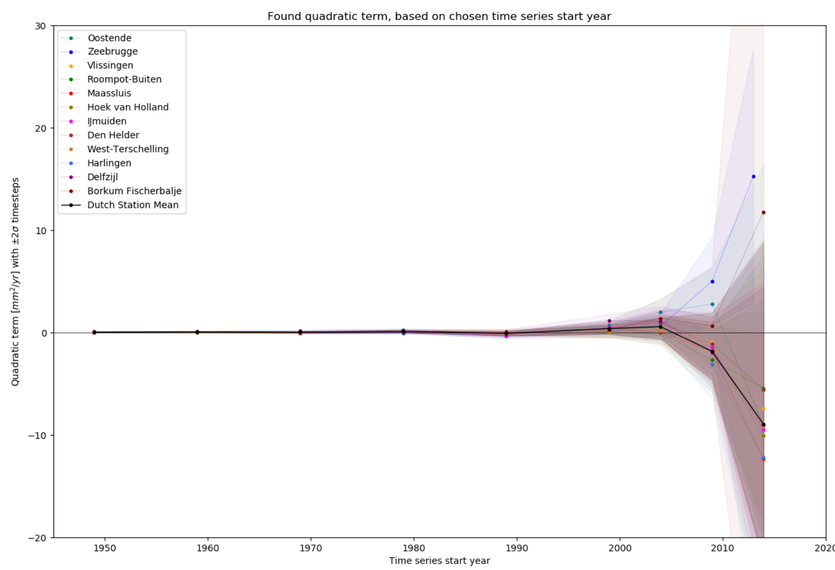


Figure 8: Found sea level acceleration, based on a quadratic model, for different model start years. Colored areas indicate the 2σ uncertainty associated with the acceleration found for each given start year. A Dutch station mean has been included in black.